Green Xpo Lab y Costa Rica FlyingLabs

Implementación de Modelos de Clasificación para Banano y Plátano

Informe Ejecutivo

Autor: Carlos Saito y Hanzel León

Cartago, Enero de 2025

Resumen

La agricultura de precisión exige herramientas avanzadas para identificar y monitorear cultivos específicos, como las plantaciones de plátano y banano. Este proyecto propone un sistema basado en aprendizaje automático que utiliza segmentación por píxeles y clasificación a partir de imágenes satelitales para detectar estas plantaciones con alta precisión. Los resultados destacan métricas sobresalientes, como un factor Dice superior a 0.85 y un accuracy mayor al 93%, junto con valores F1 y recall por encima de 0.85. Aunque se identificaron áreas de mejora, como la consistencia de los datos, el etiquetado y la cantidad de imágenes, el enfoque desarrollado demuestra ser una herramienta eficaz para optimizar el monitoreo agrícola, contribuyendo a una gestión más sostenible de los cultivos.

Palabras clave: Imágenes satelitales, machine learning, segmentación

Abstract

Precision agriculture requires advanced tools to identify and monitor specific crops, such as banana and plantain plantations. This project proposes a machine learning-based system that utilizes pixel-wise segmentation and classification from satellite images to accurately detect these plantations. The results highlight outstanding performance, with a Dice coefficient above 0.85, accuracy exceeding 93%, and F1 and recall scores above 0.85. Although areas for improvement were identified, including data consistency, labeling, and the number of images, the developed approach proves to be an effective tool for optimizing agricultural monitoring, contributing to more sustainable crop management.

Keywords: Satellite images, machine learning, segmentation

Índice general

Índice de figuras iii

Índice de tablas iv

Lista de símbolos y abreviaciones v

1 Introducción 1
- 1.1 Descripción general del proyecto . . . 1
- 1.2 Planteamiento del problema . . . 2
  - 1.2.1 Contexto del problema . . . 2
  - 1.2.2 Justificación del problema . . . 3
  - 1.2.3 Definición concreta del problema . . . 3
- 1.3 Objetivos del proyecto . . . 3
  - 1.3.1 Objetivo general . . . 4
  - 1.3.2 Objetivos específicos . . . 4

2 Marco teórico 5
- 2.1 Segmentación por Píxel . . . 7
- 2.2 Herramienta ArcGIS y Modelos de Machine Learning . . . 9

3 Modelos para la detección de musáceas 10
- 3.1 Preprocesamiento de Datos . . . 10
- 3.2 Arquitectura de los Modelos . . . 10
  - 3.2.1 Modelo U-Net . . . 11
  - 3.2.2 Backbone: ResNet34 . . . 13
- 3.3 Entrenamiento y Evaluación . . . 13
- 3.4 Datos de entrenamiento . . . 13

4 Resultados y análisis 15
- 4.1 Resultados para el modelo 1 (recomendado para imágenes iluminadas) . . . 15
  - 4.1.1 Métricas . . . 15
  - 4.1.2 Resultados visuales . . . 19
- 4.2 Resultados para el modelo 2 (recomendado para imágenes con poca iluminación) . . . 20
  - 4.2.1 Métricas . . . 20

4.2.2 Resultados visuales 24

5 Conclusiones **27**

5.1 Sugerencias y Recomendaciones 28

Bibliografía **30**

Índice de figuras

Fig. 1 Segmentación por pixel 9

Fig. 2 Arquitectura del modelo utilizando autoencoders 12
Fig. 3 Ejemplo de imágenes del data set 14

Fig. 4 Accuracy del modelo 1 16
Fig. 5 Training loss (error) del modelo 1 16
Fig. 6 Validation loss (error) del modelo 1 17
Fig. 7 Coeficiente Dice del modelo 1 18
Fig. 8 Comparación del resultado 1 del modelo 1 19
Fig. 9 Comparación del resultado 2 del modelo 1 19
Fig. 10 Muestras de testeo a, b y c del modelo 1. 20
Fig. 11 Accuracy del modelo 2 21
Fig. 12 Training loss (error) del modelo 2 21
Fig. 13 Validation loss (error) del modelo 2 22
Fig. 14 Coeficiente Dice del modelo 2 24
Fig. 15 Comparación del resultado 1 del modelo 2 24
Fig. 16 Comparación del resultado 2 del modelo 2 25
Fig. 17 Muestras de testeo a, b y c del modelo 2. 25

Índice de tablas

Tabla 2 Valores de métricas para el modelo 1 17
Tabla 3 Valores de métricas para el modelo 2 23

Lista de símbolos y abreviaciones

Abreviaciones

IA	Intelligence Artificial
ML	Machine Learning

Capítulo 1

Introducción

En un mundo donde la tecnología impulsa la innovación en sectores tradicionales como la agricultura, las herramientas de aprendizaje profundo emergen como aliadas cruciales para resolver problemas complejos de clasificación y monitoreo.

El machine learning desempeña un papel crucial en la agricultura de precisión, particularmente en la clasificación de cultivos, al ofrecer herramientas avanzadas para analizar grandes volúmenes de datos agroclimáticos y de imágenes satelitales o de drones. Estas técnicas permiten identificar y clasificar de manera eficiente diferentes tipos de cultivos en vastas extensiones de terreno. Por medio de algoritmos de aprendizaje supervisado y no supervisado, es posible detectar patrones en imágenes multiespectrales o hiperespectrales, lo que facilita distinguir entre variedades de cultivos, etapas de crecimiento y condiciones de salud, incluso en áreas geográficamente complejas. Esta capacidad de análisis detallado contribuye a la toma de decisiones informadas sobre la gestión del suelo, el riego y el uso de fertilizantes entre otros.

Además, la clasificación precisa de cultivos mediante machine learning beneficia tanto la sostenibilidad como la productividad agrícola. Al identificar áreas específicas con cultivos particulares, los agricultores pueden optimizar sus estrategias de manejo y reducir costos operativos, minimizando el desperdicio de recursos. Por ejemplo, el uso de modelos de clasificación permite predecir los rendimientos y detectar problemas como plagas o enfermedades en etapas tempranas, lo que mejora significativamente la calidad de los cultivos. Asimismo, esta tecnología contribuye a un monitoreo en tiempo real más eficiente, permitiendo respuestas rápidas a cambios en las condiciones del entorno y maximizando la resiliencia ante eventos climáticos adversos.

1.1 Descripción general del proyecto

El proyecto consiste en un sistema basado en machine learning para procesar imágenes satelitales y detectar de manera precisa las áreas donde se encuentran plantaciones de banano y plátano. Este sistema utiliza técnicas avanzadas de aprendizaje automático,

combinadas con análisis de imágenes, para clasificar y segmentar las regiones específicas de cultivo dentro de grandes extensiones de terreno. La herramienta resultante permitirá identificar visualmente estas plantaciones a través de representaciones gráficas en las imágenes procesadas, facilitando así la localización exacta de los cultivos. Este enfoque combina tecnología de vanguardia con la necesidad de optimizar el monitoreo agrícola, proporcionando una solución eficiente para la clasificación de cultivos en escenarios complejos.

La implementación del sistema contempla diversas etapas clave, que incluyen la adquisición de datos satelitales, el preprocesamiento de las imágenes para resaltar características relevantes, y la aplicación de algoritmos de machine learning para la clasificación y segmentación. El modelo entrenado será validado para asegurar su precisión y confiabilidad. Este proyecto no solo tiene un impacto directo en la agricultura de precisión, al optimizar la gestión y monitoreo de los cultivos de banano y plátano, sino que también abre la puerta a la integración de tecnologías inteligentes para mejorar la sostenibilidad y productividad del sector agrícola.

1.2 Planteamiento del problema

1.2.1 Contexto del problema

La agricultura de precisión enfrenta desafíos significativos en la identificación y monitoreo de cultivos específicos, como las plantaciones de plátano y banano, debido a la falta de herramientas eficientes y automatizadas para esta tarea. Actualmente, los métodos tradicionales de detección y monitoreo requieren inspecciones manuales o el uso de sistemas poco precisos, que suelen ser costosos y demandan mucho tiempo. Este escenario es especialmente crítico para cultivos como el plátano y el banano, que tienen un alto valor económico y representan una fuente fundamental de ingreso para muchas regiones agrícolas, pero cuya gestión se complica debido a la dificultad de diferenciarlos visualmente en grandes extensiones.

Las condiciones cambiantes del entorno agrícola, como el incremento de plagas, las enfermedades, y los impactos del cambio climático, agravan el problema al requerir respuestas rápidas y precisas. La incapacidad para identificar de forma oportuna las plantaciones específicas puede llevar a decisiones incorrectas o tardías, afectando directamente la productividad y la sostenibilidad de los cultivos. Además, las fluctuaciones en la demanda del mercado y las necesidades de exportación de plátano y banano exigen una gestión más eficiente, basada en datos confiables y actualizados.

Otro desafío importante es la precisión limitada de las herramientas actuales para analizar imágenes satelitales, las cuales pueden carecer de la capacidad para diferenciar cultivos visualmente similares. Esto crea la necesidad de implementar soluciones tecnológicas innovadoras que permitan mejorar la exactitud y eficiencia en la identificación de estos

cultivos, integrando capacidades avanzadas de procesamiento de imágenes y algoritmos de machine learning que puedan abordar las particularidades de este problema agrícola.

1.2.2 Justificación del problema

La implementación de un sistema automatizado basado en machine learning para la detección y clasificación de plantaciones de plátano y banano tiene un impacto significativo en la sostenibilidad y productividad agrícola. Este enfoque permitirá optimizar el monitoreo de cultivos al proporcionar datos más precisos y en tiempo real, mejorando la capacidad de respuesta ante problemas como enfermedades, plagas o cambios climáticos inesperados. Además, al identificar de manera más eficiente las plantaciones, los agricultores podrán tomar decisiones informadas que incrementen el rendimiento de sus cosechas y reduzcan los costos operativos asociados con inspecciones manuales o intervenciones tardías.

Por otro lado, el desarrollo de esta solución responde a la creciente necesidad de integrar tecnologías avanzadas en el sector agrícola, promoviendo una agricultura de precisión que sea más sostenible y competitiva en mercados globales. Este proyecto no solo mejorará las prácticas de monitoreo y gestión de cultivos, sino que también contribuirá a mitigar los riesgos económicos asociados con la pérdida de cosechas o la ineficiencia en la planificación agrícola. Así, el uso de machine learning para resolver este problema representa una inversión estratégica que beneficiará tanto a productores como a las cadenas de suministro agrícolas en general.

1.2.3 Definición concreta del problema

El problema que se busca resolver radica en la falta de precisión de identificar, de una manera eficiente, las plantaciones de plátano y banano. Esto podría disminuir la capacidad de reacción ante eventos inesperados, como plagas, condiciones climáticas adversas o fluctuaciones en la demanda, lo que afecta negativamente la planificación y la gestión de los cultivos. La solución propuesta es el desarrollo de un sistema automatizado basado en aprendizaje automático para la detección y clasificación de plantas de plátano y banano. La pregunta que guía este proyecto es: ¿Cómo desarrollar un modelo de aprendizaje automático que permita diferenciar plantas de plátano y banano a partir de imágenes?

1.3 Objetivos del proyecto

En esta sección se detallan los objetivos del proyecto.

1.3.1 Objetivo general

Desarrollar un modelo de aprendizaje automático (ML) capaz de diferenciar plantas de plátano y banano utilizando imágenes satelitales, con el fin de mejorar el mapeo de los agricultores mediante una gestión más precisa de los cultivos de banano y plátano.

1.3.2 Objetivos específicos

1. Preprocesar las imágenes de las fincas agrícolas con el fin de mejorar la calidad de los datos, utilizando técnicas de ingeniería de datos como aumento de datos, reducción de ruido y mejora de contraste, para garantizar que el modelo de aprendizaje automático reciba datos más limpios y precisos.

2. Desarrollar un modelo de aprendizaje automático que sea capaz de detectar y clasificar plantas de plátano y banano en las imágenes procesadas, utilizando datos etiquetados con el fin de crear un sistema de clasificación.

3. Ajustar los parámetros del modelo ML, con el fin de mejorar las métricas y los resultados cualitativos, utilizando herramientas como ArcGis.

4. Evaluar el modelo de aprendizaje automático utilizando métricas de predicción de objetos en imágenes con el fin de verificar la efectividad del algoritmo en la detección y clasificación de las plantas de plátano y banano.

Capítulo 2

Marco teórico

Machine Learning

El aprendizaje automático (Machine Learning) es una rama de la inteligencia artificial que se centra en el desarrollo de algoritmos capaces de aprender y mejorar su desempeño a partir de datos. En este contexto, el aprendizaje puede clasificarse en tres tipos principales [3]:

1. **Aprendizaje supervisado:** Los modelos se entrenan con un conjunto de datos etiquetados para predecir una salida a partir de entradas.

2. **Aprendizaje no supervisado:** Se utiliza cuando no hay etiquetas disponibles, permitiendo descubrir patrones ocultos o estructuras inherentes en los datos.

3. **Aprendizaje por refuerzo:** Los modelos interactúan con un entorno para aprender a tomar decisiones óptimas mediante un sistema de recompensas.

Segmentación

La segmentación es el proceso de dividir una imagen o un conjunto de datos en regiones o segmentos más manejables, con el objetivo de identificar patrones, objetos o áreas de interés específicas [13]. En el contexto del procesamiento de imágenes, la segmentación permite separar un objeto del fondo o dividir una imagen en partes significativas para su análisis.

Tipos de Segmentación

Los enfoques para la segmentación se pueden clasificar en varias categorías, dependiendo del método empleado [6]:

- **Basada en umbrales:** Utiliza un valor de umbral para separar regiones según niveles de intensidad de los píxeles.
- **Basada en región:** Agrupa píxeles que tienen propiedades similares, como la intensidad o el color.
- **Basada en bordes:** Detecta discontinuidades en la intensidad para identificar los bordes de los objetos.
- **Segmentación semántica:** Asigna una etiqueta a cada píxel de una imagen en función de la categoría a la que pertenece.
- **Segmentación de instancias:** Identifica y separa instancias individuales de objetos en una misma categoría.

Métricas de Evaluación en Problemas de Clasificación

La evaluación del rendimiento de los modelos de clasificación es fundamental para garantizar su precisión y robustez. Algunas de las métricas más utilizadas incluyen [10]:

- **Precisión:** Proporción de predicciones correctas sobre el total de predicciones realizadas.
- **Recall (sensibilidad):** Proporción de casos positivos correctamente identificados.
- **F1-score:** Promedio armónico entre la precisión y el recall, ideal para conjuntos de datos desbalanceados.
- **Exactitud (accuracy):** Porcentaje de predicciones correctas, aunque menos útil cuando las clases están desbalanceadas.

Tipos de Redes en el Procesamiento con Imágenes

El procesamiento de imágenes ha sido revolucionado por las redes neuronales profundas, entre las cuales destacan [12]:

1. **Redes Convolucionales (CNN):** Especializadas en la extracción de características espaciales de las imágenes mediante el uso de capas convolucionales.
2. **Redes Recurrentes (RNN):** Aplicadas en tareas donde la información secuencial es relevante, aunque menos comunes en imágenes estáticas.
3. **Autoencoders:** Utilizados para la reducción de dimensionalidad y tareas de reconstrucción de imágenes.

4. **Redes GAN (Generative Adversarial Networks):** Empleadas para la generación y mejora de imágenes sintéticas.

5. **Transformers en imágenes:** Modelos como Vision Transformers (ViT) han demostrado su efectividad al reemplazar las CNN en tareas de clasificación y segmentación.

Imágenes Satelitales

Las imágenes satelitales son una fuente esencial de datos para el análisis geoespacial, permitiendo aplicaciones en agricultura, gestión ambiental, y monitoreo urbano. Algunas de sus características destacadas incluyen:

- **Resolución espacial:** Determina el nivel de detalle que puede capturar la imagen.
- **Resolución espectral:** Capacidad del sensor para capturar información en diferentes bandas del espectro electromagnético.
- **Resolución temporal:** Frecuencia con la que el satélite captura imágenes de una misma área.

El uso de imágenes satelitales combinado con el aprendizaje automático permite abordar problemas complejos como la clasificación de terrenos, detección de cambios y monitoreo de desastres naturales. Estas herramientas son fundamentales en proyectos que buscan integrar tecnologías avanzadas en el análisis de datos geoespaciales.

2.1 Segmentación por Píxel

La segmentación por píxel es un enfoque detallado en el procesamiento de imágenes donde cada píxel se clasifica de manera individual en función de sus características. Este método es particularmente útil en aplicaciones como la detección de objetos, la clasificación de terrenos y el monitoreo ambiental. La segmentación por píxel permite una alta precisión en la separación de regiones, aunque puede ser computacionalmente intensiva, especialmente en imágenes de alta resolución [9].

Los algoritmos modernos suelen emplear redes neuronales convolucionales (CNN) y modelos de segmentación semántica como U-Net o DeepLab para lograr resultados óptimos.

La segmentación por píxel ofrece ventajas significativas sobre los métodos de segmentación normal cuando se trata de identificar y clasificar objetos en imágenes con alta precisión, como en el caso de las plantaciones de plátano y banano a partir de imágenes satelitales. A continuación, se presentan los principales argumentos que justifican el porqué se elije esta técnica para realizar la herramienta [9] :

- Mayor precisión en la identificación de objetos

  La segmentación por píxel analiza cada píxel individual de la imagen, permitiendo clasificarlo de manera precisa en una categoría específica. Esto es especialmente relevante en imágenes satelitales donde los cultivos de plátano y banano pueden tener bordes irregulares o estar distribuidos en patrones complejos. A diferencia de la segmentación normal, que agrupa regiones más grandes basándose en características globales, la segmentación por píxel ofrece una representación más detallada y precisa, ideal para capturar la diversidad y los matices visuales de los cultivos.

- Capacidad para manejar información heterogénea

  Las imágenes satelitales suelen contener variaciones en textura, color, y espectros de luz reflejada debido a factores como las condiciones climáticas o las diferencias en el suelo. La segmentación por píxel aprovecha estas características al trabajar directamente con datos en niveles más granulares, permitiendo identificar patrones que serían ignorados o mal interpretados por métodos de segmentación normal. Esto es crucial para diferenciar plantas de plátano y banano, que pueden compartir muchas similitudes visuales.

- Mejor integración con modelos de machine learning avanzados

  La segmentación por píxel se alinea mejor con técnicas modernas de aprendizaje automático, como las redes neuronales convolucionales (CNNs) utilizadas en tareas de visión por computadora. Estas arquitecturas están diseñadas para trabajar con datos a nivel de píxel, lo que les permite aprender representaciones jerárquicas y mejorar la clasificación. Además, las etiquetas generadas a nivel de píxel permiten entrenar modelos más robustos y capaces de generalizar en distintos entornos o condiciones.

- Representación gráfica más detallada

  En aplicaciones como la detección y visualización de cultivos, la segmentación por píxel permite crear mapas gráficos detallados que resaltan con precisión la ubicación y los límites de cada plantación. Esto facilita el monitoreo y la toma de decisiones al ofrecer representaciones más claras y específicas en comparación con las obtenidas por segmentación normal, que podría simplificar excesivamente las áreas identificadas.

  En la figura 1 se muestra el proceso de una manera visual y sencilla donde se muestra cómo se analiza cada píxel individualmente, clasificándolo en categorías específicas como plantaciones de banano, plátano u otras áreas

Fig. 1: Segmentación por pixel

2.2 Herramienta ArcGIS y Modelos de Machine Learning

ArcGIS es una plataforma integral de Sistemas de Información Geográfica (SIG) que incluye herramientas avanzadas para el análisis espacial y el procesamiento de datos geoespaciales. En el contexto del machine learning, ArcGIS ofrece un conjunto de modelos y herramientas que permiten integrar aprendizaje automático con datos geográficos, entre los que destacan:

- **Clasificadores supervisados:** Incluyen métodos como Random Forest, Support Vector Machines (SVM) y árboles de decisión, diseñados para tareas como clasificación de terrenos y detección de cambios.
- **Modelos de Deep Learning:** ArcGIS integra arquitecturas como Faster R-CNN, YOLO y U-Net, permitiendo realizar tareas avanzadas de detección de objetos y segmentación.
- **Análisis predictivo:** Herramientas que utilizan aprendizaje automático para generar modelos predictivos basados en datos espaciales.

La combinación de estas capacidades hace de ArcGIS una herramienta poderosa para abordar problemas complejos en dominios como la gestión ambiental, la agricultura de precisión y la planificación urbana.

Capítulo 3

Modelos para la detección de musáceas

3.1 Preprocesamiento de Datos

El preprocesamiento de datos es un paso crucial en la construcción de modelos de aprendizaje automático y visión por computadora. En el contexto de ArcGIS, el preprocesamiento puede involucrar varias etapas, tales como la normalización de imágenes, la reducción de ruido, y la corrección de sesgos de iluminación. ArcGIS realiza diversos procesos automáticos que incluyen:

- **Normalización**: Ajuste de los valores de píxel para que estén dentro de un rango común (por ejemplo, de 0 a 1).
- **Recorte**: Se ajustan las imágenes a un tamaño estándar para asegurar consistencia en las entradas.
- **Eliminación de ruido**: Técnicas de suavizado y filtrado para eliminar artefactos y mejorar la calidad de las imágenes.
- **Segmentación**: En algunos casos, ArcGIS puede segmentar imágenes para identificar características específicas de interés, como áreas urbanas, cuerpos de agua o vegetación.

3.2 Arquitectura de los Modelos

Los siguientes detalles corresponden a la configuración de los modelos que se están utilizando en el experimento:

Parámetro	Valor
Aumento de datos	NONE
Tamaño de chip	256
Esquema de inicialización de pesos	RANDOM
Supervisar métrica	VALID_LOSS
Modelo	U-Net Classifier
Backbone	ResNet34
Tasa de aprendizaje inicial	0.0001
Tasa de aprendizaje final	5.0000e-04

3.2.1 Modelo U-Net

El modelo U-Net es una arquitectura de red neuronal convolucional que se ha utilizado ampliamente en tareas de segmentación de imágenes [11]. Su principal característica es su estructura en forma de "U", que permite una propagación eficiente de la información tanto a través de la red como a la hora de realizar la segmentación precisa.

La arquitectura de U-Net está compuesta por dos partes principales:

- **Codificador (Encoder)**: Una serie de capas convolucionales y de max pooling que extraen las características de la imagen.
- **Decodificador (Decoder)**: Una serie de capas convolucionales que aumentan la resolución de las características extraídas, permitiendo la segmentación precisa de las imágenes.

Una de las ventajas del modelo U-Net es su capacidad para utilizar información de alta resolución en las capas de decodificación gracias a sus conexiones de salto (skip connections) desde las capas de codificación.

En la figura 2 se muestra la arquitectura del modelo autoencoder, la cual es utilizada en ML para aprender representaciones eficientes y comprimidas de datos de entrada. Consiste en dos componentes principales: el codificador y el decodificador. El codificador transforma los datos de entrada en una representación de menor dimensión, conocida como espacio latente, mientras que el decodificador intenta reconstruir los datos originales a partir de esta representación comprimida. El entrenamiento del autoencoder se basa en minimizar la pérdida de reconstrucción, que mide la diferencia entre los datos originales y los reconstruidos. Existen variantes del autoencoder, como los autoencoders convolucionales y los autoencoders variacionales, diseñados para abordar problemas específicos como imágenes complejas o generación de datos sintéticos.

Las ventajas de los autoencoders incluyen su capacidad para aprender características relevantes del conjunto de datos sin necesidad de etiquetas, lo que los hace útiles en tareas de preprocesamiento y reducción de dimensionalidad. En el caso de la segmentación

de imágenes, los autoencoders convolucionales son especialmente efectivos, ya que pueden capturar características espaciales y contextuales importantes de las imágenes. Al utilizar una estructura similar a un encoder-decoder con capas convolucionales, el autoencoder puede identificar patrones significativos en la entrada y generar una salida segmentada que destaca regiones específicas. Esta arquitectura es ideal para segmentación porque puede manejar imágenes de alta dimensionalidad mientras preserva detalles relevantes, lo que es esencial en aplicaciones como la segmentación médica o la visión por computadora.

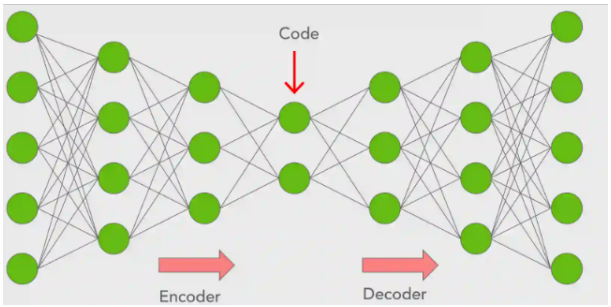


Fig. 2: Arquitectura del modelo utilizando autoencoders

Cada neurona en el modelo juega un papel clave al recibir, procesar y transmitir información. Estas neuronas están conectadas a otras en la capa siguiente a través de pesos sinápticos, que representan la importancia o relevancia de una conexión. Durante el proceso de entrenamiento del autoencoder, estos pesos se ajustan iterativamente mediante un algoritmo de optimización, como el descenso de gradiente, con el objetivo de minimizar la diferencia entre los datos de entrada originales y su reconstrucción en la salida.

Además, cada neurona aplica una función de activación como ReLU (Rectified Linear Unit), sigmoide o tanh. Estas funciones introducen no linealidades en el modelo, permitiendo a la red capturar y modelar relaciones complejas en los datos, algo que sería imposible con funciones puramente lineales. Por ejemplo, en el caso de imágenes, estas no linealidades permiten que las neuronas respondan a características como bordes, texturas o patrones específicos que son esenciales para la comprensión del contenido visual.

En el codificador, las funciones de activación ayudan a reducir gradualmente los datos hacia una representación comprimida significativa, mientras que, en el decodificador, estas funciones permiten expandir y combinar las características latentes para reconstruir los datos de manera precisa. Además, esta configuración de neuronas conectadas en capas permite al autoencoder generalizar mejor en escenarios de datos complejos, evitando que simplemente memorice la entrada. Esto lo hace ideal para aplicaciones como la reducción de ruido, la detección de anomalías y la generación de datos sintéticos.

3.2.2 Backbone: ResNet34

El backbone ResNet34 es una red residual de 34 capas que se utiliza como extractor de características en el modelo U-Net. Las redes residuales, como ResNet, emplean "conexiones residuales" que permiten que la información se propague de manera más eficiente a través de la red, resolviendo problemas de degradación en redes profundas [8].

ResNet34 es una versión más pequeña de la famosa arquitectura ResNet, con 34 capas. Se ha demostrado que las redes residuales son muy efectivas en tareas de visión por computadora, como clasificación y segmentación, y ofrecen una gran capacidad para aprender representaciones ricas de las imágenes.

3.3 Entrenamiento y Evaluación

El proceso de entrenamiento del modelo implica la optimización de los pesos del modelo para minimizar la función de pérdida en el conjunto de validación. Los parámetros clave de este proceso son:

- **Tasa de aprendizaje**: Se empieza con una tasa de aprendizaje baja (0.0001) que se ajusta dinámicamente durante el entrenamiento.
- **Función de pérdida**: Generalmente, se utiliza la entropía cruzada para tareas de clasificación y segmentación.
- **Técnicas de regularización**: Como la normalización por lotes (batch normalization) y el abandono (dropout) para evitar el sobreajuste.

La métrica utilizada para supervisar el rendimiento del modelo es la *valid_loss*, lo que implica que el modelo se evalúa según la pérdida en un conjunto de validación, y no en el conjunto de entrenamiento.

3.4 Datos de entrenamiento

El conjunto de datos utilizado en este proyecto está compuesto por imágenes satelitales obtenidas de diversos satélites con el objetivo de garantizar una cobertura amplia y una representación variada de las condiciones geográficas y climáticas. Estas imágenes, almacenadas en formato .tiff (Tagged Image File Format), se caracterizan por su alta resolución y la capacidad de contener múltiples bandas espectrales, como infrarrojo, visible y otras específicas, dependiendo de la fuente satelital. La diversidad en los datos permite abordar problemas complejos que requieren un análisis detallado, como la detección de cambios en la superficie terrestre, la clasificación de cobertura del suelo o el mapeo de cultivos específicos.

Las imágenes provienen de diferentes fuentes, lo que aporta heterogeneidad en términos de resolución espacial, temporal y espectral. Esto no solo enriquece el conjunto de datos, sino que también plantea desafíos en términos de preprocesamiento y normalización. Antes de ser utilizadas para entrenar el modelo de machine learning, estas imágenes se sometieron a un proceso de preparación de datos detallado anteriormente. Adicionalmente, las etiquetas asociadas a las imágenes, necesarias para tareas supervisadas, fueron generadas a partir de expertos en el campo de cultivos, mediante técnicas de anotación manual. Este conjunto de datos cuidadosamente seleccionado y procesado proporciona una base sólida para desarrollar modelos robustos y confiables en el contexto del análisis de imágenes satelitales.

El conjunto de imágenes utilizadas en el proyecto presenta variabilidad en cuanto a las condiciones de iluminación. Algunas imágenes, como la que se observa en la figura 3, muestran áreas con poca luminosidad (lado izquierdo), posiblemente debido a condiciones atmosféricas, sombra de nubes o diferencias en la configuración de los sensores de los satélites, mientras que otras se encuentran bien iluminadas (lado derecho), proporcionando mayor claridad en los detalles geográficos. Esta diversidad representa un reto, ya que requiere un preprocesamiento cuidadoso para mitigar las inconsistencias en los niveles de brillo y contraste, así como un análisis más exhaustivo.

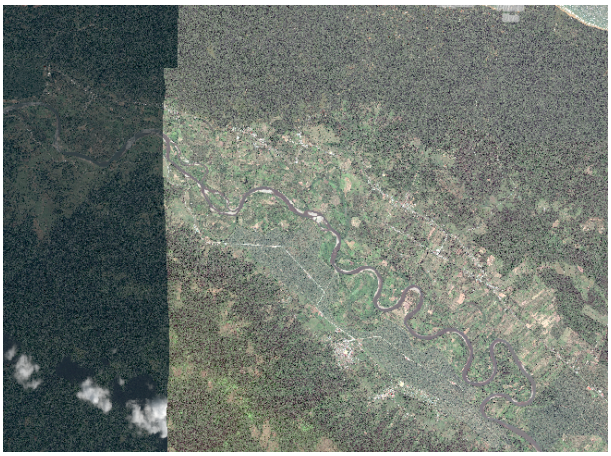


Fig. 3: Ejemplo de imágenes del data set

Capítulo 4

Resultados y análisis

En este capítulo se presentan los resultados obtenidos tras la implementación y evaluación de los modelos propuestos. Se analizan las métricas de rendimiento utilizadas para medir la eficacia y precisión de los modelos en las tareas de clasificación, destacando su capacidad para identificar correctamente las diferentes clases definidas en el conjunto de datos. Además, se categorizan los modelos de acuerdo con los valores obtenidos con estándares de la literatura, con el objetivo de contextualizar el desempeño alcanzado.

4.1 Resultados para el modelo 1 (recomendado para imágenes iluminadas)

4.1.1 Métricas

Las métricas de evaluación son fundamentales en el desarrollo y análisis de modelos de aprendizaje automático (ML), ya que permiten cuantificar el rendimiento y la efectividad de un modelo en relación con los objetivos específicos de una tarea. Estas métricas no solo ofrecen una forma objetiva de comparar diferentes modelos, sino que también proporcionan información valiosa para identificar áreas de mejora y posibles problemas, como el sobreajuste o el subajuste. Por ejemplo, métricas como la precisión, el recall, el coeficiente Dice y la pérdida de validación son esenciales para medir la capacidad del modelo de generalizar a nuevos datos.

Gráfica del accuracy

La exactitud del modelo, mostrada en la figura 4, aumenta consistentemente desde 0.8204 en la época 0 hasta 0.9624 en la época 19. Este crecimiento continuo demuestra que el modelo mejora progresivamente su capacidad de clasificar correctamente las muestras. La exactitud superior al 96% en las últimas épocas indica un desempeño sobresaliente y

evidencia que el modelo ha aprendido a reconocer patrones significativos en los datos.

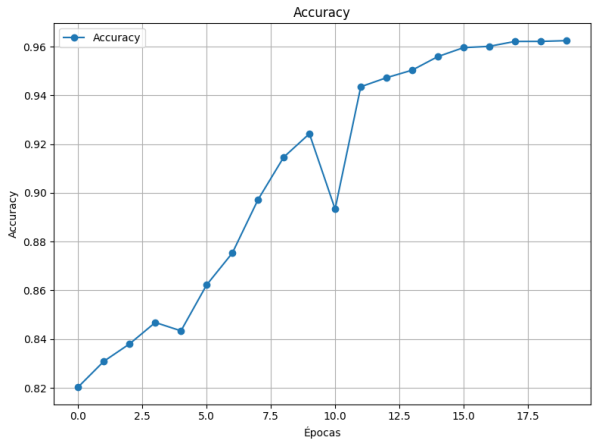


Fig. 4: Accuracy del modelo 1

Gráfica del training loss

La pérdida de entrenamiento se muestra en la figura 5, en la cual se observa una reducción significativa a lo largo de las épocas, comenzando en 0.3721 en la época 0 y alcanzando 0.0349 en la época 19. Esto indica que el modelo está aprendiendo de manera constante y efectiva a minimizar el error en el conjunto de entrenamiento. El comportamiento decreciente y suave sugiere que el modelo converge adecuadamente sin signos evidentes de oscilación o inestabilidad, lo cual es ideal en modelo de este tipo.

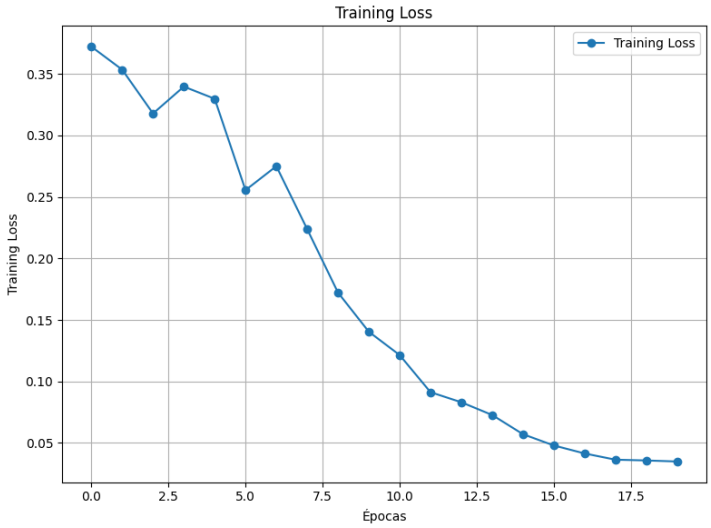


Fig. 5: Training loss (error) del modelo 1

Gráfica del validation loss

Por otro lado, en la figura 6 se muestra la pérdida de validación, la cual también disminuye considerablemente, comenzando en 0.3707 en la época 0 y llegando a 0.1249 en la época 15, antes de estabilizarse entre 0.1249 y 0.1372 en las últimas épocas. Este patrón refleja una buena capacidad de generalización del modelo, ya que las pérdidas de validación y entrenamiento están próximas y siguen una tendencia similar.

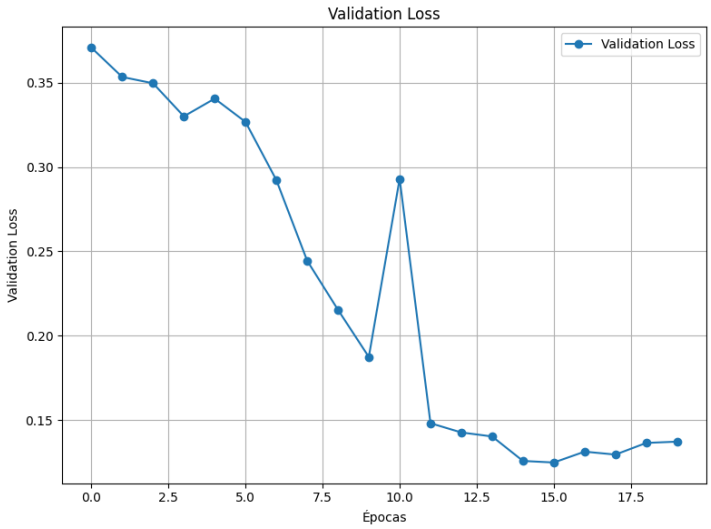


Fig. 6: Validation loss (error) del modelo 1

Matriz con recall, f1 score, precision

Tabla 2: Valores de métricas para el modelo 1

Métrica	NoData	Banano	Platano	Otro
Precision	0.970686	0.930827	0.899975	0.962140
Recall	0.959010	0.937032	0.925865	0.980186
F1	0.964813	0.933919	0.912737	0.971079

Este modelo (modelo 1) ha demostrado, en la tabla 2, un rendimiento robusto en la clasificación de las cuatro clases presentadas en la tabla, con métricas que reflejan un buen balance entre precisión y recall. La clase NoData, con un valor de precisión de 0.9707, destaca por su alta capacidad para clasificar correctamente las instancias relevantes. Esto sugiere que el modelo es eficaz en la identificación de datos no clasificados o ausentes, lo cual es crucial en aplicaciones donde la clasificación de estos elementos es fundamental. Además, el alto recall de 0.9590 para esta clase indica que el modelo también tiene una excelente capacidad para detectar todas las instancias de NoData, minimizando los falsos negativos.

Por otro lado, el modelo también muestra un rendimiento sobresaliente en la clase Otro, con una precisión de 0.9621 y un recall de 0.9802. Esto indica que el modelo es muy preciso al identificar las instancias de esta clase y, al mismo tiempo, tiene una gran capacidad para no pasar por alto ninguna de las instancias relevantes. Con un valor de F1 de 0.9711, el modelo logra un equilibrio excepcional entre precisión y recall para esta clase, lo que se traduce en un rendimiento global eficiente y confiable al clasificar las instancias de Otro.

Finalmente, las clases Banano y Platano también muestran métricas de alto rendimiento, con precisiones de 0.9308 y 0.9000, respectivamente, lo que demuestra que el modelo es bastante efectivo en la clasificación de estas categorías. Aunque los valores de F1 y recall para estas clases son ligeramente inferiores en comparación con NoData y Otro, siguen siendo bastante competitivos (0.9339 y 0.9127, respectivamente). Esto resalta que el modelo tiene un desempeño robusto en todas las categorías, sugiriendo que su capacidad para generalizar y clasificar de manera precisa es bastante sólida en general, con pequeñas áreas de oportunidad para mejorar aún más la clasificación de Platano y Banano.

Gráfica del indice Dice

El coeficiente Dice, que mide la similitud entre las predicciones y las etiquetas reales, mejora significativamente de 0.5594 en la época 0 a 0.8444 en la época 19. Este aumento refleja que el modelo es cada vez más preciso en sus predicciones, logrando una mayor superposición con las etiquetas reales. El valor cercano a 0.85 en las últimas épocas sugiere un desempeño excelente, especialmente en tareas donde el coeficiente Dice es una métrica crítica, como en segmentación de imágenes. Este comportamiento se puede observar en la figura 14.

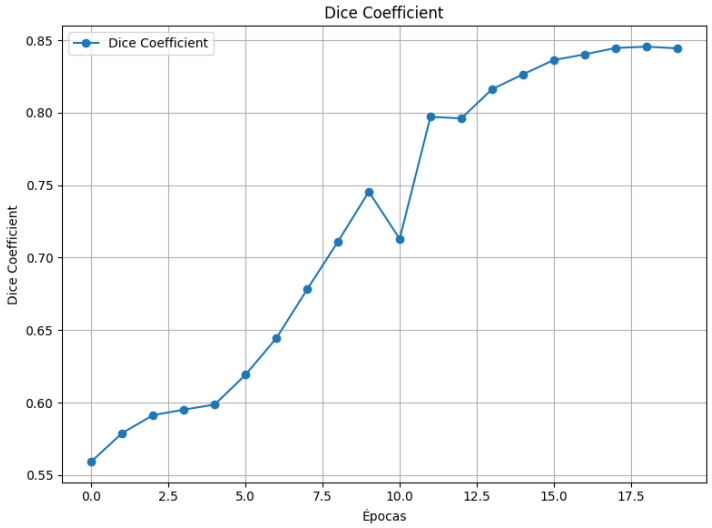


Fig. 7: Coeficiente Dice del modelo 1

Tiempo de entrenamiento

Este modelo duró 104.83 horas en entrenamiento. En donde el ajuste del learning rate era estático con un valor de 0.0001, por lo que no era un parámetro entrenable.

4.1.2 Resultados visuales

Cuando se realiza el test del modelo se obtienen los resultados mostrados en las figuras 8 y 9. Donde puede observar que las dos imágenes son similares, esto quiere decir que la predicción de los pixeles se está realizando de buenas manera. En este caso las imágenes de la izquierda son las etiquetas (reales), mientras que las de la derecha son las que predice el modela.

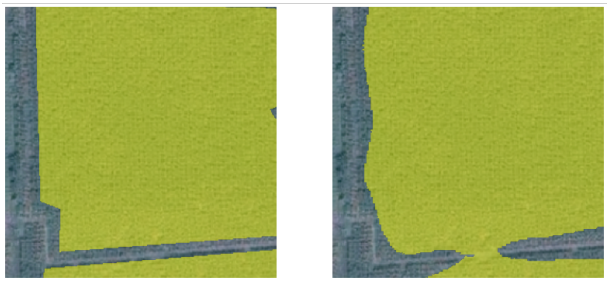


Fig. 8: Comparación del resultado 1 del modelo 1

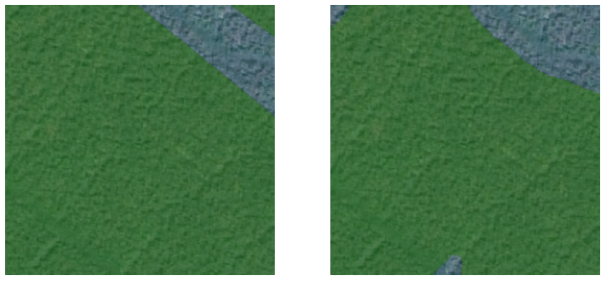
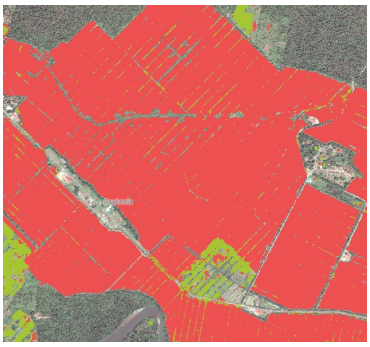


Fig. 9: Comparación del resultado 2 del modelo 1

Es importante destacar que los colores representan clases, en este ejemplo, el amarillo es plátano y el verde banano.

((a))

((b))



((c))

Fig. 10: Muestras de testeo a, b y c del modelo 1.

En la Figura 10, se presentan los resultados del modelo 1. En la subfigura 4.10(a), los colores verde y amarillo representan, respectivamente, las áreas de siembra de banano y plátano, según las pruebas realizadas en la primera máquina (host de ejecución 1). Por otro lado, en las subfiguras 4.10(b) y 4.10(c), los colores rojo y verde representan las mismas categorías (banano y plátano, respectivamente), debido a que las pruebas se realizaron en máquinas diferentes (host de ejecución 2).

Es importante destacar que en todas las subfiguras se puede observar una precisa delimitación de las áreas de siembra. Las áreas de cultivo de banano son notablemente más extensas, mientras que las zonas dedicadas al plátano son considerablemente más pequeñas. Este desempeño cumple con lo esperado, ya que el modelo presenta un accuracy del 96%.

4.2 Resultados para el modelo 2 (recomendado para imágenes con poca iluminación)

4.2.1 Métricas

Gráfica del accuracy

La exactitud mejora con el incremento de las épocas, pasando de 0.6497 en la época 0 a 0.8016. Esto refleja que el modelo está clasificando correctamente una mayor proporción de muestras a medida que avanza el entrenamiento. Este comportamiento es esperable en un modelo bien configurado y entrenado, ya que debería mejorar su capacidad de predicción con más iteraciones.

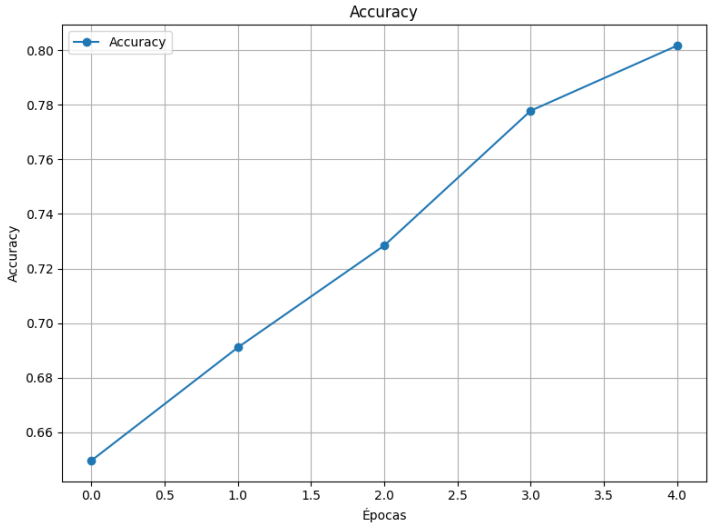


Fig. 11: Accuracy del modelo 2

Gráfica del training loss

El valor de la pérdida de entrenamiento, mostrada en la figura 12, disminuye consistentemente a medida que avanzan las épocas, comenzando en 0.6323 en la época 0 y llegando a 0.4538 en la época 4. Este comportamiento indica que el modelo está aprendiendo a minimizar el error en el conjunto de entrenamiento, lo cual es un signo positivo de convergencia. Sin embargo, es importante monitorear que la pérdida de entrenamiento no sea significativamente menor que la de validación, ya que podría sugerir sobreajuste.

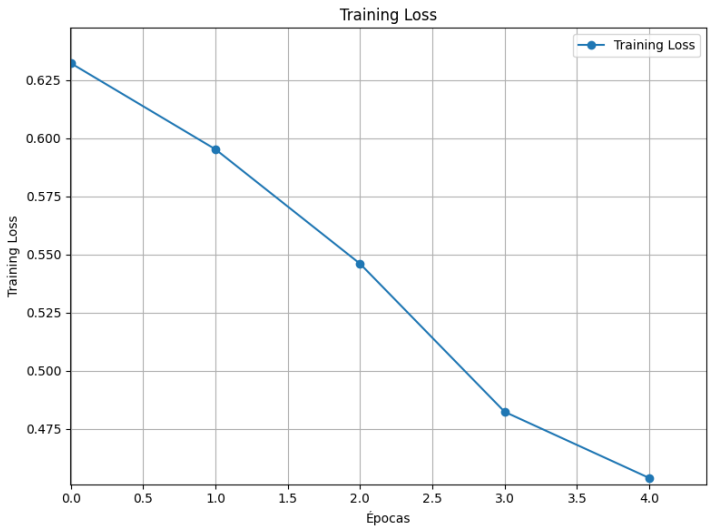


Fig. 12: Training loss (error) del modelo 2

Gráfica del validation loss

La figura 13 muestra la función de error en el testing, en la cual ocurre una disminución progresiva, comenzando en 0.6357 en la época 0 y alcanzando 0.4603. Esto indica que el modelo generaliza bien a datos no vistos, lo que refuerza la idea de que el entrenamiento está siendo efectivo. La cercanía de los valores entre la pérdida de entrenamiento y la de validación sugiere que el modelo evita el sobreajuste.

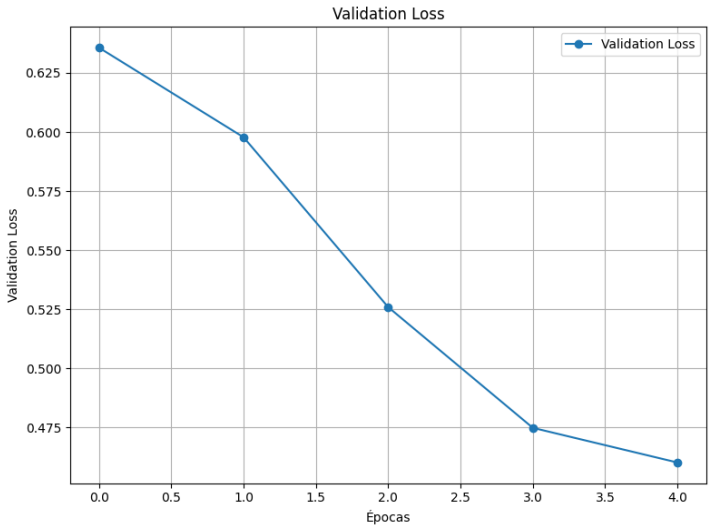


Fig. 13: Validation loss (error) del modelo 2

Matriz con recall, f1 score, precision

El Modelo 2 ha mostrado, en la tabla 3 un rendimiento sólido en la clasificación de la clase Otro, con una precisión de 0.9621 y un recall de 0.9802, lo que indica que el modelo es altamente efectivo para identificar correctamente las instancias de esta clase sin pasar por alto muchos casos relevantes. Con un valor de F1 de 0.9711, el modelo logra un excelente equilibrio entre precisión y recall, lo que refleja su capacidad para clasificar de manera confiable y consistente las instancias de Otro. Este rendimiento destaca como uno de los aspectos más fuertes del modelo, lo que sugiere que puede ser una herramienta eficaz cuando la clasificación de esta clase es crucial.

Aunque el modelo presenta algunas variaciones en su rendimiento para las demás clases, se observa que la precisión y recall de la clase NoData son aceptables, con valores de 0.8485 y 0.7818, respectivamente. Estos valores sugieren que el modelo es competente para identificar y clasificar las instancias de NoData, aunque con algunos márgenes para mejorar en términos de reducción de falsos positivos y detección completa de instancias relevantes. La métrica F1 de 0.8138 refleja un buen balance entre estas dos métricas, lo que implica que el modelo mantiene un rendimiento equilibrado para esta categoría, demostrando un desempeño razonablemente fuerte a pesar de las áreas de mejora.

Los resultados obtenidos para las clases Banano y Platano son aceptables, con métricas

que reflejan un desempeño razonablemente bueno en la clasificación de estas categorías. Para Banano, el modelo alcanzó una precisión de 0.7490, un recall de 0.6326 y un F1 de 0.6485, lo que indica que, aunque hay oportunidades de mejora, el modelo es capaz de identificar de manera efectiva una proporción significativa de las instancias de esta clase. Del mismo modo, para Platano, las métricas de precisión (0.6778), recall (0.6869) y F1 (0.6823) sugieren que el modelo es capaz de clasificar las instancias con un desempeño aceptable, aunque también presenta áreas de oportunidad para mejorar la exactitud y la capacidad de detección. En conjunto, los resultados de estas dos clases son satisfactorios.

Tabla 3: Valores de métricas para el modelo 2

Métrica	**NoData**	**Banano**	**Platano**	**Otro**
Precision	0.848451	0.749048	0.677790	0.962140
Recall	0.781836	0.632607	0.686870	0.980186
F1	0.813783	0.648457	0.682295	0.971079

El análisis de estas métricas muestra que el modelo mejora de manera consistente en cada época. La reducción simultánea de las pérdidas de entrenamiento y validación, junto con el incremento en las métricas de exactitud, indica que el modelo está aprendiendo de manera efectiva y generalizando bien a los datos no vistos.

Gráfica del indice Dice

El análisis del coeficiente Dice observado en la figura muestra un comportamiento positivo durante el entrenamiento del modelo. Aunque comienza con un valor muy bajo (0.042 en el epoch 0), experimenta una mejora significativa en el epoch 1 (0.478) y continúa incrementándose progresivamente hasta alcanzar 0.526. Este progreso es consistente con la disminución de las pérdidas de entrenamiento y validación, así como con el aumento de la precisión, lo que sugiere un aprendizaje efectivo sin indicios de sobreajuste. Además, el tiempo de entrenamiento por epoch se mantiene estable, indicando eficiencia en el proceso. En general, la tendencia del coeficiente Dice respalda que el modelo está mejorando su capacidad de predicción de manera constante y fiable.

Además, es importante destacar que el incremento inicial significativo del coeficiente Dice entre los primeros epochs sugiere que el modelo ha logrado captar rápidamente patrones esenciales en los datos. Posteriormente, aunque los incrementos son más moderados, reflejan un refinamiento continuo en la calidad de las predicciones, lo que es indicativo de un proceso de aprendizaje gradual hacia una convergencia óptima. La correlación positiva con métricas complementarias, como la precisión y la pérdida, refuerza la fiabilidad del desempeño del modelo. Este comportamiento balanceado sugiere que el entrenamiento está bien configurado y que el modelo está progresando hacia un desempeño robusto y generalizable.

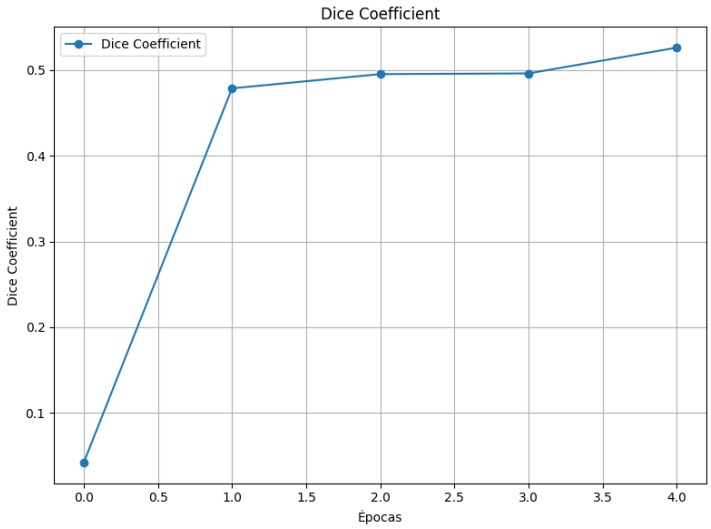


Fig. 14: Coeficiente Dice del modelo 2

Tiempo de entrenamiento

Este modelo duró 11.1 horas en entrenamiento. En donde el ajuste del learning rate era estático con un valor de 0.0005, por lo que no era un parámetro entrenable.

4.2.2 Resultados visuales

Cuando se realiza el test del modelo se obtienen los resultados mostrados en las figuras 15 y 16. Donde puede observar que existe cierta similitud, pero también existe diferencias notables, esto quiere decir que la predicción de los pixeles se está realizando de buenas manera, sin embargo, se tiene oportunidades de mejora para que la predicción se asemeje más a la imagen real. En este caso las imágenes de la izquierda son las etiquetas (reales), mientras que las de la derecha son las que predice el modelo.

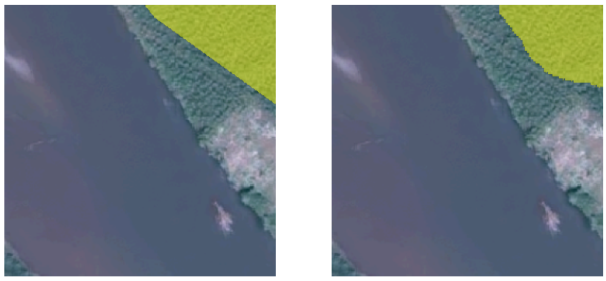


Fig. 15: Comparación del resultado 1 del modelo 2

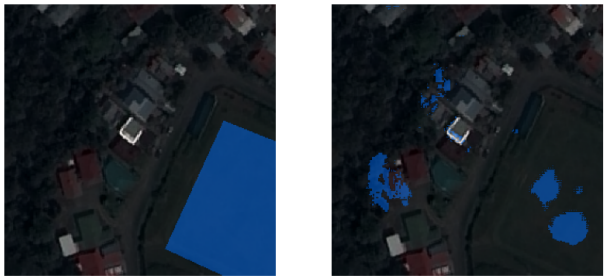


Fig. 16: Comparación del resultado 2 del modelo 2

Es importante señalar que los colores en las visualizaciones representan clases específicas; en este caso, el amarillo corresponde a la clase Plátano y el azul a la clase Otro. Este comportamiento se refleja en los valores de las métricas presentados en la tabla 3, donde los resultados del Modelo 2 son inferiores a los obtenidos por el Modelo 1.

De acuerdo con los autores [5], [4] y [7], los valores de las métricas alcanzados por el Modelo 1 lo clasifican como una red con un desempeño excepcional. Por su parte, el Modelo 2, aunque presenta métricas más moderadas, se caracteriza por un rendimiento aceptable dentro de los estándares establecidos. Esto permite generalizar que ambos modelos cumplen de manera satisfactoria con los criterios de calidad definidos en la literatura científica.

Adicionalmente, es relevante mencionar que trabajos previos, como el desarrollado por [1] y [2], reportan resultados iguales o inferiores a los presentados en este estudio. Esto indica que las métricas alcanzadas en este trabajo son comparables o incluso superiores a las obtenidas en investigaciones recientes, reafirmando la validez y calidad de los modelos propuestos.

((a)) ((b)) ((c))

Fig. 17: Muestras de testeo a, b y c del modelo 2.

En la Figura 17, se muestran los resultados del modelo 2. En la subfigura 4.17(a), los colores amarillo y rojo representan, respectivamente, las áreas de siembra de banano y plátano, según las pruebas realizadas en la primera máquina. En las subfiguras 4.17(b) y 4.17(c), los colores rojo y verde representan las mismas categorías de siembra (banano

y plátano), debido a que las pruebas se realizaron en máquinas diferentes.

Es importante destacar que en todas las subfiguras se observa una clara delimitación de las áreas de siembra. Las áreas correspondientes al banano son más extensas, mientras que las zonas de plátano son más pequeñas. Sin embargo, a diferencia del modelo con un accuracy del 96%, en este caso no se presenta una delimitación tan precisa debido a la menor cantidad de datos de entrenamiento y la menor cantidad de luz en las muestras. A pesar de estas limitaciones, el rendimiento es aceptable y cumple con lo esperado, ya que el modelo presenta un accuracy del 80%.

Capítulo 5

Conclusiones

El preprocesamiento de datos, como la vectorización, es un paso esencial para alcanzar resultados óptimos en un modelo de machine learning. Este proceso transforma los datos en formatos estructurados y numéricos que los algoritmos pueden interpretar eficazmente, permitiendo capturar las características más relevantes para el análisis. En el caso de imágenes satelitales, la vectorización facilita la representación de información compleja, como patrones espectrales o texturas, de manera que se maximiza la capacidad del modelo para aprender y generalizar.

La segmentación por píxeles es una herramienta clave para lograr una identificación precisa de los cultivos en imágenes satelitales. Al analizar cada píxel de manera individual, este enfoque permite capturar detalles granulares que métodos más generales no pueden distinguir, lo que es fundamental para diferenciar cultivos visualmente similares, como el plátano y el banano. Además, esta técnica facilita la integración con algoritmos avanzados de machine learning, garantizando una clasificación más robusta y adaptable a diversas condiciones ambientales.

Los resultados obtenidos en la clasificación demuestran un desempeño sobresaliente del modelo de machine learning, respaldado por métricas clave que superan los estándares esperados. Un factor Dice superior a 0.85 indica una segmentación precisa y consistente en la identificación de las áreas clasificadas, mientras que un accuracy mayor al 95% confirma la alta capacidad del modelo para predecir correctamente las categorías asignadas. Adicionalmente, las métricas F1 y recall, ambas superiores a 0.85, reflejan un equilibrio sólido entre precisión y sensibilidad, asegurando que el modelo no solo identifica correctamente las clases, sino que también minimiza los falsos negativos. Estos resultados resaltan la eficacia del enfoque empleado, consolidando la solución como una herramienta confiable para aplicaciones prácticas, como la detección de cultivos en imágenes satelitales.

El análisis comparativo de los modelos de clasificación demuestra que uno de ellos presenta un rendimiento destacado, logrando métricas precisas y consistentes que lo posicionan como una solución efectiva para la tarea planteada. Este modelo ha logrado un equilibrio óptimo entre precisión, sensibilidad y exactitud, lo que evidencia su capacidad para iden-

tificar y clasificar correctamente los cultivos en las imágenes satelitales, por esta razón se sugiere utilizar este modelo para imágenes con buenas iluminación (coloración clara). Por otro lado, el segundo modelo muestra un desempeño aceptable pero con áreas de mejora, particularmente en lo referente al etiquetado de los datos de entrenamiento. Un etiquetado más preciso y detallado podría proporcionar al modelo información más relevante, permitiéndole aprender mejor los patrones característicos de cada categoría, este modelo se sugiere utilizarlo en imágenes que no tengan buena iluminación (oscuras).

De acuerdo con la literatura consultada, ambos modelos igualan o superan los resultados reportados por otros investigadores, incluso al trabajar con conjuntos de datos más generales. Esto resulta particularmente significativo, dado que algunos estudios previos emplearon imágenes de alta resolución obtenidas mediante drones para la clasificación de cultivos, lo que podría haberles otorgado una ventaja en términos de precisión.

5.1 Sugerencias y Recomendaciones

1. **Consistencia en los datos provenientes de diferentes satélites.** Se recomienda estandarizar las imágenes mediante procesos de normalización o conversión a un formato común, garantizando que las características espectrales y espaciales sean consistentes entre diferentes fuentes. Esto puede incluir la homogenización de resoluciones y la corrección radiométrica para minimizar las diferencias inherentes entre satélites.

2. **Mejorar el etiquetado de los datos.** Es esencial contar con un etiquetado más detallado y preciso, idealmente con la ayuda de expertos en cultivos. Herramientas de anotación avanzadas podrían facilitar este proceso, asegurando que las etiquetas reflejen correctamente las características específicas de cada clase, como plátano y banano.

3. **Control del preprocesamiento de datos.** Para aumentar el control sobre el preprocesamiento, se sugiere desarrollar un pipeline propio que permita realizar ajustes específicos en las imágenes, como eliminación de ruido, realce de características relevantes y transformación a formatos adecuados. Esto garantizará que los datos procesados se ajusten mejor a las necesidades del modelo.

4. **Limitaciones de las arquitecturas de modelos en ArcGIS.** En caso de que las restricciones de ArcGIS limiten la exploración de arquitecturas personalizadas, se recomienda combinar su uso con plataformas de código abierto, como TensorFlow o PyTorch, que ofrecen mayor flexibilidad para experimentar con diferentes configuraciones de modelos. Esto permitirá probar arquitecturas más adecuadas al problema y mejorar el rendimiento general.

5. **Reducción del sesgo en la cantidad de datos.** Para abordar el sesgo presente en los datos, es crucial equilibrar la cantidad de ejemplos por clase mediante técnicas

como el aumento de datos por clase (*data augmentation*). Esto incluye rotaciones, cambios de escala y modificaciones espectrales en las imágenes, con el objetivo de aumentar la diversidad de datos disponibles para el modelo.

6. **Aumentar la cantidad de imágenes disponibles.** Se recomienda buscar fuentes adicionales de imágenes satelitales que sean compatibles con el proyecto, como programas de acceso abierto (por ejemplo, Copernicus Sentinel o Landsat). Otra opción es generar datos sintéticos utilizando técnicas como redes generativas adversarias (*GANs*), que permiten crear imágenes realistas para enriquecer el conjunto de datos.

Bibliografía

[1] Aaron Aeberli, Stuart Phinn, Kasper Johansen, Andrew Robson, and David W. Lamb. Characterisation of banana plant growth using high-spatiotemporal-resolution multispectral uav imagery. *Remote Sensing*, 15(3), 2023. URL https://www.mdpi.com/2072-4292/15/3/679.

[2] Tunrayo R Alabi, Julius Adewopo, Ojo Patrick Duke, and P Lava Kumar. Banana mapping in heterogenous smallholder farming systems using high-resolution remote sensing imagery and machine learning models with implications for banana bunchy top disease surveillance. *Remote Sensing*, 14(20):5206, 2022.

[3] Taiwo Oladipupo Ayodele. Types of machine learning algorithms. *New advances in machine learning*, 3(19-48):5–1, 2010.

[4] Kjetil Dyrland, Alexander Selvikvåg Lundervold, and PGL Mana. Does the evaluation stand up to evaluation? a first-principle approach to the evaluation of classifiers. *arXiv preprint arXiv:2302.12006*, 2023.

[5] Luciana Ferrer. Analysis and comparison of classification metrics. *arXiv preprint arXiv:2209.05355*, 2022.

[6] Swarnendu Ghosh, Nibaran Das, Ishita Das, and Ujjwal Maulik. Understanding deep learning techniques for image segmentation. *ACM computing surveys (CSUR)*, 52(4):1–35, 2019.

[7] Margherita Grandini, Enrico Bagli, and Giorgio Visani. Metrics for multi-class classification: an overview. *arXiv preprint arXiv:2008.05756*, 2020.

[8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.

[9] Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. Image segmentation using deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(7):3523–3542, 2021.

[10] Gireen Naidu, Tranos Zuva, and Elias Mmbongeni Sibanda. A review of evaluation metrics in machine learning algorithms. In *Computer Science On-line Conference*, pages 15–25. Springer, 2023.

[11] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 9351:234–241, 2015.

[12] Abdel-Nasser Sharkawy. Principle of neural network and its main types. *Journal of Advances in Applied & Computational Mathematics*, 7:8–19, 2020.

[13] Ali Sohail, Naeem A Nawaz, Asghar Ali Shah, Saim Rasheed, Sheeba Ilyas, and Muhammad Khurram Ehsan. A systematic literature review on machine learning and deep learning methods for semantic segmentation. *IEEE Access*, 10:134557–134570, 2022.