

**TÉCNICAS DE CLASIFICACIÓN  
SEMI-SUPERVISADA INTEGRANDO LA  
INFORMACIÓN ESPACIAL Y ESPECTRAL PARA  
EL PROCESAMIENTO DE IMÁGENES  
MULTIESPECTRALES DE ZONAS AGRÍCOLAS  
CAPTURADAS SOBRE EL ORIENTE  
ANTIOQUEÑO**

**CAMILO ACEVEDO CORREA**

MAESTRÍA EN AUTOMATIZACIÓN Y CONTROL INDUSTRIAL

DIRECTORES:

MARIA CONSTANZA TORRES MADROÑERO  
RICARDO FRANCO CEBALLOS



**Institución Universitaria**

INSTITUTO TECNOLÓGICO METROPOLITANO  
FACULTAD DE INGENIERÍAS  
MEDELLÍN, 2024

# TÉCNICAS DE CLASIFICACIÓN SEMI-SUPERVISADA INTEGRANDO LA INFORMACIÓN ESPACIAL Y ESPECTRAL PARA EL PROCESAMIENTO DE IMÁGENES MULTIESPECTRALES DE ZONAS AGRÍCOLAS CAPTURADAS SOBRE EL ORIENTE ANTIOQUEÑO

**Autor:** Camilo Acevedo Correa

**Directores:** Maria Constanza Torres Madroñero.  
Ricardo Franco Ceballos.

**Facultad:** Facultad de Ingenierías.

**Titulación:** Maestría en Automatización y Control Industrial.

**Palabras clave:** Espacial – espectral, clasificación, semi-supervisado, percepción remota, multiespectral, pseudo etiquetas, grafos convolucionales.

## Resumen

Los algoritmos de aprendizaje automático han mostrado lo fuerte que son al momento de procesar imágenes multiespectrales y al tener varias muestras etiquetadas.

El problema radica en que el procesamiento de las muestras tiende a ser computacionalmente costoso y presenta limitaciones. Para disminuir el número de muestras etiquetadas y reducir el costo de procesamiento, en el estado del arte se cuenta con el aprendizaje semi-supervisado que aprende a partir de muestras etiquetadas y no etiquetadas. El aprendizaje semi-supervisado permite la clasificación de elementos con un limitado número de muestras, al contener esta característica facilita el aprendizaje de modelos de inteligencia artificial. En particular, este aprendizaje provee métodos de procesamiento de datos como las Redes Prototípicas de Grafos Semi-Supervisadas (SSGPN) y el framework de Generación de Pseudo Etiquetas Probabilístico (Res-CP). El primero emplea una red de grafos probabilística con capas prototípicas que mejora la discriminación entre clases y la acentuación entre las mismas, optimizando los bordes de las imágenes. Además, la inmersión de los grafos (G-Norm) ayuda a acelerar la convergencia y disminuir el costo computacional. El segundo método, se

centra en la generación de pseudo-etiquetas confiable mediante un enfoque probabilístico que administra la incertidumbre, mejorando la generalización de redes neuronales con pocos datos etiquetados. Este trabajo explora las técnicas semi-supervisadas anteriores, añadiéndole extracción de características espaciales resaltando regiones de imágenes multiespectrales como Indian Pines y del Oriente Antioqueño para la clasificación de materiales de la imagen con pocas muestras de entrenamiento.

# **SEMI-SUPERVISED CLASSIFICATION TECHNIQUES INTEGRATING SPATIAL AND SPECTRAL INFORMATION FOR THE PROCESSING OF MULTISPECTRAL IMAGES OF AGRICULTURAL AREAS CAPTURED OVER EASTERN ANTIOQUIA EASTERN ANTIOQUIA**

**Author:** Camilo Acevedo Correa.

**Supervisor:** Maria Constanza Torres Madroñero.  
Ricardo Franco Ceballos.

**Department:** Faculty of Engineering.

**Degree:** Master's Degree in Automation and Industrial Control.

**Keywords:** Spatial-spectral, classification, semi-supervised, remote sensing, multispectral, pseudo labels, convolutional graphs.

## **Abstract**

Machine learning algorithms have demonstrated their strength in processing multispectral images, especially when several labeled samples are available. However, processing such samples tends to be computationally expensive and limited. To reduce the number of labeled samples and decrease processing costs, the state of the art includes semi-supervised learning, which learns from both labeled and unlabeled samples. Semi-supervised learning allows for the classification of elements with a limited number of samples, and this characteristic facilitates the learning of artificial intelligence models. In particular, this type of learning provides data processing methods such as Semi-Supervised Graph Prototypical Networks (SSGPN) and the Probabilistic Pseudo-Label Generation framework (Res-CP). The former employs a probabilistic graph network with prototypical layers that enhances class discrimination and accentuation, optimizing the edges of images. Additionally, the graph embedding (G-Norm) helps accelerate convergence and reduce computational costs. The latter method focuses on the generation of reliable pseudo-labels through a probabilistic approach that manages uncertainty, improving the generalization of neural networks with few labeled data. This work explores the aforementioned semi-supervised techniques, adding spatial feature

extraction to highlight regions of multispectral images such as Indian Pines and those from Oriente Antioqueño for to classify image materials with few training samples.

# Agradecimientos

Ésta tesis de maestría fue posible a muchas personas que creyeron en mi, como mi por eso quiero agradecerle a Maria Torres que fue mi directora en esta tesis y me propició de conocimiento para poder desarrollar el trabajo, también agradecer a los profesores que durante el posgrado me brindaron las asesorías pertinentes al proyecto. Por último, agradecer a mi familia por el apoyo.

Adicionalmente, este trabajo fue posible por la financiación recibida a través del programa 475-2020 Sistemas Inteligentes para el Monitoreo de Cultivos Permanentes y Transitorios financiado por Minciencias, y las horas cátedra proporcionadas por la Facultad de Ingenierías del Instituto Tecnológico Metropolitano.

# Acrónimos

<b>NN</b>	Red Neuronal (Neural Network)
<b>RGB</b>	Estándar Rojos Verde Azul
<b>ROI</b>	Region de Interés
<b>SVM</b>	Máquina de Soporte Vectorial (Support Vector Machine)
<b>DNN</b>	Red Neuronal Profunda (Deep Neural Network)
<b>RCN</b>	Red Residual Convolutacional (Residual Convolution Network)
<b>PLG</b>	Generador de Pseudo Etiquetas (Pseudo Label Generator)
<b>SSRN</b>	Red Residual Espectral - Espacial
<b>SSGPN</b>	Redes gráficas prototípicas semisupervisadas
<b>OAC</b>	Precisión Global (Overall Accuracy)

# Tabla de Contenido

<b>Resumen</b>	<b>II</b>
<b>Abstract</b>	<b>IV</b>
<b>Agradecimientos</b>	<b>VI</b>
<b>Acrónimos</b>	<b>VII</b>
<b>1 Introducción</b>	<b>1</b>
1.1 Justificación . . . . .	1
1.2 Problema de Investigación . . . . .	3
1.3 Objetivos . . . . .	4
1.3.1 General . . . . .	4
1.3.2 Específicos . . . . .	4
1.4 Estructura del documento . . . . .	4
1.5 Contribuciones . . . . .	5
<b>2 Estado del Arte y Marco Teórico</b>	<b>6</b>
2.1 Percepción remota . . . . .	7
2.2 Imagenología espectral . . . . .	8
2.3 Procesamiento de imágenes espectrales . . . . .	9
2.3.1 Desmezclado lineal . . . . .	9
2.3.2 Detección de objetivos . . . . .	9
2.3.3 Clasificación . . . . .	10

2.3.4	Aprendizaje semi-supervisado y procesamiento de imágenes espectrales . . . . .	11
2.3.5	Aprendizaje espacial-espectral y procesamiento de imágenes espectrales. . . . .	13
<b>3</b>	<b>Metodología</b>	<b>18</b>
3.1	Base de Datos . . . . .	18
3.2	Técnicas semi-supervisadas . . . . .	19
3.2.1	Pseudo etiquetas . . . . .	21
3.2.2	Clasificación Semi-supervisada Basada en grafos . . . . .	25
3.3	Extracción de Muestras de Entrenamiento y Pruebas . . . . .	27
3.4	Extracción de características espaciales . . . . .	28
3.4.1	Entropía . . . . .	28
3.4.2	Operadores morfológicos . . . . .	30
3.4.3	Nombramiento de los conjuntos de datos . . . . .	32
<b>4</b>	<b>Resultados y Discusión</b>	<b>33</b>
4.1	Resultados obtenidos con Indian Pines . . . . .	34
4.1.1	<b>A. Configuración de hiperparámetros y criterios de evaluación:</b> . . . . .	34
4.2	Resultados obtenidos con imagen del Oriente Antioqueño . . . . .	42
4.3	Comparación de los métodos semi-supervisados Res-CP y SSGPN	45
	<b>Conclusiones y Trabajo Futuro</b>	<b>48</b>
	<b>Bibliografía</b>	<b>54</b>

# Índice de figuras

2.1	Mapa conceptual de la taxonomía de la clasificación semi-supervisada. Adaptado de [1]. . . . .	11
2.2	Métodos de clasificación espacial-espectral de preprocesamiento, integrado y posprocesamiento. Adaptado de [2]. . . . .	14
2.3	Ejemplo de clasificación espacial-espectral. Adaptado de [3]. . . . .	15
3.1	Banda de muestra de Indian Pines lado izquierda y mapa de clasificación lado derecho. . . . .	19
3.2	A la izquierda imagen Las Palmas capturada por Landsat 8. Derecha: Mapa de clasificación. . . . .	20
3.3	Diagrama de clasificación de la metodología propuesta. Adaptado de [4] . . . . .	23
3.4	Diagrama de codificación de características mediante una Red Convolutiva Residual. Adaptado de [5] . . . . .	24
3.5	Diagrama de generación de Pseudoetiquetas. Adaptado de [5] . . . . .	24
3.6	Diagrama de metodología de la Red Prototípica basada en grafos. . . . .	26
3.7	Diagrama de bloques de extracción de características espaciales usando selección de bandas, obtención de imágenes MSI = 9 bandas espectrales, 18 bandas con información espectral y entropía (9 bandas espectrales y 9 bandas con entropía ), 18 bandas con información espectral y operadores morfológicos (9 bandas espectrales y 9 bandas con operadores morfológicos) y 27 bandas con información espectral, entropía y operadores morfológicos (9 bandas espectrales, 9 bandas con entropía y 9 bandas con operadores morfológicos). . . . .	29
3.8	Ejemplo de metodología usando la entropía. Imagen MSI sin aplicar entropía. Al lado derecho MSI con entropía. . . . .	30

---

3.9	Ejemplo de metodología usando los operadores morfológicos. . .	31
4.1	(a) Matriz de confusión de Indian Pines usando Res-Cp con información espectral y 20 muestras por clase. (b) Matriz de confusión de Indian Pines usando Res-Cp con información espacial-espectral y 20 muestras por clase. . . . .	37
4.2	(c) Matriz de confusión de Indian Pines usando SSGPN con información espectral.(d) Matriz de confusión de Indian Pines usando SSGPN con información espacial-espectral. . . . .	41
4.3	Gráfico de barras de comparación de precisión entre SSGPN y Res-CP. . . . .	47

# Índice de Tablas

3.1	Landsat-8 Bandas . . . . .	20
3.2	Tabla de los nombres de los dataset de imágenes multiespectrales creadas con la información espacial. IP equivale a Indian Pines y OA a Oriente Antioqueño 9, 18 y 27 indican el número de canales que contiene cada imagen multiespectral. . . . .	32
4.1	Número de pseudo-muestras generadas por PLG . . . . .	35
4.2	Parámetros de Ejecución del Script Python de Res-CP . . . . .	36
4.3	Resultados del experimento usando SSRN de Res-CP con las imágenes multiespectrales combinadas. . . . .	38
4.4	(a) En la parte superior mapa de clasificación. Matriz de imágenes por categoría y número de muestras . . . . .	39
4.5	Etiquetas de Indian Pines . . . . .	40
4.6	Configuración de Hiperparámetros del X-GPN . . . . .	42
4.7	Resultados del experimento usando X-GPN con cada una de las imágenes multiespectrales combinadas . . . . .	42
4.8	Resultados de Indian Pines: (a) IP_9, (b) IP_18, (c) IP_18_m, (d) IP_27 y (e) mapa de clasificación . . . . .	42
4.9	Comparación de las Matrices de Confusión Generadas con Diferentes Configuraciones Espectrales . . . . .	43
4.10	Resultados de Indian Pines: (a) Mapa de clasificación OA (Oriente Antioqueño), (b) Resultado OA con Res-CP , (c) Resultado de OA con SSGPN . . . . .	44
4.11	Comparativa de métricas para diferentes OA usando SSGPN y Res-Cp. . . . .	44

---

4.12 Comparación entre SSGPN y Res-CP para la clasificación de imágenes multiespectrales. . . . .	45
4.13 Comparativa de métricas para diferentes IPs usando ssgpn y res-cp.	46

# Capítulo 1

## Introducción

### Contenido

---

<b>1.1 Justificación</b> . . . . .	<b>1</b>
<b>1.2 Problema de Investigación</b> . . . . .	<b>3</b>
<b>1.3 Objetivos</b> . . . . .	<b>4</b>
1.3.1 General . . . . .	4
1.3.2 Específicos . . . . .	4
<b>1.4 Estructura del documento</b> . . . . .	<b>4</b>
<b>1.5 Contribuciones</b> . . . . .	<b>5</b>

---

### 1.1. Justificación

La percepción remota es una ciencia que integra modelos, técnicas y sistemas para la obtención de datos a distancia, es decir, sin tener contacto directo con los objetos muestreados [6]. Una de las aplicaciones de la percepción remota, también denominada teledetección, es la observación de la tierra, en la cual se emplean aeronaves no tripuladas, aviones, o satélites, integrando sensores pasivos o activos, que permiten medir a distancia características físicas de una superficie u objeto en la Tierra [7]. Dentro de las aplicaciones de observación de la Tierra se encuentran áreas de investigación como la agricultura de precisión, monitoreo ambiental, cambio climático, minería, planeación territorial, entre otros [8].

Particularmente en la agricultura de precisión, se reconoce la percepción remota como una de las tecnologías con mayor potencial de desarrollo para el

monitoreo de cultivos. Esto se debe a la posibilidad de medir grandes extensiones de tierra en poco tiempo. La combinación de los datos remotos y de técnicas de Big Data e inteligencia artificial, abren una posibilidad para el desarrollo de nuevas tecnologías que permita conocer el estado de un cultivo, identificar de forma temprana enfermedades, estimar la productividad, y analizar el efecto de los cultivos en las regiones [7]. Los datos capturados por sistemas remotos y la información extraída mediante inteligencia artificial puede soportar la toma de decisiones del agricultor con respecto a las variables que afectan su cultivo, tales como tipo de cultivo a plantar, fertilizantes o cantidad de pesticidas a emplear. Este tipo de decisiones van ligadas a fuentes de información, ya que necesitan un análisis sofisticado, porque afectarán directamente la calidad del cultivo [9].

La percepción remota en combinación con la inteligencia artificial brinda una posibilidad de desarrollo de tecnologías para ser implementadas en el sector agroindustrial en Colombia, donde se reconoce la falta de tecnificación. El Plan Estratégico de Ciencia, Tecnología e Innovación del Sector Agropecuario Colombiano – PECTIA 2017-2027 [10] establece la necesidad investigar desde los campos de la ingeniería y las tecnologías de la información y comunicación (TICs), para la búsqueda de soluciones para el agro. Por su parte, la investigación de las tecnologías de la información y la comunicación dentro de la agricultura se priorizan como una de las áreas de investigación de la ciudad y el área metropolitana en la Agenda Regional de I+D+i elaborada por G8 en 2020, donde se incluye las tareas de tecnificación del agro, territorios inteligentes para la vida, y tecnologías informáticas habilitadoras de la industria 4.0 [11].

Particularmente, es de interés en esta tesis de maestría contribuir en el desarrollo de metodologías de procesamiento y análisis de datos remotos. La percepción remota aplicada a la agricultura de precisión puede emplear sensores pasivos, que incluyen cámaras multiespectrales o hiperespectrales. Las imágenes multiespectral o hiperespectrales miden la energía reflejada o la radiación electromagnética en diferentes longitudes de onda [7]. Las imágenes multiespectrales capturan información tanto espectral como espacial de un objeto o superficie. El análisis de este tipo de datos converge en diferentes tareas de aprendizaje de máquina, como son la clasificación, el desmezclado lineal y la detección de objetivos. Actualmente, como se presenta en la revisión de literatura, existe un creciente interés en investigación por técnicas de clasificación semi-supervisadas, que emplean pocos datos etiquetados para entrenar un sistema que reconoce a que clase pertenece cada píxel de una imagen [12]. La importancia de las técnicas semi-supervisadas radica en la dificultad de adquirir datos etiquetados para grandes extensiones de tierra. Por otra parte, existe una tendencia al desarrollo de nuevos algoritmos que integren tanto la información espectral como espacial que son adquiridas por las imágenes, ya que varios trabajos

han demostrado la superioridad de esta combinación de información [13], sobre clasificadores espectrales. Por lo anterior, esta tesis de maestría está orientada a explorar técnicas semi-supervisadas de imágenes espectrales, integrando la información espacial y espectral, y su evaluación en datos capturados sobre cultivos agrícolas.

Este tesis de maestría hace parte del programa 475-2020 Sistemas Inteligentes para el Monitoreo de Cultivos Permanentes y Transitorios financiado por Minciencias, y en el cual participan el Instituto Tecnológico Metropolitano - ITM como entidad ejecutora, AGROSAVIA - Corporación Colombiana de Investigación Agropecuaria, y Black Square.

## 1.2. Problema de Investigación

Las imágenes multiespectrales e hiperespectrales están siendo ampliamente utilizadas en diferentes aplicaciones, como la agricultura, minería, observación de la tierra, urbanismo, entre otros. Una de las tareas de procesamiento de imágenes espectrales más relevante es la clasificación, en la cual se busca asignar una etiqueta a cada píxel de la imagen [14] que proporcione información sobre los componentes de una superficie. Clasificar este tipo de imágenes ha sido crucial para promover el uso de la percepción remota y se ha logrado un avance importante por el interés que se le ha presentado [15].

Usualmente, la clasificación de imágenes espectrales emplea los métodos tradicionales de aprendizaje de máquina, tales como máquinas de soporte vectorial, bosques aleatorios, y redes neuronales [16]. Los cuales en su mayoría son supervisados y realizan una clasificación espectral, es decir, píxel por píxel. Los métodos de clasificación supervisado requieren una gran cantidad de datos espectrales etiquetados para el entrenamiento del sistema. Sin embargo, obtener las etiquetas es un trabajo tedioso, porque requiere trabajo de campo y mano de obra. Además, los clasificadores supervisados pueden transigir de muestras o etiquetas erróneas debido a un reajuste de las etiquetas. Específicamente, para el procesamiento de imágenes espectrales remotas, la obtención de las etiquetas es más complicado, por tanto, se requiere de algoritmos que utilicen con pocas muestras etiquetadas para el entrenamiento de los sistemas [14].

De ahí, se encuentran en la literatura con métodos semi-supervisados para este tipo de imágenes [17]. Sin embargo, la mayoría de ellos están orientados a la clasificación espectral, es decir teniendo en cuenta únicamente la información a lo largo de las longitudes de onda. Las imágenes espectrales además de capturar la información espectral también contienen información espacial, que en

mucho de los procesamientos es desaprovechada [18]. Por lo tanto, es necesario el desarrollo de metodologías para la clasificación semi-supervisada de imágenes espectrales que incorporen la información espacial, considerando que esta información puede permitir obtener mejores resultados.

Lo que se quiere explorar en este trabajo es como combinar métodos semi-supervisados que se encuentran la literatura con enfoques de clasificación espacial-espectral. Se busca simplificar el trabajo que se tiene al etiquetar los datos y optimizar el proceso de clasificación.

## **1.3. Objetivos**

### **1.3.1. General**

Desarrollar una metodología de clasificación integrando técnicas semi-supervisadas con enfoques espaciales espectrales para el procesamiento de imágenes multi-espectrales de zonas agrícolas capturadas sobre el oriente Antioqueño.

### **1.3.2. Específicos**

- Comparar técnicas semi-supervisada para la clasificación de imágenes multiespectrales.
- Integrar la información espacial a la técnica semi-supervisada para la clasificación de imágenes multiespectrales en zonas agrícolas capturadas sobre el oriente Antioqueño.
- Comparar el desempeño de clasificación de la metodología propuesta con técnicas supervisadas o semi-supervisadas que emplean solamente la información espectral.

## **1.4. Estructura del documento**

Cada uno de los siguientes capítulos se presenta la siguiente estructura:

- Capítulo 1: Describe los objetivos, justificación problema de investigación y por último las contribuciones generadas con esta investigación.

- Capitulo 2: Presenta el estado del arte y marco teórico. Comenzado con percepción remota, imagenología espectral y procesamiento de imágenes espectrales.
- Capitulo 3: Expone la metodología de investigación aplicada. Exhibiendo las bases de datos utilizadas, métodos de aprendizaje de máquinas, técnicas de visión artificial y la descripción de la metodología.
- Capitulo 4: Muestra los resultados obtenidos de la investigación según los capítulos anteriores.
- Capitulo 5: Concluye los resultados de la investigación.

## 1.5. Contribuciones

- Torres-Madronero, M. C., Goez, M., Guzman, M. A., Rondon, T., Carmo-  
na, P., Acevedo-Correa, C., ... Casamitjana, M. (2022). Spectral Library of  
Maize Leaves under Nitrogen Deficiency Stress. *Data*, 8(1), 2.
- Acevedo-Correa, C.; Goez, M.M.; Torres-Madronero, M.C.; Rondon, T. Low-  
cost clamp for the measurement of vegetation spectral signatures. *Hardwa-  
reX* 2022, 1–9.

# Capítulo 2

## Estado del Arte y Marco Teórico

### Contenido

---

<b>2.1</b>	<b>Percepción remota</b>	<b>7</b>
<b>2.2</b>	<b>Imagenología espectral</b>	<b>8</b>
<b>2.3</b>	<b>Procesamiento de imágenes espectrales</b>	<b>9</b>
2.3.1	Desmezclado lineal	9
2.3.2	Detección de objetivos	9
2.3.3	Clasificación	10
2.3.4	Aprendizaje semi-supervisado y procesamiento de imágenes espectrales	11
2.3.5	Aprendizaje espacial-espectral y procesamiento de imágenes espectrales.	13

---

### Sinopsis

En este capítulo se aborda todo lo relacionado con imágenes espectrales donde se comienza con una introducción a la percepción remota en la sección 2.1. Aquí se revisan el uso de los instrumentos de teledetección con los que se obtienen los datos sobre la superficie terrestre mediante plataformas satélites.

En la sección 2.2, se profundiza la imagenología espectral, una técnica de captura de información en múltiples bandas del espectro electromagnético. Se habla de metodologías disponibles en este campo para proporcionar detalles

precisos y específicos de las imágenes multi e hiperespectrales a diferencias de las imágenes convencionales.

En la sección 2.3 se habla sobre el procesamiento de las imágenes espectrales, comenzando por el desmezclado lineal en la subsección 2.3.1, donde este proceso descompone las imágenes en componentes básicos, permitiendo una mejor interpretación de los datos.

La subsección 2.3.2 comprende la detección de objetivos, una técnica de identificación y localización de objetos dentro de una imagen espectral.

En la subsección 2.3.3, se presenta el método de clasificación, donde este categoriza los píxeles de una imagen en diferentes clases basadas en características en longitud de onda.

Por último, la sección 2.3.4 aborda algunos de los métodos semi-supervisados y su aplicación en el procesamiento de imágenes espectrales. Se analiza cómo esta técnica combina datos etiquetados y no etiquetados al momento de ser procesados en un algoritmo de inteligencia artificial y mejorar la precisión del análisis espectral.

Este capítulo muestra las técnicas y métodos usados en el procesamiento de imágenes espectrales, destacando su manejo y procesamiento para la aplicación en diversas disciplinas.

## 2.1. Percepción remota

La percepción remota es una ciencia que usa modelos, técnicas y sistemas para la caracterización de un objeto o superficie sin tener un contacto directo. Entre las áreas de la percepción remota se encuentra la observación de la tierra, que emplea sistemas integrados en aeronaves no tripuladas, satélites y aviones, para capturar información sobre la superficie terrestre [6]. Los sensores usados para la captura de datos se dividen en dos: pasivos y activos. Los sensores pasivos recolectan información generada por la interacción de una fuente de energía externa y una superficie, es decir, estos sensores no tienen su propia fuente de energía, en esta categoría se encuentran los sensores ópticos multiespectrales e hiperespectrales. Por otra parte, los sensores activos son los que usan una fuente de energía propia para una superficie, ejemplo de estos son los sensores de radar o LIDAR [7]. Los sistemas de percepción remota utilizan tecnologías de la información para poder recolectar y procesar grandes volúmenes de datos espectrales, esta información suele ser densa debido a las altas resoluciones espaciales – espectrales [2]. Para esto se han implementado técnicas de procesa-

miento de datos de inteligencia artificial, Big Data y aprendizaje de máquina [19]. Entre las aplicaciones de la percepción remota para la observación de la tierra se encuentran la agricultura de precisión, el monitoreo ambiental, estudios de cambio climático, planeación territorial, minería, y defensa [6].

## 2.2. Imagenología espectral

Las cámaras multispectrales e hiperespectrales miden la energía reflejada o emitida por una superficie a lo largo del espectro electromagnético. Usualmente este tipo de dispositivos captura longitudes de ondas desde el rango visible al infrarrojo cercano. Las imágenes multispectrales se caracterizan por capturar decenas de bandas, no continuas, a lo largo el espectro electromagnético [8, 20]. Por su parte, las imágenes hiperespectrales capturan cientos de bandas o canales, de forma casi continua a lo largo del espectro [18]. Las imágenes espectrales (multi o hiperespectrales) permiten conocer características físicas de la superficie a partir de la interacción con la luz [14]. Las imágenes espectrales se componen de varios píxeles que brindan información espacial. De esta forma, una imagen espectral se puede visualizar como un cubo, con dos dimensiones espaciales y una dimensión que corresponde al espectro. El procesamiento de las imágenes espectral busca discriminar, clasificar, caracterizar o cuantificar los elementos o materiales presentes en la superficie.

Dado que las imágenes espectrales capturan información espacial y espectral, en diferentes periodos de tiempo de acuerdo con la plataforma que se está empleando, existen diferentes resoluciones a considerar en el momento de seleccionar el tipo de cámara. Principalmente se deben considerar cuatro tipos de resoluciones: espacial, espectral, radiométrica y temporal. La resolución espacial está relacionada con la capacidad de la imagen para permitir la diferenciación de dos objetos que están cercanos entre sí, es decir, es la mínima distancia entre dos objetos o píxeles diferentes separados en una imagen [8]. Comúnmente, la resolución espacial esta relaciona con el tamaño del píxel: que porción de la superficie es capturado en un píxel. Por otra parte, la resolución espectral está relacionada al número de bandas y anchos de banda. La resolución radiométrica se define por el número de bits empleados para representar la cuantificación de la intensidad de la reflectancia; entre mayor número de bits, mayor es nivel de cuantificación. Finalmente, la resolución temporal se relaciona con la frecuencia de muestreo de una misma superficie a lo largo del tiempo. De acuerdo con la aplicación y el problema que se busca resolver, las resoluciones espaciales, espectrales, radiométricas y temporales afectan en diferente medida las capacidades de un sistema [8].

## 2.3. Procesamiento de imágenes espectrales

El procesamiento de imágenes espectrales converge en tres tareas: detección de objetivos, clasificación, y desmezclado lineal.

### 2.3.1. Desmezclado lineal

El desmezclado en una imagen hiperespectral es el proceso de separación de la información espectral de un píxel en componentes básicos denominados endmembers. Los endmembers son una colección de firmas que representan los materiales que constituyen una imagen. El análisis de desmezclado lineal busca estimar estas firmas desde los datos espectrales, así como calcular su abundancia. La abundancia de un endmember representa el porcentaje de ese elemento o componente en cada píxel [21, 22].

Los modelos de desmezclado se basan según el tipo de mezcla. Uno de los modelos más empleados es el mezclado lineal [23]. Este modelo asume que la mezcla de los elementos o materiales es microscópica en la incidencia de la luz que interactúa con un solo material, como en caso de una escena tipo tablero de ajedrez. Dicho lo anterior la mezcla ocurre dentro del instrumento de medición, debido a que la resolución del instrumento de medición no es suficiente. En el mezclado lineal la luz que pasa directamente por el material se refleja de manera separada hasta llegar al sensor, representando el mezclado lineal [21].

### 2.3.2. Detección de objetivos

En la detección de objetivos se requiere localizar objetos o materiales dentro de una escena o campo para después cuantificar, localizar o identificar [24]. Personas de algunas áreas como la geología, científicos de climas y entre otros usan este tipo de tareas para identificar y vigilar especies o en el caso del clima para localizar fuentes de metano o dióxido de carbono [25].

Adicionalmente, el tipo de procesamiento puede clasificarse también de acuerdo con la información analizada en espectral o espacial-espectral. En el procesamiento espectral no se tiene en cuenta la relación espacial de los píxeles; en este tipo de procesamiento, el cubo se transforma en una matriz bidimensional, donde las filas representan las muestras (píxeles en el cubo), y las columnas las características (bandas espectrales). En esta categoría recaen la mayoría de los algoritmos propuestos en la literatura, considerando que varios de ellos son la aplicación de técnicas de aprendizaje de máquinas para el análisis de imáge-

nes espectrales. Por otra parte, el procesamiento espacial-espectral considera la relación existen entre un píxel y sus vecinos [18]. Este tipo de procesamiento, también denominado orientado a objetos, ha demostrado mejores desempeños que las técnicas espectrales.

### 2.3.3. Clasificación

La clasificación de una imagen es la tarea de asignar etiquetas a cada uno de los píxeles. La percepción remota ha enriquecido los avances para la captura de datos espaciales y espectrales, gracias a la implementación de sensores de alta resolución. En específico, teniendo una mayor resolución espectral de imágenes hiperespectrales suministra una nueva ejecución en la indagación de datos. Dicho lo anterior, la clasificación de imágenes hiperespectrales tiende a ser compleja debido a la limitación de las muestras de entrenamiento que se conoce como el fenómeno de Hughes o maldición de la dimensionalidad [14]. Existen diferentes tipos de aprendizajes, entre ellos están el supervisado, no supervisado y el semi-supervisado.

El aprendizaje supervisado requiere de datos de entrenamiento, es decir, píxeles etiquetados, que permitan entrenar un sistema. Una vez entrenado, dado un conjunto de datos que se ingresan por una entrada, el sistema retorna las etiquetas de la imagen [1]. Por tanto, en el aprendizaje supervisado se debe tener un conocimiento previo de los datos, lo cual usualmente requiere la recolección de información en campo; para varias aplicaciones de percepción remota, el etiquetado de datos es difícil [14].

En este tipo de aprendizaje no se tienen valores etiquetados en la imagen, es decir no se conocen los posibles valores de la salida [1]. El aprendizaje no supervisado buscar inferir un patrón a partir de los datos de entrada. Por lo cual, ese tipo de aprendizaje no requiere intervención humana. Generalmente se utiliza para extraer y segmentar características ayudando a encontrar grupo de datos [14]. La mayor dificultad con este tipo de procesos es la validación de los resultados, ya que no cuenta con información a priori.

Este aprendizaje, normalmente intenta mejorar uno de los dos aprendizajes definidos en los párrafos anteriores por que usan los datos etiquetados y no etiquetados para el modelo de procesamiento de datos. Al momento de procesar los datos se abordan alternativas como: (I) en un problema de clasificación se tiene en cuenta los puntos adicionales donde la etiqueta se desconoce y así se mejora el proceso de clasificación; o (II) a partir de métodos de agrupación, el aprendizaje del sistema depende de una etiqueta que pertenezca a cada uno de las clases existentes o etiquetadas [1].

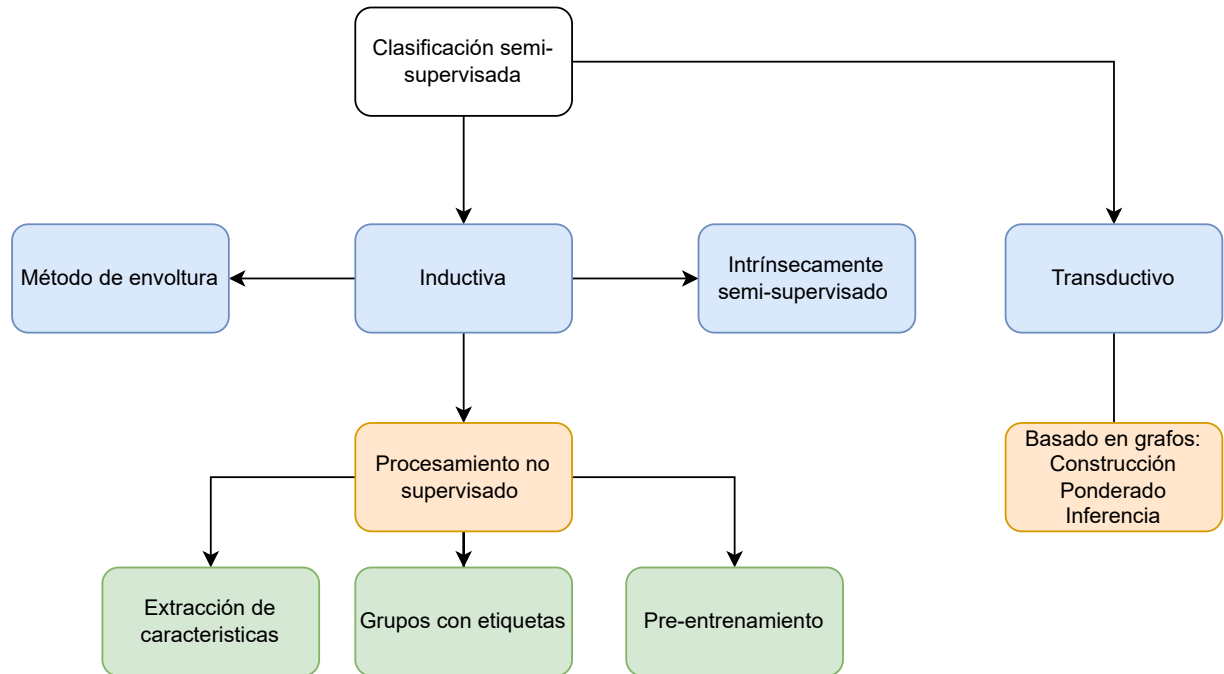


Figura 2.1: Mapa conceptual de la taxonomía de la clasificación semi-supervisada. Adaptado de [1].

#### 2.3.4. Aprendizaje semi-supervisado y procesamiento de imágenes espectrales

El aprendizaje de máquina en el caso semi-supervisado se centra en métodos de clasificación, donde los datos etiquetados son mínimos. A lo largo del tiempo se han propuesto varios algoritmos de clasificación para el aprendizaje semi-supervisado. Estos algoritmos se diferencian en la manera en que procesan los datos no etiquetados y la relación que tienen con el aprendizaje supervisado. En la figura 2.1 se presenta la taxonomía de los métodos de aprendizaje semi-supervisado [1]. En general, los métodos semi-supervisados se clasifican en inductivos y en transductivos.

En los métodos inductivos se propone crear un modelo de clasificación, mientras que en los métodos transductivos solo se preocupan en obtener predicciones de etiquetas para los datos no etiquetados. Los inductivos generalmente se

basan directamente en los datos no etiquetados ampliando los algoritmos supervisados. Por su parte, los transductivos se centran en grafos [15].

### **Métodos inductivos**

Los métodos inductivos parten directamente del entrenamiento de un modelo de clasificación o regresión, usando datos etiquetados y no etiquetados. Este modelo se utiliza para predecir nuevas muestras no vistas previamente. El objetivo es generalizar a partir del entrenamiento para hacer predicciones precisas sobre cualquier nuevo dato dentro del dominio del problema [15]. Del método inductivo se despliegan los métodos de envoltura o “wrapper”, procesamiento no supervisado y el intrínsecamente semi-supervisado. Los cuales se definen como:

#### *Método de envoltura o “wrapper”*

Este método se encuentra entre uno de los más antiguos y conocidos para el aprendizaje semi-supervisado [26]. Consiste en ajustar un algoritmo supervisado con técnicas de aprendizaje no supervisado. Generalmente, estos métodos usan la preclasificación de los datos no etiquetados usando clustering u otras técnicas de segmentación, seguidas de una fase supervisada donde se ajusta un clasificador a los datos pre-clasificados. En la primera etapa se tienen uno o más clasificadores supervisados y entrenan los datos etiquetados y posiblemente integran datos los pseudo-etiquetados de iteraciones pasadas. En el segundo paso los clasificadores presentes usan los datos para inferir etiquetas para los datos no etiquetados previamente, después de haber sido procesados los datos más confiables son pseudo etiquetados y así se predice una posible etiqueta para los datos no etiquetados [1].

Una ventaja de este método es la versatilidad al emplear los clasificadores de base supervisado. [1]. Con el método de envoltura se han trabajado algunas corrientes tales como el método de auto entrenamiento que consiste en entrenar datos etiquetados y pseudo etiquetados que previamente se etiquetaron con las iteraciones de un clasificador supervisado [1] [27]. Dicho lo anterior, este tipo de métodos se ha usado en clasificación de imágenes multispectrales, enfocados en clasificación semi-supervisada. Por ejemplo en se presentan muestras no etiquetadas y se seleccionan de manera inteligente usando un enfoque de autoaprendizaje.

También en el método de envoltura se tiene una extensión del autoaprendizaje llamado co-entrenamiento, consiste en usar múltiples clasificadores supervisados, algo similar con el auto-entrenamiento. Para que el co-entrenamiento tenga éxito, es relevante que los datos de aprendizajes no estén directamente correlacionados en sus predicciones, debido a que los datos se limitan [28]. El co-

entrenamiento como uno de los derivados del método de envoltura, también es usado para la clasificación de imágenes hiperespectrales, donde la clasificación es inspirada en el TLD (Tracking Learning Detection) [29]. El detector identifica los falsos negativos y los falsos positivos en una imagen hiperespectral, que corresponden a datos no etiquetados. Posteriormente, estos datos convergen con los datos etiquetados para llegar al clasificador espectral [30].

#### *Procesamiento no supervisado*

El procesamiento no supervisado consta de la extracción de características útiles proveniente de una base de datos no etiquetada. Aquí el método de envoltura o “wrapper” no puede ser usado, debido a que requiere de datos conocidos para poder desarrollarse [1].

#### *Métodos intrínsecamente semi-supervisados*

Después de haber definido el algoritmo de aprendizaje inductivo en general, este método enfatiza el aprendizaje de muestras etiquetadas y no etiquetadas. Este no depende de ningún agente relacionado con el aprendizaje supervisado. Por lo general depende de extensiones de los métodos supervisados para los datos etiquetados y no etiquetados. Uno de los métodos que incluye es el del margen máximo [31]. Este clasificador de margen máximo se centra en reducir la distancia de los datos y el límite de decisión [15].

### **2.3.5. Aprendizaje espacial-espectral y procesamiento de imágenes espectrales.**

En la clasificación de imágenes multi e hiperespectrales se encuentra una tendencia hacia el desarrollo de técnicas que integran la información espectral con la espacial. Estos métodos se pueden dividir en tres tipos de clasificaciones que son el preprocesamiento, integrado y el posprocesamiento, como se muestra en la figura 2.2.

*Clasificación basada en preprocesamiento:* En este método la información espacial se extrae en características espaciales para luego ser procesada en el clasificador. El preprocesamiento generalmente se divide en dos fases: 1) Extracción de características espaciales-espectrales para obtener patrones importantes de la imagen hiperespectral y 2) la clasificación de características realizada por clasificadores estándar tales como máquinas de soporte vectorial y regresiones [32]. Dicho lo anterior, la parte de extracción de características es la parte más importante para el rendimiento de las técnicas de clasificación [33].

*Clasificación integrada:* Teniendo en cuenta los conceptos anteriores en este método de clasificación la información espacial y espectral se usan para formar

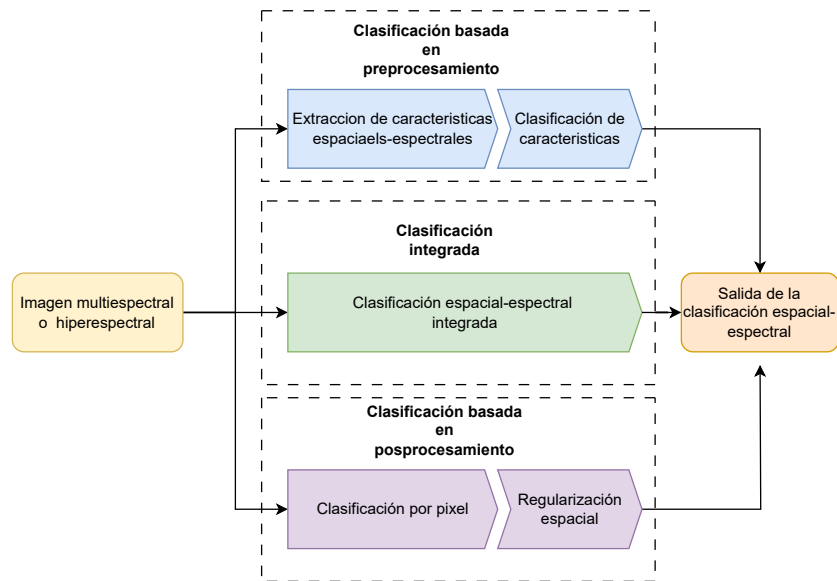


Figura 2.2: Métodos de clasificación espacial-espectral de preprocesamiento, integrado y posprocesamiento. Adaptado de [2].

un clasificador integrado. Esto quiere decir que al momento de utilizar la información espacial y espectral no se separan claramente en el proceso de clasificación [32]. Para la representación de la información se emplean estructuras como *kernels*, tensores, o superpíxeles.

*Clasificación basada en posprocesamiento:* en este método de clasificación se parte de un mapa obtenido desde una clasificación píxel por píxel. Estos métodos asumen que el vecindario de un píxel proporciona información sobre la clase del mismo. Por lo cual, estos métodos incorporan un paso de regularización espacial de los mapas de clasificación. En la regularización espacial, la clase de un píxel específico va a cambiar de acuerdo a la clase más frecuente de su vecindario. Este vecindario se puede determinar por *kernels* espaciales de diferentes tamaños o por regiones obtenidas desde una segmentación [32]) como se muestra en la figura 2.3

### Extracción de características espaciales.

Como se mostró anteriormente, uno de los enfoques para incorporar la información espacial en la clasificación de imágenes multispectrales e hiperespectrales es la extracción de características. Entre las características más em-

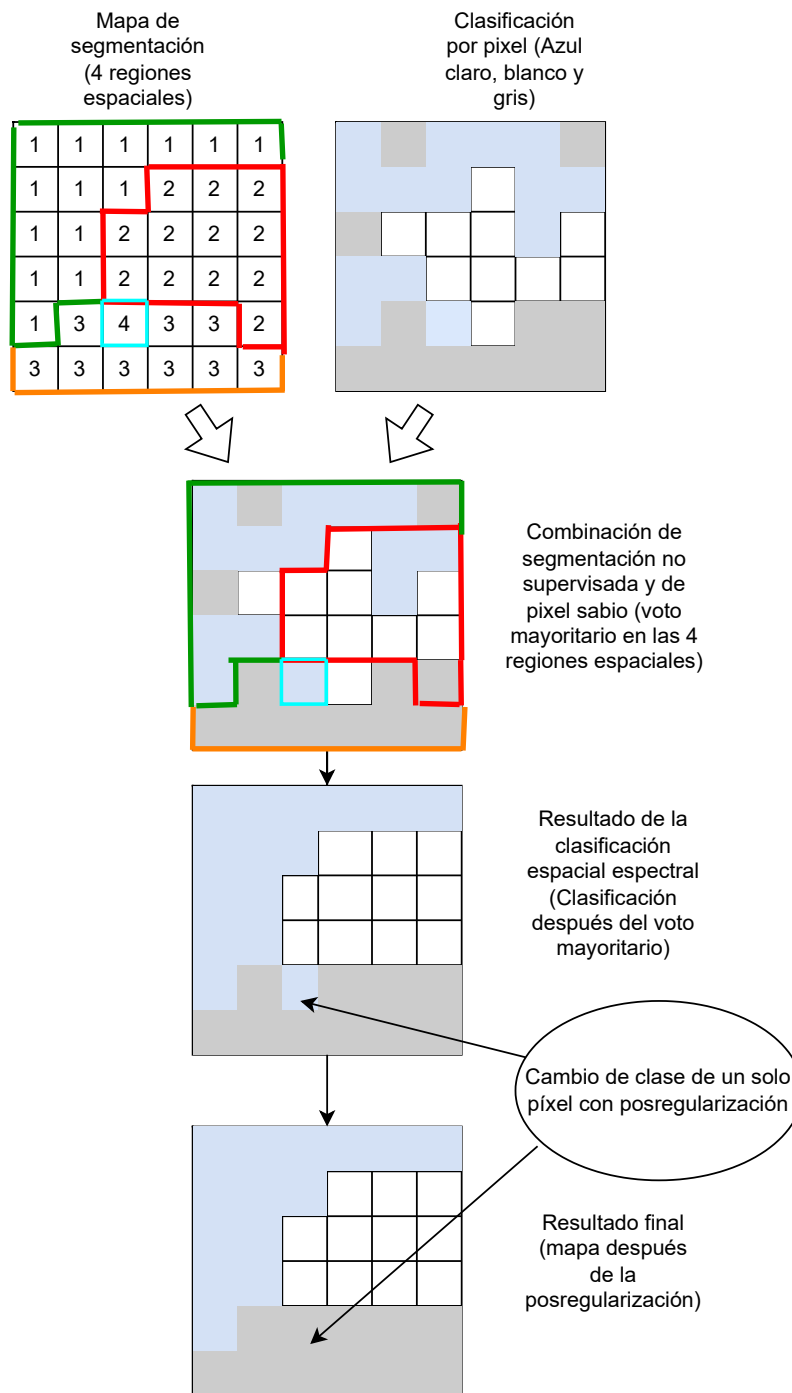


Figura 2.3: Ejemplo de clasificación espacial-espectral. Adaptado de [3].

pleadas para el análisis de este tipo de datos se encuentra la entropía y los coeficientes morfológicos.

*Entropía:* La entropía se encarga de medir la incertidumbre existente en una fuente [34]. En las imágenes se encarga de medir la cantidad de información aleatoria existente en la imagen [35]. Donde una alta entropía significa que la imagen tiene mucho desorden o complejidad, por otro lado, una baja entropía significa que la imagen tiene menos complejidad y es más predecible [36]. Al tener estas definiciones, se puede decir que hay tipos de entropías usadas en el estado del arte, donde una de ellas es la entropía de Shannon que se define como en la ecuación 2.1:

$$H(X) = - \sum_{i=1}^N p(x_i) \cdot \log_b p(x_i) \quad (2.1)$$

Donde:

- $H(X)$  es la entropía de la distribución de probabilidad de la variable aleatoria  $X$ .
- $p(x_i)$  es la probabilidad de que la variable aleatoria  $X$  tome el valor  $x_i$ .
- $N$  es el número de posibles valores que puede tomar la variable aleatoria.
- $b$  es la base del logaritmo utilizado (generalmente  $b = 2$  para bits de información).

Esta es usada para medir la incertidumbre promedio en una distribución de probabilidad [36]. Resaltando los componentes o valores que puede tomar la variable aleatoria.

*Operaciones morfológicas:* Son un conjunto de operaciones usadas para tratar imágenes teniendo en cuenta las formas de los objetos en la imagen [37]. Parecido a los *kernels* o núcleos convolucionales [38], las operaciones morfológicas usan un elemento estructurador para cambiar cada píxel de una imagen en un valor determinado por sus píxeles vecinos. Por lo tanto, existen tipos de operaciones morfológicas que se definen como [39]:

- **Erosión:** Se encarga de disminuir el tamaño de los objetos de una imagen al suprimir los píxeles de los bordes de los objetos.
- **Dilatación:** Esta operación aumenta el tamaño de los objetos en una imagen adicionando píxeles al perímetro del objeto.

- **Apertura:** Esta operación aplica la erosión seguida de la dilatación. Se usa principalmente para eliminar pequeños objetos ruidosos en primera instancia y para separar objetos conectados sin afectar sustancialmente el tamaño de los objetos grandes.
- **Cierre:** Esta operación se efectúa de manera contraria que la apertura. Se realiza una operación de dilatación seguida por una de erosión. Se usa principalmente para cerrar pequeños huecos provenientes de los objetos y para unir objetos próximos a los objetos sin afectar considerablemente a los objetos mas grandes.
- **Apertura por área:** Es una variante de la operación de apertura donde aplica un umbral por área a los objetos que estén por debajo del objetivo. Se usa para eliminar pequeños detalles u objetos en una imagen.
- **Cierre por área:** Es una variante de la operación de cierre, donde aplica a objetos una operación de cierre de agujeros dentro de los objetos característicos de la imagen, sin afectar significativamente el tamaños de los objetos grandes por debajo de un umbral de área.

# Capítulo 3

## Metodología

### Contenido

---

<b>3.1 Base de Datos</b> . . . . .	<b>18</b>
<b>3.2 Técnicas semi-supervisadas</b> . . . . .	<b>19</b>
3.2.1 Pseudo etiquetas . . . . .	21
3.2.2 Clasificación Semi-supervisada Basada en grafos . . . . .	25
<b>3.3 Extracción de Muestras de Entrenamiento y Pruebas</b> . . . . .	<b>27</b>
<b>3.4 Extracción de características espaciales</b> . . . . .	<b>28</b>
3.4.1 Entropía . . . . .	28
3.4.2 Operadores morfológicos . . . . .	30
3.4.3 Nombramiento de los conjuntos de datos . . . . .	32

---

### Sinopsis

Este capítulo presenta la elección de las bases de datos, métodos de aprendizaje semi-supervisados, extracción de características espaciales y finalmente una descripción de la metodología.

#### 3.1. Base de Datos

Este proyecto partirá de dos bases de datos de imágenes hiper o multiespectrales. La primera son imágenes procedentes de las bases de datos Indian

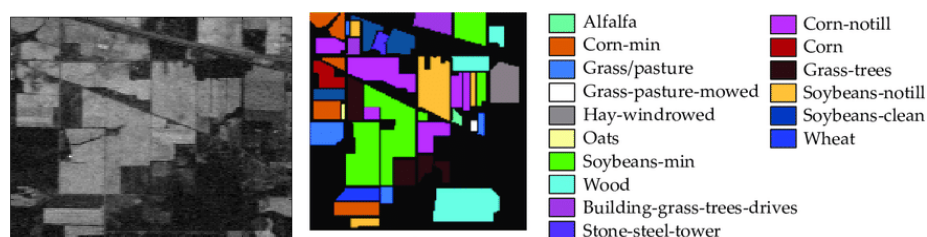


Figura 3.1: Banda de muestra de Indian Pines lado izquierdo y mapa de clasificación lado derecho.

Indian Pines la cual cuenta con imágenes hiperespectral (con más de 100 bandas) que se encuentran disponibles de forma gratuita <sup>1</sup> y que tienen mapas de clasificación (es decir, datos etiquetados). Esta imagen se empleará para realizar la comparación de las técnicas semisupervisadas basadas en pseudo etiquetas y en grafos, dado que, al tener un mapa de clasificación, se pueden calcular matrices de confusión y métricas de desempeño como la precisión general y por clases. Con esta imagen se realizará un submuestreo de la imagen hiperespectral a las 8 bandas espectrales, que corresponden con las bandas capturadas por el satélite Landsat 8. Estas bandas se seleccionan considerando que Landsat 8 es uno de los referentes internacionales en imágenes multiespectrales.

Indian Pines es una imagen hiperespectral capturada por el sensor AVIRIS de la NASA sobre Indiana, Estados Unidos. La imagen cuenta con 145 x 145 píxeles y 220 bandas espectrales, entre los 300 y 2200 nm a lo largo del espectro electromagnético. La Figura 3.1 muestra una banda de muestra de Indian Pines y el mapa de clasificación. Este mapa incluye 16 clases, entre las cuales están diferentes tipos de cultivo de maíz y soja.

Las Palmas es el conjunto de datos que brinda una visión singular al capturar una cuenca hidrográfica en la zona oriental de Antioquia, Colombia. Este conjunto, obtenido por el sensor Landsat8 en 2017, incluye 295x200 píxeles y distingue cuatro clases: áreas boscosas, pastizales, zonas urbanas y suelo desnudo. Su enfoque localizado ofrece valiosas percepciones sobre la dinámica de la cobertura terrestre en la región [40] como se observa en la Figura 3.2.

## 3.2. Técnicas semi-supervisadas

En las siguientes secciones se describen dos técnicas semi supervisadas, donde la primera es basada en pseudo etiquetas y la segunda en grafos convo-

<sup>1</sup>Enlace de descarga de la base de datos: <https://bit.ly/3ACC5SO>

Tabla 3.1: Landsat-8 Bandas

<i>Landsat-8 OLI and TIRS Bands (micro-m)</i>			
Banda	Resolución	Rango Espectral	Longitud de onda
Banda 1	30 m	Coastal/ Aerosol	0.435 - 0.451
Banda 2	30 m	Azul	0.452 - 0.512
Banda 3	30 m	Verde	0.533 - 0.590
Banda 4	30 m	Rojo	0.636- 0.673
Banda 5	30 m	NIR	0.851 - 0.879
Banda 6	30 m	SWIR-1	1.566 - 1.651
Banda 7	30 m	SWIR-2	2.107 - 2.294
Banda 8	15 m	Pan	0.503 - 0.676
Banda 9	30 m	Cirrus	1.363 - 1.384
Banda 10	100 m	TIR-1	10.60-11.19
Banda 11	100 m	TIR-2	11.50-12.51

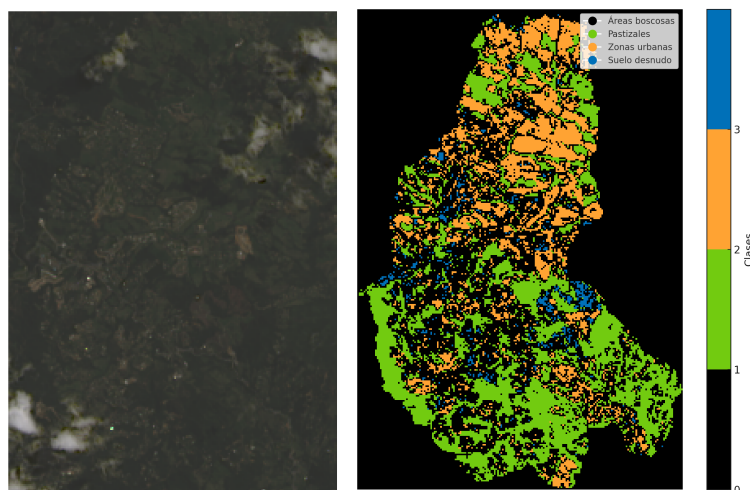


Figura 3.2: A la izquierda imagen Las Palmas capturada por Landsat 8. Derecha: Mapa de clasificación.

lucionales.

### 3.2.1. Pseudo etiquetas

En esta etapa del trabajo se exploraron técnicas inductivas semi-supervisadas para procesamiento de datos en imágenes multiespectrales, utilizando técnicas no supervisadas como los métodos de grupos de etiqueta, extracción de características y pre-entrenamiento para generar pseudo-etiquetas. La metodología de la técnica se presenta en el diagrama de bloques presentado en la Figura 3.3.

Para abordar el etiquetamiento de datos y clasificación en la imágenes multiespectrales, se aprovechan el aprendizaje semi-supervisado para mejorar la capacidad de generalización en los algoritmos de aprendizaje profundo. Por lo tanto el framework semi-supervisado, llamado Res-CP, usa una cantidad pequeña de muestras de entrenamiento por clase, para generar pseudo etiquetas en las muestras no etiquetadas. Las pseudo etiquetas posteriormente se usan con un método de clasificación para entrenar una red neuronal profunda (DNN). El framework Res-CP parte de 3 componentes: primero una red residual convolucional (RCN), segundo un generador de pseudo etiquetas (PLG) y por último una configuración de red neuronal llamada SSRN. Los anteriores pasos se describen a continuación:

1. **RCN:** Una red residual basada en centros que codifica las muestras etiquetadas y no etiquetadas en un espacio de incrustación común.
2. **PLG:** Generador probabilístico de pseudo etiquetas, este estima la distribución de la incertidumbre de las muestras codificadas en el primer paso para generar pseudo etiquetas. Es decir, asume la incertidumbre de las muestras etiquetadas codificadas de cada clase como una distribución gaussiana, donde la media de la distribución estimada representa la confiabilidad de las muestras y la varianza la incertidumbre de la estimación. Por último, la distribución se usa como referencia para evaluar las muestras no etiquetadas pertenecientes al conjunto designado seleccionando las que mas confianza para la generar las pseudo etiquetas.
3. **SSRN:** La red residual espectral-espacial en sus siglas en inglés es un marco de aprendizaje profundo diseñado de extremo a extremo para la clasificación de imágenes hiperespectrales (HSI) tomando cudros de estas imágenes sin necesidad de tratamiento de características previas.

En general, el RCN trata de codificar las clases, el PLG genera las pseudo etiquetas, y el SSRN utiliza las pseudo etiquetas para un entrenamiento super-

visado. Por lo cual, en esta ultima etapa se puede emplear cualquier otra red neuronal.

Mediante RCN , se usó un codificador de características para procesar muestras etiquetadas y no etiquetadas mediante un espacio de incrustación común utilizado. A continuación se describe el paso a paso de la RCN como se presenta en la Figura 3.4.

- Primero se comenzó con una **extracción de parches** [4] con un tamaño de  $w \times w \times C$  donde  $w$  representa ancho y altura del parche, y  $C$  la profundidad o los canales de la imagen multiespectral. En este proceso de extracción de parches se tiene un conjunto etiquetado compuesto por los fragmentos extraídos y sus etiquetas correspondientes.
- Segundo se usó un modelo de codificación determinístico que se divide en en dos subcomponentes :  $fv$  y  $gv$ . Donde  $fv$  usa una red residual espectral -espacial (SSRN) como base de extracción de características; cuenta con capas CNN 3D y conexiones residuales. Por el  $gv$  contiene dos  $FC$  que reciben la salida de  $fv$  y procesan la pérdida para actualizar los parámetros de RCN usando la pérdida conjunta que incluye la pérdida central y la pérdida softmax. Donde la primera, estimó el centro de cada clase obligando a las muestras a coincidir con el centro correspondiente en la etapa de entrenamiento. La pérdida softmax ayudó a que las características de cada clase en la muestra sea mas separables. Finalmente, se usaron  $fv$  y  $gv$  para actualizar los parámetros de  $fv$  y aprender el espacio de incrustación usando las muestras etiquetadas. Luego,  $fv$  se aplica para codificar tanto muestras etiquetadas como muestras no etiquetadas en el espacio de incrustación aprendido. Las salidas son las muestras etiquetadas y no etiquetadas codificadas.

**Generador de Pseudoetiquetas** usando **PLG**. Este enfoque se dividió en las etapas de estimación de distribución y el emparejamiento de muestras no etiquetadas. A continuación se presenta el diagrama de la metodología **PLG** en la Figura 3.5

En la primera etapa del PLG se determinó una representación probabilística de las muestras etiquetadas en el espacio latente, dado que las muestras podrían contener ruido y variar la calidad por lo que se aplicó una distribución probabilística como la Gaussiana capturando esta incertidumbre. Las características codificadas de las muestras etiquetadas de usaron para construir una distribución probabilística en el espacio latente donde esta distribución actuó como referencia para evaluar la paridad de las muestras no etiquetadas [4].

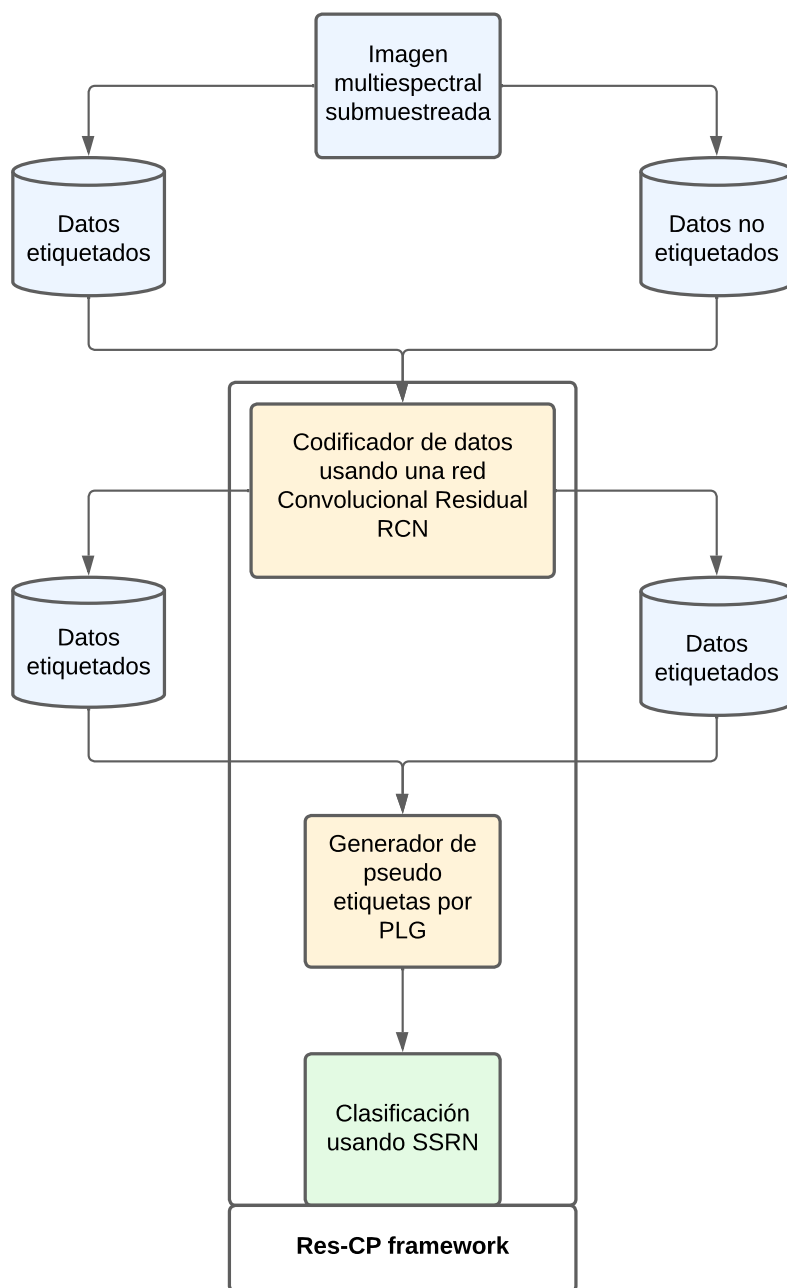


Figura 3.3: Diagrama de clasificación de la metodología propuesta. Adaptado de [4]

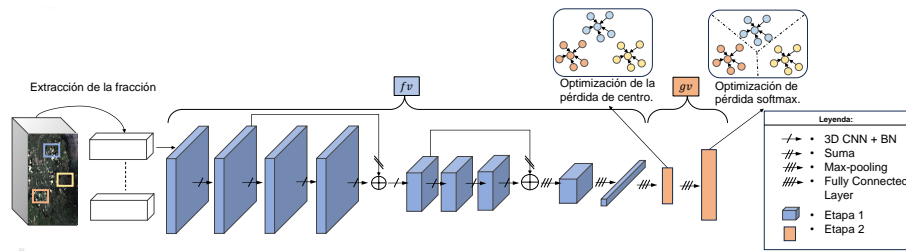


Figura 3.4: Diagrama de codificación de características mediante una Red Convolutional Residual. Adaptado de [5]

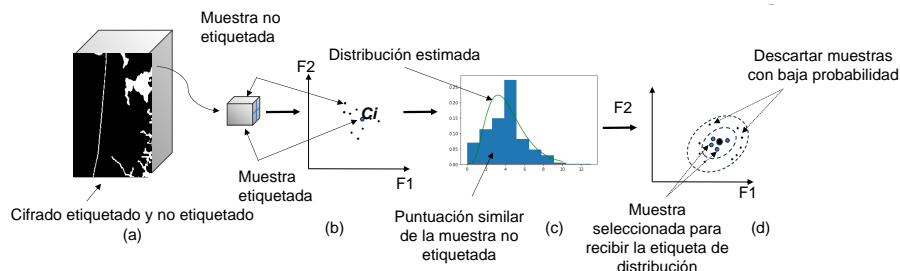


Figura 3.5: Diagrama de generación de Pseudoetiquetas. Adaptado de [5]

En la segunda, se empleó el emparejamiento de **muestras no etiquetadas**, comparando estas con la distribución de las muestras etiquetadas. En este proceso, se seleccionaron aquellas muestras no etiquetadas con mayor probabilidad de pertenecer a la clase esperada. Por lo tanto, se calcularon la similitud entre las muestras no etiquetadas con la distribución de las muestras etiquetadas donde las muestras etiquetadas que contengan un alto grado de paridad sean seleccionadas para la generación de pseudoetiquetas. Las muestras con baja probabilidad se descartan [4] [41].

Este proceso permite la mejora en clasificación semisupervisada generando etiquetas para las muestras no etiquetadas basadas en su similitud con las muestras etiquetadas. Las pseudoetiquetas permitieron que el modelo se entrenara con un conjunto de datos amplio.

Finalmente, se empleó **red residual espectral-espacial** (SSRN) compuesta por bloques residuales espectrales y espaciales que aprenden consecutivamente características discriminativas de firmas espectrales y contextos espaciales en imágenes hiperespectrales. Los bloques residuales se conectan entre cada capa convolucional 3-D a través de un mapeo de identidad, facilitando la retropropagación del gradiente y disminuyendo el desvanecimiento del gradiente [42].

En el framework propuesto por [5], se plantea una red convolucional residual

(RCN) basado en la codificación de muestras etiquetadas y no etiquetadas según un espacio común. A continuación, se presenta la arquitectura de la RCN 3.4. Donde esta configuración en su primera etapa codifica las características espaciales-espectrales. En la parte final de la red consta de dos capas totalmente conectadas (FC); que recibe la salida de la primera etapa de la RCN calculando la pérdida de los parámetros para actualizarlos usando una pérdida conjunta centro-softmax. Esta pérdida es introducida en [43] que se divide en dos: centro y softmax. Donde la primera se encarga de estimar el centro de cada clase obligando a las características de cada una de las muestras a emparejarse con el centro adecuado durante la etapa de entrenamiento [43]. Por otro lado, la pérdida softmax ayuda a que las muestras sean más separables. Teniendo en cuenta lo anterior, todos estos datos codificados se aplican para que el modelo aprenda un espacio de incrustación usando muestras etiquetadas. Toda esta descripción se puede observar en la figura 3.5

### 3.2.2. Clasificación Semi-supervisada Basada en grafos

Esta metodología de clasificación se centra en redes prototípicas de grafos para la clasificación de imágenes multiespectral (MSI). A continuación se presenta la metodología usada para el entrenamiento del modelo y pruebas del método ver figura 3.6.

La metodología aplicada con este método comenzó con los dataset descritos en la tabla 3.2 donde se usó la red prototípica de grafos de escalada cruzada implementada (X-GPN) por [41], diseñada para aprovechar la información espectral y espacial de las imágenes MSI en diferentes escalas. Los pasos implementados usados fueron:

1. **Construcción de matrices de adyacencia:** En esta primera etapa se utilizaron tres escalas de (3x3, 5x5 y 7x7) para construir matrices de adyacencia determinadas por vecindades multiescala que intervienen en las convoluciones de grafos multirrama [12].
2. **Capas de convolución de grafos (G-conv):** La red está constituida de dos capas G-Conv mencionadas en [41,44]. Las primeras y segundas unidades ocultas de las capas G-Conv son configuradas experimentalmente con 32 y 16 unidades, respectivamente [41].
3. **Capa prototípica:** Al finalizar la red, una capa prototípica que examina la información de atribución se alimenta con las características de los nodos fusionados. Los ideales prototipos se aprendan junto con las funciones de

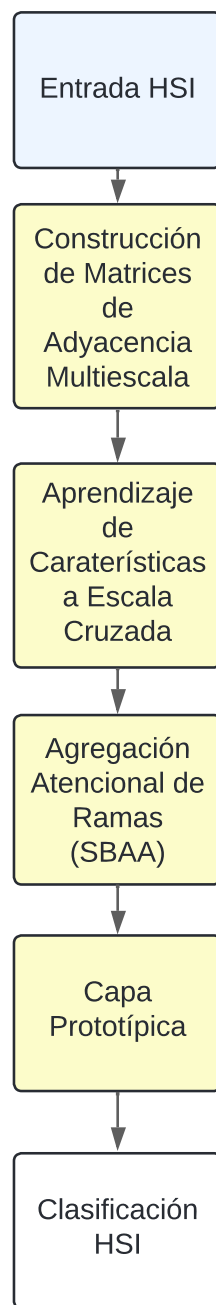


Figura 3.6: Diagrama de metodología de la Red Prototípica basada en grafos.

los nodos en el proceso de entrenamiento, y en el procedimiento de prueba, se establecen las categorías de las muestras que no se tienen etiquetado mediante el uso de estos prototipos.

4. **Regularización con etiquetas mixtas:** Al ser una clasificación semi-supervisada la discriminación de las características extraídas se mejoraron usando la regularización propuesta en [41] que involucra muestras etiquetadas y no etiquetadas; donde los prototipos mencionados anteriormente se optimizan utilizando ambos tipos de muestras, aumentando su capacidad de regularización para una clasificación mas precisa.
5. **Optimización:** El algoritmo emplea un optimizador Adam con un método del descenso de gradiente de lotes completos, donde los hiperparámetros usados se fijaron en: tasa de aprendizaje igual a 0.01 y las épocas en 500.

### 3.3. Extracción de Muestras de Entrenamiento y Pruebas

Para la partición de los datos en conjuntos de entrenamiento y pruebas, se usó la siguiente estrategia:

1. **Conjunto de datos Indian Pines (IP):** El esquema estándar de partición de datos de entrenamiento y prueba para este conjunto de datos ha sido utilizado como referencia en la literatura reciente [45] [46] por lo tanto se continuó con ese método. En general, se seleccionan el 10% de las muestras de entrenamiento de manera aleatoria para datos de validación y así desarrollar las redes. Durante el proceso de entrenamiento de la red, el 90% de las muestras de entrenamiento contienen etiquetas disponibles para optimizar los pesos de la red y el asociamiento de las muestras no etiquetadas.
2. **Conjunto de datos Las Palmas Oriente antioqueño (OA):** Se empleó la estrategia de elegir muestras aleatorias para el conjunto de entrenamiento [47]. Donde se genera un conjunto de entrenamiento y testeo según un porcentaje de muestras indicadas, en este caso el estudio se hizo con 20 muestras [4] por clase a la hora de clasificarlas en la red prototípica de grafos.

## 3.4. Extracción de características espaciales

En esta sección se usaron métodos de extracción de características espaciales empleando un submuestreo que consiste en reducir el número de bandas o canales disponibles en la HSI. El principal método de extracción de características empleado fue el de selección de bandas, aplicado a la figura 3.1, teniendo en cuenta 8 bandas del satélite Landsat 8 que contienen las siguientes longitudes de onda en la tabla 3.1. Se escogieron las bandas que se encuentran a 30 metros. Para hacer la operación de submuestreo obteniendo las longitudes de ondas objetivo para el experimento, donde la imagen 3.1 contenía 200 canales y se redujo a 9 canales. Finalmente, a cada una de las bandas seleccionadas de la MSI resultante se le aplicó una normalización escalar a los valores de las bandas en un rango de 0 a 1 y por último se usaron dos técnicas de visión por computador como la entropía y los operadores morfológicos; cada una de estas operaciones se hicieron con el lenguaje de programación Python usando librerías como *skimage*. La metodología propuesta se puede ver en diagrama 3.7.

### 3.4.1. Entropía

La entropía es usada para medir el nivel de aleatoriedad o incertidumbre en la información contenida en una imagen [34]. En este caso, se usó la entropía de Shannon para detectar bordes y texturas de interés en cada una de las bandas disponibles en la imagen multispectral, teniendo como objetivo marcar los cambios bruscos de cada uno de las clases de la imagen para definir los límites entre dos regiones separadas [48]. Como se puede observar en la figura 3.8 como ejemplo.

Durante el cálculo de la entropía, se analizó la imagen MSI mostrada en la Figura 3.8, destacando los contornos para resaltar las fronteras entre las distintas clases. Cada uno de los cambios sujetos en la imagen MSI con entropía perteneciente en la 3.8 están sujetos al hiperparámetro o elemento estructurador *disk* disponible en la librería *skimage* de Python donde este es un disco con radio especificado, que es una matriz binaria donde los píxeles en el interior del disco toman un valor de 1 y los píxeles fuera del disco son 0. El valor del parámetro  $\textit{disk}$  se ajustó entre 1 y 9, observando que a medida que aumentaba, la imagen se distorsionaba progresivamente, lo que ocasionaba la pérdida de información relevante. Por esta razón, se estableció un límite superior de 9. Finalmente, se seleccionó  $\textit{disk} = 4$ , ya que ofrecía un equilibrio óptimo entre la nitidez de los valores bajos y la opacidad de los valores altos.

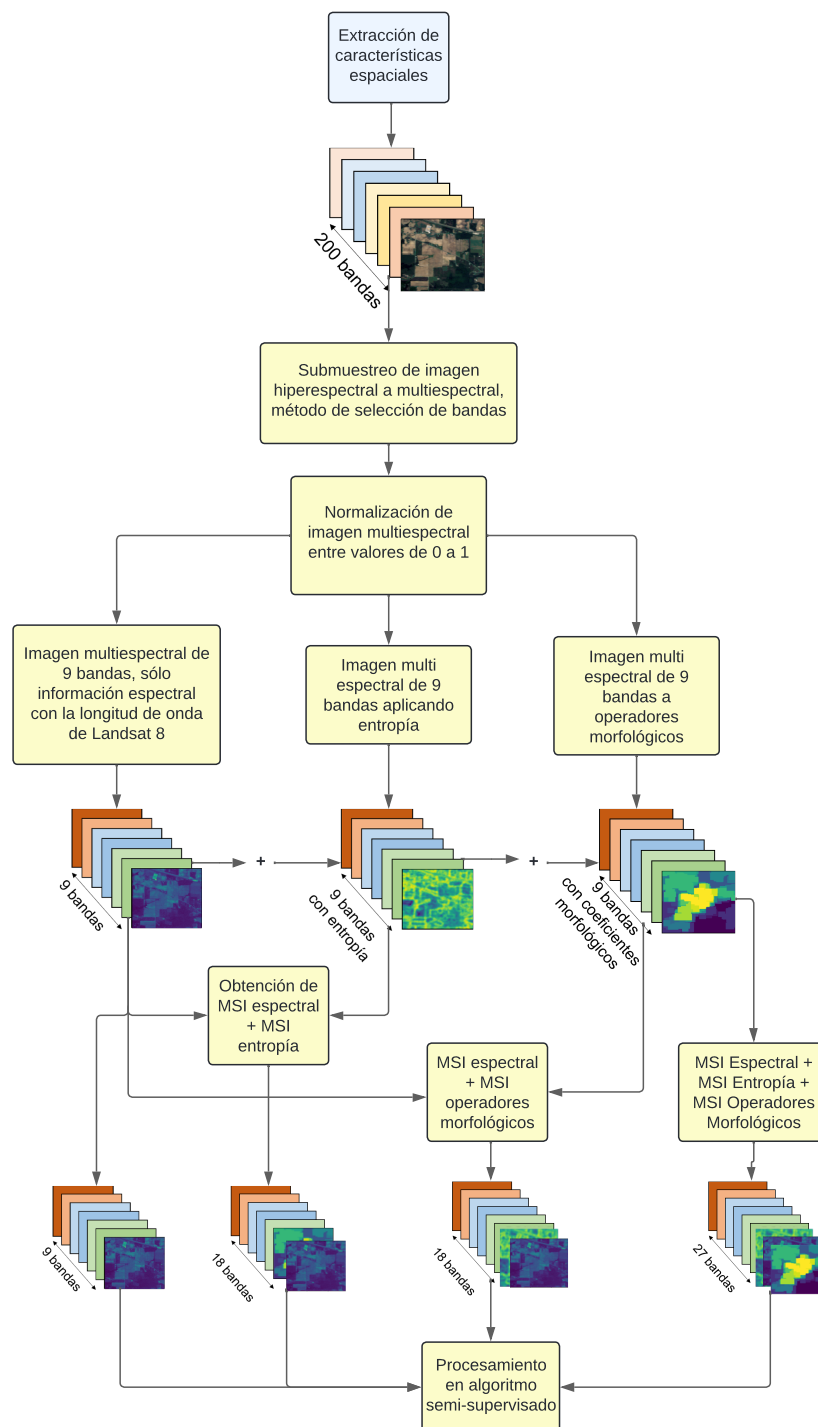


Figura 3.7: Diagrama de bloques de extracción de características espaciales usando selección de bandas, obtención de imágenes MSI = 9 bandas espectrales, 18 bandas con información espectral y entropía (9 bandas espectrales y 9 bandas con entropía), 18 bandas con información espectral y operadores morfológicos (9 bandas espectrales y 9 bandas con operadores morfológicos) y 27 bandas con información espectral, entropía y operadores morfológicos (9 bandas espectrales, 9 bandas con entropía y 9 bandas con operadores morfológicos).

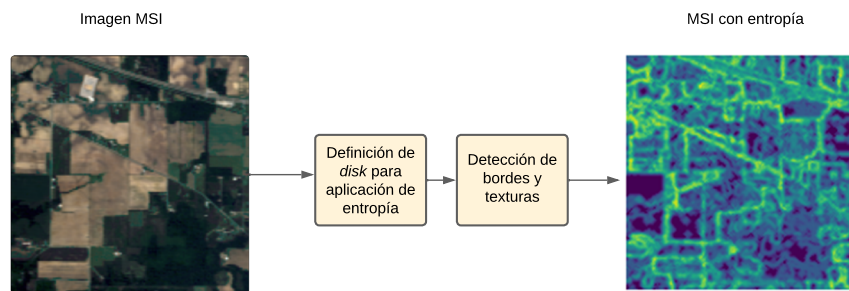


Figura 3.8: Ejemplo de metodología usando la entropía. Imagen MSI sin aplicar entropía. Al lado derecho MSI con entropía.

### 3.4.2. Operadores morfológicos

En este paso se usaron los operadores morfológicos en la imagen multiespectral usando las operaciones de dilatación, erosión, apertura, clausura. A continuación se presenta un diagrama de la metodología aplicada para la imagen multiespectral Figura 3.9

Primero, se binariza la imagen multiespectral aplicándole un bajo valor de umbral, tratando de limpiar la imagen. Después se designa un elemento estructurante o matriz de **unos** por banda de la imagen, con una dimensión de  $5 \times 5$  parecida a la convolución [49], el tamaño de la ventana se eligió con base a los elementos de la imagen para ir resaltando las características de cada banda de la imagen multiespectral, además el elemento estructurador fue mayor que los píxeles de los objetos que se deseaban suprimir y menor que los píxeles de los objetos que se querían conservar. Segundo, se aplicó sucesivas operaciones de erosión [37] para eliminar ruido de la imagen manteniendo al mismo tiempo las características claves de la imagen, para esto se usó la operación de apertura.

Tercero se aplicó una dilatación sucesiva para ir resaltando las clases de la imagen seguida de otra erosión sucesiva para ir restaurando el tamaño de las características de la imagen, aplicando el número de dilataciones sucesivas del mismo modo que se aplicaron las erosiones usando el mismo elemento estructurador garantizando la restauración óptima de la imagen llegando a que sea lo más parecida posible.

Finalmente se usaron las operaciones morfológicas de área de apertura y área de clausura que ayudarán a que la imagen quede restaurada y lo más parecido posible a la original, rellenando las áreas agujereadas y eliminando ruido sin necesidad de usar un elemento estructurador.

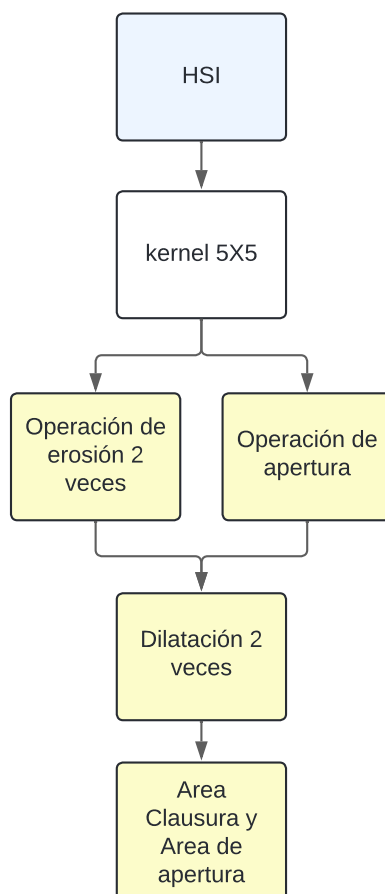


Figura 3.9: Ejemplo de metodología usando los operadores morfológicos.

### 3.4.3. Nombramiento de los conjuntos de datos

Después de haber tratado cada una de las imágenes con las operaciones de submuestreo, entropía y operadores morfológicos ver tabla 3.2, se crearon 5 conjuntos de datos con los siguientes nombres

Tabla 3.2: Tabla de los nombres de los dataset de imágenes multiespectrales creadas con la información espacial. IP equivale a Indian Pines y OA a Oriente Antioqueño 9, 18 y 27 indican el número de canales que contiene cada imagen multiespectral.

<b>Conjunto de datos (MSI)</b>	<b>Características Utilizadas</b>
IP_9	Solo información espectral
IP_18	Información espectral y entropía
IP_18_m	Información espectral y operadores morfológicos
IP_27	Información espectral, operadores morfológicos y entropía
Palmas_OA	Información espectral, operadores morfológicos y entropía

Esta metodología presenta como se parametrizaron cada una de estas técnicas para ser llevados a un estudio óptimo.

# Capítulo 4

## Resultados y Discusión

### Contenido

---

4.1	Resultados obtenidos con Indian Pines . . . . .	34
4.1.1	A. Configuración de hiperparámetros y criterios de evaluación: . . . . .	34
4.2	Resultados obtenidos con imagen del Oriente Antioqueño .	42
4.3	Comparación de los métodos semi-supervisados Res-CP y SSGPN . . . . .	45

---

### Sinopsis

Este capítulo presenta los resultados obtenidos al procesar la imagen multi-espectral Indian Pines. Se pusieron a prueba dos métodos semi-supervisados. El primero es el *framework Res-CP*, que se enfoca en una generación probabilística de pseudo-etiquetas, y el segundo es una red prototípica de grafos, cuyo objetivo es comparar su eficacia al clasificar muestras en los conjuntos de datos anteriormente mencionados. El primer método semi-supervisado, basado en pseudo-etiquetas, se evaluó analizando su eficacia en la clasificación. Se utilizó la Red Neuronal SSRN propuesta por [50], revisando tiempos de cómputo, precisión por número de muestras seleccionadas por clase y sintonización adecuada de hiperparámetros durante el entrenamiento de la red neuronal. El segundo método, basado en grafos, es una técnica que aprovecha la estructura de los datos y el relacionamiento espacial-espectral intrínseco en las imágenes MSI. Este método construye gráficos que se enlazan entre píxeles o regiones en la imagen, obteniendo la correlación y dependencia contextual entre los datos. También se

analizarán los resultados considerando solo la información espectral de las imágenes y la información espacial integrada mediante técnicas de procesamiento de imágenes.

Por último, todos los estudios realizados en este trabajo se desarrollaron utilizando la unidad de procesamiento gráfico (GPU) Nvidia Geforce RTX 4060 en un computador de escritorio con sistema operativo Windows 11. Para el primer método semi-supervisado, se utilizó Pytorch 2.2.0+cu121 y Python 3.12, mientras que para el segundo método, se usó un entorno virtual con Tensorflow 1.3.1, Keras 2.1.5 y Python 3.6.8.

## 4.1. Resultados obtenidos con Indian Pines

### 4.1.1. A. Configuración de hiperparámetros y criterios de evaluación:

A. En esta sección, el *framework* Res-CP contiene dos subcomponentes: **RCN** y **PLG**, los cuales se utilizan para generar las pseudo-etiquetas del método semi-supervisado. Además, se utilizó el optimizador Optuna [51], que optimizó la tasa de aprendizaje del modelo, la tasa de los centros de pérdida, los pesos de pérdida de centro, el tamaño del paso (*stepsize*) y la evaluación de caída del hiperparámetro gamma, logrando un buen desempeño del modelo al clasificar las pseudo-etiquetas. Asimismo, cada uno de estos experimentos se realizó con diferentes muestras de entrenamiento, mencionadas en el capítulo anterior, donde se eligieron los mejores resultados de cada uno de los subcomponentes del *framework* Res-CP.

1. Parametrización de RCN: La red utilizada en este subcomponente fue recuperada de [4], donde los hiperparámetros ajustados fueron: la tasa de aprendizaje más usada fue de 0.00001, basada en la optimización de [50]. El tamaño del *batch* y el número de épocas fueron 16 y 500, respectivamente. Se establecieron en 0.1 y 0.5 los hiperparámetros de pérdida centro-máxima suave, que regulan la variación dentro de las clases y la tasa de aprendizaje de los centros. Las funciones de pérdida centro-softmax se optimizan mediante la propagación del error cuadrático medio (RMS-Prop) [52]. Para la tarea de entrenamiento, se establece  $w$  en 9, y la capa softmax se ajusta según la cantidad de clases en cada conjunto de datos.
2. Configuración de PLG: La configuración de los hiperparámetros de la red se centra en un intervalo de confianza llamado  $Rl$ , que realiza la tarea de

emparejamiento con base en un  $w = 0.5$ . En otras palabras,  $w$  se posiciona de forma que extrae muestras no etiquetadas a una distancia casi igual a la de las muestras etiquetadas [4] en las imágenes multiespectrales trabajadas.

3. Criterios de evaluación: Los criterios de evaluación utilizados para medir el rendimiento del modelo fueron: precisión global (OAC), coeficiente kappa (k) y precisión media (AA) (cada acrónimo en paréntesis representa sus siglas en inglés). Se ejecutaron 20 estudios usando el optimizador Optuna para obtener los mejores hiperparámetros de entrenamiento y las mejores métricas de desempeño.

Los resultados descritos en la tabla 4.3 muestran cada uno de los datasets evaluados con diferentes muestras de entrenamiento. Estos datasets fueron aumentando su profundidad a medida que se integraban bandas espaciales, donde cada una de estas imágenes se tomaron un número de muestras por clase comenzando por 5, 10 y 20. Los resultados mostraron que, cuanto más información espacial y más muestras por clase se seleccionaron, mejor fue la clasificación del método semi-supervisado.

Las pseudo-etiquetas generadas por el *framework* Res-CP son efectivas para mejorar la capacidad de generalización y la precisión de los modelos en la clasificación de imágenes multiespectrales. Tomando una fracción de las muestras etiquetadas en la tabla 4.1, se puede evidenciar la cantidad de pseudo-etiquetas logradas por PLG con los conjuntos de datos, usando solo información espectral y usando información espectral y espacial. Inicialmente, Res-CP se usó únicamente con la información espectral de la imagen, obteniendo métricas de desempeño del 60 % por parte del dataset IP\_9\_1. Según la tabla 4.3, este resultado mantuvo un estándar de épocas = 200, *batch size* = 32, y el resto de los hiperparámetros de optimización se muestran en la tabla 4.2. Cabe destacar que cada uno de los datasets aumentó su porcentaje de acierto según la cantidad de pseudo-etiquetas logradas con base en las muestras etiquetadas extraídas del mapa de clasificación.

Tabla 4.1: Número de pseudo-muestras generadas por PLG

IP_9_20	Palmas_OA	Palmas_OA_21	IP_27_20
1485	2181	2596	1589

Al evaluar las pseudo-etiquetas, se obtuvo la matriz 4.3, donde se probaron varios conjuntos de datos con diferentes muestras de entrenamiento (5, 10 y 20) para evaluar el rendimiento del método semi-supervisado. Se observó que,

Tabla 4.2: Parámetros de Ejecución del Script Python de Res-CP

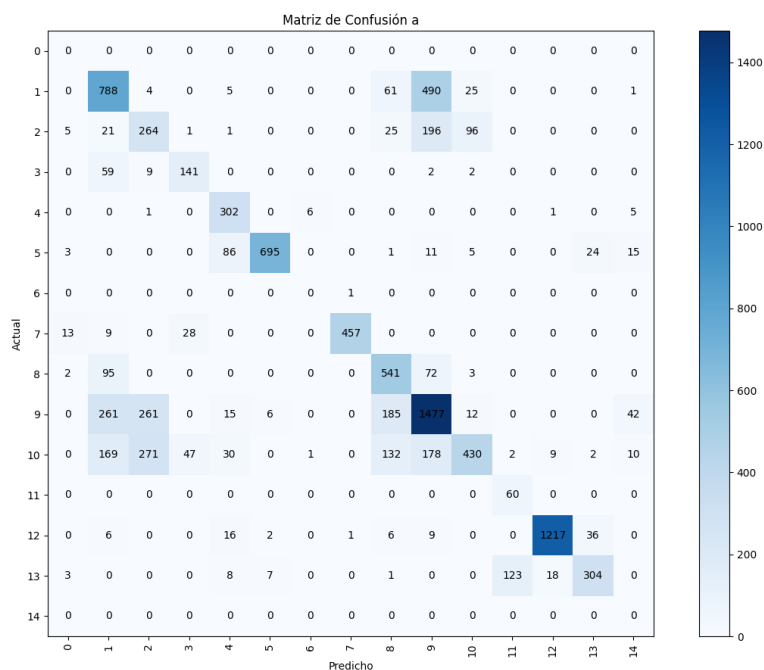
Parámetro	Valor
Dataset	Indian Pines (IP)
Máximas Épocas (max-epoch)	200
Tasa de Aprendizaje del Modelo (lr-model)	1.011e-05
Tasa de Aprendizaje del Centro (lr-cent)	0.010
Peso del Centro (weight-cent)	0.176
Gamma	0.897
Tamaño del Paso (stepsize)	25

cuanto más muestras por clase y bandas espaciales contiene la imagen, mayor es el rendimiento en clasificación. Se puede comparar con la imagen (a), que es el mapa de clasificación original.

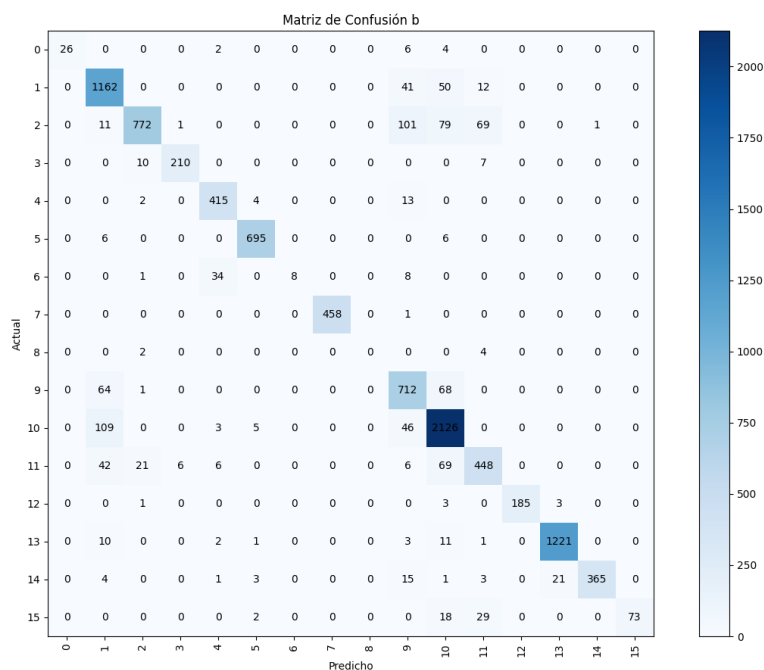
- El entrenamiento de la imagen IP\_9 contiene solo información espectral, es decir, no se le añadieron características de entropía y operadores morfológicos, por lo tanto, tras todo el procesamiento semi-supervisado, se obtuvieron resultados por debajo del 67.25 % de precisión en la clasificación, teniendo en cuenta que la imagen quedó reducida de 200 bandas a 9 bandas. Además, se puede evidenciar el comportamiento de las clases en la tabla 4.3 al momento de clasificarse. Por otro lado, al añadir la extracción de características a las 9 bandas, se notó un aumento en las métricas de desempeño, llegando a un máximo de 89.39 % en el caso de IP\_27\_20, lo que indica que este tipo de procesamiento de imágenes ayuda a la convergencia de las clases en las imágenes MSI. Res-CP no aborda específicamente la precisión en áreas de bordes, pero mejora las características discriminativas, potenciando la clasificación en las DNNs.

#### *B. Método Redes gráficas prototípicas semisupervisadas (SSGPN)*

El método basado en red prototípica de grafos semi-supervisada fue usado para la clasificación de imágenes multiespectrales. combinan técnicas de redes convolucionales de grafos (GCN) con capas prototípicas. Este método mejora la precisión de la clasificación por medio de compactación de los resultados, donde el uso de la capa prototípica permitió obtener características más discriminativas, incrementando la separación inter-clase y la compactación intra-clase, dando una clasificación precisa. Teniendo como componentes claves: la Regularización de Entropía Temporal (TER), Normalización de grafos (G-Norm) usados en [41]. Para ello, el método fue sintonizados con los siguientes hiperparámetros de entrenamiento descritos en la Tabla 4.6, donde el tamaño del vecindario ( $7 \times 7$



(a)



(b)

Figura 4.1: (a) Matriz de confusión de Indian Pines usando Res-Cp con información espectral y 20 muestras por clase. (b) Matriz de confusión de Indian Pines usando Res-Cp con información espacial-espectral y 20 muestras por clase.

Tabla 4.3: Resultados del experimento usando SSRN de Res-CP con las imágenes multiespectrales combinadas.

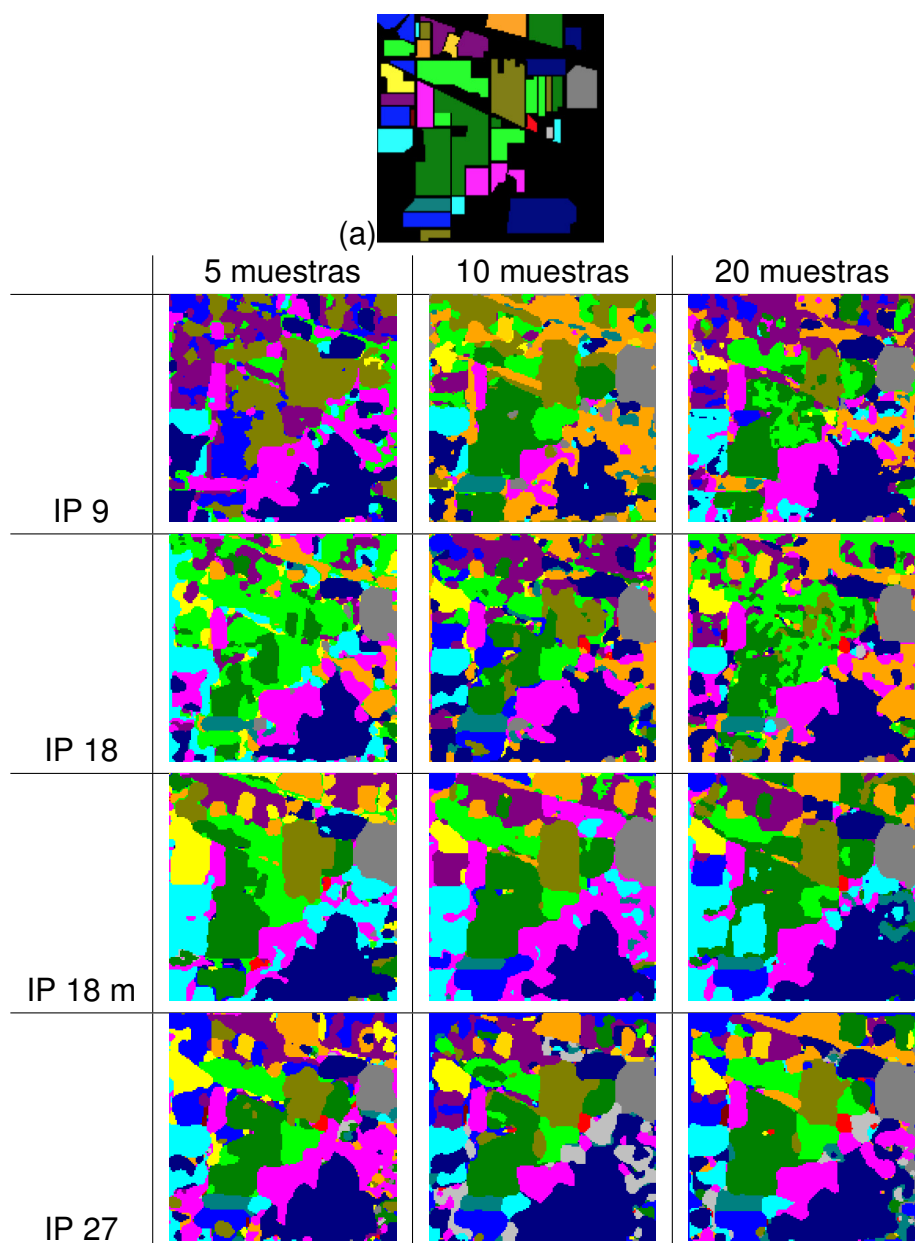
Imagen multiespectral Indian Pines	Precisión Global (%)	Coefficiente Kappa (%)	Precisión media (%)	Tiempo total transcurrido (h:m:s)
IP_9_5	63.08	57.69	60.27	0:05:09
IP_9_10	41.17	35.53	21.53	0:05:38
IP_9_20	67.25	62.64	64.67	0:06:48
IP_18_5	54.95	48.68	57.00	0:01:43
IP_18_10	71.64	68.17	62.31	0:04:45
IP_18_20	66.10	61.33	83.72	0:05:23
IP_18_m_5	64.29	59.41	52.92	0:06:27
IP_18_m_10	77.21	74.09	66.09	0:06:00
IP_18_m_20	76.54	73.31	67.25	0:06:00
IP_27_5	76.01	73.00	77.52	0:05:09
IP_27_10	78.93	76.11	80.80	0:05:36
<b>IP_27_20</b>	<b>89.39</b>	<b>87.92</b>	<b>76.95</b>	<b>0:05:09</b>

píxeles) es usado para la construcción de la matriz de adyacencia con el parámetro  $K = 40$  de vecinos. Para calcular las aristas iniciales de la matriz adyacencia el parámetro  $\tau$  se estableció en 0.1 por [41], las G-Conv se establecieron en 32 y 16 respectivamente, el parámetro de empujamiento  $\alpha$  que afecta la generación de prototipos en la capa prototípica, se estableció en 1. Finalmente, el resto de los hiperparámetros se fueron sintonizando a medida que el algoritmo tendía a las métricas necesitadas.

- En la primera etapa de entrenamiento, se utilizó el conjunto de datos IP\_9 para observar el comportamiento con las 9 bandas disponibles, donde se obtuvo una precisión global del 89.41 % (ver Tabla 4.7), lo que demuestra la buena capacidad de generalización del método, considerando el porcentaje de muestras de entrenamiento descrito en la sección 3.3.
- En la segunda etapa, se agregó la extracción de características a IP\_9, aumentando los canales de la imagen original de nueve bandas a 18 y 27 bandas, lo que incrementó significativamente la precisión del modelo X-GPN. En particular, el conjunto de datos IP\_18, que contenía características relacionadas con la entropía de las nueve bandas espectrales de la MSI, mostró un mayor nivel de precisión.

Los resultados se pueden observar directamente en la Tabla 4.8 donde el mapa de clasificación (e) se puede comparar con los resultados obtenidos al momento de la clasificación de características.

Tabla 4.4: (a) En la parte superior mapa de clasificación. Matriz de imágenes por categoría y número de muestras



Este enfoque fue implementado en los dataset de la tabla 3.2. El entrenamiento de X-GPN muestra los resultados con una alta precisión desde la imagen IP\_9\_20 que solo integra la información espectral de la imagen y el resto de los conjuntos de datos integra la información espectral, en la Tabla 4.7 se evidencia

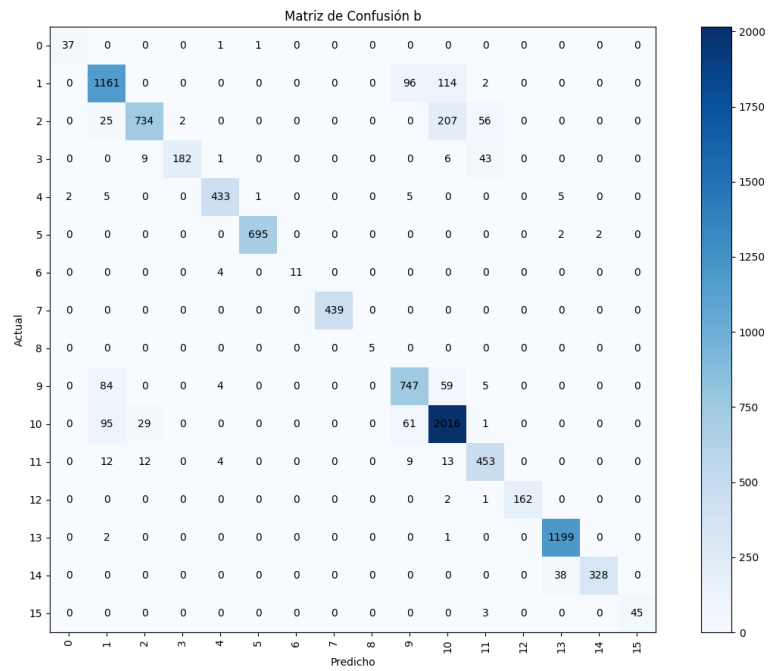
Tabla 4.5: Etiquetas de Indian Pines

Index	Class Label
1	Alfalfa
2	Corn-notill
3	Corn-mintill
4	Corn
5	Grass-pasture
6	Grass-trees
7	Grass-pasture-mowed
8	Hay-windrowed
9	Oats
10	Soybean-notill
11	Soybean-mintill
12	Soybean-clean
13	Wheat
14	Woods
15	Buildings-Grass-Trees-Drives
16	Stone-Steel-Towers

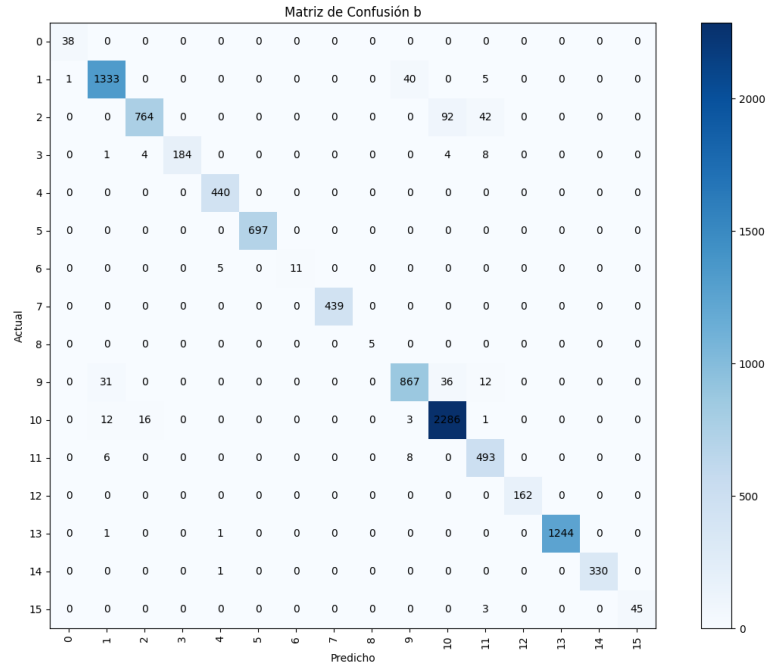
la , especialmente en regiones de bordes de las imágenes, en la Tabla 4.8 se observa la capacidad de generalización del método y como separa las clases de manera efectiva.

### C. Resultados de las matrices de confusión

Las matrices de confusión seleccionadas evaluaron los mejores desempeños de clasificación de los conjuntos de datos de IP\_9\_20, IP\_27\_20 e IP\_18. Los resultados obtenidos fueron con los marcos de trabajo Res-CP y SSGPN, donde las matrices (a), (b) fueron con el primer marco de trabajo y las (c) y (d) fueron con el segundo marco de trabajo. A continuación en la Tabla 4.9 se presenta un análisis comparativo de los resultados de las matrices de confusión donde se tuvo en cuenta la precisión general, errores notables, desempeño comparativo y las conclusiones. (Las etiquetas de las matrices de confusión (a), (b), (c) y (d) se pueden observar en la Tabla 4.5).



(c)



(d)

Figura 4.2: (c) Matriz de confusión de Indian Pines usando SSGPN con información espectral.(d) Matriz de confusión de Indian Pines usando SSGPN con información espacial-espectral.

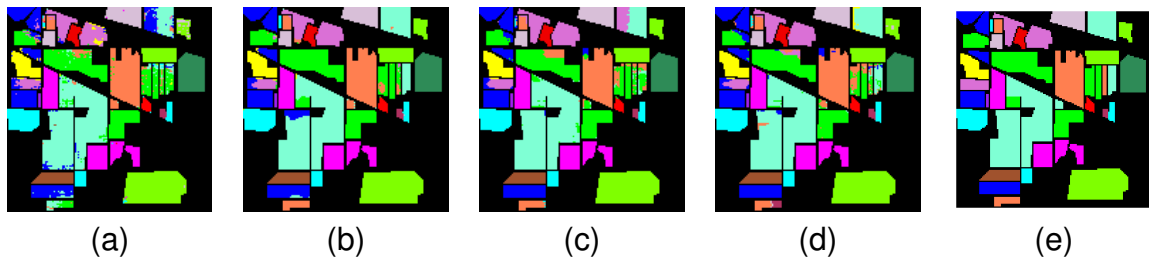
Tabla 4.6: Configuración de Hiperparámetros del X-GPN

Hiperparámetro	Valor
Tamaño del Vecindario	$7 \times 7$ píxeles
Número de Vecinos Más Cercanos ( $K$ )	40
Parámetro $\tau$	0.1
Unidades Ocultas en la 1ra G-Conv	32
Unidades Ocultas en la 2da G-Conv	16
Parámetro de Empinamiento $\alpha$	1 (Indian Pines)
Parámetro del Regularizador $\beta$	10
Algoritmo de Optimización	Adam
Tasa de Aprendizaje	0.01
Número de Épocas	500

Tabla 4.7: Resultados del experimento usando X-GPN con cada una de las imágenes multiespectrales combinadas

Imagen	Precisión Global (%)	Coefficiente Kappa (%)	Precisión Media (%)
IP_9	89.41	87.90	93.33
<b>IP_18</b>	<b>97.41</b>	<b>96.05</b>	<b>96.56</b>
IP_18_m	92.66	91.61	93.23
IP_27	93.70	92.79	94.80

Tabla 4.8: Resultados de Indian Pines: (a) IP\_9, (b) IP\_18, (c) IP\_18\_m, (d) IP\_27 y (e) mapa de clasificación



## 4.2. Resultados obtenidos con imagen del Oriente Antioqueño

En esta sección se usó la imagen del oriente antioqueño donde se le añadieron las extracciones de características espaciales, se usaron los algoritmos

Tabla 4.9: Comparación de las Matrices de Confusión Generadas con Diferentes Configuraciones Espectrales

Aspecto	Matriz (a)	Matriz (b)	Matriz (C)	Matriz (d)
<b>Precisión General</b>	<ul style="list-style-type: none"> <li>- Alta precisión en 'Corn-notill', 'Corn-mintill', 'Soybean-notill', y 'Woods'.</li> <li>- Leve confusión entre 'Soybean-clean' y 'Soybean-mintill'.</li> </ul>	<ul style="list-style-type: none"> <li>- Alta precisión en 'Corn-notill', 'Soybean-notill', 'Soybean-mintill', y 'Woods'.</li> <li>- Mayor dispersión en 'Corn', indicando una menor precisión en esta clase.</li> </ul>	<ul style="list-style-type: none"> <li>- Alta precisión en 'Corn-notill' y 'Soybean-clean'.</li> <li>- Leve confusión en 'Soybean-mintill' y 'Wheat'.</li> </ul>	<ul style="list-style-type: none"> <li>- Alta precisión en 'Corn-notill' y 'Soybean-clean'.</li> <li>- Baja confusión entre las clases, destacando 'Soybean-mintill'.</li> </ul>
<b>Errores Notables</b>	<ul style="list-style-type: none"> <li>- Algunos errores en 'Soybean-mintill' y 'Soybean-clean'.</li> <li>- Confusión en 'Oats', 'Wheat', y 'Grass-pasture-mowed'.</li> </ul>	<ul style="list-style-type: none"> <li>- Aumento en la confusión entre 'Soybean-clean' y 'Soybean-mintill'.</li> <li>- Mayor confusión en 'Corn-mintill'.</li> </ul>	<ul style="list-style-type: none"> <li>- Confusión leve en 'Hay-windrowed' y 'Wheat'.</li> <li>- 'Soybean-clean' clasificada incorrectamente en algunas instancias.</li> </ul>	<ul style="list-style-type: none"> <li>- Reducción en los errores de clasificación en 'Soybean-mintill' y 'Soybean-clean'.</li> <li>- Confusión mínima en 'Wheat' y 'Woods'.</li> </ul>
<b>Desempeño Comparativo</b>	<ul style="list-style-type: none"> <li>- Mejores resultados generales, menor confusión.</li> </ul>	<ul style="list-style-type: none"> <li>- Ligeramente aumento en la confusión, posiblemente debido a canales adicionales.</li> </ul>	<ul style="list-style-type: none"> <li>- Desempeño comparable con la matriz 1, con diferencias leves en clases específicas.</li> </ul>	<ul style="list-style-type: none"> <li>- Desempeño más consistente en todas las clases con menor confusión general.</li> </ul>
<b>Conclusión</b>	<ul style="list-style-type: none"> <li>- La configuración con 9 canales espectrales es efectiva y reduce la confusión.</li> </ul>	<ul style="list-style-type: none"> <li>- La adición de más canales no mejora necesariamente la clasificación, puede introducir ruido.</li> </ul>	<ul style="list-style-type: none"> <li>- 27 canales espectrales mantienen un buen equilibrio entre precisión y confusión.</li> </ul>	<ul style="list-style-type: none"> <li>- 18 canales espectrales proporcionan una clasificación precisa con mínima confusión, mejor que otras configuraciones.</li> </ul>

semi-supervisados mencionados en el ítem anterior y se lograron resultados de clasificación como se muestran en la Tabla 4.11.

Se evidencia que los métodos de la Tabla 4.11 (b) y (c) muestran características dispersas según el mapa de clasificación (a). Por parte de la imagen (b) refleja el resultado del método Res-CP, que en comparación con (a) presenta un alto ruido y menor coherencia en la clasificación de características. Las clases están menos definidas y se nota una sobre segmentación en algunas áreas. Posee baja continuidad espacial con las clases.

El resultado de la imagen (c) con el método SSGPN con respecto al mapa de clasificación (a), tiende a ser más coherente y se aproxima más al mapa de referencia. Presenta una distribución más continua debido a la definición de los bordes de las clases. Este método puede definir mucho mejor en términos de precisión espacial de la imagen.

Tabla 4.10: Resultados de Indian Pines: (a) Mapa de clasificación OA (Oriente Antioqueño), (b) Resultado OA con Res-CP , (c) Resultado de OA con SSGPN

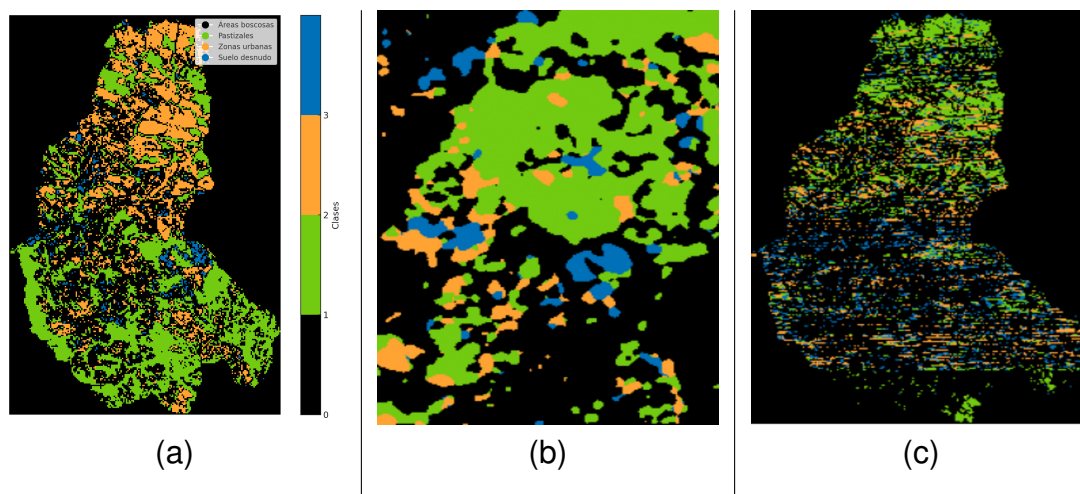


Tabla 4.11: Comparativa de métricas para diferentes OA usando SSGPN y Res-Cp.

OA	SSGPN	Res-Cp
<b>OA_7_20</b>	Precisión Global = 89.41 % Coeficiente Kappa = 87.90 % Precisión Media = 93.33 %	Precisión Global = 78.98 % Coeficiente Kappa = 61.87 % Precisión Media = 55.69 %
<b>OA_21_20</b>	Precisión Global = 70.65 % Coeficiente Kappa = 92.79 % Precisión Media = 94.80 %	Precisión Global = 82.48 % Coeficiente Kappa = 68.13 % Precisión Media = 61.16 %

En la Tabla 4.11 se presenta una comparativa de las métricas de precisión obtenidas utilizando dos métodos diferentes, SSGPN y Res-Cp, para diferentes OAs. Los resultados muestran que, para el caso de OA\_7\_20, el método SSGPN alcanza una Precisión Global de 89.41 %, un Coeficiente Kappa de 87.90 %, y una Precisión Media de 93.33 %. Estos valores son considerablemente superiores a los obtenidos con el método Res-Cp para el mismo OA, que reporta una Precisión Global de 78.98 %, un Coeficiente Kappa de 61.87 %, y una Precisión Media de 55.69 %.

En el caso de OA\_21\_20, los resultados muestran un desempeño inverso. El método SSGPN obtiene una Precisión Global de 70.65 %, un Coeficiente Kappa de 92.79 %, y una Precisión Media de 94.80 %, mientras que el método Res-Cp presenta una Precisión Global de 82.48 %, un Coeficiente Kappa de 68.13 %, y una Precisión Media de 61.16 %.

Estos resultados sugieren que, aunque el método SSGPN ofrece un desempeño superior en la mayoría de los casos, Res-Cp puede ser más efectivo en ciertos escenarios específicos, lo que destaca la importancia de seleccionar el método adecuado según el tipo de OA y las características particulares de los datos.

### 4.3. Comparación de los métodos semi-supervisados Res-CP y SSGPN

Tabla 4.12: Comparación entre SSGPN y Res-CP para la clasificación de imágenes multiespectrales.

Aspecto	SSGPN	Res-CP
<b>Enfoque Principal</b>	Uso de redes prototípicas y convolucionales de grafos para mejorar la clasificación mediante capas prototípicas.	Generación de pseudo-etiquetas mediante un marco probabilístico basado en la distribución de características latentes.
<b>Objetivo</b>	Mejorar la precisión de la clasificación mediante la compactación intra-clase y la discriminación inter-clase.	Mejorar la capacidad de generalización de las DNNs usando pseudo-etiquetas generadas con alta confianza.
<b>Componentes Clave</b>	<ul style="list-style-type: none"> <li>▪ Capa Prototípica</li> <li>▪ Regularizador de Entropía Temporal (TER)</li> <li>▪ Normalización de Grafos (G-Norm)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Red Convolucional Residual (RCN)</li> <li>▪ Generador de Pseudo-Etiquetas (PLG)</li> <li>▪ PRCLSTM II</li> </ul>
<b>Rendimiento de Clasificación</b>	Excelente rendimiento en escenarios con datos limitados, superando a otros métodos como GCN y CNN.	Significativa mejora en la clasificación de DNNs cuando se usan las pseudo-etiquetas generadas, incluso con pocas muestras etiquetadas.
<b>Precisión en Áreas de Bordes</b>	Alta precisión en áreas de bordes y detalles finos en las imágenes clasificadas.	No se aborda específicamente la precisión en áreas de bordes, pero se mejoran las características discriminativas generales.
<b>Precisión con Extracción de Características Espaciales</b>	Alta precisión en clasificación y en delimitación de clases.	Aumento en las métricas de clasificación y en la obtención de pseudo-etiquetas.
<b>Eficiencia Computacional</b>	Mejorada con la introducción de G-Norm, lo que acelera la convergencia del modelo.	Enfoque ligero y compatible con diversas DNNs, aunque no se aborda directamente la eficiencia computacional.
<b>Manejo de la Incertidumbre</b>	No aborda directamente la incertidumbre, pero mejora la discriminación de características.	Utiliza un marco probabilístico para manejar la incertidumbre en la generación de pseudo-etiquetas, filtrando aquellas con baja confianza.
<b>Aplicabilidad</b>	Adecuado para escenarios donde la precisión y la compactación de clases son críticas.	Ideal para mejorar DNNs en escenarios con datos etiquetados limitados, especialmente cuando la calidad de los datos es variable.
<b>Robustez</b>	Alta robustez en la clasificación de clases con características similares.	Alta robustez frente a la incertidumbre de las etiquetas, minimizando el sesgo de confirmación y errores propagados.

En todos los casos presentados, el método SSGPN demuestra un desempeño consistentemente superior al método Res-Cp, tanto en términos de Precisión

Tabla 4.13: Comparativa de métricas para diferentes IPs usando ssgpn y res-cp.

IP	SSGPN	Res-CP
IP_9_20	Precisión Global = 89.41 % Coeficiente Kappa = 87.90 % Precisión Media = 93.33 %	Precisión Global = 67.25 % Coeficiente Kappa = 62.64 % Precisión Media = 64.67 %
IP_18_20	Precisión Global = 97.41 % Coeficiente Kappa = 96.05 % Precisión Media = 96.56 %	Precisión Global = 66.10 % Coeficiente Kappa = 61.33 % Precisión Media = 83.72 %
IP_18_m_20	Precisión Global = 92.66 % Coeficiente Kappa = 91.61 % Precisión Media = 93.23 %	Precisión Global = 76.54 % Coeficiente Kappa = 73.31 % Precisión Media = 67.25 %
IP_27_20	Precisión Global = 93.70 % Coeficiente Kappa = 92.79 % Precisión Media = 94.80 %	Precisión Global = 89.39 % Coeficiente Kappa = 87.92 % Precisión Media = 76.95 %

Global, Coeficiente Kappa, como en Precisión Media. Esto indica que SSGPN es más efectivo para la clasificación de imágenes en estos conjuntos de datos específicos, proporcionando resultados más precisos y fiables.

En general, la gráfica 4.3 subraya la superioridad del método SSGPN en la mayoría de las métricas, lo que lo posiciona como una opción más robusta y precisa para la clasificación de imágenes en comparación con Res-Cp. No obstante, Res-Cp muestra un rendimiento competitivo en algunos casos específicos, lo que podría ser relevante en aplicaciones donde la Precisión Global es prioritaria. Este análisis demuestra que la selección del método debe considerar tanto la naturaleza del conjunto de datos como las métricas de rendimiento más críticas para la aplicación en cuestión.

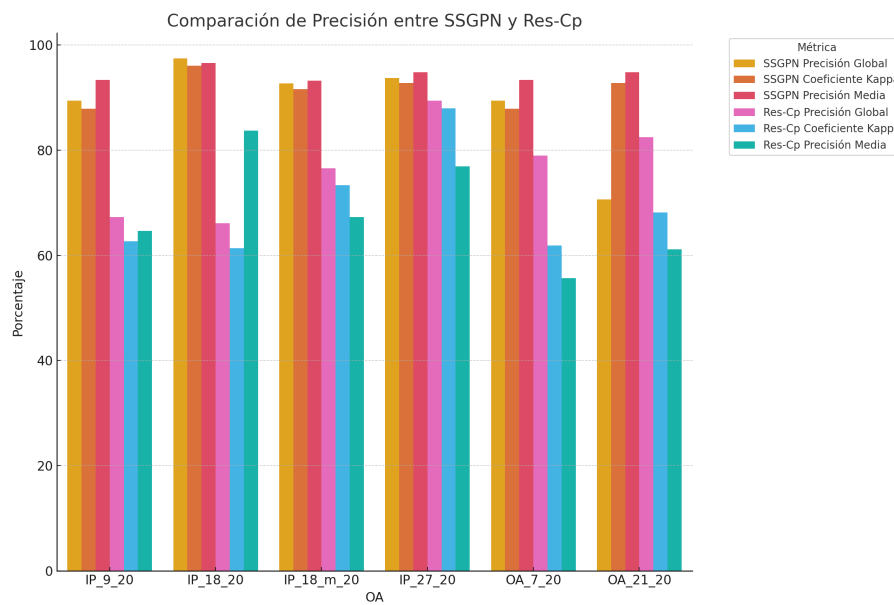


Figura 4.3: Gráfico de barras de comparación de precisión entre SSGPN y Res-CP.

# Conclusiones y Trabajo Futuro

En esta tesis de grado se propuso un estudio comparativo de técnicas semi-supervisadas para imágenes MSI, integrando información espectral-espacial. Se revisaron dos enfoques: uno basado en pseudo-etiquetas y otro en grafos probabilísticos. El primer enfoque tiende a ser más preciso a la hora de generalizar, sin requerir un ajuste fino en comparación con el segundo método. Además, se adapta más fácilmente a los conjuntos de clases en comparación con el método basado en grafos. Por otro lado, el segundo método, SSGPN, clasifica las clases de manera ordenada, conservando los bordes, a diferencia del primero, que no lo hace de manera tan precisa.

En cuanto al enfoque de características espaciales añadidas en los conjuntos de datos de Indian Pines y Oriente Antioqueño, se observa que Indian Pines, al ser un conjunto de datos más estudiado en el estado del arte, es más amigable para trabajar debido a la cantidad de investigaciones previas. En contraste, el segundo conjunto de datos tiende a ser más complejo a la hora de elegir hiperparámetros, y su capacidad de clasificación se complica. Además, al añadir la extracción de características espaciales, se mejoraron los resultados hasta en un 20% en el caso de REs-CP, mientras que en SSGPN hubo una disminución en la precisión del modelo de hasta un 20%. En general, el método de pseudo-etiquetas maneja mejor la incertidumbre de las etiquetas.

Trabajos futuros para enfoques semi-supervisados incluyen optimizar la búsqueda de hiperparámetros, logrando una metodología adecuada para explotar el potencial de estos métodos de inteligencia artificial. Además, se debería resaltar, mediante la extracción de características espaciales, las regiones de los bordes en las clases para obtener mejores resultados.

# Bibliografía

- [1] Jesper E Van Engelen and Holger H Hoos. A survey on semi-supervised learning. *Machine learning*, 109(2):373–440, 2020.
- [2] Yanbo Huang, Zhong-xin Chen, YU Tao, Xiang-zhi Huang, and Xing-fa Gu. Agricultural remote sensing big data: Management and applications. *Journal of Integrative Agriculture*, 17(9):1915–1931, 2018.
- [3] Yuliya Tarabalka, Jón Atli Benediktsson, and Jocelyn Chanussot. Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. *IEEE transactions on geoscience and remote sensing*, 47(8):2973–2987, 2009.
- [4] Majid Seydgar, Amin Alizadeh Naeini, Mengmeng Zhang, Wei Li, and Mehran Satari. 3-d convolution-recurrent networks for spectral-spatial classification of hyperspectral images. *Remote Sensing*, 11(7):883, 2019.
- [5] Majid Seydgar, Shahryar Rahnamayan, Pedram Ghamisi, and Azam Asilian Bidgoli. Semisupervised hyperspectral image classification using a probabilistic pseudo-label generation framework. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–18, 2022.
- [6] John A Richards and John A Richards. Sources and characteristics of remote sensing image data. *Remote Sensing Digital Image Analysis: An Introduction*, pages 1–37, 1993.
- [7] Rajendra P Sishodia, Ram L Ray, and Sudhir K Singh. Applications of remote sensing in precision agriculture: A review. *Remote Sensing*, 12(19):3136, 2020.
- [8] Floyd F Sabins Jr and James M Ellis. *Remote sensing: Principles, interpretation, and applications*. Waveland Press, 2020.
- [9] Benoit A Aubert, Andreas Schroeder, and Jonathan Grimaudo. It as enabler of sustainable farming: An empirical analysis of farmers’ adoption decision of

- precision agriculture technology. *Decision support systems*, 54(1):510–520, 2012.
- [10] Corporación Colombiana de Investigación Agropecuaria. Plan estratégico de ciencia, tecnología e innovación del sector agropecuario colombiano (2017-2027), 2016.
- [11] MINED VIGILADAS. Agenda regional de investigación y desarrollo tecnológico hacia la innovación. 2019.
- [12] Bobo Xi, Jiaojiao Li, Yunsong Li, and Qian Du. Semi-supervised graph prototypical networks for hyperspectral image classification. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 2851–2854, 2021.
- [13] Chen Shi and Le Wang. Incorporating spatial information in spectral unmixing: A review. *Remote Sensing of Environment*, 149:70–87, 2014.
- [14] Karbhari V Kale, Mahesh M Solankar, Dhananjay B Nalawade, Rajesh K Dhumal, and Hanumant R Gite. A research review on hyperspectral data processing and analysis algorithms. *Proceedings of the national academy of sciences, India section a: physical sciences*, 87:541–555, 2017.
- [15] Xiaorui Ma, Hongyu Wang, and Jie Wang. Semisupervised classification for hyperspectral image based on multi-decision labeling and deep feature learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 120:99–107, 2016.
- [16] Te Han, Dongxiang Jiang, Qi Zhao, Lei Wang, and Kai Yin. Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. *Transactions of the Institute of Measurement and Control*, 40(8):2681–2693, 2018.
- [17] Anyong Qin, Zhaowei Shang, Jinyu Tian, Yulong Wang, Taiping Zhang, and Yuan Yan Tang. Spectral–spatial graph convolutional networks for semisupervised hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 16(2):241–245, 2018.
- [18] David Ruiz, Bladimir Bacca, and Eduardo Caicedo. Hyperspectral images classification based on inception network and kernel pca. *IEEE Latin America Transactions*, 17(12):1995–2004, 2019.
- [19] Andreas Kamilaris, Andreas Kartakoullis, and Francesc X Prenafeta-Boldú. A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture*, 143:23–37, 2017.

- [20] Pedram Ghamisi, Javier Plaza, Yushi Chen, Jun Li, and Antonio J Plaza. Advanced spectral classifiers for hyperspectral images: A review. *IEEE Geoscience and Remote Sensing Magazine*, 5(1):8–32, 2017.
- [21] José M Bioucas-Dias, Antonio Plaza, Nicolas Dobigeon, Mario Parente, Qian Du, Paul Gader, and Jocelyn Chanussot. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE journal of selected topics in applied earth observations and remote sensing*, 5(2):354–379, 2012.
- [22] Jun Li, Inmaculada Dópido, Paolo Gamba, and Antonio Plaza. Complementarity of discriminative classifiers and spectral unmixing techniques for the interpretation of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5):2899–2912, 2014.
- [23] Nirmal Keshava and John F Mustard. Spectral unmixing. *IEEE signal processing magazine*, 19(1):44–57, 2002.
- [24] Yanzi Shi, Jiaojiao Li, Yunsong Li, and Qian Du. Sensor-independent hyperspectral target detection with semisupervised domain adaptive few-shot learning. *IEEE Transactions on Geoscience and Remote Sensing*, 59(8):6894–6906, 2020.
- [25] James Theiler, Amanda Ziemann, Stefania Matteoli, and Marco Diani. Spectral variability of remotely sensed target materials: Causes, models, and strategies for mitigation and robust exploitation. *IEEE Geoscience and Remote Sensing Magazine*, 7(2):8–30, 2019.
- [26] Xiaojin Jerry Zhu. Semi-supervised learning literature survey. 2005.
- [27] Sjoerd de Vries and Dirk Thierens. A reliable ensemble based approach to semi-supervised learning. *Knowledge-Based Systems*, 215:106738, 2021.
- [28] Islam Nassar, Samitha Herath, Ehsan Abbasnejad, Wray Buntine, and Gholamreza Haffari. All labels are not created equal: Enhancing semi-supervision via label grouping and co-training. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7241–7250, 2021.
- [29] Zdenek Kalal, Krystian Mikolajczyk, and Jiri Matas. Tracking-learning-detection. *IEEE transactions on pattern analysis and machine intelligence*, 34(7):1409–1422, 2011.

- [30] Michał Romaszewski, Przemysław Głomb, and Michał Cholewa. Semi-supervised hyperspectral classification from a small number of training samples using a co-training approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 121:60–76, 2016.
- [31] Haiqin Yang, Kaizhu Huang, Irwin King, and Michael R Lyu. Maximum margin semi-supervised learning with irrelevant data. *Neural Networks*, 70:90–102, 2015.
- [32] Lin He, Jun Li, Chenying Liu, and Shutao Li. Recent advances on spectral-spatial hyperspectral image classification: An overview and new guidelines. *IEEE Transactions on Geoscience and Remote Sensing*, 56(3):1579–1597, 2017.
- [33] Xiuping Jia, Bor-Chen Kuo, and Melba M Crawford. Feature mining for hyperspectral image classification. *Proceedings of the IEEE*, 101(3):676–697, 2013.
- [34] Carlos A Cattaneo, Ledda I Larcher, Ana I Ruggeri, Andrea C Herrera, and Enrique M Biasoni. Métodos de umbralización de imágenes digitales basados en entropía de shannon y otros. *Mecánica Computacional*, 30(36):2785–2805, 2011.
- [35] Amelia Carolina Sparavigna. Entropy in image analysis, 2019.
- [36] Stephen F Gull and John Skilling. Maximum entropy method in image processing. In *lee proceedings f (communications, radar and signal processing)*, volume 131, pages 646–659. IET, 1984.
- [37] Diya Chudasama, Tanvi Patel, Shubham Joshi, and Ghanshyam I Prajapati. Image segmentation using morphological operations. *International Journal of Computer Applications*, 117(18), 2015.
- [38] Yunho Jeon and Junmo Kim. Active convolution: Learning the shape of convolution for image classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4201–4209, 2017.
- [39] Pierre Soille et al. *Morphological image analysis: principles and applications*, volume 2. Springer, 1999.
- [40] Maria Casamitjana, Maria C Torres-Madroño, Jaime Bernal-Riobo, and Diego Varga. Soil moisture analysis by means of multispectral images according to land use and spatial resolution on andosols in the colombian andes. *Applied Sciences*, 10(16):5540, 2020.

- [41] Bobo Xi, Jiaojiao Li, Yunsong Li, Rui Song, Yuchao Xiao, Qian Du, and Jocelyn Chanussot. Semisupervised cross-scale graph prototypical network for hyperspectral image classification. *IEEE Transactions on Neural Networks and Learning Systems*, 34(11):9337–9351, 2023.
- [42] Hong Hui Tan and King Hann Lim. Vanishing gradient mitigation with deep learning neural network optimization. In *2019 7th international conference on smart computing & communications (ICSCC)*, pages 1–4. IEEE, 2019.
- [43] Yandong Wen, Kaipeng Zhang, Zhifeng Li, and Yu Qiao. A discriminative feature learning approach for deep face recognition. In *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VII 14*, pages 499–515. Springer, 2016.
- [44] Bobo Xi, Jiaojiao Li, Yunsong Li, and Qian Du. Semi-supervised graph prototypical networks for hyperspectral image classification. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 2851–2854, 2021.
- [45] Lichao Mou, Xiaoqiang Lu, Xuelong Li, and Xiao Xiang Zhu. Nonlocal graph convolutional networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 58(12):8246–8257, 2020.
- [46] Danfeng Hong, Lianru Gao, Jing Yao, Bing Zhang, Antonio Plaza, and Jocelyn Chanussot. Graph convolutional networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(7):5966–5978, 2020.
- [47] Jie Liang, Jun Zhou, Yuntao Qian, Lian Wen, Xiao Bai, and Yongsheng Gao. On the sampling strategy for evaluation of spectral-spatial methods in hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2):862–880, 2016.
- [48] Abbas Kiani and Mahmood Reza Sahebi. Edge detection based on the shannon entropy by piecewise thresholding on remote sensing images. *IET Computer Vision*, 9(5):758–768, 2015.
- [49] Fernando Giménez Palomares, Juan A Monsoriu, and Elena Alemany. Aplicación de la convolución de matrices al filtrado de imágenes. *Modelling in Science Education and Learning*, 9(1):97–108, 2016.
- [50] Zilong Zhong, Jonathan Li, Zhiming Luo, and Michael Chapman. Spectral-spatial residual network for hyperspectral image classification: A 3-d

- deep learning framework. *IEEE Transactions on Geoscience and Remote Sensing*, 56(2):847–858, 2018.
- [51] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019.
- [52] Tijmen Tieleman. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural networks for machine learning*, 4(2):26, 2012.

