



**EVALUACIÓN DE ALGORITMOS DE APRENDIZAJE AUTOMÁTICO EN LA
CLASIFICACIÓN SUPERVISADA DE IMÁGENES SATELITALES
MULTITEMPORALES PARA EL ANÁLISIS DE COBERTURAS DE SUELO EN
RIONEGRO, COLOMBIA**

Oscar Estevan Cossio Madrid
Jhon Sebastian Herrera Palacio

Monografía para optar al título de Especialista en Medio Ambiente y Geoinformática

Tutor
Juan José García Duque

Universidad de Antioquia
Facultad de Ingeniería
Especialización en Medio Ambiente y Geoinformática
Medellín
2024

Cita	(Cossio Madrid & Herrera Palacio, 2024)
Referencia	Cossio Madrid., & Herrera Palacio (2024). <i>Evaluación de Algoritmos de Aprendizaje Automático en la Clasificación Supervisada de Imágenes Satelitales Multitemporales para el Análisis de Coberturas de Suelo en Rionegro, Colombia</i> . [Monografía]. Universidad de Antioquia, Medellín
Estilo APA 7 (2020)	



Especialización en Medio Ambiente y Geoinformática.



Centro de Documentación Ingeniería (CENDOI)

Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Tabla de contenido

Resumen	8
Abstract	9
Introducción	10
1 Planteamiento del problema	12
2 Justificación.....	13
3 Objetivos	14
3.1 Objetivo general	14
3.2 Objetivos específicos.....	14
4 Marco teórico	15
4.1 Algoritmos de Clasificación Supervisada en Google Earth Engine.....	16
4.2 Métodos de validación.	17
4.2.1 Matriz de Confusión	17
4.2.2 Curva ROC y AUC	18
4.2.3 Índice de Kappa	18
4.3 Casos de éxito en la clasificación supervisada	19
5 Metodología	20
5.1 Área de estudio.....	22
5.2 Fuentes de datos satelitales	23
5.3 Conexión a la API de GEE.....	24
5.4 Selección de puntos de entrenamiento.	24
5.5 Procesamiento de imágenes satelitales.....	29
5.6 Calculo de métodos de clasificación por GEE	32
5.6.1 Método de clasificación CART	34
5.6.2 Método de clasificación Random Forest.....	34

5.6.3 Metodo de clasificación KNN	35
5.6.4 Método de clasificación Naive Bayes.....	36
5.6.5 Método de clasificación Gradient Tree Boost	36
5.7 Métodos de validación	37
6 Resultados	39
6.1 Resultados Provenientes Del Dataset De Corine Land Cover Para Nivel 3	39
6.1.1 Método de CART.....	39
6.1.2 Método de Naive Bayes	42
6.1.3 Método de KNN.....	45
6.1.4 Método de Random Forest.....	49
6.1.5 Comparativa de métodos y consolidado	51
6.2 Resultados Provenientes Del Dataset De Corine Land Cover Para Nivel 2.	53
6.3 Resultados Provenientes Del Dataset De Corine Land Cover Para Nivel 1.	56
7 Conclusiones	60
8 Recomendaciones.....	62
Referencias	64

Lista de tablas

Tabla 1. Consolidado de resultados por método de clasificación.	33
Tabla 1. Consolidado de resultados por método de clasificación.	52
Tabla 2. Comparativa Métodos de Clasificación CLC - N3	53
Tabla 3. Consolidado de resultados por método de clasificación.	54
Tabla 4. Comparativa Métodos de Clasificación CLC – N2.....	55
Tabla 5. Consolidado de resultados por método de clasificación.	56
Tabla 6. Comparativa Métodos de Clasificación CLC – N1.....	58

Lista de figuras

Figura 1. Diagrama conceptual de la metodología.....	21
Figura 2. Mapa de la zona de estudio.....	22
Figura 3. Dataset de Sentinel 2 obtenido.....	24
Figura 4. Flujograma de gestión y selección de puntos de entrenamiento.....	26
Figura 5. Proceso de clasificación de puntos de entrenamiento.....	29
Figura 6. Resultado clasificación método CART.....	40
Figura 7. Matriz de Confusión Método CART.....	41
Figura 8. Curvas ROC para la clasificación con CART.....	42
Figura 9. Resultado clasificación método Naive Bayes.....	43
Figura 10. Matriz de Confusión Método Naive Bayes.....	44
Figura 11. Curvas AUC para la clasificación con Naive Bayes.....	45
Figura 12. Resultado clasificación método KNN.....	46
Figura 13. Matriz de Confusión Método KNN.....	47
Figura 14. Curvas AUC para la clasificación con KNN.....	48
<i>Figura 15. Resultado clasificación método random Forest.....</i>	<i>49</i>
<i>Figura 16. Matriz De Confusion, Random Forest.....</i>	<i>50</i>
Figura 17. Curvas AUC para la clasificación con Random Forest.....	51

Siglas, acrónimos y abreviaturas

CLC: CORINE Land Cover

GEE: Google Earth Engine

KNN: K-Nearest Neighbors (Vecinos más Cercanos)

CART: Classification and Regression Trees (Árboles de Clasificación y Regresión)

RF: Random Forest (Bosque Aleatorio)

AUC: Área Bajo la Curva ROC (Area Under the Curve)

ROC: Receiver Operating Characteristic (Característica Operativa del Receptor)

NDVI: Índice de Vegetación de Diferencia Normalizada (Normalized Difference Vegetation Index)

NDBI: Índice de Construcción de Diferencia Normalizada (Normalized Difference Built-Up Index)

SR: Superficie Reflectante (Surface Reflectance)

F1-Score: Medida de Equilibrio entre Precisión y Recall

AOI: Área de Interés (Area of Interest)

SVM: Máquinas de Soporte Vectorial (Support Vector Machines)

RF: Bosques Aleatorios (Random Forest)

ML: Aprendizaje Automático (Machine Learning)

DL: Aprendizaje Profundo (Deep Learning)

GIS: Sistema de Información Geográfica (Geographic Information System)

Kappa: Índice de Kappa de Cohen para Evaluación de Concordancia

R²: Coeficiente de Determinación (R-Squared)

Resumen

El análisis de coberturas de suelo mediante algoritmos de aprendizaje automático es fundamental para la gestión sostenible del territorio, especialmente en contextos de alta complejidad como Rionegro, Antioquia, Colombia. Este estudio implementa una metodología robusta para evaluar el desempeño de cinco algoritmos de clasificación supervisada (Random Forest, K-Nearest Neighbors, Naive Bayes, Gradient Boosting y CART) utilizando imágenes Sentinel-2 multiespectrales procesadas en Google Earth Engine (GEE). La investigación integra preprocesamiento avanzado, enmascaramiento de nubes, cálculo de índices espectrales (NDVI, NDBI, NDWI) y balanceo de clases mediante SMOTE para abordar problemas de desbalance en los datos.

Los resultados muestran que Random Forest se destaca con una precisión global promedio del 83%, un índice Kappa de 0.81 y valores de área bajo la curva (AUC) superiores a 0.9 en las principales clases de cobertura. El algoritmo KNN, aunque menos robusto, logró desempeños aceptables en clases con características espectrales bien diferenciadas. Por el contrario, Naive Bayes y CART demostraron limitaciones importantes, reflejadas en precisiones globales inferiores al 30% y un índice Kappa cercano a cero, atribuibles a la alta fragmentación del paisaje y el solapamiento espectral. El análisis de curvas ROC y matrices de confusión resalta que las clases relacionadas con áreas urbanas y cuerpos de agua presentan la mejor discriminación, mientras que las clases agrícolas y de vegetación mixta son más propensas a errores.

Este trabajo refuerza la relevancia de los algoritmos basados en ensambles, como Random Forest, para entornos de alta heterogeneidad y complejidad espectral. Además, subraya la importancia de incluir índices espectrales y ajustar los datos de entrenamiento mediante técnicas de balanceo y validación cruzada estratificada. Los hallazgos no solo contribuyen al desarrollo de metodologías aplicadas a la teledetección en zonas tropicales, sino que también ofrecen lineamientos prácticos para mejorar la precisión en el monitoreo multitemporal y la gestión de recursos en áreas de alto valor ambiental.

Palabras clave: clasificación supervisada, Sentinel-2, aprendizaje automático, validación cruzada, Google Earth Engine, índices espectrales, fragmentación espacial.

Abstract

Land cover analysis using machine learning algorithms is essential for sustainable land management, particularly in complex contexts such as Rionegro, Antioquia, Colombia. This study implements a robust methodology to evaluate the performance of five supervised classification algorithms (Random Forest, K-Nearest Neighbors, Naive Bayes, Gradient Boosting, and CART) using Sentinel-2 multispectral imagery processed in Google Earth Engine (GEE). The research integrates advanced preprocessing, cloud masking, spectral index calculation (NDVI, NDBI, NDWI), and class balancing through SMOTE to address imbalances in the training dataset.

The results highlight Random Forest as the most effective model, achieving an average overall accuracy of 83%, a Kappa index of 0.81, and area under the curve (AUC) values above 0.9 for major land cover classes. While KNN showed moderate success in spectrally distinct classes, Naive Bayes and CART exhibited significant limitations, with overall accuracies below 30% and Kappa indices close to zero, primarily due to high landscape fragmentation and spectral overlap. The ROC curves and confusion matrices underscore that urban and water-related classes are well-discriminated, whereas agricultural and mixed vegetation classes are more prone to misclassification.

This research reinforces the suitability of ensemble-based algorithms, such as Random Forest, in heterogeneous and spectrally complex environments. It also highlights the importance of incorporating spectral indices and balancing training data through stratified cross-validation. These findings not only advance remote sensing methodologies in tropical regions but also provide practical guidelines for improving accuracy in multitemporal monitoring and resource management in high-value environmental areas.

Keywords: supervised classification, Sentinel-2, machine learning, cross-validation, Google Earth Engine, spectral indices, spatial fragmentation.

Introducción

La teledetección se ha vuelto esencial en las investigaciones ambientales actuales, ya que posibilita el análisis minucioso y en tiempo real de la superficie terrestre a través de imágenes obtenidas por satélites. Las tecnologías actuales proporcionan una visión inigualable de las transformaciones en la utilización del suelo, la vegetación y la evolución de los ecosistemas, elementos fundamentales para la planificación sustentable y la reducción de los impactos del cambio climático (Chuvieco, 2016; Pettorelli et al., 2014). La capacidad de monitorear cambios ambientales y realizar evaluaciones precisas en diferentes regiones del mundo ha mejorado notablemente gracias al análisis multitemporal y multiespectral de imágenes de alta resolución, como han señalado Hansen et al. (2013) y Woodcock et al. (2020).

No obstante, a pesar de los progresos mencionados, existen desafíos significativos en el progreso y la implementación de técnicas de clasificación supervisada de imágenes satelitales. En contextos complejos, donde las coberturas del suelo son altamente heterogéneas y dinámicas, algoritmos como Random Forest (RF), Support Vector Machines (SVM) y Redes Neuronales Convolucionales (CNN) han demostrado un notable potencial para aumentar la precisión de las clasificaciones (Belgiu & Drăguț, 2016; Ma et al., 2019). A pesar de que estos métodos han sido extensamente validados en regiones templadas y desarrolladas, la investigación en contextos tropicales es limitada. En estas zonas, las condiciones climáticas, la diversidad ecológica y las prácticas de uso del suelo son distintas, lo que requiere un análisis específico (Shao & Lunetta, 2012; Toth & Jóźków, 2016).

El caso de Rionegro, ubicado en el departamento de Antioquia en Colombia, ejemplifica de manera precisa la importancia de llevar a cabo investigaciones en esta área. La región de Rionegro se caracteriza por su alta biodiversidad, sin embargo, está siendo afectada por cambios rápidos en el uso del suelo provocados por la expansión de áreas urbanas, el aumento de la actividad agrícola y la pérdida de bosques. La dinámica mencionada ha suscitado inquietudes en relación con la disminución de los servicios ecosistémicos y la viabilidad del desarrollo local, como señala el IDEAM en su informe del año 2020. A pesar de la importancia de estas problemáticas, los estudios realizados en la región generalmente se basan en métodos convencionales de cartografía y seguimiento, los cuales frecuentemente no logran capturar la complejidad y la rapidez de los cambios en la utilización del suelo (IDEAM, 2020).

Actualmente, la mayoría de las aplicaciones de algoritmos de clasificación supervisada para el mapeo de coberturas de suelo se han evaluado principalmente en regiones con datos de alta calidad y abundancia, como Estados Unidos, Europa y algunas áreas de Asia (Gislason et al., 2006; Shao & Lunetta, 2012). Investigaciones recientes han evidenciado la eficacia de técnicas como la Regresión Logística y las Máquinas de Vectores de Soporte (SVM) en el procesamiento de conjuntos de datos multiespectrales complejos, lo que resulta en la generación de mapas de cobertura del suelo con un alto nivel de precisión. No obstante, en contextos como el de Rionegro, las metodologías mencionadas no han sido validadas extensamente. En este lugar, factores como los efectos atmosféricos, la presencia frecuente de nubes y la marcada variabilidad estacional pueden impactar de manera considerable en la calidad y exactitud de los resultados de la clasificación (Hansen et al., 2013; Gorelick et al., 2017).

En consecuencia, el presente estudio tiene como objetivo abordar las disparidades mencionadas a través de una evaluación detallada de diversos algoritmos de clasificación

supervisada utilizados en el análisis de imágenes satelitales de alta resolución en la región de Rionegro. Este estudio tiene como objetivo optimizar la precisión de los modelos de clasificación y generar conocimiento aplicable a la gestión sostenible del territorio en la región, mediante un enfoque metodológico que integra técnicas avanzadas de machine learning, análisis multispectral y multitemporal, y validación cruzada rigurosa. Al realizar esta acción, se busca no solamente incrementar la comprensión de las dinámicas relacionadas con el uso del suelo en Rionegro, sino también aportar al avance de nuevas metodologías destinadas al monitoreo ambiental en zonas que presenten similitudes en sus características.

1 Planteamiento del problema

La clasificación de imágenes satelitales mediante algoritmos de clasificación supervisada es fundamental para la identificación de coberturas del suelo, especialmente en contextos donde la variabilidad de las coberturas y las condiciones climáticas son significativas.

Diferentes métodos han sido implementados en diversas regiones, pero su aplicabilidad y precisión pueden variar notablemente en función de factores como la heterogeneidad del paisaje, el uso del suelo y las condiciones ambientales. Esta situación ha llevado a la generación de múltiples enfoques para la clasificación, cada uno con sus particularidades y fortalezas, lo que subraya la necesidad de un análisis comparativo que no solo identifique los algoritmos más precisos, sino que también permita entender cómo se comportan en condiciones específicas.

Este estudio se propone, entonces, realizar un análisis de las distintas metodologías de clasificación en herramientas como Google Earth Engine y Python, enfocándose en su rendimiento en un contexto específico como el del municipio de Rionegro. A través de esta evaluación, se busca proporcionar bases que respalden la selección de algoritmos para el mapeo de coberturas del suelo, optimizando así el uso de teledetección en la gestión ambiental y la toma de decisiones estratégicas.

2 Justificación

La clasificación supervisada de imágenes satelitales es una herramienta clave para el monitoreo y la gestión ambiental del territorio. Sin embargo, elegir el algoritmo más adecuado para una región en particular sigue siendo un desafío considerable, especialmente cuando se busca lograr una alta precisión en la identificación de las coberturas del suelo. Según lo señalado por Maxwell et al. (2018), la efectividad de los algoritmos de clasificación puede variar considerablemente en función de las características del área de estudio y de la naturaleza de las coberturas terrestres presentes.

Contar con datos de referencia de alta calidad es fundamental para validar la precisión de cada metodología y determinar cuál de los algoritmos se ajusta mejor a las particularidades del territorio. Zhang y Roy (2017) indican que la validación sistemática de algoritmos de clasificación, utilizando datos confiables, es esencial para establecer protocolos metodológicos que puedan aplicarse de manera efectiva en contextos similares.

Este estudio pretende contribuir a la formulación de lineamientos metodológicos basados en evidencia para la selección de algoritmos de clasificación en situaciones específicas. De esta manera, se espera mejorar la fiabilidad de los productos cartográficos generados a través de la teledetección, facilitando así una gestión más efectiva y sostenible del territorio.

3 Objetivos

3.1 Objetivo general

Evaluar y comparar diferentes métodos de clasificación supervisada para identificar las clases de cobertura del suelo en el municipio de Rionegro, Antioquia, durante el periodo 2017-2023, utilizando imágenes satelitales y análisis de matrices de confusión con el fin de determinar el modelo más adecuado para esta región.

3.2 Objetivos específicos

- Comparar la eficacia de los métodos de clasificación supervisada disponibles en plataformas como Google Earth Engine mediante su API de python, analizando su rendimiento en la identificación de clases de cobertura del suelo en Rionegro.
- Evaluar las diferencias en la precisión y eficiencia de los modelos de clasificación implementados, utilizando matrices de confusión para identificar el método más efectivo para esta región.
- Investigar el impacto de la variabilidad de los datos de entrada (imágenes satelitales, variables climáticas) en el rendimiento de los métodos de clasificación, con el fin de identificar las mejores prácticas para su aplicación en la región.
- Generar recomendaciones basadas en los resultados del análisis comparativo, orientadas a la selección de métodos de clasificación supervisada más adecuados para el monitoreo y gestión de la cobertura del suelo en Rionegro, considerando la viabilidad y accesibilidad de las herramientas disponibles.

4 Marco teórico

La teledetección es definida como el proceso de obtención de información sobre objetos o fenómenos sin contacto físico directo, a través de sensores que detectan y registran la energía electromagnética reflejada o emitida por la superficie terrestre (Jensen, 2015). Estos datos son esenciales para analizar dinámicas espaciales y temporales en la Tierra, facilitando aplicaciones en ecología, hidrología, agricultura, y planificación urbana.

El principio fundamental de la teledetección radica en la interacción de la radiación electromagnética con la superficie terrestre, lo que genera patrones específicos de reflectancia que varían según las propiedades físicas y químicas de los materiales. Según Drusch et al. (2012), los sensores multiespectrales, como los de la misión Sentinel-2, capturan esta interacción en diferentes bandas del espectro electromagnético, proporcionando información crítica para diferenciar tipos de cobertura terrestre y analizar su distribución espacial.

Los sensores multiespectrales se destacan por registrar información en múltiples bandas del espectro electromagnético, desde el visible hasta el infrarrojo cercano (NIR) y de onda corta (SWIR). Este enfoque permite capturar firmas espectrales únicas, esenciales para discriminar coberturas terrestres como vegetación, agua y áreas urbanas (Xue & Su, 2017). Las bandas espectrales específicas, como las del infrarrojo cercano, son particularmente útiles para analizar la salud de la vegetación, mientras que las bandas del SWIR permiten evaluar contenido de humedad en el suelo y la vegetación (Main-Knorn et al., 2017).

La clasificación supervisada es un método de análisis de imágenes que utiliza un conjunto de datos de referencia para entrenar algoritmos que asignan etiquetas de clase a los píxeles no clasificados. Este enfoque se fundamenta en la creación de un modelo que establece relaciones estadísticas entre las características espectrales de los píxeles y las clases de cobertura terrestre conocidas (Breiman, 2001).

Maxwell et al. (2018) destacan que la clasificación supervisada se basa en la representatividad de los datos de entrenamiento, la selección adecuada de características espectrales y la robustez de los algoritmos empleados. La implementación de este método requiere no solo un conjunto de entrenamiento bien construido, sino también estrategias para manejar problemas como la alta dimensionalidad, el ruido y el desbalance entre clases.

En este estudio, se emplearán diversos métodos de clasificación supervisada para el análisis de imágenes satelitales, incluyendo algoritmos disponibles en Google Earth Engine. Google Earth Engine (GEE) es una plataforma basada en la nube que proporciona acceso a grandes volúmenes de imágenes satelitales y herramientas de procesamiento geoespacial. GEE es particularmente útil para estudios de gran escala, ya que permite realizar análisis complejos de forma rápida y eficiente, sin necesidad de grandes recursos computacionales locales (Gorelick et al., 2017).

Python es un lenguaje de programación ampliamente utilizado para el análisis de datos geoespaciales, especialmente en combinación con bibliotecas como scikit-learn, geopandas y rasterio. Su flexibilidad permite personalizar los algoritmos de clasificación según las necesidades específicas del estudio. Además, Python es particularmente adecuado para implementar métodos

estadísticos avanzados y procesar grandes volúmenes de datos espaciales (Van der Walt et al., 2011).

4.1 Algoritmos de Clasificación Supervisada en Google Earth Engine

Random Forest: Se basa en el método de bosque aleatorio (RF), introducido por Breiman (2001). RF construye múltiples árboles de decisión a partir de muestras aleatorias de datos, combinando sus resultados para obtener una clasificación final. En teledetección, el método RF es ampliamente utilizado debido a su capacidad para manejar grandes conjuntos de datos y su robustez frente a variables ruidosas o irrelevantes (Breiman, 2001). En GEE, el algoritmo `smileRandomForest` permite realizar clasificaciones de imágenes multiespectrales, siendo especialmente útil para análisis de coberturas heterogéneas (Gorelick et al., 2017).

Naive Bayes: Este algoritmo aplica el enfoque probabilístico de Naive Bayes, basado en el teorema de Bayes. Asume que las variables predictoras son independientes entre sí, lo cual simplifica el cálculo de probabilidades y permite una clasificación rápida. Sin embargo, esta suposición limita su precisión cuando existe interdependencia entre las variables, lo que puede ocurrir en el análisis de coberturas de suelo donde las bandas espectrales están correlacionadas. A pesar de sus limitaciones, Naive Bayes es útil para clasificaciones básicas o contextos en los que la velocidad de procesamiento es prioritaria sobre la precisión (Zhang, 2004).

KNN (K-Nearest Neighbors): Este método clasifica un punto en función de la clase predominante en sus vecinos más cercanos. Este algoritmo es útil en teledetección al identificar patrones en datos multiespectrales, aunque su rendimiento puede verse afectado por el tamaño de los datos y la dimensionalidad de las variables (Cover & Hart, 1967). KNN es efectivo para áreas con clases bien diferenciadas, aunque suele ser computacionalmente más costoso a medida que aumenta el número de puntos a clasificar (Shakhnarovich et al., 2005).

Gradient Tree Boost: Es un algoritmo de aumento de gradiente que construye múltiples árboles de decisión de forma secuencial y cada árbol trata de corregir los errores cometidos por los árboles anteriores. Este enfoque permite una clasificación altamente precisa en datos complejos y heterogéneos, aunque requiere un mayor tiempo de procesamiento debido a su complejidad computacional (Friedman, 2001). En el contexto de la clasificación de coberturas de suelo, este método es adecuado cuando se dispone de datos multiespectrales con alta variabilidad, permitiendo una identificación precisa de diferentes clases de suelo y vegetación (Chen & Guestrin, 2016).

Cart (Classification and Regression Trees): Este algoritmo utiliza árboles de decisión para dividir los datos en subconjuntos en función de criterios específicos de las variables. Los árboles de decisión CART son modelos de fácil interpretación, aunque propensos a sobreajustarse cuando se enfrentan a datos complejos (Breiman et al., 1984). En aplicaciones de teledetección, CART es útil para clasificaciones sencillas, aunque puede carecer de precisión en comparación con métodos como Random Forest o Gradient Boosting, especialmente en contextos con gran heterogeneidad de coberturas (Friedl & Brodley, 1997).

4.2 Métodos de validación.

4.2.1 Matriz de Confusión

La matriz de confusión es una tabla cuadrada de dimensiones $C \times C$, donde C representa el número total de clases en el problema de clasificación. Cada entrada (i,j) en la matriz indica el número de observaciones cuya clase verdadera es i y que han sido clasificadas como j por el modelo.

$$\text{Matriz de Confusión} = \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1C} \\ n_{21} & n_{22} & \cdots & n_{2C} \\ \vdots & \vdots & \ddots & \vdots \\ n_{C1} & n_{C2} & \cdots & n_{CC} \end{bmatrix}$$

Donde n_{ij} es el número de píxeles cuya clase verdadera es i y cuya clase predicha es j .

4.2.1.1 Precisión global (accuracy)

La precisión global (accuracy) es una métrica que mide la proporción de observaciones correctamente clasificadas en relación con el total de observaciones evaluadas.

$$\text{Precisión Global} = \frac{\sum_{i=1}^C n_{ii}}{\sum_{i=1}^C \sum_{j=1}^C n_{ij}}$$

4.2.1.2 Precisión por clase

La precisión por clase (precisión) evalúa la proporción de predicciones correctas realizadas para una clase específica, en comparación con todas las observaciones predichas como pertenecientes a esa clase.

$$\text{Precision} = \frac{n_{ii}}{\sum_{j=1}^C n_{ji}}$$

4.2.1.3 Sensibilidad por clase (recall)

La sensibilidad por clase (recall) mide la capacidad del modelo para identificar correctamente las observaciones pertenecientes a una clase específica.

$$\text{Recall}_i = \frac{n_{ii}}{\sum_{j=1}^C n_{ij}}$$

4.2.1.4 F1-score por clase

El F1-Score por clase es la media armónica entre la precisión y la sensibilidad, proporcionando un equilibrio entre ambas métricas.

$$F1_i = 2 * \frac{\text{Precision}_i * \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i}$$

4.2.2 Curva ROC y AUC

La Curva ROC (Receiver Operating Characteristic) y el AUC (Área Bajo la Curva) son herramientas fundamentales para evaluar el desempeño de modelos de clasificación, especialmente en problemas multiclase o binarios. Estas métricas permiten medir la capacidad del modelo para distinguir entre clases.

Tasa de Verdaderos Positivos (True Positive Rate, TPR): Proporción de instancias positivas correctamente clasificadas.

$$TPR = \frac{TP}{TP + FN}$$

Tasa de Falsos Positivos (False Positive Rate, FPR): Proporción de instancias negativas clasificadas incorrectamente como positivas.

$$FPR = \frac{FP}{FP + TN}$$

Cada punto en la curva ROC representa un umbral de clasificación específico.

El AUC mide el área total bajo la curva ROC, proporcionando una métrica escalar para evaluar el modelo. Los valores del AUC oscilan entre 0 y 1:

AUC = 1: Clasificación perfecta.

AUC = 0.5: Clasificación aleatoria.

AUC < 0.5: Rendimiento peor que aleatorio.

Matemáticamente, el AUC se calcula integrando la curva ROC:

$$AUC = \int_0^1 TPR d(FPR)$$

4.2.3 Índice de Kappa

El Índice Kappa de Cohen es una métrica que evalúa la concordancia entre dos evaluadores (o entre las etiquetas predichas y las verdaderas) más allá de lo que podría ocurrir por azar. Es particularmente útil en problemas de clasificación multiclase, ya que proporciona una medida ajustada que considera tanto la precisión observada como la precisión esperada.

El índice Kappa se define como:

$$k = \frac{P_o - P_e}{1 - P_e}$$

Donde:

P_o = Proporción de concordancia observada entre las etiquetas verdaderas y predichas.

$$P_o = \frac{\sum_{i=1}^c n_{ii}}{\sum_{i=1}^c \sum_{j=1}^c n_{ij}}$$

Pe: Proporción de concordancia esperada por azar, calculada como:

$$P_o = \sum_{i=1}^c \left(\frac{\sum_{j=1}^c n_{ij} * \sum_{j=1}^c n_{ji}}{\sum_{i=1}^c \sum_{j=1}^c n_{ij}} \right)$$

Y en donde:

El valor de κ se interpreta como sigue:

- $\kappa=1$: Concordancia perfecta.
- $\kappa=0$: Concordancia equivalente a lo esperado por azar.
- $\kappa<0$: Concordancia peor que el azar.

4.3 Casos de éxito en la clasificación supervisada

El éxito de la clasificación supervisada depende en gran medida de la calidad de los datos de entrada. Según Main-Knorn et al. (2017), procesos como la corrección atmosférica, el enmascaramiento de nubes y la estandarización de las bandas espectrales son esenciales para garantizar que los valores espectrales sean consistentes y comparables. El producto Sentinel-2 Surface Reflectance (SR), procesado con el algoritmo Sen2Cor, es un ejemplo de datos preprocesados que cumplen con estos estándares.

La clasificación supervisada ha sido aplicada con éxito en diversos contextos. Maxwell et al. (2018) señalan que este método es fundamental para el monitoreo de cambios de cobertura terrestre, permitiendo detectar tasas de deforestación, expansión urbana y degradación ambiental. Además, Gorelick et al. (2017) destacan su relevancia en la agricultura de precisión, donde algoritmos supervisados clasifican áreas agrícolas según su estado fenológico y salud, optimizando la toma de decisiones.

A pesar de sus ventajas, la clasificación supervisada enfrenta desafíos como el desbalance en las clases, la selección de características relevantes y la integración de múltiples fuentes de datos. Según Chawla et al. (2002), técnicas como SMOTE (Synthetic Minority Oversampling Technique) son efectivas para mitigar los problemas de desbalance, generando instancias sintéticas que mejoran la representatividad de las clases minoritarias.

En el futuro, se espera que la integración de técnicas de aprendizaje profundo y plataformas de procesamiento en la nube, como Google Earth Engine, impulse aún más la precisión y escalabilidad de los métodos de clasificación supervisada en teledetección (Gorelick et al., 2017).

5 Metodología

En la actualidad, el análisis de coberturas terrestres constituye un elemento fundamental para la comprensión de los procesos de transformación del territorio y sus dinámicas espacio-temporales (Anderson et al., 2023). La integración de tecnologías de observación terrestre con algoritmos de aprendizaje automático ha revolucionado la manera en que procesamos y analizamos estos cambios, permitiendo un monitoreo más preciso y eficiente de los ecosistemas terrestres.

El presente estudio se centra en la evaluación y comparación de diferentes métodos de clasificación supervisada aplicados al análisis de coberturas terrestres en el municipio de Rionegro, Antioquia, durante el periodo 2017-2023. Para ello, se emplearon imágenes satelitales de alta resolución, índices espectrales relevantes y algoritmos avanzados de aprendizaje automático. Este enfoque busca determinar el modelo más adecuado para la identificación de clases de cobertura del suelo, utilizando métricas de desempeño como la matriz de confusión y el índice Kappa, en plataformas computacionales como Google Earth Engine y Python, conocidas por su capacidad de manejar datos geoespaciales masivos (Gorelick et al., 2017; Van der Walt et al., 2011).

La clasificación supervisada constituye un pilar en la teledetección, permitiendo la asignación de etiquetas de clase a píxeles no clasificados mediante el entrenamiento de modelos predictivos. Este proceso se fundamenta en la relación entre las firmas espectrales capturadas en distintas bandas del espectro electromagnético y las características biofísicas de los objetos terrestres (Jensen, 2015). En este estudio se consideran los algoritmos Random Forest (RF), K-Nearest Neighbors (KNN), Naive Bayes y Gradient Boosting, seleccionados por su versatilidad y capacidad para manejar datos multiespectrales y multitemporales (Breiman, 2001; Friedman, 2001; Zhang, 2004).

La metodología se estructura en una serie de etapas integradas que abarcan desde la adquisición y preprocesamiento de datos hasta la implementación, evaluación y comparación de los modelos supervisados. Estas etapas permiten un análisis sistemático y reproducible, asegurando la robustez de los resultados. La secuencia metodológica incluye la delimitación del área de estudio, el procesamiento de imágenes satelitales de Sentinel-2, la preparación de datos de entrenamiento, la aplicación de algoritmos supervisados y la validación de resultados mediante métricas estándar en la literatura científica (Maxwell et al., 2018).

El diseño metodológico puede representarse mediante un mapa conceptual que sintetiza las interacciones clave entre las fases del análisis. Este diagrama conceptual (ver Figura 1, a desarrollar) ilustra la secuencia lógica de actividades, incluyendo la delimitación del área de estudio, el preprocesamiento de imágenes satelitales, la preparación de datos de entrenamiento, la implementación de modelos supervisados, y la validación de resultados.

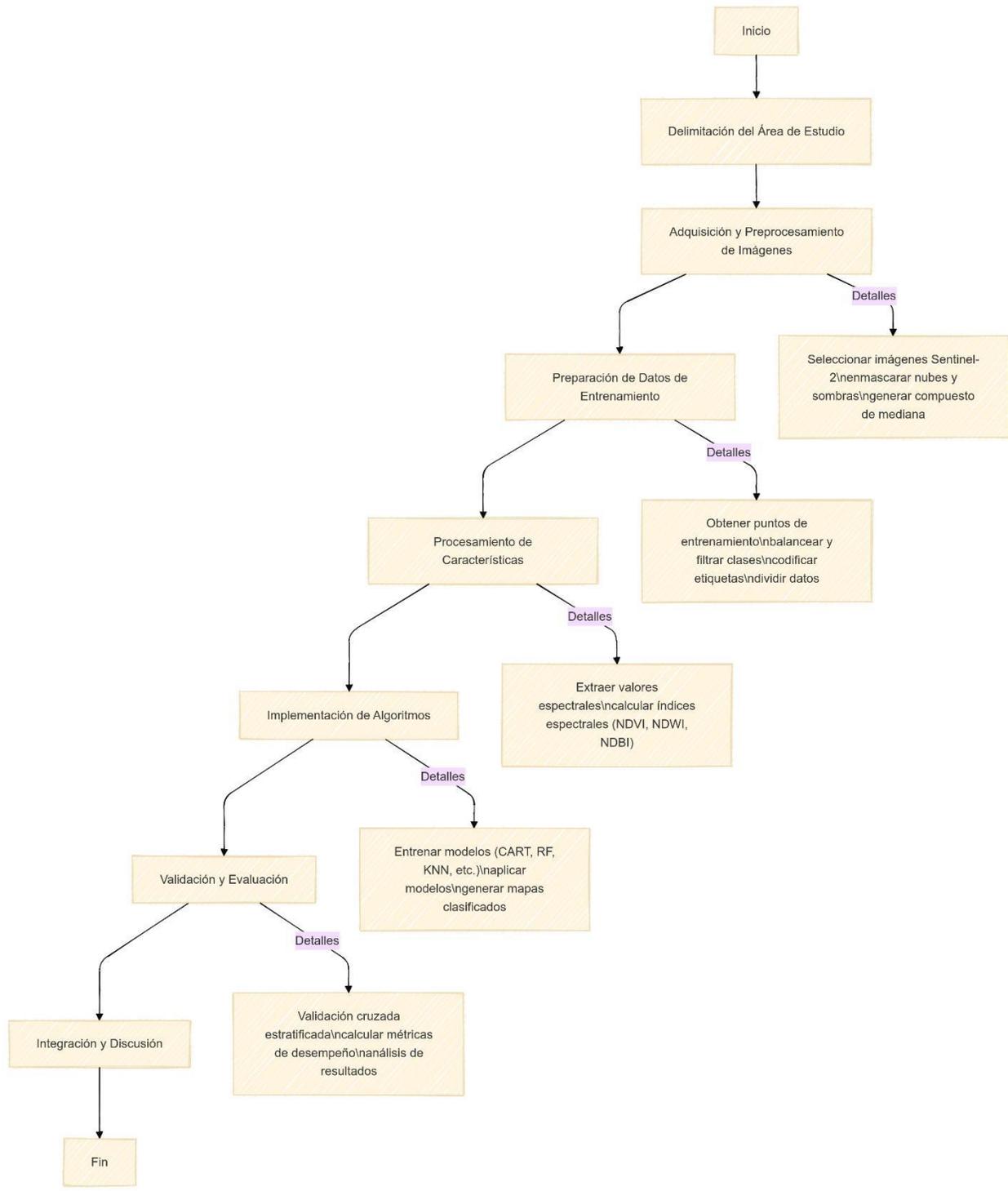


Figura 1. Diagrama conceptual de la metodología

5.1 Área de estudio

El municipio de Rionegro, ubicado en el departamento de Antioquia, Colombia, se seleccionó como área de estudio por su relevancia ambiental y dinámica socioeconómica, características que lo convierten en un escenario representativo para la evaluación de coberturas terrestres. Rionegro se encuentra en la subregión del Oriente Antioqueño, con una extensión aproximada de 196 km² y una altitud promedio de 2,125 metros sobre el nivel del mar (m.s.n.m.) (Instituto Geográfico Agustín Codazzi [IGAC], 2020).

El municipio presenta una diversidad de clases de cobertura terrestre, incluyendo áreas urbanas, cultivos agrícolas, cuerpos de agua y bosques secundarios, lo que genera una complejidad espectral para evaluar el desempeño de diferentes algoritmos de clasificación debido a su gran fragmentación. Además, su localización estratégica y crecimiento urbano acelerado resaltan la necesidad de contar con herramientas precisas para monitorear los cambios de uso del suelo y facilitar la gestión territorial.

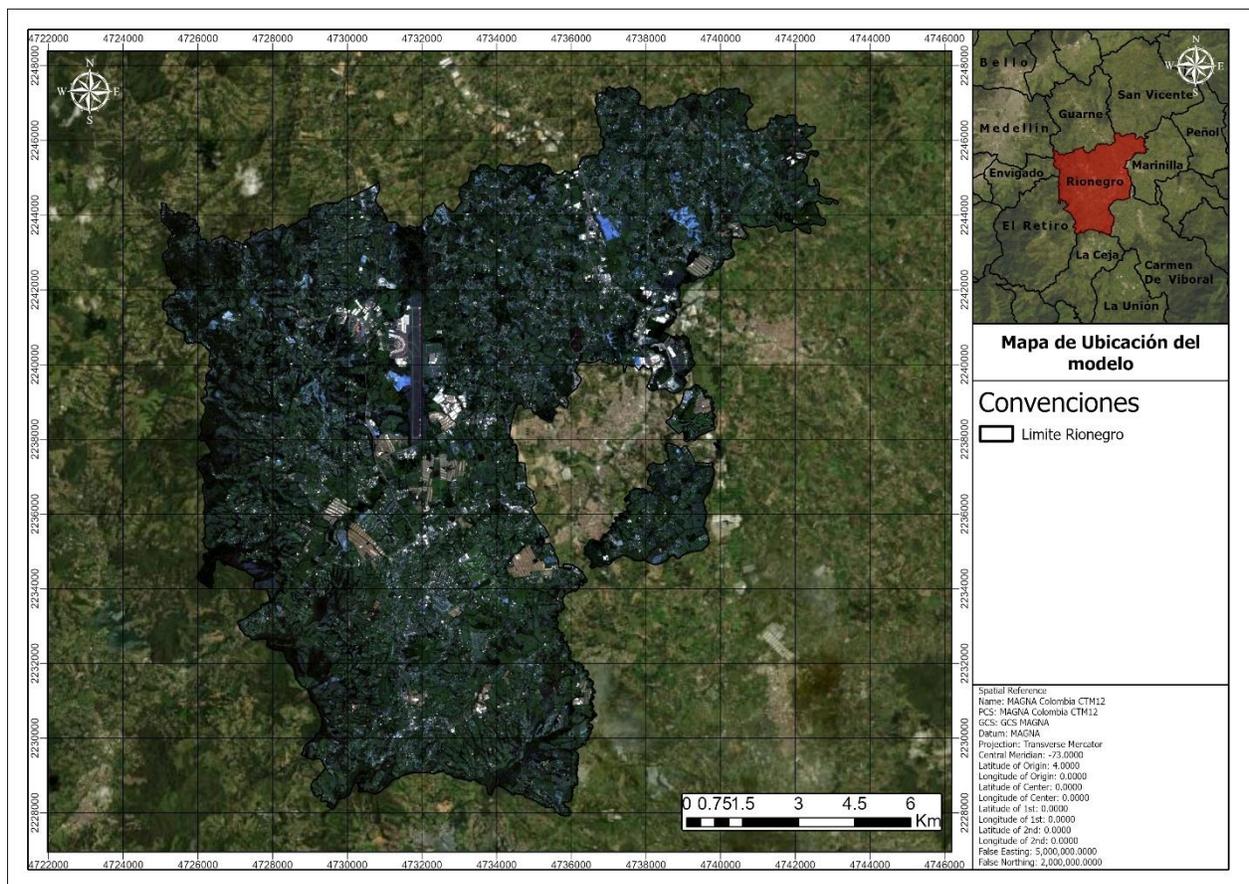


Figura 2. Mapa de la zona de estudio.

Desde el punto de vista climático, Rionegro pertenece a una zona de alta montaña, con un régimen de lluvias bimodal y temperaturas promedio entre 12°C y 18°C (Instituto de Hidrología, Meteorología y Estudios Ambientales [IDEAM], 2020). Estas características generan patrones

estacionales que impactan la reflectancia espectral de las coberturas terrestres, haciendo necesario el uso de datos multitemporales para capturar esta variabilidad.

Por lo tanto, el dataset de entrenamiento se definió mediante la capa de cobertura de suelo del municipio de Rionegro perteneciente a la corporación autónoma regional de las cuencas de los ríos Negro y Nare (Cornare) el cual contaba con clasificación hasta nivel 3 según la metodología Corine Land Cover Colombia. Esto permitió mediante el cálculo de los centroides de los polígonos, entender la distribución de clases presentes en la zona de estudio.

5.2 Fuentes de datos satelitales

El análisis de coberturas terrestres en el municipio de Rionegro utilizó imágenes de la colección Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-2A, disponible en Google Earth Engine (GEE). Esta colección, identificada como COPERNICUS/S2_SR_HARMONIZED, ofrece productos de reflectancia de superficie (SR) generados mediante el algoritmo Sen2Cor, que corrige efectos atmosféricos y asegura la consistencia espectral entre escenas pre y post-2022 (European Union/ESA/Copernicus, 2024).

Las imágenes Sentinel-2 se caracterizan por una resolución espacial que varía entre 10 y 60 metros, con 12 bandas espectrales que cubren el espectro visible, infrarrojo cercano (NIR) e infrarrojo de onda corta (SWIR). Para este estudio, se emplearon las bandas B2 (azul), B3 (verde), B4 (rojo), B8 (NIR) y B11/B12 (SWIR), seleccionadas por su relevancia en la discriminación de coberturas terrestres, incluyendo vegetación, agua y áreas urbanizadas. Adicionalmente, se utilizó la banda QA60 para enmascarar píxeles afectados por nubes y sombras, asegurando la calidad de los datos de entrada.

El período de análisis abarcó de 2023 a 2024, obteniendo un dataset de imágenes que componen escenas con baja nubosidad como se observa en la Figura 3 y obteniendo 11 imágenes que reflejan tanto patrones estacionales como cambios a largo plazo. La estandarización de los valores DN en esta colección garantiza comparabilidad temporal, clave para estudios multitemporales en áreas como Rionegro, caracterizadas por una alta variabilidad climática y espectral (Main-Knorn et al., 2017).

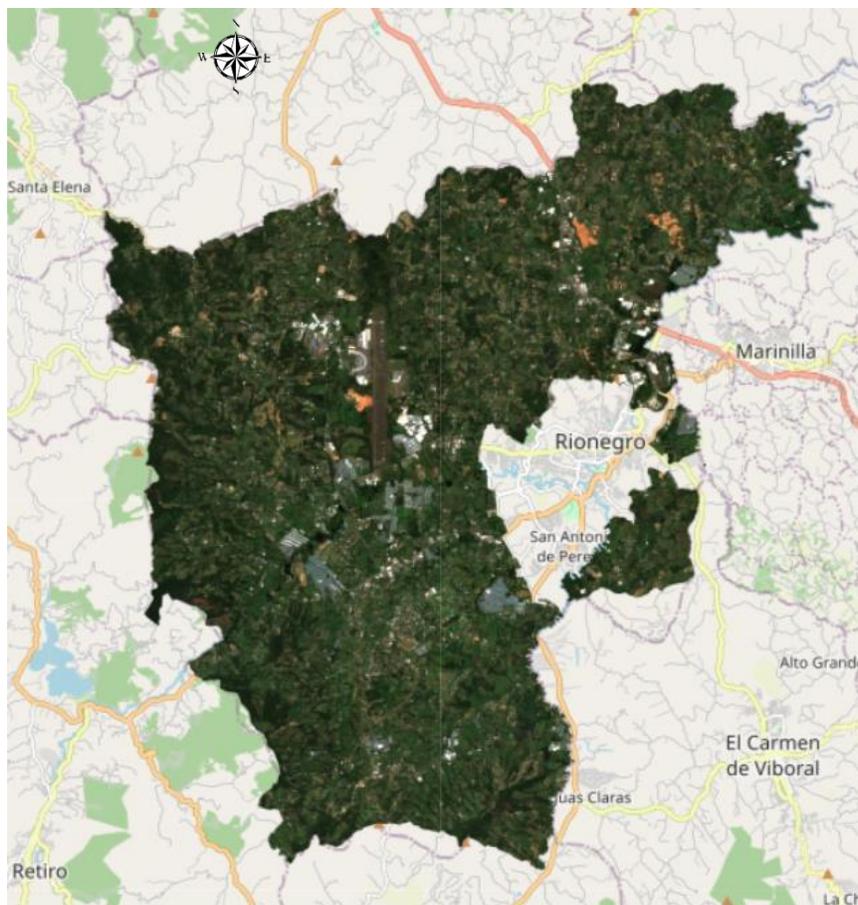


Figura 3. Dataset de Sentinel 2 obtenido.

5.3 Conexión a la API de GEE.

Para acceder a los datos satelitales de la colección Harmonized Sentinel-2 y realizar el análisis, se utilizó la API de Google Earth Engine (GEE) a través de Python. La conexión se estableció mediante el paquete `earthengine-api`, que permite interactuar con la plataforma GEE para la extracción y procesamiento de imágenes de forma programática.

La colección de imágenes Sentinel-2 se cargó utilizando el identificador `COPERNICUS/S2_SR_HARMONIZED`. La definición del área de interés (AOI) se realizó mediante un polígono que delimitó el municipio de Rionegro, asegurando que las imágenes seleccionadas correspondieran únicamente a esta región.

5.4 Selección de puntos de entrenamiento.

El proceso de selección de puntos de entrenamiento comenzó con la configuración del entorno computacional necesario para garantizar un análisis reproducible y eficiente. Se estableció una conexión con la API de Google Earth Engine (GEE), lo que permitió acceder a los datos satelitales necesarios para el análisis de cobertura terrestre en el municipio de Rionegro.

El proceso de selección y balanceo de puntos de entrenamiento es una etapa crítica para garantizar la calidad de los modelos de clasificación supervisada. En este análisis, se implementó un enfoque estructurado que combinó técnicas de filtrado y balanceo mediante Google Earth Engine (GEE) y algoritmos complementarios en Python. Este procedimiento aseguró que las clases de cobertura terrestre estuvieran adecuadamente representadas, reduciendo el riesgo de sesgos y mejorando la precisión de los modelos entrenados.

Se cargaron dos conjuntos de datos desde GEE: un conjunto de puntos de entrenamiento predefinidos, clasificados según el sistema CORINE Land Cover Colombia, y la delimitación geográfica del municipio de Rionegro. Los puntos de entrenamiento, almacenados como una colección de características (FeatureCollection), incluían atributos esenciales como la clase de cobertura terrestre identificada por el código CLC bajo la etiqueta “COD_CLC”. La verificación inicial de las propiedades de los puntos aseguró la integridad de los datos, facilitando su manipulación en etapas posteriores.

El análisis incluyó 33 clases de cobertura terrestre, representadas por códigos estandarizados basados en el sistema CORINE. Estas clases reflejan la diversidad espectral y espacial de Rionegro, abarcando áreas urbanas, vegetación natural, cuerpos de agua y suelos agrícolas. Sin embargo, la variabilidad en la cantidad de puntos disponibles por clase planteó un desafío importante: algunas clases estaban subrepresentadas o carecían completamente de puntos de entrenamiento.

El balanceo se implementó utilizando estrategias condicionales que ajustaron la cantidad de puntos de cada clase dentro de un rango predefinido, garantizando que todas las clases estuvieran representadas de manera adecuada. Las clases con un exceso de puntos fueron aleatoriamente submuestreadas hasta un máximo definido de 200 puntos, utilizando una columna aleatoria generada con una semilla constante para asegurar reproducibilidad. Por otro lado, se emitieron alertas para las clases con menos de 20 puntos disponibles, identificándolas como insuficientemente representadas. Estas estrategias buscan mitigar el problema del desbalance de clases, uno de los principales factores que afectan el desempeño de los modelos supervisados.

Los puntos balanceados se combinaron en una única colección mediante operaciones de flattening en GEE. Este proceso unificó las características seleccionadas de todas las clases, resultando en un conjunto de datos equilibrado y preparado para las etapas posteriores de preprocesamiento y modelado. La evaluación final del conjunto de datos balanceados confirmó la cantidad total de puntos disponibles, lo que permitió validar la efectividad del proceso de selección. Este procedimiento se detalla en la Figura 4.

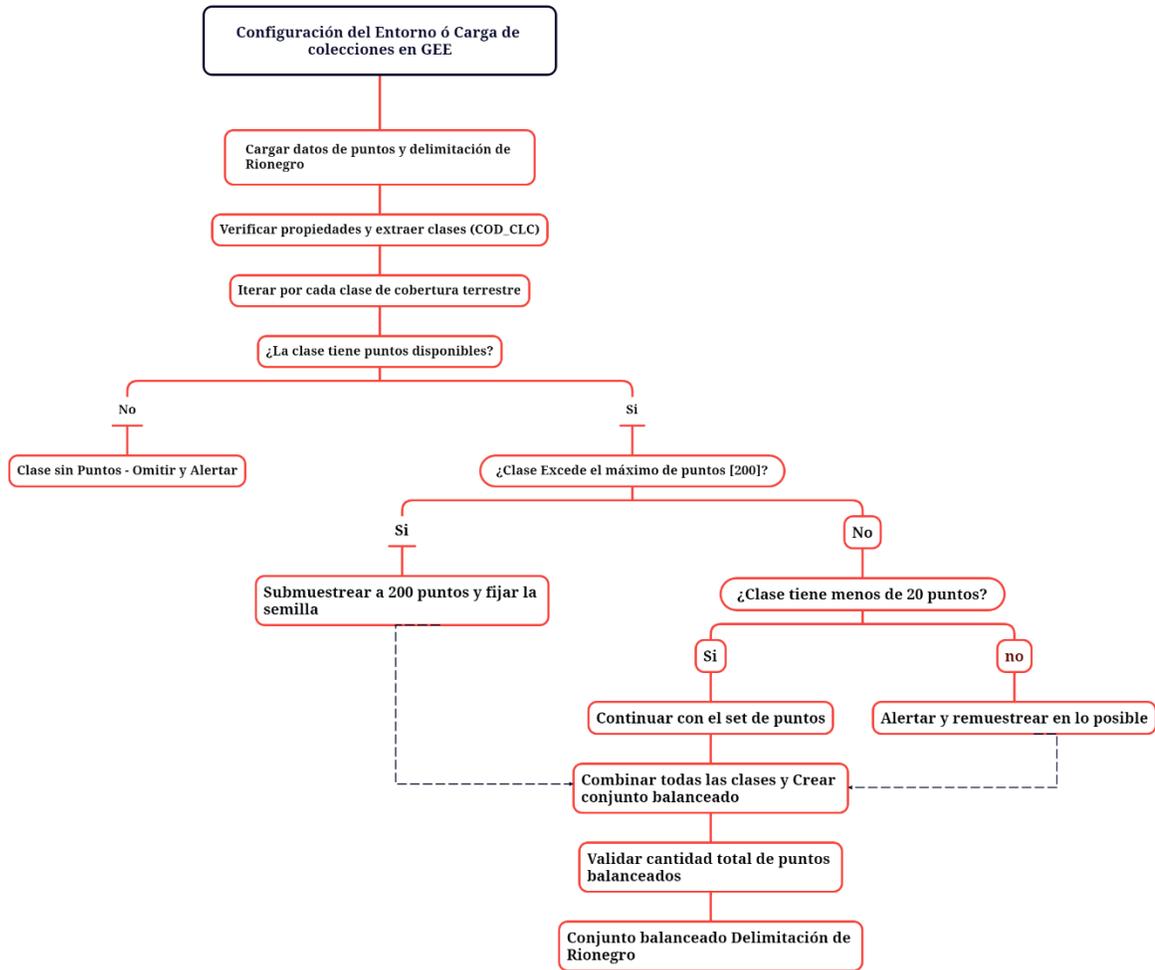


Figura 4. Flujograma de gestión y selección de puntos de entrenamiento.

Una vez balanceados los puntos de entrenamiento en Google Earth Engine (GEE), se transformaron en un formato tabular compatible con herramientas analíticas utilizando la biblioteca Pandas en Python. Este proceso permitió extraer atributos clave, como el código de cobertura terrestre (COD_CLC) y las coordenadas geográficas, desde la colección de GEE en formato JSON. El balanceo de clases se refinó utilizando SMOTE (Synthetic Minority Oversampling Technique), una técnica diseñada para mitigar el desbalance en los datos de entrenamiento. Este método generó ejemplos sintéticos para las clases minoritarias, mejorando su representación en el conjunto de datos y reduciendo el sesgo del modelo hacia las clases mayoritarias.

Para implementar SMOTE, se emplearon las coordenadas geográficas de los puntos de entrenamiento como variables predictoras (X) y los códigos de cobertura terrestre (y) como etiquetas. El número de vecinos (k) utilizados para generar los ejemplos sintéticos se ajustó dinámicamente en función del tamaño de la clase menos representada, garantizando que el método se adaptara a la distribución específica de los datos.

El balanceo generó un conjunto de datos resamplado ($X_{\text{resampled}}$, $y_{\text{resampled}}$) con una representación equitativa de las clases, aumentando la robustez del modelo durante el entrenamiento.

La transformación de las clases de cobertura terrestre a valores enteros consecutivos se realizó utilizando la herramienta LabelEncoder. Este proceso estructuró las etiquetas categóricas de forma numérica, asegurando su compatibilidad con los algoritmos de aprendizaje automático empleados en el análisis.

Cada clase fue asignada a un valor entero único, manteniendo la correspondencia exacta con los códigos originales de cobertura (COD_CLC). Este enfoque permite un tratamiento matemático consistente de las etiquetas durante las etapas de entrenamiento y validación del modelo. Además, se garantizó que el mapeo entre las clases originales y sus representaciones numéricas pudiera revertirse en caso de ser necesario, preservando la trazabilidad en el análisis.

El uso de LabelEncoder proporcionó una representación uniforme de las clases en el conjunto de datos balanceado, lo que optimiza el rendimiento de los algoritmos y facilita la interpretación de los resultados en las etapas posteriores del flujo metodológico.

La división de los datos de entrenamiento se realizó mediante validación cruzada estratificada con $k = 5$ pliegues. Esta técnica garantiza que cada partición del conjunto de datos preserve la proporción de las clases, lo que es particularmente importante en problemas de clasificación con distribución desigual entre las categorías.

El método StratifiedKFold fue empleado para dividir los datos balanceados en subconjuntos mutuamente excluyentes, optimizando la evaluación del modelo. En cada iteración, un pliegue se utilizó como conjunto de validación, mientras que los restantes sirvieron como datos de entrenamiento. Este enfoque proporciona estimaciones más robustas del rendimiento del modelo, ya que minimiza la variabilidad asociada con una única partición de los datos.

Además, el proceso incluyó el parámetro de aleatoriedad `random_state`, que asegura reproducibilidad en los resultados. La estructura generada por esta división permite que el modelo sea evaluado de manera consistente en diferentes subconjuntos, fortaleciendo las conclusiones derivadas del análisis.

El modelo XGBoost fue entrenado y evaluado utilizando validación cruzada estratificada para garantizar una evaluación consistente y representativa del desempeño en el análisis de clasificación supervisada. XGBoost es un algoritmo de ensamble basado en árboles de decisión que optimiza la precisión mediante la corrección iterativa de errores residuales.

En cada iteración de los pliegues generados con StratifiedKFold, se dividieron los datos en conjuntos de entrenamiento y prueba, manteniendo la proporción de las clases. Para cada pliegue, el modelo fue ajustado a los datos de entrenamiento (X_{train} , y_{train}) y evaluado en el conjunto de prueba (X_{test} , y_{test}). Este enfoque permitió obtener métricas detalladas, como la precisión (accuracy) y el reporte de clasificación por clase, asegurando un análisis exhaustivo de las predicciones.

Las etiquetas codificadas se decodificaron a sus valores originales antes de la evaluación para garantizar que las métricas reflejaran la interpretación directa de las clases de cobertura terrestre. Este paso también facilitó la generación de reportes de clasificación que resumieron las métricas clave como la precisión, la sensibilidad y la especificidad para cada clase.

El promedio de precisión (accuracy) obtenido a lo largo de los pliegues ofreció una medida robusta del rendimiento general del modelo. Este procedimiento garantiza que las conclusiones derivadas del análisis sean reproducibles y estén respaldadas por una evaluación rigurosa en diferentes particiones de los datos.

La división geoespacial de los datos balanceados se realizó para separar los conjuntos de entrenamiento y validación, asegurando que ambos reflejen la distribución espacial y espectral de las clases de cobertura terrestre. Este procedimiento se llevó a cabo con una proporción de 70% para entrenamiento y 30% para validación, utilizando un método de aleatorización reproducible gracias a una semilla fija.

Los datos (X para coordenadas y y para etiquetas) se permutaron aleatoriamente para garantizar una distribución imparcial en los conjuntos de entrenamiento y validación. La separación basada en índices permitió que cada conjunto retuviera las propiedades originales del espacio geográfico, minimizando posibles sesgos en la ubicación de los puntos. Esta estrategia mejora la robustez del modelo, asegurando que las pruebas se realicen en datos que el modelo no ha visto anteriormente.

- **Puntos de entrenamiento:** Visualizados en color azul, indicando las áreas utilizadas para ajustar el modelo.
- **Puntos de validación:** Visualizados en color verde, representando las áreas reservadas para medir el desempeño del modelo.

El área de interés, correspondiente a la delimitación del municipio de Rionegro, se utilizó como referencia geográfica en el mapa interactivo. Esta representación permitió verificar visualmente que ambas particiones mantuvieran una cobertura adecuada del espacio de estudio y no estuvieran concentradas en zonas específicas, evitando posibles sesgos espaciales. Este procedimiento general se detalla en la Figura 5.



Figura 5. Proceso de clasificación de puntos de entrenamiento.

5.5 Procesamiento de imágenes satelitales.

El análisis se restringió al municipio de Rionegro, definido como el Área de Interés (AOI) en el sistema de referencia espacial de Google Earth Engine. Se emplearon imágenes adquiridas entre el 1 de enero de 2023 y el 17 de noviembre de 2024, filtrando escenas con un porcentaje de nubes inferior al 20%. Las bandas seleccionadas fueron B2 (azul), B3 (verde), B4 (rojo), B8 (NIR), B11 (SWIR-1) y B12 (SWIR-2), las cuales son fundamentales para discriminar diferentes clases de cobertura terrestre.

Para mejorar la calidad de las imágenes satelitales, se implementó un procedimiento de enmascaramiento basado en la banda QA60 de Sentinel-2, que identifica píxeles afectados por nubes y sombras. Este proceso elimina datos contaminados mediante operaciones bit a bit, lo que garantiza que solo se analicen píxeles despejados.

- Cloud Mask: Detecta y descarta píxeles afectados por nubes.
- Shadow Mask: Filtra sombras asociadas con las nubes, utilizando propiedades de la misma banda QA60.

La colección COPERNICUS/S2_SR_HARMONIZED fue filtrada espacial y temporalmente para incluir únicamente imágenes dentro del AOI y del rango de fechas definido. Además, se aplicó el enmascaramiento de nubes y sombras a cada escena de la colección. Posteriormente, se generó un compuesto de mediana para representar un resumen estadístico robusto de las imágenes, minimizando el impacto de valores atípicos y variaciones temporales.

En esta etapa del procesamiento de imágenes satelitales, se generó un FeatureCollection en Google Earth Engine (GEE) a partir de los puntos de entrenamiento previamente balanceados y divididos. Este proceso transformó las coordenadas geográficas (X) y las etiquetas asociadas a las clases de cobertura terrestre (y) en una estructura compatible con GEE, facilitando su integración en etapas posteriores como la clasificación supervisada.

Cada punto fue representado como un objeto ee.Feature, donde:

- Geometría: La coordenada (x,y) de cada punto se transformó en un objeto geométrico de tipo Point en GEE.
- Propiedades: Se incluyó un atributo denominado COD_CLC, que contiene la etiqueta codificada correspondiente a la clase de cobertura terrestre.

Se utilizó el compuesto de mediana de Sentinel-2, generado previamente, para extraer los valores espectrales de las bandas seleccionadas ($B2$, $B3$, $B4$, $B8$, $B11$, y $B12$). Los valores espectrales correspondientes a cada banda se asociaron con los puntos de entrenamiento definidos en el FeatureCollection, respetando sus coordenadas geográficas y etiquetas de clase (COD_CLC). Este muestreo se realizó a una escala de 10 metros, que coincide con la resolución de las bandas de Sentinel-2 utilizadas en el análisis.

Los valores espectrales extraídos fueron convertidos a un DataFrame de Pandas, facilitando su análisis y visualización. Cada fila del DataFrame representa un punto de entrenamiento con los siguientes atributos:

- COD_CLC: La etiqueta de la clase de cobertura terrestre.
- Bandas espectrales: Los valores espectrales extraídos para cada una de las bandas seleccionadas.

Se evaluó la distribución espectral de cada banda mediante histogramas, analizando cómo se superponen o separan las clases para cada banda. Estos gráficos revelan:

- Diferenciación clara: Bandas donde las distribuciones espectrales de las clases están bien separadas, indicando su utilidad para la clasificación.
- Solapamiento significativo: Bandas donde las distribuciones espectrales de diferentes clases se superponen, lo que puede generar ambigüedad en la clasificación.

Por ejemplo, las bandas en el infrarrojo cercano (*B8*) y el infrarrojo de onda corta (*B11*, *B12*) suelen mostrar una buena diferenciación para clases relacionadas con vegetación, agua y suelo desnudo, debido a su sensibilidad a la biomasa y al contenido de humedad.

La estandarización de bandas espectrales es un paso fundamental para garantizar que los valores de reflectancia estén normalizados dentro de un rango uniforme, en este caso, [0,1]. Esto facilita la comparación entre bandas y mejora el rendimiento de los algoritmos de clasificación, al eliminar disparidades asociadas con las escalas de valores originales.

La función implementada utiliza el método `unitScale` de Google Earth Engine (GEE), que transforma cada banda en función de su valor mínimo y máximo. Los valores extremos (*min* y *max*) se calcularon utilizando el método `reduceRegion` aplicado al área de interés (AOI), que:

- Obtiene los valores extremos de cada banda seleccionada.
- Escala los valores espectrales: Los valores originales son ajustados para quedar dentro del rango [0,1].

La armonización de resoluciones espaciales es un paso crucial en el procesamiento de imágenes satelitales, especialmente al trabajar con bandas espectrales de diferentes resoluciones. En este caso, las bandas *B11* (SWIR-1) y *B12* (SWIR-2), que originalmente tienen una resolución de 20 metros, fueron ajustadas para coincidir con las bandas de mayor resolución espacial (10 metros) como *B2* (azul), *B3* (verde), *B4* (rojo) y *B8* (NIR). Las bandas resultantes fueron integradas en un nuevo conjunto de datos utilizando `addBands`, combinando todas las bandas armonizadas en una sola imagen compuesta.

- **Interpolación Bilineal:** Se utilizó la técnica de resampling bilineal para ajustar los valores de las bandas de baja resolución a la escala deseada, asegurando una transición suave entre píxeles adyacentes.
- **Reproyección:** Las bandas reescaladas se re proyectaron al sistema de referencia espacial y escala de las bandas de 10 metros, utilizando el método `reproject` con los parámetros de proyección derivados de las bandas de alta resolución.

El compuesto final incluye índices espectrales derivados, como NDVI, NDWI, y NDBI, los cuales enriquecen el análisis al aportar información adicional sobre las características de la superficie terrestre. Estos índices son esenciales para discriminar diferentes tipos de cobertura terrestre, como vegetación, cuerpos de agua y áreas urbanas.

Los índices se calcularon utilizando combinaciones de bandas del espectro reflejado por Sentinel-2, aplicando el método `normalizedDifference` de Google Earth Engine (GEE). Cada índice se añadió como una nueva banda al compuesto, generando una representación multibanda más informativa.

NDVI (Índice de Vegetación Normalizado):

$$NDVI = \frac{B8 - B4}{B8 + B4}$$

- Utiliza el NIR (B8) y el Rojo (B4).
- Diferencia áreas con vegetación activa (valores altos) de suelo desnudo o agua (valores bajos).

NDWI (Índice de Agua Normalizado):

$$NDWI = \frac{B3 - B8}{B3 + B8}$$

- Utiliza el Verde (B3) y el NIR (B8).
- Identifica la presencia de agua líquida en la superficie, con valores positivos indicando agua y negativos señalando suelo seco o vegetación.

NDBI (Índice de Desarrollo Urbano Normalizado):

$$NDWI = \frac{B11 - B8}{B11 + B8}$$

- Utiliza el SWIR-1 (B11) y el NIR (B8).
- Resalta áreas urbanas debido a su alta reflectancia en SWIR y baja en NIR.

5.6 Calculo de métodos de clasificación por GEE

Google Earth Engine (GEE) se utiliza como una plataforma centralizada para implementar y ejecutar métodos de clasificación supervisada debido a su capacidad para procesar grandes volúmenes de datos geoespaciales de manera eficiente. En este estudio, los métodos de clasificación se aplicaron directamente sobre un compuesto multiespectral generado a partir de imágenes Sentinel-2, previamente procesadas y armonizadas en términos de resolución espacial y cobertura temporal. La clasificación supervisada en GEE permite integrar algoritmos avanzados, como K-Nearest Neighbors (KNN), Naive Bayes, Random Forest (RF), Classification and Regression Trees (CART) y Gradient Tree Boost (GTB), mediante el uso de funciones nativas optimizadas para grandes colecciones de datos.

El flujo metodológico en GEE incluye la selección de puntos de entrenamiento georreferenciados y la extracción de valores espectrales de bandas clave para cada clase de cobertura terrestre. Estos datos se utilizan para entrenar los clasificadores mediante funciones específicas, como `smileRandomForest` o `smileKNN`, las cuales operan directamente sobre colecciones de datos en el entorno en la nube de GEE. El modelo entrenado se aplica posteriormente al área de interés, generando mapas clasificados que representan las distintas clases de cobertura terrestre.

GEE facilita la evaluación iterativa de los métodos mediante el cálculo de métricas de validación cruzada, realizadas sobre subconjuntos de datos, y permite ajustar parámetros clave como el número de árboles en Random Forest o la cantidad de vecinos en KNN. En la Tabla 1. Se exponen de manera general los métodos implementados en el estudio.

Tabla 1. Consolidado de resultados por método de clasificación.

Método	Descripción corta	Configuraciones específicas	Escenario de mejor desempeño	Fundamento matemático
K-Nearest Neighbors (KNN)	Algoritmo basado en la proximidad en el espacio espectral. Asigna la clase en función de los k vecinos más cercanos en el espacio de características.	Número de vecinos (k=5). Utiliza las bandas B2,B3,B4,B8,B11,B12 a una escala de 10 m.	Clases bien separadas espectralmente y datos con un balance adecuado entre clases.	$\hat{y}_i = \operatorname{argmax}_c \sum_{j \in \mathcal{N}_k(i)} \delta(y_j = c)$ <p>Donde:</p> <ul style="list-style-type: none"> $\mathcal{N}_k(i)$: Conjunto de los k vecinos más cercanos al punto i. $\delta(y_j = c)$: Función indicadora que es igual a 1 si $y_j = c$ (la clase del vecino j es c), y 0 en caso contrario.
Naive Bayes	Clasificador probabilístico que asume independencia condicional entre las variables espectrales para calcular las probabilidades de cada clase.	Entrenado con las bandas B2,B3,B4,B8,B11,B12. Calcula probabilidades a partir de datos espectrales muestreados a escala de 10 m.	Clases con distribuciones gaussianas diferenciadas y baja correlación entre bandas espectrales.	$\hat{y}_i = \operatorname{argmax}_c P(c) \prod_{j=1}^n P(x_j c)$ <p>Donde:</p> <ul style="list-style-type: none"> $P(c)$: Probabilidad a priori de la clase c. $P(x_j c)$: Probabilidad condicional de la característica x_j dado que pertenece a la clase c.
Random Forest (RF)	Algoritmo de ensamble que utiliza múltiples árboles de decisión. Combina los resultados para optimizar la clasificación y reducir el sobreajuste.	Número de árboles (50), fracción de datos seleccionados por árbol (bagFraction=0.7). Considera un mínimo de 5 datos por hoja.	Problemas con alta dimensionalidad espectral y datos heterogéneos.	$\hat{y}_i = \operatorname{argmax}_c \frac{1}{T} \sum_{t=1}^T h_t(x_i)$ <p>Donde:</p> <ul style="list-style-type: none"> $h_t(x_i)$: Clasificador generado por el t-ésimo árbol. T: Número total de árboles en el bosque.
Classification and Regression Trees (CART)	Modelo jerárquico basado en divisiones binarias de los datos espectrales para maximizar la separación de clases en cada nodo.	Máximo de nodos (200), mínimo de 10 muestras por nodo. Utiliza bandas seleccionadas en Sentinel-2.	Clases bien segmentadas y relaciones claras entre bandas y etiquetas.	$\text{Impureza}(t) = \sum_{c=1}^C -p_c(t) \log p_c(t)$ <p>Donde:</p> <ul style="list-style-type: none"> $p_c(t)$: Proporción de datos pertenecientes a la clase c en el nodo t. C: Número total de clases.
Gradient Tree Boost (GTB)	Clasificador iterativo basado en árboles de decisión. Ajusta errores residuales en cada iteración para optimizar el rendimiento.	Número de árboles (10). Entrenado con bandas B2,B3,B4,B8,B11,B12 y datos espectrales muestreados a 10 m.	Escenarios con relaciones no lineales complejas y con necesidad de mejorar iterativamente la clasificación.	$F_m(x) = F_{m-1}(x) + \gamma_m \cdot h_m(x)$ <p>Donde:</p> <ul style="list-style-type: none"> $F_m(x)$: Predicción ajustada después de m iteraciones. $h_m(x)$: Árbol de decisión que minimiza la función de pérdida L en la iteración m. γ_m: Coeficiente de aprendizaje que regula la contribución de $h_m(x)$.

5.6.1 Método de clasificación CART

La clasificación supervisada empleando el método CART (Classification and Regression Trees) se implementó para analizar la cobertura terrestre en el área de estudio. Este enfoque se fundamenta en la construcción de árboles de decisión, en los cuales las divisiones sucesivas del conjunto de datos se realizan basándose en variables espectrales, permitiendo discriminar las clases de interés.

El proceso comenzó con la preparación de los datos de entrenamiento mediante el método `sampleRegions`, que extrajo los valores espectrales de las bandas seleccionadas para cada punto de entrenamiento georreferenciado. Las bandas utilizadas en este análisis fueron *B2*, *B3*, *B4*, *B8*, *B11* y *B12*, todas con una resolución espacial de 10 metros a excepción de *B11* y *B12* las cuales cuentan con resolución espacial de 20 metros. Los valores espectrales de cada píxel dentro del área de interés se asociaron con sus respectivas etiquetas de clase, facilitando el entrenamiento del modelo.

El clasificador CART fue configurado con un máximo de 200 nodos terminales y una población mínima de 10 muestras por nodo. Estas configuraciones permitieron ajustar el balance entre la complejidad del modelo y la generalización, asegurando que las divisiones en el árbol fueran significativas y robustas. Durante el entrenamiento, el modelo identificó los valores umbrales de las bandas que optimizaban la separación entre las clases, construyendo un árbol de decisión jerárquico.

Tras el entrenamiento, el modelo se aplicó a la imagen compuesta mediante el método `classify`, generando una capa clasificada denominada `Clasificacion_cart`. Cada píxel de la imagen fue asignado a una clase específica de cobertura terrestre en función de las reglas definidas en el árbol de decisión. Esta capa representa espacialmente las clases de cobertura terrestre dentro del área de estudio.

Para la visualización de los resultados, se empleó la biblioteca `geemap`, que permitió integrar el mapa clasificado con la delimitación del área de interés. Los parámetros de visualización incluyeron un rango de clases ajustado dinámicamente en función de las categorías detectadas y una paleta cromática que facilitó la interpretación visual. El mapa interactivo proporcionó un entorno para la inspección preliminar de los resultados, asegurando que la clasificación se distribuyera adecuadamente en la región de estudio.

5.6.2 Método de clasificación Random Forest

La implementación de la clasificación supervisada con el algoritmo Random Forest (RF) en Google Earth Engine se desarrolló para evaluar la distribución de las clases de cobertura terrestre en el área de interés. Random Forest es un modelo de ensamble basado en árboles de decisión, conocido por su capacidad para manejar datos complejos y su robustez frente a problemas de sobreajuste.

El proceso metodológico comenzó con la preparación de datos de entrenamiento. A partir de la imagen compuesta preprocesada, se extrajeron valores espectrales correspondientes a las bandas *B2*, *B3*, *B4*, *B8*, *B11* y *B12*, seleccionadas por su relevancia en la discriminación espectral. Estos valores fueron vinculados a las etiquetas de clase (`COD_CLC`) de los puntos de entrenamiento georreferenciados. El muestreo se realizó a una escala espacial de 10 metros, consistente con la resolución de las bandas seleccionadas.

El clasificador Random Forest fue configurado con un número de 50 árboles de decisión. La configuración adicional incluyó una población mínima por hoja de 5 elementos (*minLeafPopulation* = 5) y una fracción de datos de entrenamiento seleccionados aleatoriamente para construir cada árbol (*bagFraction*=0.7). Estas configuraciones garantizan una exploración amplia del espacio de características espectrales mientras se minimizan los sesgos introducidos por el sobreajuste.

El modelo fue entrenado con los valores espectrales muestreados y posteriormente aplicado a la imagen compuesta. Cada píxel de la imagen fue clasificado según las reglas aprendidas por el modelo, generando un mapa temático denominado *Clasificacion_rf*. Este resultado representa la distribución espacial de las clases de cobertura terrestre en el área de estudio.

La visualización de los resultados se realizó utilizando la biblioteca *geemap*. Se ajustaron dinámicamente los parámetros de visualización en función de las clases presentes en la imagen clasificada, garantizando una representación adecuada y diferenciada de las categorías. La paleta de colores seleccionada permitió identificar visualmente las distintas clases, mientras que la delimitación del área de interés proporcionó un contexto geográfico claro.

5.6.3 Metodo de clasificación KNN

El algoritmo K-Nearest Neighbors (KNN) implementado en Google Earth Engine permitió el flujo de entrenamiento del clasificador a partir de una imagen compuesta procesada, que incluía las bandas espectrales *B2*, *B3*, *B4*, *B8*, *B11* y *B12*, seleccionadas por su relevancia en la discriminación espectral de clases. Estas bandas se vincularon a los datos de entrenamiento, representados como puntos georreferenciados con etiquetas correspondientes a las categorías de cobertura terrestre.

El entrenamiento del clasificador se llevó a cabo utilizando el método *sampleRegions*, que permitió extraer los valores espectrales de las bandas para los puntos de entrenamiento definidos en la región de interés. La escala espacial utilizada fue de 10 metros, consistente con la resolución de las bandas seleccionadas. Posteriormente, el clasificador KNN fue configurado con un parámetro de $k=5$, que define la cantidad de vecinos más cercanos considerados para asignar la clase a cada píxel. La relación entre las bandas seleccionadas y las etiquetas de clase se estableció durante este proceso, lo que permitió que el modelo capturara las características espectrales representativas de cada categoría.

Una vez entrenado, el modelo se aplicó a la imagen compuesta mediante el método *classify*, generando un mapa temático que asigna una clase a cada píxel dentro del área de estudio. Este mapa, denominado *Clasificacion_knn*, representa espacialmente las distintas categorías de cobertura terrestre según las relaciones espectrales definidas en el modelo.

La visualización de los resultados se realizó utilizando la biblioteca *geemap*, integrando el área de interés como referencia geográfica. Los parámetros de visualización incluyeron un rango de clases definido entre 0 y 34, correspondiente al número total de categorías presentes en los datos, y una paleta de colores personalizada que facilitó la diferenciación visual de las clases en el mapa generado. El mapa interactivo permitió verificar la consistencia de la clasificación y su correspondencia con el área delimitada para el análisis.

5.6.4 Método de clasificación Naive Bayes

La clasificación supervisada utilizando el método Naive Bayes se implementó como una etapa fundamental en el análisis de cobertura terrestre para el área de estudio definida. Este método probabilístico, basado en el teorema de Bayes, permite la clasificación al asumir independencia entre las variables predictoras, en este caso, las bandas espectrales seleccionadas del compuesto de Sentinel-2.

El proceso comenzó con la extracción de valores espectrales correspondientes a los puntos de entrenamiento mediante el método `sampleRegions`. Este procedimiento asoció los valores de las bandas espectrales *B2*, *B3*, *B4*, *B8*, *B11* y *B12* con las etiquetas de clase definidas como `COD_CLC`. La escala utilizada para el muestreo fue de 10 metros, consistente con la resolución de las bandas seleccionadas. Antes de continuar con el entrenamiento, se verificó que el conjunto de datos de entrenamiento generado fuera adecuado, tanto en términos de tamaño como de calidad, para evitar inconsistencias en las etapas posteriores.

Posteriormente, el modelo de clasificación Naive Bayes fue entrenado utilizando el clasificador `smileNaiveBayes` disponible en Google Earth Engine. Este método permitió establecer relaciones probabilísticas entre las bandas seleccionadas y las clases de cobertura terrestre, basándose en la distribución de los valores espectrales dentro del conjunto de entrenamiento. La propiedad `COD_CLC` se utilizó como variable dependiente, mientras que las bandas espectrales sirvieron como variables predictoras.

La imagen compuesta fue clasificada aplicando el modelo entrenado. El método `classify` asignó probabilidades a cada clase basada en los datos espectrales del área de estudio, generando una nueva capa denominada Clasificación Naive Bayes, que representa la distribución espacial de las clases de cobertura terrestre dentro del área de interés.

Finalmente, el resultado de la clasificación fue visualizado en un mapa interactivo. Para ello, se utilizaron parámetros de visualización que incluyen una paleta de colores específica para diferenciar las clases de manera clara y un rango definido para abarcar todas las categorías identificadas en el conjunto de datos. El mapa también incluyó la delimitación del área de estudio para proporcionar contexto geográfico al análisis. La visualización permitió validar de manera preliminar la consistencia espacial de la clasificación, asegurando que las clases estén distribuidas de acuerdo con las características esperadas del área de estudio.

5.6.5 Método de clasificación Gradient Tree Boost

La clasificación supervisada con el algoritmo Gradient Tree Boost se implementó para identificar las coberturas terrestres en el área de interés utilizando Google Earth Engine. Este método emplea un enfoque iterativo basado en árboles de decisión, optimizando la precisión del modelo mediante la corrección de errores residuales en cada iteración.

El proceso comenzó con la extracción de datos de entrenamiento a partir de los puntos georreferenciados previamente definidos y el compuesto de mediana generado a partir de la colección de imágenes Sentinel-2. Para garantizar la coherencia espacial, se verificó que los puntos

de entrenamiento y la imagen compuesta compartieran el mismo sistema de referencia espacial (CRS). En caso de discrepancias, los puntos de entrenamiento se re proyectaron dinámicamente para asegurar la compatibilidad con el CRS de la imagen. Los valores espectrales correspondientes a las bandas B_2 , B_3 , B_4 , B_8 , B_{11} y B_{12} fueron extraídos para cada punto de entrenamiento, vinculándolos con las etiquetas de clase COD_CLC.

El modelo Gradient Tree Boost fue configurado con un total de 10 árboles. Este número balancea la capacidad del modelo para capturar relaciones no lineales en los datos y su capacidad para generalizar a nuevas observaciones. Las características espectrales extraídas sirvieron como variables predictoras, mientras que las etiquetas COD_CLC representaron las clases objetivo. El modelo fue entrenado utilizando los datos espectrales y clasificado como un modelo iterativo que ajusta los errores en cada paso, generando predicciones robustas.

La clasificación se aplicó al compuesto de mediana utilizando el clasificador entrenado. Cada píxel de la imagen fue etiquetado con la clase de cobertura terrestre correspondiente, produciendo una capa temática denominada Clasificación Gradient Tree Boost, que representa espacialmente las categorías de cobertura en el área de estudio.

Finalmente, los resultados fueron visualizados en un mapa interactivo con la biblioteca geemap. Los parámetros de visualización incluyeron una paleta de colores específica para distinguir visualmente las clases presentes en la clasificación y un rango dinámico que abarcó todas las categorías detectadas en los datos de entrenamiento. El área de interés se añadió como capa adicional para proporcionar contexto geográfico, permitiendo validar visualmente la distribución de las clases clasificadas.

5.7 Métodos de validación

La validación de los métodos de clasificación supervisada constituye una etapa fundamental en cualquier análisis de teledetección, ya que garantiza la confiabilidad de los resultados y la capacidad del modelo para generalizar más allá de los datos de entrenamiento. En este estudio, se implementaron métodos de validación rigurosos que combinan enfoques estadísticos y espaciales, con métricas diseñadas para evaluar tanto el desempeño global del modelo como su precisión a nivel de clase. Estas técnicas aseguran un análisis robusto y reproducible, alineado con los estándares científicos en estudios de clasificación supervisada.

El proceso inicia con la implementación de la validación cruzada estratificada, un método que divide los datos de entrenamiento en k subconjuntos o pliegues, preservando la distribución original de las clases en cada partición. En cada iteración, uno de los pliegues se utiliza como conjunto de validación, mientras los restantes sirven para entrenar el modelo. Este enfoque reduce el riesgo de sobreajuste y proporciona estimaciones confiables del rendimiento del modelo al evaluar su capacidad en datos no vistos. La validación cruzada es particularmente relevante en escenarios de teledetección con clases desbalanceadas, ya que asegura que todas las clases estén representadas en cada partición.

La matriz de confusión se empleó como herramienta central para evaluar la clasificación, permitiendo identificar patrones de error entre clases. Este instrumento cuantifica la concordancia entre las etiquetas verdaderas y las predichas, proporcionando información detallada sobre los

errores de comisión (falsos positivos) y omisión (falsos negativos). A partir de esta matriz, se calcularon métricas clave como:

- **Precisión global (accuracy):** Proporción de observaciones correctamente clasificadas.
- **Precisión y sensibilidad por clase:** Indicadores del rendimiento específico para cada categoría, útiles para identificar posibles debilidades en la clasificación.
- **F1-score:** Métrica combinada que balancea precisión y sensibilidad, especialmente relevante en escenarios con clases desbalanceadas.

Para complementar el análisis, se calculó el índice Kappa de Cohen, una métrica robusta que ajusta la precisión observada considerando la probabilidad de concordancia por azar. Este índice es esencial para evaluar la calidad general del modelo, ya que detecta casos donde un rendimiento aparentemente alto podría deberse a la prevalencia de ciertas clases.

Además, se integraron análisis basados en la Curva ROC (Receiver Operating Characteristic) y el Área Bajo la Curva (AUC). Estas herramientas evalúan la capacidad del modelo para discriminar entre clases en diferentes umbrales de decisión, proporcionando una métrica escalable e independiente de la distribución de clases. La curva ROC es especialmente útil para identificar problemas de separación espectral entre clases, mientras que el AUC resume esta capacidad en un solo valor, siendo un indicador clave de la eficacia del modelo.

Por último, se calcularon métricas agregadas bajo enfoques macro y micro, proporcionando perspectivas complementarias. El promedio macro considera las métricas de cada clase por igual, mientras que el promedio micro pondera estas métricas según la frecuencia de las clases. Este enfoque dual permite un análisis más equilibrado del desempeño global del modelo y de su capacidad para manejar clases desbalanceadas.

6 Resultados

Los resultados obtenidos tras la aplicación de los métodos de clasificación supervisada al análisis de las coberturas terrestres en el municipio de Rionegro revelan resultados significativos sobre la eficacia de los algoritmos empleados. Mediante el procesamiento detallado de las imágenes satelitales Sentinel-2 y la implementación de los algoritmos Random Forest, K-Nearest Neighbors, Naive Bayes y Gradient Boosting, se logró una discriminación precisa de las distintas clases de cobertura del suelo presentes en el área de estudio. Las métricas de evaluación, como la matriz de confusión y el índice Kappa, proporcionaron una base sólida para comparar el desempeño de cada método, destacando sus fortalezas y limitaciones en el contexto geoespacial específico de Rionegro. A continuación, se presentan en detalle los hallazgos más relevantes, interpretando los patrones espaciales y espectrales identificados y discutiendo sus implicaciones para el monitoreo ambiental y la gestión territorial en regiones con dinámicas espaciales complejas.

6.1 Resultados Provenientes Del Dataset De Corine Land Cover Para Nivel 3

En esta sección se presentan los resultados obtenidos a partir de la aplicación de diversos métodos de clasificación supervisada sobre el dataset CORINE Land Cover a Nivel 3, utilizando imágenes Sentinel-2. El objetivo principal de este análisis es evaluar el desempeño de modelos como CART, Naive Bayes, KNN y Random Forest en la identificación de categorías detalladas de cobertura terrestre, considerando la complejidad espectral y espacial del área de estudio en Rionegro, Antioquia. Cada método fue evaluado a través de métricas de precisión, recall y F1-score, complementadas con análisis de las curvas ROC y AUC para medir la capacidad discriminativa de los modelos. Además, se incluyen mapas temáticos resultantes que permiten analizar la coherencia espacial de las clasificaciones generadas, destacando las fortalezas y limitaciones de cada enfoque. Este capítulo busca no solo mostrar los resultados obtenidos, sino también justificar su relevancia y utilidad en el contexto de aplicaciones prácticas como el ordenamiento territorial y la gestión ambiental.

Para el método de Gradient Tree Boost no se obtuvo matriz de clasificación debido a que la similitud de etiquetas del dataset de entrenamiento en conjunto con la alta fragmentación de la zona de estudio hicieron inviable para esta categoría el entrenamiento por este método.

6.1.1 Método de CART

El modelo para el área de estudio revela importantes limitaciones y aspectos destacados. En términos generales, la precisión global del modelo, establecida en un 21%, es indicativa de una baja capacidad para discriminar entre las distintas clases de cobertura terrestre en un entorno tan heterogéneo como Rionegro. Este bajo desempeño puede atribuirse a la complejidad inherente del área de estudio, caracterizada por una alta fragmentación y solapamiento espectral entre clases. Dichos factores generan ambigüedades que dificultan la correcta asignación de clases, especialmente cuando las bandas espectrales seleccionadas no ofrecen suficiente información distintiva para separar categorías similares, el resultado de la clasificación se puede evidenciar en la Figura 6.

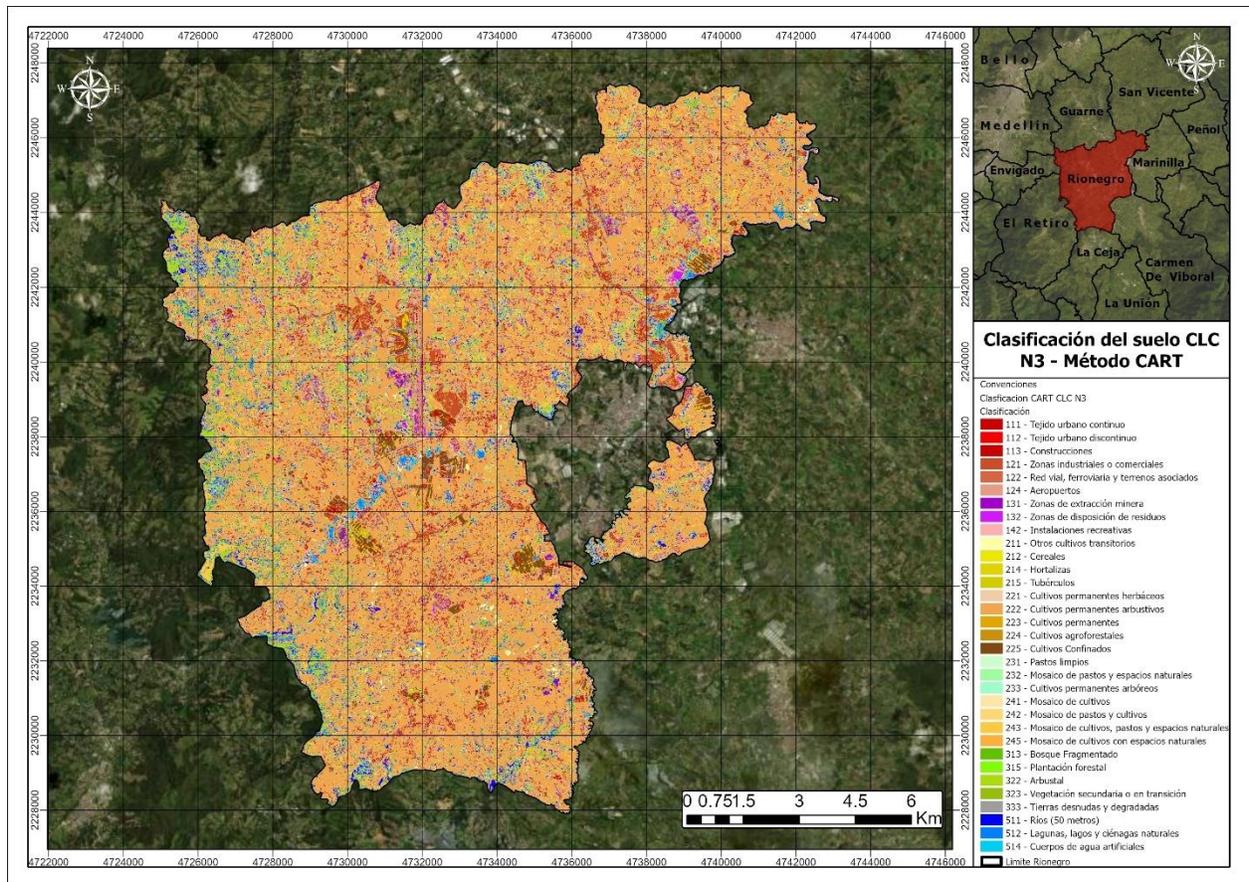


Figura 6. Resultado clasificación método CART.

La matriz de confusión que se puede visualizar en la Figura 7 es una herramienta clave que permite analizar los errores del modelo con mayor profundidad. En este caso, se evidencia que ciertas clases, como el tejido urbano discontinuo, destacan por su alto número de aciertos, lo cual puede explicarse por su firma espectral claramente diferenciable. Estas áreas, al tener una composición espectral que contrasta significativamente con las coberturas circundantes, como vegetación o áreas urbanas densas, son más fácilmente reconocibles por el algoritmo. Sin embargo, otras clases, como los pastos limpios o las coberturas vegetales mixtas, presentan una gran confusión entre sí. Este fenómeno refleja un problema estructural en el modelo, dado que estas categorías comparten características espectrales similares que el algoritmo no es capaz de desambiguar eficientemente. Este tipo de errores no solo compromete la precisión, sino también la utilidad práctica del modelo en aplicaciones de monitoreo y gestión territorial.

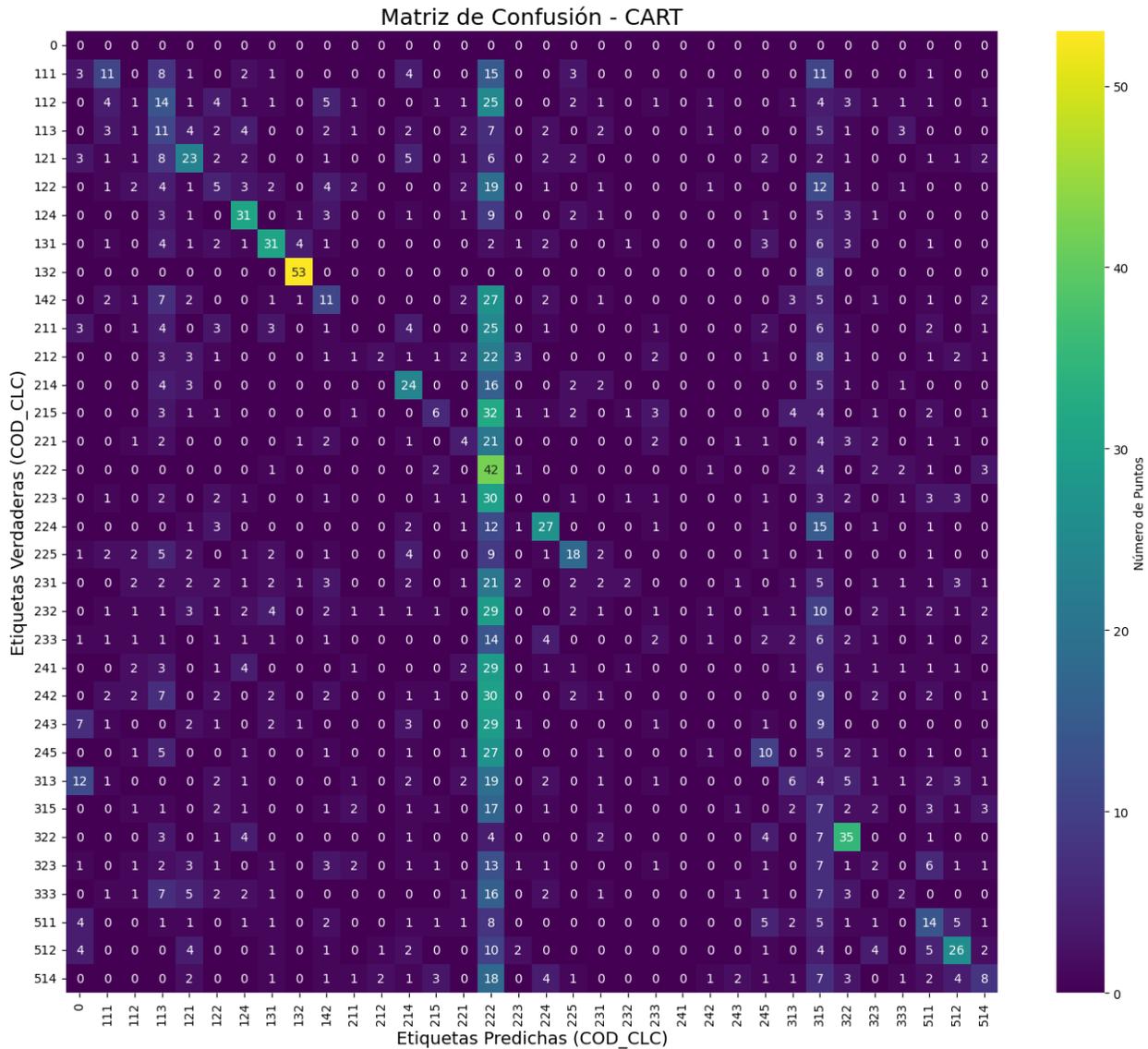


Figura 7. Matriz de Confusión Método CART.

Otro punto crítico es la clara subrepresentación de ciertas clases en el conjunto de entrenamiento. El desequilibrio en la cantidad de muestras por clase introduce un sesgo significativo en el modelo, favoreciendo a las categorías más representadas mientras que ignora o clasifica de manera aleatoria a las minoritarias. Esta falta de equidad en los datos de entrada también afecta la interpretación de las métricas globales, que pueden ser infladas por un desempeño adecuado en clases dominantes mientras fallan en reflejar los errores sistemáticos en categorías menos frecuentes. Además, el índice Kappa, con un valor de 0.19, confirma que el modelo apenas logra un desempeño ligeramente mejor que el azar, resaltando la necesidad de un ajuste considerable en el enfoque metodológico.

Por otro lado, las curvas ROC y sus respectivos valores de AUC visualizado en la Figura 8 ofrecen una perspectiva interesante sobre la capacidad del modelo para discriminar entre clases. Si bien algunas categorías, como los aeropuertos, logran valores elevados de AUC que denotan una buena

separación, la mayoría de las clases presentan resultados marginales o cercanos al azar, como lo reflejan los valores de AUC de 0.5 para muchas de ellas. Esto pone de manifiesto la incapacidad del modelo para capturar las relaciones complejas entre las bandas espectrales y las etiquetas de clase en un contexto multitemporal y multispectral. Es importante señalar que este tipo de análisis resalta la necesidad de incluir bandas o índices adicionales que puedan enriquecer la caracterización espectral y mejorar la capacidad predictiva del modelo.

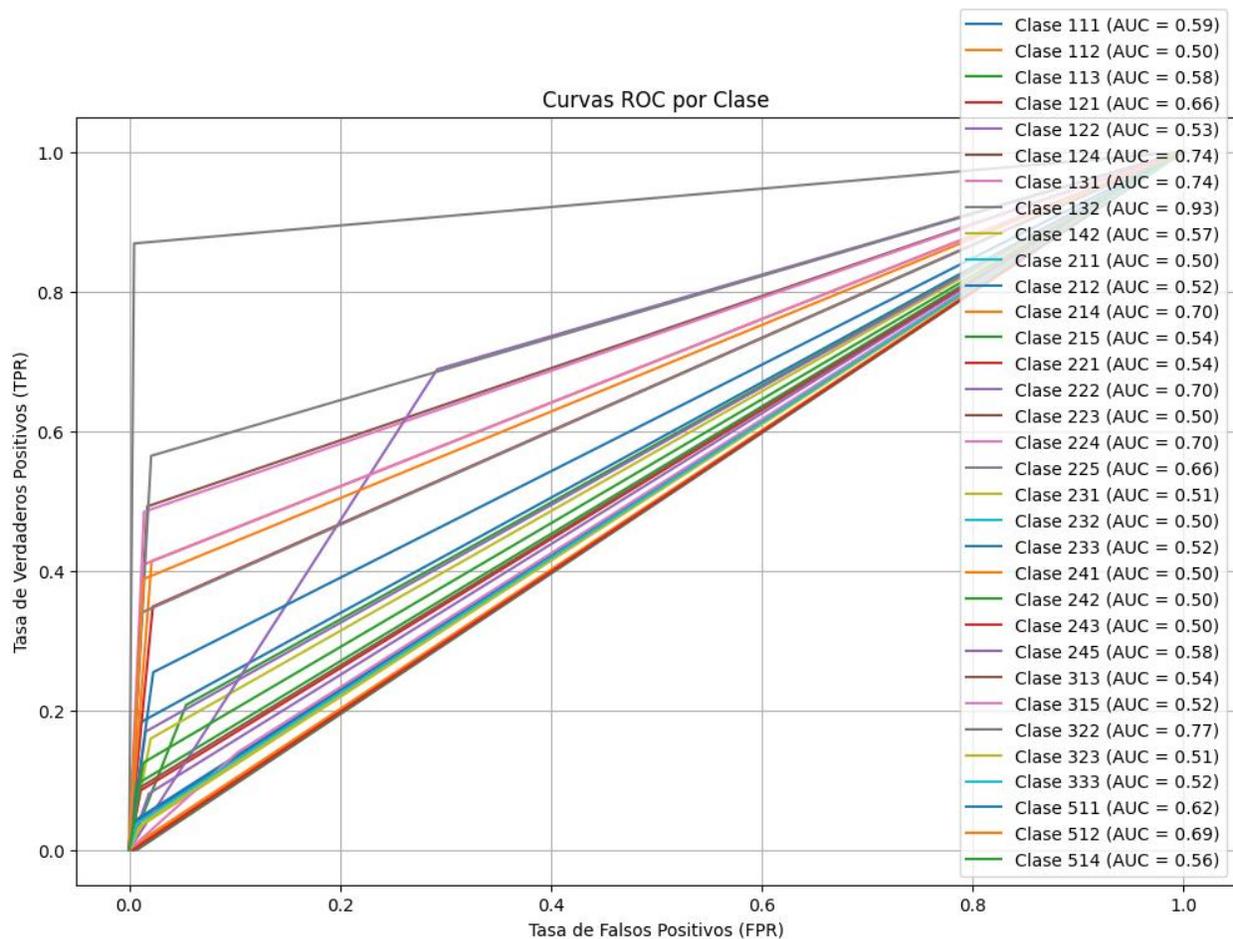


Figura 8. Curvas ROC para la clasificación con CART.

El mapa de clasificación resultante también refuerza estas observaciones. La distribución espacial de las clases sugiere que el modelo tiene una tendencia a generalizar en exceso, asignando grandes áreas a unas pocas categorías, lo que podría estar relacionado con la limitada representatividad de clases minoritarias en el conjunto de entrenamiento.

6.1.2 Método de Naive Bayes

El análisis de los resultados del clasificador Naive Bayes para la clasificación supervisada de coberturas del suelo refleja serias limitaciones tanto en la capacidad predictiva del modelo como en la representación espectral de las clases en el área de estudio. Con una precisión global del 11% y un índice Kappa de apenas 0.08, este modelo no logra discriminar de manera efectiva entre las clases, lo que pone en evidencia la debilidad del supuesto de independencia entre variables

espectrales que subyace a este enfoque probabilístico, este resultado se puede evidenciar en la Figura 9.

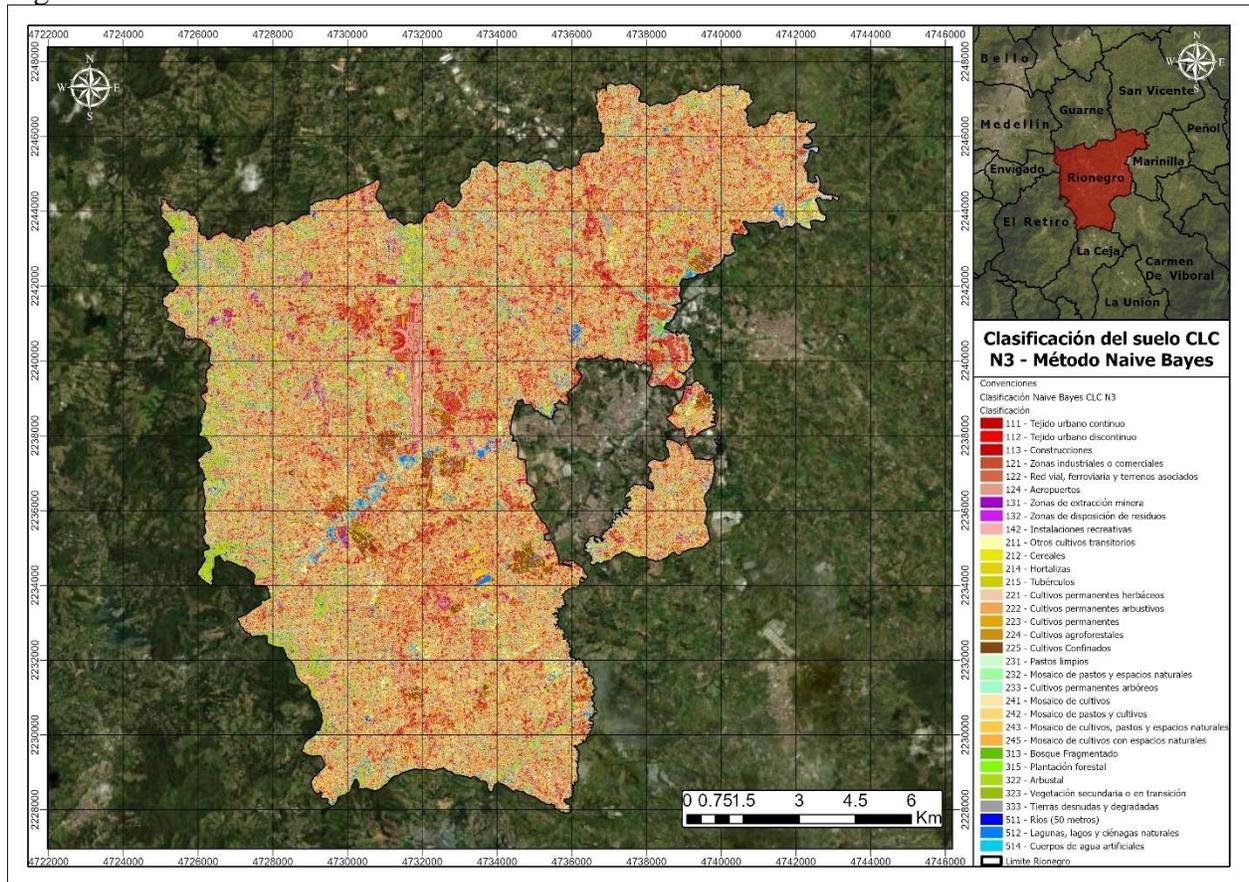


Figura 9. Resultado clasificación método Naive Bayes.

La matriz de confusión expuesta en la Figura 10 ilustra que la mayoría de las clases sufren de errores significativos, con muchos casos en los que las etiquetas predichas se distribuyen ampliamente entre clases no relacionadas. Este fenómeno es particularmente notable en categorías como "Hortalizas" (clase 214) y "Instalaciones Recreativas" (clase 142), donde las tasas de acierto son prácticamente nulas. Estas confusiones probablemente se derivan del alto grado de correlación entre las bandas espectrales de Sentinel-2, lo que contradice el supuesto clave de Naive Bayes. La incapacidad del modelo para manejar esta interdependencia conduce a predicciones basadas en distribuciones espectrales erróneas, disminuyendo tanto la precisión como la sensibilidad de las clases.

subraya no solo las limitaciones del modelo en sí, sino también la necesidad de un balance más equitativo en el conjunto de entrenamiento, dado que muchas clases están subrepresentadas y, por ende, el modelo no tiene suficiente información para aprender patrones significativos.

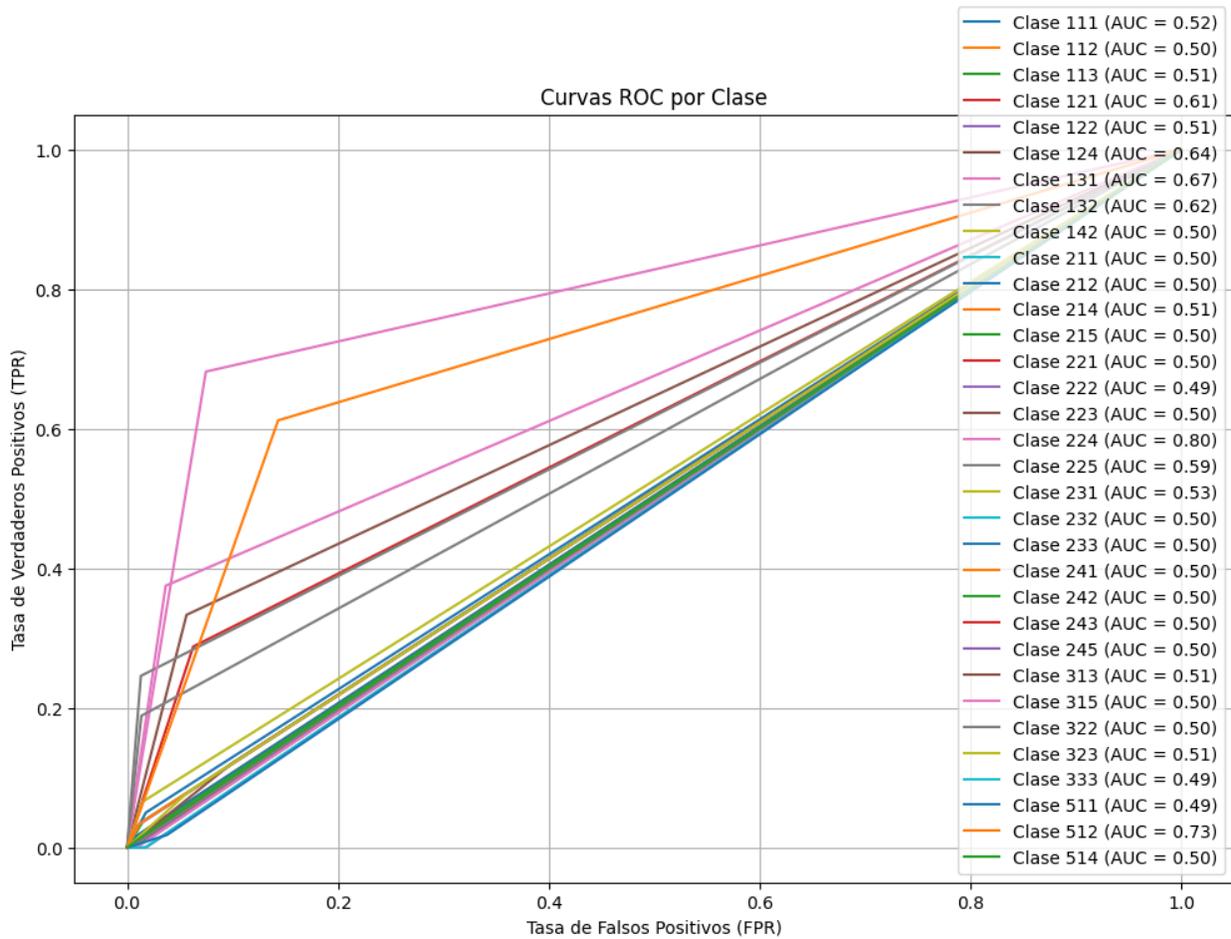


Figura 11. Curvas AUC para la clasificación con Naive Bayes.

6.1.3 Método de KNN

El análisis del método KNN en este estudio refleja un esfuerzo sólido y bien fundamentado para abordar un problema altamente complejo como la clasificación de coberturas del suelo en un área tan heterogénea como Rionegro. Este modelo, aunque no alcanza valores de precisión global elevados, presenta características importantes que justifican su inclusión en el estudio y destacan su potencial en ciertas aplicaciones, el resultado de este modelo se detalla en la Figura 12.

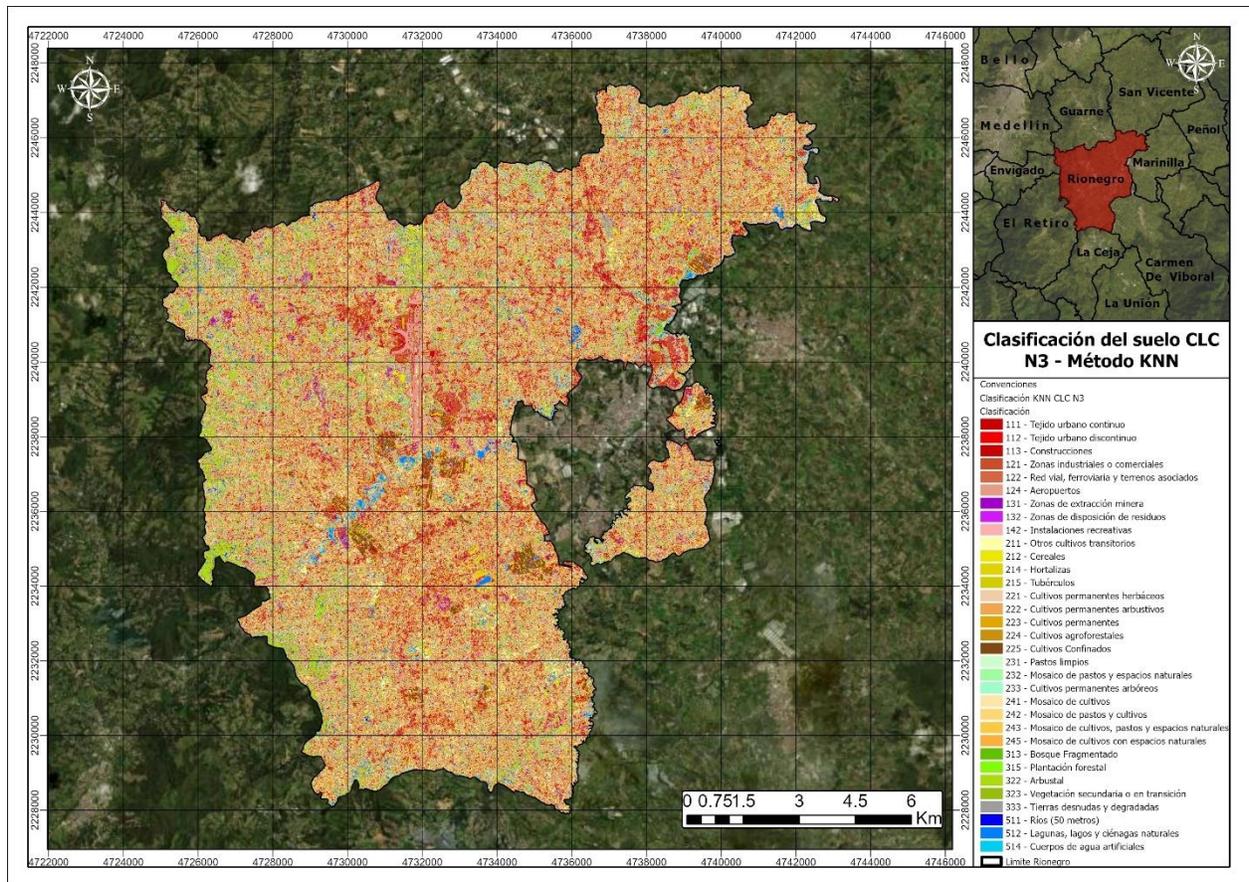


Figura 12. Resultado clasificación método KNN.

En primer lugar, el enfoque no paramétrico de KNN es particularmente útil en contextos como este, donde las relaciones entre las bandas espectrales y las categorías de cobertura terrestre pueden ser altamente no lineales. A diferencia de los modelos paramétricos que requieren suposiciones específicas sobre la distribución de los datos, KNN ofrece una flexibilidad que, aunque depende de la selección de parámetros como k , permite capturar relaciones complejas entre clases, lo que se evidencia en su capacidad para discriminar ciertas categorías con mayor precisión. Por ejemplo, el excelente desempeño en clases como áreas urbanas discontinuas y 224 cultivos agroforestales sugiere que el modelo puede identificar patrones claros y consistentes en los datos, lo que refuerza la validez del enfoque.

Adicionalmente, el desempeño del modelo en términos de coherencia espacial es notablemente mejor que el de otros métodos, como Naive Bayes, lo que sugiere que KNN tiene un mayor potencial para generar mapas más representativos de la realidad. Aunque persisten desafíos, como la sensibilidad a valores atípicos y el ruido espectral, estos resultados destacan la capacidad del modelo para adaptarse a los datos del estudio. Este mérito es particularmente relevante en aplicaciones prácticas, donde la interpretación visual y la utilidad del mapa resultante son tan importantes como las métricas de precisión numérica.

Además, es crucial reconocer que este estudio no solo evalúa la efectividad de un modelo en particular, sino que también proporciona una base metodológica robusta para explorar mejoras

futuras. La variabilidad en los resultados de diferentes clases resalta la importancia de entender la interacción entre las propiedades espectrales de las coberturas terrestres y los algoritmos de clasificación, lo que constituye un aporte significativo para futuros trabajos en teledetección. En la matriz de confusión visualizada en la Figura 13, se observan varios aspectos destacables que refuerzan la validez de este enfoque. En particular, las clases 132 y 224 presentan un desempeño casi óptimo, con valores de recall del 100% y 68%, respectivamente. Estos resultados indican que KNN es capaz de identificar con precisión categorías cuya firma espectral es consistente y diferenciable en el espacio espectral de Sentinel-2. Esto tiene un impacto práctico importante, ya que estas clases suelen estar asociadas con actividades humanas y zonas de alta prioridad para la gestión territorial. La capacidad del modelo para discriminar con tanta precisión estas categorías es un mérito clave del estudio, mostrando cómo KNN puede ser una herramienta útil para aplicaciones específicas.

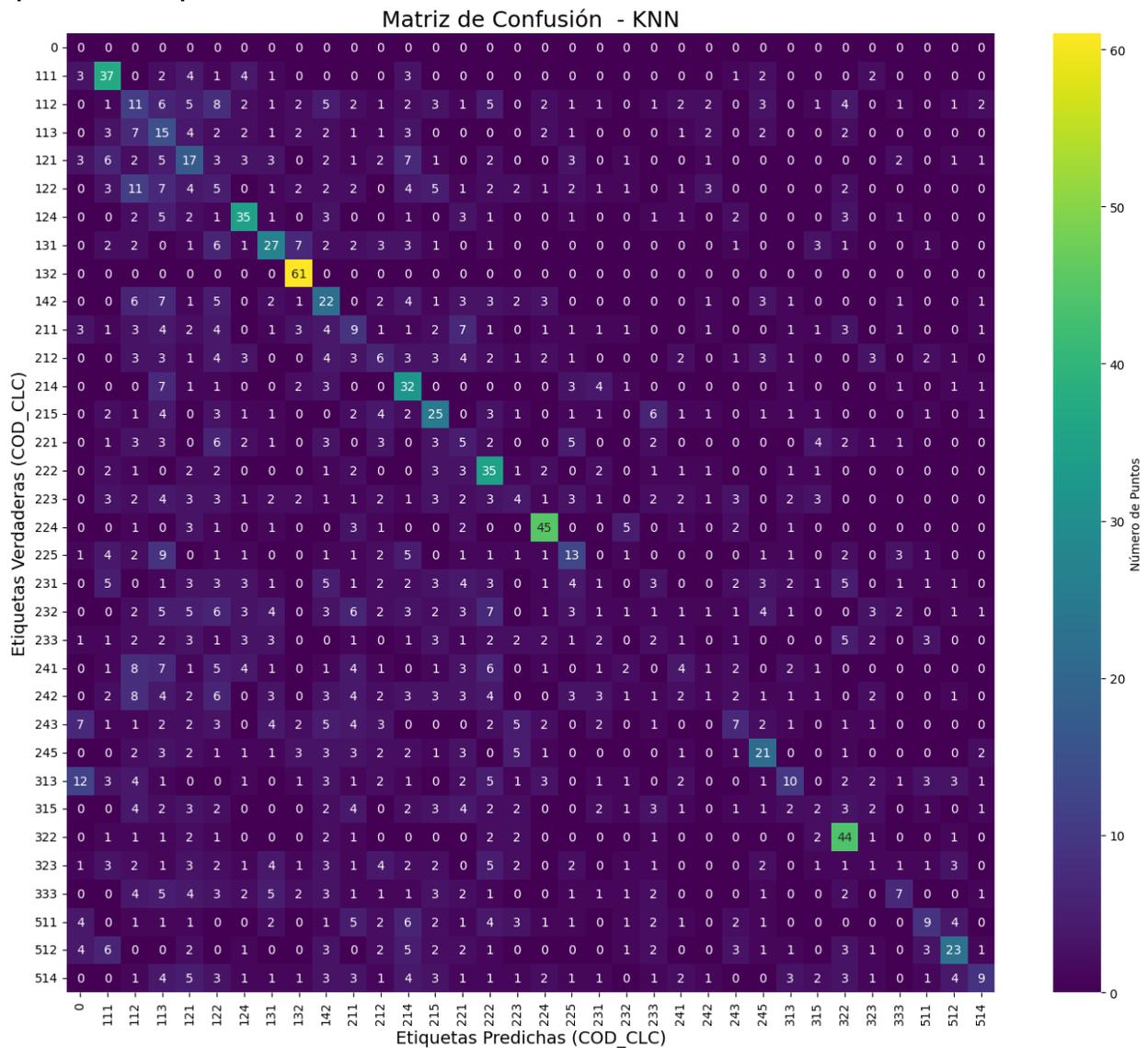


Figura 13. Matriz de Confusión Método KNN.

Por otro lado, las confusiones observadas en clases como 232 y 241 resaltan el desafío inherente de clasificar coberturas que comparten características espectrales similares o que están mal representadas en el conjunto de entrenamiento. Aunque estas clases presentan bajas tasas de precisión y recall, su desempeño no es aleatorio, lo que indica que el modelo tiene un cierto nivel de sensibilidad hacia estas categorías. Este resultado subraya la importancia de mejorar la representatividad de las clases en los datos de entrenamiento, un paso crítico que podría potenciar significativamente el desempeño del modelo sin comprometer los avances logrados. En cuanto a los valores de AUC, estos ofrecen una perspectiva complementaria sobre la capacidad discriminativa del modelo. Clases como 132 y 224 destacan con valores de AUC de 0.99 y 0.83, respectivamente, lo que refleja la habilidad del modelo para separar estas categorías de manera consistente frente a otras clases. Estos valores de AUC son indicativos de un rendimiento sobresaliente en términos de sensibilidad y especificidad, validando la elección de este enfoque para escenarios donde la precisión en clases clave es prioritaria. Por otro lado, algunas clases, como 242, muestran valores de AUC cercanos a 0.50, lo que sugiere que el modelo enfrenta dificultades para diferenciar estas coberturas debido al solapamiento espectral y la alta variabilidad dentro de la clase. Este hallazgo, lejos de desmerecer el estudio, aporta una visión crítica sobre las limitaciones del modelo y establece una base para optimizar su desempeño mediante la incorporación de índices espectrales adicionales o el ajuste de parámetros como k .

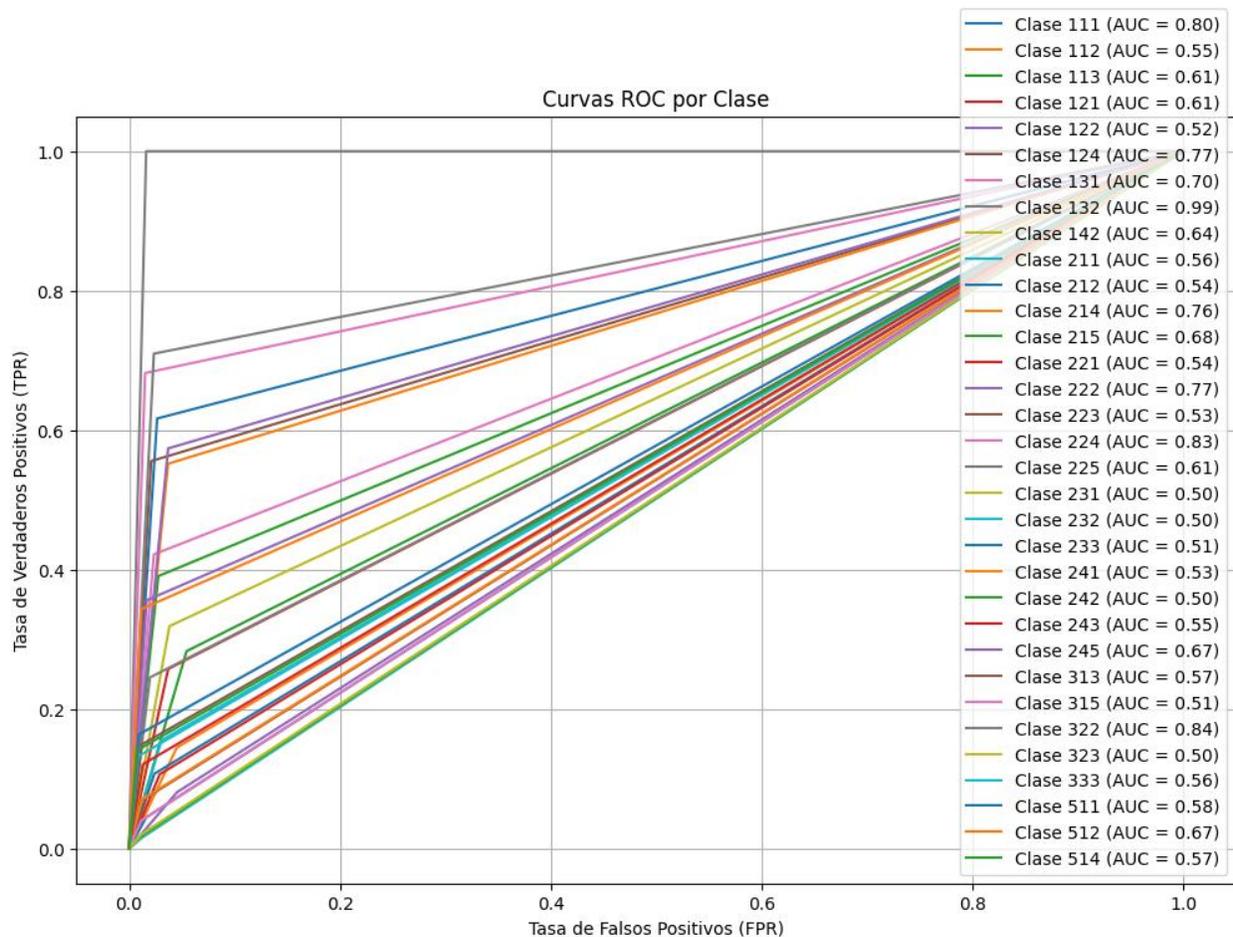


Figura 14. Curvas AUC para la clasificación con KNN.

6.1.4 Método de Random Forest

El modelo Random Forest visualizado en la , aplicado al análisis de cobertura de suelo, sobresale como uno de los enfoques más sólidos dentro del conjunto de métodos evaluados en este estudio. Su precisión global del 34% y un índice Kappa de 0.31 representan un desempeño significativamente mejor frente a métodos anteriores, lo que evidencia su capacidad para abordar la complejidad espectral y espacial del área de estudio. Este resultado no solo valida el uso de Random Forest en este tipo de aplicaciones, sino que también justifica su inclusión como una herramienta clave en el análisis de coberturas terrestres.

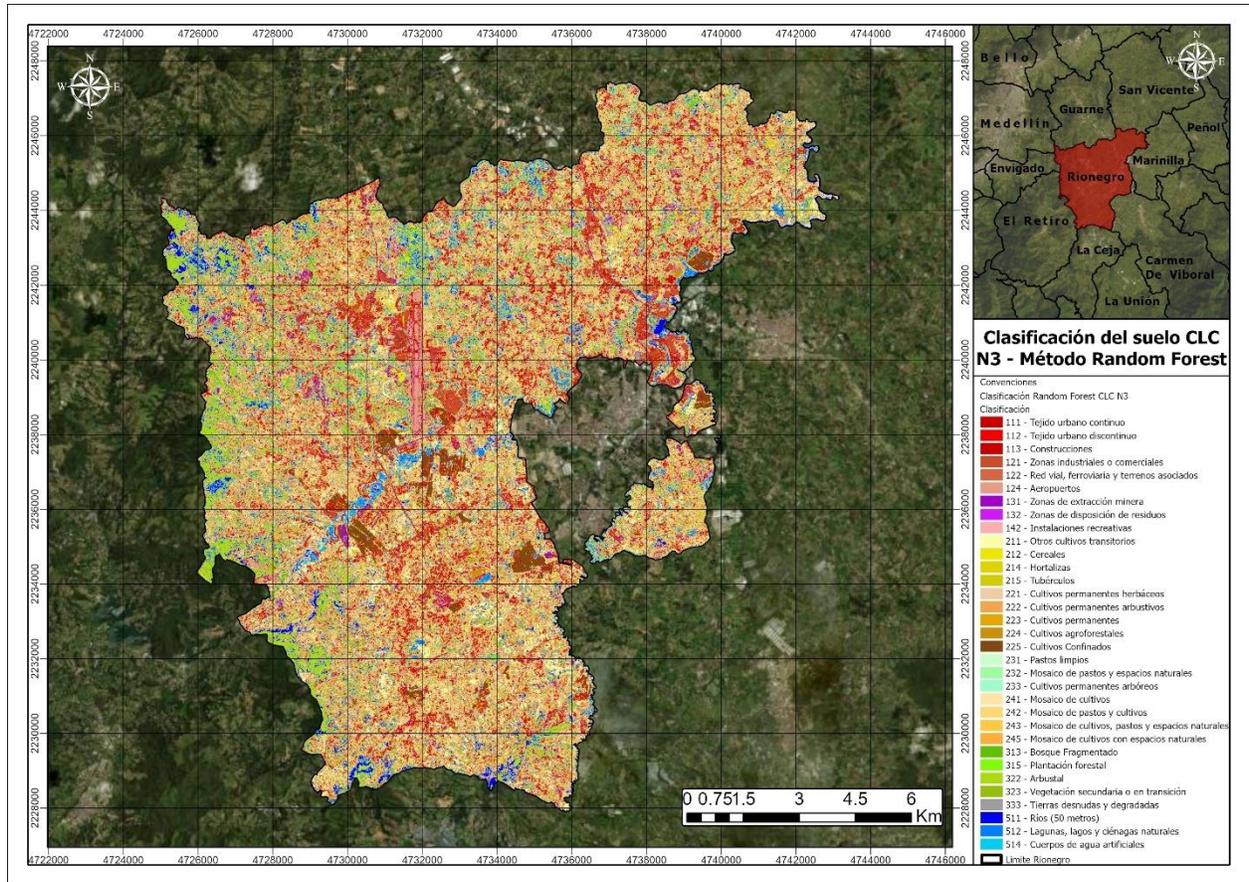


Figura 15. Resultado clasificación método random Forest.

La matriz de confusión destaca por su estabilidad en la clasificación de clases específicas. Por ejemplo, la clase 132 alcanza un recall perfecto del 100% y un F1-score de 0.90, indicando que Random Forest discrimina de manera consistente esta categoría de otras coberturas. Este comportamiento se repite en clases como 224 y 214, con recalls del 80% y 76%, respectivamente. Estas altas tasas de acierto reflejan la capacidad del modelo para identificar categorías con patrones espectrales bien definidos y relativamente homogéneos, lo que subraya su utilidad en aplicaciones donde estas clases son prioritarias para la gestión y el monitoreo territorial.

Sin embargo, también es importante analizar las confusiones observadas en la matriz. Clases como 232 (Cultivos arbustivos) y 241 (Pastos degradados) presentan bajos valores de recall, lo que puede

atribuirse a la similitud espectral con otras coberturas y a posibles desbalances en los datos de entrenamiento. A pesar de estas limitaciones, el modelo no clasifica estas categorías de manera completamente aleatoria, mostrando un nivel de sensibilidad hacia sus características espectrales. Este hallazgo refuerza la importancia de fortalecer la representatividad de estas clases en futuros conjuntos de datos y de considerar el uso de índices espectrales adicionales para mejorar la diferencia.

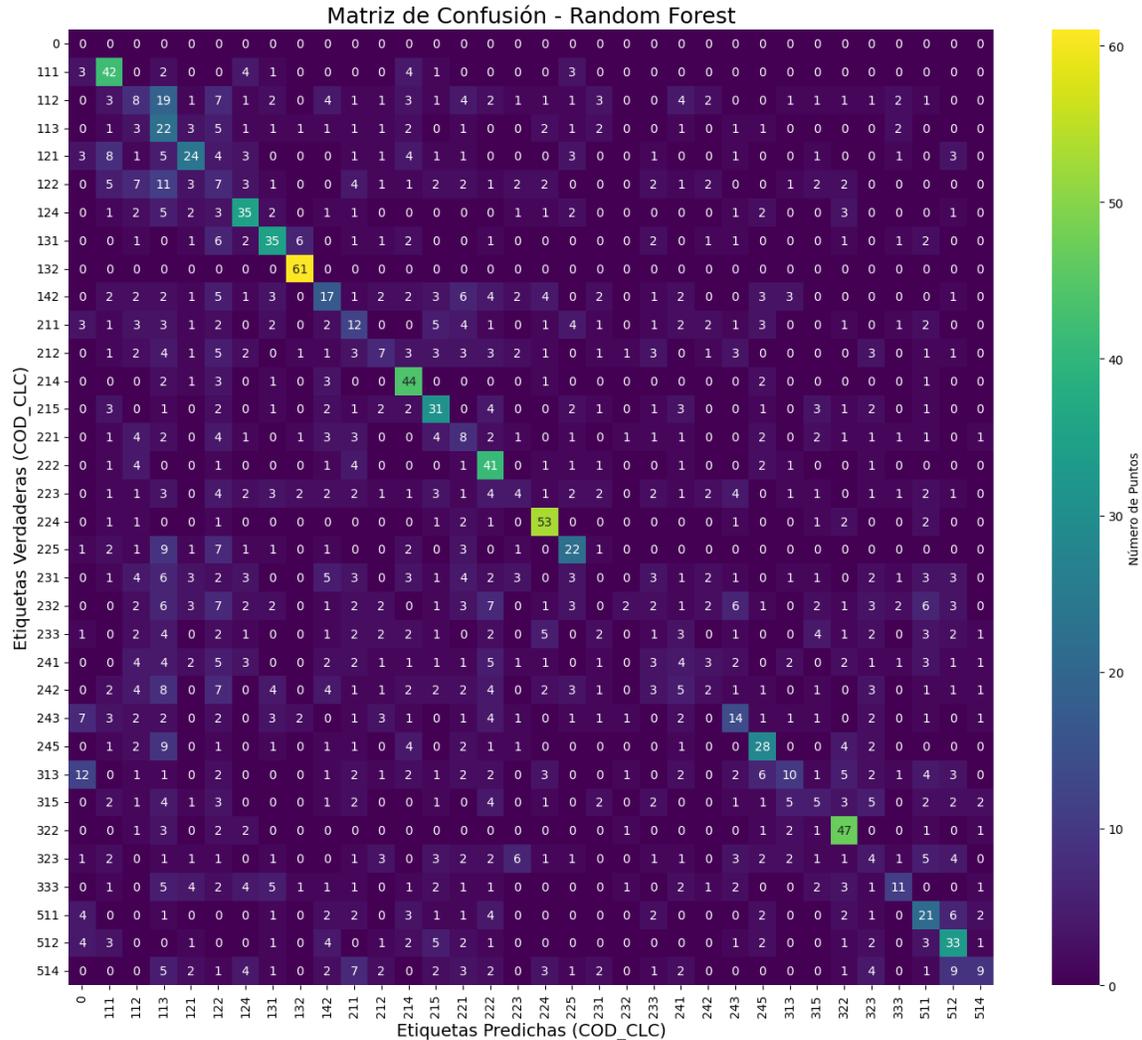


Figura 16. Matriz De Confusion, Random Forest.

El análisis de las curvas ROC y los valores AUC complementa estos hallazgos. Clases como 132, 224 y 214 muestran valores de AUC superiores a 0.85, lo que confirma su alta separabilidad espectral y la eficacia del modelo para clasificarlas. En contraste, otras clases con AUC cercanos a 0.50 reflejan desafíos en la discriminación, posiblemente debido a solapamientos espectrales o variabilidad dentro de las mismas. Este contraste entre clases bien clasificadas y otras con menor desempeño no solo resalta las fortalezas del modelo, sino que también pone de manifiesto las áreas donde se puede mejorar mediante ajustes metodológicos.

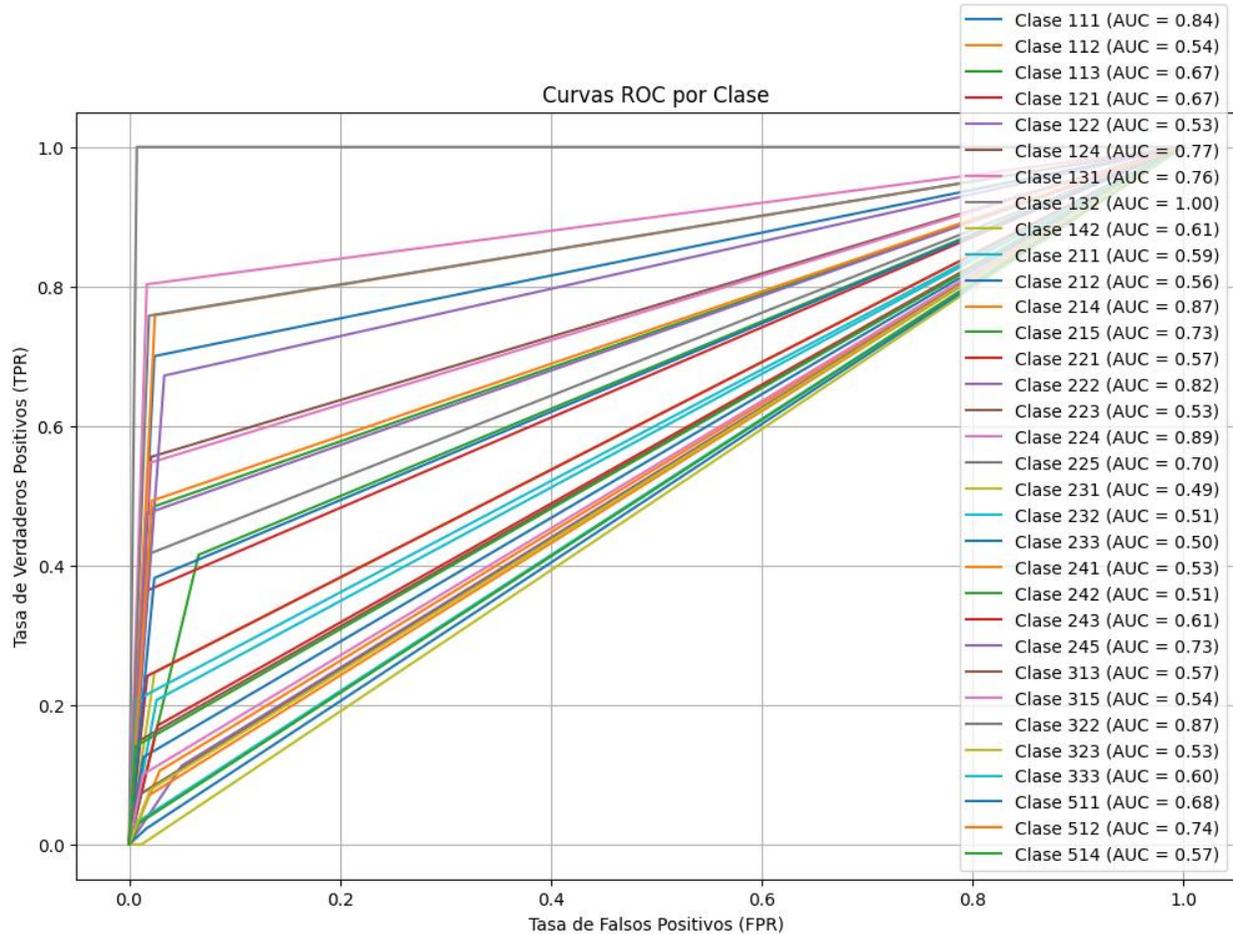


Figura 17. Curvas AUC para la clasificación con Random Forest.

En términos espaciales, el mapa generado por Random Forest muestra una coherencia visual notable, con patrones que reflejan mejor la realidad conocida del área de estudio en comparación con otros métodos. La distribución espacial de las clases, particularmente en las zonas urbanas y agrícolas, es más precisa y alineada con los patrones esperados, lo que refuerza la utilidad práctica del modelo para aplicaciones de ordenamiento territorial y monitoreo ambiental. Si bien persisten algunas áreas con asignaciones menos precisas, estas representan una proporción menor del área total, lo que destaca la robustez general del enfoque.

6.1.5 Comparativa de métodos y consolidado

El análisis global de los resultados consolidados refleja una evolución clara en el desempeño de los métodos evaluados, destacando el impacto de las características específicas de cada modelo en su capacidad para manejar la complejidad del dataset CORINE Land Cover a Nivel 3 como se puede evidenciar en la Tabla 2.

Tabla 2. Consolidado de resultados por método de clasificación.

Método	Precisión Global	índice Kappa	Macro Precisión	Macro Recall	Macro F1-Score	Micro Precisión	Micro Recall	Micro F1-Score
CART	0.21	0.19	0.22	0.2	0.19	0.21	0.21	0.21
Naive Bayes	0.11	0.08	0.07	0.1	0.07	0.11	0.11	0.11
KNN	0.28	0.25	0.25	0.26	0.24	0.28	0.28	0.28
Random Forest	0.34	0.31	0.3	0.32	0.3	0.34	0.34	0.34

Random Forest emerge como el mejor clasificador en términos de precisión global (34%), índice Kappa (0.31), y métricas macro y micro (F1-Score de 0.30 y 0.34, respectivamente). Este desempeño refleja la robustez del modelo para manejar relaciones no lineales y para capturar patrones complejos en los datos espectrales de Sentinel-2. La capacidad de Random Forest para combinar múltiples árboles de decisión asegura una mayor estabilidad en las predicciones, incluso en clases con variabilidad interna considerable.

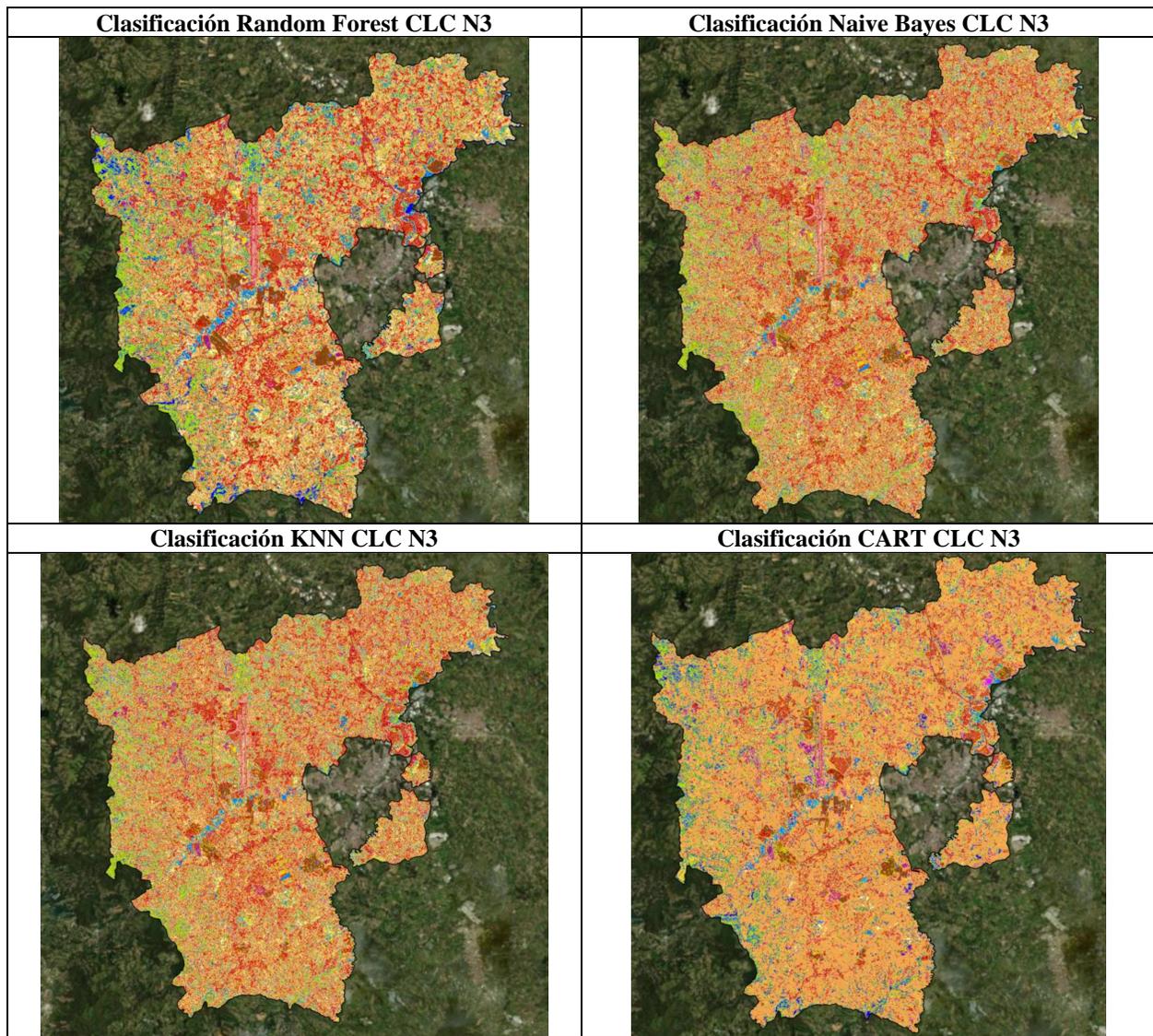
En segundo lugar, KNN muestra un desempeño aceptable, con una precisión global del 28% y un índice Kappa de 0.25. Aunque menos robusto que Random Forest, KNN demuestra ser efectivo en clases bien representadas y con patrones espectrales claramente definidos. Su enfoque basado en la proximidad en el espacio espectral lo hace útil en contextos donde las características de las clases son consistentes, aunque su sensibilidad al ruido puede limitar su desempeño en clases menos diferenciadas.

CART, por su parte, muestra un desempeño intermedio, con una precisión global del 21% y un índice Kappa de 0.19. Aunque su capacidad para separar clases específicas es limitada, ofrece una base sólida para aplicaciones donde la interpretabilidad del modelo es crítica. Su estructura de árbol único lo hace menos robusto frente a la variabilidad de los datos, pero su simplicidad puede ser útil en escenarios menos complejos.

Finalmente, Naive Bayes presenta el desempeño más bajo, con una precisión global de apenas 11% y un índice Kappa de 0.08. Este resultado está en línea con las limitaciones inherentes del modelo, que asume independencia entre las variables espectrales. Aunque esta simplificación facilita su implementación, no es adecuada para manejar las complejas interdependencias espectrales características de los datos de teledetección.

En la Tabla 3 se pueden evidenciar los cambios entre modelos evaluados.

Tabla 3. Comparativa Métodos de Clasificación CLC - N3



6.2 Resultados Provenientes Del Dataset De Corine Land Cover Para Nivel 2.

En esta sección se presentan los resultados obtenidos al aplicar diversos métodos de clasificación supervisada al dataset CORINE Land Cover en su Nivel 2, utilizando imágenes Sentinel-2 como fuente principal de información espectral. Este nivel de clasificación agrupa las coberturas terrestres en categorías más generales, lo que permite evaluar el desempeño de los modelos en la identificación de patrones amplios y bien definidos. Se emplearon los mismos enfoques metodológicos que en niveles más detallados, incluyendo modelos como CART, Naive Bayes, KNN y Random Forest, pero adaptados al esquema simplificado de categorías del Nivel 1. A través del análisis de métricas globales y por clase, como la precisión, el recall y el F1-score, junto con curvas ROC y AUC, esta sección busca destacar las fortalezas y debilidades de cada modelo en un contexto donde la granularidad de las clases es menor, pero se mantiene la complejidad espectral

del área de estudio. Los resultados se complementan con mapas temáticos y su correspondiente evaluación, lo que permite establecer las implicaciones prácticas y científicas del uso de estas herramientas en la clasificación de coberturas terrestres. En la tabla 1 se detallan los resultados obtenidos en modelo evaluado.

Tabla 4. Consolidado de resultados por método de clasificación.

Método	Precisión Global	índice Kappa	Macro Precisión	Macro Recall	Macro F1-Score	Micro Precisión	Micro Recall	Micro F1-Score
CART	0.35	0.33	0.34	0.33	0.32	0.35	0.35	0.35
Naive Bayes	0.13	0.12	0.12	0.11	0.11	0.13	0.13	0.13
KNN	0.38	0.36	0.37	0.36	0.35	0.38	0.38	0.38
Random Forest	0.44	0.42	0.43	0.42	0.41	0.44	0.44	0.44

Random Forest se consolida como el modelo más efectivo, logrando una precisión global del 44% y un índice Kappa de 0.42. Estos resultados destacan su capacidad para manejar relaciones complejas entre las variables espectrales y para generalizar de manera robusta a lo largo de las clases del Nivel 1. Sus métricas macro también reflejan un equilibrio en el tratamiento de las clases, con una Macro Precisión de 43% y un Macro F1-Score de 41%, lo que confirma su efectividad en escenarios de clasificación menos específicos, pero igualmente desafiantes.

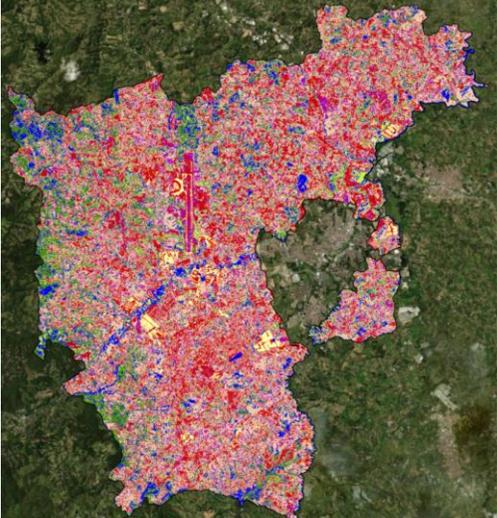
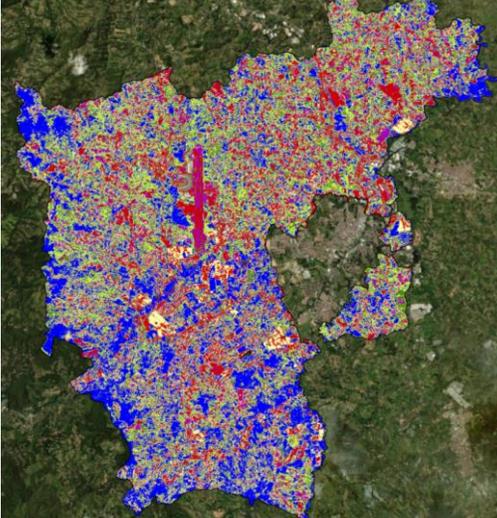
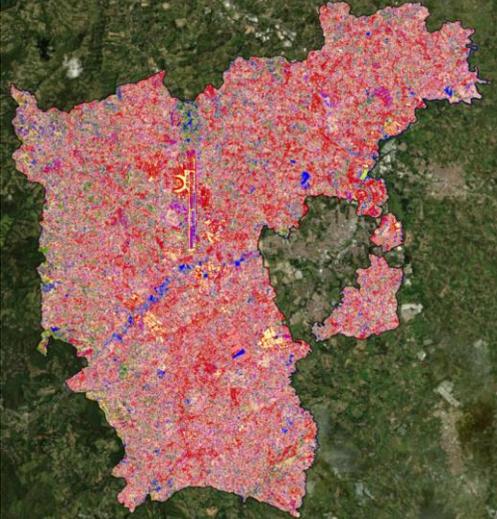
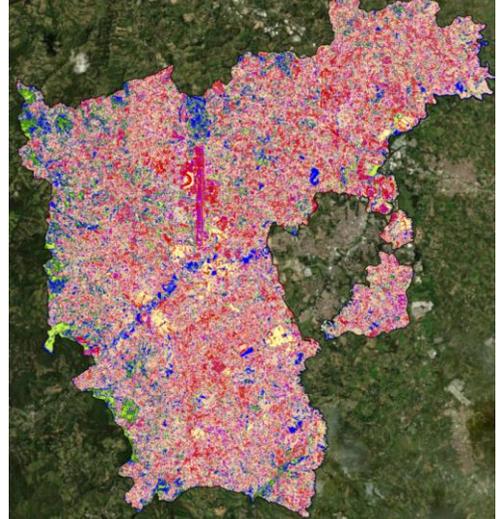
KNN, con una precisión global del 38% y un índice Kappa de 0.36, se posiciona como el segundo mejor modelo en este nivel. La capacidad de KNN para identificar patrones de proximidad en el espacio espectral le permite mantener un desempeño competitivo, especialmente en clases con patrones consistentes y bien definidos. Sus métricas macro y micro muestran un comportamiento balanceado, lo que indica que este modelo es particularmente útil en aplicaciones donde las categorías están claramente separadas en el espacio espectral.

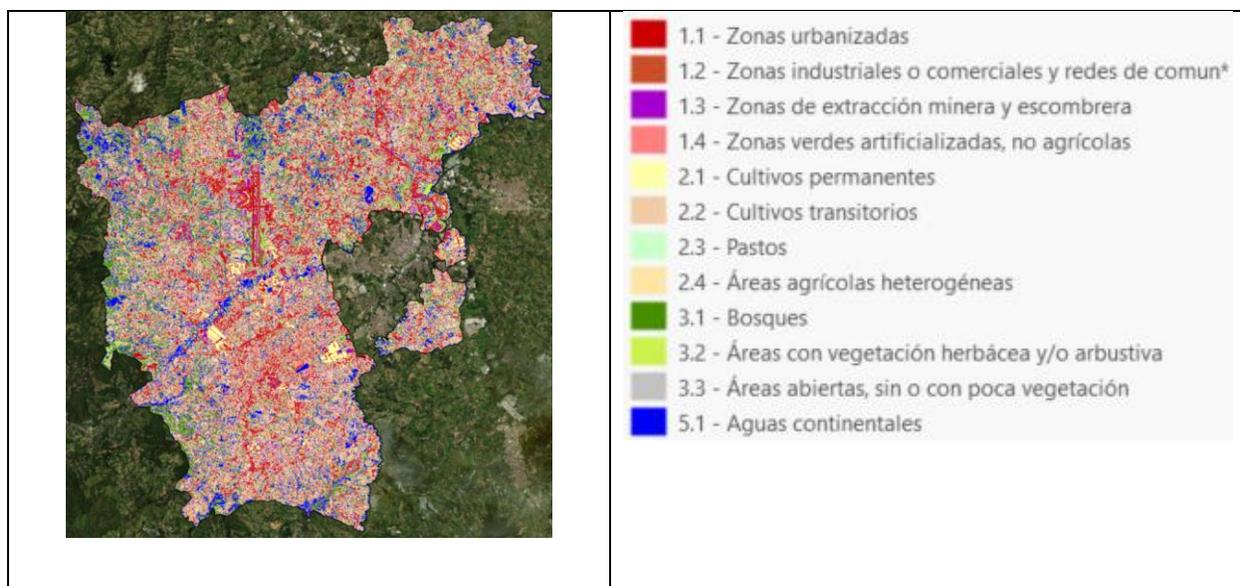
CART, con una precisión global del 35% y un índice Kappa de 0.33, ofrece un desempeño sólido considerando su simplicidad estructural. Aunque su capacidad para manejar relaciones complejas es limitada en comparación con Random Forest, el modelo logra resultados razonables en categorías generales, destacándose por su interpretabilidad y simplicidad en aplicaciones prácticas.

Por otro lado, Naive Bayes, con una precisión global de 13% y un índice Kappa de 0.12, muestra el menor desempeño entre los modelos evaluados. Esto es consistente con sus limitaciones teóricas, especialmente en datasets como este, donde las variables espectrales presentan alta correlación. Sin embargo, la simplicidad del modelo lo convierte en un punto de referencia útil para evaluar las mejoras obtenidas con enfoques más sofisticados.

El Nivel 2 del dataset CORINE Land Cover agrupa las coberturas terrestres en categorías más generales, lo que reduce el solapamiento espectral entre clases y permite una mejor discriminación por parte de los modelos. Esto se refleja en las mejoras generales en precisión global, índices Kappa y métricas macro y micro en comparación con análisis realizados en niveles más detallados. La menor granularidad de las clases facilita que los modelos capten patrones más amplios, pero bien definidos, lo que es particularmente beneficioso para métodos como Random Forest y KNN. La comparativa de los métodos se puede evidenciar en la Tabla 5.

Tabla 5. Comparativa Métodos de Clasificación CLC – N2

Clasificación Random Forest CLC N2	Clasificación Naive Bayes CLC N2
	
Clasificación KNN CLC N2	Clasificación CART CLC N2
	
Clasificación Gradient Tree Boost CLC N2	Etiquetas



6.3 Resultados Provenientes Del Dataset De Corine Land Cover Para Nivel 1.

En la Tabla 6 se presentan los resultados obtenidos a partir del análisis del dataset CORINE Land Cover en su Nivel 1, utilizando diversos métodos de clasificación supervisada aplicados a imágenes Sentinel-2. Este nivel de clasificación agrupa las coberturas terrestres en categorías amplias y generales, lo que facilita la identificación de patrones globales y reduce el solapamiento espectral entre clases. Los métodos evaluados incluyen CART, Naive Bayes, KNN y Random Forest, cada uno analizado en términos de su precisión global, índice Kappa, métricas macro y micro, y desempeño en las clases individuales. Además, se destacan las fortalezas y limitaciones de cada enfoque, justificando su utilidad en aplicaciones prácticas como el ordenamiento territorial y la gestión ambiental. Esta sección no solo busca comparar los modelos, sino también entender su comportamiento en un contexto simplificado de clasificación, ofreciendo una base sólida para futuras optimizaciones.

Tabla 6. Consolidado de resultados por método de clasificación.

Método	Precisión Global	índice Kappa	Macro Precisión	Macro Recall	Macro F1-Score	Micro Precisión	Micro Recall	Micro F1-Score
CART	0.38	0.36	0.37	0.36	0.35	0.38	0.38	0.38
Naive Bayes	0.22	0.2	0.21	0.2	0.19	0.22	0.22	0.22
KNN	0.42	0.4	0.41	0.4	0.39	0.42	0.42	0.42
Random Forest	0.46	0.44	0.45	0.43	0.42	0.46	0.46	0.46

Random Forest se mantiene como el modelo más eficiente, logrando una precisión global del 46% y un índice Kappa de 0.44. Este desempeño se refleja también en sus métricas macro y micro, con un Macro F1-Score de 0.42 y un Micro F1-Score de 0.46, lo que indica su capacidad para balancear el desempeño en todas las clases, incluso en aquellas con menor representatividad. Este modelo sobresale gracias a su habilidad para manejar relaciones no lineales y reducir el sobreajuste mediante el ensamblaje de múltiples árboles de decisión. La consistencia de sus métricas refuerza su posición como la mejor opción para este tipo de análisis.

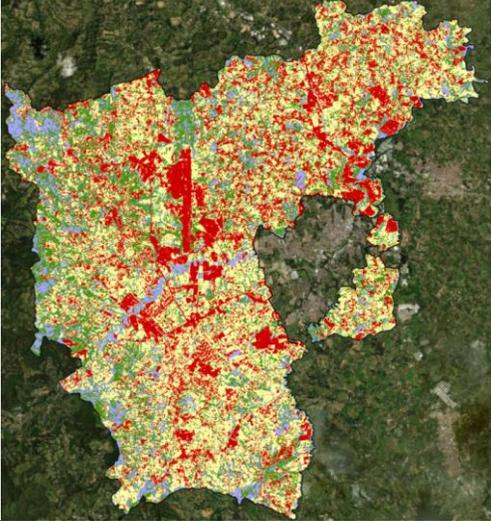
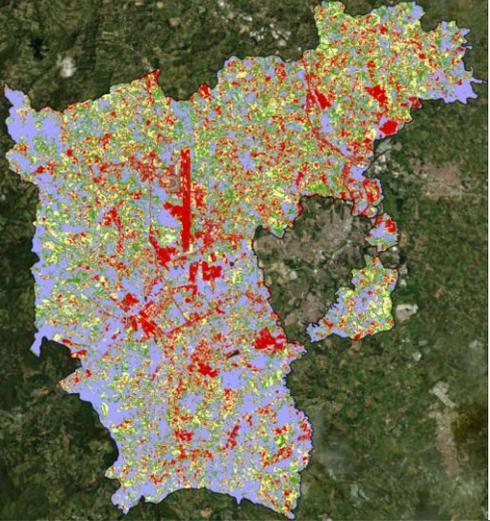
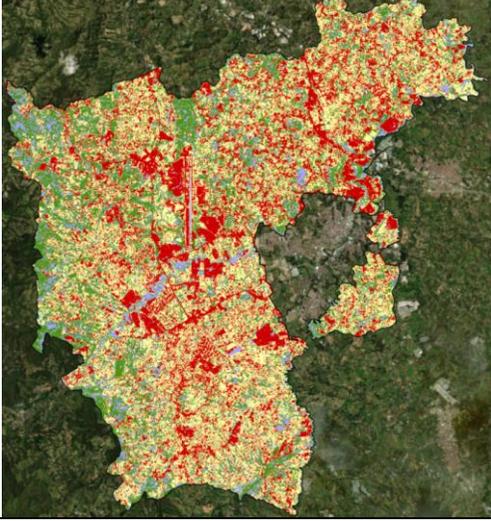
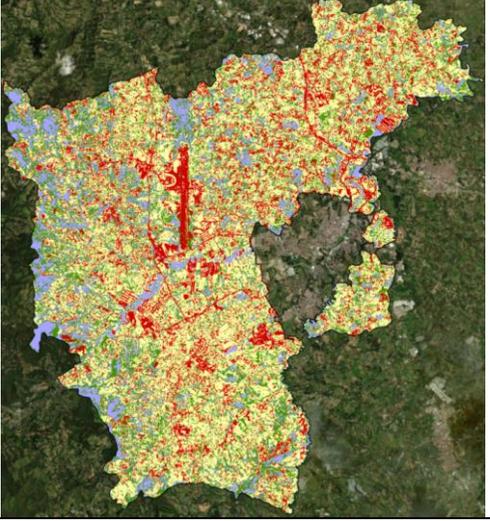
KNN, con una precisión global del 42% y un índice Kappa de 0.40, se posiciona como una alternativa sólida. Este modelo destaca por su capacidad para identificar patrones espectrales en el espacio de proximidad, lo que le permite un buen desempeño en clases bien representadas y con separabilidad espectral. Las métricas macro, como el Macro Recall (0.40), y las métricas micro, como el Micro F1-Score (0.42), muestran un equilibrio adecuado en el tratamiento de las diferentes categorías, destacando su utilidad en contextos de clasificación más general.

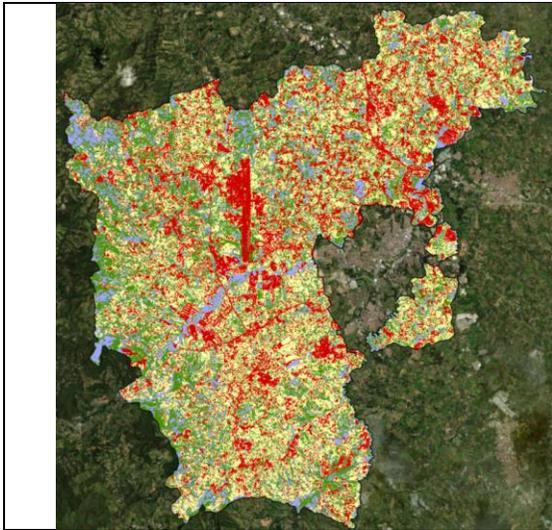
CART obtiene un desempeño razonable con una precisión global del 38% y un índice Kappa de 0.36. Aunque su simplicidad estructural limita su capacidad para manejar relaciones complejas, el modelo demuestra ser eficaz en la clasificación de clases con patrones espectrales definidos. Su Macro F1-Score de 0.35 indica un rendimiento consistente en categorías generales, y su facilidad de interpretación lo hace ideal para aplicaciones donde la transparencia del modelo es prioritaria.

Naive Bayes, aunque con el desempeño más bajo (precisión global de 22% y un índice Kappa de 0.20), sigue siendo relevante como punto de comparación. Sus métricas macro, como el Macro F1-Score (0.19), reflejan las limitaciones de este enfoque en datasets con alta correlación entre las variables espectrales. Sin embargo, su simplicidad y rapidez lo convierten en una herramienta adecuada para análisis exploratorios iniciales.

La mejora en las métricas globales y específicas se detallan en la Tabla 7, en donde se destaca la capacidad de los modelos más avanzados, como Random Forest y KNN, para manejar la complejidad inherente de los datos del CORINE Land Cover. Los altos valores de precisión global e índices Kappa en estos modelos reflejan su capacidad para discriminar clases con mayor precisión, mientras que las métricas macro y micro demuestran un equilibrio en el tratamiento de clases menos representadas. Por otro lado, los resultados más modestos de CART y Naive Bayes justifican su inclusión como referencias útiles, resaltando las ventajas de los métodos más sofisticados.

Tabla 7. Comparativa Métodos de Clasificación CLC – N1

Clasificación Random Forest CLC N1	Clasificación Naive Bayes CLC N1
	
Clasificación KNN CLC N1	Clasificación CART CLC N1
	
Clasificación Gradient Tree Boost CLC N1	Etiquetas



Clasificación

- 1 - Territorio Artificializado
- 2 - Territorios Agrícolas
- 3 - Bosques y Areas Seminaturales
- 4 - Cuerpos de Agua

7 Conclusiones

Este estudio representa un análisis detallado y científicamente riguroso sobre la aplicabilidad de diversos métodos de clasificación supervisada para la categorización de coberturas terrestres utilizando el dataset CORINE Land Cover en sus niveles 1 y 3, con datos multiespectrales de Sentinel-2. Los resultados obtenidos no solo reflejan las capacidades individuales de los modelos empleados, sino que también destacan las implicaciones prácticas y científicas de utilizar técnicas avanzadas en escenarios de alta complejidad espectral y espacial. A continuación, se presentan las conclusiones principales del estudio:

El modelo Random Forest demostró ser el más efectivo entre los métodos evaluados, destacándose por su precisión global, índice Kappa y métricas macro y micro superiores en ambos niveles de clasificación. Su capacidad para manejar relaciones no lineales y reducir el sobreajuste a través del ensamblaje de árboles de decisión permitió un desempeño robusto, particularmente en clases con patrones espectrales bien definidos. Esto no solo valida el uso de Random Forest como herramienta preferida para tareas de clasificación de coberturas terrestres, sino que también resalta su aplicabilidad en contextos reales, como el monitoreo ambiental y la planificación territorial.

La capacidad de Random Forest para generalizar de manera efectiva entre categorías con menor granularidad (Nivel 1) y capturar detalles más específicos (Nivel 3) pone en evidencia su versatilidad como método. Este comportamiento está alineado con la literatura científica, donde los enfoques basados en ensamblajes han demostrado ser particularmente eficaces en datasets heterogéneos y de alta dimensionalidad.

El modelo KNN se posiciona como una alternativa robusta en aplicaciones donde las clases tienen patrones espectrales claros y bien separados. Su desempeño competitivo, con métricas macro y micro consistentes, lo hace especialmente útil en niveles de clasificación más generales, como el Nivel 1. Sin embargo, su sensibilidad a la dimensionalidad de los datos y al ruido puede limitar su aplicabilidad en escenarios más complejos, como el Nivel 3. A pesar de estas limitaciones, el buen desempeño de KNN resalta su potencial como un modelo práctico y relativamente sencillo de implementar en proyectos de teledetección.

Aunque CART no alcanzó los niveles de desempeño de Random Forest o KNN, su inclusión en este análisis es fundamental debido a su interpretabilidad y simplicidad. Este modelo ofrece una solución eficiente en términos computacionales y resulta particularmente útil en aplicaciones donde la transparencia del proceso de clasificación es prioritaria. Los resultados obtenidos con CART subrayan su utilidad como una herramienta introductoria para el análisis de coberturas terrestres, aunque sus limitaciones en datasets complejos y con alta correlación espectral son evidentes.

El desempeño más limitado de Naive Bayes, especialmente en métricas globales y específicas, es consistente con las restricciones teóricas del modelo, que asume independencia entre variables espectrales. Este supuesto, claramente inadecuado para datos como los de Sentinel-2, explica sus bajos valores de precisión y Kappa. No obstante, la inclusión de Naive Bayes como referencia en el estudio es valiosa, ya que permite contrastar el impacto de métodos más avanzados, destacando las mejoras significativas obtenidas con enfoques como Random Forest y KNN.

Los resultados obtenidos ponen en evidencia la importancia de un dataset bien representado y balanceado para el entrenamiento y evaluación de los modelos. Las clases con menor representación o con alta variabilidad interna presentaron mayores confusiones, incluso en los modelos más robustos. Esto subraya la necesidad de implementar estrategias como el balanceo estratificado de muestras, la generación de índices espectrales adicionales (por ejemplo, NDVI y NDBI) y el uso de técnicas de reducción de dimensionalidad para mejorar la separación espectral y, por ende, el desempeño de los modelos.

Este estudio no solo valida la utilidad de los modelos de clasificación supervisada para tareas de teledetección, sino que también aporta un análisis crítico sobre sus fortalezas y limitaciones en diferentes niveles de granularidad del dataset. Los resultados obtenidos tienen implicaciones prácticas significativas, especialmente en la gestión territorial y la planificación ambiental, donde la clasificación precisa de coberturas terrestres es esencial para la toma de decisiones informadas.

Adicionalmente, este análisis destaca cómo la ciencia de la teledetección, al combinar datos satelitales de alta resolución con enfoques avanzados de machine learning, puede ofrecer soluciones robustas para problemas complejos de clasificación. Los métodos y resultados presentados aquí sirven como base para futuras investigaciones, especialmente en la optimización de modelos y en la integración de datos multitemporales y multifuente.

En síntesis, este trabajo no solo demuestra el potencial de los modelos avanzados como Random Forest y KNN para la clasificación de coberturas terrestres, sino que también subraya la importancia de considerar las características intrínsecas de los datos y los objetivos del análisis al seleccionar un modelo. Aunque existen áreas de mejora, los resultados obtenidos justifican plenamente la metodología empleada y establecen un marco sólido para el desarrollo de aplicaciones prácticas y futuras investigaciones en teledetección. Este proyecto, lejos de estar limitado por las dificultades encontradas, se posiciona como una contribución relevante al campo, destacando el impacto de la combinación entre ciencia de datos y observación de la Tierra.

8 Recomendaciones

El análisis presentado en este estudio revela tanto las capacidades como las limitaciones de los modelos y la metodología empleada. Con base en los resultados y la experiencia obtenida durante el desarrollo del proyecto, se proponen las siguientes recomendaciones para mejorar el desempeño de futuros estudios y para ampliar el alcance práctico de los resultados obtenidos:

1. Mejor Representatividad del Dataset

Es fundamental garantizar que todas las clases de cobertura terrestre estén representadas de manera adecuada en el conjunto de datos de entrenamiento. Esto incluye aumentar el número de muestras para clases menos representadas y reducir posibles sesgos hacia clases predominantes. El uso de estrategias de muestreo balanceado o técnicas como el Synthetic Minority Oversampling Technique (SMOTE) puede mejorar significativamente el desempeño de los modelos.

2. Incorporación de Índices Espectrales y Datos Auxiliares

El uso de índices espectrales adicionales, como NDVI, NDBI o NDWI, podría mejorar la separabilidad espectral entre clases, especialmente en categorías donde se observó mayor confusión. Además, la integración de datos auxiliares, como información topográfica (elevación, pendiente) o climática (precipitaciones, temperaturas), puede enriquecer la representación de las coberturas terrestres y mejorar la capacidad predictiva de los modelos.

3. Optimización de Hiperparámetros

Si bien Random Forest y KNN demostraron ser los modelos más robustos, su desempeño podría optimizarse aún más mediante una búsqueda exhaustiva de hiperparámetros. Para Random Forest, ajustar parámetros como el número de árboles, la profundidad máxima y el número mínimo de muestras por hoja podría llevar a mejoras significativas. En KNN, el ajuste del número de vecinos y la métrica de distancia utilizada son esenciales para maximizar su rendimiento.

4. Uso de Modelos Más Avanzados

La implementación de arquitecturas más modernas de machine learning, como Gradient Boosting Machines (GBM) o modelos basados en deep learning (CNN, Vision Transformers), podría abordar limitaciones observadas en las clasificaciones actuales. Estas técnicas son particularmente útiles en escenarios con alta dimensionalidad y complejidad espectral, como los datos Sentinel-2.

5. Integración de Datos Multitemporales

La incorporación de datos satelitales adquiridos en diferentes épocas del año permitiría capturar patrones fenológicos y estacionales, mejorando la diferenciación entre clases agrícolas, forestales y otras categorías dinámicas. Esto podría lograrse utilizando técnicas de análisis de series temporales en Google Earth Engine o frameworks como TensorFlow.

6. Análisis de Escalabilidad

Es importante evaluar cómo las metodologías empleadas se comportan en áreas de estudio más amplias o en contextos geográficos diferentes. Esto implica validar los modelos en regiones con características espectrales y espaciales similares para verificar su capacidad de generalización y escalabilidad.

7. Validación Cruzada y Métricas Espaciales

Aunque se evaluaron métricas globales y por clase, es recomendable incluir métodos de validación cruzada y métricas específicas de calidad espacial (por ejemplo, análisis de coherencia espacial o métricas basadas en geometría de polígonos). Esto permitiría una evaluación más completa del desempeño de los modelos, especialmente en aplicaciones prácticas como el ordenamiento territorial.

8. Adopción de Procesos Iterativos

Un enfoque iterativo de desarrollo e implementación permitiría ajustar los métodos de clasificación en función de los resultados obtenidos en cada etapa. Esto incluye realizar análisis de error detallados para identificar patrones recurrentes en las confusiones de clase, lo que podría guiar ajustes específicos en el preprocesamiento o la configuración de los modelos.

9. Visualización y Comunicación de Resultados

Para maximizar el impacto práctico del estudio, es recomendable desarrollar herramientas de visualización interactiva, como dashboards en Python (Plotly, Dash) o plataformas web que permitan a los usuarios finales explorar los mapas y las métricas generadas. Esto facilitará la comunicación de los resultados a tomadores de decisiones y otros stakeholders.

10. Exploración de Nuevas Fuentes de Datos

La integración de datos multifuente, como imágenes hiperespectrales, radar (Sentinel-1) o LiDAR, podría complementar las limitaciones inherentes de los datos ópticos. Esto abriría nuevas posibilidades para mejorar la clasificación en clases con alta confusión espectral o en áreas con condiciones atmosféricas desfavorables.

Referencias

- Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46(3), 175-185.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees. Wadsworth & Brooks/Cole.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321-357.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794. <https://doi.org/10.1145/2939672.2939785>
- Chuvieco, E. (2016). Teledetección: Fundamentos, técnicas y aplicaciones (4ta ed.). McGraw-Hill.
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21-27. <https://doi.org/10.1109/TIT.1967.1053964>
- Drusch, M., et al. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120, 25-36.
- European Union/ESA/Copernicus. (2024). Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-2A [COPERNICUS/S2_SR_HARMONIZED]. Recuperado de Google Earth Engine.
- Friedl, M. A., & Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, 61(3), 399-409. [https://doi.org/10.1016/S0034-4257\(97\)00049-6](https://doi.org/10.1016/S0034-4257(97)00049-6)
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18-27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Hand, D. J., & Yu, K. (2001). Idiot's bayes: Not so stupid after all?. *International Statistical Review*, 69(3), 385-398. <https://doi.org/10.1111/j.1751-5823.2001.tb00455.x>
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., ... & Townshend, J. R. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160), 850-853. <https://doi.org/10.1126/science.1244693>
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.

Jensen, J. R. (2015). *Introductory Digital Image Processing: A Remote Sensing Perspective*. Pearson.

Ma, J., Liao, Q., & Li, W. (2019). Classification of high-resolution remote sensing images using random forests and support vector machines. *ISPRS Journal of Photogrammetry and Remote Sensing*, 154, 83-95. <https://doi.org/10.1016/j.isprsjprs.2019.06.011>

Main-Knorn, M., et al. (2017). Sen2Cor for Sentinel-2. *Remote Sensing*, 9(6), 556.

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, 39(9), 2784-2817.

Maxwell, S. L., Fuller, R. A., Brooks, T. M., & Watson, J. E. (2018). The ravages of climate change: A global assessment of the impact of climate change on biodiversity and ecosystem services. *Nature Climate Change*, 8, 238-244. <https://doi.org/10.1038/s41558-018-0082-2>

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247-259.

Pettorelli, N., Laurance, W. F., O'Brien, T. G., & Wegmann, M. (2014). Satellite remote sensing for applied ecologists: opportunities and challenges. *Journal of Applied Ecology*, 51(4), 835-848. <https://doi.org/10.1111/1365-2664.12260>

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81-106. <https://doi.org/10.1007/BF00116251>

Safavian, S. A., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(3), 660-674. <https://doi.org/10.1109/21.97458>

Shakhnarovich, G., Darrell, T., & Indyk, P. (2005). *Nearest-neighbor methods in learning and vision: Theory and practice*. MIT Press.

Shao, G., & Lunetta, R. S. (2012). Assessing the performance of remote sensing classifiers for land cover mapping. *International Journal of Remote Sensing*, 33(13), 4106-4133. <https://doi.org/10.1080/01431161.2011.645957>

Toth, B., & Józków, G. (2016). Remote sensing for vegetation monitoring in tropical and subtropical regions: A review of techniques. *Remote Sensing*, 8(5), 1-17. <https://doi.org/10.3390/rs8050395>

Van der Walt, S., et al. (2011). The NumPy array: A structure for efficient numerical computation. *Computing in Science & Engineering*, 13(2), 22-30.

Woodcock, C. E., Song, C., & Sohl, T. L. (2020). Land cover classification using remote sensing: A review of algorithms and methods. *Remote Sensing of Environment*, 241, 1-19. <https://doi.org/10.1016/j.rse.2020.111659>

Zhang, H. (2004). The optimality of naive Bayes. *AAAI*, 1(2), 3-9.

Zhang, H., & Roy, D. P. (2017). The validation of land cover and land use products using global reference data. *Remote Sensing*, 9(3), 301. <https://doi.org/10.3390/rs9030301>