UNIVERSIDAD NACIONAL DE COLOMBIA

Detección y clasificación de antracnosis en mango usando imágenes hiperespectrales y técnicas de aprendizaje profundo.

Andres Felipe Montenegro Bermudez

Universidad Nacional de Colombia
Facultad de ingenieria, Departamento de ingenieria electrica y electronica
Bogotá, Colombia
2021

Detección y clasificación de antracnosis en mango usando imágenes hiperespectrales y técnicas de aprendizaje profundo.

Andres Felipe Montenegro Bermudez

Trabajo de grado presentado como requisito para optar al título de:
Magister en automatizacion industrial

Director:
Ing. Ph.D. Flavio Augusto Prieto Ortiz

Línea de Investigación:
Procesamiento digital de imagenes y machine learning (Automatizacion de procesos)
Grupo de Investigación:
GAUNAL

Universidad Nacional de Colombia
Facultad de ingenieria, Departamento de ingenieria electrica y electronica
Bogotá, Colombia
2021

”Sólo podemos ver poco del futuro, pero lo suficiente para darnos cuenta de que hay mucho que hacer.”

Alan Turing

Agradecimientos

A mi esposa, Laura Gonzalez, a mi hijo, Emilio Montenegro, mis padres, Olga Bermudez y Leonardo Montenegro, a quienes dedico este trabajo, ustedes han sido la inspiración y el apoyo para llegar a este punto.

A la Universidad Nacional de Colombia que me abrió sus puertas y me puso en el camino de muchas personas que han dejado en mi una huella de conocimiento y aprendizaje. A mi director, Flavio Augusto Prieto Ortiz, por toda su ayuda en este proceso de formación e investigación.

Resumen

Detección y clasificación de antracnosis en mango usando imágenes hiperespectrales y técnicas de aprendizaje profundo

Desde el surgimiento de las técnicas de inteligencia artificial han sido muchos los campos de producción que buscan usar estas técnicas para mejorar y automatizar sus procesos productivos. Uno de los campos que ha implementado dichas técnicas es la agricultura. Al combinar estos dos campos, la inteligencia artificial y la agricultura, nace una de las ramas de la agricultura de precisión, uno de los retos que se ha propuesto para esta rama es lograr diagnosticar enfermedades, o problemas en los cultivos, a través de la manipulación y entendimiento de imágenes. gracias a cambios en su morfología, color o reflectancia es posible determinar de manera temprana y poco invasiva los diferentes agentes que pueden estar afectando los cultivos. Un problema que no ha sido ampliamente estudiado es la antracnosis en cultivos de mango. Por ende este trabajo propone una solución de detección temprana de esta enfermedad, basada en el uso y análisis de imágenes hiperespectrales, mediante técnicas de aprendizaje profundo. En este estudio se mostraran conceptos acerca de la adquisición y procesamiento de este tipo de imágenes, así como el diseño y uso de algoritmos de aprendizaje profundo para la detección de antracnosis, sobre hojas de cultivos de mango en ambientes con iluminación controlada. Los resultados que se obtuvieron muestran lo útil que puede resultar aplicar esta tecnología, ya que se logró obtener un modelo basado en redes neuronales convolucionales en dos dimensiones el cual logra métricas de exactitud y precisión del 100 %, y una función de pérdida loss del 0,001 %, lo cual muestra el potencial de la fusión de la inteligencia artificial y el uso de las imágenes hiperespectrales para el desarrollo tecnológico de la agricultura de precisión.

Palabras clave: Visión por computador, agricultura de precisión, inteligencia artificial, imágenes hiperespectrales.

Abstract

Detection and classification of mango anthracnose using hyperspectral imaging and deep learning techniques

Since the emergence of artificial intelligence techniques, many production fields have sought to use these techniques to improve and automate their production processes. One of the

fields that has implemented these techniques is agriculture. By combining these two fields, artificial intelligence and agriculture, one of the branches of precision agriculture is born, one of the challenges that has been proposed for this branch is to diagnose diseases or problems in crops, through the manipulation and understanding of images. thanks to changes in morphology, color or reflectance it is possible to determine early and minimally invasive way the different agents that may be affecting crops. A problem that has not been widely studied is anthracnose in mango crops, therefore this work proposes a solution for early detection of this disease, based on the use and analysis of hyperspectral images, using deep learning techniques. This study will show concepts about the acquisition and processing of this type of images, as well as the design and use of deep learning algorithms for the detection of anthracnose on mango crop leaves in environments with controlled illumination. The results obtained show how useful it can be to apply this technology, since it was possible to obtain a model based on convolutional neural networks in two dimensions which achieves accuracy and precision metrics of 100 %, and a loss function of 0.001 %, which shows the potential of the fusion of artificial intelligence and the use of hyperspectral images for the technological development of precision agriculture.

Keywords: Computer vision, precision agriculture, artificial intelligence, hyperspectral imaging

Contenido

Agradecimientos vii

Resumen ix

Lista de Figuras xiii

1 Introducción 2
1.1 Objetivo general . . . 4
1.2 Objetivos específicos . . . 4

2 Marco teórico 5
2.1 Antracnosis . . . 5
2.2 Imágenes y firmas hiperespectrales . . . 6
2.3 Agricultura de presición (AP) . . . 7
2.4 Procesamiento digital de imágenes . . . 8
2.4.1 Espacios de color . . . 9
2.4.2 Segmentación . . . 11
2.4.3 Reducción de ruido . . . 12
2.4.4 Operaciones morfológicas . . . 12
2.5 Inteligencia artificial . . . 12
2.5.1 Aprendizaje de máquina (Machine learning) . . . 12
2.5.2 Redes neuronales artificiales (ANN) . . . 13
2.5.3 Deep learning . . . 13
2.6 Métricas de desempeño . . . 14

3 Análisis hiperespectral basado en redes neuronales profundas 16
3.1 Identificación del problema . . . 16
3.2 Solución propuesta . . . 16
3.3 Adquisición de imágenes . . . 18
3.4 Etiquetado . . . 19
3.5 Lectura y preprocesamiento . . . 20
3.5.1 Cambio de espacio de color . . . 21
3.5.2 Segmentación de las imágenes . . . 22
3.5.3 Operaciones morfológicas . . . 22

3.6 Aprendizaje profundo 24
3.6.1 Red neuronal convolucional de dos dimensiones (CNN2D) 24
3.6.2 Red neuronal convolucional de tres dimensiones (CNN3D) 26
3.7 Consideraciones previas 27

4 Resultados **28**
4.1 Modelo CNN2D aplicado a las muestras por día 28
4.2 Modelo CNN3D aplicado a las muestras por día 30
4.3 Modelo CNN2D aplicado a las muestras mezcladas 32
4.4 Modelo CNN3D aplicado a las muestras mezcladas 33

5 Conclusiones **35**

Bibliografía **36**

Lista de Figuras

2-1 Pruebas de patogenicidad de diferentes especies de mangos. A) Filipino B) Señora C) Filipino D) Keitt E) Haden F)Haden; G-Señora H-Keitt). Imagen tomada de [24] 6
2-2 Representación de una hoja en diferentes bandas espectrales 7
2-3 Firma espectral de varios píxeles de una imagen. 8
2-4 Cubo RGB, extraido de [4] 9
2-5 Imagen original en RGB vs Imagen en HSV 10
2-6 Representacion gráfica del concepto de deep learning, fuente propia 13
2-7 Representacion de una red neuronal convolucional (CNN), tomada de [3] . . 14

3-1 Diagrama del proceso de análisis hiperespectral basado en redes neuronales profundas 17
3-2 Disposición 1 de las imágenes hiperespectrales, separadas por clases 19
3-3 Disposición 2 de las imágenes hiperespectrales, separadas por dias 20
3-4 Etapas del cambio de espacio de color: A) Imagen original representada en espacio de color RGB. B) Imagen representada en espacio de color HSV. C) Plano H del espacio de color HSV. D) Plano S del espacio de color HSV. E) Plano V del espacio de color HSV 21
3-5 Histograma de una imagen del plano S 22
3-6 Preprocesamiento de la imagen: A) Plano S del espacio de color HSV. B) Imágen de una hoja después del proceso de umbralización C) Imagen de una hoja después del proceso de opening. D) Imagen de una hoja después del proceso de erosión E) Imágen de una hoja después del proceso de dilatación F) Máscara obtenida de la imagen de la hoja G) Imágen de una después de ser combinada con la máscara 23
3-7 Imagen resultante después del preprocesamiento 24
3-8 Modelo CNN2D 25
3-9 Modelo CNN3D 26

Lista de Tablas

4-1 Métricas del modelo CNN2D, experimento 1, durante 5 días 29
4-2 Métricas del modelo CNN2D, experimento 2, durante 5 días 29
4-3 Métricas del modelo CNN2D, experimento 3, durante 5 días 29
4-4 Métricas del modelo CNN3D, experimento 1, durante 5 días 31
4-5 Métricas del modelo CNN3D, experimento 2, durante 5 días 31
4-6 Métricas del modelo CNN3D, experimento 3, durante 5 días 31
4-7 Métricas del modelo CNN2D, experimento 1, mezclando los datos 32
4-8 Métricas del modelo CNN2D, experimento 2, mezclando los datos 32
4-9 Métricas del modelo CNN2D, experimento 3, mezclando los datos 32
4-10 Métricas del modelo CNN3D, experimento 1, mezclando los datos 33
4-11 Métricas del modelo CNN3D, experimento 2, mezclando los datos 34
4-12 Métricas del modelo CNN3D, experimento 3, mezclando los datos 34

1 Introducción

La antracnosis es una enfermedad fúngica que afecta a un importante grupo de cultivos dentro de los cuales se encuentran la fresa, mango, cacao, aguacate, papaya, [17] [23] entre otros. Esta enfermedad es causada por un hongo que pertenece a los géneros $Colletotrichum$, $Gloeosporium$ o la especie $Coniothyrium fuckelii$. Dependiendo de las condiciones climatológicas, esta enfermedad puede llegar a ocasionar daños hasta del cien por ciento de un cultivo, generando pérdidas económicas o problemas fitosanitarios, por lo cual se hace necesario poder identificarla de manera temprana y poco invasiva para poder tomar acciones al respecto. Uno de los síntomas más evidentes de la antracnosis es la aparición de manchas de color oscuro sobre las hojas o frutos, estas manchas en estado inicial son circulares y pequeñas y van creciendo y oscureciendo con el paso del tiempo, estas manchas se conocen como necrosis y pueden llegar a cubrir en su totalidad a los frutos u hojas. Según [2] las plantas se vuelven más propensas a sufrir antracnosis si estas han tenido golpes, o si presentan heridas causadas por insectos. Dado que esta es una enfermedad fúngica, su forma de propagación es por medio de esporas que viajan por el aire, de ahí que su propagación es sea rápida y difícil de controlar.

En la actualidad, los procesos para controlar y mitigar los efectos de la antracnosis en campo se basan en el uso de fungicidas y agentes químicos [9], esto de forma invasiva ya que no existe una herramienta de detección temprana en campo que sirva de apoyo a los productores para la toma de decisiones sobre el estado de su cultivo. Según [9] existe una gran mejoría tanto en el crecimiento como en la calidad de los mangos gracias al uso de químicos que regulan los daños ocasionados por la antracnosis, el uso de estos químicos se hace para evitar que las plantas o los cultivos sean infectados con agentes patógenos tanto en la precosecha como en la poscosecha, sin embargo, el uso excesivo de estos insumos puede llegar a ser nocivo tanto para los cultivos como para el consumidor final, razón por la cual en la actualidad, y cada vez con mas frecuencia, se ha aumentado la investigacion referente a la reducción y control del uso de estos productos. Si la antracnosis es detectada de manera temprana en los cultivos, es posible tomar decisiones que impliquen un menor uso de químicos, o inclusive, se podría pretender erradicar los frutos contaminados de forma temprana y controlada.

Por otro, lado el uso de técnicas de visión por computador y de inteligencia artificial han sido de gran interés en el sector agroindustrial ya que gracias a estas es posible detectar de manera temprana una gran cantidad de enfermedades y anomalías que se presentan en los

cultivos y que a simple vista no se podrían detectar. Recientemente se han comenzado a usar imágenes hiperespectrales, las cuales permiten visualizar características morfológicas o de color, además permiten estudiar y analizar de manera no invasiva ni destructiva las firmas espectrales características de los objetos y cómo estas se comportan a lo largo del tiempo [21]. Las imágenes hiperespectrales están compuestas por un hipercubo el cual está conformado por las dos dimensiones espaciales (X y Y) y adicional a estas una dimensión espectral λ , donde a cada pixel de la imágen hiperespectral se le conoce como una firma espectral. Esta firma muestra a lo largo del componente espacial las caracteristicas espectrales únicas y relevantes de un objeto bajo observación.

Existen varios estudios [16] [6] [15] [10] que reportan la factibilidad y el gran potencial que tiene el uso de estas tecnologias en diferentes aplicaciones en el sector agroindustrial desde la gestión de recursos forestales, pasando por la medición cuantitativa del estado de salud en masas arboladas, la cartografía de los bosques o la detección de enfermedades o plagas. En concreto, en la agricultura este tipo de tecnología se ha usado para determinar algunas variables de las plantes tales como su estado de irrigación o de salud, determinar la cantidad de fertilizantes, la presencia de enfermedades entre otras. Puntualmente sobre cultivos de mango hacer uso de esta tecnología puede representar un gran avance sobretodo si es posible llevar estos estudios a condiciones de campo, ya que podría realizarse una detección temprana y tomar decisiones preventivas a diferentes enfermedades como la antracnosis.

Las técnicas de inteligencia artificial sumadas al estudio de las imágenes hiperespectrales son una poderosa herramienta al momento de detectar de forma automática los síntomas asociados a este tipo de patógenos en plantas, ya que estas técnicas son ampliamente usadas para la clasificacion de patrones o predicción de eventos basados en conocimientos previos. Se han encontrado estudios acerca de esta enfermedad haciendo uso de imágenes hiperespectrales e inteligencia artificial en otro tipo de cultivos tales como guanábana, oliva, aguacate, tomate de árbol, entre otros [11][12]. Para el cultivo de mango se han encontrado algunos estudios recientes [20], sin embargo, no ha sido ampliamente estudiado por lo que se considera que el presente trabajo servirá como un avance importante para la identificación de esta enfermedad. La importancia de trabajar puntualmente con el mango radica en que este es una de las frutas que mas se exportan en Colombia, lo que hace que los estándares de calidad deban ser muy altos.

Este documento muestra el proceso de diseño y desarrollo de técnicas de análisis hiperespectral basado en redes neuronales profundas, en el primer capítulo se detallan los principales temas que serán abordados a lo largo del documento, dando al lector las razones por las cuales se desarrolló el trabajo con estas tecnologías y el por qué se trabaja con el cultivo de mango. En el segundo capítulo se presenta la fundamentación teórica de los conceptos básicos que se deben tener para poder entender el desarrollo del documento. En el tercer capítulo se

presentan los procesos realizados a lo largo de este trabajo de forma detallada, se explican que técnicas fueron usadas y cómo se implementaron. En el cuarto capítulo se muestran los resultados obtenidos, así como un análisis de los mismos, buscando de manera lógica una explicación al comportamiento de los modelos. En el quinto y último capítulo se dan unas conclusiones acerca de los resultados obtenidos y se presentan algunas recomendaciones para trabajos futuros.

1.1. Objetivo general

Desarrollar un algoritmo computacional basado en técnicas de aprendizaje profundo y el procesamiento de imágenes hiperespectrales para la detección de antracnosis en mango.

1.2. Objetivos específicos

- Diseñar un algoritmo computacional de procesamiento de imágenes hiperespectrales que permita la adecuada corrección y tratamiento de datos.
- Implementar técnicas de inteligencia artificial para la detección de antracnosis en mango basadas en redes neuronales convolucionales.
- Validar el rendimiento de las técnicas de inteligencia artificial propuestas mediante un análisis estadístico de los resultados, usando datos de cultivos de mango.

2 Marco teórico

Previo al tratamiento de los datos, al desarrollo de los modelos de aprendizaje profundo y a los experimentos establecidos para este trabajo, se hace necesario presentar la fundamentación teórica acerca de los conceptos relacionados con la temática referente a los datos espectrales, el procesamiento de imágenes y el aprendizaje profundo con miras a que el lector se familiarice con los términos que serán usados a lo largo de este documento.

2.1. Antracnosis

La antracnosis es considerada como la principal enfermedad que afecta a los cultivos de mango, la cual afecta de manera considerable las hojas, flores y frutos. Esta es una enfermedad fungica provocada por el hongo Colletotrichum, la cual afecta no solo a los mangos, si no también a una gran cantidad de frutas y hortalizas, por ejemplo se han encontrado estudios donde analizan sus efectos sobre papaya, guanábana, oliva, murucuyá, entre muchos otros[12][11][24][20]. La antracnosis se presenta en los cultivos en estado de maduración, manifestándose como una cantidad de puntos negros que van aumentando en cantidad y tamaño, propagándose por toda la superficie hasta el punto de afectar al fruto completo provocándole necrosis completa. Por estos motivos esta enfermedad hace que la calidad y cantidad de frutos cosechados disminuya afectando considerablemente los procesos de comercialización.

En la Figura **2-1** se observa el efecto que tiene la enfermedad sobre hojas de varios tipos de mango que fueron inoculadas con la enfermedad, iniciando como una serie de manchas pequeñas color marrón, que van creciendo gradualmente a lo largo de los días hasta llegar a cubrir la totalidad de la hoja, de manera similar se presenta en los frutos de mango, al ser esta una enfermedad fúngica su principal método de propagación es por aspersión, es decir, las partículas de este hongo viajan por el aire y se van aferrando de los frutos y hojas que encuentran.

Figura 2-1: Pruebas de patogenicidad de diferentes especies de mangos. A) Filipino B) Señora C) Filipino D) Keitt E) Haden F)Haden; G-Señora H-Keitt). Imagen tomada de [24]

2.2. Imágenes y firmas hiperespectrales

Las imágenes hiperespectrales son imágenes que recopilan la información espectral y espacial que caracteriza al objeto que está siendo capturado, es decir, muestran el objeto bajo estudio en diferentes longitudes de onda. El espectro electromagnético está compuesto por cientos de bandas que representan los diferentes tipos de energía luminosa, este espectro comprende desde las longitudes de onda más cortas (10^{-5} μm), hasta las longitudes de onda más largas (10^{6} μm). Las imágenes hiperespectrales contienen información de múltiples bandas que se encuentran una tras otra, por lo cual se puede decir que se tiene información del espectro

continuo o de la firma espectral del objeto. En la Figura **2-2** se observa una imagen representada en 4 bandas diferentes, aunque todas provienen de la misma superficie la información contenida en ellas es única con respecto a sus valores de reflectancia.

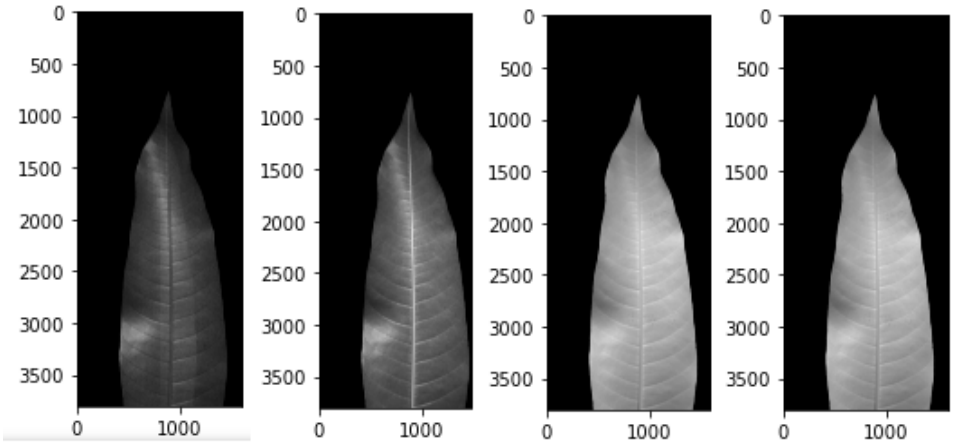

Figura 2-2: Representación de una hoja en diferentes bandas espectrales

Por otro lado, las firmas espectrales como la que se observa en la Figura **2-3** son la forma en la que se representa la radiancia a través del espectro en un píxel especifico de la imagen, cada objeto y cada superficie tiene una interacción específica con la radiación, es decir absorbe o refleja diferentes longitudes de onda. Las firmas espectrales pueden variar dependiendo del objeto de estudio, inclusive una misma superficie o un mismo objeto pueden tener firmas diferentes debido a cambios de incidencia de luz, ángulos de iluminación, falta de homogeneidad en la superficie, entre otros.

2.3. Agricultura de presición (AP)

En la actualidad existe un creciente interés por aumentar la producción agrícola, así como también mejorar la calidad de dichas producciones. Es por esto que cada vez con más frecuencia se buscan y se implementan soluciones del tipo tecnológico a esta problemática. Por lo anterior surgió la agricultura de presición, la cual consiste en usar la tecnología con el fin de mejorar la calidad y cantidad de productos cultivados . Dentro de las tecnologías que comúnmente se usan dentro de la AP están el posicionamiento geoespacial, la recolección de datos de interés (humedad, temperatura, topografía, radiación solar, entre otros) y toma de fotografías de manera terrestre o aérea, con el fin de determinar anomalías en los cultivos. Además de recolectar información también es posible tomar acciones usando los datos obtenidos, es decir activar sistemas de riego, controlar la exposición a la radiación solar, o

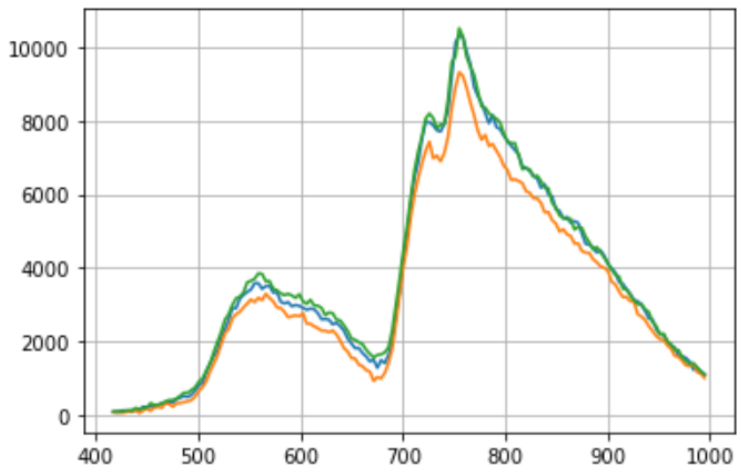

Figura 2-3: Firma espectral de varios píxeles de una imagen.

controlar malezas, plagas o enfermedades suministrando dosis variables de insumos agrícolas. [13][7].

Combinando los dos conceptos anteriormente mencionados cabe resaltar que una de las tecnologías usadas recientemente dentro de la AP son las imágenes hiperespectrales, estas imágenes permiten observar y analizar los cultivos bien sea a partir de sus hojas, frutos o cualquier otra parte, en búsqueda de problemas que se puedan presentar tales como malezas o enfermedades, que puede que no sean evidentes ni para un profesional experimentado, ya que gracias al uso de las imágenes hiperespectales es posible cuantificar las características biofísicas, bioquímicas y estructurales de los cultivos, mediante el análisis de las firmas espectrales de cada imagen o inclusive de cada píxel [18].

2.4. Procesamiento digital de imágenes

El procesamiento digital de imágenes en el campo tecnológico es conocido támbien como visión por computador. Se refiere principalmente a una serie de algoritmos y procesos usados con el fin de extraer información característica de imágenes digitales para poder analizarlas en diferentes ámbitos, bien sea en su composición de color, en búsqueda de formas características o en búsqueda de datos que puedan ser relevantes para un estudio en particular. Estos procesos van desde la toma y adquisición de las imágenes, pasando por un preprocesamiento (cambios de escala, espacios de color, corrección de ruido, filtros, segmentación, etc), y una extracción de características, la visión por computadora se ha ido masificando de manera acelerada debido a que puede ser usada en múltiples aplicaciones de la vida cotidiana, industrial, agrícola, comercial, médica, entre otras [5] [14] [8]. A continuación se dará

una breve explicación de los algoritmos usados en el preprocesamiento de las imágenes que se usaron en este estudio en particular:

2.4.1. Espacios de color

Un espacio o modelo de color es una forma de representar de manera númerica los colores presentes en una imágen, por lo general las imágenes están representadas por 3 canales en el espacio de color RGB (Red, green, blue), este es el espacio de color más común y en el que la mayoría de imágenes son representadas. Sin embargo, hay más de 150 espacios de color que tienen diferentes usos dentro de la visión por computador. La idea de hacer cambios en los modelos de color es lograr resaltar características importantes dentro de las imágenes. Algunos de los espacios de color más usados son RGB, BGR, GRAY, YCrCb, XYZ, HSV, Lab, Luv, HLS, YUV, sin embargo, para el procesamiento de las imágenes que se usan en este estudio, los espacios de color elegidos fueron el RGB para las imágenes adquiridas, y el HSV para lograr resaltar los contornos en las imágenes. El espacio de color RGB se refiere a los colores primarios rojo, verde y azul, gracias a la combinación aritmética de estos tres, es posible tener una gama de numerosos colores [4]. En la Figura **2-4** se observa el espacio de color RGB en forma de cubo, donde cada nuevo color es obtenido por la proyección de cada uno de los ejes.

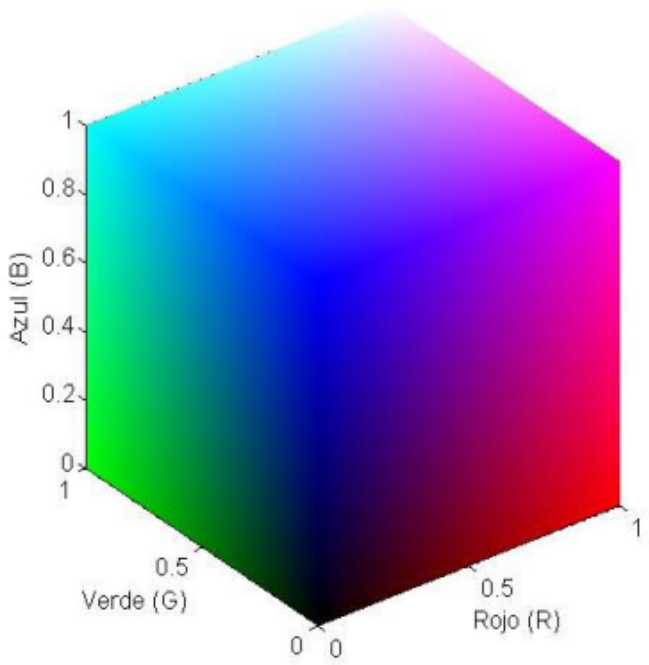

Figura 2-4: Cubo RGB, extraido de [4]

Uno de los principales inconvenientes al trabajar con el espacio de color RGB es al momento

de hacer segmentación de imágenes, ya que esta debería hacerse en 3 dimensiones debido a que los tres canales son dependientes el uno del otro, como se observa en la Figura **2-5** el cambio de espacio de color hace que la hoja presente en la imagen se resalte más y de esta manera sea más fácil reconocer sus contornos. El espacio de color HSV representa el matiz (H - Hue), saturación (S- Saturation) y Valor (V-Value), este espacio de color es comúnmente usado en procesos de segmentación.

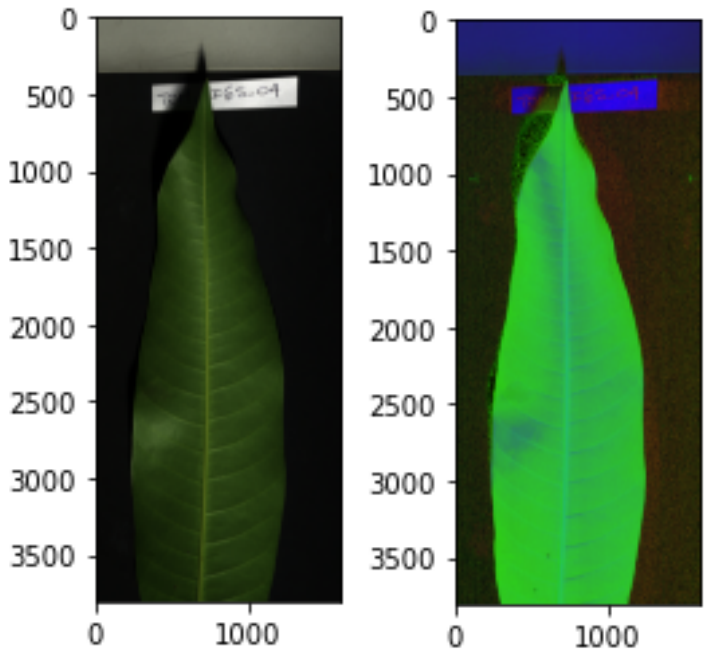

Figura 2-5: Imagen original en RGB vs Imagen en HSV

La forma matemática para convertir las imágenes del espacio de color RGB a HSV tiene en cuenta varias consideraciones, en primer lugar es necesario normalizar los valores de R,G y B esto se hace mediante:

$$R' = \frac{R}{255} \qquad G' = \frac{G}{255} \qquad B' = \frac{B}{255} \tag{2-1}$$

Seguido de esto es necesario determinar los valores máximo (MAX) y mínimo (MIN) de los valores RGB normalizados, las componentes HSV se calculan:

$$H = \begin{cases} 0 & si \quad MAX = MIN \\ (60^o * \frac{G'-B'}{MAX-MIN} + 360^o) * mod360^o & si \quad MAX = R' \\ (60^o * \frac{B'-R'}{MAX-MIN} + 120^o) & si \quad MAX = G' \\ (60^o * \frac{R'-G'}{MAX-MIN} + 240^o) & si \quad MAX = B' \end{cases} \tag{2-2}$$

$$S = \begin{cases} 0 & si \quad MAX = 0 \\ \frac{MAX-MIN}{MAX} = 1 - \frac{MIN}{MAX} & si \quad MAX \neq 0 \end{cases} \tag{2-3}$$

$$V = MAX \tag{2-4}$$

Estos cálculos se muestran de forma meramente informativa ya que varias de las librerías de PYTHON que sirven para el manejo de imágenes, traen dentro de sus funciones estos cambios de color y los hacen de manera automática usando un par de instrucciones.

2.4.2. Segmentación

El proceso de segmentación consiste en separar los objetos presentes en una imagen, esto se logra aplicando una serie de reglas para colocar cada píxel de la imagen dentro de un subgrupo, la segmentación es usada por lo general para localizar objetos o determinar los límites de un objeto dentro de una imágen, existen varios métodos para hacer la segmentación dentro de los cuales se encuentran segmentación basada en umbral, segmentación basada en bordes, segmentación basada en regiones, segmentación basada en grupos, entre otros. En este caso en particular, se hizo uso del método de umbral que consiste en asignar un valor de 0 o 1 a cada píxel, dependiendo si este es menor o mayor a un umbral establecido. Esto de forma matemática se ve como:

$$g(x,y) = \begin{cases} 0 & si \quad f(x,y) < T \\ 255 & si \quad f(x,y) \geq T \end{cases} \tag{2-5}$$

Donde $f(x,y)$ es una imagen en escala de grises, T es el valor del umbral seleccionado, y $g(x,y)$ la imagen resultando en blanco y negro o lo que es lo mismo en valores de 0 o 255.

2.4.3. Reducción de ruido

El filtrado para reducción de ruido se usa con el fin de suavizar las imágenes eliminando pequeños puntos aleatorios que son considerados ruido. El ruido es considerado una variable aleatoria con media 0, gráficamente son píxeles o vecindarios de píxeles que se encuentran de manera aleatoria y que no coinciden con la estructura de la imagen. Existen varias técnicas de filtrado, desde técnicas de desenfoque y suavizado, las cuales logran eliminar el ruido hasta cierto punto, hasta algoritmos más robustos que logran determinar y suavizar el ruido de una manera más precisa.

2.4.4. Operaciones morfológicas

La palabra morfología significa forma o estructura de un objeto. En el tratamiento de imágenes, las operaciones morfológicas son operaciones matemáticas basadas en formas, las cuales consisten en analizar la imagen con respecto a un objeto en particular llamado elemento estructurante, dentro de las principales operaciones se encuentran la dilatación, la erosión, el opening y el closing, estas operaciones buscan resaltar algunos detalles o eliminar otros, por lo que suelen ser usadas como filtros de reducción de ruido dentro de las imágenes.

2.5. Inteligencia artificial

El término de inteligencia artificial es demasiado amplio y extenso, ya que este se puede encontrar en múltiples campos de la vida, de esta misma manera, cada campo lo busca definir. Sin embargo, una definición amplia, según [22] es, “El esfuerzo para automatizar tareas intelectuales normalmente realizadas por humanos”. A través del tiempo se ha buscado dotar de esta “ inteligencia” a los computadores, máquinas o dispositivos electrónicos, a su vez se ha ido incrementando la cantidad de tareas automatizadas que son posibles. Visto de otra manera se podría llegar a interpretar como la capacidad que tienen las máquinas para tomar datos, procesarlos, obtener información y tomar decisiones a partir de la información procesada, aunque en realidad esta capacidad que tienen las máquinas es posible gracias a la inteligencia humana que las programa para que ejecuten estas acciones de aprendizaje.

2.5.1. Aprendizaje de máquina (Machine learning)

El machine learning se encuentra inmerso dentro de la inteligencia artificial como un subcampo de ella, este se puede definir como la capacidad que pueden obtener las máquinas de adquirir un conocimiento o realizar un aprendizaje sin que sean explícitamente programadas las reglas necesarias para dicho aprendizaje. A diferencia de la inteligencia artificial en general, el machine learning se encarga de aprender patrones a partir de la información de entrada y generar predicciones a partir de los patrones aprendidos. El machine learning se

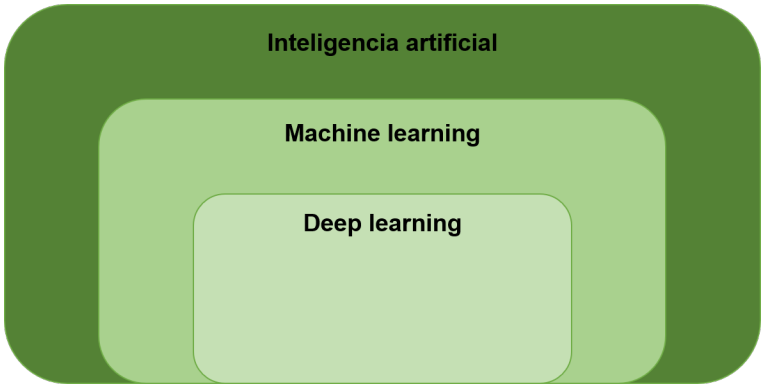

Figura 2-6: Representacion gráfica del concepto de deep learning, fuente propia

puede dividir en 3 categorías, el aprendizaje supervisado, el aprendizaje no supervisado y el aprendizaje por refuerzo. En el aprendizaje supervisado cada dato de entrada tiene asociada una etiqueta, es decir se le está diciendo explícitamente a la máquina a que clase o grupo perteneces este dato. En el aprendizaje no supervisado, los datos no se encuentran etiquetados, por lo que el trabajo del algoritmo es crear clases o grupos a partir de las similitudes o patrones que encuentre en los datos de entrada. Por último, se encuentra el aprendizaje por refuerzo, este tipo de algoritmos busca explorar espacios desconocidos generando conocimiento a partir de prueba y error, este aprenderá a partir de las penalizaciones que le vayan siendo impuestas.

2.5.2. Redes neuronales artificiales (ANN)

Las ANN son una manera de representar las redes de neuronas humanas, mediante ellas se busca imitar la manera en que los humanos adquieren conocimiento, básicamente de lo que se tratan es de redes de neurodos que tienen una entrada de datos, estos datos son procesados y entendidos por la red, la cual entrega una salida, a los datos que se obtienen en la salida por lo general se le hace una retroalimentación para que la red pueda saber o entender si la información que se esperaba a la salida coincide con la información que se obtuvo.

2.5.3. Deep learning

El deep learning se puede entender como redes con muchas capas intermedias que son las capas encargadas del procesamiento de la información, estas capas son las llamadas capas profundas. Una de las características principales del aprendizaje profundo es que gracias a la gran cantidad de capas que manejan la información de entrada, esta se procesa desde múltiples niveles de abstracción. La red comienza procesando características a nivel macro, y a estas características la red vuelve a analizarlas con el fin de encontrar nueva informa-

cion relevante y así consecutivamente hasta un criterio de parada definido. Uno de los usos principales del deep learning, y tal vez la razón por la que ha logrado tener tanta acogida es su capacidad de interpretar el lenguaje natural, bien sea en forma de texto, voz o imágenes.

Una de las técnicas usadas dentro del mundo del deep learning son las llamadas CNN o redes neuronales convolucionales, estas redes son usadas principalmente con el fin de analizar información de imágenes, mediante la aplicación de filtros en diferentes capas y niveles de abstracción, es posible ir extrayendo información cada vez más detallada, en la Figura **2-7** se puede observar la representación de una CNN en la cual la figura roja de la izquierda representa la entrada, en este caso una imagen de 3 dimensiones, alto, ancho y la profundidad que son los canales RGB. Esta imagen es descompuesta por kernels y se extraen sus características principales organizándolas en forma de vector, este proceso se repite cuantas veces lo requiera el problema a solucionar.

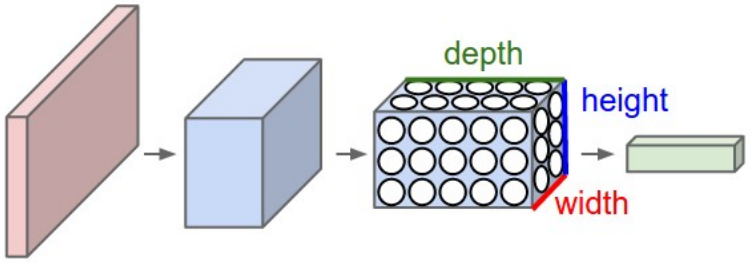

Figura 2-7: Representacion de una red neuronal convolucional (CNN), tomada de [3]

2.6. Métricas de desempeño

Para poder determinar que tan bueno es el modelo que se está usando es posible usar diferentes métricas que ayudan a medir su desempeño. Dentro de estas se encuentran metricas de exactitud, metricas probabilísticas, metricas de regresión y metricas de clasificación basadas en verdaderos/falsos positivos y negativos, dentro de este último grupo es posible encontrar métricas tales como área bajo la curva (AUC por sus siglas en ingles), recall, presición, sensitividad y especificidad, y una de las más usadas que es la matriz de confusión, esta es una tabla que describe el desempeño del modelo de manera gráfica y en la cual se puede visualizar que tantos datos fueron bien clasificados y que tantos datos fueron clasificados de manera errónea y a que clase fueron asignados.

Las metricas que se usan para evaluar el desempeño de los experimentos realizados son 4 LOSS, ACCURACY, MSE, PRECISION y RECALL.

LOSS se refiere a la función de perdida del modelo, es decir evalúa la desviación entre las predicciones realizadas y el valor real de las observaciones, para esta métrica entre menor sea el valor quiere decir que la desviación es menor y así mismo que las predicciones realizadas por el modelo coinciden con las observaciones. ACCURACY o exactitud por su traducción al español calcula la cercanía entre predicciones y las observaciones, en términos estadísticos esta está relacionada con los sesgos de una estimación, esta métrica esta dada como un porcentaje entre un 0 y 100 % por lo que a diferencia de la anterior entre mayor sea este porcentaje quiere decir que las predicciones del modelo coinciden con las observaciones realizadas. MSE o error cuadrático medio mide el error cuadrado promedio de las predicciones, para cada una de las predicciones calcula la diferencia cuadrada entre las predicciones y la observación y luego promedia estos valores, entre mas alto sea este valor, peor es el comportamiento del modelo. PRECISION o precisión por su traducción al español, evalúa la cantidad de verdaderos/falsos positivos y negativos que son clasificados correctamente, entre menor es la dispersión mayor la precisión. Por ultimo RECALL o sensitividad se refiere a la tasa de verdaderos positivos es decir cuenta la proporción de casos positivos que fueron correctamente identificados.

3 Análisis hiperespectral basado en redes neuronales profundas

3.1. Identificación del problema

Colombia es uno de los 10 exportadores más grandes del mundo de frutas tropicales, encontrando el mango como uno de sus productos con mayor influencia en mercados internacionales, por lo que se hace necesario mejorar la calidad de los productos que son exportados buscando crear técnicas que ayuden a detectar de manera temprana enfermedades en los cultivos.

Actualmente la detección de enfermedades o problemas fitosanitarios que existen en los cultivos se hace mediante inspecciones visuales realizadas por experto agrónomos o fitopatologos, debido a esto las revisiones se hacen de manera sesgada, superficial y sensible a la fatiga visual y al conocimiento previo del experto. La detección visual hace que la detección temprana no sea posible, es decir, para que la enfermedad sea perceptible al ojo humano deben pasar varios días de desarrollo y posiblemente también se haya esparcido a otros sectores del cultivo. Por lo anteriormente descrito se hace necesario crear modelos de identificación y clasificación de patrones que ayuden a la detección de enfermedades. Se propone entonces la detección de antracnosis en mango, en una fase temprana para poder mejorar la calidad de los productos exportados y mejorar la productividad del sector agrícola, adicional se busca disminuir el impacto ambiental y en general ofrecer productos libres de agentes químicos al consumidor final.

3.2. Solución propuesta

Para el desarrollo de la solución propuesta al problema identificado de como lograr hacer una detección de antracnosis en cultivos de mango de manera temprana y no invasiva, se planteó el uso de técnicas y algoritmos de visión artificial y aprendizaje profundo (DL por sus siglas en inglés). Este enfoque tiene varias ventajas, una de ellas es que gracias al uso de técnicas de aprendizaje profundo es posible saltar varios procesos previos que son necesarios cuando se hace machine learning, como la extracción de características, ya que los algoritmos de DL se encargan precisamente de ir analizando y encontrando las características principales dentro

de las imágenes. Otra de las ventajas importantes es el uso de imágenes hiperespectrales, ya que gracias al uso de estas es posible analizar la información no solo de manera espacial si no que también de manera espectral, y así poder observar de manera detallada los elementos que caracterizan cada uno de los grupos de clasificacion, ya que en este caso en particular la información espacial proporcionaría únicamente detalles acerca de la forma de las hojas y no seria posible identificar la enfermedad. Existen estudios que muestran los beneficios de usar estas técnicas en conjunto, es decir, analizar las imágenes hiperespectrales con técnicas de aprendizaje profundo [5] [25] [19], dentro de los beneficios que se muestran son la capacidad de poder analizar de manera paralela las características espaciales y espectrales propias de los objetos observados, ya que además de hacer detección de alguna característica particular es posible segmentar y clasificar imágenes dependiendo de los diferentes valores de reflectancia presentes en la imagen, entre otras muchas aplicaciones.

Gracias a una revisión de los métodos usados actualmente para la identificacion de esta enfermedad se logra evidenciar que no hay técnicas de este estilo para su detección y esta se hace de forma visual, lo cual puede que se haga en un momento en que la enfermedad ya este bastante avanzada, no tenga una solución óptima y de tenerla sea demasiado invasiva, el proceso que se propuso trabajar es el que se puede evidenciar en la Figura **3-1**.

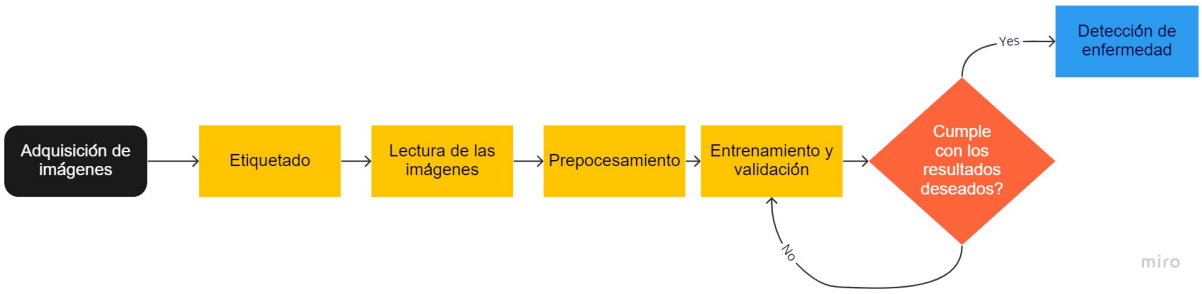

Figura 3-1: Diagrama del proceso de análisis hiperespectral basado en redes neuronales profundas

Si bien existen en la academia, estudios acerca de la detección de esta enfermedad con el uso de imágenes hiperespectrales como [12] [20] [11], estos hacen uso de técnicas de machine learning o se centran en otro tipo de cultivos, por lo que se puede afirmar que el enfoque que se está implementando no ha sido ampliamente investigado, además, puede ayudar a mejorar la calidad de los frutos cosechados y comercializados. Como alternativas a los trabajos mencionados se plantea hacer uso de imágenes hiperespectrales completas, ya que en la mayoría de los casos por falta de capacidad de procesamiento se hace reducción de dimensionalidad, lo cual hace que se pierda información que en teoría no es relevante pero igual puede ayudar

a mejorar la calidad de los modelos, ya que una de las características de los modelos de DL es que a mayor cantidad de información estos pueden mejorar los resultados de la clasificación.

En el diagrama de flujo de la Figura **3-1** se muestran las etapas que se llevaron a cabo, las cuales se detallaran en las siguientes secciones. En primer lugar está la adquisición de las imágenes, en este caso imágenes hiperespectrales, se explica por que se usan este tipo de imágenes, como se componen y como es su procesamiento mediante python. Se dará una explicación de la etapa de etiquetado de las imágenes la cual juega un papel muy importante en los siguientes pasos del proceso. Las etapas de lectura y preprocesamiento se desarrollarán en una sola sección y se detallará que algoritmos, librerías y demás fueron usadas. Por último, la etapa de entrenamiento y validación de las redes de aprendizaje profundo, en esta sección se hace un analísis de los modelos que fueron escogidos, como funcionan y que aporte proporcionan a la solución del problema anteriormente mencionado.

3.3. Adquisición de imágenes

El sistema de visión utilizado para la captura de imágenes hiperespectrales corresponde a un sensor pushbroom Hyspex VNIR- 1600 (Hyspex by Neo, Norsk Elektro Optikk, Noruega), con resolución espectral de 3.629383 nm, 1600 píxeles espaciales y rango de 400-1000 nm, obteniendo 160 bandas por escaneo. El sensor hiperespectral se colocó sobre la hoja a una distancia de 30cm, la cual se iluminó con una fuente de luz halógena VNIR de 150W a 45 grados sobre el plano donde se encuentra la muestra a fotografiar. Los archivos obtenidos por este objetivo tienen un formato propio de la marca .HYSPEX, para la manipulación de las imágenes se hizo uso del lenguaje de programación PYTHON, el cual es un lenguaje de alto nivel y de código abierto. Para el analísis principal de las imágenes y su composición se utilizó la librería SPECTRAL, esta librería tiene como finalidad la manipulación en general de imágenes hiperespectrales la cual tiene módulos de lectura, escritura, visualización y clasificación.

La toma de las muestras se realizó en un ambiente con iluminación controlada, lo cual representa una mayor facilidad al momento del procesamiento. Para este experimento se tomaron 5 muestras de hojas sanas, 5 muestras de hojas inoculadas y 5 muestras de hojas enfermas, de cada una de las hojas se hizo un registro durante 6 días, donde el día 0 es el día en que ninguna hoja estaba inoculada, el día 1 se hizo el proceso de inoculación y de allí en adelante se fue registrando su desarrollo hasta el día 6. De esta manera se obtuvieron al rededor de 90 imágenes hiperespectrales con las cuales se busca capturar la aparición de la enfermedad en distintas etapas de desarrollo.

3.4. Etiquetado

Anteriormente se mencionó que el proceso de etiquetado de las imágenes se colocaba dentro del diagrama como una etapa separada y muy importante. Aunque no es un proceso demorado, es indispensable que este se haga de manera rigurosa ya que al momento de hacer la lectura de las imágenes mediante PYTHON, estas etiquetas son las que van a definir a que clase se le está informando al programa que pertenecen las imágenes. Para este caso en particular se separaron las imágenes siguiendo varios criterios con el fin de realizar varios experimentos. En primer lugar y como se puede ver de manera gráfica en la Figura **3-2**, se separaron todas las imágenes en 3 grupos CONTROL, INOCULADAS y ENFERMAS, a su vez estos 3 grupos se encuentran separados por hoja, es decir, por cada una de las 5 muestras, y por último se encuentra el registro de cada uno de los 6 días. El grupo de imágenes que se menciona como control corresponde a las muestras sanas, las cuales van a servir de guía para que las redes neuronales entiendan como se ve una hoja sana y como se ven las hojas enfermas e inoculadas.

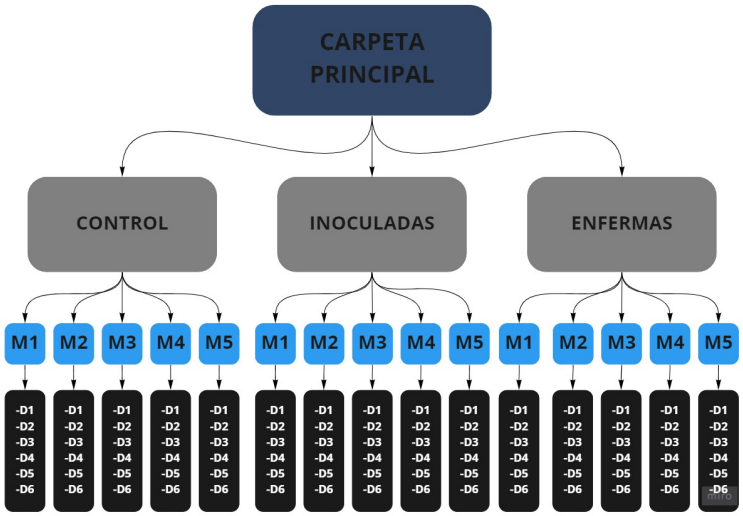

Figura 3-2: Disposición 1 de las imágenes hiperespectrales, separadas por clases

Posterior a esta separación y etiquetado se opta por hacer una reubicación con el fin de experimentar y determinar a partir de que día se puede hacer una clasificación de manera confiable, por lo que en este caso se hizo la separación de las imágenes por días, es decir, se crearon grupos de imágenes desde el día 1 hasta el día 6, cada día se separa en los 3 grupos control, inoculadas y enfermas, y por último se encuentran las 5 muestras. Con esta separación es posible observar como es el desempeño de la clasificación de las redes usadas dependiendo de el día de desarrollo de la enfermedad en las hojas.

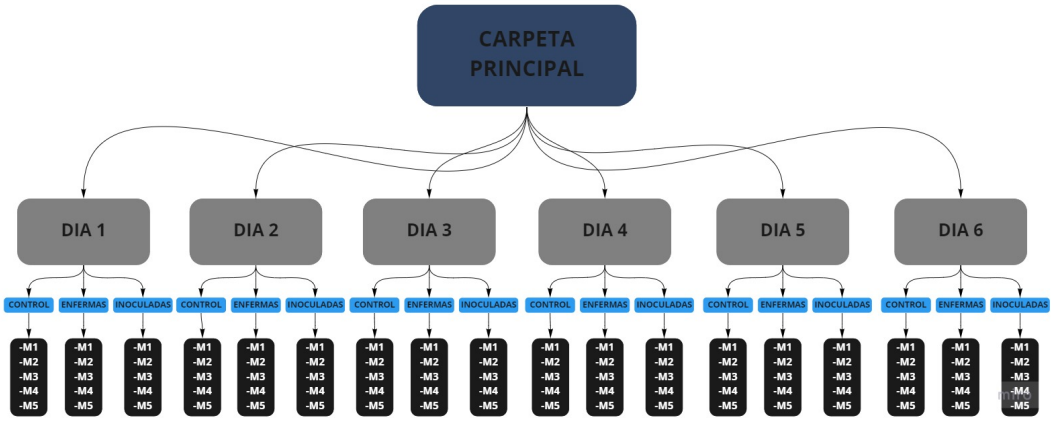

Figura 3-3: Disposición 2 de las imágenes hiperespectrales, separadas por dias

3.5. Lectura y preprocesamiento

Para la etapa de lectura de las imágenes se hizo uso de la librería SPECTRAL de python, la cual se configuró con el fin de poder manipular imágenes hiperespectrales en diferentes formatos. En el caso en particular de los experimentos realizados existen dos maneras de leer la información para luego ser procesada, el primero consiste en leer unicamente los metadatos de la imagen, esto es debido a que estos archivos pueden ser demasiado grandes, por ejemplo las imagenes que se utilizaron en los experimentos de este trabajo tienen un tamaño de 2GB cada una lo cual hace que se deba tener un espacio importante en memoria. Cuando se leen los archivos de esta manera lo que observa el programa son los siguientes datos:

RowData:
Data Source: /Inoculated/Sample 1/tom igs 01 16000 us 2x 2019-11-24T124442 corr.hyspex
Rows: 3800
Samples: 1600
Bands: 160
Interleave: BIL
Quantization: 16 bits
Data format: uint16

Al leer los archivos de esta manera se pueden hacer algunas tareas como visualizar sus firmas espectrales, visualizar los hipercubos correspondientes a esa imagen, convertirlas a RGB u otros formatos y particularmente conocer la información de la composición. Sin embargo, esta manera no proporciona como tal una imagen o las matrices correspondientes a una imagen, por lo que se hace necesario optar por leer completamente las imágenes. El programa

puede mostrar las imágenes como matrices de las dimensiones correspondientes al archivo en este caso 3800x1600x160, es decir, está visualizando un hipercubo de 160 bandas de imágenes de 3800 x 1600 píxeles.

Posterior a la lectura de los datos se lleva a cabo una etapa de preprocesamiento, la cual consiste en preparar las imágenes de tal manera que puedan ser leídas por las redes neuronales, haciendo ajustes de tamaño, color, segmentado y eliminación de ruido, este preprocesamiento será detallado a continuación.

3.5.1. Cambio de espacio de color

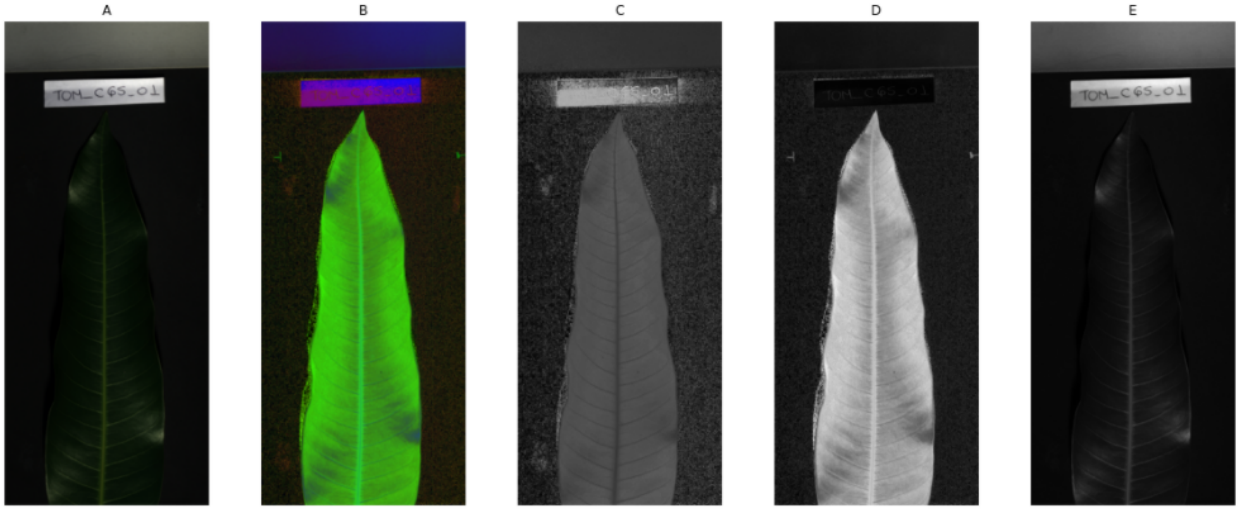

Figura 3-4: Etapas del cambio de espacio de color: A) Imagen original representada en espacio de color RGB. B) Imagen representada en espacio de color HSV. C) Plano H del espacio de color HSV. D) Plano S del espacio de color HSV. E) Plano V del espacio de color HSV

Los cambios de espacio de color en visión artificial se usan con el fin de resaltar algunas de las características más importantes dentro de la imagen, en este caso en particular y después de ensayos entre varios espacios de color, se determina que el espacio que mejor resultado tiene es el HSV. Como se ve en la Figura **3-4** este espacio de color proporciona un contraste muy alto entre el fondo de la imagen y el contorno de la hoja, lo cual va a servir más adelante para poder segmentarla y generar una mascara eliminando el fondo negro de la imagen que no aporta ninguna información de valor, en la Figura **3-4** B se muestra la imagen de una hoja representada en el espacio de color HSV y esta está compuesta por 3 canales (matiz, saturación y valor), sin embargo, al separar los canales se obtienen 3 imágenes en escala de grises de las cuales la imagen correspondiente al canal de saturación es la que proporciona el mayor contraste de las tres, razón por la cual fue la escogida para continuar con el proceso.

3.5.2. Segmentación de las imágenes

Luego de obtener la imágen adecuada con el suficiente contraste, se realiza el proceso de segmentación con el fin de separar únicamente la hoja, que es el objeto de interés de estudio, para este proceso se usa la técnica segmentación por umbrales o umbralización, este proceso es el más adecuado, ya que gracias al contraste que hay entre la hoja y el fondo, el histograma que representa esta imagen y sus niveles de grises, proporciona la información de cuales son los niveles de grises predominantes. Este histograma tiene una forma de dos lóbulos los cuales pueden ser separados por un umbral, en la Figura **3-5** se observa el histograma correspondiente a una de las imágenes que se usaron, donde se puede analizar que un valor intermedio entre los dos lóbulos está entre 80 y 140, estos valores cambian entre cada una de las imágenes pero para este caso se usó un umbral de 80 para todas, este proceso en este caso fue iterativo aunque existen ya métodos que lo hacen de manera automatizada. La intención de hacer esta segmentación es poder generar lo que se conoce en visión artificial como una máscara, la cual es una imagen binaria que se combina con otra imágen con el fin de separar únicamente la porción correspondiente a la máscara de la imagen original.

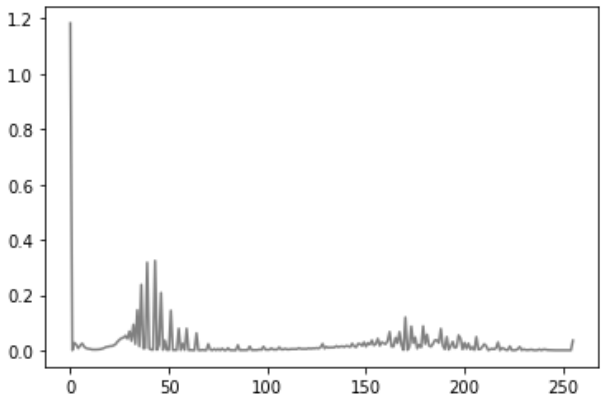

Figura 3-5: Histograma de una imagen del plano S

Como se observa en la Figura **3-6**, en el proceso de umbralización se genera una imagen con una gran cantidad de ruido, en general es un termino usado para todas aquellas modificaciones no deseadas que sufre una señal, o en este caso, una imagen.

3.5.3. Operaciones morfológicas

Como se evidencia en la Figura **3-6**, luego de hacer la segmentación por umbral quedan algunos detalles como puntos negros y blancos aislados de la imagen de la hoja, es necesario

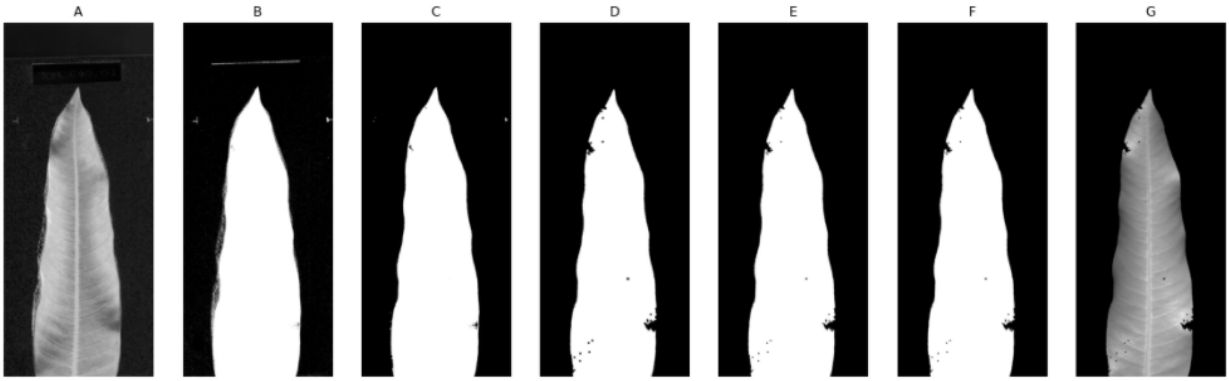

Figura 3-6: Preprocesamiento de la imagen: A) Plano S del espacio de color HSV. B) Imágen de una hoja después del proceso de umbralización C) Imagen de una hoja después del proceso de opening. D) Imagen de una hoja después del proceso de erosión E) Imágen de una hoja después del proceso de dilatación F) Máscara obtenida de la imagen de la hoja G) Imágen de una después de ser combinada con la máscara

eliminar dichos artefactos para mejorar la calidad de clasificación posterior. La primera operación que se usó fue la denominada OPENING, esta operación realiza un proceso de erosión seguido de un proceso de dilatación, al aplicar esta operacion se observa que tiene buenos resultados eliminando la parte de ruido correspondiente a la sombra generada por la hoja. Sin embargo, se observan puntos de ruido en otras partes, sobretodo puntos blancos en areas negras, por lo que seguido de esta operación se utiliza la erosión. Para esta operación se hace uso de un kernel que recorre la imagen y va revisando si todos los píxel debajo del kernel son 1 (blancos), en caso de no ser así vuelve este píxel 0 (negro). Sin embargo, al hacer esta operación se observa la aparición de otro tipo de ruido en este caso puntos negros en regiones blancas, para lo cual se usa la operación de dilatación, mediante la cual se busca expandir las áreas blancas donde sea necesario, por último se obtiene la máscara que es aplicada a la imagen original, eliminando todo aquello que no nos proporciona información importante o que puede generar clasificaciones erróneas[1].

En la Figura **3-7** se observa la imagen resultante después de los procesos de segmentación, eliminación de ruido y combinación con la máscara, imágenes como esta son la información de entrada de las redes neuronales que serán usadas.

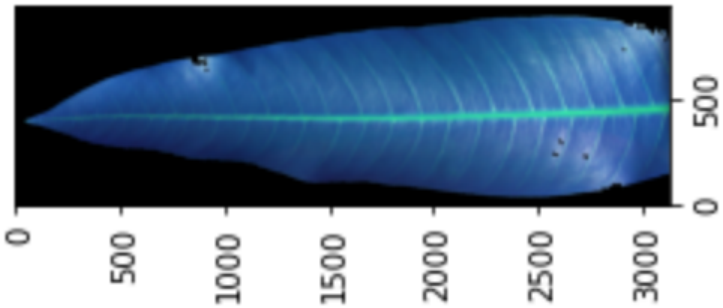

Figura 3-7: Imagen resultante después del preprocesamiento

3.6. Aprendizaje profundo

Después de tener las imágenes de las hojas como la que observa en la Figura **3-7** se comienza a trabajar en las redes neuronales que se van a usar. El primer paso es etiquetar una a una las imágenes leídas, se crearon entonces 3 etiquetas correspondientes a los grupos de control, inoculación y enfermas, y a cada imagen se le crea una etiqueta numérica dependiendo el grupo al que pertenecen. Esta información es almacenada en un vector con un tamaño igual a la cantidad de imágenes que se van a usar. Seguido a este proceso de etiquetado se hace una mezcla de todos los datos que se leyeron, para luego crear dos grupos, un grupo de entrenamiento y un grupo de test, en este caso en particular se hizo una división del 80 % para entrenamiento y 20 % para test, aunque estos valores pueden cambiar dependiendo del problema que se vaya a abarcar. Para este trabajo se hizo uso de dos modelos de deep learning basados en redes neuronales convolucionales, el primero es una red neuronal convolucional en dos dimensiones (CNN2D) y el segundo modelo se trata de una red neuronal convolucional en tres dimensiones (CNN3D), se aplicaron los dos modelos para los 6 días de los que se tienen muestras de las hojas, los resultados de esos experimentos se muestran y discuten en el siguiente capítulo.

3.6.1. Red neuronal convolucional de dos dimensiones (CNN2D)

El modelo que se usó es un modelo secuencial lo que quiere decir que solo se pueden agregar y procesar capas de manera secuencial, el modelo diseñado consta de 17 capas en total que están dispuestas como se observa en la Figura **3-8**, las imágenes que ingresan a este modelo tienen un tamaño de 640x200, dado que las imágenes originales tienen un tamaño de 3800 x 1600, se hizo necesario un redimensionamiento ya que las imágenes son muy grandes. La forma en la que se dividieron las imágenes para ingresarlas a este modelo fue por bandas, es decir, cada una de las 160 bandas de cada una de las imágenes hiperespectrales se presenta por separado.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 640, 200, 64)      1664
_________________________________________________________________
leaky_re_lu (LeakyReLU)      (None, 640, 200, 64)      0
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 320, 100, 64)      0
_________________________________________________________________
dropout (Dropout)            (None, 320, 100, 64)      0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 320, 100, 128)     204928
_________________________________________________________________
leaky_re_lu_1 (LeakyReLU)    (None, 320, 100, 128)     0
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 160, 50, 128)      0
_________________________________________________________________
dropout_1 (Dropout)          (None, 160, 50, 128)      0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 160, 50, 256)      819456
_________________________________________________________________
leaky_re_lu_2 (LeakyReLU)    (None, 160, 50, 256)      0
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 80, 25, 256)       0
_________________________________________________________________
dropout_2 (Dropout)          (None, 80, 25, 256)       0
_________________________________________________________________
flatten (Flatten)            (None, 512000)            0
_________________________________________________________________
dense (Dense)                (None, 256)               131072256
_________________________________________________________________
leaky_re_lu_3 (LeakyReLU)    (None, 256)               0
_________________________________________________________________
dropout_3 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_1 (Dense)              (None, 3)                 771
=================================================================
Total params: 132,099,075
Trainable params: 132,099,075
Non-trainable params: 0
_________________________________________________________________
```

Figura 3-8: Modelo CNN2D

En el modelo que se presenta en la Figura **3-8** se comienza por una capa convolucional 2D, en esta capa intervienen algunos parámetros como el número de filtros que para la primera capa son 64 de tamaño 5x5, con una función de activación lineal y el tamaño de las imágenes que entran en esta capa es de 640 x 200 x 1. Luego de esta capa se añade una capa con una función de activación leaky reLu, esta función de activación busca eliminar linealidad al momento de la clasificacion, por otro lado, busca eliminar sensibilidad. La siguiente capa es una capa Max Pooling de tamaño 2x2, la cual busca reducir el tamaño de las imágenes a la mitad sin perder las características más importantes que detecto en cada uno de los filtros, la capa que sigue es una capa de dropout, la cual busca eliminar el problema de overfiting o sobre entrenamiento que se presenta cuando se usan capas totalmente conectadas.Las siguientes capas son una replica de estas 4 con la diferencia que las siguientes capas convolucionales usan el doble de filtros, es decir 128 y 256, la capa flatten se usa con el fin de aplanar la información, esta capa lo que hace es volver tomar unos datos de entrada y devolverlos en forma de vector, la capa DENSE se encarga de conectar todas las salidas de la capa anterior con las entradas de la siguiente es decir hace una capa comple-

tamente conectada, en total los parámetros que va a evaluar la red neuronal son 132.099.075.

Los parámetros y capas que se usaron en este modelo se sintonizaron y se modificaron de manera empírica, se hicieron varios experimentos encontrando en este modelo una buena relación de tiempo de entrenamiento y calidad de clasificación.

3.6.2. Red neuronal convolucional de tres dimensiones (CNN3D)

El segundo modelo que se usa es un modelo basado en redes neuronales convolucionales en tres dimensiones, la diferencia principal con el anterior modelo radica en la forma en que este procesa la información espacial. El modelo CNN2D procesa la información en solo dos planos X e Y, mientras que el modelo CNN3D es capaz de interpretar la información en 3 planos X, Y y Z, gracias a esto no es necesario separar las bandas y crear imágenes de un solo canal, esta es una gran ventaja ya que se puede analizar de manera paralela la información espacial y espectral de las imágenes.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv3d (Conv3D)              (None, 160, 50, 160, 32)  4032
_________________________________________________________________
leaky_re_lu (LeakyReLU)      (None, 160, 50, 160, 32)  0
_________________________________________________________________
max_pooling3d (MaxPooling3D) (None, 80, 25, 80, 32)    0
_________________________________________________________________
dropout (Dropout)            (None, 80, 25, 80, 32)    0
_________________________________________________________________
conv3d_1 (Conv3D)            (None, 80, 25, 80, 64)    256064
_________________________________________________________________
leaky_re_lu_1 (LeakyReLU)    (None, 80, 25, 80, 64)    0
_________________________________________________________________
max_pooling3d_1 (MaxPooling3 (None, 40, 13, 40, 64)    0
_________________________________________________________________
dropout_1 (Dropout)          (None, 40, 13, 40, 64)    0
_________________________________________________________________
flatten (Flatten)            (None, 1331200)           0
_________________________________________________________________
dense (Dense)                (None, 8)                 10649608
_________________________________________________________________
leaky_re_lu_2 (LeakyReLU)    (None, 8)                 0
_________________________________________________________________
dropout_2 (Dropout)          (None, 8)                 0
_________________________________________________________________
dense_1 (Dense)              (None, 3)                 27
=================================================================
Total params: 10,909,731
Trainable params: 10,909,731
Non-trainable params: 0
_________________________________________________________________
```

Figura 3-9: Modelo CNN3D

En la Figura **3-9** se tiene el modelo que se uso como CNN3D, se busca usar las mismas capas que en el anterior modelo orientadas al análisis en tres dimensiones, con el fin de

poder comparar los resultados de los dos modelos, las capas con nombre CONV3D tienen 32 y 64 filtros respectivamente, ambas con kernels de 5x5x5, las capas de MAXPOOLING 3D usan kernels de 2x2x2, el resto de capas tienen las mismas funciones que el modelo anterior.

3.7. Consideraciones previas

Con el fin de interpretar mejor como se obtuvieron los resultados y posiblemente replicarlos se hace una explicación acerca de las herramientas que se usaron para el desarrollo de los experimentos. En primer lugar el equipo usado tiene las siguientes características: una placa madre gigabyte b450m ds3h, un procesador AMD Ryzen 5 3600 de 6 núcleos que trabaja a 3.6 GHz, una tarjeta gráfica NVIDIA Geforce GTX 10 Ti, esta GPU trabaja a 1.4 GHz y tiene una memoria interna de 4 GB GDDR5.

Por otro lado, referente al software se uso el sistema operativo windows 10 Pro de 64 bits, para hacer uso de python se hizo sobre los notebooks de Jupyter, a través de la plataforma de anaconda, la versión de python usada fue 3.8.5, además las principales librerías usadas fueron SPECTRAL, MATPLOTLIB, NUMPY, OPENCV, SKLEARN, KERAS y TENSORFLOW.

4 Resultados

En este capítulo se presentan los resultados obtenidos con los modelos descritos en el Capítulo 3. El desarrollo de los experimentos se dividió en 4 partes.

En primer lugar se usó el modelo CNN2D durante los 5 días, evaluando las métricas de LOSS, ACCURACY, MSE, PRESICION y RECALL. Este mismo experimento se repitió 3 veces, como se observa en la Figura **3-1**, el diagrama de flujo muestra que en el último paso de entrenamiento y validación, si no se llega al resultado esperado, se vuelve a entrenar y validar, razón por la cual se logró evidenciar que con 3 experimentos se podían obtener los resultados esperados y que, al momento de hacer más validaciones, los modelos mantienen las mismas metricas. La segunda parte de los experimentos se hizo con el modelo CNN3D, se realizó el experimento 3 veces igual que el modelo anterior. La tercera parte se hizo con el modelo CNN2D, pero esta vez con una mezcla de los datos del primer y segundo día, luego con el tercero y cuarto día. La última parte del desarrollo de los experimentos se realizó con el modelo CNN3D y con la misma mezcla de datos que el experimento anterior. Cada uno de estos experimentos y sus resultados serán detallados en este capítulo.

4.1. Modelo CNN2D aplicado a las muestras por día

Para este modelo, que fue explicado en la sección 3.6.1, los experimentos realizados usaron las imágenes hiperespectrales separando sus bandas y leyendo cada banda como una imagen diferente. Se estableció de manera visual, gracias a las firmas espectrales de las hojas, que las bandas que tienen una mayor variación en presencia de la enfermedad están entre la banda 100 y 140, correspondientes al rango de 775 - 925 nm, por lo que se separaron estas 40 bandas. Los resultados esperados para estos modelos son: exactitud, precision y recall superiores al 90 %, adicionalmente LOSS y MSE por debajo de 1 %.

Posterior a esto se hizo el preprocesamiento de las imágenes y se ingresaron al modelo, obteniendo los resultados consignados en las tablas 4-1, 4-2 y 4-3, cada una de estas tablas corresponde a un proceso de entrenamiento y validación diferente, donde se logra evidenciar que:

MODELO CNN 2D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DIA1	0,449	0,918	0,09	0,963	0,7174
DIA2	0,437	0,962	0,071	1	0,8813
DIA3	0,4	0,845	0,0713	0,8812	0,8042
DIA4	0,51	0,684	0,122	0,695	0,625
DIA5	1	0,13	1	0,018	1

Tabla 4-1: Métricas del modelo CNN2D, experimento 1, durante 5 días

MODELO CNN 2D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DIA1	0,188	1	0,023	1	0,9565
DIA2	0,219	0,962	0,029	0,99	0,9625
DIA3	0,1208	0,954	0,0219	0,953	0,933
DIA4	0,076	1	0,008	1	0,9643
DIA5	0,088	0,9943	0,0123	0,994	0,993

Tabla 4-2: Métricas del modelo CNN2D, experimento 2, durante 5 días

MODELO CNN 2D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DIA1	0,1074	1	0,011	1	0,9837
DIA2	0,1218	0,9624	0,016	0,962	0,962
DIA3	0,036	1	0,0049	1	1
DIA4	0,738	0,642	0,167	0,642	0,6429
DIA5	0,054	1	0,057	1	1

Tabla 4-3: Métricas del modelo CNN2D, experimento 3, durante 5 días

- Al realizar el proceso de entrenamiento y validación repetidas veces las métricas de desempeño mejoran, siendo 3 el número de veces suficiente para lograr unas métricas acordes a los resultados esperados.

- El día 1 después de la inoculación y al cabo de los 3 experimentos presenta en conjunto muy buenos resultados, ya que su precisión y exactitud son del 100 %, quiere decir que está clasificando correctamente todas las muestras, su función de pérdida es baja y el error cuadrático medio también, esto quiere decir que la certeza con la que clasifica es muy alta.

- El día 2 después de la inoculación, aunque la enfermedad está más desarrollada el proceso de clasificación al cabo de los 3 experimentos tuvo en general un desempeño más bajo que el día 1, los resultados bajan al rededor de un 5 % en precisión y efectividad, la función de pérdida LOSS y el MSE mantienen buenos resultados por debajo del 1 %, como se espera.
- El día 3 después de la inoculación vuelve a presentar unas buenas métricas ya que la función de pérdida LOSS y el MSE disminuyeron con respecto al día 1, los valores de precisión y exactitud siguen siendo del 100 %, en este caso, el valor de recall también es del 100 %.
- El día 4 después de la inoculación vuelve a tener una disminución de los valores de las métricas, sin embargo, esto se presenta únicamente en el tercer experimento ya que en el segundo experimento se tuvo un desempeño incluso mejor que en los días anteriores, esta disminución puede deberse al fenómeno de overfiting o sobre entrenamiento.
- El día 5 después de la inoculación tiene un desempeño mejor que el del día anterior en el último experimento, aunque no es el mejor de todos los días.

En general se puede observar que, en conjunto, el mejor desempeño se obtuvo en el día 3 después de la inoculación y en el tercer experimento, sin embargo lo que se espera es que la detección se haga lo mas temprano posible, por lo que los resultados del primer día pueden ser considerados como solución al problema propuesto.

4.2. Modelo CNN3D aplicado a las muestras por día

Para este modelo, el cual fue explicado en la Sección 3.6.2, los experimentos que se realizaron consisten en usar las imágenes hiperespectrales sin separarlas por bandas, ya que, como se explicó anteriormente, gracias a esta arquitectura es posible evaluar las imágenes tanto espacial como espectralmente, por lo que se evaluaron las imágenes completas con sus 160 bandas. Al igual que el modelo anterior se entrenó y evaluó 3 veces para poder comparar sus resultados y determinar cual tiene un mejor desempeño. Los resultados esperados son los mismos descritos para el modelo en 2 dimensiones. Los resultados obtenidos se consignaron en las tablas 4-4, 4-5 y 4-6, y se logró determinar que:

- Al realizar el proceso de entrenamiento y validación las mismas 3 veces que en el modelo anterior, las métricas de desempeño mejoran, sin embargo, se tienen unos resultados muy lejanos a los esperados.
- En general no existe un buen desempeño usando este modelo, los valores de la función de pérdida LOSS son muy altos, lo que quiere decir que la incertidumbre en las predicciones es alta, en ningún caso inferiores a 1, el valor de exactitud en el mejor de los

MODELO CNN 3D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DIA1	8,725	0,285	0,409	0,333	0,285
DIA2	16,279	0	0,664	0	0
DIA3	14,745	0,555	0,296	0,555	0,555
DIA4	9,827	0,285	0,476	0,285	0,285
DIA5	2,193	0,428	0,328	0,428	0,428

Tabla 4-4: Métricas del modelo CNN3D, experimento 1, durante 5 días

MODELO CNN 3D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DIA1	13,638	0,285	0,426	0,285	0,285
DIA2	8,77	0	0,602	0	0
DIA3	7,66	0,555	0,278	0,556	0,555
DIA4	1,368	0,714	0,163	0,714	0,714
DIA5	3,175	0,571	0,238	0,571	0,571

Tabla 4-5: Métricas del modelo CNN3D, experimento 2, durante 5 días

MODELO CNN 3D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DIA1	11,317	0,2857	0,3666	0,285	0,285
DIA2	6,42	0,166	0,507	0,167	0,166
DIA3	10,7	0,555	0,296	0,555	0,555
DIA4	2,65	0,714	0,173	0,714	0,714
DIA5	2,191	0,714	0,199	0,714	0,714

Tabla 4-6: Métricas del modelo CNN3D, experimento 3, durante 5 días

casos fue de 0,714 o 71,4 %, el MSE tuvo valores elevados nunca inferiores a 0,163, las métricas de PRECISION y RECALL al igual que la exactitud tuvieron valores bajos, lo cual muestra que el modelo no funciona bien.

En general se puede observar que esta arquitectura no funciona bien para el objetivo propuesto. Esto puede deberse a que para usar arquitecturas de deep learning, como en este caso las redes neuronales convolucionales, se hace necesario tener una gran cantidad de información o de imágenes. En este caso en particular, como se mencionó anteriormente, por día se tienen entre 20 y 25 imágenes hiperespectrales.

4.3. Modelo CNN2D aplicado a las muestras mezcladas

Teniendo en cuenta los resultados obtenidos por los experimentos anteriores, donde las métricas de desempeño del modelo CNN3D no son las mejores, y habiendo evidenciado que este error se debe a la poca cantidad de imágenes hiperespectrales, se decidió aumentar la cantidad de imágenes que iba a tener como entrada cada uno de los modelos. Por esto se combinaron las imágenes de los días 1 y 2 y posteriormente las imágenes de los días 3 y 4, para el modelo CNN2D, al igual que el primer experimento, se separaron las bandas de las imágenes hiperespectrales para ser analizadas por separado, por lo que la cantidad de información se duplicó. Los resultados obtenidos se consignaron en las tablas 4-8 y 4-9 y se logra evidenciar que:

MODELO CNN 2D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DÍAS 1 Y 2	0,326	0,895	0,06	0,924	0,819
DÍAS 3 Y 4	0,224	0,975	0,032	0,987	0,938

Tabla 4-7: Métricas del modelo CNN2D, experimento 1, mezclando los datos

MODELO CNN 2D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DÍAS 1 Y 2	0,123	1	0,016	1	0,967
DÍAS 3 Y 4	0,102	1	0,011	1	0,9804

Tabla 4-8: Métricas del modelo CNN2D, experimento 2, mezclando los datos

MODELO CNN 2D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DÍAS 1 Y 2	0,064	1	0,0063	1	0,997
DÍAS 3 Y 4	0,065	1	0,0068	1	0,9804

Tabla 4-9: Métricas del modelo CNN2D, experimento 3, mezclando los datos

- El modelo presenta unas buenas métricas, desde el segundo experimento la exactitud y la precisión son del 100 %, lo cual quiere decir que de manera más eficiente que el primer experimento se logra que las predicciones sean siempre acertadas. Mientras que las métricas de función de pérdida y MSE son bajas y disminuyen hasta el tercer experimento, es decir, la certeza con la que hacen las predicciones es alta.

- Al tener a la entrada del modelo más datos, este tiene un mejor desempeño en un tiempo más corto, y en menor cantidad de experimentos.

En general este experimento muestra como a mayor cantidad de datos es posible obtener una mejor clasificación en una menor cantidad de entrenamientos, sin embargo, las métricas no son las que mejor comportamiento presentan. Una posible solución a la falta de datos que le ingresan a los modelos es usar las técnicas de data augmentation, sin embargo, por falta de capacidad de procesamiento y memoria en el equipo usado no es posible trabajar con más imágenes.

4.4. Modelo CNN3D aplicado a las muestras mezcladas

Con el fin comparar los resultados obtenidos en la sección anterior se usa el mismo conjunto de datos ya mezclados con el modelo CNN3D, para poder confirmar lo que se evidenció en la Sección 4.2, donde a falta de datos de entrada el modelo no se tiene un buen desempeño.

Nuevamente se usaron las imágenes hiperespectrales completas para lograr hacer un análisis espacial y espectral mediante el modelo en tres dimensiones. Los resultados obtenidos fueron consignados en las tablas 4-10, 4-11 y 4-12. De estos resultados se puede observar que:

- Aun haciendo la mezcla de los datos, obteniendo el doble de estos para cada experimento, las metricas de desempeño no presentan una mejoría.
- Si bien el desempeño entre un grupo de días y el otro es notorio y significativo, los resultados entre experimentos del mismo modelo y el mismo grupo de días no son favorables, ya que las metricas que se obtuvieron no son las esperadas.

MODELO CNN 3D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DÍAS 1 Y 2	14,624	0,153	0,534	0,153	0,153
DÍAS 3 Y 4	6,549	0,25	0,464	0,25	0,25

Tabla 4-10: Métricas del modelo CNN3D, experimento 1, mezclando los datos

En general, se puede observar que, a pesar que existan más muestras, no es posible que el modelo tenga mejor desempeño ya que aunque son más muestras, sigue siendo un número limitado de estas. Sin embargo, lo que se esperaría es poder aumentar cada vez más esta cantidad de muestras para lograr un mejor desempeño de este modelo, pero en este caso no es posible ya que las características del equipo que se usó en los experimentos son limitadas y aumentar aun más los datos requiere espacios en memoria con los que no se cuentan.

MODELO CNN 3D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DÍAS 1 Y 2	10,53	0,23	0,495	0,23	0,23
DÍAS 3 Y 4	1,347	0,6875	0,2	0,687	0,687

Tabla 4-11: Métricas del modelo CNN3D, experimento 2, mezclando los datos

MODELO CNN 3D					
	LOSS	ACCURACY	MSE	PRECISION	RECALL
DÍAS 1 Y 2	11,32	0,384	0,383	0,384	0,384
DÍAS 3 Y 4	4,37	0,3125	0,4333	0,3125	0,3125

Tabla 4-12: Métricas del modelo CNN3D, experimento 3, mezclando los datos

5 Conclusiones

Para poder obtener un buen desempeño y unos buenos resultados dentro del campo de deep learning, se hace necesario que los datos tengan un procesamiento previo ya que estos deben tener siempre las mismas características para que las redes neuronales puedan procesar características como dimensiones espaciales, cantidad de canales, entre otras

Gracias a la revisión de la literatura que existe de modelos de redes neuronales para la clasificacion de imágenes, se logró evidenciar cuáles redes son las que más se adaptan a la necesidad y al problema expuesto en este trabajo. Aunque existen diversas variaciones de redes neuronales convolucionales, las dos que se implementaron en este trabajo fueron las que más se adaptaron. Por su lado, el modelo de red neuronal convolucional en dos dimensiones o CNN2D proporciona una clasificacion confiable, esto se puede afirmar a partir de las métricas que fueron presentadas, donde se muestra que este modelo llega a predecir el 100 % de los datos con unos porcentajes de error pequeños.

Por otro lado, el modelo de red convolucional en tres dimensiones no logró proporcionar los resultados esperados ya que, en teoría, por el hecho de analizar las imágenes de forma espacial y espectral al mismo tiempo debería poder hacer una mejor clasificación que el modelo en dos dimensiones. Sin embargo, la limitación en la cantidad de datos, y adicional las limitaciones en el equipo usado, hicieron que no fuera posible aumentar la cantidad de datos a usar. Teniendo en cuenta lo anteriormente mencionado y como es esperado en todo desarrollo las técnicas y modelos usados en este trabajo, están sujetos a una validación mas extensa, con el fin de evaluar y mejorar los resultados obtenidos.

Como recomendaciones y trabajo futuro se plantea el uso de estos y otros modelos de aprendizaje profundo, haciendo uso de herramientas computacionales más robustas ya que la gran limitante, como se ha mencionado, es la capacidad de memoria y procesamiento. Aumentar dichas características permitiría pues hacer modelos más robustos y hacer uso de una mayor cantidad de imágenes para el procesamiento de las mismas. Debido a la poca cantidad de imágenes que pueden ser adquiridas, se recomienda hacer uso de técnicas de data augmentation, estas técnicas buscan hacer modificaciones morfológicas a las imágenes, sin afectar su composición.

Bibliografía

[1] https://docs.opencv.org/4.x/df/d9d/tutorial$_p y_c olorspaces.html$.

[2] https://www.deccoiberica.es/que-es-la-antracnosis-y-como-afecta-a-los-cultivos/, 2018.

[3] https://cs231n.github.io/convolutional-networks/, 2021.

[4] E. Alegre, G. Pajares, and A. D. L. Escalera. *Conceptos y Métodos en visión por computador*. 2016.

[5] L. Annala, M. A. Eskelinen, J. Hämäläinen, A. Riihinen, and I. Pölönen. Practical approach for hyperspectral image processing in python. volume 42, pages 45–52. International Society for Photogrammetry and Remote Sensing, 4 2018. doi: 10.5194/isprs-archives-XLII-3-45-2018.

[6] A. Bannari, A. Pacheco, K. Staenz, H. McNairn, and K. Omari. Estimating and mapping crop residues cover on agricultural lands using hyperspectral and ikonos data. *Remote Sensing of Environment*, 104:447–459, 10 2006. ISSN 00344257. doi: 10.1016/j.rse.2006.05.018.

[7] R. Bongiovanni, E. C. Mantovani, S. Best, and Álvaro Roel. Agricultura de precisiÓn: Integrando conocimientos para una agricultura moderna y sustentable, 2006.

[8] G. Bradski and A. Kaehler. *Learning OpenCV*. 2008.

[9] J. Carrillo, J. Armando, S. Estrada, D. Rangel, A. Barajas, I. Zequera, R. Molar, Z. D. la Garza Ruiz, M. Vera, and E. Fentanes. Control biológico de antracnosis [colletotrichum gloeosporioides(penz.) penz. y sacc.] y su efecto en la calidad poscosecha delmango (mangifera indica l.) en sinaloa, méxico. *Revista Mexicana de Fitopatología*, page 10, 2005. ISSN 0185-3309. URL `http://www.redalyc.org/articulo.oa?id=61223104`.

[10] G. Casal, J. Dominguez, N. Sanchez, and J. Freire. Utilización del sensor de imagen airborne hyperspectralscanner (ahs) para la cartografía de bosques de sargassum muticumen la ría de vigo (galicia). *Revista de teledetección*, page 6, 2008.

[11] J. D. P. Díaz. Aplicación de las técnicas de aprendizaje automático para la detección temprana de antracnosis en hojas de guanábana, 2020.

[12] A. Fazari, O. J. Pellicer-Valero, J. Gómez-Sanchıs, B. Bernardi, S. Cubero, S. Benalia, G. Zimbalatti, and J. Blasco. Application of deep convolutional neural networks for the detection of anthracnose in olives using vis/nir hyperspectral images. *Computers and Electronics in Agriculture*, 187, 8 2021. ISSN 01681699. doi: 10.1016/j.compag.2021.106252.

[13] E. García and F. Flego. Agricultura de precisión.

[14] R. C. Gonzalez and R. E. R. E. Woods. *Digital image processing*. 2001. ISBN 0201180758.

[15] M. L. Guillen-Climent, H. Mas, A. Fernández-Landa, N. Algeet-Abarquero, and J. L. Tomé. Using hipersepctral images for decay detection in pinus halepensis (mill.) in the mediterranean forest. *Revista de Teledeteccion*, 2020:59–69, 2020. ISSN 19888740. doi: 10.4995/raet.2020.13289.

[16] R. L. Lawrence, S. D. Wood, and R. L. Sheley. Mapping invasive plants using hyperspectral imagery and breiman cutler classifications (randomforest). *Remote Sensing of Environment*, 100:356–362, 2 2006. ISSN 00344257. doi: 10.1016/j.rse.2005.10.014.

[17] B. A. Lemus Soriano and D. A. Pérez Aguilar. Manejo de la antracnosis del aguacate con biofungicidas. page 5, 2017.

[18] J. L. Perez. Imágenes hiperespectralesy sus aplicacionesen estudios de suelos, cultivosy bosques, en la era de la cuarta revolución industrial. *REVISTA UD Y LA GEOMÁTICA*, pages 40–70, 2021. ISSN 2344-8407. URL `http://revistas.udistrital.edu.co/ojs/index.php/UDGeo/index`.

[19] S. Prasad and J. Chanussot. *Advances in Computer Vision and Pattern Recognition Hyperspectral Image Analysis Advances in Machine Learning and Signal Processing*. 2020. URL `http://www.springer.com/series/4205`.

[20] L. Ramirez. Desarrollo de un sistema para la identificación temprana de la antracnosis en frutos de mangobasado en visión de máquina, 2021.

[21] A. Roman-Gonzalez and N. I. Vargas-Cuentas. Análisis de imágenes hiperespectrales, 2015. URL `https://www.researchgate.net/publication/259840849`.

[22] J. Torres. *Python Deep Learning*. Primera edition, 2020.

[23] L. Urdaneta, M. Sanabria, D. Rodríguez, and M. P. de Camacaro. Antracnosis causada por colletotrichum acutatum simmonds en frutos de fresa en los estados lara y trujillo, venezuela, 2013.

[24] L. A. Valdés, D. L. C. Consuegra, A. Gómez, M. E. Carballo, M. Capote, I. González, J. M. Alvarez, and W. Rohde. CaracterizaciÓn morfolÓgica, cultural y patogÉnica de aislados de colletotrichum sp. produciendo antracnosis en mango (mangifera indica l.). *La Granja*, 26: 38, 9 2017. ISSN 1390-3799. doi: 10.17163/lgr.n26.2017.04.

[25] S. Yu, S. Jia, and C. Xu. Convolutional neural networks for hyperspectral image classification. *Neurocomputing*, 219:88–98, 1 2017. ISSN 18728286. doi: 10.1016/j.neucom.2016.09.010.