



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ingeniería de Sistemas e Informática

Escuela Profesional de Ingeniería de Sistemas

**Aplicación de modelos de transferencia de aprendizaje
para la detección de casos de catarata**

TESIS

Para optar el Título Profesional de Ingeniero de Sistemas

AUTOR

Piero Yahir CURAY CHACÓN

ASESOR

Mg. Santiago MOQUILLAZA HENRIQUEZ

Lima, Perú

2025



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Curay, P. (2025). *Aplicación de modelos de transferencia de aprendizaje para la detección de casos de catarata*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería de Sistemas e Informática, Escuela Profesional de Ingeniería de Sistemas]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios autor/ asesor

Datos de autor	
Nombres y apellidos	Piero Yahir Curay Chacón
Tipo de documento de identidad	DNI
Número de documento de identidad	71528950
URL de ORCID	https://orcid.org/0009-0006-2210-4064
Datos de asesor	
Nombres y apellidos	Santiago Domingo Moquillaza Henriquez
Tipo de documento de identidad	DNI
Número de documento de identidad	08280889
URL de ORCID	https://orcid.org/0000-0001-9531-881X
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Ana María Huayna Dueñas
Tipo de documento	DNI
Número de documento de identidad	06017183
Miembro del jurado 1	
Nombres y apellidos	Jorge Guerra Guerra
Tipo de documento	DNI
Número de documento de identidad	08473333
Datos de investigación	
Línea de investigación	No aplica

Grupo de investigación	No aplica
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	País: Perú Departamento: Lima Provincia: Lima Distrito: Cercado de Lima Jr. Carlos Amezaga No. 375 Universidad Nacional Mayor de San Marcos Latitud: -12.0564232 Longitud: -77.0843327
Año o rango de años en que se realizó la investigación	2025
URL de disciplinas OCDE	2.02.04 -- Ingeniería de sistemas y comunicaciones https://purl.org/pe-repo/ocde/ford#2.02.04



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
FACULTAD DE INGENIERIA DE SISTEMAS E INFORMATICA
Escuela Profesional de Ingeniería de Sistemas

Acta Virtual de Sustentación de Tesis

Siendo las 15:00 horas del día 07 de julio del año 2025, se reunieron virtualmente los docentes designados como miembros de Jurado de Tesis, presidido por la Mg. Ana María Huayna Dueñas, Mg. Jorge Guerra Guerra (Miembro) y el Mg. Santiago Moquillaza Henriquez (Miembro Asesor), usando la plataforma Meet (<https://meet.google.com/jna-kcfr-opq>), para la sustentación Virtual de la tesis Intitulada: **“APLICACIÓN DE MODELOS DE TRANSFERENCIA DE APRENDIZAJE PARA LA DETECCIÓN DE CASOS DE CATARATA”**, del Bachiller: Piero Yahir CURAY CHACÓN; para obtener el Título Profesional de Ingeniero de Sistemas.

Acto seguido de la exposición de la Tesis, la Presidenta invitó al Bachiller a responder las preguntas formuladas por los Miembros del Jurado.

El Bachiller, en el curso de sus intervenciones demostró pleno dominio del tema, al responder con acierto y fluidez las preguntas formuladas por los señores miembros del Jurado.

Finalmente habiéndose efectuado la calificación correspondiente por los miembros del Jurado, el bachiller obtuvo la nota de (...15...)

A continuación, la Presidenta del Jurado Mg. Ana María Huayna Dueñas, declara al Bachiller **Ingeniero de Sistemas**.

Siendo ...16:21.: horas, se levantó la sesión.

Mg. Ana María Huayna Dueñas
Presidente

Mg. Jorge Guerra Guerra
Miembro

Mg. Santiago Moquillaza Henriquez
Miembro Asesor



CERTIFICADO DE SIMILITUD

Yo, SANTIAGO DOMINGO MOQUILLAZA HENRIQUEZ en mi condición de asesor acreditado con el Dictamen de la UPG o de la EP N° 000010-2024-MP-FISI/UNMSM de la tesis/ monografía/ informe de investigación/trabajo académico, cuyo título es **APLICACIÓN DE MODELOS DE TRANSFERENCIA DE APRENDIZAJE PARA LA DETECCIÓN DE CASOS DE CATARATA**, presentado por el bachiller/ magíster/ egresado/ licenciado/ estudiante **CURAY CHACON, PIERO YAHIR** para optar el grado* de/ **título***/ especialidad en **INGENIERÍA DE SISTEMAS**, CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el **10 %** de similitud, nivel **PERMITIDO** para continuar con los trámites correspondientes y para su publicación en el repositorio institucional.

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del grado*/ **título***/ especialidad* **INGENIERO DE SISTEMAS**.

Lima, 11 de agosto 2025

Firma del Asesor

Nombres y Apellidos del asesor: Santiago Domingo Moquillaza Henriquez_

DNI: 08280889

* Según corresponda: bachiller, magíster, doctor; título profesional o título de segunda especialidad

La vigencia del presente certificado corresponde a un período de noventa (90) días desde la fecha de su

emisión; siempre que no haya observaciones en el texto durante el proceso de revisión por el jurado informante o surja alguna contingencia.

Dedicatoria

Este trabajo está dedicado a mi familia, amigos, profesores y compañeros como muestra de mi agradecimiento por la guía, la confianza, el tiempo y el cariño brindado en estos años.

Agradecimientos

Expreso mi gratitud al profesor Santiago Moquillaza por su apoyo, guía y compartir su valiosa experiencia durante el proceso de elaboración de este proyecto de tesis.

También, deseo agradecer a la Facultad de Ingeniería de Sistemas e Informática.

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS
APLICACIÓN DE MODELOS DE TRANSFERENCIA DE APRENDIZAJE PARA
LA DETECCIÓN DE CASOS DE CATARATA

Autor: CURAY CHACÓN, Piero Yahir
Asesor: MOQUILLAZA HENRÍQUEZ, Santiago Domingo
Título: Tesis para optar el Título Profesional de Ingeniero de Sistemas
Fecha: Abril, 2025

Resumen

En países de bajos y medianos ingresos, la detección de cataratas es fundamental debido a su alta incidencia como causa de ceguera y la falta de acceso a métodos de detección adecuados. En este estudio, se utilizó Transferencia de Aprendizaje con modelos como InceptionV3, VGG-19 y ResNet50, junto con la metodología CRISP-DM, para crear un eficiente modelo de detección de cataratas. Se evaluaron estos modelos utilizando un conjunto de datos de imágenes oculares que abarcaba diversas enfermedades oculares (ODIR) como casos de cataratas. Se encontró que VGG-19 tenía la mayor precisión, con un 97%, superando a InceptionV3 (91%) y ResNet50 (94%). Asimismo, se implementó con éxito en una aplicación web desarrollada en Python y Flask para una interacción amigable con los usuarios. En resumen, este estudio demostró que la Transferencia de Aprendizaje es efectiva y respaldada por CRISP-DM, y la aplicación web resultante promete mejorar la calidad de vida de quienes enfrentan cataratas.

Palabras claves: Detección de catarata, modelo pre entrenado, fondo ocular, transferencia de aprendizaje, aplicación

MAJOR UNIVERSITY OF SAN MARCOS
FACULTY OF SYSTEMS ENGINEERING
PROFESSIONAL SCHOOL OF SYSTEMS ENGINEERING
APPLICATION OF TRANSFER LEARNING MODELS FOR THE DETECTION
OF CATARACT CASES

Author: CURAY CHACÓN, Piero Yahir
Adviser: MOQUILLAZA HENRÍQUEZ, Santiago Domingo
Title: Thesis, to choose the Professional Title System Engineer
Date: April 2025

Abstract

In low and middle-income countries, the detection of cataracts is crucial due to their high incidence as a cause of blindness and the lack of access to adequate detection methods. In this study, Transfer Learning was used with models such as InceptionV3, VGG-19, and ResNet50, along with the CRISP-DM methodology, to create an efficient cataract detection model. These models were evaluated using a dataset of ocular images that encompassed various eye diseases (ODIR), including cases of cataracts. It was found that VGG-19 had the highest accuracy, at 97%, surpassing InceptionV3 (91%) and ResNet50 (94%). Furthermore, it was successfully implemented in a web application developed in Python and Flask for user-friendly interaction. In summary, this study demonstrated the effectiveness of Transfer Learning, supported by CRISP-DM, and the resulting web application holds promise for improving the quality of life for those facing cataracts.

Keywords: Cataract detection, pre-trained model, ocular background, transfer learning, application

Índice

Dedicatoria.....	V
Agradecimientos.....	VI
Resumen.....	VII
Abstract.....	VIII
Índice.....	X
Lista de tablas.....	XIII
Lista de figuras.....	XIV
Introducción	1
1 Capítulo I – Planteamiento metodológico.....	2
1.1 Planteamiento del problema.....	2
1.1.1 Problema principal	2
1.1.2 Problemas específicos	2
1.2 Objetivos de la Investigación	3
1.2.1 General.....	3
1.2.2 Específicos	3
1.3 Justificación	4
1.3.1 Justificación teórica	4
1.3.2 Justificación práctica.....	4
1.4 Alcances.....	4
1.5 Limitaciones.....	5
1.6 Identificación de variables	6
1.6.1 Variable independiente	6
1.6.2 Variable dependiente	6
1.7 Universo y Muestra.....	6
1.7.1 Universo.....	6
1.7.2 Muestra	6
1.8 Metodología y diseño de investigación	6
1.8.1 Tipo de investigación.....	6
1.8.2 Nivel de investigación	7
1.8.3 Métodos de la investigación	7
1.8.4 Diseño de la investigación	7
1.9 Organización de la tesis	8
2 Capítulo II – Marco teórico.....	9
2.1 Catarata	9
2.2 Imagen de fondo ocular	10
2.3 Preprocesamiento de imágenes.....	11
2.4 Machine Learning.....	12
2.5 Deep Learning.....	12

2.6	Convolutional Neural Network.....	12
2.7	Transferencia de Aprendizaje	13
2.8	VGGNet.....	14
2.9	InceptionV3	15
2.10	ResNet 16	
2.11	Función de Activación Sigmoidal.....	17
2.12	Adam Optimization.....	17
2.13	Binary Cross Entropy.....	18
2.14	Dropout	18
2.15	Batch Normalization.....	19
2.16	Metodología CRISP-DM.....	19
2.17	Matriz de Confusión	20
2.18	Reporte de Clasificación.....	22
2.19	Coeficiente V de Aiken.....	22
3	Capítulo III – Estado del Arte metodológico.....	23
3.1	Metodología de la revisión	23
3.1.1	Planificación.....	23
3.1.2	Palabras claves	23
3.1.3	Inclusión.....	24
3.1.4	Exclusión.....	24
3.2	Implementación/Revisión de artículos.....	24
3.3	Resultados de la revisión	34
3.4	Herramientas comerciales para el diagnóstico automatizado de enfermedades oculares	42
4	Capítulo IV – Implementación	43
4.1	Aporte del estudio.....	43
4.2	Estrategia para el diseño de la solución y selección de tecnologías	44
4.3	Metodología de la solución.....	46
4.4	Herramientas.....	48
4.4.1	Hardware.....	48
4.4.2	Software	48
4.5	Validación del Sistema	50
4.6	Desarrollo de la solución	51
4.6.1	Comprensión del problema	51
4.6.2	Comprensión de los datos	52
4.6.3	Preparación de los datos.....	54
4.6.4	Modelización.....	59
4.6.5	Evaluación.....	63
4.6.6	Resultados experimentales.....	64
4.6.7	Validación y Comparativa	69
4.6.8	Implementación o despliegue del modelo.....	75
5	Capítulo V – Discusión, Conclusiones y recomendaciones.....	78
5.1	Discusión	78
5.2	Conclusiones y recomendaciones	79

6	Referencias.....	82
7	Anexos	A
	ANEXO 1: Publicaciones seleccionadas y preguntas de investigación.....	A
	ANEXO 2: Matriz de Consistencia.....	E
	ANEXO 3: Matriz de Operacionalización.....	G
	ANEXO 4: Cuestionarios de Juicio de Expertos	H

Lista de tablas

Tabla 1 Desempeño de los modelos de Afroze y su equipo	27
Tabla 2 Bases de Datos o Editoriales asociadas a las preguntas de investigación	35
Tabla 3 Modelos de Transferencia de Aprendizaje para detectar enfermedades oculares ..	35
Tabla 4 Técnicas de preprocesamiento de imágenes	37
Tabla 5 Metodologías de los artículos revisados	39
Tabla 6 Métricas para evaluar el desempeño de los modelos	40
Tabla 7 Herramientas para el Desarrollo de modelos de Transferencia de Aprendizaje	48
Tabla 8 Herramientas para el despliegue del modelo en una aplicación web	49
Tabla 9 Reporte de clasificación del VGG-19	65
Tabla 10 Reporte de clasificación del InceptionV3	67
Tabla 11 Reporte de clasificación del ResNet50	68
Tabla 12 Resumen comparativo del desempeño de los modelos	69
Tabla 13 Resumen de resultados de la validación	71
Tabla 14 Benchmarking de resultados propios vs otros autores	72
Tabla 15 Resumen de respuestas de los expertos	73
Tabla 16 Escala de respuestas de juicio de expertos	74
Tabla 17 Cálculo V de Aiken	74

Lista de figuras

Figura 1 Niveles de evaluación de cataratas en imagen de fondo ocular	10
Figura 2 Ejemplos de preprocesamiento de imágenes.....	11
Figura 3 Diferentes procesos de Aprendizaje entre el Machine Learning Tradicional (a) y el Transfer Learning (b).....	14
Figura 4 Arquitectura VGG19.....	14
Figura 5 Diagrama de alto nivel del Inception v3	15
Figura 6 Arquitectura ResNet.....	16
Figura 7 Función Sigmoidal	17
Figura 8 Modelo de la metodología CRISP-DM.....	20
Figura 9 Plantilla de una Matriz de Confusión.....	21
Figura 10 Cantidad de artículos por pregunta de investigación	34
Figura 11 Modelos de Transferencia de Aprendizaje para detectar enfermedades oculares	37
Figura 12 Técnicas de preprocesamiento de imágenes	39
Figura 13 Metodologías para desarrollar un modelo de transferencia de aprendizaje	40
Figura 14 Métricas de desempeño más usados.....	41
Figura 15 Arquitectura de la aplicación web de detección de casos de cataratas	46
Figura 16 Logo Jupyter Notebook.....	48
Figura 17 Logo Python	48
Figura 18 Logo TensorFlow	49
Figura 19 Logo Scikit Learn.....	49
Figura 20 Logo Visual Studio Code	49
Figura 21 Logo Python + Flask	49
Figura 22 Logo HTML + CSS + JavaScript.....	50
Figura 23 Logo Conda.....	50
Figura 24 Imagen de fondo ocular de un caso normal	53
Figura 25 Imagen de fondo ocular de un caso de catarata.....	54
Figura 26 Casos de catarata en dataset ODIR	55
Figura 27 Casos sanos en dataset ODIR.....	55
Figura 28 Cantidad de casos sanos y casos de catarata del dataset ODIR	56
Figura 29 Almacenando en dataset.....	56
Figura 30 Transformaciones de Image Rotate e Image Flip	58
Figura 31 Transformaciones de Grayscale y Contrast Enhancement	58
Figura 32 Muestras del nuevo dataset elaborado.....	59
Figura 33 Partición del dataset para el entrenamiento y evaluación	60
Figura 34 Instancia del modelo preentrenado VGG-19.....	61
Figura 35 Resumen del modelo preentrenado VGG-19	61
Figura 36 Compilación y checkpoint del modelo preentrenado VGG-19.....	62
Figura 37 Entrenamiento del modelo	63
Figura 38 Predicciones del modelo VGG19	64
Figura 39 Accuracy score del VGG-19	64
Figura 40 Matriz de Confusión del VGG-19.....	65
Figura 41 Accuracy score del InceptionV3	66
Figura 42 Matriz de Confusión del InceptionV3	66
Figura 43 Accuracy score del ResNet50	67
Figura 44 Matriz de Confusión del ResNet50	68
Figura 45 Muestras del nuevo dataset para la validación	70

Figura 46 Seleccionando imágenes para la validación.....	70
Figura 47 Predicciones del modelo VGG19 en la Validación.....	71
Figura 48 Exportación del modelo con mejor desempeño	75
Figura 49 Despliegue del modelo VGG-19 en aplicación web	75
Figura 50 Caso positivo de catarata en la aplicación web	76
Figura 51 Caso negativo de catarata en la aplicación web	76
Figura 52 Versión colectiva de la aplicación web	77

Introducción

La deficiencia visual, según un informe de la Organización Mundial de la Salud, es un problema que padecen alrededor de 2,200 millones de personas alrededor del mundo (OMS, 2020).

Esta alteración suele originarse cuando afecciones oculares como las cataratas o el glaucoma no se tratan oportunamente, derivando en deficiencia visual leve, moderada o ceguera. Asimismo, las consecuencias en adultos con deficiencia visual (Brown & Barrett, 2011) suelen presentar tasas bajas de participación y productividad en el ámbito laboral, al mismo tiempo, que son más propensos a accidentes automovilísticos.

En los últimos años se han observado una serie de trabajos científicos acompañados de avances informáticos para la detección y prevención de deficiencias visuales (Ting et al., 2019). Estos avances informáticos que involucran algoritmos y modelos de Deep Learning han cumplido exitosamente en sus respectivos objetivos como la detección de casos de cataratas (Zhang et al., 2022), así como otras enfermedades oculares.

Por consiguiente, en el presente trabajo de tesis se propone el uso de modelos pre entrenados a través de la técnica de Transferencia de Aprendizaje, como InceptionV3, VGG19 y ResNet50, bajo la metodología CRISP-DM para crear un modelo que detecte casos de catarata. Tras comparar los desempeños de los modelos, se selecciona el que cuenta con el mejor rendimiento para implementarlo en el despliegue de una aplicación web que detecte este mal tras subir una imagen de fondo ocular.

Capítulo I – Planteamiento metodológico

1.1 Planteamiento del problema

1.1.1 Problema principal

En el Perú, la principal causa de ceguera es la catarata (58%), seguida por el glaucoma, y la degeneración macular relacionada a la edad (Minsa, 2022). La ceguera en el Perú se ve afectada por el acceso limitado a la atención oftalmológica, al envejecimiento de la población los cuales demuestran una alta demanda de estos servicios y a otros factores como los cambios de estilo de vida; además, otras barreras identificadas fueron que las personas afectadas no eran conscientes que el tratamiento es posible o el temor de realizarse la cirugía y obtener malos resultados (Campos et al., 2014). Por ello, resulta necesario el desarrollo de programas destinados a la detección y tratamiento de estas afecciones para evitar el deterioro de la agudeza visual. Si bien se han incrementado los avances en la atención ocular, aún hay necesidades mundiales que van desde ineficientes sistemas de salud hasta la falta de trabajos de investigación. (OMS, 2020) Entre ellas se destaca la necesidad de mejorar la capacidad de las encuestas para detectar causas evitables de ceguera en países de bajos y medianos ingresos.

Ante lo planteado se formula el problema:

- ¿Cómo se puede aplicar un modelo de transferencia de aprendizaje para la detección de casos de cataratas?

1.1.2 Problemas específicos

- ¿Qué tipo de imágenes médicas son aplicables en un modelo de transferencia de aprendizaje para la detección de cataratas?
- ¿Con qué precisión el modelo de transferencia de aprendizaje detectará un

caso de catarata en una imagen de fondo ocular?

- ¿Qué metodología se usará de guía para el desarrollo del modelo de Transferencia de Aprendizaje destinado a la detección de catarata?
- ¿Cómo se implementará el modelo de Transferencia de Aprendizaje detector de catarata para que esté disponible para el manejo del usuario final?

1.2 Objetivos de la Investigación

1.2.1 General

- Desarrollar un modelo basado en transferencia de aprendizaje que permita detectar casos de cataratas a partir de imágenes médicas oculares.

1.2.2 Específicos

- Encontrar un *dataset* de imágenes médicas oculares de un repositorio confiable destinado a la detección de cataratas mediante un modelo de transferencia de aprendizaje.
- Alcanzar una precisión mínima del 85% en la detección de casos de cataratas en imágenes médicas a través un modelo de transferencia de aprendizaje.
- Determinar la metodología a usar como guía para el desarrollo de un modelo de transferencia de aprendizaje destinada a la detección de casos de catarata.
- Desplegar el modelo de Transferencia de Aprendizaje para la detección de casos de cataratas en una aplicación web para el manejo del usuario final.

1.3 Justificación

1.3.1 Justificación teórica

El presente trabajo de tesis tiene como propósito contribuir al campo de la inteligencia artificial aplicada a la salud visual, específicamente en el área de diagnóstico automatizado de enfermedades oculares. Se plantea la aplicación de modelos de Transferencia de Aprendizaje como una alternativa eficiente para el reconocimiento de cataratas a partir de imágenes oculares, evaluando su precisión y viabilidad. Este estudio busca ofrecer una base metodológica y técnica que pueda ser utilizada como referencia para futuras investigaciones relacionadas con la detección temprana de patologías visuales mediante técnicas de aprendizaje profundo.

1.3.2 Justificación práctica

Esta investigación surge como respuesta a la necesidad de mejorar el acceso a diagnósticos oftalmológicos en regiones con recursos limitados, donde la catarata continúa siendo una causa predominante de ceguera prevenible. La propuesta de un sistema automatizado basado en modelos preentrenados permite apoyar la labor médica mediante una herramienta accesible y de rápida implementación. De esta forma, se busca fortalecer la prevención y detección oportuna de cataratas, optimizando los recursos disponibles y reduciendo la brecha tecnológica en el sector salud.

1.4 Alcances

El presente trabajo de tesis tiene como objetivo desarrollar modelos de redes neuronales convolucionales capaces de detectar casos de cataratas a partir de imágenes del fondo ocular (*fundus*), mediante la aplicación de técnicas de Transferencia de Aprendizaje.

Esta metodología permite acelerar el proceso de entrenamiento y mejorar el rendimiento en comparación con modelos contruidos desde cero.

Para ello, se utilizarán modelos preentrenados como VGG19, ResNet50 e InceptionV3, los cuales serán entrenados y ajustados utilizando un conjunto de imágenes oftalmológicas recopiladas de una institución médica china. El desempeño de cada modelo será evaluado a través de métricas estándar como precisión (*precision*), exactitud (*accuracy*), sensibilidad (*recall*), puntuación F1 (*f1-score*) y análisis de la matriz de confusión. Adicionalmente, se contempla incorporar una validación cualitativa mediante el juicio de expertos en salud visual.

Finalmente, el modelo que obtenga el mejor desempeño será implementado en una aplicación web de uso sencillo, destinada a apoyar a profesionales de la salud como médicos generales, oftalmólogos en formación o técnicos especializados en atención primaria, brindando así una herramienta complementaria para el tamizaje temprano de cataratas.

1.5 Limitaciones

- Empleo de conjuntos de datos o *datasets* públicos de casos de cataratas de fondo ocular.
- Tamaño de los *datasets* de los casos de cataratas.
- Empleo de herramientas informáticas accesibles y gratuitas.
- Limitaciones en la potencia de hardware para realizar el entrenamiento de los modelos de Transferencia de Aprendizaje.

1.6 Identificación de variables

1.6.1 Variable independiente

- Modelos de transferencia de aprendizaje

1.6.2 Variable dependiente

- Detección de casos de catarata.

1.7 Universo y Muestra

1.7.1 Universo

El universo es una serie de imágenes de fondo ocular en el dataset ODIR, que abarca patologías oculares provenientes de 5000 pacientes atendidos de instituciones médicas de China.

1.7.2 Muestra

La muestra será 1094 imágenes clasificadas entre casos sanos (500) y con cataratas (594).

1.8 Metodología y diseño de investigación

1.8.1 Tipo de investigación

El tipo de investigación es **aplicada-predictiva**, ya que combina un enfoque práctico con el objetivo de anticipar resultados específicos (Huaman et al, 2022). En el nivel aplicado se busca resolver un problema real en el ámbito de la oftalmología mediante el desarrollo e implementación de un modelo de Transferencia de Aprendizaje para la detección de cataratas en imágenes médicas, ofreciendo así una solución tecnológica

innovadora para mejorar la salud visual. Por otro lado, el nivel predictivo se caracteriza por la capacidad del modelo desarrollado para anticipar la presencia de cataratas en nuevas imágenes oculares con base en patrones aprendidos durante el entrenamiento.

1.8.2 Nivel de investigación

El nivel de investigación es **explicativo-exploratorio**, ya que combina el análisis de relaciones causa-efecto entre las variables independiente y dependiente, con un enfoque en la exploración inicial de recursos y datos relevantes (Huaman et al, 2022). La parte explicativa se centra en comprender cómo las modificaciones realizadas en las imágenes médicas influyen en la precisión del modelo para detectar cataratas. Por su parte, el enfoque exploratorio abarca la búsqueda, selección y análisis de datasets de imágenes oculares utilizados para el entrenamiento y evaluación de los modelos de transferencia de aprendizaje.

1.8.3 Métodos de la investigación

Se usó métodos **cuantitativos** (Hernández-Sampieri & Mendoza, 2006), basados en el análisis de datos numéricos derivados de métricas de desempeño (precisión, sensibilidad, F1-Score). Adicionalmente, se emplearon métodos **experimentales** para evaluar los efectos de las transformaciones de los datos (imágenes) y configuraciones del modelo sobre el rendimiento del sistema.

1.8.4 Diseño de la investigación

El diseño de investigación empleado en este estudio es **cuasi-experimental**, ya que involucra la manipulación de imágenes médicas modificadas mediante técnicas de *data augmentation*. Este enfoque permite observar y evaluar cómo dichas modificaciones

influyen en el rendimiento de los modelos de Transferencia de Aprendizaje utilizados para la detección de cataratas. Aunque no se cuenta con un control total de las condiciones experimentales, este diseño es adecuado para analizar los efectos de los cambios introducidos en un entorno aplicado (Hernández-Sampieri & Mendoza, 2006).

1.9 Organización de la tesis

Este trabajo de tesis está organizado en seis capítulos, que se resumen de la siguiente manera:

El primer capítulo, que es el presente, expone el planteamiento del problema y objetivos de la investigación, los cuales están destinados a ser resueltos durante la ejecución del proyecto.

El segundo capítulo abarca el marco teórico, donde se explican los conceptos y términos relevantes para el tema de la tesis, que se utilizarán a lo largo del documento.

El tercer capítulo analiza el estado del arte metodológico, revisando investigaciones previas relacionadas con el tema propuesto. En este apartado, se exploran diversos métodos, modelos, técnicas e imágenes médicas para abordar las problemáticas planteadas en el estudio.

El cuarto capítulo se dedica a presentar el aporte del trabajo y la implementación, ofreciendo un detallado relato sobre la construcción de la solución propuesta.

Por último, el quinto capítulo engloba la discusión de los resultados del trabajo, las conclusiones obtenidas y las recomendaciones sugeridas.

Capítulo II – Marco teórico

El presente capítulo se preparó un glosario de términos que son usados durante el trabajo expuesto. Las definiciones fueron rescatadas de fuentes como artículos científicos y sitios web confiables.

2.1 Catarata

La catarata es una enfermedad ocular que involucra la opacidad del cristalino del ojo que causa una visión nublada y dificulta tareas como leer, conducir o reconocer expresiones faciales. Si bien la mayoría de las cataratas se desarrollan de forma progresiva y no afectan significativamente la visión en etapas tempranas, con el tiempo, pueden alterar la visión de manera significativa (Mayo Clinic). En etapas tempranas, se pueden usar anteojos y una iluminación más intensa para contrarrestar las cataratas. Sin embargo, si la visión se deteriora lo suficiente, se puede requerir de una cirugía de cataratas.

La Organización Mundial de la Salud (OMS) señala que la catarata constituye la principal causa de ceguera a nivel global, siendo responsable de aproximadamente 94 millones de casos en todo el mundo (OMS, 2022). La incidencia de catarata se ha relacionado con la edad, pero también puede estar relacionado con factores genéticos, ambientales y de estilos de vida.

En la actualidad, el procedimiento más efectivo para la catarata es la cirugía de reemplazo de lente. De acuerdo con un estudio publicado por la revista científica *The Lancet*, la cirugía de catarata es un procedimiento seguro y efectivo que puede mejorar significativamente la calidad de vida de los pacientes (Bourne et al, 2017). Sin embargo, el acceso a la cirugía sigue siendo limitado en muchas partes del mundo, especialmente en países de bajos ingresos.

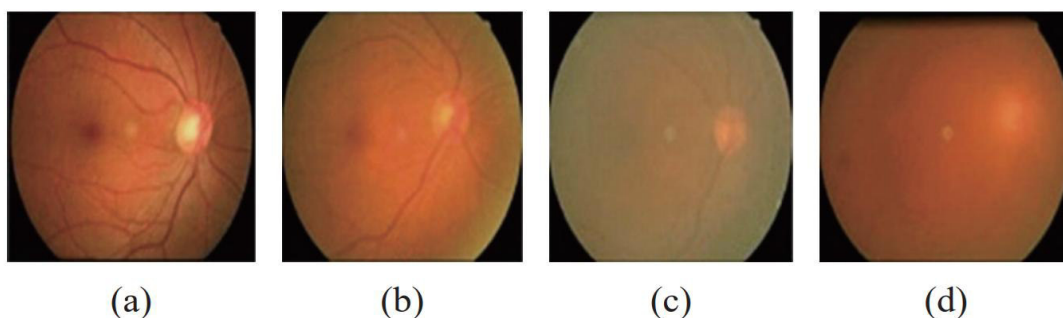
2.2 Imagen de fondo ocular

Estas imágenes, también conocidas como *fundus images*, son un tipo de imagen ocular altamente especializado que permite visualizar el revestimiento interno del ojo y las estructuras situadas en la parte posterior del mismo (Zhang et al., 2022). Este tipo de imágenes se obtienen usando una cámara de fondo ocular junto con un microscopio de baja potencia, generalmente, operada por oftalmólogos u operadores profesionales.

Mediante este tipo de imágenes se pueden detectar y clasificar casos de cataratas como lo demuestra Xu y su equipo, ellos proponen un sistema de clasificación de cataratas basados en imágenes de fondo ocular que se centra en observar el nivel de desenfoque (Xu et al, 2010). Para ello se emplea cinco niveles de evaluación mediante el análisis del grado de desenfoque en las imágenes del fondo de ojo como se muestra en la **Figura 1**: Grado 0: Claro; Grado 1: Borrosidad en la región de los vasos sanguíneos pequeños; Grado 2: Borrosidad en las ramas más grandes de la vena o arteria retiniana; Grado 3: Borrosidad en la región del disco óptico; Grado 4: Borrosidad en toda la imagen del fondo de ojo (Zhang et al., 2022).

Figura 1

Niveles de evaluación de cataratas en imagen de fondo ocular



Nota. (a) Normal; (b) Inmaduro; (c) Maduro; (d) Hiper maduro. Tomado de *Machine Learning for Cataract Classification/Grading on Ophthalmic Imaging Modalities: A Survey* por Zhang et al, 2022.

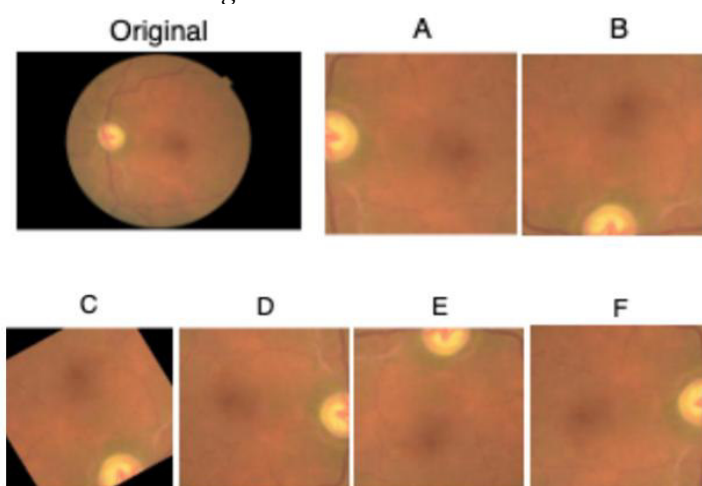
2.3 Preprocesamiento de imágenes

El preprocesamiento de imágenes retinianas desempeña un papel crucial, ya que mejora la calidad de las imágenes y facilita una visualización más efectiva (Sarki et al., 2020). Entre los métodos de preprocesamiento más recomendados están:

- **Image Enhancement:** Extrae el canal verde de una imagen RGB para obtener información relevante. Luego, se mejora el contraste mediante CLAHE (*Contrast Limited Adaptive Histogram Equalization*), se ajusta la iluminación para aumentar el brillo y se elimina el ruido con un filtro gaussiano, logrando así imágenes más suaves y de mayor calidad.
- **Image Augmentation:** La aumentación de imágenes es una técnica que amplía un conjunto de datos de entrenamiento al crear variaciones artificiales de imágenes, como se visualiza en la **Figura 2**. Esto mejora la capacidad de los modelos de aprendizaje profundo para aplicar su conocimiento a nuevas imágenes mediante transformaciones geométricas como rotación, redimensión, espejeo, volteo, entre otras.

Figura 2

Ejemplos de preprocesamiento de imágenes



Nota. Original; (A) Crop; (B) Rotación 90°; (C) Rotación 120°; (D) Rotación 180°; (E) Rotación 270°; (F) Espejo. Tomado de *Early detection of diabetic eye disease from fundus images with deep learning* por Sarki et al., 2020

2.4 Machine Learning

También conocido como Aprendizaje Automático, es un campo de estudio que se centra en el diseño y desarrollo de algoritmos capaces de aprender y realizar predicciones o decisiones basadas en datos, sin estar programados explícitamente (Fenner, 2022).

Un modelo de aprendizaje automático es un programa informático que aprende de una serie de datos y puede hacer predicciones o tomar decisiones en base en ellos. El modelo se construye a partir del entrenamiento de datos para producir una función de predicción o de decisión (Jordan & Mitchell, 2015). Asimismo, es pertinente el uso de métricas para evaluar el rendimiento del modelo de acuerdo con la problemática que buscamos resolver siendo las comunes los problemas de regresión o toma de decisiones (Fayrix).

2.5 Deep Learning

El Deep Learning (DL) o Aprendizaje Profundo se fundamenta en los principios teóricos de las redes neuronales artificiales clásicas, pero se distingue por el uso de un gran número de neuronas y capas ocultas. Asimismo, incorpora nuevos modelos y paradigmas de entrenamiento que le permiten adaptarse con mayor eficacia y extraer características de conjuntos de datos altamente complejos (Sarmiento-Ramos, 2020).

2.6 Convolutional Neural Network

También conocidas como redes neuronales convolucionales (CNN, por sus siglas en inglés), estas arquitecturas están inspiradas en el funcionamiento de la corteza visual del cerebro, la cual opera mediante mapas de campos receptivos locales. A diferencia de las redes neuronales multicapas tradicionales (DMLP), en las CNN cada neurona se conecta únicamente a una región específica de la capa siguiente, lo que permite procesar

información visual de forma más eficiente (Sarmiento-Ramos, 2020).

Estas redes aprovechan la estructura jerárquica de los datos para detectar características simples que, combinadas progresivamente, permiten reconocer patrones complejos. El término "convolucional" proviene del operador matemático de convolución, utilizado para facilitar y optimizar operaciones dentro de la red.

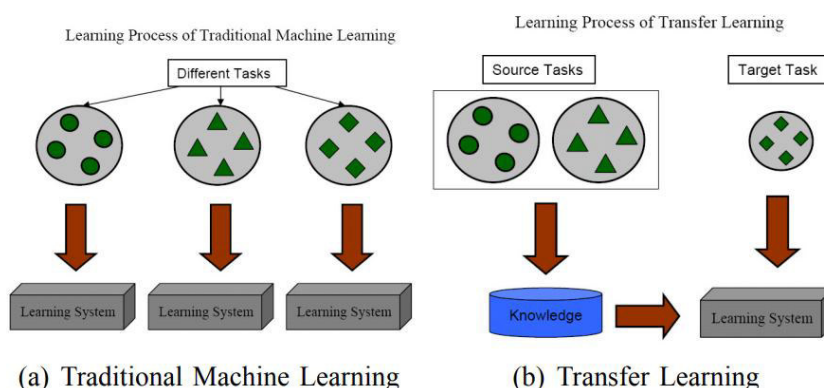
En una CNN, los datos de entrada son transformados por filtros que se aplican repetidamente a lo largo de las distintas capas, reutilizando los mismos parámetros en múltiples posiciones, lo que contribuye a reducir la cantidad total de parámetros del modelo. Este tipo de redes es ampliamente utilizado en el campo de la visión por computadora, particularmente en tareas como la clasificación y segmentación de imágenes (Sarmiento-Ramos, 2020).

2.7 Transferencia de Aprendizaje

También conocido como *Transfer Learning*, se refiere a la técnica de aprendizaje automático en la que se emplea un modelo pre entrenado para resolver un problema o una nueva tarea (Pan & Yang, 2010). El modelo preentrenado ha sido entrenado previamente con un conjunto amplio de datos para reconocer patrones generales útiles para tareas relacionadas al problema de interés. El uso de esta técnica permite reducir la cantidad de datos de entrenamiento y el tiempo requerido para la nueva tarea significativamente, asimismo, que conlleva a un mejor desempeño que entrenar un modelo desde cero. El enfoque de Transferencia de Aprendizaje se emplea frecuentemente en tareas como el reconocimiento de imágenes y el procesamiento del lenguaje natural. Esta técnica se compara con el método de *Machine Learning* tradicional en la **Figura 3**.

Figura 3

Diferentes procesos de aprendizaje entre el machine learning tradicional (a) y el transfer learning (b)



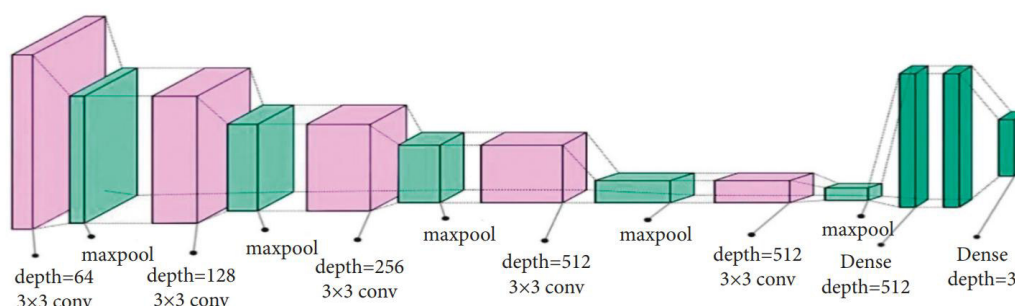
Nota. Tomado de *A survey on transfer learning* por Pan & Yang, 2010.

2.8 VGGNet

Por lo general, el modelo VGGNet se ha desarrollado para lidiar con grandes volúmenes de datos que incluyen imágenes, y se reconoce como una parte de Imagenet. (Vijayaraghavan, 2022). De igual manera, se le considera una de las técnicas más destacadas en la construcción de componentes fundamentales de redes neuronales convolucionales (CNN). Su característica distintiva es su capacidad para aprender de antemano las características genéricas de las imágenes y procesarlas de manera eficaz. La estructura de dicho modelo se ilustra en la **Figura 4**.

Figura 4

Arquitectura VGG19



Nota. Tomado de *Deep Learning for Ocular Disease Recognition: An Inner-Class Balance* por Khan et al., 2022

Entre los tipos más comunes se encuentran el VGG16 y el VGG19, ambos se diferencian por las capas de procesamiento, en el caso del último cuenta con 16 capas convolucionales, 1 clasificador *softmax* y 2 capas completamente conectadas. (Vijayaraghavan, 2022).

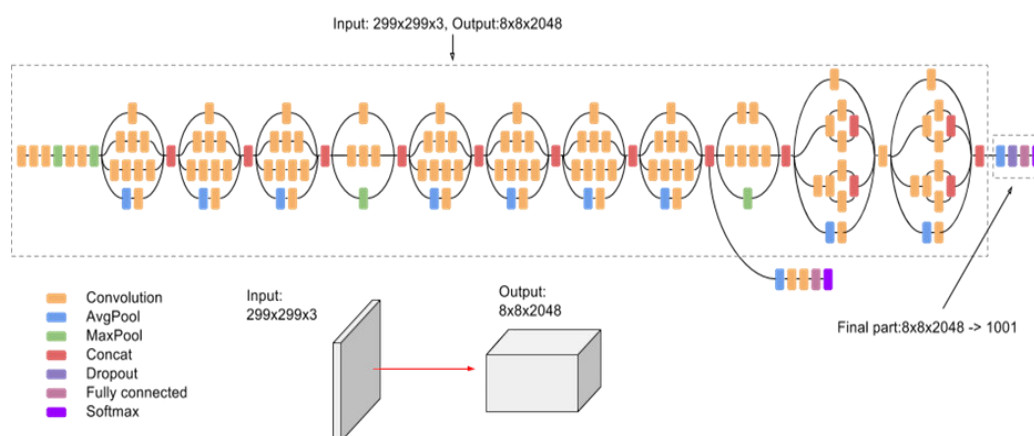
2.9 InceptionV3

Inception v3 es un modelo avanzado de detección de imágenes que ha alcanzado una precisión superior al 78.1% en el conjunto de datos de *ImageNet* (Google Cloud, 2023). Este modelo representa una síntesis de múltiples ideas desarrolladas por diversos investigadores a lo largo de los años.

Su arquitectura está compuesta por bloques de construcción simétricos y asimétricos que incluyen operaciones como convoluciones, reducciones promedio y máximas, concatenaciones, *dropout* y capas completamente conectadas, como se ilustra en la **Figura 5**. Asimismo, el modelo utiliza ampliamente la normalización por lotes, aplicada a las entradas de activación, para mejorar el entrenamiento. La función de pérdida se calcula mediante la función *softmax*.

Figura 5

Diagrama de alto nivel del Inception v3



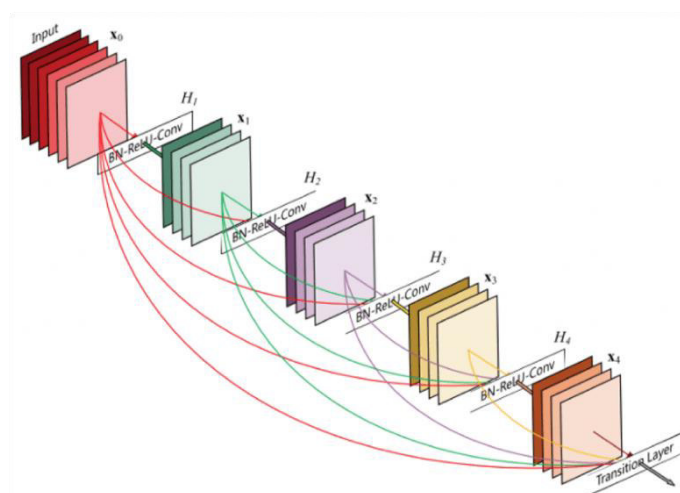
Nota. Tomado de *Guía avanzada de Inception v3* por Google Cloud, 2023

2.10 ResNet

Las redes neuronales residuales, o ResNet, son una variante destacada de las redes neuronales artificiales que se inspiran en la estructura de la corteza cerebral. Utilizan conexiones de salto entre capas, conocidas como atajos, para superar problemas como la desaparición de gradientes durante el entrenamiento (Data Science Team, 2020). Estos modelos suelen incorporar saltos de dos o tres capas, normalización por lotes y no linealidad, y en ocasiones emplean matrices de peso adicionales, denominadas "Redes de autopistas", para aprender los pesos de los saltos como se visualiza en la **Figura 6**. Esta arquitectura simplifica la red durante el entrenamiento inicial, acelera el aprendizaje y, a medida que avanza el entrenamiento, permite que las capas se vuelvan más profundas y eficientes en el aprendizaje. Sin embargo, su eficacia puede depender de la linealidad de las capas intermedias.

Figura 6

Arquitectura ResNet



Nota. Tomado de *Redes neuronales residuales - Lo que necesitas saber (ResNet)* por Data Science Team, 2020

Por otro lado, las redes neuronales sin saltos tienen más flexibilidad para explorar el espacio de características, pero son más susceptibles a perturbaciones y requieren una

mayor cantidad de datos de entrenamiento para generalizar correctamente (Data Science Team, 2020). En resumen, las *ResNet* son una estrategia ingeniosa para superar problemas de gradiente desvanecido, simplificar las redes durante el entrenamiento inicial y luego permitir una expansión más efectiva de las capas para un aprendizaje más eficiente.

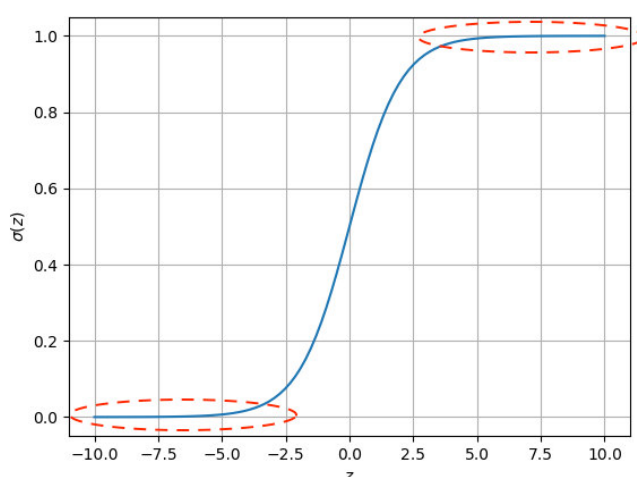
2.11 Función de Activación Sigmoidal

Entre las funciones de activación más utilizadas en proyectos de *Deep Learning* se encuentra aquella que mapea cualquier rango de valores de entrada a un rango de salida entre 0 y 1 (Sotaquirá, 2018).

Actualmente, su uso es limitado y se aplica principalmente en tareas de clasificación binaria (Sotaquirá, 2018). Esto se debe a un problema de saturación: la función alcanza el valor de 1 para entradas altas y 0 para entradas bajas, lo que genera gradientes extremadamente pequeños durante el entrenamiento con gradiente descendente, como se muestra en la **Figura 7**. Esta saturación dificulta la convergencia del algoritmo.

Figura 7

Función sigmoidal



Nota. Tomado de *La función de activación. Codificando Bits* por Sotaquirá, 2018

2.12 Adam Optimization

El método de optimización Adam es un enfoque de descenso de gradiente estocástico que se apoya en la estimación adaptable de momentos de primer y segundo orden (Kingma & Ba, 2014). Este método es eficiente desde el punto de vista computacional, no exige mucha memoria, es insensible a la reescala diagonal de gradientes y resulta particularmente apropiado para abordar problemas que involucran grandes cantidades de datos o parámetros.

2.13 Binary Cross Entropy

La entropía cruzada binaria o *binary cross entropy* es una función de pérdida utilizada en la clasificación binaria que puede aplicarse eficazmente a problemas de clasificación múltiple en redes neuronales convolucionales, destacando su relevancia en este contexto (Usha et. al, 2020).

La fórmula de *binary cross entropy* es:

$$\text{Binary Cross Entropy} = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i))$$

Fuente: (Roberts, 2023)

Donde:

- N es la cantidad de ejemplares en el *dataset*
- $p(y_i)$ es la probabilidad de uno
- $1 - p(y_i)$ es la probabilidad de cero

2.14 Dropout

Dropout es una técnica utilizada en redes neuronales para mitigar el problema del sobreajuste. En lugar de entrenar múltiples modelos de gran tamaño, este método consiste en desactivar aleatoriamente un porcentaje de neuronas durante el proceso de

entrenamiento, lo que permite simular el comportamiento combinado de varias redes de forma eficiente (Lindholm et al, 2019). Esto fomenta una mayor generalización y robustez del modelo al forzarlo a aprender de manera más equitativa. Durante la inferencia, todas las neuronas están activas, y las predicciones se basan en la combinación de conocimientos de todas ellas, reduciendo así el riesgo de sobreajuste y mejorando el rendimiento general del modelo.

2.15 Batch Normalization

La normalización por lotes es una técnica que estandariza las entradas de una capa en cada mini-lote, lo que estabiliza el proceso de aprendizaje y reduce considerablemente la cantidad de épocas necesarias para entrenar redes profundas (Browniee, 2019).

2.16 Metodología CRISP-DM

La metodología CRISP-DM es una de las más utilizadas en proyectos de minería de datos, promovida principalmente por IBM (Espinosa-Zuñiga, 2020). Ha sido ampliamente adoptada como el modelo de referencia predominante para la ejecución de proyectos orientados al análisis de datos y la extracción de conocimiento (Marbán, 2018).

Son un total de seis fases que constituyen la metodología CRISP-DM como se muestra en la Figura 8.

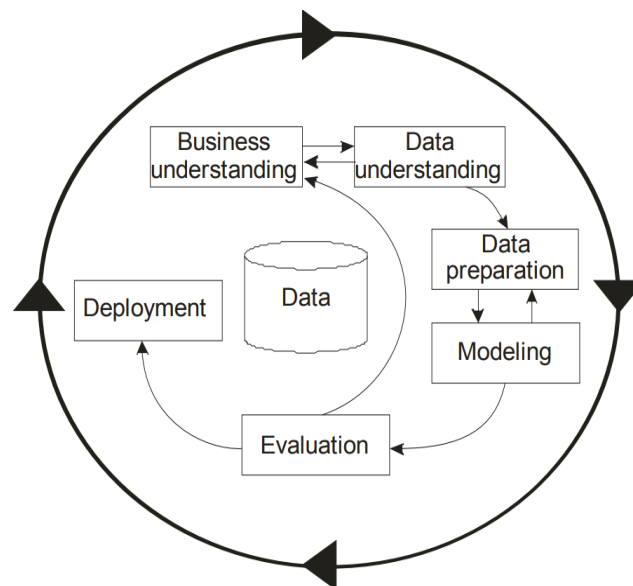
1. **Comprender el negocio o problema**, considerada la fase más crítica, implica la identificación del problema, determinar los objetivos y evaluación de la situación actual del caso.
2. **Comprender los datos**, lo cual implica la recolección y descripción de los datos, así como su exploración a través de pruebas estadísticas básicas.
3. **Preparación de los datos**, implica la selección y transformación de los datos para

ser usados en la etapa de modelado.

4. **Modelización**, en esta fase se selecciona la técnica apropiada, se divide el conjunto de entrenamiento y prueba, y se ajusta el modelo con una serie de parámetros.
5. **Evaluación del modelo**, se determina la calidad del modelo mediante el análisis de medidas estadísticas y la comparación de los resultados con trabajos previos o con expertos.
6. **Implementación del modelo**, en esta última fase se aprovecha el conocimiento obtenido mediante acciones concretas como despliegues y se realiza la documentación del proceso para su revisión y mejora futura.

Figura 8

Modelo de la metodología CRISP-DM



Nota. Tomado de *CRISP-DM 1.0 step-by-step data mining guide* por Chapman et al., 2000

2.17 Matriz de Confusión

Los modelos de clasificación con múltiples salidas categóricas pueden presentar desafíos al evaluar su rendimiento, ya que las medidas de error tradicionales calculan el error total sin identificar errores individuales en las categorías (Simplilearn, 2020). Esto puede llevar a situaciones en las que un modelo obtenga una alta precisión al predecir la

clase mayoritaria en un conjunto de datos desequilibrado, mientras ignora las clases minoritarias. En tales casos, las matrices de confusión se vuelven fundamentales, ya que permiten un análisis detallado de las predicciones por categoría, revelando posibles desequilibrios y deficiencias en la capacidad del modelo para clasificar cada clase por separado.

Figura 9

Plantilla de una matriz de confusión

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

Nota. Tomado de *What is a confusion matrix in machine learning?* por Simplilearn, 2020

Como se aprecia en la Figura 9 se pueden obtener cuatro distintas combinaciones partiendo de las predicciones y los valores reales de una clasificación:

- **Verdadero Positivo (True Positive):** Corresponde a los casos en los que el modelo predice correctamente una condición positiva, es decir, cuando la clase real y la clase predicha coinciden como positivas.
- **Falso Positivo (False Positive):** Se refiere a las ocasiones en las que el modelo clasifica incorrectamente una instancia negativa como positiva; predice una condición que en realidad no está presente.

- **Verdadero Negativo (True Negative):** Representa los casos en los que el modelo acierta al identificar una condición negativa, coincidiendo la predicción con el valor real negativo.
- **Falso Negativo (False Negative):** Ocurre cuando el modelo no detecta una condición positiva y la clasifica erróneamente como negativa, es decir, se omite una instancia que realmente es positiva.

2.18 Reporte de Clasificación

Se refiere a un informe que resume la calidad de rendimiento del modelo de la clasificación realizada por un modelo de aprendizaje automático construido. Está compuesto principalmente por cinco columnas y un total de $(N+3)$ filas. (Pintusaini, 2022). La primera columna contiene el nombre de la etiqueta de clase, seguida de los valores de las métricas *Precision*, *Recall*, *F1-Score* y *Support*. Las N filas corresponden a las N etiquetas de clase, mientras que las otras tres filas abordan la precisión global, el promedio macro y el promedio ponderado.

2.19 Coeficiente V de Aiken

El coeficiente V de Aiken es una medida estadística empleada para evaluar el nivel de consenso entre un grupo de expertos respecto a la relevancia, claridad o representatividad de los ítems que conforman un instrumento de evaluación (Aiken, 1980). Esta técnica consiste en asignar puntuaciones ordinales a cada ítem y calcular un índice que varía entre 0 y 1; cuanto más cercano esté el valor a 1, mayor será el grado de acuerdo entre los evaluadores sobre la validez del contenido. Una de sus principales fortalezas es que permite realizar una validación cuantitativa, objetiva y sencilla de instrumentos en fases iniciales de investigación, siendo especialmente útil cuando se dispone de un número reducido de jueces.

Capítulo III – Estado del Arte metodológico

Este capítulo detalla la metodología empleada para la revisión de literatura científica, así como los principales resultados obtenidos a partir del análisis de diversas fuentes académicas.

3.1 Metodología de la revisión

Para la revisión de la literatura se utilizó la metodología de Kitchenham y Charters, *Systematic Literature Review* o SLR (Kitchenham & Charters, 2007) que propone un proceso estructurado y planificado que consta de las siguientes fases:

3.1.1 Planificación

Esta etapa consistió en desarrollar preguntas de investigación, protocolos de búsqueda y palabras clave, de igual manera, de evaluar fuentes relevantes y establecer criterios de inclusión y exclusión.

- P1: ¿Qué modelos de transferencia de aprendizaje son aplicables para la detección de enfermedades oculares?
- P2: ¿Qué técnicas de preprocesamiento de imágenes son usadas para mejorar el rendimiento del modelo de transferencia de aprendizaje para la detección de enfermedades oculares?
- P3: ¿Qué metodologías son usadas para el desarrollo de un modelo de transferencia de aprendizaje para la detección de enfermedades oculares?
- P4: ¿Qué métricas son usadas para evaluar el desempeño de un modelo de transferencia de aprendizaje para la detección de enfermedades oculares?

3.1.2 Palabras claves

(“classification” OR “detection”) AND (“transfer learning” OR “machine learning” OR “deep learning”) AND (“eye disease” OR “cataract”) AND “method”.

3.1.3 Inclusión

- Aquellas publicaciones que sean del tipo *Journal*.
- Revistas reconocidas.
- Publicaciones relacionadas a diversas enfermedades oculares incluyendo la catarata.

3.1.4 Exclusión

- Cualquier publicación que sea distinto al inglés o español.
- Publicaciones internacionales que estén fuera del rango 2019 - 2024.

3.2 Implementación/Revisión de artículos

Para llevar a cabo la investigación, se realizó una revisión sistemática de la literatura. Las fuentes de datos utilizadas incluyeron bases de datos científicas que contienen publicaciones relevantes en el ámbito de la investigación, tales como *Scopus*, *IEEE Xplorer*, *Springer*, *Elsevier*, *Mendeley*, *Science Direct*, *Scholar Google*, *Wiley* y *Alicia Concytect*.

Se seleccionaron 15 publicaciones que se muestran en resumen a continuación y en el Anexo 1.

1. El estudio “***Deep Learning for Ocular Disease Recognition: An Inner-Class Balance***” realizado por Khan et al. (2022) propuso un enfoque para abordar el desequilibrio de clases en un conjunto de datos de 5,000 imágenes de fondo de ojo, transformando el problema de clasificación multiclase en uno binario. Para ello, se utilizó el modelo VGG-19, que entrenó clasificadores para diferenciar entre ojos sanos y aquellos con alguna enfermedad ocular. Las pruebas experimentales se llevaron a cabo en la plataforma *Google Colab*, utilizando el optimizador Adam junto con la función de pérdida binaria basada en entropía cruzada.

Los resultados mostraron un rendimiento notable. Se obtuvo una precisión

del 98.10 % al distinguir entre un ojo normal y uno con miopía, mientras que la prueba de ojos normales versus cataratas arrojó una precisión del 94.03 %. En el caso de glaucoma, la precisión fue del 90.94 %, demostrando la efectividad del modelo en diferentes patologías oculares.

2. La investigación de Vijayaraghavan (2022), “***Detection and Revelation of Multiple Ocular Diseases using Transfer Learning Techniques***”, tuvo como objetivo detectar múltiples enfermedades oculares utilizando técnicas de aprendizaje profundo. Para ello, se emplearon varios modelos de Transferencia de Aprendizaje, entre los que se incluyeron VGG-16, VGG-19, RestNet50, MobileNet, AlexNet e InceptionNetV3. Los resultados mostraron diferencias en la precisión de cada modelo, destacando InceptionNetV3 con un 92.1% de precisión, mientras que los otros modelos obtuvieron las siguientes precisiones: VGG-16 (88.3%), AlexNet (86.9%), VGG-19 (88%), RestNet50 (86.9%) y MobileNet (90.7%).

En conclusión, se identificó a InceptionNetV3 como el modelo más preciso, alcanzando una precisión final de 92.1%. Conjuntamente, la investigación resaltó la importancia de la metodología CRISP-DM como una herramienta clave en el desarrollo del proyecto de detección de enfermedades oculares mediante aprendizaje profundo.

3. Zhang et al. (2022) realizaron una revisión exhaustiva de la literatura con el fin de recopilar los avances recientes en el uso de técnicas de aprendizaje automático para la clasificación y gradación de cataratas a partir de imágenes oftálmicas. A lo largo del estudio, “***Machine Learning for Cataract Classification/Grading on Ophthalmic Imaging Modalities: A Survey***”, se identificaron varias limitaciones, como la anotación automática de imágenes

y los objetivos clínicos de diagnóstico, lo que evidenció la necesidad de perfeccionar los métodos existentes para alcanzar mayor precisión. Como conclusión, los autores destacaron la importancia de mejorar la interpretabilidad y comprensión de las técnicas de aprendizaje profundo en el diagnóstico de cataratas. Asimismo, se propuso la creación de un estándar para las imágenes oftálmicas que facilite la clasificación de cataratas, aportando mejoras significativas en este campo.

4. En el estudio "***Glaucoma Detection Using Inception Convolutional Neural Network***", Afroze et al. (2021) desarrollaron un modelo basado en la red neuronal InceptionV3 con el objetivo de optimizar la detección de glaucoma, reduciendo tanto los tiempos como los costos asociados. El conjunto de datos utilizado incluyó 2,336 imágenes de *fundus* con glaucoma y 3,736 imágenes normales, extraídas de las bases de datos ACRIMA y LAG. En total, se emplearon 5,640 imágenes para el entrenamiento y pruebas del modelo, con 612 imágenes adicionales para la evaluación. Los resultados mostraron una precisión de 0.8529 y un área bajo la curva (AUC) de 0.9387. También, al comparar el rendimiento del modelo con otros algoritmos, como ResNet50 y DenseNet121, se obtuvieron precisiones de 0.7761 y 0.8153, respectivamente, confirmando la superioridad del enfoque basado en InceptionV3 para la detección de glaucoma como se aprecia en la **Tabla 1**.

Tabla 1

Desempeño de los modelos de Afroze y su equipo

Model	Test accuracy	Class name	Precision	Recall	F1 Score
Inception V3	0.8529	Glaucoma	0.86	0.85	0.85
		Normal	0.85	0.86	0.86
DenseNet 121	0.8153	Glaucoma	0.74	0.96	0.84
		Normal	0.95	0.68	0.79
ResNet50	0.7761	Glaucoma	0.83	0.68	0.75
		Normal	0.74	0.87	0.80

Nota. Tomado de *Glaucoma Detection Using Inception Convolutional Neural Network* por Afroze et al, 2021.

- En el estudio "*Cataract detection based on ocular B-ultrasound images by collaborative monitoring deep learning*", Wang et al. (2021) propusieron un enfoque basado en aprendizaje profundo, denominado Monitoreo Colaborativo para la Detección de Cataratas (CMDL), aplicado a imágenes de ecografía ocular tipo B. Este enfoque incorporó un módulo de aprendizaje colaborativo, diseñado para mejorar la precisión en la identificación de cataratas.

Los resultados experimentales demostraron la efectividad de CMDL, el cual combinó una red de detección de objetos y tres módulos de extracción de características, logrando una mejora significativa en la precisión de detección. Este método innovador ofrece un gran potencial para futuras investigaciones en el campo de la detección de cataratas utilizando ecografías oculares tipo B.

- En el estudio "*ACCV: automatic classification algorithm of cataract video based on deep learning*", Hu et al. (2021) desarrollaron un algoritmo para la clasificación automática en tiempo real de cataratas a partir de videos. Los resultados obtenidos reflejaron un alto nivel de precisión (0.9400), junto con

un área bajo la curva (AUC) de 0.9880 y un valor F1 de 0.9388, lo que evidencia un desempeño sólido del modelo. Por otra parte, cada cuadro de video fue procesado en solo 29 microsegundos gracias al uso de un método de reconocimiento del espacio de color, lo que mejoró significativamente la eficiencia del proceso.

El uso de la red neuronal convolucional YOLOv3 se demostró eficaz en la simulación del diagnóstico de cataratas basado en videos, lo que destaca su potencial para la detección temprana de esta enfermedad, mejorando tanto la precisión como la rapidez en la identificación.

7. En el estudio "*Automatic Detection of Diabetic Eye Disease Through Deep Learning Using Fundus Images: A Survey*", Sarki et al. (2020a) realizaron una revisión sistemática de enfoques automatizados para la detección de la enfermedad ocular relacionada con la diabetes (*Diabetic Eye Disease*, DED). Utilizando conjuntos de datos provenientes de repositorios como Kaggle y Messidor, los autores compararon diversas metodologías, destacando la clasificación binaria como el enfoque predominante para distinguir entre ojos sanos y aquellos afectados por DED. Los resultados del análisis mostraron una alta efectividad en la identificación de casos graves de DED. Sin embargo, se enfatiza la necesidad de desarrollar clasificadores más precisos que puedan detectar las etapas tempranas o leves de la enfermedad, lo que sería clave para mejorar las intervenciones preventivas.
8. En el artículo "*Early detection of diabetic eye disease from fundus images with deep learning*", Sarki et al. (2020b) aplicaron técnicas de aprendizaje profundo para mejorar la detección temprana de enfermedades oculares

diabéticas, enfrentando las limitaciones actuales. El estudio subraya la importancia del preprocesamiento de imágenes para desarrollar clasificadores altamente precisos, resaltando que tanto la arquitectura de la red como la calidad de los datos son factores cruciales en la creación de un sistema eficaz.

Los autores concluyen que es fundamental mejorar la clasificación temprana de males como la retinopatía diabética, el glaucoma y el edema macular diabético, mediante métodos que prevengan la ceguera permanente. Estas mejoras son clave para lograr un diagnóstico más oportuno y eficaz de estas condiciones.

9. En el estudio "***Automatic Cataract Severity Detection and Grading Using Deep Learning***", Yadav y Yadav (2023) propusieron un método innovador para la detección y clasificación de cataratas utilizando espectrogramas 2D-DFT de imágenes de fondo de ojo en lugar de las imágenes originales. Este enfoque, basado en redes neuronales convolucionales (CNN), permite capturar detalles de pequeños vasos sanguíneos como componentes de alta frecuencia, lo que facilita su uso como características distintivas. El conjunto de datos fue organizado a partir de fuentes de código abierto, mejorando la calidad de las imágenes a través de procesos de selección y preprocesamiento. Los espectrogramas obtenidos se analizaron con la CNN y se clasificaron mediante un clasificador *softmax*. Los resultados mostraron que este método superó a modelos pre entrenados como AlexNet, VGGNet y ResNet50, logrando una precisión del 93.10%, una sensibilidad del 93.13%, y una especificidad del 97.71%, con menores tiempos de computación debido a la simplicidad del modelo.

10. En el artículo "***Deep Learning in Ophthalmology: The Technical and Clinical Considerations***", Ting et al. (2019) llevaron a cabo una exhaustiva revisión de las tendencias en aplicaciones de Inteligencia Artificial, Machine Learning y Deep Learning en oftalmología. El estudio destaca el potencial de estas tecnologías, impulsado por el avance en unidades de procesamiento gráfico, mejoras en los modelos matemáticos y la disponibilidad de grandes conjuntos de datos (*Big Data*). El análisis revela la eficacia del *Deep Learning* en la detección de condiciones como glaucoma, retinopatía diabética (RD), errores de refracción y otros factores de riesgo cardiovascular, mediante el análisis de *fundus* o imágenes de fondo de ojo y tomografías ópticas. Asimismo, el artículo identifica áreas clínicas no satisfechas y condiciones oftálmicas relevantes en la salud pública, donde la aplicación de Inteligencia Artificial y *Deep Learning* podría tener un impacto significativo
11. En el artículo "***An Annotation-free Restoration Network for Cataractous Fundus Images***", Heng Li et al. (2022) presentan una solución innovadora para la restauración de imágenes de fondo ocular afectadas por cataratas sin necesidad de anotaciones detalladas. La propuesta, denominada ArcNet, se centra en extraer componentes de alta frecuencia de las imágenes, evitando la segmentación convencional y preservando las estructuras esenciales de la retina.

ArcNet fue entrenado inicialmente con imágenes sintetizadas y luego adaptado para imágenes reales afectadas por cataratas. Los resultados experimentales demostraron que ArcNet supera a otros algoritmos en términos de desempeño en el diagnóstico automático de enfermedades de

fondo ocular, como las cataratas. La capacidad de ArcNet para restaurar imágenes sin datos anotados subraya su potencial clínico, facilitando su implementación en entornos médicos con recursos limitados.

12. En el estudio "*Use of Machine Learning to Assess Cataract Surgery Skill Level with Tool Detection*", Ruzicki et al. (2022) desarrollaron un modelo de aprendizaje automático para analizar habilidades en cirugías de cataratas mediante la detección de instrumentos quirúrgicos en videos. Utilizando la red de clasificación de imágenes ResNet 152, el modelo logró identificar con alta precisión los instrumentos utilizados durante las cirugías, alcanzando un área bajo la curva (AUC) entre 0.933 y 0.998 para la detección de herramientas.

Sin embargo, la clasificación de las habilidades de los cirujanos resultó ser más desafiante. El modelo mostró un AUC de 0.550 y una precisión del 54.3% para fragmentos individuales, un AUC de 0.570 y una precisión del 57.8% para cirugías individuales, y un AUC de 0.692 con una precisión del 63.3% para evaluaciones globales de usuarios. Estos hallazgos indican que, aunque el aprendizaje automático puede identificar instrumentos quirúrgicos con precisión, la evaluación exacta de las habilidades de los cirujanos sigue siendo un reto considerable.

13. En la tesis de Robalino (2021), "*Diagnóstico precoz de la retinopatía diabética utilizando visión computacional*", se aborda la detección temprana de la retinopatía diabética mediante el uso de visión computacional y redes neuronales convolucionales para analizar imágenes de fondo de ojo y detectar esta patología. Utilizando un *dataset* extraído de Kaggle, el estudio alcanzó una precisión del 80% en la clasificación de imágenes, lo que es considerado

aceptable para este tipo de proyectos. El sistema desarrollado tiene el potencial de ser una herramienta útil en la identificación de la retinopatía diabética, contribuyendo a la prevención de esta enfermedad a través de chequeos visuales regulares.

Entre las recomendaciones del autor se presenta mejorar la base de datos mediante *crowdsourcing* para corregir imágenes mal clasificadas y colaborar con clínicas especializadas para evitar errores de clasificación. También, sugiere explorar metodologías y algoritmos adicionales para optimizar la detección de enfermedades, incluyendo el ajuste de redes neuronales con capas y neuronas más profundas. Finalmente, aconseja utilizar computadoras con unidades de procesamiento gráfico (GPU), como las tarjetas GeForce, para acelerar el aprendizaje del modelo.

14. La tesis titulada “*Prevención y detección temprana de glaucoma a través del procesamiento digital de imágenes del fondo del ojo humano y de la variación de la presión intraocular y su correlación con el espesor central de la córnea*” de Pinos (2019) aborda la prevención y detección precoz del glaucoma mediante el uso de procesamiento digital de imágenes del fondo ocular y el análisis de la relación entre la presión intraocular (PIO) y el espesor central de la córnea (CCT). El autor propone desarrollar una herramienta de soporte para especialistas utilizando técnicas de procesamiento de imágenes para analizar la relación copa-disco y la regla ISNT. La investigación, realizada con software como MATLAB y Python, generó una base de datos propia y empleó simulaciones 3D para estudiar el impacto del aumento de la PIO en el nervio óptico y la estructura del ojo. Los resultados mostraron que, a pesar de las limitaciones y la falta de software

específico para simulaciones del ojo, el enfoque propuesto es prometedor para la detección y diagnóstico del glaucoma, especialmente en contextos con recursos limitados.

Pino concluye que la disponibilidad de imágenes del ojo humano es limitada, y la simulación del ojo presenta desafíos significativos, adicionalmente, recomienda corregir imágenes mal clasificadas mediante *crowdsourcing* y trabajar con clínicas especializadas para mejorar la calidad de los datos. Asimismo, sugiere continuar explorando metodologías y algoritmos avanzados, así como utilizar software libre especializado para superar las limitaciones de software comercial. A pesar de los resultados alentadores, se enfatiza que el diagnóstico definitivo debe ser realizado por especialistas médicos, y se propone la integración de técnicas adicionales como lógica difusa y redes neuronales para mejorar el diagnóstico y tratamiento del glaucoma.

15. La tesis de Ñuflo y Mecca (2022), “***Sistema para diagnosticar la catarata mediante el uso de redes neuronales convolucionales***”, propone un sistema para diagnosticar cataratas utilizando redes neuronales convolucionales (CNN) a través de una aplicación móvil la cual emplea inteligencia artificial para analizar imágenes del fondo del ojo, ofreciendo una solución accesible y portátil, destinada a facilitar el diagnóstico en comunidades con recursos limitados.

El sistema se construyó en cuatro fases: preprocesamiento de imágenes, entrenamiento del modelo, diseño de la aplicación móvil y validación. Durante las pruebas, el modelo logró un promedio de precisión del 87%, con un 92% de acierto en la detección de cataratas y un 82% en ojos normales.

Estos resultados demuestran que la aplicación puede diagnosticar cataratas de manera precisa, lo que la convierte en una herramienta útil y práctica, especialmente para zonas alejadas donde no hay acceso inmediato a especialistas en oftalmología.

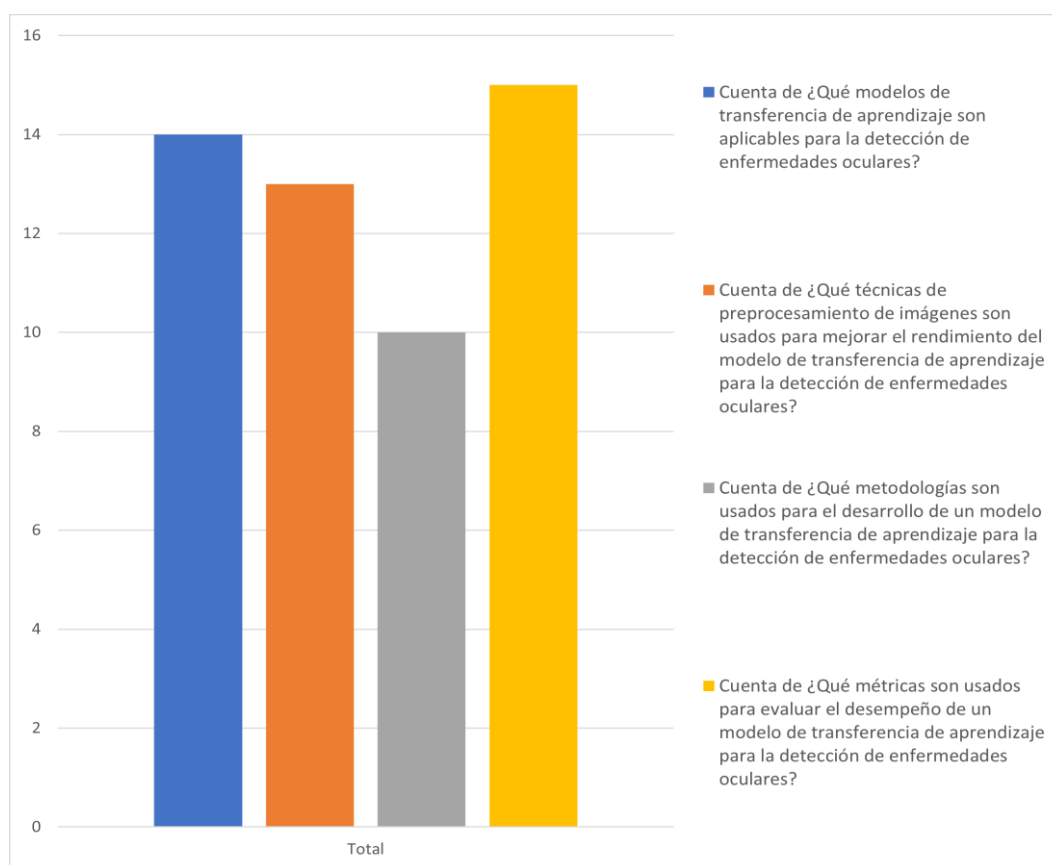
Para llevar a cabo la investigación, se realizó una revisión sistemática de la literatura.

3.3 Resultados de la revisión

En la **Figura 10** se muestra el conteo de los artículos revisados que contestan las preguntas de investigación planteadas durante la planificación.

Figura 10

Cantidad de artículos por pregunta de investigación



Nota. Elaboración propia.

Las bases de datos y editoriales correspondientes a los artículos revisados según las preguntas de investigación se muestran en la **Tabla 2**. Asimismo, los resultados de las revisiones según las preguntas de investigación se detallan en las tablas indicadas.

Tabla 2

Bases de datos o editoriales asociadas a las preguntas de investigación

PREGUNTAS	BASE DE DATOS O EDITORIAL								
	SCOPUS	SPRINGER	SCIENCE DIRECT	MENDELEY	IEEE EXPLORER	SCHOLAR GOOGLE	WILEY	ELSEVIER	ALICIA CONCYTEC
Modelos de transferencia de aprendizaje	[1], [4], [10]	[3], [6]	[5]	[8]	[7], [11]	[2]	[9]	[12]	[13], [15]
Pre-procesamiento de imágenes	[1], [4]	[3], [6]	[5]	[8]	[7], [11]	[2]	[9]		[13], [14], [15]
Metodología	[1], [4]	[3]	[5]		[11]	[2]	[9]		[13], [14], [15]
Métricas de desempeño	[1], [4], [10]	[3], [6]	[5]	[8]	[7], [11]	[2]	[9]	[12]	[13], [14], [15]

Nota. Elaboración propia.

P1 ¿Qué modelos de transferencia de aprendizaje son aplicables para la detección de enfermedades oculares?

Las contribuciones de los distintos autores se encuentran resumidas en la **Tabla 3**.

Tabla 3

Modelos de transferencia de aprendizaje para detectar enfermedades oculares

ID	Modelos	Referencias
M01	VGG-16	[2], [5], [8]
M02	VGG-19	[1], [2], [6], [8], [9]
M03	InceptionNetV3	[2], [4], [6], [10]
M04	MobileNet	[2], [6], [8]

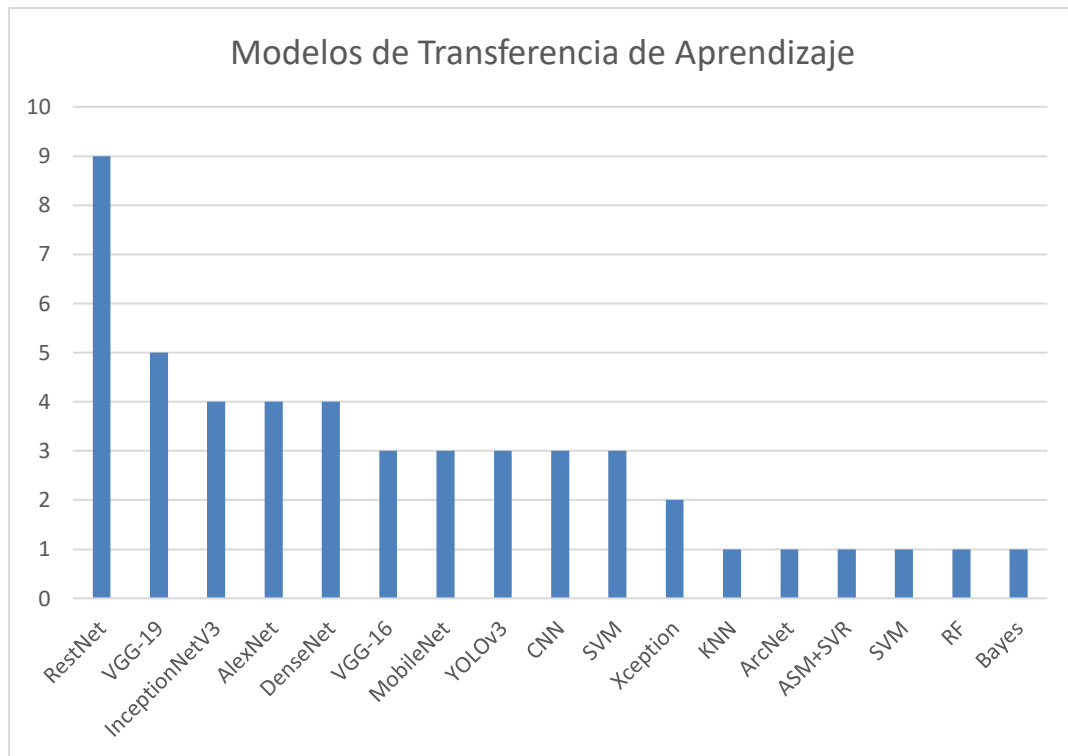
M05	RestNet	[2], [4], [5], [6], [8], [9], [10], [12], [15]
M06	AlexNet	[2], [5], [9], [10]
M07	SVM	[3], [9], [15]
M08	KNN	[3]
M09	CNN	[3], [9], [13]
M10	DenseNet	[4], [5], [8], [10]
M11	YOLOv3	[5], [6], [7]
M12	Xception	[6], [8]
M13	ArcNet	[11]
M14	ASM+SVR	[3]
M15	SVM	[3]
M16	RF	[3]
M17	Bayes	[3]

Nota. Elaboración propia.

En la **Figura 11**, se puede apreciar los Modelos de transferencia de aprendizaje para detectar enfermedades oculares.

Figura 11

Modelos de transferencia de aprendizaje para detectar enfermedades oculares



Nota. Elaboración propia.

P2 ¿Qué técnicas de preprocesamiento de imágenes son usados para mejorar el rendimiento del modelo de transferencia de aprendizaje para la detección de enfermedades oculares?

Los aportes de los autores se pueden evidencian en la **Tabla 4**.

Tabla 4

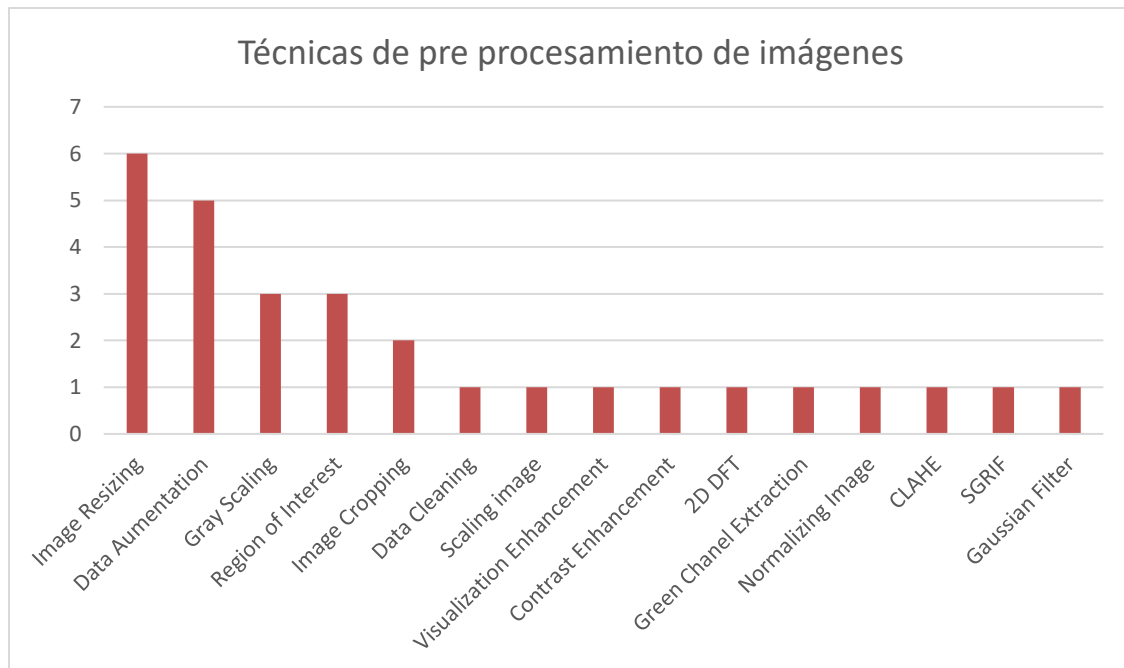
Técnicas de preprocesamiento de imágenes

ID	Técnicas	Referencias
T01	Image Resizing	[1], [2], [3], [6], [9], [13]
T02	Data Cleaning	[2]
T03	Data Augmentation	[2], [4], [8], [9], [15]
T04	Image Cropping	[2] [13]
T05	Scaling image	[2]

T06	Gray Scaling	[5], [13], [14]
T07	Visualization Enhancement	[7]
T08	Contrast Enhancement	[7]
T09	Region of Interest	[8], [13], [14]
T10	2D DFT	[9]
T11	Green Chanel Extraction	[9]
T12	Normalizing Image	[9]
T13	CLAHE	[11]
T14	SGRIF	[11]
T15	Gaussian Filter	[14]

Nota. Elaboración propia.

Se puede apreciar en la **Figura 12**, las técnicas de preprocesamiento de imágenes para los modelos de transferencia de aprendizaje destinados a la detección de enfermedades oculares.

Figura 12*Técnicas de preprocesamiento de imágenes**Nota.* Elaboración propia.

P3 ¿Qué metodologías son usados para el desarrollo de un modelo de transferencia de aprendizaje para la detección de enfermedades oculares?

Se puede apreciar en la **Tabla 5**, lo que expresan los autores:

Tabla 5*Metodologías de los artículos revisados*

ID	Metodologías	Referencias
ME01	Metodología personalizada	[1], [3],[9], [11], [13], [14]
ME02	CRISP-DM	[2]
ME03	Collaborative Monitoring Deep Learning	[5]
ME04	SCRUM	[15]

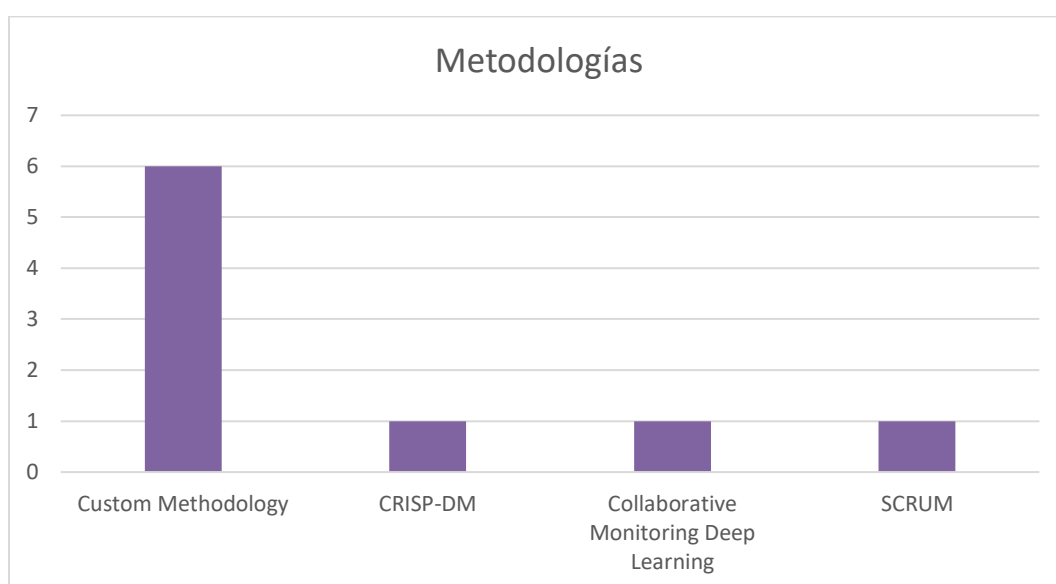
Nota. Elaboración propia.

Cabe destacar que las metodologías identificadas como ME01 presentan fases similares a las del enfoque CRISP-DM, no obstante, al no denominarlas como tal por sus respectivos autores se están considerando como métodos personalizados.

Por otro lado, estos resultados se están contabilizando en un gráfico de barras verticales como se muestra en la **Figura 13**.

Figura 13

Metodologías para desarrollar un modelo de transferencia de aprendizaje



Nota. Elaboración propia.

P4. ¿Qué métricas son usadas para evaluar el desempeño de un modelo de transferencia de aprendizaje para la detección de enfermedades oculares?

Se puede apreciar en la **Tabla 6**, lo que expresan los investigadores.

Tabla 6

Métricas para evaluar el desempeño de los modelos

ID	Métricas de desempeño	Referencias
MD01	Recall	[1], [2], [4], [15]
MD02	Precision	[1], [2], [3], [4], [13], [15]

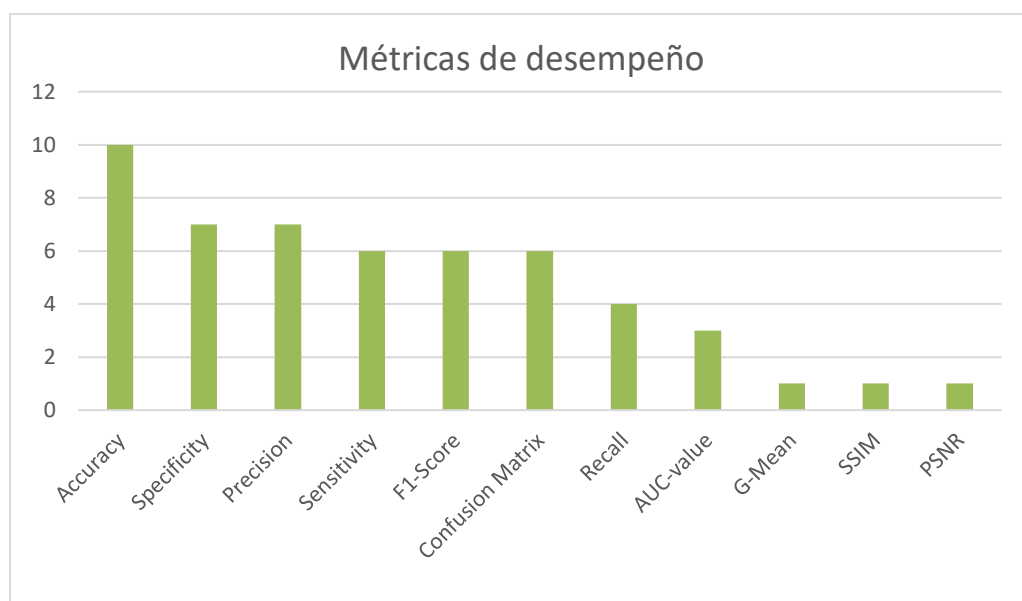
MD03	Accuracy	[1], [2], [3], [4], [5], [6], [7], [8], [9], [12]
MD04	F1-Score	[1], [3], [4], [6], [9], [15]
MD05	Matriz de Confusión	[1], [2], [4], [9], [14], [15]
MD06	AUC-Value	[2], [10], [12]
MD07	Sensitivity	[3], [6], [8], [9], [10], [14]
MD08	Specificity	[3], [4], [6], [8], [9], [10], [14]
MD09	G-Mean	[3]
MD10	SSIM	[11]
MD11	PSNR	[11]

Nota. Elaboración propia.

En la **Figura 14** se muestran las métricas de desempeño utilizadas con mayor frecuencia.

Figura 14

Métricas de desempeño más usados



Nota. Elaboración propia.

3.4 Herramientas comerciales para el diagnóstico automatizado de enfermedades oculares

Además de los avances investigativos, actualmente existen soluciones comerciales consolidadas que integran inteligencia artificial para el diagnóstico clínico de enfermedades oculares. Entre ellas, se encuentra EyeArt, un sistema aprobado por la Administración de Alimentos y Medicamentos de los Estados Unidos (FDA), que permite detectar retinopatía diabética y otras afecciones a partir del análisis automatizado de imágenes del fondo de ojo (Eyenuk Inc., 2023). De forma similar, IDx-DR, también autorizado por la FDA, emplea algoritmos de inteligencia artificial para identificar signos tempranos de daño ocular en pacientes diabéticos sin necesidad de intervención médica directa (Digital Diagnostics, 2022). Por su parte, RetinaNet, aunque inicialmente desarrollada para la detección de retinopatía diabética, ha sido adaptada por diversos grupos de investigación para identificar otras patologías como cataratas y glaucoma mediante aprendizaje profundo (Gulshan et al., 2016). Estas herramientas reflejan el potencial clínico y diagnóstico de los modelos automatizados en la identificación oportuna de enfermedades visuales.

Capítulo IV – Implementación

El presente capítulo señala el aporte del presente trabajo, así como las estrategias y herramientas utilizadas para cumplir con los objetivos trazados, asimismo, relata el desarrollo de la propuesta siguiendo las fases de la metodología CRISP-DM desde la comprensión del problema y procesamiento de los datos, culminando en la implementación del modelo de mejor rendimiento en una app web.

4.1 Aporte del estudio

El aporte del presente trabajo será la implementación de un modelo de Transferencia de Aprendizaje orientado a la detección de casos de cataratas en una aplicación web. Dicha propuesta permitirá un diagnóstico eficiente y preciso en pacientes con cataratas y beneficiará la calidad de atención médica y calidad de vida de los pacientes, así como reducir costos de atención médica.

Este estudio se basó en una revisión exhaustiva de la literatura disponible en diversos repositorios de información, con el objetivo de identificar las tendencias y avances más recientes en la aplicación de técnicas de aprendizaje automático para el reconocimiento de casos de cataratas. A partir de los resultados obtenidos, se adoptó la metodología CRISP-DM como guía para el desarrollo del proyecto, destacando que esta metodología es ampliamente utilizada en el ámbito del *Machine Learning*.

Otro punto en la revisión de la literatura a considerar son los diversos algoritmos de aprendizaje automático son aplicables para este trabajo, por ello, se optaron por los modelos de Inceptionv3, ResNet-50 y VGG-19 por ser los más acudidos en la literatura y cuyas métricas de desempeño resultan altamente eficientes. Asimismo, se elaborará un análisis comparativo de los tres modelos para determinar cuál de ellos cuentan con el mejor

valor de una serie de métricas de evaluación, de este modo, el modelo ganador será desplegado en una aplicación web elaborado con *Python* y *Flask* donde el usuario final, un médico de atención primaria o un oftalmólogo novato, puede ingresar una imagen médica ocular de un paciente y determinar si tiene un caso positivo de catarata.

Se espera que este trabajo sirva como referencia y apoyo para futuras investigaciones en el campo de la detección de enfermedades oculares, como la catarata, el glaucoma, entre otras.

4.2 Estrategia para el diseño de la solución y selección de tecnologías

Para el diseño de la solución se sigue la siguiente estrategia:

- Definir las etapas en las que se dividirá la solución.
- Elegir un *dataset* que contengan imágenes de enfermedades oculares que incluyan casos de cataratas y casos sanos o normales.
- Seleccionar las técnicas, métodos y algoritmos pertinentes para lograr los objetivos planteados.

Se analizaron diversas propuestas de autores en el estado del arte, cuya selección adecuada permitirá alcanzar resultados óptimos en la detección de casos de cataratas. Para el diseño de la solución, se adoptará la metodología CRISP-DM, conocida por sus seis fases específicas orientadas a proyectos de Machine Learning o Data Mining. Esta metodología guiará tanto la implementación de técnicas de preprocesamiento de imágenes como el entrenamiento y la evaluación de modelos de aprendizaje automático.

En este contexto:

- Se mantendrán como guía las seis fases de la metodología CRISP-DM:

Comprensión de la problemática, comprensión de los datos, preparación o preprocesamiento de datos, modelización, evaluación del modelo e implementación del modelo.

- El conjunto de datos seleccionado para este proyecto es ODIR u *Ocular Disease Intelligent Recognition*, una base oftálmica que contiene fotografías a color del fondo del ojo, recopilada por Shanggong Medical Technology Co., Ltd. Este *dataset* etiqueta numerosas enfermedades oculares, como Glaucoma, Cataratas, Miopía y otras.
- Entre las diversas técnicas de preprocesamiento de imágenes revisadas del estado de arte, se aplicarán *image resizing* (redimensión de imagen), *image flip* (vuelta de imagen), *image rotating* (rotación de imagen), *grayscale* (escala de grises) y *contrast enhancement* (mejora de contraste).
- La selección del modelo con mejor rendimiento se basará en los resultados del reporte de clasificación, y dicho modelo será exportado para su despliegue en una aplicación web. Esta aplicación permitirá la detección de casos de cataratas mediante la subida de imágenes de fondo ocular.
- Tanto el desarrollo de los modelos de aprendizaje automático como la implementación del aplicativo web se realizarán utilizando herramientas clave. Para los modelos, se emplearán las librerías de Keras, facilitando la configuración de parámetros, capas, entrenamiento y evaluación. En el lado del aplicativo web, se utilizará el lenguaje Python junto con el *framework* Flask.

4.3 Metodología de la solución

La **Figura 15** presenta la arquitectura de la solución, explicando el sistema propuesto como parte del aporte práctico.

Figura 15

Arquitectura de la aplicación web de detección de casos de cataratas



Nota. Elaboración propia.

Como se puede observar en la Figura 15, la solución comprende ocho bloques. Inicia con la obtención de imágenes de fondo ocular del *dataset* ODIR (1) o del conjunto *Ocular Disease Intelligent Recognition* recuperado del repositorio de Kaggle. Luego, se realiza un filtrado de las imágenes (2), seleccionando casos de ojos sanos y con cataratas. Posteriormente, se procede al preprocesamiento de las imágenes (3) mediante técnicas como *Image Rotate*, *Image Flip*, *Grayscale* y *Contrast Enhancement*. Estas nuevas imágenes generadas o aumentadas se utilizan para crear un nuevo *dataset* (4), que será empleado en el desarrollo de modelos de Transferencia de Aprendizaje destinados a la detección de casos de cataratas.

En el bloque de modelización (5), se inicia con la partición de este nuevo *dataset*, utilizando el 70% para el entrenamiento y el 30% para la evaluación o *testing* de los modelos VGG-19, Inception V3 y ResNet50. Estos modelos se compilan mediante el optimizador Adam en compañía de la función de pérdida denominada *binary crossentropy*. Se evalúan utilizando la métrica de *accuracy*. Posteriormente, se realiza la evaluación de estos modelos (6) mediante la librería *scikit-learn*, utilizando funciones como *score_accuracy* y *confusion_matrix*, que proporcionan el puntaje de *accuracy* y la matriz de confusión, respectivamente. Igualmente, se emplea la función *classification_report*, que retorna los valores de *Precision*, *Recall*, *F1-Score* y *Support* en una tabla.

El siguiente bloque implica la selección del modelo con el mejor rendimiento (7) comparando los resultados obtenidos en la evaluación. Después de determinar el mejor modelo, se exportan sus pesos para su despliegue (8) en una aplicación web desarrollada utilizando el lenguaje Python y el *framework* Flask.

4.4 Herramientas

4.4.1 Hardware

Para el desarrollo del proyecto propuesto se utilizó una laptop Lenovo IdeaPad 5 15ALC05 con las siguientes características:

- Procesador: AMD Ryzen 7 5700U with Radeon Graphics - 1.80 GHz
- RAM instalada: 16.0 GB
- Sistema Operativo: Windows 11 Home Single Language

4.4.2 Software

La **Tabla 7** señalan las herramientas empleadas para el desarrollo de los modelos de Transferencia de Aprendizaje en el proyecto, mientras que las herramientas para el despliegue del modelo en una aplicación web se muestra en la **Tabla 8**.

Tabla 7

Herramientas para el desarrollo de modelos de transferencia de aprendizaje

Desarrollo de modelos de Transferencia de Aprendizaje		
Jupyter Notebook	Entorno de desarrollo popularmente empleado en ciencia de datos y aprendizaje automático mediante la ejecución del lenguaje Python	Figura 16 <i>Logo Jupyter Notebook</i> 
Python	Lenguaje de programación de alto nivel con una extensa variedad de bibliotecas destinadas al aprendizaje automático	Figura 17 <i>Logo Python</i> 

TensorFlow	Biblioteca desarrollado por Google que permite diseñar y construir modelos de aprendizaje automático.	Figura 18 <i>Logo TensorFlow</i>
		 TensorFlow

Scikit learn	Biblioteca destinada al aprendizaje automático cuyas tareas de métrica de desempeño son usadas en el presente trabajo	Figura 19 <i>Logo Scikit Learn</i>
		

Nota. Elaboración propia.

Tabla 8

Herramientas para el despliegue del modelo en una aplicación web

Despliegue del modelo en aplicación web	
---	--

Visual Code	Studio	Editor de código fuente empleado para el desarrollo de la aplicación web con el <i>microframework</i> de Flask	Figura 20 <i>Logo Visual Studio Code</i>
			

Python + Flask	<i>Microframework</i> con soporte en Python para la creación de aplicaciones web	Figura 21 <i>Logo Python + Flask</i>
		

Figura 22

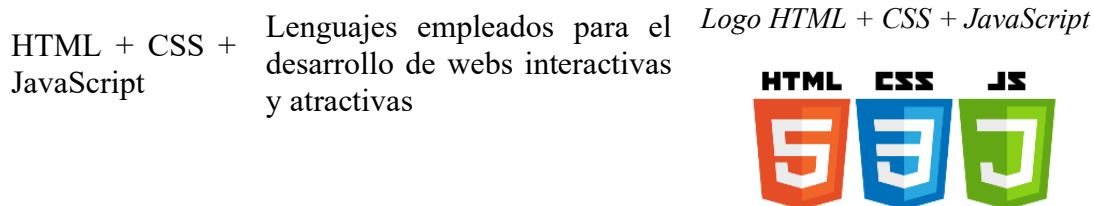
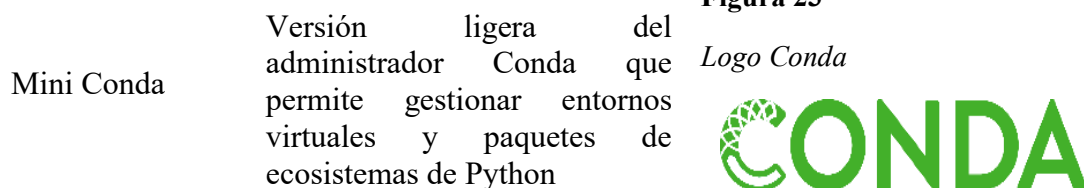


Figura 23



Nota. Elaboración propia.

4.5 Validación del Sistema

Para validar la efectividad del modelo propuesto se considerará las siguientes métricas de desempeño:

- *Accuracy:*

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Donde:

TP: Verdadero positivo
TN: Verdadero negativo
FP: Falso positivo
FN: Falso negativo

Nota. Tomado de *Precision, recall, F1, accuracy en clasificación* por Martinez, 2020

- *Precision:*

$$Precision = \frac{TP}{TP + FN}$$

Nota. Tomado de *Probabilistic extension of precision, recall, and F1 score for more thorough evaluation of classification models* por Yacouby & Axman, 2020

- *Recall*

$$Recall = \frac{TP}{TP + FN}$$

Nota. Tomado de *Probabilistic extension of precision, recall, and F1 score for more thorough evaluation of classification models* por Yacouby & Axman, 2020

- *F1-Score*

$$F1Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Nota. Tomado de *Precision, recall, F1, accuracy en clasificación* por Martinez, 2020

4.6 Desarrollo de la solución

La metodología seleccionada es CRISP-DM la cual consta de seis fases:

- Comprensión del problema.
- Comprensión de los datos.
- Preparación de los datos.
- Modelado.
- Evaluación del modelo.
- Implementación o despliegue del modelo.

4.6.1 Comprensión del problema

En la industria médica, los oftalmólogos suelen emplear las imágenes de fondo ocular o lámpara de hendidura para el diagnóstico de múltiples enfermedades oculares. También, con la aplicación de modelos de transferencia de aprendizaje acompañado con este tipo de imágenes médicas, en especial los de fondo ocular o *fundus*, se pueden detectar casos de cataratas para lograr un diagnóstico automatizado y brindar una mejor atención a los pacientes y reducir costos médicos. Para ello se emplean tres (3) modelos de transferencia de aprendizaje (InceptionV3, VGG-19 y Resnet50) que han sido entrenados y evaluados bajo una serie de métricas de desempeño, siendo el *accuracy* la métrica principal. El principal valor de la propuesta es que el diagnóstico temprano de catarata conlleva en un impacto positivo para el paciente ya que puede evitar el aumento de la

severidad, acudir a cirugías o la pérdida de la visión.

4.6.2 Comprensión de los datos

4.6.2.1 Recolección de datos

Resulta necesario para el estudio la recolección del *dataset* denominado ODIR recuperado del sitio Kaggle (Larxel, 2019)

Enlace del *dataset* <https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k>

4.6.2.2 Descripción del Dataset

El conjunto de datos ODIR u *Ocular Disease Intelligent Recognition* es una base de datos oftálmica que consta de fotografías a color del fondo de ojo de pacientes (Larxel, 2019). Este *dataset*, recopilado por *Shanggong Medical Technology Co., Ltd.*, etiqueta numerosas enfermedades oculares como Glaucoma, Catarata, Miopía y otros más, además, cuenta con los siguientes atributos:

- ID del paciente.
- Edad del paciente.
- Sexo del paciente.
- Fondo ocular derecho.
- Fondo ocular izquierdo.
- Diagnóstico ocular izquierdo.
- Diagnóstico ocular derecho.
- N (Caso normal).
- D (Diabetes).
- C (Catarata).
- A (Degeneración Macular relacionada con la edad).

- H (Hipertensión).
- M (Miopía patológica).
- O (Otras anomalías).

Del presente *dataset* ODIR se tomará los casos positivos a Catarata y una porción de casos normales destinados para los entrenamientos y las pruebas de los modelos.

4.6.2.3 Población de Estudio

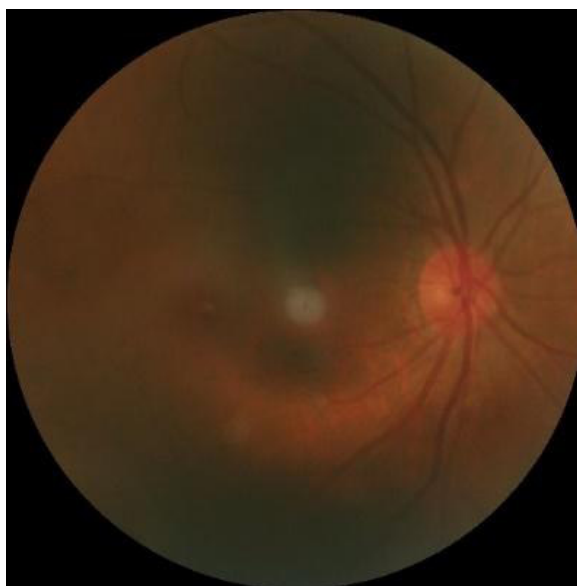
Fotografías captadas de los ojos izquierdos y derechos de personas de China.

4.6.2.4 Tamaño de muestra

- Población de 5000 personas de China con diversas enfermedades oculares.
- 594 casos de cataratas pertenecientes al presente *dataset* como la **Figura 24**.
- 500 casos normales pertenecientes al presente *dataset* como la **Figura 25**.

Figura 24

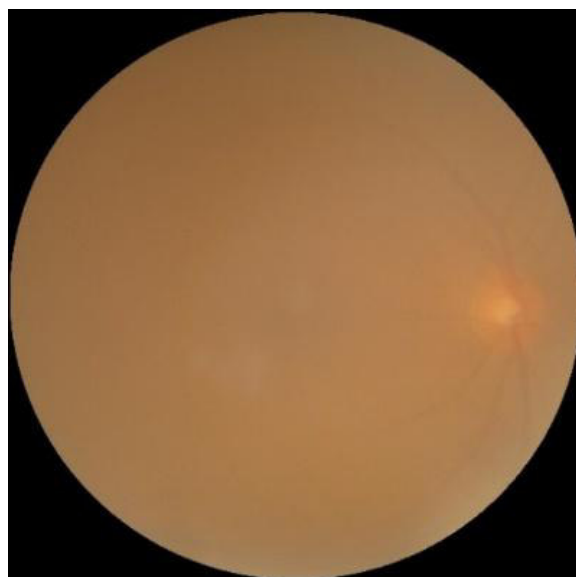
Imagen de fondo ocular de un caso normal



Nota: Imagen corresponde al ojo derecho del paciente con ID 0. Tomado de Kaggle

Figura 25

Imagen de fondo ocular de un caso de catarata



Nota: Imagen corresponde al ojo derecho del paciente con ID 112. Tomado de Kaggle

4.6.3 Preparación de los datos

En esta fase, se realizan una serie de operaciones a las imágenes del *dataset* previo a su implementación en el entrenamiento de los modelos. Estas operaciones se pueden agrupar en dos sub-fases o etapas:

- i. **Segmentación del dataset:** Del total de imágenes presentes del *dataset* de *Kaggle* se seleccionan los casos sanos y de catarata, los cuales serán almacenados en un nuevo *dataset*.

Para ello, elaboramos la función *tiene_cataract()* que busca la cadena “*cataract*” en las etiquetas de diagnóstico de ambos ojos con el fin de guardar los nombres de las imágenes de casos cataratas en listas acorde al ojo correspondiente: *catarata_izquierda* y *catarata_derecha* como se muestra en la **Figura 26**.

Figura 26*Casos de catarata en dataset ODIR*

```
def tiene_cataract(text):
    if "cataract" in text:
        return 1
    else:
        return 0

df["catarta_izquierda"] = df["Left-Diagnostic Keywords"].apply(lambda x: tiene_cataract(x))
df["catarata_derecha"] = df["Right-Diagnostic Keywords"].apply(lambda x: tiene_cataract(x))

catarata_izquierda = df.loc[(df.C == 1) & (df.catarta_izquierda == 1)][["Left-Fundus"].values
catarata_izquierda[:15] #mostrando solo 15 nombres

array(['0_left.jpg', '81_left.jpg', '103_left.jpg', '119_left.jpg',
      '254_left.jpg', '294_left.jpg', '330_left.jpg', '448_left.jpg',
      '465_left.jpg', '477_left.jpg', '553_left.jpg', '560_left.jpg',
      '594_left.jpg', '611_left.jpg', '625_left.jpg'], dtype=object)

catarata_derecha = df.loc[(df.C == 1) & (df.catarata_derecha == 1)][["Right-Fundus"].values
catarata_derecha[:15] #mostrando solo 15 nombres

array(['24_right.jpg', '81_right.jpg', '112_right.jpg', '188_right.jpg',
      '218_right.jpg', '345_right.jpg', '354_right.jpg', '477_right.jpg',
      '553_right.jpg', '560_right.jpg', '625_right.jpg', '726_right.jpg',
      '769_right.jpg', '949_right.jpg', '955_right.jpg'], dtype=object)
```

Nota. Elaboración propia.

Para el caso de ojos sanos se busca la cadena “*normal fundus*” en las etiquetas de los diagnósticos de ambos ojos para guardarlos en las listas de `normal_izquierdo` y `normal_derecho` como se muestra en la **Figura 27**.

Figura 27*Casos sanos en dataset ODIR*

```
normal_izquierdo = df.loc[(df.C == 0) & (df["Left-Diagnostic Keywords"] == "normal fundus")][["Left-Fundus"].sample(250, random_state=42)]
normal_izquierdo[:15]

array(['3332_left.jpg', '4059_left.jpg', '69_left.jpg', '2415_left.jpg',
      '4176_left.jpg', '2711_left.jpg', '4614_left.jpg', '3174_left.jpg',
      '2862_left.jpg', '2424_left.jpg', '3193_left.jpg', '3306_left.jpg',
      '2961_left.jpg', '2425_left.jpg', '2858_left.jpg'], dtype=object)

normal_derecho = df.loc[(df.C == 0) & (df["Right-Diagnostic Keywords"] == "normal fundus")][["Right-Fundus"].sample(250, random_state=42)]
normal_derecho[:15]

array(['2964_right.jpg', '680_right.jpg', '500_right.jpg',
      '2368_right.jpg', '2820_right.jpg', '2769_right.jpg',
      '2696_right.jpg', '2890_right.jpg', '940_right.jpg',
      '2553_right.jpg', '3371_right.jpg', '3042_right.jpg',
      '919_right.jpg', '3427_right.jpg', '379_right.jpg'], dtype=object)
```

Nota: Elaboración propia.

A continuación, concatenamos las dos listas de los nombres de cada ojo en una lista acorde a su diagnóstico (catarata o normal) correspondiente para su posterior almacenamiento como se aprecia en la **Figura 28**.

Figura 28

Cantidad de casos sanos y casos de catarata del dataset ODIR

```
#Concatenamos los resultados de ambos ojos para los dos casos
catarata = np.concatenate((catarata_izquierda, catarata_derecha),axis=0)
normal = np.concatenate((normal_izquierdo, normal_derecho),axis=0)

print("Numero de casos con catarata: {}".format(len(catarata)))
print("Numero de casos sanos o normales: {}".format(len(normal)))
```

```
Numero de casos con catarata: 594
Numero de casos sanos o normales: 500
```

Nota. Elaboración propia.

Para finalizar esta etapa, almacenamos los casos de catarata con la etiqueta o *label* '1' en un *dataset* mediante la función creada *crear_dataset()*, mientras que los casos sanos se les asigna la etiqueta '0' como en la **Figura 29**.

Figura 29

Almacenando en dataset

```
from tensorflow.keras.preprocessing.image import load_img, img_to_array

ruta_dataset = "preprocessed_images/"
image_size=224
labels = []
dataset = []

def crear_dataset(imagen_categoria, label):
    for img in tqdm(imagen_categoria):
        ruta_imagen = os.path.join(ruta_dataset, img)
        try:
            image = cv2.imread(ruta_imagen, cv2.IMREAD_COLOR)
            image = cv2.resize(image, (image_size, image_size))
        except:
            continue

        dataset.append([np.array(image), np.array(label)])
    random.shuffle(dataset)
    return dataset

dataset = crear_dataset(catarata, 1)

100% |████████████████████████████████████████████████████████████████████████████████| 594/594 [00:02<00:00, 220.68it/s]

dataset = crear_dataset(normal, 0)

100% |████████████████████████████████████████████████████████████████████████████████| 500/500 [00:07<00:00, 69.02it/s]
```

Nota. Elaboración propia.

- ii. **Preprocesamiento de imágenes seleccionadas:** Se aplican las técnicas de redimensión de imagen (*Image Resizing*), volteo de imagen (*Image Flip*), rotación de imagen (*Image Rotate*), escala de grises (*GrayScale*) y mejora de contraste (*Contrast Enhancement*) para generar imágenes aumentadas, es decir, nuevas imágenes variantes de las ya existentes las cuales serán añadidas al nuevo *dataset*.

En la **Figura 30** definimos un nuevo directorio denominada “imagenes_aumentadas” donde guardaremos las nuevas imágenes generadas. Asimismo, creamos en dos listas para definir las transformaciones de que se aplicarán para las técnicas de rotación (*rotate*) y vuelta (*flip*) los cuales serán utilizados en un bucle que aplicará dichas transformaciones a cada una de las imágenes de los casos sanos y de cataratas ya separadas en la carpeta “imagenes_seleccionadas”. De manera similar se aplican transformaciones de escalado de grises (*Grayscale*) y mejora de contraste (*Constrast Enhancement*) como se ve en la **Figura 31**.

Figura 30*Transformaciones de image rotate e image flip*

```

# Define el directorio que contiene tus imágenes originales
img_dir = "imagenes_seleccionadas/"

# Define el directorio donde se guardarán las nuevas imágenes generadas
output_dir = "imagenes_aumentadas/"

# Define las transformaciones que deseas aplicar
rot_angles = [30, 60, 90, -30, -60, -90] #angulos de rotacion
flip_modes = [0, 1, -1] # vueltas: 0 ninguno, 1 horizontal, -1 vertical

gray = True # Convertir a escala de grises

contrast_enhance = True # Mejora de contraste

# Itera sobre todas las imágenes originales
for filename in os.listdir(img_dir):
    # Carga la imagen
    img = cv2.imread(os.path.join(img_dir, filename))

    # Aplica las transformaciones
    # Image Rotate
    for angle in rot_angles:
        # Rota la imagen
        M = cv2.getRotationMatrix2D((img.shape[1] / 2, img.shape[0] / 2), angle, 1)
        rotated_img = cv2.warpAffine(img, M, (img.shape[1], img.shape[0]))

        # Guarda la imagen rotada
        new_filename = filename[:-4] + "_rot" + str(angle) + ".jpg"
        cv2.imwrite(os.path.join(output_dir, new_filename), rotated_img)

    # Image Flip
    for mode in flip_modes:
        # Voltea la imagen
        flipped_img = cv2.flip(img, mode)

        # Guarda la imagen volteada
        new_filename = filename[:-4] + "_flip" + str(mode) + ".jpg"
        cv2.imwrite(os.path.join(output_dir, new_filename), flipped_img)

```

Nota. Elaboración propia.**Figura 31***Transformaciones de grayscale y contrast enhancement*

```

# Grayscale
if gray:
    gray_img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
    new_filename = filename[:-4] + "_gray.jpg"
    cv2.imwrite(os.path.join(output_dir, new_filename), gray_img)

# Contrast Enhancement
if contrast_enhance:
    # Aplica la mejora de contraste utilizando la ecualización del histograma
    img_yuv = cv2.cvtColor(img, cv2.COLOR_BGR2YUV)
    img_yuv[:, :, 0] = cv2.equalizeHist(img_yuv[:, :, 0])
    enhanced_img = cv2.cvtColor(img_yuv, cv2.COLOR_YUV2BGR)

    new_filename = filename[:-4] + "_enhanced.jpg"
    cv2.imwrite(os.path.join(output_dir, new_filename), enhanced_img)

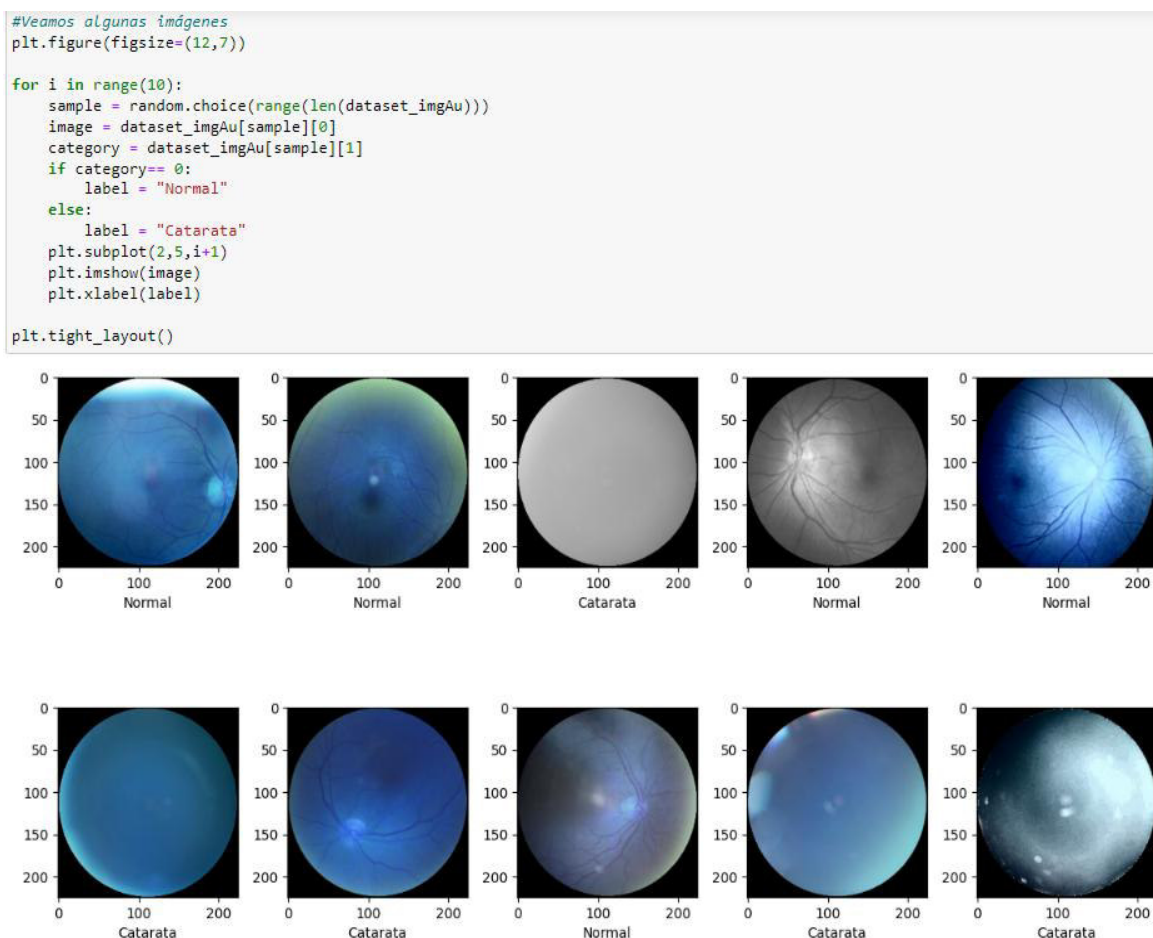
```

Nota. Elaboración propia.

Guardamos estas imágenes aumentadas obtenidas en el *dataset* previamente creado de manera similar como se explicó en la fase anterior, por último, veamos algunas muestras de manera aleatoria con la siguiente celda de código como en la **Figura 32**.

Figura 32

Muestras del nuevo dataset elaborado



Nota. Elaboración propia.

4.6.4 Modelización

Esta fase comienza con la partición del nuevo *dataset* en un 70% destinado al entrenamiento, mientras que el otro 30% será usado durante la evaluación como se muestra en la Figura 33.

Figura 33

Partición del dataset para el entrenamiento y evaluación

```
from sklearn.model_selection import train_test_split  
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.3)
```

Nota: Elaboración propia.

A través de la librería *tensorflow.keras.applications* se importan las arquitecturas de las redes VGG-19, Inception V-3 y ResNet50. Al momento de instanciar una arquitectura, por ejemplo, el modelo VGG-19, se señala que se usarán los pesos preentrenados del conjunto de datos *Imagenet* en el parámetro *weights*, para el parámetro *include_top* se le da valor *false* para excluir la capa de clasificación final de la red y el parámetro *input_shape* corresponde a las dimensiones de las imágenes de entrada.

Además, se establecen que las capas del modelo instanciado no sean entrenables para evitar que los pesos importados de *Imagenet* sean alterados durante el entrenamiento. Luego, se crea una variable de nombre *model* que importa la clase *Sequential* de *tensorflow.keras* para indicar que se trata de una secuencia de capas de un modelo que toma de base el VGG-19 anterior sumado a las capas *Flatten* y *Dense*, el primero se usa para aplanar los datos de salida de la red VGG-19 en un arreglo unidimensional, mientras que la segunda capa añadida se usa como capa de salida como una unidad y función de activación sigmoide para la clasificación binaria como vemos en las **Figuras 34 y Figura 35**.

Figura 34*Instancia del modelo preentrenado VGG-19*

```

from tensorflow.keras.applications.vgg19 import VGG19

vgg = VGG19(weights="imagenet", include_top = False, input_shape=(image_size, image_size, 3))

for layer in vgg.layers:
    layer.trainable = False

from tensorflow.keras import Sequential
from tensorflow.keras.layers import Flatten, Dense

model = Sequential()
model.add(vgg)
model.add(Flatten())
model.add(Dense(1, activation="sigmoid"))
# model.add(Dense(1, activation="softmax"))

```

Nota. Elaboración propia.**Figura 35***Resumen del modelo preentrenado VGG-19*

```

model.summary()

```

Model: "sequential"

Layer (type)	Output Shape	Param #
vgg19 (Functional)	(None, 7, 7, 512)	20024384
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 1)	25089

```

Total params: 20,049,473
Trainable params: 25,089
Non-trainable params: 20,024,384

```

Nota. Elaboración propia.

Después, se procede a ajustar la compilación del modelo con el optimizador Adam, para

la clasificación binaria se emplea la función de pérdida *binary_crossentropy* y la métrica *accuracy* para conocer el rendimiento del modelo. En la siguiente celda se importan e instancian las clases *ModelCheckpoint* y *EarlyStopping*, las cuales serán usadas como puntos de guardado del modelo con la mejor precisión durante el entrenamiento y detener el entrenamiento si no se observa mejoría en la precisión durante un número determinado de *epochs*. Estos ajustes se muestran en la **Figura 36**.

Figura 36

Compilación y checkpoint del modelo preentrenado VGG-19

```
model.compile(optimizer="adam",loss="binary_crossentropy",metrics=["accuracy"])

from tensorflow.keras.callbacks import ModelCheckpoint,EarlyStopping

checkpoint = ModelCheckpoint("vgg19.h5", monitor="val_accuracy", verbose=1, save_best_only=True,
                             save_weights_only=False, save_freq='epoch')

earlystop = EarlyStopping(monitor="val_accuracy",patience=5,verbose=1)
```

Nota. Elaboración propia.

Por último, en la **Figura 37** se realiza el entrenamiento del modelo estableciendo un total de 10 épocas, así como un tamaño de *batch* de 32, además, los datos de validación serán comprendidos por las variables *x_test* y *y_test* para evaluar el rendimiento del modelo en cada época. Para visualizar el progreso de cada época fue necesario establecer el parámetro *verbose* con un valor 1. Por otro lado, el argumento *callbacks* usa los objetos *checkpoint* y *earlystop* creados en el paso anterior como parte de las funciones de devolución de llamada durante el entrenamiento.

Figura 37*Entrenamiento del modelo*

```

history = model.fit(x_train,y_train,batch_size=32,epochs=10,validation_data=(x_test,y_test),
                    verbose=1,callbacks=[checkpoint,earlystop])

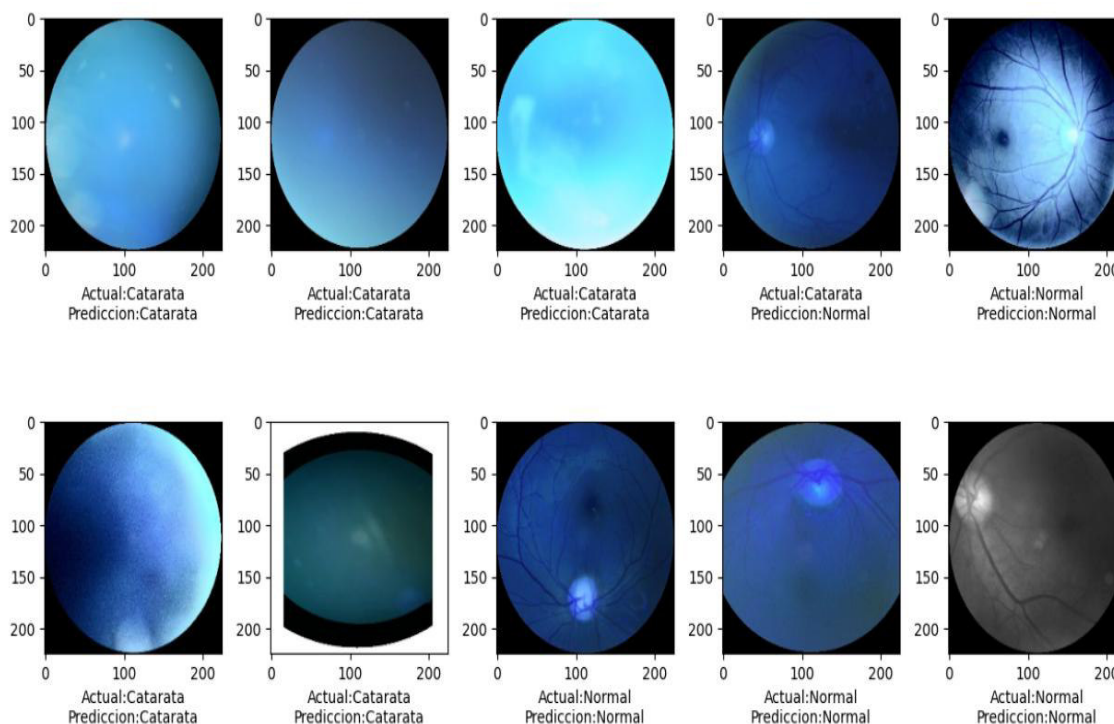
Epoch 1/10
262/262 [=====] - ETA: 0s - loss: 1.0920 - accuracy: 0.9224
Epoch 1: val_accuracy improved from -inf to 0.95405, saving model to vgg19.h5
262/262 [=====] - 1168s 4s/step - loss: 1.0920 - accuracy: 0.9224 - val_loss: 0.4612 - val_accuracy:
0.9541
Epoch 2/10
262/262 [=====] - ETA: 0s - loss: 0.3405 - accuracy: 0.9628
Epoch 2: val_accuracy improved from 0.95405 to 0.96714, saving model to vgg19.h5
262/262 [=====] - 1146s 4s/step - loss: 0.3405 - accuracy: 0.9628 - val_loss: 0.4463 - val_accuracy:
0.9671
Epoch 3/10
262/262 [=====] - ETA: 0s - loss: 0.1983 - accuracy: 0.9786
Epoch 3: val_accuracy improved from 0.96714 to 0.96798, saving model to vgg19.h5
262/262 [=====] - 1104s 4s/step - loss: 0.1983 - accuracy: 0.9786 - val_loss: 0.4205 - val_accuracy:
0.9680
Epoch 4/10
262/262 [=====] - ETA: 0s - loss: 0.0617 - accuracy: 0.9899
Epoch 4: val_accuracy improved from 0.96798 to 0.97522, saving model to vgg19.h5
262/262 [=====] - 1023s 4s/step - loss: 0.0617 - accuracy: 0.9899 - val_loss: 0.2937 - val_accuracy:
0.9752
Epoch 5/10
262/262 [=====] - ETA: 0s - loss: 0.1390 - accuracy: 0.9848
Epoch 5: val_accuracy did not improve from 0.97522
262/262 [=====] - 984s 4s/step - loss: 0.1390 - accuracy: 0.9848 - val_loss: 0.3995 - val_accuracy: 0.
9702

```

Nota. Elaboración propia.

4.6.5 Evaluación

Se emplea el otro 30% del *dataset* para evaluar el desempeño de los modelos, para ello se usan las métricas desempeño de *score_accuracy*, *confusion_matrix* y *classification_report* de la librería de Scikit Learn. En la **Figura 38** se muestran predicciones de imágenes aleatorios durante la evaluación.

Figura 38*Predicciones del modelo VGG19**Nota.* Elaboración propia.

4.6.6 Resultados experimentales

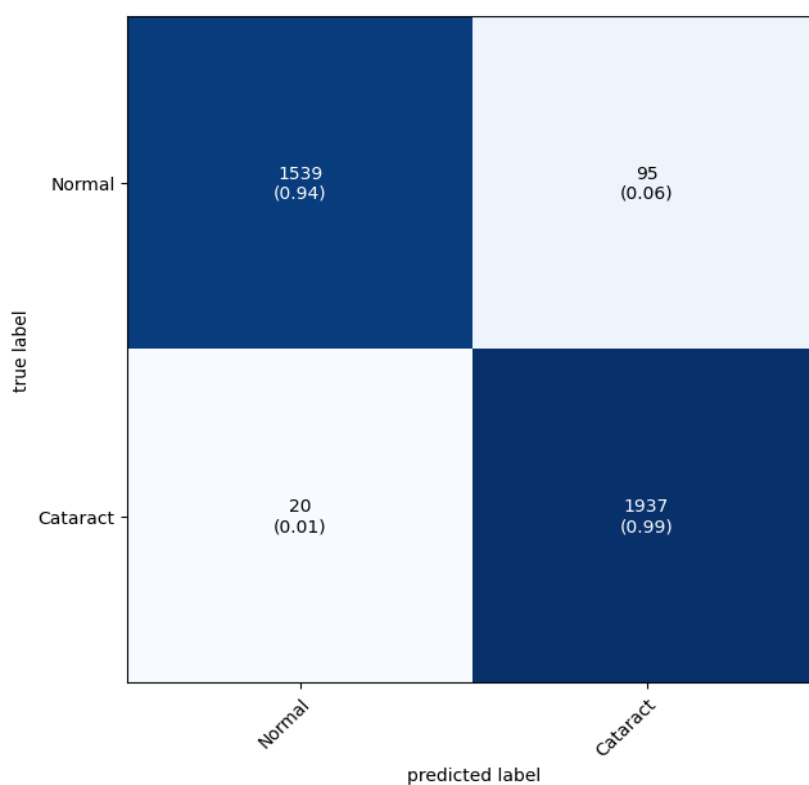
A continuación, se presentan las puntuaciones en las métricas de desempeño señaladas de los tres modelos entrenados durante el proyecto (**Figuras 39 al 44**), así como una tabla comparativa a modo de resumen respectivamente (**Tablas 9 al 11**).

Figura 39*Accuracy score del VGG-19*

```
accuracy_score(y_test,y_pred_bin)
```

```
0.9679754942912837
```

Nota. Elaboración propia.

Figura 40*Matriz de confusión del VGG-19**Nota.* Elaboración propia.**Tabla 9***Reporte de clasificación del VGG-19*

	precision	recall	f1-score	support
0	0.99	0.94	0.96	1634
1	0.95	0.99	0.97	1957
accuracy			0.97	2938
macro avg	0.97	0.97	0.97	2938
Weighted avg	0.97	0.97	0.97	2938

Nota. Elaboración propia.

Figura 41*Accuracy score del InceptionV3*

```
accuracy_score(y_test,y_pred_bin)
```

```
0.9075466443887497
```

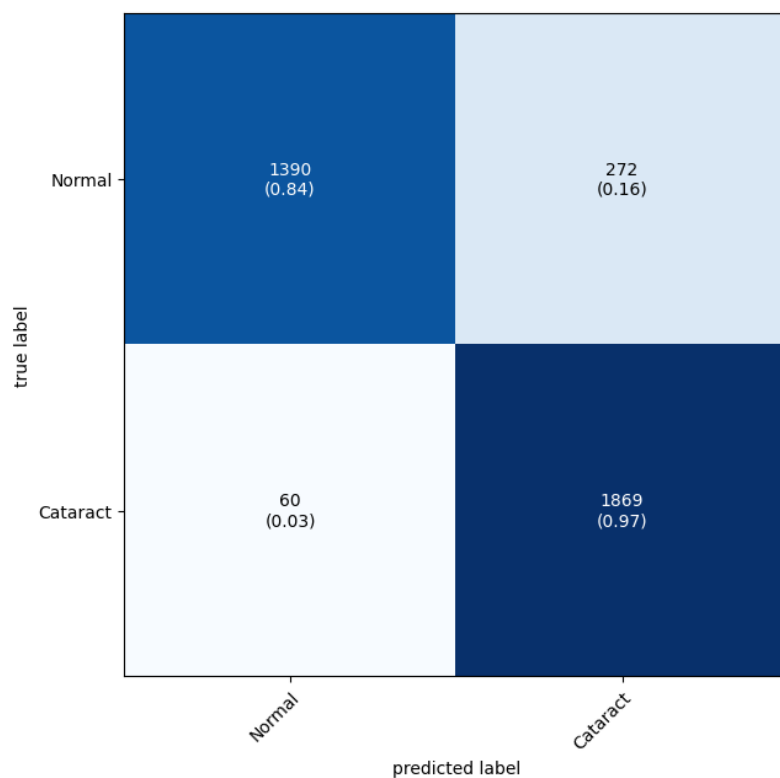
Nota. Elaboración propia.**Figura 42***Matriz de confusión del InceptionV3**Nota.* Elaboración propia.

Tabla 10*Reporte de clasificación del InceptionV3*

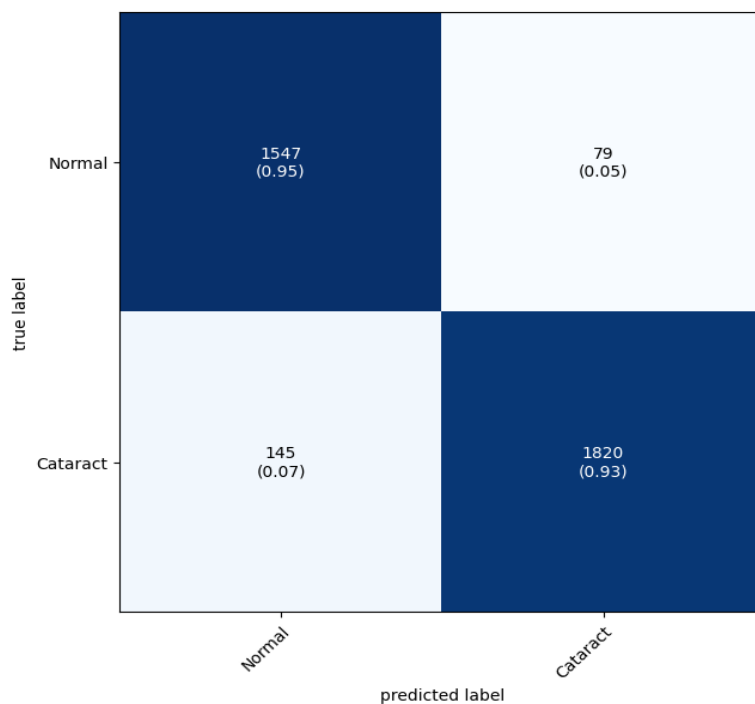
	precision	recall	f1-score	support
0	0.96	0.84	0.89	1662
1	0.87	0.97	0.92	1929
accuracy			0.91	3591
macro avg	0.92	0.90	0.91	3591
Weighted avg	0.91	0.91	0.91	3591

Nota. Elaboración propia.**Figura 43***Accuracy score del ResNet50*

```
accuracy_score(y_test,y_pred)
```

0.9376218323586745

Nota. Elaboración propia.

Figura 44*Matriz de confusión del ResNet50**Nota.* Elaboración propia.**Tabla 11***Reporte de clasificación del ResNet50*

	precision	recall	f1-score	support
0	0.91	0.95	0.93	1626
1	0.96	0.93	0.94	1965
accuracy			0.94	3591
macro avg	0.94	0.94	0.94	3591
Weighted avg	0.94	0.94	0.94	3591

Nota. Elaboración propia.

Tabla 12*Resumen comparativo del desempeño de los modelos*

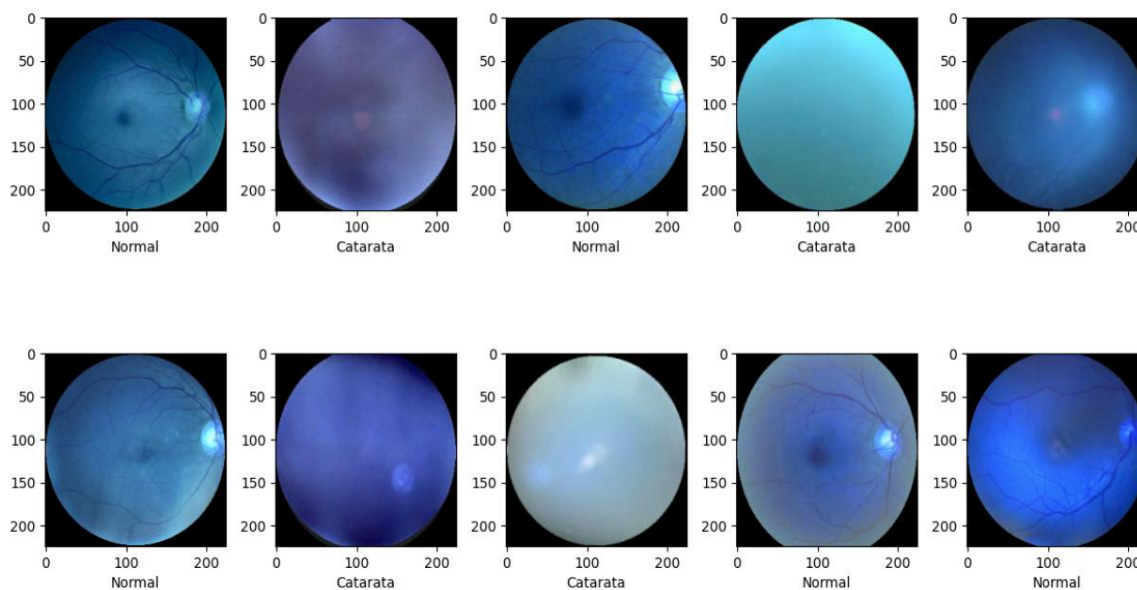
	Accuracy	Precision		Recall		F1-Score	
		Catarata	Sano	Catarata	Sano	Catarata	Sano
VGG-19	0.97	0.99	0.95	0.99	0.94	0.97	0.98
InceptionV3	0.91	0.87	0.96	0.97	0.84	0.92	0.89
ResNet50	0.94	0.96	0.91	0.93	0.95	0.94	0.93

Nota. Elaboración propia.

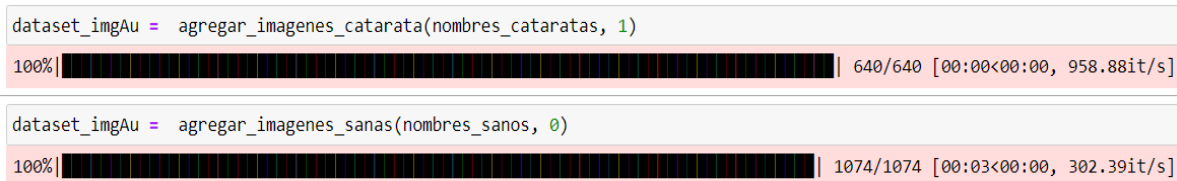
Como se observa en la **Tabla 12**, el modelo VGG-19 alcanzó una precisión del 97%, superando a InceptionV3 (91%) y ResNet50 (94%).

4.6.7 Validación y Comparativa

Aparte de la evaluación realizada en la Tabla 12, a modo de validación, se procede a probar los modelos obtenidos con un nuevo *dataset* de imágenes (Venkat, 2021) obtenido en Kaggle: <https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification>. La **Figura 45** presenta este nuevo conjunto de imágenes consta de múltiples clases de enfermedades oculares recopiladas de fuentes como IDRID, Ocular Recognition, HRF, entre otros.

Figura 45*Muestras del nuevo dataset para la validación**Nota.* Elaboración propia.

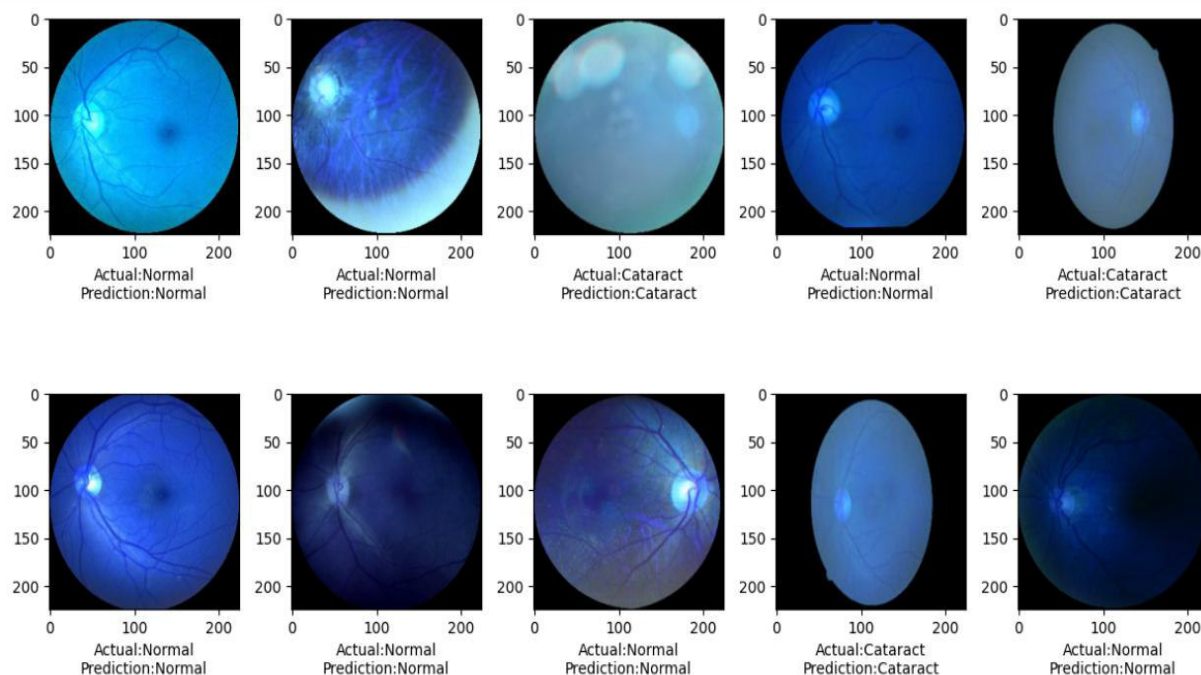
Se seleccionaron un total de 1074 imágenes (640 de cataratas y 434 sanos) para etiquetar los casos positivos como ‘1’ mientras que los sanos como ‘0’ en una variable llamada ‘*dataset_imgAu*’ como se aprecia en la **Figura 46**.

Figura 46*Seleccionando imágenes para la validación**Nota.* Elaboración propia.

Se instanciaron los tres tipos de modelos propuestos y se les importó los pesos entrenados durante la modelización (5.1.4) para aplicarles la función *predict()* considerando la nueva variable como entrada. En la **Figura 47** expone algunos ejemplos de las predicciones resultantes del modelo VGG19.

Figura 47

Predicciones del modelo VGG19 en la validación



Nota. Elaboración propia.

De este modo, se obtuvieron los siguientes resultados de la prueba de este nuevo *dataset* frente a los tres modelos VGG19, InceptionV3 y ResNet50 en la **Tabla 13**.

Tabla 13

Resumen de resultados de la validación

	Accuracy	Precision		Recall		F1-Score	
		Catarata	Sano	Catarata	Sano	Catarata	Sano
VGG-19	0.92	0.88	0.94	0.91	0.92	0.89	0.93
InceptionV3	0.85	0.73	0.96	0.95	0.8	0.82	0.87
ResNet50	0.91	0.93	0.9	0.81	0.96	0.87	0.93

Nota. Elaboración propia.

Como se observa de los resultados obtenidos de la validación presentan puntuaciones poco menores a lo visto en la **Tabla 12**, no obstante, cumplen con uno de los objetivos

planteados del proyecto: alcanzar una precisión de 85% como mínimo.

Por último, otra manera de validar nuestros resultados es usando *Benchmarking*, es decir, mediante una evaluación comparativa entre mi aporte con otros resultados sobre los modelos detectores de cataratas vistos en el estado del arte en la **Tabla 14**.

Tabla 14

Benchmarking de resultados propios vs otros autores

	<i>Accuracy</i>				
	<i>Propio</i>	<i>Validación</i>	(Khan et al, 2022)	(Vijayaraghavan, 2022)	(Hu et al., 2021)
VGG-19	0.97	0.92	0.94	0.88	0.87
InceptionV3	0.91	0.85	-	0.921	0.81
ResNet50	0.94	0.91	-	0.869	0.86
MobileNet	-	-	-	0.907	0.88

Nota. Elaboración propia.

A través de la validación por *Benchmarking* se concluye que el modelo VGG19 obtenido durante la modelización (5.1.4) ha tenido la puntuación más alta (97%), mientras que el InceptionV3 del autor Vijayaraghavan presenta el *accuracy* más alto (92.1%) para este tipo de modelo, por último, el modelo propio de ResNet50 supera al resto de sus homónimos con un 94%.

Adicionalmente, se llevó a cabo un proceso de juicio de expertos mediante un cuestionario como instrumento de recolección de datos, contando la participación de tres médicos especialistas en oftalmología. Cada experto evaluó los ítems del cuestionario en función de su relevancia, claridad y pertinencia para el objetivo del estudio. Como evidencia del juicio realizado, se adjuntan las respuestas y firmas de los expertos respectivos, las cuales se encuentran en el **Anexo 2**. Un resumen de las respuestas se expone en la **Tabla 15**.

Tabla 15*Resumen de respuestas de los expertos*

Ítem/Pregunta	Dr. Carmona	Dr. Limaylla	Dr. Lu
1. El tema abordado en el proyecto es relevante para la práctica médica en oftalmología.	De acuerdo	De acuerdo	Totalmente de acuerdo
2. El uso de modelos basados en inteligencia artificial puede mejorar el diagnóstico de cataratas.	De acuerdo	De acuerdo	De acuerdo
3. La detección automatizada de cataratas mediante imágenes oculares es clínicamente útil.	De acuerdo	Totalmente de acuerdo	De acuerdo
4. Este tipo de herramientas podrían implementarse en centros médicos de bajos recursos.	De acuerdo	Totalmente de acuerdo	Totalmente de acuerdo
5. Un sistema como el propuesto podría contribuir a reducir la ceguera por catarata en zonas vulnerables.	De acuerdo	Totalmente de acuerdo	De acuerdo

Nota. Elaboración propia.

Por último, se aplicó el coeficiente V de Aiken para cuantificar el grado de acuerdo entre los evaluadores, para ello las respuestas del cuestionario fueron escaladas del 1 al 5 como se visualiza en la **Tabla 16**. Los resultados obtenidos fueron satisfactorios, con valores superiores a 0.7 para todos los ítems o preguntas evaluadas, demostrando una adecuada validez del contenido del cuestionario como se identifica en la **Tabla 17**.

Tabla 16*Escala de respuestas de juicio de expertos*

Valor	Respuesta
1	Totalmente en desacuerdo
2	En desacuerdo
3	Ni de acuerdo ni en desacuerdo
4	De acuerdo
5	Totalmente de acuerdo

Nota. Elaboración propia.**Tabla 17***Cálculo V de Aiken*

Item/Pregunta	V de Aiken
1	0.83333333
2	0.75
3	0.83333333
4	0.91666667
5	0.83333333
Promedio	0.83333333

Nota. Elaboración propia.

4.6.8 Implementación o despliegue del modelo

Figura 48

Exportación del modelo con mejor desempeño

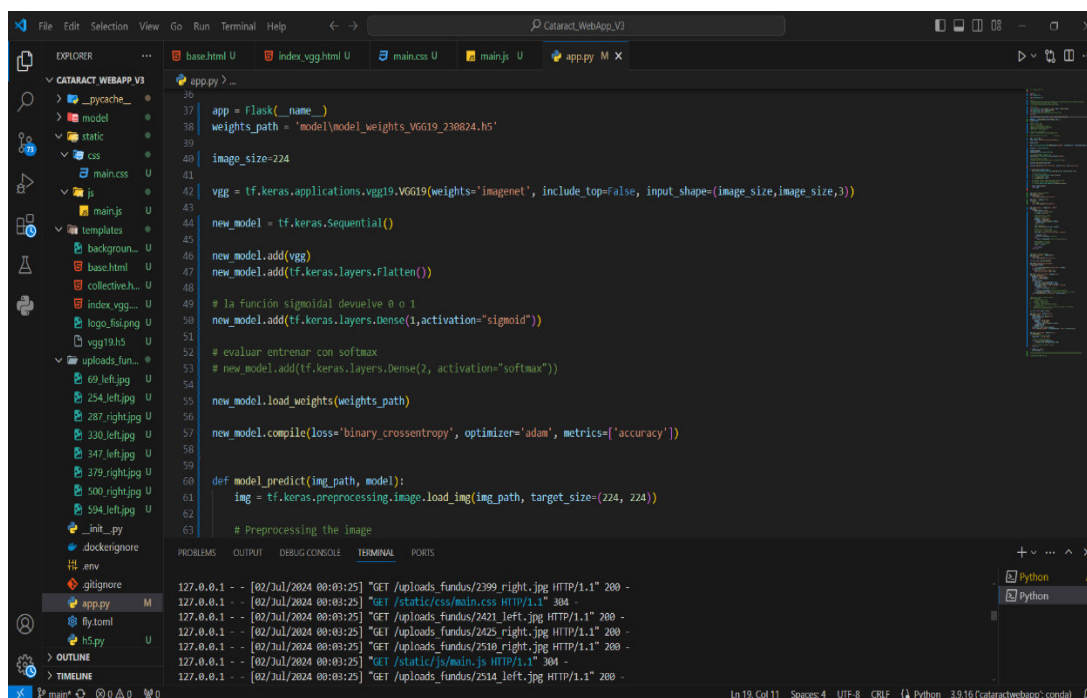
```
## guardar el modelo y sus pesos
model_json = model.to_json()
with open("model.json", "w") as json_file:
    json_file.write(model_json)
model.save_weights("model_weights_VGG19_230824.h5")
```

Nota. Elaboración propia.

Como se muestra en el **Figura 48**, tras seleccionar el modelo con mejor desempeño, se procede a exportar un archivo de extensión *h5* los pesos del entrenamiento para usarlas en una aplicación web desarrollada en *Python* junto con el *framework* de *Flask*. En la **Figura 49**, dicha aplicación web instancia un modelo del tipo ganador que importará los pesos del entrenamiento almacenado en el archivo *h5*.

Figura 49

Despliegue del modelo VGG-19 en aplicación web



Nota. Elaboración propia.

La interfaz del usuario de la aplicación web está diseñado de forma intuitiva para que un médico de primera línea o medicina general o un oftalmólogo novato puede navegar y subir una imagen para conocer el diagnóstico de un paciente ante un posible caso positivo de catarata en la versión individual como en la **Figura 50** o un caso negativo en la **Figura 51**.

Figura 50

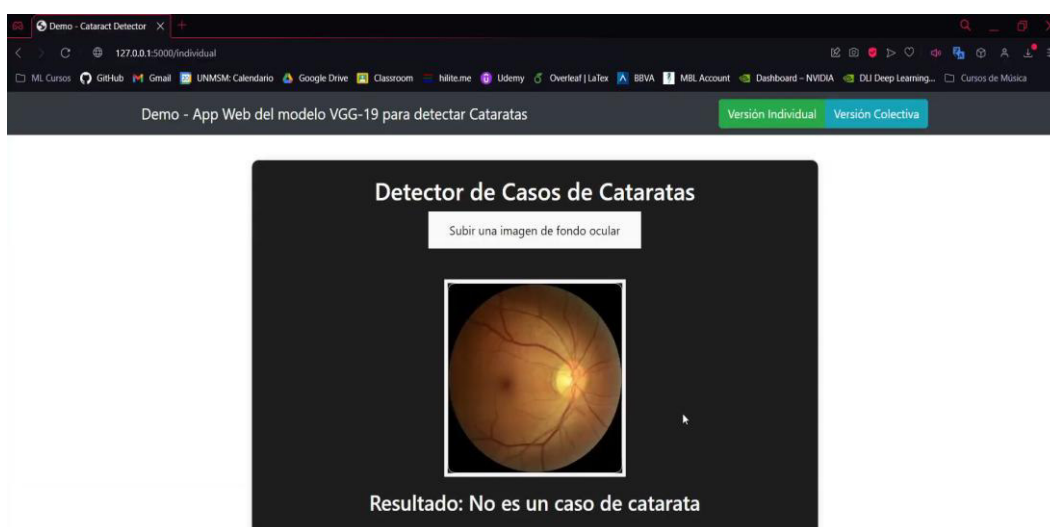
Caso positivo de catarata en la aplicación web



Nota. Elaboración propia.

Figura 51

Caso negativo de catarata en la aplicación web



Nota. Elaboración propia.

En la versión colectiva, la aplicación web permite al usuario final subir múltiples imágenes para brindar sus respectivos diagnósticos, de este modo se emula un potencial sistema de diagnóstico automatizado para ser empleado en campañas de salud como se aprecia en la **Figura 52**.

Figura 52

Versión colectiva de la aplicación web

Demo - App Web del modelo VGG-19 para detectar Cataratas
Versión Individual
Versión Colectiva

Detector de Casos de Cataratas

Seleccionar imágenes
Cargar imágenes

Detectar imágenes cargadas

Resultados

Imagen	Resultado
 (Paciente 1)	No es un caso de catarata
 (Paciente 2)	No es un caso de catarata
 (Paciente 3)	No es un caso de catarata
 (Paciente 4)	Catarata
 (Paciente 5)	Catarata

Nota. Elaboración propia.

Capítulo V – Discusión, Conclusiones y recomendaciones

5.1 Discusión

- Durante el entrenamiento de los modelos propuestos, se evaluó la posibilidad de descartar la técnica de preprocesamiento *image cropping*, debido a su impacto negativo en los resultados de evaluación.
- Al comparar los resultados del presente estudio con el estado del arte, se observa que el modelo VGG-19 entrenado alcanzó una precisión del 97% en la detección de cataratas, superando los valores reportados en investigaciones similares. Por ejemplo, Khan et al. (2022) obtuvieron una precisión del 94.03% para cataratas utilizando VGG-19, mientras que Vijayaraghavan (2022) reportó un 88% con el mismo modelo para múltiples enfermedades oculares. Modelos alternativos como InceptionV3 y ResNet50 alcanzaron precisiones de 91% y 94% respectivamente en este estudio, valores comparables pero inferiores al desempeño de VGG-19. Estos resultados evidencian que el modelo propuesto es eficaz, tanto en términos de desempeño como en estabilidad.
- Desde el enfoque metodológico, la adopción de la metodología CRISP-DM fue validada por autores como Vijayaraghavan (2022), quien destacó su utilidad para estructurar proyectos de aprendizaje profundo en el ámbito médico. La implementación de esta metodología en la presente investigación permitió una organización clara y reproducible del desarrollo del modelo, desde la comprensión del problema hasta su despliegue en una aplicación web funcional.
- Por otro lado, el juicio de expertos recopilado en este trabajo mediante cuestionarios permitió validar la aplicabilidad y pertinencia del modelo desde una perspectiva profesional. Las respuestas indicaron una alta valoración respecto a la utilidad clínica del sistema, especialmente como apoyo diagnóstico en contextos

con acceso limitado a especialistas. Este componente no suele estar presente en muchas investigaciones revisadas, por lo que representa un valor agregado relevante en términos de evaluación externa y viabilidad en entornos reales.

5.2 Conclusiones y recomendaciones

Como resultado del desarrollo de esta investigación, se alcanzaron todos los objetivos planteados, permitiendo formular las siguientes conclusiones:

- i. Se logró desarrollar un modelo basado en transferencia de aprendizaje capaz de detectar casos de catarata a partir de imágenes médicas oculares, alcanzando una precisión del 97% con el modelo VGG-19. Este resultado no solo valida la efectividad de la técnica implementada, sino que también supera el umbral mínimo de 85% propuesto como meta, cumpliendo con el objetivo general del estudio y evidenciando el éxito del enfoque planteado en el segundo objetivo específico.
- ii. Para el entrenamiento del modelo, se seleccionó y utilizó de forma adecuada el conjunto de datos ODIR, proveniente de un repositorio confiable. Este *dataset* permitió trabajar con una variedad representativa de imágenes oculares, lo que contribuyó significativamente al rendimiento final del sistema, en concordancia con el primer objetivo específico relacionado con la recolección de datos de calidad.
- iii. Durante el desarrollo de la solución, se adoptó la metodología CRISP-DM como guía de desarrollo, lo que permitió organizar de forma estructurada todas las fases del proyecto. Esta elección resultó efectiva desde la comprensión del problema hasta la evaluación del modelo, dando cumplimiento al tercer objetivo específico centrado en definir un enfoque metodológico adecuado.
- iv. El modelo entrenado fue implementado satisfactoriamente en una aplicación web, diseñada con una interfaz accesible para profesionales como médicos generales,

técnicos en atención primaria y oftalmólogos en formación. Este resultado responde directamente al cuarto objetivo específico, al permitir que la herramienta desarrollada sea utilizada por el usuario final de manera práctica y sencilla.

- v. Finalmente, se concluye que la propuesta para la detección automatizada de cataratas no solo es viable desde el punto de vista técnico, sino que posee un alto potencial de aplicación social. Esta solución podría contribuir significativamente a mejorar la detección temprana de cataratas, especialmente en contextos con acceso limitado a especialistas, reforzando así la pertinencia del objetivo general del estudio.

Respecto a las recomendaciones, se sugiere desplegar el sistema en la nube a través de una aplicación móvil, la cual permitiría mayor portabilidad y accesibilidad en diversos entornos. No obstante, se recomienda mantener una versión local del sistema, especialmente en aquellas zonas donde existan limitaciones de conectividad.

Por otro lado, se aconseja ampliar la base de datos utilizada para el entrenamiento del modelo, incorporando casos representativos del contexto nacional. En ese sentido, si bien se realizó un intento de solicitud de imágenes oculares al Instituto Nacional de Oftalmología del Perú (INO), este no tuvo éxito debido a las restricciones impuestas por la Ley General de Salud N.º 26842, la cual, en su artículo 15, establece que toda historia clínica debe contar con reserva y protección (Ministerio de Salud, 1997). Ante ello, se sugiere gestionar una nueva solicitud mediante el Vicerrectorado de Investigación, lo cual podría facilitar el acceso a datos bajo convenios institucionales formales.

Asimismo, se recomienda integrar imágenes obtenidas mediante lámparas de hendidura, según lo sugerido por expertos en oftalmología. Sin embargo, se debe considerar que los equipos

necesarios para su captura tienen un costo elevado, lo que puede representar una barrera operativa. Además, dado que los repositorios abiertos como Kaggle presentan una limitada variedad de imágenes especializadas, se sugiere incorporar conjuntos de datos que incluyan etiquetas específicas sobre el grado de severidad de las cataratas, como se expone en un artículo revisado (Zhang et al, 2022), con el fin de mejorar el modelo con nuevas capacidades de clasificación.

Adicionalmente, se propone la realización de pruebas piloto del sistema en centros médicos locales, con el fin de evaluar su desempeño en contextos reales de atención. Estas pruebas permitirán validar su utilidad clínica, identificar posibles mejoras y recibir retroalimentación directa de los profesionales que lo utilizarían.

Finalmente, se recomienda considerar las brechas logísticas y técnicas que podrían surgir en el despliegue del sistema automatizado en campañas oftalmológicas en regiones del interior del país, como la falta de conectividad a internet, el alto costo de instrumentos médicos especializados (por ejemplo, retinógrafos), o la necesidad de capacitación técnica del personal. Identificar y planificar estrategias para superar estas barreras es clave para lograr una implementación efectiva y sostenible del sistema propuesto, asimismo, la asesoría de un especialista en oftalmología para brindar continuidad y soporte en este proyecto.

Referencias

- Afroze, T., Akther, S., Chowdhury, M. A., Hossain, E., Hossain, M. S., & Andersson, K. (2021). *Glaucoma detection using Inception Convolutional Neural Network V3*. Communications in Computer and Information Science, 1435.
- Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, 40(4), 955–959. <https://doi.org/10.1177/001316448004000419>
- Bourne, R. R., Flaxman, S. R., Braithwaite, T., Cicinelli, M. V., Das, A., Jonas, J. B., et al. (2017). Magnitude, temporal trends, and projections of the global prevalence of blindness and vision impairment: a systematic review and meta-analysis. *The Lancet Global Health*, 5(9), e888–e897. [https://doi.org/10.1016/S2214-109X\(17\)30293-0](https://doi.org/10.1016/S2214-109X(17)30293-0)
- Brown, R. L. & Barrett, A. E. (2011). Visual impairment and quality of life among older adults: an examination of explanations for the relationship. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 66(3), 364–373.
- Browniee, J. (2019). *A gentle introduction to batch normalization for deep neural networks*. Recuperado de <https://machinelearningmastery.com/batch-normalization-for-training-of-deep-neural-networks/>
- Campos, B., Montjoy, A., Dulanto, E., Gonzales, V., Tecse, C., et al. (2014). Prevalencia y causas de ceguera en Perú: encuesta nacional. *Pan American Journal of Public Health*, 36(5), 283–289.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0 step-by-step data mining guide*. Technical report, CRISP-DM.
- Data Science Team. (2020). *Redes neuronales residuales - Lo que necesitas saber (ResNet)*. Recuperado de <https://datascience.eu/es/aprendizaje-automatico/una->

[vision-general-de-resnet-y-sus-variantes/](#)

Digital Diagnostics. (2022). *IDx-DR: Intelligent diabetic retinopathy detection*.

Recuperado de <https://www.digitaldiagnostics.com/idx-dr/>

Espinosa-Zúñiga, J. J. (2020). Aplicación de metodología CRISP-DM para segmentación geográfica de una base de datos pública. *Ingeniería Investigación y Tecnología*, 21(1), 1–13. <https://doi.org/10.22201/fi.25940732e.2020.21n1.008>.

Eyenuk Inc. (2023). *EyeArt AI Eye Screening System*. Recuperado de <https://www.eyenuk.com/eyeart>

FAYRIX. (s.f.). *Selección de métricas para aprendizaje automático*. Recuperado de <https://fayrix.com/machine-learning-metrics>

Fenner, M. (2022). *Machine learning with Python for everyone*. Pearson Education.

Google Cloud. (2023). *Guía avanzada de Inception v3*. Recuperado de <https://cloud.google.com/tpu/docs/inception-v3-advanced?hl=es-419>

Guidelines for performing Systematic Literature Reviews in Software Engineering. (2007).

Version 2.3 (EBSE-2007-01). School of Computer Science and Mathematics, Keele University.

Recuperado de

https://www.elsevier.com/_data/promis_misc/525444systematicreviewsguide.pdf

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., et al. (2016). *Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs*. JAMA, 316(22), 2402–2410. <https://doi.org/10.1001/jama.2016.17216>

Hernández-Sampieri, R., & Mendoza, C. (2006). *Metodología de la investigación*. McGraw Hill Education.

Hu, S., Luan, X., Wu, H., Yang, Y., & Huang, D. (2021). ACCV: automatic classification algorithm of cataract video based on deep learning. *BioMed Engineering OnLine*,

20(1), 78. <https://doi.org/10.1186/s12938-021-00906-3>

- Huaman, E., Anicama, E., González, E., Félix, H., & Chu, W. (2022). *Metodología de la investigación científica: Guía práctica para la elección, diseño y desarrollo de la investigación*. Universidad Autónoma de Ica. Recuperado de <https://repositorio.autonomadeica.edu.pe/bitstream/20.500.14441/2558/2/METODOLOGIA%20DE%20LA%20INVESTIGACIÓN%20CIENTIFICA.pdf>
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>
- Khan, M. S., Tafshir, N., Alam, K. N., Dhruva, A. R., et al. (2022). Deep learning for ocular disease recognition: An inner-class balance. *Computational Intelligence and Neuroscience*, 2022.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint*, arXiv:1412.6980.
- Larxel. (2019). *Ocular disease recognition*. Recuperado de <https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k>
- Li, H., Liu, H., Hu, Y., Fu, H., Zhao, Y., Miao, H., & Liu, J. (2022). An annotation-free restoration network for cataractous fundus images. *arXiv preprint*. Recuperado de <http://arxiv.org/abs/2203.07737>
- Lindholm, A., Wahlström, N., Lindsten, F., & Schön, T. B. (2019). *Supervised machine learning*. Aimlcommunity.com. Recuperado de <https://aimlcommunity.com/wp-content/uploads/2019/09/Supervised-ML.pdf>
- Marbán, M. S. (2018). *A data mining & knowledge discovery process model*. Recuperado de http://cdn.intechopen.com/pdfs/5937/InTech-A_data_mining_amp_knowledge_discovery_process_model.pdf
- Martínez, J. (2020). *Precision, recall, F1, accuracy en clasificación*. IArtificial.net.

Recuperado de https://github.com/jmartinezheras/2018-MachineLearning-Lectures-ESA/blob/master/iartificial_net/IArtificial_net.pdf

Mayo Clinic. (s.f.). *Cataratas*. Recuperado de <https://www.mayoclinic.org/es-es/diseases-conditions/cataracts/symptoms-causes/syc-20353790>.

Ministerio de Salud. (1997). *Ley General de Salud - Ley N° 26842*. Recuperado de <https://www.gob.pe/institucion/minsa/normas-legales/256661-26842>

Minsa. (2022). *Plan de gestión de salud ocular y prevención de la ceguera 2022–2026*.
Ministerio de Salud del Perú.

Ñuflo, I. & Mecca, F. (2022). *Sistema para diagnosticar la catarata mediante el uso de redes neuronales convolucionales* (Tesis de licenciatura, Universidad Nacional Mayor de San Marcos). <https://hdl.handle.net/20.500.12672/18348>

Organización Mundial de la Salud. (2020). *Informe mundial sobre la visión*. OMS, Ginebra.

Organización Mundial de la Salud. (2022). *Ceguera y discapacidad visual*. Recuperado de <https://www.who.int/es/news-room/fact-sheets/detail/blindness-and-visual-impairment>.

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.
<https://doi.org/10.1109/TKDE.2009.191>

Pinos, E. (2019). *Prevención y detección temprana de glaucoma a través del procesamiento digital de imágenes del fondo del ojo humano y de la variación de la presión intraocular y su correlación con el espesor central de la córnea* (Tesis doctoral en Ingeniería con mención en Automatización, Universidad de Piura).
<https://hdl.handle.net/11042/4528>

Pintusaini. (2022). *Compute classification report and confusion matrix in Python*.

- GeeksforGeeks. Recuperado de <https://www.geeksforgeeks.org/compute-classification-report-and-confusion-matrix-in-python/>
- Rainey, L., Elsmann, E. B. M., Van Nispen, R. M. A., Van Leeuwen, L. M., & Van Rens, G. H. M. B. (2016). Comprehending the impact of low vision on the lives of children and adolescents: a qualitative approach. *Quality of Life Research*, 25, 2633–2643.
- Robalino, H. (2021). *Diagnóstico precoz de la retinopatía diabética utilizando visión computacional* (Tesis de licenciatura, Universidad Ricardo Palma). <https://hdl.handle.net/20.500.14138/4530>
- Roberts, A. (2023). *Binary cross entropy: Where to use log loss in model monitoring*. Arize AI. Recuperado de <https://arize.com/blog-course/binary-cross-entropy-log-loss/>
- Ruzicki, J., Holden, M., Cheon, S., Ungi, T., Egan, R., & Law, C. (2022). Use of machine learning to assess cataract surgery skill level with tool detection. *Ophthalmology Science*, 3(1), 100235. <https://doi.org/10.1016/j.xops.2022.100235>
- Sarki, R., Ahmed, K., Wang, H., & Zhang, Y. (2020a). Automatic detection of diabetic eye disease through deep learning using fundus images: A survey. *IEEE Access*, 8, 151133–151149.
- Sarki, R., Ahmed, K., Wang, H., Michalska, S., & Zhang, Y. (2020b). Early detection of diabetic eye disease from fundus images with deep learning. En R. Borovica-Gajic, J. Qi, & W. Wang (Eds.), *Databases Theory and Applications. ADC 2020. Lecture Notes in Computer Science* (Vol. 12008). Springer. https://doi.org/10.1007/978-3-030-39469-1_20
- Sarmiento-Ramos, J. L. (2020). Aplicaciones de las redes neuronales y el deep learning a la ingeniería biomédica. *Revista UIS Ingenierías*, 19(4), 1–18. <https://doi.org/10.18273/revuin.v19n4-2020001>
- Simplilearn. (2020). *What is a confusion matrix in machine learning?* Recuperado de

<https://www.simplilearn.com/tutorials/machine-learning-tutorial/confusion-matrix-machine-learning>

- Sotaquirá, M. (2018). *La función de activación*. Codificando Bits. Recuperado de <https://www.codificandobits.com/blog/funcion-de-activacion/>
- Ting, D. S. W., Peng, L., Varadarajan, A. V., Keane, P. A., et al. (2019). Deep learning in ophthalmology: The technical and clinical considerations. *Progress in Retinal and Eye Research*, 72.
- Usha, R., Theerthagiri, P., Jeena, I., & Vamsidhar, Y. (2020). Binary cross entropy with deep learning technique for image classification. *International Journal of Advanced Trends in Computer Science and Engineering*, 9. <https://doi.org/10.30534/ijatcse/2020/175942020>
- Venkat, G. (2022). *Eye diseases classification*. Recuperado de <https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification/data>
- Vijayaraghavan, U. (2022). *Detection and revelation of multiple ocular diseases using transfer learning techniques* (Tesis de maestría, National College of Ireland). <https://norma.ncirl.ie/id/eprint/6336>
- Wang, Y., Tang, C., Wang, J., Sang, Y., & Lv, J. (2021). Cataract detection based on ocular B-ultrasound images by collaborative monitoring deep learning. *Knowledge-Based Systems*, 231, 107442. <https://doi.org/10.1016/j.knosys.2021.107442>
- Xu, L., Yang, C. W., Yang, H., Wang, S., Shi, Y. Y., & Song, D. X. (2010). The study of predicting the visual acuity after phacoemulsification according to the blur level of fundus photography. *Ophthalmology in China*, 19(2), 81–83.
- Yacouby, R., & Axman, D. (2020). Probabilistic extension of precision, recall, and F1 score for more thorough evaluation of classification models. *Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems*.

Yadav, S., & Yadav, J. K. P. S. (2023). Automatic cataract severity detection and grading using deep learning. *Journal of Sensor*, 2023, 1–17.

<https://doi.org/10.1155/2023/2973836>

Zhang, X.-Q., Hu, Y., Xiao, Z.-J., Fang, J.-S., Higashita, R., & Liu, J. (2022). Machine learning for cataract classification/grading on ophthalmic imaging modalities: A survey. *Machine Intelligence Research*, 19(3), 184–208.

<https://doi.org/10.1007/s11633-022-1329-0>.

Anexos

ANEXO 1: Publicaciones seleccionadas y preguntas de investigación

ID	TÍTULO	AUTOR(ES)	AÑO	PAIS	JOURNAL /PUBLISHER	REPOSITORIO	P1: ¿QUÉ MODELOS DE TRANSFERENCIA DE APRENDIZAJE SON APLICABLES PARA LA DETECCIÓN DE ENFERMEDADES OCULARES?	P2: ¿QUÉ TÉCNICAS DE PREPROCESAMIENTO DE IMÁGENES SON USADOS PARA MEJORAR EL RENDIMIENTO DEL MODELO DE TRANSFERENCIA DE APRENDIZAJE PARA LA DETECCIÓN DE ENFERMEDADES OCULARES?	P3: ¿QUÉ METODOLOGÍAS SON USADOS PARA EL DESARROLLO DE UN MODELO DE TRANSFERENCIA DE APRENDIZAJE PARA LA DETECCIÓN DE ENFERMEDADES OCULARES?	P4: ¿QUÉ MÉTRICAS SON USADOS PARA EVALUAR EL DESEMPEÑO DE UN MODELO DE TRANSFERENCIA DE APRENDIZAJE PARA LA DETECCIÓN DE ENFERMEDADES OCULARES?
1	Deep Learning for Ocular Disease Recognition: An Inner-Class Balance	Khan, M. S., Tafshir, N., Alam, K. N., Dhruba, A. R., Khan, M. M., Albraikan, A. A., & Almalki, F. A.	2022	Arabia Saudita	Hindawi	Scopus	1	1	1	1
2	Detection and Revelation of Multiple Ocular Diseases using Transfer Learning Techniques	Uppiliappan Vijayaraghavan	2022	Irlanda	-	Scholar Google	1	1	1	1
3	Machine Learning for Cataract Classification/Grading on Ophthalmic Imaging Modalities: A Survey	Xiao-Qing Zhang, Yan Hu, Zun-Jie Xiao, Jian-Sheng Fang, Risa Higashita, Jiang Liu	2022	China	Machine Intelligence Research	Springer	1	1	1	1
4	Glaucoma Detection Using Inception	Afroze, T., Akther, S.,	2021	Bangladés	-	Scopus	1	1	1	1

	Convolutional Neural Network V3	Chowdhury, M. A., Hossain, E., Hossain, M. S., & Andersson, K.								
5	Cataract detection based on ocular B-ultrasound images by collaborative monitoring deep learning	Yong Wang, Chenwei Tang, Jian Wang, Yongsheng Sang, Jiancheng Lv, Shenming Hu, Xinze Luan, Hong Wu, Xiaoting Wang, Chunhong Yan, Jingying Wang, Guantong Liu & Wei He	2021	China	Knowledge-Based Systems	Science Direct	1	1	1	1
6	ACCV: automatic classification algorithm of cataract video based on deep learning	Sarki, R., Ahmed, K., Wang, H., & Zhang, Y.	2021	China	BioMedical Engineering OnLine	Springer	1	1	-	1
7	Automatic Detection of Diabetic Eye Disease Through Deep Learning Using Fundus Images: A Survey	Rubina Sarki, Khandakar Ahmed & Yanchun Zhang	2020	Australia	IEEE Access	IEEE Xplorer	1	1	-	1
8	Early detection of diabetic eye disease through deep learning using fundus images	Yadav, S., & Yadav, J. K. P. S	2020	Australia	EAI Endorsed Transactions on Pervasive Health and Technology	Mendeley	1	1	-	1
9	Automatic Cataract Severity Detection and Grading Using Deep Learning	Ting, D. S. W., Peng, L., Varadarajan, A. V., Keane, P.	2023	Malasia	Journal of Sensor	Wiley	1	1	1	1
10	Deep Learning in Ophthalmology: The Technical and Clinical Considerations		2019	Reino Unido	Progress in Retinal and Eye Research	Scopus	1	-	-	1

		A., Burlina, P. M., Chiang, M. F., Schmetterer, L., Pasquale, L. R., Bressler, N. M., Webster, D. R., Abramoff, M., & Wong, T. Y.								
11	An Annotation-free Restoration Network for Cataractous Fundus Images	Heng Li, Haofeng Liu, Yan Hu, Huazhu Fu, Yitian Zhao, Hanpei Miao, Jiang Liu	2022	China	IEEE Transactions on Medical Imaging	IEEE Explorer	1	1	1	1
12	Use of Machine Learning to Assess Cataract Surgery Skill Level with Tool Detection	Ruzicki J., Holden M., Cheon S., Ungi T., Egan R. & Law C.	2022	Canadá	Ophthalmology Science	Elsevier	1	-	-	1
13	Diagnóstico precoz de la retinopatía diabética utilizando visión computacional	Robalino, H.	(2021)	Perú	Universidad Ricardo Palma	Alicia Concytec	1	1	1	1
14	Prevención y detección temprana de glaucoma a través del procesamiento digital de imágenes del fondo del ojo humano y de la variación de la presión intraocular y su correlación con el espesor central de la córnea	Pinos, E.	(2019)	Perú	Universidad de Piura	Alicia Concytec	-	1	1	1
15	Sistema para diagnosticar la catarata	Ñuflo, I. & Mecca, F.	(2022)	Perú	Universidad Nacional	Alicia	1	1	1	1

	mediante el uso de redes neuronales convolucionales	Mayor de San Marcos	Concytec	
--	---	------------------------	----------	--

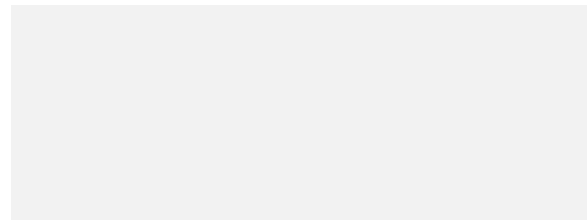
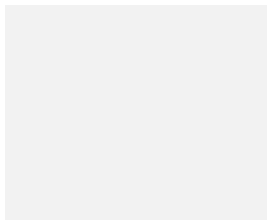
Nota. Elaboración propia.

ANEXO 2: Matriz de Consistencia

PROBLEMA	OBJETIVOS	VARIABLES E MÉTRICAS	POBLACIÓN Y MUESTRA	METODOLOGÍA	INSTRUMENTOS
Problema general: ¿Cómo se puede aplicar un modelo de transferencia de aprendizaje para la detección de casos de cataratas?	Objetivo general: Desarrollar un modelo basado en transferencia de aprendizaje que permita detectar casos de cataratas a partir de imágenes médicas oculares.	Variable independiente: Modelos de transferencia de aprendizaje (VGG-19, InceptionV3, ResNet50)	Población: Imágenes médicas de fondo ocular (<i>dataset</i> ODIR).	Tipo de investigación: Aplicada - predictiva	
Problemas específicos: 1. ¿Qué tipo de imágenes médicas son aplicables en un modelo de transferencia de aprendizaje para la detección de cataratas? 2. ¿Con qué precisión el modelo de transferencia de aprendizaje detectará un caso de catarata en una imagen de fondo ocular? 3. ¿Qué metodología se usará de guía para el desarrollo del modelo de Transferencia de Aprendizaje destinado a la detección de catarata? 4. ¿Cómo se implementará el modelo de Transferencia de Aprendizaje detector de catarata para que esté	Objetivos específicos: 1. Encontrar un <i>dataset</i> de imágenes médicas oculares de un repositorio confiable destinado a la detección de cataratas mediante un modelo de transferencia de aprendizaje. 2. Alcanzar una precisión mínima del 85% en la detección de casos de cataratas en imágenes médicas a través un modelo de transferencia de aprendizaje. 3. Determinar la metodología a usar como guía para el desarrollo de un modelo de transferencia de aprendizaje destinada a la detección de casos de catarata.	Variable dependiente: Detección de casos de catarata. Métricas: 1. Accuracy 2. Precision 3. Recall 4. F1-Score	Muestra: Subconjunto de imágenes divididas en categorías (con catarata / casos sanos).	Diseño de investigación: Cuasi -experimental Metodología: CRISP-DM para el desarrollo del sistema. Comparación de modelos y evaluación experimental.	1. Reporte de clasificación 2. Matriz de confusión 3. Cuestionario de validación con expertos en oftalmología (juicio experto)

disponible para el manejo del usuario final?

4. Desplegar el modelo de Transferencia de Aprendizaje para la detección de casos de cataratas en una aplicación web para el manejo del usuario final.



Nota. Elaboración propia.

ANEXO 3: Matriz de Operacionalización

VARIABLE	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	DIMENSIONES	INDICADORES	ESCALA DE MEDICIÓN	UNIDAD DE ANÁLISIS
MODELOS DE TRANSFERENCIA DE APRENDIZAJE	Técnica de <i>Deep Learning</i> que reutiliza modelos entrenados previamente para mejorar el rendimiento en nuevas tareas, con menos datos y menor costo computacional.	Implementación de los modelos VGG-19, InceptionV3 y ResNet50 sobre el conjunto de datos ODIR para detectar cataratas en imágenes de fondo ocular.	Arquitectura del modelo	- Tipo de modelo seleccionado - Tiempo de entrenamiento - Precisión alcanzada	- Categorica (modelo) - Ratio (precisión)	Modelos computacionales entrenados
DETECCIÓN DE CASOS DE CATARATA	Proceso automático mediante el cual se identifica la presencia o ausencia de cataratas en imágenes de fondo ocular utilizando métricas de evaluación.	Evaluación de las predicciones del modelo sobre el conjunto de datos de prueba mediante métricas de clasificación.	- Sensibilidad - Especificidad - Precisión - Exactitud - F1 Score	- Porcentaje de verdaderos positivos (TP) - Porcentaje de verdaderos negativos (TN) - Precisión total	Porcentaje (%)	Imágenes clasificadas

Nota. Elaboración propia.

ANEXO 4: Cuestionarios de Juicio de Expertos

FICHA DE MEDICOS (DNI, CMP, CELULAR)

	Médico 1	Médico 2	Médico 3
Nombres y Apellidos	Luis Washington Lu Basauri	Christian Paul Limaylla Guerrero	Antonio Mariano Carmona Huaman
DNI	10277015	45047420	47836394
Código CMP	009625	068874	076858
Correo electrónico	ljlub5015@gmail.com	limaylla.christian@gmail.com	antonio_mariano_04@hotmail.com

Nota. Elaboración propia.

Cuestionario 1:

FICHA DE VALIDACIÓN – JUICIO DE EXPERTO

TÍTULO DE TESIS: Aplicación de modelos de Transferencia de Aprendizaje para la detección de casos de catarata

OBJETIVO DE LA EVALUACIÓN: Recoger la opinión de especialistas en salud ocular sobre la pertinencia, aplicabilidad e impacto del proyecto desarrollado, con el fin de fortalecer la calidad y relevancia del trabajo de tesis.

Nombre completo: Dr. Luis J. Lú Basauri

DNI: 10277015

Grado Académico: Superior

Tiempo laboral: 34 años

Centro médico u hospital donde labora:
Clínica Perseus de la Visión

Correo electrónico: ljlub5015@gmail.com

Firma: 

Fecha: 4.4.26

Nota. Elaboración propia.

Cuestionario para el Juicio de Expertos

Instrucciones: Por favor, marque con una "X" la opción que mejor represente su opinión. También puede incluir observaciones adicionales al final del formulario.

Pregunta	Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo
El tema abordado en el proyecto es relevante para la práctica médica en oftalmología.	☐	☐	☐	☐	☒
El uso de modelos basados en inteligencia artificial puede mejorar el diagnóstico de cataratas.	☐	☐	☐	☒	☐
La detección automatizada de cataratas mediante imágenes oculares es clínicamente útil.	☐	☐	☐	☒	☐
Este tipo de herramientas podrían implementarse en centros médicos de bajos recursos.	☐	☐	☐	☐	☒
Un sistema como el propuesto podría contribuir a reducir la ceguera por catarata en zonas vulnerables.	☐	☐	☐	☒	☐

Observaciones o sugerencias adicionales:

Nota. Elaboración propia.

Cuestionario 2:

FICHA DE VALIDACIÓN – JUICIO DE EXPERTO

TÍTULO DE TESIS: Aplicación de modelos de Transferencia de Aprendizaje para la detección de casos de catarata

OBJETIVO DE LA EVALUACIÓN: Recoger la opinión de especialistas en salud ocular sobre la pertinencia, aplicabilidad e impacto del proyecto desarrollado, con el fin de fortalecer la calidad y relevancia del trabajo de tesis.

Nombre completo: Christian Paul Limaylla Guerrero

DNI: 45047420

Grado Académico: Cirujano Oftalmólogo

Tiempo laboral: 6 años

Centro médico u hospital donde labora:

Clinica Peruana de la Vision

Correo electrónico: limaylla.christian@gmail.com

Firma: 

Fecha: 04/04/25

Nota. Elaboración propia.

Cuestionario para el Juicio de Expertos

Instrucciones: Por favor, marque con una "X" la opción que mejor represente su opinión. También puede incluir observaciones adicionales al final del formulario.

Pregunta	Totalmente en desacuerdo	En desacuerdo	NI de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo
El tema abordado en el proyecto es relevante para la práctica médica en oftalmología.	☐	☐	☐	☒	☐
El uso de modelos basados en inteligencia artificial puede mejorar el diagnóstico de cataratas.	☐	☐	☐	☒	☐
La detección automatizada de cataratas mediante imágenes oculares es clínicamente útil.	☐	☐	☐	☐	☒
Este tipo de herramientas podrían implementarse en centros médicos de bajos recursos.	☐	☐	☐	☐	☒
Un sistema como el propuesto podría contribuir a reducir la ceguera por catarata en zonas vulnerables.	☐	☐	☐	☐	☒

Observaciones o sugerencias adicionales:

El diagnóstico de catarata es únicamente con lámpara de hendidura; el fondo de ojo poco visible ayudaría a sospechar catarata, pero requiere una foto clínica que evidencie la catarata. Puede obtenerse con lámpara de hendidura o con dispositivos Smart.

Nota. Elaboración propia.

Cuestionario 3:

FICHA DE VALIDACIÓN – JUICIO DE EXPERTO

TÍTULO DE TESIS: Aplicación de modelos de Transferencia de Aprendizaje para la detección de casos de catarata

OBJETIVO DE LA EVALUACIÓN: Recoger la opinión de especialistas en salud ocular sobre la pertinencia, aplicabilidad e impacto del proyecto desarrollado, con el fin de fortalecer la calidad y relevancia del trabajo de tesis.

Nombre completo: Antonio Carrasco Huamán

DNI: 47836394

Grado Académico: Superior

Tiempo laboral: 9 meses

Centro médico u hospital donde labora:

Clinica Poma de la Vision

Correo electrónico: antonio - merino 04@hotmail.com

Firma: 

Fecha: 15/04/25

Nota. Elaboración propia.

Cuestionario para el Juicio de Expertos

Instrucciones: Por favor, marque con una "X" la opción que mejor represente su opinión.
También puede incluir observaciones adicionales al final del formulario.

Pregunta	Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo
El tema abordado en el proyecto es relevante para la práctica médica en oftalmología.	☐	☐	☐	☒	☐
El uso de modelos basados en inteligencia artificial puede mejorar el diagnóstico de cataratas.	☐	☐	☐	☒	☐
La detección automatizada de cataratas mediante imágenes oculares es clínicamente útil.	☐	☐	☐	☒	☐
Este tipo de herramientas podrían implementarse en centros médicos de bajos recursos.	☐	☐	☐	☒	☐
Un sistema como el propuesto podría contribuir a reducir la ceguera por catarata en zonas vulnerables.	☐	☐	☐	☒	☐

Observaciones o sugerencias adicionales:

Nota. Elaboración propia.