INSTITUTO TECNOLÓGICO SUPERIOR DE MISANTLA

CLASIFICACIÓN DEL GRADO DE RETINOPATÍA DIABÉTICA CON APRENDIZAJE AUTOMÁTICO

T E S I S

PARA OBTENER EL GRADO DE MAESTRO EN SISTEMAS COMPUTACIONALES

P R E S E N T A:

ANGEL SALAZAR ALVAREZ

DIRECTOR DE TESIS

DR. SIMÓN PEDRO ARGUIJO HERNÁNDEZ

CODIRECTORES

DR. DAVID LARA ALAVAZARES
DR. ROBERTO ÁNGEL MELÉNDEZ ARMENTA

MISANTLA, VER. AGOSTO 2024

INSTITUTO TECNOLÓGICO SUPERIOR DE MISANTLA

AUTORIZACIÓN DE IMPRESIÓN DE TRABAJO DE TITULACIÓN MAESTRÍA

FECHA:
28 DE AGOSTO DE 2024

ASUNTO: **AUTORIZACIÓN DE IMPRESIÓN: TESIS PROFESIONAL**

A QUIEN CORRESPONDA:

Por medio de la presente se hace constar que el (la) C:

ÁNGEL SALAZAR ÁLVAREZ

Estudiante de la **MAESTRÍA EN SISTEMAS COMPUTACIONALES** con No. de Control **222T0541** ha cumplido satisfactoriamente con lo estipulado por el **Lineamiento de Posgrado para la obtención del grado de Maestría** mediante **TESIS PROFESIONAL**

Por tal motivo se **Autoriza** la impresión del **Tema** titulado:

"CLASIFICACIÓN DEL GRADO DE RETINOPATÍA DIABÉTICA CON APRENDIZAJE AUTOMÁTICO"

Dándose un plazo no mayor de un mes de la expedición de la presente a la solicitud del examen para la obtención del grado de maestría.

ATENTAMENTE

SEP SEV INSTITUTO TECNOLÓGICO SUPERIOR DE MISANTLA

COORDINACIÓN DE ESTUDIOS PROFESIONALES

DR. SIMÓN PEDRO ARGUIJO HERNÁNDEZ
Presidente

DR. DAVID LARA ALABAZARES
Secretario

DR. ROBERTO ÁNGEL MELENDEZ ARMENTA
Vocal

Archivo.

Agradecimientos

Quiero expresar mi más profundo agradecimiento al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) por el invaluable apoyo brindado durante el transcurso de mi maestría, ya que los recursos proporcionados fueron fundamentales para la realización de este proyecto y el avance en mi formación académica y profesional.

Asimismo, extiendo mi sincero agradecimiento al Instituto Tecnológico Superior de Misantla, donde llevé a cabo este proceso. Su infraestructura, recursos académicos y el constante apoyo de su personal docente y administrativo fueron muy importantes para el desarrollo de este trabajo.

Resumen

La retinopatía diabética (DR) es una complicación microvascular de la diabetes mellitus, que afecta la retina y puede llevar a la ceguera si no se detecta y es tratada a tiempo. Con el creciente número de pacientes diabéticos a nivel mundial, la detección temprana y precisa de DR se ha convertido en una prioridad crítico en la salud pública. Las tecnologías de imagen modernas y los algoritmos de aprendizaje automático han abierto nuevas posibilidades para el diagnóstico automatizado y la evaluación del grado de severidad de DR.

En esta investigación, nos enfocamos en comparar diferentes métodos de extracción de características y técnicas para la detección y clasificación de DR. Utilizando el dataset APTOS 2019 Blindness Detection, el cual incluye imágenes de retina con diversas condiciones y clasificaciones de DR en cinco etapas de severidad, evaluamos la eficacia y eficiencia de los métodos basados en análisis de textura (LBP y GLCM) y los basados en redes neuronales profundas (DenseNet 121 y Ademxapp Model A).

El análisis se realiza en dos enfoques de clasificación: binaria y multiclase. La clasificación binaria distingue entre la presencia y ausencia de DR, mientras que la clasificación multiclase identifica los diferentes grados de severidad de la enfermedad. Además, evaluamos la influencia de diferentes espacios de color (RGB, LAB, YCbCr) y técnicas de mejora de imagen en la extracción de características y el rendimiento de los clasificadores.

Palabras clave: Retinopatía Diabética, extracción de características, LBP, GLCM, DenseNet121, Ademxapp Model A, RGB, LAB, YCbCr.

Abstract

Diabetic Retinopathy (DR) is a microvascular complication of diabetes mellitus, which affects the retina and can lead to blindness if not detected and treated early. With the increasing number worldwide of diabetic patients, early DR detection DR has become a critical public health priority. Modern imaging technologies and machine learning algorithms have opened up new possibilities for automated diagnosis and assessment of severity DR degree.

In this research, we focus on comparing different feature extraction methods and techniques for DR detection and classification. Using the APTOS 2019 Blindness Detection dataset, which includes retinal images with different conditions and DR classifications in five stages of severity, we evaluate the effectiveness and efficiency of texture analysis-based (LBP and GLCM) and deep neural network-based (DenseNet-21 and Ademxapp Model A) methods.

The analysis is performed in two classification approaches: binary and multiclass. Binary classification distinguishes between the presence and absence of DR, while multiclass classification identifies different degrees of disease severity. Furthermore, we evaluate the influence of different colour spaces (RGB, LAB, YCbCr) and image enhancement techniques on feature extraction and classifier performance.

Keywords: Diabetic Retinopathy, feature extraction, LBP, GLCM, DenseNet-121, Ademxapp Model A, RGB, LAB, YCbCr.

Índice general

Resumen **IV**

Abstract **V**

1. Generalidades. **1**

1.1. Antecedentes. 1

1.2. Planteamiento del problema. 3

1.3. Justificación. 3

1.4. Preguntas de investigación. 4

1.5. Hipótesis. 5

1.6. Objetivo General. 5

1.6.1. Objetivos específicos 6

1.7. Metodología. 6

2. Marco Teórico. **8**

2.1. Retinopatía Diabética. 8

2.2. Imagen Digital. 9

2.3. Espacios de Color. 10

2.3.1. Espacio de color RGB 11

2.3.2. Espacio de color LAB 12

2.3.3. Espacio de color YCbCr 13

2.4. Filtros. 14

2.5. CLAHE. 15

2.6. Textura. 17
2.6.1. Patrones Locales Binarios (LBP). 18
2.6.2. Características de Haralick. 19
2.7. Redes Neuronales Profundas (DNN). 24
2.7.1. DenseNet-121 26
2.7.2. Ademxapp Model A 27
2.8. Máquina de Vector Soporte (SVM). 28
2.9. Random Forest (RF). 28
2.10. Gradient Boosted Trees (GBT). 29
2.11. Trabajos Relacionados. 30

3. Desarrollo. 32
3.1. Conjunto de datos. 32
3.2. Procesamiento de imágenes. 34
3.2.1. Mejoramiento de imágenes en espacio de color LAB. 35
3.2.2. Mejoramiento de imágenes espacio de color YCbCr 37
3.3. Extractores de Características. 38
3.3.1. LBP. 39
3.3.2. Características de Haralick. 40
3.3.3. Ademaxpp Model A. 41
3.3.4. DenseNet-121. 41
3.4. Clasificación. 42
3.4.1. Binaria. 42
3.4.2. Multiclase. 43
3.4.3. Muestreo. 44
3.4.4. Clasificación de características. 45

4. Resultados. 46
4.1. Mejoramiento de imágenes 46
4.2. Extracción de características 47
4.3. Clasficación binaria y multiclase 48

4.4. Análisis comparativo 59

5. Conclusiones. 63

Índice de figuras

1.1. Lesiones de DR, imagen tomada de [5]. 2

2.1. Imagen de 16 pixeles. 9
2.2. Espectro eletromagnético y espectro visible. 10
2.3. Espacio de color RGB. 11
2.4. Espacio de color LAB. 13
2.5. Imagen de fondo de retina antes de después de aplicar CLAHE. 16
2.6. Regiones en que se divide una imagen por medio de CLAHE. 16
2.7. Operador básico LBP. 18
2.8. Vecindario circular (P,R)$\Longrightarrow$(8,2) para vecindarios mayores de 3 x 3, imagen tomada de [19]. 19
2.9. GLCM de imagen, un píxel de referencia y su vecino inmediato. 20
2.10. Direcciones para GLCM completa. 20
2.11. Representación general de una ANN. 25
2.12. Estructura Dense-Net 121, 427 capas, imagen tomada de [24]. 26
2.13. Estructura Ademxapp Model A, 142 capas, imagen tomada de [24]. 27
2.14. Ejemplo de un problema separable en dos dimensiones. Los vectores de soporte están marcados con cuadros color gris y definen el margen óptimo de separación entre clases. El hiperplano se marca con la línea punteada. 28
2.15. Marco conceptual de Random Forest. 29
2.16. Conjunto de árboles de decisión impulsados por grandiente. 30

3.1. Etapas DR Aptos 2019 en espacio de color RGB. Muestra representativa de cada grado de retinopatía. 34

3.2. Proceso para mejorar imágenes en espacio de color LAB propuesto en [35]. . 36
3.3. Muestra de imágenes mejoradas de espacio de color RGB (primera fila) a LAB (segunda fila) propuesto en [35]. 36
3.4. Proceso para mejorar imágenes en espacio de color YCbCr propuesto en [36]. 37
3.5. Muestra de imágenes mejoradas de espacio de color RGB (primera fila) a YCbCr (segunda fila) propuesto en [36]. 38
3.6. Organización de datasets. 39
3.7. La primera fila muestra las imágenes en espacio de color RGB, mejoradas en espacio de color LAB y YCbCr, la segunda muestra la conversión respectiva en niveles de gris. 40
3.8. DNN's como extractores de características. 42

4.1. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores LBP, RGB. 49
4.2. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores LBP, LAB. 49
4.3. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores LBP, YCbCr. 49
4.4. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores GLCM, RGB. 49
4.5. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores GLCM, LAB. 50
4.6. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores GLCM, YCbCr. 50
4.7. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores DenseNet121, RGB. 50
4.8. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores DenseNet121, LAB. 50
4.9. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores DenseNet121, YCbCr. 51

4.10. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores Ademxapp Model A, RGB. 51
4.11. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores Ademxapp Model A, LAB. 51
4.12. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores Ademxapp Model A, YCbCr. 51
4.13. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores LBP, RGB. 52
4.14. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores LBP, LAB. 52
4.15. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores LBP, YCbCr. 53
4.16. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores GLCM, RGB. 53
4.17. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores GLCM, LAB. 54
4.18. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores GLCM, YCbCr. 54
4.19. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores DenseNet 121, RGB. 55
4.20. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores DenseNet 121, LAB. 55
4.21. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores DenseNet 121, YCbCr. 56
4.22. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores Ademxapp Model A, RGB. 56
4.23. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores Ademxapp Model A, LAB. 57
4.24. Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores Ademxapp Model A, YCbCr. 57

Índice de tablas.

3.1. Clasificación y composición APTOS 2019. 33
3.2. Resoluciones espaciales del conjunto de imágenes APTOS 2019. 34
3.3. Reorganización de clases para clasificación binaria. 42
3.4. Identificadores de clase para clasificación multiclase. 43
3.5. Muestreo aleatorio propuesto 80/20. 44

4.1. Imagen origianl en RGB y su mejoramiento en LAB y YCbCr. 46
4.2. Tiempo requerido para extracción de características. 47
4.3. Exactitud en espacio de color RGB. 58
4.4. Exactitud en espacio de color LAB. 58
4.5. Exactitud en espacio de color YCbCr. 58
4.6. Exactitud en espacio de color RGB. 58
4.7. Exactitud en espacio de color LAB. 58
4.8. Exactitud en espacio de color YCbCr. 59

Capítulo 1

Generalidades.

1.1. Antecedentes.

Diabetes Mellitus o Diabetes, es una condición médica que ocurre cuando los niveles de glucosa (azúcar) en la sangre son elevados en el cuerpo. Es un desorden metabólico crónico que afecta cómo el cuerpo procesa la glucosa. Este padecimiento puede ser causado por varios factores, entre ellos, el factor genético, estilos de vida tales como, hábitos alimenticios con alto contenido de azúcares refinadas, grasas y sodio, el sedentarismo, obesidad además de ciertas condiciones médicas [1]. El INEGI reportó que en el 2022 se registraron más de 115,000 defunciones causadas por diabetes, y en 2021 fue la tercera causa de muerte superada únicamente por problemas cardíacos y COVID-19, respectivamente[2].

Existen tres tipos de diabetes [3]:

- Tipo 1. Es una enfermedad en la que el sistema inmunológico ataca y destruye las células productoras de insulina en el páncreas. Esto resulta en una producción nula o limitada de insulina, lo que lleva a niveles altos de azúcar en la sangre. Este tipo de diabetes regularmente se diagnostica en la infancia o la adolescencia.

- Tipo 2. Es la forma mas común de diabetes mellitus y ocurre cuando el cuerpo se vuelve resistente a los efectos de la insulina o no produce la suficiente para mantener los niveles normales de azucar en la sangre y se presenta generalmente en personas adulta.

- Diabetes gestacional. Este tipo de diabetes ocurre durante el embarazo, con valores de glucosa por encima de lo normal, pero por debajo de los diagnósticos de diabetes. Presenta un mayor riesgo en complicaciones durante la gestación y parto, además se incrementa la posibilidad de adquirir diabetes tipo 2 para la madre e hijos.

Cuando la diabetes no se controla adecuadamente, pueden generarse complicaciones a largo plazo que afectan seriamente la salud de la persona. En terminología médica, las complicaciones asociadas con la diabetes se agrupan como micro y macrovasculares, las cuales se pueden dividir en enfermedades del corazón, problemas en riñones, problemas en extremidades del cuerpo, pies y manos, enfermedades en los ojos, afecciones dentales, en la piel, digestivas, sexuales, de salud mental y debilidad general en el sistema inmunológico, por lo que el riesgo a infecciones para las personas que padecen esta condición médica aumenta [4].

Esta investigación se enfoca en Retinopatía Diabética (DR, por sus siglas en inglés),la cual es una complicación ocular crónica que se desarrolla como resultado del daño en los vasos sanguíneos de la retina, lo que conduce a la perdida parcial de la visión en escenarios tempranos, evolucionando hasta su pérdida total en las etapas finales.

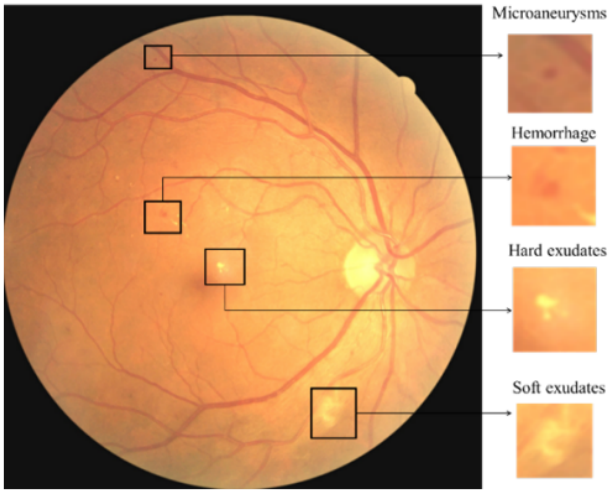

Figura 1.1: Lesiones de DR, imagen tomada de [5].

DR se presenta como una problema microvascular progresivo de la retina y se manifiesta clínicamente a través de varias lesiones: lesiones rojas tales como microaneurismas (MA) y hemorragias (HM), lesiones brillantes llamadas exudados de lípidos (EX) que pueden ser blandos o duros, y neovascularización, tal como se muestra en la figura 1.1. De acuerdo con el

número de lesiones presentes en la retina, DR se clasifica en 5 grados o etapas de severidad: No DR, leve, moderada, grave y proliferativa. Es importante el diagnóstico temprano de la afección ya que permite considerar tratamientos y medidas para disminuir o corregir el deterioro de la vista [6].

1.2. Planteamiento del problema.

DR presenta un desafío significativo en su diagnóstico y clasificación. Incluso los especialistas encuentran dificultades para discernir entre las distintas etapas de la enfermedad y se enfrentan a una tarea compleja al intentar categorizar con exactitud los diversos grados: desde la ausencia de retinopatía hasta los niveles moderados, severos y proliferativos. Este proceso se torna más desafiante para aquellos en formación de esta práctica.

La complejidad del diagnóstico radica en la subjetividad inherente a la evaluación de DR. La distinción entre los grados de afección no siempre es clara ni consistente, incluso entre profesionales experimentados. Esta falta de unanimidad en la clasificación puede llevar a discrepancias significativas en el correcto diagnóstico lo que puede afectar directamente los posibles tratamientos y eficacia de la intervención médica, tratamientos inadecuados o tardíos, comprometiendo la visión del paciente y afectando su calidad de vida. Además, puede generar un derroche de recursos médicos al no abordar adecuadamente la progresión de la enfermedad.

La necesidad de establecer métodos de diagnóstico más precisos y estandarizados para DR es evidente, superar las dificultades asociadas con la clasificación de los diferentes grados del padecimiento es crucial para garantizar una atención efectiva y oportuna, por lo tanto, desarrollar una herramienta de diagnóstico de DR tiene beneficios significativos.

1.3. Justificación.

DR representa una complicación oftalmológica crítica en pacientes con diabetes, con impactos significativos en su calidad de vida y salud visual. Sin embargo, el diagnóstico preciso de los diferentes grados de esta afección se enfrenta a desafíos considerablemente

complejos, evidentes por la falta de claridad en la clasificación incluso entre profesionales experimentados lo cual genera discrepancias y puede acarrear consecuencias directas en la efectividad de los tratamientos y en la calidad de vida de los pacientes. Esta falta de precisión diagnóstica no solo compromete la atención médica, sino que también conlleva a posibles tratamientos inadecuados, empeorando el impacto en la visión y la salud general de las personas afectadas.

Ante este escenario, es necesario desarrollar métodos de diagnóstico más precisos y estandarizados de DR. El enfoque planteado se centra en la utilización de técnicas de mejoramiento de imágenes, la extracción de características de textura de Haralick y Patrones Binarios Locales (LBP, son sus siglas en inglés) así como de características profundas, y la clasificación y comparación de estas con Bosques Aleatorios, Árboles Potenciados por Gradiente y Máquinas de Soporte Vectorial, RF, GBT y SVM por sus siglas en inglés, respectivamente.

1.4. Preguntas de investigación.

Las preguntas de investigación a contestar en este trabajo se orientan en la extracción y clasificación de características de textura, así como características profundas de imágenes mejoradas en los espacios de color YCbCr y LAB.

- ¿Cómo pueden las técnicas de análisis de imágenes contribuir a mejorar la precisión y eficacia del diagnóstico de DR?
- ¿Cuál es la influencia de los extractores de características en la capacidad de discernir entre los diferentes grados de DR?
- ¿Cuál es el extractor de características para las imágenes de retina que mejor contribuyen a la clasificación precisa del grado de DR?
- ¿Cuál es el impacto de las técnicas de mejoramiento de imágenes y cómo influyen en la clasificación los espacios de color LAB y YCbCr?
- ¿Qué diferencias significativas se observan en los resultados de clasificación al utilizar diferentes clasificadores RF, GBT y SVM, para evaluar las características extraídas de

las imágenes de retina en el diagnóstico de ausencia o presencia de DR y sus grados de severidad?

- ¿Cuál es el impacto en términos de costo computacional de cada enfoque de extracción de características y clasificación?
- ¿Existe alguna combinación óptima de extractores de características que maximice la precisión en la clasificación de imágenes de retina para cada grado de DR?

1.5. Hipótesis.

Las técnicas de análisis de imágenes basadas en textura, como Haralick y LBP, así como las características profundas, ofrecen un rico conjunto de características que pueden ser relevantes para la clasificación de los grados de DR a partir de imágenes de retina. Se plantea que la aplicación de algoritmos de aprendizaje supervisado, específicamente SVM, GBT y RF, permitirán alcanzar una precisión de al menos el 95 % en la clasificación del conjunto de imágenes planteados aplicando técnicas de mejoramiento de imágenes en los espacios de color LAB y YCbCr para detectar la presencia o ausencia de DR.

Es importante destacar que lo resultados esperados no solo se basan en la efectividad de las técnicas de análisis, mejoramiento, extracción y clasificación de las características, sino también en la capacidad para superar desafíos inherentes a los conjuntos de datos limitados. Incluso ante la pobreza de los conjuntos de imágenes disponibles, con muestras de diferentes dimensiones y desbalance en las clases representativas, se espera que la combinación de las técnicas mencionadas logre una precisión considerable en la detección de los grados de DR.

1.6. Objetivo General.

Implementar un método de clasificación de imágenes de retina, basado en aprendizaje automático, para diagnosticar con precisión la retinopatía diabética.

1.6.1. Objetivos específicos

- Aplicar técnicas de mejoramiento de imágenes en los espacios de color LAB y YCbCr que permitar resaltar algunos detalles que no se observan en el conjunto de datos originales.
- Extraer los vectores de características de textura de Haralick y LBP, así como características profundas, para su posterior clasificación en la detección de DR y sus diversas etapas.
- Entrenar y evaluar el rendimiento de diversos clasificadores en su capacidad para distinguir la ausencia o presencia de DR, así como sus grados de severidad, a través de diversas métricas
- Evaluar el impacto en términos de costo computacional de cada enfoque de extracción de características y su clasificación.

1.7. Metodología.

Para identificar la ausencia o presencia de DR y determinar su grado, se seguirán los pasos que se describen a continuación.

1. Preprocesamiento del conjunto de imágenes.
2. Extracción de caracterısticas con LBP, Haralick y DNNś.
3. Clasificación de los vectores de características extraídas con SVM, GBT y RF.
4. Análisis y comparación de las matrices de confusión obtenidas.

El preprocesamiento del conjunto de imágenes conlleva una serie de pasos propios de las técnicas de mejoramiento de imágenes, tales como el aumento del brillo y contraste en diversos espacios de color.

La extracción de características consiste en obtener vectores representativos de las imágenes, cada extractor proporciona vectores de longitudes distintas, lo cual influye en distintos requerimientos computacionales y tiempo para su extracción.

Para realizar la clasificación de los vectores, se consideraron razones de entrenamiento y prueba para los modelos de clasificación, de tal manera que los vectores de características correspondieran a las mismas imágenes para cada método de extracción utilizado.

Los modelos de clasificación proporcionan la información para realizar el análisis entre los algoritmos y técnicas utilizadas, a través de la matriz de confusión se obtienen las métricas necesarias para conocer el desempeño de los modelos propuestos y realizar las observaciones y comparaciones correspondientes.

Capítulo 2

Marco Teórico.

2.1. Retinopatía Diabética.

DR es la principal complicación ocular de la diabetes de tipo 1 y 2, y afecta aproximandamente entre un 30 % y 40 % de las personas que viven con diabetes. En las últimas décadas se han logrado avances en diversas disciplinas que ayudan a prevenir y disminuir las tasas de pérdida de visión ocasionadas por la Diabetes, además se han mejorado la evaluación y detección de enfermedades oculares, obtención de imágenes y se han desarrollado diversos tratamientos, por ejemplo, los sistemas de clasificación de RD, tales como el Estudio de Tratamiento Temprano de la Retinopatía Diabética (ETDRS) y las escalas de gravedad de la Clasificación Internacional de Retinopatía Diabética (ICDR) que ayudan a pronosticar los riesgos de progresión de esta enfermedad, la fotocoagulación láser panretinal (PRP) y el láser de escaneo de patrones (PASCAL) que ayudan a prevenir la pérdida de visión causada por la Retinopatía Diabética Proliferativa (PDR), la tomografía de coherencia óptica (OCT) para la obtención de imágenes no invasivas en los pacientes, además de tratamientos intravítreos con agentes anti-factor de crecimiento endotelial vascular (anti-VEGF) para edema macular diabético (DME) [7].

Campos de rápido crecimiento como lo son la Inteligencia Artificial (AI), son utilizados para elaboración de diversas aplicaciones y el desarrollo de algoritmos aplicados en la predicción de la diabetes y sus variadas complicaciones, aportanto utilidad en la prevención, diagnóstico y tratamiento general de la diabetes. Aprovechando la potencia de cálculo de

unidades de procesamiento gráfico (GPU), han surgido métodos, modelos y arquitecturas con grandes capacidades en el procesamiento de datos, capaces de aprender e identificar patrones significativos a partir de grandes conjuntos de datos, demostrando gran potencial en el área médica. Se ha demostrado que algunas prácticas basadas en Aprendizaje Profundo (DL, por sus siglas en inglés) alcanzan el mismo nivel que los expertos clínicos, a pesar de esto, la mayoría de estas herramientas no han sido evaluadas en estudios clínicos controlados debido a que su aplicación en entornos reales se obstaculiza por varias cuestiones prácticas, como la transparencia de algoritmos, el intercambio de registros, la privacidad y la normalización de los datos de los pacientes, además, los algoritmos y modelos grandes cantidades de datos médicos para realizar estudios a gran escala.

2.2. Imagen Digital.

Una imagen digital es una representación bidimensional de una escena tridimensional definida por su resolución espacial, su profundidad de color y otras características técnicas. Esencialmente, se trata de una matriz en la que a cada posición le corresponde un cierto color o nivel de gris, tal como se muestra en la figura 2.1.

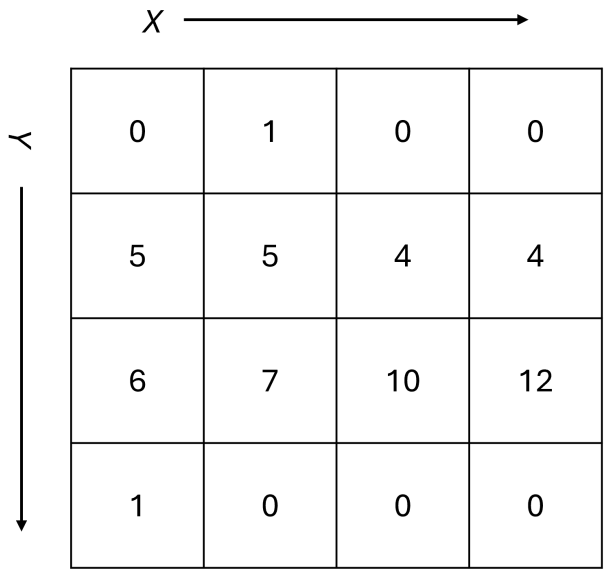

Figura 2.1: Imagen de 16 pixeles.

Matemáticamente puede se define como una función bidimensional de la intensidad de

la luz $f(x, y)$, donde x e y representan las coordenadas espaciales y el valor de f, en un punto cualquiera, es proporcional al brillo o nivel de gris en dicho punto de la imagen; puede considerarse como una matriz de M filas y N columnas, cuyos índices de fila y columna identifican un punto en la imagen y el valor correspondiente del pixel [8].

2.3. Espacios de Color.

La definición de color puede variar de acuerdo al contexto en cuestión, desde un punto de vista científico se puede interpretar como una respuesta del sistema visual humano que se genera en el cerebro a partir de las ondas electromagnéticas reflejadas por los objetos y que son captadas por los ojos. Cuando un haz de luz visible incide sobre un objeto, ésta se transmite, se absorbe o se refleja por éste. La luz reflejada por el objeto es la que se asocia con el color de éste y el cual nos permite identificarlo y clasificarlo. La figura 2.2 pesenta el espectro electromagnético mostrando las diferentes longitudes de onda, así como el espectro visible.

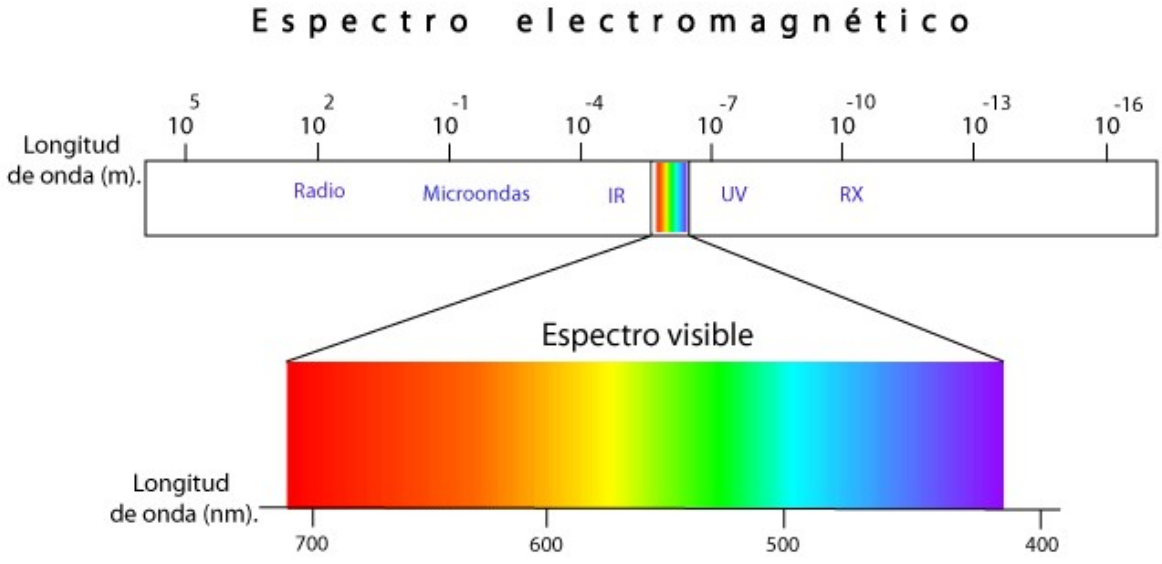

Figura 2.2: Espectro eletromagnético y espectro visible.

Cuando se trata el color desde un punto de vista de procesamiento digital de imágenes,

es necesario recurrir a los denominados espacios de color, las cuales permiten describir el color en diferentes sistemas de coordenadas. Matemáticamente, un espacio de color puede ser definido por una base de N componentes vectoriales cuya combinación lineal o no, genera todos los posibles colores que pueden ser representados por dicho espacio de color [9].

Existen diversos espacios de color, cada uno ofrece diferentes posibilidades y opciones, así que son utilizados de acuerdo a los resultados buscados o metodologías específicas empleadas. Se dividen en dos tipos: los orientados a dispositivos, que se basan en la fisiología de la retina ocular y los orientados a reproducir la percepción humana, basados en fenómenos perceptuales tales como la luminosidad, tono y la saturación.

2.3.1. Espacio de color RGB

RGB es un modelo de color aditivo en el que los colores se crean a partir de la combinación lineal de los tres colores primarios rojo (R), verde (G) y azul (B). Este espacio de color se utiliza para desplegar las imágenes en dispositivos electrónicos.

Geométricamente, el color está compuesto por la combinación lineal de sus tres componentes, ver figura 2.3. Cada eje del espacio de colo RGB representa la intensidad del componente RGB del píxel en esa dirección. Los vértices de cada eje, es decir, (255,0,0) ó normalizados de 0 a 1 como (1,0,0) es totalmente rojo, (0,255,0) ó (0,1,0) es totalmente verde y (0,0,255) ó (0,0,1) es totalmente azul.

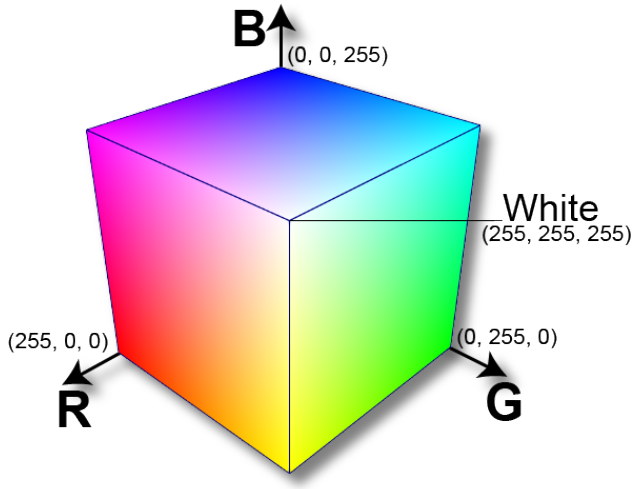

Figura 2.3: Espacio de color RGB.

La imagen en este espacio de color está definida como sigue:

$$X_{ij} = [r_{ij}, g_{ij}, b_{ij}] \tag{2.1}$$

Donde ij indica la coordenada del píxel correspondiente, r_{ij} es el valor del componente en R, g_{ij} es el valor del componente en G y b_{ij} es el valor del componente en B [10]. Si bien el espacio de color RGB es el más común en dispositivos electrónicos, diversos espacios de color se definen con respecto a éste a través de transformaciones lineales o no. Es importante señalar que cada espacio de color tiene sus propias características y ventajas que lo hacen ideal para determinadas aplicaciones.

2.3.2. Espacio de color LAB

El espacio de color L*a*b* o CIELAB es un estándar internacional para la medición de color desarrollado por la Comisión Internacional de Iluminación (CIE) en 1976. Los asteriscos (*) que siguen cada letra forman parte del nombre, donde, L*, a* y b* representan los parámetros o componentes que integran este espacio de color [11].

- L*: Es el componente de luminancia o luminosidad, donde los píxeles pueden tomar valores desde 0, que representa negro, hasta 100, que representa blanco. Este componente mide el brillo o la intensidad del color.

- a*: Indica la posición entre rojo y verde. Los píxeles pueden tomar valores desde -128 hasta 127 (aunque comúnmente se menciona de -120 a 120). Los valores negativos corresponden al verde y los positivos al rojo.

- b*: Indica la posición entre azul y amarillo, con valores desde -128 hasta 127 (similarmente, a veces mencionado de -120 a 120). Los valores negativos corresponden al azul y los positivos al amarillo.

Los parámetros a* y b* son los componentes de crominancia o cromáticos del modelo, ver figura 2.4.

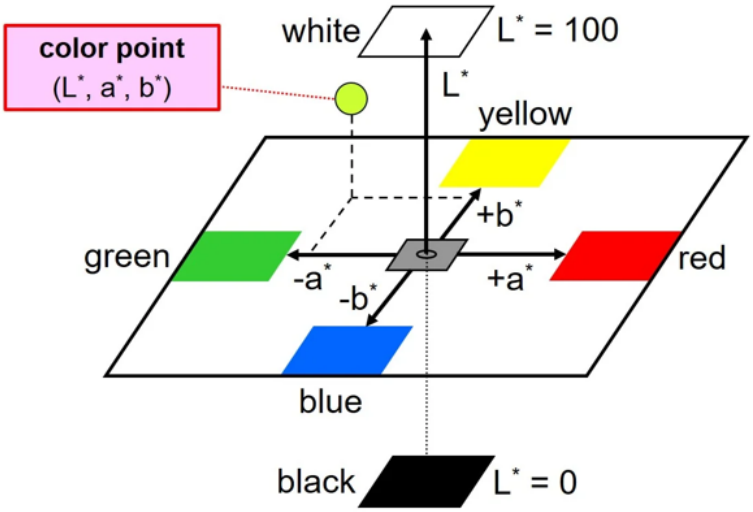

Figura 2.4: Espacio de color LAB.

El espacio de color L*a*b* es utilizado para describir todos los colores que puede percibir el ojo humano y puede ser representado adecuadamente en un espacio tridimensional. Es ampliamente utilizado en contextos donde es importante tener una representación perceptualmente uniforme de los colores.

2.3.3. Espacio de color YCbCr

El espacio de color YCbCr es un estándar ampliamente utilizado en sistemas de video y compresión de imágenes, como JPEG y MPEG. Fue desarrollado por la Unión Internacional de Telecomunicaciones (UIT) [12]. Los componentes Y, Cb y Cr representan los parámetros o componentes que integran este espacio de color:

- Y: Es el componente de luminancia o brillo, que representa la intensidad de luz en la imagen Y se calcula como una combinación ponderada de los componentes RGB.
- Cb: Es el componente de crominancia que indica la diferencia entre el componente azul y el valor de luminancia (Y). Especifica cuánto azul debe ser añadido o sustraído de la luminancia para obtener el color correcto.
- Cr: Es el componente de crominancia que indica la diferencia entre el componente rojo y el valor de luminancia (Y). Especifica cuánto rojo debe ser añadido o sustraído de la

luminancia para obtener el color correcto.

Las ecuaciones de transformación del espacio de color RGB al espacio de color YCbCr son:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0{,}2568 & 0{,}5041 & 0{,}0979 \\ -0{,}1482 & -0{,}2910 & 0{,}4392 \\ 0{,}4392 & -0{,}3678 & -0{,}0714 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \tag{2.2}$$

El espacio de color YCbCr se utiliza para separar la información de luminancia de la información de crominancia, lo que permite una mayor eficiencia en la compresión de imágenes y videos. Esto se debe a que el ojo humano es más sensible a los cambios en la luminancia que a los cambios en la crominancia, permitiendo que los componentes de crominancia (Cb y Cr) puedan ser subsampleados con poca pérdida perceptible de calidad [13].

El espacio de color YCbCr facilita la manipulación y transmisión de señales de video, ya que permite una representación más eficiente y compatible con los sistemas de transmisión y almacenamiento de video. Es una herramienta esencial en la ingeniería de video y procesamiento de imágenes.

2.4. Filtros.

El filtrado espacial de las imágenes se utiliza para modificar el valor de intensidad de un píxel considerando los valores de intensidad de sus píxeles vecinos. Esta vecindad, también conocida como kernel, determina el tamaño y la forma de la región que se considera para el cálculo del nuevo valor del píxel. Al modificar el valor de intensidad de un píxel en función de sus vecinos, se pueden lograr diversos efectos en la imagen, como suavizado, realce de bordes, detección de características, entre otros. Cabe mencionar que el filtrado espacial tiene una amplia gama de aplicaciones en el procesamiento de imágenes, desde la reducción de ruido hasta la segmentación de imágenes y la detección de objetos [14].

El proceso de filtrado consiste en la convolución de una imagen con una matriz de filtrado, también conocida como máscara o kernel (ecuación 2.3), de tamaño variable $N \times N$ de tamaño impar. El proceso de la convolución de la imagen con el kernel genera un nuevo valor que

combina el valor original del píxel y los valores de los píxeles vecinos.

$$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{2.3}$$

Existen diversas técnicas de filtrado de imágenes, cada una ofreciendo aplicaciones y características específicas:

- Suavizado y Eliminación de Ruido: El kernel de estas máscaras de convolución suman a uno; esto permite que ciertos tipos de ruidos presentes en las imágenes se minimicen o eliminen; sin embargo, al aumentar el tamaño del filtro, los bordes de la imagen tienen a ser más burdos. Entre los filtros que se utilizan están el promediador y sus variantes gaussianas.

- Realce de Bordes y Detalles: Estos filtros enfatizan las altas frecuencias presentes en las imágenes, que corresponden a cambios rápidos de intensidad de los píxeles, como bordes y detalles finos, su kernel suma a cero. Ejemplos incluyen los filtros Laplaciano, Sobel y Prewitt.

Como previamente se mencionó, el filtrado se aplica en diversas situaciones para resolver diversos problemas en el procesamiento de imágenes, desde mejorar la calidad visual hasta facilitar el análisis y la detección de características importantes.

2.5. CLAHE.

La Ecualización Adaptativa de Histograma Limitada por el Contraste (CLAHE, por sus siglas en inglés) es una técnica utilizada para mejorar el contraste de una imagen. Ha demostrado buenos resultados en imágenes médicas para la detección de microaneurismas y segmentación de vasos sanguíneos retinianos [15]. Esta técnica permite una mejora del contraste que es más específica a las características locales de la imagen, resultando en una imagen mejorada en términos de detalles y visibilidad, tal como se muestra en la figura 2.5.

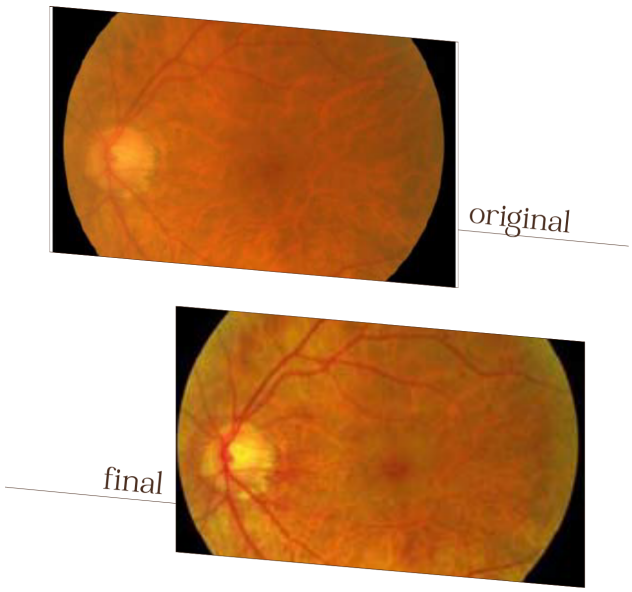

Figura 2.5: Imagen de fondo de retina antes de después de aplicar CLAHE.

CLAHE trabaja en pequeños bloques o regiones de la imagen, lo que permite una mejora del contraste más localizada y controlada de la siguiente forma [16]:

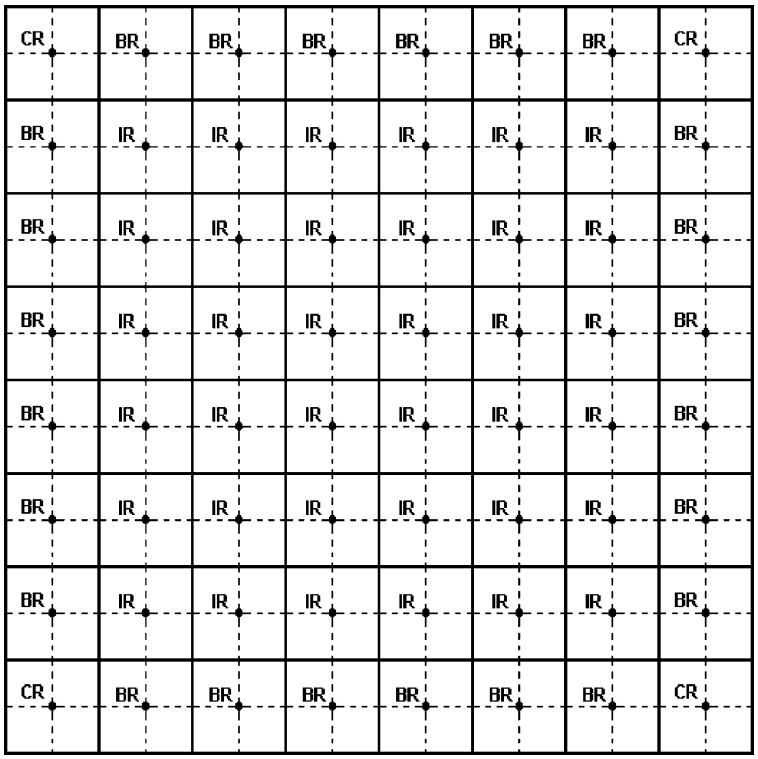

Figura 2.6: Regiones en que se divide una imagen por medio de CLAHE.

1. División en bloques: La imagen se divide en 64 bloques no superpuestos llamados

regiones o tiles (figura 2.6).

2. Ecualización de histograma por bloque: Se realiza la ecualización del histograma en cada bloque de manera independiente.

3. Limitación del contraste: Se aplica una limitación al contraste para evitar la amplificación excesiva del ruido.

4. Interpolación de los bloques: Los bloques procesados son combinados suavemente para eliminar las discontinuidades visibles entre ellos.

2.6. Textura.

La textura puede describirse como una estructura formada por muchos elementos o patrones similares que están organizados de manera ordenada. Ningún elemento individual destaca de manera significativa por encima de los demás, lo que resulta en una percepción visual coherente y uniforme para el observador. Esta organización y repetición de elementos crean una impresión general unificada de la superficie o área en cuestión [17].

El análisis de texturas tiene como objetivo encontrar una forma única de representar las características implícitas de las texturas y representarlas de una forma única y simple de tal manera que puedan ser aprovechadas para una clasificación sólida y precisa de objetos.

La textura ocupa un papel importante en el análisis de imágenes y el reconocimiento de patrones. Las características de textura se calculan a través de la distribución estadística de las combinaciones de intensidades en posiciones específicas relativas entre sí en la imagen. Se considera una propiedad innata de prácticamente todas las superficies la cual ofrece información importante sobre la disposición estructural de las superficies y su relación con el entorno, y se categoriza como: fina, gruesa, lisa, mollar, irregular o lineada. Para el ojo humano es muy sencillo reconocerla y describirla en términos empíricos, pero para procesamiento digital ha sido extremadamente reacia a una definición y análisis preciso. Dado que las propiedades texturales de las imágenes contienen información útil, es importante conocer cómo extraer sus características.

2.6.1. Patrones Locales Binarios (LBP).

El término Patrones Binarios Locales se refiere a una técnica descriptiva utilizada en la extracción de características de textura. Esta técnica se basa en comparar el valor de intensidad de un píxel con los valores de intensidad de píxeles vecinos y asignar un valor binario a cada comparación. Es no paramétrica, computacionalmente simple y muy eficiente para el análisis de texturas. LBP se define como una medida de textura invariante en escala de grises y se deriva de un valor general de textura en una vecindad local de píxeles. Para cada píxel de una imagen, se genera un código binario al comparar su valor con el valor del píxel central. Si el valor del píxel vecino es igual o mayor al valor del píxel central, se asigna un 1, de lo contrario, se asigna un 0. Los valores binarios obtenidos se concatenan para formar un número binario (en vecindades de 8 vecinos, se obtiene un número binario de 8 bits). Este número binario se convierte en un valor decimal que representa el valor LBP del píxel central, tal como se muestra en la figura 2.7. Finalmente, se construye un histograma de los valores LBP para una región o para toda la imagen, que se utiliza como descriptor de textura.

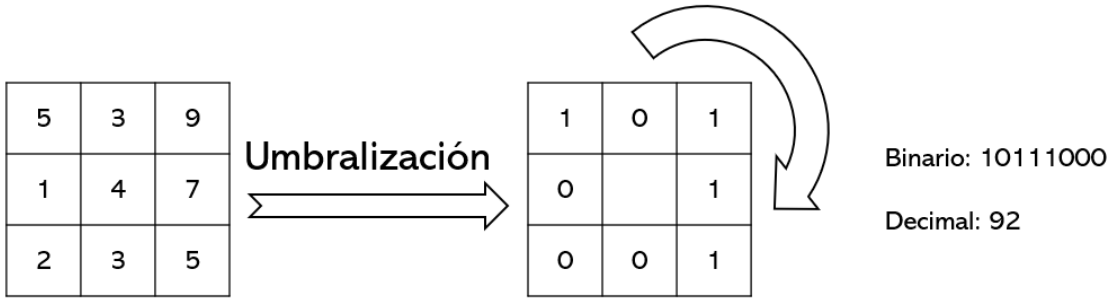

Figura 2.7: Operador básico LBP.

La versión básica del operador LBP considera vecindades de 3 x 3, es decir, los 8 vecinos de un píxel. Sin embargo, la definición se ha ampliado para incluir todos los vecindarios circulares con cualquier número de píxeles y se ha convertido en una medida realmente potente de la textura de imágenes, demostrando excelentes resultados en términos de precisión y complejidad computacional en muchos escenarios, tales como la inspección visual, la recuperación de imágenes, el análisis de imágenes biomédicas, análisis de imágenes faciales, análisis de movimiento, modelado de entornos y análisis de escenas exteriores [18].

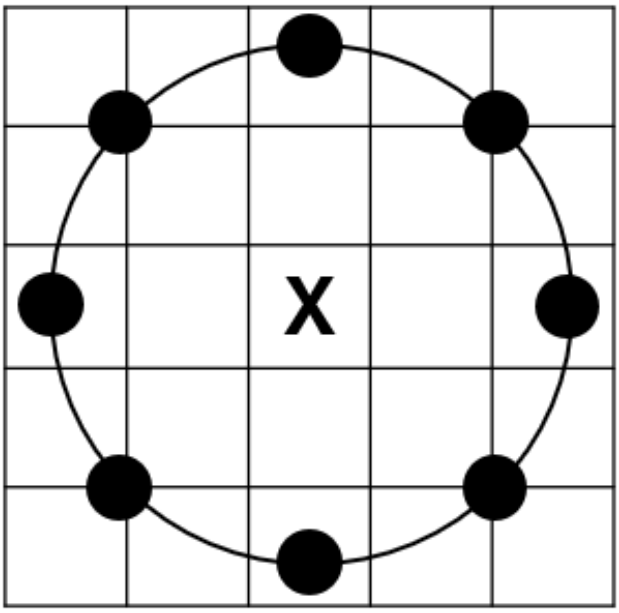

Figura 2.8: Vecindario circular (P,R)$\Longrightarrow$(8,2) para vecindarios mayores de 3 x 3, imagen tomada de [19].

2.6.2. Características de Haralick.

La Matriz de Co-ocurrencia de Niveles de Gris (GLCM) es una técnica ampliamente utilizada para el análisis de texturas de segundo grado. GLCM es un histograma bidimensional de la frecuencia con la que aparecen en una imagen diferentes combinaciones de los valores de los píxeles (niveles de gris), indica la frecuencia de aparición de pares de píxeles. Se define con respecto a una fila y una columna determinadas, un elemento P(i,j) denotado por C_{ij} representa el número de veces que un punto que tiene un nivel de gris j aparece en relación con un punto que tiene un nivel de gris i, (ver figura 2.9) [20].

Una GLCM es una matriz en la que el número de filas y columnas es igual al número de niveles de gris (G) de la imagen. El elemento de de la matriz $P(i, j|\Delta x, \Delta y)$ es la frecuencia relativa con la que dos píxeles contiguos aparecen dentro de un determinado vecindario, uno con intensidad i y el otro con intensidad j.

También se puede decir que el elemento de la matriz $P(i, j|d, \theta)$ contiene los valores de probabilidad entre los niveles de gris i y j a una distancia de desplazamiento particular d y en un ángulo específico θ.

Imagen

20	5	10	5	5
5	20	10	5	20
10	5	5	10	5
2	20	20	10	5
10	5	10	10	5

GLCM de imagen

i/j	5	10	20
5	2	3	3
10	7	1	0
20	1	2	1

Figura 2.9: GLCM de imagen, un píxel de referencia y su vecino inmediato.

Ahora bien, es necesario considerar cada combinación de $(\Delta x, \Delta y)$ ó (d, θ) de tal manera que se contemplen todas las distribuciones de probabilidad, una GLCM completa que incluya todas direcciones, es decir, que se cubran las cuatro principales direcciones (horizontal, diagonal ascendente, vertical y diagonal ascendente) en la que los píxeles de una imagen pueden relacionarse, esto es, $\theta = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$ ó 0°, 45°, 90° y 135°, ver figura 2.10.

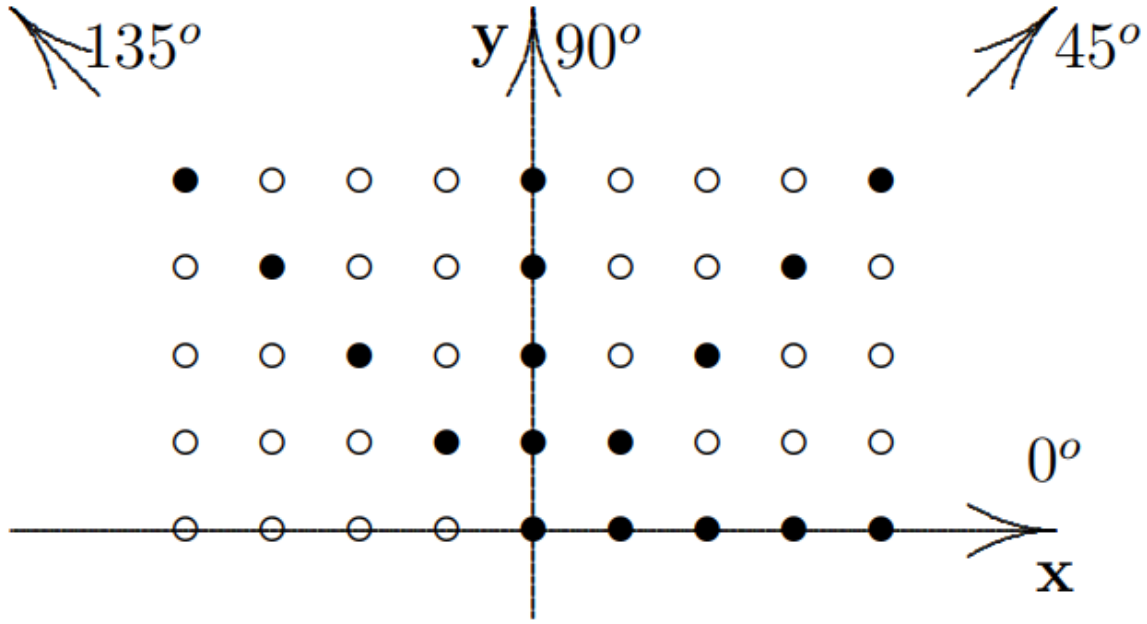

Figura 2.10: Direcciones para GLCM completa.

Existen relaciones sencillas entre ciertos pares de las distribuciones de probabilidad esti-

madas $P(d,\theta)$. $P^t(d,\theta)$ es la transposición de la matriz $P(d,\theta)$, entonces [21]:

$$P(d,0°) = P^t(d,180°) \tag{2.4}$$

$$P(d,45°) = P^t(d,225°) \tag{2.5}$$

$$P(d,90°) = P^t(d,270°) \tag{2.6}$$

$$P(d,135°) = P^t(d,315°) \tag{2.7}$$

Robert M. Haralick definió características texturales aplicables a muchos tipos de datos de imágenes, a partir de GLCM para cuantificar y analizar textura; propuso 14 características diseñadas para cuantificar diferentes aspectos de la textura de una imagen [22]:

Notación

$p(i,j)$ es el valor de la GLCM en la posición (i,j).

$p_x(i)$ es la suma de la i-ésima fila de la GLCM.

$p_y(j)$ es la suma de la i-ésima columna de la GLCM.

N_g es número de los difeterentes niveles de gris cuantificados en la imagen.

$$p_x(i) = \sum_{j=1}^{N_g} p(i,j) \tag{2.8}$$

$$p_y(j) = \sum_{i=1}^{N_g} p(i,j) \tag{2.9}$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ i+j=k}}^{N_g}, \qquad k = 2,3,\cdots,2N_g. \tag{2.10}$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ |i-j|=k}}^{N_g}, \qquad k = 0,1,\cdots,N_g-1. \tag{2.11}$$

Características Texturales

1. *Segundo Momento Angular o Energía:*

$$f_1 = \sum_i \sum_j p(i,j)^2. \tag{2.12}$$

2. *Constraste:*

$$f_2 = \sum_{n=0}^{N_{g-1}} n^2 \left\{ \sum_{\substack{i=1 \\ |i-j|=n}}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\}. \tag{2.13}$$

3. *Correlación:*

$$f_3 = \frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}. \tag{2.14}$$

   donde μ_x, μ_y, σ_x y σ_y son la media y desviación estándar de p_x y p_y.

4. *Suma de cuadrados o Varianza:*

$$f_4 = \sum_i \sum_j (i-\mu)^2 p(i,j). \tag{2.15}$$

5. *Momento Diferencial Inverso u Homogeneidad:*

$$f_5 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j). \tag{2.16}$$

6. *Suma de promedio:*

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i). \tag{2.17}$$

7. *Suma de Varianza:*

$$f_7 = \sum_{i=2}^{2N_g} (i - f_8)^2 p_{x+y}(i). \tag{2.18}$$

8. *Suma de Entropía:*

$$f_8 = -\sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\}. \tag{2.19}$$

Dado que algunas de las probabilidades pueden ser 0 y log(0) no está definido, es recomendable utilizar el término $\log(p + \epsilon)$ (ϵ es una constante positiva pequeña) en lugar de $log(p)$ en los cálculos de entropía.

9. *Entropía:*

$$f_9 = -\sum_i \sum_j p(i,j) \log(p(i,j)). \tag{2.20}$$

10. *Diferencia de Varianza:*

$$f_{10} = \text{varianza de } p_{x-y}. \tag{2.21}$$

11. *Diferencia de Entropía:*

$$f_{11} = -\sum_{i=0}^{N_{g-1}} p_{x-y}(i) \log\{p_{x-y}(i)\} \tag{2.22}$$

12. *Medidas de Correlación 1:*

$$f_{12} = \frac{HXY - HXY1}{max\{HX, HY\}} \tag{2.23}$$

13. *Medidas de Correlación 2:*

$$f_{13} = \sqrt{1 - e^{[-2,0(HXY2-HXY)]}} \tag{2.24}$$

$$HXY = -\sum_i \sum_j p(i,j) \log(p(i,j)) \tag{2.25}$$

donde HX y HY se refiere a la Entropía de p_x y p_y, y

$$HXY1 = -\sum_i \sum_j p(i,j) \log\{p_x(i)p_y(j)\} \tag{2.26}$$

$$HXY2 = -\sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\}. \tag{2.27}$$

14. *Coeficiente máximo de Correlación:*

$$f_{14} = \sqrt{\text{Segundo eigenvalor mayor de } \mathcal{Q}} \tag{2.28}$$

donde

$$\mathcal{Q}(i,j) = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}. \tag{2.29}$$

Estas medidas de coorrelación tienen algunas propiedades deseables que no aparecen en la medida de coorrelación f_3.

2.7. Redes Neuronales Profundas (DNN).

Una Red Neuronal Artificial (ANN) es un modelo computacional inspirado en las redes neuronales biológicas, compuesto por unidades o neuronas, también conocidas como perceptrones, cuya función se define por la conexión entre ellas. Estas unidades no suelen estar conectadas directamente entre sí, sino que están organizadas en diferentes capas: una capa de entrada, una capa de salida y una o más capas ocultas. Las neuronas en la capa de entrada reciben y procesan las señales de entrada, y envían una señal de salida a otras neuronas en la red. Cada neurona puede conectarse a otras y posee una función de activación y una función

umbral, las cuales pueden ser funciones continuas, lineales o no lineales. La señal que pasa por una neurona se transforma mediante pesos que modifican las funciones y, por lo tanto, alteran la señal de salida que llega a la siguiente neurona en la red, la figura 2.11 muestra un esquema general de una ANN.

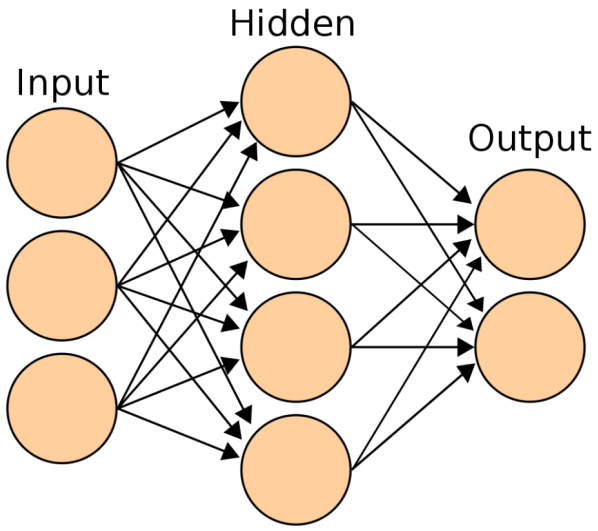

Figura 2.11: Representación general de una ANN.

Las redes neuronales que emplean una gran cantidad de capas ocultas se denominan profundas (DNN). Las DNN no necesitan una fase previa de extracción de características para su entrenamiento, ya que los procesos de convolución involucrados en las capas ocultas realizan este proceso junto con otras operaciones como lo son el pooling y la rectificación de las salidas. Cabe mencionar que, en particular, las convolucionales (CNN) aprenden automáticamente jerarquías de características más relevantes de las imágenes a través de múltiples capas de convolución. Las capas superiores de las CNN contienen información semántica y describen las características globales de las imágenes, mientras que las capas intermedias describen características locales y las capas inferiores contienen información de bajo nivel para la descripción de la textura, los bordes, etc. Las características aprendidas automáticamente por las DNN se denominan características profundas. Debido a la forma jerárquica en que aprenden características, las DNN se utilizan como extractores de características a cualquier grado de profundidad.

Entrenar CNN profundas desde cero para la clasificación de imágenes requiere una gran cantidad de datos de entrenamiento etiquetados, amplios recursos computacionales y de memoria y una gran cantidad de experiencia para garantizar una convergencia adecuada, lo que resulta en un proceso que consume demasiado tiempo.

En este trabajo de investigación consideramos como extractores de características profundas las arquitecturas pre-entrenadas Ademxapp Model A y DenseNet-121.

2.7.1. DenseNet-121

La arquitectura DenseNet ha sido utilizada en diversas áreas debido a su gran rendimiento, se ha empleado para el diagnóstico de varias enfermedades por ejemplo para la detección de Alzheimer, COVID-19, enfermedades de plantas, incluso predecir condiciones climáticas [23]. Estas redes convolucionales profundas, conectan cada capa con todas las demás de forma feed-forward, donde cada capa utiliza como entrada los mapas de características de todas las capas anteriores, y sus propios mapas de características se utilizan como entrada en todas las capas posteriores. DenseNet-121 logró una exactitud de 75 % en el conjunto de validación original de ILSVR 2012. La figura 2.12 muestra una represetación de la estructura de Dense-Net 121.

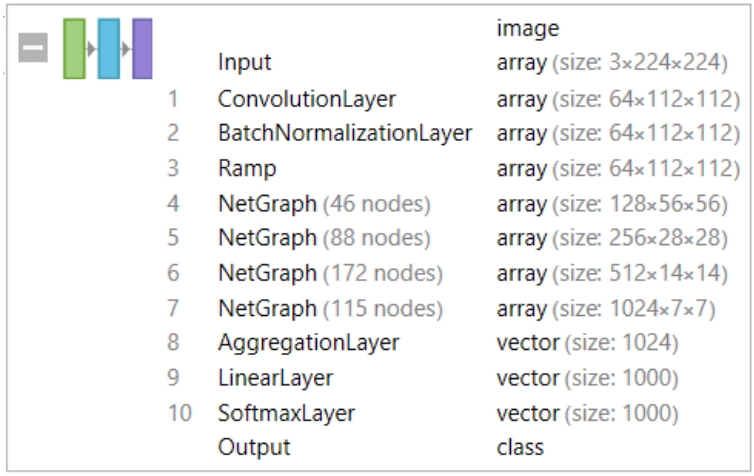

		image
	Input	array (size: 3×224×224)
1	ConvolutionLayer	array (size: 64×112×112)
2	BatchNormalizationLayer	array (size: 64×112×112)
3	Ramp	array (size: 64×112×112)
4	NetGraph (46 nodes)	array (size: 128×56×56)
5	NetGraph (88 nodes)	array (size: 256×28×28)
6	NetGraph (172 nodes)	array (size: 512×14×14)
7	NetGraph (115 nodes)	array (size: 1024×7×7)
8	AggregationLayer	vector (size: 1024)
9	LinearLayer	vector (size: 1000)
10	SoftmaxLayer	vector (size: 1000)
	Output	class

Figura 2.12: Estructura Dense-Net 121, 427 capas, imagen tomada de [24].

2.7.2. Ademxapp Model A

Es una CNN presentada por investigadores de la Universidad de Adelaida en 2016, es una arquitectura de redes residuales, menos profunda, que supera el rendimiento de modelos más profundos como ResNet-200 en el conjunto de entrenamiento de ImageNet [25]. Crearon 6 modelos, A, B, C, D, E y F de esta arquitectura, con diferentes tamaños de entrada, diferentes kernels y los evaluaron con el dataset de clasificación ImageNet Large Scale Visual Recognition Challenge (ILSVR) 2012 frente a otras arquitecturas, VGG16, ResNet-50, ResNet101, ResNet1-52, ResNet200, Inception demostrando mayor eficiencia desde el punto de vista espacial y mejor rendimiento no solo en clasificación del conjunto de datos sino también en la segmentación de las imágenes. ILSVR es un conjunto de imágenes que consta de 1.2 millones de imágenes de entrenamiento con 1000 clases de objetos. La figura 2.13 muestra una representación de la estructura del modelo A.

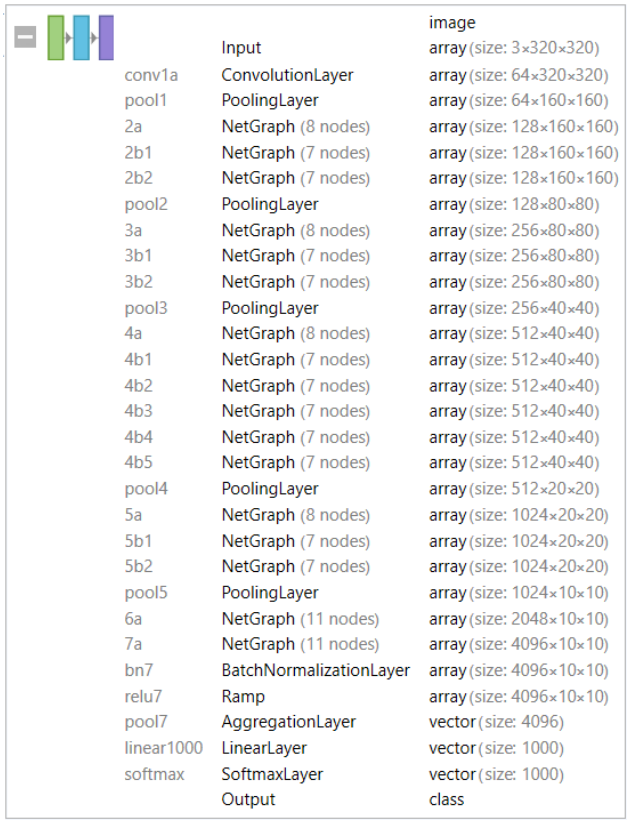

		image
	Input	array (size: 3×320×320)
conv1a	ConvolutionLayer	array (size: 64×320×320)
pool1	PoolingLayer	array (size: 64×160×160)
2a	NetGraph (8 nodes)	array (size: 128×160×160)
2b1	NetGraph (7 nodes)	array (size: 128×160×160)
2b2	NetGraph (7 nodes)	array (size: 128×160×160)
pool2	PoolingLayer	array (size: 128×80×80)
3a	NetGraph (8 nodes)	array (size: 256×80×80)
3b1	NetGraph (7 nodes)	array (size: 256×80×80)
3b2	NetGraph (7 nodes)	array (size: 256×80×80)
pool3	PoolingLayer	array (size: 256×40×40)
4a	NetGraph (8 nodes)	array (size: 512×40×40)
4b1	NetGraph (7 nodes)	array (size: 512×40×40)
4b2	NetGraph (7 nodes)	array (size: 512×40×40)
4b3	NetGraph (7 nodes)	array (size: 512×40×40)
4b4	NetGraph (7 nodes)	array (size: 512×40×40)
4b5	NetGraph (7 nodes)	array (size: 512×40×40)
pool4	PoolingLayer	array (size: 512×20×20)
5a	NetGraph (8 nodes)	array (size: 1024×20×20)
5b1	NetGraph (7 nodes)	array (size: 1024×20×20)
5b2	NetGraph (7 nodes)	array (size: 1024×20×20)
pool5	PoolingLayer	array (size: 1024×10×10)
6a	NetGraph (11 nodes)	array (size: 2048×10×10)
7a	NetGraph (11 nodes)	array (size: 4096×10×10)
bn7	BatchNormalizationLayer	array (size: 4096×10×10)
relu7	Ramp	array (size: 4096×10×10)
pool7	AggregationLayer	vector (size: 4096)
linear1000	LinearLayer	vector (size: 1000)
softmax	SoftmaxLayer	vector (size: 1000)
	Output	class

Figura 2.13: Estructura Ademxapp Model A, 142 capas, imagen tomada de [24].

2.8. Máquina de Vector Soporte (SVM).

Máquina de Vector de Soporte es un algoritmo propuesto en los años 90 para resolver problemas de clasificación binaria, aunque actualmente se utiliza para problemas de regresión, agrupamiento y clasificación multiclase en diversas áreas, tales como visión, artificial, reconocimiento de objetos, procesamiento de lenguaje natural, entre otros [26]. La idea principal de SVM es encontrar separadores lineales, mejor conocidos como hiperplanos, que dividan las diferentes clases en el conjunto de datos en el espacio de características. Los puntos que se encuentran más cerca del hiperplano se denominan vectores de soporte, de ahí el nombre de SVM, tal como se observa en la imagen 2.14.

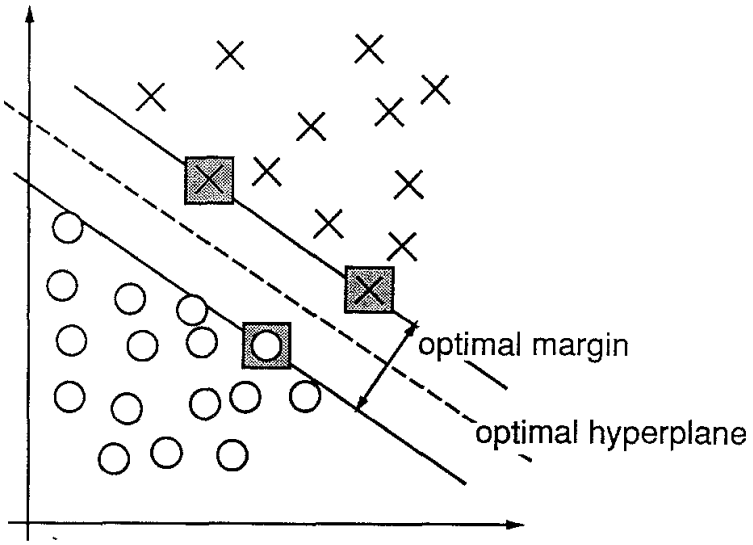

Figura 2.14: Ejemplo de un problema separable en dos dimensiones. Los vectores de soporte están marcados con cuadros color gris y definen el margen óptimo de separación entre clases. El hiperplano se marca con la línea punteada.

2.9. Random Forest (RF).

Bosque Aleatorio es un algoritmo de aprendizaje profundo muy utilizado como método de clasificación y regresión. Se basa en la predicción conjunta de varios árboles de decisión creados aleatoriamente completamente independientes entre sí [27]; ha demostrado excelentes resultados en entornos en donde el número de variables es significativamente mayor que el número de observaciones, ver Figura 2.15.

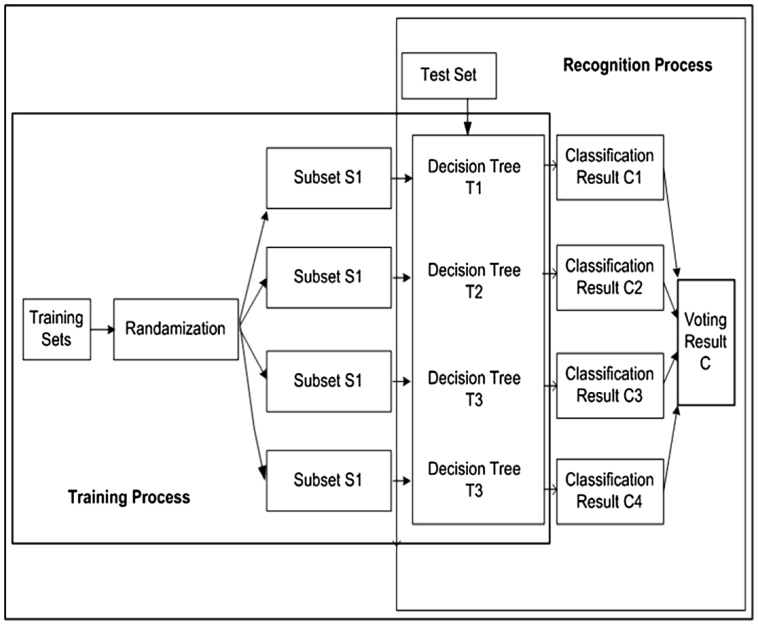

Figura 2.15: Marco conceptual de Random Forest.

El proceso implica introducir una muestra de prueba en los clasificadores, y la etiqueta de clase se decide considerando una votación mayoritaria de las clasificaciones individuales de cada árbol de decisión.

RF es eficaz para resolver el problema de sobreajuste y tiene amplias aplicaciones en diversas áreas, como la clasificación de textos y la clasificación de imágenes, entre otros. Además, es lo suficientemente versátil para aplicarse a problemas a gran escala y devuelve la importancia de las varibles consideradas de un conjunto de datos.

2.10. Gradient Boosted Trees (GBT).

Los árboles impulsados por gradiente son una técnica de aprendizaje automático utilizada en muchas aplicaciones, como clasificación multiclase, modelado de procesos, transferencias de aprendizaje y predicción. Este enfoque combina múltiples modelos más simples, normalmente árboles de decisión, para mejorar la precisión y la generalización del modelo final, ver figura 2.16.

La idea principal de este algoritmo es entrenar árboles de decisión secuencialmente, de tal forma que cada nuevo árbol se enfoca en corregir los errores del modelo anterior, las

funciones de pérdida aplicadas pueden ser arbitrarias y dependen del investigador o la tarea específica en cada situación [28].

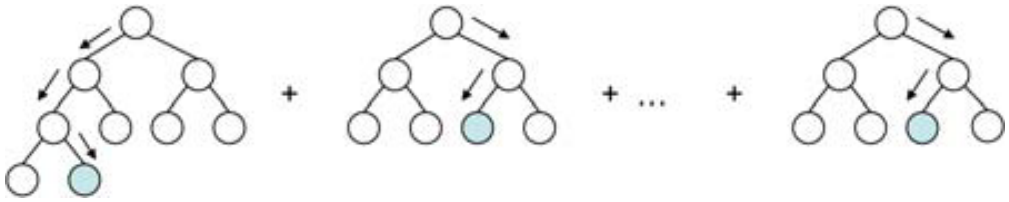

Figura 2.16: Conjunto de árboles de decisión impulsados por grandiente.

2.11. Trabajos Relacionados.

La Inteligencia Artificial (AI) ha logrado avances significativos en años recientes con una notable evolución en las técnicas de aprendizaje automático (ML). El aprendizaje profundo ha demostrado notables resultados en el reconocimiento de imágenes, reconocimiento del habla y procesamiento del lenguaje natural de tal forma que se ha ido adoptando en muchos dominios, incluidos los sociales, telecomunicaciones, ciberseguriad y la medicina. Para el análisis de imágenes médicas a logrado resultados importantes en varias especialidades; particularmente en oftalmología, se han propuesto una amplia gama de algoritmos y técnicas para la detección y diagóstico de DR [29], algunas de estas tecnologías disponibles comercialmente.

Entre las arquitecturas de aprendizaje profundo utilizadas para extraer características profundas y/o clasificar imágenes con transferencia de conocimiento de encuentran: AlexNet, Visual Geometry Group Network (VGG-16 y VGG-19), GoogleNet, Residual Network (ResNet), Densley Connected Network (DenseNet), Inception Convolutional Recurrent Neural Network, Generative Adversarial Network (GAN), Autoencoder, Restricted Boltzmann Machine (RBM), Long Short-Term Memory (LSTM), Deep Reinforcement Learning (DRL), entre otras [30].

Otras técnicas de extracción de características consisten en análsis de textura que permiten los patrones distintivos característicos a cada imagen, una vez extraídas las características de textura se entrena un clasificador con la finalidad de detectar DR [31].

De igual manera, existen técnicas para identificar y extraer patrones o características únicas de una imágen, estas técnicas se basan en encontrar y extraer característcas morfológicas únicas que describen las lesiones en el caso de DR [32].

Para este tipo de análisis existen conjuntos de imágenes autorizados estandarizados para su utilización, muchos de ellos de acceso libre, clasificados y autorizados por instituciones expertas en esta condición médica [33].

Capítulo 3

Desarrollo.

3.1. Conjunto de datos.

El conjunto de datos empleado en esta investigación corresponde al APTOS 2019 Blindness Detection, el cual fue dado a conocer por la Asia Pacific Tele-Ophthalmology Society. Este conjunto de datos es de acceso abierto y tiene como propósito el desarrollo de modelos de aprendizaje automático orientados a la detección temprana de retinopatía diabética (DR) en áreas rurales [34]. El conjunto de datos consta de un un total 3662 de imágenes de retina capturadas en diversas condiciones de iluminación y, además, con diferentes resoluciones espaciales. Tal como se detalla en la Tabla 3.1, las imágenes de DR están clasificadas en cinco clases: No DR, DR leve, DR moderada, DR grave y DR proliferativa. En la Tabla 3.1 se muestra la cantidad y proporción de imágenes correspondientes a cada clase. Como se puede observar, el conjunto de datos está desbalanceado, es decir, la cantidad de imágenes o muestras para cada clase difiere. La clase mayoritaria, No DR, corresponde a aproximadamente la mitad del total de muestras del conjunto de imágenes. Mientras que, la clase minoritaria, asociada a DR grave, solamente consta de 193 muestras. Esta desproporción en la distribución de las muestras entre las diferentes clases representa un desafío importante en el desarrollo de modelos de aprendizaje automático robustos y precisos para la detección temprana de DR.

Como se mencionó, la resolución espacial de las imágenes del APTOS 2019 varían significativamente, desde los 474×358 hasta los 4288×2848 píxeles. Como se muestra en la Tabla

3.2, existen un total de 17 diferentes resoluciones espaciales; por ejemplo, hay 638 imágenes con una resolución de 2416x1736 píxeles y una imagen con una resolución de 2146x1764. Entre los problemas de iluminación que presenta el dataset se pueden mencionar, por ejemplo, la subexposición; lo que dificulta la visualización de detalles finos de la retina, sobreexposición; lo que lleva a la saturación de los píxeles y la pérdida de información, por mencionar algunos.

Grado DR	**Lesiones**	**Muestras**	**%**
No DR	No hay lesiones.	1805	49.3
DR leve	Solo Microaneurismas.	370	10.1
DR moderada	Más que MA, pero menos que DR grave.	999	27.2
DR grave	Cualquiera de las siguientes: ▪ Más de 20 microaneurismas intrarretinales cada 4 cuadrantes. ▪ Vasos sanguíneos bien definidos en más de dos cuadrantes. ▪ Anormalidades microvasculares intrarretinales en al menos dos cuadrantes sin señales de DR proliferativa.	193	5.2
DR proliferativa	Una o más de las siguientes: ▪ Neovascularización. ▪ Hemorragias pre-retinales.	295	8.2

Tabla 3.1: Clasificación y composición APTOS 2019.

Como se puede deducir, al presentar una amplia gama de variaciones en términos de resolución y condiciones de iluminación, el APTOS 2019 ofrece un desafío tanto para los algoritmos de visión por computadora como de clasificación, por ejemplo, estos últimos no pueden llegar a generalizar correctamente. La figura 3.1 muestra algunas imágenes que componen el dataset, destacando estas variaciones en calidad y resolución.

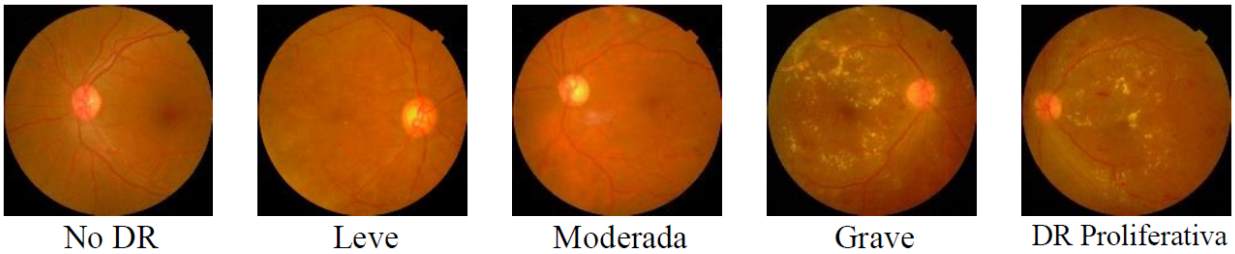

Figura 3.1: Etapas DR Aptos 2019 en espacio de color RGB. Muestra representativa de cada grado de retinopatía.

No. Muestras	**Resolución espacial**
2	474x358
42	640x840
287	819x614
974	1050x1050
2	1467x1110
14	1476x1117
92	1504x1000
61	1844x1226
351	2048x1536
28	2144x1424
1	2146x1764
638	2416x1736
533	2588x1958
34	2896x1944
410	3216x2136
141	3388x2555
52	4288x2848
3662 Muestras	17 resoluciones

Tabla 3.2: Resoluciones espaciales del conjunto de imágenes APTOS 2019.

3.2. Procesamiento de imágenes.

El uso de diversas cámaras, así como inconsistencias en la iluminación de la retina, provoca poca uniformidad en el conjunto de imágenes del APTOS 2019. Para reducir dicha heterogeneidad en las imágenes, que en última instancia afecta el rendimiento del modelo de clasificación, así como para resaltar algunos detalles finos de las imágenes, se realiza un mejoramiento tanto de brillo y contraste de las imágenes con CLAHE; que, como se mencionó,

es uno de los métodos de preprocesamiento más utilizados. Este mejoramiento con CLAHE permite resaltar aquellas partes de la retina que presentan algún tipo de daños debido a DR. En las siguientes secciones de indica cómo se mejoraron las imágenes con CLAHE en los espacios de color LAB y YCbCr.

3.2.1. Mejoramiento de imágenes en espacio de color LAB.

La figura 3.2 muestra el proceso utilizado para el mejoramiento de las imágenes de retina.

Las imágenes se convierten del espacio de color RGB al espacio de color LAB. En este espacio, se selecciona únicamente el componente B, se realiza un mejoramiento tanto de brillo y contraste de las imágenes con CLAHE; que, como se mencionó, es uno de los métodos de preprocesamiento más utilizados. Este mejoramiento con CLAHE permite resaltar aquellas partes de la retina que presentan algún tipo de daños debido a DR. En las siguientes secciones de indica cómo se mejoraron las imágenes con CLAHE en los espacios de color LAB y YCbCr.

Después de la ecualización, se vuelven a combinar los componentes en el espacio LAB y se transforma la imagen de regreso al espacio de color RGB. En el espacio RGB, separan nuevamente los componentes R, G y B seleccionándose exclusivamente el componente verde (G) y se aplica a este canal nuevamente CLAHE. Luego, se combinan los componentes rojo (R) y azul (B) con el componente verde mejorado para obtener imágenes con un contraste optimizado.

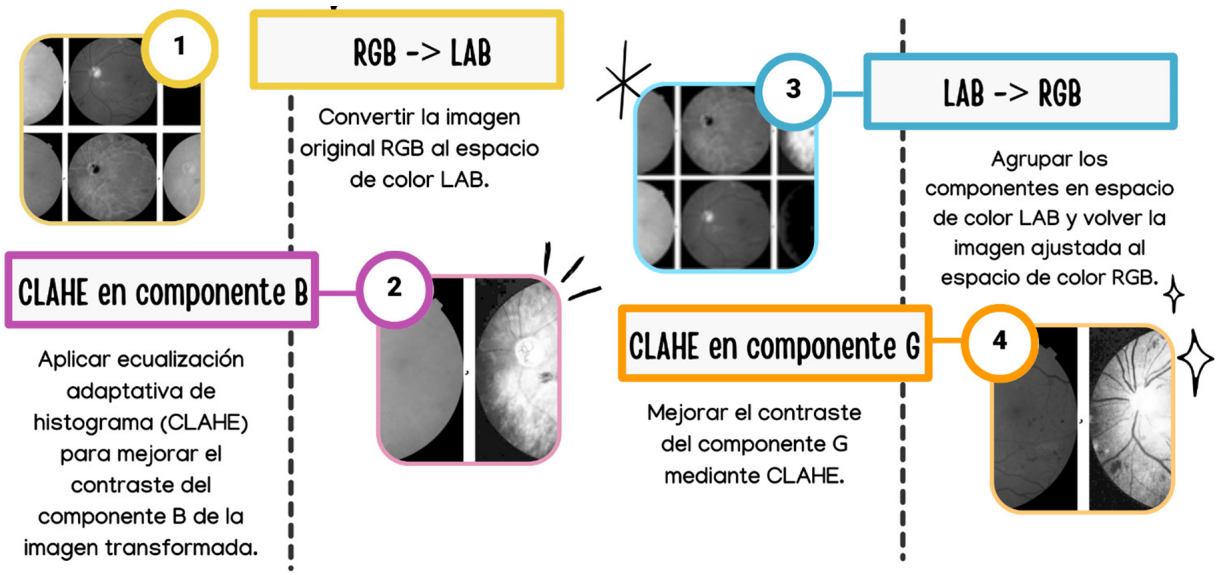

Figura 3.2: Proceso para mejorar imágenes en espacio de color LAB propuesto en [35].

Este proceso de mejora en el espacio de color LAB permite una manipulación más eficaz del contraste y el brillo, ya que separa la luminancia de la información de color, lo que es beneficioso para muchas aplicaciones de procesamiento de imágenes, incluyendo la visión por computadora y el reconocimiento de patrones. La imagen 3.3 presenta un muestra del mejoramiento aplicado en espacio de color LAB.

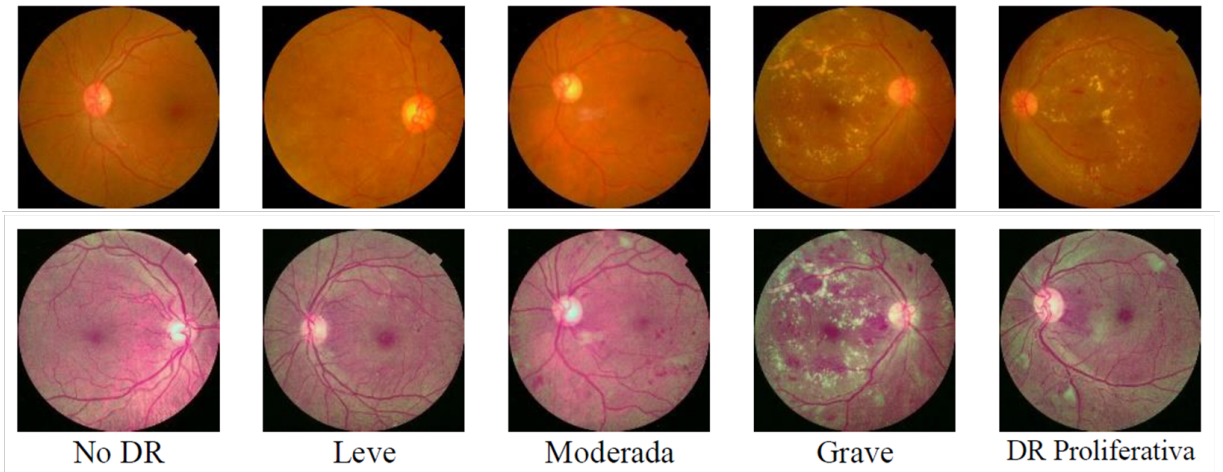

Figura 3.3: Muestra de imágenes mejoradas de espacio de color RGB (primera fila) a LAB (segunda fila) propuesto en [35].

3.2.2. Mejoramiento de imágenes espacio de color YCbCr

La segunda técnica utilizada para el mejoramiento del conjunto de imágenes se basa en trabajar en el espacio de color YCbCr. Este espacio de color separa la información de luminancia de la información de crominancia, lo que permite un procesamiento más específico y eficiente, ver Figura 3.4.

En esta metodología, las imágenes se convierten del espacio de color RGB a YCbCr. Una vez en este espacio, se aplica un filtro de media al componente Y para suavizar y reducir el ruido. Este filtro es eficaz para atenuar variaciones rápidas de brillo que pueden ser interpretadas como ruido, mejorando la calidad de la imagen sin afectar los detalles significativos.

Posteriormente, se realiza una ecualización del histograma del componente Y para que la distribución de sus valores de intensidad sean lo más uniformes posible. La ecualización del histograma aumenta el contraste de la imagen, mejorando la visibilidad de los detalles en áreas oscuras o brillantes y proporcionando una representación más equilibrada de la luminancia.

Después de estos ajustes, el componente Y mejorado se vuelve a combinar con los componentes Cb y Cr originales. Finalmente, las imágenes se convierten de regreso al espacio de color RGB, completando el proceso de mejoramiento en el espacio de color YCbCr.

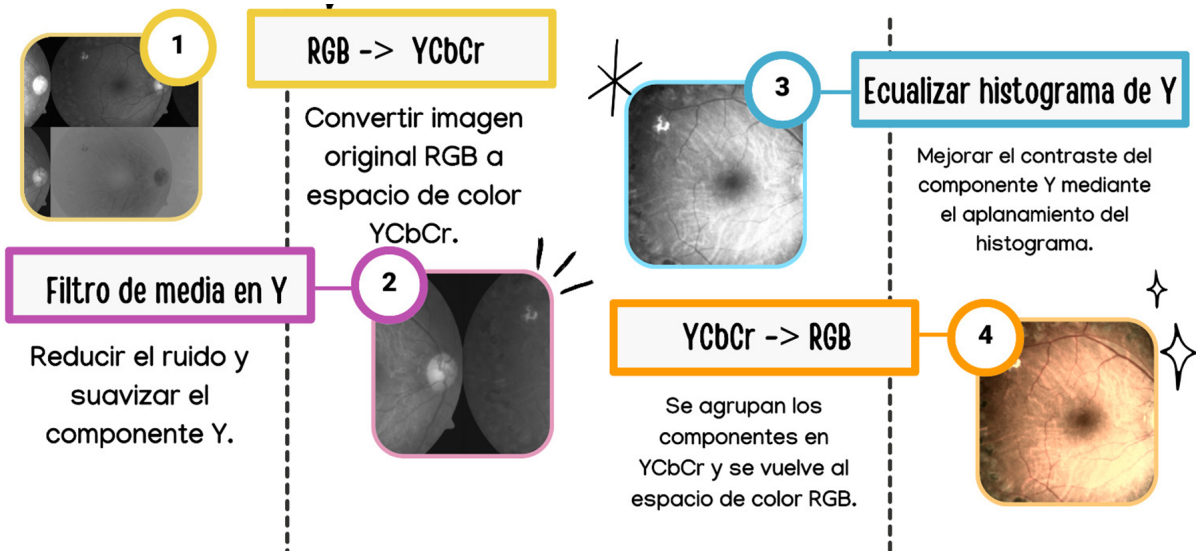

Figura 3.4: Proceso para mejorar imágenes en espacio de color YCbCr propuesto en [36].

Trabajar en el espacio YCbCr permite un procesamiento más focalizado de la luminancia, que es crucial para la percepción de detalles en una imagen, mientras que se preserva la información de color, resultando en imágenes mejoradas con mayor claridad y contraste. La imagen presenta una muestra del mejoramiento aplicado en espacio de color YCbCr.

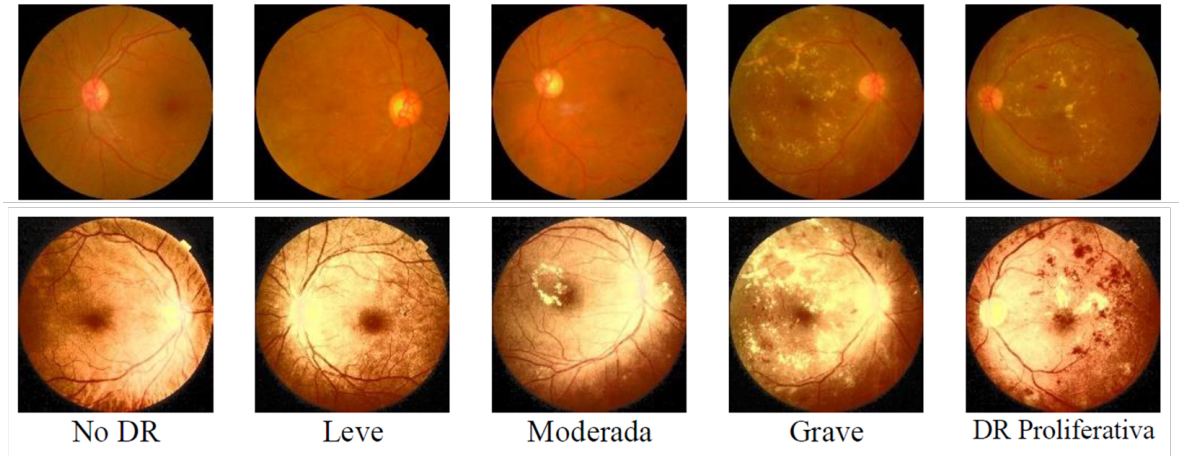

Figura 3.5: Muestra de imágenes mejoradas de espacio de color RGB (primera fila) a YCbCr (segunda fila) propuesto en [36].

3.3. Extractores de Características.

Con las imágenes mejoradas, se cuenta con tres conjuntos de datos: RGB, LAB y YCbCr. Cada espacio de color ofrece diferentes visualizaciones de las lesiones de la retina. En el espacio LAB, los vasos sanguíneos se distinguen mucho mejor, mientras que en el espacio YCbCr, algunos cuadrantes muestran una iluminación completamente diferente en comparación con el RGB, incluso en comparación con las imágenes mejoradas en el espacio de color LAB. Esto permite que cada espacio de color destaque diferentes aspectos o perspectivas sobre el estado de DR.

De acuerdo con la metodología planteada, se procede a la extracción de características utilizando GLMC y LBP, considerando las imágenes en el espacio de color RGB. Además, se realiza la extracción de características profundas mediante los modelos preentrenados de Ademaxpp Model A y DenseNet 121 para el conjunto de datos en cada espacio de color planteado.

La extracción de características se realizó de manera ordenada y en orden alfabético para

cada uno de los conjuntos de datos. La figura 3.6 muestra la organización utilizada para agrupar las imágenes del dataset y realizar la extracción.

Figura 3.6: Organización de datasets.

De este modo, al finalizar la extracción, podemos asegurar que los vectores característicos de cada imagen correspondan a la misma imagen. Al concluir el proceso, se obtiene un conjunto de datos en forma lista formado por los vectores característicos ordenados y etiquetados de acuerdo a su clase correspondientes a la organización del conjunto de imágenes.

3.3.1. LBP.

Los Patrones Locales Binarios (LBP) permiten extraer las características de textura del conjunto de imágenes propuesto. Siendo un algoritmo sencillo y eficaz, LBP ofrece una alternativa importante para considerar tanto en la detección (DR) como en la identificación de su grado de severidad.

Para realizar la aplicación de LBP, se requiere un preprocesamiento de los conjuntos de imágenes en cualquiera de los espacios de color seleccionados. Dado que LBP se basa en la comparación de la intensidad de los píxeles, resulta más sencillo computacionalmente trabajar en un solo canal; por lo tanto, es necesario convertir las imágenes a niveles de grises, como se muestra en la figura 3.7.

Figura 3.7: La primera fila muestra las imágenes en espacio de color RGB, mejoradas en espacio de color LAB y YCbCr, la segunda muestra la conversión respectiva en niveles de gris.

Se utilizó un operador básico de LBP, es decir, vecindades de 3×3, obteniendo vectores característicos de imágenes a 8 bits, resultando en 256 elementos.

Se aplica el algoritmo LBP en cada imagen para la obtención de los vectores de características, que posteriormente serán almacenados y utilizados en la etapa de clasificación. La función representativa del algoritmo LBP se implementó en GNU Octave [37].

3.3.2. Características de Haralick.

Las Matrices de Co-ocurrencia de Nivel de Gris representan otra técnica que permiten extraer las características texturales del conjunto de imágenes propuesto.

Igual que LBP, para hacer la comparación de los valores de los píxeles, resulta más sencillo aplicar GLCM en un solo canal, por lo que también se realiza la conversión del conjunto de imágenes a niveles de grises.

GLCM se calcula para diferentes ángulos y distancias entre los píxeles para capturar la información direccional y de escala de la textura. Para cada imagen, se aplica GLCM para la obtención de los vectores de características, que posteriormente serán almacenados

y utilizados en la etapa de clasificación. La función representativa del algoritmo GLCM se implementó en GNU Octave.

A partir de la GLCM calculada previamente, se derivaron las características de Haralick, como lo propuso Tommy Löfstedt. Löfstedt presentó 21 características para representar textura, algunas incluyen medidas como contraste, correlación, energía y homogeneidad, las cuales pueden ser representadas ordenadamente mediante vectores de 21 elementos útiles para detectar y clasificar las lesiones de DR [38].

3.3.3. Ademaxpp Model A.

Utilizar una red neuronal profunda como extractor de características implica procesar imágenes a través de la red y extraer las activaciones de una o varias capas intermedias. Estas activaciones se consideran vectores de características que representan la información esencial de las imágenes.

Ademxapp Model A es una arquitectura de red neuronal profunda diseñada para tareas de clasificación y reconocimiento. En este contexto, se utiliza como extractor de características, proporcionando vectores característicos detallados para cada imagen procesada.

La arquitectura se secciona en la capa de agregación, lo que permite generar vectores característicos de 4096 elementos por cada imagen. Esta gran capacidad de extracción de características es crucial para capturar detalles finos y complejos de las imágenes de entrada, facilitando una mejor discriminación y análisis en tareas subsecuentes de aprendizaje automático y visión por computadora, ver Figura 3.8.

3.3.4. DenseNet-121.

DenseNet-121 es una arquitectura de red neuronal profunda conocida por su eficiencia y capacidad de reutilización de características. Al ser utilizada como extractor de características, DenseNet-121 genera vectores característicos de 1024 elementos por cada imagen procesada, ver Figura 3.8.

La arquitectura de DenseNet-121 está diseñada para permitir conexiones directas entre todas las capas. Esta capacidad de densidad en las conexiones mejora significativamente el

rendimiento en la extracción de características, proporcionando vectores que capturan de manera efectiva patrones y detalles importantes de las imágenes.

DenseNet-121 es particularmente útil en aplicaciones de visión por computadora donde la precisión y la eficiencia son cruciales, como en la clasificación de imágenes médicas y la detección de DR.

Figura 3.8: DNN's como extractores de características.

3.4. Clasificación.

3.4.1. Binaria.

La clasificación binaria tiene la finalidad de discernir las imágenes que presentan lesiones de retinopatía de las que no, es decir, detectar la ausencia o presencia de la enfermedad.

Dado que el etiquetado porporcionado por el autor del conjunto de imágenes está compuesto de 5 clases, es necesario reorganizar las etiquetas para formar únicamente dos clases, ver tabla 3.3.

Clase (Autor)	**Clase propuesta**
No DR	0
Leve	1
Moderada	1
Grave	1
DR Proliferativa	1

Tabla 3.3: Reorganización de clases para clasificación binaria.

Las imágenes correspondientes a la clase No DR, donde no hay lesiones, se consideran la primera clase. Las muestras de las clases Leve, Moderada, Grave y Proliferativa se agrupan para formar la segunda clase, que indica la presencia de lesiones de DR.

De esta manera, se crean dos clases:

- ***Clase 0***: Contiene las imágenes sin lesiones de retinopatía diabética (No DR).
- ***Clase 1***: Engloba todas las imágenes que contienen lesiones, incluyendo las categorías Leve, Moderada, Grave y Proliferativa.

En resumen, la clase 0 se utiliza para identificar la ausencia de retinopatía diabética, mientras que la clase 1 se utiliza para identificar la presencia de retinopatía diabética. En este punto se tiene un listado de vectores característicos ordenados y organizados en dos clases únicamente.

3.4.2. Multiclase.

La clasificación multiclase tiene como objetivo identificar la etapa o grado de severidad de las imágenes del conjunto de datos, manteniendo las etiquetas originales proporcionadas por el autor. Las clases y sus correspondientes identificadores se muestran en la tabla 3.4.

Clase (Autor)	**Identificador**
No DR	0
Leve	1
Moderada	2
Grave	3
DR Proliferativa	4

Tabla 3.4: Identificadores de clase para clasificación multiclase.

De este modo se genera un listado de vectores característicos para cada clase ordenados de acuerdo a la extracción de características y etiquetados con el identificador correspondiente a su clase.

3.4.3. Muestreo.

Para realizar la clasificación binaria y multiclase, se utilizó un muestreo aleatorio del 80/20, es decir, el 80 % de los datos vectores se emplearon para el entrenamiento de los clasificadores y el 20 % para la validación de los modelos.

Cuando se aplica un muestreo aleatorio general para separar conjuntos de datos en porciones de entrenamiento y validación en datasets desbalanceados, es común que las clases minoritarias no incluyan muestras representativas durante el proceso de entrenamiento o validación debido a la aleatoriedad de la selección. Esto puede resultar en una clasificación errónea durante la evaluación de los modelos una vez entrenados.

En este caso, las clases 1, 3 y 4, que se conforman por 370, 193 y 295 muestras respectivamente, al realizar un muestreo, es muy probable que no incluyan o incluyan una cantidad significativamente pequeña de muestras para la correcta representación de su clase durante el entrenamiento o la validación de los modelos. Debido a esto, se propuso considerar un muestreo aleatorio estratificado, es decir, un 80/20 para cada una de las clases. De esta forma, se asegura que cada clase esté representada tanto en el conjunto de entrenamiento como en el de validación. La distribución del muestreo se presenta en la tabla 3.5

Id. de clase	**80 % muestras**	**20 % muestras**	**total**
0	1444	361	1805
1	296	74	370
2	799	200	999
3	154	39	193
4	236	59	295

Tabla 3.5: Muestreo aleatorio propuesto 80/20.

Para la clasificación binaria, se agrupan las muestras de las clases con Ids 1, 2, 3 y 4 para formar una única clase. En la clasificación multiclase, se consideran las cantidades mostradas en la tabla 3.5.

Para garantizar la comparación del desempeño entre las diferentes técnicas de extracción de características, se utilizaron las mismas imágenes en cada experimento realizado. Por ejemplo, para la clase 0, se aseguró que los 1444 vectores característicos sean los mismos vectores representativos de cada imagen en cada uno de los datasets utilizados.

3.4.4. Clasificación de características.

La clasificación binaria y multiclase se realizó con 3 algoritmos de aprendizaje automático, SVM, RF y GBT para los conjuntos de vectores característicos correspondientes a las imágenes en espacio de color RGB, LAB y YCbCr.

Support Vector Machine (SVM).

Se empleó SVM con un kernel Radial Basis Function (RBF) para la clasificación de los vectores característicos. SVM es conocido por su capacidad para manejar eficazmente conjuntos de datos de alta dimensionalidad y no lineales. El kernel RBF permite la transformación no lineal de los datos para encontrar fronteras de decisión más complejas.

Random Forest (RF).

Se utilizó Random Forest para la clasificación, configurado con 100 árboles de decisión. Este método es robusto frente a datasets desbalanceados y capaz de manejar características no lineales en los conjuntos de datos.

Gradient Boosting Trees (GBT).

Se aplicó Gradient Boosting con 100 árboles de decisión y una tasa de aprendizaje de 0.1. Este algoritmo construye árboles de decisión de manera secuencial, donde cada árbol corrige los errores de predicción de los árboles anteriores. La baja tasa de aprendizaje ayuda a controlar la contribución de cada árbol, mejorando así la capacidad de generalización del modelo.

Estos algoritmos fueron seleccionados debido a su capacidad para manejar conjuntos de datos complejos y desbalanceados como el presente, buscando evaluar y comparar su rendimiento en la clasificación de vectores los característicos extraídos de imágenes de retina para la detección y diagnóstico del grado de DR.

Capítulo 4

Resultados.

En esta sección, se presentan los resultados obtenidos a partir de la clasificación binaria y multiclase del dataset de imágenes de retina APTOS 2019 utilizando diferentes técnicas de mejoramiento de imágenes y algoritmos de extracción de características.

Se evaluó el rendimiento de los modelos de clasificación a través de matrices de confusión, tiempos de extracción de características y tiempos de entrenamiento.

4.1. Mejoramiento de imágenes

Se aplicaron técnicas de mejoramiento de imágenes en los espacios de color LAB y YCbCr para mejorar el contraste y reducir el ruido. A continuación, se muestran ejemplos de imágenes antes y después del mejoramiento en la figura 4.1.

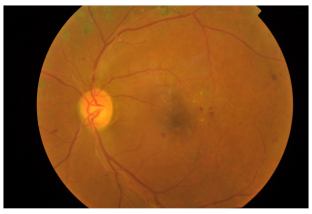	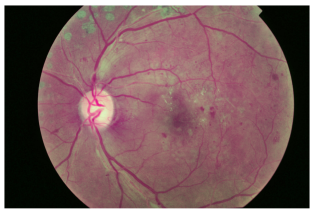	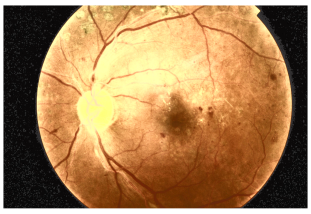
RGB	LAB	YCbCr

Tabla 4.1: Imagen origianl en RGB y su mejoramiento en LAB y YCbCr.

En el estudio se obtuvieron dos conjuntos de datos mejorados, uno en el espacio de color LAB y otro en el espacio de color YCbCr. Cada uno de estos datasets aporta características

y perspectivas únicas para la detección y diagnóstico del grado de DR.

El espacio de color LAB ofrece una representación más precisa de las variaciones de color perceptibles por el ojo humano, lo que facilita la identificación de anomalías sutiles en las imágenes retinianas. Por otro lado, el espacio de color YCbCr, ampliamente utilizado en procesamiento de imágenes y compresión, permite una separación eficaz de la información de luminancia y crominancia, mejorando la detección de patrones específicos relacionados con la DR. La combinación de estos dos enfoques proporciona una visión más completa y robusta para el análisis y diagnóstico de esta enfermedad ocular.

4.2. Extracción de características

Se utilizaron los algoritmos LBP y GLCM para realizar un análisis basado en textura, además de las redes neuronales profundas DenseNet 121 y Ademxapp Model A para extraer características de los conjunto de imágenes propuestos. Los tiempos de extracción de características se presentan en la tabla 4.2.

Algoritmo	**RGB** *(min)*	**LAB** *(min)*	**YCbCr** *(min)*
LBP	67	66	71
GLCM	63	59	66
DenseNet 121	8	10	9
Ademxapp Model A	81	87	77

Tabla 4.2: Tiempo requerido para extracción de características.

La tabla muestra los tiempos en minutos requeridos para extraer características de los espacios de color RGB, LAB y YCbCr utilizando cuatro algoritmos: LBP, GLCM, DenseNet 121 y Ademxapp Model A como extractores.

- El algoritmo LBP requiere tiempos de extracción de características similares en los espacios de color RGB y LAB. Aunque la disparidad es de apenas 4 y 5 min, el espacio de color YCbCr muestra ser menos eficiente respecto al tiempo para la extracción de vectores caractísticos.

- GLCM demuestra ser superior en el espacio de color LAB, no así en YCbCr.

- DenseNet 121 es significativamente más rápido en la extracción de características comparado con los otros algoritmos, ya que requiere menos de 10min en cada espacio de color

- Ademxapp Model A es el algoritmo más lento de los cuatro algoritmos, requiriendo mayor tiempo en espacio de color LAB y el menor en espacio de color YCbCr, demuestra ser el algoritmo más efectivo como extractor de características.

La variabilidad en los tiempos de extracción de características puede influir en la elección del algoritmo dependiendo de los requisitos de tiempo y precisión para la detección y diagnóstico de DR.

4.3. Clasficación binaria y multiclase

La clasificación binaria se realizó agrupando las clases de la siguiente manera: clase 0 representa una clase y las clases 1 a 4 representan la segunda clase, como se mostró en la Tabla 3.3. En la clasificación multiclase, se consideraron las etiquetas proporcionadas por el autor del dataset. Se ejecutó un muestreo aleatorio por clase (80 % entrenamiento, 20 % pruebas) para asegurar la inclusión de muestras representativas de cada clase, tal como se muestra en la tabla 3.5.

Para comprender el desempeño de los modelos de clasificación, se consideraron las matrices de confusión de cada espacio de color y las extactitudes obtenidas. La matriz de confusión es una herramienta valiosa que permite evaluar el rendimiento al mostrar la distribución de las predicciones correctas e incorrectas de cada técnica de extracción empleada. Además, proporciona un análisis más profundo debido a las diversas métricas de evaluación que se se se pueden obtener, tales como exactitud que , F1 Score, Recall, entre otras.

Clasificación binaria

A continuación se muestran las matrices de confusión para clasificación binaria.

Actual class \ Predicted class	No DR	DR	
No DR	336	25	361
DR	33	339	372
	369	364	

Actual class \ Predicted class	No DR	DR	
No DR	334	27	361
DR	43	329	372
	377	356	

Actual class \ Predicted class	No DR	DR	
No DR	340	21	361
DR	31	341	372
	371	360	

Figura 4.1: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores LBP, RGB.

Actual class \ Predicted class	No DR	DR	
No DR	333	28	361
DR	45	327	372
	378	355	

Actual class \ Predicted class	No DR	DR	
No DR	339	22	361
DR	34	338	372
	373	360	

Actual class \ Predicted class	No DR	DR	
No DR	339	22	361
DR	34	338	372
	373	360	

Figura 4.2: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores LBP, LAB.

Actual class \ Predicted class	No DR	DR	
No DR	337	24	361
DR	41	331	372
	378	355	

Actual class \ Predicted class	No DR	DR	
No DR	342	19	361
DR	38	334	372
	380	353	

Actual class \ Predicted class	No DR	DR	
No DR	340	21	361
DR	33	338	372
	373	360	

Figura 4.3: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores LBP, YCbCr.

Actual class \ Predicted class	No DR	DR	
No DR	319	42	361
DR	29	343	372
	348	385	

Actual class \ Predicted class	No DR	DR	
No DR	328	33	361
DR	40	332	372
	368	365	

Actual class \ Predicted class	No DR	DR	
No DR	326	35	361
DR	41	331	372
	367	366	

Figura 4.4: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores GLCM, RGB.

Actual class \ Predicted class	No DR	DR	
No DR	327	34	361
DR	35	337	372
	362	371	

Actual class \ Predicted class	No DR	DR	
No DR	339	22	361
DR	44	328	372
	368	350	

Actual class \ Predicted class	No DR	DR	
No DR	325	36	361
DR	44	334	372
	363	370	

Figura 4.5: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores GLCM, LAB.

Actual class \ Predicted class	No DR	DR	
No DR	338	23	361
DR	37	335	372
	375	358	

Actual class \ Predicted class	No DR	DR	
No DR	343	18	361
DR	39	333	372
	382	351	

Actual class \ Predicted class	No DR	DR	
No DR	336	25	361
DR	33	339	372
	369	364	

Figura 4.6: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores GLCM, YCbCr.

Actual class \ Predicted class	No DR	DR	
No DR	356	5	361
DR	14	358	372
	370	363	

Actual class \ Predicted class	No DR	DR	
No DR	354	7	361
DR	19	353	372
	373	360	

Actual class \ Predicted class	No DR	DR	
No DR	353	8	361
DR	19	353	372
	372	361	

Figura 4.7: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores DenseNet121, RGB.

Actual class \ Predicted class	No DR	DR	
No DR	353	8	361
DR	13	359	372
	366	367	

Actual class \ Predicted class	No DR	DR	
No DR	357	4	361
DR	17	355	372
	374	359	

Actual class \ Predicted class	No DR	DR	
No DR	350	11	361
DR	17	355	372
	367	366	

Figura 4.8: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores DenseNet121, LAB.

Actual class \ Predicted class	No DR	DR	
No DR	353	8	361
DR	13	359	372
	366	367	

Actual class \ Predicted class	No DR	DR	
No DR	357	4	361
DR	17	355	372
	374	359	

Actual class \ Predicted class	No DR	DR	
No DR	350	11	361
DR	17	355	372
	367	366	

Figura 4.9: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores DenseNet121, YCbCr.

Actual class \ Predicted class	No DR	DR	
No DR	356	5	361
DR	11	361	372
	367	366	

Actual class \ Predicted class	No DR	DR	
No DR	352	9	361
DR	18	354	372
	370	363	

Actual class \ Predicted class	No DR	DR	
No DR	354	7	361
DR	13	359	372
	367	366	

Figura 4.10: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores Ademxapp Model A, RGB.

Actual class \ Predicted class	No DR	DR	
No DR	357	4	361
DR	10	362	372
	367	366	

Actual class \ Predicted class	No DR	DR	
No DR	353	8	361
DR	23	349	372
	376	357	

Actual class \ Predicted class	No DR	DR	
No DR	352	9	361
DR	17	355	372
	369	364	

Figura 4.11: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores Ademxapp Model A, LAB.

Actual class \ Predicted class	No DR	DR	
No DR	356	5	361
DR	12	360	372
	368	365	

Actual class \ Predicted class	No DR	DR	
No DR	355	6	361
DR	19	353	372
	374	359	

Actual class \ Predicted class	No DR	DR	
No DR	357	4	361
DR	18	354	372
	375	358	

Figura 4.12: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación binaria de vectores Ademxapp Model A, YCbCr.

Las matrices de confusión siguientes muestran el desempeño de los clasificadores cuando se busca identificar el grado de DR.

Clasificación multiclase

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	346	4	11	0	0	361
Leve	8	45	20	1	0	74
Moderada	28	35	136	1	0	200
Grave	5	12	20	2	0	39
Proliferativa	11	13	35	0	0	59
	398	109	222	4	0	

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	349	1	11	0	0	361
Leve	8	23	38	1	4	74
Moderada	21	15	156	1	7	200
Grave	5	2	22	2	5	39
Proliferativa	6	6	38	1	8	59
	398	50	265	5	24	

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	346	2	13	0	0	361
Leve	10	29	30	1	4	74
Moderada	23	13	158	0	6	200
Grave	2	2	30	4	1	39
Proliferativa	9	6	36	0	8	59
	390	52	267	5	19	

Figura 4.13: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores LBP, RGB.

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	344	0	17	0	0	361
Leve	11	5	57	0	1	74
Moderada	33	1	166	0	0	200
Grave	8	1	28	0	2	39
Proliferativa	15	1	40	0	3	59
	411	8	308	0	6	

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	349	2	10	0	0	361
Leve	10	28	27	1	8	74
Moderada	24	14	153	1	8	200
Grave	6	3	26	1	3	39
Proliferativa	8	6	34	1	10	59
	397	53	250	4	29	

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	348	2	11	0	0	361
Leve	8	32	29	1	4	74
Moderada	20	19	157	1	3	200
Grave	4	8	25	1	1	39
Proliferativa	8	9	38	1	3	59
	388	70	260	4	11	

Figura 4.14: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores LBP, LAB.

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	345	0	16	0	0	361
Leve	9	6	59	0	0	74
Moderada	30	1	168	1	0	200
Grave	4	2	32	1	0	39
Proliferativa	11	3	45	0	0	59
	399	12	320	2	0	

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	353	1	7	0	0	361
Leve	10	30	25	1	8	74
Moderada	29	21	135	1	14	200
Grave	4	4	23	3	5	39
Proliferativa	11	3	36	0	9	59
	407	59	226	5	36	

x|

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	346	2	13	0	0	361
Leve	10	29	30	1	4	74
Moderada	23	13	158	0	6	200
Grave	2	2	30	4	1	39
Proliferativa	9	6	36	0	8	59
	390	52	267	5	19	

Figura 4.15: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores LBP, YCbCr.

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	344	1	16	0	0	361
Leve	15	35	33	1	0	74
Moderada	29	20	150	1	0	200
Grave	6	3	28	2	0	39
Proliferativa	12	6	41	0	0	59
	406	55	268	4	0	

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	344	0	16	0	1	361
Leve	13	24	33	2	2	74
Moderada	26	24	139	4	7	200
Grave	8	4	23	3	1	39
Proliferativa	11	4	35	3	6	59
	402	56	246	12	17	

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	337	2	22	0	0	361
Leve	12	25	32	1	4	74
Moderada	29	37	137	2	5	200
Grave	6	4	27	2	0	39
Proliferativa	13	7	37	1	1	59
	397	65	255	6	10	

Figura 4.16: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores GLCM, RGB.

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	345	1	15	0	0	361
Leve	14	24	35	1	0	74
Moderada	26	16	157	1	0	200
Grave	6	2	29	2	0	39
Proliferativa	9	6	34	1	9	59
	403	47	279	4	0	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	346	1	15	0	0	361
Leve	14	22	27	1	10	74
Moderada	24	19	140	3	14	200
Grave	8	3	21	4	3	39
Proliferativa	9	6	34	1	9	59
	401	50	237	9	36	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	337	1	23	0	0	361
Leve	12	28	31	1	2	74
Moderada	23	34	146	2	5	200
Grave	6	5	21	3	4	39
Proliferativa	10	6	35	1	7	59
	388	64	256	7	18	

Figura 4.17: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores GLCM, LAB.

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	346	1	14	0	0	361
Leve	11	13	49	1	0	74
Moderada	26	19	154	1	0	200
Grave	1	1	35	2	0	39
Proliferativa	11	3	45	0	0	59
	395	37	297	4	0	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	346	1	11	2	1	361
Leve	11	26	41	2	4	74
Moderada	28	14	150	3	5	200
Grave	1	3	28	2	4	39
Proliferativa	11	3	36	0	9	59
	397	38	266	9	23	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	343	4	13	1	0	361
Leve	11	18	43	2	0	74
Moderada	22	18	153	3	4	200
Grave	1	2	31	3	2	39
Proliferativa	9	3	41	0	6	59
	386	45	281	9	12	

Figura 4.18: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores GLCM, YCbCr.

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	357	3	1	0	0	361
Leve	10	37	25	1	1	74
Moderada	10	12	166	3	9	200
Grave	0	0	17	15	7	39
Proliferativa	0	1	31	4	23	59
	377	53	240	23	40	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	355	3	3	0	0	361
Leve	9	42	21	1	1	74
Moderada	7	10	167	8	8	200
Grave	0	0	15	15	9	39
Proliferativa	0	0	28	7	24	59
	371	55	234	31	42	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	355	2	4	0	0	361
Leve	11	28	33	1	1	74
Moderada	12	3	179	2	4	200
Grave	0	1	22	5	11	39
Proliferativa	1	1	40	0	17	59
	379	35	278	8	33	

Figura 4.19: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores DenseNet 121, RGB.

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	355	2	4	0	0	361
Leve	9	38	26	0	1	74
Moderada	7	15	168	3	7	200
Grave	0	0	16	15	8	39
Proliferativa	0	2	31	6	20	59
	371	57	245	24	36	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	356	1	4	0	0	361
Leve	13	23	38	0	0	74
Moderada	15	4	174	1	6	200
Grave	0	0	23	3	13	39
Proliferativa	0	0	39	0	20	59
	372	54	255	25	27	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	354	4	4	0	0	361
Leve	9	35	29	1	0	74
Moderada	10	13	171	3	3	200
Grave	0	1	17	15	6	39
Proliferativa	0	1	34	6	18	59
	372	54	255	25	27	

Figura 4.20: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores DenseNet 121, LAB.

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	354	6	1	0	0	361
Leve	9	37	25	0	3	74
Moderada	8	10	173	4	5	200
Grave	0	0	21	11	7	39
Proliferativa	0	2	35	5	17	59
	371	55	255	20	32	

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	355	1	5	0	0	361
Leve	14	18	40	0	2	74
Moderada	15	4	175	1	5	200
Grave	0	0	28	3	8	39
Proliferativa	1	1	42	0	15	59
	385	24	290	4	30	

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	356	3	2	0	0	361
Leve	11	36	26	0	1	74
Moderada	8	10	170	4	8	200
Grave	0	0	23	11	5	39
Proliferativa	0	2	35	4	18	59
	375	51	256	19	32	

Figura 4.21: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores DenseNet 121, YCbCr.

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	357	2	2	0	0	361
Leve	7	42	25	0	0	74
Moderada	6	10	173	5	6	200
Grave	0	0	21	10	8	39
Proliferativa	0	1	28	6	24	59
	370	55	249	21	38	

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	357	2	2	0	0	361
Leve	13	10	50	0	1	74
Moderada	16	1	177	0	6	200
Grave	2	0	29	1	7	39
Proliferativa	3	0	40	0	16	59
	391	13	298	1	30	

Predecited class

Actual class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	356	3	2	0	0	361
Leve	10	30	34	0	0	74
Moderada	8	12	174	0	6	200
Grave	2	0	28	4	5	39
Proliferativa	0	0	35	2	22	59
	376	45	273	6	33	

Figura 4.22: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores Ademxapp Model A, RGB.

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	357	3	1	0	0	361
Leve	7	45	22	0	0	74
Moderada	6	12	176	3	3	200
Grave	0	0	21	12	6	39
Proliferativa	0	2	24	5	28	59
	370	62	244	20	37	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	356	2	3	0	0	361
Leve	13	22	39	0	0	74
Moderada	20	2	172	1	5	200
Grave	1	0	27	0	11	39
Proliferativa	2	0	40	0	17	59
	392	26	281	1	33	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	355	2	4	0	0	361
Leve	13	30	30	0	1	74
Moderada	13	10	167	3	7	200
Grave	0	0	25	7	7	39
Proliferativa	0	1	35	2	21	59
	381	43	261	12	36	

Figura 4.23: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores Ademxapp Model A, LAB.

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	356	4	1	0	0	361
Leve	7	39	26	1	1	74
Moderada	4	7	179	5	5	200
Grave	0	0	20	12	7	39
Proliferativa	0	1	30	2	26	59
	367	51	256	20	39	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	355	2	3	0	1	361
Leve	16	15	42	0	1	74
Moderada	17	4	170	1	8	200
Grave	0	0	32	1	6	39
Proliferativa	2	1	47	0	9	59
	390	22	294	2	25	

Actual class \ Predecited class	No DR	Leve	Moderada	Grave	Proliferativa	
No DR	355	4	2	0	0	361
Leve	12	25	35	0	2	74
Moderada	13	9	172	3	3	200
Grave	0	0	25	8	6	39
Proliferativa	2	2	42	3	10	59
	382	40	276	12	23	

Figura 4.24: Matrices de confusión SVM, RF y GBT respectivamente. Clasificación multiclase de vectores Ademxapp Model A, YCbCr.

Las tablas a continuación muestran la exactitud obtenida en cada modelo entrenado para detección e identificación del grado de DR en cada espacio de color trabajado.

Clasificación binaria

Clasificador	***LBP***	***GLCM***	***Ademxapp Model A***	***DenseNet-121***
RF	0.9222	0.9004	0.9631	0.9645
SVM	0.9113	0.9031	0.9781	0.9740
GBT	0.9249	0.8963	0.9727	0.9672

Tabla 4.3: Exactitud en espacio de color RGB.

Clasificador	***LBP***	***GLCM***	***Ademxapp Model A***	***DenseNet-121***
RF	0.9236	0.9099	0.9577	0.9771
SVM	0.9236	0.9058	0.9809	0.9771
GBT	0.9255	0.8990	0.9645	0.9618

Tabla 4.4: Exactitud en espacio de color LAB.

Clasificador	***LBP***	***GLCM***	***Ademxapp Model A***	***DenseNet-121***
RF	0.9222	0.9222	0.9658	0.9713
SVM	0.9113	0.9181	0.9768	0.9713
GBT	0.9249	0.9208	0.9699	0.9618

Tabla 4.5: Exactitud en espacio de color YCbCr.

Clasificación multiclase

Clasificador	***LBP***	***GLCM***	***Ademxapp Model A***	***DenseNet-121***
RF	0.7339	0.7039	0.7653	0.8226
SVM	0.7216	0.7244	0.8267	0.8226
GBT	0.7435	0.6848	0.7994	0.7967

Tabla 4.6: Exactitud en espacio de color RGB.

Clasificador	***LBP***	***GLCM***	***Ademxapp Model A***	***DenseNet-121***
RF	0.7380	0.7010	0.7735	0.7858
SVM	0.7066	0.7326	0.8431	0.8130
GBT	0.7380	0.7010	0.7912	0.8090

Tabla 4.7: Exactitud en espacio de color LAB.

Clasificador	*LBP*	*GLCM*	*Ademxapp Model A*	*DenseNet-121*
RF	0.7230	0.7271	0.7503	0.7721
SVM	0.7094	0.7025	0.8349	0.8076
GBT	0.7435	0.7135	0.7776	0.8062

Tabla 4.8: Exactitud en espacio de color YCbCr.

4.4. Análisis comparativo

Al comparar los diferentes espacios de color propuestos por medio de sus vectores de carácterísticas para realizar tanto clasificación binaria y multiclase, es posible considerar varios enfoques, por ejemplo la exactitud u otra métrica de interés, la eficiencia en relación a costo computacional o el tiempo que conlleva el proceso total para identificar la ausencia o presencia de DR y el diagnóstico del DR.

Clasificación binaria y técnicas basadas en textura

LBP y GLCM son las técnicas utilizadas para realizar un análisis de texturas, considerando los resultados obtenidos se puede obervar:

- GBT es el modelo más robusto y eficaz para ambas técnicas de extracción de características, alcanzando el mejor desempeño en el espacio de color YCbCr. Tanto LBP y GLCM en espacio de color RGB alcanzó una exactitudes superiores a 92 %.

- SVM muestra un menor desempeño en comparación con GBT y RF, especialmente en el espacio de color LAB cuando se utiliza GLCM. Cuando se evaluaron los vectores característicos de imagenes en espacio de color LAB por medio de SVM se logró una exactitud de 89.90 %.

- El espacio de color YCbCr proporciona una mayor consistencia en los vectores característicos extraídos por GLCM ya que logra exactitudes de 92.22 % con cualquiera de los clasificadores propuestos.

- En general, las técnicas de extracción de características basadas en textura LBP y GLCM son más eficaces cuando se combinan con el modelo GBT y cuando se utilizan imágenes en el espacio de color RGB o YCbCr.

Clasificación binaria y características profundas

Ademxapp Model A y DenseNet 121 son las arquitecturas propuestas para extracción de características, después de la evaluación es importante señalar:

- SVM mostró ser el modelo más robusto y eficaz para ambas arquitecturas alcanzando el mejor desempeño en los espacios de color LAB mediante Ademxapp Model A y YCbCr por medio de DenseNet 121 logrando 98.09 % y 97.71 % de exactitud respectivamente.

- En relación a la exactitud obtenida con los demás algoritmos, RF muestra un desempeño ligeramente mas bajo en espacio de color LAB con ambas arquitecturas logrando exactitudes de 95.77 % en Ademxapp Model A y 96.18 % en DenseNet 121 en YCbCr. A pesar de obtener los peores desempeños, siguen siendo relativamente altos considerando tiempo de extracción y entrenamiendo de los modelos.

- RF exhibió un desempeño inferior con relación a los otros modelos en prácticamente todos los escenarios evaluados, especialmente en el espacio de color LAB por medio de características extraídas de GLCM, sin embargo, resulta una alternativa significativa considerando el tiempo de entrenamiento y extracción de característcas en combinación con DensetNet 121 en espacio de color YCbCr.

- Ademxapp Model A ofrece un desempeño ligeramente superior en el mejor de los casos con 98.09 % de exactitud en LAB pero a costa de tiempos de extracción y entrenamiento significativamente más largos en comparación con DenseNet 121.

- Por otra parte, DenseNet 121 proporciona tiempos de extracción y entrenamiento más eficientes, lo que puede ser una ventaja significativa en aplicaciones donde el tiempo es crítico.

- En general, ambas arquitecturas de características profundas son eficaces para la clasificación binaria de imágenes retina, con Ademxapp Model A mostrando un ligero margen de mejora en desempeño a cambio de tiempos de extracción más largos, mientras que DenseNet 121 ofrece un balance más favorable entre desempeño y tiempo de extracción.

Clasificación binaria y LBP-GLCM

- GBT en combinación con YCbCr mediante las características extraídas por LBP ofrecen el mejor desempeño en clasificación multiclase, alcanzando una exactitud de 74.35
- SVM muestra un rendimiento robusto a través de GLCM en el espacio de color LAB, alcanzando una exactitud de 72.03
- RGB es un espacio de color consistentemente LBP, al eliminar la necesidad de mejora en las imágenes y ofreciendo resultados sólidos con exactitudes entre 71 % y 73 % con cualquiera de los clasificadores utilizados.
- LAB demuestra ser el espacio de color más consistente para GLCM, con exactitudes estables entre 71 % y 72 %, además requiere menos tiempo para extracción de características.
- En relacion a los de tiempos de extracción, GLCM es más eficiente que LBP.
- Es importante considerar que tanto LBP como GLCM presentan dificultades para la identificación de las clases Grave y Proliferativa de DR con el conjuntos de datos y muestreo propuesto de acuerdo a las matrices de confusión obtenidas.

Clasificación binaria y DNN

- Ademxapp Model A ofrece el mejor desempeño en términos de exactitud cuando se clasifica por medio de SVM en el espacio de color LAB, alcanzando el 84.31 % de exactitud en la identificación del grado de DR. Sin embargo, sus tiempos de extracción son considerablemente largos, lo que presenta una desventaja importante cuando el tiempo es una variable crucial.
- DenseNet 121 es consistente y eficiente, mostrando tiempos de extracción mucho más cortos en comparación con Ademxapp Model A. Su mejor desempeño se observa en la cuando se extraen características de espacio de color RGB y se clasifican mediante SVM, lo grando una exactitud del 82.26 %.

- La consistencia de DenseNet 121 en el espacio de color RGB, con exactitudes entre 79 % y 82 % en cualquier clasificador, la hace una opción robusta para la extracción de características profundas.

Los extractores de características profundas, particularmente Ademxapp Model A y DenseNet 121, ofrecen una superioridad clara en términos de exactitud tanto para la clasificación binaria como multiclase. Aunque las técnicas basadas en textura (LBP y GLCM) son menos precisas, siguen siendo métodos viables y más sencillos de implementar. DenseNet 121 destaca por su eficiencia en tiempo de extracción, mientras que Ademxapp Model A muestra el mejor desempeño global en términos de exactitud.

Capítulo 5

Conclusiones.

El objetivo de este análisis fue comparar diversos métodos de extracción de características aplicados en la clasificación binaria y multiclase para la detección y evaluación de DR con la finalidad de encontrar el proceso más eficiente. La clasificación binaria se centró en detectar la presencia o ausencia de retinopatía, mientras que la clasificación multiclase buscó identificar el grado de severidad de la misma.

En la clasificación binaria, los extractores de características profundas Ademxapp Model A y DenseNet 121 mostraron un desempeño significativamente superior en comparación con las técnicas basadas en textura, LBP y GLCM:

- Ademxapp Model A obtuvo la mejor exactitud cuando se extraen características del espacio de color LAB y se clasifican a través de SVM, demostrando ser la técnica más eficaz para la presencia o ausencia de DR del conjunto de datos propuesto. Sin embargo, su tiempo de extracción de características fue considerablemente alto.

- DenseNet 121 también mostró un excelente desempeño en espacio de color YCbCr y SVM, además destacó por su eficiencia en el tiempo de extracción de características. LBP y GLCM, aunque menos precisos, alcanzaron exactitudes superiores a 90 %, demostrando ser métodos viables y relativamente consistentes en todos los espacios de color.

- Los extractores de características profundas son claramente superiores, ofreciendo una precisión alta en la detección de la presencia o ausencia de DR. DenseNet 121 es

particularmente notable por su rapidez en la extracción de características.

Para la clasificación multiclase, que evalúa el grado de severidad de DR, los resultados también favorecieron a los extractores de características profundas:

- Ademxapp Model A nuevamente lideró con una exactitud más alta. A pesar de que requiere de mucho tiempo en la extracción de características, la clasificación de la severidad de la DR lo convierte en una opción robusta.
- DenseNet 121 logró un 82.26 % a traves de espacio de color RGB y SVM, además mantuvo tiempos de extracción rápidos, siendo una opción eficiente y precisa.
- GLCM y LBP obtuvieron desempeños más modestos, con exactitudes de 68 % y 74 % en general en los espacios de color trabajados. No obstante, demostraron consistencia en sus respectivos espacios de color y clasificadores, siendo métodos más sencillos y menos demandantes en términos de recursos computacionales.
- DNN's también dominan en la identificación del grado de severidad de DR. Ademxapp Model A proporciona la mayor exactitud, aunque a costa de tiempos de procesamiento más largos. DenseNet 121 ofrece un buen equilibrio entre precisión y eficiencia.

Aunque los métodos basados en textura como LBP y GLCM pueden ser útiles y más fáciles de implementar, los extractores de características profundas son preferibles para aplicaciones clínicas donde la precisión es crítica, como en la detección y clasificación de la retinopatía diabética. La elección del método dependerá del balance deseado entre exactitud y tiempo de procesamiento, con una ligera ventaja para DenseNet 121 en escenarios donde la velocidad es crucial.

Es esencial considerar diversos factores que pueden influir significativamente en los resultados y desempeño de los modelos. Uno de los desafíos más importantes en este estudio fue el desbalance del dataset, el cual puede afectar negativamente el desempeño de los clasificadores, especialmente en clases minoritarias. Para abordar esta problemática se pueden implementar técnicas como aumento de datos, selección y discrimnación de muestras.

La reducción de la dimensionalidad de los vectores de características es otra consideración importante, métodos como selección de características pueden ayudar a disminuir la complejidad computacional y mejorar el desempeño de los modelos de clasificación.

Además de experimentar otras técnicas de análisis de textura, es posible considerar la modificación de la resolución espacial del conjunto de imágenes ya que este puede afectar el análsis texturial, lo que puede influir tanto en el tiempo como en el desempeño de los algoritmos basados en textura.

Finalmente es importante considerar otras arquitecturas de redes neuronales, ya que ofrecen diferentes perspectivas y posibles mejoras en la extracción de caracteristicas.

Referencias

[1] Gobierno. *Diabetes en México*. 2021. URL: https://www.gob.mx/promosalud/acciones-y-programas/diabetes-en-mexico-284509 (visitado 15-10-2023).

[2] INEGI. *Estadísticas de Defunciones Registradas (EDR) de enero a junio de 2023 (preliminar)*. 2024. URL: https://www.inegi.org.mx/app/buscador/default.html?q=diabetes (visitado 15-10-2023).

[3] Organización Panamericana de la Salud. *Diabetes OPS/OMS*. 2023. URL: https://www.paho.org/es/temas/diabetes (visitado 15-10-2023).

[4] Helen D Nickerson y Sanjoy Dutta. "Diabetic complications: current challenges and opportunities". en. En: *J. Cardiovasc. Transl. Res.* 5.4 (ago. de 2012), págs. 375-379.

[5] Wejdan L Alyoubi, Maysoon F Abulkhair y Wafaa M Shalash. "Diabetic retinopathy fundus image classification and lesions localization system using deep learning". en. En: *Sensors (Basel)* 21.11 (mayo de 2021), pág. 3704.

[6] Samer Ellahham. "Artificial intelligence: The future for diabetes care". en. En: *Am. J. Med.* 133.8 (ago. de 2020), págs. 895-900.

[7] Tien-En Tan y Tien Yin Wong. "Diabetic retinopathy: Looking forward to 2030". en. En: *Front. Endocrinol. (Lausanne)* 13 (2022), pág. 1077669.

[8] Ulises Román Concha Nora La Serna Palomino. "Técnicas de Segmentación en Procesamiento Digital de Imágenes". En: *Revista de investigación de Sistemas e Informática* 6 (dic. de 2009), págs. 9-16.

[9] D. Ochoa C. Crespo. "Evaluación Cuantitativa de la Influencia de los Espacios de Color para la Detección Automática de Células". es. En: *Revista Tecnológica ESPOL* 20.1 (oct. de 2007), págs. 73-78.

[10] Shota FURUKAWA, Rei NAGASAKI y Yoshiaki UEDA. "Brightness and Saturation Enhancement by Using Color Vector on Equi-Hue Plane in RGB Color Space". En: *Procedia Computer Science* 207 (2022), págs. 966-975. DOI: https://doi.org/10.1016/j.procs.2022.09.152.

[11] Katherine León et al. "Color measurement in Lab units from RGB digital images". En: *Food Research International* 39.10 (2006). Physical Properties VI, págs. 1084-1091. ISSN: 0963-9969. DOI: https://doi.org/10.1016/j.foodres.2006.03.006.

[12] Balkrishan Ahirwal, Mahesh Khadtare y Rakesh Mehta. "FPGA based system for color space transformation RGB to YIQ and YCbCr". En: *2007 International Conference on Intelligent and Advanced Systems*. 2007, págs. 1345-1349. DOI: 10.1109/ICIAS.2007.4658603.

[13] N.I. Zaidi et al. "Fire recognition using RGB and YCbCr color space". En: *ARPN Journal of Engineering and Applied Sciences* 10 (ene. de 2015), págs. 9786-9790.

[14] Research Paper Research Paper, R. C. Singh Chandel y Gaurav Gupta. "Image Filtering Algorithms and Techniques: A Review". En: *International Journal of Advanced Research in Computer Science and Software Engineering* 10 (2013).

[15] Agung W. Setiawan et al. "Color retinal image enhancement using CLAHE". En: *International Conference on ICT for Smart Society*. 2013, págs. 1-3. DOI: 10.1109/ICTSS.2013.6588092.

[16] Ali M Reza. "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement". En: *J. VLSI Sign. Process. Syst. Sign. Image Video Technol.* 38.1 (ago. de 2004), págs. 35-44.

[17] L. Van Gool, P. Dewaele y A. Oosterlinck. "Texture analysis Anno 1983". En: *Computer Vision, Graphics, and Image Processing* 29.3 (1985), págs. 336-357. ISSN: 0734-189X. DOI: https://doi.org/10.1016/0734-189X(85)90130-6.

[18] Matti Pietikäinen. “Image Analysis with Local Binary Patterns”. En: *Image Analysis*. Ed. por Heikki Kalviainen, Jussi Parkkinen y Arto Kaarna. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, págs. 115-118. ISBN: 978-3-540-31566-7.

[19] Timo Ahonen, Abdenour Hadid y Matti Pietikäinen. “Face Recognition with Local Binary Patterns”. En: *Computer Vision - ECCV 2004*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, págs. 469-481. ISBN: 978-3-540-24670-1.

[20] S Sreehari Sastry et al. “Transition temperatures of thermotropic liquid crystals from the Local Binary Gray Level Cooccurrence Matrix”. En: *Adv. Condens. Matter Phys.* 2012 (2012), págs. 1-9.

[21] Fritz Albregtsen. “Statistical Texture Measures Computed from Gray Level Coocurrence Matrices”. En: *Image Processing Laboratory Department of Informatics University of Oslo*. 2008, págs. 1-14. URL: https://api.semanticscholar.org/CorpusID:18927855.

[22] Robert M Haralick, K Shanmugam e Its’hak Dinstein. “Textural Features for Image Classification”. En: *IEEE Trans. Syst. Man Cybern.* SMC-3.6 (nov. de 1973), págs. 610-621.

[23] Saleh A Albelwi. “Deep architecture based on DenseNet-121 model for weather image recognition”. en. En: *Int. J. Adv. Comput. Sci. Appl.* 13.10 (2022).

[24] Wolfram Mathematica. *Wolfram Neural Net Repository of Neural Network Models*. 2024. URL: https://resources.wolframcloud.com/NeuralNetRepository/ (visitado 01-02-2024).

[25] Zifeng Wu, Chunhua Shen y Anton van den Hengel. “Wider or Deeper: Revisiting the ResNet Model for Visual Recognition”. En: *Pattern Recogn.* 90.C (jun. de 2019), págs. 119-133. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2019.01.006. URL: https://doi.org/10.1016/j.patcog.2019.01.006.

[26] Corinna Cortes y Vladimir Vapnik. “Support-vector networks”. en. En: *Mach. Learn.* 20.3 (sep. de 1995), págs. 273-297.

[27] Aakash Parmar, Rakesh Katariya y Vatsal Patel. "A Review on Random Forest: An Ensemble Classifier". En: Springer International Publishing, ene. de 2019, págs. 758-763. ISBN: 978-3-030-03145-9. DOI: 10.1007/978-3-030-03146-6_86.

[28] Alexey Natekin y Alois Knoll. "Gradient boosting machines, a tutorial". En: *Front. Neurorobot.* 7 (2013).

[29] Andrzej Grzybowski et al. "Artificial intelligence for diabetic retinopathy screening: a review". en. En: *EYE* 34.3 (mar. de 2020), págs. 451-460.

[30] Dolly Das, Saroj Kr Biswas y Sivaji Bandyopadhyay. "A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning". en. En: *Multimed. Tools Appl.* 81.18 (mar. de 2022), págs. 25613-25655.

[31] Mohamed A Berbar. "Features extraction using encoded local binary pattern for detection and grading diabetic retinopathy". en. En: *Health Inf. Sci. Syst.* 10.1 (dic. de 2022), pág. 14.

[32] Maryam Monemian y Hossein Rabbani. "Detecting red-lesions from retinal fundus images using unique morphological features". en. En: *Sci. Rep.* 13.1 (mar. de 2023), pág. 3487.

[33] Luis Filipe Nakayama et al. "Diabetic retinopathy classification for supervised machine learning algorithms". en. En: *Int. J. Retina Vitreous* 8.1 (ene. de 2022), pág. 1.

[34] kaggle. *APTOS 2019 Blindness Detection*. 2019. URL: https://www.kaggle.com/competitions/aptos2019-blindness-detection/discussion/103344 (visitado 03-09-2023).

[35] Priyadharsini C y Jagadeesh Kannan R. "Retinal image enhancement based on color dominance of image". En: *Scientific Reports* 13.1 (mayo de 2023). ISSN: 2045-2322. DOI: 10.1038/s41598-023-34212-w. URL: http://dx.doi.org/10.1038/s41598-023-34212-w.

[36] Anas Bilal et al. "Diabetic Retinopathy Detection and Classification Using Mixed Models for a Disease Grading Database". En: *IEEE Access* PP (feb. de 2021), págs. 1-1. DOI: 10.1109/ACCESS.2021.3056186.

[37] Jhon W. Eaton. *GNU Octave*. 2024. URL: https://octave.org (visitado 01-11-2023).

[38] Tommy Löfstedt et al. "Gray-level invariant Haralick texture features". En: *PLOS ONE* 14.2 (feb. de 2019). Ed. por Mathieu Hatt, e0212110. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0212110. URL: http://dx.doi.org/10.1371/journal.pone.0212110.