



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO



DESARROLLO DE UN SISTEMA DE IDENTIFICACIÓN
Y RECONOCIMIENTO DE PLANTAS USANDO
TÉCNICAS DE APRENDIZAJE MÁQUINA

T E S I S

QUE PARA OBTENER EL GRADO DE:
INGENIERA EN COMPUTACIÓN

PRESENTA:

LAURA YADIRA DOMINGUEZ JALILI

DIRECTOR DE TESIS:

DR. EN C. JAIR CERVANTES CANALES

JULIO, 2015

M. en C. E. VIRIDIANA BANDA ARZATE
SUBDIRECTOR ACADEMICO DEL
CENTRO UNIVERSITARIO UAEM TEXCOCO.
PRESENTE:

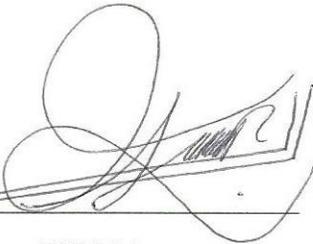
AT'N L. EN I.A. CINTHYA TERESITA ISLAS RODRÍGUEZ
RESPONSABLE DEL DEPARTAMENTO DE TITULACION.

Con base en las revisiones efectuadas al trabajo escrito titulado "Desarrollo de un sistema de identificación y reconocimiento de plantas usando técnicas de aprendizaje máquina" que para obtener el título de Licenciado en Ingeniería en Computación presenta la sustentante Domínguez Jalili Laura Yadira, con número de cuenta 0921028 respectivamente, se concluye que cumple con los requisitos teórico-metodológicos por lo que se le otorga el voto aprobatorio para su sustentación, pudiendo continuar con la etapa de digitalización del trabajo escrito.

Atentamente



FIRMA
DR. EN C.C. FARID GARCÍA
LAMONT

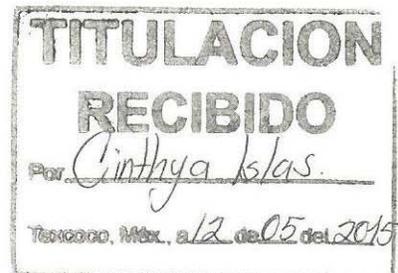


FIRMA
M. EN C.A. JOSÉ SERGIO RUIZ
CASTILLA



DR. EN C. JAIR CERVANTES CANALES

c.c.p. LAURA YADIRA DOMÍNGUEZ JALILI
c.c.p. DR. EN D. RICARDO COLÍN GARCÍA
c.c.p. M. EN P.P. ANTONIO INOUE CERVANTES



Desarrollo de un Sistema de Identificación y Reconocimiento de Plantas Usando Técnicas de Aprendizaje Máquina

por

Laura Yadira Dominguez Jalili

Tesis presentada para obtener el grado de

Ingeniera en Computación

en la

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO

. Julio, 2015

Índice de figuras

3-1. Operador cruzados de Roberts	31
4-1. SVM	46
5-1. Ejemplo de códigos de cadena	54
5-2. Transformada de Hough	56
6-1. Metodología propuesta	80
6-2. Conjunto de hojas utilizado	81
6-3. Técnicas de Segmentación	83
7-1. Subconjuntos de especies de plantas utilizados.	87
7-2. Subconjuntos de especies de plantas utilizados.	88

Índice de tablas

2-1. Tipos de ecualización	20
7-1. Resultados de clasificación de características Cromáticas, Texturales y Geométricas por separado	90
7-2. Resultados de clasificación usando características Cromáticas-Texturales, Cromáticas-Geométricas, Texturales-Geométricas	91
7-3. Resultados de clasificación con las características Cromáticas, Texturales y Geométricas.	91

Agradecimientos

Saludos! Les agradezco a mis revisores al Dr. en. C.C Farid y al M. en C.A. Sergio por su apoyo y dedicación para la revisión de mi tesis. También le quiero agradecer a mi director de tesis, el Dr. en. C. Jair por su tolerancia, comprensión, apoyo y dedicación para que se llevara a cabo esta tesis, que sin duda sin su apoyo esta tesis no sería un éxito. Además deseo agradecerles por su amistad que me brindaron siempre.

Primeramente deseo agradecerles a mis padres, mi abuelita, mis hermanas por todo su apoyo indispensable y amor que siempre me han brindado durante mi desarrollo académico y a mis sobrinitos Johan, Denisse y el que viene en camino que también forman parte de mi motivación personal y profesional.

Índice general

1. Introducción	1
1.1. Planteamiento del problema	3
1.2. Justificación	4
1.3. Objetivos y alcances del proyecto	6
1.3.1. Objetivo General	6
1.3.2. Objetivos específicos	6
1.4. Hipótesis	6
1.5. Estado del arte	7
2. Procesamiento Digital de Imágenes	11
2.1. Operaciones Puntuales	11
2.1.1. Operador Identidad	12
2.1.2. Operador Inverso o Negativo	12
2.1.3. Operador Umbral	12
2.1.4. Operador Intervalo de Umbral Binario	12
2.1.5. Sustracción de la media.	12
2.1.6. Filtros basados en las derivadas.	13
2.1.7. Operador Intervalo de Umbral Binario Invertido	13
2.1.8. Operador Umbral de la Escala de Grises	13
2.1.9. Operador Umbral de la Escala de Grises Invertido	14
2.1.10. Operador Extensión	14
2.2. Operadores aritméticos	14

2.2.1.	Suma	14
2.2.2.	Resta	15
2.3.	Operaciones grupales	15
2.3.1.	Filtro de paso bajo.	15
2.3.2.	Filtro de la Media	16
2.3.3.	Filtro de la Media Ponderada	16
2.3.4.	Filtro de la Mediana	16
2.3.5.	Filtro Gaussiano	16
2.3.6.	Filtros Pasa Altos	17
2.3.7.	Transformaciones de Intensidad	17
2.3.8.	Transformaciones lineales	17
2.3.9.	Transformaciones no lineales	17
2.3.10.	Histogramas	18
2.3.11.	Expansión del histograma	18
2.3.12.	Contracción del histograma	19
2.3.13.	Desplazamiento del histograma	19
2.3.14.	Ecualización de Histograma	19
2.3.15.	Normalización de Histograma	20
2.3.16.	Expansión de Contraste	20
2.3.17.	Contraste	20
2.3.18.	Brillo	21
2.3.19.	Interpolación	21
2.3.20.	Interpolación por vecino más próximo	21
2.3.21.	Interpolación bilineal	22
2.3.22.	Interpolación bicúbica	22
2.3.23.	Detección de bordes, algoritmo de Canny	23
3.	Segmentación	27
3.1.	Métodos basados en detección de bordes	27
3.1.1.	Operador de Sobel	27
3.1.2.	Operador de Prewitt	29

3.1.3. Operador de Roberts	30
3.2. Método de Otsu	32
3.3. Segmentación PCA	33
4. Clasificadores	36
4.1. Redes neuronales	37
4.2. El Perceptrón	41
4.3. BackPropagation	43
4.4. Máquinas de Vectores de Soporte (SVM)	44
4.4.1. Fase de Entrenamiento.	45
4.4.2. Fase de prueba	47
5. Extracción de características	49
5.1. Características geométricas	49
5.2. Características texturales	65
5.3. Características Cromáticas	75
6. Metodología	79
6.1. Modelo propuesto	80
6.2. Preprocesamiento	80
6.3. Segmentación	82
6.4. Extracción de Características	84
6.4.1. Características Texturales:	84
6.4.2. Características Geométricas:	84
6.4.3. Características Cromáticas:	84
6.5. Clasificación e identificación	85
7. Resultados experimentales	86
7.1. Conjunto de datos utilizados	86
7.2. Validación	89
7.3. Resultados experimentales	89

Desarrollo de un Sistema de Identificación y Reconocimiento de Plantas Usando Técnicas de Aprendizaje Máquina

por

Laura Yadira Dominguez Jalili

Resumen

El reconocimiento de plantas a partir de técnicas de aprendizaje automático es un campo de investigación muy activo en los últimos años. Los algoritmos de detección y reconocimiento de plantas juegan un rol vital en muchas aplicaciones importantes. Estos algoritmos pueden ayudar a agilizar procesos y tareas importantes en aduanas, agricultura, investigación, etc. Los algoritmos más importantes que se han desarrollado utilizan extractores de características a partir de imágenes de hojas, cuyos sistemas de identificación consisten de varios procesos como: pre-procesamiento, segmentación, extracción de características y reconocimiento. Aunque todos los procesos son importantes en el sistema, la extracción de características y el método de clasificación utilizado influyen bastante en una correcta identificación. Esta tesis direcciona dos problemas de investigación importantes en reconocimiento, 1) extracción de características y 2) Clasificación. Los resultados obtenidos son analizados en función del conjunto de técnicas de extracción de características y de los métodos de clasificación utilizados.

Desarrollo de un Sistema de Identificación y Reconocimiento de Plantas Usando Técnicas de Aprendizaje Máquina

by

Laura Yadira Dominguez Jalili

Abstract

In recent years, the use of machine learning techniques to plant recognition is a very active research's field. Detection and recognition algorithms play a vital role in many important applications. These algorithms can help to speed up processes and important tasks like customs, agriculture, research, etc. The most important research on plant identification have been developed using feature extractors from leave's images. These systems are integrated by several processes such as pre-processing, segmentation, feature extraction and recognition. Although all the processes are important in the system, feature extraction and classification have a pretty influence in a correct identification. This thesis addresses two important research problems in recognition, 1) feature extraction and 2) classification. The results are analyzed in light of the feature extraction techniques and classification methods used.

Capítulo 1

Introducción

En los últimos años, el incremento del poder de cómputo y miniaturización de componentes ha permitido la fabricación de cada vez mejores computadores con mayores capacidades de almacenamiento y poder de procesamiento en menores espacios y tiempo. Esto ha permitido el desarrollo de algoritmos que requieren un alto costo computacional. En la actualidad los sistemas de identificación basados en visión son un conjunto enorme de procesos que contienen desde técnicas de preprocesamiento de imágenes, segmentación, extracción de características, identificación, clasificación, y validación entre otras. Estos sistemas requieren de algoritmos muy complejos que sin el poder de cómputo actual sería imposible de implementarlos. Los sistemas de identificación han sido implementados en diversos campos de la ciencia como: medicina

[27],[36],[14],[37],[31],[38], seguridad [5], [24], [18], [6], [21], [10], botánica [2], [12], [16], [15], [1], [72]. En medicina se han desarrollado múltiples sistemas de identificación en diversas áreas como; medicina forense (sistemas de identificación para determinar cadáver reciente y cadáver antiguo), odontología forense (Sistema que identifique los restos óseos y determinar sexo, raza, edad, etc.) oftalmología (Sistema de identificación en enfermedades oculares), Radiología (sistemas de identificación en placas de rayos x para detectar caries, para detectar fisuras craneales, etc.), Alergología (sistemas de identificación de alergias para determinar un análisis de causa), Angiología (sistemas de identificación de problemas en el sistema linfático de las arterias), etc.

Sin embargo otras áreas como agricultura no son la excepción. En agricultura han sido implementados sistemas de visión para identificación de sembradíos [95], [44], identificación de

plagas y enfermedades de plantas [12], [15], [68], [70], sistemas de control de crecimiento de plantas [3], [42], sistemas de visión para identificación automática de tipo de suelo [57], [2], [48], sistemas de control de temperatura en invernaderos, identificación de tipo de planta en base a imágenes de su hoja [40], [32], [74], etc.

En los últimos años se han desarrollado diversos sistemas de identificación de plantas en base a imágenes. Estos explotan diversas características que contienen las hojas como textura, forma geométrica, color y venación.

Los métodos de análisis de textura en el dominio del espacio pueden ser clasificados por diferentes unidades de análisis como el pixel, los elementos de borde y los extremos de los tonos de gris. No existen algoritmos específicos para cada tipo de textura. Algunos algoritmos de clasificación de textura o extracción de esta es mediante algoritmo de Lloyd, Algoritmo *fuzzy clustering*, algoritmos bayesiano (complemento con otro algoritmo), matriz de co-ocurrencia de niveles de gris (matriz que utiliza el pixel como unidad de análisis), operadores de gradiente (utilizados para detectar los bordes), entre otros. Los bordes se caracterizan por su dirección y tamaño y están relacionados con la rugosidad (tosquedad) de la textura y la dirección de textura se mide con el histograma de la dirección del borde.

Para la extracción de características geométricas es muy variable en sus métodos ya que en casi todo se utilizan características geométricas, la altura, lo ancho, radio, rectangularidad, área, perímetro, circularidad, excentricidad, etc. Nos apoyamos en técnicas de distancia y medición para rectificar estas características de la imagen. Algunos algoritmos utilizados son los momentos geométricos de Hu, momentos de Zernike, método híbrido Mean Shift con clasificación a distancia, hacen uso del contorno de la hoja y sus propiedades morfológicas etc.

El color se distingue bajo tres atributos básicos: longitud de onda dominante o cromaticidad, pureza o saturación, brillantez o intensidad. Algunos algoritmos utilizados para extraer las características en los espacios perceptuales de color (HSI) ó en Espacios de escala de cromaticidad uniforme (UCS) son la Transformadas de Fourier, Transformada de Wevelet, White-Patch, Gray-World, Gray-Edges, etc. Algunos espacios de color han sido desarrollados para describir la información de color de forma más parecida a como la percibe el ser humano, se separa la información de intensidad de la cromática, para buscar variaciones locales de color en la imagen.

Las características de nervadura o venación son fáciles de clasificar dado a sus posiciones y

dependen unas de otras, siguiendo un mismo patrón de lineamiento. Es una de las características más tomadas en cuenta para conjuntos grandes de plantas, el clasificar las plantas mediante la nervadura no es lo ideal, ya que son características muy utilizadas pero se deben apoyar en otros parámetros de identificación. Los algoritmos más utilizados para extraer las características de nervadura es el *Unconstrained Hit-or-Miss-Transform* (UHMT) y los descriptores de Fourier con un clasificador de vecino cercano, para la distancia entre las nervaduras [20], [81], [78].

El objetivo de esta tesis, es desarrollar un sistema que obtenga las características más discriminantes ya sean texturales, cromáticas, geométricas o de color. Para ello, se implementaron diferentes técnicas de segmentación y se obtuvo un análisis estadístico de los resultados obtenidos.

1.1. Planteamiento del problema

Los sistemas de identificación actuales identifican a la planta en base a un conjunto de características, generalmente este conjunto de características pertenece a un solo grupo (texturales, cromáticas, geométricas o de venación). Estos sistemas de identificación son eficientes, sin embargo en la literatura actual, no existe un estudio que ilustre la importancia de cada grupo de características y su factor discriminante. Bajo condiciones más generales la eficiencia del clasificador podría reducirse debido a factores como:

1. Las hojas tienen múltiples formas: elíptica, acintada, falciforme, lobulada, lanceolada, acorazonada, escuamiforme, acicular, flabelada, ovada, ovalada, alesnada, acorazonada, sagitada y linear. Sin embargo, existen muchas plantas con clasificación diferente que poseen la misma forma o formas de hoja muy similares.
2. Plantas con nervadura o de venación en las hojas iguales. Las hojas son divididas en tres tipos de venación: la venación paralelinervia, venación peninervia y venación palminervia. La paralelinervia tiene los nervios paralelos unos de otros, la peninervia tienen un nervio

principal del que salen los nervios secundarios y la palminervia tiene un punto central a inicio del tallo de donde se desprenden las líneas de venación.

3. Plantas diferentes con misma textura; la textura puede variar en las caras de la hoja, muchas veces tiene la hoja una textura en la parte de adelante y otra textura atrás de la hoja y muchas veces la textura es la misma y esa textura es un factor común en las especies de las plantas, hay texturas lisas, rugosas, peluda o vellosa, aterciopelada, áspera, escamosa, etc.
4. Plantas de misma especie con diferente color, la cromaticidad de la hoja es un aspecto muy variante, ya que el color de un conjunto de hojas de la misma especie puede cambiar de color debido a diversos factores, siendo la cromaticidad en las plantas dependiente del entorno como es la temperatura, el clima, el tipo de suelo en donde esta plantada, vitaminas, edad de la hoja, enfermedades de la planta, plagas en la planta, etc. Plantas diferentes con misma cromaticidad, el color de las hojas de las plantas tienden a ir bajo un mismo patrón, en general las hojas cambian con los cambios climáticos del año, en primavera y verano las hojas obtienen un tono verde de la clorofila, en otoño los arboles sintetizan la clorofila y toman pigmentos de carotenoides y antocianinas, por eso toman una tonalidad amarilla, anaranjado y rojizos. El pigmento de las hojas ha sido verde claro, verde amarillento, verde medio, verde oscuro, verde azulado, verde argénteo o blanquecino, púrpura, rojizo, anaranjado, variegado (amarillo-verde o verde-blanco), etc.

Debido a la gran cantidad de variables que intervienen en la clasificación de una hoja, el análisis de sistemas que utilicen más de un grupo de características son necesarios.

1.2. Justificación

Desarrollar un sistema que identifique de forma eficiente el tipo de planta a partir de su hoja, es un reto importante debido a sus múltiples aplicaciones como las que se enumeran a continuación:

1. En botánica los sistemas de identificación son de gran utilidad, la botánica se dedica al estudio de las plantas como la forma de las plantas, se encarga de la clasificación

y ordenamiento con nomenclatura de plantas, estudia las funciones de las plantas, su evolución, sus enfermedades e incluso estudia la aplicación que las personas les damos a las plantas. Mediante el desarrollo y uso de software para la identificación de las plantas, todas las tareas de investigación y estudio sería más fácil, consumiría menor número de recursos en tiempo y dinero, permitiendo además la transmisión del conocimiento e investigación. El desarrollo de sistemas automáticos de identificación de plantas puede ayudar a resolver problemas actuales en el área de botánica.

2. El uso de plantas medicinales y su reconocimiento es un proceso de identificación delicado, para que una planta sea considerada medicinal tiene que tener propiedades curativas o preventivas para algún padecimiento, sin embargo no necesariamente se debe usar toda la planta muchas veces solo son necesarias las hojas, el tallo o corteza. Mediante el uso de sistemas de identificación es posible facilitar el proceso para reconocer de forma adecuada y tener mayor seguridad en la información de las plantas medicinales. Muchas instituciones y organizaciones interesadas en el estudio de las plantas buscan tener la facilidad de tener acceso al tipo de planta, información de las propiedades curativas para el desarrollo de tratamientos medicinales en enfermedades, para realizar estudios de los suelos de las plantas para especies nuevas o extintas, sistemas de identificación de factores ambientales que afectan algunas especies de plantas, muchas veces los resultados de sistemas de identificación de plantas ayudan a construir una imagen local, nacional y mundial de la flora del planeta. Además de conseguir un control de las especies de las plantas.
3. En las aduanas se realizan sistemas de identificación en control de acceso físico de personal, de mercaderías comercializadas internacionalmente a los efectos del cumplimiento de las funciones aduaneras, en la aduana se deben cumplir políticas, procedimientos y normas de registro en géneros y mercancías que se importan o exportan por lo que los sistemas de identificación permiten gran manejo de revisión y control. Los sistemas automáticos de identificación de plantas ayudarían a disminuir el tiempo de revisión y control, permitiendo reducir costos y pérdidas.
4. Las plagas muchas veces no son identificadas a simple vista y son determinadas por un conjunto de parámetros y características que estas cumplan, parámetros que con los sistemas

nos permiten identificar de forma más rápida y eficiente. Los sistemas de identificación de plantas, podrían a su vez permitir mejoras en identificación de plagas, ya que existen plagas que únicamente atacan a determinado cultivo.

1.3. Objetivos y alcances del proyecto

1.3.1. Objetivo General

Desarrollar un sistema para identificación de plantas a partir de imágenes de hojas.

1.3.2. Objetivos específicos

1. Investigar y determinar las técnicas necesarias para el preprocesamiento de las imágenes de la hoja.
2. Aplicar las técnicas de segmentación para la separación de la hoja a partir del fondo
3. Identificar y emplear las técnicas para la extracción de características geométricas, texturales y venales de cada hoja.
4. Implementar los algoritmos de clasificación
5. Identificar la tasa de error de cada clasificador para determinar el algoritmo de clasificación óptimo en nuestro sistema.
6. Realizar la validación de los algoritmos desarrollados
7. Realizar las pruebas finales al sistema

1.4. Hipótesis

¿Es factible desarrollar un sistema de identificación eficiente para conjuntos grandes de imágenes de hojas de plantas?

1.5. Estado del arte

La importancia de las plantas radica en que son un elemento importante para el equilibrio de la vida en el planeta, en cualquier lugar que nos encontremos se hace presente una gran variedad de plantas así como en regiones que son inhabitables para el ser humano. Una gran variedad de plantas contienen cualidades únicas que servirían para el desarrollo humano y que a veces algunas de ellas están en riesgo de extinción, de aquí radica la importancia para la identificación de hojas de plantas de forma eficiente y eficaz. Los algoritmos existentes pueden ser clasificados acorde a las técnicas de extracción de características utilizada y estas podrían ser características texturales [23], [8], [29], cromáticas [15], [6], geométricas [34], [17], [40] o de venación [20], [32].

En la actualidad se han desarrollado varios sistemas de identificación de plantas. El departamento de Automatización de la Universidad de Ciencia y Tecnología de China, realizó un estudio “Reconocimiento de especies de plantas basado en la forma de la hoja” [17]. En este proyecto primero se construyó una base de datos de hojas de diferentes plantas y utilizó un método de clasificación referido como *Move Median Centers* (MMC). En éste estudio los investigadores tomaron en cuenta las características de forma, color, textura y estructura de las hojas, en la mayoría de las hojas su color es verde pero este depende de las sombras, cambios del agua, nutrientes, atmosfera, y la temporada por lo que se contemplaron las fotos en color gris. Para la creación de la base de datos arrancaron la hoja de la planta, y tomaron la imagen a color de la hoja con una cámara digital, de tal forma que en la imagen incluían primero la hoja y el fondo de la imagen era borroso. Su base de datos incluía 20 especies de diferentes plantas, cada especie incluía 20 imágenes de muestra, con un total de 400 imágenes en la base de datos. Las características geométricas consisten en el aspecto de la proporción, rectangularidad, área, perímetro, esfericidad, circularidad, excentricidad y factor de forma, etc. por lo que consideran en este trabajo el uso de MMC. El algoritmo calcula la mediana multidimensional de la clase, el radio que abarca los puntos de la clase, además de realizar un número determinado de iteraciones para eliminar el centro de la hiperesfera para ampliarla y abarcar el mayor número posible de puntos. Al encontrar la hiperesfera de mayor tamaño los puntos dentro de ésta son eliminados y se realiza mismo proceso para todos los puntos siguientes.

En una publicación en la revista internacional de ciencias de Ingeniería y Tecnología, en la Universidad Noida de la India[39], se publica el artículo “Sistema Automático de reconocimiento de hojas para identificación de plantas utilizando tecnología de una máquina de visión”. En este proyecto se contemplan tres fases, pre-procesamiento (técnica de mejora para la imagen antes del proceso computacional), extracción de características (fase donde se derivan las características basadas en color y forma de la hoja de la imagen) y clasificación (las características obtenidas son usadas como entradas para la clasificación, y el resultado es utilizado como prueba y comparado mediante una red neuronal artificial y un clasificador de distancia Euclidiana). Para realizar la comparación de las hojas contemplaron las características de color, características de forma, características geométricas como el diámetro, largo, ancho, área y perímetro de la hoja, también características morfológicas como rectangularidad, factor de forma, relación de aspecto, factor de estrecho, relación de perímetro, de diámetro, relación de perímetro, de largo y ancho y características de la vena. Este sistema fue probado en un conjunto de datos Flavia. El clasificador Nearest Neighbor (KNN) se basa en la distancia Euclidiana y es uno de los clasificadores más sencillos, crece de forma lineal con el tamaño del conjunto de datos, sin limitar la escalabilidad del clasificador [39]. El algoritmo KNN requiere que se especifique una métrica para poder medir la proximidad. Las predicciones se basan en los ejemplos más parecidos al que hay que predecir. Entre las ventajas del modelo las más importantes son: no se necesita hacer alguna suposición sobre los conceptos a aprender, se pueden aprender conceptos complejos usando funciones sencillas como aproximaciones locales y es muy tolerante al ruido [9].

Otro artículo importante “Clasificación de la imagen de una hoja de una planta con fondo complicado”, los autores utilizan un método de segmentación de transformación de cuencas controlado, (*Marker Controlled Watershed Transformation (MCWT)*) combinado con pre-segmentación y operaciones morfológicas que son introducidas para segmentar la imagen de la hoja con fondo complicado, se utilizaron siete momentos geométricos de Hu y diecisiete momentos de Zernike, son extraídos como características de forma después de segmentar las imágenes binarias y de remover el tallo, también se utilizó un clasificador de centro de movimiento de una hiper-esfera (MCH) para reducir características eficientemente. La tasa de clasificación para hojas de plantas con fondo complicado con el algoritmo es de 92.6% [40].

En [3] los autores desarrollan un sistema para clasificación de arboles, los autores contemplan atributos como la localización de los árboles, el diámetro, la altura del árbol y el volumen de la madera, teniendo estas características generales se propone en este proyecto la clasificación de cinco especies de árboles diferentes basándose en el análisis geométrico 3D de la textura de la corteza. Las características de textura se calculan utilizando una combinación de transformaciones complejas de Wavelet (CWT) y la transformada de contorno (CT). Para la clasificación se utiliza el clasificador Random Forest (RF). [4] Presenta un novedoso método de clasificación basada en color y características de textura y utilizando una máquina de soporte vectorial de multi-clases (MSVM) .

En [23] se propone un algoritmo basado en texturas para la segmentación de imágenes médicas tridimensionales, se utilizan indicadores de textura calculados a partir de estadísticas de segundo orden, recabados mediante matrices de coocurrencia de niveles de gris. [8] Implementa una modificación del algoritmo *LBP (Local Binary Pattern)* para el análisis textural de imágenes multispectrales que dé como resultado una correcta interpretación de una zona de estudio como es el Valle de México, mediante imágenes satelitales. [29] Realiza la detección de tumores en imágenes de resonancia magnética (MRI) tridimensionales, para este trabajo se propone un algoritmo de crecimiento de regiones en dos etapas basado en características texturales obtenidas a partir de matrices de coocurrencia. En [17] se desarrolla un proyecto “Uso de hojas de color para identificar falta de nitrógeno y potasio en los tomates”. Los autores desarrollan un sistema para diagnóstico de enfermedades de plantas y deficiencia en nutrientes. El sistema extrae características de las hojas como textura y color, además obtiene la composición de intensidad mediante histogramas. El sistema utiliza algoritmos genéticos (GA) para seleccionar las características que obtengan mayor información al diagnosticar enfermedades en la planta. [17] Desarrolla un sistema de reconocimiento de hojas de plantas para facilitar rápidamente una clasificación de las plantas, utilizan un método de clasificación referenciado como un clasificador hiperesfera de centro de movimiento mediana (MMC), el sistema utiliza características morfológicas digitales y extracción de características geométricas tales como: proporción, rectangularidad, área, perímetro, esfericidad, circularidad, excentricidad. Las características son extraídas desde el contorno de la hoja. [40] Desarrolla un clasificador de plantas utilizando la foto de la hoja con un fondo complicado, Se utilizan momentos geométricos de Hu, momen-

tos de Zernike para extraer la forma de la hoja desde una segmentación de imágenes binarias y adicionalmente utilizaron un clasificador de centros de movimientos de hiperesfera (MCH). En [20] los autores diseñan un sistema que obtiene patrones distintivos en las nervaduras de distintos cultivares de soja, que permitan el reconocimiento automático aún si el resto de las características como forma, color y textura de las hojas sean similares. El sistema trabaja con las imágenes en escala de grises, la segmentación de las nervaduras se realiza mediante un sistema de segmentación especial que, permite extraer todos los píxeles con un determinado patrón en configuración de sus vecinos. En [32] se diseña un sistema de reconocimiento de hojas de especies leñosas de Europa Central, sistema que incluye 151 especies con 50 hojas por especie, antes de realizar la segmentación se realizó el trazo del contorno de la hoja, se tomaron en cuenta las características de textura y nervadura, se realizó la normalización de la translación, rotación, escalamiento y puntos de partida límite con descriptores de Fourier, el tamaño de la hoja se conoció usando características de separación y para clasificar se eligió un clasificador de vecino más cercano.

Aunque los métodos utilizados en el estado del arte son muy importantes y reportan un alto desempeño. No existe en la literatura un estudio sobre la influencia en el desempeño de las características utilizadas. En esta tesis se aborda este problema.

Capítulo 2

Procesamiento Digital de Imágenes

El procesamiento digital de imágenes (PDI), se refiere a procesar las imágenes del mundo real de manera digital por medio de una computadora. El PDI se realiza principalmente en dos áreas: en el mejoramiento de la información pictórica para la interpretación humana; y en el procesamiento de datos de la imagen para su almacenamiento. El PDI se agrupa en nivel bajo y nivel medio el nivel bajo se refiere a operaciones básicas como reducción de ruido, aumento de contraste y mejora de nitidez y el nivel medio se refiere a operaciones como reconocimiento o extracción de características como la segmentación (regiones, objetos), caracterización de objetos, clasificación y la salida son atributos de objetos (bordes, contornos, identidades de objetos individuales) [96], [63]. En este Capítulo se definen las técnicas de pre-procesamiento más elementales en PDI.

2.1. Operaciones Puntuales

1. Las operaciones puntuales son transformaciones al imagen de uno a uno de los pixeles, es decir el nuevo valor de un pixel (q) en la posición (i, j) , esta en función de otra imagen de un pixel (p) pero en la misma posición (i, j) . Una operación puntual se expresa matemáticamente como: Donde E es la imagen de entrada y S la imagen de salida. La función f especifica el mapeo del nivel de gris de la entrada al nivel de gris de la salida [96].

$$S(x, y) = f(E[x, y]) \quad (2-1)$$

2.1.1. Operador Identidad

Este operador crea una imagen idéntica a la imagen de entrada.

$$Q(i, j) = P(i, j) \quad (2-2)$$

2.1.2. Operador Inverso o Negativo

Este operador crea una imagen inversa de la imagen de entrada.

$$Q(i, j) = 255 - P(i, j) \quad (2-3)$$

2.1.3. Operador Umbral

Este operador crea una imagen de salida binaria, a partir de una imagen en tonos de grises, el parámetro se da acorde al valor de umbral u .

$$Q = \begin{cases} 0, & P \leq u \\ 255, & P > u \end{cases} \quad (2-4)$$

2.1.4. Operador Intervalo de Umbral Binario

Crea una imagen de salida binaria, a partir de una imagen con tonos de grises, donde los valores que esten dentro del intervalo de u_1 y u_2 su valor será 0 y si estan fuera del rango su valor es 255.

$$Q = \begin{cases} 255, & u_2 \leq P \leq u_1 \\ 0, & u_1 < P < u_2 \end{cases} \quad (2-5)$$

2.1.5. Sustracción de la media.

Si se considera que un filtro de paso bajo sirve para resaltar componentes a gran escala eliminando la variabilidad local, si a la imagen original se le resta el resultado de pasarle un filtro pasa bajo se consigue resaltar esa variabilidad local. La matriz de filtrado de este filtro

menos media puede calcularse directamente restando a la matriz de filtrado identidad la matriz de filtrado de la media.

$$N1 = media_O \quad (2-6)$$

$$f(npa) = media_O - N1. \quad (2-7)$$

2.1.6. Filtros basados en las derivadas.

La derivada de una función $y = f(x)$ es el incremento de y para cada incremento infinitesimal de x . En el caso de Modelo Digital de Elevaciones la derivada es la pendiente. La segunda derivada es la derivada de la derivada [82].

$$f(x) = Ne^{-ax^2} \quad (2-8)$$

donde N es una constante de normalización que depende de a pero no de x . Hay varias posibilidades para el exponente a . En la forma estándar $a = 1/2\sigma^2$

2.1.7. Operador Intervalo de Umbral Binario Invertido

Crea una imagen de salida binaria, a partir de una imagen en grises, donde los valores que estén dentro del intervalo de u_1 y u_2 su valor será 255 y si están fuera del rango su valor es 0.

$$Q = \begin{cases} 255, & u_1 < P < u_2 \\ 0, & u_1 \geq P \geq u_2 \end{cases} \quad (2-9)$$

2.1.8. Operador Umbral de la Escala de Grises

Crea una imagen de salida con los valores de nivel de gris que estén comprendidos entre el intervalo de u_1 y u_2 y el resto se cambia a 255.

$$Q = \begin{cases} 255, & P < u_1 \text{ o } P > u_2 \\ P, & u_1 < P < u_2 \end{cases} \quad (2-10)$$

2.1.9. Operador Umbral de la Escala de Grises Invertido

Crea una imagen de salida con los valores de grises propios pero invertidos que estén comprendidos en el intervalo de u_1 y u_2 , para el resto el valor será 255.

$$Q = \begin{cases} 255, & u_2 < P < u_1 \\ 255 - P, & u_1 < P < u_2 \end{cases} \quad (2-11)$$

2.1.10. Operador Extensión

Este operador crea una imagen de salida con una escala de grises completa de acuerdo al intervalo de u_1 y u_2 y suprimir todos los valores fuera de ese rango.

$$Q = \begin{cases} 0, & u_2 < P < u_1 \\ (P - u_1) \times \left(\frac{255}{u_2 - u_1}\right), & u_1 < P < u_2 \end{cases} \quad (2-12)$$

2.2. Operadores aritméticos

Este tipo de transformación utiliza la información contenida en la misma posición de los pixeles de dos imágenes dando como resultado una tercera [61], [60], [96].

$$C(i, j) = f(A(i, j), B(i, j)) \quad (2-13)$$

2.2.1. Suma

Esta técnica se usa para reducir el ruido en las imágenes y para producir el efecto de movimiento entre imágenes iguales. Donde: A, B son las imágenes entrada y K es el número de imágenes.

$$C(i, j) = \frac{f(A(i, j) + B(i, j))}{K} \quad (2-14)$$

2.2.2. Resta

Esta Técnica es útil para detectar el cambio producido en dos imágenes que han sido captadas en dos instantes de tiempo diferentes. Donde: A, B son las imágenes entrada y K es el número de imágenes.

$$C_{(i,j)} = f(A_{(i,j)} - B_{(i,j)}) \times K \quad (2-15)$$

2.3. Operaciones grupales

Las operaciones de vecindad nos permiten mejorar el contraste de una imagen que no era muy buena, además de reducir ruido que se presente en las imágenes, para ello se utiliza una máscara. Donde I es la máscara con valores constantes y W son los pesos de los vecinos del pixel que se está analizando. Las técnicas de filtrado nos permiten resaltar u ocultar de forma selectiva la información contenida en una imagen a diferentes escalas espaciales. El proceso de filtrado consiste en aplicar a cada uno de los pixeles de la imagen de una matriz de filtrado de tamaño $N \times N$, compuesta por números enteros y que genera un nuevo valor mediante una función del valor original, el resultado final se divide entre un escalar [26], [84], [52], [79], [96].

$$\begin{bmatrix} W_{1,1} & W_{1,2} & W_{1,3} \\ W_{2,1} & W_{2,2} & W_{2,3} \\ W_{3,1} & W_{3,2} & W_{3,3} \end{bmatrix} \begin{bmatrix} I_1 & I_2 & I_3 \\ I_4 & I_5 & I_6 \\ I_7 & I_8 & I_9 \end{bmatrix}$$

El cálculo de un nuevo pixel R_{ij} es definido mediante R

$$R_{ij} = W_{1,1}I_{i-1,j-1} + W_{1,2}I_{i,j-1} + W_{1,3}I_{i+1,j-1} + W_{2,1}I_{i-1,j} + \dots + W_{3,3}I_{i+1,j+1} \quad (2-16)$$

2.3.1. Filtro de paso bajo.

Su objetivo es suavizar la imagen, útil para cuando la imagen tiene mucho ruido y se quiere eliminar o también se utiliza para resaltar la información correspondiente a una determinada

escala que es el tamaño de la matriz de filtrado [26].

2.3.2. Filtro de la Media

En este filtro se sustituye cada pixel en $I(x, y)$ por el promedio (o media) de sus vecinos. El filtro media no es adecuado para remover ruido impulsivo de sal y pimienta, pero funciona mejor para ruido Gaussiano y ruido de pecas (speckle) [26], [56].

2.3.3. Filtro de la Media Ponderada

Los elementos de la matriz de filtrado no son todos 1 sino que se da más peso a uno de ellos (generalmente el central) para obtener un resultado más parecido a la imagen original y evitar que se vea borrosa [26].

2.3.4. Filtro de la Mediana

Los pixeles de la nueva imagen se generan calculando la mediana del conjunto de píxeles del entorno de vecindad del píxel correspondiente a la imagen original. Filtro se utiliza bastante para cuando se tiene ruido aleatorio como sal y pimienta[26], [56].

2.3.5. Filtro Gaussiano

Este tipo de filtro reduce especialmente el ruido tipo gaussiano, que produce variaciones en la imagen, originado por las diferencias de ganancias del sensor, ruido de digitalización, etc.

El filtro Gaussiano en general da mejores resultados que un promedio o media y se argumenta a que la vista humana hace un filtrado de este tipo, el filtro gaussiano tiene la desventaja de suavizar las orillas o discontinuidades, generando que se emborronen.

El valor máximo aparece en el pixel central y disminuye hacia los extremos tanto más rápido cuanto menor sea el parámetro de desviación típica, el resultado será un conjunto de valores entre 0 y 1 para transformar la matriz a una matriz de números enteros se divide toda la matriz por el menor de los valores obtenidos [56].

$$g(x, y) = e^{-\frac{x^2 + y^2}{2s^2}} \quad (2-17)$$

$$G(x, y) = \frac{g(x, y)}{\min_{x,y}(g(x, y))} \quad (2-18)$$

2.3.6. Filtros Pasa Altos

Se utilizan para detectar cambios de luminosidad, para la detección de patrones como bordes o para resaltar detalles finos de una imagen.

2.3.7. Transformaciones de Intensidad

Una transformación de intensidad consiste en mapear los valores de intensidad de cada pixel a otros valores de acuerdo a cierta función de transformación, las funciones pueden ser de dos tipos: Lineales y No lineales.

2.3.8. Transformaciones lineales

Se tiene una relación lineal de los valores de pixel de la imagen de salida respecto de la imagen de entrada. Los tipos de transformaciones lineales son utilizadas para:

- Obtener el negativo de una imagen.
- Aumentar o disminuir la intensidad (brillo de la imagen)
- Aumento de contraste.

2.3.9. Transformaciones no lineales

Son funciones que mantienen la estructura básica de la imagen, se utiliza en:

- Expansión (o aumento) de contraste: Se incrementa el contraste, en forma diferente para distintos rangos de intensidades.
- Compresión de rango dinámico: Se reduce el rango de niveles de gris o intensidades de la imagen.

Intensificación de un rango de niveles: Se aumenta la intensidad de un rango de niveles de gris de la imagen.

2.3.10. Histogramas

El histograma de una imagen es una función discreta que representa el número de píxeles en la imagen en función de los niveles de intensidad g . La probabilidad $P(g)$ de ocurrencia de un determinado nivel g se define como [96]:

$$P(g) = \frac{N(g)}{M} \quad (2-19)$$

Donde M es el número de píxeles en la imagen y $N(g)$ es el número de píxeles en el nivel de intensidad g .

Un Histograma contiene el número de píxeles que tienen el mismo nivel de gris. Se representa como un gráfico de barras en el que las abscisas son los distintos colores de la imagen y las ordenadas la frecuencia relativa con la que cada color aparece en la imagen. El histograma proporciona información sobre el brillo y el contraste de la imagen.

2.3.11. Expansión del histograma

Esta operación consiste en distribuir las frecuencias de la apariencia de los píxeles en todo el ancho del histograma, esto equivale a expandir el histograma de manera que el valor de la intensidad más baja sea cero y el de la intensidad más alta sea el valor máximo. Es posible aumentar el contraste de una imagen, si una imagen es demasiado oscura se volverá más visible.

Una función para expandir los niveles de gris de un histograma se puede definir de la siguiente manera:

Donde $f(i, j)$ es el nivel de gris de la imagen de entrada; $f(i, j)_{MAX}$ es el mayor valor del nivel de gris en la imagen de entrada f . $f(i, j)_{MIN}$ es el menor valor del nivel de gris en la imagen de entrada f . MAX y MIN corresponden al máximo y mínimo valores posibles de los niveles de gris (para una imagen de 8 bits sería 0 y 255 respectivamente).

$$G(i, j) = \left[\frac{f(i, j) - f(i, j)_{MIN}}{f(i, j)_{MAX} - f(i, j)_{MIN}} \right] [MAX - MIN] + MIN \quad (2-20)$$

2.3.12. Contracción del histograma

Esta técnica no produce realmente un realzado de la imagen, produce una disminución del contraste de la imagen.

Donde $f(i, j)$ es el nivel de gris de la imagen de entrada, $f(i, j)_{MAX}$ es el mayor valor del nivel de gris en la imagen de entrada f . $f(i, j)_{MIN}$ es el menor valor del nivel de gris en la imagen de entrada f . C_{MAX} y C_{MIN} corresponden al máximo y mínimo valores posibles de los niveles de gris.

$$G(i, j) = \left[\frac{C_{MAX} - C_{MIN}}{f(i, j)_{MAX} - f(i, j)_{MIN}} \right] [f(i, j) - f(i, j)_{MIN}] + C_{MIN} \quad (2-21)$$

2.3.13. Desplazamiento del histograma

El desplazamiento del histograma se usa para aclarar y oscurecer una imagen pero manteniendo la relación entre los valores de los niveles de gris. Esta operación puede llevarse a cabo por la simple adición o sustracción de un número fijo a todos los valores del nivel de gris. Donde DES es un valor constante utilizado para llevar a cabo el desplazamiento del histograma.

$$G(i, j) = f(i, j) + DES \quad (2-22)$$

2.3.14. Ecuación de Histograma

Mediante la ecualización se pretende armonizar la distribución del nivel de luminosidad de la imagen, este proceso modifica el histograma de tal modo que en la nueva imagen todos los niveles tengan la misma frecuencia, se distribuyen los valores de intensidad a lo largo del espectro [94], [65].

Considerando los valores de intensidad máximo y mínimo en el rango de niveles de gris de la imagen dados, g_{max} y g_{min} . Con $P_g(g) = \sum_{g=0}^g p(g)$ y α un parámetro a definir.

Tabla 2-1: Tipos de ecualización

Ecualización	Expresión Matemática
Uniforme	$F(g) = [g_{\text{máx}} - g_{\text{mín}}] P_g(g) + g_{\text{mín}}$
Exponencial	$F(g) = g_{\text{mín}} - \frac{1}{\alpha} \ln [1 - P_g(g)]$
Reyleigh	$F(g) = g_{\text{mín}} + \left[2\alpha^2 \ln \left(\frac{1}{1 - p_g(g)} \right) \right]^{\frac{1}{2}}$
Hipercúbica	$F(g) = \left([\sqrt[3]{g_{\text{máx}}} - \sqrt[3]{g_{\text{mín}}}] P_g(g) + \sqrt[3]{g_{\text{mín}}} \right)^3$
Logaritmo hiperbólica	$F(g) = g_{\text{mín}} \left[\frac{g_{\text{máx}}}{g_{\text{mín}}} \right] P_g(g)$

2.3.15. Normalización de Histograma

Para evitar que los valores de un histograma sean muy dispares, se puede normalizar dicho histograma. Donde N_{max} es el nuevo máximo valor, N_{min} es el nuevo mínimo valor.

$$N_{x,y} = \frac{N_{\text{max}} - N_{\text{min}}}{O_{\text{máx}} - O_{\text{mín}}} \times (O_{x,y} - O_{\text{mín}}) + N_{\text{min}} \quad (2-23)$$

2.3.16. Expansión de Contraste

Se incrementa el contraste, en forma diferente para distintos rangos de intensidades. Donde $a = 2^n - 1$; $b = 0$; $C = \text{pixel con intensidad máxima}$ y $d = \text{pixel con intensidad mínima}$.

$$N_{x,y} = (O_{x,y} - C) \left(\frac{a - b}{a - d} \right) + a \quad (2-24)$$

2.3.17. Contraste

Estirar el histograma lo suficiente, para el mejor aprovechamiento de los niveles disponibles y un mejor diferenciamiento entre objetos.

Lograr un punto C que tomará el punto medio del histograma, y lo distribuirá de la mejor forma en el rango de intensidades. Donde K es el factor de aumento o disminución de contraste, y se encuentra en el rango $0 \leq K \leq \text{valor}$, para valor se recomienda un dato de 2 o 3. El C es un valor que puede ser la media, el promedio, la moda o la mitad del rango de intensidades.

$$R(x, y) = \{[1(x, y) - C] \times K\} + C \quad (2-25)$$

Cuando utilizamos variación de contraste, y deseamos llevar la escala actual de histograma a la escala 0-255, llamamos a éste proceso Contraste Automático.

$$R_{(x,y)} = [I_{(x,y)} - I_{(\text{mín})}] \times \left[\frac{255}{(I_{(\text{máx})} - I_{(\text{mín})})} \right] \quad (2-26)$$

2.3.18. Brillo

Aumentar el brillo de una imagen consiste en sumar o restar una constante a los colores que constituyen un píxel, cuidando siempre de nunca rebasar los límites 0 y 255. Donde K puede ser negativo o positivo.

$$R(x, y) = I(x, y) + K \quad (2-27)$$

2.3.19. Interpolación

La interpolación se considera como el cálculo del valor de intensidad de un píxel, en una posición cualquiera. Donde $P(i, j)$ es el valor de un píxel de la imagen original en la posición espacial (i, j) , $P(x, y)$ es el valor del píxel imagen resultante de la interpolación a partir de $p(i, j)$ y $h(x, y)$ es lo que se conoce como núcleo de interpolación.

Existen diferentes núcleos de interpolación, de los mas utilizados son vecino más próximo, bilineal y bicúbico.

$$P(x, y) = \sum_{i=-n}^n \sum_{j=-m}^m p(i, j)h(x - i, y - j) \quad (2-28)$$

2.3.20. Interpolación por vecino más próximo

Una forma de hacerlo es suponer que el píxel, para ser interpolado, toma el mismo valor que el más cercano de entre los cuatro que lo rodean; para decidir cuál es el más cercano, se puede utilizar la distancia euclídea. Otra posibilidad es asignarle la intensidad media asociada a los

dos píxeles más cercanos, uno en la dirección horizontal x y el otro en la dirección vertical y . El núcleo de interpolación $h(x, y)$ estaría definido como:

$$h(x) = \begin{cases} 1, & 0 < |x| < 0,5 \\ 0, & \text{de otro modo} \end{cases} \quad (2-29)$$

2.3.21. Interpolación bilineal

Asigna al píxel en cuestión un valor medio ponderado de las luminancias de los cuatro píxeles que le rodean. Los factores de ponderación vienen dados por las distancias entre el píxel y los del entorno.

$$h(x) = \begin{cases} 1 - |x|, & 0 < |x| < 1 \\ 0, & \text{de otro modo} \end{cases} \quad (2-30)$$

2.3.22. Interpolación bicúbica

La interpolación bicúbica tiene no sólo los cuatro píxeles diagonales más cercanos, sino sus puntos más cercanos también, en la que intervienen 16 puntos vecinos del píxel que se está interpolando.

$$h(x) = \begin{cases} 1 - 2|x|^2 + |x|^3, & 0 \leq |x| < 1 \\ 4 - 8|x| + 5|x|^2 - |x|^3, & 1 \leq |x| < 2 \\ 0, & \text{de otro modo} \end{cases} \quad (2-31)$$

2.3.23. Detección de bordes, algoritmo de Canny

Un borde, dentro de una imagen, es considerado un tipo de discontinuidad, por lo que este puede ser detectado usando derivadas de primer y segundo orden [55]. El detector Canny es considerado el más efectivo, consta de 4 fases:

1. Obtención del gradiente
2. Supresión no máxima al resultado del gradiente
3. Se busca conservar la propiedad (histéresis) de umbral a la supresión no máxima
4. Cerrar los contornos.

Para obtener el gradiente se tiene que suavizar la imagen. Aplicar un suavizado con filtro gaussiano o algún otro filtro pero cuidando el no suavizar tanto la imagen porque podría perder detalles.

- 1.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (2-32)$$

El gradiente de una imagen $f(x, y)$ en un punto (x, y) se define como un vector bidimensional dado por :

$$G[f(x, y)] = \begin{bmatrix} J_x \\ J_y \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} f(x, y) \\ \frac{\partial}{\partial y} f(x, y) \end{bmatrix} \quad (2-33)$$

Calcular el gradiente:

Para cada pixel de la imagen $J(i, j)$; Calcular las componentes del gradiente J_x y J_y ; Calcular la magnitud de los bordes.

$$e_s(i, j) = \sqrt{J_x^2(i, j) + J_y^2(i, j)} \quad (2-34)$$

Estimar la orientación del gradiente de los bordes.

$$e_o(i, j) = \tan^{-1} \frac{J_y}{J_x} \quad (2-35)$$

Supresión no máxima al resultado del gradiente, el objetivo de este paso es obtener bordes de 1 píxel de grosor al considerar únicamente píxeles cuya magnitud es máxima en bordes gruesos y descartar aquellos cuyas magnitudes no alcancen ese máximo. Se consideran cuatro direcciones identificadas por las orientaciones de 0° , 45° , 90° y 135° con respecto al eje horizontal. Para cada píxel se encuentra la dirección que mejor se aproxime a la dirección del ángulo de gradiente.

1. Para todo punto se obtiene la dirección más cercana dk a 0° , 45° , 90° y 135° en $Ea(i, j)$
2. Si $Em(i, j)$ es menor que uno de sus dos vecinos en la dirección dk , $IN(i, j) = 0$. Si no $IN(i, j) = Em(i, j)$.

Posteriormente se observa si el valor de la magnitud de gradiente es más pequeño que al menos uno de sus dos vecinos en la dirección del ángulo obtenida en el paso anterior. De ser así se asigna el valor 0 a dicho píxel, en caso contrario se asigna el valor que tenga la magnitud del gradiente. La salida de este segundo paso es la imagen I con los bordes adelgazados, es decir, $Em(ij)$, después de la supresión no máxima de puntos de borde.

3. Histéresis de umbral a la supresión no máxima. La imagen obtenida en el paso anterior suele contener máximos locales creados por el ruido. Una solución para eliminar dicho ruido es la histéresis del umbral.

El proceso consiste en tomar la imagen obtenida del paso anterior, tomar la orientación de los puntos de borde de la imagen y tomar dos umbrales, el primero más pequeño que el segundo. Para cada punto de la imagen se debe localizar el siguiente punto de borde no explorado que sea mayor al segundo umbral. A partir de dicho punto seguir las cadenas de

máximos locales conectados en ambas direcciones perpendiculares a la normal del borde siempre que sean mayores al primer umbral. Así se marcan todos los puntos explorados y se almacena la lista de todos los puntos en el contorno conectado. Es así como en este paso se logra eliminar las uniones en forma de Y de los segmentos que confluyan en un punto.

Entrada I , Ea (orientación del gradiente), y dos umbrales $T1$ y $T2(T2 > T1)$

Considerar cuatro direcciones $d1$, $d2$, $d3$, $d4$ identificadas por las direcciones de 0° , 45° , 90° y 135° con respecto al eje horizontal

a) Para cada píxel (i, j) :

- 1) Encontrar la dirección d_k que mejor se aproxima a la dirección $Ea(i, j)$, que viene a ser la perpendicular al borde.
- 2) Si $Em(i, j)$ es más pequeño que al menos uno de sus dos vecinos en la dirección d_k , al píxel (i, j) de I se le asigna el valor 0, $I(i, j) = 0$ (supresión), de otro modo $I(i, j) = Em(i, j)$.

b) Devolver I .

c) Cerrar los contornos.

Un método muy utilizado es el algoritmo de Deriche y Cocquerez. Este algoritmo utiliza como entrada una imagen binarizada de contornos de un píxel de ancho. El algoritmo busca los extremos de los contornos abiertos y sigue la dirección del máximo gradiente hasta cerrarlos con otro extremo abierto.

El procedimiento consiste en buscar para cada píxel uno de los ocho patrones posibles que delimitan la continuación del contorno en tres direcciones posibles. Esto se logra con la convolución de cada píxel con una máscara específica. Cuando alguno de los tres puntos es ya un píxel de borde se entiende que el borde se ha cerrado, de lo contrario se elige el píxel con el valor máximo de gradiente y se marca como nuevo píxel de borde y se aplica nuevamente la convolución. Estos pasos se repiten para todo extremo abierto hasta encontrar su cierre o hasta llegar a cierto número de iteraciones determinado.

Entrada: imagen I obtenida del paso anterior

imagen Eo de la orientación del gradiente

umbral $T1$

umbral $T2$

donde $T1 < T2$

Salida: imagen G con los bordes conectados de contornos

1. Para todos los puntos de I y explorando I en orden fijo:
 - a) Localizar el siguiente punto de borde no explorado previamente, $I(i, j)$, tal que $I(i, j) > T2$
 - b) Comenzar a partir de $I(i, j)$, seguir las cadenas de máximos locales conectados en ambas direcciones perpendiculares a la normal de borde, siempre que $I > T1$
 - c) Marcar todos los puntos explorados y, salvar la lista de todos los puntos en el entorno conectado encontrado.
2. Devolver G formada por el conjunto de bordes conectados de contornos de la imagen, así como la magnitud y orientación, describiendo las propiedades de los puntos de borde.

Capítulo 3

Segmentación

Este capítulo se mostrarán algunas técnicas de segmentación de imágenes, estas técnicas se basan en el proceso de dividir una imagen digital en varios grupos de píxeles, separando un objeto en específico, realizando un trabajo de simplificación para una imagen más significativa y fácil de analizar.

En la segmentación de una imagen encontramos procesos como extracción de bordes, regiones y puntos de interés; se usa para localizar objetos y para encontrar los límites de estos objetos dentro de una imagen.

3.1. Métodos basados en detección de bordes

3.1.1. Operador de Sobel

El gradiente de una imagen $f(x, y)$ en un punto (x, y) se define como un vector bidimensional dado por la ecuación, siendo un vector perpendicular al borde:

$$G[f(x, y)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} f(x, y) \\ \frac{\partial}{\partial y} f(x, y) \end{bmatrix} = \begin{bmatrix} \frac{f(x+1, y) - f(x, y)}{\Delta x} \\ \frac{f(x, y+1) - f(x, y)}{\Delta y} \end{bmatrix} \quad (3-1)$$

donde el vector G apunta en la dirección de variación máxima de f en el punto (x, y) por unidad de distancia con la magnitud y dirección dadas por:

$$|G| = \sqrt{G_x^2 + G_y^2} \quad (3-2)$$

$$\phi(x, y) = \tan^{-1} \frac{G_y}{G_x} \quad (3-3)$$

Es una práctica habitual aproximar la magnitud del gradiente con valores absolutos:

$$|G| = |G_x| + |G_y| \quad (3-4)$$

Finalmente, la imagen obtenida resulta ser una imagen binaria obtenida con el siguiente criterio, donde T es un valor de umbral no negativo.

$$g(x, y) = \begin{cases} 1, & G[f(x, y)] \geq T \\ 0, & G[f(x, y)] < T \end{cases} \quad (3-5)$$

Indicado anteriormente, los valores de G_x y G_y de la ecuación:

$$G[f(x, y)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} f(x, y) \\ \frac{\partial}{\partial y} f(x, y) \end{bmatrix} = \begin{bmatrix} \frac{f(x+1, y) - f(x, y)}{\Delta x} \\ \frac{f(x, y+1) - f(x, y)}{\Delta y} \end{bmatrix}$$

pueden obtenerse por convolución de la imagen con las máscaras 3×3 , conocidas como operadores de Sobel:

Máscara usada para obtener G_x en el punto central de la región 3×3 .

$$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad (3-6)$$

Máscara usada para obtener G_y en el mismo punto:

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad (3-7)$$

Basado en la $\begin{bmatrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{bmatrix}$ las derivadas basadas en los operadores de Sobel son:

$$G_x = (z_1 + 2z_4 + z_7) - (z_3 + 2z_6 + z_9) \quad (3-8)$$

$$G_y = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \quad (3-9)$$

donde los distintos valores de z en la región de la matriz de $z_1 \dots z_9$ son los niveles de gris de los pixeles solapados por las máscaras en cualquier localización de la imagen. Para obtener los valores de las componentes del vector gradiente en el punto definido por el pixel central de la región, se utilizan las expresiones $G_x = (z_1 + 2z_4 + z_7) - (z_3 + 2z_6 + z_9)$ y $G_y = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$ con lo que la magnitud y el ángulo se pueden obtener a partir de $|G| = \sqrt{G_x^2 + G_y^2}$ con $\phi(x, y) = \tan^{-1} \frac{G_y}{G_x}$ y $|G| = |G_x| + |G_y|$ es decir, obtenemos un valor del gradiente en dicho punto.

3.1.2. Operador de Prewitt

El operador de Prewitt es similar al de Sobel diferenciándose en los coeficientes de las máscaras:

1. Máscara usada para obtener G_x en el punto central de una región de dimensión 3×3

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \quad (3-10)$$

2. Máscara usada para obtener G_y en el mismo punto

$$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad (3-11)$$

3. La magnitud y dirección del gradiente se obtienen como en el caso de Sobel por medio de las ecuaciones.

$$|G| = \sqrt{G_x^2 + G_y^2} \quad (3-12)$$

$$\phi(x, y) = \tan^{-1} \frac{G_y}{G_x} \quad (3-13)$$

$$|G| = |G_x| + |G_y| \quad (3-14)$$

3.1.3. Operador de Roberts

El operador de Roberts a diferencia de los anteriores, marca solamente los puntos de borde sin informarnos sobre la orientación de éstos. Es un operador muy simple que trabaja muy bien en imágenes binarias. Opera según las dos diagonales perpendiculares mostradas, se definen las diagonales para el operador de Roberts [96]:

y definidas por la ecuación:

$$R = \sqrt{D_1^2 + D_2^2} \quad (3-15)$$

$$R = |D_1| + |D_2| \quad (3-16)$$

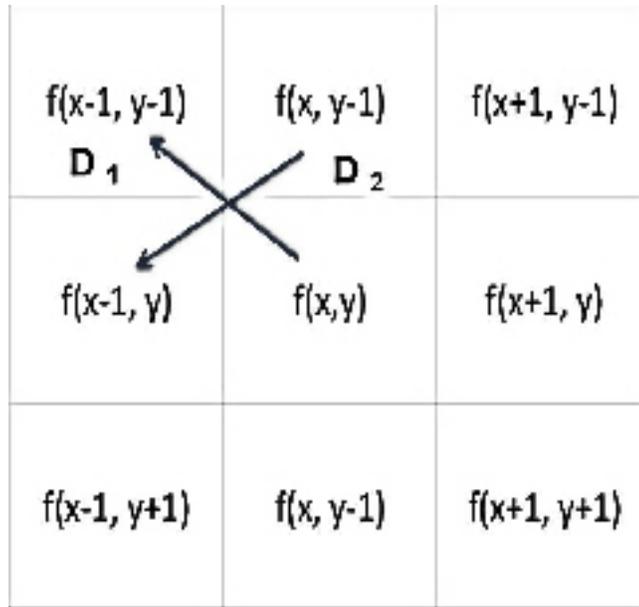


Figura 3-1: Operador cruzados de Roberts

$$D_1 = f(x, y) - f(x - 1, y - 1) \quad (3-17)$$

$$D_2 = f(x, y - 1) - f(x - 1, y) \quad (3-18)$$

Existen dos formas del operador de Roberts, dadas por $R = \sqrt{D_1^2 + D_2^2}$; $R = |D_1| + |D_2|$

:

1. La raíz cuadrada de la suma de las diferencias de los vecinos diagonales al cuadrado
2. Suma de la magnitud de las diferencias de los vecinos diagonales 3-1. Definición de las diagonales para el operador de Roberts.

3.2. Método de Otsu

Dada una imagen con L niveles de intensidad y asumiendo que el umbral buscado es T , las posibilidades acumuladas hasta T y desde T hasta L resultan ser [69], [96]:

$$w_1(t) = \sum_{z=1}^T P(z) \quad (3-19)$$

$$w_2(t) = \sum_{z=T+1}^L P(z) \quad (3-20)$$

A continuación se obtienen las medias y varianzas asociadas:

$$\mu_1(t) = \sum_{z=1}^T zP(z) \quad (3-21)$$

$$\mu_2(t) = \sum_{z=T+1}^L zP(z) \quad (3-22)$$

$$\sigma_1^2(t) = \sum_{z=1}^T (z - \mu_1(t))^2 \frac{P(z)}{w_1(t)} \quad (3-23)$$

$$\sigma_2^2(t) = \sum_{z=T+1}^L (z - \mu_2(t))^2 \frac{P(z)}{w_2(t)} \quad (3-24)$$

Finalmente se obtiene la varianza ponderada

$$\sigma_w^2(t) = w_1(t)\sigma_1^2(t) + w_2(t)\sigma_2^2(t) \quad (3-25)$$

Se elige el umbral T correspondiente al nivel de intensidad que proporcione la mínima varianza ponderada ya definida.

3.3. Segmentación PCA

El análisis de componentes principales (Principal Components Analysis) es una herramienta matemática muy conocida, con aplicaciones en diversas áreas de la ciencia y como técnica tradicional de proyección sobre un subespacio para reconocimiento de caras, es probablemente la más utilizada también para encontrar patrones en los datos de alta dimensión. Esta técnica de análisis estadístico de datos multidimensionales, fue introducida originalmente por Pearson en 1901 y desarrollada posteriormente por Hotelling en psicometría [93]. La técnica PCA es también conocida como la transformada de Karhunen-Loeve Transformation (KLT) ó Transformada de Hotelling.

Transformada de Karhunen-Loeve (KLT): Dado un vector aleatorio $x = [x_0 x_1 \dots x_{m-1}]^T \in R^m$, con momentos de segundo orden finitos, se define la KLT como la transformación lineal de x :

$$y = V^T x \quad (3-26)$$

donde $y = [y_0 y_1 \dots y_{p-1}]^T \in R^p$ ($p \leq m$) es el vector transformado, y la matriz $V = [v_0 v_1 \dots v_{p-1}]^T \in R^{m \times p}$ se forma colocando en las columnas los primeros p autovectores normalizados de la matriz covarianzas $C_{xx} = E[(x - \bar{x})(x - \bar{x})^T]$, ordenados por orden decreciente de sus autovalores no nulos asociados.

Consideramos la matriz X formada por M vectores columna X_k de n componentes: $x = [x_1, x_2, \dots, x_M]$, $x_j = [x_{jk}]$, $k = 1, 2, \dots, n$; $j = 1, 2, \dots, M$. Cada vector puede representar, por ejemplo, los píxeles de una imagen, por lo que n representa el número de píxeles de la imagen, si la imagen consiste en K filas y L columnas, cada vector tendrá $n = K \times L$ componentes. La media m_x de los vectores y la matriz de covarianza C_x están dadas por:

$$m_x = \frac{1}{M} \sum_{k=1}^M X_k \quad (3-27)$$

$$C_x = \frac{1}{M-1} X_m X_m^t \quad (3-28)$$

donde $X_m = [x_1 - m_x, \dots, X_M - m_x]$, la matriz de covarianza es real y simétrica, por lo

que siempre es posible encontrar un conjunto de n autovectores ortonormales. Sean e_i y λ_i , $i = 1, 2, \dots, n$, los autovectores y los correspondientes autovalores de la matriz C_x , ordenados convenientemente en orden descendente, de modo que $\lambda_j \geq \lambda_{j+1}$, para $j = 1, 2, \dots, n - 1$. Por definición, los autovectores y autovalores de una matriz C de dimensión $N \times N$ satisfacen la relación $Ce_i = \lambda_i e_i$ para $i = 1, 2, \dots, n$. Sea A una matriz cuyas filas están formadas por los autovectores de C_x , ordenados de forma que la primera fila de A es la transpuesta del autovector asociado con el autovalor de mayor valor y la última fila es la transpuesta del autovector asociado con el autovalor más pequeño.

La transformada PCA está dada por:

$$y = A(x - m_x) \quad (3-29)$$

La media de los vectores y resultantes es cero, Además la matriz de covarianza de estos vectores, Cy , es una matriz diagonal cuyos elementos a lo largo de la diagonal principal son los autovalores de Cx , esto es:

$$Cy = \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \\ & & \cdot & \\ & & & \cdot & 0 \\ 0 & & & 0 & \lambda_n \end{bmatrix} \quad (3-30)$$

Los elementos fuera de la diagonal principal son 0, de forma que los elementos de los vectores y no están correlacionados. Recordar que los λ_j son los autovalores de Cx y que los elementos a lo largo de la diagonal principal de una matriz diagonal son sus autovalores. Por tanto, Cx y Cy tienen los mismos autovalores. Lo mismo se mantiene para los autovectores.

Otra propiedad de PCA es la reconstrucción de x a partir de y , puesto que las filas de A son vectores ortonormales $A^{-1} = A^t$, y cualquier vector x puede recuperarse a partir de su correspondiente y y utilizando la relación:

$$x = A^t y + m_x \quad (3-31)$$

Supongamos que en lugar de utilizar todos los autovectores de Cx formamos una matriz A_d a partir de los d autovectores correspondientes a los d autovalores de mayor valor, para obtener un matriz de orden $D \times N$. Los vectores y serían d dimensionales y la reconstrucción dada en la ecuación anterior, no sería exacta. El vector reconstruido utilizando A_d quedaría:

$$\hat{x} = A_d^t y + m_x \quad (3-32)$$

Se puede demostrar que el error al cuadrado medio e_{cm} entre x y \hat{x} está dado por:

$$e_{cm} = \sum_{j=1}^n \lambda_j - \sum_{j=1}^d \lambda_j = \sum_{j=d+1}^n \lambda_j \quad (3-33)$$

De la ecuación anterior se define que el error es cero si $d = n$, es decir si todos los autovectores se utilizan en la recuperación. Puesto que los λ_j decrecen monótonamente, y también se muestra que el error puede minimizarse seleccionando los d autovectores asociados con los autovalores de mayor valor.

Capítulo 4

Clasificadores

La necesidad de desarrollar máquinas inteligentes, que sean capaces de comportarse como el cerebro humano, tal que emulen actividades que un ser humano cotidianamente realiza, nos ha llevado al desarrollo de redes neuronales artificiales interconectadas entre sí de forma paralela, que intentan interactuar con objetos del mundo real de la misma forma que un sistema nervioso biológico, para esto se ha tratado de entender como funcionan las neuronas del cerebro humano y de que forma se originan, además del como interactúan para realizar las actividades. Tener un conocimiento pleno de una planta puede ser muy difícil, ya que su tamaño, forma, textura y cromaticidad es variable bajo diferentes condiciones. Las hojas de las plantas pueden sufrir desgaste de color o degradación de sus características por las condiciones climáticas, etapa de la hoja, etc. Esto puede ocasionar incertidumbre en el sistema, por ello el uso de la técnica de redes neuronales.

En esta tesis se realizará el desarrollo de un sistema que emule el conocimiento humano, para llevar a cabo la identificación de las plantas mediante la técnica de reconocimiento de patrones con redes neuronales artificiales, para ello se analizarán las hojas de las plantas a identificar, de tal forma que teniendo una fotografía de la hoja, el sistema realizará una comparación de las características a validar, tales como el tamaño, la pigmentación, forma, etc, de tal forma que el sistema arrojará el nombre de la planta que coincide con las características de la fotografía. El desarrollo de técnicas para el procesamiento digital de imágenes nos ha permitido tanto la mejora como la extracción de características de una imagen, lo cual a dado camino al reconocimiento de patrones, y con la ayuda de la inteligencia artificial podemos emplear

todas estas técnicas en sistemas donde se requiere identificar las diferencias de las imágenes y determinar su identificación.

4.1. Redes neuronales

Las redes neuronales artificiales (RNA) pretenden imitar en una pequeña escala la forma del funcionamiento de las neuronas que forman al cerebro humano, las RNA almacenan conocimientos de forma experimental, de esta manera el conocimiento retenido es utilizado por la misma para llevar a cabo las actividades. Una RNA comparte con el cerebro humano que el conocimiento se adquiere por medio de un proceso de aprendizaje y los pesos sinápticos o fuerza con que están interconectadas las neuronas se utilizan para almacenar la información. Lo que se realiza inicialmente es entrenar el sistema, se aplican determinadas entradas a la red y se observa la salida, si la salida no se adecua a la que se esperaba entonces se ajustan los pesos de cada neurona y de esta manera conforme se van haciendo interacciones la red va aprendiendo, en las RNA no es necesario que se programe ni el conocimiento ni las reglas para el procesamiento de la información, las RNA obtienen el conocimiento mediante una regla de aprendizaje que adapta o cambia los pesos finales conforme los ejemplos que se le coloquen [79], [80], [86].

Las RNA almacenan la información de forma distribuida en todas las conexiones, capas ocultas de tal forma que si tiene una entrada incompleta o errónea, la red no pierde su camino, esta sigue respondiendo adecuadamente. Para construir una red neuronal se debe tomar en cuenta un algoritmo de entrenamiento acorde a la red neuronal que se elija (Adaline y Madaline, Perceptron, Back-Propagation, entre otras), el número de neuronas a ocupar por capa, y forma de representación de los datos.

La neurona artificial fue diseñada para emular las características del funcionamiento básico de la neurona biológica, se define según la siguiente ecuación donde Net es la salida, X el vector de entrada y W el vector de pesos [96].

$$Net = X \times W \tag{4-1}$$

Para llevar a cabo el entrenamiento de una RNA se utiliza un algoritmo de aprendizaje, estos se clasifican en entrenamiento supervisado y no supervisado, nosotros utilizaremos el algoritmo de entrenamiento Back-Propagation el cual será bajo entrenamiento supervisado, se le coloca un patrón de entrada bajo un patrón de salida deseado y los pesos son ajustados en cada iteración de tal forma que el error de salida disminuya hasta lograr el objetivo de salida deseada. La red Back-Propagation tiene gran capacidad para trabajar en problemas de clasificación, su propia característica de capas ocultas le permite una mayor eficiencia en sus resultados, esta red trabajo en dos fases en la cual recorre primero desde la entrada-capas ocultas-salida, esto arroja un margen de error, bajo este margen se realiza hacia atrás, de tal forma que ahora se inicia por la salida hasta llegar a la entrada hasta ir disminuyendo el rango de error de nuestra red [75].

El uso de RNA en diversas aplicaciones ha sido de gran utilidad mejorando la calidad de diversos sistemas, ya que con el paso del tiempo el uso e investigación de las RNA han dado mejoras a estas y por lo tanto mayor exactitud en los resultados. Acorde al problema a resolver se hace la elección de la red neuronal que sea más competente y adecuada para llevar a cabo las tareas de solución.

Los principios característicos de las redes neuronales artificiales según las autoras Hilera y Martínez son [97] :

1. Aprendizaje adaptativo: Las RNA pueden comportarse en función de un entrenamiento mediante ejemplos, lo cual le permite adaptarse a nuevas condiciones de trabajo de acuerdo a las modificaciones constantes que esta reciba.
2. Auto-organización: Enlazada a la generalización, característica que le permite a la red responder a datos o situaciones que no había experimentado con anterioridad, donde la entrada puede ser poco clara o incompleta, la cual puede interferir sobre su entrenamiento.
3. Tolerancia a fallos: Las redes neuronales son capaces de reconocer información con distorsión o incompleta, pero su tolerancia a fallos les permite seguir trabajando aunque halla destrucción de una parte de la red, esta característica se puede apreciar en el algoritmo

back-propagation ya que el realiza el almacenamiento de la información de forma distribuida en todo su dimensionamiento de la red, ya que este utiliza capas ocultas donde al perder una parte de la información, este sigue funcionando ya que cuenta con la mayor cantidad de datos.

4. Operación en tiempo real: Todas las redes neuronales, mencionando todos los métodos existentes cuentan con el reconocimiento de patrones en tiempo real debido a su forma de trabajar paralelamente y sus simultaneas iteraciones.
5. Fácil inserción en la tecnología: El fácil uso de chips para redes neuronales, lo cual permite la integración modular en los sistemas.

De esta forma las RNA son identificadas por su gran número de características que realizan el apoyo en los sistemas para un mejor desempeño y mayor utilidad en todo ámbito.

En 1943, los científicos estadounidenses Walter Pitts y el psicólogo Warren Sturgis McCulloch publicaron el primer modelo matemático de neurona artificial llamada neurona McCulloch-Pitts [73]. En este modelo cada neurona estaba dotada de entradas y salidas, para lo cual cada entrada es afectada por un peso y la salida está en función del cambio que realicen los pesos en la entrada, de tal forma que los pesos pueden ser utilizados para darle un control de la salida, para realizar la activación de la neurona se calcula la suma de los productos de cada entrada teniendo así una función de activación [30], [22].

En 1957 el psicopatólogo Frank Rosenblatt inventó el perceptron [71] [19], [22] tomando como base la neurona de McCulloch-Pitts, el perceptron es el modelo matemático mas simple de una neurona, la neurona es una célula especializada y caracterizada por poseer una cantidad indefinida de entradas llamadas dentritas y un canal de salida llamado axón. En 1960, Bernard Widrow y Marcial Ho desarrollaron el modelo Adaline [19]. En 1974 Paul Werbos desarrollo el algoritmo de aprendizaje Back-Propagation [25], [19], [30]. En la actualidad las investigaciones y la utilidad que se le está dando a esta tecnología de reconocimiento de patrones a sido masiva, estas se agrupan según varios criterios, el autor Deboeck, Guido J. Pattern [Cita] propone la siguiente agrupación:

1. Modelación financiera y económica.

2. Perfiles de mercado y clientes.
3. Aplicaciones médicas.
4. Gerencia del conocimiento y descubrimiento de datos.
5. Optimización de procesos industriales y control de calidad.
6. Investigación científica.

Con el propósito de fundamentar la presente investigación, a continuación se muestran las principales ideas que hoy existen sobre algoritmos de Clasificación, algunos tocantes al tema de un identificador de plantas. El área del procesamiento de imágenes a incrementado continuamente para diferentes áreas de aplicación tales como la morfología matemática en imágenes digitales, redes neuronales artificiales, reconocimiento de imágenes, de tales técnicas que se desprenden los sistemas de análisis basados en conocimiento, entre otras, por ejemplo [33], [13], [7], [28]:

1. En los bancos de Perú utilizan un sistema para el otorgamiento de un crédito, bajo el uso del algoritmo de árbol de clasificación CART de tal forma que reduzcan el riesgo y el tiempo de proceso de evaluación de créditos para clientes sin empleo formal e ingresos irregulares, este sistema busca características o patrones de comportamiento que permita clasificar a los clientes en solventes o insolventes y bajo un historial crediticio del cliente.
2. Los investigadores de Allied Signal entrenaron una red para reconocer objetivos submarinos a partir de un sonar.
3. Dentro del área financiera, las RNA se han usado para modelar el comportamiento del mercado de capitales, para evaluación de inversiones, para simulación de problemas macroeconómicos o pronostico de bancarrotas.
4. La Universidad de Pennsylvania creó una red neuronal óptica para detectar objetos mediante un radar.
5. El soft NestorWriter de Nestor, Inc. utiliza un diseño de red neuronal para interpretar letras manuscritas sobre un pad sensible.

6. En la Universidad de Toronto, Hinton y Sejnowski han desarrollado una máquina llamada Boltzman que consiste en la red de Hopfield.

Dentro de los métodos de clasificación, están los basados en redes neuronales como son el perceptrón, el perceptrón multicapa, algoritmos de agrupamiento como el Lloyd, máquinas de soporte vectorial, la red backpropagation, etc.

4.2. El Perceptrón

En su forma mas básica, el aprende una función discriminante del tipo $fd(x)$; el perceptrón para dos clases, $c1$ y $c2$ se desarrolla bajo este esquema [96].

$$fd(x) = \sum_{i=1}^n w_i x_i + w_{n+1} \quad (4-2)$$

Indica una función de decisión lineal con respecto a las componentes de los vectores x_i . Los coeficientes w_i , $i = 1, 2, \dots, n, n + 1$, son llamados pesos, modifican las entradas antes de que sean sumadas y suministradas al elemento de umbral, los pesos son similares a la sinapsis en el sistema neuronal humano. La función que transforma la salida correspondiente a la suma en la salida final se denomina función de activación y toma el valor 1 si su argumento es verdadero y -1 si es falso. Cuando $fd(x) > 0$ el elemento umbral produce una salida del perceptrón cuyo valor es $+1$, indicando que el patrón x es reconocido como perteneciente a la clase $c1$ y a la inversa, cuando $fd(x) \leq 0$, el límite de decisión implementado por el perceptrón se obtiene con la ecuación igualada a 0.

$$fd(x) = \sum_{i=1}^n w_i x_i + \theta w_{n+1} = 0 \quad (4-3)$$

$$w_1 x_1 + w_2 x_2 + \dots + w_n x_n + \theta w_{n+1} = 0 \quad (4-4)$$

La salida del umbral se determina :

$$O = \begin{cases} -1, & fd(x) \leq 0 \\ +1, & fd(x) > 0 \end{cases} \quad (4-5)$$

El objetivo consiste en obtener los pesos w_1 para discriminar entre dos clases c_1 o c_2 , se dispone de un conjunto de muestras de entrenamiento, denominadas de aprendizaje $X = \{x_1, x_2, \dots, x_n\}$ cuya pertenencia a una determinada clase es conocida, para obtener efectividad en el aprendizaje es necesario que el número de muestras n deben ser mayor a la cantidad de clases.

1. Inicialización de los pesos y del umbral. Inicialmente se asignan valores aleatorios a cada uno de los pesos $w_i, i = 1, 2, \dots, n, n + 1$. y al bias θ .

2. Presentación de un nuevo par (entrada, salida esperada). Presentar un nuevo patrón de entrada $x_p = \{x_1, x_2, \dots, x_n\}$ junto con la salida esperada $fd_i(k)$.

3. Cálculo de la salida actual. Siendo en este caso f la función de transferencia escalón:

$$y(k) = f[w^t x_i - w_{n+1}] \quad (4-6)$$

4. Adaptación de los pesos

$$w_i(k + 1) = w_i k + \alpha(k)[fd_i(k) - y(k)]x_i \quad (4-7)$$

La salida esperada $fd_i(k)$ es 1 si el patrón pertenece a la clase A y -1 si es de la clase B .

5. Si no hay convergencia (los pesos cambian), retornar a paso 2.

Se necesita converger hacia una distribución de funciones discriminantes. De tal forma que para las dos clases dadas c_1 y c_2 la solución al problema de clasificar un vector x desconocido es asociarlo a la clase cuya función discriminante $f(d)x$ es máxima, por consiguiente habrá 2 funciones discriminantes fd_1 y fd_2 .

4.3. BackPropagation

Este modelo de Red Neuronal es utilizado para problemas no lineales separables multiclase, basado en varias neuronas, lo cual nos sirve mucho para el reconocimiento de patrones [96].

Las capas de neuronas que conforman la arquitectura de la red, son colocadas de forma que su salida de cada neurona en una capa, proporciona la entrada de las siguientes neuronas de la capa próxima, donde:

A es el nombre predeterminado que se le da a la capa primera, y el número de neuronas colocadas en la capa A son representadas como N_A , y $N_A = d$ por lo que d es la dimensión de los vectores de entrada (patrones); la capa de salida (final de todas las subcapas) es llamada capa Q y su cantidad de neuronas se representa como N_Q , los valores finales se determinan como clases C , por lo cual N_Q es igual al número de clases C para las que la red se va a entrenar y sobre las que se llevara a cabo la clasificación; cada neurona tiene la misma forma que el perceptrón y su función de activación deber ser derivable para poder desarrollar la regla de entrenamiento con backpropagation. Se pueden utilizar diferentes tipos de funciones de activación para las diferentes capas o diferentes nodos de la misma capa de la red, pero es práctico utilizar el mismo tipo de función para todas las neuronas de la red neuronal.

Para llevar a cabo el entrenamiento, se debe desarrollar una regla de entrenamiento como la regla Delta de un perceptrón, la cual nos permita ajustar los pesos en cada una de las capas, buscando un mínimo a una función de error.

Calculamos el error desde la capa de salida.

$$E_Q = \frac{1}{2} \sum_{q=1}^{N_Q} (r_q - O_q)^2 \quad (4-8)$$

donde r_q son las respuestas deseadas y las respuestas actuales O_q . N_Q es el número de nodos en la capa de salida Q y por conveniencia en la notación para tomar la derivada mas tarde se coloca el $1/2$.

$$\Delta w_{qp} = -\alpha \frac{\partial E_Q}{\partial w_{qp}} \quad (4-9)$$

La capa P precede a la capa Q , Δw_{qp} se utiliza para actualizar los pesos y α es un incremento de corrección positivo; el error E_Q es una función para las salidas O_q , mediante la regla de la cadena de derivación podemos evaluar la derivada parcial de E_Q .

Ecuación para calcular las entradas netas para neuronas ocultas procedentes de las neuronas de entrada. Donde índice h se refiere a magnitudes de la capa oculta, subíndice p , al p -ésimo vector de entrenamiento, j a la j -ésima neurona oculta. Theta es opcional es una entrada más.

$$net_{pj}^h = \sum_{i=1}^N w_{ji}^h x_{pi} + \theta_j^h \quad (4-10)$$

Se realizan los mismos cálculos para obtener las salidas de las neuronas de salida.

$$net_{pk}^o = \sum_{j=1}^N w_{kj}^o y_{pj} + \theta_k^o \quad (4-11)$$

Dado un conjunto de patrones de entrada a cada uno de ellos se le asigna un objetivo de salida y se definen los parámetros de entrenamiento relativos a los siguientes valores: número de iteraciones, tolerancia, error cuadrático medio y razón de aprendizaje.

4.4. Máquinas de Vectores de Soporte (SVM)

Técnica que se introdujo en el paradigma de redes neuronales para la clasificación de patrones, inicialmente en la clasificación biclase, contemplando el proyectar los datos en espacios de más dimensiones que los originales y así lograr una mayor separabilidad de las clases, las SVM realizan sus procesos en dos fases: entrenamiento y decisión.

Las SVM están constituidas por algoritmos de aprendizaje supervisado, estas aparecieron en los años noventa, anteriormente ya se tenían avances que tomaron como ideas para su desarrollo, como fue el uso de los kernels, su interpretación geométrica y la construcción de un hiperplano de separación óptimo en un contexto no paramétrico [51], [96].

4.4.1. Fase de Entrenamiento.

Basada en la observación de un conjunto X de n muestras, las salidas del sistema son dos valores simbólicos $y \in \{+1, -1\}$ de forma que el conjunto de entrenamiento está formado por los pares $(x_i, y_i), i = 1 \dots n$ donde cada vector x_i se corresponde con un vector de entrenamiento y los valores $y_i \in \{+1, -1\}$ indican la clase a la que pertenece cada vector. El objetivo del proceso de entrenamiento consiste en encontrar una función de decisión capaz de separar las dos clases, en caso de que las clases no sean separables los vectores de entrenamiento se proyectan a un espacio de dimensión superior mediante el uso de funciones de transformación no lineales, de ser así, la función de decisión se sitúa en el hiperplano de esa dimensión, se define la función de decisión:

$$F(x) = \sum_{i=1}^n \alpha_i y_i H(x_i, x) - b \quad (4-12)$$

donde b es una constante, $\alpha_i, i = 1, \dots, n$.

Los patrones x_i asociados con valores de α_i distintos de cero se denominan vectores soporte, determinados los vectores soporte la función de decisión se muestra:

$$f(x) = \sum_{\text{support vectors}} \alpha_i y_i H(x_i, x) - b \quad (4-13)$$

Existen diferentes kernels para mapear el conjunto de datos de entrada:

1. Lineal

$$H(x, y) = x^t y \quad (4-14)$$

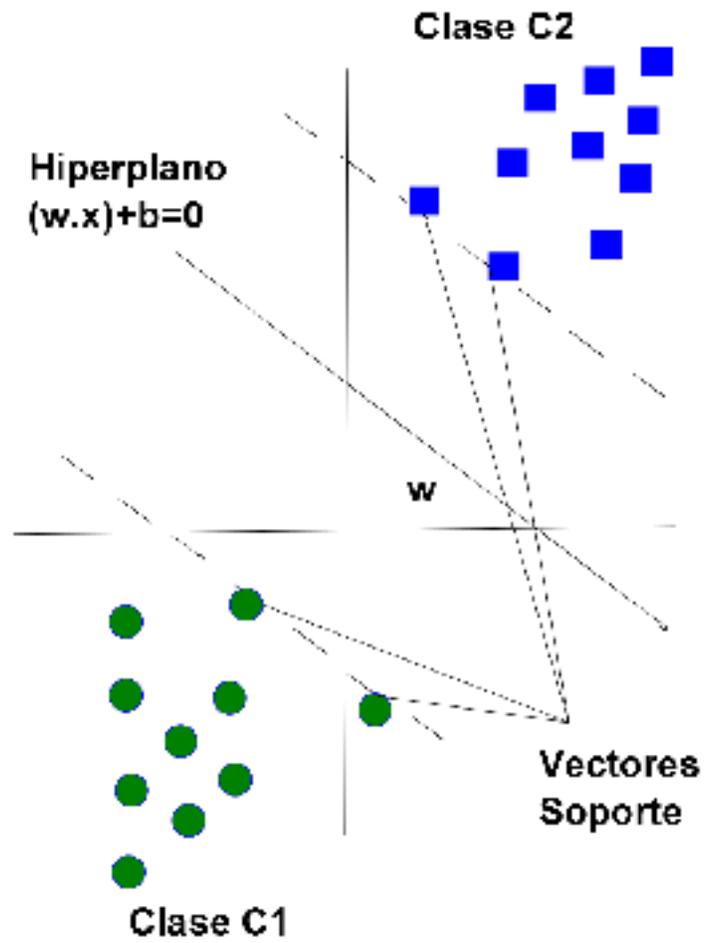
2. Funciones de base radial Gaussiana

$$H(x, y) = \exp\left\{-\frac{\|x - y\|^2}{2\sigma^2}\right\} \quad (4-15)$$

3. Exponencial

$$H(x, y) = \exp\left\{-\frac{\|x - y\|}{2\sigma^2}\right\} \quad (4-16)$$

optimo



2.png

Figura 4-1: SVM

4. Polinomios con grado d :

$$H(x, y) = (1 + \langle x, y \rangle)^d \quad (4-17)$$

5. Función sigmoide

$$H(x, y) = \tanh(p \langle x, y \rangle + \gamma) \quad (4-18)$$

donde $\langle x, y \rangle$ se refiere al producto interno con ρ y γ parámetros de ajuste.

4.4.2. Fase de prueba

Se determina la clase a la que pertenece de acuerdo con el signo (polaridad) de $f(x)$. La magnitud puede considerarse como una medida de certidumbre sobre la decisión.

Los vectores de soporte son los datos más representativos de todos, ya que los vectores de soporte son los patrones con mayor grado de información y que almacenando sólo estos vectores, se llega a la misma solución que con todos los patrones; esta ventaja se refleja en el ahorro de espacio en el almacenamiento de los resultados de aprendizaje.

La distancia mínima desde el hiperplano que separa las clases al patrón más cercano se denomina margen τ . Un hiperplano de separación se denomina óptimo si el margen es máximo; la distancia entre el hiperplano de separación y un patrón z es $y_k |f(x)| / \|w\|$ donde w está dado por:

$$w = \sum_{i=1}^n \alpha_i y_i x_i \equiv \sum_{\text{vectores soporte}} \alpha_i y_i x_i \quad (4-19)$$

la constante b se obtiene:

$$\alpha_i \{y_i [(w \cdot x_i) + b] - 1\} = 0 \quad (4-20)$$

Suponiendo que existe un margen τ todas las muestras de entrenamiento obedecen la desigualdad;

$$\frac{y_k f(x_k)}{\|w\|} \geq \tau, k = 1, \dots, n \quad (4-21)$$

donde $y_k = \{+1, -1\}$.

Para disminuir el número infinito de soluciones que difieren sólo en el escalado de w , se debe fijar la escala según el producto de τ y la norma de w . $\tau \|w\| = 1$. Maximizar el margen τ es equivalente a minimizar la norma de w .

Capítulo 5

Extracción de características

El sistema de reconocimiento visual humano trabaja a partir de la extracción de características particulares de un objeto, las cuales sirven para clasificar futuras imágenes de ese objeto u objeto similar. En el proceso de reconocimiento visual por computadora el proceso no es diferente. Para que la computadora pueda clasificar objetos, debe tener características del objeto a identificar, estas características son medidas que se obtienen de los píxeles que definen al objeto en una imagen. Las características se determinan de acuerdo a la imagen que se necesita reconocer y que tal imagen posea esas características. En los últimos años se han propuesto una gran cantidad de algoritmos para extracción de características. de forma general, estos se pueden clasificar en tres grandes grupos: geométricas, texturales y cromáticas. En éste Capítulo se estudian los métodos de extracción más comunes en cada grupo.

5.1. Características geométricas

Las características geométricas son aquellas propiedades de tamaño (área, perímetro), asimetría de la hoja acorde a un eje central, circularidad o redondez de la hoja, irregularidad del borde (forma que predomina en el borde del contorno de la hoja), diámetro y alargamiento de la hoja [76], [87].

Las propiedades métricas de las formas dependen de la distancia Euclídea, por lo tanto, una

propiedad métrica cambiará si el plano de la figura se distorsiona, propiedades métricas como el área, el perímetro y el centro de gravedad.

Perímetro:

Es el número de píxeles que forman el contorno de la forma, el contorno lo vamos a obtener binarizando la imagen, después se realiza el conteo de los píxeles.

Área:

Es el número de píxeles de la forma en la imagen binaria.

Diámetro:

Es la distancia entre dos puntos cualesquiera de la forma, también llamado como Eje mayor.

Circularidad o redondez:

Es la característica que nos permite determinar forma de la hoja, obteniendo el nivel de desviación de un círculo perfecto.

Contorno:

Línea de frontera que limita la forma de una imagen.

Alargamiento:

Distancia entre el punto del extremo superior al punto de extremo inferior de la hoja [41].

El área A es el número de píxeles contenidos dentro de su frontera. El perímetro P es la longitud de su frontera.

$$P = \sum \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \quad (5-1)$$

El área es el número de píxeles donde $I(x, y) = 1$

$$A = \sum_{x=1}^N \sum_{y=1}^M I(x, y) \quad (5-2)$$

El Centro de Gravedad (\bar{X}, \bar{Y}) es un único punto representativo de la región (mismo peso específico a todos los píxeles) y se obtiene:

$$\bar{X} = \frac{1}{A} \sum_i x_i \quad (5-3)$$

$$\bar{Y} = \frac{1}{A} \sum_i y_i \quad (5-4)$$

Si se necesita calcular el centro de gravedad teniendo en cuenta el valor de la intensidad en cada punto:

$$\bar{X} = \frac{\sum_i f(x_i, y_i) x_i}{\sum_i f(x_i, y_i)} \quad (5-5)$$

$$\bar{Y} = \frac{\sum_i f(x_i, y_i) y_i}{\sum_i f(x_i, y_i)} \quad (5-6)$$

Otra propiedad es la elongación o razón de aspecto, si fuese el caso de un rectángulo sería la razón de su longitud a su ancho. Un valor de uno sería un cuadrado y próximo a uno sería un rectángulo. Para generalizar este concepto a regiones de cualquier tipo, circunscribimos un rectángulo a la figura y calculamos la razón de aspecto del rectángulo circunscrito.

Los ejes mayor y menor de una región se definen en términos de su frontera y sirven para obtener la orientación de un objeto. La redondez de una región se define como la razón entre el área y el eje mayor al cuadrado y la compacidad como el cociente entre la raíz cuadrada del área y el eje mayor.

Para ajustar una recta a la región utilizando el ajuste por mínimos cuadrados:

Dado un conjunto de puntos en el plano $(x_i, y_i); i = 1 \dots n$, se trata de encontrar dos números a y b tales que la función de error sea mínima:

Función de Error.

$$\sum_{i=1}^n [(c_0 + c_1 x_i) - y_i]^2 \quad (5-7)$$

Mediante el uso de la pseudoinversa de una matriz cuyo origen está en la teoría de mínimos cuadrados, dado el conjunto de puntos (x_i, y_i) podemos escribir la siguiente ecuación matricial de la forma $Y a = b$.

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & x_n \\ Y \end{bmatrix} \begin{bmatrix} C_0 \\ C_1 \\ a \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \\ b \end{bmatrix} \quad (5-8)$$

Considerando $\|Ya - b\|^2$ debe ser mínimo, siendo la solución final.

$$\begin{bmatrix} C_0 \\ C_1 \end{bmatrix} = Y'b \quad (5-9)$$

Esta línea determina también la orientación de la región, mientras que el ángulo de esta línea se obtiene a partir:

$$S_x = \sum x_i \quad (5-10)$$

$$S_y = \sum y_i \quad (5-11)$$

$$S_{xx} = \sum x_i^2 \quad (5-12)$$

$$S_{yy} = \sum y_i^2 \quad (5-13)$$

$$S_{xy} = \sum x_i y_i \quad (5-14)$$

$$M_{xx} = S_{xx} - \frac{S_x^2}{A} \quad (5-15)$$

$$M_{yy} = S_{yy} - \frac{S_y^2}{A} \quad (5-16)$$

$$M_{xy} = S_{xy} - \frac{S_x S_y}{A} \quad (5-17)$$

Empleando la sig. ecuación se le dara la orientación:

$$\phi = \tan^{-1} \left\{ \frac{M_{xx} - M_{yy} + \sqrt{(M_{xx} - M_{yy})^2 + 4M_{xy}^2}}{2M_{xy}} \right\} \quad (5-18)$$

Para determinar la circularidad, partimos de perímetro y el área.

$$\text{Circularidad} : \frac{P^2}{4\pi A} \quad (5-19)$$

Compacidad.

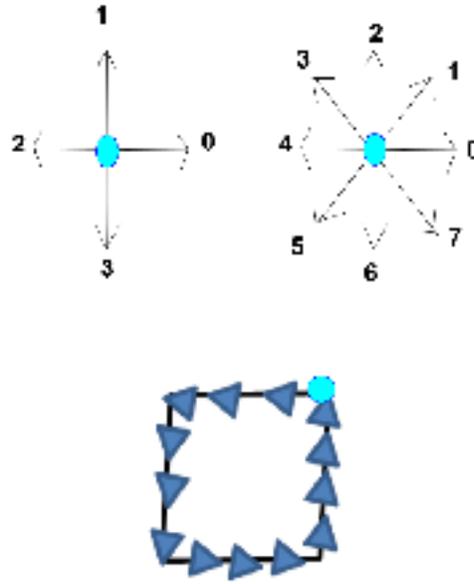
$$\frac{P^2}{A} \quad (5-20)$$

La detección de contorno u orillas de una imagen es primordial para su reconocimiento, la visión humana utiliza las orillas de manera jerárquica y la experiencia visual para reconocer objetos, ocasionalmente nuestra vista completa los bordes del objeto que están faltantes o distorsionados. Los descriptores de contornos se utilizan para describir la forma de los contornos de los objetos.

Códigos de Cadena:

El código de cadena se utiliza para representar un conjunto de puntos, que constituyen una línea recta o no, también se usa para representar una frontera como un conjunto de segmentos con longitud y dirección especificadas.

El código de cadenas describe los objetos mediante una sucesión de píxeles, se define partiendo de un pixel cualquiera del contorno y encadenando la dirección en que se encuentran los



Ejemplo: 2,2,2,3,3,3,0,0,0,1,1,1

Figura 5-1: Ejemplo de códigos de cadena

puntos adyacentes del contorno, con una conectividad de 4 o 8 pixeles con una determinada orientación; para comenzar en el primer pixel se tomara la información de la posición, es decir la dirección que sigue el contorno [92].

Cuando se utiliza un entorno de 4 vecinos, tendremos cuatro orientaciones (0,1,2,3) para los segmentos, por lo tanto para un entorno de 8 vecinos utilizaremos ocho orientaciones posibles, debes seleccionar el entorno que mejor te sirva para tu imagen. El código de cadena genera una secuencia de números que representan las orientaciones de los segmentos del contorno, se tomara un punto del contorno y se seguirá la dirección de las manecillas del reloj. Cuando nuestro objeto está formado por regiones no conectadas o con agujeros se necesita más de una cadena para representarlos, y se debe especificar si la cadena es de la parte exterior del objeto o a un agujero; cuando se haya terminado de obtener los arreglos o cadena del contorno, y nuestro objeto tiene agujeros, se puede reconstruir nuestro objeto mediante operaciones de relleno. Para cada segmento se construye su histograma de código de cadena, cada barra correspondiente

a cada número de código representa la frecuencia de aparición del número de código.

$C(i, j)$ es la frecuencia de aparición del número de un determinado código de cadena.

i designa el correspondiente código de cadena.

N es el cardinal de A

$$\text{Código de Cadena} = A = \{i | C(i)\} \quad (5-21)$$

$$\text{Número de Barras del histograma} = N = \text{cardinal}(A) \quad (5-22)$$

Transformada de Hough:

Dedicado para encontrar la ecuación de una línea que pase por un conjunto de n puntos en el plano xy . La transformada considera las relaciones globales entre los píxeles de borde, de tal forma que podemos encontrar patrones en la imagen como líneas o círculos [67].

Un conjunto en línea de píxeles se define:

$$y = ax + b \quad (5-23)$$

Se considera un punto (x_i, y_i) , se tienen diversas líneas que pasan por ese punto y que satisfacen la ecuación:

$$Y_i = ax_i + b \quad (5-24)$$

para valores variables de a y b ; de tal manera que si tomamos consideramos el plano ab o bien el espacio de parámetros y la ecuación:

$$b = -x_i a + y_i \quad (5-25)$$

tendríamos la ecuación de una única línea para un valor fijo de (x_i, y_i) ; para un segundo punto (x_j, y_j) se tendrá también una línea en el espacio de parámetros asociada, y esta línea interceptará con la línea asociada a (x_i, y_i) en (a', b') , donde a' es la pendiente y b' la altura a la que la recta que contienen a los dos puntos (x_i, y_i) y (x_j, y_j) corta al eje xy .

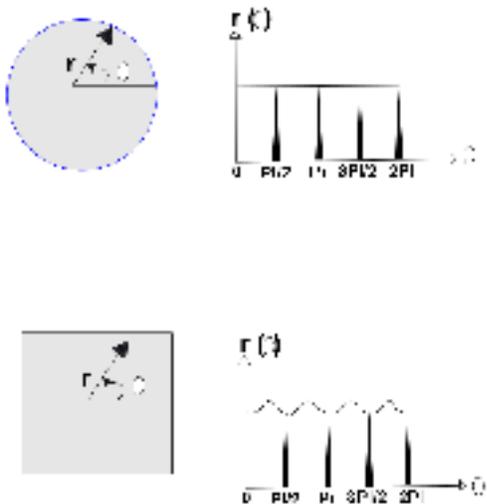


Figura 5-2: Transformada de Hough

Se debe subdividir el espacio de parámetros en las celdas de acumulación $A(i, j)$ que corresponde con un cuadrado asociado a las coordenadas (a_j, b_j) , se deben iniciar las celdas en 0; para cada punto (x_k, y_k) del plano de la imagen, se tomara como a , cada una de las subdivisiones del eje a y se calcula el b con la ecuación, una vez obtenidos los valores de b , estos se redondearan al valor más cercano del eje b .

Descriptores de Fourier:

Los descriptores de Fourier tienen la característica de ser invariantes a transformaciones geométricas y tolerantes ante ruido; se ha probado que los descriptores de Fourier son robustos e invariantes a las rotaciones, traslaciones y escalas.

El método de descriptores de Fourier (DFs) consiste en describir una forma de una figura plana cerrada a partir de su contorno, siendo el contorno en forma continua como una función en el tiempo considerando a éste como un par de señales $(x(t), y(t))$, donde la parte real se representa por la señal $x(t)$ y la parte imaginaria se representa por la señal $y(t)$, la curva de

contorno viene dada por las coordenadas paramétricas $(x(t), y(t))$, que son funciones periódicas, donde el parámetro t describe la longitud del camino de la línea de contorno calculada desde el punto de partida, primero debemos seleccionar N puntos equidistantes del contorno (muestreo), y a partir de estos valores obtenemos el vector complejo [59], [91].

$$(x(t), y(t)), n = 0, 1, 2, \dots, N - 1 \quad (5-26)$$

$$x(n) = x_n + jy_n \quad (5-27)$$

$$z = x + iy \quad (5-28)$$

siendo $x = (x(0), x(1), \dots, x(N - 1))$, $y = (y(0), y(1), \dots, y(N - 1))$

Empezamos en un punto arbitrario del contorno y siguiéndolo alguna dirección de avance, a favor de las manecillas del reloj, se tiene una secuencia de números complejos. La transformada discreta de Fourier $X = F(x)$ viene dada por la expresión:

$$X(k) = \sum_{n=0}^{N-1} x(n) \exp\left(\frac{-j2\pi kn}{N}\right), \quad 0 \leq k \leq N - 1 \quad (5-29)$$

$$Z(u) = \frac{1}{N} \sum_{n=0}^{N-1} z(n) \exp\left(-\frac{2\pi nu i}{N}\right), \quad u = 0, 1, 2, \dots, N - 1 \quad (5-30)$$

Estos valores se conocen con el nombre de descriptores de Fourier Cartesianos del contorno. De esta forma los descriptores de Fourier pueden representar un contorno cerrado arbitrario, generalmente se logra una buena descripción en base a pocos términos.

A partir de ellos se puede reconstruir el contorno mediante la transformación inversa:

$$z(n) = \frac{1}{N} \sum_{k=0}^{N-1} z(k) \exp\left(\frac{2\pi nk i}{N}\right), \quad n = 0, 1, 2, \dots, N - 1 \quad (5-31)$$

$$z(0) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) + i \frac{1}{N} \sum_{n=0}^{N-1} y(n) \quad (5-32)$$

$z(0)$ es el coeficiente que representa el punto medio o centro de gravedad del contorno del objeto, llamado centroide. Los coeficientes de Fourier $z(k)$ representan las variaciones en las tendencias del contorno para valores de k pequeños y las variaciones para los valores de k grandes. Por lo tanto los componentes de alta frecuencia tienen en cuenta los detalles más finos del contorno mientras que las componentes de baja frecuencia determinan la forma global del contorno, obteniéndose detalles más finos del contorno del objeto conforme mayor sea N .

Los coeficientes de Fourier se calculan directamente a partir de la posición del trazado del contorno: (Propiedades de los Descriptores de Fourier)

Normal:

$$Z(u) \quad (5-33)$$

Traslación

$$Z(u) = Z(u) + z_0 \delta(n) \quad (5-34)$$

Escala o Zoom

$$Z(u) = \alpha Z(u) \quad (5-35)$$

Punto de comienzo

$$Z(u) = Z(u) e^{-\frac{2\pi n u i}{N}} \quad (5-36)$$

Rotación

$$Z(u) = Z(u) e^{j\theta_0} \quad (5-37)$$

Reflexión

$$Z(u) = Z(-u) e^{(j2\theta_0)} + 2\gamma \delta(u) \quad (5-38)$$

Momentos:

Los momentos nos ayudan a determinar medidas estadísticas que permiten describir a los objetos, como tamaño, posición, orientación y la forma. Los momentos espaciales de un objeto

son medidas estadísticas que permiten describir regiones, acorde a sus puntos interiores [89].

Tenemos $f(x, y)$ como la intensidad del punto (x, y) en una región, el momento de orden $(p + q)$ para una región se define: donde la sumatoria se toma sobre todas las coordenadas espaciales (x, y) de puntos de la región.

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad (5-39)$$

El momento central de orden $(p + q)$ viene dado por:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q I(x, y) \quad (5-40)$$

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad (5-41)$$

$$\bar{y} = \frac{m_{01}}{m_{00}} \quad (5-42)$$

Los momentos centrales normalizados de orden $(p + q)$ se definen:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad (5-43)$$

donde, para $(p + q) = 2, 3, \dots$

$$\gamma = \frac{p + q}{2} + 1 \quad (5-44)$$

El momento espacial de orden cero es la suma de los valores de brillo de los píxeles. Corresponde al área en imágenes binarias.

$$m_{00} = \sum_x \sum_y I(x, y) \quad (5-45)$$

Los momentos espaciales de orden uno, es la sumatoria en la dirección x, y , de los valores de brillo de los píxeles del objeto, sucesivamente de la sumatoria se multiplica por su valor de coordenada x, y en la imagen.

$$m_{10} = \sum_x \sum_y xI(x, y) \quad (5-46)$$

$$m_{01} = \sum_x \sum_y yI(x, y) \quad (5-47)$$

Los momentos de orden dos, nos dan información sobre la orientación del objeto. El conjunto de momentos de Hu se pueden obtener usando los momentos centrales normalizados de ordenes 2 y 3, son invariantes para translaciones, rotaciones y escalados.

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + 3\eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})^2(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - (3\eta_{21} + \eta_{03})^2] + \\
&\quad 3(\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= [(3\eta_{21} + \eta_{03})(\eta_{30} + \eta_{12})(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - \\
&\quad (\eta_{30} + 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - \eta_{21} + \eta_{03})^2]
\end{aligned} \quad (5-48)$$

Los siete momentos obtenidos son los valores que conforman el vector de características.

Elipse:

Las características se derivan de una elipse que se aproxima al contorno de la región, se busca una elipse que sea lo más parecida al borde de la región, para después extraer las características de la elipse tales características como eje mayor, eje menor, orientación y posición de la elipse.

Los bordes de una región se definen como (x_i, y_i) para $i = 1, \dots, L$; la ecuación de una elipse se define:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (5-49)$$

Se busca un vector a que haga que la ecuación anterior, sea cercana a cero, esto se logra minimizando de esta forma:

$$\|Xa\| \rightarrow \text{mín} \quad (5-50)$$

donde X es una matriz de L filas cuya fila i es x_i por lo tanto se restringe $\|a\| = 1$, a es la última columna de V , donde $X = USV^T$, donde S es una matriz diagonal de la misma dimensión que X con elementos no negativos ordenados en forma descendente, y las matrices U y V son matrices unitarias.

Las características de la eclipse se obtienen:

$$\left(\frac{x - x_0}{a_e}\right)^2 + \left(\frac{y - y_0}{b_e}\right)^2 = 1 \quad (5-51)$$

donde:

$$a_e = \frac{1}{\sqrt{sa_p}} \quad (5-52)$$

$$b_e = \frac{1}{\sqrt{sb_p}} \quad (5-53)$$

$$s = \frac{1}{v - f} \quad (5-54)$$

$$v = t^T T t \quad (5-55)$$

$$T = \begin{matrix} a & b/2 \\ b/2 & c \end{matrix} \quad (5-56)$$

$$t = \frac{1}{2} T^{-1} \begin{bmatrix} d \\ e \end{bmatrix} \quad (5-57)$$

$$a_p = a \cos^2(\alpha) + b \cos(\alpha) \sin(\alpha) + c \sin^2(\alpha) \quad (5-58)$$

$$b_p = a \sin^2(\alpha) - b \cos(\alpha) \sin(\alpha) + c \cos^2(\alpha) \quad (5-59)$$

$$\alpha = \frac{1}{2} \tan^{-1}\left(\frac{b}{a-c}\right) \quad (5-60)$$

$$x_0 = t_1, y_0 = t_2 \quad (5-61)$$

Definimos los ejes del elipse como a_e y b_e , el centro de origen (x_0, y_0) y la orientación de la elipse es α . Partiendo de los ejes definimos la excentricidad como:

$$e_x = \frac{\min(a_e, b_e)}{\max(a_e, b_e)} \quad (5-62)$$

Para regiones circulares o redondez la excentricidad sera 1, y para regiones más ovaladas la excentricidad disminuye su valor $0 \leq e_x \leq 1$.

Invariantes de Flusser

Los momentos invariantes de Flusser son derivadas de los momentos invariantes de Hu; pueden ser usados para el reconocimiento de objetos deformados [53], [64].

Su estructura matemática de los momentos de Flusser y Suk se expresa:

$$u = a_0 + a_1x + a_2y \quad (5-63)$$

$$v = b_0 + b_1x + b_2y \quad (5-64)$$

donde (x, y) y (u, v) son las coordenadas en el plano de la imagen antes y después de la transformación.

Momentos de Flusser y Suk:

$$M_{F1} = \frac{1}{\mu_{00}^4} (\mu_{20}\mu_{02} - \mu_{11}) \quad (5-65)$$

$$M_{F2} = \frac{1}{\mu_{00}^{10}} (\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{03}\mu_{12}^3 + 4\mu_{03}\mu_{21}^3) - \frac{1}{\mu_{00}^{10}} 3\mu_{21}^2\mu_{12}^2 \quad (5-66)$$

$$M_{F3} = \frac{1}{\mu_{00}^7} (\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12})) - \frac{1}{\mu_{00}^7} \mu_{02}(\mu_{30}\mu_{12} - \mu_{12}^2) \quad (5-67)$$

$$M_{F4} = \frac{1}{\mu_{00}^{11}} (\mu_{20}^3\mu_{03}^3 - 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}^2\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^2\mu_{02}\mu_{12}^2) + \frac{1}{\mu_{00}^{11}} (12\mu_{20}\mu_{11}^2\mu_{21}\mu_{03} + 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03} - 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12}) + \frac{1}{\mu_{00}^{11}} (-8\mu_{11}^3\mu_{30}\mu_{03} - 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03} - 9\mu_{20}\mu_{02}^2\mu_{21}^2) \quad (5-68)$$

$$M_{F5} = \frac{1}{\mu_{00}^6} (\mu_{40}\mu_{04} - 4\mu_{31}\mu_{12} + 3\mu_{22}^2) \quad (5-69)$$

$$M_{F6} = \frac{1}{\mu_{00}^9} (\mu_{40}\mu_{04}\mu_{22} + 2\mu_{31}\mu_{22}\mu_{12} - \mu_{40}\mu_{12}^2 - \mu_{04}\mu_{31}^2 - \mu_{22}^3) \quad (5-70)$$

donde μ_{pq} :

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (5-71)$$

Histograma de Gradientes Orientados (HOG).

Descriptor de características, utilizando la detección de bordes, está técnica cuanta las ocurrencias de las orientaciones de los gradientes en partes específicas de una imagen, calcula un espacio de celdas superpuestas con el fin de normalizar las muestras locales y aumentar la precisión [50].

Un descriptor de HOG, se debe dividir la imagen en un cierto número de sub-ímagenes del mismo tamaño, denominadas celdas, las cuales se agrupan en bloques con un mismo número de celdas de ancho y alto todos, el avance en bloques horizontalmente se realiza eliminando

la columna de celdas de la izquierda y añadiendo la columna de la derecha y verticalmente eliminando la fila de celdas de arriba y añadiendo la fila de celdas de abajo.

Una imagen A de tamaño $W \times H$, tamaño de celda $C_W \times C_H$ con $W \bmod C_w = 0$ y $H \bmod C_H = 0$ y un tamaño de bloque en celdas $B_W \times B_H$, el ancho y el alto de una imagen en celdas W_C y H_C , y el número de bloques distribuidos horizontalmente y verticalmente N_{BW} y N_{BH} :

$$W_c = \frac{W}{C_W} \quad (5-72)$$

$$H_c = \frac{H}{C_H} \quad (5-73)$$

$$N_{BW} = 1 + W_c - B_W \quad , \quad N_{BH} = 1 + H_c - B_H \quad (5-74)$$

N_c es el número total de celdas y N_b es el número total de bloques resultantes de la imagen A se definen como :

$$N_C = W_c \times H_c \quad (5-75)$$

$$N_B = N_{BW} \times N_{BH} \quad (5-76)$$

El descriptor calcula el Hog de cada celda y cada bloque agrupa los Hogs de sus celdas, se define el número de Hogs que contiene un descriptor:

$$N_{HOG} = B_W \times B_H \times N_B \quad (5-77)$$

Al dividir el Hog en n clases, dado que cada bloque contiene $B_W \times B_H$, para el número total de valores N_v será:

$$N_v = n \times N_{HOG} \quad (5-78)$$

Por cada píxel dentro de una celda se acumulan los histogramas de orientación de los gra-

dientes, estos histogramas capturan propiedades (bordes) de forma local dentro de la celda, estos histogramas son invariantes a chicas deformaciones, al finalizar se debe normalizar cada histograma de las celdas con respecto de la energía del gradiente en un vecindario alrededor de la celda, la energía del gradiente se obtiene multiplicando cada posición del vector de características, la energía se calcula:

$$N_{\delta\gamma}(i, j) = (C(i, j)^2 + C(i, \delta, j)^2 + C(i, j + \gamma)^2 + C(i + \delta, j + \gamma)^2)^{\frac{1}{2}} \quad (5-79)$$

donde $\delta, \gamma \in \{-1, 1\}$.

HOG Pirámide

Consiste en hacer varias iteraciones de una misma imagen con distintas resoluciones, donde cada nivel de resolución se extraen las características más representativas extraídas por el descriptor de Hog.

5.2. Características texturales

Los estudios de la textura de una imagen se han dado mediante técnicas como funciones de autocorrelación, mallas de Markov de primero y segundo orden restringidas, espectro de potencia, frecuencias relativas de varios niveles de gris sobre una imagen no normalizada, técnica que asciende al modelo estadístico de segundo orden con run length para calcular 5 descriptores en las direcciones 0, 45, 90 y 135 grados, entre otras.

La textura es una característica utilizada para la segmentación, identificación de objetos o regiones de interés en una imagen y obtención de forma.

“Una textura está definida por la uniformidad, densidad, grosor, rugosidad, regularidad, intensidad y direccionalidad de medias discretas del tono en los píxeles y de sus relaciones espaciales” según [46] en las imágenes digitales las texturas están dadas por:

1. La frecuencia de cambio de los tonos de los píxeles.
2. La dirección o direcciones de cambio.

3. El contraste entre un píxel y sus vecinos.
4. El tono o el color es un factor muy independiente de la textura.

Para caracterizar las texturas de una imagen es necesario procesar la imagen y extraer su información mediante descriptores de frecuencia, descriptores estructurales, descriptores basados en modelos y descriptores probabilísticos.

1.- Descriptores de Frecuencia:

Pretenden obtener la transformada en frecuencia de la imagen y a partir de ésta se obtendrán ciertas características, características fáciles de obtener del espectro en coordenadas polares.

Estos descriptores se basan en el análisis del espectro de Fourier mediante la transformada de Fourier; el espectro de Fourier describe la direccionalidad de patrones bidimensionales periódicos de la imagen, estos patrones de textura son distinguibles como concentraciones altas de energía en el espectro, tales características vienen dadas por una Magnitud de picos prominentes en frecuencia que proporcionan información sobre la direccionalidad de los patrones de textura y por una Localización de los picos que proporcionan información sobre el periodo espacial de los patrones.

Cuando las texturas son heterogéneas o de carácter aleatorio este descriptor se restringe por su poca información que puede arrojar, ya que realiza una descomposición de las señales en sus componentes frecuenciales.

Descriptor espectral de Fourier da buenos resultados para imágenes de textura homogéneas pero sólo es útil para texturas con patrones periódicos y su análisis solo lo realiza en frecuencia no en temporal [90].

Descriptor espectral Wavelet es el único que permite el análisis multi resolución que permite localizar en espacio y frecuencia, otra ventaja es que no incrementa el número de datos; tiene un elevado costo computacional debido a que tiene varios niveles de descomposición [58].

Transformada de Gabor

La transformada de Gabor o Filtro de Gabor entrega información textural mediante las frecuencias espaciales de la imagen, al igual que su orientación; utiliza la introducción de ventanas para poder observar la parte de la señal en algún punto de estudio, se considera eficiente

el uso de una función gaussiana para utilizarse de ventana, ya que la función gaussiana es no nula donde se alcanza la máxima concentración simultanea en tiempo y frecuencia. Los filtros de Gabor se definen como ondas planas que están restringidas por una función gaussiana, se obtiene la medida del filtro para una frecuencia dada en un área local de la imagen con ayuda de la transformada de Wavelet Gabor [93], [62], [88], [77], [66].

$$\Psi_f(\vec{x}) = \frac{k_j^2}{\sigma^2} \exp\left(-\frac{k_j^2 x^2}{2\sigma^2}\right) \left[\exp(i\vec{k} \cdot \vec{x}) - \exp\left(\frac{\sigma^2}{2}\right) \right] \quad (5-80)$$

$i\vec{k}$ donde \vec{x} denota la dirección del vector, restringida por una función envolvente gaussiana de la forma:

$$\exp\left(-\frac{k_j^2 x^2}{2\sigma^2}\right) \quad (5-81)$$

2.- Descriptores Estructurales:

Son técnicas que describen las texturas como la composición de elementos de textura bien definidos, como líneas paralelas espaciadas, o concatenaciones de figuras geométricas de formas conocidas como pentágonos u otras formas geométricas; se basan en obtener patrones primitivos de la textura, la textura puede analizarse bajo la perspectiva de computar propiedades estadísticas de los elementos extraídos o bien extraer la regla de colocación que describa la textura. Dada la naturaleza de las texturas es difícil trabajar con métodos estructurales ya que no todas las texturas tienen un elemento que se repita.

Obtienes resultados rápidos si la textura está formada por un elemento estructural reconocible por el sistema, pero este descriptor solo es útil para texturas con patrones periódicos y sólo realiza análisis en frecuencia, no el análisis temporal.

3.- Descriptores basados en modelos:

Estas técnicas se fundamentan en la construcción de un modelo empírico de cada píxel en la imagen basado en una suma ponderada de las intensidades de los píxeles en la vecindad del primero. Los parámetros estimados de los modelos así generados se utilizan como descriptores de las características de la textura. Los parámetros estimados sirven para describir y sintetizar

una textura, ya que reúnen las características esenciales que caracterizan a la textura. Como ejemplo de estos descriptores son los modelos auto regresivos, los campos aleatorios de Markov o los modelos fractales.

3.1.-Cadenas de Markov:

Los métodos que utilizan los campos aleatorios de Markov gaussianos caracterizan la relación estadística entre un píxel y sus vecinos; la intensidad de un píxel depende de las intensidades de los píxeles vecinos, el modelo resultante consta de un número de parámetros igual al tamaño de la máscara de la vecindad, los parámetros se pueden estimar mediante un algoritmo de mínimos cuadrados sobre cada posición de la máscara en la imagen.

La máscara se forma tomando una serie de píxeles vecinos, siendo la mas sencilla la de primer orden, con los cuatro píxeles directamente conectados y la de segundo orden es formada por los ocho píxeles que lo rodean.

Sus resultados con cadenas de Markov son muy buenos para micro texturas e interrelaciona píxeles vecinos; pero tiene un nivel de complejidad alta y necesita de definir una máscara adecuada.

3.2.- Modelos Fractales:

Muestra la correlación existente entre la dimensión fractal de una textura y su tosquedad, la descripción fractal se basa en la determinación de la dimensión fractal.

La propiedad de Auto-similaridad indica que la dimensión fractal de una imagen es independiente de la escala; un conjunto acotado A se dice que es auto-similar si está constituido por la unión de N copias distintas de si mismo, cada una de ellas escalada con un factor r , la dimensión fractal correspondiente a dicha región viene dada por [49]:

$$D = \frac{\log(Nr)}{\log(r^{-1})} \quad (5-82)$$

La dimensión fractal D muestra un panorama de la rugosidad de una superficie, mientras mayor sea D mas rugosa será la superficie.

Da buenos resultados para superficies rugosas y sólo es útil para texturas rugosas.

4.- Descriptores Probabilísticos:

Los descriptores probabilísticos son metodos simples y se apoyan en estadísticas de histogramas de niveles de gris de una imagen, estos estadísticos pueden ser de primer, segundo o tercer orden.

Primer Orden: las medidas texturales de primer orden son calculadas a partir de los valores de gris originales de la imagen y su frecuencia, como la media, varianza, desviación estándar; no se considera la relación entre píxeles. Son sencillos y rápidos y necesitan poca memoria, pero no tienen en cuenta la información dada por la posición relativa de los píxeles.

Segundo Orden: son las medidas que consideran la relación de co-ocurrencia entre grupos de dos píxeles de la imagen original y a una distancia dada. Es adecuada para texturas de origen aleatorio y que no respondan a un patrón determinado y tiene en cuenta la posición relativa de los píxeles entre sí, a pesar de su alta carga computacional.

Tercer Orden y más: se consideran las relaciones entre 3 o más píxeles. Su cálculo es posible pero no se implementa porque requiere mucho tiempo y su resultado es difícil de interpretar.

Matriz de Co-ocurrencia:

La Matriz de Co-ocurrencia describe la frecuencia de un nivel de gris que aparece en una relación espacial específica con otro valor de gris, dentro del área de una ventana determinada, esta matriz indica la forma en que los valores de los píxeles ocurren al lado de otro valor [83].

La matriz de co-ocurrencia considera la relación espacial entre dos píxeles, llamados píxeles de referencia y píxel vecino, si se escoge el píxel vecino que está situado a la derecha de cada píxel de referencia, se expresaría como $(1,0)$, donde 1 es el píxel en la dirección x y 0 el píxel en la dirección y. Cada píxel se va convirtiendo sucesivamente en el píxel de referencia, empezando por el ubicado arriba a la izquierda y finaliza abajo en la derecha, los píxeles ubicados en el margen derecho de la imagen original, no tienen vecino a la derecha por lo tanto no son usados (píxel vecino, píxel de referencia). Después se realiza la normalización de la matriz de co-ocurrencia.

Normalización de la matriz.

En la matriz precedente, se cuenta cada píxel de referencia con su vecino a la derecha. Si el cálculo se realiza solo de este modo, usando solo una dirección, la matriz no es simétrica respecto

de la diagonal. Sin embargo la simetría es necesaria para el cálculo, esto se logra si cada par de píxeles se cuentan dos veces, una vez a la derecha y otra vez a la izquierda (intercambiando los píxeles de referencia y vecino en el segundo cálculo).

Otra forma mas sencilla es sumarle a esta matriz su matriz traspuesta. La matriz traspuesta se logra intercambiando las filas y columnas de la matriz de co-ocurrencia, sumando cada elemento de la matriz original y su traspuesta se llega a la matriz simétrica para una relación horizontal (derecha+izquierda). El paso siguiente es expresar esta matriz como probabilidad.

Probabilidad es el número de veces que un evento ocurre, dividido por el número total de posibles eventos, se cálculo bajo la siguiente ecuación:

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i,j=0}^{N=1} V_{i,j}} \quad (5-83)$$

Donde:

i es el número de filas y j el número de columnas

V es el valor de la celda i, j

$P_{i,j}$ es la probabilidad en la celda i, j

N es el número de filas o columnas

Se realiza la ecuación en la matriz simétrica, tomando en cuenta que el número de pares totales $\sum_{i,j=0}^{N=1} V_{i,j}$ es tomada para una relación a la izquierda (1,0), considerando también la relación a la derecha (-1,0) para construir la matriz horizontal, los pares se duplican, de esta forma se ira llenando una nueva matriz nombrada, matriz normalizada horizontal de la matriz original

De la matriz normalizada pueden derivarse diferentes medidas texturales como la homogeneidad, contraste, disimilaridad, desviación estándar, correlación, energía, entropía, etc.

Las matrices de co-ocurrencia capturan propiedades de texturas pero no pueden ser utilizadas directamente para comparar dos texturas. Utilizando la matriz de co-ocurrencia normalizada se pueden calcular ciertas características que describen las texturas de manera más compacta (descriptores de Texturas). Un problema al derivar mediciones de texturas utilizando

la matriz de co-ocurrencia es cómo elegir el tamaño del desplazamiento d . Se sugiere utilizar la prueba estadística de X^2 para seleccionar los valores de d que tengan la mayoría de la estructura (maximizar el valor):

$$X^2(d) = \left(\sum_i \sum_j \frac{N_d^2[i, j]}{N_d[i]N_d[j]} - 1 \right) \quad (5-84)$$

Donde:

$$N_d[i] = \sum_j N_d[i, j].$$

$$N_d[j] = \sum_i N_d[i, j].$$

Las técnicas de análisis de textura describen la distribución de niveles de gris, calculando variables estadísticas que describan propiedades como la variabilidad, el contraste, valor medio, entre otras. Una forma de discernir entre diferentes texturas es comparar sus estadísticas del nivel de gris de primer orden. Por primer orden se entienden las estadísticas en las que se ven involucrados píxeles simples en contraposición a las estadísticas de más de un píxel (pares, tripletes). Se puede utilizar el histograma de nivel de gris de la textura, donde la normalización proporciona la función de densidad de probabilidad de la imagen de acuerdo a la textura, se pueden comparar los histogramas normalizados del nivel de gris de imágenes de texturas o utilizar medidas derivadas como la media, la mediana o la varianza.

Sea z una variable aleatoria que representa la intensidad discreta de la imagen, sea $p(z_i)$ donde $i = 1, 2, \dots, L$ su histograma correspondiente, donde L es el número de niveles de intensidades diferentes. El momento n -ésimo de z respecto de la media se define como:

$$\mu_n(z) = \sum_{i=1}^L (z_i - m)^n p(z_i) \quad (5-85)$$

donde m es el valor medio de z y la intensidad media de la imagen es:

$$m = \sum_{i=1}^L z_i p(z_i) \quad (5-86)$$

El momento de segundo orden, denominado varianza $\sigma^2(z)$, es gran importancia para la descripción de texturas, ya que la varianza es una medida del contraste de intensidad que se puede usar para obtener descriptores de suavidad relativa;

$$R = 1 - \frac{1}{1 + \sigma^2(z)} \quad (5-87)$$

es 0 para áreas con intensidad constante $\sigma^2(z) = 0$ y para valores grandes de $\sigma^2(z) = 1$.

El momento de tercer grado es una medida de la oblicuidad del histograma, mientras que el de cuarto orden es una medida de cuán plano es el histograma, A partir del momento de quinto orden no se es tan fácil relacionar los momentos con la forma del histograma, pero son utilizados para obtener más información cuantitativa de discriminación de texturas.

A partir de estadísticas del nivel de gris de segundo orden se obtienen las matrices de dependencia espacial del nivel de gris. Sea P un operador de posición definido por $P = (d\cos\theta, d\sin\theta)$ y sea A una matriz de dimensión $k \times k$ cuyo elemento a_{ij} es el número de veces que los pixeles cuya intensidad es z_i aparecen (en la posición especificada por P) en relación a puntos cuya intensidad es z_j con $1 \leq i, j \leq k$, Tomaremos a n como el número total de pares de puntos de la imagen que satisfacen a P . Si definimos una matriz C como la formada de dividir cada elemento de A por n , entonces cada elemento de la matriz C es una estimación de la probabilidad compuesta de que un par de puntos que satisfagan P tengan valores (z_i, z_j) . La matriz C se le nombra matriz de coocurrencia del nivel de gris, dado que C depende de P , es posible detectar la presencia de unos patrones de textura dados eligiendo adecuadamente el operador de posición. Se pretende analizar una matriz C , para clasificar la textura de la región sobre la cual se calculó C . Propusieron un conjunto de descriptores utilizando los coeficientes de la matriz C [47].

Descriptores de Haralick

Los descriptores de textura definidos por Haralick, son un conjunto de medidas de textura basadas en la matriz de co-ocurrencia, para emplear estos descriptores es necesario que la información textural de la imagen este contenida en relaciones espaciales que se dan entre los niveles de gris de un objeto. Los descriptores son homogeneidad, contraste, disimilaridad,

entropía, asm, media, varianza y correlación.

1. Probabilidad máxima:

$$\max_{i,j}(c_{ij}) \tag{5-88}$$

2. Energía:

Cuanto mas suave es la textura, mayor valor toma la energía, (intensidad constante, *energía* = 1, para distribución uniforme de parejas, *energía* = $p < 1$).

$$\sum_i \sum_j^{255} |c_{ij}|^2 \tag{5-89}$$

3. Momento de distinción de elementos de orden k

$$\sum_i \sum_j (i - j)^k c_{ij} \tag{5-90}$$

4. Momento inverso de distinción de elementos de orden k

$$\sum_i \sum_j \frac{c_{ij}}{(i - j)^k} \quad i \neq j \tag{5-91}$$

5. Entropía:

Toma valores bajos cuanto mas suave es la textura (menos aleatoría).

$$-\sum_i \sum_j c_{ij} \log c_{ij} \tag{5-92}$$

6. Uniformidad:

$$\sum_i \sum_j c_{ij}^2 \tag{5-93}$$

7. Correlación:

Utilizando la matriz de co-ocurrencia normalizada los resultados de la correlación es entre -1

y 1. Algunas propiedades de la correlación son: Un objeto tiene mas alta correlación dentro de el que entre otros adyacentes y los píxeles cercanos están mas correlacionados entre sí que los píxeles mas distantes.

$$\frac{1}{\sigma_x \sigma_y} \sum_i \sum_j (i - \mu_x)(j - \mu_y) c_{ij} \quad (5-94)$$

donde:

$$\mu_x = \sum_i i \sum_j c_{ij} \quad (5-95)$$

$$\mu_y = \sum_j j \sum_i c_{ij} \quad (5-96)$$

$$\sigma_x^2 = \sum_i (i - \mu_x)^2 \sum_j c_{ij} \quad (5-97)$$

$$\sigma_y^2 = \sum_j (j - \mu_y)^2 \sum_i c_{ij} \quad (5-98)$$

8. Inercia:

Cuanto mayor es el contraste de la imagen, mayor es el valor de Inercia al ponderar las diferencias $(i - j)$ con $(i - j)^2$.

$$\sum_i \sum_j (i - j)^2 c_{ij} \quad (5-99)$$

9. Homogeneidad local:

Cuanto menor es el contraste, más alto es la homogeneidad debido a la ponderación inversa de $(i - j)^2$.

$$\sum_i \sum_j \frac{1}{1 + (i - j)^2} c_{ij} \quad (5-100)$$

10. Disimilaridad:

Similar al contraste o inercia, es una medida de la variación local en la imagen; tiene un valor alto cuando la región dentro de la escala de la ventana tiene un alto contraste.

$$\sum_i \sum_{j=0}^{N=1} (i-j)c_{ij} \quad (5-101)$$

11. Desviación Estándar:

Mismos resultados para i como para j pues la matriz de co-ocurrencia es simétrica.

$$\sigma_i^2 = \sum_i \sum_{j=0}^{N=1} c_{ij} (i - \mu_i)^2 \quad (5-102)$$

$$\sigma_j^2 = \sum_i \sum_{j=0}^{N=1} c_{ij} (i - \mu_j)^2 \quad (5-103)$$

finalmente desviación estándar, se basa en la media y la dispersión alrededor de la media de los valores de las celdas de la matriz de co-ocurrencia. La varianza utiliza la combinación entre los pixeles de referencia y vecino, no es la misma que la varianza de los niveles de grises de la imagen original.

$$\sigma_i = \sqrt{\sigma_i^2} \quad (5-104)$$

$$\sigma_j = \sqrt{\sigma_j^2} \quad (5-105)$$

5.3. Características Cromáticas

Todos los colores son vistos como combinaciones variables de los tres colores primarios: red, green y blue, las características de brillo, matíz y saturación son utilizadas para determinar una distinción entre los colores.

El matiz lo asociamos con la longitud de onda dominante en la mezcla de longitudes de onda de la luz (color que la persona percibe); El brillo nos indica la intensidad (luz) de una color; La saturación es la cantidad de luz blanca mezclada con un color, lo cual nos indica la pureza del color y si un color está muy saturado este se muestra con un color vivo e intenso, y uno menos saturado se ve descolorido y gris (al no tener la imagen saturación, sus colores se convierten en tonalidad de gris).

Al tomar de manera conjunta al matiz y la saturación se le denomina cromaticidad, entonces una imagen se distingue su color por la cromaticidad y su brillo [76].

Descriptor de la Distribución del Color (DDC).

Es una representación del color de resolución invariante, está diseñado para capturar la distribución espacial del color en un imagen, para realizar el proceso de extracción de características de color, primero realizamos un análisis del color mas representativo o la rejilla basada en la selección del color más representativo y después realizamos la Transformada Discreta del Coseno mediante la cuantización.

Este descriptor se divide en cuatro etapas: división de la imagen, selección del color más representativo, transformada DCT, y la exploración zig-zag.

Se recomienda que en la división de la imagen, se divida la imagen en 64 bloques para obtener la invariabilidad de la resolución o escalado.

La selección del color, se realizara en un único color como representante de cada bloque, para seleccionar el color más representativo se obtendrá la media de los colores de los píxeles en un mismo bloque; una vez realizada la selección se guardaran en una imagen con tamaño 8x8.

Se emplea la transformada DCT y se realiza la exploración en zig-zag, iniciando en el primer píxel de la matriz, partiendo a la derecha un píxel, y después se realiza con orientación en diagonal hacia abajo, aproximamos al píxel que esta abajo del píxel final de la diagonal y realizamos el mismo proceso pero con la diagonal hacia arriba.

Contraste.

El contraste se refiere a la variación de intensidades en un objeto, si hay mucha variación de intensidades se determina que el contraste es alto y si hay poca variación de intensidades

determinamos que el contraste es bajo. El contraste en una medida para la diferencia de color entre la región y su entorno, la región y el entorno no tienen píxeles comunes, la agrupación de la región y el entorno forman una zona definida [85]:

$$g[i, j] = x[i + i_r, j] + j_r \quad (5-106)$$

$i = 1, \dots, 2h + 1$; $j = 1, \dots, 2w + 1$; donde h y w son la altura y ancho de la región; los puntos centrales de esa zona se define como:

$$i_r = \bar{i} - h - 1 ; j_r = \bar{j} - w - 1 \quad (5-107)$$

(\bar{i}, \bar{j}) corresponden al centro de gravedad de la región. Entre más pequeña sea la diferencia de la variable de color en la región con respecto a su entorno, más pequeño será el contorno.

$$K_1 = \frac{G - G_e}{G_e} \quad (5-108)$$

Se definió el contraste en k_1 donde G y G_e representan el promedio de la variable de color en la región y en el entorno.

DCT- Transformada discreta del coseno para la distribución del color

Es una transformada basada en la Transformada de Fourier discreta (DFT), aunque solo actúa sobre funciones periódicas con simetría par y el resultado es una secuencia de números reales; la DCT expresa una secuencia finita de varios puntos como resultado de la suma de distintas señales cosenoidales en frecuencias múltiples, la DFT trabaja con exponenciales complejos y la DCT trabaja con una serie de números finitos; por lo tanto la transformada de coseno discreta es una función lineal invertible.

$$f = R^N \rightarrow R^N \quad (5-109)$$

Donde R indica al conjunto de los números reales; además de ser una forma equivalente a una matriz cuadrada de $N \times N$. Formula para calcular la DCT:

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N} \quad (5-110)$$

$$0 \leq p \leq M-1, 0 \leq q \leq N-1$$

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \leq p \leq M-1 \end{cases} \quad (5-111)$$

$$\alpha_q = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leq q \leq N-1 \end{cases} \quad (5-112)$$

Capítulo 6

Metodología

En años recientes, el análisis de imágenes ha atraído la atención de muchos investigadores en el área de procesamiento de imágenes y reconocimiento de patrones. Los tipos de tareas en análisis de imágenes son complicados. Estas tareas incluyen de forma general a métodos de pre-procesamiento para mejorar la imagen y aumentar la precisión de segmentación, métodos de segmentación, extractores de características y técnicas de clasificación. Cada uno de estas fases de un sistema es muy importante para un buen desempeño, sin embargo en esta tesis se realizara un especial énfasis en dos tareas, extracción de características y clasificación. Para analizar una imagen es necesario obtener las características de la región que la componen. Una característica puede estar en función de una o más medidas, calculadas de tal forma que estas miden un rasgo significativo del objeto en la imagen.

Los métodos de extracción de características más utilizados en los últimos años son los métodos de extracción de características geométricas, sin embargo en la metodología propuesta se analiza la influencia de las características texturales y cromáticas en el desempeño de los clasificadores. En este capítulo se detallan tanto la metodología propuesta como las técnicas de estudio utilizadas para el análisis de las hojas, y finalmente la solución para la identificación de las plantas a la que se llegó.

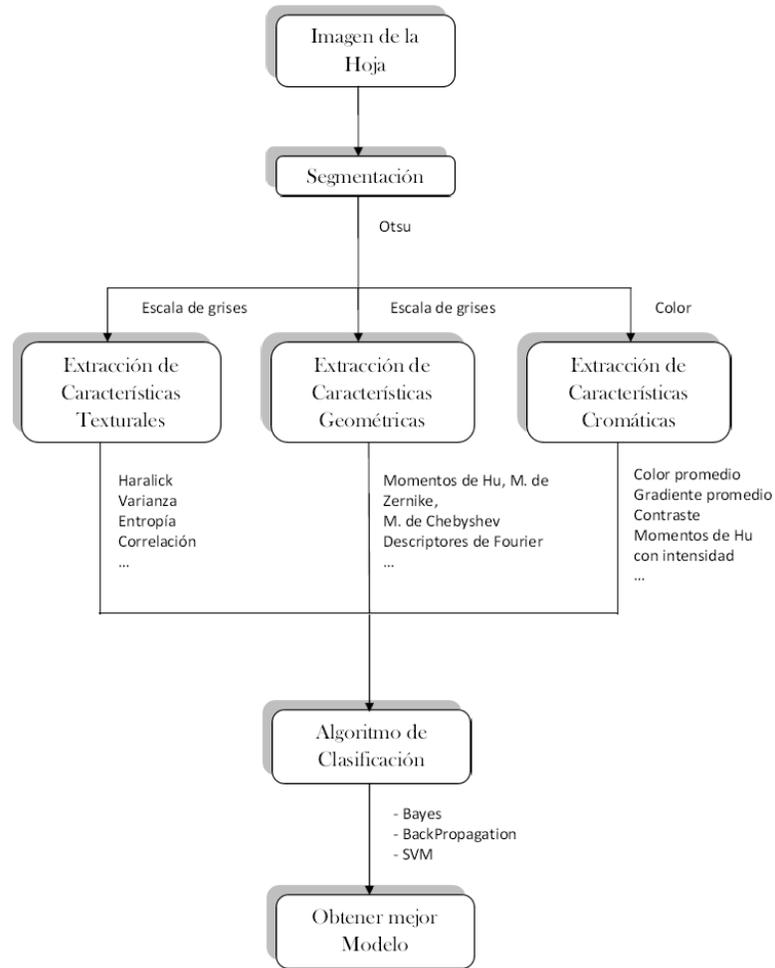


Figura 6-1: Metodología propuesta

6.1. Modelo propuesto

El modelo propuesto consiste de cuatro pasos. El diagrama de bloques del método propuesto es mostrado en la Figura 6-1. Las siguientes Sub-secciones muestran en detalle los pasos llevados a cabo en la metodología propuesta.

6.2. Preprocesamiento

El primer paso fué la selección del conjunto de datos. La colección de datos utilizada en esté trabajo fue obtenida de la Universidad de Hefei, conjunto de datos ICL que contiene 16849



Figura 6-2: Conjunto de hojas utilizado

imágenes de hojas, obtenidas de 220 especies. Algunas de las imágenes de hojas utilizadas son mostradas en la Figura 6-2, como se puede apreciar están fueron obtenidas bajo condiciones específicas como: color de fondo de la imagen e iluminación.

El preprocesamiento de una imagen digital es utilizado para la mejora de la imagen, eliminar ruido, eliminar borrosidad, aumentar brillo o reducirlo, etc; son muchas las técnicas de mejora de imagen que se le puede realizar a los objetos. Para esta investigación se utilizaron un filtro gaussiano de máscara 3×3 y un filtro de la mediana. Esto ayuda a reducir el ruido en imágenes que presentan este problema, como pueden ser pequeños puntos en la imagen y también para generar un efecto de suavizado, el suavizado es la mezcla ligera de los colores vecinos de cada pixel. Sin embargo, debido a que las imágenes utilizadas fueron tomadas bajo estrictos controles de iluminación y color de fondo, se pudo concluir a partir de los experimentos realizados que el

uso del los filtros no influía significativamente en nuestro caso.

6.3. Segmentación

Para este trabajo se utilizaron las técnicas de segmentación de Otsu, Frontera Adaptativa, Pca y K-means. La segmentación que mayor precisión en la forma de la hoja fue la segmentación de K-means, la cual nos permitió trabajar con imágenes de hojas con forma compleja, como es el caso de la tercer imagen que nos permite visualizar la mejor segmentación.

Para la validación de las técnicas de segmentación únicamente se realizó una verificación visual, sin embargo se escogieron las imágenes de formas complejas para validar la eficiencia de la segmentación. La frontera adaptativa dio buenos resultados aproximados a lo deseado pero no lo suficiente por lo tanto esta segmentación al igual que las demás fueron descartadas y solo se tomaron como muestra para la elección de nuestra segmentación que utilizamos.

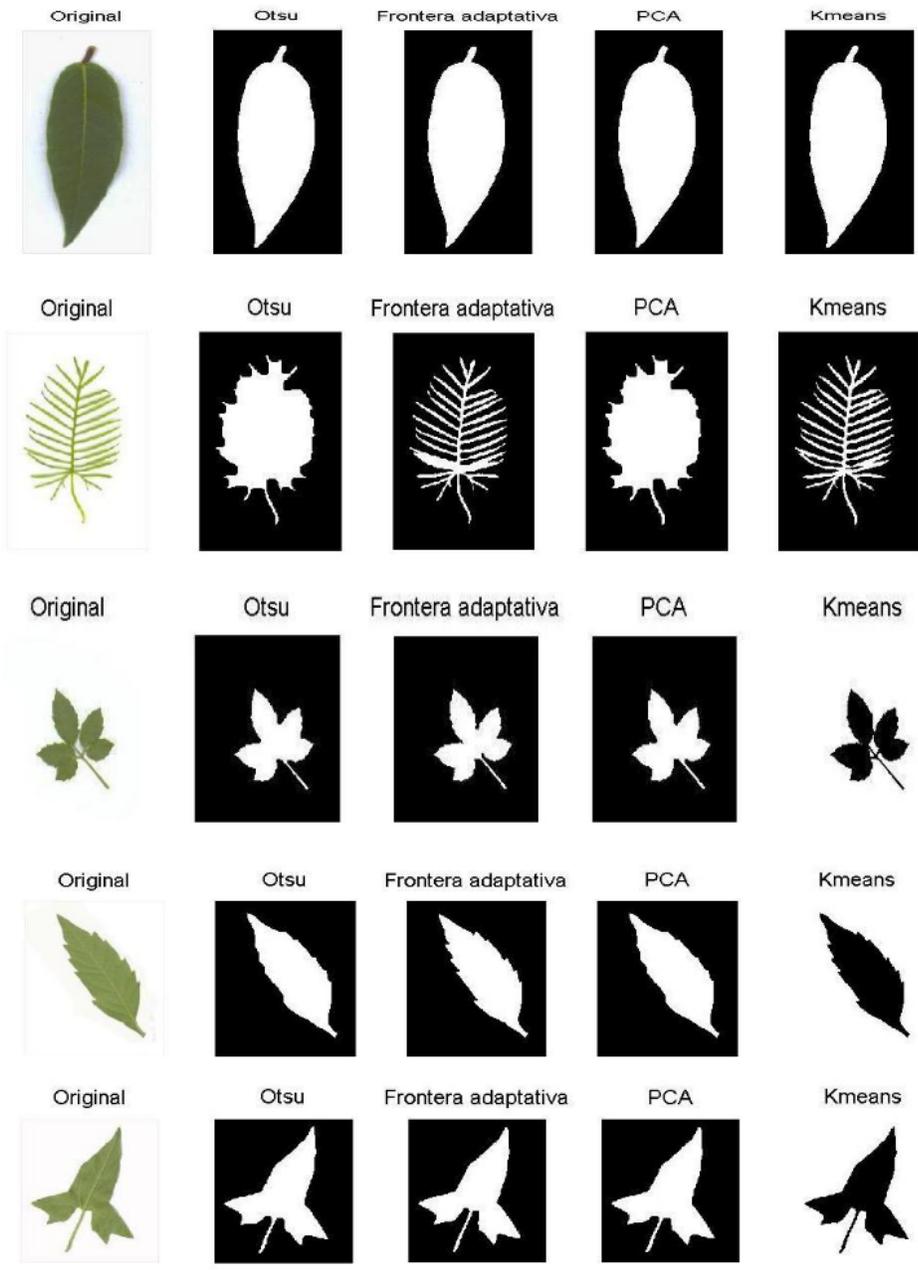


Figura 6-3: Técnicas de Segmentación

6.4. Extracción de Características

Las técnicas de extracción de características son clasificadas en cromáticas, texturales y geométricas. Cada característica de una imagen, al extraerse se va agregando a un vector que contiene todas las características de la imagen. Las características pueden ser variables únicas o estar definidas por más variables. Una vez que se han definido el número de características a extraer estas conforman un vector $\hat{x}_i = \{x_l\}_{l=1}^t$, donde \hat{x}_i representa el i ésimo vector de características del conjunto de datos, $i = 1, 2, \dots, m$. l representa la l ésima característica, el número de características es definido por t , mientras que el número de imágenes en el conjunto de datos está definido por m . Un vector de características de una imagen podría estar definido por el vector siguiente:

$$\hat{x}_i = [2,1, 1,7, 6,1, 9, 0,2, 4,5, 7,8, \dots, 2,1] \quad (6-1)$$

Las técnicas utilizadas en esta tesis, se mencionan a continuación.

6.4.1. Características Texturales:

Se utilizan los descriptores texturales de Haralick, se toman en cuenta las características de segundo momento angular, contraste, correlación, suma de cuadrados, momento de diferencia inversa, suma promedio, suma de entropía, suma de varianzas, entropía, diferencia de varianzas, diferencia de entropías, medidas de correlación y máximo coeficiente de correlación.

6.4.2. Características Geométricas:

Se utilizan los descriptores de Fourier, los momentos de Hu, Momentos, Gupta en el cual tomamos el centro de gravedad, la altura, el ancho, área, perímetro, redondez, diámetro, longitud de los ejes, entre otras, también obtuvimos las características geométricas básicas, el descriptor de flusserl y la elipsidad de la imagen.

6.4.3. Características Cromáticas:

Para extraer las características cromáticas se utilizaron los momentos de Hu intensidad para los colores rojo, verde y azul, el descriptor Hog Piramidal, descriptor Gabor y Fourier.

6.5. Clasificación e identificación

Una vez obtenidas las características mas relevantes de las hojas para esta investigación, se utilizan tres clasificadores diferentes para llevar a cabo un análisis comparativo de los resultados. Los clasificadores utilizados son: maquinas de soporte vectorial (SVM), clasificador bayesiano y una red neuronal backpropagation.

Las pruebas sobre el conjunto de datos con los métodos previamente mencionados se realizaron con el software de Weka, por su aceptación y uso en investigaciones de problemas de clasificación de conjunto de datos, minería de datos entre otros y a la facilidad y fiabilidad que supone su uso, razón por la cual se optó por este software. Con los resultados obtenidos de cada clasificador se realizó un análisis comparativo del desempeño de cada uno sobre el conjunto de datos de las hojas y así determinar cual de los tres se desempeña mejor bajo las condiciones de esta investigación.

Capítulo 7

Resultados experimentales

Los resultados experimentales obtenidos nos permiten decidir y manejar el entorno de clasificación que mejor satisfaga nuestro proyecto. En este sistema de identificación se manejaron los datos bajo el análisis de tres clasificadores de manera general, para realizar el estudio de precisión e índice de variación para cada módulo de características usadas, tales como las cromáticas, texturales y geométricas de nuestro conjunto de imágenes.

Las tablas de los resultados experimentales fueron obtenidas bajo combinaciones de características diferentes para determinar los posibles cambios y alteraciones de los resultados con el objetivo de comparar y crear un análisis más eficaz en los resultados de los datos.

7.1. Conjunto de datos utilizados

Como se menciona anteriormente, la colección de datos utilizada en este trabajo fue obtenida de la Universidad de Hefei, conjunto de datos ICL que contiene 16849 imágenes de hojas, obtenidas de 220 especies. El conjunto de imágenes se subdividió en grupos, estos grupos eran hechos de acuerdo a su semejanza de las hojas de diferente especie; para estos grupos se pretendía analizar sus resultados de clasificación, partiendo de la idea de semejanza a simple vista de las hojas y que efecto produciría esto en los clasificadores. Otro grupo fue realizado inversamente, se colocaron imágenes con características diferentes, a simple vista no se compartía semejanza entre las hojas, todos los grupos fueron sometidos a los clasificadores y así obtener los resultados de ambos grupos.

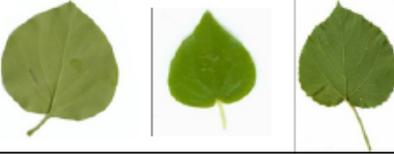
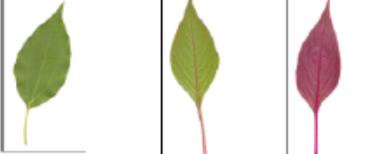
Subconjunto	Familias Asociadas	Ejemplos	Razon de Similitud
1	19		Orbicular
2	8		Lineal
3	14		Lanceolada
4	27		Elíptica
5	12		Aovada
6	3		Lacerada con forma de pentágono

Figura 7-1: Subconjuntos de especies de plantas utilizados.

Subconjunto	Familias Asociadas	Ejemplos	Razon de Similitud
7	7		Lineal dentada
8	9		Espatulada
9	13		Aovada con cuspide
10	20		Elongada
11	37		Obovada

Figura 7-2: Subconjuntos de especies de plantas utilizados.

7.2. Validación

La validación cruzada es un método que permite garantizar que el resultado obtenido. La técnica es utilizada para evaluar los resultados y garantizar que son independientes de la partición entre datos de entrenamiento y prueba. La técnica consiste en repetir y calcular la media aritmética obtenida de las medidas de evaluación sobre diferentes particiones. Se utiliza en entornos donde el objetivo principal es la predicción y se quiere estimar cómo la precisión se verá afectada al utilizar un modelo en la práctica. En todos los experimento llevados a cabo se utilizó una validación cruzada (crossvalidation), con $k=10$.

7.3. Resultados experimentales

Después de llevar a cabo los procesos mencionados anteriormente se obtuvo el desempeño de varios clasificadores utilizando un conjunto individual de características. La Tabla 1 muestra los resultados obtenidos con tres métodos de clasificación y utilizando un solo tipo de características (Cromáticas o Texturales o Geométricas).

En las tablas de resultados, CH_ i representa los subconjuntos de datos utilizados y los clasificadores utilizados son representados por sus siglas, Máquina de Soporte Vectorial (SVM), Bayesiano (Bayes) y Backpropagation (BP).

La Tabla7-1 muestra que el clasificador con mayor número de aciertos para los conjuntos en las características cromáticas y geométricas es SVM, i.e. en todos los casos el clasificador con mejor desempeño fue SVM. También se puede observar que el desempeño del clasificador BP fue muy alto, mostrando en algunos conjuntos precisiones del 100%, para el caso particular de las características texturales clasificadas con SMO se observan resultados similares a BP de baja precisión en comparación con las características Cromáticas y Geométricas.

La Tabla 2 muestra los resultados obtenidos con combinaciones de características, en este caso pares de tipos de características. Los pares de tipos de características que se utilizaron son: Cromáticas-Texturales, Cromáticas-Geométricas y Texturales-Geométricas.

En la Tabla7-2, podemos observar que la combinación de características fue buena, ya que se pueden observar precisiones hasta del 100% con el conjunto CH_6 al igual que en la Tabla anterior para este mismo conjunto y de manera general todos los clasificadores (exceptuando al

Tabla 7-1: Resultados de clasificación de características Cromáticas, Texturales y Geométricas por separado

	Cromáticas			Texturales			Geométricas		
	Bayes	BP	SVM	Bayes	BP	SVM	Bayes	BP	SVM
CH_1	88.7	94.214	94.913	36.283	47.785	56.902	81.893	93.566	95.408
CH_2	81.57	84.502	88.288	38.873	68.244	73.053	85.588	90	92.941
CH_3	88.22	94.736	95.035	37.180	74.067	77.518	78.088	89.320	91.918
CH_4	86.25	94.566	95.628	36.729	67.324	75.881	74.703	90.747	93.409
CH_5	93.67	95.560	95.727	28.303	71.042	72.841	85.444	95.111	97.289
CH_6	99.37	100	100	75.534	91.710	93.308	93.081	99.371	99.742
CH_7	95.56	96.704	97.457	48.256	82.396	85.082	97.875	98.752	98.733
CH_8	95.85	97.830	98.561	56.406	84.239	87.875	95.266	98.316	98.041
CH_9	88.38	90.648	90.872	38.393	65.776	69.670	84.592	93.807	94.410
CH_10	86.31	95.854	96.365	39.549	78.893	82.871	80.216	93.105	94.182
CH_11	77.62	92.333	93.912	33.721	63.418	68.161	74.832	88.356	91.324
CH_12	79.37	95.722	97.754	28.238	61.558	65.585	79.985	86.933	89.095

clasificador bayesiano en algunos conjuntos) muestran precisiones superiores al 90 %, algo que no se aprecia en la primera tabla, y se debe tomar en cuenta que en las tres combinaciones el clasificador con mejores resultados es SVM, con un promedio de 94,86 % para el primer conjunto, 96,53 % para el segundo y 94,50 % para el último

Comparando la Tabla7-1 con la Tabla7-2 que tiene las características combinadas, se concluye que la combinación de características para la clasificación de hojas mejora la precisión del clasificador y por lo tanto mejora el poder de discriminación entre los diferentes grupos de hojas.

Por último, se muestran los resultados obtenidos utilizando todas las características obtenidas. La Tabla 3 muestra los desempeños obtenidos. es posible apreciar que la tendencia de los resultados combinando pares de características, en la tabla 7-3 se combinaron las tres características de las hojas obteniendo una precisión de clasificación de 98,20 % mediante SVM

Por lo tanto se concluye de la Tabla7-3 que los resultados pertenecen a un parámetro promedio, los cuales no son bajos, y su nivel de aciertos de los tres clasificadores es bueno.

Tabla 7-2: Resultados de clasificación usando características Cromáticas-Texturales, Cromáticas-Geométricas, Texturales-Geométricas

	Cromáticas-Texturales			Cromáticas-Geométricas			Texturales-Geométricas		
	Bayes	BP	SVM	Bayes	BP	SVM	Bayes	BP	SVM
CH_1	87.357	93.426	94.264	92.371	97.242	97.593	86.472	93.805	95.276
CH_2	82.061	84.732	86.253	87.058	90.588	92.476	88.122	91.869	91.421
CH_3	90.413	95.864	96.421	92.51	96.948	97.93	85.150	93.045	94.285
CH_4	87.261	94.005	95.218	91.423	97.262	95.825	84.332	92.370	94.720
CH_5	92.242	95.952	96.301	96.333	95.111	96.015	90.219	94.603	95.086
CH_6	98.742	100	100	98.742	100	100	96.226	99.371	99.371
CH_7	96.861	95.48	97.289	98.853	96.566	97.717	96.380	96.533	98.760
CH_8	95.052	96.135	96.211	96.844	96.211	97.823	96.614	96.697	96.958
CH_9	87.772	91.484	92.108	94.41	95.072	95.238	87.554	92.139	90.174
CH_10	87.748	96.926	95.332	89.77	95.296	96.163	90.505	92.628	91.939
CH_11	80.537	92.569	93.169	83.815	93.889	94.925	84.039	92.655	92.817
CH_12	81.078	95.342	95.865	87.785	95.233	96.721	86.340	89.204	93.204

Tabla 7-3: Resultados de clasificación con las características Cromáticas, Texturales y Geométricas.

	Cromáticas-Texturales-Geométricas		
	Bayes	BP	SVM
CH_1	91.909	96.207	98.524
CH_2	87.739	90.038	94.941
CH_3	93.797	96.992	98.882
CH_4	90.940	96.253	97.183
CH_5	95.784	98.145	99.153
CH_6	98.742	100	100
CH_7	98.535	98.954	99.062
CH_8	96.354	98.177	99.828
CH_9	95.196	93.886	97.893
CH_10	90.027	98.292	98.531
CH_11	86.482	93.807	95.917
CH_12	88.488	96.19	98.526

Capítulo 8

Conclusiones y discusiones

Para llevar a cabo el desarrollo de un sistema de reconocimiento de imágenes de hojas en esta tesis, se trabajó con el conjunto de imágenes de la Universidad de Hefei con el objetivo de tomar para pruebas este conjunto, el proceso para llevar a cabo este sistema se resume en las siguientes fases: obtención de imágenes, mejora de la imagen, segmentación de las imágenes, análisis de características específicas y discriminativas de la imagen de interés, una vez analizadas y elegidas las características se realizó la extracción de características, el conjunto de datos obtenidos se clasificaron por tres métodos distintos teniendo en cuenta el orden de las clases en los conjuntos utilizando características de manera independiente y de manera combinada, los resultados se normalizaron y se realizó la validación cruzada, después se analizaron los resultados para determinar el clasificador con mejor precisión para los conjuntos de imágenes, para finalizar los resultados se evaluaron.

En la mejora de imagen, se aplicó un filtro gaussiano y un filtro mediana para disminuir ruido, pero se determinó que para el conjunto de imágenes utilizado no era necesario aplicar un mejoramiento de imágenes, dado que el conjunto de imágenes, estaba realizado bajo ciertas condiciones tales como que el fondo de todas las imágenes es de color blanco, la resolución de las imágenes es buena al igual que la iluminación y sin errores de enfoque en la toma de la fotografía, por tales motivos no influyó el que se aplicará o no una mejora de imagen ya que no era necesario.

En la fase de segmentación se contempló segmentación PCA, Otsu, K-means y Frontera

Adaptativa, concluyendo que la segmentación mas eficiente y precisa en la forma de la hoja es la segmentación k-means, pero se verían afectadas y con un nivel de variación alto si el fondo de la imagen fuera difícil de separar, por ejemplo si la fotografía de la hoja tuviera otras hojas en la parte de fondo, esto causaría confusión en la separación de la hoja del fondo porque la segmentación se basa en las variaciones de tonalidad en intensidades y si estamos hablando que en el fondo hay otras hojas del mismo tipo que la de interés se corre el riesgo de que se segmente la imagen de la hoja con la agrupación de contornos de las hojas de fondo, dando así una imagen segmentada no deseada y deformada.

Para el análisis de las características que eran convenientes y necesarias tomar, se llevó a cabo una selección discriminante tomando todas las características de forma general, al terminar de contemplar las características que creemos necesitar, se comienza a delimitar esta decisión, primero separar las características que pertenezcan a texturales, cromáticas y geométricas, después de ver analizado y asumido las características que se necesitaban se realizó una búsqueda de técnicas que nos pudieran ser de apoyo para extraer las propiedades necesarias. Para esta fase de extracción de características se dividió en tres parámetros principales la textura, la cromaticidad y las propiedades geométricas de las imágenes segmentadas como propiedades únicas de cada especie de hojas. Es necesario mencionar que se utilizaron las siguientes técnicas para la extracción de características: descriptores de Haralick, contraste, correlación, momento de diferencia inversa, entropía, varianza, máximo coeficiente de correlación, descriptores de Fourier, momentos de Hu, Gupta, área, perímetro, redondez, diámetro, longitud de ejes, descriptor de Flusser, elipsidad, descriptores Hog Piramidal, descriptores de Gabor y Fourier; no son las únicas técnicas mas sin embargo para este trabajo fueron las más adecuadas y pertinentes.

Al analizar las tres tablas de resultados se concluye que si se obtiene solo una de las tres características de la imagen se pueden obtener resultados aceptables si se controla en las imágenes factores como el fondo que sea contrastante al objeto de interés como en esta investigación, sin embargo al combinar dos o las tres características de las imágenes se obtienen resultados muy superiores a los obtenidos solo con una característica, también es de destacar que el clasificador SVM fue el que mejor desempeño tuvo de todos los clasificadores utilizados bajo los 3 aspectos utilizados que fueron clasificación sin combinación de características, utilizando combinación de dos y combinación de tres características.

La utilización de los sistemas de reconocimiento de imágenes en la botánica, es de suma importancia e interés debido a las ventajas de identificación de las hojas y la precisión de una confiable clasificación de tales especies de plantas. El trabajo de los botánicos para la separación, almacenamiento y clasificación de las plantas es pesado por su agrumada cantidad de especies que existen, por ello se ha determinado el desarrollo y uso de estos sistemas de reconocimiento de patrones o bien para botánica reconocimiento de hojas.

Con este trabajo se pretende fomentar y favorecer un mayor y mejor entendimiento del uso de las redes neuronales en las técnicas de visión por computador para sistemas de identificación de objetos, ya que se ha contemplado en esta tesis el proceso general para cualquier sistema de reconocimiento. Se realizó con la finalidad de aclarar y mostrar la necesidad de estos sistemas en el área de botánica, favoreciendo en tiempo, trabajo, precisión y confiabilidad de una clasificación de plantas mediante la fotografía de su hoja.

Bibliografía

- [1] A. Kutya-Olesiuk, P.Ciosek, E. W. Effect of lead accumulation in maize leaves on their chemical images created by aflow-through electronic tongue *Talanta*, Octubre 2012
- [2] A. Kötschau, G. Büchel, J. E. R. M. W. v. T. D. Element pattern recognition and classification in sunflowers (*Helianthus annuus*) grown on contaminated and non-contaminated soil *Microchemical Journal*, Diciembre 2013
- [3] Ahlem Othmani, A. P. Single tree species classification from Terrestrial Laser Scanning data *Oficina Nacional de Arboles de Francia*, 2013.
- [4] Andrés Castrillejo, Natalia da Silva, G. I. Consistencia de Random Forests. *Universidad de la República*, 2010
- [5] Bahaa-Eldin, A. M. A medium resolution fingerprint matching system *Ain Shams Engineering Journal* , 2013, 4, 393 - 408
- [6] Boles, W. A security system based on human iris identification using wavelet transform *Engineering Applications of Artificial Intelligence*, 1998
- [7] en C. José Luis Calderón O., M. *Redes Neuronales Artificiales 1. Teoría y Diseño*. 2002
- [8] Chávez, J. L. “Textura multispectral a partir de una variación del algoritmo LBP” *Instituto de Geofísica, UNAM*, 2011
- [9] Cristina García Cambronero, I. G. M. Algoritmos de Aprendizaje:KNN y MEANS *Inteligencia en Redes de Telecomunicación*, 2000, 8

- [10] D. Cordell, M. Jackson, S. W. Phosphorus flows through the Australian food system: Identifying intervention points as a roadmap to phosphorus security Environmental science a policy, Marzo 2013
- [11] Deny L. Hernández Rabadán, F. R. U. C. Reconocimiento de Hojas de Plantas Instituto Tecnológico y de Estudios Superiores de Monterrey, 2009
- [12] Elio Ramos, D. S. F. Classification of leaf epidermis microphotographs using texture features Ecological Informatics, 2009
- [13] Flores, J. W. S. Reconocimiento de patrones de morosidad para un producto crediticio usando la técnica de árbol de clasificación CART. Universidad Nacional Mayor de San Marcos. Ingeniería Industrial., 2005
- [14] Gianluigi Pillonetto, F. D. Kernel methods in system identification, machine learning and function estimation: A survey Pattern Recognition, 2014
- [15] Guili Xu, Fengling Zhang, S. G. S. Y. Y. H. M. Use of leaf color images to identify nitrogen and potassium deficient tomatoes Pattern Recognition Letters, 2011
- [16] H. Muhammad Asraf, M. T. Nooritawati, M. S. R. A Comparative Study in Kernel-Based Support Vector Machine of Oil Palm Leaves Nutrient Disease Procedia Engineering, 2012
- [17] Ji-Xiang Du, Xiao-Feng Wang, G.-J. Z. Leaf shape based plant species recognition Department of Automation, University of Science and Technology of China, Hefei, Anhui 230027, China, Intelligent Computing Lab, Institute of Intelligent Machines, Chinese Academy of Sciences., 2006
- [18] Junghoi Kim, J. C. Digital holographic security system based on random phase encoded reference beams and fingerprint identification Optics Communications, 2004
- [19] Kyal, F. H. Principales tipos de Redes Neuronales, Universidad Tecnológica de Pereira. 2003
- [20] Larese, R. M. Craviotto, M. R. A. C. G. P. M. G. Identificación automática de cultivares de soja basado en el análisis de nervaduras ECIImag, 2012

- [21] LI Chao, YANG Yi-xian, N. X.-x. Biometric-based personal identity authentication system and security analysis The Journal of China Universities of Posts and Telecommunications, 2006
- [22] López, L. E. J. C. Simulación paralela de redes neuronales aplicadas al proceso de visión temprana. Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, 2009
- [23] Manterola, H. L., O. J. I. F. E. d. F. M. Extracción de objetos en imágenes médicas tridimensionales basada en características texturales Congreso Argentino de Informática y Salud, CAIS, 2013
- [24] Manu N. Lakshmanan, A. J. K. An X-ray scatter system for material identification in cluttered objects: A Monte Carlo simulation study Nuclear Instruments and Methods in Physics Research B, 2014
- [25] Marco Antonio Valencia Reyes, Cornelio Yáñez Márquez, L. P. S. F. Algoritmo Backpropagation para Redes Neuronales: conceptos y aplicaciones Instituto Politécnico Nacional Centro de Investigación en Computación, 2006
- [26] Márquez, J. Procesamiento y Análisis de Señales e Imágenes 2010, 7
- [27] Nassar, D. E. M. & Ammar, H. H. A neural network system for matching dental radiographs Pattern Recognition, 2007, 40, 65 - 79
- [28] Aguilera Fernández Oscar A, Carbajal Coronado Jaime, F. C. M. D. L. A. L. A. K. G. M. d. O. L. Redes Neuronales, su Implementación en Hardware. 2000
- [29] Orlando, José Ignacio, M. H. L. d. F. M. Detección de tumores en MRI basada en características texturales ECIImag, 2012
- [30] Perez, M. A. Redes Neuronales: Alfa-Beta y Morfológicas. Centro de Investigación en Computación del Instituto Politécnico Nacional de México 2004
- [31] Petasyuk, G. System and criterial method of the identification and quantitative estimation of the geometrical shape of the abrasive powder grain projection Powder Technology, 2014

- [32] Petr Novotny, T. S. Leaf recognition of woody species in Central Europe SciVerse ScienceDirect, 2013
- [33] María del Pilar Gómez Gil, A. M. V. Redes Neuronales Artificiales para Calificar la Capacidad de Crédito de Entidades Mexicanas de Gobierno Komputer Sapiens, Julio - Diciembre 2010, Año II, Vol. II., 26 / 32
- [34] Revollo, G. N. Revollo, C. D. G. M. P. Clasificación y Extracción de Características en Imágenes de Playa ECIImag, 2012
- [35] Rodriguez Gallano, M. R. Clasificación de imágenes de Satélite mediante software libre: nuevas tendencias en algoritmos de Inteligencia Artificial XV Congreso Nacional de Tecnologías de la Información Geográfica, 2012 .
- [36] Steven L. Clark, MD; Janet A. Meyers, R. D. R. F. R. K. M. M. J. B. P. M. A systematic approach to the identification and classification of near-miss events on labor and delivery in a large, national health care system Ajog Patient Safety Series, 2012
- [37] Susan R. Snyder, Alessandra M. Favoretto, J. H. D. Effectiveness of barcoding for reducing patient specimen and laboratory testing identification errors: A Laboratory Medicine Best Practices systematic review and meta-analysis Clinical Biochemistry, 2014
- [38] T. Chomdej, W. Pankaowa, S. C. Intelligent dental identification system (IDIS) in forensic medicine Forensic Science International, 2006
- [39] Vijay Satti, Anshul Satya, S. S. An Automatic Leaf Recognition System for Plant Identification using machine Vision Technology International Journal of Engineering Science and Technology (IJEST)ASET, Amity University Noida, Uttar Pradesh, India,ASET, Amity University Noida, Uttar Pradesh, India, 2013, Vol. 5 No.04, 7
- [40] Xiao-Feng Wang, De-Shuang Huang, J.-X. D. H. X. L. H. Classification of plant leaf images with complicated background Applied Mathematics and Computation, 2008
- [41] Castillo, F. D. Tolerancias Geométricas Facultad de Estudios Superiores Cuautitlan, UNAM, Departamento de Ingeniería., 2008

- [42] Shannwen Zhang, Yingke Lei, T. D. Label propagation based supervised locality projection analysis for plant leaf classification *Pattern Recognition*, El Sevier, 2013, 46, 1891-1907
- [95] Yongming Chen, Ping Lin, Y. H. Velocity representation method for description of contour shape and the classification of weed leaf images *Science Direct*, El Sevier, 2011, 109, 186-195
- [44] Guillaume Cerutti, Laure Tougne, J. M. Understanding leaves in natural images, A model based approach for tree species identification *Computer Vision and Image Understanding*, El Sevier, 2013, 117, 1482-1501
- [81] Pascual, J. E. S. Mecanismos de Captura y Procesado de Imágenes de Venas para Identificación Personal Universidad Carlos III de Madrid, 2011
- [46] Corona, A. C. Texturas Escuela Superior de Ingenieros, 2004
- [47] Ciriza, Albizua, A. Análisis de la utilidad de los descriptores texturales de Haralickk para la localización arranques de frutal en ortofoto *Teledetección: Agua y Desarrollo Sustentable*, 2009, XIII, 23-26
- [48] Alberto Cugnetto, Laura Santagostini, L. R. Tracing the terroirs via the elemental composition of leaves, grapes and derived wines in cv Nebbiolo *Scientia Horticulturae*, 2014, 172, 101-108
- [49] Andre Backes, D. C. Color texture analysis based on fractal descriptors *Pattern Recognition*, El Sevier, 2012, 45, 1984-1992
- [50] Arranz Aranda, Liu Yin, L. C. Interacción persona computador basada en el reconocimiento visual de manos Universidad Complutense de Madrid, 2012
- [51] Bishop Introducción al Reconocimiento de Patrones Capitulo 7. Maquinas de Vectores de Soporte (SVM) 2013, 2
- [52] Bishop Introducción al Reconocimiento de Patrones Capitulo 7. Maquinas de Vectores de Soporte (SVM) 2013, 2
- [53] Carrillo, H. C. Conversión de texto manuscrito a formato digital utilizando Máquinas de Soporte Vectorial Instituto Politécnico Nacional, 2007

- [54] Chuan Min Zhain, J. X. D. Recognition of plant leaf image based on fractal dimension features Neurocomputing, El sevier, 2013, 116, 150-156
- [55] Currás, M. Detección de Bordes 2012
- [56] Dego, G. A. Procesamiento de Imágenes Universidad de Salamanca, 2005
- [57] Edwin Caraballo, Rita Ávila Predicción del Rendimiento de un Cultivo de Plátano mediante Redes Neuronales Artificiales de Regresión Generalizada Ciencias y Tecnología, 2012, 6, 31-40
- [58] Elkin Eduardo, German Méndez, G. P. Clasificador de Música por Género utilizando Redes Neuronales Artificiales Pontificia Universidad Javeriana, Facultad de Ingeniería, Bogotá, 2003
- [59] Emilce Moler, J. P. Técnicas de Procesamiento Digital de Imágenes: Una Aplicacion para Identificación de Personas a través de los Senos Frontales Mecánica Computacional, 2010, XXIX, 6285-6294
- [60] Enrique Sucar, G. G. Visión Computacional Instituto Nacional de Astrofísica , Optica y Electrónica, 2007
- [61] Francisco Ascacibar, Alpha Pernía, F. A. M. C. de La Rioja, U. (ed.) Técnicas y Algoritmos Básicos de Visión Artificial Técnicas y Algoritmos Básicos de Visión Artificial, 2006
- [62] Garay, J. P. Transformadas Matemáticas en Teoría de Señales Academia de Ciencias Exactas, Fisicas, Quimicas y Naturales de Zaragoza, 1998
- [63] García, A. Procesamiento y Análisis de Imágenes Digitales I Univerdidad de Chile, Universidad de Chile, 2011
- [64] Gegen, L. K. Propuesta de nuevos criterios de vida de herramienta de corte y determinación del desgaste a partir de la clasificación de imágenes digitales caracterizadas utilizando descriptores de región simples y basados en momentos. Universidad de León, 2007

- [65] Gualteros, D. M. B. Sistema de Visión por computadoras para detectar hierba no deseada en proptotipo de cultivo de frijol usando ambiente controlado Universidad Católica de Colombia, 2014
- [66] Hang Jun Wang, H. N. Q. A new Gabor based approach for wood recognition Neurocomputing, El sevier, 2013, 116, 192-200
- [67] Him, Q. Procesamiento Global empleando la Transformada de Hough 2005
- [68] James Cope, D. C. Plant species identification using digital morphometrics: A review Expert System with Applicatiions, 2012, 39, 7562-7573
- [69] Jesus Cid, J. A. Extracción de Características 2006
- [70] Kutyla Olesiuk, C. Effect of lead accumulation in maize leaves on their chemical images created by a flow- through electronic tongue Talanta, El Sevier, 2013, 103, 179-185
- [71] Luis Alonso Romero, T. C. Redes Neuronales y Reconocimiento de Patrones Universidad de Valladolid, España., 2011
- [72] Martínez Corral, M. R. Desarrollo de una base de datos para caraterización de alfalfa en un sistema de Visión Artificial Revista Internacional de Botánica Experimental, 2009, 78, 43-47
- [73] Matich, D. J. REdes Neuronales: Conceptos y Aplicaciones Universidad Tecnológica Nacional, 2001
- [74] McKenzie, J. Characterisation of tea leaves according to their total mineral content by means of probabilistic neural networks Food Chemistry, El Sevier, 2010, 123, 850-864
- [75] Merelo, Prieto, C. Diseño de Redes Neuronales Artificiales mediante Algoritmos Evolutivos Iberoamericana de Inteligencia Artificial, 2001, 5, 2-32
- [76] Mery, D. Extracción de Características Universidad Católica de Chile, 2006
- [77] Miranda, G. M. Evaluación de Algoritmos supervisados de extracción de características para clasificacion de texturas. Escuela Politécnica Superior, 2009

- [78] Monica Larese, R. N. Análisis de Imágenes para el estudio de soja CIFASIS, 2010, 27, 303-308
- [79] Pablo, P. G. Reconocimiento de Imágenes utilizando Redes Neuronales Artificiales Universidad Complutense de Madrid, 2013
- [80] Parra, S. T. Redes Neuronales Universidad Politécnica de Querétaro, 2006
- [81] Pascual, J. E. S. Mecanismos de Captura y Procesado de Imágenes de Venas para Identificación Personal Universidad Carlos III de Madrid, 2011
- [82] Platero, C. Vision Artificial Capitulo 3. Procesamiento Digital de Imágenes Visión Artificial, 2012, 91-113
- [83] Presutti, M. La matriz de coocurrencia en la clasificación multiespectral: para la enseñanza de medidas texturales Educao em Sensoramiento Remoto no Ambito de Mercosul, 2004, 11, 113-118
- [84] Rivas Araiza, Jorge Mendiola, R. M. Mejora de Contraste y Compensación en Cambios de Iluminación Computación y Sistemas, 2007, 10, 357-371
- [85] Riveira, I. S. Apuntes Teledetección 2009
- [86] Rubira, E. A. Introducción al reconocimiento de patrones mediante redes neuronales Computación y Sistemas, 2004, 23, 14-19
- [87] Sampallo, G. Reconocimiento de Tipos de Hojas Revista Iberoamericana de Inteligencia Artificial, 2003, 21, 55-62
- [88] Serrano, E. P. Introducción a la Transformada Wavelet y sus aplicaciones al Procesamiento de Señales de Emisión Acústica Universidad Nacional de General SAN Martín, 2011
- [89] Torres, G. P. Control Visual Flexible empleando Momentos Invariantes Universidad de Alicante, 2009
- [90] Valenzuela, S. P. Identificación de especies vegetales utilizando dispositivos móviles Universidad de Chile, 2013

- [91] Vicente Feliu Batle, A. O. Descriptores de Fourier para Identificación y posicionamiento de objetos en entornos 3D *Ciencias y Tecnología*, 2004, 10, 33-41
- [92] Vélez, J. F. S. *Visión por Computador* 2002
- [93] William Castrillon, D. A. Técnicas de Extracción de Características en Imágenes para el Reconocimiento de Expresiones Faciales *Scientia et Technica*, 2008, 38, 116-121
- [94] Xavier Felip Leon, A. M. C. Modificación del Histograma de una Imagen *Estudios de Informativa de la UJI*, 2005, 5, 8
- [95] Yongming Chen, Ping Lin, Y. H. Velocity representation method for description of contour shape and the classification of weed leaf images *Science Direct, El Sevier*, 2011, 109, 186-195
- [96] Gonzalo Pajares Martinsanz, J. M. d. I. C. G. Editor, A. G. (ed.) *Ejercicios Resueltos de Visión por Computador RA-MA Editorial*, 2008, 354
- [97] Martínez, J. R. H. V. J. Rama, E. (ed.) *Redes Neuronales Artificiales. Fundamentos, Modelos y Aplicaciones* 1995