



**Universidad Autónoma del Estado de México**

**CENTRO UNIVERSITARIO UAEM TEXCOCO**

**Detección Automática de los Grados de Severidad de  
Retinopatía Diabética y Glaucoma Utilizando Técnicas  
Clásicas**

*Para obtener el grado de Maestra en Ciencias de la Computación  
presenta:*

Dalia Berenice Luna Coria

Director de tesis:  
Dr. Jair Cervantes Canales

Noviembre 2023

## Abstract

The use of machine learning algorithms for detecting visual diseases such as Diabetic Retinopathy and Glaucoma is of utmost importance. These algorithms enable efficient and accurate processing of medical images, extracting relevant features from them. This image processing stage is fundamental as it provides a solid foundation for the subsequent classification of cases.

In the mentioned text, it is highlighted that machine learning algorithms were used for the classification of features extracted from images. The results show high performance in the detection of these diseases, with accuracies ranging from 89.9% to 94.3%. This indicates that the algorithms are capable of generalizing well from the training data, which is essential for their applicability in real-world situations.

In summary, the use of machine learning algorithms in the detection of visual diseases such as Diabetic Retinopathy and Glaucoma offers a powerful tool for medicine. These algorithms enable efficient processing of medical images and precise classification of cases, which can lead to early detection and timely treatment of these diseases, thereby improving the visual health of patients. The obtained results support the utility and effectiveness of these algorithms.

## RESUMEN

El uso de algoritmos de aprendizaje automático para detectar enfermedades visuales como la Retinopatía Diabética y el Glaucoma es de suma importancia. Estos algoritmos permiten procesar imágenes médicas de manera eficiente y precisa, extrayendo características relevantes de las mismas. Esta etapa de procesamiento de imágenes es fundamental ya que proporciona una base sólida para la clasificación posterior de los casos.

En el texto mencionado, se destaca que los algoritmos de aprendizaje automático se utilizaron para la clasificación de las características extraídas de las imágenes. Los resultados muestran un alto rendimiento en la detección de estas enfermedades, con precisiones que oscilan entre el 89.9 % y el 94.3% empleando SVM y Redes Neuronales Aritificales respectivamente. Esto indica que los algoritmos son capaces de generalizar bien a partir de los datos de entrenamiento, lo que es esencial para su aplicabilidad en situaciones del mundo real.

En resumen, el uso de algoritmos de aprendizaje automático en la detección de enfermedades visuales como la Retinopatía Diabética y el Glaucoma ofrece una herramienta poderosa para la medicina. Estos algoritmos permiten un procesamiento eficiente de imágenes médicas y una clasificación precisa de casos, lo que puede llevar a una detección temprana y un tratamiento oportuno de estas enfermedades, mejorando así la salud visual de los pacientes. Los resultados obtenidos respaldan la utilidad y la eficacia de estos algoritmos.

# Índice general

<b>1. INTRODUCCIÓN</b>	<b>9</b>
1.1. PROBLEMÁTICA	10
1.2. JUSTIFICACIÓN	11
1.3. OBJETIVOS	11
1.4. HIPÓTESIS	12
1.5. ANTECEDENTES	13
1.5.1. Sistema de Reconocimiento de Objetos	13
1.5.2. Enfermedades de los Ojos	14
1.5.3. Glaucoma	14
1.5.4. Retinopatía	15
<b>2. ESTADO DEL ARTE</b>	<b>17</b>
2.1. Trabajos relacionados	17
2.1.1. Combinación de múltiples características profundas para la clasificación del glaucoma	17
2.1.2. Síntesis de imágenes retinales y aprendizaje semisupervisado para la evaluación del glaucoma	18
2.1.3. Clasificación de la retinopatía diabética basada en patrones de exudados duros, mediante procesamiento de imágenes y SVM	18
2.1.4. Diseño y desarrollo de un sistema de detección automática de Glaucoma a partir de la segmentación y el estudio de las capas de la retina en OCT Circumpapilares	20
2.1.5. Una base de datos a gran escala y un modelo CNN para la detección de glaucoma basada en la atención	20
2.1.6. Clasificación de los ángulos de la cámara anterior en imágenes OCT del segmento anterior a través de redes neuronales convolucionales de regiones multi escala huaying	22
2.1.7. Conjunto de datos para la detección de glaucoma asistida por computadora	22
2.1.8. Clasificación automatizada de las etapas del glaucoma utilizando la transformada de wavelet analítica flexible a partir de imágenes del fondo de la retina	23
<b>3. MARCO TEÓRICO</b>	<b>25</b>
3.1. Técnicas de preprocesamiento	25
3.1.1. Escala de grises	25
3.1.2. Espacio de color RGB	25
3.1.3. Espacio de color HSI	27
3.1.4. Normalización de Histograma	27
3.1.5. Ecuilización de histograma adaptativo limitado por contraste (CLAHE)	28
3.1.6. Realzado de bordes	32
3.2. Técnicas de segmentación	32
3.2.1. Método Otsu	33
3.2.2. Método basado en Clustering (DBSCAN)	34
3.3. Operaciones Morfológicas	36
3.3.1. Erosión	37
3.3.2. Dilatación	37
3.4. Extracción de características	38
3.4.1. Descriptor Haralick	38
3.4.2. Descriptor LBP (Patrón Binario Local)	42
3.4.3. Descriptor HOG (Histograma de gradientes orientados)	46
3.5. Métodos de validación	48
3.5.1. K-Fold Cross Validation	49

3.6. Algoritmos de clasificación	49
3.6.1. Redes Neuronales Artificiales (RNA)	50
3.6.2. Árboles de decisión	52
3.6.3. Algoritmos Bayesianos	53
3.6.4. Máquinas de Soporte Vectorial (SVM)	55
3.6.5. KNN (K-Nearest Neighbors)	56
3.6.6. Búsqueda de Malla (Grid search)	56
<b>4. METODOLOGÍA</b>	<b>59</b>
4.1. Conjunto de datos	59
4.2. Fase de Segmentación	62
4.2.1. Pre-procesamiento	63
4.2.2. Segmentación	64
4.3. Fase de Identificación (Clasificación)	66
4.3.1. Extracción de características	66
<b>5. RESULTADOS EXPERIMENTALES</b>	<b>69</b>
5.1. Proceso para la extracción de características	69
5.1.1. Segmentación	70
5.1.2. Selección y extracción de características	70
5.2. Resultados de la clasificación	73
5.2.1. Clasificación utilizando Redes Neuronales Aritificiales (RNA)	74
5.2.2. Clasificación utilizando Naive Bayes	76
5.2.3. Clasificación utilizando K-NN (K-n vecinos cercanos)	78
5.2.4. Clasificación utilizando Árboles de decisión	80
5.2.5. Clasificación utilizando Máquinas de Soporte Vectorial (SVM)	82
5.2.6. Clasificación para el conjunto de Retinopatía Diabética (RD)	83
5.2.7. Clasificación para el conjunto de Glaucoma	86
5.3. Discusión de los resultados	88
5.3.1. Efectividad en problemas de clasificación lineal y no lineal	89
<b>6. CONCLUSIONES</b>	<b>91</b>
6.1. Trabajos a futuro	92
6.1.1. Mejorar la precisión del modelo	92
6.1.2. Predicción en la progresión de Glaucoma y RD	93
6.1.3. Validación en poblaciones diversas	93
6.1.4. Creación de un conjunto de datos	93

# Índice de figuras

1.1. Ojo humano con diagnóstico de Glaucoma [44]	15
2.1. Proceso de detección de Glaucoma [45]	17
2.2. Diversos métodos empleados para la detección de Glaucoma [27]	18
2.3. Disco óptico: (a) Imagen original, (b) Canal verde (c) Canal cyan, (d) Diferencia de las imágenes b y c, (e) Binarización, (f) Complemento de la imagen e. [22]	19
2.4. Vasos Sanguíneos: (a) Imagen original, (b) Componente cyan, (c) Procesamiento morfológico, (d) Diferencia de las imágenes b y c, (e) Binarización, (af) Reducción de ruido. [22]	19
2.5. Modelo empleado de una Red Neuronal Perceptrón Multicapa, para trabajar con una validación interna. [51]	20
2.6. Ejemplos de fondos de imagen de Glaucoma. [46]	21
2.7. Descripción general de la arquitectura de red propuesta para la clasificación de los ángulos de la cámara anterior. [34]	23
3.1. Espacio de color RGB	26
3.2. Imagen en escala de grises con su respectivo histograma, mostrando la distribución píxeles concentrada en el centro de la imagen.	30
3.3. Proceso resultante de equalización aplicada a una imagen en escala de grises, mostrando una distribución más plana en el histograma.	30
3.4. Tipos de puntos detectados por el algoritmo DBSCAN	35
3.5. Elementos estructurados	37
3.6. a) Conjunto A, b) Elemento estructurado B, c) A eriosionada por B.	38
3.7. a) Texto de muestra de baja resolución con caracteres rotos (vista ampliada), b) Elemento estructurado, c) Dilatación a) por b). Se unieron los caracteres rotos.	39
3.8. Representación del cálculo de la Matriz de co-ocurrencia de niveles de gris. Donde se puede observar que, en la parte superior hay dos ejemplos de construcción de la matriz, tomando en cuenta 1 píxel adyacente, así como, la dirección horizontal y diagonal. Y, en la parte inferior, se observa los offset correspondientes a los diferentes ángulos [51].	40
3.9. Conjuntos de vecinos circularmente simétricos [26].	44
3.10. Las rotaciones invariantes de patrones binarios pueden ocurrir en el conjunto de ocho vecinos que tienen simetría circular. [33]	46
3.11. Ejecuciones de K-Fold Cross Validation con K=5	49
3.12. Metodología de clasificación de aprendizaje automático	50
3.13. Estructura básica de una RNA	51
3.14. Funcionamiento general del algoritmo Random Forest	54
3.15. Ejemplo sencillo de red bayesiana. Modelo que analiza la influencia de la lluvia en caso de que el rociador esté activado, y la influencia conjunta de la lluvia y el rociador en caso de que la hierba esté húmeda.	54
4.1. Componentes de un sistema de visión.	59
4.2. a) Cataratas, b) Retinopatía diabética, c) Glaucoma, d) Ojo sano. Clases del conjunto de datos utilizado.	60
4.3. Conjunto de datos de RD extraído de Kaggle.	61
4.4. Conjunto de datos de Glaucoma extraído de EyePACS.	62
4.5. Proceso general de la fase de Segmentación para obtener los vasos sanguíneos de la retina.	62
4.6. Proceso general de la fase de Identificación para extraer características de vasos sanguíneos de la retina.	66
5.1. Métodos en phyton para aplicar pre-procesamiento a las imágenes.	71

5.2. Método en python para la obtención de las características del conjunto de datos, utilizando la librería de OpenCV. . . . .	72
5.3. Interfaz de Matlab, mostrando un ejemplo de las pruebas con los diferentes algoritmos de aprendizaje automático . . . . .	74
5.4. Matriz de confusión para una Red Neuronal Amplia, mostrando los patrones mejor clasificados correspondientes a la clase 2. . . . .	75
5.5. Gráfica ROC para una Red Neuronal Amplia, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad. . . . .	76
5.6. Matriz de confusión para el algoritmo Naive Bayes, mostrando los patrones mejor clasificados correspondientes a cada clase. . . . .	77
5.7. Gráfica ROC para el algoritmo Naive Bayes, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad. . . . .	78
5.8. Matriz de confusión para el algoritmo K-NN, mostrando los patrones mejor clasificados correspondientes a cada clase. . . . .	79
5.9. Gráfica ROC para el algoritmo K-NN, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad. . . . .	80
5.10. Matriz de confusión para el algoritmo Random Forest, mostrando las características mejor clasificadas correspondientes a cada clase. . . . .	81
5.11. Matriz de confusión para el algoritmo Random Forest, mostrando las características mejor clasificadas correspondientes a cada clase. . . . .	82
5.12. Matriz de confusión para el algoritmo SVM Cúbico, mostrando las características mejor clasificadas correspondientes a cada clase. . . . .	83
5.13. Gráfica ROC para el algoritmo SVM Cúbico, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad. Clasificando en la clase de RD. . . . .	84
5.14. Matriz de confusión de RNA perceptrón multicapa. Clasificando en la clase 2 (RD Leve). . . . .	85
5.15. Gráfica ROC para el algoritmo RNA perceptrón multicapa, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad. Clasificando en la clase 2 (RD Leve). . . . .	86
5.16. Matriz de confusión la clasificación de glaucoma utilizando una SVM. Clasificando en la clase 0 (GPAA). . . . .	87
5.17. Gráfica ROC para el algoritmo SVM, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad. Clasificando en la clase 0 (GPAA). . . . .	88

## INTRODUCCIÓN

En diversas áreas de investigación se ha visto una gran simplificación de tareas gracias al avance en los sistemas de visión. En los últimos años se han creado una gran variedad de estos sistemas, de los cuales, resultan ser bastante útiles en diferentes aplicaciones prácticas. En el campo de la medicina, por ejemplo, se están desarrollando técnicas de visión artificial que permiten detectar enfermedades de manera automática y asistir en el diagnóstico de pacientes, lo cual ha llevado al desarrollo de múltiples sistemas que abordan diversas ramas de la medicina. [6], [77], [42], [57]. Los sistemas de visión han demostrado ser herramientas valiosas en la resolución de problemas en la vida real, ya que pueden procesar grandes cantidades de información visual de manera rápida y precisa, lo que permite automatizar tareas que antes eran tediosas o imposibles para los humanos. Un ejemplo de cómo los sistemas de visión han ayudado a resolver problemas de la vida real es en la industria automotriz. Los sistemas de visión se utilizan para identificar y reconocer objetos en la carretera, lo que permite la detección de peatones, ciclistas, señales de tráfico y otros vehículos. Esta tecnología ha permitido la implementación de sistemas avanzados de asistencia al conductor y ha mejorado significativamente la seguridad en la carretera [70].

En esta tesis se desarrollará un sistema de visión para detectar de forma automática el glaucoma y retinopatía diabética. El glaucoma se refiere a un grupo de trastornos oculares que dañan el nervio óptico y provocan pérdida de campo visual. Este trastorno se produce debido a una alteración en el drenaje del humor acuoso, lo que provoca su acumulación en el ojo, aumentando la presión ocular y afectando las fibras del nervio óptico [8]. Por otro lado, la retinopatía diabética es una afección ocular que ocurre en personas que tienen diabetes, y se produce debido al daño en los vasos sanguíneos que proveen de nutrientes a la retina en la parte posterior del ojo. Estos vasos debilitados pueden permitir la fuga de líquido o sangre, lo que daña o forma tejidos fibrosos en la retina y produce una imagen borrosa. También pueden aparecer ramificaciones frágiles similares a los cepillos y agrandamientos en ciertas áreas. Además, la presencia de retinopatía diabética en sus distintos grados puede aumentar el riesgo de complicaciones graves que afectan irreversiblemente la función visual [32]. El desarrollo de sistemas automáticos para la detección de estas enfermedades es muy importante, esto debido a que los sistemas automáticos no requieren la intervención de un profesional del área para realizar un diagnóstico. El desarrollo de sistemas de visión es importante, además de brindar un diagnóstico,

también permiten reconocer el grado de severidad con el que ciertos padecimientos afectan la salud visual de los pacientes [70].

## 1.1. PROBLEMÁTICA

De acuerdo a los estudios realizados por [20], se estima que en México hay alrededor de 1.5 millones de personas que padecen glaucoma, con hasta 50 mil casos de ceguera debido a la detección tardía de esta enfermedad. En América Latina, se estima que 15.5 millones de personas padecen glaucoma, de los cuales 10.6 millones son en México. Esto indica que aproximadamente el 10 por ciento de la población en riesgo de padecer retinopatía diabética (RD) [77]. Ambas enfermedades son asintomáticas y no tienen cura, lo que puede llevar a ceguera irreversible si no se detectan a tiempo. Sin embargo, el tratamiento temprano en la primera etapa del glaucoma de ángulo abierto puede retrasar la progresión de la enfermedad. Por esta razón, es crucial que el glaucoma se diagnostique oportunamente. Para diagnosticar el glaucoma en la actualidad, el médico debe ordenar la toma de fotografías del ángulo abierto o cerrado del ojo. Sin embargo, el diagnóstico certero factible solo para un médico con un entrenamiento de varios años en el área. El proceso tradicional de detección de enfermedades oculares, suele emplear un procesamiento de imágenes que utiliza diferentes algoritmos, por lo tanto, la precisión en el diagnóstico puede variar. Por otra parte, los algoritmos de aprendizaje automático pueden extraer con mayor exactitud las regiones de la imagen para posteriormente, clasificar los síntomas y brindar un diagnóstico más preciso.

Actualmente las herramientas que han sido desarrolladas pueden generar diagnósticos a través del procesamiento de imágenes o empleando algoritmos de aprendizaje automático, aunque la precisión es alta, no existe un sistema que determine el grado de avance que tiene un paciente que padece RD o glaucoma. Es importante que, además de diagnosticar, también sea posible especificar el grado en el que un paciente con RD o glaucoma se encuentra.

El objetivo de este trabajo de tesis es emplear métodos automáticos para el diagnóstico de enfermedades visuales, específicamente la retinopatía diabética y el glaucoma. Además, en caso de detectar la presencia de estas enfermedades, el método utilizado será capaz de evaluar el grado de gravedad o complicación que presentan los pacientes diagnosticados. Con esto se logrará reducir los costos de recursos computacionales y, al mismo tiempo, aumentar la precisión en el diagnóstico de glaucoma y retinopatía diabética (RD).

## 1.2. JUSTIFICACIÓN

El glaucoma es una enfermedad que ocasiona la ceguera irreversible y es la causa principal de este padecimiento a nivel mundial. Se estima que alrededor de 70 millones de personas en todo el mundo padecen glaucoma, de las cuales el 10 por ciento tienen ceguera en ambos ojos [20]. Para enfermedades oculares como el glaucoma y la retinopatía diabética, es crucial un tratamiento rápido para reducir su progreso y sus efectos. Un diagnóstico temprano es esencial para lograr este objetivo, pero solo puede ser hecho con precisión por un médico especialista que requiere varios años de entrenamiento y capacitación. Una buena identificación depende de la experiencia, entrenamiento del médico y el diagnóstico puede llevar tiempo en ser obtenido debido al glaucoma. Es por ello que son necesarias técnicas o dispositivos automáticos más eficientes y capaces de obtener un diagnóstico en menor tiempo.

La propuesta de solución desarrollada en este trabajo de investigación, permitirá detectar mediante un sistema de visión, la presencia de rasgos anormales en una imagen de la retina humana, con ello, será más sencillo diagnosticar el padecimiento examinado (RD o glaucoma).

Por lo anterior, el sistema de visión se va a encargar de optimizar las características con respecto al diagnóstico, además de brindar una detección, también especificará los diferentes grados de severidad de los padecimientos mencionados. Se realizará un análisis de diferentes métodos de procesamiento de imágenes para determinar cuál de ellos ofrece el diagnóstico más preciso, contribuyendo a disminuir la cantidad de diagnósticos erróneos tanto negativos como positivos.

## 1.3. OBJETIVOS

El propósito principal de esta tesis es crear un sistema automático de visión que emplee diversas técnicas de procesamiento digital de imágenes para clasificar enfermedades oculares como glaucoma y retinopatía diabética. Para alcanzar este objetivo general, se han establecido varios objetivos específicos, que incluyen:

1. Evaluar el rendimiento de distintos algoritmos de procesamiento digital de imágenes para detectar RD y Glaucoma.
2. Comparar con las técnicas clásicas de extracción de características y clasificación, empleadas en el estado del arte..
3. Clasificar el diagnóstico de RD y Glaucoma de acuerdo al grado de severidad.

4. Realizar pruebas para el funcionamiento del sistema de visión automático..
5. Discutir los resultados obtenidos.

## 1.4. HIPÓTESIS

¿Es posible mejorar el desempeño de detección automática de glaucoma y retinopatía diabética, utilizando técnicas clásicas de procesamiento de imágenes?

Las propiedades de la imagen del ojo humano que son relevantes para que un oftalmólogo determine la enfermedad que padece un paciente incluyen el tamaño, la forma y el color del nervio óptico, el espesor de la capa de fibras nerviosas de la retina y la presencia de lesiones o anomalías en la retina. También se pueden evaluar otras propiedades de la imagen, como la presencia de vasos sanguíneos anormales o de fluidos en la retina [9]. La capacidad del oftalmólogo para analizar estas propiedades se basa en su formación y experiencia en el examen de la imagen del ojo, y puede ser asistido por herramientas de diagnóstico como lámparas de hendidura, lentes de aumento y equipos de imagen avanzados .

Por otra parte, el sistema de visión utiliza determinadas técnicas de procesamiento de imágenes para mejorar las características visuales de una imagen para su análisis, posteriormente, se obtiene una imagen nueva donde su contenido ha sido adecuado para su procesamiento computacional. Al aplicar métodos clásicos a imágenes para detectar RD o glaucoma, se hace posible la adquisición de rasgos específicos, por ejemplo, las venas de la retina y el área general del ojo, lo que facilita al sistema de visión definir qué anomalías existen en cada imagen.

## 1.5. ANTECEDENTES

Para resolver el problema planteado en la sección 1.1 es necesario profundizar en dos áreas: sistemas de visión y medicina. En esta sección se exponen los fundamentos del problema que se aborda en esta tesis y las técnicas existentes en el estado del arte para solucionarlo.

### 1.5.1. Sistema de Reconocimiento de Objetos

La identificación de objetos es un campo en el ámbito de la visión artificial, que se dedica a desarrollar estrategias para detectar y clasificar objetos en una imagen o video. Mientras que los humanos podemos reconocer una gran variedad de objetos en imágenes con facilidad, incluso cuando estos difieren en perspectiva, tamaño, posición o incluso están parcialmente ocultos, esto representa un reto para los sistemas de visión artificial. A lo largo de varias décadas, se han implementado diversos métodos que imitan la habilidad humana de reconocimiento de objetos, permitiendo el reconocimiento independiente de la escala, rotación, posición, perspectiva y visión obstruida del objeto. Un aspecto fundamental de la identificación de objetos en imágenes involucra la extracción de información para establecer las características de objetos físicos o abstractos en una imagen [71].

El proceso empieza por definir el objeto a detectar, una vez definido se delimita el objeto en la imagen y se extraen las características que lo definen. Este conjunto de características puede ser obtenido mediante cualquiera de las técnicas en el estado del arte. Cada grupo de características posee una capacidad discriminativa diferente.

En términos generales, un sistema completo de reconocimiento de objetos involucra la utilización de sensores para capturar las observaciones a clasificar, un sistema que extrae las características de la información observada y las transforma en valores numéricos o simbólicos. El proceso de selección de características es fundamental para la clasificación de cualquier problema y permite la eliminación de características ruidosas que puedan inducir a error, características irrelevantes que no aportan información valiosa, o características redundantes que contienen la misma información que otras. Este proceso resulta en una disminución en los tiempos de procesamiento de los datos, una menor necesidad de espacio de almacenamiento, una reducción en el costo de la obtención de datos, y lo más importante, la selección de un subconjunto de las características originales que proveen la mayor cantidad de información para un problema específico. Un sistema de clasificación o descripción basado en

las características extraídas clasifica la medición de acuerdo a las características seleccionadas [1].

En la actualidad, uno de los mayores retos del reconocimiento de objetos es la creación de algoritmos capaces de extraer características altamente discriminativas y definir descriptores que permitan una representación precisa de dichas características, lo cual es fundamental para la visión computacional. A lo largo de décadas, se han implementado diversos enfoques para solucionar este desafío, ya que es uno de los problemas fundamentales en este campo.

### 1.5.2. Enfermedades de los Ojos

Antes de explicar el funcionamiento del ojo en relación al glaucoma y la retinopatía, es necesario definir algunos conceptos. El ojo cuenta con una serie de componentes, como la córnea que es la primera lente y brinda protección al ojo, el iris y la pupila que regulan la cantidad de luz que entra al ojo, el cristalino que es la segunda lente, la retina que es sensible a la luz, el humor vítreo que es un líquido transparente con consistencia gelatinosa entre el iris y la retina, y el humor acuoso que es un líquido transparente que nutre el ojo y va desde el cristalino hasta la córnea. El humor acuoso tiene un proceso de drenaje que permite regular la presión ocular y para que el ojo pueda percibir imágenes del mundo exterior, la luz entra por la pupila y es enfocada por el cristalino en la retina, que a su vez envía información al cerebro a través del nervio óptico. El desafío es entender cómo estos componentes se relacionan con enfermedades como el glaucoma y la retinopatía [68].

### 1.5.3. Glaucoma

El glaucoma es un conjunto de afecciones oculares que perjudican al nervio óptico. La presión intraocular elevada es la causa principal del daño al nervio óptico en la mayoría de los casos. [7], [44]. Los tipos de glaucoma pueden clasificarse dentro de los siguientes tipos:

- Glaucoma crónico de ángulo abierto: Es la variante más común de esta enfermedad, que se origina por la obstrucción de los conductos de drenaje en el ojo, impidiendo que los líquidos fluyan adecuadamente y provocando un aumento en la presión intraocular. En sus primeras etapas, este tipo de glaucoma puede no manifestar síntomas, por lo que su detección suele ocurrir durante exámenes visuales de rutina. Los síntomas de esta variante incluyen una pérdida gradual de la visión periférica y la sensación de ver a través de un túnel.

- Glaucoma agudo de ángulo cerrado: Se produce rápidamente cuando el flujo de líquidos en el ojo se bloquea súbitamente, lo que puede provocar dolores intensos, cambios en la visión, como halos alrededor de las luces, y malestar general, como dolores de cabeza, náuseas y vómitos.
- Glaucoma secundario: Puede ser causado por golpes, medicamentos o cirugías.
- Glaucoma congénito: Se presenta desde el nacimiento o poco después, y los síntomas pueden incluir ojos llorosos y turbios, así como aversión a la luz. Aunque es poco frecuente, este tipo de glaucoma es grave y se produce por una anomalía en el ojo.

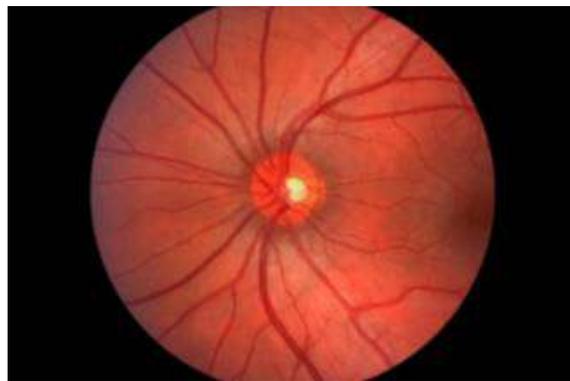


Figura 1.1: Ojo humano con diagnóstico de Glaucoma [44]

#### 1.5.4. Retinopatía

La retinopatía diabética es la enfermedad vascular más común que afecta la retina y se presenta en pacientes diabéticos con un mal control de la glucemia [7]. Es responsable de una importante pérdida de visión en personas entre 20 y 70 años en el mundo occidental, y es una de las causas más significativas de ceguera a nivel mundial. Para tratarla, se requiere un enfoque multidisciplinario que involucre al endocrinólogo y a los médicos que atienden al paciente, así como un seguimiento oftalmológico para considerar el tratamiento con láser de la retina si es necesario. Además, es crucial supervisar la aparición de glaucoma y el desarrollo de cataratas, ya que estos pacientes pueden requerir cirugía más temprana.

A continuación, se mencionan los diferentes grados de severidad de la retinopatía diabética:

- Retinopatía diabética no proliferativa: También conocida como retinopatía diabética leve o moderada, se caracteriza por la presencia de microaneurismas, exudados y tortuosidades en los vasos sanguíneos de la retina. Aunque esta

forma de retinopatía no es perjudicial para la visión, puede ser un indicador de una complicación más grave en el futuro.

- Retinopatía diabética proliferativa: Es una forma más avanzada de retinopatía diabética que puede causar la formación de nuevos vasos sanguíneos anormales en la retina, lo que se conoce como neovascularización. Estos nuevos vasos sanguíneos son frágiles y pueden sangrar y causar una descamación de la retina, lo que puede resultar en una pérdida severa de la visión.
- Microaneurismas: Son pequeñas protuberancias en los vasos sanguíneos de la retina que pueden causar la filtración de sangre en la retina.
- Exudados: Son manchas de líquido en la retina que pueden ser causadas por la filtración de sangre de los vasos sanguíneos dañados.

Encalada y Gamboa [37] llevaron a cabo un proyecto de investigación para identificar el glaucoma basándose en las características del nervio óptico. Para lograrlo, utilizaron técnicas de pre-procesamiento y segmentación para identificar la copa y el disco óptico. Sin embargo, encontraron que algunas personas tienen nervios ópticos más grandes, lo que podría afectar los resultados. Para solucionar este problema, crearon una regla que permite medir la relación entre la copa y el disco óptico. En individuos que padecen diabetes, los elevados niveles de glucosa en la sangre ocasionan daño a los vasos sanguíneos de la retina. Este daño puede dar lugar a inflamación, obstrucción o incluso la muerte de estos vasos. Como respuesta, el cuerpo intenta recuperar los vasos dañados generando nuevos vasos sanguíneos anómalos y frágiles, lo que puede desencadenar su ruptura, hemorragia y obstrucción de la visión.

Wong y Bressler [75] mencionan el sistema de detección de retinopatía basado en la tecnología de aprendizaje profundo de inteligencia artificial de Google. Esta tecnología ya se utiliza para identificar grandes volúmenes de datos, como las búsquedas de los usuarios o las imágenes cargadas en Google Photos. En esta ocasión, se aplicó en el campo de la medicina para procesar y evaluar imágenes de pacientes con retinopatía, con éxito en las pruebas. No obstante, Wong y Bressler también señalan algunas limitaciones en el estudio. En particular, el conjunto de datos con retinopatía severa es demasiado pequeño para que los resultados sean representativos de este tipo de casos, en los que realmente existe un riesgo de ceguera. Además, el algoritmo solo puede detectar la retinopatía, lo que excluye otros padecimientos oculares, como el glaucoma o la degeneración macular.

## ESTADO DEL ARTE

### 2.1. Trabajos relacionados

Varios investigadores se han interesado en las causas principales de las enfermedades visuales para detectarlas. Además, hay tecnologías que ayudan a los especialistas a tratar estos padecimientos. El propósito de los métodos diseñados para detectar enfermedades como RD y Glaucoma es proporcionar un diagnóstico temprano.

Algunos trabajos relacionados a la detección de enfermedades visuales como la RD y Glaucoma son los siguientes:

#### 2.1.1. Combinación de múltiples características profundas para la clasificación del glaucoma.

La autora Annan Li y sus colaboradores definen el glaucoma como una enfermedad que es la principal causa de ceguera en la población. En los últimos diez años se han creado diversas herramientas para diagnosticar el glaucoma de forma más precisa. No obstante, estos autores sugieren una nueva clasificación del glaucoma que combina múltiples características extraídas a través de las Redes Neuronales Convolucionales. Para llevar a cabo esta investigación, utilizaron el conjunto de datos Origa, el cual está disponible públicamente para su análisis. [45].

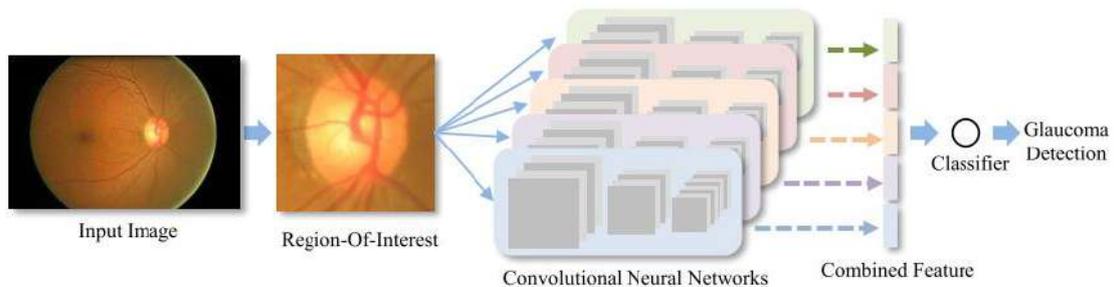


Figura 2.1: Proceso de detección de Glaucoma [45]

### 2.1.2. Síntesis de imágenes retinales y aprendizaje semisupervisado para la evaluación del glaucoma.

En su artículo, Andrés Díaz y su equipo de investigación emplean Redes Adversariales Generativas Convolucionales Profundas (DCGAN) para detectar el glaucoma. Entrenan un generador de imágenes de la retina y utilizan un método de aprendizaje semisupervisado para evaluar automáticamente el glaucoma, ya que han descubierto que se puede monitorizar el glaucoma analizando el disco óptico y su entorno. La innovación de esta propuesta es que ofrece un nuevo generador de imágenes de la retina que puede ser entrenado con un método de aprendizaje semisupervisado [27]. Han utilizado alrededor de 86926 imágenes para el entrenamiento, lo que no solo permite generar imágenes sintéticas, sino también asignar etiquetas a cada imagen. Por lo tanto, el generador de imágenes de la retina propuesto en este estudio puede utilizarse en futuros trabajos para generar un número ilimitado de imágenes de la retina con etiquetas de glaucoma.

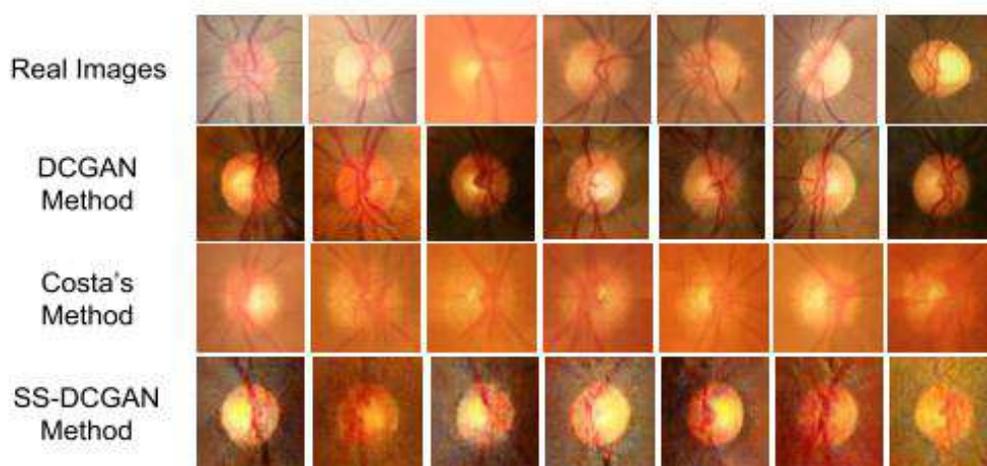


Figura 2.2: Diversos métodos empleados para la detección de Glaucoma [27]

### 2.1.3. Clasificación de la retinopatía diabética basada en patrones de exudados duros, mediante procesamiento de imágenes y SVM.

En su trabajo de investigación, Fernanda Cisneros y su equipo destacan la retinopatía diabética como una de las principales complicaciones de la diabetes, que puede dañar los vasos sanguíneos de la retina en ausencia de un tratamiento adecuado. Para mejorar la detección de esta enfermedad, proponen una clasificación automática de imágenes con exudados en el ojo afectado.

En el análisis, se utilizó la segmentación de imágenes para obtener características aisladas, y se aplicó una máquina de soporte vectorial para determinar la presencia de exudados. Con una base de datos de 130 imágenes, los resultados demostraron una alta sensibilidad del 87,3 por ciento y una especificidad del 84,6 por ciento. Estas características tienen un gran potencial para la detección y clasificación de la retinopatía diabética, y la máquina de soporte vectorial muestra resultados superiores a otros algoritmos de aprendizaje automático, con una sensibilidad cercana al 90 por ciento y una exactitud del 92 por ciento [22].

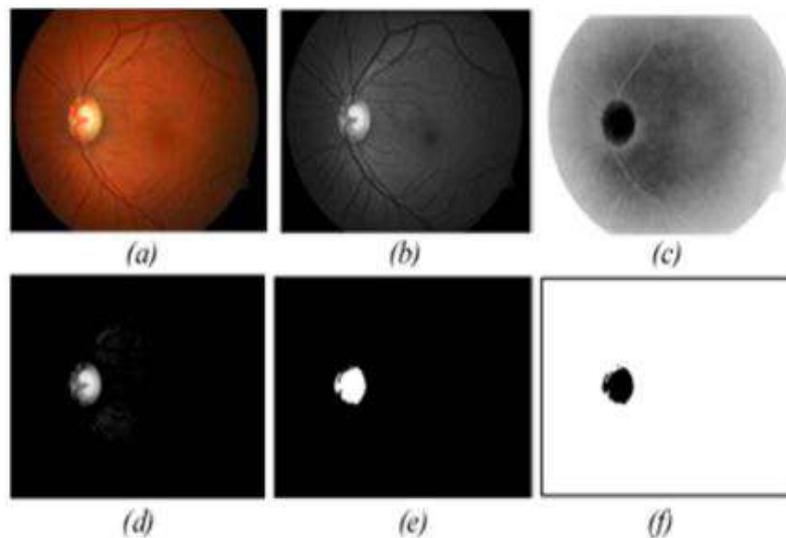


Figura 2.3: Disco óptico: (a) Imagen original, (b) Canal verde (c) Canal cyan, (d) Diferencia de las imágenes b y c, (e) Binarización, (f) Complemento de la imagen e. [22]

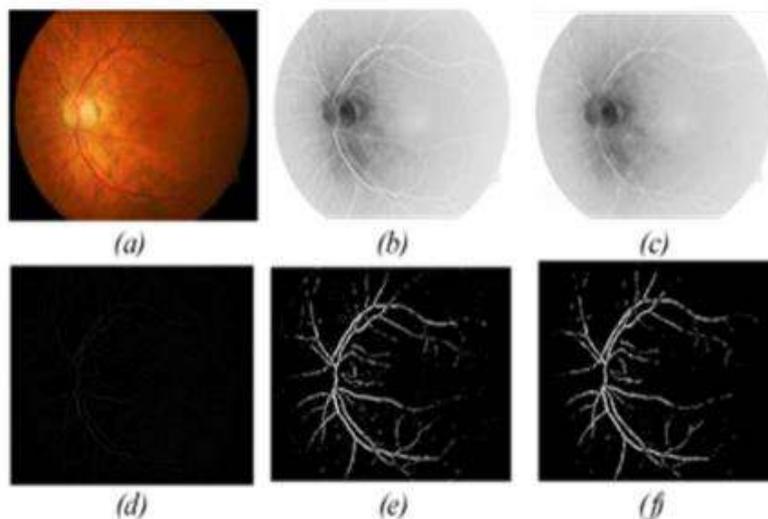


Figura 2.4: Vasos Sanguíneos: (a) Imagen original, (b) Componente cyan, (c) Procesamiento morfológico, (d) Diferencia de las imágenes b y c, (e) Binarización, (af) Reducción de ruido. [22]

#### 2.1.4. Diseño y desarrollo de un sistema de detección automática de Glaucoma a partir de la segmentación y el estudio de las capas de la retina en OCT Circumpapilares

Para desarrollar los modelos, se emplean técnicas de inteligencia artificial basadas en el aprendizaje automático. Inicialmente, se realiza una extracción de características relacionadas con el grosor y la textura de la capa RNFL de la retina en las imágenes médicas. Estas características extraídas son sometidas a un análisis estadístico detallado para identificar las más pertinentes para este estudio. Se utiliza un conjunto de datos que consta de 376 muestras, compuestas por 199 de individuos sanos y 177 de pacientes con enfermedades oculares. Para garantizar la robustez y confiabilidad de los modelos, se aplica una estrategia de partición de datos adecuada. Luego, se aborda el problema de la clasificación mediante la implementación de clasificadores no lineales, como Support Vector Machine, Perceptrón Multicapa y K-Nearest Neighbor [51].

Finalmente, se comparan los resultados obtenidos a través de estos modelos con los resultados previamente publicados en la literatura con el objetivo de aportar nueva información, identificar posibles mejoras y explorar nuevas áreas de investigación. Los resultados de sensibilidad y AUC obtenidos a través de los clasificadores de aprendizaje automático en este estudio son ligeramente menores en comparación con los obtenidos en el modelo propuesto.

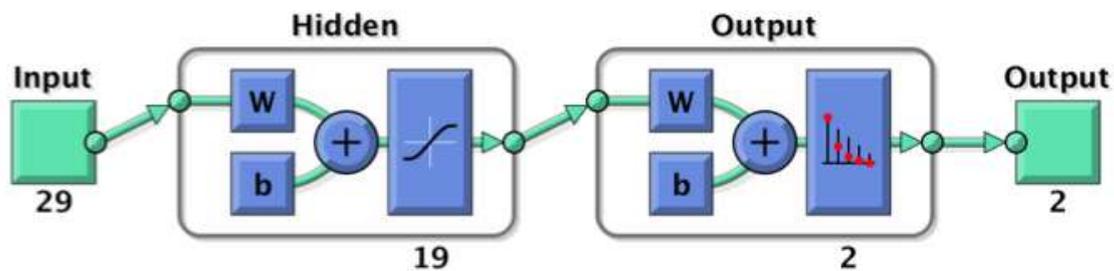


Figura 2.5: Modelo empleado de una Red Neuronal Perceptrón Multicapa, para trabajar con una validación interna. [51]

#### 2.1.5. Una base de datos a gran escala y un modelo CNN para la detección de glaucoma basada en la atención

Esta investigación realizada por Li Liu et. al, hace mención del daño irreversible que se presenta en la visión de un paciente que padece glaucoma. Aunque ya existen técnicas que permiten diagnosticar este padecimiento, no se puede evitar la presencia

en la redundancia de imágenes del fondo del ojo, es por ese motivo que se genera una precisión no tan confiable al momento de brindar un diagnóstico. Para evitar esta desventaja, los autores proponen una red neuronal convolucional (CNN) que está basada en la atención para detectar glaucoma, la cual, se denominada AG-CNN.

Primero se define una base de datos de glaucoma basada en la atención (LAG) a gran escala, dicha base de datos incluye al rededor de 11 760 imágenes del fondo de ojo etiquetadas como glaucoma positivo (4878) o glaucoma negativo (6882) [46].

Luego, se diseña una nueva estructura de AG-CNN, que incluye una subred de predicción de atención, una subred de localización de áreas patológicas y una subred de clasificación de glaucoma.

A diferencia de otros métodos de CNN basados en la atención, las características también se visualizan como el área patológica localizada, que se agregan aún más en nuestra estructura AG-CNN para mejorar el rendimiento de detección de glaucoma.

El enfoque AG-CNN propuesto por los autores ha demostrado avances significativos en la detección de glaucoma de vanguardia, según los resultados obtenidos en las pruebas realizadas en la base de datos de los autores LAG y otra base de datos pública de glaucoma.

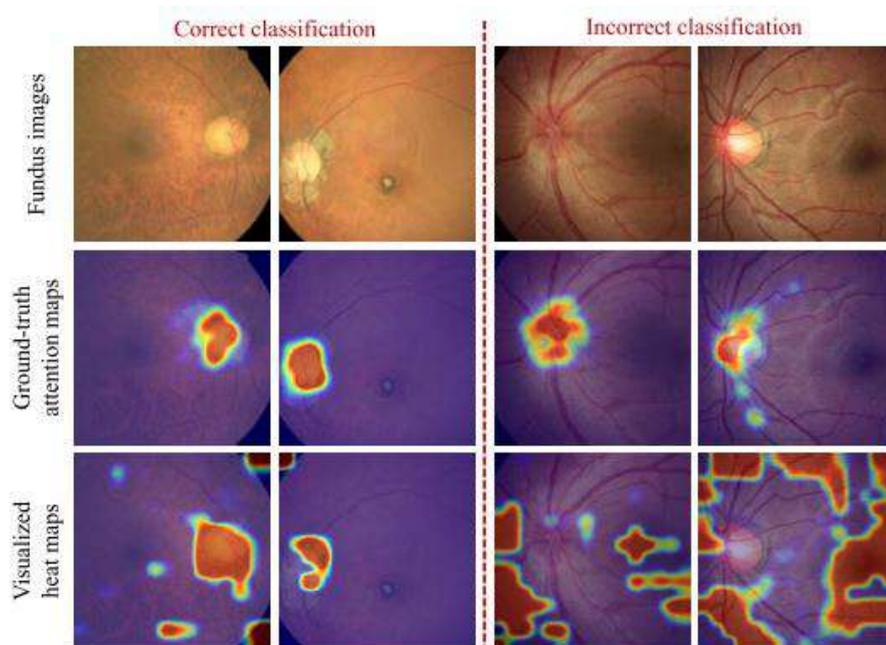


Figura 2.6: Ejemplos de fondos de imagen de Glaucoma. [46]

### **2.1.6. Clasificación de los ángulos de la cámara anterior en imágenes OCT del segmento anterior a través de redes neuronales convolucionales de regiones multi escala huaying**

Los autores de este artículo consideran al glaucoma de ángulo cerrado como una de las principales causas de ceguera en Asia y proponen un nuevo enfoque para clasificar los ángulos de la cámara anterior en abiertos, estrechos y cerrados en la tomografía de coherencia óptica del segmento anterior (AS-OCT). Para mejorar la eficacia de este trabajo y evaluar el glaucoma de ángulo cerrado, se aprenden las anotaciones manuales de la gonioscopia [34].

En este texto se describe un marco propuesto que localiza automáticamente la región del ángulo de la cámara anterior, lo cual es importante para identificar clínicamente el glaucoma. Luego, se aplican tres redes neuronales convolucionales de regiones multiescala (MSRCNN) en tres escalas de imágenes del ángulo de la cámara recortadas en la arquitectura definida por los autores.

Estas redes extraen representaciones de características, las cuales se apilan en una capa completamente conectada para clasificar el tipo de glaucoma. El método se evaluó con un conjunto de datos de aproximadamente 9,728 imágenes del ángulo de la cámara anterior y los resultados muestran que el método propuesto supera a los métodos de vanguardia existentes en términos de aplicabilidad, eficacia y precisión.

### **2.1.7. Conjunto de datos para la detección de glaucoma asistida por computadora**

En esta investigación se explica que la detección automatizada del glaucoma se ve obstaculizada por la falta de imágenes de grandes fondos de retina disponibles públicamente, lo que dificulta el estudio y diagnóstico para la inteligencia artificial. Además, los conjuntos de datos disponibles suelen tener imágenes poco nítidas y criterios estrictos, y a menudo no se permite su uso en la investigación. Por esta razón, la selección de conjuntos de datos existentes es limitada y hace que sea más difícil crear un sistema que funcione en un entorno real. Los autores del artículo presentan un conjunto de datos de imágenes de fondo de retina llamado G1020, que está disponible públicamente y se puede utilizar para la clasificación del glaucoma [5].

Por lo anterior, se presenta un conjunto de datos que se ha diseñado de acuerdo a las prácticas más comunes de la oftalmología de rutina y que se espera que sirva como un conjunto de datos estándar para la detección de glaucoma. La base de datos se

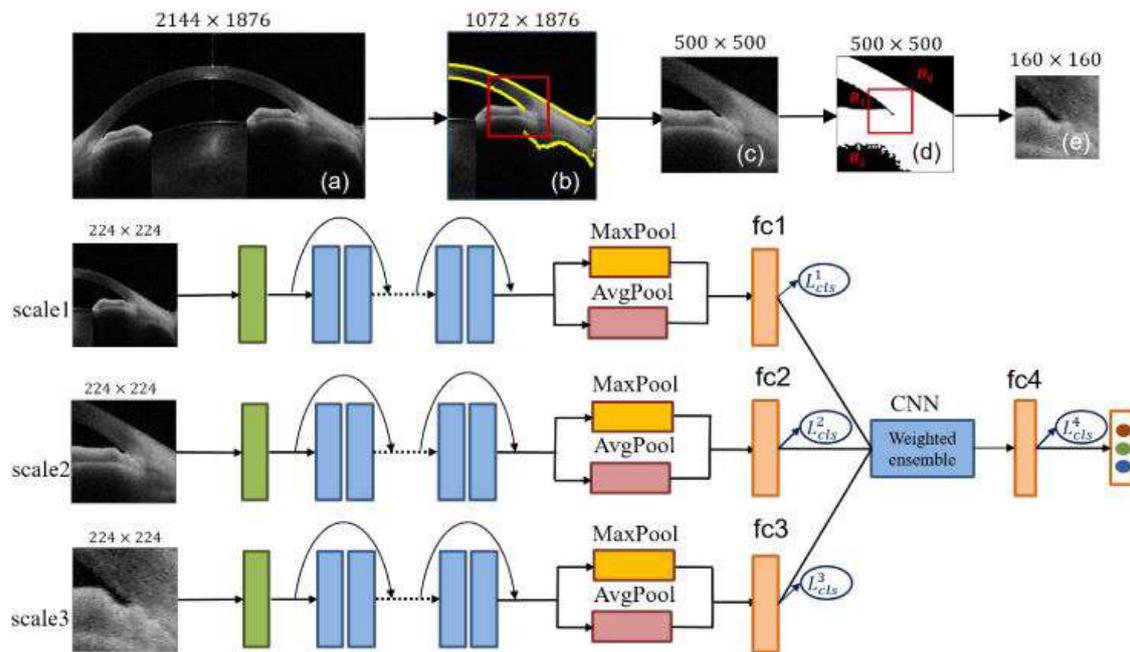


Figura 2.7: Descripción general de la arquitectura de red propuesta para la clasificación de los ángulos de la cámara anterior. [34]

compone de alrededor de 1020 imágenes en color de fondo de ojo con alta resolución, lo que permite proporcionar anotaciones precisas para el diagnóstico de glaucoma, la segmentación del disco óptico y la copa óptica, la relación copa-disco vertical, el tamaño del borde neuroretiniano en los cuadrantes inferior, superior, nasal y temporal, y la ubicación de la caja para el disco óptico. Los autores señalan que se han realizado extensos experimentos y que los resultados muestran que el diagnóstico de glaucoma puede automatizarse y que la segmentación del disco óptico y la copa óptica pueden agilizar el proceso.

### 2.1.8. Clasificación automatizada de las etapas del glaucoma utilizando la transformada de wavelet analítica flexible a partir de imágenes del fondo de la retina

El glaucoma es uno de los padecimientos visuales más comunes, sus síntomas suelen ser silenciosos, aunque, en casos más severos, la presencia de intenso dolor y pérdida progresiva de la visión, hacen que las personas acudan a la realización de estudios con un oftalmólogo. Por otra parte, la detección temprana del glaucoma suele ser crucial ya que puede ayudar a desacelerar el progreso. Sin embargo, como desventaja de la detección de este padecimiento es a través de métodos e instrumentos tradicionales que suelen ser manuales, por lo cual, la mayoría de ellos consumen mucho tiempo

y son menos precisos. Por esta razón, se necesita el diagnóstico automatizado de glaucoma para la detección oportuna en una etapa temprana con alta precisión. Los autores proponen un novedoso método basado en la transformada de wavelet analítica flexible (FAWT) para la clasificación de las etapas del glaucoma.

En el método propuesto, se ha empleado FAWT para desintegrar las imágenes pre-procesadas en más imágenes de subbanda. Luego, se aplican los algoritmos ReliefF y de conteo de cajas secuenciales (SBC) para extraer las diferentes características que están basadas en entropías y dimensiones fractales (FD), respectivamente. Una vez realizado el primer paso del trabajo propuesto, los valores de características que han sido extraídos se clasifican mediante una reducción dimensional del análisis discriminante lineal (LDA) de Fisher [56].

Por último, en los resultados obtenidos, se puede apreciar las características de mayor rango que se han utilizado para clasificar las etapas de glaucoma utilizando máquinas de soporte vectorial (SVM) de mínimos cuadrados (LS-SVM). Es por ello que este proceso ha sido estudiado y analizado en una base de datos de glaucoma grande y diversa disponible públicamente. La precisión de clasificación del método propuesto es del 93,40 por ciento utilizando una validación cruzada de diez veces. El método diseñado y probado por los autores ha demostrado un mejor rendimiento para la clasificación del glaucoma en comparación con los métodos existentes, por lo cual, se define que podrá funcionar como un asistente que brinde apoyo a los oftalmólogos en la detección de glaucoma.

## MARCO TEÓRICO

En esta sección del documento se presentan las técnicas y algoritmos específicos de procesamiento de imágenes digitales que se utilizaron para llevar a cabo esta tesis. El capítulo se divide en seis secciones, cada una detallando los materiales y métodos que se utilizaron para desarrollar la metodología de esta investigación.

### 3.1. Técnicas de preprocesamiento

El objetivo principal del preprocesamiento de imágenes es mejorar las características de una imagen original para obtener una imagen de mayor calidad que facilite la búsqueda de información relevante o mejore la interpretación de la misma. En la siguiente sección se describen en detalle los filtros utilizados para mejorar la imagen.

#### 3.1.1. Escala de grises

La técnica descrita consiste en eliminar la información de tonalidad y saturación de una imagen, manteniendo solo los valores de luminosidad. Las imágenes en blanco y negro tienen múltiples tonos de gris y no contienen colores, lo que reduce el tamaño del archivo al convertir una imagen en color a escala de grises. Es importante destacar que esta técnica reduce significativamente la cantidad de información de la imagen original.

La ponderación de cada canal de color en relación a las frecuencias del espectro cercanas al rojo, verde y azul, es representada por la ecuación de la luminancia [43]:

$$Y = R * 0,3 + G * 0,59 + B * 0,11 \quad (3.1)$$

Para obtener una imagen en escala de grises, es necesario aplicar la ecuación 3.1 a cada uno de los píxeles de la imagen a color.

#### 3.1.2. Espacio de color RGB

El espacio de color utilizado en este estudio se basa en una síntesis aditiva, que es un modelo matemático que describe cómo se puede obtener cualquier color a partir de la suma de otros colores. Este modelo se basa en el uso de los colores rojo, verde

y azul para generar todos los demás colores [19]. Cada color en el modelo RGB se representa mediante una combinación de valores de intensidad de cada uno de los tres colores primarios.

La figura 3.1 ejemplifica una representación gráfica tridimensional del espacio de color RGB, en el cual el color negro corresponde al punto de origen y el color blanco se encuentra en el vértice más alejado del origen. Para el procesamiento adecuado de imágenes a color, los valores de Rojo, Verde y Azul deben encontrarse en el rango  $[0-255]$ . Es importante asegurar que las imágenes a color se encuentren en este rango para evitar problemas en el procesamiento de las mismas y obtener resultados precisos y fiables.

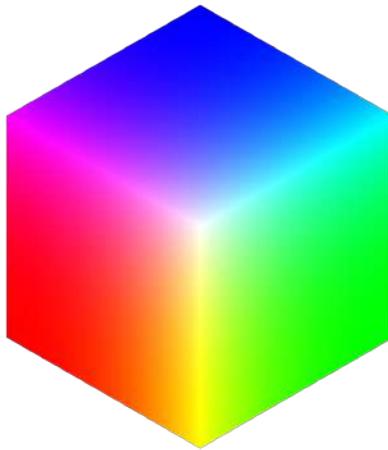


Figura 3.1: Espacio de color RGB

Este modelo simple de color, es muy utilizado para el procesamiento del color, ya que cuando se utiliza siempre que se pretenden transformar colores a otro espacio de color diferente al RGB.

### 3.1.3. Espacio de color HSI

La representación HSI (Hue, Saturation, Intensity) se refiere a un modelo de color que utiliza tres componentes para definir un color en términos de su matiz, saturación e intensidad. Esta representación se puede visualizar como un cono doble que tiene como vértices los colores blanco y negro. El matiz se representa mediante un ángulo, la saturación mediante la distancia al eje y la intensidad mediante la distancia al eje blanco-negro. La representación en forma gráfica de este modelo es una figura tridimensional que muestra la relación entre estos tres componentes:

- **Matiz:** Es la longitud de onda dominante en la imagen, o también se puede entender como cada uno de los colores en estado puro. Se calcula con la siguiente fórmula:

$$H = \begin{cases} \theta & \text{si } B \leq G \\ 360 - \theta & \text{si } B > G \end{cases}$$

con:

$$\theta = \cos^{-1} \frac{\frac{1}{2}(R - G) + (R - B)}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \quad (3.2)$$

Se puede estandarizar el valor de la matiz al intervalo  $[0,1]$  al dividir todos los valores que resultan de la fórmula 3.2 por  $360^\circ$ .

- **Saturación:** La saturación es una medida de la pureza o intensidad de un color, donde los valores más altos corresponden a colores puros y los valores más bajos a tonos de gris. Su cálculo se realiza mediante una fórmula específica:

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)] \quad (3.3)$$

- **Intensidad:** La luminosidad se refiere a la cantidad de luz que emite o refleja un objeto y determina la claridad u oscuridad de un color. Se puede calcular mediante la siguiente fórmula:

$$I = \frac{1}{3}(R + G + B) \quad (3.4)$$

### 3.1.4. Normalización de Histograma

Se conoce como histograma a una representación gráfica de los niveles de intensidad de una imagen, en la que se muestra la relación entre el número de píxeles presentes

en la imagen y la intensidad de cada color. En este tipo de gráfica, el valor de gris mínimo corresponderá a cero y el valor máximo a 255, lo que significa que se puede visualizar la distribución de los tonos de gris presentes en la imagen. Es decir, un histograma muestra la cantidad de píxeles en una imagen que tienen un valor determinado de intensidad de color y permite analizar la luminosidad de una imagen de manera precisa.

Cuando se habla de normalizar, se realiza una suma de todos los bins (rangos de números), que, propiamente, es la probabilidad que tiene una variable  $X$  de tomar ciertos valores, dicha suma tiene que ser igual a 1 y, posteriormente se hace una división por la cantidad total de datos. La normalización permite que los valores de un histograma sean las más equivalente posible, en caso de que dichos valores sean dispares, se puede normalizar dicho histograma de la siguiente manera:

$$N_f(r) = \frac{h_f(r)}{\text{size}} \quad (3.5)$$

Donde size es una matriz de tamaño  $N \times M$ .

La normalización de un histograma es importante para el preprocesamiento de imágenes en muchas aplicaciones de visión por computadora, especialmente en técnicas que requieren una representación uniforme de la intensidad de la imagen, por ejemplo, en el caso de la segmentación de imágenes o la identificación de características.

### 3.1.5. Ecuación de histograma adaptativo limitado por contraste (CLAHE)

La ecualización de histograma es una técnica utilizada en el procesamiento digital de imágenes para mejorar su calidad visual. Su objetivo es aumentar el rango dinámico de una imagen digital, es decir, mejorar el contraste y la claridad de la imagen. Esta técnica se basa en una operación sobre funciones continuas en un dominio determinado. [31].

Para profundizar en la explicación anterior, se define la variable  $s$  como una variable aleatoria que puede asumir valores en un rango de 0 a 1. La densidad de probabilidad de la variable  $s$  es representada por  $f(s)$  y su función de distribución es  $F(s)$ . Si se desea definir una nueva variable continua,  $r$ , se debe tener en cuenta la siguiente ecuación:

$$r = F(s) = \int_0^s f(t) dt \quad (3.6)$$

Cuando los valores para la variable  $s$  se encuentran demasiado ajustados, es necesario que se extienda el rango de manera uniforme sobre el intervalo definido,  $[0 \ 1]$ .

Para el caso de variables continuas, donde  $f(s)$  es la densidad ligada al histograma, se emplea la solución anterior. Por otra parte, en el caso de variables discretas, se discretiza la integral, realizando la división del intervalo  $[0-1]$  en 256 partes iguales, con esta solución se espera obtener un resultado correspondiente a una escala de 256 niveles de gris [31]. Por lo tanto, para el caso discreto se aplica la siguiente ecuación:

$$F(k) = \sum_{i=0}^K f \frac{i}{256} * \frac{1}{256} \quad (3.7)$$

En la práctica suele trabajarse con valores entre 0 y 255, por esta razón, se emplea la siguiente ecuación:

$$F(k) = 255 \sum_{i=0}^K P_i \quad (3.8)$$

En la parte izquierda de la Figura 10 se encuentra la imagen original, donde se puede apreciar que la imagen es oscura, por lo tanto, hay una mayor concentración de píxeles que están en el nivel de intensidad cerca del 0, esto quiere decir que el que prevalece el color negro, mientras que, en la derecha se aprecia el histograma, el cual permite visualizar los canales rojo, verde y azul, por tratarse de una imagen en escala de grises. Los píxeles lucen un poco más concentrados en el centro de la imagen, lo que se busca al aplicar una ecualización es distribuir la intensidad de los píxeles.

Ahora en la parte izquierda de la Figura 3.2 se representa el proceso de ecualización en una imagen en escala de grises (imagen original). Primero se ajusta la distribución de los niveles de intensidad, posteriormente se mejora el contraste en la imagen original, y, finalmente, se obtiene una lucidez más precisa en la imagen nueva. El histograma luce un poco más plano que en la Figura 3.3, debido al ajuste de las intensidades.

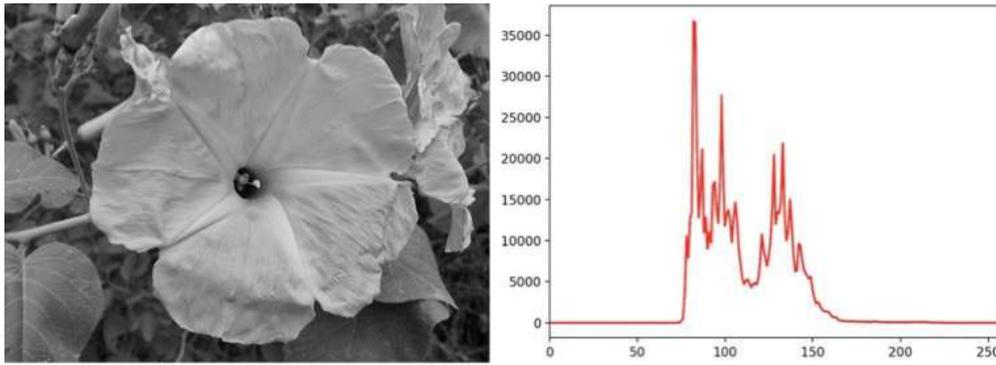


Figura 3.2: Imagen en escala de grises con su respectivo histograma, mostrando la distribución píxeles concentrada en el centro de la imagen.

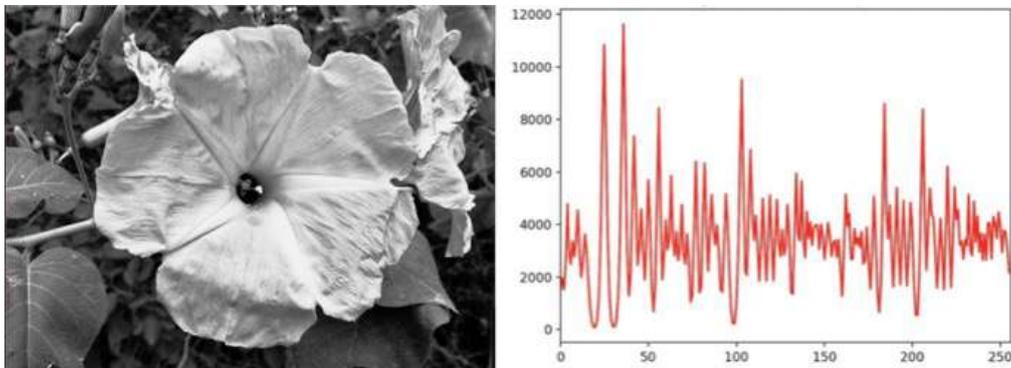


Figura 3.3: Proceso resultante de equalización aplicada a una imagen en escala de grises, mostrando una distribución más plana en el histograma.

Por otra parte, CLAHE aumenta el contraste de una imagen y mejora la precisión de la extracción de detalles. Para lograrlo, compara el histograma local de la imagen y redistribuye el contraste, lo que produce una mayor cantidad de detalles en la imagen. En comparación con la Ecuación de Histograma Adaptativo (AHE), que redistribuye el valor de contraste en diferentes regiones de la imagen, CLAHE define la forma del histograma para obtener una mejor calidad de imagen [50]. Además, este algoritmo puede reducir significativamente el ruido y mejorar su desempeño a través del uso de la interpolación, un proceso que estima valores desconocidos basados en los valores conocidos de la muestra. El proceso del algoritmo CLAHE consiste en:

1. Se adquiere una imagen con ruido.
2. Se determinan los valores de entrada:
  - a) Número total de las regiones por separado, tomando en cuenta filas y columnas.
  - b) Número total de bandejas usadas en la función de transformación del his-

tograma, también conocidas como rango dinámico.

- c) Definir el tipo de parámetros de distribución.
3. Dividir la imagen adquirida en diferentes secciones para que las entradas sean procesadas.
  4. Se aplica un proceso sobre la región de interés.
  5. Se realiza una creación de un mapa que considera los niveles de grises en la imagen y el histograma recortado. Luego, se divide el total de píxeles dentro de la sección de interés de manera equitativa en cada nivel de gris, determinando el nivel de gris promedio con la siguiente fórmula:

$$N_{avg} = \frac{N_{CR} - X_p * N_{CR} - Y_p}{N_{gray}} \quad (3.9)$$

Donde:

$N_{avg}$ , que es el número de niveles de gris presentes en la región de interés, y  $N_{CR} - X_p$  y  $N_{CR} - Y_p$ , que son el número de píxeles en las direcciones  $X$  e  $Y$  de la región de interés, respectivamente.

Posteriormente, se realiza el cálculo del límite actual del clip.

$$N_{CL} = N_{CLIP} * N_{avg} \quad (3.10)$$

6. Por último, se realiza una interpolación utilizando un mapa de escala de grises, lo que resulta en una nueva imagen modificada.

CLAHE trabaja en pequeñas regiones de la imagen, conocidas como mosaicos, en lugar de procesar la imagen completa. Se calcula una función de transformación de contraste para aplicarla a cada mosaico, con el objetivo de mejorar su contraste individual y lograr una salida que se asemeje al histograma de distribución especificado. Luego, los mosaicos vecinos se combinan mediante interpolación bilineal para evitar la aparición de límites artificiales. Por último, se limita el contraste en las zonas homogéneas para prevenir la amplificación del ruido existente en la imagen. Este proceso resulta en una imagen mejorada con mayor claridad y detalle en las regiones de interés. Es importante destacar que este método es particularmente útil para mejorar imágenes en las que se requiere un alto contraste local, como en radiografías o imágenes médicas en general [50].

Por lo anterior, el objetivo principal del algoritmo CLAHE es ajustar el contraste de la imagen original para mejorar su clasificación, lo que lo hace útil en una variedad

de algoritmos. No obstante, su función es particularmente notable cuando se utiliza en una red neuronal convolucional. [76].

### 3.1.6. Realzado de bordes

Para el proceso de segmentación, es importante identificar los bordes o fronteras de los objetos de la imagen. Estas fronteras se localizan en las regiones donde la intensidad de la imagen presenta cambios notables. El realzado de bordes permite ampliar dichas variaciones de la intensidad, exactamente en el espacio donde se producen, de esta manera se realzan los detalles de una imagen que hayan podido quedar borrosos. Se parte del principio de que mientras más brusco sea el cambio de intensidad, más fácil es detectar el borde [74].

## 3.2. Técnicas de segmentación

La segmentación de imágenes es el proceso mediante el cual se divide una imagen en varios segmentos. El propósito principal de esta técnica es modificar la representación de una imagen en partes más significativas y fáciles de analizar. En la mayoría de los casos, la segmentación se utiliza para localizar objetos y establecer límites claros entre ellos en una imagen.

Para nada es recomendable procesar una imagen completa, porque muchas partes de la imagen pueden contener información que no es útil para el problema que se está resolviendo, por lo tanto, al segmentar una imagen, se puede hacer uso solo de los segmentos importantes para el procesamiento. Por tal motivo es un paso importante dentro del procesamiento de imágenes, y es uno de los pasos que se realizan antes de extraer las características relevantes de la imagen y posteriormente, reconocer los patrones.

Existen tres principios básicos en los que se basa la segmentación:

1. **Similitud:** Los píxeles de un objeto tienen valores similares con respecto a alguna propiedad específica. En este contexto, se estudia la regularidad de los valores de nivel de gris presentes en la imagen.
2. **Discontinuidad:** Los objetos y las regiones que se desean segmentar suelen destacar del entorno y presentar contornos o bordes más definidos que los objetos que se encuentran en su entorno.
3. **Conectividad:** Los píxeles que pertenecen a un mismo objeto o región tienden

a estar adyacentes, es decir, a agruparse y formar regiones homogéneas en la imagen.

En otras palabras se podría entender a la segmentación como un proceso de clasificación por pixel en donde se asigna una clase a cada pixel de la imagen analizada, generando una imagen del mismo tamaño que la analizada, en la que cada pixel expresa una etiqueta. El problema de la segmentación se puede dividir en diferentes problemas especializados:

- Segmentación por color
- Segmentación por texturas
- Segmentación semántica
- Segmentación de instancias

La imagen resultado de una segmentación esta preparada para ser consumida por un algoritmo de extracción de características, pero no para que pueda ser visualizada. Es decir, para presentar el resultado de una segmentación es necesario crear una imagen con etiquetas o representar el resultado mediante contornos, que son líneas cerradas que realzan los bordes de los segmentos encontrados.

Es importante señalar, que existen muchos algoritmos de segmentación, y se crean nuevos cada día, por lo tanto se ha desarrollado una gran variedad de enfoques aplicando conocimientos específicos para realizar esta tarea. Para este trabajo de investigación se van a utilizar el método Otsu.

### 3.2.1. Método Otsu

Esta técnica se basa principalmente en una varianza o dispersión de los valores, su objetivo es buscar que dicha varianza se encuentre dentro de las clases definidas. Primero se debe obtener el cálculo del histograma normalizado de la imagen, representado por:

- $M \times N$  se define como una función  $f(x, y)$ .
- Definir el total de niveles de gris en la imagen desde 0, 1, 2, . . . , L.
- La probabilidad de ocurrencia del nivel en la intensidad  $i$ :

$$p_i = \frac{n_i}{MN} \quad (3.11)$$

Calcular las sumas acumulativas:

$$w_1(k) = \sum_{i=0}^k p_i \quad (3.12)$$

Si se establece un valor umbral  $T$  en la intensidad  $K(T = K)$ , se divide el conjunto de píxeles en dos clases distintas:

- C1= Son aquellos píxeles que poseen una intensidad de  $[0, k]$ .
- C2= Son aquellos píxeles con una intensidad de  $[k+1, L]$ .

$$w_1(t) = \sum_{i=0}^T p_i \quad (3.13)$$

$$w_2(t) = \sum_{z=T+1}^L p_i \quad (3.14)$$

A continuación, se calculan las medias acumulativas y las varianzas entre las dos clases divididas:

$$\mu_1(t) = \sum_{i=1}^T \frac{ip(i)}{w_1(t)} \quad (3.15)$$

$$\mu_2(t) = \sum_{i=1}^T \frac{ip(i)}{w_2(t)} \quad (3.16)$$

$$\sigma_1^2(t) = \sum_{i=1}^T (i - \mu_2(t))^2 \frac{p(i)}{w_1(t)} \quad (3.17)$$

$$\sigma_2^2(t) = \sum_{i=1}^T (i - \mu_2(t))^2 \frac{p(i)}{w_2(t)} \quad (3.18)$$

Por último, la varianza entre las clases se define como:

$$\sigma_w^2 = w_1(t)\sigma_1^2(t) + w_2\sigma_2^2(t) \quad (3.19)$$

Después de aplicar las ecuaciones mencionadas anteriormente, se selecciona un valor umbral  $T$  que corresponde a la varianza existente entre las dos clases de píxeles.

### 3.2.2. Método basado en Clustering (DBSCAN)

El algoritmo de agrupamiento DBSCAN (Agrupamiento espacial basado en densidad de aplicaciones con ruido) es un método de aprendizaje no supervisado que busca identificar grupos o clústeres distintivos en los datos. Su funcionamiento se basa en la idea de que un grupo en el espacio de datos se define como una región contigua de

alta densidad de puntos, que está separada de otros grupos por regiones contiguas de baja densidad de puntos [29].

Este algoritmo es capaz de descubrir grupos de diferentes formas o tamaños a partir de una gran cantidad de datos, incluso si contienen ruido y valores atípicos. Para lograr esto, DBSCAN utiliza dos parámetros fundamentales:

1. **minPts:** Se refiere al número mínimo de puntos que deben estar agrupados para que una región se considere densa. Es decir, se trata de un umbral que determina cuántos puntos deben estar presentes en una región para que esta sea considerada un grupo.
2. **eps( $\epsilon$ ):** Es una medida de distancia, se emplea para ubicar los puntos cercanos a cualquier punto en su entorno.

Los parámetros anteriores pueden comprenderse a partir de dos conceptos:

- **Accesibilidad:** La definición establece que un punto es considerado alcanzable desde otro si está ubicado dentro de una distancia específica,  $\epsilon$ , de este último.
- **Conectividad:** Para determinar si los puntos pertenecen a un grupo en particular, se utiliza un enfoque de encadenamiento basado en la transitividad.

Existen tres tipos de puntos, una vez que se ha ejecutado el algoritmo de agrupamiento DBSCAN.

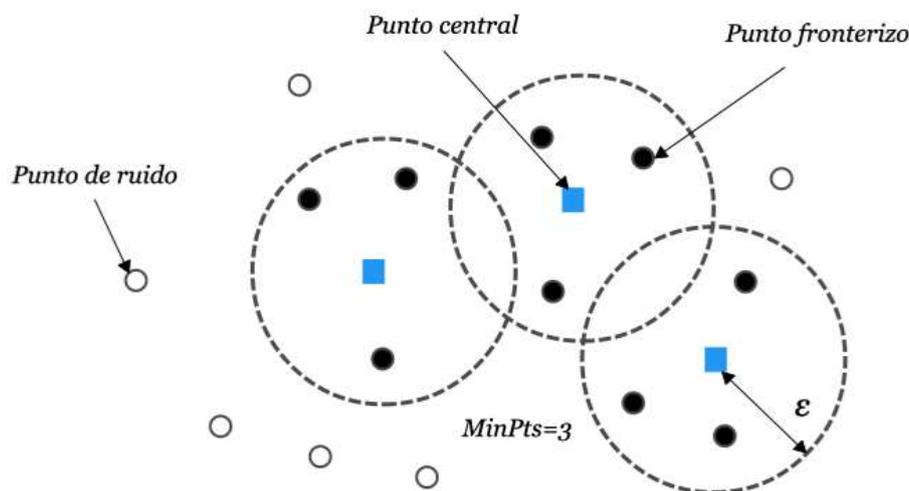


Figura 3.4: Tipos de puntos detectados por el algoritmo DBSCAN

1. **Punto central:** Se define como un punto que tiene ciertos  $m$  puntos dentro de una distancia  $n$  de sí mismo.
2. **Punto fronterizo:** Es un punto que tiene al menos un punto central a una

distancia  $n$ .

3. **Punto de ruido:** Se trata de un puntos no central ni fronterizo. Tiene menos de  $m$  puntos a una distancia  $n$  de sí mismo.

### Algoritmo DBSCAN paso a paso

- El algoritmo inicia seleccionando arbitrariamente un punto en el conjunto de datos(hasta que se hayan seleccionado todos los puntos).
- Si hay al menos ' $minPts$ ' puntos dentro de un radio de ' $\epsilon$ ' al punto seleccionado, entonces se considera que todos esos puntos son parte del mismo grupo.
- Los grupos se expanden repitiendo recursivamente el cálculo de la vecindad para cada punto vecino.

### 3.3. Operaciones Morfológicas

La morfología matemática es una herramienta que se utiliza para analizar y tratar las estructuras de los objetos en las imágenes digitales. Es la base del procesamiento de imágenes morfológicas, que utiliza un conjunto de operadores que pueden ser aplicados en imágenes binarias o en escala de grises, para transformarlas de acuerdo a las características específicas indicadas [67].

En esta teoría matemática, una imagen es vista como un subconjunto de un espacio de números enteros en 2 dimensiones. Luego, para aplicar estos operadores morfológicos es necesario definir un elemento estructural denotado por  $B$ , que se puede definir como una subimagen utilizada para sondear a la imagen de estudio [60]. Es esencial que este componente estructural esté compuesto por disposiciones rectangulares, que se crean al agregar la menor cantidad necesaria de elementos base para constituir una matriz rectangular.

Los puntos representados en la figura 3.5 denotan los centros de los elementos estructurales. Los operadores morfológicos se encargan de simplificar y mantener las características relevantes en la estructura de los objetos.

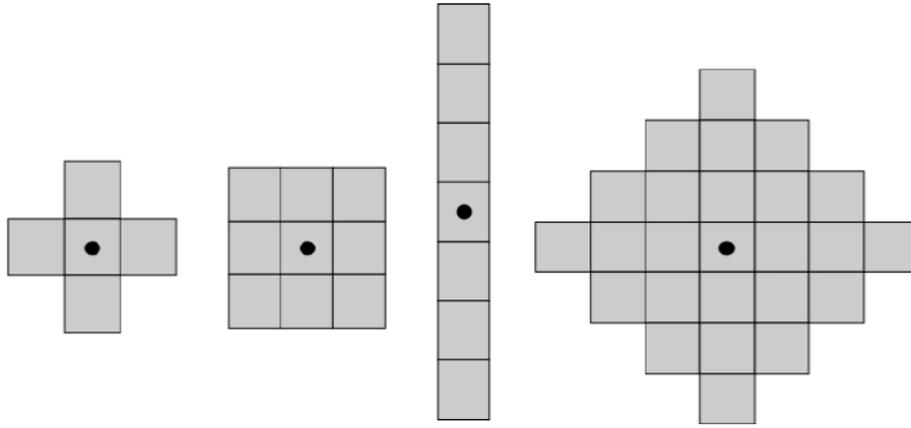


Figura 3.5: Elementos estructurados

### 3.3.1. Erosión

La erosión dada por  $A$  y  $B$ , denotada como  $A \ominus B$  se define como la operación que se realiza en conjuntos de  $A$  y  $B$  en  $Z^2$ :

$$A \ominus B = \{z | (B)_z \subseteq A\} \quad (3.20)$$

La ecuación anterior representa la erosión de  $A$  por  $B$ , resultando un conjunto donde todos los puntos  $z$  tales que  $B$ , trasladados por  $z$ , están contenidos en  $A$ . Cuando el elemento estructural  $B$  tiene un centro y este centro se encuentra en el origen de  $A$ , entonces la erosión de  $A$  por  $B$  se puede ver como los puntos alcanzados por el centro de  $B$  cuando  $B$  se mueve dentro de  $A$ .

Ejemplo:

### 3.3.2. Dilatación

La dilatación de  $A$  por  $B$  en  $Z^2$ , denotada como  $A \oplus B$ , se define como el conjunto de todas las sumas de un elemento en  $A$  y un elemento en  $B$ , donde la suma se realiza:

$$A \oplus B = \{z | (\hat{B}) \cup A \neq \emptyset\} \quad (3.21)$$

La dilatación de  $A$  por  $B$  es un conjunto de elementos que se obtiene al sumar cada elemento de  $A$  con cada elemento de  $B$ , de manera modular. Para entender esto, se puede imaginar reflejar  $B$  sobre su origen y desplazar esta reflexión en  $z$ . Además, se pueden considerar todos los desplazamientos de  $z$  que permiten que  $B$  y  $A$  se superpongan en al menos un elemento.

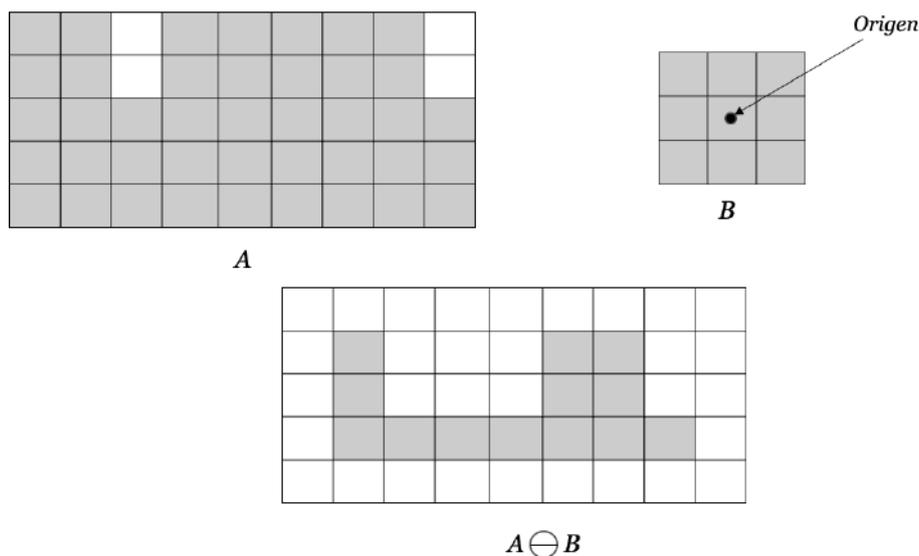


Figura 3.6: a) Conjunto A, b) Elemento estructurado B, c) A erosionada por B.

En caso de que  $B$  tenga un centro en el origen, la dilatación se puede interpretar como los puntos cubiertos por  $B$  cuando el centro de  $B$  se mueve dentro de  $A$ .

### 3.4. Extracción de características

La extracción de características es un proceso fundamental en el campo de la inteligencia artificial y el aprendizaje automático. Se refiere a la identificación y selección de aspectos específicos e informativos de los datos originales que son relevantes para resolver un problema o realizar una tarea específica. [12], [13].

La extracción de características es crucial porque permite reducir la dimensionalidad de los datos y eliminar información redundante o poco relevante, lo que a menudo mejora el rendimiento de los modelos de aprendizaje automático. En resumen, es el proceso de identificar y resaltar aspectos clave de los datos que ayudan a los algoritmos a tomar decisiones más precisas y efectivas en una amplia gama de aplicaciones, desde la visión por computadora hasta el procesamiento de texto y más allá [11], [40].

#### 3.4.1. Descriptor Haralick

Según [35], los descriptores de Haralick son medidas estadísticas que se usan para describir las propiedades texturales de una imagen, y son ampliamente empleados en la visión por computadora y el análisis de imágenes en general.

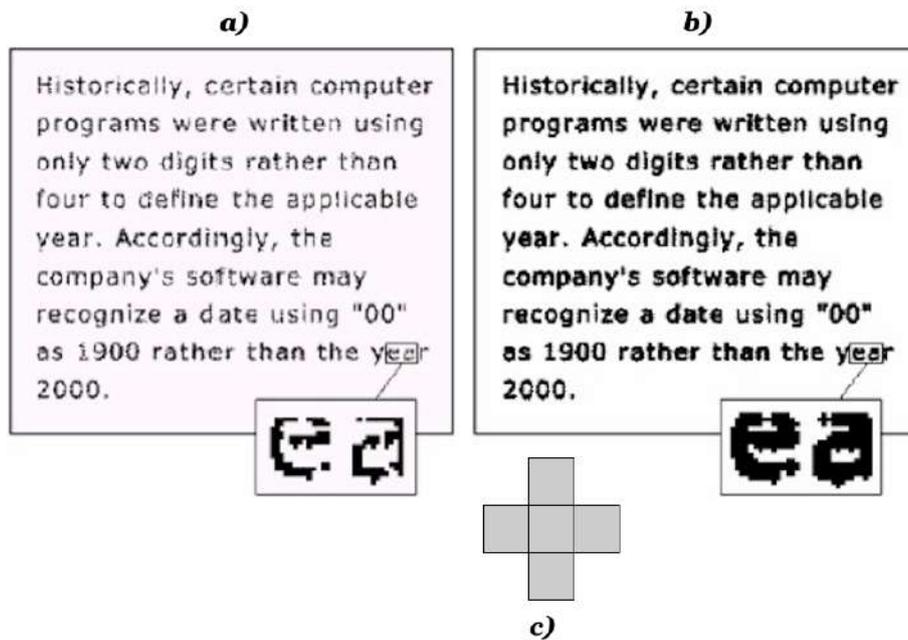


Figura 3.7: **a)** Texto de muestra de baja resolución con caracteres rotos (vista ampliada), **b)** Elemento estructurado, **c)** Dilatación a) por b). Se unieron los caracteres rotos.

La matriz de co-ocurrencia, cuando se aplica a imágenes, es una representación de cuántas veces ciertos píxeles o colores aparecen juntos en una imagen. En lugar de palabras, estamos analizando píxeles en una imagen. La matriz de co-ocurrencia aplicada a imágenes es una herramienta que ayuda a analizar la distribución y las relaciones entre píxeles o colores en una imagen para extraer información sobre patrones y texturas [51].

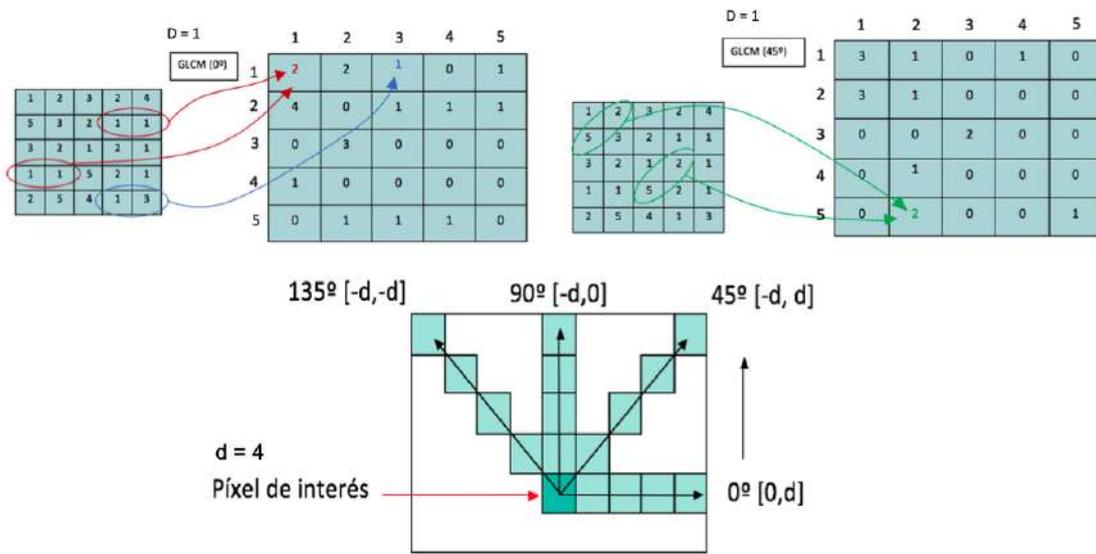


Figura 3.8: Representación del cálculo de la Matriz de co-ocurrencia de niveles de gris. Donde se puede observar que, en la parte superior hay dos ejemplos de construcción de la matriz, tomando en cuenta 1 píxel adyacente, así como, la dirección horizontal y diagonal. Y, en la parte inferior, se observa los offset correspondientes a los diferentes ángulos [51].

El fundamento estadístico de los descriptores de Haralick se basa en la premisa de que toda la información textural de una imagen se agrupa en las relaciones espaciales entre los distintos niveles de gris de un elemento. Para obtener los descriptores, se utilizan los diferentes momentos estadísticos del histograma, que contiene los niveles de intensidad de una región específica. En este sentido, se define  $px(i)$  con  $i = 0, 1, 2, 3, \dots, N - 1$  como el histograma correspondiente, donde  $N$  representa los distintos niveles de intensidad. Si se aplica a la representación en bytes,  $N$  tomará el valor de 256 [24], [21], [66].

Es posible calcular el valor medio  $m$  de  $z$  empleando la siguiente ecuación:

$$m = \sum_{i=0}^{L-1} z_i p(z_i) \quad (3.22)$$

Donde la ecuación 3.22 representa el nivel medio de la intensidad en una región determinada, la cual está definida por los valores de  $z_i$  (intensidad) y su respectivas probabilidades  $p(z_i)$ .

Para alcanzar el  $n$ -ésimo momento de  $z$  con respecto a la media se toma en cuenta la ecuación:

$$\mu_n(z) = \sum_{i=0}^{L-1} z_i^n p(z_i) \quad (3.23)$$

Los tipos descriptores de Haralick incluyen:

- Contraste: se encarga de medir la preferencia de intensidad en una imagen.
- Energía: mide la homogeneidad de una imagen.
- Entropía: mide la aleatoriedad en una imagen.
- Homogeneidad: mide la similitud entre intensidades vecinas en una imagen.
- Correlación: mide la relación lineal entre intensidades vecinas en una imagen.
- ASM (Media de la Suma de los cuadrados): mide la complejidad textural en una imagen.
- Diferencia de promedios: mide la diferencia entre intensidades vecinas en una imagen.
- Diferencia de varianzas: mide la seleccionada en intensidades vecinas en una imagen.

Es posible combinar los descriptores de Haralick de diversas maneras para crear un perfil textural más completo y complejo. Esta técnica se emplea en aplicaciones como la detección de objetos, la segmentación de imágenes y la clasificación de imágenes.

Según [35] y [72], existen diferentes descriptores Haralick, a continuación, se muestran los más utilizados:

$$f_1 = \sum_i \sum_j [p(i, j)^2] \quad (3.24)$$

$$f_2 = \sum_i^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) |i - j| = n \right\} \quad (3.25)$$

$$f_3 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [ijp(i, j) - \mu_x \mu_y]}{\sigma_x \sigma_y} \quad (3.26)$$

$$f_4 = \sum_i \sum_j (i - \mu_x)^2 p(i, j) \quad (3.27)$$

$$f_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \quad (3.28)$$

$$f_6 = \sum_{i=2}^{2N_g} iP_{x+y}(i) \quad (3.29)$$

$$f_7 = \sum_{i=2}^{2N_g} (i - f_8)^2 P_{x+y(i)} \quad (3.30)$$

$$f_8 = \sum_{i=2}^{2N_g} P_{x+y(i)} \log\{P_{x+y(i)}\} \quad (3.31)$$

$$f_9 = - \sum_i \sum_j p(i, j) \log\{p(i, j)\} \quad (3.32)$$

$$f_{10} = - \sum_{i=0}^{N_g-1} (i - f_8)^2 P_{x-y}(i) \quad (3.33)$$

$$f_{11} = - \sum_{i=0}^{N_g-1} (i - f_8)^2 P_{x-y}(i) \quad (3.34)$$

$$f_{12} = \frac{H_{xy} - H_{xy1}}{\max\{H_x, H_y\}} \quad (3.35)$$

$$f_{13} = (1 - e^{[-2(H_{xy2} - H_{xy})]})^{1/2} \quad (3.36)$$

Donde:

- $p(i, j)$ : Se define el valor de la matriz de co-ocurrencia en la posición  $(i, j)$ .
- $N_g$ : Se refiere a los niveles de intensidad presentes en la región segmentada de la imagen.
- $p_x(i)$ : Se refiere a la  $i$ -ésima entrada en la matriz de co-ocurrencia y es el resultado de la suma de las columnas de  $p(i, j)$ .
- $R$ : Es el número total de los elementos en la matriz de co-ocurrencia.
- $\mu_x, \mu_y$ : Se refieren al promedio de  $P_x$  y de  $P_y$  respectivamente.
- $\sigma_x, \sigma_y$ : Representan la desviación estándar de  $P_x$  y de  $P_y$  respectivamente.
- $H_x, H_y$ : Representan la entropía de  $P_x$  y de  $P_y$  respectivamente.

### 3.4.2. Descriptor LBP (Patrón Binario Local)

El algoritmo de patrón binario local se utiliza para extraer características de textura de una imagen. Su objetivo es etiquetar los píxeles de una imagen en función de

su vecindad de umbral, asignando el valor del píxel central a cada uno de ellos. El resultado se considera como un número binario. [33]. Debido a su bajo costo computacional, LBP se ha utilizado en diversas soluciones para obtener características de textura. Además, su robustez ante las variaciones de intensidad de los niveles de gris en una imagen permite aplicaciones cada vez más realistas. En la actualidad, la forma en que se utiliza LBP suele ser diferente a su versión básica. Originalmente, se trabajaba con un conjunto de píxeles vecinos arbitrarios dispuestos en una configuración circular, pero se han desarrollado versiones adicionales del algoritmo. En general, LBP genera un código binario que describe un patrón de textura local, generado a partir del umbral de los píxeles vecinos en relación al valor de gris del píxel central [26].

La introducción de LBP como algoritmo de textura se debe a [53]. El operador LBP se utiliza para etiquetar los píxeles de una imagen mediante matrices de 3x3, comparando cada píxel con el valor del centro de la matriz y generando un resultado binario [26].

La derivación del algoritmo LBP parte de la definición de  $T$ , que es la textura de la imagen, compuesta por una distribución de niveles de gris en los píxeles. Para ser válida, esta distribución debe cumplir la siguiente condición:

$$P + 1(P > 0) \quad (3.37)$$

Ahora, para definir a  $T$  se emplea la siguiente fórmula:

$$T = t(g_c, g_0, g_1, g_2, \dots, g_p - 1) \quad (3.38)$$

El valor de  $(g_c)$  se refiere al nivel de gris del píxel ubicado en el centro de la matriz  $n \times n$ , mientras que  $(gp)$  se refiere a los valores de gris de los píxeles que se encuentran en un círculo definido por un radio  $R(R > 0)$  y que están en un conjunto circular de los vecinos respecto al centro [26]. Si las coordenadas de  $(gc)$  son iguales a  $(0, 0)$ , entonces las coordenadas de  $(gp)$  estarán dadas por:

$$\left(-R \sin\left(\frac{2\pi * p}{P}\right), R \cos\left(\frac{2\pi * p}{P}\right)\right) \quad (3.39)$$

En la siguiente imagen se pueden apreciar tres diferentes conjuntos de vecinos, los cuales se ajustan circularmente para los distintos valores de  $P$  y  $R$ :

Posteriormente, para tener presente la información, se pueden restar los puntos de

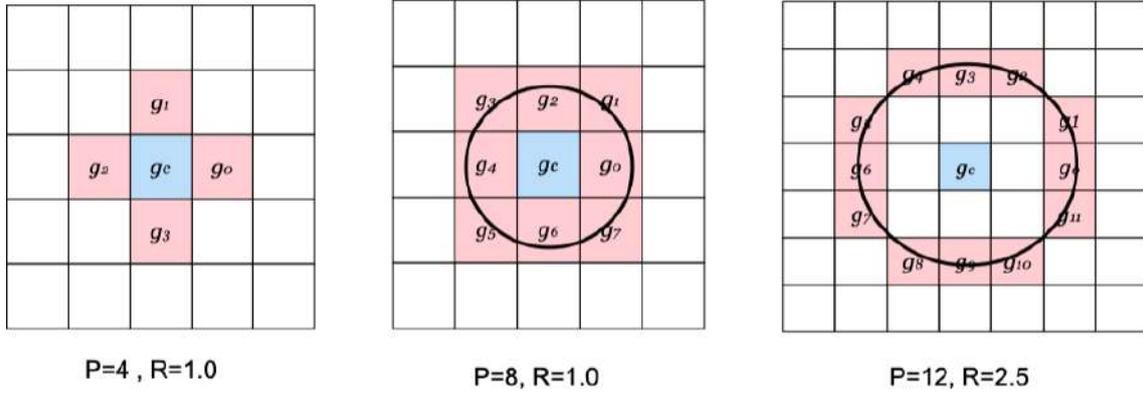


Figura 3.9: Conjuntos de vecinos circularmente simétricos [26].

$(g_c)$  junto con  $(g_p)$ , esto se realiza de la siguiente manera:

$$T = t(g_c, g_0 - g_c, g_1 - g_c, g_2 - g_c, \dots, g_{p-1} - g_c) \quad (3.40)$$

Asimismo, por la ecuación anterior, se asume entonces que las diferencias son independientes de la distribución, por lo tanto, la expresión se puede factorizar de la siguiente manera:

$$T \approx t(g_c)(g_0 - g_c, g_1 - g_c, g_2 - g_c, \dots, g_{p-1} - g_c) \quad (3.41)$$

Debido a que  $t(g_c)$  describe la luminosidad de una imagen, que no está relacionada con la textura de la imagen local, es posible discriminarla. En consecuencia, se tiene lo siguiente:

$$T \approx (g_0 - g_c, g_1 - g_c, g_2 - g_c, \dots, g_{p-1} - g_c) \quad (3.42)$$

Para lograr invariancia frente a cualquier transformación constante de escala de grises [26] [53], se consideran únicamente los signos de las diferencias, de manera que las diferencias no se vean afectadas por dichos cambios de escala:

$$T \approx t(s(g_0 - g_c), s(g_1 - g_c), s(g_2 - g_c), \dots, s(g_{p-1} - g_c)) \quad (3.43)$$

Por lo cual,  $S$  define dos condiciones:

$$S(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x \leq 0 \end{cases}$$

Ahora, a cada signo  $S(g_p - g_c)$  se le brinda un peso  $2^p$ , transformando la diferencia en el muestreo a un único código LBP que se define como:

$$LBP = \sum_{p=0}^{p-1} S(g_p - g_c)^{2^p} \quad (3.44)$$

### Invariancia a la rotación

La invariancia a la rotación se refiere a la rotación de la imagen [33], cuando esta acción se ejecuta, los valores correspondientes de gris ( $g_p$ ) suelen moverse respecto a la longitud del círculo, al rededor de ( $g_0$ ).

Para asignar un identificador único que permita eliminar el efecto de rotación, se define la siguiente ecuación:

$$LBP_{P,R}^{r_i} = \min\{ROR(LBP_{P,R}, i) | i = 0, 1, 2, \dots, P - 1\} \quad (3.45)$$

En la que  $LBP_{P,R}(g_c)$  es el patrón LBP para el píxel en la posición  $(c, r)$  con radio  $R$  y  $P$  vecinos. El operador de rotación  $R$  representa la rotación en sentido antihorario. El operador  $\min$  representa el valor mínimo entre dos números binarios. De esta manera, se obtiene un patrón LBP único e invariante ante la rotación

Donde  $ROR(x, i)$  se encarga de desplazarse circularmente a nivel de bit, dirigiéndose hacia la derecha, es decir,  $P$ -bits de  $x$   $i$ -veces. En resumen, la ecuación anterior aplica un giro al conjunto de vecinos hacia la derecha respetando al bit más significativo respecto a  $g_{p-1}$ .

Por otra parte,  $LBP_{P,R}^{r_i}$  se encarga de cuantificar las rotaciones que ocurren para los patrones que corresponden a características de la imagen, por lo tanto, puede considerarse un proceso de detección de rasgos [26] [53].

La siguiente imagen muestra las 36 rotaciones invariantes de los patrones binarios locales que se generan para el caso donde  $P=8$ , es decir,  $LBP_{8,R}^{r_i}$ . Esto significa que existen 36 valores diferentes posibles:

La imagen previa muestra los patrones binarios locales resultantes de LBP, donde los círculos blancos y negros representan bits con valores de 0 y 1, respectivamente, en la salida de 8 bits del operador. En la primera fila se muestran los nueve patrones uniformes y los números dentro de ellos representan su único valor  $LBP_{8,R}^{r_i}$ .

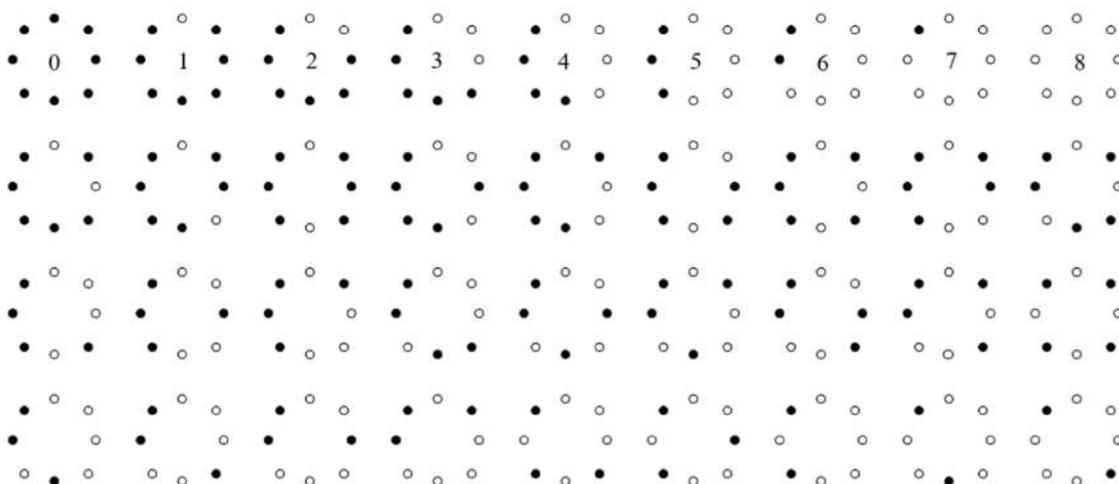


Figura 3.10: Las rotaciones invariantes de patrones binarios pueden ocurrir en el conjunto de ocho vecinos que tienen simetría circular. [33]

### 3.4.3. Descriptor HOG (Histograma de gradientes orientados)

Los descriptores HOG son utilizados para detectar y describir las características locales de una imagen, específicamente la orientación del gradiente en estas áreas. Su finalidad es aprovechar la información del gradiente y partir de esta información, combinar los datos en forma de histogramas orientados locales, los cuales, se calculan en celdas de un tamaño más reducido, la distribución de las celdas debe estar de manera uniforme en la imagen [63]. Por lo tanto, la información analizada permitirá distinguir las formas de cada uno de los objetos en la imagen.

Posteriormente, una vez que se adquiere la información, los histogramas que han sido calculados en cada celda generada, se agrupan en bloques un poco más grandes. Estos bloques se encargan para normalizar la imagen resultante. Los descriptores HOG brindan información respecto a los cambios de intensidad o los bordes de una imagen [63].

Según [25] y [63], las etapas que integran al algoritmo de aplicación de descriptores HOG son las siguientes:

1. **Normalización del espacio de color:** Esta etapa también es conocida como pre-procesamiento, consta de dos sub etapas que se encargan de ecualizar el color de la imagen para reducir la iluminación.
  - a) Imagen en escala de grises: Toma la imagen a color, donde sus componentes

se convierten en a escala de grises a partir de la siguiente ecuación:

$$Y = R * 0,3 + G * 0,59 + B * 0,11 \quad (3.46)$$

- b) Corrección gamma: Tiene la finalidad de aumentar o disminuir el brillo general de una imagen. El método más común es el método de la raíz cuadrada, donde la ecuación se representa:

$$Y(x, y) = I(x, y)^\gamma \quad (3.47)$$

2. **Cálculo de gradientes:** Para calcular un gradiente en el punto central, se aplica la siguiente relación:

$$dx = A(x + 1, y) - A(x - 1, y) \quad (3.48)$$

$$dy = A(x, y + 1) - A(x, y - 1) \quad (3.49)$$

Donde:

- a) dx: Es la resta horizontal entre el píxel anterior y el siguiente.  
 b) dy: Es la resta vertical entre el píxel anterior y el siguiente.

3. **Histograma de la dirección del gradiente:** En esta etapa, se generan varias secciones de la imagen, formando un número específico de celdas. Posteriormente, de cada celda se obtiene un histograma de las direcciones de los gradientes.

- a) Primero, se debe definir el tamaño de la cada celda en la imagen.  
 b) Posteriormente, se debe tomar en cuenta el rango de orientaciones en un intervalo fijo. Cuando se toma la orientación del gradiente con signo, el rango será de 0 a 360°. De lo contrario, si se toma la orientación del gradiente sin signo, el rango irá desde 0 a 180°.  
 c) Por último, el valor del intervalo seleccionado en el histograma final, será adquirido para una concentración de la magnitud de los gradientes asignados a ese intervalo. Formalmente el cálculo del histograma se representa de la siguiente manera:

$$h(k) = \sum_{(x,y) \in C} \omega(k)(x, y)g(x, y) \quad (3.50)$$

4. **Normalización del bloque:** En esta etapa las celdas en conjunto generan grupos en grandes bloques espacialmente conectados. Dichos bloques, generalmente se juntan provocando que cada celda contribuya en más de una vez en el descriptor final, lo que aumenta rendimiento del sistema. Posteriormente, se asigna un vector a cada bloque para que éste sea normalizado, en el trabajo de Dalal [25] se mencionan tres expresiones para normalizar un vector:

$$L2 - norm : f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}} \quad (3.51)$$

$$L1 - norm : f = \frac{v}{\|v\|_1 + e} \quad (3.52)$$

$$L1 - sqrt : f = \sqrt{\frac{v}{\|v\|_1 + e}} \quad (3.53)$$

Como se observa en las tres expresiones anteriores,  $v$  define al vector que no está normalizado, el cual, contiene a todos los histogramas que son parte del grupo de un bloque, por otra parte,  $\|v_k\|$  representa al vector  $k$ -normalizado de los histogramas correspondientes a cada bloque, donde el valor es igual a  $k=1,2$ , por último,  $e$  es definida como una constante que contiene un valor muy cercano a cero.

5. **Agrupación de los descriptores:** Una vez que los bloques hayan sido creados, éstos se enlazan y, el descriptor resultante adquiere una descripción completa en la ventana de detección, que posteriormente se emplea en un clasificador.

### 3.5. Métodos de validación

Los métodos de validación, son técnicas que permiten conseguir una clasificación más certera, la idea central es que el banco de datos se divide en dos subconjuntos: conjunto de entrenamiento y conjunto prueba.

El conjunto de entrenamiento son los datos con los que el algoritmo de clasificación va a aprender, y el conjunto de prueba son los datos con los que se va a evaluar el resultado de la clasificación, esto con ayuda de una o más métricas de rendimiento. Para este trabajo de investigación, se ha decidido utilizar el método de validación cruzada (K-Fold Cross Validation).

### 3.5.1. K-Fold Cross Validation

El método de validación cruzada está resumido por los siguientes pasos:

- Definir el valor de  $K$ , con el que el algoritmo va a trabajar.
- Dividir en  $K$  subconjuntos el banco de datos que se va a trabajar.
- El Conjunto de Entrenamiento queda formado por  $K - 1$  subconjuntos del banco de datos.
- El Conjunto de Prueba queda formado por uno de los  $K$  subconjuntos del banco de datos.
- Se repite este proceso con un total de  $K$  ciclos o iteraciones, con cada uno de los posibles subconjuntos usados para el Conjunto de Prueba.
- El algoritmo de clasificación también se repite  $K$  iteraciones, utilizando un Conjunto de Entrenamiento y Prueba diferente en cada repetición.
- Al finalizar todas las iteraciones, se promedian los resultados obtenidos para obtener un único valor mediante la media aritmética.

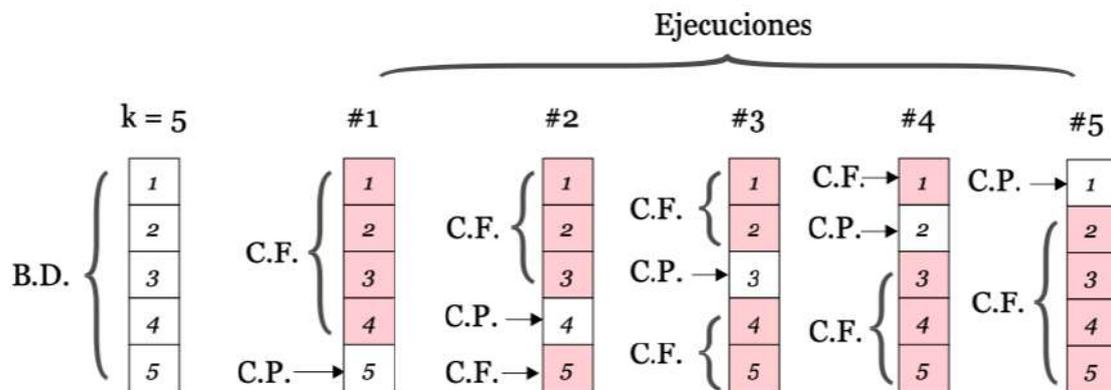


Figura 3.11: Ejecuciones de K-Fold Cross Validation con  $K=5$

## 3.6. Algoritmos de clasificación

Se emplean principalmente cuando el resultado que se desea obtener es discreto.

Estos algoritmos se enfocan en la clasificación general de la información brindada por un conjunto de datos, los cuales pertenecen a un entrenamiento que es utilizado para representar clases, asimismo, conservan algunas etiquetas para ser referidos a la clase correcta [73], [55], [41], [39]. A este conjunto de prototipos correctamente

etiquetados se les llama conjunto de entrenamiento, este conocimiento servirá a las nuevas muestras para realizar la clasificación correcta. La clasificación se divide en dos tipos:

- **Clasificación Supervisada:** Su objetivo principal es determinar, según lo que se tenga conocimiento, cual es la clase a la que debería concernir una nueva muestra, teniendo en cuenta la información que se pueda extraer.
- **Clasificación no Supervisada:** A diferencia de la clasificación supervisada, en la clasificación no supervisada no se asignan categorías de manera predefinida, en lugar de eso, se basa en la identificación de patrones y agrupaciones basadas en probabilidades.

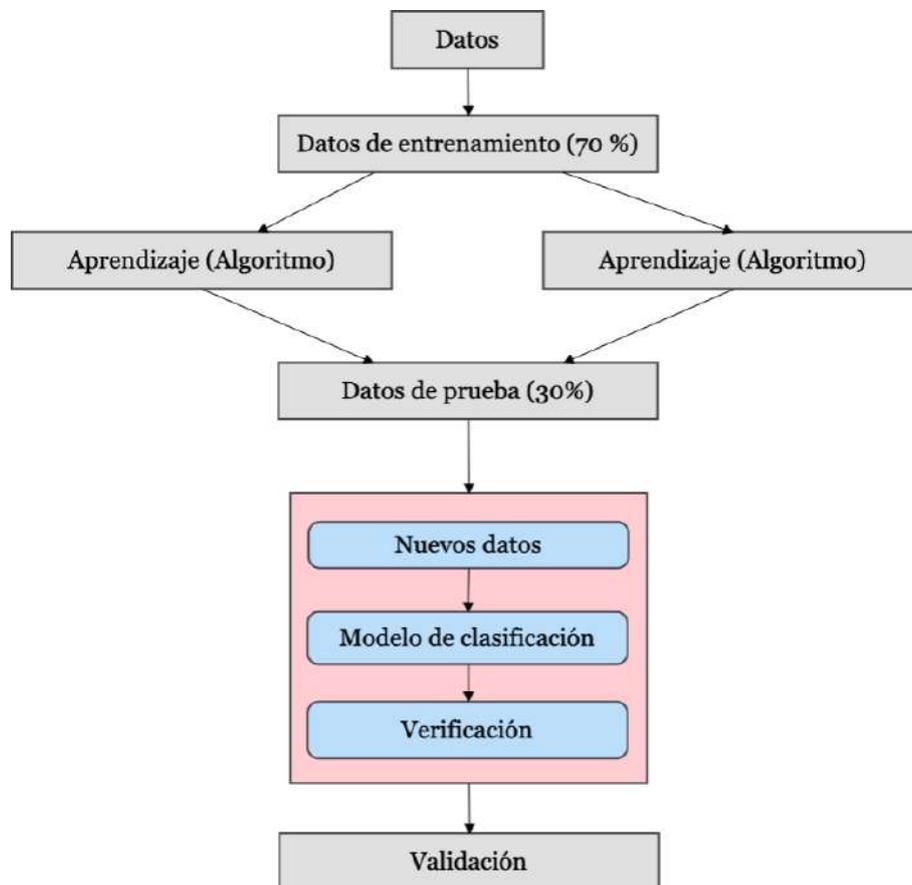


Figura 3.12: Metodología de clasificación de aprendizaje automático

### 3.6.1. Redes Neuronales Artificiales (RNA)

En el ámbito de la inteligencia artificial, existen algoritmos como las redes neuronales artificiales. Estas redes son entrenadas a través de diversas entradas que son adquiridas mediante impulsos internos o externos en un sistema. Durante el entre-

namiento, estas entradas se multiplican por pesos que se asignan de forma aleatoria. Las redes neuronales artificiales presentan dos tipos de entrenamiento: durante la etapa de aprendizaje, la red es entrenada para que los pesos sinápticos se ajusten adecuadamente a la red. La segunda etapa es la de ejecución, donde la red se vuelve operativa y toma un valor de funcionamiento real [3]. La característica principal de las redes neuronales artificiales es su estructura computacional distribuida, la cual imita la organización del sistema nervioso humano [28], [16].

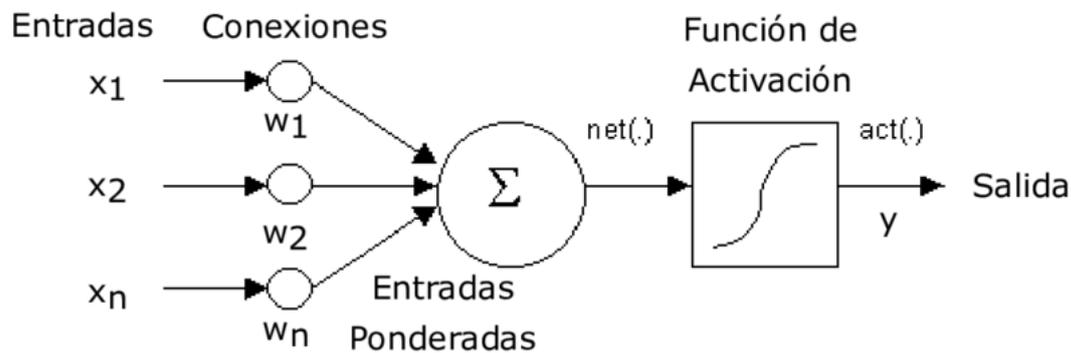


Figura 3.13: Estructura básica de una RNA

### Algoritmo Backpropagation

El backpropagation es un algoritmo de aprendizaje supervisado utilizado para entrenar redes pre-entrenadas. Este algoritmo consta de dos fases. Primero, se proporciona un patrón de entrada a la red y se obtiene una salida a través de las capas sucesivas. Después, se compara la salida obtenida con la deseada y se calcula una señal de error para cada salida. Estas señales de error se propagan hacia atrás, desde la capa de salida hacia las neuronas de la capa oculta que contribuyen directamente a la salida. Cada neurona de la capa oculta solo recibe una fracción de la señal total de error, en función de su contribución relativa a la salida original. Este procedimiento se repite en cada capa hasta que todas las neuronas de la red hayan recibido una señal de error que indique su contribución al error total. [2], [65].

A grandes rasgos, para utilizar el algoritmo de Backpropagation, se debe seguir el siguiente proceso:

1. Para completar el proceso, se debe calcular la salida de la red  $\sigma^2$  usando uno de los conjuntos proporcionados para las pruebas  $x$ .
2. A continuación, se compara la salida  $t$  y se determina el error mediante la

siguiente fórmula:

$$E(x; W^{(1)}, W^{(2)}) = \frac{1}{2} \sum_{i=1}^m (t_i - o_i^2)^2 \quad (3.54)$$

3. Para continuar con el proceso, es necesario calcular las derivadas parciales del error en función de los pesos  $W^{(2)}$  que conectan la última capa oculta con la salida correspondiente.
4. Además, se deben calcular las derivadas parciales del error respecto a los pesos  $W^{(1)}$  que conectan la capa de entrada con la capa oculta.
5. Luego, se ajustan los pesos de cada neurona para minimizar el error. Este ajuste de pesos se repite varias veces hasta que se logre reducir el error a un nivel deseado, a partir de la siguiente ecuación:

$$O = X_j * F(net_j) \quad (3.55)$$

Donde:

- $\sigma^2 = (o_1, o_2, \dots, o_m)^T$ : Define a la capa de salida, también conocido como el vector que acerca el valor deseado  $t = (t_1, t_2, \dots, t_m)^T$ .
- $W^{(1)}$ : define una matriz de tamaño  $(n + 1) \times k$ , sus valores  $W_{i,j}^{(2)}$  representan los pesos de la conexión ente las unidades  $\bar{o}^{(0)}$  y  $o_j^{(1)}$ .
- $W^{(2)}$ : define una matriz de tamaño  $(k + 1) \times m$ , sus valores  $W_{i,j}^{(2)}$  representan los pesos de la conexión ente las unidades  $\bar{o}^{(1)}$  y  $o_j^{(2)}$ .

### 3.6.2. Árboles de decisión

Un árbol de decisión es una herramienta que se utiliza para representar gráficamente diferentes escenarios y posibles resultados que pueden surgir a partir de una decisión tomada. A través de algoritmos, estos árboles permiten la toma de decisiones óptimas, basándose en probabilidades y considerando diferentes eventos posibles. Además, son una forma efectiva de visualizar y organizar el trabajo necesario para resolver un problema [30]. Asimismo, los árboles de decisión sirven de guías jerárquicas con múltiples vías, donde las características para evaluar determinado evento, dependen de las opciones que responden a un problema de decisión.

### Random Forest

Se trata de un método de aprendizaje automático que permite realizar diversas tareas de regresión, así como, de clasificación. Por otra parte, Random Forest trabaja con valores perdidos, valores atípicos y exploración de datos. Este método de aprendizaje emplea conjuntos, donde cada modelo que alberga un grupo, es combinado para formar un nuevo modelo con características más completas [61], [49].

Random Forest es una técnica que combina varios árboles de decisión para analizar diferentes porciones de los datos. Es importante destacar que cada árbol no tiene acceso completo a los datos de entrenamiento, sino que se entrena con diferentes muestras de datos. De esta manera, cada árbol aborda el mismo problema de una manera ligeramente diferente. Al combinar los resultados de los árboles individuales, los errores se compensan mutuamente, lo que resulta en una predicción más precisa y óptima [69].

El procedimiento por el que cada árbol es construido se describe a continuación:

1. Supongamos que hay  $N$  casos de prueba y  $M$  variables en el clasificador.
2. En cada nodo, solo se considerarán  $m$  variables de entrada para tomar una decisión, y  $m$  debe ser significativamente menor que  $M$ .
3. Seleccionar un conjunto de entrenamiento específico para este árbol y utilizar los demás casos de prueba para evaluar el rendimiento.
4. En cada nodo del árbol, se elegirán  $m$  variables de forma aleatoria para tomar una decisión. Luego, se calculará la mejor forma de dividir el conjunto de entrenamiento según las  $m$  variables seleccionadas.

#### 3.6.3. Algoritmos Bayesianos

Los algoritmos bayesianos forman parte de los clasificadores de Machine Learning (ML), estos modelos son definidos formalmente como grafos dirigidos acíclicos, donde sus nodos representan variables aleatorias, las aristas definen aquellas dependencias bajo ciertas condiciones, los nodos que no cuentan con alguna conexión representan las variables que son condicionalmente independientes de las demás. A cada nodo se le asigna una función de probabilidad, la cual se encarga de tomar un conjunto de entrada y devolver la probabilidad de la variable que representa al nodo en cuestión [54].

A partir de ciertos datos adquiridos, por ejemplo,  $x$ , y el parámetro  $\theta$ , un análisis

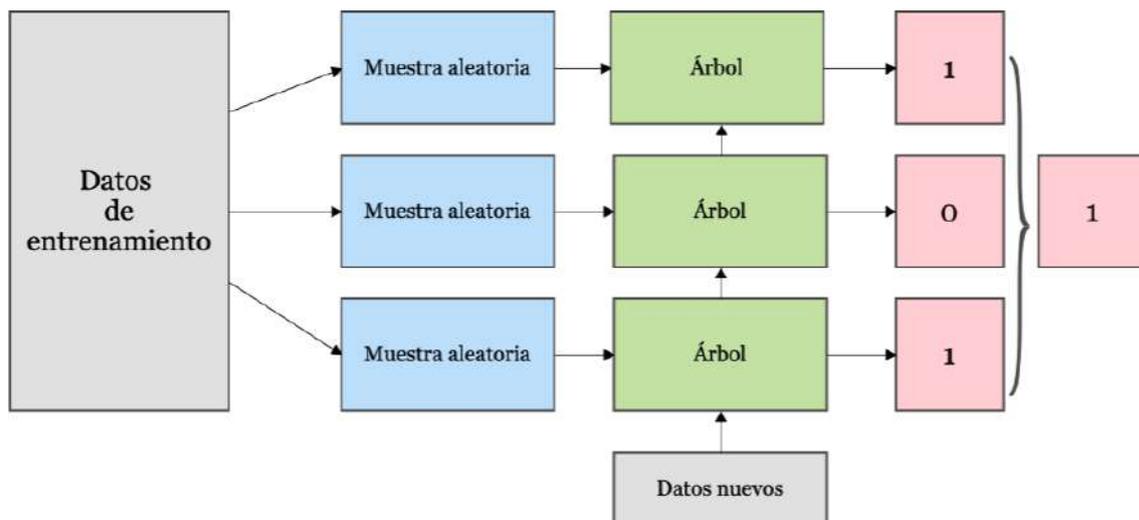


Figura 3.14: Funcionamiento general del algoritmo Random Forest

sencillo de una red bayesiana inicia con una probabilidad previa  $p(\theta)$  y una probabilidad  $p(x|\theta)$  para calcular una probabilidad final  $p(\theta|x) \propto p(x|\theta)p(\theta)$ . Este análisis simple de un modelo jerárquico bayesiano se repite determinadas veces, hasta que los parámetros previos no dependan de otros parámetros que no sean mencionados [23], [59].

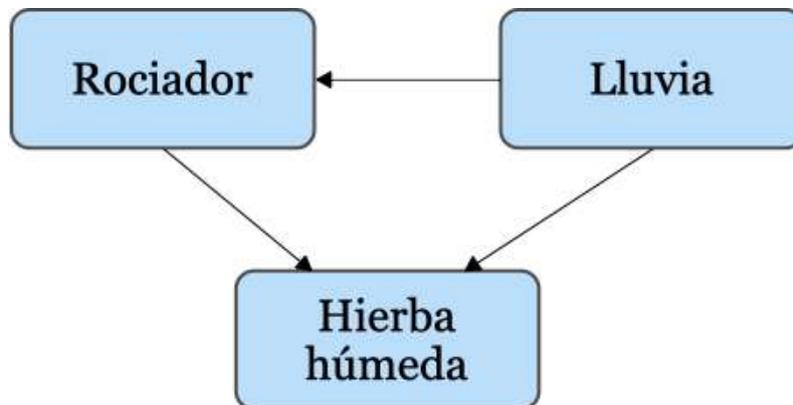


Figura 3.15: Ejemplo sencillo de red bayesiana. Modelo que analiza la influencia de la lluvia en caso de que el rociador esté activado, y la influencia conjunta de la lluvia y el rociador en caso de que la hierba esté húmeda.

### Naive Bayes

El clasificador Naive Bayes es un clasificador probabilístico que se refiere al Teorema de Bayes. Este algoritmo se basa en probabilidades condicionales con suposiciones independientes entre sus características [62]. El clasificador Naive Bayes asume etiquetas de clase categóricas y categoriza los datos según el conjunto de entrenamiento

y los valores en los datos de prueba. Naive Bayes encuentra aplicaciones en filtrado de spam, clasificación de texto, sistema de recomendación híbrido, y, además, se encarga de categorizar la polaridad de la oración como positiva o negativa [4]. El clasificador bayesiano matemáticamente se define a continuación:

$$P(X|E_1, E_2, \dots, E_n) = \frac{P(E_1, E_2, \dots, E_n|X)(P(X))}{P(E_1, E_2, \dots, E_n)} \quad (3.56)$$

Donde:

- $X$ : Es la probabilidad de un evento.
- $E$ : Es la certeza asignada.
- $P(E_1, E_2, \dots, E_n|X)$ : Representa la probabilidad.
- $P(X)$ : Previo.
- $E_1, E_2, \dots, E_n$ : Normalización constante.

#### 3.6.4. Máquinas de Soporte Vectorial (SVM)

Las máquinas de soporte vectorial son un método de clasificación inicialmente concebido para solucionar problemas binarios. Este método se enfoca en encontrar un hiperplano óptimo que maximice el margen de separación entre las dos clases, lo que resulta en una mejor clasificación [52], [18], [14], [15].

Las SVM representan fronteras para definir decisiones relacionadas a pequeños subconjuntos, asimismo, analizan la capacidad de clasificación entre clases, a esto se le conoce como vectores de soporte [10], [47], [48], [17].

Para realizar el cálculo utilizando un clasificador multiclase, se utiliza la configuración cuadrática siguiente:

$$w_r \in H, \epsilon^r \in R^m, b_r \in \frac{1}{2} \sum_{r=1}^M \|w_r\|^2 + \frac{c}{m} \sum_{i=1}^m \sum_{r \neq y_1} \epsilon_j^r \quad (3.57)$$

Sujeto a:

$$w_r \in H, \epsilon^r \in R^m, b_r \in \frac{1}{2} \sum_{r=1}^M \|w_r\|^2 + \frac{c}{m} \sum_{i=1}^m \sum_{r \neq y_1} \epsilon_j^r \quad (3.58)$$

### 3.6.5. KNN (K-Nearest Neighbors)

El algoritmo de K-Nearest Neighbors (KNN) es un método de aprendizaje automático supervisado utilizado para clasificar puntos de datos o predecir valores basados en la similitud con los datos de entrenamiento existentes. En esencia, KNN busca la cercanía en el espacio de características entre los puntos de datos.

En la fase de entrenamiento, KNN simplemente almacena los datos de entrenamiento sin realizar ningún procesamiento adicional. Estos datos de entrenamiento consisten en características (atributos) y las etiquetas correspondientes para la clasificación, o los valores objetivo para la regresión [36].

Cuando se realiza una clasificación o predicción con KNN, se busca un número definido de vecinos más cercanos en función de una medida de distancia, como la distancia euclidiana. Este número, denotado como "K", es un parámetro ajustable que se elige antes de aplicar el algoritmo.

Una vez que se encuentran los K vecinos más cercanos, en el caso de clasificación, se asigna la etiqueta más común entre los vecinos al nuevo punto de datos. En la regresión, se realiza un promedio de los valores objetivo de los K vecinos más cercanos para obtener la predicción final [58].

Es importante tener en cuenta que KNN asume que los puntos de datos cercanos en el espacio de características tienen etiquetas o valores objetivo similares. Por lo tanto, la elección adecuada de la medida de distancia y el valor de K es fundamental para el éxito del algoritmo.

### 3.6.6. Búsqueda de Malla (Grid search)

El algoritmo de búsqueda de cuadrícula es un método de optimización de parámetros utilizado comúnmente en aprendizaje automático y en otras áreas de la ciencia de datos.

El objetivo del algoritmo de búsqueda de cuadrícula es encontrar la mejor combinación de valores de los parámetros de un modelo de aprendizaje automático. En lugar de probar valores aleatorios para los parámetros, el algoritmo de búsqueda de cuadrícula define una cuadrícula de valores para cada uno de los parámetros y evalúa el rendimiento del modelo en cada combinación posible de valores de parámetros [38], [64].

A continuación, se describe cómo funciona el algoritmo de búsqueda de cuadrícula en cuatro pasos:

1. Definir la cuadrícula de parámetros: Para cada parámetro del modelo, se define una cuadrícula de valores para probar. Por ejemplo, si el modelo tiene dos parámetros, se puede definir una cuadrícula de 10 valores para cada parámetro, lo que resulta en 100 combinaciones de valores de parámetros a probar.
2. Entrenar y evaluar el modelo: Para cada combinación de valores de parámetros, se entrena y se evalúa el modelo utilizando un conjunto de datos de entrenamiento y un conjunto de datos de prueba.
3. Seleccionar el mejor modelo: Se selecciona la combinación de valores de parámetros que producen el mejor rendimiento del modelo en el conjunto de datos de prueba.
4. Validar el modelo: Se valida el modelo seleccionado en un conjunto de datos independiente para asegurarse de que el modelo sea generalizable a nuevos datos.

En resumen, el algoritmo de búsqueda de cuadrícula es una técnica de optimización de parámetros que se utiliza para encontrar la mejor combinación de valores de los parámetros de un modelo de aprendizaje automático. El algoritmo funciona evaluando el rendimiento del modelo en cada combinación posible de valores de parámetros en una cuadrícula predefinida y seleccionando la combinación de valores de parámetros que produzcan el mejor rendimiento del modelo en el conjunto de datos de prueba.



## METODOLOGÍA

En el presente capítulo, se describe el proceso metodológico empleado para optimizar la clasificación de los grados de severidad de RD y glaucoma.

Es importante mencionar, que la metodología propuesta en este trabajo de tesis está dividida en dos etapas, segmentación e identificación, para ambas fases, se utilizan componentes primordiales que pertenecen a un sistema de visión, el cual, permite interpretar y comprender la información visual capturada por una cámara o sensor de imagen. Estos sistemas utilizan técnicas de procesamiento de imágenes y aprendizaje automático para extraer características, reconocer patrones y realizar tareas de análisis visual. En la figura 4.1 se muestra el proceso de un sistema de visión por computadora:

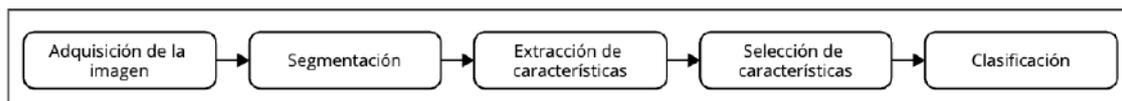


Figura 4.1: Componentes de un sistema de visión.

### 4.1. Conjunto de datos

La adquisición de imágenes es un paso crucial en la realización de una correcta clasificación de datos. En el contexto de un proyecto, se puede utilizar un conjunto de datos que contenga imágenes con una región local amplia para su análisis. En el caso específico de esta tesis, el primer conjunto de datos con el que se trabajó consta de cuatro clases diferentes, con un total de 4,217 imágenes, en la figura 4.2 se muestra la organización del conjunto de datos: la fila a) representa la clase de Cataratas, la fila b) corresponde a la clase de Retinopatía diabética, la fila c) muestra la clase de Glaucoma, y, por último, la fila d) representa la clase de Ojo sano. Cada una de estas clases tiene características distintas que permiten su diferenciación, lo que hace que sea esencial contar con un conjunto de datos amplio y bien definido para poder llevar a cabo una clasificación precisa y efectiva.

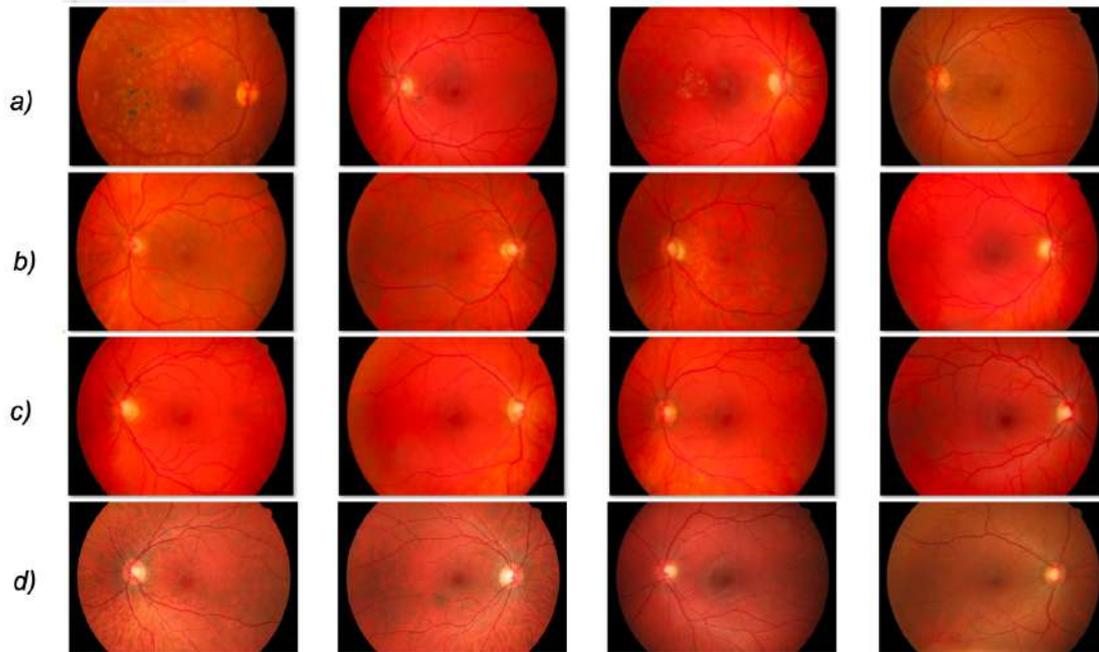


Figura 4.2: **a)** Cataratas, **b)** Retinopatía diabética, **c)** Glaucoma, **d)** Ojo sano. Clases del conjunto de datos utilizado.

El segundo conjunto de datos con el que se trabajó en esta investigación se obtuvo del repositorio Kaggle, y contiene un total de 35,122 imágenes de las cuales 25,806 corresponden a clase 0 (RD Ausente), 2,443 corresponden a la clase 1 (RD Leve), 5,292 pertenecen a la clase 2 (RD Moderada), 873 pertenecen a la clase 3 (RD Severa) y, por último, 708 corresponden a la clase 4 (RD Proliferativa), esta organización del conjunto de datos es bastante extensa, ya que, contiene cada uno de los grados de severidad para el caso de RD.

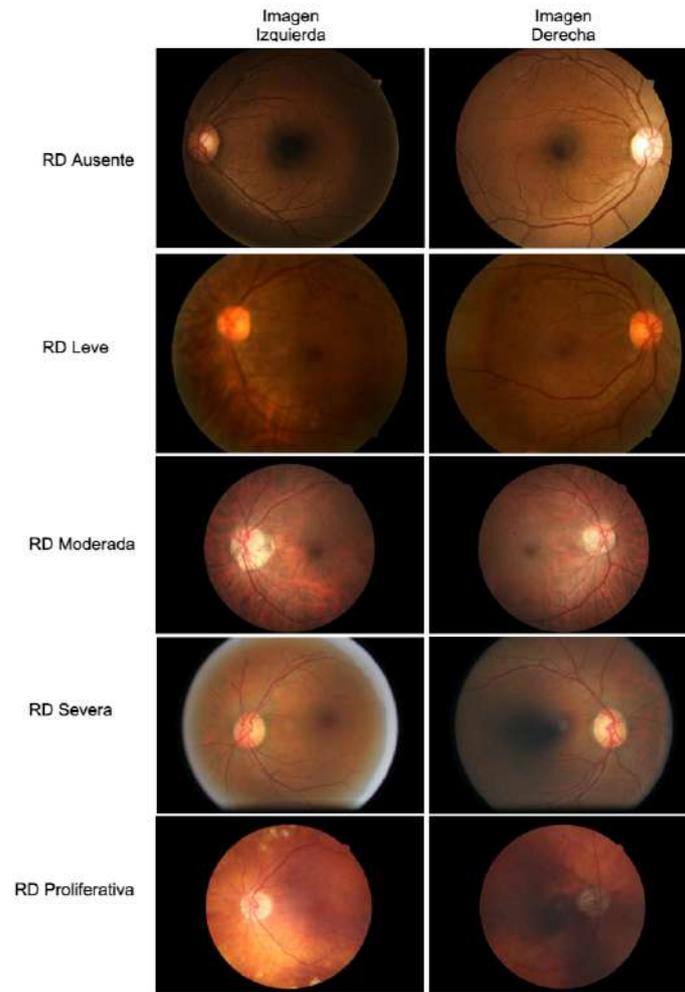


Figura 4.3: Conjunto de datos de RD extraído de Kaggle.

En el caso del tercer conjunto de datos con el que se trabajó en este trabajo de investigación, se obtuvo del repositorio de EyePACS y, corresponde a imágenes de Glaucoma, su totalidad de imágenes consta de 1,020 de las cuales 710 corresponden a la clase 0 (Glaucoma primario de ángulo abierto) y 310 corresponden a la clase 1 (Glaucoma de ángulo cerrado). Este conjunto de imágenes únicamente incluye dos de los cuatro grados de severidad de glaucoma, sin embargo, la calidad de las imágenes resulta ser bastante buena.

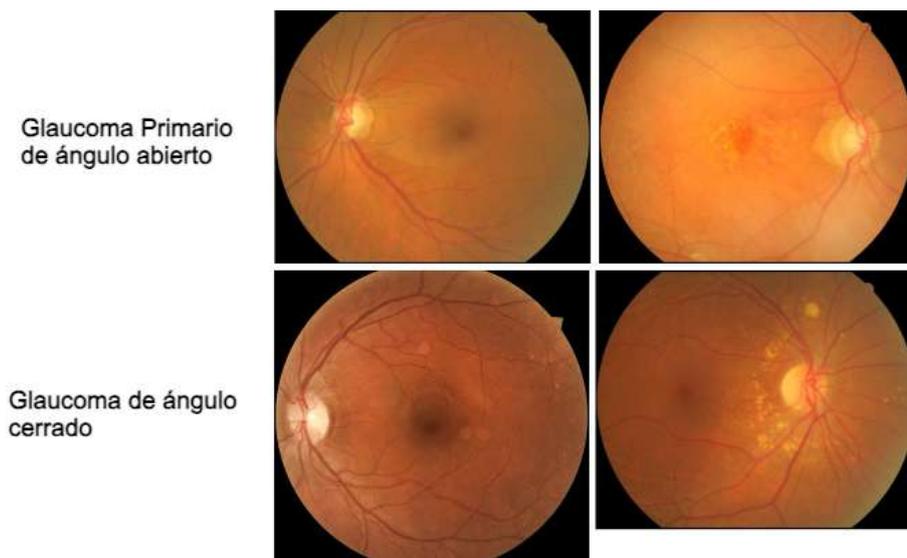


Figura 4.4: Conjunto de datos de Glaucoma extraído de EyePACS.

## 4.2. Fase de Segmentación

En esta fase se tiene la intención de reconocer los vasos sanguíneos de la retina, este proceso se realiza a partir de la obtención del banco de datos, trabajando y procesando cada una de las imágenes brindadas al sistema de visión.

En esta primera fase, se emplean diferentes técnicas y parámetros para lograr una correcta segmentación, por lo que ha sido mencionado en el capítulo 3, al segmentar una imagen, se pueden lograr diferentes niveles de granularidad, desde una segmentación a nivel de píxeles hasta una segmentación a nivel de objetos completos.

En la figura 4.5 se muestra el flujo general de la etapa de Segmentación:

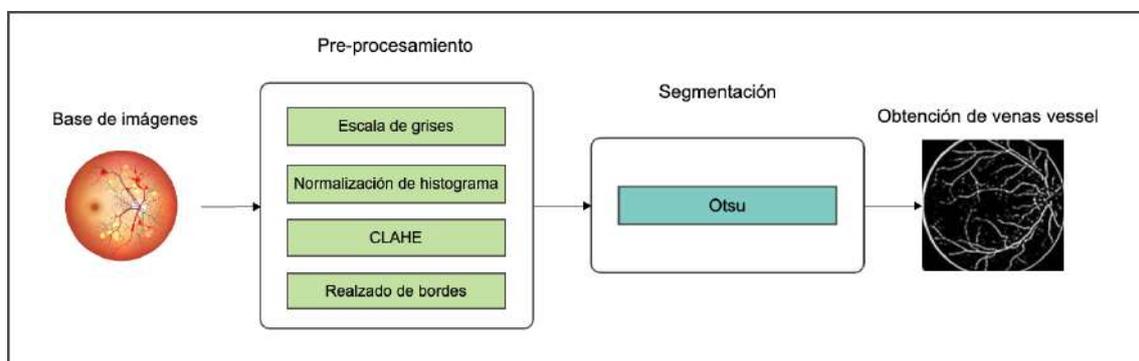


Figura 4.5: Proceso general de la fase de Segmentación para obtener los vasos sanguíneos de la retina.

### 4.2.1. Pre-procesamiento

Una vez que se obtienen las imágenes en la etapa de adquisición de la imagen, se realizan diversas operaciones para mejorar la calidad y el contenido de las imágenes, con el fin de hacerlas más adecuadas para su posterior procesamiento y análisis. Los algoritmos empleados en este proceso son los siguientes:

1. Algoritmo para convertir una imagen RGB a escala de grises: Esta técnica se utiliza para reducir la información de la imagen, lo que facilita su análisis. Para los propósitos de esta investigación, se utiliza el método de ponderación para realizar la conversión de una imagen RGB a escala de grises, es importante mencionar que se realizó un diminuto ajuste en los decimales de la ecuación original planteada en la sección 3.1.1, por lo que, la expresión resulta de la siguiente manera:

$$Y = (R * 0,299) + (G * 0,587) + (B * 0,114) \quad (4.1)$$

2. Normalización del histograma: En este proceso se distribuyen los valores de los píxeles de las imágenes para mejorar el contraste y resaltar sus principales características. Cabe mencionar, que se trabaja con una normalización estándar, por motivos de que los parámetros calculados se encuentran en distintos rangos, por tal motivo, es necesario realizar una transformación de los datos.

Lo descrito anteriormente, se representa por la siguiente ecuación:

$$z = \frac{(x - \mu)}{(\sigma)} \quad (4.2)$$

- $x$ : Corresponde al valor de una variable.
- $\mu$ : Representa la media de los valores asignados a la variable.
- $\sigma$ : Es la desviación estándar de los valores asignados a la variable.
- $z$ : Es el valor normalizado de la variable.

La normalización estándar tiene como objetivo transformar los valores de una variable en una escala común, eliminando las diferencias en la escala y la unidad de medida entre diferentes variables, lo que permite compararlas de manera más adecuada.

3. Resaltar el contraste de la imagen: Se corrige la iluminación y se mejora la visibilidad de los detalles en la imagen, por medio del algoritmo CLAHE. El

proceso empleado para aplicar esta técnica se describe a continuación:

1. El primer paso es detectar si la imagen es en escala de grises o a color. Para este caso, como se están trabajando con imágenes a color, entonces se mejora la imagen en el canal  $I$  del espacio de color  $HSI$ . Posteriormente, cuando se hace el cambio a escala de grises, se aplica la técnica CLAHE directamente sobre la imagen.
2. Los parámetros ajustables de CLAHE son el número de regiones contextuales  $R_x \times R_y$  y el valor de clip, que está en el rango de 0 a 1.
4. Realzado de bordes: Mejorar la visibilidad de los bordes y los detalles relevantes de una imagen.

En el caso de esta investigación, se utilizará la fórmula gaussiana, la cual, es ampliamente utilizada debido a su capacidad para suavizar la imagen y resaltar cambios graduales en la intensidad, lo que ayuda a identificar y analizar los bordes de manera más efectiva.

La fórmula de Gauss se representa de la siguiente manera:

$$g(x, y) = \frac{1}{(2\pi\sigma^2)} * e^{\frac{-(x^2+y^2)}{(2\sigma^2)}} \quad (4.3)$$

Donde:

$f(x, y)$ : Representa el valor de la función gaussiana en las coordenadas  $(x, y)$ .

Asimismo, la función gaussiana está determinada por dos parámetros:

$\sigma$ : Es la desviación estándar de la distribución gaussiana. Este parámetro controla la amplitud o la cantidad de difuminado aplicado a la imagen. Un valor mayor de  $\sigma$  produce un difuminado más fuerte, mientras que un valor menor produce un difuminado más suave.

$2\pi$ : Es una constante que se utiliza en la fórmula y se relaciona con la forma de la curva de la distribución gaussiana.

#### 4.2.2. Segmentación

En este proceso se divide en diferentes regiones la imagen, estas secciones tienen características similares, lo que permite resaltar áreas específicas. El algoritmo de Otsu es una técnica de umbralización automática que se utiliza para identificar el umbral óptimo en una imagen en escala de grises. Su objetivo principal es encontrar

el umbral que minimice la varianza dentro de cada clase, lo que significa que busca maximizar la homogeneidad dentro de cada región de la imagen y, al mismo tiempo, maximizar la diferencia entre estas regiones.

Una vez obtenido el umbral óptimo, se aplican operadores morfológicos, como la erosión y dilatación, para refinar la segmentación de la imagen. La erosión disminuye el tamaño de los objetos y elimina detalles pequeños, mientras que la dilatación aumenta el tamaño de los objetos y rellena huecos.

1. Erosión: La operación de erosión es utilizada en esta sección para eliminar los píxeles de las regiones de la imagen que tienen una intensidad menor que el valor de los píxeles vecinos.

Para aplicar una fórmula de erosión a una imagen, se utiliza un operador de máscara que se desplaza sobre la imagen y compara los valores de los píxeles en la región de la máscara.

$$g(x, y) = f(x, y) \cap M \quad (4.4)$$

Donde:

$g(x, y)$ : Corresponde a la imagen de salida.

$f(x, y)$ : Corresponde a la imagen de entrada.

Los parámetros de la fórmula de erosión son la forma y el tamaño de la máscara utilizada para comparar los valores de los píxeles en la imagen original.

2. Dilatación: Este operador morfológico será aplicado a imágenes en escala de grises con el objetivo de aumentar la región de los objetos en la imagen. La fórmula para la dilatación consiste en deslizar una ventana o máscara sobre la imagen y tomar el valor máximo de los píxeles dentro de la ventana.

El píxel en el centro de la ventana se reemplaza por este valor máximo, lo que hace que los objetos en la imagen se expandan y aumenten su tamaño.

Los parámetros de la dilatación son la forma y tamaño de la máscara. La forma de la máscara puede ser un cuadrado, un círculo, una cruz, entre otras formas.

El tamaño de la máscara se refiere al número de píxeles que abarca la ventana.

Al combinar el algoritmo de Otsu con operadores morfológicos, se puede lograr una segmentación más precisa de las regiones de interés en las imágenes de la retina. La umbralización de Otsu permite separar automáticamente las regiones basándose

en la información de intensidad, mientras que los operadores morfológicos ayudan a refinar la segmentación y eliminar artefactos no deseados.

### 4.3. Fase de Identificación (Clasificación)

Por otra parte, en la fase de Identificación se pretende realizar una clasificación adecuada relacionada al área general del ojo, es decir, clasificar los vasos sanguíneos con su respectivo padecimiento (RD o glaucoma) y el grado de severidad.

Es importante mencionar que, en esta fase, también se incluyen las mismas tareas que en la fase de Segmentación, sin embargo, ésta posee un proceso más, denominado Extracción de características:

En la figura 4.6 se ilustra el funcionamiento general en la fase de Identificación:

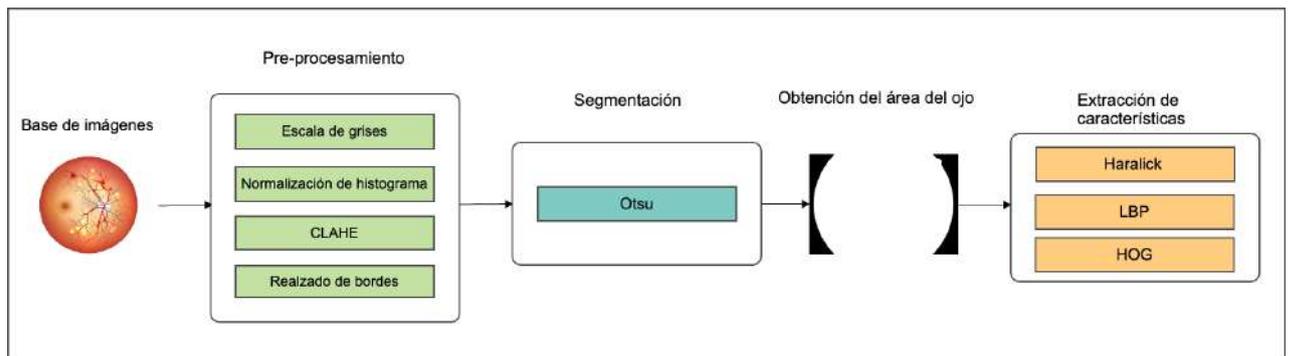


Figura 4.6: Proceso general de la fase de Identificación para extraer características de vasos sanguíneos de la retina.

#### 4.3.1. Extracción de características

Esta tarea permite identificar y extraer patrones, formas, texturas y otros aspectos relevantes de las imágenes del banco de datos, con el fin de representarlas de manera más compacta y resumida. Asimismo, los algoritmos que serán en esta tarea son:

1. Descriptores de Haralick: Este descriptor de textura, será utilizado para la extracción de características de imágenes de la retina humana, con el fin de identificar los padecimientos mencionados anteriormente.

Este descriptor se basa en la matriz de co-ocurrencia de niveles de gris, la cual se construye a partir de la imagen original. Esta matriz describe la distribución espacial de los pares de niveles de gris en la imagen.

A continuación, se muestra el diseño de la matriz de co-ocurrencia de los niveles de grises en una imagen de la retina humana:

2. Descriptores HOG: Este descriptor se basa en la idea de que la apariencia local y la forma de un objeto pueden ser descritas por la distribución de las direcciones del gradiente o las orientaciones de los bordes.

Para utilizar el descriptor HOG en esta investigación, se deben extraer, primero, las características de la imagen segmentada. Para ello, se crearán secciones la imagen, formando celdas pequeñas, y se calculan los histogramas de gradientes orientados en cada celda. Luego, se concatenan los histogramas para formar un vector de características.

3. Descriptores LBP: Para identificar los patrones en una imagen, se utilizará el descriptor LBP, pues se encargará de analizar las relaciones de contraste entre píxeles adyacentes en una imagen de la retina humana.

LBP tomará un píxel de la imagen para comparar su valor de intensidad con el de sus vecinos, generando una cadena binaria que representa la distribución local de contraste. Esta cadena binaria seerá utiliza como característica para identificar patrones en la imagen.

Los descriptores mencionados, permiten describir las propiedades texturales de las imágenes del banco de datos, con el fin de obtener las características similares de cada una de ellas y, posteriormente, clasificarlas.



## RESULTADOS EXPERIMENTALES

En este capítulo, se mencionan los experimentos realizados con los conjuntos de datos descritos en la metodología, asimismo, se explican los resultados que fueron obtenidos a partir de los métodos y herramientas utilizadas para el análisis de su desempeño.

Los experimentos generados en este proyecto parten de la propuesta de solución planteada en la sección 1.2, del capítulo 1.

Es importante mencionar que, se emplearon diferentes técnicas y métodos para la clasificación de los grados de severidad que corresponden a las características que presente cada imagen.

Como se mencionó en el capítulo 3, sección 3.6, los algoritmos empleados para realizar cada una de las pruebas van desde, Redes Neuronales Artificiales, Árboles de Decisión, SVM (Máquinas de Soporte Vectorial), Naive Bayes y K-NN (k-nearest neighbors o K-vecinos más cercanos), empleando además, métodos de validación como K-fold Cross Validation (Validación Cruzada).

Para cada método, es relevante mencionar que, los parámetros que involucran para su proceso de análisis, fueron modificados de acuerdo a las necesidades del trabajo de investigación, con la finalidad de obtener el mejor rendimiento posible.

### 5.1. Proceso para la extracción de características

Como se mencionó en el capítulo 4, sección 4.2, primero es necesario obtener los vasos sanguíneos de la retina, este resultado se obtiene mediante el proceso de segmentación.

Para los conjuntos de datos empleados en este trabajo, se utilizó el mismo procedimiento para la adquisición de las venas de vessel. Una vez que se detectan las características, se procede a su clasificación empleando los algoritmos de aprendizaje automático mencionados.

A continuación, se describen los parámetros utilizados para el proceso de la extracción de características.

### 5.1.1. Segmentación

Para detectar las características importantes de las imágenes se empleó el algoritmo Otsu, este método se utiliza a imágenes que ya fueron sometidas a un pre-procesamiento, mediante las técnicas mencionadas en la sección 4.2.1 de este documento. Por lo tanto, es importante mencionar que en este proceso se deben ajustar ciertos parámetros, con el fin de obtener las mejores regiones de la imagen.

Los parámetros que fueron ajustados para realizar el proceso de segmentación, son los siguientes:

1. Tamaño del kernel para el suavizado: El tamaño varía de entre (3, 3) o (5, 5). Estos tamaños de kernel suelen funcionar bien para suavizar la imagen y reducir el ruido antes de aplicar la segmentación.
2. Método de umbralización: Se utiliza el método `cv2.THRESH_BINARY` de la biblioteca OpenCV. Selecciona el método de umbralización adecuado para obtener una imagen binaria donde las venas estén resaltadas.
3. Postprocesamiento: Este proceso incluye un tamaño del kernel para operaciones morfológicas, para los fines de este trabajo, el tamaño está entre (3, 3) o (5, 5).
4. Tamaño del área para eliminación de objetos pequeños: Se utilizó un tamaño de 50 píxeles. Estos tamaños de kernel y umbrales de área pueden ser ajustados según el grosor de las venas y el ruido presente en la imagen.

El siguiente fragmento de código programado en python, empleando librerías de OpenCV, se encarga de aplicar algoritmos de pre-procesamiento para preparar las imágenes de ambos conjuntos de datos. Una de las ventajas de emplear librerías en python, es que, los métodos ya están programados, pero, para el caso de esta investigación, sí fue importante ajustar algunos parámetros.

### 5.1.2. Selección y extracción de características

Una vez que las imágenes de ambos conjuntos de datos ya han sido pre-procesadas, como se mencionó en la sección 5.1.1 de este documento; ahora deben seleccionarse y extraerse las características importantes de cada imagen, que posteriormente, serán clasificadas de acuerdo a sus rasgos.

A continuación, se describe el proceso general para extraer las características, tomando en cuenta, los parámetros necesarios para lograr la adquisición de los rasgos importantes:

```

import cv2
import numpy as np

def aplicar_escala_grises(imagen):
    imagen_gris = cv2.cvtColor(imagen, cv2.COLOR_BGR2GRAY)
    return imagen_gris

def aplicar_segmentacion(imagen_gris):
    # Aplicar algoritmo de segmentación (ejemplo: umbralización)
    _, imagen_segmentada = cv2.threshold(imagen_gris, 127, 255, cv2.THRESH_BINARY)
    return imagen_segmentada

def aplicar_clahe(imagen_gris):
    clahe = cv2.createCLAHE(clipLimit=2.0, tileGridSize=(8, 8))
    imagen_clahe = clahe.apply(imagen_gris)
    return imagen_clahe

# Cargar la imagen
imagen_original = cv2.imread('normales.png')

# Aplicar escala de grises
imagen_gris = aplicar_escala_grises(imagen_original)

# Aplicar segmentación
imagen_segmentada = aplicar_segmentacion(imagen_gris)

# Aplicar CLAHE
imagen_clahe = aplicar_clahe(imagen_gris)

# Guardar las imágenes por separado
cv2.imwrite('/Users/dalialuna/Downloads/imagen_gris1.jpg', imagen_gris)
cv2.imwrite('/Users/dalialuna/Downloads/imagen_segmentada1.jpg', imagen_segmentada)
cv2.imwrite('/Users/dalialuna/Downloads/imagen_clahe3.jpg', imagen_clahe)

```

Figura 5.1: Métodos en python para aplicar pre-procesamiento a las imágenes.

1. Cargar las imágenes preprocesadas: Se comienza por cargar las imágenes de la retina que ya han sido preprocesadas.
2. Extraer características locales: Se utilizaron métodos como Patrones binarios locales (LBP) y el Histograma de gradientes ordenados (HOG) para extraer características locales de las imágenes. Estos métodos permiten capturar información sobre la textura y las formas presentes en las imágenes de la retina.

El descriptor LBP codifica patrones locales de variaciones de intensidad, mientras que el HOG detecta gradientes y direcciones de cambio en la intensidad de los píxeles.

3. Calcular características globales: Además de las características locales, también se calculan las características globales, las cuales, se encargan de proporcionar una representación resumida de toda la imagen. Para el análisis de texturas de Haralick, se utilizan matrices de co-ocurrencia de niveles de gris para extraer características estadísticas, como la energía, la entropía y la homogeneidad.

4. Organizar las características: Una vez que se han extraído las características de cada imagen en los conjuntos de entrenamiento y prueba, se organizan en una estructura de datos adecuada, para las necesidades de esta investigación, las características fueron guardadas en una matriz.

En la figura 5.2, se muestra el método empleado para obtener las características importantes de cada conjunto de datos, también utilizando librerías de OpenCV en python.

```
# Calcular el descriptor LBP (Local Binary Patterns)
lbp_radius = 3
lbp_n_points = 8 * lbp_radius
lbp_histogram = np.histogram(local_binary_pattern(imagen, lbp_n_points, lbp_radius), bins=np.arange(0,
                                                    lbp_n_points + 3), range=(0, lbp_n_points + 2))[0]

# Calcular el descriptor Haralick
haralick_features = mahotas.features.haralick(imagen).mean(axis=0)

# Calcular el descriptor HOG (Histogram of Oriented Gradients)
hog_orientaciones = 9
hog_pixeles_por_celda = (8, 8)
hog_celdas_por_bloque = (2, 2)
hog_norma = 'L2-Hys'
hog_features = hog(imagen, orientations=hog_orientaciones, pixels_per_cell=hog_pixeles_por_celda,
                   cells_per_block=hog_celdas_por_bloque, block_norm=hog_norma, visualize=False, feature_vector=True)

# Mostrar los resultados
print('Características LBP:')
print(lbp_histogram)

print('Características Haralick:')
print(haralick_features)

print('Características HOG:')
print(hog_features)
```

Figura 5.2: Método en python para la obtención de las características del conjunto de datos, utilizando la librería de OpenCV.

## 5.2. Resultados de la clasificación

Como se ha mencionado en la sección 5.1 de este documento, se describe el proceso detallado y general para la obtención de las características, con las cuales se realizará la clasificación correspondiente. Es importante mencionar que, una vez que se han adquirido las características importantes, éstas fueron guardadas en un archivo que correspondiente al conjunto de datos utilizado, es importante mencionar, se obtuvo un total de 2,387 para cada uno, respetando su totalidad de patrones respectivamente.

En la validación cruzada con  $k = 5$ , en cada iteración,  $1/5$  (aproximadamente el 20 %) de los datos se utilizarán como conjunto de prueba, y los  $4/5$  (aproximadamente el 80 %) restantes se utilizarán como conjunto de entrenamiento. Esto significa que los conjuntos de entrenamiento y prueba resultaron de la siguiente manera:

- Para el primer conjunto de datos:

Conjunto de prueba:  $(1/5) * 4,217 = 843$  patrones.

Conjunto de entrenamiento:  $(4/5) * 4,127 = 3,374$  patrones.

- Para el segundo conjunto de datos (RD):

Conjunto de prueba:  $(1/5) * 35,122 = 7,025$  patrones.

Conjunto de entrenamiento:  $(4/5) * 35,122 = 28,097$  patrones.

- Para el tercer conjunto de datos (Glaucoma):

Conjunto de prueba:  $(1/5) * 1,020 = 204$  patrones.

Conjunto de entrenamiento:  $(4/5) * 1,020 = 816$  patrones.

El objetivo principal para la clasificación de esta investigación, es identificar aquellos atributos que sean más significativos para clasificar los datos en un conjunto de categorías específicas, utilizando diversas técnicas de selección de atributos y algoritmos de clasificación. El resultado de este trabajo podría ser útil para mejorar la eficiencia y precisión de los modelos de clasificación utilizados en diversos ámbitos.

Los experimentos se realizaron en Matlab, ya que suele ser una herramienta completa para realizar las pruebas necesarias, de acuerdo con los objetivos establecidos en este trabajo de tesis.

Matlab permite programar de manera fácil y rápida, gracias a que cuenta con un lenguaje de programación orientado a matrices y vectores, que simplifica la escritura de código. Es por esta razón, es que se emplearon diferentes técnicas de clasificación

para trabajar con las características extraídas, que en conjunto, formaron un nuevo banco de datos, el cual se ha mencionado anteriormente.

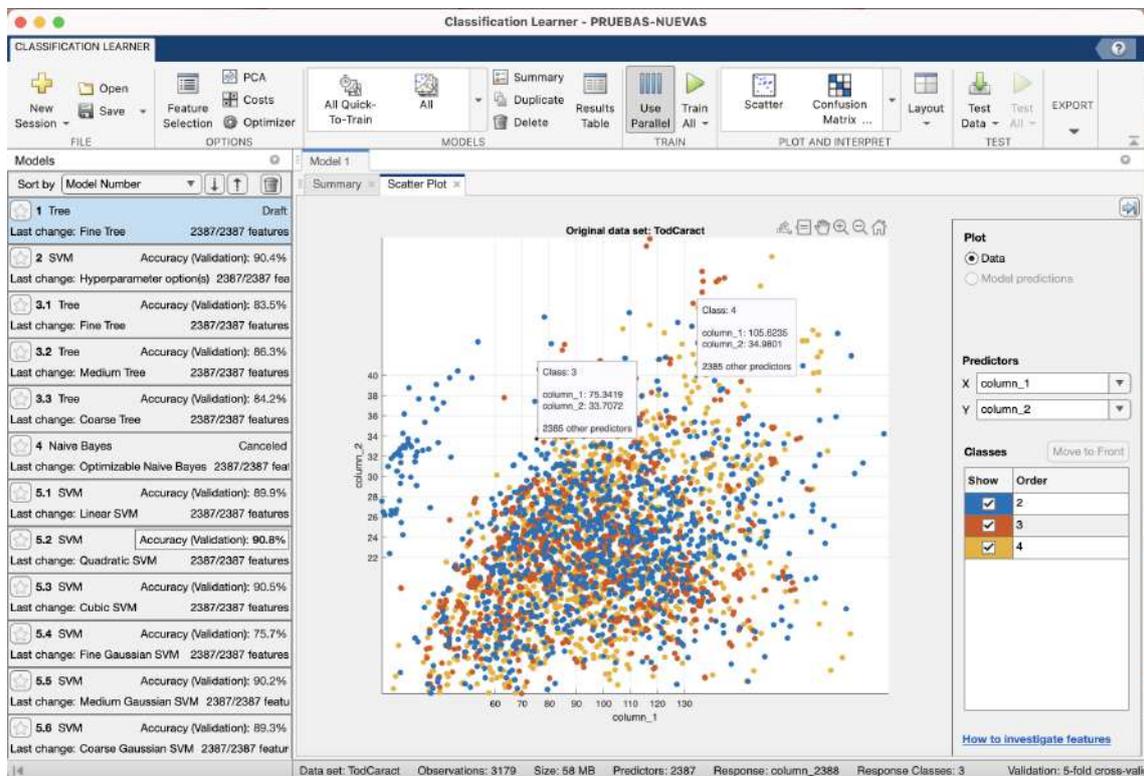


Figura 5.3: Interfaz de Matlab, mostrando un ejemplo de las pruebas con los diferentes algoritmos de aprendizaje automático

A continuación, se muestran los resultados obtenidos de acuerdo a las técnicas de clasificación mencionadas en la sección 5.

### 5.2.1. Clasificación utilizando Redes Neuronales Artificiales (RNA)

La clasificación realizada para el primer conjunto de datos CD 1 DRIVE, se realizó mediante los algoritmos de aprendizaje mencionados previamente, los resultados obtenidos se describen a continuación:

Se utilizaron diversas arquitecturas de Redes Neuronales en Matlab, y de todas ellas, la mayor precisión se obtuvo con una Red Neuronal Amplia, alcanzando una precisión del 88.7%. En la figura 5.4, se muestra la matriz de confusión con los patrones que fueron mejor clasificados.

**Red Neuronal Amplia**

	1	2	3	4
1	921		75	42
2	3	1085	2	8
3	79	4	786	138
4	45	7	121	901
	1	2	3	4

Clase verdadera

Clase predicha

Figura 5.4: Matriz de confusión para una Red Neuronal Amplia, mostrando los patrones mejor clasificados correspondientes a la clase 2.

Asimismo, en la figura 5.5, se visualiza una gráfica de ROC, que muestra la relación entre la tasa de verdaderos positivos y la tasa de falsos positivos, para el caso del modelo de la Red Neuronal Amplia:

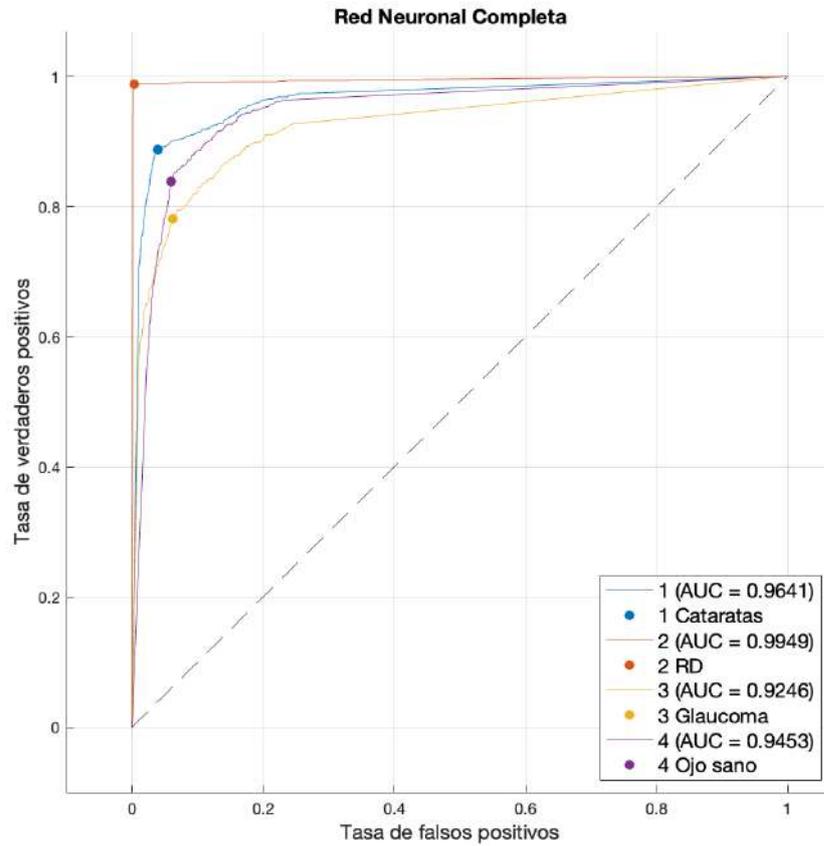


Figura 5.5: Gráfica ROC para una Red Neuronal Amplia, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad.

### 5.2.2. Clasificación utilizando Naive Bayes

En este modelo de clasificación, la precisión no resultó ser bastante satisfactoria, las diferentes arquitecturas o tipos en este algoritmo fallaron al momento de ejecutarse, resultando una precisión del 42.6 %, siendo el resultado más bajo.

En la figura 5.6, se muestra la matriz de confusión para el modelo Naive Bayes, indicando el total de patrones clasificados en cada clase.

**Naive Bayes**

Clase verdadera \ Clase predicha	1	2	3	4
1	217	717	45	59
2	8	1090		
3	43	775	112	77
4	24	470	205	375

Figura 5.6: Matriz de confusión para el algoritmo Naive Bayes, mostrando los patrones mejor clasificados correspondientes a cada clase.

La gráfica ROC, aplicada al modelo Naive Bayes, figura 5.8, permite resaltar la clasificación correcta para la tasa de verdaderos positivos y, para la tasa de falsos positivos.

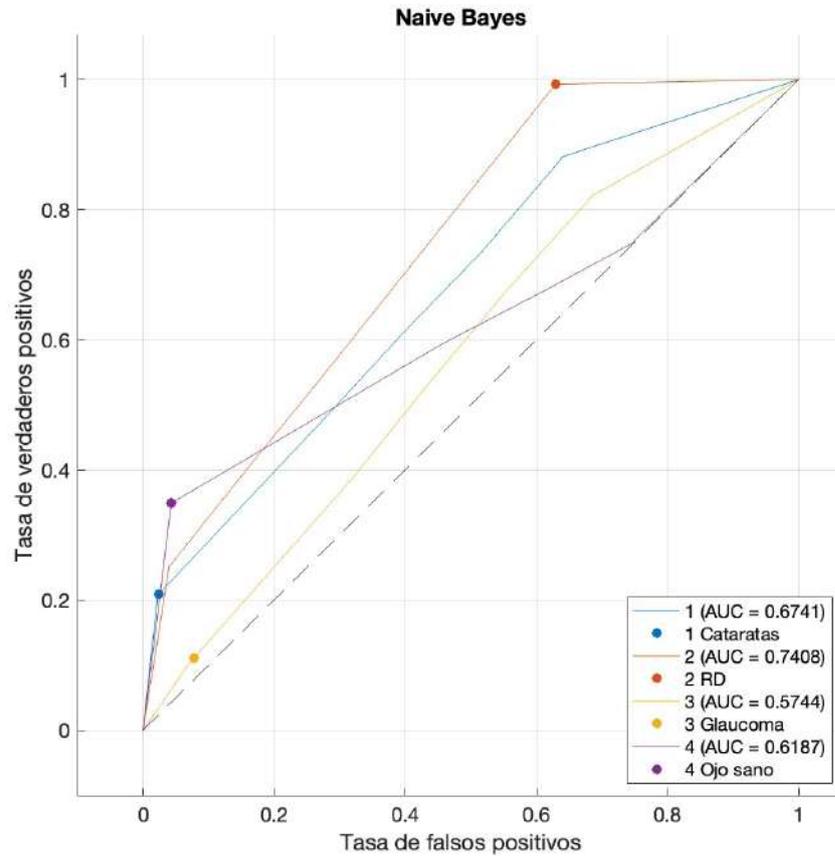


Figura 5.7: Gráfica ROC para el algoritmo Naive Bayes, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad.

### 5.2.3. Clasificación utilizando K-NN (K-n vecinos cercanos)

A pesar de que K-NN suele ser un algoritmo simple, en el caso de este trabajo de investigación, resultó ser bastante costoso computacionalmente, el motivo fue por el tamaño de los datos que ejecutó, ya que se debían calcular las distancias de los patrones o instancias. Sin embargo, el resultado obtenido no fue tan bajo, ya que se alcanzó una precisión del 84.1 %.

La matriz de confusión en la figura 5.8, representa las características mejor identificadas y asignadas de acuerdo a la clase, empleando K-NN.

**KNN**

Clase verdadera	1	798	5	139	96
	2	9	1065	9	15
	3	83	19	673	232
	4	24	18	26	1006
		1	2	3	4
		Clase predicha			

Figura 5.8: Matriz de confusión para el algoritmo K-NN, mostrando los patrones mejor clasificados correspondientes a cada clase.

Como se observa en la figura 5.9, la clase que mejor fue clasificada de acuerdo a las características, fue la clase 2, correspondiente a RD (Retinopatía Diabética).

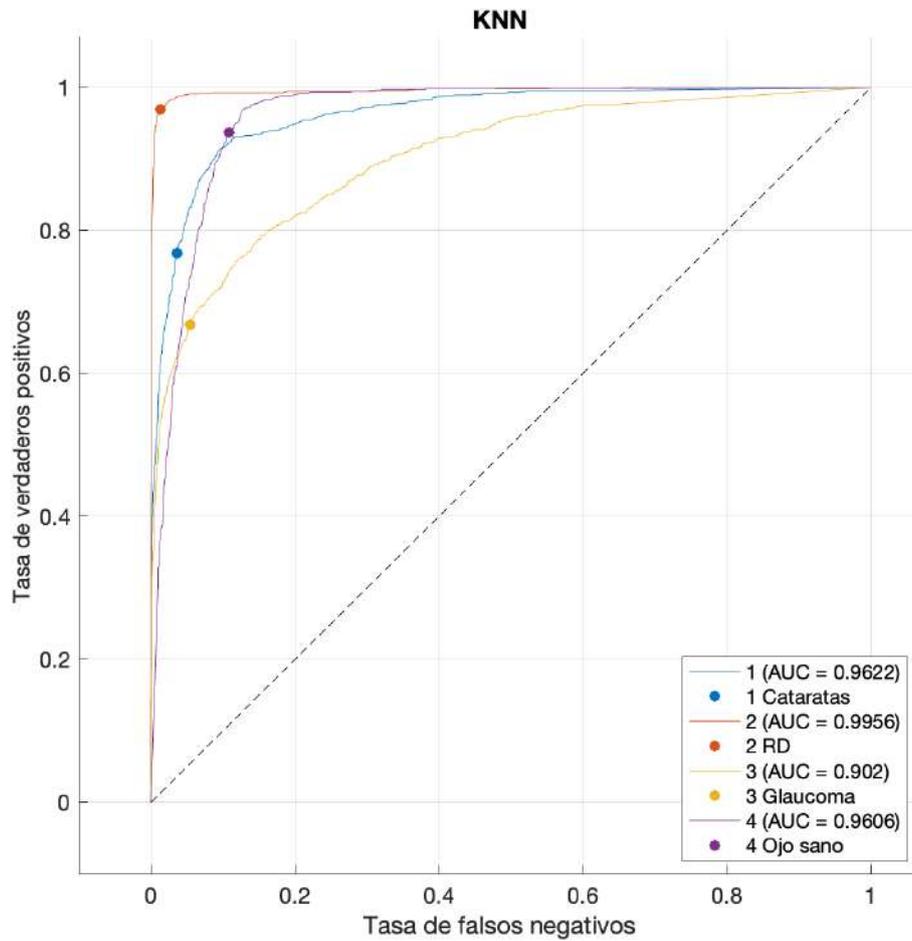


Figura 5.9: Gráfica ROC para el algoritmo K-NN, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad.

#### 5.2.4. Clasificación utilizando Árboles de decisión

Para este algoritmo, al igual que los anteriores, se emplearon cada una de las arquitecturas en Matlab para clasificar las características, el costo computacional no resultó ser bastante costoso, y el resultado de la precisión alcanzó un 82.9%, empleando el modelo Random Forest.

En la matriz de confusión para el algoritmo Random Forest, se muestra el total de patrones clasificados correctamente, con su respectiva clase.

**Árbol de decisión**

1	814	8	138	78
2	12	1032	20	34
3	128	16	677	186
4	41	26	120	887
	1	2	3	4

Clase verdadera

Clase predicha

Figura 5.10: Matriz de confusión para el algoritmo Random Forest, mostrando las características mejor clasificadas correspondientes a cada clase.

El resultado de clasificación, que corresponde a las características del conjunto de datos, se asigna también a la clase 2, es decir, a RD. En la gráfica ROC, de la figura 5.11, se puede observar esta distribución empleando el modelo Random Forest.

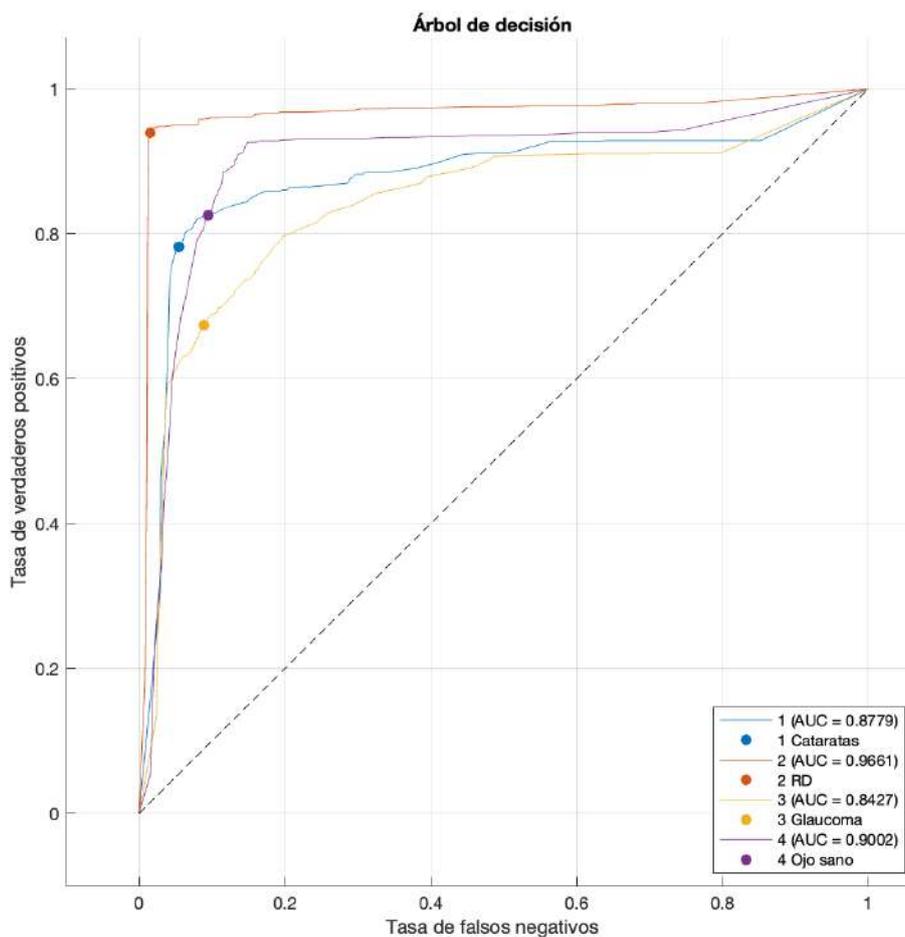


Figura 5.11: Matriz de confusión para el algoritmo Random Forest, mostrando las características mejor clasificadas correspondientes a cada clase.

### 5.2.5. Clasificación utilizando Máquinas de Soporte Vectorial (SVM)

Las Máquinas de Soporte Vectorial, al igual que los métodos anteriores, también tienen varias arquitecturas en el entorno de Matlab, lo que permite visualizar la ejecución y el resultado generado. Es algoritmo de aprendizaje automático, brindó una precisión del 89.9 %.

En la matriz de confusión, figura 5.12, se muestra una óptima clasificación de características en la clase de RD, grado 2, utilizando el modelo SVM cúbico.

Asimismo, la sensibilidad y especificidad de la SVM, plasman una distribución específica para la clase de RD, lo que significa que la clasificación para la padecimiento detectado aumenta en la tasa de verdaderos positivos, es decir, el mayor porcentaje de las características se clasifican correctamente.

True Class	1	2	3	4
1	929	2	61	46
2	4	1088	2	4
3	50	4	778	175
4	26	7	77	964

Figura 5.12: Matriz de confusión para el algoritmo SVM Cúbico, mostrando las características mejor clasificadas correspondientes a cada clase.

Por otra parte, es importante mencionar verificar cómo se clasificaron los otros dos conjuntos de datos, en esta ocasión, solo se describirán aquellos algoritmos que resultaron con un mejor precisión al momento de brindar un diagnóstico:

### 5.2.6. Clasificación para el conjunto de Retinopatía Diabética (RD)

Para la clasificación del conjunto de datos de RD, se emplearon los mismos algoritmos de aprendizaje automático (ML), esta tarea es importante para verificar la clasificación de los patrones en las 5 clases.

El algoritmo que mostró una mejor exactitud de clasificación fue una Red neuronal perceptrón multicapa, brindando un 92.1 %.

A continuación se muestra la matriz de confusión, mostrando los patrones clasificados en las diferentes clases del conjunto de datos con base a sus características:

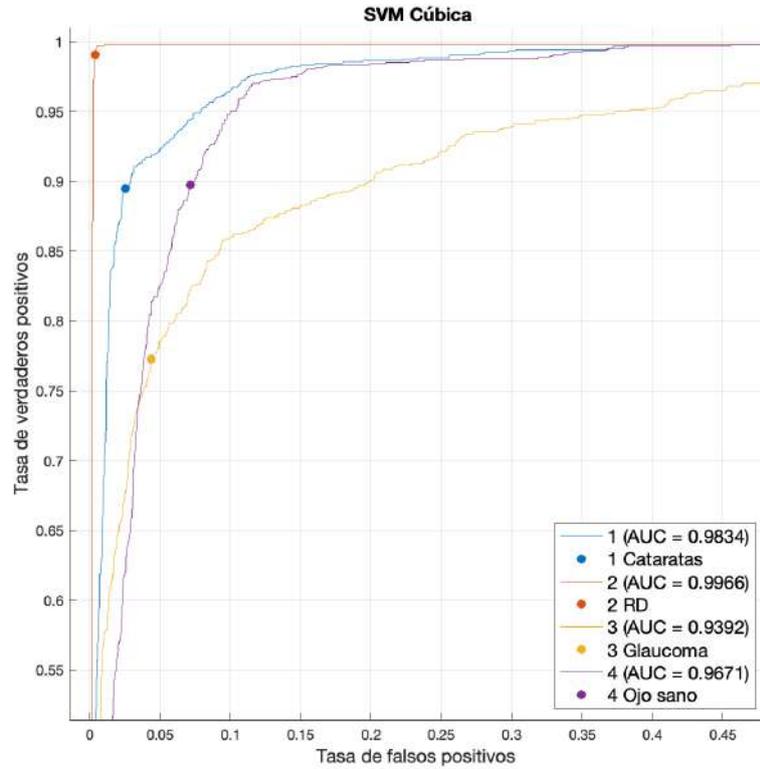


Figura 5.13: Gráfica ROC para el algoritmo SVM Cúbico, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad. Clasificando en la clase de RD.

Asimismo, la sensibilidad y especificidad de la RNA, plasman una distribución específica para la clase de RD Leve, lo que significa que la clasificación para el grado de severidad detectado de RD, aumenta en la tasa de verdaderos positivos, es decir, el mayor porcentaje de las características se clasifican correctamente.

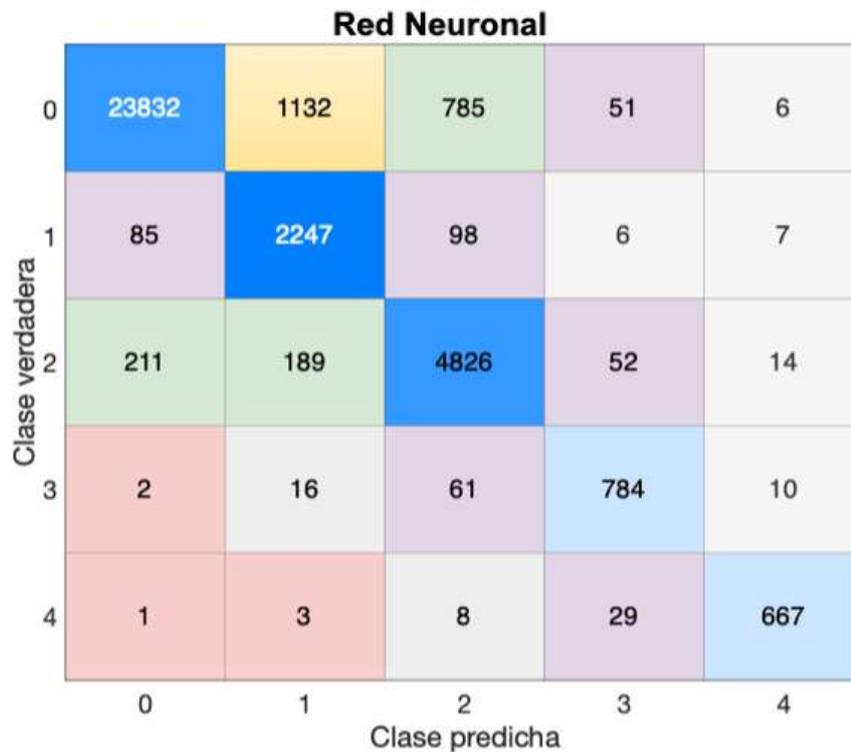


Figura 5.14: Matriz de confusión de RNA perceptrón multicapa. Clasificando en la clase 2 (RD Leve).

Técnica de clasificación	Precisión	Aciertos
NN	92.1 %	32,356
SVM	91.3 %	29,245
Bayes Naive	42.6 %	20,613
Árboles de decisión	76.2 %	26,784
KNN	83.2 %	29,245

Cuadro 5.1: Resultados experimentales en Matlab para el conjunto de datos de RD

Como se puede observar en la tabla 5.1, se muestran las precisiones correspondientes al conjunto de imágenes de RD, detallando las clasificaciones realizadas por los diferentes algoritmos de aprendizaje automático utilizado. Además, se muestra el total de patrones clasificados correctamente en su respectiva clase (aciertos).

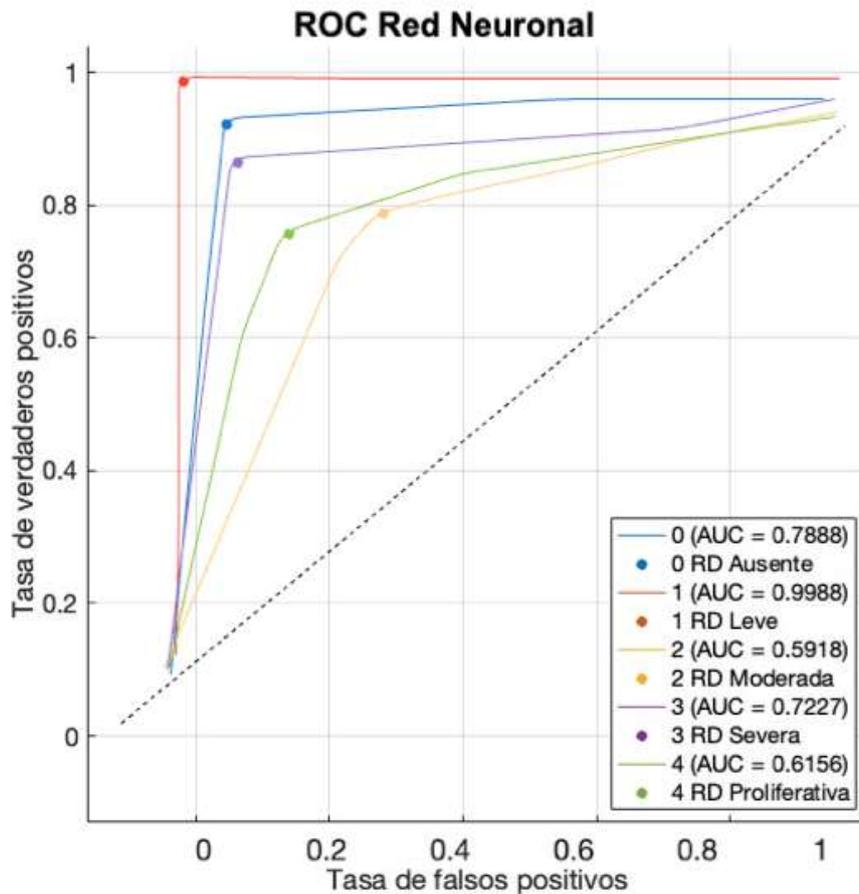


Figura 5.15: Gráfica ROC para el algoritmo RNA perceptrón multicapa, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad. Clasificando en la clase 2 (RD Leve).

### 5.2.7. Clasificación para el conjunto de Glaucoma

Como se realizó en el conjunto de RD, para en caso del conjunto de Glaucoma, se describirá exactamente igual el algoritmo que brindó una mejor precisión.

Para la clasificación del conjunto de datos de Glaucoma, se emplearon los mismos algoritmos de aprendizaje automático (ML), esta tarea es importante para verificar la clasificación de los patrones en las 2 clases correspondientes al grado de severidad.

El algoritmo que mostró una mejor exactitud de clasificación fue una SVM, brindando un 94.3%.

A continuación se muestra la matriz de confusión, mostrando los patrones clasificados en las diferentes clases del conjunto de datos con base a sus características:

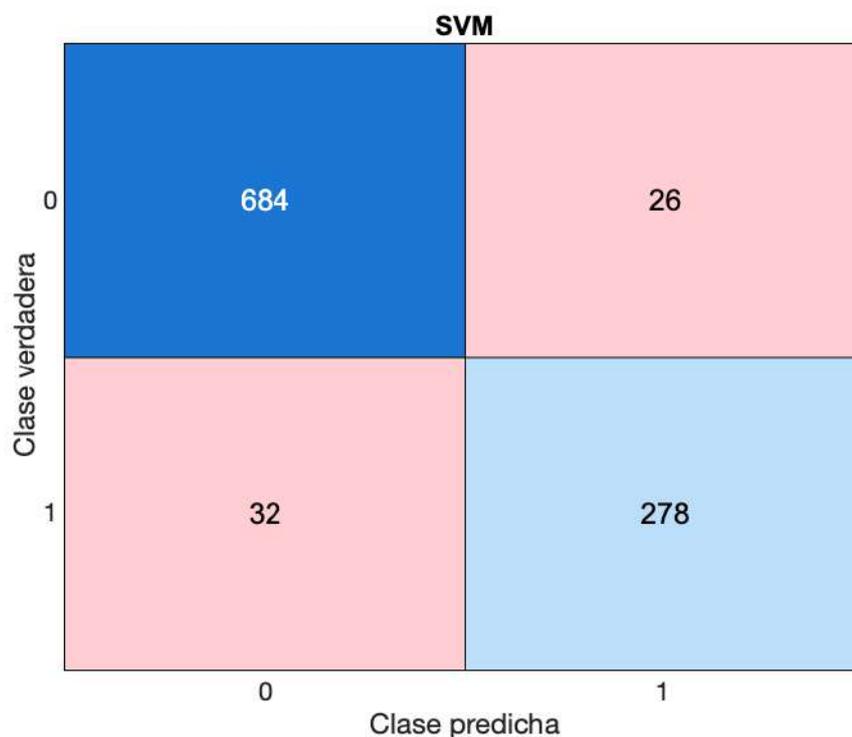


Figura 5.16: Matriz de confusión la clasificación de glaucoma utilizando una SVM. Clasificando en la clase 0 (GPAA).

Técnica de clasificación	Precisión	Aciertos
NN	93.2 %	951
SVM	94.3 %	962
Bayes Naive	74.8 %	763
Árboles de decisión	65.9 %	673
KNN	87.4 %	892

Cuadro 5.2: Resultados experimentales en Matlab para el conjunto de datos de Glaucoma

Asimismo, la sensibilidad y especificidad de la SVM, plasman una distribución específica para la clase de 0 (GPAA), lo que significa que la clasificación para el grado de severidad detectado de glaucoma, aumenta en la tasa de verdaderos positivos, es decir, el mayor porcentaje de las características se clasifican correctamente.

Como se puede observar en la tabla 5.2, se muestran las precisiones correspondientes al conjunto de imágenes de glaucoma, detallando las clasificaciones realizadas por los diferentes algoritmos de aprendizaje automático utilizado. Además, se muestra el total de patrones clasificados correctamente en su respectiva clase (aciertos).

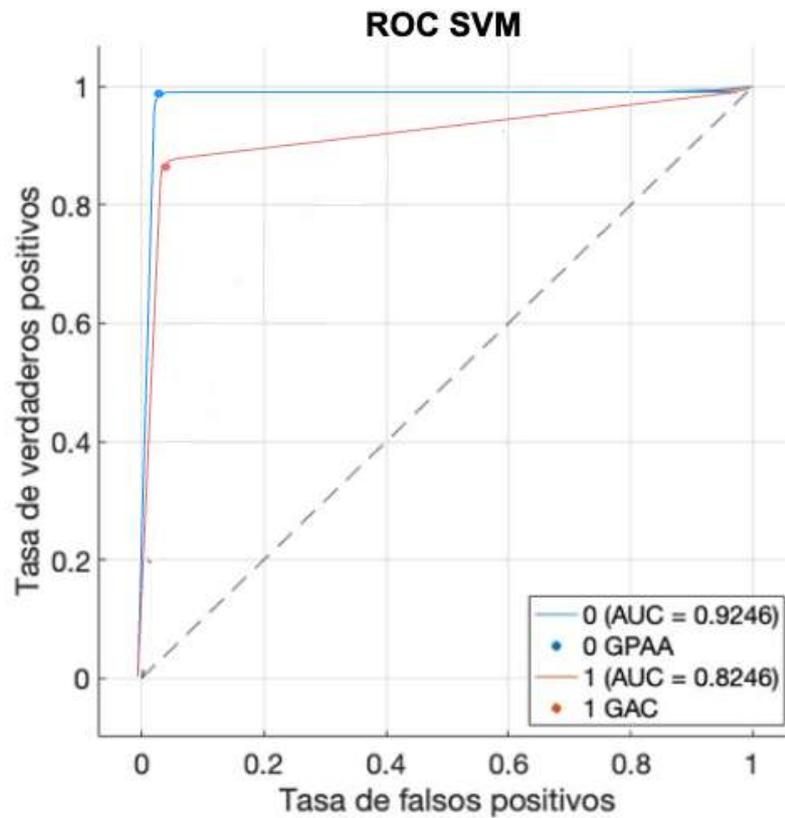


Figura 5.17: Gráfica ROC para el algoritmo SVM, mostrando la distribución de las clases de acuerdo a la sensibilidad y especificidad. Clasificando en la clase 0 (GPAA).

### 5.3. Discusión de los resultados

Los resultados obtenidos mediante el uso de algoritmos de aprendizaje automático, brindó resultados satisfactorios para detectar RD y glaucoma, así como, los grados de severidad para ambos padecimientos.

En la siguiente tabla, se hace una comparación del modelo propuesto y los métodos desarrollados recientemente, los cuales fueron mencionados en el estado del arte.

Autor	Métodos empleados	Precisión
Liu Li, et al. (2019).	Red neuronal convolucional (CNN) llamada AG-CNN.	90 %.
Huaying Hao, et al. (2019)	Base de datos pública para diagnóstico (CAD).	88.4 %
CisnerosFernanda, et al. (2019)	SVM	87.3 % y 84.6 %
Moreno, Sandra, et al. (2019)	RNA y KNN	86 % y 84 %
Huaying Hao, et al. (2019)	CNN	91 %
Deepak Parashar, et al. (2020)	FAWT	92.4 %
Luna, Dalia (2023)	ML y procesamiento de imágenes	92.1 % y 94.3 %

Cuadro 5.3: Resultados experimentales en Matlab para el conjunto de datos de Glaucoma

En MATLAB, la implementación de SVM cúbica se realiza mediante la función "fitcsvm", que permite entrenar un modelo SVM utilizando diferentes tipos de kernel, incluyendo la función kernel cúbica. Esta función proporciona una manera eficiente y sencilla de utilizar SVM cúbica en proyectos de clasificación. El modelo de Red Neuronal Artificial (RNA) empleado para clasificar en este estudio ha demostrado un rendimiento significativo, alcanzando una precisión del 92.1 % en la tarea de clasificación de características. Este resultado es prometedor y nos lleva a considerar varios aspectos clave relacionados con el modelo y la tarea en cuestión.

### 5.3.1. Efectividad en problemas de clasificación lineal y no lineal

Las SVM pueden manejar tanto problemas de clasificación lineal como no lineal. Esto significa que son versátiles y pueden adaptarse a una amplia variedad de problemas de aprendizaje automático, independientemente de si los datos son separables linealmente o no.

La SVM cúbica muestra un buen rendimiento en MATLAB porque es capaz de separar datos no linealmente separables en un espacio de mayor dimensión. Además, el uso de esta función kernel permite evitar el sobreajuste de los datos, ya que es menos propenso a caer en problemas de sesgo y varianza.

Las SVM pueden manejar tanto problemas de clasificación lineal como no lineal. Esto significa que son versátiles y pueden adaptarse a una amplia variedad de problemas de aprendizaje automático, independientemente de si los datos son separables linealmente o no.

Es importante destacar que el modelo se ha entrenado con un conjunto de datos lo suficientemente grande y diverso para representar adecuadamente las categorías objetivo. El tamaño del conjunto de datos y su calidad son factores críticos para el éxito del modelo.

Dado el éxito alcanzado con este modelo, los pasos futuros podrían incluir la ex-

pansión del conjunto de datos, la exploración de arquitecturas más complejas y la evaluación del rendimiento en situaciones de datos desequilibrados. Además, podría considerarse la interpretación de las características aprendidas por el modelo para mejorar la comprensión del problema.

## CONCLUSIONES

En primer lugar, se ha realizado una búsqueda en la literatura sobre el aprendizaje automático y la fisiología ocular relacionada con el glaucoma y la retinopatía diabética, así como sobre los fundamentos de los algoritmos utilizados. Esto condujo al descubrimiento de diversas técnicas basadas en imágenes para diagnosticar estas enfermedades. Además, se ha llevado a cabo una revisión exhaustiva de los estudios existentes para comprender el estado actual de la investigación en la clasificación de imágenes mediante el aprendizaje automático para el diagnóstico del glaucoma y de la retinopatía diabética.

Asimismo, se realizó un análisis de los diferentes algoritmos de aprendizaje automático, para establecer aquellos que, con base en sus características, brindaran un proceso de trabajo más óptimo para trabajar con el procesamiento y extracción de características de los conjuntos de datos. Se dedujo que la precisión, surge de realizar ciertos ajustes a los parámetros, sobre todo en el caso de las Redes Neuronales y SVM.

Los conjuntos de imágenes que se obtuvieron de DRIVE, Kaggle y EyePACS resultaron contener imágenes en diferentes ángulos de la retina, lo que hace más sencillo al sistema de visión, rescatar la gran mayoría de los rasgos, gracias a su alta calidad digital, es por ello que, al trabajar con un conjunto de datos que contiene imágenes con una definición consirebale, es posible que el proceso de extracción y selección de características sea más ágil, y permite reducir la discriminación de características relevantes.

Como resultado del procesamiento de las imágenes, los algoritmos empleados en esta fase, generaron la obtención de características, las cuales fueron procesadas a una etapa de clasificación con diferentes algoritmos de aprendizaje automático. La gran cantidad de características y ejemplos proporciona una base sólida para que el sistema de visión haya realizado una clasificación robusta y efectiva. En resumen, se ha realizado una preparación exhaustiva de datos que debería respaldar un rendimiento confiable en la tarea de clasificación.

Para los objetivos planteados en esta investigación, la SVM ha obtenido un rendimiento del 89.9 %, una RNA generó una precisión del 92.1 %, y un 94.3 % empleando nuevamente una SVM. Por lo tanto, los resultados obtenidos en comparación con otros métodos, sugiere que se puede generalizar bien a partir de los datos de entre-

namiento. En otras palabras, es capaz de hacer predicciones precisas en datos que no ha visto antes, lo que es esencial para la aplicabilidad en situaciones del mundo real.

## 6.1. Trabajos a futuro

A continuación, se mencionan algunos trabajos a futuro, los cuales pueden ser implementados al modelo propuesto en este trabajo de tesis, con el fin de brindar apoyo a la sociedad y mejores resultados a largo plazo.

### 6.1.1. Mejorar la precisión del modelo

Explorar formas de mejorar la precisión, incluyendo la optimización de hiperparámetros o, si es necesario, incorporar más datos.

1. Optimización de Hiperparámetros: Una estrategia inicial para mejorar la precisión será explorar la optimización de hiperparámetros. Esto implicará ajustar los parámetros del modelo, como la tasa de aprendizaje, el número de capas y neuronas en la red neuronal, las funciones de activación y otros hiperparámetros relevantes.
2. Regularización: Además de la optimización de hiperparámetros, se investigarán técnicas de regularización, como la adición de términos de penalización en la función de pérdida. La regularización puede ayudar a prevenir el overfitting y mejorar la generalización del modelo.
3. Aumento de Datos: El aumento de datos es otra estrategia importante para mejorar la precisión. Se explorará la posibilidad de generar datos adicionales mediante técnicas de aumento de datos, como la rotación, la traslación, el cambio de escala y la inversión. Esto permitirá que el modelo se entrene con un conjunto de datos más diverso y robusto.
4. Adquisición de Datos Adicionales: Si se considera necesario, se buscarán fuentes adicionales de datos para enriquecer el conjunto de datos de entrenamiento. La adquisición de datos adicionales puede implicar la recopilación de nuevas muestras o la búsqueda de conjuntos de datos públicos que sean relevantes para la tarea.

### 6.1.2. Predicción en la progresión de Glaucoma y RD

Además de detectar el grado de severidad actual, también predecir cómo cambiará la severidad con el tiempo.

El trabajo a futuro se centrará en avanzar más allá de la simple detección del grado de severidad actual del glaucoma y la retinopatía diabética (RD) para abordar la cuestión crucial de predecir cómo cambiará la severidad de estas enfermedades con el tiempo.

1. Modelos de Series Temporales: Se desarrollarán modelos de series temporales que puedan capturar y predecir la progresión de la enfermedad con el tiempo. Esto implica el seguimiento de pacientes a lo largo del tiempo y el análisis de cómo cambian las mediciones clave, como la presión intraocular en el caso del glaucoma o la evolución de las lesiones en el caso de la RD.
2. Datos Longitudinales: Será esencial recopilar y utilizar datos longitudinales de pacientes que hayan sido evaluados en múltiples momentos a lo largo de un período prolongado.

### 6.1.3. Validación en poblaciones diversas

Asegurar que el modelo propuesto sea aplicable a diversos grupos de la población, tomando en cuenta factores como la edad, género y grupos sociales.

Esta tarea se enfoca en garantizar que cualquier modelo, sistema o enfoque desarrollado sea válido y aplicable en una amplia variedad de poblaciones, lo que implica la inclusión de diversidad étnica, cultural, geográfica y socioeconómica. Esta tarea es esencial para garantizar la equidad, la representatividad y la utilidad de las soluciones en un mundo diverso y globalizado.

### 6.1.4. Creación de un conjunto de datos

Implica la generación de un conjunto de datos específico para el caso de glaucoma, debiera estar personalizado para abordar un problema particular de investigación, desarrollo de modelos o aplicación de algoritmos. La creación de un conjunto de datos es fundamental para garantizar que haya datos relevantes y de alta calidad disponibles para llevar a cabo análisis y entrenamiento de modelos.

Para el caso del glaucoma, sería importante incluir los cuatro grados de severidad,

posteriormente se aplicaría el modelo propuesto en esta tesis, para verificar los resultados de precisión que brinde.

## Bibliografía

- [1] Yhary Estefanía Arias Trillos. Sistema web de reconocimiento y clasificación de patologías a través de imágenes médicas basado en técnicas de aprendizaje de máquina. 2019.
- [2] Devansh Arpit, Stanisław Jastrzębski, Nicolas Ballas, David Krueger, Emmanuel Bengio, Maxinder S Kanwal, Tegan Maharaj, Asja Fischer, Aaron Courville, Yoshua Bengio, et al. A closer look at memorization in deep networks. In *International conference on machine learning*, pages 233–242. PMLR, 2017.
- [3] Wilmer Rivas Asanza and Bertha Mazón Olivo. Redes neuronales artificiales aplicadas al reconocimiento de patrones. *Editorial UTMACH*, 2018.
- [4] Elie Azeraf, Emmanuel Monfrini, and Wojciech Pieczynski. Using the naive bayes as a discriminative classifier. *arXiv preprint arXiv:2012.13572*, 2020.
- [5] Muhammad Naseer Bajwa, Gur Amrit Pal Singh, Wolfgang Neumeier, Muhammad Imran Malik, Andreas Dengel, and Sheraz Ahmed. G1020: A benchmark retinal fundus image dataset for computer-aided glaucoma detection. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7. IEEE, 2020.
- [6] M Batouche, R Benlamri, and MK Kholadi. A computer vision system for diagnosing scoliosis using moiré images. *Computers in Biology and Medicine*, 26(4):339–353, 1996.
- [7] Bradley Bowling. *Kanski. Oftalmología clínica: Un enfoque sistemático*. Elsevier, 2016.
- [8] Bárbara Burgos Blasco. Estudio del perfil de citoquinas en la lágrima y el humor acuoso de pacientes con glaucoma. 2022.
- [9] Christian Marcelo Cajamarca Bueno and Marlon Fernando Chazi Solis. Desarrollo de una herramienta para la identificación y prevención del glaucoma a través de procesamiento digital de imágenes biomédicas mediante la relación de volumen de fibras nerviosas. B.S. thesis, 2018.
- [10] Gaspar Cano, Jose García-Rodríguez, S Orts, Alberto García-García, J Peña-García, Alfonso Pérez-Garrido, and Horacio Pérez-Sánchez. Predicción de solubilidad de fármacos usando máquinas de soporte vectorial sobre unidades de procesamiento gráfico. *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería*, 33(1-2):97–102, 2017.

- [11] Jair Cervantes, Farid García-Lamont, Lisbeth Rodríguez-Mazahua, Arturo Yee Rendon, and Asdrúbal López Chau. Recognition of mexican sign language from frames in video sequences. In *Intelligent Computing Theories and Application*, pages 353–362. Springer International Publishing, 2016.
- [12] Jair Cervantes, Farid Garcia Lamont, Lisbeth Rodriguez Mazahua, Alfonso Zarco Hidalgo, and José S. Ruiz Castilla. Complex identification of plants from leaves. In *Intelligent Computing Methodologies*, pages 376–387. Springer International Publishing, 2018.
- [13] Jair Cervantes, Xiaoou Li, and Yu Wen. Support vector classification for large data sets by reducing training data with change of classes. In *2008 IEEE International Conference on Systems, Man and Cybernetics*. IEEE, oct 2008.
- [14] Jair Cervantes, Xiaoou Li, and Wen Yu. Support vector machine classification based on fuzzy clustering for large data sets. In *Lecture Notes in Computer Science*, pages 572–582. Springer Berlin Heidelberg, 2006.
- [15] Jair Cervantes, Xiaoou Li, and Wen Yu. SVM classification for large data sets by considering models of classes distribution. In *2007 Sixth Mexican International Conference on Artificial Intelligence, Special Session (MICAI)*. IEEE, nov 2007.
- [16] Jair Cervantes, Xiaoou Li, and Wen Yu. Splice site detection in DNA sequences using a fast classification algorithm. In *2009 IEEE International Conference on Systems, Man and Cybernetics*. IEEE, oct 2009.
- [17] Jair Cervantes, Xiaoou Li, and Wen Yu. Imbalanced data classification via support vector machines and genetic algorithms. *Connection Science*, 26(4):335–348, jun 2014.
- [18] Jair Cervantes, Xiaoou Li, Wen Yu, and Kang Li. Support vector machine classification for large data sets via minimum enclosing ball clustering. *Neurocomputing*, 71(4-6):611–619, jan 2008.
- [19] Jair Cervantes, Jesús Taltempa, Farid García-Lamont, José S. Ruiz Castilla, Arturo Yee Rendon, and Laura D. Jalili. Análisis comparativo de las técnicas utilizadas en un sistema de reconocimiento de hojas de planta. *Revista Iberoamericana de Automática e Informática Industrial RIAI*, 14(1):104–114, jan 2017.
- [20] M Chethan, Chandrashekar Dasari, Gaurav V Uttarkar, and DN Sachin. Diagnosis of glaucoma using machine learning-a survey. In *2019 Third International*

- Conference on Inventive Systems and Control (ICISC)*, pages 210–214. IEEE, 2019.
- [21] Raquel Ciriza, L Albizua, and M Gonzales. Análisis de la utilidad de los descriptores texturales de haralick para la localización arranques de frutal en ortofoto. In *Teledetección: Agua y desarrollo sostenible. XIII Congreso de la Asociación Espanola de Teledetección*, pages 597–600, 2009.
- [22] Fernanda Cisneros-Guzmán, Saúl Tovar-Arriaga, Carlos Pedraza, and Arturo González-Gutierrez. Classification of diabetic retinopathy based on hard exudates patterns, using images processing and svm. In *2019 IEEE Colombian Conference on Applications in Computational Intelligence (ColCACI)*, pages 1–5. IEEE, 2019.
- [23] Cynthia Lorena Corso, Calixto Maldonado, Florencia Pereyra, Gimena Martínez, and Matías Donnet. Fusión de algoritmos bayesianos y árboles de clasificación como propuesta para la clasificación supervisada de fallos de equipos en un laboratorio de cómputos. In *XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires)*, 2017.
- [24] Alfredo R Cuello, Miriam E Antes, and Ailin S Ortone Lois. Teledetección y geomática aplicada al monitoreo de los recursos forestales nativos en el ámbito de la defensa.
- [25] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, pages 886–893. Ieee, 2005.
- [26] Enrique José De Ramón Balmaseda. Transformaciones basadas en el algoritmo local binary pattern de imágenes capturadas con la kinect para clasificación facial. Master's thesis, 2011.
- [27] Andres Diaz-Pinto, Adrián Colomer, Valery Naranjo, Sandra Morales, Yanwu Xu, and Alejandro F Frangi. Retinal image synthesis and semi-supervised learning for glaucoma assessment. *IEEE transactions on medical imaging*, 38(9):2211–2218, 2019.
- [28] Josué Espejel-Cabrera, Jair Cervantes, Farid García-Lamont, José Sergio Ruiz Castilla, and Laura D. Jalili. Mexican sign language segmentation using color based neuronal networks to detect the individual skin color. *Expert Systems with Applications*, 183:115295, nov 2021.

- [29] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231, 1996.
- [30] Christian Layme Fernández, José Manuel Suri Canaza, David Jose Peña Ugarte, and Jhon Yoset Luna Quispe. Aplicación de los árboles de decisión en la identificación de sitios web fraudulentos. *Innovación y Software*, 3(1):6–16, 2022.
- [31] Luis A Fernández, Daniel Diaz, and Roberto Depaoli. Optimización de la equalización del histograma en el procesamiento de imágenes digitales. In *VII Workshop de Investigadores en Ciencias de la Computación*, 2005.
- [32] Jason Fung. *El código de la diabetes: Prevenir y revertir la diabetes tipo-2 de manera natural*. Editorial Sirio SA, 2018.
- [33] Óscar García-Olalla Olivera, Enrique Alegre Gutiérrez, et al. Descripción de textura en imágenes utilizando local binary pattern (lbp).
- [34] Huaying Hao, Yitian Zhao, Huazhu Fu, Qiaoling Shang, Fei Li, Xiulan Zhang, and Jiang Liu. Anterior chamber angles classification in anterior segment oct images via multi-scale regions convolutional neural networks. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 849–852. IEEE, 2019.
- [35] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621, 1973.
- [36] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.
- [37] Francisco Javier Moreno Hernández, Raúl Reina Vaíllo, V Luis, Jesús Sebastián Damas Arroyo, and Rafael Sabido Solana. Desarrollo de un sistema tecnológico para el registro del comportamiento de jugadores de tenis y tenis en silla de ruedas en situaciones de respuesta de reacción. *European Journal of Human Movement*, (10):165–199, 2003.
- [38] Chih-Wei Hsu and Chih-Jen Lin. A practical guide to support vector classification. *Technical report*, 2003.
- [39] JH Santiago JE Cabrera Adrian Trueba J Cervantes, FG Lamont. Clasificación del lenguaje de señas mexicano con svm generando datos artificiales -. *Revista vínculos*, 10(1), 2013.

- [40] Laura Jalili, Jair Cervantes, Farid García-Lamont, and Adrián Trueba. Emotion recognition from facial expressions using a genetic algorithm to feature extraction. In *Intelligent Computing Theories and Application*, pages 59–71. Springer International Publishing, 2021.
- [41] Laura D. Jalili, Alfredo Morales, Jair Cervantes, and José S. Ruiz-Castilla. Improving the performance of leaves identification by features selection with genetic algorithms. In *Communications in Computer and Information Science*, pages 103–114. Springer International Publishing, 2016.
- [42] Omar Khan, Jetan H Badhiwala, Giovanni Grasso, and Michael G Fehlings. Use of machine learning and artificial intelligence to drive personalized medicine approaches for spine care. *World neurosurgery*, 140:512–518, 2020.
- [43] Tarun Kumar and Karun Verma. A theory based on conversion of rgb image to gray image. *International Journal of Computer Applications*, 7(2):7–10, 2010.
- [44] Andrew G Lee. Clinical pathways in neuro-ophthalmology. *An evidence-based approach*, 2003.
- [45] Annan Li, Yunhong Wang, Jun Cheng, and Jiang Liu. Combining multiple deep features for glaucoma classification. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 985–989. IEEE, 2018.
- [46] Liu Li, Mai Xu, Hanruo Liu, Yang Li, Xiaofei Wang, Lai Jiang, Zulin Wang, Xiang Fan, and Ningli Wang. A large-scale database and a cnn model for attention-based glaucoma detection. *IEEE transactions on medical imaging*, 39(2):413–424, 2019.
- [47] Xiaou Li, Jair Cervantes, and Wen Yu. Two-stage svm classification for large data sets via randomly reducing and recovering training data. In *2007 IEEE International Conference on Systems, Man and Cybernetics*. IEEE, 2007.
- [48] Xiaou Li, Jair Cervantes, and Wen Yu. A novel SVM classification method for large data sets. In *2010 IEEE International Conference on Granular Computing*. IEEE, aug 2010.
- [49] Xiaou Li, Jair Cervantes, and Wen Yu. Fast classification for large data sets via random selection clustering and support vector machines. *Intelligent Data Analysis*, 16(6):897–914, nov 2012.
- [50] Miguel López Montiel, Yoshio Rubio, Moises Sánchez-Adame, and Ulises Orozco Rosas. Algoritmo de procesamiento de imágenes para señales de tránsito. 2019.

- [51] Sandra Moreno Peramo. *Diseño y desarrollo de un sistema de detección automática de Glaucoma a partir de la segmentación y el estudio de las capas de la retina en imágenes OCT circumpapilares*. PhD thesis, Universitat Politècnica de València, 2019.
- [52] Rodolfo Mosquera, Omar D Castrillón, and Liliana Parra. Máquinas de soporte vectorial, clasificador naïve bayes y algoritmos genéticos para la predicción de riesgos psicosociales en docentes de colegios públicos colombianos. *Información tecnológica*, 29(6):153–162, 2018.
- [53] Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996.
- [54] Bryon Wladimir Oviedo Bayas and Cristian Zambrano Vega. Redes bayesianas aplicadas a la predicción de errores en las redes definidas por software. *Revista Universidad y Sociedad*, 13(2):419–429, 2021.
- [55] Mercedes Eugenia Paoletti, Juan Mario Haut, Javier Plaza, and Antonio Plaza. Estudio comparativo de técnicas de clasificación de imágenes hiperespectrales. *Revista Iberoamericana de Automática e Informática Industrial*, 16(2):129–137, 2019.
- [56] Deepak Parashar and Dheeraj Kumar Agrawal. Automated classification of glaucoma stages using flexible analytic wavelet transform from retinal fundus images. *IEEE Sensors Journal*, 20(21):12885–12894, 2020.
- [57] Vimla L Patel, Edward H Shortliffe, Mario Stefanelli, Peter Szolovits, Michael R Berthold, Riccardo Bellazzi, and Ameen Abu-Hanna. The coming of age of artificial intelligence in medicine. *Artificial intelligence in medicine*, 46(1):5–17, 2009.
- [58] P. Patil and A. Kumar. K-nearest neighbors algorithm using python. *International Journal of Advanced Research in Computer Science*, 8(5):946–949, 2017.
- [59] Judea Pearl. Theoretical impediments to machine learning with seven sparks from the causal revolution. *arXiv preprint arXiv:1801.04016*, 2018.
- [60] Carrillo Peñaloza and Yersica Paola. Desarrollo de un software para determinar el resultado obtenido por unos dados en el juego craps mediante procesamiento de imágenes y técnicas de visión artificial para personas no videntes o con baja capacidad visual. 2022.

- [61] LN Romaña Peralta and M Sc Mario Gregorio Saavedra Rodríguez. Modelo predictivo para la selección de docentes a través de un aprendizaje supervisado. *Fundación Universitaria Los Libertadores*, 2020.
- [62] Anjuman Prabhat and Vikas Khullar. Sentiment classification on big data using naïve bayes and logistic regression. In *2017 International Conference on Computer Communication and Informatics (ICCCI)*, pages 1–5. IEEE, 2017.
- [63] José Antonio Ramos Alonso. Generación de algoritmos para clasificación de obstáculos. B.S. thesis, 2016.
- [64] Sebastian Raschka and Vahid Mirjalili. *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2*. Packt Publishing, 2019.
- [65] Rodrigo Rojas-Moraleda, Nektarios A Valous, Aoife Gowen, Carlos Esquerre, Steffen Härtel, Luis Salinas, and Colm O’donnell. A frame-based ann for classification of hyperspectral images: assessment of mechanical damage in mushrooms. *Neural Computing and Applications*, 28(1):969–981, 2017.
- [66] Ángel Ruiz Alonso. Comportamiento y análisis de descriptores de texturas en imágenes modis. 2011.
- [67] David Sánchez Alvarez. *Detección de exudados en imágenes digitales de fondo de ojo*. PhD thesis, Universidad Central "Marta Abreu" de Las Villas, Facultad de Ingeniería . . . , 2021.
- [68] Arturo Navarro Saucedo. Desarrollo de un prototipo de cámara de fondo de ojo no midriática. 2019.
- [69] Anderson Solano Bernal et al. Scoring de riesgo de lavado de activos y financiación del terrorismo. 2022.
- [70] Vasil Teigens, Peter Skalfist, and Daniel Mikelsten. *Inteligencia artificial: la cuarta revolución industrial*. Cambridge Stanford Books, 2020.
- [71] Arturo Torres Imán. Diseño de guiado de un robot móvil por visión artificial. 2019.
- [72] Maria Fernanda URDAÑEZ CARBAJAL. Detección automática de glaucoma y retinopatía diabética usando características cromáticas y texturales.
- [73] Gabriela Isamar Viera Maza. Procesamiento de imágenes usando opencv aplicado en raspberry pi para la clasificación del cacao. 2017.

- [74] Hongzhi Wang and John Oliensis. Generalizing edge detection to contour detection for image segmentation. *Computer Vision and Image Understanding*, 114(7):731–744, 2010.
- [75] Tien Yin Wong and Neil M Bressler. Artificial intelligence with deep learning technology looks into diabetic retinopathy screening. *Jama*, 316(22):2366–2367, 2016.
- [76] Garima Yadav, Saurabh Maheshwari, and Anjali Agarwal. Contrast limited adaptive histogram equalization based enhancement for real time video system. In *2014 international conference on advances in computing, communications and informatics (ICACCI)*, pages 2392–2397. IEEE, 2014.
- [77] Maciel Zortea, Thomas R Schopf, Kevin Thon, Marc Geilhufe, Kristian Hindberg, Herbert Kirchesch, Kajsa Møllersen, Jörn Schulz, Stein Olav Skrøvseth, and Fred Godtlielsen. Performance of a dermoscopy-based computer vision system for the diagnosis of pigmented skin lesions compared with visual evaluation by experienced dermatologists. *Artificial intelligence in medicine*, 60(1):13–26, 2014.