

Universidad Autónoma del Estado de México
Centro Universitario UAEM Texcoco
Maestría en Ciencias en Computación



Identificación y cuantificación de daño de la enfermedad de la Roya del café en hojas

Tesis que presenta
Oscar Eder Flores Colorado

Para obtener el Grado de
Maestro en Ciencias en Computación

Asesor de tesis
Dr. Jair Cervantes Canales

Texcoco, Estado de México, Noviembre de 2021

Abstract

Coffee is a crop with a tradition of many years in Mexico, thanks to this, the country is among the main producers and exporters of coffee. However, in recent years coffee crops have been affected by many diseases and pests which has severely affected the production of this bean, bringing with it multiple consequences such as economic crisis and emigration of families that depend on this crop in Mexico. Pattern recognition is a technological tool that can help agronomists and plant pathologists in the early detection of diseases and pests in coffee. In this study, apparently healthy leaves and four diseases of economic importance in Mexico are identified, which are coffee rust, coffee leaf miner, phoma leaf spot and cercospora leaf spot. Different segmentation algorithms were tested, including Otsu, Principal Component Analysis (PCA) and Global Frontier Method. Chromatic, geometric and textural characteristics were extracted from the leaves of the coffee plant. Four classification algorithms were tested, including Support Vector Machine with SMO (Sequential Minimal Optimization), Random Forest, Naive Bayes and Artificial Neural Networks Backpropagation. The best precision obtained was 83 % with PCA segmentation and Artificial Neural Networks Backpropagation algorithm.

Resumen

El café es un cultivo con una tradición de muchos años en México lo que lo sitúa dentro de los principales productores y exportadores de café. Sin embargo, en los últimos años los cultivos de café han sido afectados por muchas enfermedades y plagas, que han causado severas afectaciones en la producción de este grano, trayendo consigo múltiples consecuencias como crisis económicas y emigración de familias que dependen de este cultivo en las regiones productoras de este cultivo en México. El reconocimiento de patrones es una herramienta tecnológica que puede ayudar a agrónomos y fitopatólogos en la pronta detección de enfermedades y plagas en el café. En este trabajo se identifican las hojas aparentemente sanas y cuatro enfermedades de importancia económica en México, las cuales son la Roya del café, el minador de la hoja, phoma quema y la mancha de hierro o cercospora. Se probaron diferentes algoritmos de segmentación entre los cuales se encuentran Otsu, *Principal Component Analysis*(PCA) y Método de Frontera Global. Se extrajeron características cromáticas, geométricas y texturales de las hojas de la planta de café. Se probaron cuatro algoritmos de clasificación entre los que se encuentran *Support Vector Machine* con SMO (Sequential Minimal Optimization), *Random Forest*, *Naive Bayes* y Redes Neuronales Artificiales *Backpropagation*. La mejor precisión que se obtuvo fue de 83 % con segmentación PCA y usando el algoritmo de Redes Neuronales Artificiales *Backpropagation*.

Índice general

Abstract	VII
Resumen	IX
Índice general	XI
Índice de figuras	XVII
Índice de cuadros	XIX
1. Introducción	1
1.1. Planteamiento del problema	3
1.2. Justificación	4
1.3. Objetivos	4
1.4. Hipótesis	5
1.5. Metodología	5
1.6. Estado del arte	6
2. Preliminares	9
2.1. La planta del café	9
2.1.1. La estructura de la planta del café	9
La raíces	9
El tallo	10
Las ramas	10
2.1.2. La hoja del café	10
2.1.3. El fruto del café	11
2.1.4. Variedades de la planta del café en México	11
<i>Coffea Arabica</i>	11
<i>Coffea canephora</i>	12
2.2. Aprendizaje Supervisado	12
2.3. Aprendizaje No Supervisado	13

2.4.	Reconocimiento de patrones	14
2.5.	Pre-procesamiento	15
2.5.1.	Escalado de imagen	16
2.5.2.	Espacios de color	16
	RGB	16
2.5.3.	Histograma de una imagen	17
	Estructura del Histograma	18
	Distribución homogénea	18
	Distribución con zonas oscuras	19
	Distribución con sobreexposición	19
	Distribución con contraluz	20
	Ecuilización del Histograma	20
2.5.4.	Binarización de imágenes	20
2.5.5.	Detección de bordes	21
2.6.	Segmentación	22
2.6.1.	Segmentación por umbralización	22
	Método Otsu para umbral	23
2.6.2.	Segmentación por clustering	25
	Distancia Cuadrada Euclidiana	25
	Inercia de los clústeres	25
2.6.3.	PCA (Principal Component Analysis)	26
2.7.	Extracción de Características	27
2.7.1.	Características Geométricas	27
	Área	27
	Perímetro	27
	Altura y Ancho	27
	Excentricidad	28
	Inercia	28
	Rectangularidad	28
	Momentos de Hu	28
2.7.2.	Características Cromáticas	30
	Media Aritmética	30
	Varianza	31
	Curtosis	31
	Leptocúrtica	31

	Mesocúrtica	31
	Platicúrtica	32
	Cálculo de la curtosis	32
	Histograma de gradientes orientados HOG	32
	Filtros de Gabor	34
	Patrones Binarios Locales LBP	34
2.7.3.	Características Texturales	36
	Características Texturales de Haralick	36
	Matriz de Coocurrencia GLCM(Grey Level Co-occurrence Matrix)	37
	Segundo Momento Angular	37
	Contraste	38
	Correlación	38
	Suma de cuadrados - varianza	38
	Momento de Diferencia Inversa(Homogeneidad)	38
	Suma promedio	38
	Suma de Varianza	39
	Suma de Entropía	39
	Entropía	39
	Diferencia de Varianza	39
	Diferencia de Entropía	39
	Medidas de información de Correlación 1	40
	Medidas de información de Correlación 2	40
	Máximo Coeficiente de Correlación	40
2.8.	Clasificación	40
2.8.1.	Redes Bayesianas	40
	Inferencia	41
2.8.2.	Naive Bayes	42
	Estructura del sistema de aprendizaje	42
2.8.3.	Support Vector Machines(SVM)	42
	Caso linealmente separable	43
2.8.4.	Redes Neuronales Artificiales	45
	Perceptrón	46
	Funciones de activación	47
	Función umbral(Escalón)	47
	Función Sigmoide(Logística)	47

Función Softmax	48
3. Metodología	49
3.1. Conjunto de datos utilizado	50
3.2. Pre-procesamiento	52
3.3. Segmentación	52
3.4. Extracción de características	53
3.4.1. Características Geométricas	54
3.4.2. Características Cromáticas	54
3.4.3. Características Texturales	54
3.5. Clasificación	55
3.5.1. <i>Support Vector Machine</i>	55
3.5.2. Redes Neuronales Artificiales <i>Backpropagation</i>	55
3.5.3. <i>Random Forest</i>	55
3.5.4. <i>Naive Bayes</i>	56
4. Resultados experimentales	57
4.1. Variables Estadísticas del clasificador	57
4.1.1. Verdadero positivo	57
4.1.2. Verdadero negativo	57
4.1.3. Falso negativo	57
4.1.4. Falso positivo	58
4.1.5. Sensibilidad	58
4.1.6. Especificidad	59
4.1.7. Tasa de falsos positivos(False Positives Rate en inglés)	59
4.1.8. Precisión	59
4.1.9. Valor F	59
4.1.10. Curva ROC	59
4.2. Desempeño de clasificadores utilizando características cromáticas	60
4.2.1. Matrices de confusión - características cromáticas	61
4.3. Desempeño de clasificadores utilizando características geométricas	63
4.3.1. Matrices de confusión - características geométricas	64
4.4. Desempeño de clasificadores utilizando características texturales	65
4.4.1. Matrices de confusión - características texturales	66
4.4.2. Diagramas de caja de la sensibilidad de los clasificadores por tipo de características	67

4.5. Desempeño de clasificadores utilizando características texturales y cromáticas	67
4.5.1. Matrices de confusión - características texturales y cromáticas	68
4.6. Desempeño de clasificadores utilizando características texturales y geométricas	69
4.6.1. Matrices de confusión - características texturales y geométricas	70
4.7. Desempeño de clasificadores utilizando características cromáticas y geométricas	70
4.7.1. Matrices de confusión - características cromáticas y geométricas	71
4.8. Desempeño de clasificadores utilizando características cromáticas, geométricas y texturales	72
4.8.1. Matrices de confusión - características cromáticas, geométricas y texturales	73
4.8.2. Diagramas de caja de la precisión de los clasificadores	74
5. Conclusiones	75

Índice de figuras

1.1. Planta sana y Planta dañada por la roya del café	4
2.1. Estructura de la planta del café	10
2.2. Hoja de la planta	11
2.3. Clasificación de dos clases con aprendizaje supervisado	13
2.4. Agrupamiento de datos de acuerdo a sus características	14
2.5. Espacio de color RGB	17
2.6. Estructura del Histograma	18
2.7. Distribución homogénea	18
2.8. Distribución con zonas oscuras	19
2.9. Distribución con sobreexposición	19
2.10. Distribución con contraluz	20
2.11. Binarización de una imagen	21
2.12. Detección de bordes por Canny	22
2.13. Los momentos de Hu ayudan a identificar o clasificar una hoja sin im- portar su tamaño, traslación o inversión	29
2.14. Distribución Leptocúrtica	31
2.15. Distribución Mesocúrtica	31
2.16. Distribución Platicúrtica	32
2.17. Histograma de gradientes orientados	33
2.18. Ejemplo del calculo del valor LBP del pixel central 123	35
2.19. Comparación de características de una hoja, una hoja rotada y un paisaje aleatorio	36
2.20. Matriz de Coocurrencia	37
2.21. Ilustración de una SVM para clasificar 2 clases	43
2.22. Utilización de Kernels	45
2.23. Composición de una neurona artificial	46
2.24. Red neuronal con capas de entrada, capas ocultas y capas de salida	47
3.1. Metodología propuesta	49

3.2. Hojas del data set	51
3.3. Hojas del data set	52
3.4. Segmentación por tres métodos	53
4.1. Clasificación de hojas de café	58
4.2. Curva ROC y línea de no-discriminación	60
4.3. Matriz de confusión características cromáticas	63
4.4. Matriz de confusión características geométricas	65
4.5. Matriz de confusión características texturales	66
4.6. Matriz de confusión características cromáticas y texturales	68
4.7. Matriz de confusión características geométricas y texturales	70
4.8. Matriz de confusión características geométricas y texturales	71
4.9. Matriz de confusión características geométricas y texturales	73

Índice de cuadros

Capítulo 1

Introducción

El reconocimiento de patrones en imágenes digitales es una tecnología que se ha presentado como una herramienta interesante para apoyar y ayudar en áreas de cualquier índole. En el ámbito de la salud por ejemplo, se han desarrollado herramientas para la detección de tumores cancerígenos en el cuerpo (Wang, 2019) y la detección de cancer en la piel (Jain, 2015); otra área donde se ha avanzado bastante es en el sector automotriz donde la conducción autónoma que con ayuda de varios sensores y cámaras pueden trazar un recorrido con poca intervención humana (Ma, 2020); El reconocimiento de texto (*Optical Character Recognition*, en inglés) es otra área donde se ha progresado mucho ya que actualmente existen un amplio abanico de herramientas para la identificación de texto por medio de una cámara. Pueden incluso reconocer texto escrito a mano (Sharma, 2013). En resumen, el reconocimiento de patrones tiene un potencial enorme en un sin número de ámbitos de la vida cotidiana.

Los sistemas de reconocimiento de patrones intentan emular lo que el ojo humano interpreta y relaciona cuando mira colores, patrones y texturas en un objeto de su interés. Se debe tener en cuenta que al desarrollar un sistema de reconocimiento de imágenes se busca que este extraiga información útil de imágenes digitales, procese sus características y ofrezca una clasificación de acuerdo con el objetivo con el que fue desarrollado.

El café es un cultivo que forma parte de una cadena productiva muy importante para la actividad agropecuaria en México. Se estima que hay alrededor de 500 mil productores que se dedican a la producción de este grano en 15 estados del país con un aproximado de 675 mil ha de superficie plantada (Flores, 2015). Teniendo a Chiapas, Veracruz y Puebla como los principales estados productores. México es conocido por producir café de excelente calidad porque es beneficiado por una gran diversidad de climas y superficies a una buena altitud que posibilitan la instalación de cultivos de la

variedad de mejor calidad (*Coffea Arabica*) (SADER, 2018). No obstante, el café es afectado de manera permanente a lo largo de todo el año por plagas y enfermedades lo cual merma la producción y provocan lamentables pérdidas económicas y acarrear consecuencias como la emigración, pauperización de poblaciones y el abandono definitivo de la cafeticultura (Henderson, 2019).

Los sistemas de reconocimiento de patrones se pueden aplicar en los cultivos del café para ayudar a agrónomos y fitopatólogos a identificar plagas y enfermedades de manera más precisa dado que la visión humana incide un margen de error intrínseco dependiendo del nivel de experiencia de la persona. El propósito de realizar este proceso de manera automatizada y precisa es atender el problema rápidamente y así evitar la dispersión de estos problemas a predios o parcelas contiguas.

En este trabajo de investigación se desarrolló una herramienta de reconocimiento de patrones que clasifica la hoja sana y cuatro plagas y enfermedades de importancia económica presentes en la planta del café especie (*Coffea Arabica*). Las cuatro plagas y enfermedades son el minador del café, la roya del café, la cercospora o mancha de hierro y la *phoma*, *Quema*. Para la realización de este trabajo se utilizaron 739 imágenes de hojas de café especie *Coffea Arabica* de un conjunto de datos obtenido en Mendeley Data llamado BRACOL - «A Brazilian Arabica Coffee Leaf images dataset to identification and quantification of coffee diseases and pests» (Krohling, 2019). Se extrajeron características cromáticas, geométricas y texturales y haciendo combinaciones entre ellas se logró una precisión máxima de 83 % con redes neuronales artificiales. En adición a ese algoritmo de clasificación, se utilizó una Máquina de Vectores de Soporte (Support Vector Machine, en inglés) con el algoritmo SMO (Sequential Minimal Optimization), el algoritmo de clasificación Bosque aleatorio (Random Forest, en inglés) y el clasificador bayesiano ingenuo (Naive Bayes, en inglés).

Esta tesis se divide en cinco capítulos. En el primer capítulo se presenta la introducción donde se habla de que trata esta investigación, además se puede encontrar el planteamiento del problema, la justificación del problema, la metodología de la investigación, entre otros. En el estado del arte se discuten diversos trabajos relacionados al reconocimiento de patrones en plantas para la identificación de plagas y enfermedades. En el segundo capítulo, se presentan los preliminares sobre el pre-procesamiento, la segmentación, la extracción de características y la clasificación de datos. En el tercer capítulo, se aborda los pasos propuestos que se siguieron para desarrollar la herramienta de reconocimiento de patrones, esto constituye la metodología que se empleó para esta investigación. En el cuarto capítulo se muestran los resultados y métricas de

rendimiento que se obtuvieron ordenados por tipo de características (cromáticas, geométricas y texturales). Finalmente, en el capítulo cinco se discuten los resultados, los apartados que se deben de mejorar de la herramienta y se sientan las bases para un trabajo futuro.

1.1. Planteamiento del problema

Las enfermedades del café impactan negativamente los cultivos de esta planta, ya que es el sustento de muchas familias en condiciones de pobreza en el sur y sureste de México (Läderach, 2011), lo que provoca que la producción de sus cultivos se vea significativamente reducida y por ende, también su ingreso económico.

Identificar las enfermedades del café resulta a veces complicado por el hecho de que algunas enfermedades se manifiestan de manera muy similar en la hoja. Para diagnosticar una enfermedad específica, a menudo, se necesita la intervención de un agrónomo experto para que analice la planta y sus hojas y en base a esto identifique particularidades de una enfermedad con respecto a otra que provoca síntomas parecidos. Aunado a eso, hay otros factores que agregan complejidad a la identificación de plagas y enfermedades del café entre los que se encuentran la nutrición de la planta; el estrés de la planta causado por la aplicación de pesticidas y herbicidas; y el cambio climático (Anderson, 2008).

Entre las enfermedades y plagas de la planta del café que destacan por su capacidad para reducir significativamente la producción se encuentra la roya del café. Esta enfermedad hace que la planta pierda la mayoría de sus hojas en un proceso llamado defoliación. En pocas palabras, si no hay hojas, no hay fotosíntesis, y esto hace que la planta no desarrolle frutos o estos no alcancen su maduración. En la FIGURA 1.1 se muestran dos imágenes; a la izquierda una planta sana y a la derecha una planta con muy pocas hojas que ha sido dañada por la roya del café.



FIGURA 1.1: Planta sana y Planta dañada por la roya del café

1.2. Justificación

Actualmente, el café representa el 0.66 % del PIB agrícola del país (SADER, 2018). Es una cadena productiva muy importante para estados del sur y sureste de México. Los cultivos de café mexicano son afectados por muchas enfermedades cada año. La venta y producción de café que se elabora en los diferentes estados productores representa para muchas familias (gran parte de ellas indígenas) su único salario (Berlanga, 2011). Los cultivos de café sufrieron una crisis de producción entre el año 2012 y 2013 a causa de la roya del café que devastó una gran cantidad de superficie de dicho cultivo (Escamilla, 2016). Es por eso que este trabajo busca coadyuvar en la identificación de enfermedades y plagas del café con una herramienta de reconocimiento de patrones de fácil acceso a agrónomos y fitopatólogos para detectar tempranamente las enfermedades que afectan al café y, por consiguiente, aplicar las medidas fitosanitarias pertinentes para prevenir su desarrollo y dispersión.

1.3. Objetivos

Desarrollar una herramienta de reconocimiento de patrones en imágenes digitales para la identificación de la planta sana y cuatro enfermedades y plagas de importancia

económica de la planta del café especie *Coffea Arabica*.

1. Obtener o crear un conjunto de datos y con una buena representatividad para cada una de las enfermedades y plagas que se buscan clasificar.
2. Evaluar tres diferentes técnicas de segmentación: Otsu, *Principal Component Analysis* y método de frontera global para identificar cuál es la técnica que mejor resultado genera.
3. Realizar la extracción de características cromáticas, texturales y geométricas y combinaciones entre ellas.
4. Analizar cuatro diferentes algoritmos de clasificación (*Naive Bayes*, *SVM con SMO*, *Random Forest* y redes neuronales artificiales) y determinar cuál es el algoritmo que obtiene mejores métricas de rendimiento.
5. Analizar y determinar bajo qué condiciones el sistema alcanza su mayor rendimiento.
6. Constituir la metodología para el desarrollo del estudio.
7. Determinar parámetros a tomar para cada una de las tareas (Pre-procesamiento, segmentación, extracción de características y clasificación)

1.4. Hipótesis

La utilización de una herramienta de reconocimiento de patrones en imágenes digitales puede ayudar a los productores y agricultores del cultivo de café a identificar enfermedades en la planta y, en consecuencia, aplicar medidas fitosanitarias preventivas para evitar la dispersión de estas.

1.5. Metodología

El diseño de sistemas de reconocimiento de objetos requiere diversos pasos y procesos. La fiabilidad y la eficiencia de un sistema de reconocimiento de objetos dependen directamente de cada uno de los pasos llevados a cabo, desde las técnicas de pre-procesamiento utilizadas, los métodos de segmentación, las características extraídas y hasta los clasificadores utilizados.

A continuación los pasos que se llevarán a cabo para la investigación:

1. Obtener un conjunto de datos representativo cada una de las clases
2. Utilizar diferentes técnicas de segmentación y elegir la que mejor resultados genere
3. Obtener las características texturales, cromáticas y geométricas de cada imagen
4. Probar diferentes algoritmos de clasificación y determinar cuál es el que mejor resultado entrega
5. Ajustar y definir los parámetros óptimos de funcionamiento de los algoritmos de clasificación
6. Comparar los diferentes algoritmos de clasificación

1.6. Estado del arte

El procesamiento de imágenes digitales está siendo cada vez más utilizado en la agricultura para la identificación de plagas y enfermedades de diferentes cultivos. Entre los principales usos están la identificación de diferentes especies y tipos de hojas de plantas (Sampallo, 2003); la clasificación de diferentes enfermedades en diferentes cultivos (Barbedo, 2016) y la estimación de la severidad o cuantificación del daño provocado por enfermedades en hojas y frutos (Manso, 2019).

Dentro de los algoritmos de clasificación de enfermedades en plantas que se han utilizado ampliamente se encuentran las máquinas de vectores de soporte (SVM, por sus siglas en inglés), Redes Neuronales Artificiales (ANN, por sus siglas en inglés), clasificadores bayesianos, árboles de decisión (Decision Trees en inglés) y clasificadores del vecino más cercano (KNN, por sus siglas en inglés).

Entre los autores que han trabajado con el algoritmo de Máquina de Vectores de Soporte (SVM) se encuentran Qin et al. (2016), quienes en busca de diagnosticar y clasificar cuatro enfermedades de la alfalfa, extrajeron 129 características texturales, cromáticas y geométricas. Compararon los algoritmos SVM, KNN y Bosque aleatorio (Random Forest en inglés). El que mejor resultados obtuvo fue el SVM teniendo 94.74 % de precisión. (Bhange, 2015) Estos autores experimentaron con cámaras de 10, 5 y 3 megapíxeles para clasificar si una granada está infectada o no subiendo imágenes a un sistema web para extraer características, como el color, morfología y el vector de coherencia de color obteniendo precisiones con SVM del 85 % para imágenes de 10 megapíxeles, 82 % para

las de 5 megapíxeles y 79 % para las de 3 megapíxeles. (Gavhale, 2014) Trabajaron en la identificación de enfermedades de las plantas de los cítricos que incluyen toronja, limón, lima y naranjas atacadas por el cancro y antracnosis. El modelo propuesto incluye como pre-procesamiento la conversión de RGB a diferentes espacios de color; la segmentación se hizo con agrupamiento por *K-means*; se extrajeron características texturales utilizando el método de la matriz de co-ocurrencia del nivel de gris (*GLCM*, por sus siglas en inglés) y también se extrajeron características de color por el valor promedio. Los resultados muestran una tasa del 95 % de precisión con *SVM*. Pujari (2016) clasificaron diversos cultivos como trigo, girasol, uva, maíz, pepino, algodón, col y tomate para determinar si están infectados por algún hongo o bacteria o estaban siendo atacados por nematodos. Aplicaron una reducción de dimensionalidad y descubrieron que de 24 solo 8 características fueron significativas para la clasificación de enfermedades con *SVM* y *ANN* obteniendo 92.17 % y 87.48 % respectivamente.

Las redes neuronales no son algo nuevo, sin embargo, su amplia utilización se dio en años recientes con la mejora en la capacidad de procesamiento de los recursos computacionales. Un trabajo interesante reciente sobre redes neuronales es el de Mohanty (2016), realizó diversos experimentos con 14 cultivos y 26 enfermedades usando redes neuronales convolucionales profundas. Notaron que para el entrenamiento de este tipo de redes se necesita mucho trabajo computacional pero que al final vale la pena ya que se obtuvo una precisión máxima de 99.35 % en la clasificación de manera muy rápida. Barbedo (2019) trabajó con redes neuronales convolucionales para varios cultivos. En su trabajo logró clasificar los cultivos sanos con 89 % de exactitud, un poco enfermos con 31 %, enfermos moderadamente 87 % y enfermos graves con un 94 %. Afirma que una de las limitaciones que impiden que aparezcan sistemas integrales para la clasificación de enfermedades en cultivos es la cantidad y representatividad del conjunto de datos. Dhakate (2015) Crearon un sistema para reconocer y clasificar cuatro enfermedades de la granada. Se utilizó segmentación por *K-means clustering*. Las características texturales fueron extraídas por medio de la matriz de co-ocurrencia del nivel de gris. Se usó una red neuronal retro-propagada y obtuvieron resultados muy cercanos a 90 %. Por su parte, Ramakrishnan (2015) segmentaron y obtuvieron características cromáticas para después clasificarlos con una red neuronal retro-propagada para identificar cuatro enfermedades de la planta del cacahuete. En los resultados de su experimento encontraron que de 100 imágenes de hojas de plantas enfermas se obtuvo un 97 % de precisión para cuatro enfermedades.

Capítulo 2

Preliminares

En este capítulo se presentan aspectos teóricos básicos sobre la planta del café con el objetivo de entender mejor los problemas que rodean a la planta del café. Además, se describe más a detalle sobre el pre-procesamiento, la segmentación de imágenes, extracción de características y los algoritmos de clasificación.

2.1. La planta del café

La planta del café proviene de Africa, específicamente de la región donde hoy se encuentra Etiopía y su introducción en México fue en 1790 (Medina et al., 2016). Su descubrimiento fue hecho por un monje llamado Kaldhi que observó como sus cabras se mostraban muy inquietas al comer las bayas de este arbusto de hojas perennes. La planta del café tiene diferentes estados fenológicos durante el año como la floración, el amarre del fruto, la fructificación y formación del café cereza para su cosecha. Los productores y agricultores llaman 'beneficio' a un conjunto de etapas donde los granos del café son despulpados, secados, tostados, entre otros procesos. Para culminar, los granos de café son empacados para su distribución y venta.

2.1.1. La estructura de la planta del café

La planta del café está constituida por las raíces, el tallo, las ramas, las hojas, las flores y los frutos.

La raíces

Se encuentran debajo de la tierra y sirven para proveer agua y nutrientes a las plantas. Las raíces de la planta del café pueden alcanzar más de cincuenta centímetros si la planta tiene más de 5 años. Hay una raíz principal de las cuales salen dos tipos de

raíces: las fuertes y las raicillas. Las primeras sirven como anclaje para dar estabilidad a la planta. La longitud de las raíces es normalmente la misma que la longitud del largo de las ramas (Monroig, 2010)

El tallo

El tallo de la planta del café sirve para sostener prácticamente toda la estructura superficial de la planta. El tallo se encarga de distribuir los nutrientes a las ramas que, a su vez, la distribuyen a las hojas, flores, y frutos.

Las ramas

Las ramas de la planta del café son las encargadas de sostener las hojas, flores y frutos. Las ramas tienen su origen en el tallo. A medida que el tallo crece, también lo hacen las ramas. En la FIGURA 2.1 se puede observar la estructura de la planta del café.

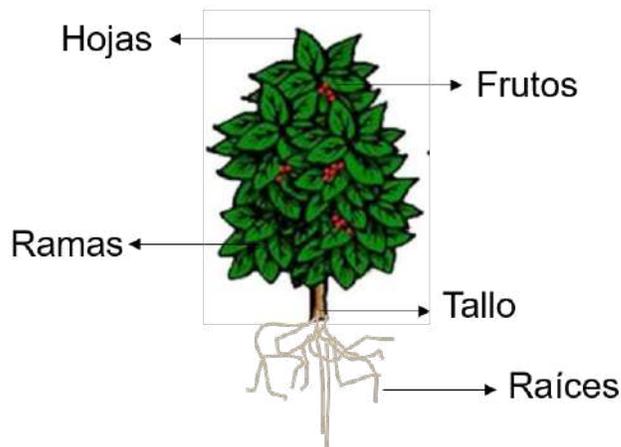


FIGURA 2.1: Estructura de la planta del café

2.1.2. La hoja del café

Es el órgano verde de la planta con diferentes tamaños y formas dependiendo de la especie y variedad del café. El color verde se debe a la clorofila que está contenida en los cloroplastos de las células y que hacen posible la fotosíntesis. Generalmente, la hoja del café mide entre cinco y diez centímetros de largo y su vida ronda entre los siete y ocho meses. En la FIGURA 2.2 se muestra una hoja de café y sus características.

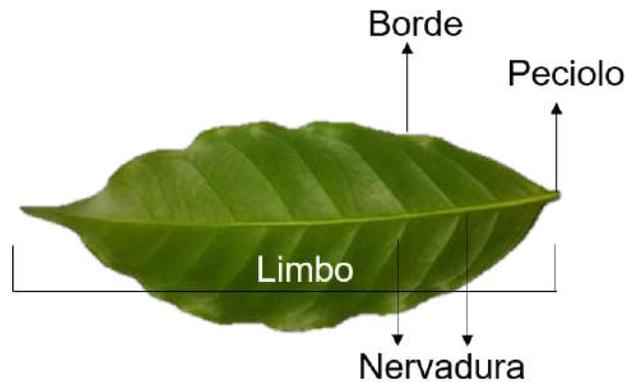


FIGURA 2.2: Hoja de la planta

2.1.3. El fruto del café

El fruto del café se llama drupa. Es un fruto que se asemeja mucho a una cereza cuando este está maduro. El fruto cambia de color dependiendo del estado fenológico en el que se encuentra la planta. Del fruto se extrae el grano el cual es el que todo el mundo consume una vez que ha pasado por diversos procesos.

2.1.4. Variedades de la planta del café en México

En México se encuentran principalmente dos especies de importancia comercial que son la *Coffea arabica* y la *Coffea canephora*.

Coffea Arabica

Tiene un sabor de mucho mayor calidad que otras especies y se desarrolla en zonas templadas con temperaturas medias mayores a 10 %, idealmente rondando los 18 ° C. Su metabolismo se alenta a medida que disminuye la temperatura de su entorno provocando que el fruto forme más compuestos precursores del sabor y consiguiendo más intensidad en su sabor.

La *Coffea Arabica* es una planta autopolinizable lo cual quiere decir que muy a menudo el polen fecunda su propia flor. Esto resulta una desventaja para la planta ya que la variabilidad genética es pobre. Esto trae consecuencias como la poca adaptabilidad a los cambios de su entorno como los climáticos, además de ser más susceptibles al ataque de plagas y enfermedades.

Los tipos de *Coffea arabica* son:

- *typica* o criollo
- *bourbon*
- *caturra*
- *maragogipe*
- *pluma Hidalgo*
- *catuai*

Coffea canephora

A diferencia de la especie *Coffea Arabica*, esta se desarrolla mejor en zonas de baja altitud y de mayor temperatura ambiental. Necesita de polinización cruzada, por lo que tienen una variabilidad genética mayor comparada con otras especies lo que la hace más adaptable a cambios en su entorno y es más resistente a plagas y enfermedades. La variedad más conocida de esta especie es la robusta.

2.2. Aprendizaje Supervisado

El aprendizaje supervisado busca relacionar un conjunto de datos de entrada que comunmente se les conoce como x con un conjunto de variables denominadas y con el objetivo de predecir salidas para datos de entrada nuevos para el sistema de aprendizaje supervisado.

El aprendizaje supervisado permite crear una función que generaliza nuevos datos en base a datos ya procesados llamados datos de entrenamiento. Estos datos pueden ser imágenes, audios, documentos de texto, entre otros y normalmente son convertidos en vectores para ser procesados por diferentes algoritmos como regresiones lineales, redes bayesianas, máquinas de vectores de soporte, árboles de decisión, etc.

La base de datos del aprendizaje supervisado es un conjunto de datos entrada-salida $(x_n, y_n)_{1 \leq n \leq N}$ con $x_n \in X$ y $y_n \in Y$

De manera general, se puede decir que con el aprendizaje supervisado se pueden resolver dos tipos de problemas:

- Regresión: $Y \in R$, la salida que se busca es un valor dentro de un conjunto continuo de reales.

- Clasificación: $Y = 1, \dots, I$, cuando el conjunto de valores de salida es finito y se tiene que atribuir una etiqueta a cada valor de entrada.

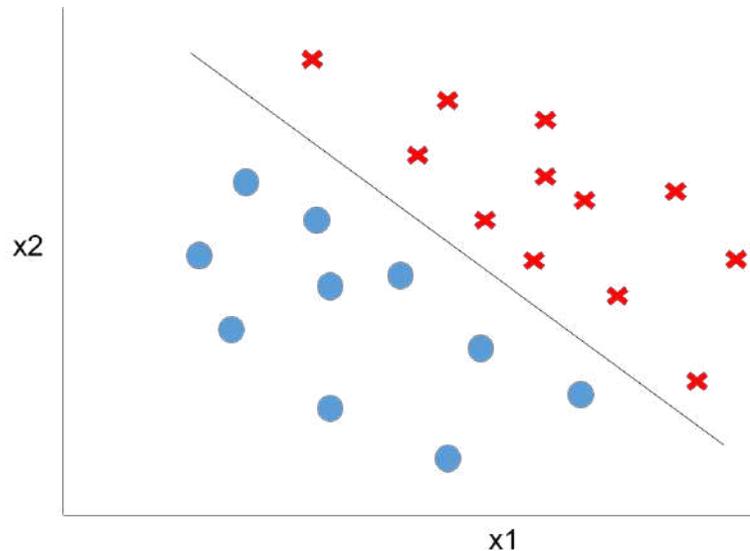


FIGURA 2.3: Clasificación de dos clases con aprendizaje supervisado

En la FIGURA 2.3, se muestra como hay una línea que divide los elementos que son de la clase círculo y los otros elementos que son de la clase cruz. La función que resulta de este aprendizaje se llama función de predicción, hipótesis o modelo.

El aprendizaje supervisado es ampliamente utilizado y se pueden observar muchos ejemplos en la vida cotidiana. Por ejemplo, en fábricas modernas pueden haber sistemas que detecten, por medio de una cámara si una fruta o vegetal se encuentran en mal estado o no (Arakeri, 2016). Estos sistemas, son entrenados con una gran cantidad de imágenes que representan frutos sanos y frutos en mal estado. De esa manera, el sistema aprende y puede clasificar los frutos mucho más rápido y con poca intervención humana.

2.3. Aprendizaje No Supervisado

El aprendizaje no supervisado permite ordenar datos que no están etiquetados y clasificarlos de acuerdo a sus características. Una de las principales diferencias entre el aprendizaje supervisado y el no supervisado es que el primero recibe como entradas

datos ya etiquetados y el segundo debe construir agrupamientos de acuerdo a las características más relevantes de estos datos. Los principales métodos del aprendizaje supervisado son: *Principal Component Analysis* y *Cluster Analysis*.

La utilización del aprendizaje no supervisado es utilizado para resolver problemas como los siguientes:

- Particionamiento o agrupamiento de datos utilizando el algoritmo de *K-means* o agrupamiento jerárquico.
- Estimación de densidad de distribución con algoritmos como *Kernel density estimation*
- La reducción de dimensión con técnicas como PCA(*Principal Component Analysis*)

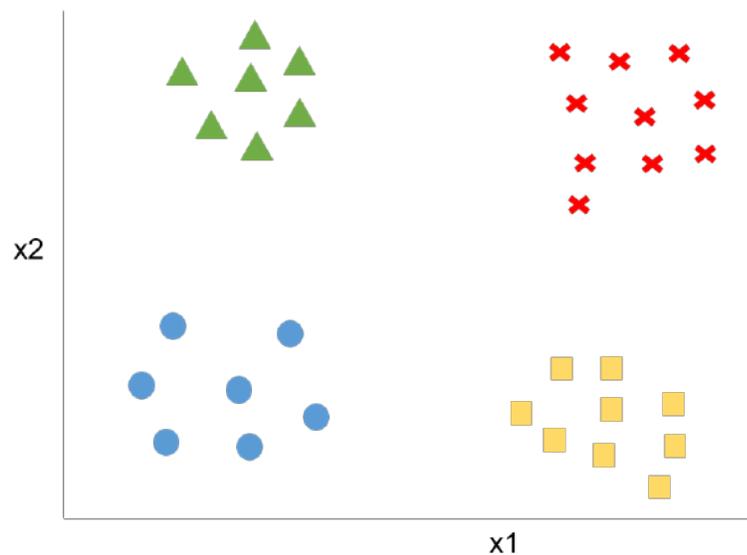


FIGURA 2.4: Agrupamiento de datos de acuerdo a sus características

En la FIGURA 2.4, se muestra como los elementos se agrupan de acuerdo a sus características comunes o aquellas más relevantes. En el ejemplo, el clasificador logró encontrar cuatros grupos o clases.

2.4. Reconocimiento de patrones

Se trata del reconocimiento automatizado de patrones y características comunes en datos. El reconocimiento de patrones intenta proveer una respuesta razonable a todos

los valores de entrada y ejecutar un proceso de coincidencia con esos valores tomando en cuenta su variación estadística. El reconocimiento de patrones tiene un gran número de aplicaciones entre las que se encuentran el análisis de datos, procesamiento de imágenes, bioinformática, compresión de datos, gráficas de computadora, entre muchas otras aplicaciones.

El reconocimiento de patrones es una tecnología que ha permitido desarrollar aplicaciones para las más diversas áreas. Por ejemplo, en la medicina se han desarrollado sistemas para asistir a los médicos en los diagnósticos a sus pacientes. Estos sistemas se llaman *computer-aided diagnosis (CAD)* (Fujita, 2010). El reconocimiento de patrones en la voz es otra aplicación que ha pasado a formar parte esencial de los asistentes virtuales inteligentes (Intelligent Virtual Assistant, IVA en inglés). Las implementaciones más utilizadas actualmente son las desarrolladas por Google, Apple, Amazon y otras empresas para reconocer e interpretar la voz como comandos para ejecutar ciertas acciones en los sistemas operativos o aplicaciones móviles (Tulshan, 2019). Por otra parte, el reconocimiento de texto *Optical Character Recognition (OCR)*, ha permitido extraer información y digitalizarla a partir de textos escritos a máquina o incluso a mano en una hoja o en cualquier otra superficie (Islam, 2017).

2.5. Pre-procesamiento

El pre-procesamiento es el conjunto de técnicas que se aplican a las imágenes digitales que permiten por ejemplo reducir el ruido y resaltar ciertos detalles visuales con el propósito de facilitar la búsqueda de información.

El pre-procesamiento de imágenes intenta resolver los siguientes problemas:

- Escalado de imágenes, para tener imágenes del mismo tamaño ya que las fotos pueden provenir de diferentes aparatos como cámaras de teléfonos y cámaras fotográficas lo cual resulta en imágenes de tamaños variables.
- Reducción de Ruido, con el objetivo de eliminar o reducir efectos de iluminación, desenfoque y difuminación no deseados que pueden presentarse en una imagen a consecuencia del proceso de captura, digitalización y transmisión.
- Conversiones de espacio de color dependiendo del propósito del estudio.

2.5.1. Escalado de imagen

El escalado de imágenes consiste en el cambio de tamaño de una imagen digital, es decir, reducir o ampliar el número de píxeles. Esto es particularmente útil cuando por ejemplo se desea que todos los pasos habituales del procesamiento y clasificación de imágenes se ejecuten de manera más rápida debido a que las imágenes son de menor tamaño. Sin embargo, se debe de poner atención a la pérdida de precisión resultado de esas modificaciones.

2.5.2. Espacios de color

Un espacio de color es un modelo matemático abstracto que permite la representación de colores con códigos únicos. Los espacios de color más utilizados en trabajos de procesamiento digital son: RGB, HSV, HSI, $L^*a^*b^*$, $L^*u^*v^*$, YUV y YCbCr.

RGB

El modelo de color RGB se basa en la teoría Young-Helmholtz de la visión cromática tricromática que fue desarrollada por Thomas Young y Hermann Helmholtz a mediados del siglo XIX, y en la teoría del triángulo de color que se elaboró por James Clerk Maxwell.

Este espacio de color está representado por los tres colores: verde, rojo y azul. Es ampliamente utilizado en sistemas como la televisión, cámaras digitales, almacenamiento de imágenes, etc. Para formar un color, los tres principales colores son sobrepuestos y cada uno de ellos puede tener una intensidad en la mezcla de colores. La intensidad cero de cada componente asigna el valor más oscuro y la intensidad máxima asigna un blanco. Un color secundario está formado por la suma de dos principales colores de igual intensidad: cian es verde y azul, magenta es azul y rojo y amarillo es rojo y verde. Cada color secundario complementa un color primario. Cian complementa al rojo, magenta complementa al verde y amarillo complementa al azul. Cuando todos los colores primarios son mezclados en iguales intensidades el resultado es el blanco.

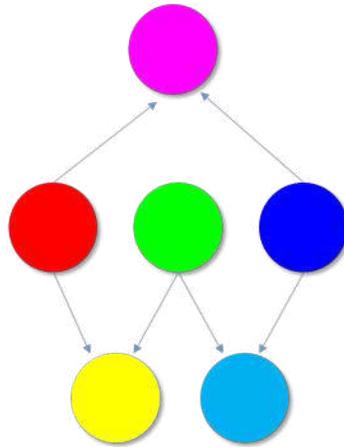


FIGURA 2.5: Espacio de color RGB

En la FIGURA 2.5 se puede ver una representación visual de los colores primarios y secundarios del modelo RGB.

2.5.3. Histograma de una imagen

El histograma de una imagen consiste en una gráfica de frecuencias donde se muestra el número de píxeles por cada nivel de color de la imagen. El histograma es útil para comparar contrastes e intensidades entre imágenes.

Cuando el rango de niveles de un color que toma la imagen se encuentra en una zona del intervalo, la imagen puede poseer poco contraste. Algunos de los problemas que se pueden detectar con el histograma de la imagen son:

- La sobreexposición resulta cuando el sensor de la cámara recibe mucha luz.
- La subexposición resulta cuando el sensor de la cámara recibe poca luz. Es lo contrario a la sobreexposición.

Estructura del Histograma

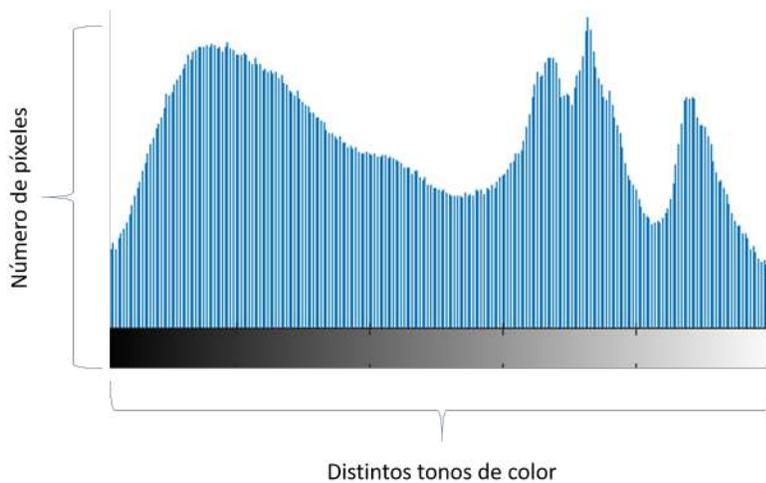


FIGURA 2.6: Estructura del Histograma

En la FIGURA 2.6 se muestra la estructura del histograma de una imagen. En el eje horizontal se encuentran los distintos tonos de la imagen y en el eje vertical el número de píxeles por cada tono.

Distribución homogénea



FIGURA 2.7: Distribución homogénea

En la FIGURA 2.7, se muestra una imagen y su histograma. Se puede observar que el histograma se distribuye a lo largo de todo el eje horizontal. La imagen no cuenta con zonas excesivamente oscuras y zonas sobreexpuestas lo cual se refleja en el histograma ya que no hay grandes 'montañas' en la gráfica. El hecho de que un histograma sea homogéneo no quiere decir que la imagen está correctamente expuesta ya que esto depende del objetivo buscado durante la captura de la imagen.

Distribución con zonas oscuras



FIGURA 2.8: Distribución con zonas oscuras

En la FIGURA 2.8, se muestra una imagen con zonas muy oscuras, este tipo de imágenes se caracteriza por tener un pico muy alto en la parte izquierda del histograma que demuestra la concentración de colores oscuros.

Distribución con sobreexposición

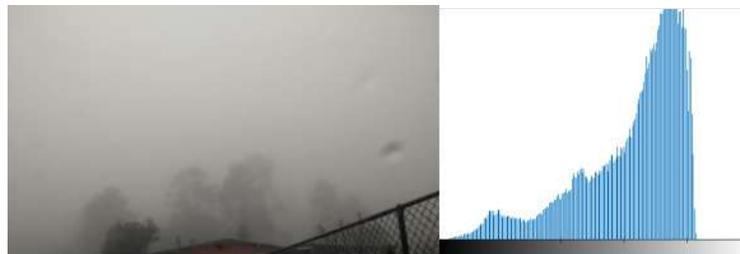


FIGURA 2.9: Distribución con sobreexposición

En la FIGURA 2.9, se puede observar que hay una imagen con una alta exposición o luminosidad que se refleja con un pico muy grande en la parte derecha donde se encuentran los tonos claros. Esta imagen fue tomada en un lugar con alta neblina lo cual causa que los tonos claros se concentren en la parte derecha. Sin embargo, esto no quiere decir que la imagen sea de baja calidad, todo depende del objetivo buscado.

Distribución con contraluz

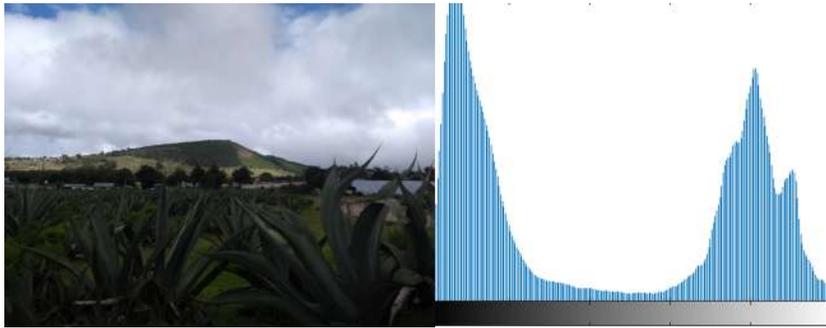


FIGURA 2.10: Distribución con contraluz

En la FIGURA 2.10, se encuentra una imagen que refleja altos niveles de concentración de tonos oscuros y también una alta concentración de tonos claros. Este tipo de imágenes a contra luz forman una letra U en el centro de su histograma. En este tipo de imágenes se puede aplicar una ecualización del histograma para hacer más evidentes los objetos o cosas capturadas en la imagen.

Ecualización del Histograma

La ecualización del histograma es una operación global que permite acentuar y mejorar el contraste visual de una imagen repartiendo de manera uniforme los valores que se encuentran contenidos en la imagen.

$$O_{i,j} = \text{floor}\{(L - 1) \times H'(I_{i,j})\} \quad (2.1)$$

En la ecuación 2.1, I es la imagen de entrada, O es la imagen una vez ecualizada, H es el histograma acumulado, L es el número de niveles de intensidad e i,j son la posición del pixel a evaluar.

Cuando el gráfico del histograma de una imagen es muy estrecho quiere decir que presenta bajo contraste pues se utilizan pocos tonos de intensidad y los valores son muy cercanos. En cambio, si una imagen presenta muchos tonos de intensidad repartidos a lo largo de todo el histograma se puede decir que la imagen presenta un buen contraste.

2.5.4. Binarización de imágenes

La binarización de la imagen trata de separar un objeto de interés del resto de la imagen o del fondo haciendo que los pixeles que conforman la imagen tomen un valor

entre dos posibles. La técnica de binarización hace un recorrido secuencial por toda la imagen para determinar que píxeles tomarán el color Negro(0) o qué píxeles tomarán el color Blanco(255) en base a un umbral T . Un píxel tomará el color Negro si el valor de su píxel está por debajo del umbral T y tomará el color Blanco si el valor de su píxel es mayor al umbral T .

Los métodos de binarización se dividen en dos tipos: Globales y Locales. Entre los métodos globales, destacan los de por umbral y el de otsu. Estos buscan encontrar un umbral para aplicar a toda la imagen. Por el contrario, los métodos locales buscan un umbral para cada píxel en la imagen haciendo uso de los valores de sus vecinos.



FIGURA 2.11: Binarización de una imagen

En la FIGURA 2.11 se muestra la foto original de la famosa Lena del lado izquierdo y en el lado derecho se encuentra la foto una vez binarizada.

2.5.5. Detección de bordes

La detección de bordes es una tarea muy importante ya que permite encontrar objetos dentro de la imagen y segmentar regiones. Los contornos son detectados donde hay zonas de píxeles con un cambio del nivel de gris muy evidente.

Dentro de los algoritmos de detección de bordes se pueden encontrar el algoritmo de Sobel, Robert Edge detector, LoG Edge detector, Declivity Edge detector, Canny Edge detector, entre otros. El algoritmo de Canny es uno de los algoritmos con mejor desempeño (Bhardwaj, 2012). El algoritmo de Canny es el siguiente:

- 1) Es recomendable que a la imagen se le aplique un filtro para eliminar el ruido. Se puede utilizar un filtro gaussiano.
- 2) Se obtiene la magnitud del gradiente con la ecuación 2.2

$$|G| = \sqrt{Gx^2 + Gy^2} \approx |Gx| + |Gy| \quad (2.2)$$

3) Se obtiene la dirección del borde usando el gradiente de dirección x y y con la ecuación 2.3

$$\theta = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad (2.3)$$

4) Eliminar cualquier valor de píxel que no sea considerado como contorno. Esto resulta en una línea delgada para el contorno.

5) Se aplica una función de histéresis basada en dos umbrales con el objetivo de reducir la aparición de contornos falsos.

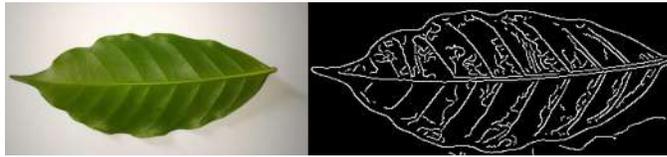


FIGURA 2.12: Detección de bordes por Canny

En la FIGURA 2.12, del lado izquierdo se muestra la imagen de una hoja y del lado derecho la obtención de sus bordes por el método de Canny.

2.6. Segmentación

La segmentación es una parte del proceso que tiene mucha relevancia ya que se separan el o los objetos de interés del resto de la imagen. A continuación se explicarán las técnicas de segmentación más utilizadas.

2.6.1. Segmentación por umbralización

Los métodos de valor umbral fueron de los primeros metodos que se utilizaron en el tratamiento de imágenes digitales. Entre los métodos más conocidos está el método Otsu que fue publicado en el año 1979 por Nobuyuki Otsu (Otsu, 1979). Los métodos de valor umbral son relativamente sencillos lo cual permite su fácil y rápida implementación. Sin embargo, este tipo de segmentación es muy sensible al ruido que pueda contener la imagen. Además, se debe tener en cuenta que el histograma de la imagen debe ser bimodal y el objeto puede ser separado del fondo usando un valor umbral (Alireza, 2016).

La segmentación por umbralización trata de resaltar regiones de características similares, mediante la clasificación de grupos de píxeles. Estos métodos buscan un valor de intensidad T llamado umbral que tiene la función de separar las regiones deseadas. La segmentación ocurre cuando se agrupan todos los píxeles cuyo valor de intensidad es menor que el umbral y , por otro lado, los píxeles con valores de intensidad mayor que el umbral T . La formula 2.4 representa de forma matemática lo que se acaba de describir.

$$T_{global}(g) = \begin{cases} 0 & \text{si } g < t \\ 1 & \text{si } g \geq t \end{cases} \quad (2.4)$$

Los métodos de umbralización recorren todos los píxeles de la imagen secuencialmente para etiquetar aquellos que son menores o mayores al umbral de intensidad que se ha fijado. Lo importante con esta técnica es encontrar el valor adecuado para el umbral.

Método Otsu para umbral

El método Otsu calcula el valor umbral con el objetivo de que la dispersión de cada segmento sea lo más pequeña posible y también que la dispersión sea lo más alta posible entre diferentes segmentos. Cabe destacar que este método no requiere de información previa de la imagen antes de su procesamiento y tampoco requiere supervisión humana para el calculo de umbrales.

El algoritmo de Otsu es el siguiente:

1) Calcular el histograma normalizado de la imagen. La ecuación 2.5 representa la probabilidad de ocurrencia del nivel de gris de una imagen según Otsu:

$$P_i = \frac{n_i}{MN} \quad (2.5)$$

2) La ecuación 2.6 permite calcular las sumas acumulativas P_1 donde $k = 0,1,\dots,L-1$.

$$P_i(k) = \sum_{k=0}^k p_i \quad (2.6)$$

El umbral T se fija en un nivel de intensidad k por lo tanto $T = k$ y se crean dos clases:

$C_1 =$ píxeles con nivel de intensidad en $[0, k]$

$C_2 =$ píxeles con nivel de intensidad en $[k+1, L-1]$

Entonces, $P_1(k)$ indica la probabilidad de que un píxel pertenezca a la clase C_1

3) Calcular las medias acumulativas $m(k)$ con la ecuación 2.7 donde $k = 0, 1, \dots, L-1$:

$$m(k) = \sum_{i=0}^k ip_i \quad (2.7)$$

4) Calcular la media global m_G :

$$m_G = \sum_{i=0}^{L-1} ip_i \quad (2.8)$$

5) Calcular la varianza entre clases:

$$\sigma_B^2(k) = \frac{[m_G P_1(k) - m(k)]^2}{P_1(k)[1 - P_1(k)]} \quad (2.9)$$

Esto permite medir la dispersión entre valores de intensidad de las dos clases C_1 y C_2

6) Maximizar la varianza anterior con la ecuación 2.10:

$$k^* = \max_{0 \leq k \leq L-1} \sigma_B^2(k) \quad (2.10)$$

Con esto se logra la dispersión máxima posible de los niveles de intensidad de las dos clases, lo cual era uno de los objetivos iniciales. Si el máximo no es único, definir k^* como la media de todos los máximos posibles.

7) Se obtiene la medida de separabilidad. Esto es el cociente entre la varianza de clases y la varianza global.

$$\mu^* = \frac{\sigma_B^2(k^*)}{\sigma_G^2} \quad (2.11)$$

Donde $f(x, y)$ es la imagen de $M \times N$. $0, 1, 2, \dots, L-1$ son los niveles de gris y n_i son el número de píxeles con nivel de intensidad i .

Cabe destacar que mientras mayor diversidad de niveles de intensidad haya entre las clases, mayor será la varianza global ya que esta es constante.

2.6.2. Segmentación por clustering

Estos algoritmos pertenecen a los métodos de aprendizaje no supervisado e intentar agrupar y presentar los datos por sus características principales. Estos métodos no necesitan datos de entrenamiento para segmentar la imagen a diferencia de los métodos de clasificación entre los que se encuentra *KNN* (*K-nearest Neighbour*, por sus siglas en inglés).

K-means es un algoritmo de agrupamiento que es ampliamente utilizado ya que es relativamente sencillo y ofrece un buen resultado en segmentación binaria. Para implementar K-means se tiene que tener un conjunto de n elementos y k grupos. Cada elemento k es asignado al grupo k más cercano a la media calculada.

El algoritmo de *K-means* consta de 4 pasos :

1. Se eligen k -centros de clusters, ya sea de forma aleatoria o siguiendo algún método heurístico,
2. Se asigna cada píxel de la imagen al cluster que minimiza la varianza entre el píxel y el centroide.
3. Recalcular los centroides calculando la media de todos sus píxeles.
4. El paso 2 y 3 se repite hasta que los píxeles no presenten cambios en sus respectivos centroides.

Distancia Cuadrada Euclidiana

La distancia más utilizada en K-means es la distancia cuadrada euclidiana. La ecuación 2.12 muestra como calcular la distancia entre dos puntos x e y en un espacio m -dimensional.

$$d(x, y)^2 = \sum_{j=1}^m (x_j - y_j)^2 = \|x - y\|_2^2 \quad (2.12)$$

Donde j es la dimensión de los puntos de muestra x e y .

Inercia de los clústeres

También llamado Suma de Errores Cuadrados

$$SSE = \sum_{i=1}^n \sum_{j=1}^k w^{(i,j)} |x^{(i)} - u^{(i)}|_2^2 \quad (2.13)$$

Donde $u(j)$ es el centroide de cluster j , $yw(i, j)$ es 1 si la muestra $x(i)$ el cluster j y 0 en contrario.

2.6.3. PCA (Principal Component Analysis)

PCA es un método estadístico que simplifica la complejidad de espacios muestrales con muchas dimensiones conservando su información. Es usualmente utilizado para la extracción de características de la imagen, para la visualización de datos complejos o para capturar la variación de los datos.

PCA encuentra un número adecuado de *clusters* que son requeridos para la segmentación ya que si hay muchos clusters hace que haya pérdida de datos y un número bajo de *clusters* hace que haya pérdida de regiones de interés.

El algoritmo de PCA es el siguiente:

- 1) Toma todos los píxeles de la imagen y los representa en una matriz. Esto se hace tomando el primer píxel de todas las imágenes acomodados en una columna.
- 2) Calcular la media de todos los valores de píxeles entre el número de píxeles de esa columna. Eso permite saber la variación del píxel con respecto a la media, pero dado que una imagen es de dos dimensiones, se tiene que sacar la covarianza.

$$Media = \frac{\text{Valor total de todos los píxeles}}{\text{Número de píxeles de la columna}} \quad (2.14)$$

- 3) La covarianza se calcula con la ecuación 2.13

$$v(x, y) = \frac{\sum_{i=1}^n (xi - X)(yi - Y)}{(n - 1)} \quad (2.15)$$

4) Calcular Eigen vector. Permite saber acerca del patrón de los datos. Los vectores Eigen son perpendiculares el uno del otro. Dibujar la línea de Eigen Vector a través del centro de los datos. El vector Eigen muestra como los datos son relacionados con respecto a la línea trazada. Después, se hace el proceso de tomar los vectores de Eigen de la matriz de covarianza que se extrajeron y que caracterizan los datos. El vector Eigen con el valor más alto es el primer componente principal(CP_1)

5) Una vez que los vectores Eigen son extraídos de la matriz de covarianza, el próximo paso es ordenar los vectores Eigen de mayor a menor. De esta manera, se obtiene los componentes principales por importancia ya que se eliminaron los datos de poca

relevancia y se mantuvieron las regiones de interés(ROI). El número de vectores Eigen es igual al número de clusters.

2.7. Extracción de Características

En este apartado se explican brevemente cuales son los tipos de características y las técnicas utilizadas para extraerlas de las imágenes.

2.7.1. Características Geométricas

Las formas geométricas son una de las características más importantes para el reconocimiento de imágenes. En muchas plantas, las enfermedades se manifiestan como figuras circulares irregulares. Sin embargo, la cromaticidad y la textura son características que diferencian una enfermedad de otra. Los descriptores geométricos básicos para esta tarea son el área, perímetro, circularidad, elipsoidad, entre otros. (Sampallo, 2003). Proveen información sobre la forma de la hoja y también sobre el tamaño del área afectada por alguna enfermedad. Es una de las características más utilizadas para la clasificación de enfermedades en plantas.

Área

El área de la región de interés está dada por el número de píxeles que la conforman.

Perímetro

Hace referencia a calcular el número de píxeles que se encuentran en los bordes de los objetos segmentados. La longitud exacta es el número de componentes horizontales y verticales más $\sqrt{2}$ veces el número de componentes diagonales siempre que se tengan las cadenas de la curva borde.

Altura y Ancho

Se obtiene calculando la altura(H) y ancho(W) de la región de interés que ha sido segmentada. Las expresiones que definen el ancho y la altura de la región segmentada son las siguientes:

$$H = j_{max} - j_{min} \quad \text{y} \quad W = i_{max} - i_{min} \quad (2.16)$$

Donde i representa el eje mayor y j el eje menor.

Excentricidad

La excentricidad permite saber la relación que hay entre la longitud del eje mayor y la longitud del eje menor. Para calcularla se utiliza el método de rectángulo de delimitación mínima (*boundingbox*)

La formula para el método de delimitación mínima es:

$$E = \frac{W - H}{H} \quad (2.17)$$

Donde W es la anchura y H la altura del objeto

Inercia

Es la sumatoria de las diferencias $(i - j)$ con $(i - j)^2$ ponderadas del valor de la inercia. Si la región segmentada presenta un alto contraste, mayor será la inercia.

$$\sum_i \sum_j (i - j)^2 c_{ij} \quad (2.18)$$

Rectangularidad

Es la forma rectangular de la imagen. Se calcula con la siguiente formula:

$$R = \frac{A_s}{A_R} \quad (2.19)$$

Donde:

A_s = área de la región de interés segmentada

A_R = área del rectángulo de delimitación mínima

Momentos de Hu

Los momentos de Hu (Hu, 1962) capturan información básica de la forma de un objeto contenido en una imagen, como su área, las coordenadas de su centroide, su orientación, etc.

Son 7 momentos de Hu. Estos ayudan a identificar por ejemplo si una hoja es la misma aunque hayan habido cambios en la rotación, traslación o inversión de la misma. Lo cual es muy importante ya que aún cuando dos hojas de la misma variedad que queremos comparar se encuentren en diferentes posiciones o tamaños, estas serán clasificadas dentro de la misma clase.

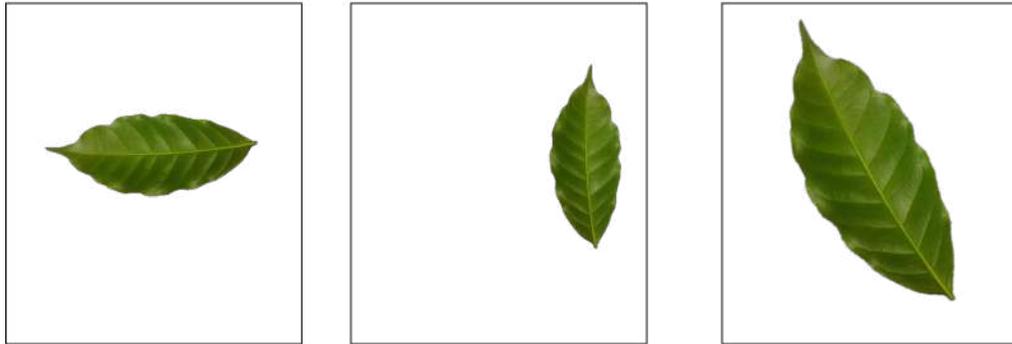


FIGURA 2.13: Los momentos de Hu ayudan a identificar o clasificar una hoja sin importar su tamaño, traslación o inversión

A continuación se muestran los 7 momentos de Hu:

$$H_1 = \mu_{20} + \mu_{02}$$

$$H_2 = (\mu_{20} - \mu_{02})^2 + 4(\mu_{11})^2$$

$$H_3 = (\mu_{30} - 3\mu_{12})^2 + (\mu_{03} - 3\mu_{21})^2$$

$$H_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{03} + \mu_{21})^2$$

$$H_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})((\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2) + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})(3(\mu_{30} + \mu_{12})^2 - (\mu_{03} + \mu_{21})^2)$$

$$H_6 = (\mu_{20} - \mu_{02})((\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2) + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03})$$

$$H_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})((\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2) + (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})(3(\mu_{30} + \mu_{12})^2 - (\mu_{03} + \mu_{21})^2)$$

2.7.2. Características Cromáticas

Estas características permiten saber la intensidad de color en el área de interés. Pueden ser calculadas por cada canal de intensidad.

Entre los descriptores de cromaticidad se encuentra el color promedio del área de interés, el gradiente promedio en el borde, similitud de color entre dos imágenes, etc. (Cervantes, 2017).

A continuación se muestran algunas características que se pueden obtener con el color:

Media Aritmética

Es una medida de tendencia central que calcula el punto medio de un conjunto de valores, es decir, la mitad de los valores estará por encima y la otra mitad por debajo de esta medida de tendencia central.

$$x = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (2.21)$$

Varianza

Es una medida de dispersión que muestra la variabilidad entre un conjunto de datos y su media.

$$Var(X) = \frac{\sum_1^n (x_i - \bar{X})^2}{n} \quad (2.22)$$

Curtosis

Es una medida estadística que permite saber como están concentrados los datos alrededor de la media.

Hay diferentes tipos de curtosis:

Leptocúrtica

Se llama así cuando hay una gran concentración de datos en torno a su media.

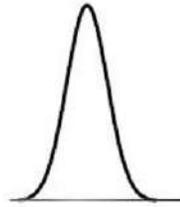


FIGURA 2.14: Distribución Leptocúrtica

Mesocúrtica

La concentración de datos es normal en torno a su media, es decir, hay una distribución normal de los datos.

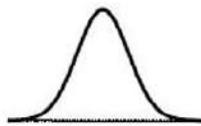


FIGURA 2.15: Distribución Mesocúrtica

Platicúrtica

Existe una concentración de datos en torno a su media.



FIGURA 2.16: Distribución Platicúrtica

Cálculo de la curtosis

Para datos sin agrupar:

$$g_2 = \frac{1}{N} \frac{\sum (x_i - \bar{x})^4}{\sigma^4} \quad (2.23)$$

Para datos agrupados en tablas de frecuencias:

$$g_2 = \frac{1}{N} \frac{\sum f_i (x_i - \bar{x})^4}{\sigma^4} \quad (2.24)$$

Para datos agrupados en intervalos:

$$g_2 = \frac{1}{N} \frac{\sum f_i (mx_i - \bar{x})^4}{\sigma^4} \quad (2.25)$$

Donde:

g_2 = es el coeficiente de curtosis

N = es el número de datos

x_i = es el valor i -ésimo de los datos

σ = es la desviación estandar de la distribución

f_i = es la frecuencia absoluta del dato i -ésimo

mx_i = es la marca de clase

Histograma de gradientes orientados HOG

Es un descriptor de características cromáticas que utiliza un único vector de características para describir un objeto en lugar de muchos vectores de característica que representan partes más pequeñas de ese objeto. El gradiente sirve para detectar los

cambios de intensidad en el color de cada píxel. Los histogramas de gradientes orientados han sido ampliamente utilizados en la detección de rostros y peatones (Zhu, 2006). El algoritmo para calcular el HOG es:

1. Seleccionar una imagen de un objeto de interés y convertirla a blanco y negro ya que no se necesita el color para este proceso
2. Dividir la imagen en celdas de tamaño fijo
3. Calcular el histograma de orientaciones de cada celda
4. Crear el descriptor global con los histogramas de todas las celdas

Para el cálculo del histograma de orientaciones en cada celda se debe de

1. Definir el rango de orientaciones en un número de intervalos fijos
2. Asignar cada píxel de la celda a un intervalo en función de la orientación del gradiente
3. Sumar la magnitud del gradiente de todos los píxeles asignados a un intervalo

En la FIGURA 2.17 se muestra una imagen con su histograma de gradientes orientados superpuesto.

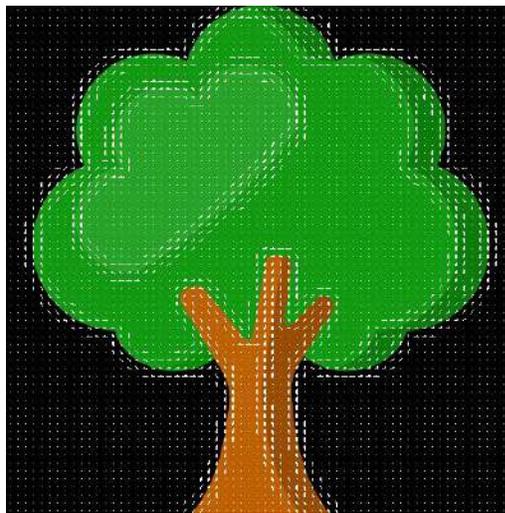


FIGURA 2.17: Histograma de gradientes orientados

Filtros de Gabor

Los filtros de gabor son ampliamente utilizados para el análisis de texturas. Se tratan de filtros paso banda que dan información de las frecuencias espaciales que contienen las imágenes. Matemáticamente se representan como una función lineal que se compone de una función sinusoidal y se multiplica por una función gaussiana.

$$g(x, y) = s(x, y)w_r(x, y) \quad (2.26)$$

donde $s(x, y)$ es la senoide compleja, llamada «portadora», y $w_r(x, y)$ es la gaussiana o «envolvente»

El filtro tiene un componente real y un componente imaginario representando direcciones ortogonales.

Complejo

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right) \quad (2.27)$$

Real

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi\frac{x'}{\lambda} + \psi\right) \quad (2.28)$$

Imaginario

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \sin\left(2\pi\frac{x'}{\lambda} + \psi\right) \quad (2.29)$$

Dónde:

$$x' = x \cos \theta + y \sin \theta$$

$$y' = -x \sin \theta + y \cos \theta$$

λ representa la longitud de onda del factor sinusoidal, θ representa la orientación de la normal a las franjas paralelas de una función de Gabor, ψ es el desplazamiento de fase, σ es la desviación estandar de la envolvente gaussiana y γ es la relación de aspecto espacial y especifica la elipticidad del soporte de la función de Gabor.

Patrones Binarios Locales LBP

El principio general es comparar el nivel de luminiscencia de un píxel con los niveles de sus vecinos, por lo tanto sirven como descriptores para clasificación de texturas.

Dependiendo de la escala de vecinos utilizada, los LBP puede ser utilizados para encontrar zonas de interés como esquinas o bordes.

El algoritmo consiste en:

1. Seleccionar el píxel central el cual se comparará con los vecinos
2. Seleccionar una escala de vecindad 8x8, 16x16
3. Calcular la diferencia con sus vecinos, si el resultado de la diferencia del píxel central con uno de sus vecinos es mayor o igual a cero, se le asigna un valor binario de 1, de lo contrario, si es menor a 0, se le asigna 0
4. Los bits activos(1) contarán para una sumatoria de 2^n descendiendo de izquierda a derecha exceptuando el píxel central

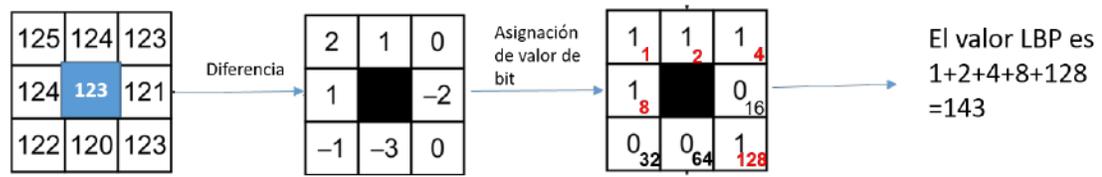


FIGURA 2.18: Ejemplo del calculo del valor LBP del píxel central 123

Para ilustrar un poco más cómo se puede diferenciar una imagen de otra gracias a las características obtenidas con los patrones binarios locales. En la FIGURA 2.19, se muestran 3 imágenes: una hoja, la misma hoja rotada y un paisaje aleatorio. Se extrajeron las características LBP y se compararon con una gráfica. Las barras azules muestran la comparación entre la hoja y la hoja rotada, se puede observar que el error es mínimo. Por otro lado, la barra naranja muestra el error que es mayor al comparar la hoja con el paisaje.

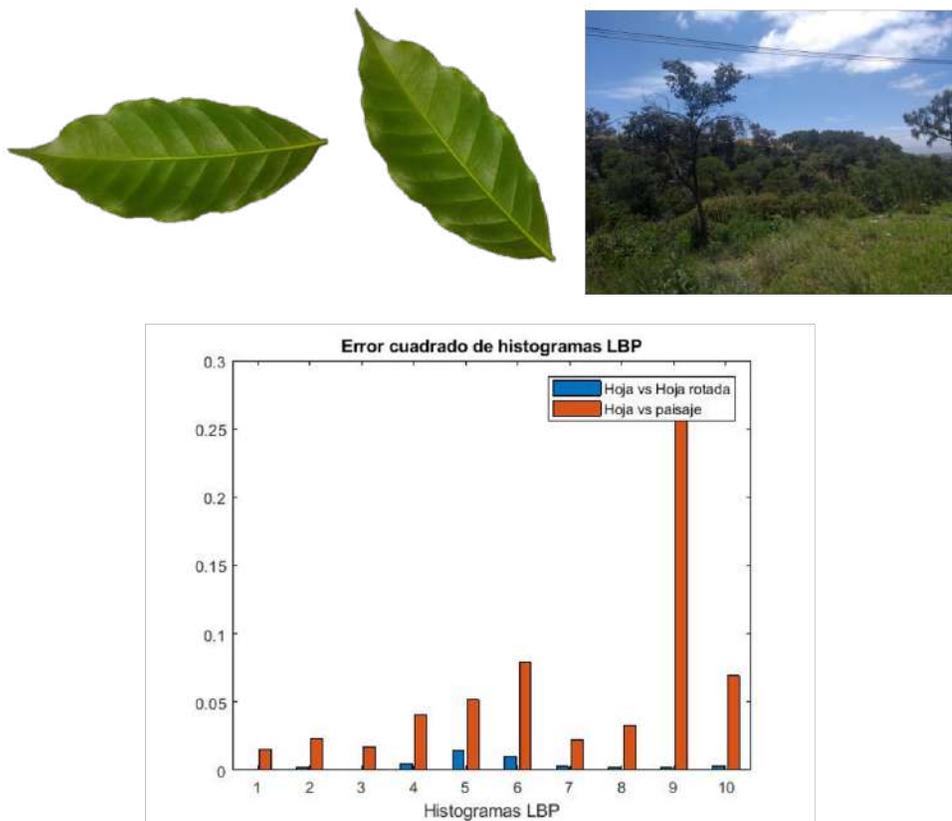


FIGURA 2.19: Comparación de características de una hoja, una hoja rotada y un paisaje aleatorio

2.7.3. Características Texturales

Las características texturales suelen definirse como patrones que se repiten en una imagen. Una textura está definida por la uniformidad, densidad, grosor, regularidad, intensidad y direccionalidad de medidas discretas del tono en los píxeles y de sus relaciones espaciales.

En el caso de las hojas, las características texturales se refieren a las características de su superficie como la nervadura, el relieve del daño causado por alguna enfermedad, encojimiento de la hoja, rugosidad, aspereza, granulación, etc.

Características Texturales de Haralick

Las características de Haralick son reglas estadísticas que permiten saber la distribución y la relación espacial de nivel de gris en una imagen. Son catorce descriptores

de textura basados en la matriz de coocurrencia del nivel de gris la cual permite discriminar patrones.

Matriz de Coocurrencia GLCM(Grey Level Co-ocurrence Matrix)

Esta matriz está basada en estadísticas de 2do orden y aproxima la probabilidad de distribución conjunta de un par de píxeles. Es utilizada para analizar texturas y discriminar patrones. Además, calcula la relación espacial que los tonos de gris tienen entre ellos. Las relaciones son calculadas en direcciones específicas(0°, 45°, 90°, 135°) con los píxeles vecinos dentro de una ventana móvil que se desplaza dentro de la imagen. La siguiente figura muestra una ventana móvil de 3x3, el píxel de referencia X y sus vecinos.

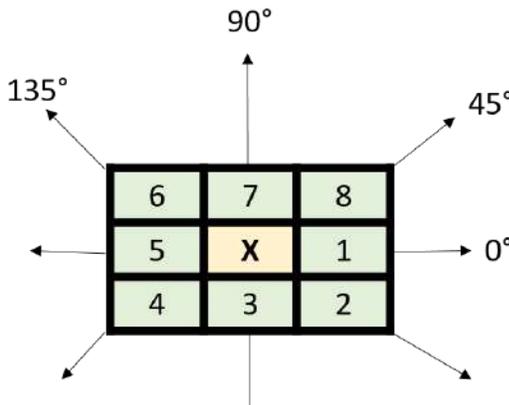


FIGURA 2.20: Matriz de Coocurrencia

$p(i, j)$: es la GLCM normalizada.

N : es el número de niveles de gris en $p(i, j)$

A continuación se muestran algunas de las características de Haralick y las ecuaciones para obtenerlas.

Segundo Momento Angular

Este descriptor permite medir la uniformidad y puede indicar la suavidad de una textura.

$$f1 = \sum_{k=1}^N \sum_{j=1}^N (p(k, j))^2 \quad (2.30)$$

Contraste

Cuantifica las variaciones de los tonos de gris de una región de interés. Entre más variación de grises, mayor es el contraste. Si el contraste es 0, los niveles de gris son constantes en la imagen.

$$f2 = \sum_{k=1}^{N-1} n^2 \left(\sum_{k=1}^N \sum_{j=1}^N (p(k, j)) \right) \quad (2.31)$$

Correlación

La correlación es la dependencia lineal de los tonos de gris en la imagen, si la correlación es 0, no existe correlación lineal entre los niveles de gris.

$$f3 = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (2.32)$$

Suma de cuadrados - varianza

$$f4 = \sum_i \sum_j (i - \mu)^2 p(i, j) \quad (2.33)$$

Momento de Diferencia Inversa(Homogeneidad)

Proporciona información sobre la regularidad local de la textura.

$$f5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \quad (2.34)$$

Suma promedio

Mide la relación entre zonas claras y densas de la imagen lo cual quiere decir que es una medida del promedio de los niveles de gris en la región de interés.

$$f6 = \sum_{j=2}^{2N} j p_{x+y}(i) \quad (2.35)$$

Suma de Varianza

Provee información final acerca de los descriptores de varianza.

$$f7 = \sum_{j=2}^{2N} (j - f_s)^2 p_{(x+y)}(i) \quad (2.36)$$

Suma de Entropía

Provee información final acerca de los descriptores de entropía.

$$f8 = - \sum_{j=2}^{2N} p_{(x+y)}(i) \log_{10} ((p_{(x+y)}(i))) \quad (2.37)$$

Entropía

Registra la heterogeneidad de una imagen, tomando valores altos en texturas con alta aleatoriedad. La entropía es alta cuando los elementos de la matriz de coocurrencia tienen relativamente valores iguales.

$$f9 = - \sum_i \sum_j p(i, j) \log\{p(i, j)\} \quad (2.38)$$

Diferencia de Varianza

Mide la variación existente en las magnitudes de las transiciones de intensidad.

$$f10 = \text{varianza de } p_{(x-y)} \quad (2.39)$$

Diferencia de Entropía

Provee información de la diferencia de los descriptores de entropía.

$$f11 = - \sum_{i=0}^{N-1} p_{(x-y)}(i) \log_{10} ((p_{(x-y)}(i))) \quad (2.40)$$

Medidas de información de Correlación 1

Mide la semejanza de la imagen consigo misma desplazada.

$$f_{12} = \frac{HXY - HXY1}{\max(HX, HY)} \quad (2.41)$$

Medidas de información de Correlación 2

$$f_{13} = (1 - \exp[-0,2(HXY2 - HXY)])^{1/2} \quad (2.42)$$

Donde HXY se deduce en la ecuación:

$$HXY = - \sum_i \sum_j p(i, j) \log\{p(i, j)\} \quad (2.43)$$

$HXY1$ y $HXY2$ están descritas en las siguientes ecuaciones:

$$HXY1 = - \sum_{i=1}^N \sum_{j=1}^N p(i, j) \log_{10} ((p_x(i)p_y(j))) \quad (2.44)$$

$$HXY2 = - \sum_{i=1}^N \sum_{j=1}^N p_x(i)p_y(j) \log_{10} ((p_x(i)p_y(j))) \quad (2.45)$$

Máximo Coeficiente de Correlación

El máximo coeficiente de correlación indica la medida de la dependencia que tengan dos variables aleatorias del sistema. La raíz cuadrada del segundo mayor valor propio de Q esta descrita en la ecuación:

$$Q(i, j) = \sum_k \frac{p(i, k)p(i, k)}{p_x(1)p_y(k)} \quad (2.46)$$

2.8. Clasificación

2.8.1. Redes Bayesianas

Son útiles para modelar un fenómeno entre un conjunto de variables y sus relaciones de dependencia. A partir de este modelo, se puede aplicar la inferencia bayesiana. La inferencia bayesiana permite estimar la probabilidad posterior de las variables que

no son conocidas, tomando como base las variables que ya son conocidas por el modelo. Las redes bayesianas son muy utilizadas para clasificación, predicción, diagnóstico, entre otras.

A las redes bayesianas se les define como una representación gráfica de dependencias para razonamiento probabilístico. Todas las relaciones de independencia condicional representadas en el grafo corresponden a relaciones de independencia en la distribución de probabilidad. La ley de Bayes es mostrada en la ecuación 2.46.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (2.47)$$

Esta ley es reescrita haciendo uso de la notación de los *targets* y el vector de características como en la ecuación 2.47 la cual predice la probabilidad t_j dada la observación de las características del vector x .

$$P(t_j|x) = \frac{P(x|t_j)P(t_j)}{\sum_{i=1}^m P(x|i)P(t_i)} \quad (2.48)$$

Dado que x es un vector con n características, queda como la ecuación 2.48.

$$P(x|t_j)P(t_j) = \prod_{i=1}^n P(x_i|t_j)P(t_j) \quad (2.49)$$

Inferencia

La propagación de probabilidades se refiere a alimentar la red bayesiana con la evidencia para conocer la probabilidad *a posteriori* de las variables. De esta manera, se le dan valores a ciertas variables y se obtiene la probabilidad posterior de las variables no conocidas dadas las variables conocidas.

Los tipos de algoritmos de inferencia

1. Una variable, cualquier estructura: Algoritmo de eliminación
2. Todas las variables, estructuras sencillamente conectadas: propagación
3. Todas las variables, cualquier estructura: Agrupamiento, Simulación estocástica y condicionamiento.

2.8.2. Naive Bayes

Naive Bayes es un algoritmo relativamente simple que utiliza la regla de Bayes en conjunto con una asunción de que los atributos son condicionalmente independientes dada la clase. Naive Bayes utiliza un mecanismo para usar información en un conjunto de datos para estimar la probabilidad $P(y|x)$ de cada clase y dado un objeto o dato x .

Estructura del sistema de aprendizaje

Naive Bayes es una forma de clasificador de redes bayesianas basados en la regla de Bayes.

$$P(y|x) = P(y)P(x|y)P(x) \quad (2.50)$$

Con la asunción de que los atributos son condicionalmente independientes dada la clase.

$$P(x|y) = \prod_{i=1}^n P(x_i|y) \quad (2.51)$$

Donde x_i es el valor de i -ésimo atributo en x , y n es el número de atributos.

$$P(x) = \prod_{i=1}^k P(c_i)P(x|c_i) \quad (2.52)$$

Donde k es el número de clases y c_i es la i -ésima clase. Por lo tanto, la primera ecuación puede ser calculada normalizando los numeradores de la ecuación después del signo =.

2.8.3. Support Vector Machines(SVM)

Las máquinas de vectores de soporte fueron desarrollados por Vladimir Vapnik en los laboratorio de AT&T. Estos algoritmos de clasificación han sido ampliamente utilizadas en años recientes por su precisión en la clasificación, que es incluso comparable a las redes neuronales artificiales (Soon, 2017).

Las SVM son clasificadores lineales, lo cual quiere decir que clasifican entre dos conjuntos de datos separados por una línea en el hiperplano. Son un modelo que representa los puntos de muestra en el espacio, separando las clases a dos espacios lo más

amplios posibles mediante un hiperplano de separación y los vectores de soporte. La solución del hiperplano óptimo es la combinación de unos cuantos puntos de entrada que son llamados vectores de soporte.

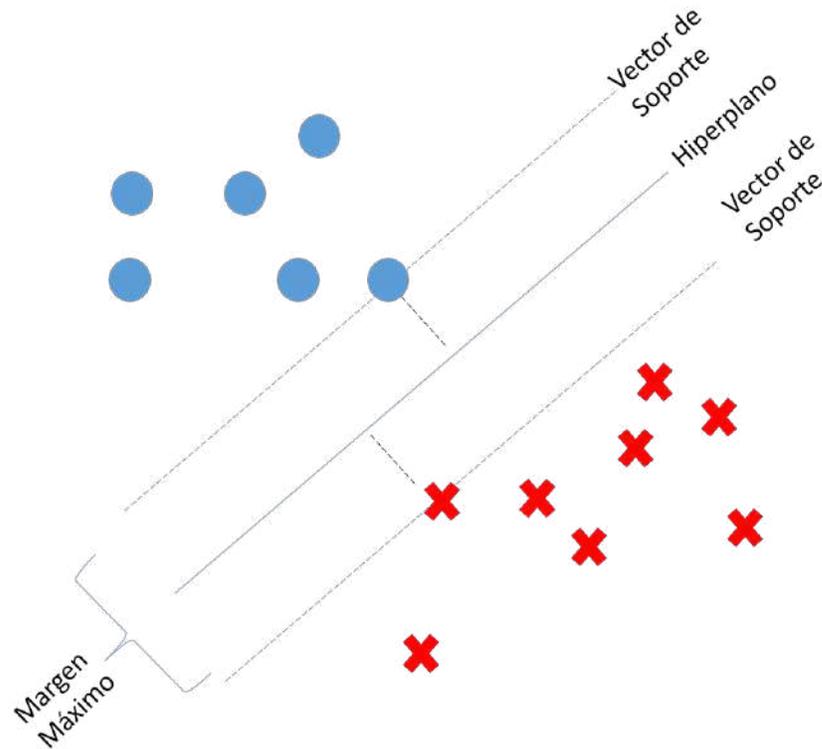


FIGURA 2.21: Ilustración de una SVM para clasificar 2 clases

Cuando no es posible separar los datos en 2 clases dada la naturaleza de estos, se pueden utilizar la función *kernel* que transforma el espacio de entrada en un espacio altamente dimensional para así poder ser separados linealmente.

Caso linealmente separable

Cada punto de entrenamiento $x_i \in \mathbb{R}^N$ pertenece a alguna de las dos clases y se le da una etiqueta $y_i \in \{-1, 1\}$ para $i = 1, \dots, l$. Comúnmente, la búsqueda de un hiperplano adecuado en un espacio de entrada es muy restrictivo para ser de uso práctico. Una solución es mapear el espacio de entrada en un espacio de características de una dimensión mayor y buscar el hiperplano óptimo en esa dimensión. Sea $z = \varphi(x)$ la

notación del correspondiente vector en el espacio de características con un mapeo φ de \mathfrak{R}^N a un espacio de características z . Encontrar el hiperplano:

$$w \cdot z + b = 0 \quad (2.53)$$

Definido por el par (w, b) , tal que se puede separar el punto x_i de acuerdo a la función

$$f(x_i) = \text{sign}(w, z + b) = \begin{cases} 1 & y_i = 1 \\ -1 & y_i = -1 \end{cases}$$

(2.54)

Donde w y $b \in \mathfrak{R}$. El conjunto S se dice que es linealmente separable si existe (w, b) tal que las inecuaciones:

$$\begin{cases} (w \cdot z_1 + b) \geq 1, & y_1 = 1 \\ (w \cdot z_1 + b) \leq -1, & y_1 = -1 \end{cases} \quad (2.55)$$

Sean válidas para todos los elementos del conjunto S . Para el caso linealmente separable de S , se puede encontrar un único hiperplano óptimo, para el cual, el margen entre las proyecciones de los puntos de entrenamiento de dos diferentes clases es maximizado.

En la FIGURA 2.22, en la imagen izquierda se encuentra un conjunto de datos que no puede ser separado linealmente pero haciendo uso de los *kernels* de SVM, ese conjunto de datos puede ser transformado a un espacio altamente dimensional como en la imagen de la derecha donde sí se puede separar el conjunto de datos.

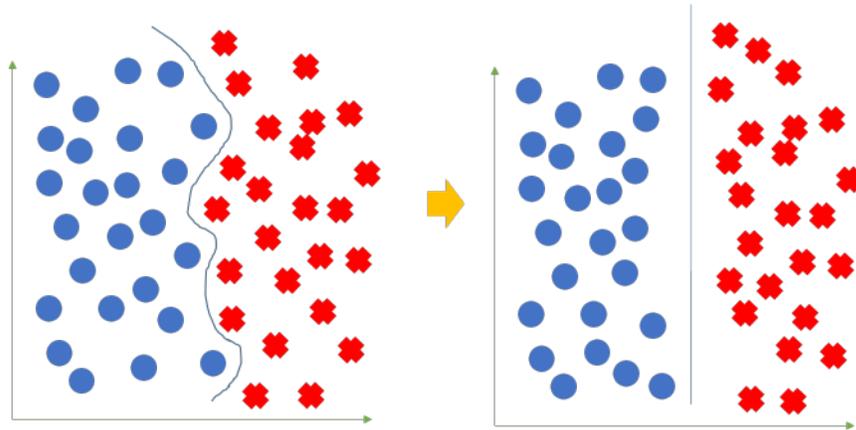


FIGURA 2.22: Utilización de Kernels

2.8.4. Redes Neuronales Artificiales

Las Redes neuronales artificiales comenzaron a desarrollarse a partir del trabajo de Warren McCulloch y Walter Pitts en 1943 que publicaron un modelo informático para redes neuronales que basaron en las matemáticas y los algoritmos. Este trabajo fue un parteaguas en esa época ya que fue en ese momento que el estudio de las redes neuronales se dividió en dos enfoques distintos. Por un lado un enfoque hacia los procesos biológicos en el cerebro y por otro lado, un enfoque hacia la aplicación de la redes neuronales para la inteligencia artificial.

Las redes neuronales artificiales intentan emular el comportamiento de las neuronas cerebrales del ser humano. Las redes neuronales artificiales son conjuntos de neuronas artificiales que están interconectadas entre sí y tiene como finalidad procesar información por medio de modelos matemáticos.

Aunque el concepto de redes neuronales artificiales no es algo nuevo, su utilización se vió limitada debido a que los recursos computacionales de esa época no era suficiente para ejecutar algoritmos de este tipo. Esto provocó que las redes neuronales artificiales quedaran únicamente en teoría.

Una neurona artificial recibe entradas, las cuales se procesan en el nodo y genera una salida ya sea hacia otras neuronas o como resultado final.

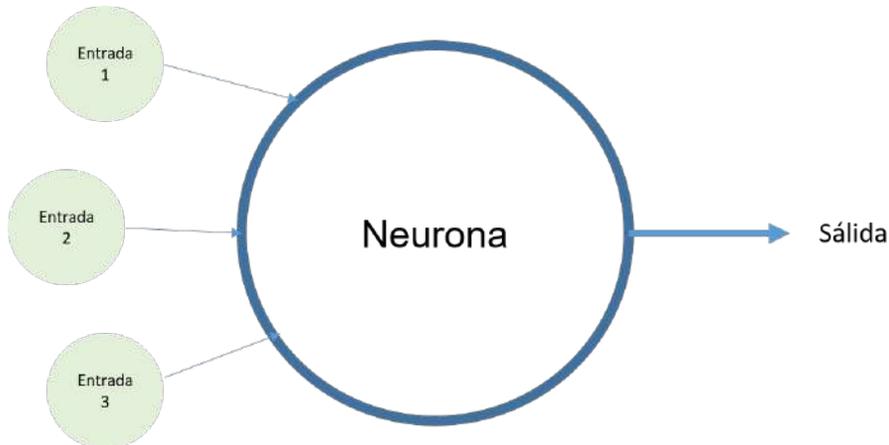


FIGURA 2.23: Composición de una neurona artificial

Perceptrón

El perceptrón fue creado por Frank Rosenblatt (Rosenblatt, 1960). Se refiere a la unidad básica de inferencia en forma de discriminador lineal capaz de separar con un hiperplano únicamente los elementos buscados y discriminando los que no se requieren. El valor de $f(x)$ se usa para clasificar x como positivo o negativo.

$$f(x) = \begin{cases} 1 & \text{si } w \cdot x - u > 0 \\ 0 & \text{en otro caso} \end{cases} \quad (2.56)$$

Donde w es un vector de pesos reales y $w \cdot x$ es el producto escalar. u es el umbral, que es una constante que no depende del valor de la entrada.

Las múltiples conexiones entre las neuronas artificiales hacen que estas tengan un peso que va cambiando mediante un algoritmo de aprendizaje.

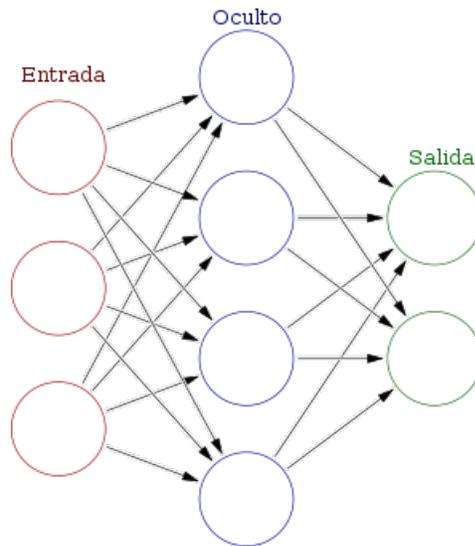


FIGURA 2.24: Red neuronal con capas de entrada, capas ocultas y capas de salida

Funciones de activación

Las entradas de la red neuronal pasan por funciones de activación que devuelven una salida. A cada capa se le asigna una función de activación para poder procesar el dato, y en base a eso poder predecir.

Función umbral(Escalón)

Se utiliza para clasificar. Devuelve cero cuando x es menor que cero y 1 si x es mayor o igual a cero.

$$\begin{aligned} y &= \text{sign}(x) \\ y &= H(x) \end{aligned} \tag{2.57}$$

Función Sigmoide(Logística)

Se utiliza para probabilidades ya que su rango de salidas está entre 0 y uno. Si x es menor a 0, la salida será 0. Si x es 0 la salida será 0.5. En valores más altos su valor será aproximadamente 1.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2.58}$$

Función Softmax

Se utiliza para clasificación múltiple y asignación de probabilidades. Puede tomar como entrada una imagen de algún animal doméstico y predecir de cuál se trata. Softmax devuelve porcentajes, por lo que el resultado será el que tenga mayor probabilidad. La suma de todas las probabilidades será 1.

$$\sigma(x_j) = \frac{e^{x_j}}{\sum_i e^{x_i}} \quad (2.59)$$

Capítulo 3

Metodología

En este capítulo se explican los pasos seguidos para implementar el sistema de reconocimiento de patrones por visión artificial aplicado a la identificación de enfermedades en la planta del café. En la FIGURA 3.1, se aprecia que el primer paso fue conseguir un conjunto de datos. Después se hizo un re-escalamiento de las imágenes. Posterior a eso, se procedió a segmentar las regiones de interés con diferentes técnicas. El siguiente paso fue extraer las características cromáticas, geométricas y texturales de las hojas. Finalmente, los vectores de características fueron procesados por diferentes algoritmos de clasificación.

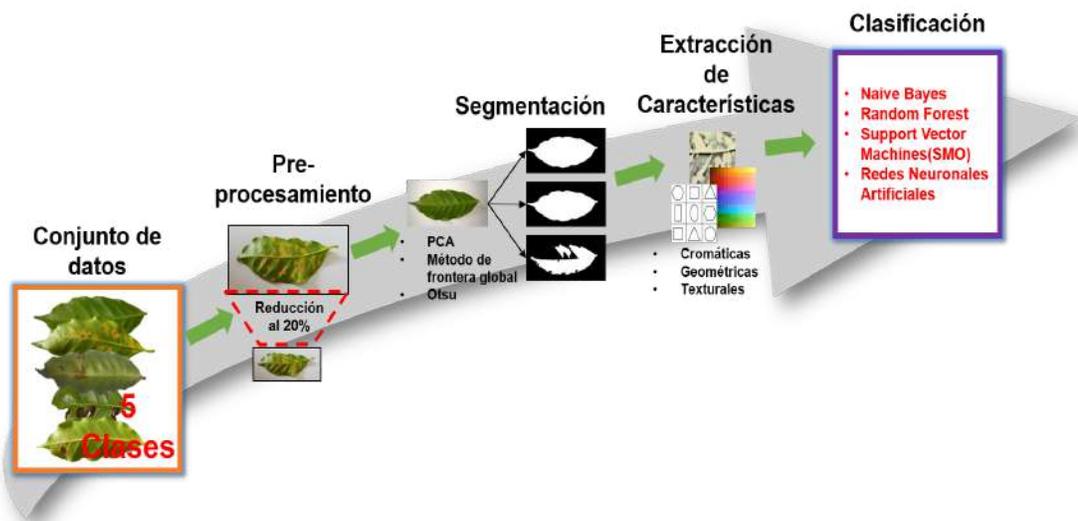


FIGURA 3.1: Metodología propuesta

3.1. Conjunto de datos utilizado

El conjunto de datos que se utilizó para este trabajo fue obtenido del sitio web data.mendeley.com, el trabajo se llama «A brazilian Arabica Coffee Leaf images dataset to identification and quantification of coffee diseases and pests». Contiene 739 imágenes de hojas sanas y cuatro enfermedades comunes de la especie *coffea arabica* que fueron tomadas en cultivos de café en Santa María de Marechal, Floreano, Brasil. Todas las fotos se tomaron con teléfonos como ASUS Zenfone 2, Xiaomi Redmi 5A, Xiaomi S2, Galaxy S8, y iPhone 6S en un ambiente controlado ya que fueron puestas en un fondo blanco. Las imágenes originales del *dataset* tienen un resolución de 2048 píxeles de ancho por 1024 de píxeles de alto.

Las imágenes del conjunto de datos no estaban organizadas en carpetas sino solo etiquetadas con la enfermedad y la severidad a la que pertenecen. Por tal motivo, se creó una carpeta por cada enfermedad y las imágenes que pertenecieran a una enfermedad dada se colocaron ahí sin importar la gravedad de la enfermedad.

Cabe destacar que en cada hoja se presenta síntomas de una sola enfermedad, por lo tanto reconocer más de una enfermedad por hoja está fuera del alcance de este trabajo debido a la dificultad para el sistema para detectar la enfermedad predominante.

En la FIGURA 3.2, se exhiben algunas hojas con las enfermedades del conjunto de datos.



FIGURA 3.2: Hojas del data set

A continuación en la FIGURA 3.3, se presenta la cantidad de imágenes para las 5 clases. Se puede observar que es un conjunto de datos desbalanceado ya que hay grandes diferencias en cuanto a cantidad de imágenes entre la clase sana y la clase de *Phoma, Quema*. Es necesario que se capturen o se obtengan más imágenes para las clases que tienen menos imágenes.

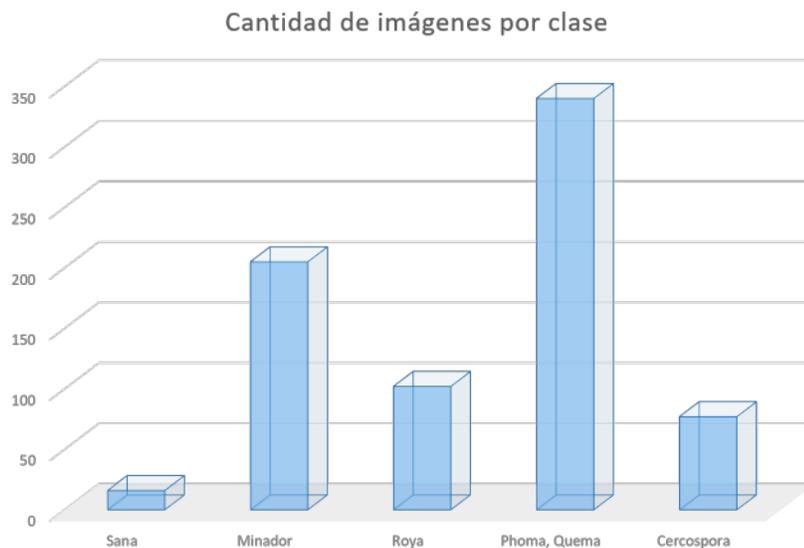


FIGURA 3.3: Hojas del data set

3.2. Pre-procesamiento

Como pre-procesamiento, se redimensionaron las imágenes originales que tenían un tamaño de 2048 píxeles de ancho y 1024 píxeles de alto, las cuales pasaron a 410 píxeles de ancho y 205 píxeles de alto. Esto quiere decir que se redujo al 20 % por ciento de su tamaño original. La reducción de tamaño no resultó en una pérdida de información de las características de las hojas. Por otro lado, se consigue que el sistema funcione con un mejor desempeño en cuanto a tiempo de ejecución.

3.3. Segmentación

La segmentación es el proceso donde se busca separar una región de interés (ROI, por sus siglas en inglés) de su fondo. Esto con el objetivo de aprovechar y mejorar el análisis sobre el área que interesa. Las técnicas que se utilizaron para este trabajo fueron Otsu (Otsu, 1979), Análisis de componentes principales PCA (Hotelling, 1933) y Método de Frontera Global. La técnica que mejor resultados generó al segmentar las imágenes fue PCA.

Al segmentar la imagen, el sistema de reconocimiento de patrones se enfoca únicamente en la región de interés que se determina por su contorno y propiedades que se usarán en el siguiente paso de la metodología que es la extracción de características.

En la FIGURA 3.4, se muestran hojas afectadas por las 4 diferentes enfermedades del café. Cada hoja fue segmentada por 3 métodos diferentes. Se puede observar que el método de frontera global tiene dificultades para segmentar algunas hojas de café. Otsu y PCA por su parte obtienen buenos resultados. No obstante, el hecho de que algunas imágenes contengan sombras derivadas de los ángulos y horas en que se tomaron las fotos, a veces estas sombras confunden a los algoritmos de segmentación implementados y terminan por segmentar esas pequeñas áreas como si fueran parte de la hoja. Sin embargo, los algoritmos de segmentación toman el área más grande segmentada, que en este caso es el área de la hoja, ignorando las sombras aisladas.

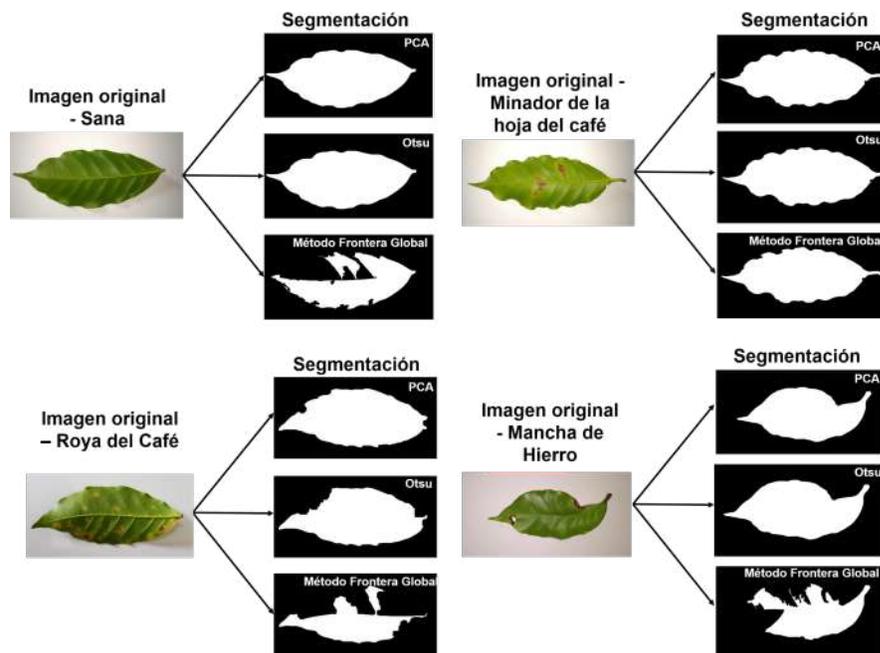


FIGURA 3.4: Segmentación por tres métodos

La segmentación que finalmente se utilizó fue PCA, ya que en la mayoría de las imágenes obtuvo resultados satisfactorios.

3.4. Extracción de características

La extracción de características es una de las tareas más importantes en el procesamiento digital de imágenes. En este paso se obtienen las formas, colores, y relieves plasmados en las imágenes y que son ocasionados por los síntomas de las plagas y enfermedades en las hojas del café. Este proceso genera vectores de características que

sirven como entradas para los algoritmos de clasificación.

El principal objetivo de la extracción de características es obtener la información más relevante de las imágenes. Un buen conjunto de características contiene información que permite discriminar una clase de otra, es decir, es representativa y discriminativa.

3.4.1. Características Geométricas

Se utilizaron momentos geométricos como los momentos centrales, que son funciones que ayudan a reconocer una imagen dentro de un eje de coordenadas independientemente de su transformación geométrica como la traslación, escalado y rotación. De igual manera, se extrajeron los momentos de Hu (Hu, 1962) que consisten en 7 momentos que ayudan también a precisar si una hoja es la misma sin importar su orientación, localización o tamaño.

También se utilizaron los primeros 8 descriptores de Fourier para el cálculo del contorno de la hoja. Por último, se utilizaron descriptores de Flusser los cuales permiten saber sobre la difuminación de la imagen (Flusser, 2009).

3.4.2. Características Cromáticas

Las características cromáticas permiten saber información sobre la intensidad del color en una región de interés. En este trabajo se obtuvo lo siguiente: Contraste utilizando una máscara de 5x5, Características de Gabor (Gabor, 1946) y momentos de Hu con intensidad (Hu, 1962).

3.4.3. Características Texturales

La textura es una propiedad innata de todas las superficies. Por ejemplo, la superficie de la madera, el patrón de una plantación, la superficie de una hoja, etc. La textura provee información importante de la organización estructural de superficies y la relación con su ambiente que lo rodea. Las características Haralick (Haralick, 1973) son obtenidas de la matriz de co-ocurrencia del nivel de gris (GLCM, por sus siglas en inglés). Se calcularon 14 características de Haralick, entre las que se encuentran: el segundo momento angular, contraste, correlación, suma de cuadrados, momento de diferencia inversa, suma promedio, suma entropía, suma de varianzas, entropía, diferencia de varianzas, diferencia de entropías, información de medidas de correlación y coeficiente máxima de correlación.

3.5. Clasificación

3.5.1. *Support Vector Machine*

Las *SVM* son técnicas muy utilizadas para la clasificación de enfermedades debido a que pueden entregar resultados muy competitivos y con un menor consumo de recursos en comparación con otros algoritmos como las redes neuronales artificiales cuando se trata de conjuntos de datos pequeños (Pujari, 2016). Una de las características fundamentales de las *SVM* es que utilizan *kernels* cuando la clasificación binaria del conjunto de datos no puede ser separado linealmente. Entonces se recurre a los *kernels* para transformar los espacios de entrada en un espacio multi-dimensional donde los conjuntos pueden ser separados linealmente una vez transformados.

Después de realizar pruebas con los diferentes *kernels* de *SVM* disponibles en Weka 3.8. Se decidió utilizar el PolyKernel con el valor de la variable compleja de 2.0. Para la validación cruzada se utilizaron 10 *folds*. El calibrador que se utilizó fue *Logistic*. El número de iteraciones se dejó por defecto (hasta que converja).

3.5.2. *Redes Neuronales Artificiales Backpropagation*

Las redes neuronales artificiales son conjuntos de neuronas artificiales interconectadas que utilizan algoritmos y modelos matemáticos para procesar conjuntos de datos para aprender de ellos y ofrecer una clasificación de acuerdo al objetivo con el que fueron desarrolladas. Una red neuronal artificial comprende a menudo de tres tipos de capa de neuronas: capa de entrada, capa oculta y capa de salida.

Las red neuronal es una red *backpropagation* ajustada a 500 épocas, una tasa de aprendizaje α que es igual a 0.1. Además de la variable *Momentum* que su valor es de 0.1.

3.5.3. *Random Forest*

Es un algoritmo que construye múltiples árboles de decisión. Tiene mejor rendimiento que los árboles individuales y es más eficiente que los algoritmos de clasificación tradicionales como las redes neuronales artificiales, especialmente cuando se trata de conjuntos de datos grandes (XinHai, 2013). Random forest puede manejar miles de variables explicativas o independientes y puede modelar interacciones complejas entre variables.

En este estudio, *Random Forest* utilizó 200 iteraciones. El valor de la variable *depth* se fijo en 0, lo cual quiere decir que el algoritmo automáticamente controlará la profundidad del árbol.

3.5.4. *Naive Bayes*

El clasificador Naive Bayes está basado en la suposición de que los valores de los atributos son condicionalmente independientes dado el valor objetivo.

Naive Bayes ignora las posibles dependencias o correlaciones entre las entradas y reduce un problema multivariable a un grupo de problemas univariable.

Capítulo 4

Resultados experimentales

En esta sección se muestran los resultados de las pruebas experimentales realizadas con weka 3.8 para el conjunto de imágenes de hojas sanas y enfermas. Se utilizaron 4 algoritmos de clasificación para comparar sus resultados: SVM con el algoritmo de Optimización Mínima Secuencial(SMO), Redes Neuronales *Backpropagation*, Random Forest y Naive Bayes.

4.1. Variables Estadísticas del clasificador

4.1.1. Verdadero positivo

Un verdadero positivo en el resultado de una medición o clasificación donde se concluye que el elemento o imagen pertenece a una clase y ha sido correctamente clasificado.

4.1.2. Verdadero negativo

Un verdadero negativo en el resultado de una medición o clasificación donde se concluye que el elemento o imagen no pertenece a una clase y ha sido correctamente clasificado.

4.1.3. Falso negativo

Un falso negativo en el resultado de una medición o clasificación donde se concluye que el elemento o imagen no pertenece a una clase determinada cuando en realidad si pertenece a dicha clase. Por lo tanto es una clasificación errónea.

4.1.4. Falso positivo

Un falso positivo en el resultado de una medición o clasificación donde se concluye que el elemento o imagen pertenece a una clase determinada cuando en realidad no pertenece a dicha clase. Por lo tanto es una clasificación errónea.



		Clasificación de hojas de café	
		Positivos	Negativos
Condición	+	Verdaderos positivos 	Falsos Negativos 
	-	Falsos positivos 	Verdaderos negativos 

FIGURA 4.1: Clasificación de hojas de café

En la figura se muestra un recuadro de las variables que se acaban de describir con un ejemplo de clasificación de hojas de café. Las hojas de café clasificadas correctamente aparecen en el recuadro de verdaderos positivos; las hojas de frambuesa clasificadas correctamente aparecen en el recuadro de verdaderos negativos. Por otro lado, las hojas de café que fueron clasificadas de manera incorrecta aparecen en el recuadro de falsos negativos y las hojas de frambuesa que fueron clasificadas incorrectamente aparecen en el recuadro de falsos positivos.

4.1.5. Sensibilidad

La sensibilidad indica la capacidad del clasificador de dar como positivos los elementos que realmente lo son. También es conocida como Tasa de Verdaderos Positivos.

$$\text{Sensibilidad} = \frac{VP}{VP + FN} \quad (4.1)$$

Donde VP son los verdaderos positivos y FN son los falsos negativos. En el caso de las hojas, los verdaderos positivos son las hojas que están realmente enfermas.

4.1.6. Especificidad

La especificidad indica la capacidad del clasificador de dar como negativos los elementos que realmente lo son. También es conocida como Tasa de Verdaderos Negativos.

$$\text{Especificidad} = \frac{VN}{VN + FP} \quad (4.2)$$

Donde FP son los falsos positivos y VN los verdaderos negativos. En el caso de las hojas, los verdaderos negativos son las hojas que están realmente sanas.

4.1.7. Tasa de falsos positivos(False Positives Rate en inglés)

Es la proporción de casos negativos que la prueba detecta como positivos.

$$FPR = \frac{FP}{FP + VN} \quad (4.3)$$

Donde FP son los falsos positivos y VN los verdaderos negativos.

4.1.8. Precisión

Se refiere a la dispersión del conjunto de valores obtenidos de mediciones repetidas de una magnitud, esto quiere decir que cuánto menor es la dispersión mayor es la precisión.

4.1.9. Valor F

Es la medida de precisión que tiene un test.

$$F_1 = 2 \cdot \frac{\text{Precisión} \cdot \text{Exhaustividad}}{\text{Precisión} + \text{Exhaustividad}} \quad (4.4)$$

4.1.10. Curva ROC

Proporciona una representación de la sensibilidad y especificidad para un umbral definido. De este modo, permite comparar el rendimiento del clasificador en función de su capacidad discriminante.

Para representar la curva ROC, se colocan valores del 0 al 1 en eje de la x para la especificidad y valores del 0 al 1 en el eje de la y para la sensibilidad. Se traza una línea llamada línea de no-discriminación que va del punto 0,0 al punto 1,1. Esta línea representa el umbral del clasificador donde este no aporta mayor información ya que las poblaciones de positivos y negativos están traslapados. Si la curva ROC del clasificador está al mismo nivel que la línea de discriminación, se dice que el clasificador tiene un valor de área bajo la curva de (AUC, en inglés) de 0.5.

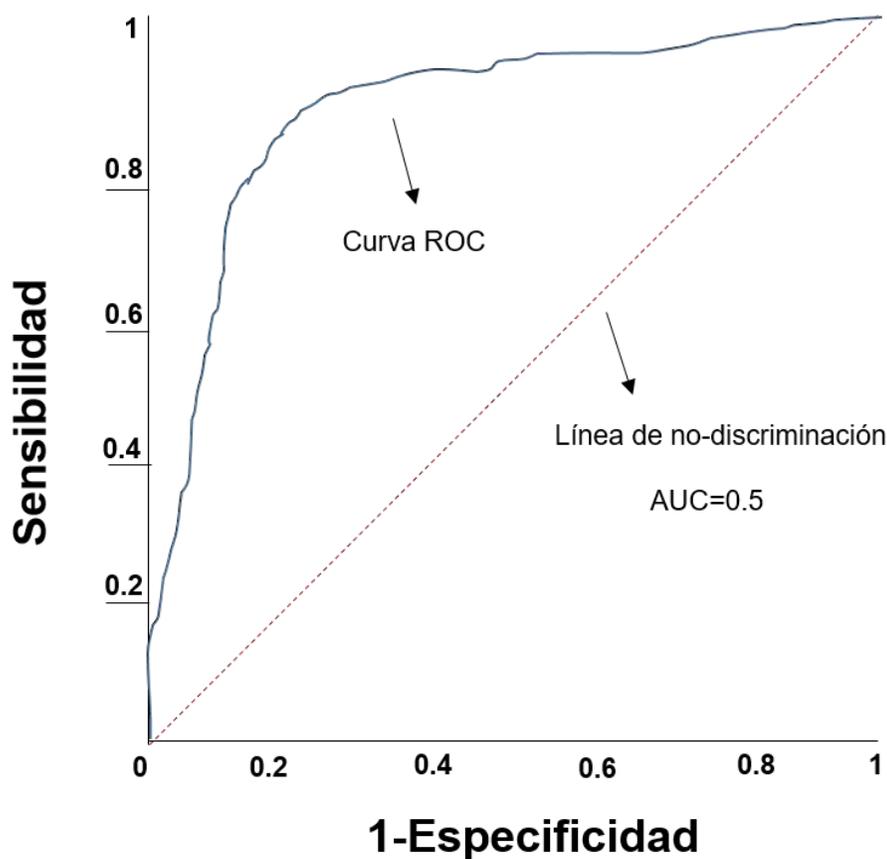


FIGURA 4.2: Curva ROC y línea de no-discriminación

4.2. Desempeño de clasificadores utilizando características cromáticas

En esta Sección se muestran los resultados obtenidos con el conjunto de imágenes de hojas utilizando únicamente las características cromáticas que se segmentaron con

PCA.

A continuación se pueden observar los resultados que fueron obtenidos por los cuatro diferentes algoritmos para las características cromáticas. Se puede ver que la sensibilidad para este tipo de características es buena, y un poco baja con el algoritmo de Naive Bayes que tiene una sensibilidad del 64 %. Es importante remarcar que la tasa de falsos positivos es baja con todos los algoritmos exceptuando a Naive Bayes que presenta una tasa más alta con respecto a los demás algoritmos. Las curvas ROC para los cuatro algoritmos es buena, ya que todos están por arriba de la línea de no-discriminación.

Parámetros	Redes Neuronales	Naive Bayes	Random Forest	SVM
Sensibilidad	0.83	0.648	0.794	0.828
FP Rate	0.059	0.139	0.092	0.078
Precisión	0.82	0.652	0.775	0.812
Valor F	0.831	0.635	0.773	0.814
Roc Area	0.940	0.839	0.936	0.895

4.2.1. Matrices de confusión - características cromáticas

Para las matrices de confusión, se agregaron las siguientes abreviaturas para los diferentes algoritmos:

- Algoritmo Support Vector Machine(SMO) es SMO
- Algoritmo Random Forest es RF
- Algoritmo Naive Bayes es NB
- Algoritmo Redes Neuronales Artificiales es RNA

Para las clases, se agregaron las siguientes abreviaturas:

- Clase Sana es S
- Clase Minador del café es M
- Clase Roya del café es R
- Clase Phoma, Quema es P

- Clase Cercospora es C

El algoritmo SMO clasifica la mayoría de las instancias de manera correcta con todas las clases. Pero tiene su mejor desempeño a la hora de clasificar la enfermedad phoma, quema.

Por su parte, Random Forest clasifica de manera correcta la mayoría de las instancias de todas las clases, exceptuando en la cercospora, es ahí donde se confunde con todas las clases menos con la clase sana.

El algoritmo Naive Bayes clasifica muy bien las instancias de la clase sana y empieza a tener problemas con la clase minador de la hoja, la clase Roya del café, la clase Phoma y termina teniendo un desempeño bajo con la clase cercospora.

El algoritmo de redes neuronales artificiales clasifica de manera correcta el 75% de las instancias de la clase sana y el otro 25% lo confunde con la clase minador de la hoja. Se desempeña bien clasificando el minador de la hoja, la roya del café, phoma quema y empieza a tener problemas con la clase cercospora.

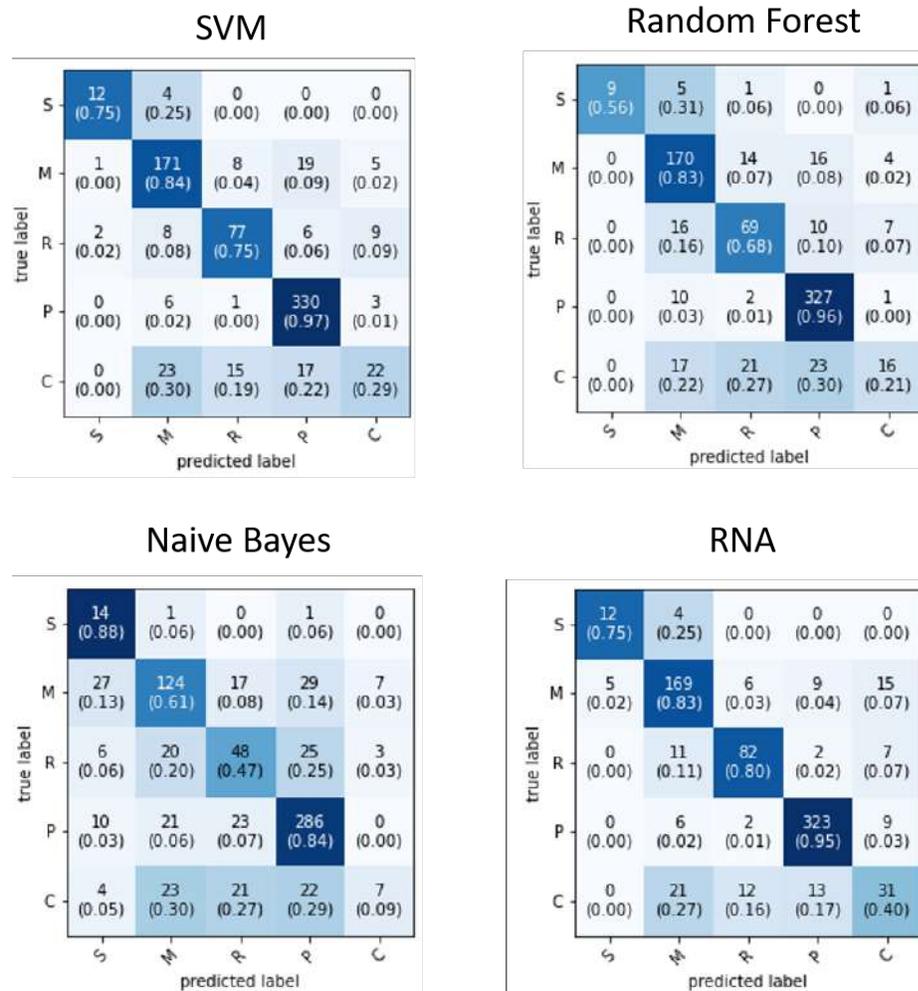


FIGURA 4.3: Matriz de confusión características cromáticas

4.3. Desempeño de clasificadores utilizando características geométricas

En esta Sección se muestran los resultados obtenidos con el conjunto de imágenes de hojas utilizando únicamente las características geométricas que se segmentaron con PCA.

Los cuatro algoritmos utilizados en este trabajo tienen un desempeño más modesto con las características geométricas, ya que la sensibilidad más alta con las características geométricas es de 73% que lo obtiene Random Forest. Las características cromáticas tuvieron una sensibilidad máxima de 83%.

Parámetros	Redes Neuronales	Naive Bayes	Random Forest	SVM
Sensibilidad	0.708	0.430	0.732	0.714
FP Rate	0.123	0.107	0.128	0.151
Precisión	0.685	0.639	0.703	0.683
Valor F	0.694	0.461	0.707	0.680
Roc Area	0.872	0.788	0.898	0.799

4.3.1. Matrices de confusión - características geométricas

El algoritmo SVM(SMO) tiene problemas para distinguir entre la clase sana y la clase minador de la hoja. Además de confundir en gran medida la clase minador con la clase roya del café. La cercospora es confundida con todas las demás enfermedades menos la clase sana.

El algoritmo Random Forest clasifica correctamente la mayoría de las instancias de las clase sana, la clase minador del café y la clase phoma.

El algoritmo Naive Bayes se desempeña bien identificando la clase sana. Al clasificar la clase minador de la hoja, el algoritmo lo confunde con la cercospora. Al clasificar la clase roya del café, el algoritmo lo confunde en gran medida con la clase cercospora. Al clasificar la clase Phoma, Naive Bayes encuentra muchas similitudes con la cercospora. Después, sucede algo peculiar ya que Naive Bayes parece tener un buen desempeño al clasificar la cercospora.

Las redes neuronales artificiales clasifican bien la mayoría de las instancias en todas las clases pero si se trata de la clase cercospora es ahí cuando tiene problemas para distinguirla de las otras enfermedades.

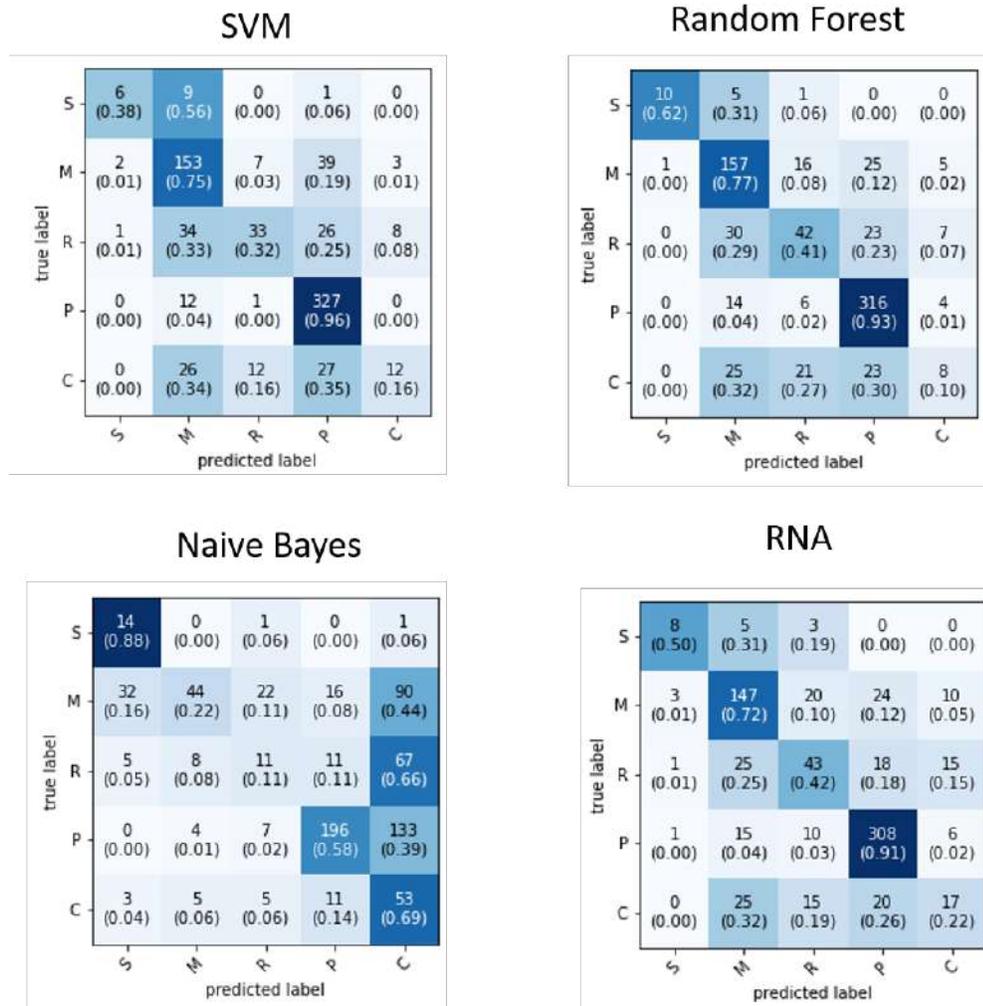


FIGURA 4.4: Matriz de confusión características geométricas

4.4. Desempeño de clasificadores utilizando características texturales

En esta Sección se muestran los resultados obtenidos con el conjunto de imágenes de hojas utilizando únicamente las características texturales que se segmentaron con PCA.

Para las características texturales se tienen resultados de sensibilidad promediando el 72%. Las redes neuronales artificiales, Random Forest y SVM(SMO) tienen resultados de sensibilidad parecidos. Al algoritmo Naive Bayes le sucede lo contrario ya que

termina con 62 % de sensibilidad. En general, los algoritmos tienen bajas tasas de falsos positivos y las curvas ROC indican un buen desempeño al clasificar.

Parámetros	Redes Neuronales	Naive Bayes	Random Forest	SVM
Sensibilidad	0.774	0.622	0.752	0.760
FP Rate	0.078	0.100	0.099	0.102
Precisión	0.763	0.655	0.736	0.728
Valor F	0.766	0.632	0.735	0.724
Roc Area	0.916	0.862	0.924	0.857

4.4.1. Matrices de confusión - características texturales

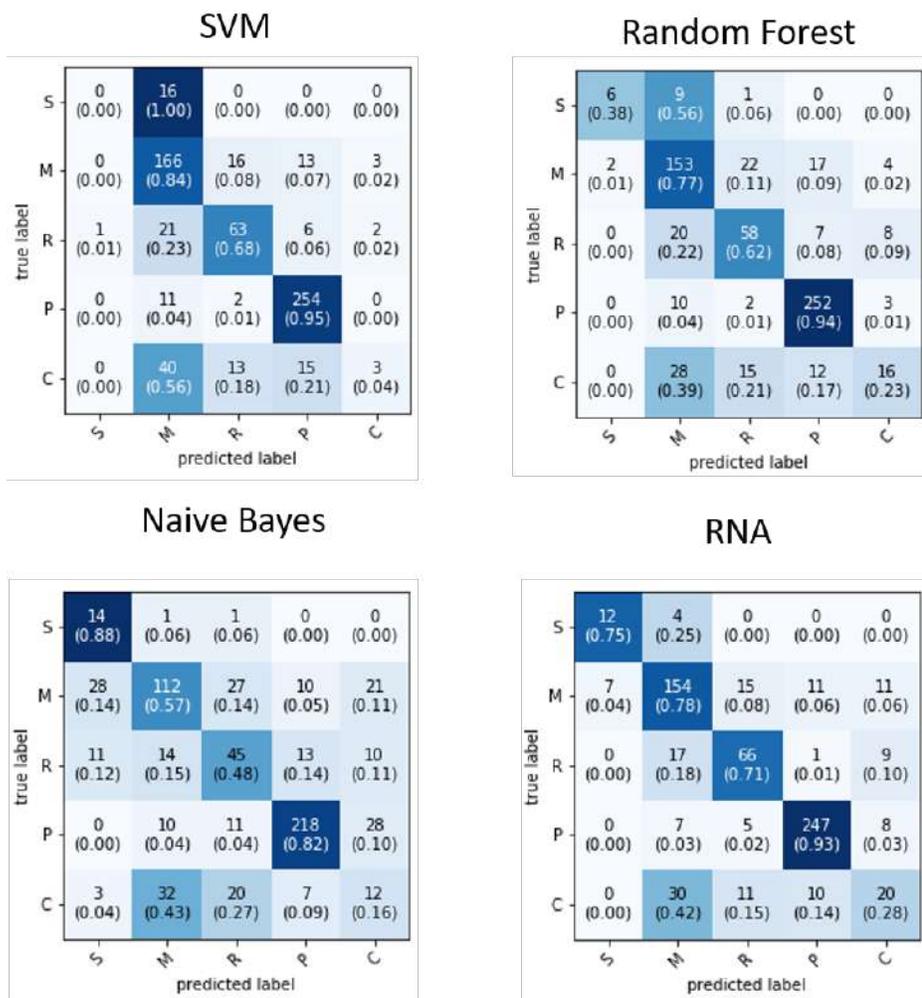
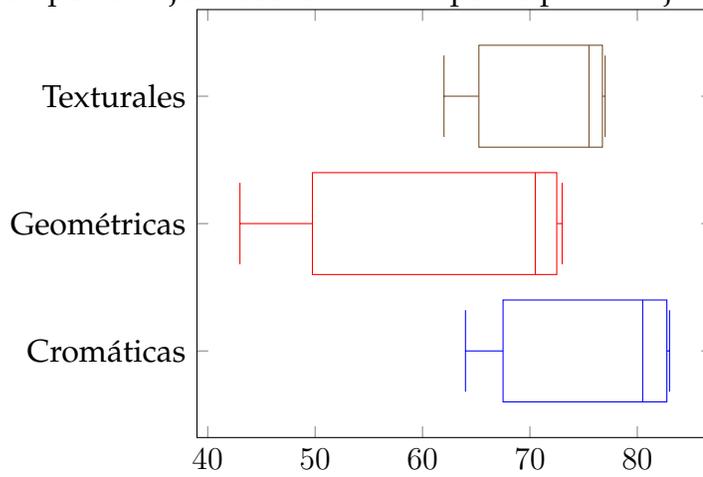


FIGURA 4.5: Matriz de confusión características texturales

4.4.2. Diagramas de caja de la sensibilidad de los clasificadores por tipo de características

Los diagramas de caja que se muestran permiten analizar cómo se comporta la sensibilidad por cada tipo de características siendo las cromáticas las que tienen una mediana del 80%. Las características geométricas parecen tener el peor desempeño dado que Naive Bayes entregó resultados muy pobres (43%). Las características texturales tienen porcentajes de sensibilidad apenas por debajo de las características cromáticas.



4.5. Desempeño de clasificadores utilizando características texturales y cromáticas

En esta Sección se muestran los resultados obtenidos con el conjunto de imágenes de hojas utilizando únicamente las características texturales y cromáticas que se segmentaron con PCA.

Para este tipo de características combinadas, nuevamente los algoritmos Redes Neuronales Artificiales, SVM(SMO) y Random Forest vuelven a tener porcentajes altos y muy parecidos en cuanto a sensibilidad. Naive Bayes vuelve a ser el algoritmo con el desempeño más bajo.

Las curvas ROC indican una buena clasificación ya que todos los algoritmos excepto Naive Bayes se acercan al punto óptimo de clasificación.

Parámetros	Redes Neuronales	Naive Bayes	Random Forest	SVM
Sensibilidad	0.791	0.650	0.771	0.802
FP Rate	0.071	0.097	0.098	0.080
Precisión	0.790	0.685	0.750	0.791
Valor F	0.790	0.660	0.748	0.793
Roc Area	0.937	0.876	0.936	0.899

4.5.1. Matrices de confusión - características texturales y cromáticas

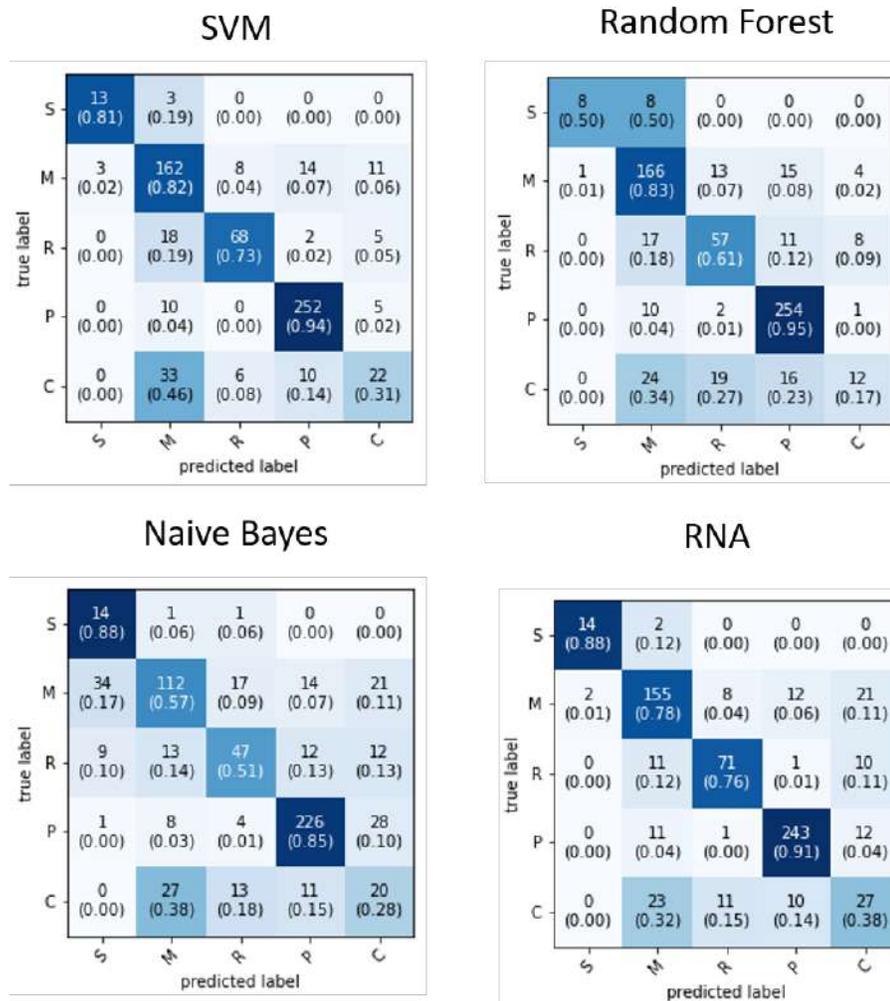


FIGURA 4.6: Matriz de confusión características cromáticas y texturales

4.6. Desempeño de clasificadores utilizando características texturales y geométricas

En esta Sección se muestran los resultados de precisión obtenidos en el conjunto de imágenes de hojas utilizando únicamente las características texturales y geométricas que se segmentaron con PCA.

Parámetros	Redes Neuronales	Naive Bayes	Random Forest	SVM
Sensibilidad	0.736	0.629	0.753	0.743
FP Rate	0.094	0.146	0.106	0.107
Precisión	0.725	0.615	0.733	0.704
Valor F	0.729	0.601	0.726	0.713
Roc Area	0.915	0.844	0.927	0.859

4.6.1. Matrices de confusión - características texturales y geométricas

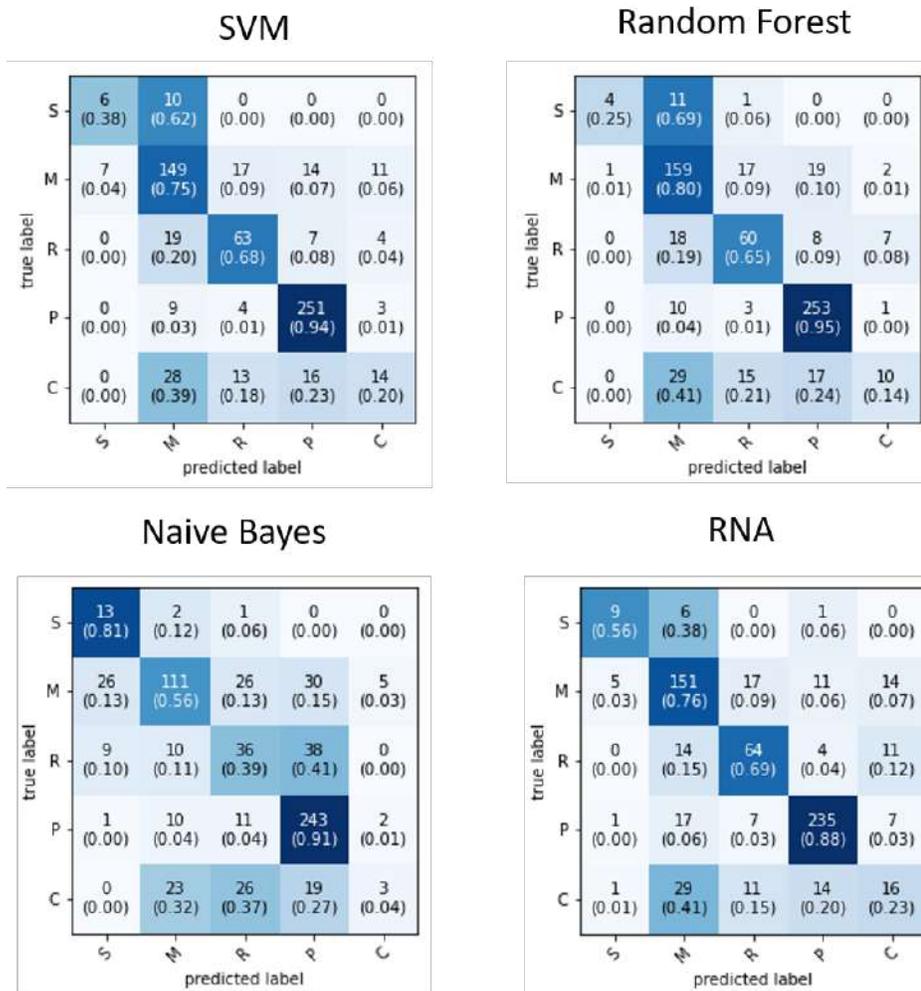


FIGURA 4.7: Matriz de confusión características geométricas y texturales

4.7. Desempeño de clasificadores utilizando características cromáticas y geométricas

En esta Sección se muestran los resultados de precisión obtenidos en el conjunto de imágenes de hojas utilizando unicamente las características cromaticas y geometricas que se segmentaron con PCA.

4.7. Desempeño de clasificadores utilizando características cromáticas y geométrica 71

Parámetros	Redes Neuronales	Naive Bayes	Random Forest	SVM
Sensibilidad	0.813	0.655	0.806	0.829
FP Rate	0.065	0.125	0.084	0.066
Precisión	0.800	0.687	0.787	0.818
Valor F	0.806	0.647	0.787	0.822
Roc Area	0.944	0.861	0.829	0.905

4.7.1. Matrices de confusión - características cromáticas y geométricas

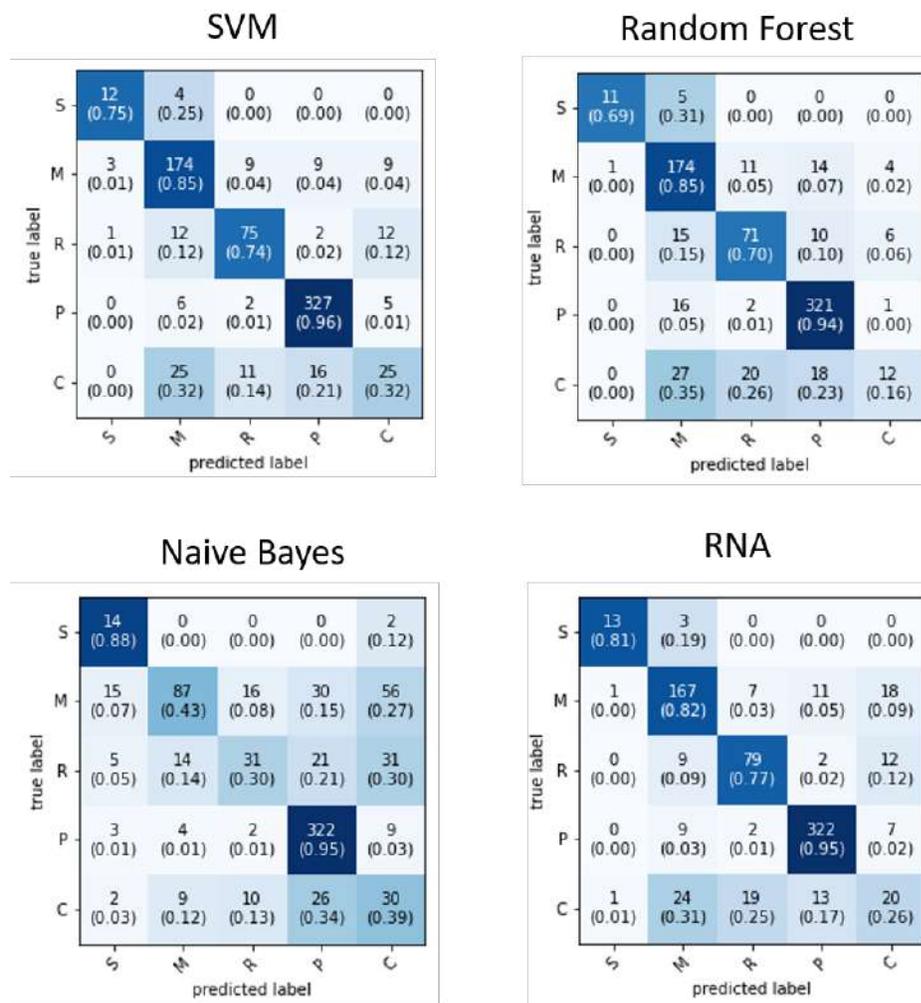


FIGURA 4.8: Matriz de confusión características geométricas y texturales

4.8. Desempeño de clasificadores utilizando características cromáticas, geométricas y texturales

En esta Sección se muestran los resultados de precisión obtenidos en el conjunto de imágenes de hojas utilizando las características cromáticas, geométricas y texturales que se segmentaron con PCA.

Parámetros	Redes Neuronales	Naive Bayes	Random Forest	SVM
Sensibilidad	0.769	0.651	0.772	0.802
FP Rate	0.079	0.094	0.099	0.075
Precisión	0.762	0.693	0.758	0.793
Valor F	0.765	0.649	0.752	0.796
Roc Area	0.925	0.859	0.934	0.896

4.8.1. Matrices de confusión - características cromáticas, geométricas y texturales

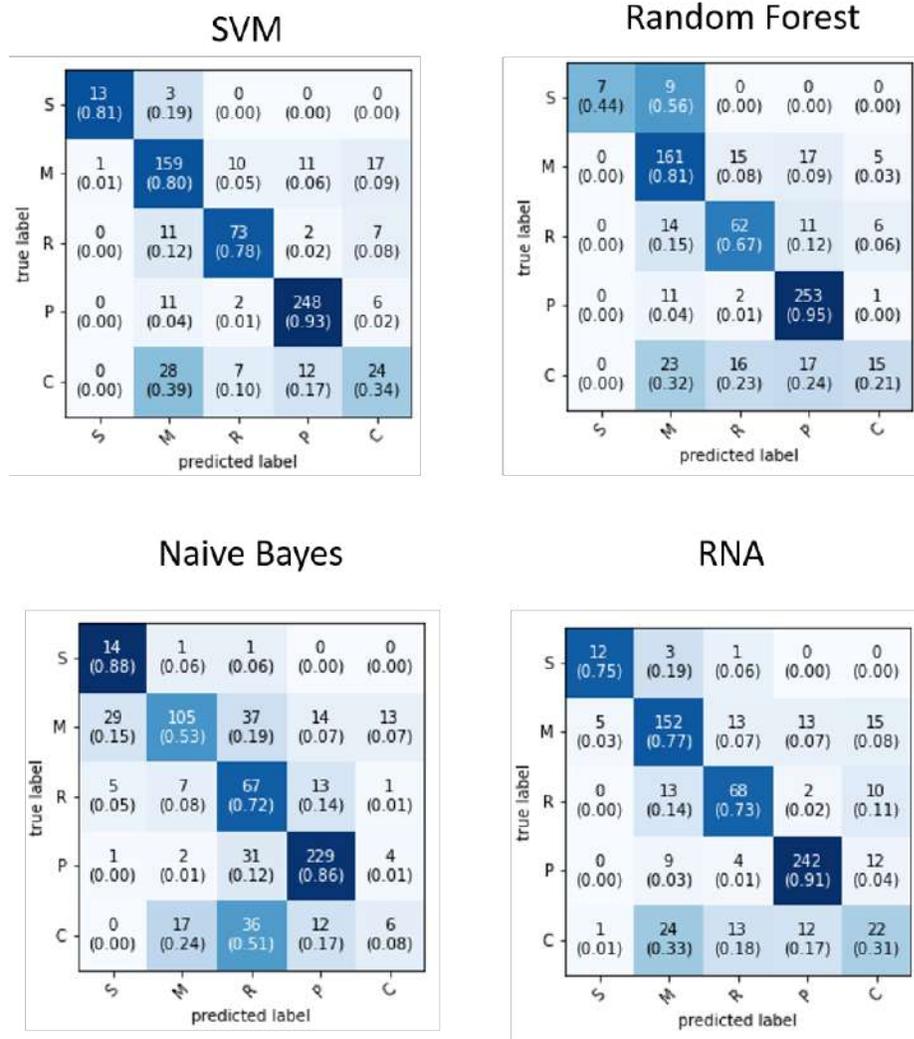
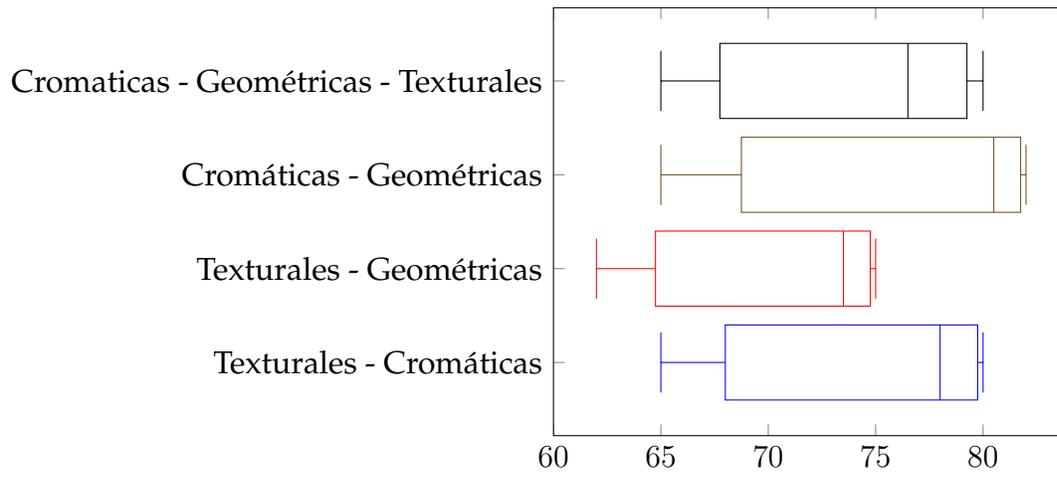


FIGURA 4.9: Matriz de confusión características geométricas y texturales

4.8.2. Diagramas de caja de la precisión de los clasificadores



Capítulo 5

Conclusiones

En este trabajo se probaron cuatro métodos de clasificación: *Naive Bayes*, *Random Forest*, Redes Neuronales Artificiales *Backpropagation* y *Support Vector Machine* con el algoritmo *SMO*. En términos generales, tanto SVM como *backpropagation* obtuvieron porcentajes que rondan el 80 % de sensibilidad.

Este trabajo muestra resultados competitivos ya que se obtiene un promedio de sensibilidad de todos los algoritmos de 72 %. Esto quiere decir que la mayoría de los algoritmos entregaron resultados satisfactorios, a excepción de *Naive Bayes* que obtuvo resultados pobres en casi todos los tipos de características. Como menciona (Barbedo, 2019), la mayoría de las veces, la buena precisión de los sistemas de reconocimiento patrones en plantas está altamente ligada a conjuntos de datos variados y con alta representatividad. El reto es obtener o mejorar el conjunto de datos para dar mayor representatividad a cada una de las clases y, por ende, mejorar la clasificación.

Como trabajo futuro, este experimento pretende ser desarrollado como una herramienta práctica y accesible ya sea por medio de una aplicación móvil o un programa lo cual representaría una gran ayuda para fitopatólogos y agrónomos para la identificación de enfermedades en la planta del café de la especie *coffea arabica*. Además del desarrollo de la herramienta de fácil acceso, se tiene pensado el desarrollo de una función para la estimación del daño causado por alguna de las enfermedades que se estudian en este trabajo para determinar en qué momento es accionable para aplicar mecanismos de control (productos químicos para atacar a una enfermedad en particular).

Bibliografía

- [1] Wang, S., Yang, D. M., Rong, R., Zhan, X., Fujimoto, J., Liu, H., Minna, J., Wistuba, I. I., Xie, Y., Xiao, G. (2019). Artificial Intelligence in Lung Cancer Pathology Image Analysis. *Cancers*, 11(11), 1673. <https://doi.org/10.3390/cancers11111673>
- [2] Zhao, C., Chan, S. S., Cham, W.-K. Chu, L. (2015). Plant identification using leaf shapes—A pattern counting approach. *Pattern Recognition*, 2015, 48, 3203-3215
- [3] Jain, S., Jagtap, V. Pise, N. (2015). Computer Aided Melanoma Skin Cancer Detection Using Image Processing. *Procedia Computer Science*, 2015, 48, 735-740
- [4] Y. Ma, Z. Wang, H. Yang L. Yang, (2020). Artificial intelligence applications in the development of autonomous vehicles: a survey. *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 2, pp. 315-329,
- [5] Sharma, O. P., Ghose, M. K., Shah, K. B., Thakur, B. K. (2013). Recent trends and tools for feature extraction in OCR technology. *International Journal of Soft Computing and Engineering*, 2(6), 220-223.
- [6] Läderach, P., Hagggar, J., Lau, C., Eitzinger, A., Ovalle, O., Baca, M., Jarvis, A. Lundy, M. (2011). *Café mesoamericano: Desarrollo de una estrategia de adaptación al cambio climático. CIAT Políticas en Síntesis no. 2. Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia. 4 p*
- [7] Anderson, E., Cherrington, E., Flores, A., Perez, J., Carrillo, R. Sempris, E. (2008). *Potential Impacts of Climate Change on Biodiversity in Central America, Mexico, and the Dominican Republic 2008*
- [8] Medina, M. J. A., Ruiz, N. R. E., Gómez, C. J. C., Sánchez, Y. J. M., Gómez, A. G. Pinto, M. O. (2016). Estudio del sistema de producción de café (*Coffea arabica* L.) en la región Frailesca, Chiapas. *CienciaUAT*. 10(2):33-43.

- [9] Miguel F. Monroig. Morfología del cafeto. En línea. <http://academic.uprm.edu/mmonroig/id53.htm>
- [10] Escamilla E. P. (2016). Las variedades de café en México ante el desafío de la roya. Breves de Políticas Públicas. Boletín Informativo [en línea]
- [11] Krohling, R. ; Esgario, J. ; Ventura, Jose A. (2019). BRACOL - A Brazilian Arabica Coffee Leaf images dataset to identification and quantification of coffee diseases and pests. Mendeley Data, V1, doi: 10.17632/yy2k5y8mxg.1
- [12] Berlanga, H. (2011). Los productores de café en México: problemática y ejercicio del presupuesto. Washington DC: Woodrow Wilson International Center for Scholars. Mexican Rural Development Research Reports, 14.
- [13] Manso, G.L., Knidel, H., Krohling, R., Ventura, J. (2019). A smartphone application to detection and classification of coffee leaf miner and coffee leaf rust. ArXiv, abs/1904.00742.
- [14] Sampallo, G. (2003). Reconocimiento de tipos de hojas. Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial, 2003
- [15] Barbedo, J. G. A. (2019). Plant disease identification from individual lesions and spots using deep learning. Biosystems Engineering, Elsevier BV, 2019, 180, 96-107
- [16] Bhange, M. Hingoliwala, H. (2015). Smart Farming: Pomegranate Disease Detection Using Image Processing. Procedia Computer Science, Elsevier BV, 2015, 58, 280-288
- [17] Gavhale, K. R., Gawande, U. Hajari, K. O. (2014). Unhealthy region of citrus leaf detection using image processing techniques. International Conference for Convergence for Technology-2014, IEEE, 2014
- [18] Devashish Pujari, R. Yakkundimath, A. S. Byadgi IJIMAI, ISSN-e 1989-1660, Vol. 3, N°. 7, 2016, págs. 6-14
- [19] Mohanty, S. P., Hughes, D. P. Salathé, M. (2016). Using Deep Learning for Image-Based Plant Disease Detection. Frontiers in Plant Science, 2016, 7, 1419
- [20] Arnal Barbedo, J. G. (2019). Plant disease identification from individual lesions and spots using deep learning. Biosystems Engineering, 2019, 180, 96-107

- [21] M. Dhakate Ingole A. B. (2015). Diagnosis of pomegranate plant diseases using neural network. 2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), Patna, India, 2015, pp. 1-4, doi: 10.1109/NCVPRIPG.2015.7490056.
- [22] Ramakrishnan M. Sahaya Anselin Nisha A. (2015). Groundnut leaf disease detection and classification by using backpropagation algorithm 2015 International Conference on Communications and Signal Processing (ICCSP), Melmaruvathur, India, 2015, pp. 0964-0968, doi: 10.1109/ICCSP.2015.7322641.
- [23] Arakeri, M. Lakshmana (2016). Computer Vision Based Fruit Grading System for Quality Evaluation of Tomato in Agriculture industry *Procedia Computer Science*, 2016, 79, 426-433
- [24] Fujita, H., Zhang, X., Kido, S., Hara, T., Zhou, X., Hatanaka, Y., Xu, R. (2010). An introduction and survey of computer-aided detection/diagnosis (CAD). In *Proceedings of the 2010 IEEE International Conference on Future Computer, Control and Communication* (pp. 200-205).
- [25] Tulshan, A. S., Dhage, S. N. Thampi, S. M., Marques, O., Krishnan, S., Li, K.-C., Ciunzo, D. Kolekar, M. H. (2019) *Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa Advances in Signal Processing and Intelligent Recognition Systems*, Springer Singapore, 2019, 190-201
- [26] Islam, N., Islam, Z. Noor, N. (2017). A Survey on Optical Character Recognition System. *ArXiv* 2017
- [27] Bhardwaj, S. Mittal, A. (2012). A Survey on Various Edge Detector Techniques *Procedia Technology*, 2012, 4, 220-226
- [28] N. Otsu. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, vol.9, no.1, pp.62-66, Jan.1979, doi:10.1109/TSMC.1979.4310076.
- [29] Ming-Kuei Hu. (1962). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179-187, February 1962, doi: 10.1109/TIT.1962.1057692.
- [30] Zhu, Q., Yeh, M., Cheng, K., Avidan, S. (2006). Fast Human Detection Using a Cascade of Histograms of Oriented Gradients. 2006 IEEE Computer Society

Conference on Computer Vision and Pattern Recognition (CVPR'06), 2, 1491-1498.

- [31] Cervantes, J., Taltempa, J., García Lamont, F., Ruiz Castilla, J., Yee Rendon, A., Jalili, L. (2017). Análisis Comparativo de las técnicas utilizadas en un Sistema de Reconocimiento de Hojas de Planta. *Revista Iberoamericana de Automática e Informática industrial*, 14(1), 104-114. doi:<https://doi.org/10.1016/j.riai.2016.09.005>
- [32] F. Rosenblatt (1960) Perceptron Simulation Experiments. in *Proceedings of the IRE*, vol. 48, no. 3, pp. 301-309, March 1960, doi: 10.1109/JRPROC.1960.287598.
- [33] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417-441. <https://doi.org/10.1037/h0071325>
- [34] Flusser, J., Zitova, B., Suk, T. (2009). *Moments and moment invariants in pattern recognition*. John Wiley Sons.
- [35] D. Gabor (1973). Theory of communication. *Journal of Institution of Electrical Engineers*, vol. 93, pp. 429-457, 1946
- [36] R. M. Haralick, K. Shanmugam I. Dinstein (1973). Textural Features for Image Classification. in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621, Nov. 1973, doi: 10.1109/TSMC.1973.4309314.
- [37] Alireza Norouzi, Mohd Shafry Mohd Rahim, Ayman Altameem, Tanzila Saba, Abdolvahab Ehsani Rad, Amjad Rehman Mueen Uddin (2014) *Medical Image Segmentation Methods, Algorithms, and Applications*. *IETE Technical Review*, 31:3, 199-213, DOI: 10.1080/02564602.2014.906861