



Facultad de Ingeniería
Tópicos Avanzados de Analítica

CamilaAndrea Arias Vargas
Felipe Clavijo Acosta
Joel Alfredo Márquez Álvarez
Juan Camilo Ramírez Restrepo
Juan Pablo Cuellar Solano

Contenido

1	Entendimiento del Negocio.....	2
2	Objetivos de Negocio.....	2
3	Objetivo de minería	3
4	Entendimiento de los datos	3
4.1.	Dataset de entrenamiento:	3
4.2.	Imagen Objetivo:	4
5	Preparación de los datos.....	5
6	Modelos.....	6
6.1.	Red Neuronal Sencilla.....	6
6.2.	Red Neuronal Multi – Capa.....	6
6.3.	Red Neuronal Convolutiva.....	7
6.4.	Red Neuronal Convolutiva con Transferencia VGG-16	8
6.5.	Red Neuronal Convolutiva con VGG-16 y optimizador Optuna.....	8
6.6.	Red Neuronal Convolutiva con ResNet-50.....	9
7	Evaluación y selección.....	10
8	Resultados	11

FRAILEJON DETECTION (a.k.a "Big Monks Detection")

1 Entendimiento del Negocio

El proyecto se plantea como eje central la conservación y el monitoreo del ecosistema del páramo, un entorno natural muy importante por su regulación del recurso hídrico y su biodiversidad singular. Los páramos de Chingaza y Cruz Verde, en particular, son áreas que albergan la especie Espeletia, conocida comúnmente como frailejones. Estas plantas desempeñan un papel esencial en la captación y almacenamiento de agua, lo que las convierte en un elemento clave para la sostenibilidad ambiental y el bienestar de las comunidades que dependen de los recursos hídricos que el páramo provee.

El objetivo principal del proyecto es desarrollar e implementar una metodología innovadora que emplee redes neuronales para la detección automática de frailejones en imágenes aéreas de alta resolución. Este enfoque permitirá una identificación precisa de las áreas donde la densidad de frailejones es baja, lo que es fundamental para la planificación de acciones de restauración ecológica. La tecnología propuesta busca automatizar el proceso de monitoreo, lo que no solo mejorará la eficiencia y precisión de la detección, sino que también reducirá la dependencia de métodos manuales, que suelen ser más costosos y menos efectivos.

El criterio de éxito del proyecto se definirá por la capacidad del modelo de redes neuronales para identificar con una precisión superior al 80% las áreas de baja densidad de frailejones en las imágenes aéreas analizadas.

El impacto esperado de este proyecto es considerable, ya que no solo optimiza los recursos destinados a la conservación del páramo, sino que también ofrece un modelo replicable que puede ser aplicado en otros páramos de Colombia y del mundo. De este modo, el proyecto no solo fortalece las estrategias de conservación local, sino que también aporta al esfuerzo global de preservar estos ecosistemas vitales, con beneficios que se extienden más allá de la región estudiada.

2 Objetivos de Negocio

1. Reducir los costos operativos asociados a la identificación, cuantificación y localización de la especie Espeletia (frailejón) en los páramos de Chingaza y Cruz Verde mediante la implementación de un sistema automatizado de procesamiento de imágenes aéreas capturadas por drones, comparando los costos actuales del proceso manual con los del nuevo enfoque automatizado.

$$\text{Reducción de costos} = (\text{Costos del método tradicional} - \text{Costos del método automatizado}) / \text{Costos del método tradicional} \times 100\%$$

2. Identificar las áreas del páramo con menor densidad poblacional de frailejones para priorizar y ejecutar acciones de restauración y conservación de esta especie clave para el ecosistema.

$$\text{Factor de densidad} = (\text{Densidad poblacional de la subárea} / \text{Densidad poblacional media del páramo}) * 100\%$$

3 Objetivo de minería

1. Diseñar un modelo de aprendizaje supervisado usando arquitectura de redes convolucionales para el procesamiento de imágenes con el fin de identificar y cuantificar la presencia de frailejones en el páramo de Chingaza.

El modelo desarrollado debe alcanzar un AUC (Área Bajo la Curva ROC) mínimo de 0.89, asegurando así su capacidad para distinguir correctamente los frailejones

4 Entendimiento de los datos

4.1. Dataset de entrenamiento:

Cantidad de imágenes totales: 250.

Tamaño de píxeles: 14700 para todas las imágenes aplanadas.

Canales: Las imágenes están en formato de 3 canales RGB.

Imágenes: Muestra de ambas categorías.

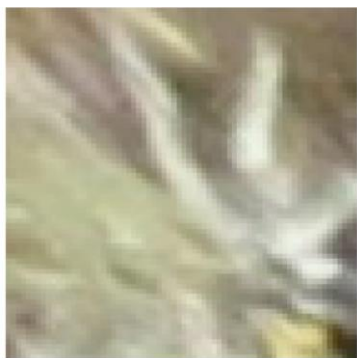
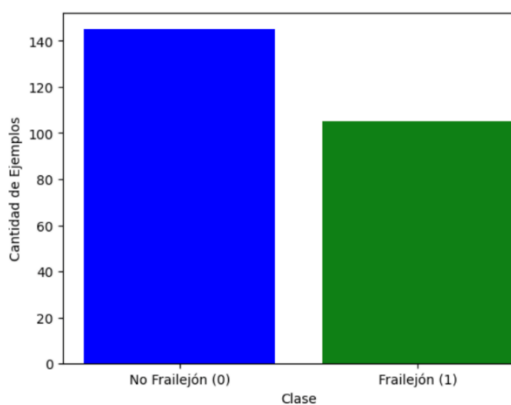


Imagen 1. Clase 1.

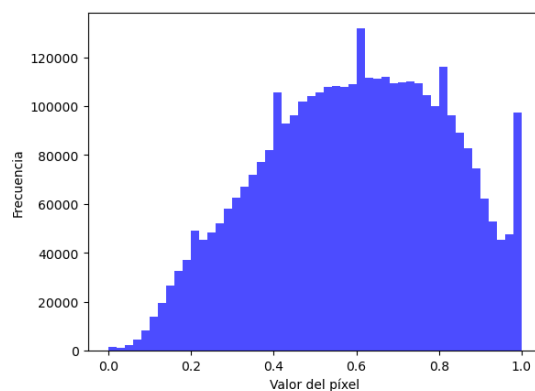


Imagen 2. Clase 0.

Balanceo de la data: Se realizó un análisis de las distribuciones de las etiquetas mostrando lo siguiente: Frailejones 105 (clase 1), No frailejones 145 (clase 0). Por lo que se evidencia un desbalance entre clases.

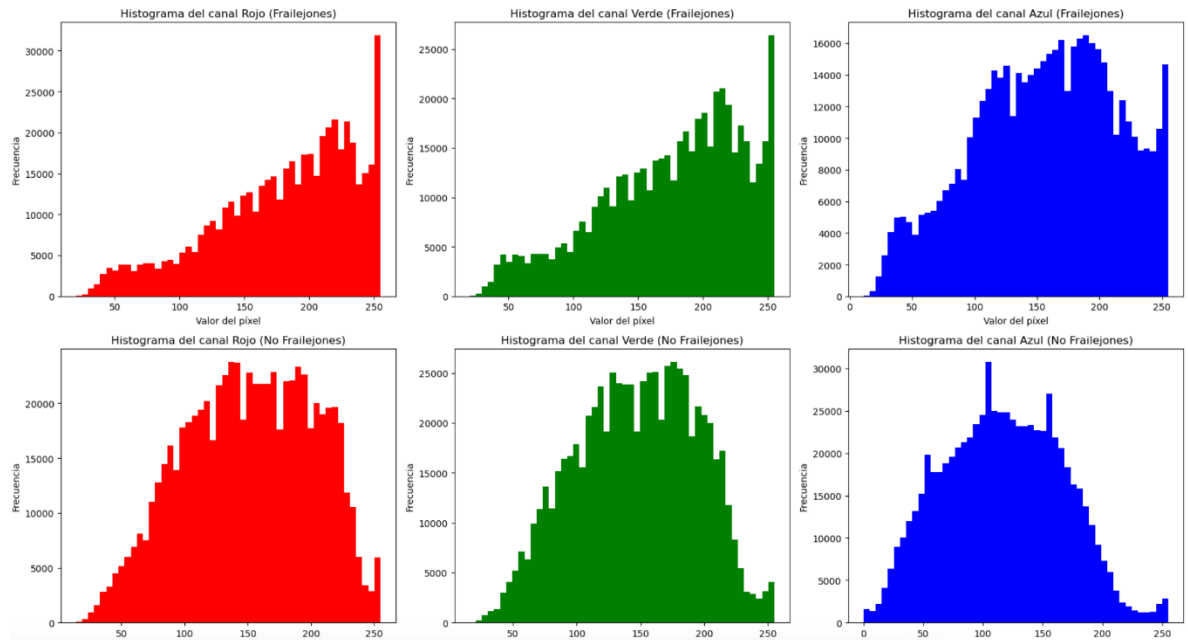


Gráfica 2. Distribución de las etiquetas.



Gráfica 3. Distribución de las clases en el conjunto de datos.

La mayor parte de los píxeles se encuentra en tonos medios en un rango de 0.4 y 0.8, los que sugiere que los frailejones y el resto de vegetación tiene una mezcla de tonos intermedios típicos del ecosistema donde se encuentra la especie.



Gráfica 4. Distribución por canal (RGB) para cada categoría.

Se realizó un análisis de los histogramas de ambas clases, identificando que en las imágenes de frailejones (clase 1) hay una predominancia de tonos brillantes en los canales rojo, verde y azul. Los píxeles tienden a concentrarse en valores altos, lo que sugiere la presencia de luz solar directa sobre las hojas y otras partes de los frailejones. En particular, el canal azul muestra una mayor variedad de tonalidades, con presencia tanto de áreas brillantes como oscuras, lo que podría corresponder a zonas de sombra y cielo en el entorno.

Por otro lado, las imágenes de no frailejones (clase 0) presentan una distribución más equilibrada de los colores, con menos concentración de píxeles en los valores más brillantes. En los canales rojo y verde, los tonos están más repartidos en los rangos medios, lo que indica una mayor diversidad en la intensidad del color. El canal azul muestra un pico más notable en los tonos medios, lo que refleja áreas más homogéneas en las zonas con tonalidades azules.

La concentración de tonos brillantes en las imágenes de frailejones podría ser un factor clave para que la red neuronal aprenda a distinguir entre ambas clases. Sin embargo, para evitar que el modelo se sesgue hacia estas zonas más brillantes, se debería aplicar una normalización en los datos. Esto ayudaría a equilibrar la intensidad de los píxeles y garantizar que la red no se enfoque excesivamente en las áreas brillantes de los frailejones, logrando así un mejor rendimiento general.

4.2. Imagen Objetivo:

Dimensiones: 4000x3000

Canales: Las imágenes están en formato de 3 canales RGB

Modelo del dispositivo: Canon EOS Rebel T5i

Marca del dispositivo: Canon

Longitud focal: 50mm

Perfil de color: sRGB IEC61966-2.1

Ojos rojos: No

Número f/: f/4

Modo de exposición: Manual

Programa de medición: Patrón

Tiempo de exposición: 1/1000 sec.

Latitud: 40°34'26.79"N

Longitud: 74°1'22.859"O

Imagen:



Imagen 3. Vista aérea de páramo, capturada por dron.

La vegetación predominante está compuesta por frailejones, que son plantas altas y robustas con rosetas de hojas gruesas, adaptadas a ambientes de gran altitud. El paisaje presenta diversas tonalidades de verde y marrón, indicando diferentes tipos de vegetación y suelo. No se observan caminos ni estructuras hechas por el hombre, lo que sugiere un hábitat natural y sin perturbaciones.

5 Preparación de los datos

Para la preparación de los datos se inicia con la extracción y separación de los datos, donde se extraen dos conjuntos de datos que se encuentran presentes en el archivo *Data_F* del repositorio en GitHub, importando la información que se tiene disponible en formato H5PY. En el proceso de extracción se divide la información en 2, una correspondiente a las imágenes de frailejones y la segunda a los casos en que no se tiene frailejones.

Después de la extracción de los conjuntos se crean los arreglos donde se almacenarán las imágenes que se procesen, que se segmentan en grupos de imágenes más pequeñas para procesar trabajar con mayor detalle, buscando aumentar la cantidad de datos a entrenar. Este paso es importante, ya que multiplica la cantidad de datos y características que se pueden extraer de las imágenes.

Una vez se tienen las diferentes extracciones de imágenes, se arreglan para que puedan manipularse de manera más eficiente, pasando de listas a arreglos. Estas listas se concatenan, consolidando la información de los arreglos de los frailejones y no frailejones, para reestructurar el arreglo y lograr aplanar las imágenes contenidas en estos, para preparar los datos para que puedan usarse dentro de los modelos. Con las imágenes aplanadas, se normalizan los datos dividiendo en 255, para llevar la escala de pixeles a un rango entre 0 y 1. De igual forma con el fin de identificar los frailejones de los no frailejones, se realiza la etiqueta de estos, etiquetando con 1 a los frailejones y 0 a los no frailejones.

Adicionalmente se efectuó una partición de datos en entrenamiento 80% y prueba 20% de los datos, para luego realizar una segunda partición y obtener una muestra del 10% de la data de entrenamiento para realizar el proceso de validación de las redes neuronales.

6 Modelos

Para el ejercicio de procesamiento de imágenes y la identificación de frailejones en la imagen objetivo, se implementaron los siguientes métodos, acompañados de sus respectivas métricas:

6.1. Red Neuronal Sencilla

La Red Neuronal Sencilla consiste en una capa de entrada, una capa oculta y una capa de salida. Este tipo de red es adecuada para problemas donde la complejidad de los datos no requiere múltiples capas o gran capacidad de procesamiento. Su uso es común en tareas simples de clasificación.

Arquitectura	Entrenamiento y Optimización	Métricas
Capa Oculta 128 neuronas Activación tanh Capa de Salida 1 neurona Activación sigmoide	Tasa Aprendizaje: 0.01 Función de pérdida: Entropía cruzada. Optimización: Gradiente descendente, con retropropagación para actualizar los pesos y sesgos.	Coste Train: 0.4056 Coste Validation: 0.5484 Accuracy Train: 0.8889 Accuracy Validation: 0.75 AUC Test: 0.8302

6.2. Red Neuronal Multi – Capa

La Red Neuronal Multicapa consta de múltiples capas ocultas entre la entrada y la salida, lo que le permite aprender representaciones más complejas y patrones no lineales. Estas redes son ideales para tareas que requieren una mayor capacidad de ajuste y son efectivas en una amplia gama de problemas de clasificación y regresión.

Arquitectura	Entrenamiento y Optimización	Métricas
Capa Oculta 1	Tasa Aprendizaje: 0.01	Coste Train: 0.4238

128 neuronas Activación tanh Capa Oculta 2 128 neuronas Activación tanh Capa Oculta 3 64 neuronas Activación tanh Capa de Salida 1 neurona Activación sigmoide	Función de pérdida: Entropía cruzada. Optimización: Gradiente descendente, con retropropagación para actualizar los pesos y sesgos.	Coste Validation: 0.5632 Accuracy Train: 0.8833 Accuracy Validation: 0.75 AUC Test: 0.8727
---	--	---

6.3. Red Neuronal Convolutacional

La Red Neuronal Convolutacional está diseñada para procesar datos estructurados en forma de cuadrícula, como imágenes. Utiliza capas convolucionales para extraer características espaciales y patrones jerárquicos, lo que la hace extremadamente efectiva para tareas como la clasificación de imágenes y el reconocimiento de objetos.

Arquitectura	Entrenamiento y Optimización	Métricas
Capas Convolucionales Conv 1: 64 filtros, Kernel 7x7, BatchNorm y activación ReLU. Conv 2: 128 filtros, Kernel 3x3, BatchNorm y activación ReLU. Conv 3: 128 filtros, Kernel 3x3, BatchNorm y activación ReLU. Conv 4: 256 filtros, Kernel 3x3, BatchNorm y activación ReLU. Conv 5: 256 filtros, Kernel 3x3, BatchNorm y activación ReLU. Capas de Pooling MaxPooling de 2x2 después de conv1, conv3 y conv5, para reducir la dimensionalidad.	Función de pérdida: Se utiliza BCEWithLogitsLoss para tareas de clasificación binaria. Optimizador: Adam, con una tasa de aprendizaje inicial de 0.001. Ajustes de Datos: Se aplicaron transformaciones de datos como rotación aleatoria, ajuste de color y recorte aleatorio para aumentar la diversidad del conjunto de entrenamiento.	Coste Validation: 0.3423 AUC Validation: 0.9823 Coste Test: 0.3423 AUC Test: 0.9823

Capas Totalmente Conectadas: FC1: 128 neuronas, activación ReLU. FC2: 64 neuronas, activación ReLU. FC3: 1 neurona (salida binaria). Dropout: Con un factor de 0.5 en las capas FC1 y FC2 para prevenir el sobreajuste.		
--	--	--

6.4. Red Neuronal Convolutiva con Transferencia VGG-16

La Red Neuronal Convolutiva con Transferencia VGG16 aprovecha los pesos preentrenados de VGG16, una arquitectura profunda con 16 capas, para reutilizar características previamente aprendidas en tareas de clasificación de imágenes. Esto acelera el entrenamiento y mejora el rendimiento al aplicar el modelo a nuevos datos.

Arquitectura	Entrenamiento y Optimización	Métricas
Capas Convolucionales Preentrenadas (VGG16): Las capas convolucionales están congeladas, por lo que no se actualizan durante el entrenamiento. Estas capas ya han sido preentrenadas en el conjunto de datos ImageNet. Capa Totalmente Conectada Modificada: La última capa de la red VGG16 se reemplaza por una capa completamente conectada con una única salida (1 neurona) para la clasificación binaria.	Optimización de Hiperparámetros: Tasa de aprendizaje (lr): Buscada entre $1e-5$ y $1e-2$. Momentum: Sugerido en el rango 0.5 - 0.9. Step Size: Entre 3 y 10 épocas para el ajuste de la tasa de aprendizaje. Gamma: Factor de reducción de la tasa de aprendizaje entre 0.1 - 0.9. Optimizador: Stochastic Gradient Descent (SGD). Scheduler: StepLR ajusta la tasa de aprendizaje cada <code>step_size</code> épocas, multiplicando por el factor gamma. Función de pérdida: Entropía cruzada binaria (BCEWithLogitsLoss).	Coste Validation: 0.2868 Validation AUC: 0.9583 Coste Test: 0.1689 AUC Test: 1.0000

6.5. Red Neuronal Convolutiva con VGG-16 y optimizador Optuna

El uso de VGG-16 preentrenada aprovecha características visuales aprendidas previamente, reduciendo el tiempo de entrenamiento y mejorando la precisión del modelo. Al combinarla con

Optuna, que optimiza automáticamente los hiperparámetros, se logra un ajuste eficiente que maximiza el rendimiento del modelo y evita el sobreajuste, facilitando una clasificación más precisa y rápida.

Arquitectura	Entrenamiento y Optimización	Métricas
<p>Capas Convolucionales Preentrenadas (VGG16): Las capas convolucionales están congeladas, por lo que no se actualizan durante el entrenamiento. Estas capas ya han sido preentrenadas en el conjunto de datos ImageNet.</p> <p>Capa Totalmente Conectada Modificada: La última capa de la red VGG16 se reemplaza por una capa completamente conectada con una única salida (1 neurona) para la clasificación binaria.</p>	<p>Optimización de Hiperparámetros (Optuna):</p> <p>Tasa de aprendizaje (lr): 0.0008502420354823822</p> <p>Momentum: 0.8392910118706614</p> <p>Step Size: 5</p> <p>Gamma: 0.7694952780723903</p> <p>Optimizador: Stochastic Gradient Descent (SGD) con los parámetros ajustados por Optuna.</p> <p>Scheduler: StepLR ajusta la tasa de aprendizaje cada step_size épocas, multiplicando por el factor gamma.</p> <p>Función de pérdida: Entropía cruzada binaria (BCEWithLogitsLoss).</p>	<p>Coste Validation: 0.2806</p> <p>Validation AUC: 0.9583</p> <p>Coste Test: 0.1601</p> <p>Test AUC: 1.0000</p>

6.6. Red Neuronal Convolutacional con ResNet-50

La arquitectura preentrenada ResNet-50 aprovecha una estructura de red profunda con bloques residuales, que permiten entrenar redes muy profundas sin sufrir problemas de desvanecimiento de gradientes. Al estar preentrenada en bases de datos extensas como ImageNet, ResNet-50 ya ha aprendido a extraer características útiles de las imágenes, lo que reduce significativamente el tiempo de entrenamiento y mejora la precisión. Además, su capacidad para capturar tanto características de bajo como de alto nivel la hace altamente efectiva en tareas complejas de clasificación.

Arquitectura	Entrenamiento y Optimización	Métricas
<p>Capas Convolucionales Preentrenadas (RESNET50): Las capas convolucionales están congeladas, por lo que no se actualizan durante el entrenamiento. Estas capas ya han sido preentrenadas en el</p>	<p>Optimización de Hiperparámetros:</p> <p>Tasa de aprendizaje (lr): Buscada entre 1e-5 y 1e-2.</p> <p>Momentum: Sugerido en el rango 0.5 - 0.9.</p> <p>Step Size: Entre 3 y 10 épocas para el ajuste de la tasa de aprendizaje.</p>	<p>Validation Loss: 0.4077</p> <p>Validation AUC: 0.9167</p> <p>Test Loss: 0.2651</p> <p>Test AUC: 0.9924</p>

conjunto de datos ImageNet. Capa Totalmente Conectada Modificada: La última capa de la red RESNET50 se reemplaza por una capa completamente conectada con una única salida (1 neurona) para la clasificación binaria.	Gamma: Factor de reducción de la tasa de aprendizaje entre 0.1 - 0.9. Optimizador: Stochastic Gradient Descent (SGD) con los parámetros. Scheduler: StepLR ajusta la tasa de aprendizaje cada step_size épocas, multiplicando por el factor gamma. Función de pérdida: Entropía cruzada binaria (BCEWithLogitsLoss).	
---	---	--

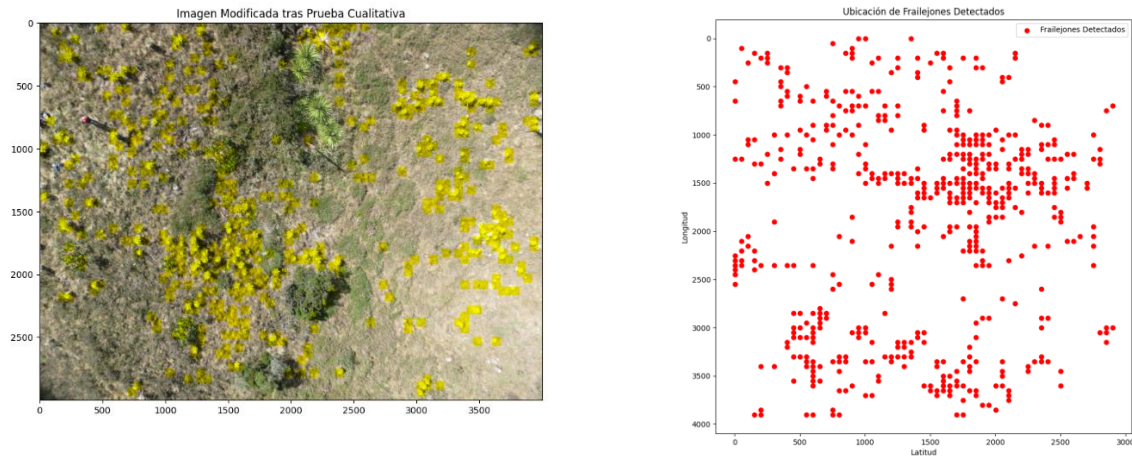
7 Evaluación y selección

Modelo	AUC Test
Red Neuronal Sencilla	0.8302
Red Neuronal Multi – Capa	0.8727
Red Neuronal Convolutacional	0.9823
Red Neuronal Convolutacional con Transferencia VGG-16	1.0000
Red Neuronal Convolutacional con VGG-16 y optimizador Optuna	1.0000
Red Neuronal Convolutacional con ResNet-50	0.9924

El mejor modelo evaluado en este trabajo es el que utiliza VGG16 pre-entrenado con optimización de hiperparámetros mediante Optuna, alcanzando un AUC de 1.0000 en validación, lo que demuestra una excelente capacidad para discriminar entre clases. Este rendimiento superior se debe principalmente al uso de transferencia de aprendizaje, que aprovecha características previamente aprendidas en el conjunto de datos ImageNet, mejorando la generalización y reduciendo el tiempo de entrenamiento. Además, la optimización de hiperparámetros clave, como la tasa de aprendizaje y el momentum, a través de Optuna permitió maximizar el rendimiento del modelo. La combinación de estas técnicas, junto con la regularización mediante StepLR y L2, ayudó a controlar el sobreajuste y mejorar la estabilidad durante el entrenamiento, lo que hace que este modelo sea el más eficiente y efectivo comparado con los otros evaluados.

8 Resultados

Tras seleccionar el mejor modelo (Red Neuronal Convolutacional con VGG-16 y optimizador Optuna), se realiza una prueba cualitativa para calcular el total de frailejones en la imagen, lo que permitirá determinar su densidad. Para ello, se ejecuta nuevamente el loop, ajustando ligeramente los parámetros para contar con precisión el número de frailejones y ubicarlos espacialmente según su latitud y longitud.



El objetivo de este proceso es identificar cuántos grupos de frailejones hay en una imagen, analizando sus posiciones. Para ello, se prueban varias opciones de agrupamiento y se selecciona la que mejor se ajuste, sin hacer el modelo innecesariamente complicado. Esto permite obtener una visión clara de cómo están distribuidos los frailejones, lo que puede ser útil para medir su densidad o estudiar su comportamiento en el terreno, encontrando un total de 681 frailejones en la imagen objetivo.

