



UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES

**CENTRO DE CIENCIAS BÁSICAS
DEPARTAMENTO
DE CIENCIAS DE LA COMPUTACIÓN**

TÍTULO

SISTEMA AUTOMÁTICO PARA LA INTERPRETACIÓN INMEDIATA DE MAMOGRAFÍAS PARA
LA DETERMINACIÓN DEL RIESGO DE CÁNCER.

PRESENTA

ING. CÉSAR EDUARDO MUÑOZ CHAVEZ.

**PARA OPTAR POR EL GRADO EN
MAESTRÍA EN CIENCIAS CON OPCIONES A LA COMPUTACIÓN, MATEMÁTICAS
APLICADAS**

TUTORES:

DR HERMILO SÁNCHEZ CRUZ.
DR. JUAN HUMBERTO SOSSA AZUELA.

ASESOR:

DR. JULIO CÉSAR PONCE GALLEGOS.

Aguascalientes, Ags., 28 de marzo de 2023

CARTA DE VOTO APROBATORIO
INDIVIDUAL

DR. JORGE MARTÍN ALFÉREZ CHÁVEZ
DECANO (A) DEL CENTRO DE CIENCIAS BASICAS

PRESENTE

Por medio del presente como **TUTOR** designado del estudiante **CÉSAR EDUARDO MUÑOZ CHÁVEZ** con ID **165289** quien realizó la tesis titulada: **SISTEMA AUTOMÁTICO PARA LA INTERPRETACIÓN INMEDIATA DE MAMOGRAFÍAS PARA LA DETERMINACIÓN DEL RIESGO DE CÁNCER**, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el VOTO APROBATORIO, para que él pueda proceder a imprimirla así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE
"Se Lumen Proferre"

Aguascalientes, Ags., a día de 4 de abril del 2023.



Dr. Hermilo Sánchez Cruz
Tutor de tesis

c.c.p.- Interesado
c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado.
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad.
Aprobado por: Depto. Control Escolar/ Depto. Apoyo al Posgrado.

Código: DO-SEE-FO-07
Actualización: 01
Emisión: 17/05/19

CARTA DE VOTO APROBATORIO
INDIVIDUAL

DR. JORGE MARTÍN ALFÉREZ CHÁVEZ
DECANO (A) DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como **TUTOR** designado del estudiante **CÉSAR EDUARDO MUÑOZ CHÁVEZ** con ID **165289** quien realizó la tesis titulada: **SISTEMA AUTOMÁTICO PARA LA INTERPRETACIÓN INMEDIATA DE MAMOGRAFÍAS PARA LA DETERMINACIÓN DEL RIESGO DE CÁNCER**, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el VOTO APROBATORIO, para que él pueda proceder a imprimirla así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE
"Se Lumen Proferre"

Aguascalientes, Ags., a día de 4 de abril del 2023.



Dr. Juan Humberto Sossa Azuela
Tutor de tesis

El nombre completo que aparece en el Voto Aprobatorio debe coincidir con el que aparece en el documento pdf. No se puede abreviar, ni omitir nombres

c.c.p.- Interesado
c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado.
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad.
Aprobado por: Depto. Control Escolar/ Depto. Apoyo al Posgrado.

Código: DO-SEE-PO-07
Actualización: 01
Emisión: 17/05/19

CARTA DE VOTO APROBATORIO
INDIVIDUAL

DR. JORGE MARTÍN ALFÉREZ CHÁVEZ
DECANO (A) DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como **ASESOR** designado del estudiante **CÉSAR EDUARDO MUÑOZ CHÁVEZ** con ID **165289** quien realizó la tesis titulada: **SISTEMA AUTOMÁTICO PARA LA INTERPRETACIÓN INMEDIATA DE MAMOGRAFÍAS PARA LA DETERMINACIÓN DEL RIESGO DE CÁNCER**, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el **VOTO APROBATORIO**, para que él pueda proceder a imprimirla así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE
"Se Lumen Proferre"

Aguascalientes, Ags., a día de 4 de abril del 2023



Dr. Julio César Ponce Gallegos
Asesor de tesis

c.c.p.- interesado
c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad
Aprobado por: Depto. Control Escolar/Depto. Apoyo al Posgrado

Código: DG-SEE-1000
Actualización: 01
Fecha: 17/05/19



DICTAMEN DE LIBERACIÓN ACADÉMICA PARA INICIAR LOS TRÁMITES DEL EXAMEN DE GRADO



Fecha de dictaminación dd/mm/aaaa: 13-04-2023

NOMBRE: CESAR EDUARDO MUÑOZ CHAVEZ ID: 165289

PROGRAMA: MAESTRÍA EN CIENCIAS, CON OPCIONES A LA COMPUTACIÓN, MATEMÁTICAS APLICADAS UGAC (del posgrado): COMPUTACIÓN-INTELIGENCIA ARTIFICIAL

TIPO DE TRABAJO: Tesis Trabajo Práctico

TITULO: SISTEMA AUTOMÁTICO PARA LA INTERPRETACIÓN INMEDIATA DE MAMOGRAFÍAS PARA LA DETERMINACIÓN DEL RIESGO DE CÁNCER.

IMPACTO SOCIAL (señalar el impacto logrado): Un sistema automático para la interpretación de mamografías podría ayudar a los profesionales de la salud a detectar signos tempranos de cáncer de mama, lo que podría llevar a un diagnóstico más rápido y a un tratamiento más efectivo.

INDICAR SI NO N.A. (NO APLICA) SEGÚN CORRESPONDA:

INDICAR	SI	NO	N.A. (NO APLICA)	SEGÚN CORRESPONDA:
<i>Elementos para la revisión académica del trabajo de tesis o trabajo práctico:</i>				
SI				El trabajo es congruente con los UGAC del programa de posgrado
SI				La problemática fue abordada desde un enfoque multidisciplinario
SI				Existe coherencia, continuidad y orden lógico del tema central con cada apartado
SI				Los resultados del trabajo dan respuesta a las preguntas de investigación o a la problemática que aborda
SI				Los resultados presentados en el trabajo son de gran relevancia científica, tecnológica o profesional según el área
SI				El trabajo demuestra más de una aportación original al conocimiento de su área
SI				Las aportaciones responden a los problemas prioritarios del país
N.A.				Generó transferencia del conocimiento o tecnológica
SI				Cumple con la ética para la investigación (reporte de la herramienta antiplagio)
<i>El egresado cumple con lo siguiente:</i>				
SI				Cumple con lo señalado por el Reglamento General de Docencia
SI				Cumple con los requisitos señalados en el plan de estudios (créditos curriculares, optativos, actividades complementarias, estancia, predoctoral, etc)
SI				Cuenta con los votos aprobatorios del comité tutelar, en caso de los posgrados profesionales si tiene solo tuto(a) podrá liberar solo el tutor
N.A.				Cuenta con la carta de satisfacción del usuario
SI				Coincide con el título y objetivo registrado
SI				Tiene congruencia con cuerpos académicos
SI				Tiene el CVU del Conacyt actualizado
N.A.				Tiene el artículo aceptado o publicado y cumple con los registros institucionales (en caso que proceda)
<i>En caso de Tesis por artículos científicos publicados:</i>				
N.A.				Aceptación o publicación de los artículos según el nivel del programa
N.A.				El estudiante es el primer autor
N.A.				El autor de correspondencia es el Tutor del Nucleo Académico Básico
N.A.				En los artículos se ven reflejados los objetivos de la tesis, ya que son producto de este trabajo de investigación
N.A.				Los artículos integran los capítulos de la tesis y se presentan en el idioma en que fueron publicados
N.A.				La aceptación o publicación de los artículos en revistas indexadas de alto impacto

Con base a estos criterios, se autoriza se continúen con los trámites de titulación y programación del examen de grado:

SI
No

FIRMAS

Elaboró:

* NOMBRE Y FIRMA DEL CONSEJERO SEGÚN LA UGAC DE ADSCRIPCIÓN:

Dr. Hermilio Sánchez Cruz

NOMBRE Y FIRMA DEL SECRETARIO TÉCNICO:

Dr. Hermilio Sánchez Cruz

* En caso de conflicto de intereses, firmará un revisor miembro del SAB de la UGAC correspondiente distinto al tutor o miembro del comité tutelar, designado por el Decano

Revisó:

NOMBRE Y FIRMA DEL SECRETARIO DE INVESTIGACIÓN Y POSGRADO:

Dr. Juan Jáuregui Ruzón

Autorizó:

NOMBRE Y FIRMA DEL DECANO:

M. en C. Jorge Martín Álvarez Chávez

Nota: procede el trámite para el Depto. de Apoyo al Posgrado

Te cursillamos con el Art. 105C del Reglamento General de Docencia que a la letra señala entre las funciones del Consejo Académico: "Cuidar la eficiencia técnica del programa de posgrado y el Art. 105F las funciones del Secretario Técnico, tener el seguimiento de los alumnos."

Agradecimientos

Quiero expresar mi profunda gratitud al Consejo Nacional de Ciencia y Tecnología (CONACYT) por brindarme la beca que hizo posible estudiar en la Universidad Autónoma de Aguascalientes (UAA) y concluir mi maestría en ciencias.

También agradezco a la UAA por ser una excelente institución académica, que me brindó la oportunidad de desarrollarme profesionalmente y lograr mis objetivos.

Agradezco de manera especial a mis asesores, el Dr. Juan Humberto Sossa Azuela y el Dr. Julio César Ponce Gallegos, por su invaluable guía y asesoría en los diferentes temas que abordé en esta tesis.

No puedo dejar de mencionar a mi director de Tesis, el Dr. Hermilo Sánchez Cruz, por su serenidad, dedicación y disposición para enseñar y ayudar en todo momento.

Quiero agradecer a mi familia por su inmenso apoyo durante todo el camino, por sus consejos, comprensión y motivación en los momentos difíciles, y por la confianza que depositaron en mí.

Finalmente, agradezco a la vida por brindarme la oportunidad de concluir con éxito este logro personal y profesional. ¡Gracias a todos!

Dedicatoria

Quiero dedicar esta tesis a mis padres, quienes han sido mi mayor apoyo durante toda mi maestría. Todo lo que he logrado es gracias a su amor y constante motivación para ser una mejor persona cada día. Agradezco profundamente su inquebrantable apoyo y comprensión en los momentos más difíciles. Sin su guía y aliento, no habría llegado hasta aquí.

Gracias de todo corazón.



Índice de contenidos

RESUMEN	11
ABSTRACT	12
INTRODUCCIÓN	13
1.1 Problema de investigación.....	16
1.2 Justificación.....	18
1.3 Objetivos.....	19
1.3.1 Objetivo general.....	19
1.3.2 Objetivos específicos.....	19
1.4 Motivación	20
1.5 Estructura de la tesis.....	21
ANTECEDENTES	23
2.1 Aplicaciones similares	23
2.2 Aplicaciones en la investigación.....	26
MARCO TEÓRICO.....	31
3.1 Cáncer de seno.....	31
3.2 Definición de mamografía	32
3.3 Daños posibles de las mamografías de detección	34
3.4 Sistema de Base de Datos e Informes de Imágenes del Seno (BI-RADS)	
35	
3.5 Inteligencia artificial.....	38
3.6 Aprendizaje supervisado.....	39

- 3.7 Algoritmos de extracción de características de las imágenes 40
 - 3.7.1 La matriz de co-ocurrencia de nivel de gris (GLCM)..... 41
 - 3.7.2 Estadísticas de primer orden (FOF)..... 44
 - 3.7.3 Matriz de longitud de ejecución de nivel gris (GLRLM)..... 46
- 3.8 Aprendizaje automático 47
 - 3.8.1 Support Vector Machine (SVM) 48
 - 3.8.2 Árbol de decisión 49
 - 3.8.3 K-vecinos más cercanos..... 51
 - 3.8.4 Bosque aleatorio..... 52
 - 3.8.5 Clasificador de aumento de gradiente 53
- 3.9 Aprendizaje profundo 53
 - 3.9.1 Redes neuronales..... 54
 - 3.9.2 Redes neuronales convolucionales..... 57
 - 3.9.2.1 Capa convolucional 58
 - 3.9.2.2 Capa de rectificación (ReLU) 59
 - 3.9.2.3 Capa de agrupamiento 60
 - 3.9.2.4 Capa Dropout 62
 - 3.9.2.5 Capa de salida (Capa de clasificación)..... 62
 - 3.9.3 Conceptos importantes al utilizar CNN 65
 - 3.9.3.1 Sobreajuste (Overfitting)..... 65
 - 3.9.3.2 Aumento de datos (Data augmentation): 66
 - 3.9.3.3 Parada temprana (Early Stopping): 66
 - 3.9.3.4 Abandono (Dropout): 67
 - 3.9.3.5 Optimizadores 68
 - 3.9.3.6 Aprendizaje de transferencia..... 70
 - 3.9.4 Redes neuronales convolucionales en la medicina 72
- 3.10 Arquitecturas de redes neuronales convolucionales 75
 - 3.10.1 AlexNet..... 75

3.10.2	MobileNet	76
3.10.3	MobileNetV3 Small y Large	77
3.10.4	VGGNet.....	78
3.10.5	ResNet	80
3.10.6	Inception-V3	82
3.10.7	DenseNet	83
3.10.8	EfficientNet.....	84
3.11	Métricas de evaluación.	86
3.12	Bases de datos utilizadas	89
3.12.1	Mini-MIAS.....	89
3.12.2	CBIS-DDSM	91
3.13	Especificaciones de hardware	93
3.14	Especificaciones de software.....	93
3.14.1	Python	93
3.14.2	Tensorflow.....	94
3.14.3	Keras.....	95
3.15	Otras librerías utilizadas.....	96
3.15.1	Numpy.....	96
3.15.2	Pandas	97
3.15.3	Pillow	98
3.15.4	OpenCV (Open-Source Computer Vision Library).....	98
3.15.5	Scikit-learn.....	99
3.15.6	Matplotlib	100
3.15.7	Albumentations.....	100
3.15.8	Pydicom.....	101

METODOLOGÍA..... 102

4.1 Modalidad de la investigación 102

4.2 Recolección de Información 103

4.3 Procesamiento y análisis de los datos 104

4.4 Desarrollo del proyecto 104

4.5 Técnica de análisis de Imágenes 104

4.6 Metodología propuesta..... 105

4.6.1 Metodología 1..... 105

4.6.1.1 Preprocesamiento de imágenes..... 108

4.6.1.2 Arquitecturas DCNN..... 112

4.6.2 Metodología 2..... 113

RESULTADOS 116

5.1 División de conjunto de datos de prueba y de entrenamiento..... 117

5.2 Resultados Metodología 1..... 117

5.3 Clasificación usando algoritmos de aprendizaje automático..... 125

5.4 Resultados Metodología 2..... 132

CONCLUSIONES Y TRABAJO FUTURO 138

6.1 Conclusión 138

6.2 Trabajo futuro..... 140

6.3 Preguntas de investigación 141

BIBLIOGRAFÍA 144

ANEXO..... 159

8.1 Clasificación usando aprendizaje profundo en la primera metodología. 159

8.2 Clasificación usando aprendizaje profundo en la segunda metodología. 184

Índice de tablas

Tabla 1 Clasificación BI-RADS [10].....	37
Tabla 2 El modelo MobileNetV3-Small aplicado en cada base de datos y técnica.	119
Tabla 3 El modelo MobileNetV3-Large aplicado en cada base de datos y técnica.	119
Tabla 4 El modelo MobileNet aplicado en cada base de datos y técnica.	120
Tabla 5 El modelo ResNet50 aplicado en cada base de datos y técnica.	120
Tabla 6 El modelo ResNet152 aplicado en cada base de datos y técnica.	121
Tabla 7 El modelo Inception-V3 aplicado en cada base de datos y técnica.	122
Tabla 8 El modelo DenseNet121 aplicado en cada base de datos y técnica.	122
Tabla 9 El modelo VGGNet16 aplicado en cada base de datos y técnica.	123
Tabla 10 El modelo VGGNet19 aplicado en cada base de datos y técnica.	124
Tabla 11 El modelo AlexNet12 aplicado en cada base de datos y técnica.	125
Tabla 12 El modelo SVM aplicado en cada base de datos y técnica.	126
Tabla 13 El modelo MLP aplicado en cada base de datos y técnica.....	127
Tabla 14 El modelo KNN aplicado en cada base de datos y técnica.	128
Tabla 15 El modelo RF aplicado en cada base de datos y técnica.	129
Tabla 16 El modelo DT aplicado en cada base de datos y técnica.	130
Tabla 17 El modelo GBC aplicado en cada base de datos y técnica.	131
Tabla 18 Usando la técnica de aumento de datos en el conjunto de datos de entrenamiento de la etapa 1.	132
Tabla 19 Usando la técnica de aumento de datos en el conjunto de datos de entrenamiento de la etapa 2.	133
Tabla 20 Usando la técnica de aumento de datos en el conjunto de datos de entrenamiento de la etapa 3.	134
Tabla 21 Usando la técnica de aumento de datos en el conjunto de datos de entrenamiento de la etapa 4.	135

Tabla 22 Los valores obtenidos como resultado del entrenamiento se muestran en esta tabla. El final resultado de cada etapa, así como de cada DCNN, solo se observa con imágenes normalizadas (NO). 136

Tabla 23 Los valores obtenidos como resultado del entrenamiento se muestran en esta tabla. Cada resultado final de la etapa se observa junto con cada DCNN utilizando la técnica CLAHE. 137

Índice de ilustraciones

Ilustración 1 Mamografías. <https://www.actedi.cat/es/actualidad/radiographer-journal-club/rango-de-interpretacion-del-radiologo-en-un-programa-de-cribado-de-1-millon-de-mamografias-digitales-id-87.html>..... 32

Ilustración 2 Mamografía MLO. <https://noticias.medsbla.com/noticias-medicas/oncologia-medica-oncologia-radioterapica/la-mamografia-digital-incrementa-las-biopsias-innecesarias-en-estados-unidos/>..... 34

Ilustración 3 Clasificación Bi-RADS y sus categorías. <https://www.diagnosticorojas.com.ar/blog/salud/que-debo-saber-sobre-el-bi-rads/> 37

Ilustración 4 Referencia sobre el algoritmo de GLCM. https://www.researchgate.net/figure/Feature-extraction-methods-in-image-processing-GLCM-and-HOG-a-An-example-of_fig1_337940559..... 42

Ilustración 5 Representación gráfica del modelo SVM. https://es.wikipedia.org/wiki/M%C3%A1quinas_de_vectores_de_soporte..... 49

Ilustración 6 Representación gráfica de un árbol de decisiones. <https://seremprendedores.com/que-son-los-arboles-de-decisiones/> 50

Ilustración 7 Representación gráfica de KNN. <https://www.mql5.com/en/articles/11678> 51

Ilustración 8 Representación gráfica de bosque aleatorio. <https://medium.com/@hpumah/bosques-aleatorios-482163ace92e>..... 52

Ilustración 9 Diferencia entre aprendizaje automático y aprendizaje profundo 54

Ilustración 10 Modelo de una red neuronal artificial básica con 2 capas ocultas.
<https://www.atriainnovation.com/que-son-las-redes-neuronales-y-sus-funciones/> 55

Ilustración 11 Modelo de un perceptrón. <http://www.cs.us.es/~fsancho/?e=72> 57

Ilustración 12 Modelo estándar para una red neuronal convolucional.
<https://www.diegocalvo.es/red-neuronal-convolucional/> 58

Ilustración 13 Imágenes sacadas de la capa de convolución.
<https://datascience.aero/aviation-function-deep-learning/> 59

Ilustración 14 Función ReLU 60

Ilustración 15 Agrupamiento máximo (Maxpooling)..... 61

Ilustración 16 Agrupamiento promedio (Averagepooling)..... 61

Ilustración 17 función de dropout. 62

Ilustración 18 Función sigmoideal. 63

Ilustración 19 Función softmax..... 64

Ilustración 20 Descripción grafica de un modelo de transferencia de aprendizaje.
<https://www.bbvaiafactory.com/es/applying-transfer-learning-to-natural-language-models/> 72

Ilustración 21 Topología de la red neuronal convolucional: AlexNet.
https://www.researchgate.net/figure/Scheme-of-the-AlexNet-network-used_fig1_320052364..... 76

Ilustración 22 Topología de la red neuronal convolucional: MobileNet.
<https://wikidocs.net/165429> 77

Ilustración 23 Bloque de MobileNet V3. <https://medium.com/ai%C2%B3-theory-practice-business/nextgen-mobilenetv3-definitions-243e396a4c9a> 78

Ilustración 24 Topología de la red neuronal convolucional: VGGNet.
<https://ai.plainenglish.io/vggnet-with-tensorflow-transfer-learning-with-vgg16-included-7e5f6fa9479a> 79

Ilustración 25 Grafica de toda la estructura del modelo ResNet.
<https://towardsdatascience.com/review-resnet-winner-of-ilsvrc-2015-image-classification-localization-detection-e39402bfa5d8> 82

Ilustración 26 Topología de la red neuronal convolucional: Inception-V3.
<https://cloud.google.com/tpu/docs/inception-v3-advanced?hl=es-419> 83

Ilustración 27 Estructura de la formación de las capas de DenseNet.
https://pytorch.org/hub/pytorch_vision_densenet/ 84

Ilustración 28 Topología de la red neuronal convolucional: EfficientNet.
<https://ai.googleblog.com/2019/05/efficientnet-improving-accuracy-and.html> 85

Ilustración 29 Matriz de confusión estructura. 87

Ilustración 30 Matriz de confusión para problemas de varias categorías. 87

Ilustración 31 Grafica ROC con la métrica AUC. 89

Ilustración 32 Imágenes de Mini-MIAS recortadas. 91

Ilustración 33 Imágenes de CBIS-DDSM recortadas. 92

Ilustración 34 Logo Python (https://es.wikipedia.org/wiki/Historia_de_Python). 94

Ilustración 35 Tensorflow logo. 95

Ilustración 36 Keras logo. 96

Ilustración 37 Numpy logo. 97

Ilustración 38 Pandas logo. 98

Ilustración 39 OpenCV logo. 99

Ilustración 40 Scikit-learn logo. 99

Ilustración 41 Graficas Matplotlib. 100

Ilustración 42 Primera metodología utilizando aprendizaje automático y aprendizaje profundo. 107

Ilustración 43 Recorte de las mamografías. 108

Ilustración 44 Imágenes incompletas. 109

Ilustración 45 Técnicas de preprocesamiento para la metodología 1. 111

Ilustración 46 Aumento de datos metodología 1. 112

Ilustración 47 Segunda metodología utilizando aprendizaje profundo. 114

Ilustración 48 Aumento de datos para la metodología 2. 115

Ilustración 49 Preprocesamiento de imágenes para la metodología 2. 115

Ilustración 50 Topología de la etapa de clasificación usando redes neuronales artificiales. 118

Ilustración 51 Grafica mostrando la métrica de exactitud para la etapa de entrenamiento, comparando los modelos de aprendizaje automatico en una clasificación binaria utilizando el conjunto de datos CBIS-DDSM. 159

Ilustración 52 Grafica mostrando la métrica de precisión para la etapa de entrenamiento, comparando los modelos de aprendizaje automático en una clasificación binaria utilizando el conjunto de datos CBIS-DDSM. 160

Ilustración 53 Grafica mostrando la métrica de recall para la etapa de entrenamiento, comparando los modelos de aprendizaje automático en una clasificación binaria utilizando el conjunto de datos CBIS-DDSM. 161

Ilustración 54 Grafica mostrando la métrica de F1-score para la etapa de entrenamiento, comparando los modelos de aprendizaje automático en una clasificación binaria utilizando el conjunto de datos CBIS-DDSM. 162

Ilustración 55 Grafica mostrando el tiempo de entrenamiento para la etapa de entrenamiento, comparando los modelos de aprendizaje automático en una clasificación binaria utilizando el conjunto de datos CBIS-DDSM. 163

Ilustración 56 Grafica mostrando la métrica de exactitud para la etapa de entrenamiento, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS..... 164

Ilustración 57 Grafica mostrando la métrica de exactitud para la etapa de prueba, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS..... 165

Ilustración 58 Grafica mostrando la métrica de exactitud para la etapa de entrenamiento en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS y CBIS-DDSM. (Las cuatro etapas). 184

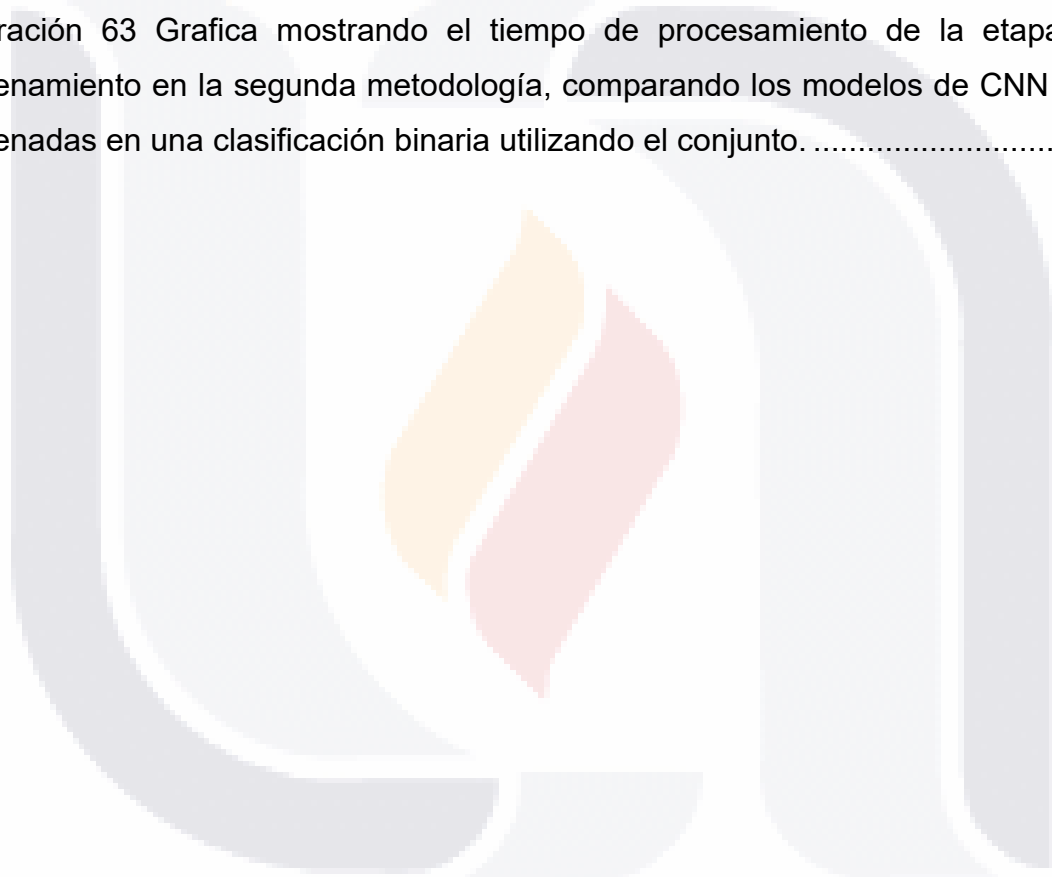
Ilustración 59 Grafica mostrando la métrica de exactitud para la etapa de prueba en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS y CBIS-DDSM. (Las cuatro etapas). 185

Ilustración 60 Grafica mostrando la métrica de AUC promedio para la etapa de entrenamiento en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS y CBIS-DDSM. (Las cuatro etapas). 186

Ilustración 61 Grafica mostrando la métrica de f1-score para la etapa de entrenamiento en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS y CBIS-DDSM. (Las cuatro etapas). 187

Ilustración 62 Grafica mostrando la métrica de recall para la etapa de entrenamiento en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS y CBIS-DDSM. (Las cuatro etapas). 188

Ilustración 63 Grafica mostrando el tiempo de procesamiento de la etapa de entrenamiento en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto. 189



RESUMEN

En México, el cáncer de mama es una de las principales causas de muerte entre las mujeres. Sus altas tasas de mortalidad pueden prevenirse con una detección temprana y efectiva. Las lesiones de cáncer de mama son áreas anormales en el tejido mamario, como microcalcificaciones, masas, distorsión arquitectónica y asimetría bilateral. Todos estos son ejemplos de lesiones de cáncer de mama. Actualmente, la mamografía es una fuente de información esencial que utilizan los especialistas para diagnosticar esta enfermedad. Sin embargo, identificar las características del cáncer de mama es un desafío debido a las dificultades de interpretar las imágenes de las mamografías. Para respaldar esta interpretación, en este documento se propone un enfoque de detección asistido por computadora. MobileNetV3 Small, MobileNetV3 Large, MobileNet, ResNet50, ResNet152, VGG16, VGG19, DenseNet121, InceptionV3 y AlexNet12 son algunos de los modelos de red neuronal convolucional profunda (DCNN) investigados. Se utilizan dos bases de datos importantes, como Mini-MIAS y CBIS-DDSM. Para extraer la región de interés (ROI) de las mamografías, la base de datos Mini-MIAS requiere un algoritmo de recorte. Después de eso, se utiliza el procesamiento previo de la imagen para mejorar la calidad de la imagen. No se requirió un algoritmo de recorte para la base de datos CBIS-DDSM porque las imágenes obtenidas ya estaban recortadas y redimensionadas, pero se requirió un procesamiento previo de mejora. Se utilizaron la ecualización de histograma adaptativa limitada de contraste (CLAHE), la ecualización de histograma (HE), el enmascaramiento de enfoque (UM) y el estiramiento de contraste para mejorar las fotos (CS). Este estudio empleó 334 imágenes de la base de datos Mini-MIAS y 2620 estudios de mamografía de película escaneada del CBIS-DDSM. Finalmente, se utilizaron cinco medidas para encontrar el rendimiento más alto en cada modelo: exactitud, precisión, recuperación, puntaje F1 y área bajo la curva ROC (AUC). Tanto en problemas (bi-clase y multi-clase) como en bases de datos (Mini-MIAS y CBIS-DDSM), nuestra metodología propuesta logra resultados de alto rendimiento.

ABSTRACT

In Mexico, breast cancer is one of the major causes of death among women. Its high mortality rates can be prevented with early and effective detection. Breast cancer lesions are abnormal areas in the breast tissue such as microcalcifications, masses, architectural distortion, and bilateral asymmetry. These are all examples of breast cancer lesions. A mammogram is currently an essential source of information used by specialists to diagnose this illness. However, identifying breast cancer features is challenging due to the difficulties of interpreting mammogram images. To support this interpretation, a computer-aided detection approach is proposed in this paper. MobileNetV3 Small, MobileNetV3 Large, MobileNet, ResNet50, ResNet152, VGG16, VGG19, DenseNet121, InceptionV3 and AlexNet12 are some of the Deep Convolutional Neural Network (DCNN) models investigated. Two important databases are used such as Mini-MIAS and CBIS-DDSM. To extract the region of interest (ROI) from mammograms, the database Mini-MIAS requires a cropping algorithm. After that, image pre-processing is used to improve image quality. A cropping algorithm wasn't required for the database CBIS-DDSM because the images obtained were already cropped and resized, but pre-processing of enhancement was required. Contrast Limited Adaptive Histogram Equalization (CLAHE), Histogram Equalization (HE), Unsharp Masking (UM), and Contrast Stretching were used to improve the photos (CS). This study employed 334 pictures from the Mini-MIAS database and 2,620 scanned film mammography studies from the CBIS-DDSM. Finally, five measures were utilized to find the highest performance on each model: accuracy, precision, recall, F1-score, and area under the ROC curve (AUC). In both problems (bi-class and multi-class) and databases (Mini-MIAS and CBIS-DDSM), our proposed methodology achieves high performance results.

CAPITULO 1.

El primer capítulo se expone el problema a tratar y se proporciona información fundamental acerca del tema. Se presenta una descripción detallada del problema de investigación junto con la motivación que impulsó el estudio, y se brindó una visión general de los objetivos que se buscaron alcanzar a través de la investigación. Estos objetivos fueron definidos de manera clara y coherente con el problema de investigación abordado.

INTRODUCCIÓN

Las mujeres latinoamericanas presentan tasas de incidencia y mortalidad por cáncer de mama más elevadas que las mujeres de países desarrollados. Para las mujeres mayores de 44 años, la incidencia de cáncer de mama es del 80%, mientras que la tasa de mortalidad es del 86% [1]. Las mujeres jóvenes también se ven afectadas por esta enfermedad, que puede afectar al 15% de las personas en países menos desarrollados como México [2]. En comparación con los países de ingresos altos, más del 45%, los países de ingresos bajos y medios tienen una tasa de mortalidad a incidencia que es considerablemente más alta, entre el 60% y el 75%. Para aquellos en grupos socioeconómicos específicos, este es un problema importante [3].

La detección temprana del cáncer de mama aumenta la posibilidad de que todos estos pacientes con cáncer puedan sobrevivir y lograr una mayor tasa de supervivencia. Por lo tanto, los radiólogos a menudo utilizan la mamografía de detección digital para el análisis, diagnóstico y categorización del cáncer de mama [4], [5].

Cuando el cáncer se detecta en una etapa temprana, a menudo responde mejor al tratamiento y es posible que aún no se haya propagado a otras partes del cuerpo. Esto significa que el tratamiento puede ser menos extenso y menos

invasivo, lo que puede generar menos complicaciones y una mejor calidad de vida para el paciente.

Por ejemplo, si el cáncer de mama se detecta y trata a tiempo, la tasa de supervivencia es mucho mayor que si se detecta en una etapa posterior. De manera similar, si el cáncer de piel se detecta temprano, a menudo se puede tratar con una cirugía simple y es posible que no requiera tratamiento adicional. Por el contrario, si el cáncer de piel no se detecta hasta que ha progresado a una etapa más avanzada, puede requerir una cirugía más extensa, radioterapia o quimioterapia, lo que puede tener un mayor impacto en la calidad de vida del paciente.

Además, la detección temprana a menudo también puede resultar en un menor costo de tratamiento. Cuando el cáncer se detecta temprano, a menudo es más tratable y menos costoso que cuando se detecta en una etapa posterior. Esto puede ser especialmente importante para las personas que tal vez no tengan acceso a costosos tratamientos contra el cáncer o que no puedan pagar el costo del tratamiento.

La mamografía es una técnica de rayos X de baja dosis que nos permite examinar la estructura interna del seno, la mejor prueba disponible para los médicos para la detección temprana del cáncer de seno es una mamografía de rutina. A pesar de que la mamografía es actualmente uno de los métodos más fiables para detectar el cáncer de mama [5]. Los expertos médicos tienen dificultades para interpretar las imágenes de las mamografías, lo que podría conducir a un diagnóstico incorrecto.

Se investigó los métodos actuales de clasificación de imágenes, y los que utilizan inteligencia artificial fueron los que más se destacaron en la actualidad [6], [7].

En una variedad de sectores, la inteligencia artificial (IA) ha avanzado rápidamente en los últimos años. Dado que las imágenes son una de las fuentes de información más cruciales para las actividades que involucran la inteligencia humana, IA se ha involucrado profundamente en el procesamiento de imágenes en

TESIS TESIS TESIS TESIS TESIS

innumerables ocasiones. Las técnicas de IA fueron beneficiosas y fueron una excelente opción para esta investigación debido a esto [8], [9].

Los avances en IA en los últimos años ayudan en la creación de nuevas tecnologías y esto mismo hace que estas aplicaciones puedan ramificarse a varios campos de la ciencia. Estos algoritmos dan soluciones a todo tipo de problemas subjetivos donde solamente el ser humano podría resolver. (i. e. Como se mencionaba en las clasificaciones de imágenes, el poder detectar anomalías en las imágenes, etc.)

Se recomendó el diagnóstico asistido por computadora (CAD, por sus siglas en inglés) como una alternativa al procedimiento de lectura doble inexacto y que consume mucho tiempo de los radiólogos [10], [11]. Algunos de estos CAD utilizan algoritmos de aprendizaje automático que toman información de patrones complejos. Estos algoritmos se basan en clasificadores tradicionales, que utilizan funciones hechas a mano para resolver tareas de aprendizaje automático. Como resultado, especialmente para las tareas que involucran la extracción y selección de características, estas técnicas se consideran complejas y requieren mucho tiempo [11], [12].

Según estudios recientes, se ha demostrado que los algoritmos de aprendizaje profundo son eficientes en tareas que incluyen la segmentación, detección y clasificación de imágenes en una variedad de disciplinas de procesamiento de imágenes y visión artificial [13]–[15].

El aprendizaje profundo es un tipo de aprendizaje automático inspirado en la estructura y función del cerebro [16]. El aprendizaje profundo ha recibido últimamente mucha atención por problemas de clasificación, particularmente en el campo del análisis de imágenes médicas[6], [17]–[19].

Por lo tanto, se analizaron los métodos actuales de clasificación de imágenes. Las DCNN se desempeñaron bien en la tarea de clasificación de imágenes según el estado del arte. Este modelo requiere una gran cantidad de imágenes, razón por la cual buscamos conjuntos de datos conocidos y de

vanguardia. Dos conjuntos de datos se destacaron del resto: CBIS-DDSM (el subconjunto de imágenes curadas de mama de DDSM) [20] y Mini-MIAS (la sociedad de análisis de imágenes mamográficas) [21].

Se utilizó el aprendizaje por transferencia con ajuste fino para llevar a cabo esta investigación [22], [23]. El aprendizaje por transferencia ocurre cuando un modelo creado para una tarea se usa para otra, y el ajuste fino ocurre cuando la salida del modelo se modifica para cumplir con la nueva tarea [24], [25].

Este trabajo presenta las siguientes contribuciones:

- Para esta investigación, se compararon varios modelos preentrenados tales como: MobileNetV3 Small [26], MobileNetV3 Large [26], MobileNet [27], ResNet50 [28], ResNet152 [28], VGG16 [29], VGG19 [29], DenseNet121 [30], InceptionV3 [31], AlexNet12 [32] y EfficientB7 [33].
- Se mejoró la calidad de cada imagen recortada. Aplicamos CLAHE (Contrast Limited Adaptive Histogram Equalization), enmascaramiento de enfoque, ecualizador de histograma y estiramiento de contraste [34], [35].
- Se usó un algoritmo de aumento de datos de la biblioteca de albumentation para aumentar la cantidad de imágenes de entrenamiento y hacer que la CNN sea más sólida [36].
- Y al final se comparó el rendimiento de cada modelo con cinco métricas (exactitud, precisión, recall, f1-score y AUC).

1.1 Problema de investigación.

Antes de poder hablar del cáncer de mama es conveniente saber que es el cáncer en primer lugar. El cáncer es un resultado de mutaciones, o algún cambio anómalo en los genes, que regulan el crecimiento de las células y las mantienen sanas. Los genes se hallan en el núcleo de las células. Por lo general, las células del cuerpo se renuevan mediante un proceso específico, llamado crecimiento celular, donde las células nuevas y sanas ocupan el lugar de las células viejas que mueren. El cuerpo

está formado por billones de células que crecen y se dividen según sea necesario, pero cuando estas células se vuelven anormales o envejecen, por lo general mueren. El cáncer comienza cuando algo sale mal en este proceso, lo que hace que las células sigan dividiéndose y no mueran cuando deberían. Como resultado, las células cancerosas crecen sin control y pueden desplazar a las células normales, lo que dificulta que el cuerpo funcione correctamente [37].

El cáncer de mama es un problema relevante de salud pública. Varios países han podido lograr una tendencia decreciente, pero en otros países la tendencia continúa de una manera ascendente.

El cáncer de mama es una enfermedad complicada que es la principal causa de muerte por cáncer en mujeres en todo el mundo, con un estimado de 500.000 muertes cada año. El 70% de estas muertes ocurren en países en desarrollo, donde las mujeres tienen menos acceso a los servicios de salud, lo que lleva a una detección temprana, tratamiento y control del cáncer deficientes. A pesar del mayor riesgo de enfermarse entre las mujeres de los países socioeconómicos altos, el riesgo de morir es mayor para las mujeres de los países pobres. Las mujeres de 25 a 45 años son generalmente las más afectadas y el cáncer de mama es la principal causa de muerte en las mujeres [38].

En México, se estimó que en 2020 hubo una incidencia total de cáncer de 9,2 millones de casos y una mortalidad de 4,4 millones de muertes. El cáncer de mama alcanzó el 24,5% de la incidencia y el 15,5% de las muertes totales [39].

Muchas de estas muertes podrían haber sido evitadas con el hecho de una detección temprana del problema. Esto es porque en caso de que fuera un tumor que fuera cancerígeno este se dividiría en cuatro etapas. La primera es que el tumor se ha propagado a otros tejidos en un área que se considera pequeña. En la segunda etapa el tumor mide 220 a 50 milímetros y algunos o ningún ganglio linfático está afectado. La tercera etapa ya tiene una propagación regional, incluso este podría ya ser visible y puede haberse propagado a la piel o la pared torácica. En la última etapa también llamada metástasis el cáncer se ha propagado más allá del área de origen y se puede encontrar en otras partes del cuerpo.

Existen dificultades que se presentan al detectar y clasificar lesiones en las imágenes mamográficas lo que provoca un mal diagnóstico del cáncer de mama.

Esto es un gran problema ya que entre más tiempo tarde, más es probable que esta enfermedad se pueda propagar impidiendo la recuperación de la persona afectada. Por esto mismo se necesita utilizar alguna técnica para el reconocimiento de imágenes que pueda ser de utilidad y puede tener obtener resultados los más certeros posible con la poca cantidad de imágenes que se tienen.

Para esto se probó los métodos convencionales de inteligencia artificial y métodos más modernos utilizados en arquitecturas más complejas de redes neuronales convolucionales. Estos métodos más modernos son más complejos y costosos que los métodos convencionales, pero han demostrado que entre más datos de entrada obtengan pueden lograr mejores resultados que los mismos.

En esta investigación no se utilizó algún tipo de GPU o TPU externo como los que se pueden obtener mediante la suscripción de Google Colab. Pero, se utilizó una computadora lo suficientemente potente para el entrenamiento de estos problemas de clasificación. Como podría ser imágenes benignas contra imágenes malignas.

1.2 Justificación.

En México, esta enfermedad es una de las principales causas de muerte en mujeres en 2020. Hubo 8000 muertes por tumores malignos de mama [40]. El cáncer de mama temprano generalmente no causa ningún dolor en sus primeras etapas, cuando es más fácil de tratar, que es la razón principal por la que es importante tener una detección temprana.

Este proyecto reforzará el área de investigación e innovación científica sobre el análisis automático de imágenes mamográficas para así reconocer y poder identificar microcalcificaciones y poder mediante el algoritmo desarrollado ayudar con la detección y diagnóstico temprano.

Este trabajo de investigación intenta poder contribuir en con una detección temprana del cáncer de mama, donde se buscarán las mejores técnicas y se

desarrollarán los mejores algoritmos capaces de poder hacerlo de manera rápida y eficiente así ayudando a especialista radiólogo que esté encargado para que puede realizar una detección eficaz y confiable de microcalcificaciones.

Se pretende al momento de realizar este proyecto poder contribuir en el desarrollo de algoritmos que puedan ayudar a los especialistas, a detectar eficazmente las lesiones usando una o varias técnicas de análisis de imágenes en mamografías, y los patrones que sean representativos que podrían ser microcalcificaciones.

1.3 Objetivos.

En esta sección se plantea el objetivo general, así como los objetivos particulares a alcanzar a lo largo del desarrollo de esta investigación.

1.3.1 Objetivo general

- Desarrollar un algoritmo de detección automática de lesiones en mamografías mediante el uso de inteligencia artificial. El algoritmo debe ser capaz de identificar y localizar posibles anomalías en las imágenes mamográficas con alta precisión y eficiencia. De esta forma, se busca mejorar el diagnóstico temprano del cáncer de mama, lo que puede ser fundamental para el tratamiento y la recuperación de las pacientes.

1.3.2 Objetivos específicos

- Definir una combinación adecuada de algoritmos que permita diseñar la metodología adecuada para la detección de lesiones en las mamografías. Esta combinación de algoritmos deberá ser capaz de realizar tareas de extracción de características y clasificación de imágenes médicas.

- Seleccionar los algoritmos que permitan identificar las características relevantes a detectar en las imágenes mamográficas.
- Implementar los algoritmos seleccionados y probar su desempeño. Es importante realizar pruebas con diferentes parámetros y configuraciones para evaluar su rendimiento y determinar cuáles son los más efectivos.
- Se debe seleccionar o crear una base de datos de imágenes médicas que incluya una variedad de imágenes mamográficas que permitan probar el desempeño de los algoritmos implementados. La base de datos debe incluir imágenes con y sin lesiones para entrenar y validar el algoritmo.
- Integrar todos los módulos del sistema para su validación correspondiente. Esto implica que los algoritmos desarrollados deben funcionar juntos para realizar la detección automática de lesiones en mamografías. Se deben realizar pruebas exhaustivas y ajustes necesarios para garantizar la precisión y eficiencia del sistema.

1.4 Motivación

La creciente importancia de las redes neuronales en el mundo actual, especialmente en el cuidado de la salud, se debe a la cantidad sin precedentes de datos disponibles. Las redes neuronales pueden extraer características y patrones que ayudan en una variedad de tareas, como la clasificación, que puede volverse compleja para los humanos. En el campo de la salud, las redes neuronales son valiosas para detectar enfermedades con precisión y prevenir problemas a largo plazo. El campo de la medicina también carece de expertos que puedan hacer diagnósticos confiables, por lo que las redes neuronales son una herramienta útil en el diagnóstico de patologías que requieren personas altamente especializadas.

1.5 Estructura de la tesis

En el capítulo 1 se presentó el problema que se abordó en la tesis y se brindó información básica sobre el tema. Se describió el problema de investigación y la motivación detrás de la investigación, y se proporcionó una descripción general de los objetivos de la investigación. Los objetivos fueron claramente definidos y alineados con el problema de investigación.

En el capítulo 2 se presentó una revisión exhaustiva de la literatura. La revisión proporcionó una visión general de la investigación previa que se había llevado a cabo en el área de estudio y proporcionó un análisis crítico de la literatura existente. Este capítulo resaltó las lagunas en la literatura que el estudio pretendía llenar y proporcionó una base para el marco teórico del estudio. El capítulo también proporcionó un resumen del estado del arte en el campo y describió el conocimiento actual sobre el tema. Se buscó una explicación sobre cada modelo utilizado para la clasificación de imágenes como las Redes neuronales convolucionales o en sus siglas en inglés CNN, y modelos más clásicos como SVM, árbol de decisión, bosque aleatorio, etc. Esta lectura del estado del arte fue para entender cada uno de los algoritmos y sus usos en el ámbito de la medicina, en especial los que estaban enfocados en la clasificación de mamografías, que era el tema de investigación.

En el capítulo 3 se proporcionaron los fundamentos teóricos para el estudio. El marco teórico se basó en la revisión de la literatura y proporcionó una explicación lógica de los objetivos del estudio y las preguntas de investigación. El capítulo proporcionó una descripción general de las teorías, modelos y conceptos que se utilizaron para analizar los datos de la investigación y desarrollar las hipótesis de la investigación. El capítulo también proporcionó un resumen del diseño de la investigación y los métodos que se utilizaron para recopilar y analizar los datos.

En el capítulo 4 se proporcionó una descripción detallada de la metodología para la clasificación de las mamografías. Se proporcionó una explicación paso a paso de las técnicas que se utilizaron para analizar imágenes de mamografía y clasificarlas en categorías normales o anormales. El capítulo incluyó una descripción de las técnicas de análisis de imágenes que se utilizaron, como el

análisis de texturas, la segmentación, la extracción de características, el aprendizaje automático y el aprendizaje profundo. También se proporcionó una descripción de los algoritmos y modelos específicos que se utilizaron para la clasificación de mamografías, incluido el tipo de red neuronal profunda y el procedimiento de entrenamiento.

En el capítulo 5 se presentaron los resultados del estudio de clasificación de la mamografía. Los resultados se organizaron y presentaron de manera clara y concisa, y este capítulo estuvo acompañado de tablas, figuras y gráficos apropiados. El capítulo proporcionó una discusión sobre el rendimiento de las metodologías de clasificación de mamografías y la medición de las métricas que se utilizaron, incluida la exactitud, la precisión, la exhaustividad y el área bajo la curva característica operativa del receptor (ROC).

En el capítulo 6 se presentarán las conclusiones y los trabajos futuros de este estudio que se enfoca en la detección del cáncer de mama utilizando modelos de redes neuronales convolucionales. El cáncer de mama es la forma más común de cáncer en mujeres en todo el mundo. Para este trabajo, se utilizó una metodología de preprocesamiento de imágenes mamográficas que incluye la eliminación de manchas negras y la recopilación de imágenes de interés (ROI) mediante un algoritmo. Además, se compararon once redes neuronales convolucionales preentrenadas y se ajustó su clasificación utilizando técnicas de ajuste fino en las capas ocultas y la capa de salida.

TESIS TESIS TESIS TESIS TESIS

CAPITULO 2

El segundo capítulo de la tesis es una revisión exhaustiva de la literatura sobre el tema de investigación, con un enfoque en la clasificación de mamografías. Se discuten diferentes modelos utilizados para la clasificación de imágenes, incluyendo redes neuronales convolucionales y modelos más clásicos. Esta revisión permite entender los algoritmos y sus usos en el ámbito médico, proporcionando una base sólida para el marco teórico del estudio.

ANTECEDENTES

2.1 Aplicaciones similares

Este artículo siguiente escrito por Mohiyuddin et al. [39] presenta un sistema de detección de tipo de tejido mamario utilizando un conjunto de datos de mamografía de MIAS. El sistema incluye etapas de preprocesamiento y cálculo de características de textura utilizando matrices de co-ocurrencia de nivel de gris. Se utilizó el clasificador de análisis discriminante lineal de Fisher en cuatro casos de clasificación diferentes y se logró una precisión de clasificación máxima del 82,48% con características en cascada. El sistema identificó que los tipos de tejido glandular graso se confunden fácilmente con los tipos de tejido glandular graso y denso.

Preetha et al. [40] se centran en el uso de imágenes de mamografía preprocesadas para mejorar el diagnóstico del cáncer de mama al reducir las predicciones de falsos positivos. El estudio extrae varias características de las imágenes, incluyendo la forma, la intensidad, la matriz de coocurrencia del nivel de gris y la matriz de longitud de ejecución del nivel de gris, que se utilizan como entradas para el clasificador. Los resultados indican que la combinación de las funciones GLCM y GLRLM es la más eficiente y logra la máxima tasa de precisión de clasificación en comparación con otras funciones. Se destaca la importancia del diagnóstico temprano del cáncer de mama, se explica brevemente los métodos de

preprocesamiento y las técnicas de extracción de características, y se presentan claramente los resultados y la identificación de las características más eficientes.

Mohanty et al. [41] propone un sistema de clasificación asistido por computadora para la detección del cáncer de mama a partir de mamografías digitales, que utiliza un enfoque de tres pasos. El estudio encuentra que el uso de un conjunto de 19 características extraídas de la matriz de longitud de ejecución de nivel de gris y la matriz de coocurrencia de nivel de gris podría distinguir masas malignas de masas benignas con una precisión del 94,9%. El resumen destaca el potencial del sistema propuesto para la detección del cáncer a partir de mamografías digitales, aunque podría proporcionar más información sobre el conjunto de datos y el rendimiento en comparación con otros estudios similares.

Dai et al. [42] utiliza el algoritmo de bosque aleatorio para el diagnóstico del cáncer de mama, destacando los beneficios del aprendizaje automático en el cuidado de la salud y la capacidad del algoritmo para producir resultados precisos mediante el aprendizaje conjunto de múltiples características. El estudio informa una alta precisión de predicción para el diagnóstico del cáncer de mama utilizando este método.

Kolchev et al. [43] comparan el rendimiento de dos modelos de procesamiento de imágenes, el modelo YOLOv4 y el modelo NCA, en el análisis de imágenes de mamografía. El estudio encuentra que el modelo NCA supera al modelo YOLOv4 en términos de tasas de verdaderos positivos y falsos positivos, tasas de falsos negativos y sensibilidad para ciertos tipos de lesiones mamarias. Además, las decisiones de los radiólogos también se vieron influenciadas por el modelo NCA, pero no por el modelo YOLOv4. En general, el resumen proporciona una descripción clara y concisa de la metodología, los resultados y las conclusiones del estudio.

Li et al. [44] muestran un estudio sobre la segmentación automática de imágenes para la detección de masas mamarias. El método propuesto utiliza aprendizaje profundo y combina una U-Net densamente conectada con puertas de atención. El estudio evalúa la efectividad del método utilizando varias métricas y la

base de datos DDSM para las pruebas. Los resultados experimentales muestran que el método propuesto logra mejores resultados de segmentación que otros métodos del estado del arte. El resumen proporciona una visión general clara del estudio y parece ser valioso para aplicaciones clínicas. Rajalakshmi et al. [45] presentan un modelo de aprendizaje profundo, el "modelo U-Net profundamente supervisado", para la segmentación de masas en mamografías que ayuda en el diagnóstico del cáncer de mama. El modelo utiliza supervisión profunda y campos aleatorios condicionales densos para abordar los desafíos de baja relación señal-ruido y límites de masa indiscernibles. Se evalúa en dos conjuntos de datos de referencia y supera a otros métodos en términos de puntuación de datos. El resumen proporciona información adecuada sobre el modelo, los desafíos y los resultados, pero se podrían haber incluido más detalles sobre la arquitectura del modelo y las características del conjunto de datos.

Patgiri et al. [46] presentan un artículo que cuestiona la idea de que los algoritmos de aprendizaje automático puedan lograr una precisión del 99% en el conjunto de datos de diagnóstico de cáncer de mama de Wisconsin. Los autores argumentan que alcanzar una precisión del 100% no es ético y que han aprovechado las debilidades de los algoritmos de aprendizaje automático para demostrar que las afirmaciones de alta precisión no son necesariamente correctas. El documento informa sobre tres resultados clave de los experimentos, que son la corrección de las precisiones, la importancia de la precisión mínima y la corrección de los algoritmos de aprendizaje automático. Sin embargo, el resumen no proporciona detalles sobre los métodos utilizados para aprovechar las debilidades de los algoritmos de aprendizaje automático o sobre los experimentos realizados para validar la afirmación. Además, no queda claro lo que los autores quieren decir con "corrección de la precisión" y "corrección de los algoritmos de aprendizaje automático". En general, el resumen carece de suficiente detalle y claridad para evaluar la metodología y los resultados del estudio.

2.2 Aplicaciones en la investigación

Los modelos DCNN han ganado popularidad recientemente como resultado de su alto rendimiento. Se han propuesto varios modelos a lo largo de los años, y algunos de ellos se han utilizado. Algunos investigadores han obtenido resultados diferentes, y analizarlos es fundamental para encontrar áreas en las que se pueden mejorar.

El siguiente artículo presenta un método para la segmentación automática del músculo pectoral de la mama en mamografías oblicuas medio-laterales utilizando una red neuronal convolucional (CNN) inspirada en la red de detección de bordes holísticamente anidada (HED). La CNN puede aprender características jerárquicas complejas para resolver la ambigüedad espacial al estimar el límite del músculo pectoral y puede encontrar objetos "similares a contornos" en las mamografías. MIAS, INBreast, BCDR y CBIS-DDSM se encontraban entre los conjuntos de datos utilizados en este estudio [47].

Se empleó un enfoque de conjunto, en el que se usó la función de Gompertz para construir clasificaciones difusas de las técnicas de clasificación base, y las puntuaciones de decisión de los modelos base se combinaron de forma adaptativa para construir predicciones finales. Usando Inception V4, ResNet 16, VGG 11 y DenseNet 121, así como otros modelos de CNN profunda [48].

Se utilizó un enfoque de aprendizaje profundo mediante redes neuronales convolucionales (CNN) para clasificar las imágenes histopatológicas de cáncer de mama del conjunto de datos de BreakHis. Se propuso un método basado en la extracción de parches de imagen para el entrenamiento de la CNN y la combinación de estos parches para la clasificación final que permita el uso de imágenes histopatológicas de alta resolución como entrada a las CNN existentes sin requerir adaptaciones que darían como resultado una imagen más compleja y compleja. arquitectura computacionalmente costosa[49].

Adicionalmente, se entrenó y validó una red utilizando una base de datos de imágenes que contienen microcalcificaciones clasificadas en benignas y malignas, las cuales fueron creadas a partir de imágenes mamográficas de pacientes peruanos [50].

Un estudio reciente introdujo el "aprendizaje de transferencia de doble disparo", que es un método revolucionario basado en la idea del aprendizaje de transferencia. Esta estrategia mejoró significativamente la precisión de la categorización [51].

Otro estudio reciente propone el uso de una red neuronal convolucional (CNN) para detectar el cáncer de mama en imágenes de mamografía clasificándolas como anomalías no cancerosas o cancerosas utilizando el conjunto de datos DDSM. Se preprocesa un conjunto de imágenes de mamografía mediante la ecualización de histogramas y las imágenes resultantes se utilizan como fuente de entrenamiento para la CNN. El sistema propuesto, llamado BCDCNN, se compara con el sistema MCCANN y los resultados muestran que BCDCNN tiene una mayor precisión de clasificación y una mayor resolución en comparación con otros sistemas existentes [52].

Esta investigación utiliza aprendizaje profundo y redes neuronales para clasificar la detección de senos normales y anormales en imágenes de mamografía utilizando el conjunto de datos Mammogramas-MIAS, que contiene 322 mamografías con 189 imágenes de senos normales y 133 imágenes de senos anormales. El estudio utilizó una red neuronal convolucional (CNN) y obtuvo resultados experimentales prometedores que sugieren la eficacia del aprendizaje profundo para la detección de cáncer de mama en imágenes de mamografía. La investigación está en curso y se están realizando más desarrollos para optimizar la arquitectura de la CNN y emplear redes preentrenadas para mejorar la precisión [53].

El siguiente estudio utiliza redes neuronales convolucionales (CNN) previamente entrenadas para detectar y clasificar tumores de mama en el conjunto de datos de INbreast mediante imágenes de mamografía. El modelo propuesto preprocesa las imágenes para mejorar la calidad de imagen y reducir el tiempo de cálculo, y luego transfiere los parámetros aprendidos de las CNN para mejorar los resultados de la clasificación. El rendimiento del modelo propuesto se evalúa utilizando cuatro métricas: precisión, sensibilidad, especificidad y el área bajo la

TESIS TESIS TESIS TESIS TESIS

curva característica operativa del receptor (AUC). El modelo obtuvo una precisión de 97,1%, sensibilidad de 96,3%, especificidad de 97,9% y AUC de 0,988% [54].

El uso de modelos previamente entrenados es bueno para descubrir nuevos modelos para tareas específicas; Se presenta un nuevo método que utiliza el modelo VGG-16 ajustado llamado BreasNet-18. Usando el conjunto de datos CBIS-DDSM para esta investigación, se aplicó un preprocesamiento a estas imágenes. La eliminación de artefactos era el nombre de la etapa inicial del preprocesamiento. Hizo uso de técnicas como el enmascaramiento binario, la apertura morfológica y la detección del contorno más grande. "Eliminar línea" es el nombre de la segunda etapa. Se usaron las siguientes técnicas en ciertas imágenes que incluían una línea recta brillante adherida al contorno del seno: operación en rango, filtro Gabor, operaciones morfológicas y máscara de inversión. Algunos algoritmos, como la corrección gamma, CLAHE y Green Fire Blue, mejoran las imágenes en la tercera etapa. Después del preprocesamiento, la técnica de aumento de datos se utiliza para adquirir imágenes adicionales y resolver problemas de ajuste excesivo o insuficiente. Finalmente, se evaluó el problema de clasificación de cuatro clases utilizando el modelo ajustado de BreastNet 18 [55].

Además, se recomienda un método de diagnóstico asistido por computadora (CAD) para clasificar 3 clases para la metodología usando ciertos métodos tradicionales (cáncer, sin cáncer y no canceroso). Las redes neuronales convolucionales (CNN), las máquinas de vectores de soporte (SVM) y los bosques aleatorios fueron los tres clasificadores que emplearon (RF). Utilizaron un conjunto de datos con un total de 1000 imágenes del sitio web de Kaggle. Tras la prueba de cada clasificador, se descubrió que CNN se desempeñó mejor, logrando una precisión del 99,6 % por ciento [56].

En la implementación se utilizaron MVGG, un modelo VGG preentrenado, una red residual y una red móvil. Después de comparar el rendimiento de varios modelos, los autores descubrieron que su modelo VGG16 ajustado con aumento de datos y pesos de ImageNet tenía el mejor rendimiento. Se utilizó el conjunto de

TESIS TESIS TESIS TESIS TESIS

datos DDSM y esta metodología arrojó resultados de precisión del 88,30 % y un valor de AUC del 93,30 % [57].

Esta metodología de aprendizaje profundo de conjunto presenta otro método para detectar con precisión tejido canceroso y no canceroso utilizando CNN preentrenada. Los diseños VGG16 y VGG19 se utilizaron tal cual y con modificaciones. En la división de patología de la Universidad Colsanitas Colombia en Bogotá, Colombia, se recolectaron 544 imágenes de diapositivas completas (WSI) de 80 pacientes con cáncer de mama. Después de normalizar las imágenes, dividieron los datos con un 80 % de entrenamiento y un 20 % de prueba. En el conjunto de datos de entrenamiento, pudieron usar una estrategia de aumento de datos para alcanzar una precisión y F1 del 95,29 por ciento [58].

El siguiente artículo describe un algoritmo de aprendizaje profundo para detectar el cáncer de mama en las mamografías. El algoritmo es único en el sentido de que utiliza un enfoque de entrenamiento de "extremo a extremo", que minimiza la dependencia de las anotaciones de lesiones y permite el uso de etiquetas de nivel de imagen. El algoritmo se entrena en dos conjuntos de datos independientes, la base de datos digital para mamografía de detección (CBIS-DDSM) y la base de datos INbreast, y los resultados muestran un rendimiento excelente con alta precisión en ambas plataformas de mamografía heterogénea. Los hallazgos son prometedores para mejorar las herramientas clínicas y reducir los resultados de detección falsos positivos y falsos negativos [59].

Las tres etapas de la metodología Multi-View Feature Fusion (MVFF) son las siguientes: en el primer paso, la mamografía se clasifica binariamente como anormal o normal, y en el segundo, se clasifican la masa y la calcificación. El paso final es clasificar la condición como maligna o benigna. Utilizaron el conjunto de datos CBIS-DDSM para esta investigación. Los valores de AUC para masa y categorización fueron 93,20 % y 0,84 para tumores malignos y benignos, respectivamente [60].

Se propone un nuevo sistema de detección asistida por computadora (CAD) para clasificar tumores masivos benignos y malignos en imágenes de mamografía mamaria utilizando técnicas de segmentación y aprendizaje profundo. El sistema

TESIS TESIS TESIS TESIS TESIS

CAD utiliza dos enfoques de segmentación, uno que implica la determinación manual de la región de interés (ROI) y el otro que utiliza técnicas basadas en umbrales y regiones. Se utiliza una red neuronal convolucional profunda (DCNN) llamada AlexNet para la extracción de características y se ajusta para clasificar dos clases [61].

Sin embargo, AlexNet no siempre puede ser el mejor modelo para usar. Los modelos de VGGNet que se han ajustado pueden funcionar mejor si las clasificaciones de masas y calcificaciones de la mamografía se realizan mediante el aprendizaje de transferencia [62].

Un marco completamente automatizado para la detección de masas en mamografías digitales de campo completo está preestablecido por Este modelo se basa en la Red neuronal convolucional basada en regiones más rápidas (Faster-RCNN) en esta investigación. En este documento, se utilizaron los conjuntos de datos OMI-DB e INbreast [63].

La clasificación de anomalías benignas y malignas en imágenes de mamografía es muy fácil con el modelo de red neuronal convolucional (CNN). Una etapa de preprocesamiento fuerte es esencial para un mejor desempeño, como ya se mencionó. Para Mini-MIAS y BCDR, use recortes, aumentos y equilibrio para lograr una precisión de 68 y 52, respectivamente [64].

TESIS TESIS TESIS TESIS TESIS

CAPITULO 3

El tercer capítulo de la tesis proporciona los fundamentos teóricos para el estudio, incluyendo una descripción de las teorías y conceptos utilizados, así como el diseño de la investigación y los métodos de recopilación y análisis de datos. El objetivo es establecer una base sólida para la metodología empleada y comprender los resultados presentados en los capítulos posteriores.

MARCO TEÓRICO

3.1 Cáncer de seno

El siguiente artículo muestra una comprensión actual y el progreso en la prevención del cáncer de mama [65]. Los autores señalan que el cáncer de mama es la segunda causa principal de muerte por cáncer entre las mujeres, pero el diagnóstico y la prevención tempranos pueden mejorar la tasa de supervivencia relativa a 5 años. En la última década, ha habido un progreso significativo en la comprensión del cáncer de mama, incluido el descubrimiento de genes relacionados y células madre del cáncer de mama, así como en el desarrollo de métodos preventivos como la quimioterapia y la prevención biológica [65].

El cáncer de mama es una enfermedad en la que se forman células malignas (cancerosas) en los tejidos de la mama. Es el tipo de cáncer más común en las mujeres de todo el mundo y el segundo cáncer más común en general. Los principales tipos de cáncer de mama son el carcinoma ductal, que comienza en el revestimiento de los conductos lácteos, y el carcinoma lobulillar, que comienza en los lóbulos de la mama.

Los síntomas del cáncer de mama pueden incluir un bulto o engrosamiento en la mama o la axila, cambios en el tamaño o la forma de la mama, cambios en la piel, como hoyuelos o enrojecimiento, y secreción del pezón. Sin embargo, muchos

cánceres de mama no causan ningún síntoma y se detectan mediante exámenes de rutina, como las mamografías [66].

Las opciones de tratamiento para el cáncer de mama incluyen cirugía, radioterapia, quimioterapia, terapia dirigida y terapia hormonal. La elección del tratamiento dependerá de la etapa y el tipo de cáncer, así como de la salud general y las preferencias del paciente.

La prevención y la detección temprana son clave en la lucha contra el cáncer de mama, y se alienta a todas las mujeres a realizarse autoexámenes, realizarse exámenes clínicos de mama con regularidad y someterse a mamografías en los intervalos apropiados para su edad y factores de riesgo.

3.2 Definición de mamografía

La mamografía es una técnica utilizada para producir imágenes de la estructura interna de los senos. Utiliza rayos X de baja dosis para crear imágenes detalladas del tejido mamario [5], que pueden usarse para detectar y diagnosticar el cáncer de mama.

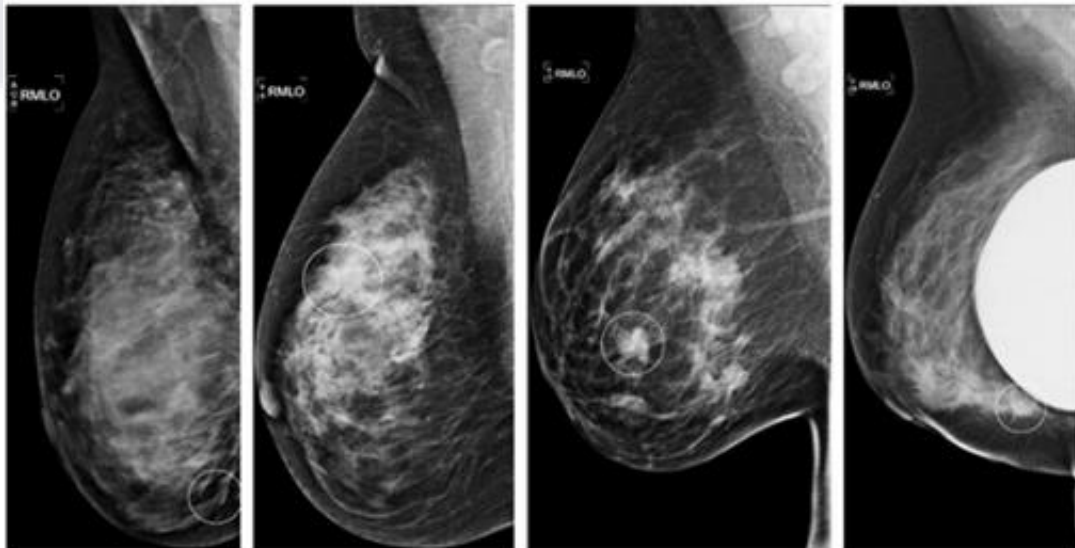


Ilustración 1 Mamografías. <https://www.actedi.cat/es/actualidad/radiographer-journal-club/rango-de-interpretacion-del-radiologo-en-un-programa-de-cribado-de-1-millon-de-mamografias-digitales-id-87.html>

La prueba generalmente se realiza en mujeres mayores de 50 años, pero también se puede recomendar para mujeres que tienen un alto riesgo de desarrollar cáncer de mama, como aquellas con antecedentes familiares de la enfermedad o ciertas mutaciones genéticas. Ilustración 1

Durante una mamografía, el seno de la mujer se coloca entre dos placas, que comprimen el seno para extender el tejido y facilitar la detección de anomalías. Luego, la máquina de rayos X toma varias imágenes del seno, que se utilizan para crear una vista detallada del tejido.

Hay dos tipos de mamografía: mamografía digital y mamografía de pantalla de película. En la mamografía digital, las imágenes de rayos X se capturan electrónicamente y se pueden ver en la pantalla de una computadora, lo que permite un análisis más preciso. La mamografía con pantalla de película, el tipo tradicional de mamografía, captura las imágenes de rayos X en una película y luego las lee un radiólogo. Las mamografías pueden detectar cambios en el tejido mamario que pueden no ser visibles durante un examen físico, como un pequeño bulto o calcificación. También se usan para controlar los cambios en el tejido mamario a lo largo del tiempo, lo que puede ser útil para detectar signos tempranos de cáncer de mama. El objetivo de la mamografía es poder obtener imágenes que aporten la máxima visibilidad de la anatomía, la mama y también de los signos de enfermedad si someter a la paciente en concreto a una exposición innecesaria a la radiación lo que podría provocar algún problema a largo plazo [67].

Existen dos tipos comunes de mamografías que se tiene que tomar en cuenta para una posterior decisión de cual escoger: CC y MLO.

CC, o cráneo-caudal, es un tipo de mamografía que se toma colocando el seno de dos maneras diferentes: una con el seno comprimido y colocado más alto en el tórax, y la otra con el seno comprimido y colocado más bajo en el tórax. Este tipo de mamografía proporciona dos vistas diferentes del seno, lo que permite un examen más completo y aumenta las posibilidades de detectar anomalías.

MLO, u oblicua mediolateral, es otro tipo de mamografía que se toma en un ligero ángulo con respecto al seno, con el seno comprimido y colocado en un ángulo de 45 grados. Este tipo de mamografía proporciona una vista diferente del seno que la vista CC y es particularmente útil para examinar las partes externa y superior del seno, donde es más probable que ocurra cáncer.

Cada base de datos puede variar en el tipo de mamografía, pero por ejemplo Mini-MIAS maneja las imágenes MLO. No afecta como tal al momento de usarlas si es que se usa la metodología que utilizamos donde se hace el recorte de al área específica de la imagen. Ilustración 2.

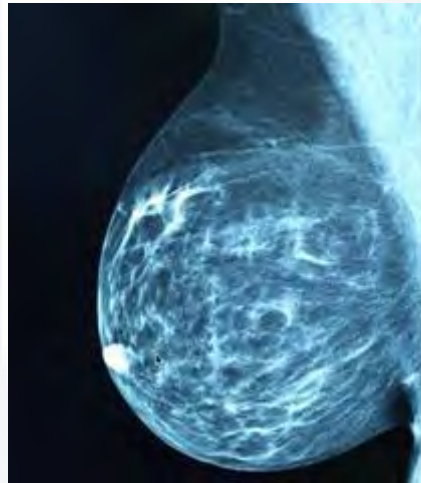


Ilustración 2 Mamografía MLO. <https://noticias.medsbla.com/noticias-medicas/oncologia-medica-oncologia-radioterapica/la-mamografia-digital-incrementa-las-biopsias-innecesarias-en-estados-unidos/>

3.3 Daños posibles de las mamografías de detección

La mamografía es una herramienta importante en la detección temprana y el diagnóstico del cáncer de mama. La principal ventaja de la mamografía es su capacidad para detectar pequeños cambios en el tejido mamario que pueden no ser visibles durante un examen físico. Esto puede ayudar a detectar el cáncer de mama en sus primeras etapas, cuando es más tratable. Además, la mamografía también se puede usar para controlar los cambios en el tejido mamario a lo largo del tiempo, lo que puede ser útil para detectar signos tempranos de cáncer de mama.

Otra ventaja de la mamografía es que está ampliamente disponible y es relativamente no invasiva. El procedimiento generalmente es bien tolerado por la mayoría de las mujeres y se puede realizar rápida y fácilmente en la mayoría de los entornos de atención médica.

Sin embargo, también existen algunas limitaciones para la mamografía. Una de las principales desventajas es que las mamografías pueden tener una alta tasa de falsos positivos, lo que significa que la prueba puede indicar la presencia de cáncer de mama cuando en realidad no lo está. Esto puede generar ansiedad innecesaria y pruebas adicionales, como biopsias, que pueden ser costosas e invasivas.

Otra limitación de la mamografía es que no siempre es precisa y es posible que no detecte todos los casos de cáncer de mama. Esto es particularmente cierto para las mujeres con tejido mamario denso, ya que es posible que las imágenes de rayos X no puedan diferenciar claramente entre tejido normal y tejido canceroso.

Además, la mamografía no se recomienda para mujeres menores de 50 años, ya que su tejido mamario es más denso y la sensibilidad de la prueba es menor.

Es importante tener en cuenta que la mamografía no debe usarse como la única herramienta de diagnóstico, sino en combinación con otras pruebas, como el examen clínico de los senos y el autoexamen de los senos. También es importante hablar sobre los posibles riesgos y beneficios de la mamografía con un proveedor de atención médica y hacerse chequeos regulares [68].

3.4 Sistema de Base de Datos e Informes de Imágenes del Seno (BI-RADS)

El sistema **BI-RADS (Breast Imaging Reporting and Data System)** es un método estandarizado para informar los resultados de las mamografías que se utiliza en los Estados Unidos. El sistema fue desarrollado por el Colegio Americano de Radiología (ACR) y se utiliza para comunicar los resultados de una mamografía tanto al paciente como al proveedor de atención médica que lo remite [68].

El sistema BI-RADS utiliza un léxico estandarizado y un sistema de clasificación para describir los resultados de una mamografía. El sistema incluye cinco categorías, cada una de las cuales corresponde a un nivel específico de sospecha de cáncer de mama:

- **Categoría 0:** examen incompleto: se necesitan imágenes o información adicional antes de poder realizar una evaluación final.
- **Categoría 1:** Negativo: no se observan anomalías en la mamografía y la paciente debe realizarse una mamografía de seguimiento de rutina en un año.
- **Categoría 2:** benigna (no cancerosa): se observa una anomalía en la mamografía, pero se determina que no es cancerosa según los resultados de una biopsia o de imágenes adicionales.
- **Categoría 3:** probablemente benigna: se observa una pequeña anomalía en la mamografía y la paciente debe repetir la mamografía en seis meses a un año para confirmar los resultados.
- **Categoría 4:** Sospechosa: se observa una anomalía en la mamografía que es sospechosa de cáncer y la paciente debe someterse a una biopsia para confirmar el diagnóstico.
- **Categoría 5:** muy sugestivo de malignidad: es muy probable que la anomalía observada en la mamografía sea cáncer y se debe realizar una biopsia a la paciente para confirmar el diagnóstico.
- **Categoría 6:** malignidad comprobada por biopsia conocida: la anomalía observada en la mamografía se confirma como maligna mediante una biopsia realizada previamente.

El sistema BI-RADS también incluye un sistema de subcategorías, que proporcionan información adicional sobre la naturaleza y ubicación de una anomalía. Además, el sistema incluye recomendaciones para la atención de seguimiento, como el momento de repetir las mamografías o las biopsias.

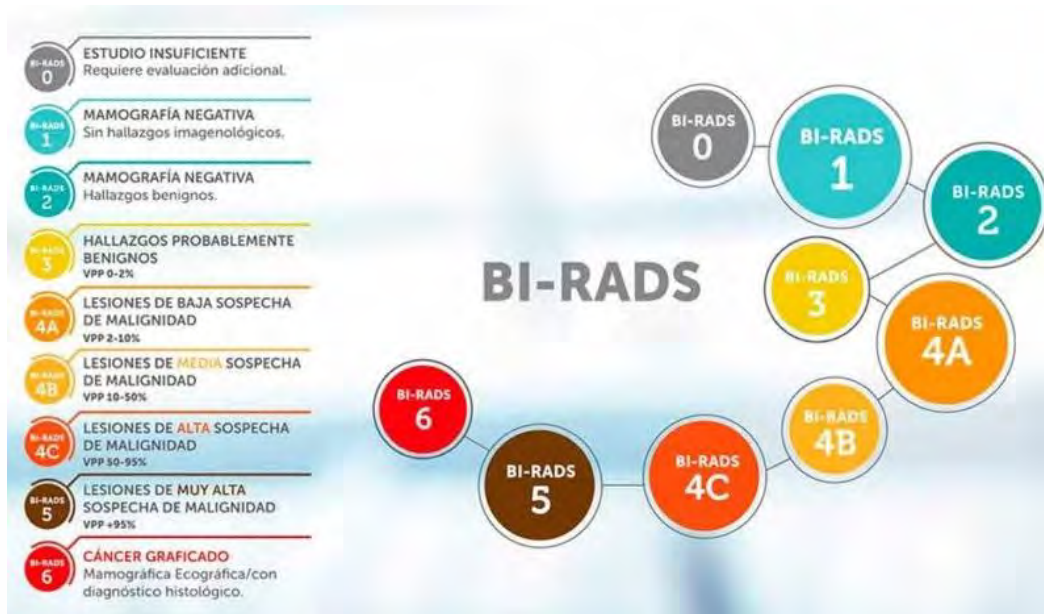


Ilustración 3 Clasificación Bi-RADS y sus categorías. <https://www.diagnosticorojas.com.ar/blog/salud/que-debo-saber-sobre-el-bi-rads/>

El sistema BI-RADS es ampliamente utilizado en los Estados Unidos y se considera un método confiable y eficiente para comunicar los resultados de una mamografía. Ayuda a estandarizar la interpretación de las mamografías y asegura que los pacientes reciban información clara y consistente sobre la salud de sus senos. Es importante tener en cuenta que el sistema BI-RADS es un sistema de informes y no un sistema de diagnóstico. Es el radiólogo quien interpreta la mamografía y asigna la categoría BI-RADS [69]. Ilustración 3.

La mamografía, al momento de detectar una lesión sospechosa de cáncer, esta entra en la clasificación de una de las categorías de BI-RADS, donde se muestra que la etapa I y II son benignas la III es probablemente benigna, mientras la IV y V aumenta la probabilidad de que sean malignas.

Tabla 1 Clasificación BI-RADS [10]

Categoría	Evaluación	Seguimiento
0	No concluyente	Requiere más estudios con imágenes para poder asignar una categoría
1	Mama normal	Examen normal. Continuar las mamografías regulares

		de detección (para mujeres mayores de 40 años de edad)
2	Hallazgos benignos (no cancerosos)	Continuar las mamografías regulares de detección (para mujeres mayores de 40 años de edad)
3	Hallazgos probablemente benignos (< 2% de riesgo de malignidad)	Hacerse una mamografía de seguimiento a los seis meses
4	Anormalidad sospechosa; considerar una biopsia	Puede requerir biopsia
5	Anormalidad que sugiere un hallazgo maligno	Requiere biopsia
6	Lesión maligna ya demostrada	Biopsia confirma la presencia de cáncer antes de iniciar tratamiento

3.5 Inteligencia artificial

La inteligencia artificial (IA) es una rama de la informática que se ocupa del desarrollo de algoritmos y sistemas que pueden realizar tareas que normalmente requerirían inteligencia humana, como comprender el lenguaje natural, reconocer imágenes y tomar decisiones. Los sistemas de IA están diseñados para realizar una amplia gama de tareas, desde tareas simples y repetitivas, como ordenar y categorizar datos, hasta tareas más complejas, como reconocimiento de imágenes y procesamiento de lenguaje natural. La IA se utiliza en una amplia gama de aplicaciones, incluidos los automóviles autónomos, el diagnóstico médico y la previsión financiera [70].

La IA también se utiliza en el aprendizaje automático, que es un subconjunto de la IA que se centra en el desarrollo de algoritmos y sistemas que pueden aprender de los datos y mejorar con el tiempo. El aprendizaje automático se utiliza

en aplicaciones como el reconocimiento de imágenes, el procesamiento del lenguaje natural y el análisis predictivo.

Es importante tener en cuenta que, si bien la IA tiene el potencial de mejorar muchos aspectos de nuestras vidas, también plantea preocupaciones éticas y sociales. Los sistemas de IA se pueden utilizar para automatizar tareas y tomar decisiones que pueden tener un impacto significativo en la vida humana. Es importante considerar las posibles consecuencias de la IA y desarrollar políticas y regulaciones que ayuden a garantizar que la tecnología se use de manera responsable y ética.

En varios campos, la inteligencia artificial (IA) se ha desarrollado rápidamente en los últimos años. Dado que las imágenes son una de las fuentes de información más importantes para las actividades que involucran la inteligencia humana, la inteligencia artificial ha estado profundamente involucrada en el procesamiento de imágenes en innumerables ocasiones. La tecnología de inteligencia artificial es ventajosa y, por lo tanto, una excelente opción para esta investigación [71], [72].

3.6 Aprendizaje supervisado

Para la investigación vio el concepto de aprendizaje supervisado, el cual es un método para el análisis de datos que resuelve problemas en base a un conjunto de datos etiquetados y estos se entrenan para una posterior clasificación.

El aprendizaje supervisado es un tipo de aprendizaje automático en el que el modelo se entrena en datos etiquetados para hacer predicciones sobre datos nuevos e invisibles. El objetivo del aprendizaje supervisado es aprender una asignación de variables de entrada (también conocidas como características o predictores) a una variable de salida (también conocida como objetivo o etiqueta). Este mapeo se puede usar para hacer predicciones sobre datos nuevos e invisible[73], [74].

El proceso de aprendizaje supervisado normalmente implica varios pasos:

- **Recopilación de datos:** se recopila un conjunto de datos etiquetados, que se utilizará para entrenar el modelo.
- **Preprocesamiento de datos:** los datos se limpian y transforman según sea necesario para garantizar que sean adecuados para su uso en el modelo.
- **Selección del modelo:** se elige una arquitectura de modelo adecuada, que se utilizará para entrenar el modelo.
- **Entrenamiento del modelo:** el modelo se entrena utilizando el conjunto de datos etiquetado. El modelo aprende a asociar patrones de entrada con etiquetas de salida.
- **Evaluación del modelo:** el modelo se evalúa en un conjunto de datos de prueba separado, para medir su rendimiento en datos no vistos.
- **Implementación del modelo:** el modelo se implementa en un escenario del mundo real, donde se utilizará para hacer predicciones sobre nuevos datos.

3.7 Algoritmos de extracción de características de las imágenes

Un concepto importante es entender la textura de la imagen y que se puede obtener de ella, que patrones son realmente importantes para tomar en cuenta y usarlos en estos modelos inteligentes.

La textura, que está determinada por la distribución espacial de los niveles de intensidad en un vecindario, es un medio para clasificar las imágenes en zonas de interés. La resolución a la que se ve una imagen tiene un impacto en el sentido de la textura, ya que una imagen grande puede no mostrar patrones finos, mientras que un examen más detallado puede exponerlos. La textura se describe como un patrón recurrente de cambios de intensidad que, en la resolución actual, son demasiado pequeños para distinguirlos como objetos individuales. Cualquier grupo

conectado de píxeles con una escala de grises constante se denomina "región texturizada".

Cuando se busca la clasificación de imágenes por ejemplo tenemos que buscar los mejores descriptores para dicha tarea. Los algoritmos clásicos del aprendizaje automático tienen buena reputación en la clasificación de problemas, pero, para esto se necesita buscar los datos para clasificar. En este caso lo que buscamos es una clasificación de mamografías primero se define el problema y dividimos las imágenes dependiendo de su clase. Cuando al final tenemos la división completa y cada imagen en su correspondiente categoría buscamos los descriptores que mejores se puedan adaptar a los algoritmos de clasificación anteriormente propuestos. En el estado del arte encontramos varios de estos descriptores, pero tres de ellos fueron los que más llamaron la atención por la simplicidad al poder obtenerlos y estos son Matriz de co-ocurrencia de nivel de gris también conocida por sus siglas en inglés como GLCM, Estadística de primer orden y la Matriz de longitud de ejecución de nivel de gris también conocida como GLRLM[75], [76].

3.7.1 La matriz de co-ocurrencia de nivel de gris (GLCM)

GLCM (Gray-Level Co-Occurrence Matrix) es un método estadístico utilizado para describir la textura de una imagen. Se basa en la idea de que la textura de una imagen se puede caracterizar por la relación espacial de los niveles de gris de los píxeles. El algoritmo GLCM funciona analizando los niveles de gris de los píxeles en una imagen y creando una matriz que describe la probabilidad de que diferentes niveles de gris ocurran juntos. Ilustración 4.

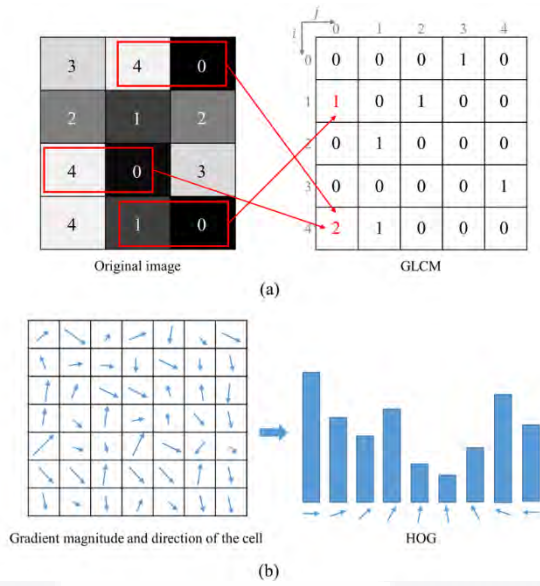


Ilustración 4 Referencia sobre el algoritmo de GLCM. https://www.researchgate.net/figure/Feature-extraction-methods-in-image-processing-GLCM-and-HOG-a-An-example-of_fig1_337940559

Esta matriz se conoce como Matriz de Co-Ocurrencia de Nivel de Gris (GLCM) y contiene información sobre la textura de la imagen [39]. El método GLCM se puede aplicar en varias direcciones y desplazamientos angulares, proporcionando una descripción más detallada de la textura. Se usa comúnmente en el procesamiento de imágenes y aplicaciones de visión por computadora, como el reconocimiento de objetos, la coincidencia de imágenes y la recuperación de imágenes. También se utiliza en imágenes médicas para detectar anomalías como tumores y en aplicaciones industriales para detectar defectos en los materiales [40], [75].

Es importante tener en cuenta que GLCM brinda una descripción estática de la textura, lo que significa que no puede tener en cuenta el movimiento o el cambio en la textura con el tiempo. Además, GLCM es sensible a la elección del desplazamiento y el tamaño de la ventana, lo que significa que los resultados pueden variar según los parámetros elegidos.

La correlación es considerada una medida de la dependencia lineal del nivel de gris entre los píxeles en las ubicaciones designadas de cada uno. Esta característica se puede representar de la siguiente forma:

Correlación: Mide la ocurrencia de la probabilidad conjunta de cada uno de los pares de píxeles especificados. El valor de la correlación es de 1 o -1. Esta característica puede ser presentada con la siguiente ecuación (1).

$$Correlation = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{\{i \times j\} \times P(i, j) - \{\mu_x \times \mu_y\}}{\sigma_x \times \sigma_y} \tag{1}$$

Contraste: Esta medición muestra las variaciones locales en la matriz de co-ocurrencia de niveles de gris. Esta característica puede ser presentada con la siguiente ecuación (1).

$$Contrast = \sum_{i=1}^G \sum_{j=1}^G |i - j|^2 P(i, j) \tag{2}$$

Segundo momento angular (ASM): Esta característica también se le conoce como uniformidad o energía. Es una medida de homogeneidad de la imagen. Esta característica puede ser presentada con la siguiente ecuación (3).

$$ASM = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{P(i, j)\}^2 \tag{3}$$

Entropía: La entropía es la suma de la diferencia de valor de intensidad en la vecindad. Esta característica puede ser presentada con la siguiente ecuación (4).

$$Entropy = \sum_{i=1}^{G-1} \sum_{j=1}^{G-1} P(i, j) \times \log(P(i, j)) \tag{4}$$

Homogeneidad local o momento de diferencia inversa (IDM): Esta medida devuelve un valor mide la proximidad de la distribución de cada elemento en el GLCM. Esta característica puede ser presentada con la siguiente ecuación (5).

$$IDM = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1 + (i + j)^2} P(i, j)$$

(5)

3.7.2 Estadísticas de primer orden (FOF)

La estadística de primer orden es un método para describir las propiedades estadísticas de una señal o imagen mediante el análisis de sus propiedades básicas, como su media, varianza e histograma. Estas estadísticas se denominan "de primer orden" porque solo involucran el análisis de píxeles o puntos de datos individuales, sin considerar la relación entre ellos [76].

Las estadísticas de primer orden son simples y computacionalmente eficientes, y se usan comúnmente en el procesamiento de imágenes y aplicaciones de visión por computadora, como la compresión de imágenes, la mejora de imágenes y el análisis de imágenes. También se utilizan en otros campos, como el procesamiento de señales y el análisis de video [77].

Es importante tener en cuenta que las estadísticas de primer orden solo brindan una descripción global de la imagen o señal y no capturan las relaciones espaciales o temporales entre píxeles o puntos de datos. Por lo tanto, no son adecuados para describir la textura o estructura de una imagen o señal. Por lo tanto, las estadísticas de segundo orden y las estadísticas de orden superior que tienen en cuenta las relaciones entre píxeles o puntos de datos también se utilizan para extraer más información de la imagen o señal.

Media: Es la suma de los valores de los píxeles dividida por el número total de píxeles. Esta característica puede ser presentada con la siguiente ecuación (6).

$$Media = \frac{1}{N} \sum_{i=1}^N X(i)$$

(6)

Desviación estándar: Es una medida de que tan dispersos están los datos en relación con la media obtenida. Esta característica puede ser presentada con la siguiente ecuación (7).

$$Desviacion\ estandar = \left(\frac{1}{N-1} \sum_{i=1}^N (X(i) - \bar{X})^2 \right)^{\frac{1}{2}}$$

(7)

Asimetría: La asimetría es una medida que mide la falta de simetría. La distribución de un conjunto de datos es simétrica si este se puede ver igual a la izquierda como a la derecha desde un punto central. Esta característica puede ser presentada con la siguiente ecuación (8).

$$Asimetria = \frac{\frac{1}{N} \sum_{i=1}^N (X(i) - \bar{X})^3}{\left(\sqrt{\frac{1}{N} \sum_{i=1}^N (X(i) - \bar{X})^2} \right)^3}$$

(8)

Curtois: Es una medida estadística que determina el grado de concentración que presentan los valores de una variable alrededor de la zona central de la distribución de frecuencias. Esta característica puede ser presentada con la siguiente ecuación (9).

$$Curtosis = \frac{\frac{1}{N} \sum_{i=1}^N (X(i) - \bar{X})^4}{\left(\sqrt{\frac{1}{N} \sum_{i=1}^N (X(i) - \bar{X})^2} \right)^2}$$

(9)

3.7.3 Matriz de longitud de ejecución de nivel gris (GLRLM)

La matriz de longitud de ejecución de nivel gris o también conocida por sus siglas en inglés (GLRLM) es una matriz donde se pueden extraer características de las imágenes para un posterior análisis de textura. El método GLRLM se considera una buena forma para la extracción de estas características y se ve los diferentes tipos de descriptores que se obtuvieron de ella.

La matriz de longitud de ejecución de nivel de grises (GLRLM) es un método estadístico utilizado en el análisis de imágenes, particularmente en imágenes médicas, para cuantificar la textura de una imagen. El método GLRLM se usa para analizar la distribución de niveles de gris en una imagen y se puede usar para identificar patrones y características que pueden no ser visibles a simple vista [40].

En GLRLM, la imagen se convierte primero en una matriz, donde cada elemento representa el nivel de gris de un píxel en la imagen. Luego, la matriz se escanea fila por fila y, para cada píxel, se registra el número de píxeles consecutivos con el mismo nivel de gris. Este proceso se repite para cada nivel de gris de la imagen, lo que da como resultado una matriz de longitudes de ejecución para cada nivel de gris.

Luego, la matriz resultante se analiza para extraer características estadísticas que describen la textura de la imagen. Estas características pueden incluir el número total de ejecuciones, la longitud media de las ejecuciones, la desviación estándar de las longitudes de las ejecuciones, el coeficiente de variación de las longitudes de las ejecuciones y el énfasis en las ejecuciones cortas [75].

Short-run emphasis (SRE): *Esta métrica aumenta cuando dominan las tiradas cortas, como en texturas finas.* Esta característica puede ser presentada con la siguiente ecuación (10).

$$SRE = \sum_{g=0}^{G-1} \sum_{l=1}^L \frac{P(g,l)}{l^2}$$

(10)

Long-run emphasis (LRE): *Esta métrica aumenta cuando dominan los plazos largos, como en grandes áreas homogéneas o texturas gruesas.* Esta característica puede ser presentada con la siguiente ecuación (11).

$$LRE = \sum_{g=0}^{G-1} \sum_{l=1}^L P(g, l) l^2 \quad (11)$$

Low gray-level run emphasis (LGRE): La métrica sube cuando se domina la textura por varias corridas de alto valor de gris. Esta característica puede ser presentada con la siguiente ecuación (12).

$$LGRE = \sum_{g=0}^{G-1} \sum_{l=1}^L \frac{P(g, l)}{(g + 1)^2} \quad (12)$$

High gray-level run emphasis (HGRE): *La métrica sube cuando se domina la textura por varias corridas de alto valor de gris.* Esta característica puede ser presentada con la siguiente ecuación (13).

$$HGRE = \sum_{g=0}^{G-1} \sum_{l=1}^L P(g, l) (g + 1)^2 \quad (13)$$

3.8 Aprendizaje automático

Un subconjunto de inteligencia artificial (IA) llamado aprendizaje automático se enfoca en desarrollar algoritmos que pueden aprender de los datos y hacer predicciones o juicios basados en ellos. Hacer posible que los ordenadores

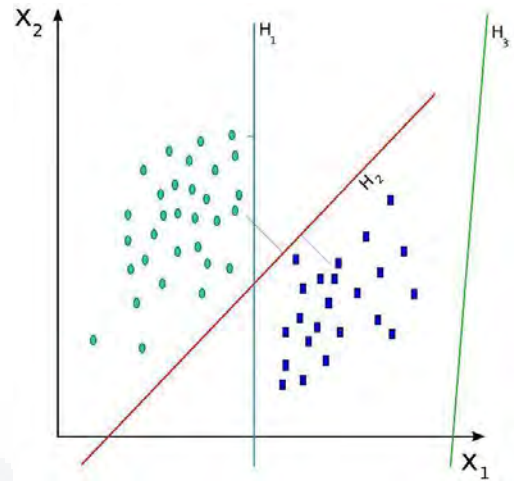
aprendan de los datos sin estar expresamente programados para realizar una determinada actividad es el objetivo del aprendizaje automático [78].

El concepto de enseñar a las computadoras a aprender de los datos fue explorado inicialmente por investigadores en la década de 1950, que es donde comenzó el aprendizaje automático. Sin embargo, la disciplina del aprendizaje automático no despegó realmente hasta finales de la década de 1980 y principios de la de 1990, cuando las mejoras en la tecnología informática permitieron el procesamiento de cantidades masivas de datos y el desarrollo de algoritmos más sofisticados [79], [80].

Debido a la enorme cantidad de datos producidos en la industria de la salud, incluidos registros médicos electrónicos, pruebas de imágenes y datos genéticos, los algoritmos de aprendizaje automático han mejorado considerablemente la investigación médica. Estos datos se pueden procesar y analizar con la ayuda de algoritmos de aprendizaje automático, lo que brinda nuevas perspectivas sobre el pronóstico, el diagnóstico y el tratamiento de enfermedades. La capacidad del aprendizaje automático para identificar patrones en los datos que pueden ser difíciles de notar para las personas es uno de los beneficios clave de emplear el aprendizaje automático en la investigación médica. La capacidad del aprendizaje automático para crear predicciones basadas en datos es otro beneficio para la investigación médica [79]. Inicialmente, estos algoritmos debían probarse, lo que requería que primero se extrajeran los descriptores de la imagen y luego buscaran el mejor modelo para sus respectivas clasificaciones. Esto fue necesario para el estudio que se hizo en la clasificación de las mamografías.

3.8.1 Support Vector Machine (SVM)

Un enfoque de aprendizaje automático supervisado popular y de uso frecuente para aplicaciones de clasificación y regresión es la máquina de vectores de soporte (SVM). Ilustración 5.



*Ilustración 5 Representación gráfica del modelo SVM.
https://es.wikipedia.org/wiki/M%C3%A1quinas_de_vectores_de_soporte.*

Se utiliza para identificar el hiperplano en un conjunto de datos cuyo margen entre clases se maximiza. El margen es la separación entre los puntos de datos más cercanos, llamados vectores de soporte, y el hiperplano [78].

A principios de la década de 1990, los investigadores buscaban un método para dividir los puntos de datos en dos clases en un espacio de alta dimensión. Fue entonces cuando surgió SVM por primera vez. El algoritmo SVM ahora se ha mejorado para manejar la clasificación de clases múltiples además de la clasificación binaria, para la que se diseñó inicialmente. SVM se usa ampliamente en una variedad de aplicaciones, incluida la clasificación de texto, la clasificación de imágenes y la bioinformática. En la investigación médica, SVM se ha utilizado para una variedad de tareas, incluido el diagnóstico y pronóstico de enfermedades, el descubrimiento de fármacos y el análisis de expresión génica [74].

3.8.2 Árbol de decisión

Un árbol de decisiones es un tipo de algoritmo de aprendizaje automático que se utiliza tanto para problemas de regresión como de clasificación. Es un modelo basado en árbol que representa un conjunto de decisiones y sus posibles consecuencias. El objetivo de un árbol de decisión es dividir los datos de entrada

en subconjuntos según los valores de las características de entrada y luego hacer una predicción según la partición a la que pertenecen los datos [78].

Los árboles de decisión se originaron durante las décadas de 1960 y 1970 cuando los académicos buscaban enfoques novedosos para la representación del conocimiento y la toma de decisiones. Inicialmente creados para problemas de clasificación binaria, los métodos de árboles de decisión ahora se han mejorado para admitir aplicaciones de regresión y clasificación de clases múltiples.

En un árbol de decisión, los datos se dividen en subconjuntos más pequeños según los valores de las características de entrada. Esto se hace creando nodos en el árbol, donde cada nodo representa una prueba de una característica. Si la prueba es verdadera, los datos se envían a una rama del árbol, y si la prueba es falsa, los datos se envían a otra rama. Este proceso continúa hasta que los datos llegan a un nodo hoja, que representa una predicción [81]. Ilustración 6.

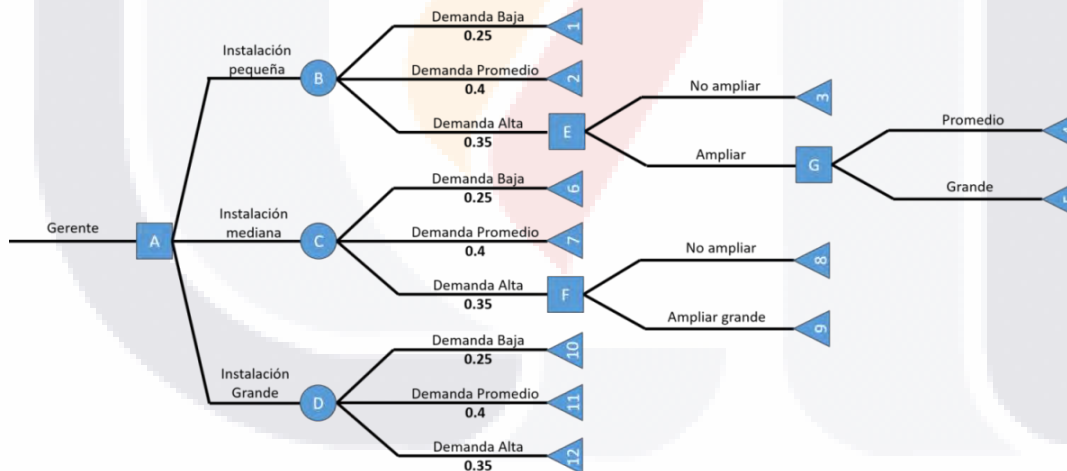


Ilustración 6 Representación gráfica de un árbol de decisiones. <https://seremprendedores.com/que-son-los-arboles-de-decisiones/>

3.8.3 K-vecinos más cercanos

Un algoritmo de aprendizaje automático simple y también muy adecuado para problemas de clasificación y regresión se llama K-vecinos más cercanos o en sus siglas en inglés, KNN. En este tipo de aprendizaje basado en instancias, el algoritmo mantiene los ejemplos de entrenamiento y predice los resultados de las nuevas instancias en función de la similitud de esas instancias con las instancias de entrenamiento que se guardaron previamente [78].

Para anticipar el resultado de una nueva instancia en función de la clase mayoritaria de sus vecinos más cercanos, la premisa principal de KNN es que es probable que una instancia sea similar a sus vecinos en el espacio de funciones. El término "K-vecinos más cercanos" se refiere al número de vecinos más cercanos utilizados para producir una predicción, que está determinada por el parámetro K.

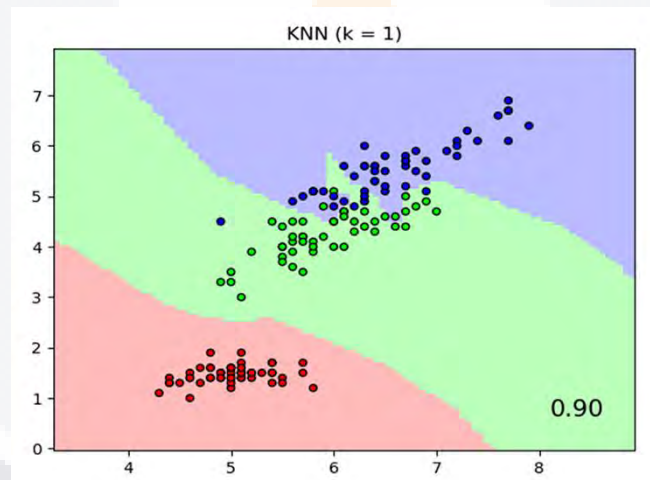


Ilustración 7 Representación gráfica de KNN. <https://www.mql5.com/en/articles/11678>

KNN se presentó por primera vez como una técnica de reconocimiento de patrones simple y eficaz en la década de 1950. Se ha mejorado y ampliado a lo largo del tiempo, y ahora se utiliza ampliamente en muchas aplicaciones diferentes, como identificación de imágenes, diagnóstico médico y sistemas de recomendación.

La sencillez y facilidad de uso de KNN son dos de sus principales características. Se puede modificar fácilmente para manejar conjuntos de datos más complicados y solo necesita una cantidad modesta de código para implementar.

KNN también tiene la ventaja de manejar datos categóricos y numéricos, lo que lo hace adecuado para una amplia gama de aplicaciones [82]. Ilustración 7

3.8.4 Bosque aleatorio

Bosque aleatorio es un algoritmo de aprendizaje automático conjunto que se utiliza para problemas de clasificación y regresión. Es una extensión de los árboles de decisión, donde varios árboles se combinan para formar un bosque, donde cada árbol hace su propia predicción y la predicción final se hace combinando las predicciones de todos los árboles [78].

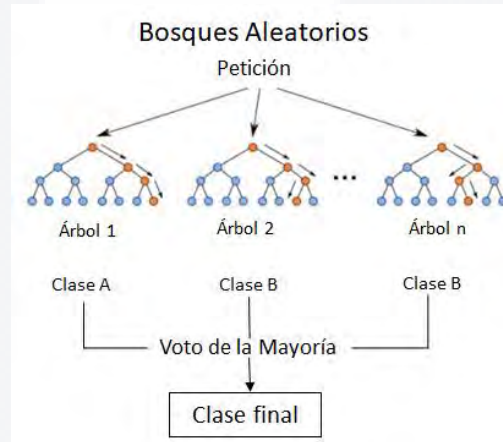


Ilustración 8 Representación gráfica de bosque aleatorio. <https://medium.com/@hpumah/bosques-aleatorios-482163ace92e>

El enfoque del algoritmo de bosque aleatorio está diseñado para evitar el sobreajuste, mejorar la generalización y aumentar la precisión al integrar las predicciones de varios árboles. Esto se debe al hecho de que cada árbol del bosque se creó utilizando un subconjunto único de los datos de entrenamiento, y cada árbol tiene una forma única de dividir los datos. Como resultado, habrá menos sobreajuste y mayor precisión porque es menos probable que los árboles del bosque tengan fuertes correlaciones entre sí.

Desde su introducción inicial en la década de 1990, Bosque aleatorio se ha convertido en uno de los algoritmos de aprendizaje automático más populares.

Funciona de manera efectiva para problemas que involucran interacciones no lineales complejas entre las características y el objetivo, así como grandes conjuntos de datos dimensionales [42]. Ilustración 8.

3.8.5 Clasificador de aumento de gradiente

Clasificador de aumento de gradiente es un algoritmo de aprendizaje automático que se utiliza para problemas de clasificación binarios y de clases múltiples. Es un tipo de aprendizaje conjunto en el que se combinan múltiples modelos débiles para formar un modelo fuerte. Clasificador de aumento de gradientes un algoritmo iterativo que mejora el modelo con cada iteración al agregar un nuevo alumno débil que intenta corregir los errores cometidos por modelos anteriores [83].

El concepto de "descenso de gradiente", en el que la función objetivo se minimiza iterativamente para obtener la mejor solución, es la base del clasificador potenciador de gradiente. La función de pérdida, que calcula la discrepancia entre el objetivo proyectado y el objetivo real, sirve como función objetivo en esta situación. Para reducir la función de pérdida, el algoritmo incorpora un nuevo alumno débil en cada iteración.

Una de las ventajas clave del clasificador potenciador de gradientes es su capacidad para manejar datos con ruido y valores atípicos. Además, también es resistente al sobreajuste, que es un problema común en los algoritmos basados en árboles de decisión. Los clasificadores que aumentan el gradiente también pueden manejar conjuntos de datos con un número desigual de ejemplos positivos y negativos [84].

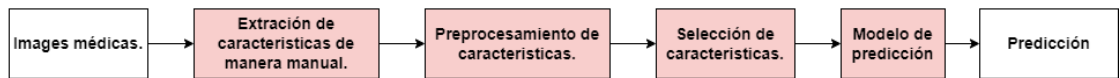
3.9 Aprendizaje profundo

Los modelos de inteligencia artificial (IA) están tomando un papel importante en la investigación y creación de servicios en el cuidado de la salud. Para esto se tiene que buscar los mejores modelos que puedan llevar a mejores resultados y con esto un subconjunto del aprendizaje automático ha tomado mucha importancia en estos

últimos años llamado aprendizaje profundo. El aprendizaje profundo está logrando grandes avances en resolver problemas que se habían resistido entre la comunidad de IA por varios años [85], [86].

Y una grande ventaja que tiene en comparación a modelos clásicos del aprendizaje automático, es la extracción de características de manera automática. Esto ayuda mucho en cierto de tipos de tareas donde no se tiene claro que tipo de descriptor podría ser el indicado al usar al igual que entre más datos se tenga estos modelos profundos pueden lograr tener mucho mejor desempeño [85]. Ilustración 9.

(a) Modelo de aprendizaje automático.



(b) Modelo aprendizaje profundo.

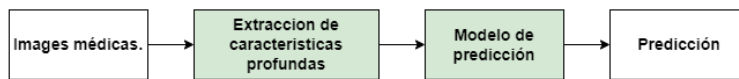


Ilustración 9 Diferencia entre aprendizaje automático y aprendizaje profundo

3.9.1 Redes neuronales

Uno de las ramas más interesantes y estudiadas de la inteligencia artificial son las redes neuronales artificiales o también conocidas por sus siglas en inglés como (ANN). Las redes neuronales artificiales son algoritmos matemáticos generados por computadora. Estas redes están diseñadas para aprender de los datos y aprender el conocimiento que contienen. Estos modelos son digitalizaciones del cerebro biológico [85].

Las redes neuronales han sido ampliamente utilizadas en las ciencias y en la tecnología con aplicaciones en varias ramas como podría ser la química, la física y la biología. Estas redes tratan de una técnica de aprendizaje supervisado y están formadas por varias capas con varias neuronas (Aunque existe los casos que son solo un perceptrón es suficiente para ser llamado una red neuronal).

De la manera más básica su topología está constituida de la siguiente forma: una capa de entrada donde entran los datos para la alimentación de la red, una capa de salida que va a representar la estimación del algoritmo y finalmente las capas internas que pueden variar dependiendo la complejidad del algoritmo que se esté utilizando. Esto será una transición donde los valores de entradas y las capas internas se les darán unos pesos y bias que se irán adaptando en la etapa del entrenamiento. Cada una de estas neuronas tendrá una función de activación que será defina en base a las necesidades de la investigación o proyecto [87].

Los algoritmos de aprendizaje profundo pueden superar a los métodos tradicionales de aprendizaje automático debido a su capacidad para manejar grandes cantidades de datos, aprender representaciones jerárquicas y extraer automáticamente funciones que son relevantes para el problema en cuestión. Esto permite que el modelo aprenda patrones más complejos en los datos y haga predicciones más precisas en comparación con los modelos superficiales. Además, los algoritmos de aprendizaje profundo también pueden manejar relaciones no lineales entre entradas y salidas, mientras que los métodos tradicionales de aprendizaje automático pueden tener problemas con este tipo de problemas [85], [87]. Ilustración 10.

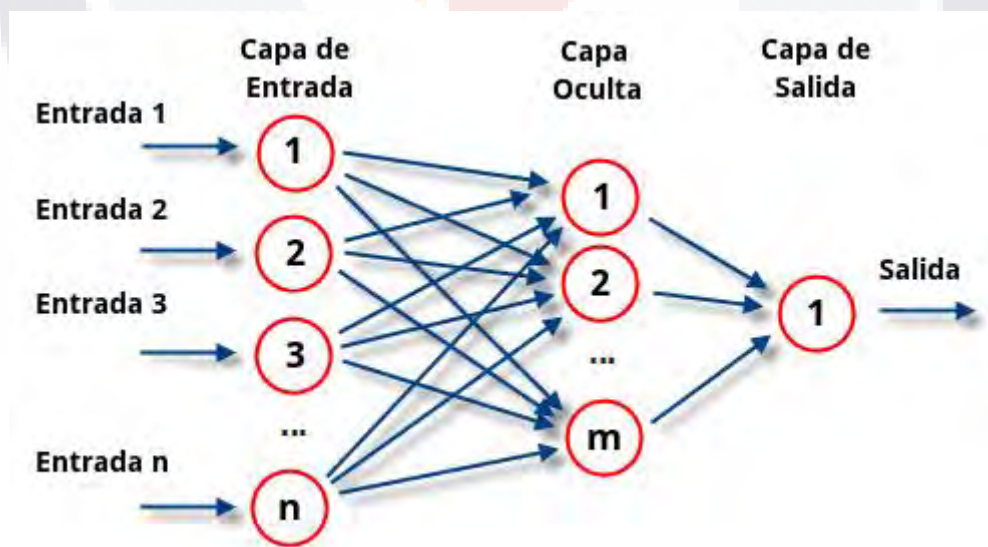


Ilustración 10 Modelo de una red neuronal artificial básica con 2 capas ocultas.
<https://www.atriainnovation.com/que-son-las-redes-neuronales-y-sus-funciones/>

Otra ventaja del aprendizaje profundo es su capacidad para aprender a partir de datos no estructurados, como imágenes y audio, con los que luchan los métodos tradicionales de aprendizaje automático. Mediante el uso de redes neuronales convolucionales (CNN) o redes neuronales recurrentes (RNN), los algoritmos de aprendizaje profundo pueden aprender representaciones significativas de este tipo de datos y hacer predicciones con gran precisión [88].

Las neuronas de cada capa están conectadas a todas las neuronas de la capa anterior y así sucesivamente, pero en cada capa independiente las neuronas que la conforman no tienen ninguna conexión interna [87]

Un modelo simple de neurona artificial básica que será la base para comprender una más compleja consiste en:

- Un conjunto de entradas.
 (X_1, \dots, X_n)
- Los pesos sinápticos que corresponden a cada entrada.
 (W_1, \dots, W_n)
- Una función de agregación.

$$\sum_{i=1}^n (W_i X_i)$$

- Una función de activación.
- Una salida.

Este tipo de redes resultaban útiles y habían dado buenos resultados en diferentes tipos de investigaciones. Incluso, muchos proyectos de regresión o clasificación más básicos podían obtener buenos resultados. Sin embargo, el problema surgía al utilizar muchos más datos de entrada, como en el caso del

análisis de imágenes. Una de las formas de utilizar estas redes neuronales en el análisis de imágenes era convertir cada pixel de la imagen en una entrada de la red, lo que implicaba una gran cantidad de entradas. Por ejemplo, una imagen a color de 224x224x3 requeriría 150,528 neuronas de entrada, lo que podría provocar problemas de escalabilidad.

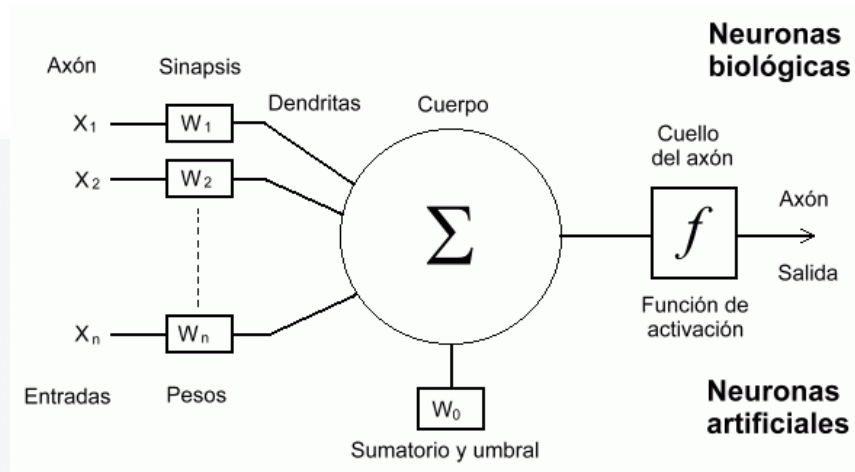


Ilustración 11 Modelo de un perceptrón. <http://www.cs.us.es/~fsancho/?e=72>

Cabe mencionar que, si se utilizaron las redes neuronales convencionales, pero sin utilizar las entradas de los píxeles de la imagen sino al igual que los otros modelos convencionales del aprendizaje automático. Se usaron los descriptores anteriormente mencionados como GLCM, FOF, y GLRLM. Esto para intentar un mejor desempeño con muchas menos entradas. Sin embargo, en esta investigación se decidió ir principalmente por unas redes neuronales más complejas llamadas redes neuronales convolucionales (CNN). Las CNN son conocidas por encontrar patrones en imágenes con mayor índice de exactitud y tener mejor desempeño general que sus antecesores.

3.9.2 Redes neuronales convolucionales.

Las redes neuronales convolucionales o CNN (Convolutional Neuronal Network por sus siglas en inglés) son un tipo de red neuronal artificial que aprenden directamente de los datos eliminando la necesidad de extracción manual de las características como sería en el aprendizaje automático [85].

Las CNN están diseñadas para imitar la función la corteza del cerebro humano. Y son particularmente buenas para encontrar patrones en las imágenes y así poder reconocer objetos o caras. CNN también son muy efectivas para las clasificaciones datos que no estén en las imágenes como podría ser el audio. La razón del buen funcionamiento de las CNN son las capas encargadas en la extracción de las características, capas que no se encuentran en una red neurona artificial convencional, y estas son las siguientes: Capa convolucional, capa de rectificación (ReLU), capa de agrupamiento, capa dropout, capa de clasificación. Ilustración 12.

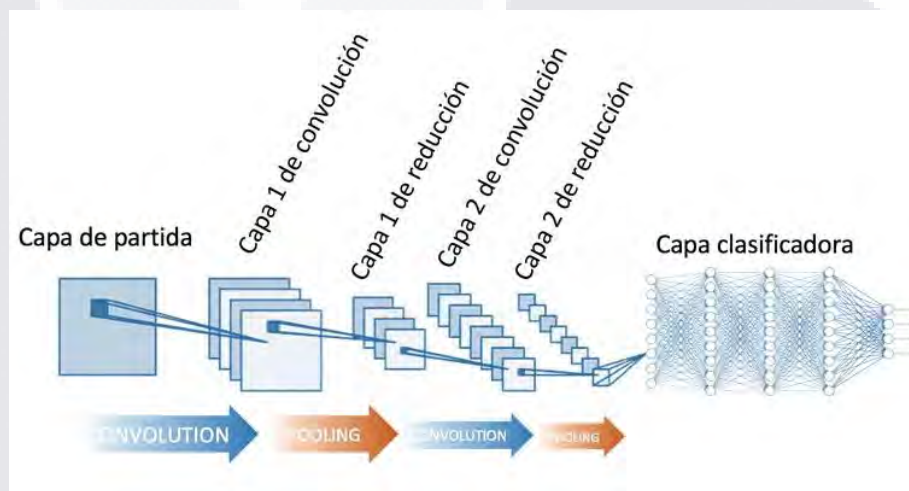


Ilustración 12 Modelo estándar para una red neuronal convolucional. <https://www.diegocalvo.es/red-neuronal-convolucional/>

3.9.2.1 Capa convolucional

Esta capa pone a la imagen de entrada a través de múltiples filtros de convolución, donde cada uno muestra diferentes características de la imagen (Pueden ser una cantidad n de filtros).

Las capas de convolución toman de entrada una imagen (Donde la imagen es tomada como una matriz de $m \times n$) y esta misma da como salida múltiples imágenes nuevas que son en base a los filtros que se le fueron dados de manera aleatoria.

Para esto se tiene que especificar el número de filtros con el cual se quiere trabajar, el tamaño de estos mismos y los pasos para su desplazamiento.

En la siguiente ilustración se puede observar como por medio de una imagen se pueden obtener en la salida diferentes imágenes con su respectivo filtro. Ilustración 13.



Ilustración 13 Imágenes sacadas de la capa de convolución. <https://datascience.aero/aviation-function-deep-learning/>

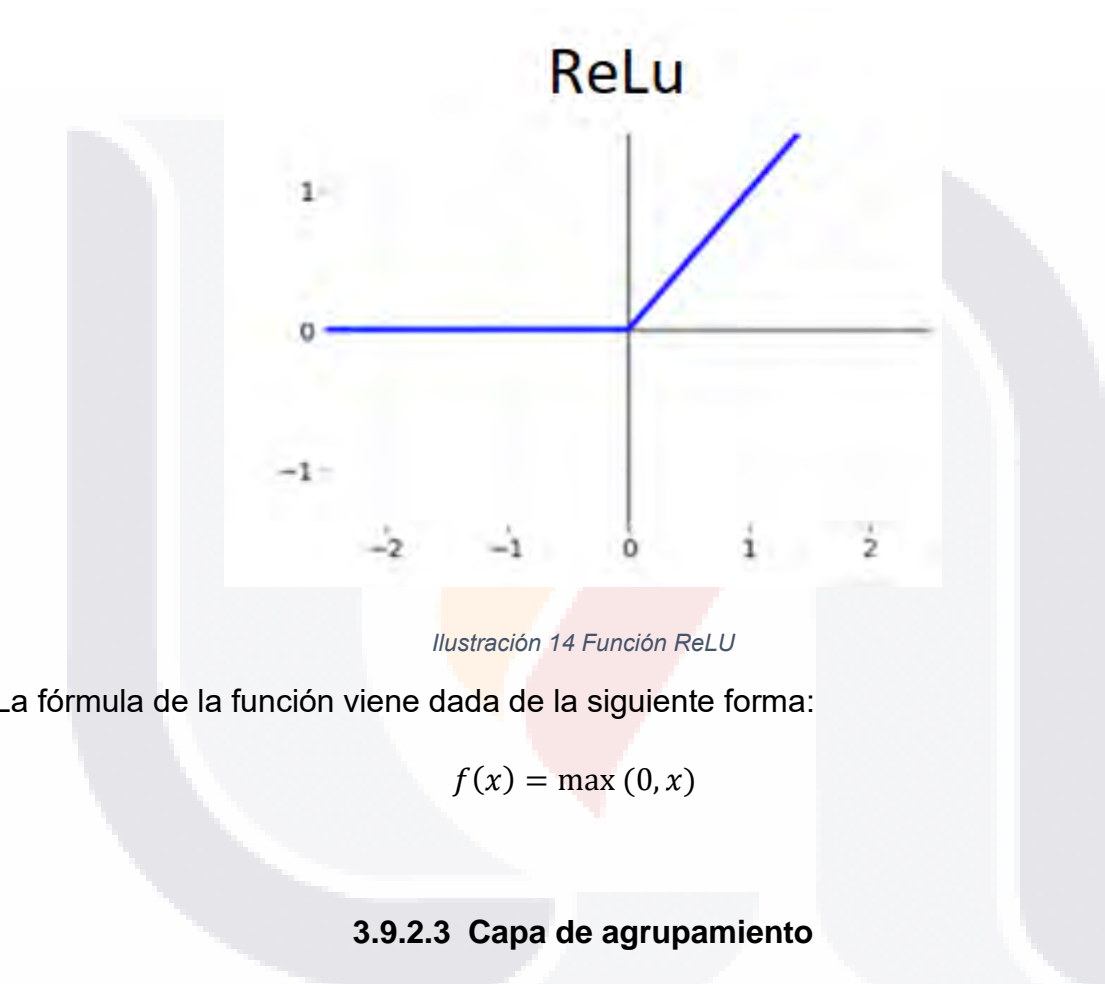
Cuando se está diseñando la topología de la red que se esté utilizando es importante saber que las capas de convolución detectan patrones complejos en base a la imagen de entrada que se le está proporcionando. En otras palabras, no es necesario elegir o buscar los filtros para detectar estos patrones, sino que estos filtros se generan con el mismo entrenamiento. Lo que quiere decir que las mismas capas buscan los mejores filtros para así obtener los patrones más característicos.

3.9.2.2 Capa de rectificación (ReLU)

Las ReLU son neuronas utilizadas para una función rectificadora. Las ReLU permiten un entrenamiento más rápido y efectivo donde los valores negativos los

pasa como 0 y mantienen todos los valores positivos. Esta capa es también conocida como activación y hay varios tipos que se pueden usar en este tipo de redes (como la función sigmoid y la función tangente hiperbólica). Pero se menciona ReLU ya que es uno de lo más utilizados en las arquitecturas actuales [85], [89].

Ilustración 14.



La fórmula de la función viene dada de la siguiente forma:

$$f(x) = \max(0, x)$$

3.9.2.3 Capa de agrupamiento

Esta capa divide la imagen de entrada por lo general después de la capa de convolución. Simplifica la salida donde se reduce el número de parámetros. Esta capa selecciona los rangos más representativos de cada región [85], [87]. Esta capa ayuda por igual a disminuir el coste computacional para las siguientes capas, sin embargo, habrá pérdida de información en la imagen. Hay varios tipos de capas de agrupamiento, pero las más conocidas son (Maxpooling y Averagepooling).

En la Ilustración 15 se muestra el concepto de agrupamiento máximo (Maxpooling) donde se obtiene el número mayor del núcleo asignado. En el siguiente ejemplo los valores serían 9 y 7 respectivamente.

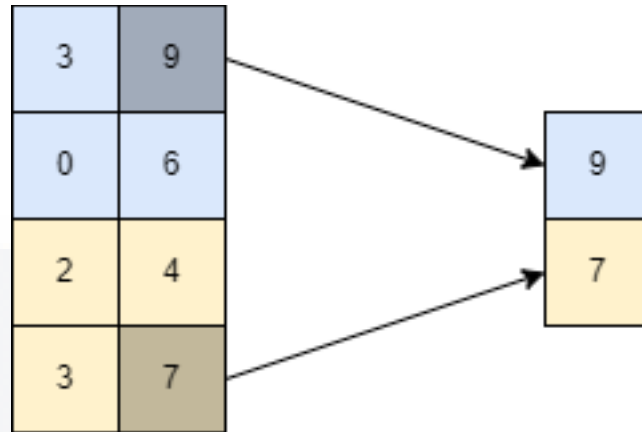


Ilustración 15 Agrupamiento máximo (Maxpooling).

En la Ilustración 16 el concepto de agrupamiento es similar pero no se obtiene el número mayor del núcleo, sino que se hace la suma de todos los valores dentro de ese mismo núcleo. Y los valores obtenidos pueden ser enteros, pero en el caso que sean decimales se redondean.

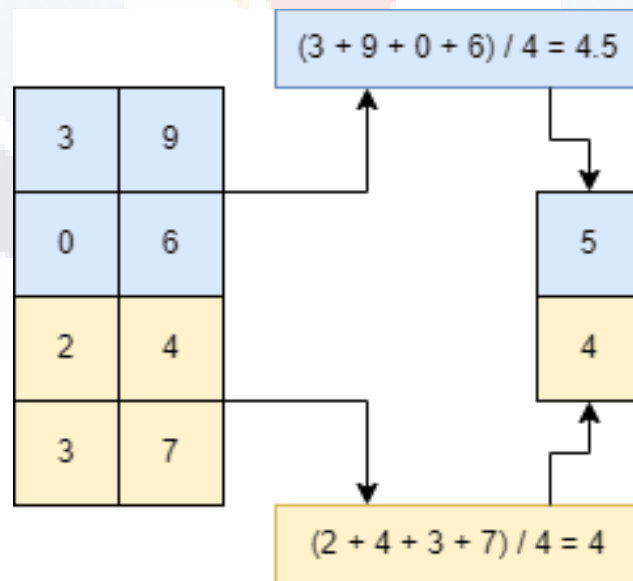


Ilustración 16 Agrupamiento promedio (Averagepooling).

3.9.2.4 Capa Dropout

Aunque *dropout* no es considerada como una capa como tal sino como una modificación que se puede añadir, aun así, se puede agregar en la topología como si fuera una capa más (Tensorflow lo permite) [88], [90]. Ilustración 17.

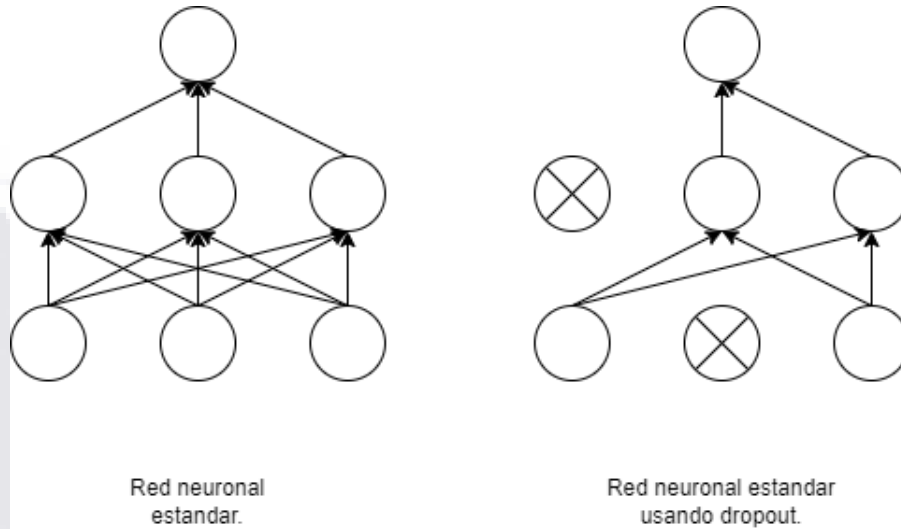


Ilustración 17 función de dropout.

El *dropout* permite desconectar de manera aleatoria un porcentaje de sus neuronas, esto ayuda en el entrenamiento a ignorar parte de la información que está disponible. De esta forma la red es obligada a aprender patrones más robustos y generales, el *dropout* es muy utilizado para el sobreajuste. En la Ilustración 17. se muestra cómo sería la red un antes y después de utilizar el *dropout*.

3.9.2.5 Capa de salida (Capa de clasificación)

Esta capa que es la última parte de toda la topología está encargada de darle los valores de salida. Estos mismos son interpretados con probabilidades de pertenecer a cada clase que se esté utilizando. Existen dos formas que son de las utilizadas en la tarea de clasificación: Clasificación binaria y clasificación multiclase [85].

La función logística tiene forma de "S", por lo que también se la conoce como función "sigmoidea". Esta forma significa que la función puede tomar cualquier valor

de entrada y asignarlo a un valor entre 0 y 1, lo que la hace útil para modelar la probabilidad. Esta es la razón por la que la función sigmoidea se usa en el aprendizaje profundo como una función de activación en la red neuronal artificial. Ilustración 18.

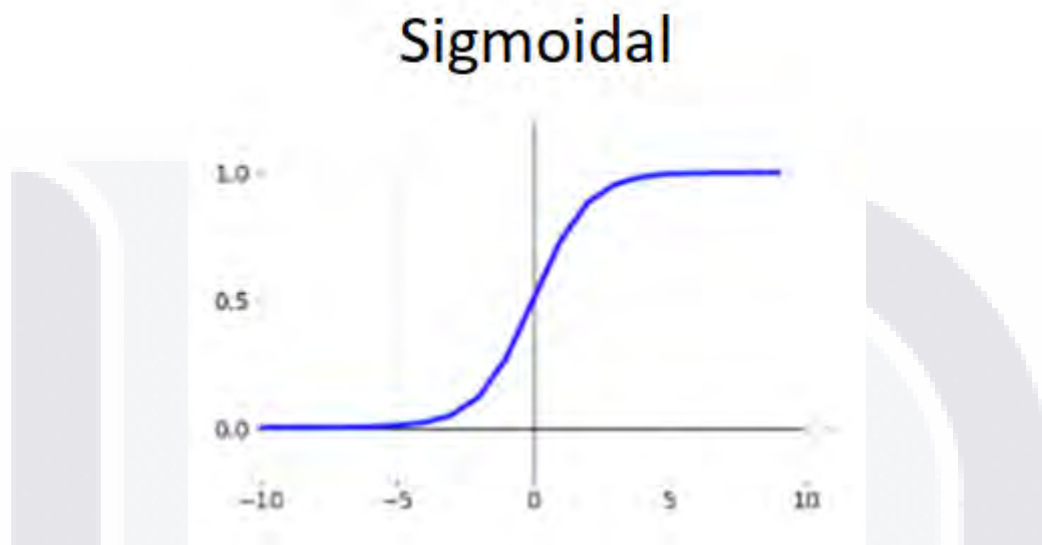


Ilustración 18 Función sigmoideal.

La función sigmoidea es diferenciable, lo que la hace útil para algoritmos de optimización como el descenso de gradiente, que se utilizan para entrenar redes neuronales. La derivada de la función sigmoidea es fácil de calcular y se puede utilizar para calcular el gradiente de la función de costo con respecto a los pesos y sesgos de la red.

La función sigmoidea tiene algunas otras propiedades que la hacen útil para las redes neuronales. Por ejemplo, es una función no lineal, lo que significa que puede modelar relaciones complejas entre entradas y salidas. También tiene un rango de salida fijo de [0, 1], lo que lo hace útil para modelar la probabilidad [91], [92].

En la primera después de sacar las características el clasificador se encarga de decidir si pertenece a una clase u a otra (por lo general se utilizada una función sigmoide para obtener valores entre 0 y 1). La siguiente formula es la representación matemática de la función sigmoide (14):

$$\text{Sigmoid } \mathcal{S}(x) = \frac{1}{1 + e^{-x}}$$

(14)

En la segunda la clasificación es muy parecida, pero en este ejemplo después de obtener las características se buscará por un porcentaje a que clase los datos pertenecen entre 3 o más clases (para esta tarea de clasificación se utilizaba la función softmax)

La función softmax se define por elementos, lo que significa que se aplica por separado a cada elemento del vector de entrada. Convierte la entrada en la distribución de probabilidad de la salida mediante la normalización de los valores de entrada. La salida de la función softmax se denomina softmax de la entrada. La función softmax tiene algunas propiedades que la hacen útil para las redes neuronales. Por ejemplo, es diferenciable, lo que lo hace útil para algoritmos de optimización como el descenso de gradiente, que se utilizan para entrenar redes neuronales. La derivada de la función softmax con respecto a sus entradas se puede calcular fácilmente [93]. Ilustración 19.

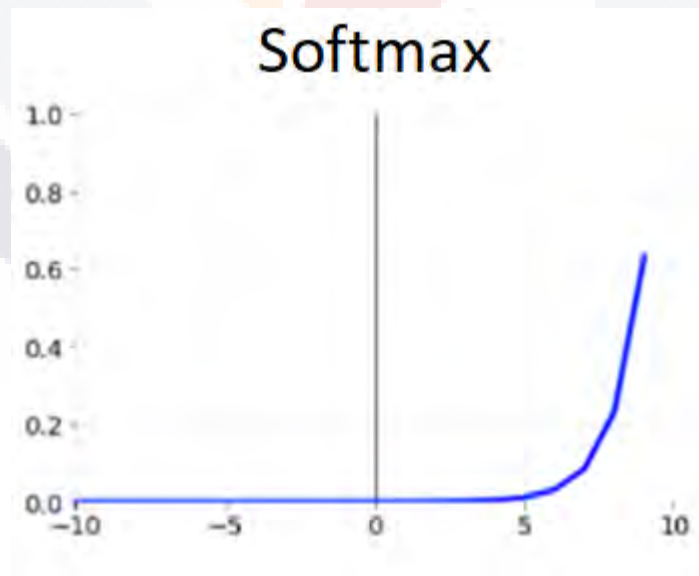


Ilustración 19 Función softmax.

Otra propiedad importante es que la salida de la función softmax está centrada en cero, lo que significa que está más equilibrada en términos de valores positivos y negativos. Esto puede ayudar a mitigar el problema de sesgo que puede causar la función sigmoidea. La siguiente fórmula es la representación matemática de la función softmax (15):

$$\text{Softmax } \sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \cdot (15)$$

3.9.3 Conceptos importantes al utilizar CNN

3.9.3.1 Sobreajuste (Overfitting)

El sobreajuste en el aprendizaje profundo ocurre cuando un modelo se entrena demasiado bien en un conjunto de datos de entrenamiento y, como resultado, se desempeña mal en datos nuevos e invisibles. Esto sucede porque el modelo ha aprendido el ruido y las variaciones aleatorias presentes en los datos de entrenamiento, en lugar de los patrones y relaciones subyacentes que se pueden generalizar a nuevos datos [85].

En el aprendizaje profundo, el sobreajuste puede ocurrir cuando un modelo tiene demasiados parámetros, lo que lo hace más complejo de lo necesario. Esta complejidad permite que el modelo se ajuste muy bien a los datos de entrenamiento, pero también aumenta la probabilidad de que el modelo se ajuste al ruido de los datos, en lugar de a los patrones subyacentes.

Una de las formas más comunes de evitar el sobreajuste en el aprendizaje profundo es utilizar técnicas de regularización. Estas técnicas agregan un término de penalización a la función de pérdida del modelo, lo que desalienta al modelo a asignar pesos altos a ciertos parámetros. Las formas más comunes de regularización para el aprendizaje profundo son la regularización L1 y L2, que agregan penalizaciones basadas en el valor absoluto o el cuadrado de los valores de los parámetros, respectivamente [94].

Otra forma de evitar el sobreajuste es mediante el uso de más datos. Tener más datos permite que el modelo aprenda más características generales y menos ruido. Sin embargo, recopilar más datos puede ser difícil o costoso, por lo que no siempre es una solución práctica.

3.9.3.2 Aumento de datos (Data augmentation):

El aumento de datos es una técnica utilizada para aumentar artificialmente el tamaño de un conjunto de datos mediante la creación de versiones modificadas de los datos existentes. Esto se hace aplicando varias transformaciones aleatorias a las imágenes, como rotación, escalado, recorte y volteo. El objetivo del aumento de datos es aumentar la diversidad de los datos de entrenamiento, lo que puede ayudar a reducir el sobreajuste y mejorar el rendimiento de generalización de los modelos de aprendizaje profundo.

En el aprendizaje profundo, el aumento de datos se usa a menudo en combinación con redes neuronales convolucionales (CNN) para tareas de clasificación de imágenes. Las CNN pueden aprender características de las imágenes de entrada, pero pueden ser sensibles a pequeños cambios en la entrada, como la traducción y la rotación. El aumento de datos se puede utilizar para exponer el modelo a estas variaciones durante el entrenamiento, lo que puede mejorar su solidez y capacidad para generalizar a nuevos datos.

3.9.3.3 Parada temprana (Early Stopping):

El **Early Stopping** o también conocida como detención anticipada es una técnica utilizada para evitar el sobreajuste en el aprendizaje profundo al interrumpir el proceso de capacitación cuando el rendimiento en un conjunto de validación comienza a degradarse. El sobreajuste ocurre cuando un modelo se especializa demasiado en los datos de entrenamiento, lo que resulta en un rendimiento de generalización deficiente en datos nuevos [95].

En el aprendizaje profundo, la detención anticipada se implementa al monitorear el rendimiento del modelo en un conjunto de validación durante el proceso de capacitación. El proceso de entrenamiento se detiene cuando el rendimiento en el conjunto de validación comienza a degradarse, como lo indica una métrica como la pérdida de validación o la precisión de la validación. Al detener el proceso de entrenamiento antes de que ocurra el sobreajuste, el modelo puede generalizar mejor los datos nuevos.

Existen diferentes estrategias para determinar cuándo detener el proceso de entrenamiento mediante la detención anticipada. Un enfoque común es usar un número fijo de épocas de entrenamiento. En este enfoque, el proceso de entrenamiento se detiene después de un cierto número de épocas, independientemente del rendimiento en el conjunto de validación.

3.9.3.4 Abandono (Dropout):

Capas de *dropout*: Con las capas de *dropout* se reduce pesos grandes lo que reduce la complejidad del modelo. Como se había mencionado anteriormente de esta capa o método, se apagan aleatoriamente neuronas para prevenir la complejidad del modelo y así prevenir el sobreajuste.

El *dropout* es una técnica utilizada para evitar el sobreajuste en el aprendizaje profundo mediante el abandono aleatorio (es decir, la configuración a cero) de un cierto porcentaje de neuronas durante el proceso de entrenamiento. Al descartar neuronas aleatoriamente, el modelo se ve obligado a aprender representaciones más sólidas y redundantes, lo que puede ayudar a mejorar el rendimiento de generalización del modelo.

El *dropout* generalmente se aplica a las capas completamente conectadas de una red neuronal profunda, pero también se puede aplicar a las capas convolucionales. La tasa de abandono es un hiperparámetro que determina la proporción de neuronas a abandonar. Una tasa de *dropout* de 0.5 significa que se abandonará la mitad de las neuronas de la capa. El *dropout* se puede implementar

TESIS TESIS TESIS TESIS TESIS

aplicando una máscara binaria a las activaciones de una capa durante el paso hacia adelante. La máscara se genera mediante el muestreo de una distribución de Bernoulli con probabilidad p , donde p es la tasa de abandono. Las neuronas correspondientes a un valor de 1 en la máscara se mantienen, mientras que las neuronas correspondientes a un valor de 0 se eliminan [85], [90].

3.9.3.5 Optimizadores

El optimizador es una parte fundamental para el entrenamiento de las CNN. Estos algoritmos o métodos son usados para cambiar los atributos de las redes neuronales que se está entrenando. Puede ser para red neuronal estándar o una CNN. Los atributos que cambian son los pesos y la tasa de aprendizaje [96].

En el aprendizaje profundo, un optimizador es un algoritmo que se utiliza para actualizar los parámetros del modelo durante el entrenamiento, con el fin de minimizar la función de pérdida. El objetivo del optimizador es encontrar el conjunto de parámetros que minimizan la pérdida, para que el modelo pueda hacer predicciones precisas sobre datos no vistos [85].

Hay varios tipos diferentes de optimizadores que se pueden usar en el aprendizaje profundo, cada uno con sus propias fortalezas y debilidades [97]. Algunos de los optimizadores más populares incluyen:

- **Descenso de gradiente estocástico / Stochastic Gradient Descent (SGD):** este es uno de los optimizadores más básicos y más utilizados. Utiliza el gradiente de la función de pérdida con respecto a los parámetros del modelo para actualizar los parámetros en la dirección opuesta al gradiente.
- **Momentum:** este optimizador es una extensión de SGD que utiliza el impulso para suavizar las actualizaciones de los parámetros. Utiliza los gradientes pasados acumulados para actualizar los parámetros.
- **Nesterov Momentum:** este optimizador es similar a Momentum, pero usa el gradiente de la posición futura en lugar de la posición actual. Esto puede ayudar a mejorar la estabilidad del proceso de entrenamiento.

- **Estimación de momento adaptativo / Adaptive Moment Estimation (Adam):** este es un optimizador más avanzado que cambia de forma adaptativa la tasa de aprendizaje para cada parámetro en función de la información de gradiente histórica.
- **Propagación de raíz cuadrática media / Root Mean Square Propagation (RMSProp):** este optimizador también es una extensión de SGD, pero utiliza un promedio móvil del gradiente cuadrático para escalar la tasa de aprendizaje.
- **Adagrad:** este optimizador adapta la tasa de aprendizaje de cada parámetro en función de la información de gradiente histórica y requiere menos memoria en comparación con otros optimizadores.

Es importante tener en cuenta que la elección del optimizador depende de la tarea específica y el tipo de datos. Cada optimizador tiene sus propias propiedades únicas y es importante experimentar con diferentes optimizadores para encontrar el mejor para la tarea específica.

Para esta investigación se utilizó el optimizador Adam por su gran desempeño y capaz de obtener muy buenos resultados [98].

Adam también tiene la ventaja de ser computacionalmente eficiente y requiere menos memoria en comparación con otros optimizadores, porque solo almacena un promedio móvil de tamaño fijo del gradiente al cuadrado, en lugar del historial completo. Este optimizador se puede usar con una variedad de redes neuronales y es resistente a la elección de la tasa de aprendizaje inicial. También puede manejar gradientes escasos y regularización L1/L2.

Adam se considera uno de los mejores optimizadores para el aprendizaje profundo porque combina las ventajas de varios otros algoritmos de optimización, como:

- Utiliza el gradiente de los parámetros actuales para actualizar los parámetros, como en el optimizador Stochastic Gradient Descent (SGD).

- También utiliza la información histórica del gradiente para adaptar la tasa de aprendizaje de cada parámetro, como en el optimizador de Adagrad.
- Utiliza el impulso para suavizar las actualizaciones de los parámetros, como en los optimizadores Momentum y Nesterov Momentum.

3.9.3.6 Aprendizaje de transferencia.

El aprendizaje por transferencia es una técnica de aprendizaje automático en la que un modelo entrenado en una tarea se utiliza como punto de partida para un modelo en una segunda tarea. Esto puede generar tiempos de entrenamiento más rápidos y un mejor rendimiento en la segunda tarea, especialmente si la segunda tarea tiene datos limitados disponibles [99], [100].

Hay varios tipos diferentes de aprendizaje de transferencia, que incluyen:

- Aprendizaje de transferencia basado en instancias: donde un modelo se entrena en un conjunto de datos y luego se ajusta en un segundo conjunto de datos. Esto se usa a menudo en tareas de visión por computadora, donde un modelo entrenado en un gran conjunto de datos de imágenes se ajusta con precisión en un conjunto de datos de imágenes más pequeño para una tarea específica.
- Aprendizaje de transferencia basado en características: donde las características aprendidas por un modelo entrenado en una tarea se utilizan como entrada para un modelo entrenado en una segunda tarea. Esto se usa a menudo en tareas de procesamiento de lenguaje natural, donde las incrustaciones de palabras aprendidas por un modelo entrenado en un gran conjunto de datos de texto se usan como entrada para un modelo entrenado en un conjunto de datos de texto más pequeño.
- Aprendizaje de transferencia de modelo preentrenado: donde un modelo preentrenado se utiliza como punto de partida para una nueva tarea. Esto se usa a menudo en tareas de visión por computadora, donde un modelo

previamente entrenado se usa como extractor de características y las características extraídas se usan como entrada para un nuevo modelo.

En general, el aprendizaje de transferencia puede ser beneficioso cuando la tarea en cuestión tiene datos etiquetados limitados, o cuando la tarea es muy similar a la tarea en la que se entrenó previamente el modelo, lo que puede ser el caso en muchas aplicaciones del mundo real.

El aprendizaje de transferencia es una técnica en la que un modelo previamente entrenado, generalmente entrenado en un conjunto de datos grande, se usa como punto de partida para entrenar un nuevo modelo en un conjunto de datos más pequeño. La idea es que el modelo previamente entrenado ya haya aprendido características útiles del gran conjunto de datos, que se pueden reutilizar en el nuevo modelo para mejorar su rendimiento [101].

Hay dos tipos principales de transferencia de aprendizaje: **basado en características y ajuste fino**.

- En el aprendizaje de transferencia basado en características, el modelo preentrenado se usa para extraer características del nuevo conjunto de datos, y se entrena un nuevo clasificador sobre estas características. Esto normalmente se hace cuando el nuevo conjunto de datos tiene un número diferente de clases o una estructura diferente que el conjunto de datos original.
- El ajuste fino, por otro lado, implica entrenar el modelo previamente entrenado en el nuevo conjunto de datos, normalmente con una tasa de aprendizaje menor que en el entrenamiento original.

Los modelos preentrenados están disponibles para la mayoría de las arquitecturas y ya están entrenados en grandes conjuntos de datos. Estos modelos están entrenados en grandes conjuntos de datos como imagenet y se pueden usar como un extractor de características o una herramienta de ajuste. Se pueden usar las capas convolucionales de estos modelos como un extractor de características y

usar una nueva capa completamente conectada para clasificar el nuevo conjunto de datos.

El ajuste fino de un modelo previamente entrenado se puede realizar descongelando las capas del modelo y entrenándolo más en el nuevo conjunto de datos. Esto se puede hacer descongelando las últimas capas del modelo y entrenándolas en el nuevo conjunto de datos. Esto puede generar mejores resultados que entrenar el modelo desde cero en un conjunto de datos más pequeño. Para la investigación se usó el aprendizaje de transferencia y aplicamos un ajuste fino, pero en este caso solamente cambiando el modelo de clasificación con uno que se ajustara al problema en cuestión. Tanto para el problema binario como el problema trinario (Multiclase) Ilustración 20.



Ilustración 20 Descripción grafica de un modelo de transferencia de aprendizaje.
<https://www.bbvaiafactory.com/es/applying-transfer-learning-to-natural-language-models/>

3.9.4 Redes neuronales convolucionales en la medicina.

Las redes neuronales convolucionales (CNN) son un tipo específico de algoritmo de aprendizaje profundo que se ha utilizado ampliamente en diversas aplicaciones médicas, incluido el análisis de imágenes, el soporte de diagnóstico y el análisis de datos [102], [103].

- **Análisis de imágenes médicas:** las CNN se pueden utilizar para analizar imágenes médicas como radiografías, tomografías computarizadas e

imágenes de resonancia magnética. Pueden ser entrenados para clasificar diferentes tipos de imágenes médicas, detectar tumores y segmentar órganos. Por ejemplo, se puede entrenar a una CNN para detectar cáncer de pulmón en tomografías computarizadas aprendiendo a identificar patrones asociados con tumores cancerosos. Otro ejemplo es que se puede entrenar una CNN para segmentar órganos en una imagen de resonancia magnética, que se puede usar para ayudar en la planificación quirúrgica.

- **Registro de imágenes médicas:** las CNN se pueden usar para alinear diferentes imágenes médicas, como imágenes pre y posoperatorias, para mejorar la precisión del diagnóstico médico. Las CNN pueden aprender a hacer coincidir las características de las imágenes y usar esta información para alinear las imágenes, lo que puede ayudar a mejorar la precisión de las intervenciones diagnósticas o terapéuticas.
- **Soporte de diagnóstico médico:** las CNN se pueden usar para ayudar en el diagnóstico de enfermedades como el cáncer de piel, el cáncer de pulmón y la retinopatía diabética. Las CNN pueden entrenarse para clasificar imágenes de lesiones cutáneas como benignas o malignas, por ejemplo, lo que puede ayudar a reducir la cantidad de biopsias innecesarias y mejorar la precisión de los diagnósticos.
- **Análisis de datos médicos:** las CNN se han utilizado para analizar registros de salud electrónicos y otros datos médicos para identificar patrones que pueden usarse para predecir los resultados de los pacientes. Por ejemplo, se puede entrenar una CNN para predecir la probabilidad de que un paciente desarrolle una determinada afección en función de su información demográfica, resultados de laboratorio y otros datos.
- **Generación de imágenes médicas:** las CNN se pueden usar para generar imágenes médicas sintéticas, como tomografías computarizadas, que se pueden usar para aumentar el conjunto de datos de entrenamiento o para simular escenarios médicos. Esto puede ser útil para entrenar modelos en condiciones raras donde hay datos limitados disponibles.

- TESIS TESIS TESIS TESIS TESIS
- **Procesamiento de lenguaje natural médico (NLP):** las CNN también se utilizan para procesar datos de texto médico de registros de salud electrónicos, notas clínicas y otros documentos médicos. Se pueden usar para tareas como el reconocimiento de entidades nombradas, la extracción de relaciones y el análisis de sentimientos.

Es importante tener en cuenta que el uso de CNN en medicina requiere una gran cantidad de datos médicos anotados para entrenar los modelos de manera efectiva y garantizar que los modelos sean sólidos y generalizables. Además, es importante tener en cuenta las consideraciones éticas y legales, como la privacidad de los datos y el consentimiento informado, cuando se trabaja con datos médicos.

En la mamografía, las CNN pueden ayudar a identificar los cánceres de mama que los radiólogos pueden pasar por alto. Las CNN también pueden ayudar a reducir la cantidad de resultados falsos positivos, lo que puede dar lugar a biopsias innecesarias.

Las razones por las que las CNN son buenas para la clasificación de mamografías son:

- Las CNN están diseñadas para procesar y analizar datos de imágenes, lo que las hace muy adecuadas para el análisis de mamografías.
- Las CNN pueden aprender relaciones complejas entre las características de la imagen y el diagnóstico de cáncer y pueden usar esta información para hacer predicciones precisas.
- Las CNN pueden manejar grandes cantidades de datos, lo que las hace adecuadas para su uso con grandes conjuntos de datos de imágenes de mamografía.

Sin embargo, es importante tener en cuenta que, si bien las CNN han mostrado resultados prometedores en la clasificación de mamografías, aún no reemplazan a los radiólogos humanos y deben usarse junto con la experiencia humana. Se necesita más investigación y desarrollo para mejorar la precisión y confiabilidad de las CNN en el análisis de imágenes médicas.

3.10 Arquitecturas de redes neuronales convolucionales

Teniendo los conceptos básicos de las redes neuronales convolucionales ahora se sabe que la diferencia entre una red artificial tradicional a una CNN son las capas que se encargarán de extraer las características esenciales de las imágenes las cuales serán analizadas y clasificadas posteriormente. Con esta información es sencillo crear una propia CNN y usarla para un proyecto. Aunque el trabajo en modificar sus hiperparámetros puede ser largo y extenuante, por lo que se decidió realizar un proceso conocido como transferencia de aprendizaje donde se utilizan CNN pre-entrenadas anteriormente y son utilizadas con cambios menores al problema que se esté utilizando en cuestión.

Las siguientes CNN preentrenadas fueron las que se utilizaron para esta tarea de clasificación de mamografías: AlexNet, MobileNet, MobileNetV3 Small, MobileNetV3 Large, VGGNet, ResNet, Inception, DenseNet, EfficientNet.

3.10.1 AlexNet

AlexNet es una red neuronal convolucional (CNN) que fue desarrollada por Alex Krizhevsky, Ilya Sutskever y Geoffrey Hinton en 2012 [32]. Fue el primer modelo de aprendizaje profundo en ganar el ImageNet Large Scale Visual Recognition Challenge (ILSVRC) en 2012, logrando un top -5 tasa de error del 15,3 %, lo que supuso una mejora significativa con respecto a los modelos de última generación anteriores.

La arquitectura de AlexNet consta de varias capas convolucionales y de agrupación, seguidas de varias capas completamente conectadas. Tiene un total de 8 capas, con 5 capas convolucionales y 3 capas completamente conectadas.

Las capas convolucionales se utilizan para extraer características de las imágenes de entrada y las capas de agrupación se utilizan para reducir la resolución espacial de los mapas de características. Las capas totalmente conectadas se utilizan para clasificar las imágenes en función de las características extraídas. Una de las innovaciones clave de AlexNet fue el uso de la Unidad Lineal Rectificada

(ReLU) como función de activación de las capas ocultas, lo que mejoró la velocidad de entrenamiento y el rendimiento del modelo Ilustración 21.

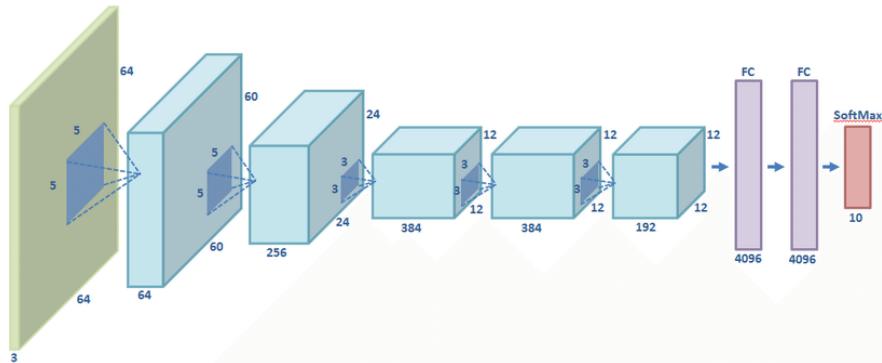


Ilustración 21 Topología de la red neuronal convolucional: AlexNet.
https://www.researchgate.net/figure/Scheme-of-the-AlexNet-network-used_fig1_320052364

3.10.2 MobileNet

MobileNet es una red neuronal convolucional (CNN) liviana desarrollada por Google en 2017 [27]. Su objetivo principal es crear un modelo que sea lo suficientemente pequeño como para ejecutarse en dispositivos móviles e integrados, al tiempo que logra un buen rendimiento en las tareas de clasificación de imágenes.

MobileNet se basa en el concepto de convoluciones separables en profundidad, que pueden reducir en gran medida la cantidad de parámetros y el costo computacional del modelo. En una convolución separable en profundidad, la operación de convolución se divide en dos etapas: una convolución en profundidad que aplica un solo filtro a cada canal de entrada y una convolución en punto que combina la salida de la convolución en profundidad. Esta separación permite un uso mucho más eficiente de la computación y reduce el número de parámetros.

La arquitectura de MobileNet consiste en una serie de capas de convolución separables en profundidad, seguidas de capas de convolución en puntos y agrupación de promedio global. También utiliza la convolución separable de Depthwise en lugar de la convolución tradicional, lo que ayuda a reducir la cantidad de parámetros y cálculos.

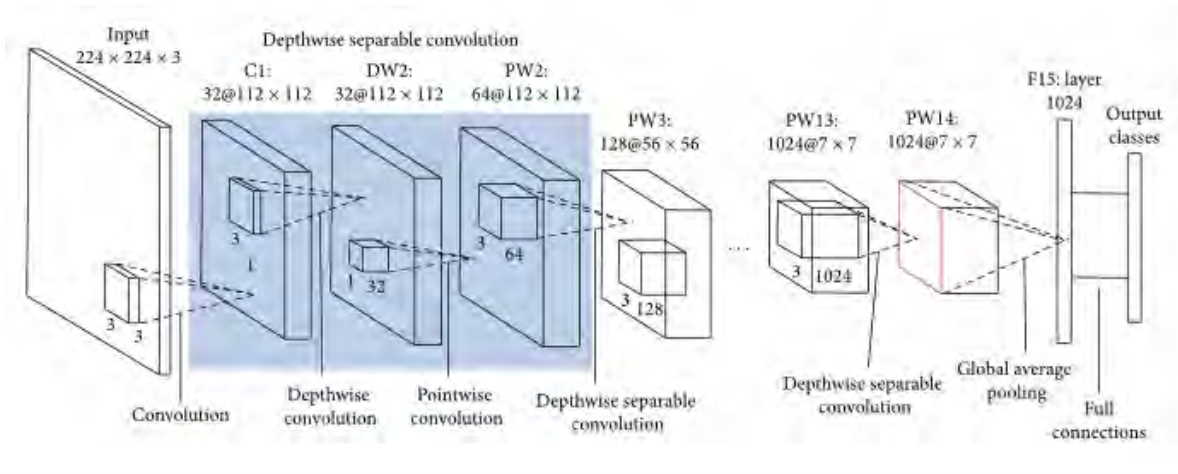


Ilustración 22 Topología de la red neuronal convolucional: MobileNet. <https://wikidocs.net/165429>

Los experimentos en el conjunto de datos de ImageNet muestran que las MobileNet logran una alta precisión con hasta 50 veces menos cálculos y una reducción de 4 veces en el uso de la memoria en comparación con las CNN tradicionales. El documento también demuestra que las redes móviles son adecuadas para aplicaciones de visión móvil en tiempo real, ya que pueden ejecutarse de manera eficiente en dispositivos móviles con recursos informáticos limitados. Ilustración 22.

3.10.3 MobileNetV3 Small y Large

MobileNetV3 es la última versión de la serie MobileNet y Google la presentó en 2019 [26]. El objetivo principal de MobileNet V3 es mejorar la eficiencia y la precisión de las versiones anteriores de MobileNet, sin dejar de ser adecuado para dispositivos móviles e integrados. MobileNetV3 viene en dos versiones: MobileNetV3-Small y MobileNetV3-Large. Ambas versiones utilizan una nueva estructura de red denominada "MobileNetV3 Block", que mejora la eficiencia y la precisión del modelo. Ilustración 23.

El documento explora diferentes opciones arquitectónicas y principios de diseño, como el uso de bloques de construcción más eficientes y la incorporación de técnicas de búsqueda de arquitectura neuronal (NAS). Estas mejoras dan como resultado que MobileNetV3 brinde una mayor precisión y una mayor eficiencia en comparación con las versiones anteriores de MobileNet y otros modelos de última generación. Se evalúa MobileNetV3 en varios conjuntos de datos de referencia y demuestra su superioridad en términos de precisión y eficiencia. Además, el documento muestra que MobileNetV3 se puede aplicar a una amplia gama de tareas de visión móvil, como la clasificación de imágenes y la detección de objetos.

La versión MobileNetV3-Small está diseñada para ser más eficiente y tiene un tamaño de modelo más pequeño, mientras que la versión MobileNetV3-Large está diseñada para ser más precisa y tiene un tamaño de modelo más grande.

La versión MobileNetV3-Small tiene 3,4 millones de parámetros, que es más pequeña que otras versiones como V1 y V2. También logra una mayor precisión que MobileNetV2 en las tareas de clasificación de imágenes.

La versión MobileNetV3-Large tiene 9,2 millones de parámetros y logra una mayor precisión que la versión V3-Small y otros modelos como V1 y V2.

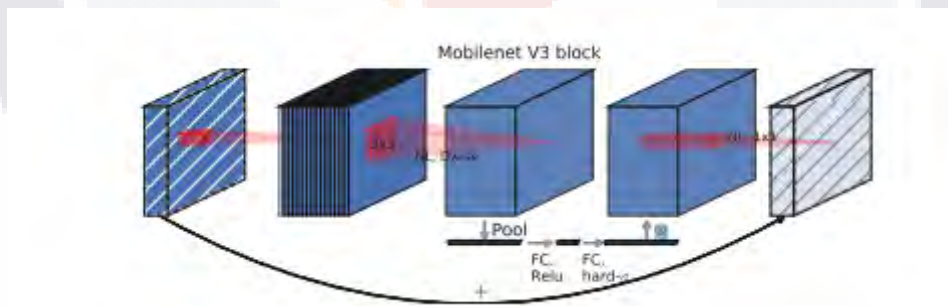


Ilustración 23 Bloque de MobileNet V3. <https://medium.com/ai/C2%B3-theory-practice-business/nextgen-mobilenetv3-definitions-243e396a4c9a>

3.10.4 VGGNet

VGGNet es una red neuronal convolucional profunda (CNN) desarrollada por Visual Geometry Group (VGG) de la Universidad de Oxford en 2014 [29]. La arquitectura

VGGNet es conocida por su alta precisión y solidez en las tareas de clasificación de imágenes.

Los autores muestran que aumentar la profundidad de la red puede mejorar la precisión, pero también presenta desafíos en términos de capacitación y eficiencia computacional. Para superar estos desafíos, los autores proponen varias innovaciones arquitectónicas, como el uso de filtros muy pequeños y operaciones de agrupación para reducir el cálculo, y el uso de conexiones residuales para facilitar la optimización.

La arquitectura de VGGNet consta de varias capas convolucionales y de agrupación, seguidas de varias capas completamente conectadas. Se caracteriza por su arquitectura profunda, con múltiples capas convolucionales apiladas una encima de otra, y el uso de pequeños filtros convolucionales, con un tamaño de campo receptivo de 3x3. También utiliza una técnica llamada "redes convolucionales muy profundas" (VDCNs) las cuales se caracterizan por su profundidad, esta profundidad permite que la red extraiga características en diferentes niveles de abstracción y mejora el rendimiento del modelo.

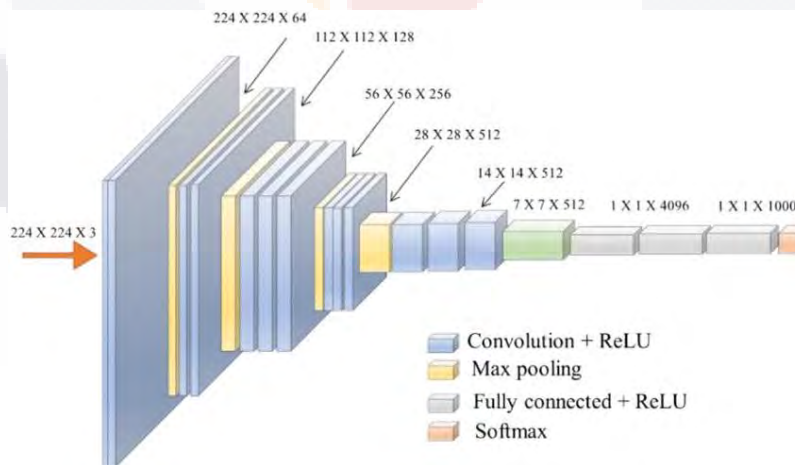


Ilustración 24 Topología de la red neuronal convolucional: VGGNet. <https://ai.plainenglish.io/vggnet-with-tensorflow-transfer-learning-with-vgg16-included-7e5f6fa9479a>

La arquitectura VGGNet utiliza capas de agrupación máximas para reducir la resolución espacial de los mapas de características y reducir el costo computacional del modelo. También utiliza ReLU (Unidad Lineal Rectificada) como función de activación de las capas ocultas. Ilustración 24.

3.10.5 ResNet

He et al. [28] presentan un marco de aprendizaje residual. ResNet (Residual Network) es una arquitectura de red neuronal convolucional profunda (CNN) desarrollada por Microsoft Research Asia. La idea principal detrás de la arquitectura ResNet es utilizar una técnica llamada "aprendizaje residual" para aliviar el problema de los gradientes que se desvanecen en las CNN profundas. En las CNN profundas tradicionales, a medida que aumenta el número de capas, los gradientes tienden a ser cada vez más pequeños, lo que dificulta el aprendizaje de la red. ResNet aborda este problema mediante la introducción de una conexión de acceso directo, o una conexión "residual", que permite que los gradientes pasen por alto una o más capas. Esto permite que la red aprenda de manera más eficiente y efectiva.

Propusieron varios modelos de ResNet, como ResNet-34 A, ResNet-34 B, ResNet-34 C, ResNet-50, ResNet-101 y ResNet-152. Un conjunto de estas redes residuales logra un error del 3,57% en el conjunto de prueba de ImageNet. El resultado de estos modelos ganó el 1er lugar en la tarea de clasificación ILSVRC 2015. ResNet-50 y ResNet-152 son los modelos utilizados para las pruebas que hicimos.

ResNet 152 es una extensión de ResNet 50, que tiene 152 capas. Al igual que ResNet 50, también se entrenó en el conjunto de datos de ImageNet. ResNet 152 tiene las mismas ventajas que ResNet 50, pero también tiene algunas ventajas adicionales: tiene incluso más capas que ResNet 50, lo que le permite aprender funciones más complejas y lograr un mejor rendimiento.

- Se ha entrenado previamente en un gran conjunto de datos, lo que lo convierte en un buen punto de partida para muchas tareas de clasificación de imágenes.
- También se puede ajustar en un conjunto de datos diferente con menos datos, ya que ya ha aprendido características sólidas del conjunto de datos de ImageNet.
- Tiene una arquitectura más profunda que ResNet 50, lo que le permite aprender funciones más complejas y lograr un mejor desempeño en ciertas tareas.
- Vale la pena señalar que cuanto más profunda sea la red, más poder de cómputo se necesitará para entrenar, y también será más propensa al sobreajuste, por lo que es importante utilizar técnicas como el aumento de datos, el abandono y la regularización para evitarlo.

En resumen, ResNet 152 es una extensión de ResNet 50, tiene 152 capas y se entrenó en el conjunto de datos de ImageNet, al igual que ResNet 50, tiene una arquitectura relativamente simple y utiliza conexiones residuales para aliviar el problema de los gradientes que se desvanecen en las CNN profundas. Es ampliamente utilizado y está bien documentado, lo que lo convierte en un buen punto de partida para muchas tareas de clasificación de imágenes. Sin embargo, tiene la ventaja adicional de tener más capas, lo que le permite aprender funciones más complejas y lograr un mejor rendimiento. Ilustración 25.

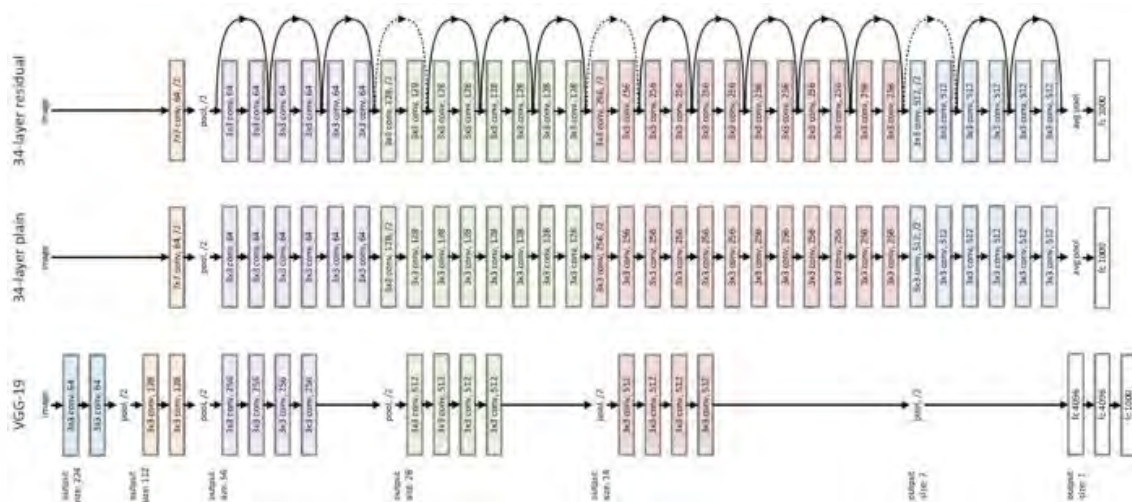


Ilustración 25 Grafica de toda la estructura del modelo ResNet. <https://towardsdatascience.com/review-resnet-winner-of-ilsvrc-2015-image-classification-localization-detection-e39402bfa5d8>

3.10.6 Inception-V3

Szegedy et al. [31] presentan un modelo de arquitectura de 48 capas de profundidad llamado Inception-V3, un modelo de reconocimiento de imágenes que ha demostrado alcanzar una precisión superior al 78,1% en el conjunto de datos de ImageNet.

Inception v3 es una arquitectura de red neuronal convolucional profunda (CNN) que fue desarrollada por Google en 2015. Es una extensión de la arquitectura Inception, que se introdujo en 2014. La idea principal detrás de la arquitectura Inception es usar múltiples capas convolucionales paralelas con diferentes tamaños de filtro, llamados "módulos de inicio", para extraer características a diferentes escalas. Inception v3 es una CNN profunda que utiliza una combinación de capas convolucionales, capas de agrupación máxima y módulos de Inception para extraer características de una imagen de entrada. La arquitectura se divide en varias partes, denominadas "bloques Inception", cada una de las cuales contiene varios módulos Inception. Cada módulo de Inception contiene múltiples capas convolucionales paralelas con diferentes tamaños de filtro y diferentes pasos, lo que permite que la red extraiga características a diferentes escalas. Ilustración 26.

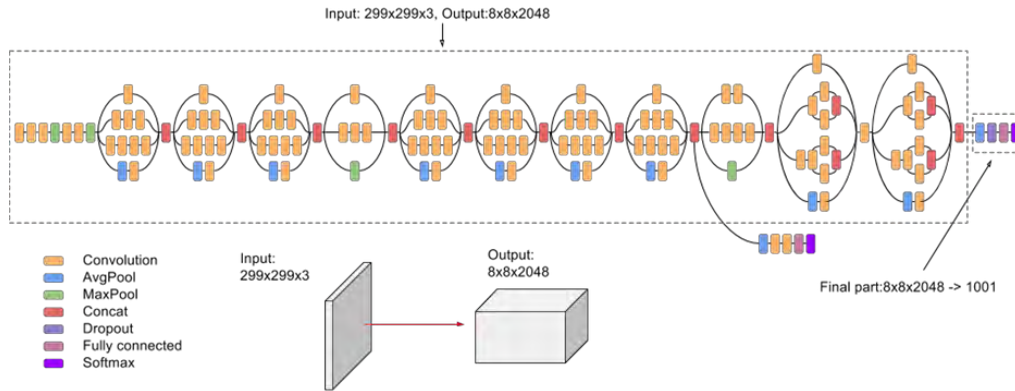


Ilustración 26 Topología de la red neuronal convolucional: Inception-V3.
<https://cloud.google.com/tpu/docs/inception-v3-advanced?hl=es-419>

3.10.7 DenseNet

Huang et al. [30] presentan un nuevo modelo llamado Red Convolucional Densa (DenseNet). Esta arquitectura conecta cada capa con todas las demás capas en forma de avance. Esta arquitectura propuesta se evalúa en cuatro tareas comparativas de reconocimiento de objetos altamente competitivas, como CIFAR-10, CIFAR-100, SVHN e ImageNet.

DenseNet (redes convolucionales densamente conectadas) es una arquitectura de red neuronal convolucional profunda (CNN) desarrollada por Gao Huang et al. en 2016. La idea principal detrás de DenseNet es utilizar conexiones densas, que son conexiones entre todas las capas de la red, en lugar del enfoque tradicional de conectar solo las capas adyacentes. Esto permite que la red propague la información de manera más eficiente y efectiva, y también ayuda a mitigar el problema de los gradientes que se desvanecen. Ilustración 27.

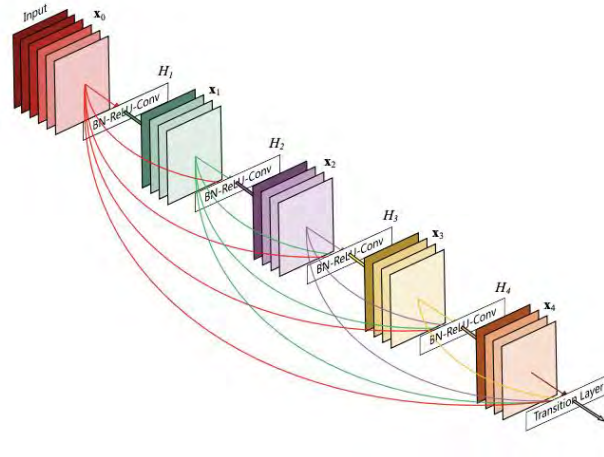


Ilustración 27 Estructura de la formación de las capas de DenseNet.
https://pytorch.org/hub/pytorch_vision_densenet/

DenseNet se compone de bloques densos, que son bloques de capas que están densamente conectadas entre sí. Cada bloque denso contiene varias capas convolucionales, seguidas de normalización por lotes y activación de ReLU. La salida de cada capa se concatena con la entrada de la siguiente capa, lo que permite que la red propague la información de manera más eficiente. Entre los bloques densos, se utilizan capas de transición para reducir la cantidad de mapas de características y las dimensiones espaciales de los mapas de características.

3.10.8 EfficientNet

Tan et al. [33] introdujeron una familia de modelos llamada EfficientNets. EfficientNet es uno de los modelos eficientes que puede lograr una precisión de última generación en ImageNet y tareas de aprendizaje de transferencia de clasificación de imágenes de uso común. Esta arquitectura se formó aprovechando una búsqueda de arquitectura neuronal multiobjetivo que optimiza la precisión y los FLOPS. Ilustración 28.

EfficientNet es una familia de arquitecturas de redes neuronales convolucionales (CNN) que fue desarrollada por Google AI en 2019. La idea principal detrás de EfficientNet es mejorar la eficiencia de las CNN escalando la

arquitectura de la red y la resolución de la imagen de entrada mientras se mantiene un nivel similar de precisión. EfficientNet utiliza una técnica llamada "escalado compuesto" para escalar la arquitectura de la red, lo que implica escalar la profundidad, el ancho y la resolución de la red simultáneamente. La profundidad, el ancho y la resolución de la red están controlados por tres coeficientes de escala: `profundidad_coeficiente`, `ancho_coeficiente` y `resolución_coeficiente`.

Además del escalado compuesto, EfficientNet también utiliza otras técnicas para mejorar la eficiencia, como:

- AutoML: esta es una técnica que encuentra automáticamente la arquitectura de red óptima para un conjunto de datos y una tarea determinados.
- Convolución separable en profundidad: esta es una técnica que separa la operación de convolución estándar en una convolución en profundidad y una convolución en punto, lo que reduce los requisitos de computación y memoria de la red.
- Bloques de compresión y excitación (SE): estos bloques se utilizan para recalibrar de manera adaptativa las respuestas de características de canal de una capa convolucional mediante el uso de información global.

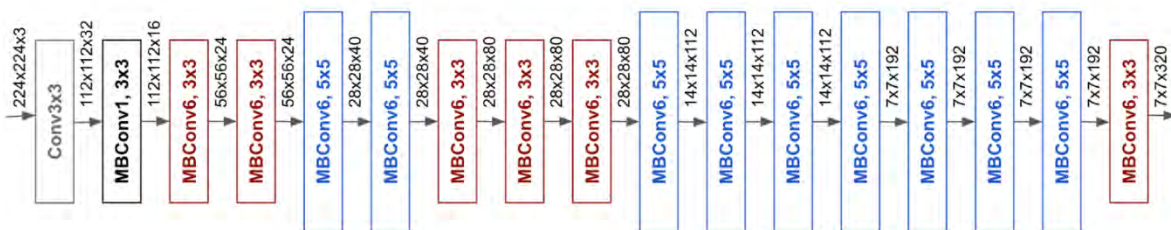


Ilustración 28 Topología de la red neuronal convolucional: EfficientNet.
<https://ai.googleblog.com/2019/05/efficientnet-improving-accuracy-and.html>

3.11 Métricas de evaluación.

Para lograr una labor de análisis y comprensión de cada método que se utilizó y hacer una evaluación de las soluciones obtenidas, es necesario usar diferentes métricas de evaluación. A continuación, se hará una descripción de las que se estuvieron usando en este trabajo.

Para llevar a cabo la evaluación con estas métricas, era necesario que el conjunto de datos estuviera etiquetado previamente y, lo más importante, que fuera de manera correcta. En esta etapa, se llevaba a cabo una clasificación en la que cada instancia del conjunto de datos tenía un etiquetado con un valor real, y luego, mediante cualquier algoritmo utilizado previamente, se obtenía el valor predicho. Esto permitía conocer qué tan bueno era el modelo para hacer una clasificación. En primer lugar, se determinaba cuántos valores el clasificador había identificado correctamente y cuáles no, dividiéndolos en las dos categorías previamente mencionadas: los valores reales y los valores predichos, donde se tenía la categoría positiva P y la categoría negativa N [104].

Así los valores que se pueden obtener pueden ser los cuatro siguientes:

- **TP (True positive):** El valor de etiquetado real y el valor predicho por el clasificador son ambos en este caso positivos.
- **FP (False negative):** El valor del etiquetado real es negativo, pero el valor predicho por el clasificador fue positivo.
- **TN (True negative):** El valor del etiquetado real y el valor predicho por el clasificador son ambos en este caso negativos.
- **FN (False negative):** El valor del etiquetado real es positivo, pero el valor predicho por el clasificador fue negativo.

El clasificador realizaba dos tipos de predicciones, de las cuales dos eran correctas y las otras dos eran errores. La métrica presentada en la siguiente imagen mostraba una matriz de confusión que permitía observar cómo estaban acomodados los valores anteriormente mencionados. Ilustración 29.

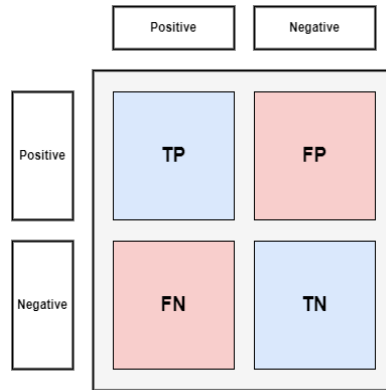


Ilustración 29 Matriz de confusión estructura.

En esta investigación, no solo se limitó el problema a una clasificación binaria, sino que también se buscó la clasificación en tres clases. En este caso, la matriz de confusión fue algo diferente a cómo sería para dos categorías. En la Ilustración 30 se puede observar un ejemplo de cómo se tendría que hacer el cálculo para cada clase.

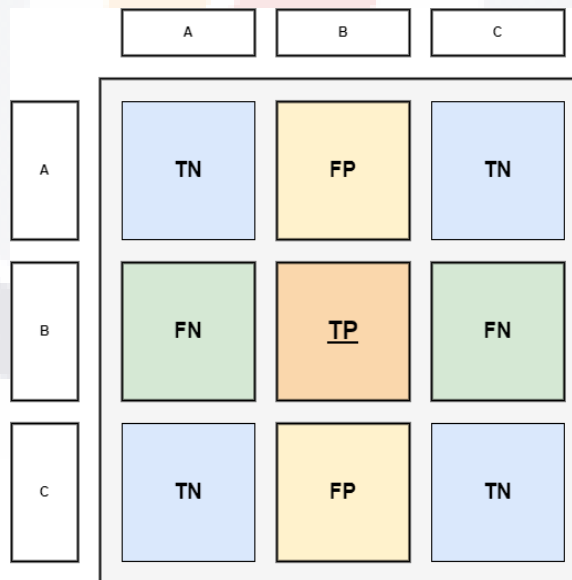


Ilustración 30 Matriz de confusión para problemas de varias categorías.

- **Exactitud:** La precisión es la relación entre el número de predicciones correctas realizadas por el modelo y el número total de entradas.

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

- **Precisión:** La precisión es el indicador que muestra la calidad de una predicción positiva realizada por el modelo.

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

- **Sensibilidad (Recall):** La sensibilidad mide la capacidad del modelo para detectar muestras positivas.

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

- **F1-score:** El F1-score es una combinación de precisión y recall. Esta métrica es buena para modelos que no están equilibrados en la misma cantidad de datos de cada clase.

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

- **Curva ROC y AUC:** La curva ROC (Receiver Operating Characteristic) es una representación gráfica que muestra el desempeño del modelo de clasificación en todos los umbrales de la clasificación. Esta curva muestra dos parámetros: La sensibilidad y la especificidad. La AUC (Area under the ROC Curve) es una representación numérica del desempeño de la clasificación. Ilustración 31.

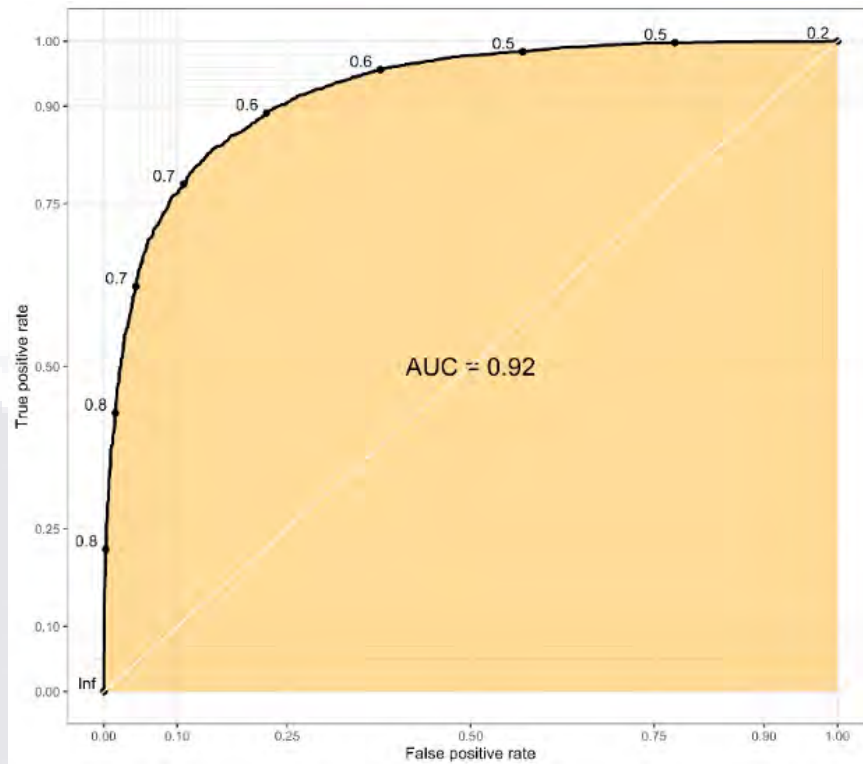


Ilustración 31 Grafica ROC con la métrica AUC.

<https://sinyi-chou.github.io/classification-auc/>

3.12 Bases de datos utilizadas

Para obtener un diagnóstico preciso en el problema de clasificación de mamografías, el preprocesamiento de datos es una de las tareas más importantes. Sin embargo, antes de eso, es necesario obtener los datos de un repositorio confiable para obtener resultados precisos. En este sentido, se seleccionaron dos bases de datos ampliamente utilizadas en el estado del arte: Mini-MIAS y CBIS-DDSM.

3.12.1 Mini-MIAS

La Sociedad de Análisis de Imágenes de Mamografía (MIAS) es un grupo de investigación británico interesado en comprender la mamografía y ha creado una base de datos de mamografía digital. La película tomada del Programa Nacional de

TESIS TESIS TESIS TESIS TESIS

Detección de Senos del Reino Unido ha sido digitalizada a 50 micrones. Hay 322 películas digitalizadas en la base de datos. También incluye las marcas "verdaderas" del radiólogo en cualquier área inusual que pueda estar presente. La base de datos se reduce a un borde de píxeles de 200 micras y se rellena/recorta para que todas las imágenes sean de 1024 x 1024.

Las imágenes en la base de datos Mini-MIAS son de calidad variable y algunas imágenes contienen artefactos como desenfoque de movimiento y ruido. Las imágenes también contienen una variedad de anomalías mamarias, incluidas masas, calcificaciones y asimetrías. Cada imagen está etiquetada con el tipo y la ubicación de la anomalía mamaria, así como la edad y el estado menopáusico de la paciente. La base de datos Mini-MIAS se usa ampliamente en la investigación y el desarrollo de sistemas CAD para la detección del cáncer de mama, ya que proporciona un conjunto diverso de imágenes con una amplia gama de anomalías mamarias y calidad de imagen. La base de datos también es útil para evaluar el rendimiento de diferentes técnicas de procesamiento de imágenes y algoritmos de aprendizaje automático para la detección del cáncer de mama [21]. Ilustración 32.

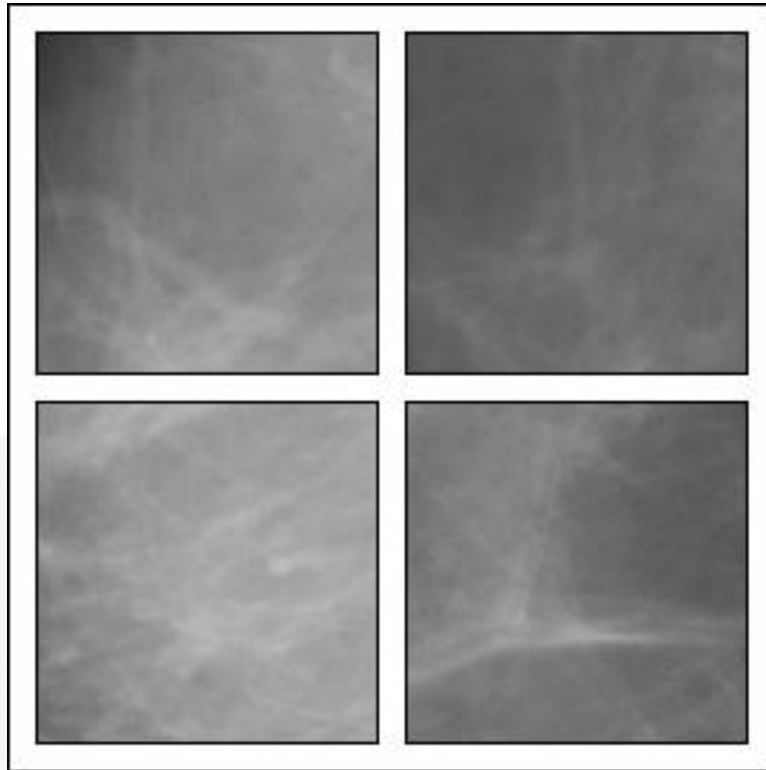


Ilustración 32 Imágenes de Mini-MIAS recortadas.

3.12.2 CBIS-DDSM

El subconjunto de imágenes curadas de mama de DDSM es una versión mejorada de DDSM (CBIS-DDSM). Radiólogos con capacitación especializada han seleccionado y curado esta muestra de imágenes del conjunto de datos DDSM original. Para estas imágenes comprimidas se ha utilizado el formato DICOM estándar. DICOM es un estándar mundial para almacenar imágenes médicas como tomografías computarizadas (CT) e imágenes de resonancia magnética (MRI).

Las imágenes en la base de datos CBIS-DDSM incluyen casos normales y anormales, con una amplia gama de anomalías mamarias, incluidas masas, calcificaciones y asimetrías. Cada imagen está etiquetada con el tipo y la ubicación de la anomalía mamaria, así como la edad y el estado menopáusico de la paciente. Las anotaciones han sido realizadas por radiólogos expertos en mamografía.

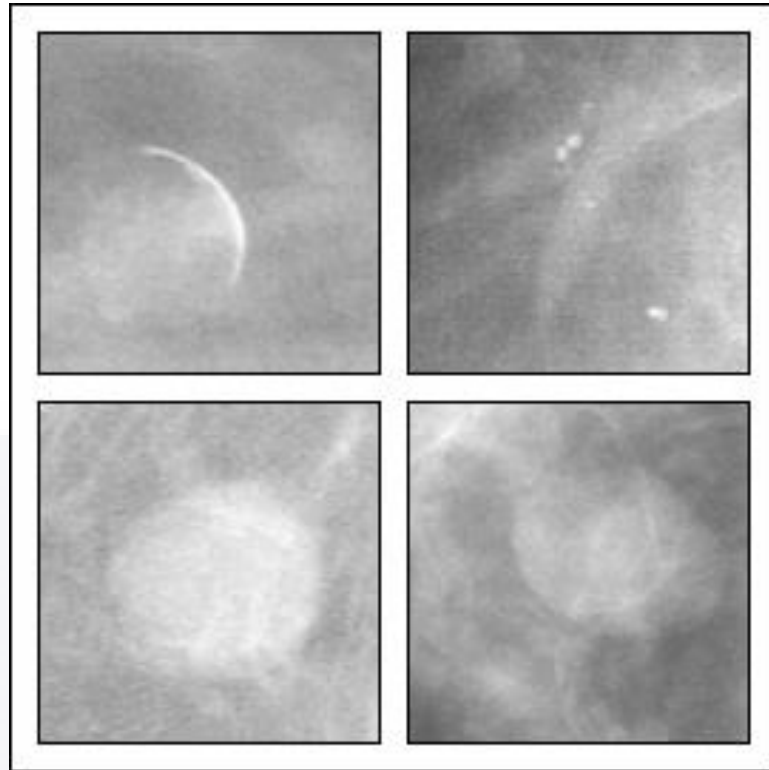


Ilustración 33 Imágenes de CBIS-DDSM recortadas.

La base de datos CBIS-DDSM se usa ampliamente en la investigación y el desarrollo de sistemas CAD para la detección del cáncer de mama, ya que proporciona un conjunto diverso de imágenes con una amplia gama de anomalías mamarias y calidad de imagen. La base de datos también es útil para evaluar el rendimiento de diferentes técnicas de procesamiento de imágenes y algoritmos de aprendizaje automático para la detección del cáncer de mama.

Una de las ventajas de la base de datos CBIS-DDSM es que incluye imágenes DM y SFM, lo que permite a los investigadores evaluar el rendimiento de los sistemas CAD en diferentes modalidades. Además, la base de datos es relativamente grande, lo que permite una evaluación más sólida del rendimiento de diferentes algoritmos [20]. Ilustración 33.

3.13 Especificaciones de hardware

En esta investigación, se utilizó Python como lenguaje de programación y se requirió una estación de trabajo con especificaciones técnicas específicas. La estación de trabajo utilizada contaba con una tarjeta gráfica NVIDIA Geforce RTX 3070ti de 8GB, un procesador Ryzen R9 5900x, 32GB de RAM DDR4 3200Mhz y un SSD XPG Spectrix de 512GB.

3.14 Especificaciones de software

Después de haber mencionado las especificaciones del hardware, se utilizaron varios lenguajes de programación y librerías para lograr los objetivos de la investigación.

3.14.1 Python

Creado por Guido Van Rossum a principios de los años 90. Python es un lenguaje de alto nivel diseñado para ser legible, reutilizable y mantenible.

Python es un lenguaje de programación de alto nivel que se usa ampliamente para el desarrollo web, el análisis de datos, la inteligencia artificial y la computación científica. Algunas de las características clave del lenguaje Python incluyen su sintaxis simple y fácil de leer, su amplia biblioteca estándar y su soporte para múltiples paradigmas de programación.

Una de las principales ventajas del lenguaje Python es su legibilidad y simplicidad. La sintaxis de Python está diseñada para ser intuitiva y fácil de entender, lo que la convierte en una excelente opción tanto para principiantes como para desarrolladores experimentados. El lenguaje también tiene una comunidad grande y activa que ha desarrollado una amplia gama de bibliotecas y marcos, lo que facilita la realización de muchas tareas sin tener que escribir código complejo. Otra ventaja de Python es su versatilidad. Python se puede utilizar para una amplia gama de tareas, desde desarrollo web y análisis de datos hasta inteligencia artificial y computación científica. El lenguaje también admite múltiples paradigmas de programación, incluida la programación orientada a objetos, funcional y de

procedimientos. Esto lo convierte en una excelente opción para una amplia gama de proyectos y casos de uso [105].

Python también tiene una comunidad grande y activa. Esto significa que hay una gran cantidad de información, tutoriales y bibliotecas disponibles para los desarrolladores. Además, la comunidad es muy activa en el desarrollo de nuevas bibliotecas y herramientas que facilitan el uso de Python para diferentes tareas.

Además, Python tiene un gran soporte para la ciencia de datos y el aprendizaje automático, tiene muchas bibliotecas como Numpy, pandas, scikit-learn y Tensorflow que se pueden usar para realizar tareas de análisis de datos y aprendizaje automático. Ilustración 34.



Ilustración 34 Logo Python (https://es.wikipedia.org/wiki/Historia_de_Python).

3.14.2 Tensorflow

Biblioteca diseñada por Google de código abierto para el desarrollo de sistemas de aprendizaje automático. Es una biblioteca matemática simbólica que utiliza flujo de datos y programación diferenciable.

TensorFlow es una biblioteca de software de código abierto para el aprendizaje automático desarrollada por Google. Se utiliza para una amplia gama

de tareas, incluida la creación y el entrenamiento de redes neuronales para tareas como el reconocimiento de imágenes y voz, el procesamiento del lenguaje natural y el análisis predictivo. TensorFlow proporciona un conjunto de herramientas para crear, entrenar e implementar modelos de aprendizaje automático, incluida una API de bajo nivel para crear gráficos computacionales y una API de alto nivel para crear modelos utilizando capas preexistentes.

Uno de los principales beneficios de TensorFlow es su flexibilidad. TensorFlow permite a los desarrolladores definir gráficos de cálculo personalizados, lo que hace posible implementar capas personalizadas, funciones de pérdida y optimizadores. Esta flexibilidad permite a los desarrolladores crear y experimentar con modelos complejos, como redes neuronales profundas, con relativa facilidad. Además, TensorFlow se puede usar tanto para investigación como para producción, ya que ha sido probado en batalla en muchas aplicaciones de aprendizaje automático a gran escala [106]. Ilustración 35.



Ilustración 35 Tensorflow logo.

3.14.3 Keras

Keras es una API de aprendizaje profundo de alto nivel desarrollado por Google para la implementación de redes neuronales. La ventaja que tiene Keras es poder realizar los prototipos de manera rápida y sencilla, poder crear la CNN que se necesitan y poder ejecutarlas tanto en el CPU con el GPU. Keras es una API de redes neuronales de alto nivel, escrita en Python y capaz de ejecutarse sobre TensorFlow. Fue desarrollado para permitir la experimentación rápida con redes

neuronales profundas. Keras proporciona una interfaz simple y consistente para la biblioteca TensorFlow, lo que facilita la creación y el entrenamiento de modelos de aprendizaje profundo. Keras abstrae la complejidad de construir una red neuronal profunda, lo que permite a los desarrolladores centrarse en la arquitectura del modelo en lugar de los detalles de la biblioteca subyacente. También proporciona un conjunto de capas y modelos preconstruidos que se pueden integrar fácilmente en un nuevo modelo, lo que puede ahorrar mucho tiempo y esfuerzo [107]. Ilustración 36.

Para obtener resultados más rápidos, se utilizó la GPU para esto. Para usar la GPU, instalamos la biblioteca Deep Neural Network (cuDNN), Compute Unified Device Architecture (CUDA) y Visual Studio 2019 para sacar el mayor provecho de CUDA.



Ilustración 36 Keras logo.

3.15 Otras librerías utilizadas.

3.15.1 Numpy

NumPy es una biblioteca de Python que se utiliza para la computación científica y el análisis de datos. Proporciona una forma poderosa y eficiente de trabajar con grandes arreglos y matrices de datos numéricos. NumPy proporciona una amplia gama de funciones matemáticas, incluidas operaciones de álgebra lineal, transformadas de Fourier y generación de números aleatorios. También proporciona herramientas para la integración con otras bibliotecas populares, como SciPy y Matplotlib. NumPy se usa ampliamente en las comunidades científicas y de ciencia

de datos y es una parte fundamental de muchas otras bibliotecas y marcos [108].
Ilustración 37.



Ilustración 37 Numpy logo.

3.15.2 Pandas

Pandas es una biblioteca de Python que se usa comúnmente para la manipulación y el análisis de datos. Proporciona estructuras de datos y herramientas de análisis de datos fáciles de usar y de alto rendimiento para manejar y manipular tablas numéricas y datos de series temporales. Las dos estructuras de datos principales en Pandas son Series (unidimensional) y DataFrame (bidimensional). Estas estructuras pueden manejar una amplia variedad de tipos de datos, incluidos datos numéricos, de texto y de fecha/hora.

Pandas también proporciona una amplia gama de funciones de manipulación de datos, como fusionar, remodelar y agrupar datos. También tiene poderosas capacidades de preprocesamiento y limpieza de datos, como el manejo de datos faltantes y valores atípicos. Pandas también puede leer y escribir datos en una variedad de formatos de archivo, incluidas bases de datos CSV, Excel y SQL. Es ampliamente utilizado en ciencia de datos y es una parte fundamental de muchas otras bibliotecas y marcos [109]. Ilustración 38.



Ilustración 38 Pandas logo.

3.15.3 Pillow

Pillow es una biblioteca de imágenes de Python de código abierto que permite a los usuarios manejar y manipular archivos de imágenes. Es una bifurcación de Python Imaging Library (PIL) y agrega soporte para Python 3 e introduce nuevas características. Algunas de las características de Pillow incluyen el cambio de tamaño, recorte y rotación de imágenes, mejora de imágenes, filtrado de imágenes y soporte para abrir y guardar una amplia gama de formatos de archivo de imagen, incluidos PNG, BMP, JPEG y TIFF. También se puede utilizar para tareas de procesamiento de imágenes, como convertir entre espacios de color, ajustar el contraste y el brillo de la imagen y aplicar filtros de imagen [110].

3.15.4 OpenCV (Open-Source Computer Vision Library)

OpenCV (Biblioteca de visión artificial de código abierto) es una biblioteca de aprendizaje automático y visión artificial multiplataforma de código abierto. Proporciona una amplia gama de funciones de procesamiento de imágenes y videos, que incluyen detección y reconocimiento de objetos, registro de imágenes, segmentación de imágenes y más.

OpenCV fue desarrollado originalmente por Intel en 1999 y ahora lo mantiene una organización sin fines de lucro. Está escrito en C++, con interfaces disponibles para C++, Python, Java y MATLAB [111]. Ilustración 39.

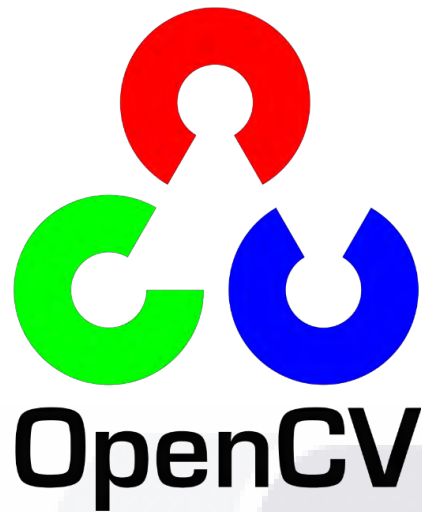


Ilustración 39 OpenCV logo.

3.15.5 Scikit-learn

Scikit-learn, también conocido como sklearn, es una biblioteca de Python para el aprendizaje automático. Proporciona una amplia gama de herramientas para el análisis y modelado de datos, incluidos algoritmos para el aprendizaje supervisado y no supervisado, así como herramientas para la evaluación y selección de modelos.

Scikit-learn se usa ampliamente en una variedad de aplicaciones, incluido el reconocimiento de imágenes y voz, el procesamiento del lenguaje natural y la bioinformática. Está diseñado para ser fácil de usar e integrar con otras bibliotecas, como NumPy y pandas [112]. Ilustración 40.



Ilustración 40 Scikit-learn logo.

3.15.6 Matplotlib

Matplotlib es una biblioteca de visualización de datos de código abierto para Python. Proporciona una API orientada a objetos para incrustar gráficos en aplicaciones utilizando kits de herramientas GUI de uso general como Tkinter, wxPython, Qt o GTK [113]. Ilustración 41.

Algunas de las características clave de matplotlib incluyen:

- Trazado: matplotlib proporciona una amplia gama de opciones de trazado, incluidos gráficos de líneas, gráficos de dispersión, gráficos de barras, histogramas, gráficos 3D y más.
- Personalización: matplotlib permite una personalización completa de los gráficos, incluida la capacidad de cambiar colores, estilos de línea, etiquetas de ejes y más.
- Formato: matplotlib se puede usar para generar gráficos en una variedad de formatos, incluidos PNG, PDF, SVG y JPG.

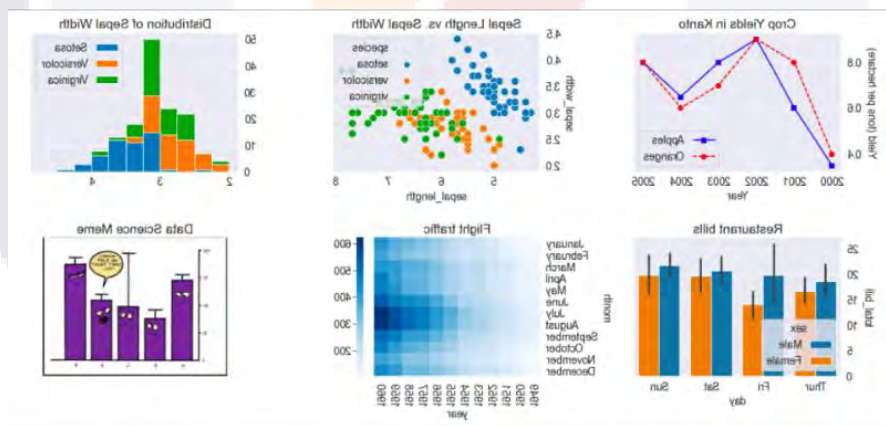


Ilustración 41 Graficas Matplotlib.

3.15.7 Albumentations

Albumentations es una biblioteca Python de código abierto para el aumento de datos de imágenes y videos. Proporciona una interfaz simple y consistente para una amplia gama de técnicas de aumento de datos de imágenes y videos. El aumento

de datos es el proceso de aplicar transformaciones aleatorias a imágenes, como rotación, volteo, escalado y más, para aumentar el tamaño y la diversidad de un conjunto de datos, lo que a su vez puede ayudar a mejorar el rendimiento de los modelos de aprendizaje automático [114].

Algunas de las características clave de Albumentations incluyen:

- Una gran variedad de técnicas de aumento de datos de imágenes y videos: Albumentations admite una amplia gama de técnicas de aumento de datos de imágenes y videos, incluidas transformaciones geométricas, de espacio de color y basadas en la intensidad.
- Interfaz consistente y simple: Albumentations proporciona una interfaz consistente y simple para aplicar técnicas de aumento de datos, lo que facilita su uso e integración con otras bibliotecas.
- Compatibilidad con puntos clave y cuadros delimitadores: Albumentations incluye compatibilidad con puntos clave y cuadros delimitadores, que se pueden utilizar para la detección de objetos y otras tareas.
- Soporte para múltiples backends: Albumentations admite múltiples backends, incluidos OpenCV, Pillow y más, lo que le permite trabajar con una amplia gama de formatos de imagen y video.

3.15.8 Pydicom

Pydicom es un paquete de Python puro para trabajar con archivos DICOM. DICOM (Imágenes digitales y comunicaciones en medicina) es un estándar para manejar, almacenar, imprimir y transmitir información de imágenes médicas. pydicom permite a los desarrolladores leer, escribir, modificar y mostrar archivos DICOM con código Python. También proporciona una interfaz simple e intuitiva para interactuar con archivos DICOM y admite una amplia gama de tipos de datos DICOM y sintaxis de transferencia. Esto lo convierte en una herramienta útil para trabajar con datos de imágenes médicas en una variedad de aplicaciones, como radiología, oncología e investigación médica [115].

CAPITULO 4

El cuarto capítulo de la tesis describe la metodología utilizada para clasificar las mamografías, explicando las técnicas de análisis de imágenes empleadas, la extracción de características, el aprendizaje automático y profundo, así como los algoritmos y modelos específicos utilizados. El objetivo de este capítulo es proporcionar una comprensión completa de la metodología utilizada en la investigación y sentar las bases para la interpretación de los resultados y conclusiones presentadas en capítulos posteriores.

METODOLOGÍA

4.1 Modalidad de la investigación

Esta investigación se basará en una metodología de investigación aplicada. Consiste en el análisis de las imágenes proporcionadas por el especialista médico en este caso sacando esas imágenes de base de datos públicas como Mini-MIAS y CBIS-DDSM para así de ellas poder detectar de una forma automática los diferentes tipos de síntomas representativos de patologías. Una vez diseñados los correspondientes algoritmos que se necesitan se usará para las pruebas correspondientes con las imágenes obtenidas. Las modalidades de investigación que se utilizaran en este proyecto son: bibliográfica, de campo y explicativa.

- La bibliográfica es utilizada ya que se necesita recurrir a fuentes de carácter científico para así poder adquirir la fundamentación teórica.
- De campo ya que el estudio de los hechos es en el lugar donde estarán los acontecimientos.
- Explicativa porque se dará una solución y se comprobará el funcionamiento en este caso del algoritmo a desarrollar el cual pretende resolver el problema.

4.2 Recolección de Información

La recopilación de información es un paso fundamental en el desarrollo de una metodología para la clasificación de las mamografías. Se trata de recopilar datos e información relevante sobre las imágenes de mamografía y las características del cáncer de mama para diseñar y evaluar un modelo de clasificación. La siguiente información generalmente se recopila para la clasificación de la mamografía:

- **Imágenes de mamografía:** se recopila una gran cantidad de imágenes de mamografía, tanto normales como anormales, para entrenar y probar el modelo de clasificación. Estas imágenes deben ser representativas de la población de mamografías e incluir una amplia gama de densidades mamarias, grosor del tejido y condiciones de imagen.
- **Anotación de datos:** las imágenes de mamografía deben incluir anotaciones con información sobre la presencia o ausencia de cáncer de mama, el tipo y la ubicación del cáncer y otra información relevante. Estos datos de anotación se utilizan para entrenar y evaluar el modelo de clasificación.
- **Información del paciente:** la información del paciente, como la edad, el historial médico y el historial familiar, también se puede recopilar y utilizar para mejorar el modelo de clasificación.
- **Preprocesamiento de imágenes:** también se debe recopilar y documentar información sobre los pasos de preprocesamiento de imágenes, como la mejora, la normalización y la segmentación de imágenes. Estos pasos son fundamentales para mejorar la precisión del modelo de clasificación.
- **Métricas de evaluación:** También se deben definir y recopilar las métricas utilizadas para evaluar el rendimiento del modelo de clasificación. Estas métricas, como la precisión, la sensibilidad y la especificidad, ayudan a determinar la eficacia del modelo de clasificación.

4.3 Procesamiento y análisis de los datos

En el proyecto se planea desarrollar diversos algoritmos de tratamiento de imágenes para el problema que se está planteando. Ya cuando las imágenes obtenidas sean procesadas por dichos algoritmos, estos resultados automáticos se compararon por las detecciones manuales del experto en cuestión para saber si el algoritmo desarrollado hace su funcionamiento correctamente.

4.4 Desarrollo del proyecto

- **Análisis de todas las imágenes de la base de datos.**

Para este paso del análisis se realizará una extracción general de las características de las imágenes, el proceso se tratará de realizar diferentes filtros a las imágenes, hacer segmentaciones, disminución de ruido, extraer regiones específicas donde están las anomalías, etc.

- **Definición de todos los algoritmos para la identificación de características al detectar las imágenes.**

En esta parte se plantea escoger el algoritmo óptimo para resolver los procesos de las imágenes analizadas con anterioridad.

- **Evaluar la calidad de la detección automática obtenida al momento de usar el algoritmo.**

Para una evaluación de calidad se necesitarán usar las imágenes que fueron sacadas de las Base de Datos de Mini-MIAS y CBIS-DDSM en este caso.

4.5 Técnica de análisis de Imágenes

Para este proyecto de investigación, era esencial analizar algunas de las técnicas de procesamiento más utilizadas en el campo de la imagenología médica. Se están creando programas y los procesadores se están volviendo más avanzados y rápidos. Como resultado, los hospitales los están empleando con más frecuencia,

ayudando a miles de pacientes que necesitan atención y permitiendo que los médicos brinden una mejor atención a los pacientes que la necesitan [15].

- **Análisis de texturas:** el análisis de texturas es una técnica que se centra en los patrones y texturas dentro de una imagen. Esto se puede usar para identificar la presencia de estructuras anormales, como masas y calcificaciones, que tienen texturas diferentes en comparación con el tejido normal.
- **Segmentación:** la segmentación es el proceso de separar las diferentes regiones de una imagen en función de sus características, como la intensidad y la textura. Esto se puede usar para separar el tejido de la mamografía de otros objetos en la imagen y para identificar las regiones de interés para un análisis posterior.
- **Extracción de características:** la extracción de características es una técnica utilizada para identificar y extraer características relevantes de una imagen que se puede utilizar para clasificar la imagen. Estas características pueden incluir formas, intensidades, texturas y patrones.

4.6 Metodología propuesta

Para estas investigaciones se hicieron dos metodologías y se probaron cada una de ellas para ver los resultados. Esta sección describe los pasos de preprocesamiento de datos, el proceso de selección de datos, la estrategia de aumento de datos y la arquitectura CNN utilizada en la tarea de clasificación utilizada en este estudio.

4.6.1 Metodología 1.

Para la primera metodología se hizo una búsqueda entre diferente tipo de transformaciones para las imágenes mejorando la calidad de las mismas. Y se hicieron pruebas de cada conjunto preprocesado usando cada modelo. Haciendo

una permutación de preprocesamiento y modelo de red neuronal convolucional pre-entrenada.

Los modelos de redes neuronales convolucionales profundas o por sus siglas en inglés (DCNN) requieren una gran cantidad de datos para su entrenamiento para lograr un alto rendimiento. Por esta razón, se tuvo que utilizar una base de datos confiable, ya que es necesario tener diagnósticos precisos para obtener mejores resultados[116]. Para esta investigación se utilizaron las imágenes Mini-MIAS y CBIS-DDSM, Mini-MIAS consta de un total de 322 imágenes, de las cuales 133 presentan anomalías (63 benignas y 51 malignas), y el resto, 208, no presentan anomalías. Por otro lado, CBIS-DDSM es un estudio de mamografía de película escaneada de 2.620. El tamaño de las imágenes de Mini-MIAS es un estándar de 1024x1024, mientras que las imágenes recortadas de CBIS-DDSM que obtuvimos tienen diferentes tamaños. Se investigaron cuatro problemas como biclase y multiclase usando la base de datos previamente. Ilustración 42.

Esta metodología, como se muestra en la Ilustración 35, se ha implementado con diferentes fases involucradas para alcanzar el objetivo planteado. La única diferencia es que CBIS-DDSM no requiere un algoritmo de imagen recortada para obtener las imágenes. Las imágenes de ROI ya se descargaron de los archivos de la página web.

La mamografía asistida por computadora es un problema desafiante que no se puede resolver simplemente mediante la clasificación de imágenes. La razón principal detrás de esto es que las anomalías en una mamografía de tamaño completo de 1024x1024, como Mini-MIAS, o incluso más grandes, como las detectadas en CBIS-DDSM, solo están presentes en un área pequeña de la imagen, y la imagen completa no es requerido. Varios modelos de DCNN requieren escalar la imagen completa a 224x224 píxeles. La mayoría de las DCNN para el análisis de imágenes mamográficas están diseñadas con este tamaño de entrada, que normalmente es de 224x224 o 299x299 píxeles. Si cambia el tamaño de una mamografía completa, corre el riesgo de perder información de las imágenes [117].

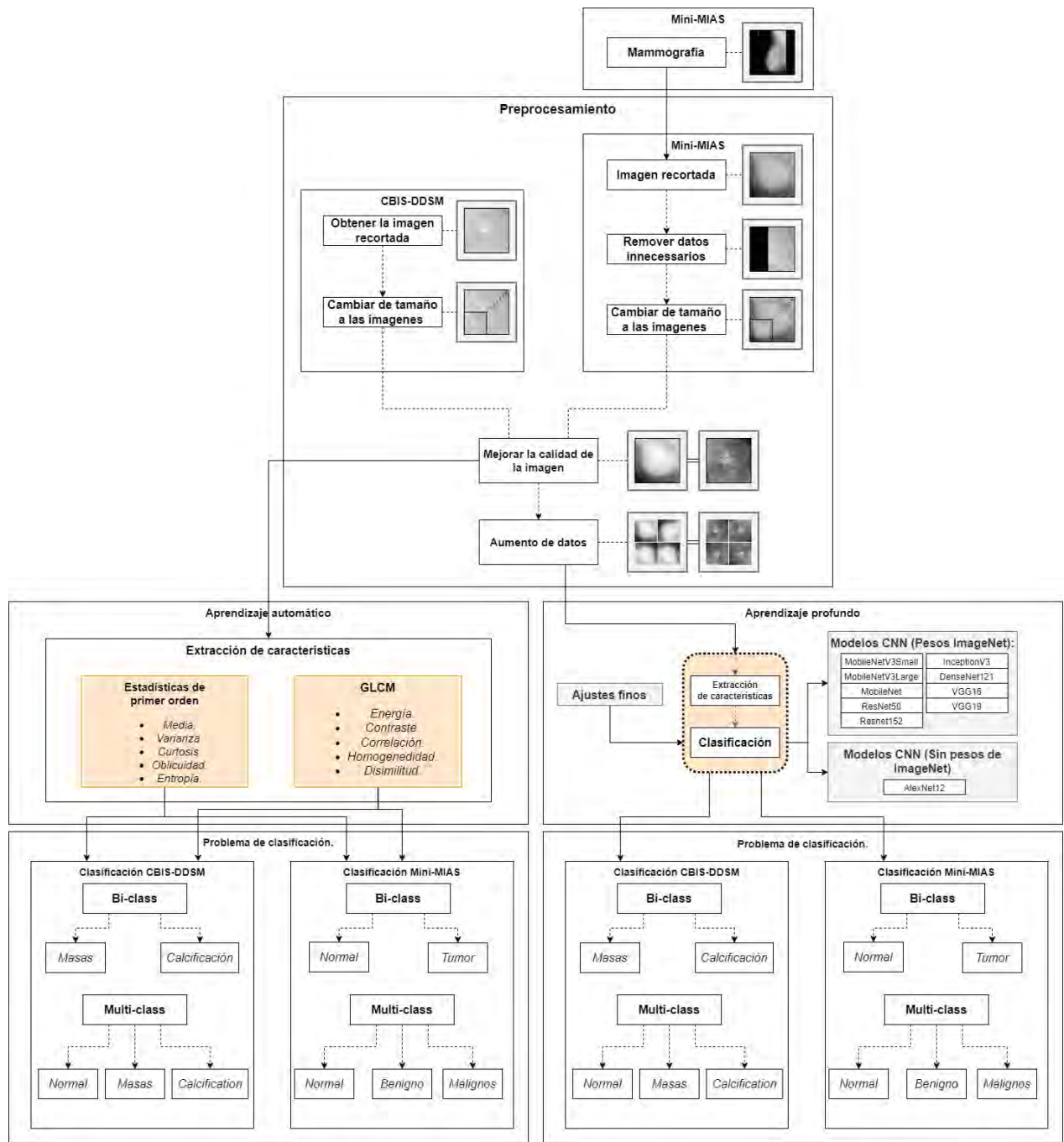


Ilustración 42 Primera metodología utilizando aprendizaje automático y aprendizaje profundo.

4.6.1.1 Preprocesamiento de imágenes.

El conjunto de datos Mini-MIAS fue utilizado en la primera fase del preprocesamiento. Los pasos involucrados en este proceso se muestran en la Figura 1. Inicialmente, todas las imágenes se descargaron del sitio web en formato PGM. En este estudio, se decidió guardar las imágenes en formato PNG debido a que este formato no pierde información y mantiene la calidad de la imagen original. Como las mamografías digitales tienen un tamaño muy grande (1024x1024), se requiere un procesamiento extenso y, en ocasiones, se pueden encontrar secciones innecesarias en ellas. Afortunadamente, el sitio web de Mini-MIAS proporciona información completa para identificar y localizar las anomalías, pero era necesario implementar una técnica para recortar las imágenes. Se creó un algoritmo para recortar las imágenes basándose en sus coordenadas, lo que permitió obtener una imagen recortada de 112x112 píxeles. En la Ilustración 43 se puede ver que después de adquirir la Región de Interés (ROI), las imágenes se redimensionaron a un tamaño estándar de 224x224 píxeles utilizando un algoritmo de interpolación intercúbica [118]. Además, se desarrolló un algoritmo para eliminar automáticamente las manchas negras presentes en las imágenes, ya que estas áreas mostraban restricciones en la mamografía.

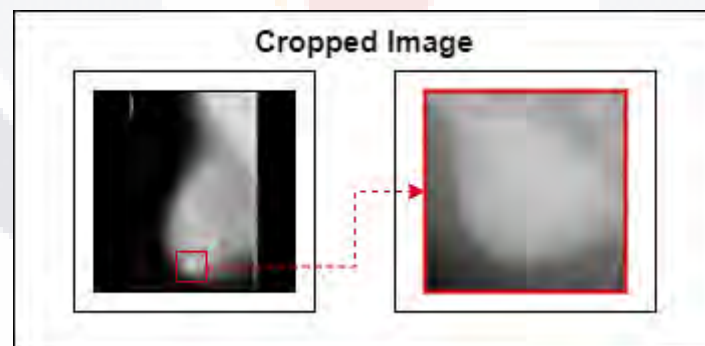


Ilustración 43 Recorte de las mamografías.

Ciertos tipos de malformaciones, como manchas negras, se muestran en algunas imágenes. Estos se eliminaron para que el aprendizaje fuera lo más eficiente posible. El algoritmo K-means y el método estadístico de Gray-Level Co-Occurrence Matrix (GLCM), que extrae características de textura estadísticas de segundo orden, se utilizaron para excluir estas imágenes [41].

Tan et al. [119], se eliminaron las imágenes que contenían manchas negras debido a que estas áreas presentaban artefactos de mamografía. Sin embargo, dado que este proceso parecía haberse realizado manualmente, se desarrolló un algoritmo para eliminar estas áreas de manera automática.

K-means es un algoritmo no supervisado que hace inferencias a partir de bases de datos usando solo vectores de entrada sin referirse a variables conocidas. Este algoritmo divide una colección de datos en múltiples grupos [120]

Las variables de clasificación fueron energía, correlación, homogeneidad, disimilitud, contraste y ASM. Se utilizaron los siguientes parámetros para esta función, graycomatrix, de la biblioteca de imágenes scikit: 1 para la distancia y 0 para el ángulo. Ilustración 44.

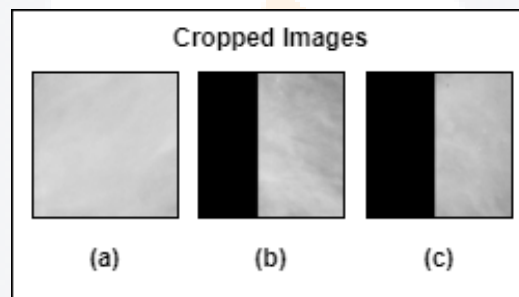


Ilustración 44 Imágenes incompletas.

Se descargó el conjunto de datos CBIS-DDSM, que contenía imágenes recortadas de mamografías que mostraban masas y calcificaciones. El formato DICOM era el utilizado por la base de datos, un estándar global para almacenar imágenes médicas como tomografías computarizadas (CT) e imágenes de resonancia magnética (MRI). Para poder trabajar con las imágenes, fue necesario cambiar la extensión de todas las mamografías a formato PNG.

Una vez descargadas las imágenes de CBIS-DDSM, se observó que tenían diferentes tamaños. Para solucionarlo, se desarrolló un algoritmo que extraía las imágenes de las carpetas y las redimensionaba automáticamente a un tamaño estándar de 244x244 mediante la utilización de un algoritmo de interpolación intercúbica [118].

TESIS TESIS TESIS TESIS TESIS

Después de que las imágenes de esta base de datos fueran descargadas, se fusionaron las imágenes benignas y malignas de cada lesión (masas y calcificaciones) para investigar si las arquitecturas utilizadas en este estudio podían distinguir entre diferentes niveles de gravedad. En total, se contaron con 1331 imágenes de calcificación y 1555 imágenes de masa.

Después de recortar las imágenes de las bases de datos, se normalizaron las imágenes. La resolución de la imagen mamográfica original era de 16 bits, con valores de píxel que oscilaban entre 0 y 65.535. Cuando se normalizaron las imágenes, el valor de cada píxel se cambió a un rango de 0 a 255. (UINT 8)

Para distinguir las lesiones en las mamografías, se tuvo que mejorar el contraste de los tejidos blandos en las imágenes Mini-MIAS y CBIS-DDSM. Como resultado, se utilizaron cuatro técnicas para mejorar la calidad de cada imagen: ecualización de histograma adaptativo limitado de contraste (CLAHE), estiramiento de contraste (CS), enmascaramiento de enfoque (UM) y ecualización de histograma (HE). Ilustración 45.

CLAHE es una variante de la ecualización adaptativa de histogramas en la que se limita la amplificación del contraste para reducir el problema de la amplificación del ruido. CS es una técnica de mejora de imagen simple que intenta mejorar el contraste en una imagen "estirando" el rango de valores de intensidad que contiene para abarcar un rango de valores deseado; UM es un método para mejorar la nitidez de la imagen al identificar los bordes y aumentar el contraste entre ellos, y HE es un método que ajusta automáticamente los valores de intensidad de los píxeles de la imagen. [34], [35]

Se compararon los resultados obtenidos de cada uno de los cuatro procedimientos para mejorar la calidad de la imagen mencionados anteriormente: CLAHE, UM, HE y CS. En este proyecto, CLAHE utilizó un límite de recorte de 0.02, mientras que UM trabajó con una fuerza de nitidez de 1.

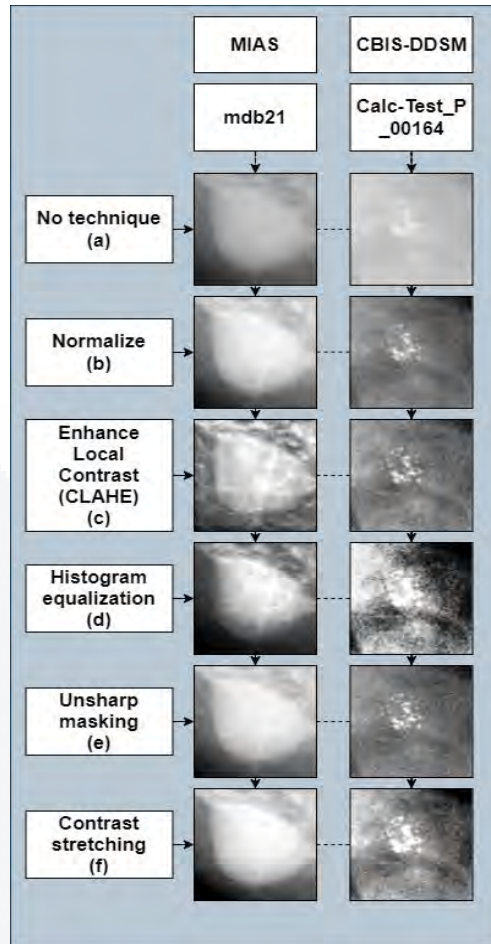


Ilustración 45 Técnicas de preprocesamiento para la metodología 1.

Después de que se etiquetaron las imágenes recortadas preprocesadas en clases, se aplicaron enfoques de mejora a cada base de datos. En el caso de un problema de dos clases, se dividieron las imágenes en las clases 0 y 1, mientras que, para un problema de varias clases, se dividieron en 0, 1 y 2. Posteriormente, se implementó el aumento de datos (DA)[121] para aumentar la cantidad de datos disponibles. El aumento de datos es una técnica que implica agregar copias ligeramente modificadas de los datos existentes o crear datos sintéticos a partir de los datos existentes.

Para esto, se desarrolló un algoritmo utilizando la biblioteca altamente funcional albumentation [36]. Este algoritmo aumenta el número de imágenes según la cantidad deseada por el investigador. La cantidad aumenta en la prueba a través

de funciones de volteo horizontal y vertical, y también puede agregar rotaciones. Sin embargo, el número de rotaciones se debe indicar antes de su implementación. El algoritmo se encarga de realizar estas rotaciones aleatorias y automáticas.

4.6.1.2 Arquitecturas DCNN.

Como se mencionó anteriormente, el aprendizaje por transferencia resulta beneficioso para aprender un nuevo problema al transferir conocimientos de uno relacionado. Se utiliza un modelo ya entrenado para realizar una tarea específica similar. En este estudio, los modelos MobileNetV3Small, MobileNetV3Large, MobileNet, ResNet50, ResNet152, VGG16, VGG19, DenseNet121, InceptionV3 y AlexNet12 fueron entrenados utilizando bases de datos preprocesadas. Las capas de clasificación final de cada modelo se reemplazaron con nuevas capas totalmente conectadas (FC) que contenían 128 neuronas y una capa de salida con dos o tres neuronas, dependiendo de si la tarea era biclase o multiclase. Ilustración.

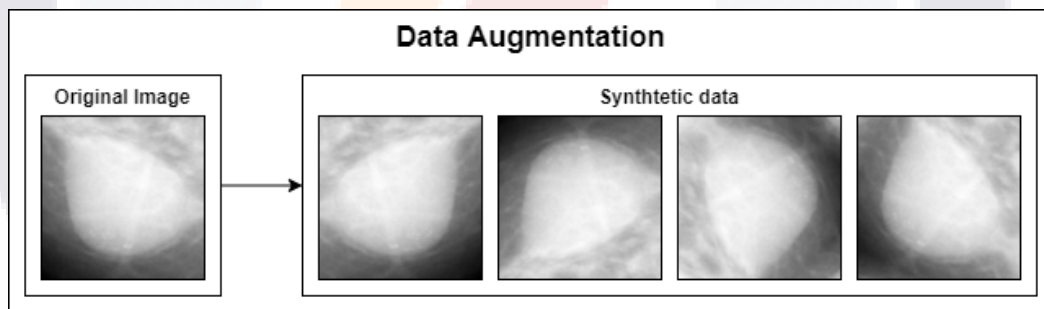


Ilustración 46 Aumento de datos metodología 1.

Obtendríamos malos resultados si no usáramos los pesos de ImageNet ya que previamente habían sido entrenados para sus arquitecturas específicas [32].

La cantidad de canales es de vital importancia a considerar ya que las imágenes utilizadas para cada DCNN preentrenado son RGB. Debemos transformar las imágenes que tenemos que añadir, aunque sean en blanco y negro, en

imágenes con los mismos tres canales para utilizar los pesos que nos proporciona TensorFlow.

4.6.2 Metodología 2.

Para esta segunda metodología los pasos fueron muy similares, pero a diferencia de la primera, el preprocesamiento cambio a uno más simple y el aumento de datos se distribuyó de manera diferente. Para el preprocesamiento solamente se cambió el rango de las imágenes de la misma manera a unos valores de 0-255 y después estas imágenes normalizadas se les aplico la técnica de CLAHE para aumentar la calidad y visibilidad de estas mismas. Es importante mencionar que tanto las imágenes normalizadas como las que se les tuvo que aplicar el preprocesamiento fueron usadas para pruebas y ver los resultados que se podían obtener de ellas.

En la parte de aumento de datos para esta metodología se hizo primero la división de las imágenes entre: entrenamiento, validación y prueba. Para al final solamente hacer el aumento de datos en las imágenes de entrenamiento. La razón de esto fue sencillo a diferencia de la primera donde se pusieron los datos de manera aleatoria, para esta parte las imágenes tendrían un mejor control y los modelos podrían generalizar más en la etapa de entrenamiento y mejorar la robustez del modelo que se está entrenando.

La ilustración 40 muestra los pasos que se hicieron para la segunda metodología la cual es más simple como se puede observar en comparación de la primera. Por los resultados de la primera se decidió solamente utilizarla en el aprendizaje profundo por los buenos resultados que estaba obteniendo y para un mejor control se dividió en 4 etapas en la parte de la clasificación del problema. Ilustración 47.

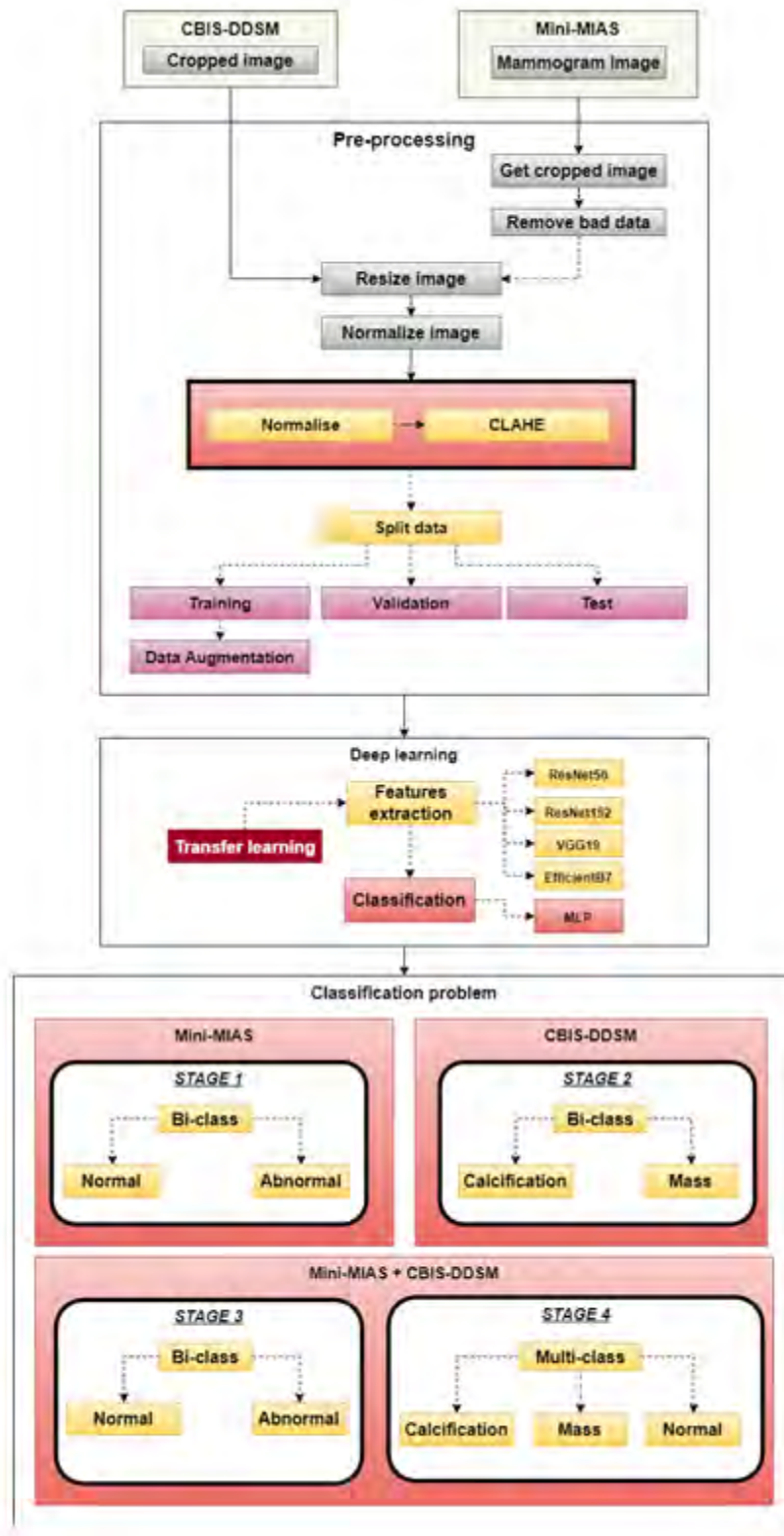


Ilustración 47 Segunda metodología utilizando aprendizaje profundo.

La Ilustración 48 se visualiza el mismo aumento de datos en las imágenes que se utilizó en la primera metodología.

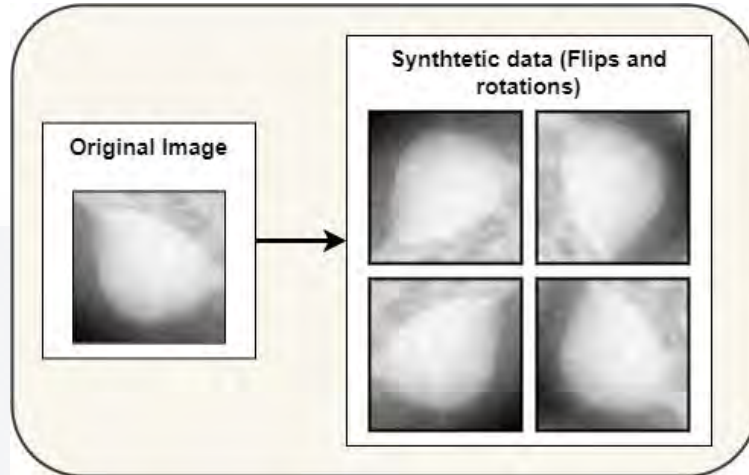


Ilustración 48 Aumento de datos para la metodología 2.

En la Ilustración 49 en cambio solamente fue un preprocesamiento más sencillo donde las imágenes recortadas fueron normalizadas y después se aplicó la técnica de CLAHE.

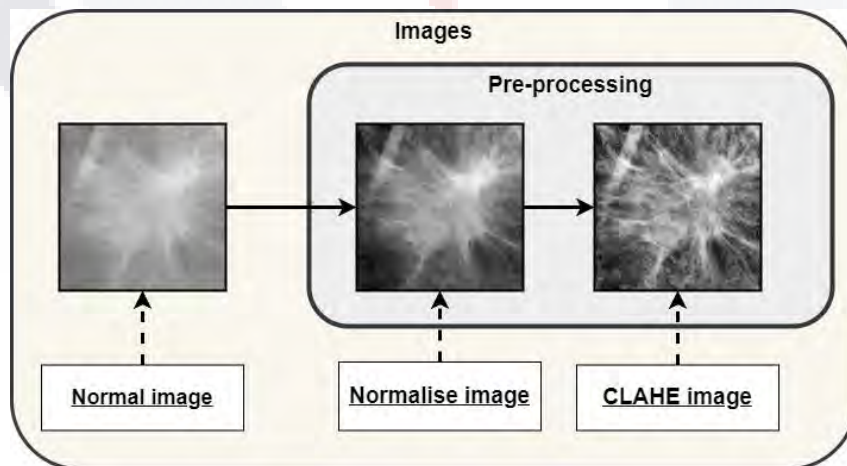


Ilustración 49 Preprocesamiento de imágenes para la metodología 2.

TESIS TESIS TESIS TESIS TESIS

CAPITULO 5

El capítulo 5 de la tesis presenta los resultados del estudio de clasificación de mamografías de manera clara y concisa, acompañados de tablas, figuras y gráficos relevantes. Se discuten las metodologías utilizadas y las métricas empleadas para medir el rendimiento. El objetivo principal del capítulo es presentar los hallazgos obtenidos y responder a las preguntas de investigación planteadas en la tesis. La discusión de los resultados se centra en identificar fortalezas y debilidades de la metodología y proporcionar recomendaciones para futuras investigaciones en este campo.

RESULTADOS

En este capítulo presentamos los resultados de dos metodologías utilizadas para clasificar imágenes mamográficas. En ambas metodologías, el conjunto de datos preprocesados se dividió en datos de entrenamiento, validación y prueba en una proporción de 80-10-10. Para el proceso de clasificación se utilizaron cuatro redes neuronales VGG19, ResNet50, ResNet151 y EfficientNetB7, empleando el optimizador Adam con una tasa de aprendizaje de 0.0001, tamaño de lote de 32 y pérdida de entropía cruzada binaria y categórica. El proceso de entrenamiento duró 20 épocas en ambas metodologías. Sin embargo, la metodología 2 incluyó el uso de aumento de datos para las imágenes del conjunto de entrenamiento, lo que aumentó la solidez del aprendizaje automático y redujo los gastos operativos. En este capítulo se muestran los resultados de ambas metodologías y se compara la eficacia de cada una para clasificar imágenes normales y tumorales en el conjunto de datos Mini-MIAS y CBIS-DDSM. En los Anexos se va a referenciar las tablas de cada metodología y las métricas que se obtuvieron.

5.1 División de conjunto de datos de prueba y de entrenamiento

Un concepto importante que se tendrá es la separación de los datos de entrenamiento, validación y prueba [122].

Train-Test Split es una técnica popular de aprendizaje automático para evaluar el rendimiento del modelo. Implica dividir un conjunto de datos en dos partes: un conjunto de entrenamiento y un conjunto de prueba. El conjunto de entrenamiento se usa para entrenar el modelo, mientras que el conjunto de prueba se usa para evaluar el rendimiento del modelo.

El conjunto de entrenamiento se usa para entrenar el modelo, el conjunto de validación se usa para ajustar los hiperparámetros del modelo y el conjunto de prueba se usa para evaluar el rendimiento del modelo en datos nuevos e invisibles. El conjunto de entrenamiento y el conjunto de validación generalmente consisten en 60-80% del total de datos.

El conjunto de prueba consta de una parte del conjunto de datos que se utiliza para evaluar el rendimiento del modelo. El modelo no está entrenado con estos datos, y el conjunto de prueba proporciona una forma de evaluar la capacidad del modelo para hacer predicciones sobre datos nuevos e invisibles. El tamaño del conjunto de prueba suele ser alrededor del 20-30% del conjunto de datos total.

5.2 Resultados Metodología 1

Dividimos el conjunto de datos preprocesados 80% en datos de entrenamiento, 10% en datos de validación y 10% en datos de prueba. Usamos los siguientes parámetros para las redes neuronales VGG19, ResNet50, ResNet151 y EfficientNetB7 para la clasificación: usamos el optimizador Adam con una tasa de aprendizaje de 0.0001, establecimos el tamaño del lote en 32 y usamos entropía cruzada binaria y categórica como la pérdida funciones. El proceso de entrenamiento duró 20 épocas. Al igual que la primera metodología. Para los dos conjuntos de datos (Mini-MIAS y CBIS-DDSM), así mismo para el problema binario y multiclase se dividió el 80% de todas las imágenes para el entrenamiento, 10%

para validación y 10% para prueba. Al entrenar cada modelo de DCNN, establecimos los siguientes parámetros: Adam es un algoritmo de optimización que inicializamos con un valor de 0.0001. Cuando se utilizó el aprendizaje profundo se realizaron 20 épocas. Para la red neuronal de clasificación, empleamos estas capas para mejorar el resultado. Ilustración 50.

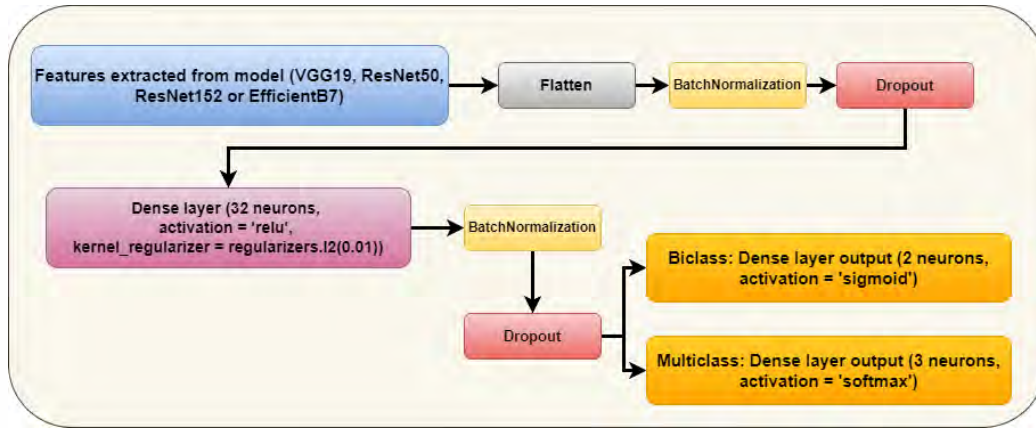


Ilustración 50 Topología de la etapa de clasificación usando redes neuronales artificiales.

La Tabla 2 presenta los resultados de una comparación entre la técnica MobileNetV3 Small y varias otras técnicas de procesamiento de imágenes que no involucran preprocesamiento. La tabla muestra varias métricas de rendimiento y los resultados logrados por cada técnica. Los resultados indican que, si bien la técnica MobileNetV3 Small produjo resultados similares en todas las técnicas de imágenes,

se encontró que los resultados normalizados fueron ligeramente superiores a los demás.

Tabla 2 El modelo MobileNetV3-Small aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
MNV3S	NT	0.9399	0.9343	0.9421	0.9382	0.9883	0.9123	0.9145	0.9123	0.9126	0.9336
MNV3S	NO	0.9531	0.9525	0.9505	0.9515	0.9917	0.9449	0.9449	0.9449	0.9449	0.9589
MNV3S	CLAHE	0.945	0.9296	0.9589	0.944	0.9874	0.9286	0.9309	0.9286	0.9286	0.9465
MNV3S	HE	0.9577	0.9549	0.9579	0.9564	0.9915	0.9434	0.9433	0.9434	0.9432	0.9581
MNV3S	UM	0.9526	0.9663	0.9348	0.9502	0.9928	0.9357	0.9358	0.9357	0.9357	0.9519
MNV3S	CS	0.9597	0.9599	0.9568	0.9583	0.9923	0.9322	0.9327	0.9322	0.9322	0.9487
Model	Technique	CBIS-DDSM									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
MNV3S	NT	0.8902	0.8982	0.8775	0.8878	0.9585	0.9256	0.9247	0.9249	0.9247	0.944
MNV3S	NO	0.8816	0.8568	0.9134	0.8842	0.9591	0.9269	0.9263	0.9263	0.9261	0.945
MNV3S	CLAHE	0.8797	0.8508	0.9179	0.8831	0.9512	0.9153	0.9154	0.9153	0.9153	0.9362
MNV3S	HE	0.8639	0.8828	0.8358	0.8587	0.9417	0.9054	0.9072	0.9059	0.9061	0.9288
MNV3S	UM	0.8874	0.8935	0.8769	0.8851	0.9593	0.9151	0.9181	0.9134	0.9136	0.9355
MNV3S	CS	0.8775	0.8664	0.8897	0.8779	0.9483	0.9156	0.9144	0.9145	0.9144	0.9362

La Tabla 3 muestra los resultados de un experimento que evaluó el rendimiento de la técnica MobileNetV3 Large en comparación con otras técnicas de imagen sin preprocesamiento. La tabla presenta los resultados en términos de varias métricas. Los resultados muestran que MobileNetV3 Large logró mejores resultados cuando se usó el estiramiento de contraste. Para el escenario multiclase, los resultados normalizados se desempeñaron mejor entre las técnicas evaluadas.

Tabla 3 El modelo MobileNetV3-Large aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
MNV3L	NT	0.944	0.9468	0.9369	0.9418	0.9884	0.9072	0.9108	0.9072	0.9067	0.9307
MNV3L	NO	0.975	0.9611	0.9884	0.9745	0.9969	0.9638	0.9641	0.9638	0.9638	0.9728
MNV3L	CLAHE	0.974	0.9736	0.9726	0.9731	0.9969	0.9571	0.9572	0.9571	0.9571	0.968
MNV3L	HE	0.9679	0.9567	0.9779	0.9672	0.9961	0.9678	0.9679	0.9678	0.9678	0.976
MNV3L	UM	0.977	0.9758	0.9768	0.9763	0.9968	0.9683	0.9685	0.9683	0.9683	0.9765
MNV3L	CS	0.9776	0.9871	0.9663	0.9766	0.9976	0.9668	0.9669	0.9668	0.9668	0.9754
Model	Technique	CBIS-DDSM									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
MNV3L	NT	0.8978	0.904	0.8878	0.8958	0.9649	0.9313	0.9304	0.9304	0.9303	0.9481
MNV3L	NO	0.9121	0.9279	0.8916	0.9094	0.9731	0.9416	0.941	0.9409	0.941	0.9559
MNV3L	CLAHE	0.9048	0.8977	0.9115	0.9045	0.9673	0.9311	0.9303	0.9304	0.9303	0.9481
MNV3L	HE	0.8928	0.8739	0.9153	0.8941	0.9595	0.9178	0.9201	0.9181	0.9188	0.938
MNV3L	UM	0.9175	0.9047	0.9314	0.9178	0.9739	0.9409	0.9403	0.9401	0.9402	0.9553
MNV3L	CS	0.9	0.9136	0.8814	0.8972	0.9637	0.9293	0.9287	0.9287	0.9284	0.9468

La Tabla 4 presenta los resultados de un experimento que evaluó el desempeño de la técnica MobileNet en comparación con otras técnicas de imagen sin preprocesamiento. La tabla muestra los resultados de cada técnica en términos de varias métricas. Los resultados muestran que MobileNet se desempeñó mejor cuando se usó un histograma de equalización. Para el escenario multiclase, los

resultados de la técnica CLAHE fueron los mejores entre las técnicas evaluadas. Se encontró que el conjunto de datos sin preprocesamiento funcionaba mal en comparación con las otras técnicas.

Tabla 4 El modelo MobileNet aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
MN	NT	0.7216	0.6546	0.899	0.7576	0.88	0.4653	0.7325	0.4653	0.3823	0.6115
MN	NO	0.9063	0.8974	0.9106	0.9039	0.9684	0.8695	0.8737	0.8695	0.8693	0.9037
MN	CLAHE	0.9124	0.9036	0.9169	0.9102	0.9738	0.8348	0.8692	0.8348	0.8359	0.8742
MN	HE	0.9531	0.9423	0.9621	0.9521	0.9892	0.9332	0.9334	0.9332	0.9331	0.9501
MN	UM	0.918	0.9247	0.9043	0.9144	0.9717	0.8516	0.8539	0.8516	0.8511	0.8903
MN	CS	0.947	0.9462	0.9442	0.9452	0.9848	0.9189	0.9192	0.9189	0.9188	0.9398
Model	Technique	CBIS-DDSM									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
MN	NT	0.8169	0.7753	0.8871	0.8275	0.8945	0.8624	0.8603	0.8615	0.8591	0.8965
MN	NO	0.836	0.8906	0.7621	0.8214	0.9262	0.8697	0.8684	0.8681	0.8676	0.9015
MN	CLAHE	0.8617	0.8477	0.8782	0.8627	0.9338	0.8854	0.8885	0.8852	0.8822	0.9141
MN	HE	0.856	0.8388	0.8775	0.8577	0.9285	0.8717	0.8817	0.8679	0.8688	0.9008
MN	UM	0.8468	0.8471	0.8423	0.8447	0.9235	0.873	0.8716	0.8723	0.8702	0.9045
MN	CS	0.8363	0.8728	0.7833	0.8256	0.9171	0.8743	0.8716	0.8732	0.872	0.9052

La Tabla 5 muestra los resultados de un experimento que evaluó el rendimiento de la técnica ResNet50 en comparación con otras técnicas de imagen sin preprocesamiento. La tabla presenta los resultados en términos de varias métricas. Los resultados muestran que ResNet50 logró los mejores resultados al usar la normalización. Para el escenario multiclase, también se encontró que la normalización funciona mejor entre las técnicas evaluadas. También se calculó el resultado promedio para cada técnica y la arquitectura ResNet50 superó a las arquitecturas anteriores que se evaluaron.

Tabla 5 El modelo ResNet50 aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
RN50	NT	0.9857	0.9812	0.9894	0.9853	0.9989	0.9704	0.9705	0.9704	0.9704	0.978
RN50	NO	0.9862	0.9863	0.9852	0.9857	0.9987	0.9689	0.9689	0.9689	0.9689	0.9767
RN50	CLAHE	0.975	0.9727	0.9758	0.9742	0.9966	0.9658	0.9671	0.9658	0.9659	0.9741
RN50	HE	0.9806	0.974	0.9863	0.9801	0.9967	0.9699	0.9699	0.9699	0.9699	0.9772
RN50	UM	0.9816	0.977	0.9852	0.9811	0.9988	0.9709	0.9712	0.9709	0.9709	0.978
RN50	CS	0.9826	0.9781	0.9863	0.9821	0.9987	0.9709	0.9709	0.9709	0.9709	0.9781
Model	Technique	CBIS-DDSM									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
RN50	NT	0.913	0.9208	0.9019	0.9112	0.9694	0.9375	0.9368	0.9369	0.9367	0.9529
RN50	NO	0.9216	0.9065	0.9384	0.9222	0.9747	0.9409	0.9402	0.9404	0.9401	0.9555
RN50	CLAHE	0.9045	0.9302	0.8724	0.9004	0.9692	0.9347	0.9341	0.9337	0.9338	0.9506
RN50	HE	0.9051	0.8967	0.9134	0.905	0.966	0.935	0.9381	0.9346	0.9358	0.9506
RN50	UM	0.92	0.923	0.9147	0.9188	0.9733	0.9398	0.9395	0.9389	0.939	0.9545
RN50	CS	0.8972	0.8614	0.9442	0.9009	0.9667	0.9282	0.9324	0.9282	0.9271	0.9463

La Tabla 6 presenta los resultados de un experimento que evaluó el desempeño de la técnica ResNet152 en comparación con otras técnicas de imagen sin preprocesamiento. La tabla muestra los resultados de cada técnica en términos de varias métricas. Los resultados muestran que ResNet152 funcionó mejor cuando se usó la normalización. Para el escenario multiclase, la técnica CLAHE produjo los mejores resultados para el conjunto de datos MIAS, mientras que la normalización funcionó mejor para el conjunto de datos CBIS-DDSM. También se calculó el resultado promedio para cada técnica y se encontró que la arquitectura ResNet151 superó a las arquitecturas anteriores que se evaluaron, lo que la convirtió en la mejor arquitectura entre las estudiadas.

Tabla 6 El modelo ResNet152 aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
RN152	NT	0.9786	0.9709	0.9852	0.978	0.998	0.9663	0.967	0.9663	0.9664	0.975
RN152	NO	0.9842	0.9761	0.9915	0.9838	0.9982	0.9704	0.9704	0.9704	0.9704	0.978
RN152	CLAHE	0.9791	0.9809	0.9758	0.9783	0.9981	0.9765	0.9765	0.9765	0.9765	0.9825
RN152	HE	0.9837	0.9862	0.98	0.9831	0.9991	0.9729	0.9734	0.9729	0.9729	0.98
RN152	UM	0.9811	0.9903	0.9705	0.9803	0.9984	0.9719	0.972	0.9719	0.9719	0.9792
RN152	CS	0.9816	0.982	0.98	0.981	0.9967	0.976	0.976	0.976	0.976	0.9819
Model	Technique	CBIS-DDSM									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
RN152	NT	0.9067	0.9042	0.9076	0.9059	0.963	0.9396	0.9396	0.9392	0.9389	0.9546
RN152	NO	0.9232	0.9235	0.9211	0.9223	0.9769	0.944	0.9436	0.9432	0.9433	0.9576
RN152	CLAHE	0.9169	0.9092	0.9243	0.9167	0.9727	0.9365	0.9368	0.9354	0.9356	0.9518
RN152	HE	0.9032	0.9103	0.8923	0.9012	0.9637	0.929	0.9297	0.9298	0.9297	0.9469
RN152	UM	0.9241	0.9094	0.9403	0.9246	0.9782	0.9416	0.9409	0.9411	0.941	0.956
RN152	CS	0.9073	0.9064	0.9064	0.9064	0.9695	0.9305	0.9297	0.93	0.9297	0.9477

La Tabla 7 muestra los resultados de un experimento que evaluó el rendimiento de la técnica InceptionV3 en comparación con otras técnicas de imagen sin preprocesamiento. La tabla presenta los resultados en términos de varias métricas. Los resultados muestran que InceptionV3 se desempeñó mejor cuando se utilizó la técnica CLAHE para el conjunto de datos MIAS. Para el escenario multiclase, el histograma de ecualización produjo los mejores resultados para el conjunto de datos MIAS, mientras que las imágenes sin preprocesamiento obtuvieron los mejores resultados para el conjunto de datos CBIS-DDSM. También se calculó y comparó el resultado promedio de cada técnica.

Tabla 7 El modelo Inception-V3 aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
IV3	NT	0.8814	0.9532	0.7939	0.8663	0.9465	0.6146	0.6294	0.6146	0.5918	0.7167
IV3	NO	0.9318	0.9406	0.9169	0.9286	0.9814	0.8883	0.89	0.8883	0.8882	0.9171
IV3	CLAHE	0.9343	0.9409	0.9221	0.9314	0.9877	0.9001	0.9005	0.9001	0.8997	0.9245
IV3	HE	0.9384	0.9108	0.9674	0.9382	0.9858	0.9199	0.9236	0.9199	0.92	0.9386
IV3	UM	0.9251	0.9729	0.8696	0.9183	0.9852	0.8893	0.8909	0.8893	0.8895	0.9176
IV3	CS	0.9465	0.9424	0.9474	0.9449	0.9846	0.9128	0.915	0.9128	0.913	0.934
Model	Technique	CBIS-DDSM									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
IV3	NT	0.8414	0.7937	0.9179	0.8513	0.919	0.8828	0.8808	0.8818	0.8811	0.9117
IV3	NO	0.8534	0.8457	0.8608	0.8532	0.9319	0.8467	0.8463	0.8456	0.8453	0.8845
IV3	CLAHE	0.8658	0.8498	0.8852	0.8671	0.9374	0.8537	0.8532	0.8526	0.8506	0.8897
IV3	HE	0.8464	0.8183	0.8865	0.851	0.9226	0.8067	0.8128	0.8	0.7923	0.851
IV3	UM	0.8528	0.8234	0.8942	0.8574	0.9316	0.8544	0.8526	0.8541	0.8528	0.8907
IV3	CS	0.843	0.8209	0.873	0.8462	0.9236	0.8237	0.8286	0.8242	0.8169	0.8682

La Tabla 8 ilustra los resultados de un experimento que evaluó el rendimiento de la arquitectura DenseNet 121 en comparación con otras técnicas de imagen sin preprocesamiento. La tabla muestra los resultados en términos de varias métricas. Los resultados muestran que la arquitectura DenseNet 121 logró los mejores resultados al usar el histograma de ecualización para el conjunto de datos MIAS. Las imágenes sin preprocesamiento funcionaron mal en comparación con las otras técnicas. Para el escenario multiclase, la técnica de enmascaramiento de enfoque produjo los mejores resultados para el conjunto de datos CBIS-DDSM. Estos resultados demuestran la eficacia de diferentes técnicas de preprocesamiento para mejorar el rendimiento de la arquitectura DenseNet 121.

Tabla 8 El modelo DenseNet121 aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
DN121	NT	0.7475	0.6633	0.9716	0.7883	0.9301	0.7104	0.7667	0.7104	0.6841	0.7841
DN121	NO	0.9567	0.9529	0.9579	0.9554	0.9922	0.9204	0.9208	0.9204	0.9205	0.9406
DN121	CLAHE	0.948	0.9953	0.8969	0.9435	0.9941	0.9317	0.9337	0.9317	0.9315	0.9482
DN121	HE	0.9643	0.9691	0.9568	0.9629	0.9914	0.949	0.9496	0.949	0.949	0.9617
DN121	UM	0.9511	0.9692	0.9284	0.9484	0.9889	0.9225	0.9236	0.9225	0.9224	0.9428
DN121	CS	0.9735	0.9807	0.9642	0.9724	0.9951	0.9332	0.9338	0.9332	0.933	0.9501
Model	Technique	CBIS-DDSM									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
DN121	NT	0.8715	0.868	0.873	0.8705	0.9417	0.9097	0.9094	0.909	0.9084	0.9321
DN121	NO	0.8705	0.9096	0.8198	0.8624	0.95	0.8986	0.9029	0.8985	0.8969	0.9241
DN121	CLAHE	0.8912	0.8726	0.9134	0.8925	0.9562	0.9099	0.9087	0.9089	0.9087	0.932
DN121	HE	0.8671	0.8318	0.9166	0.8722	0.9491	0.902	0.9045	0.9005	0.901	0.9252
DN121	UM	0.8829	0.8805	0.8833	0.8819	0.9549	0.912	0.913	0.9116	0.9103	0.934
DN121	CS	0.8788	0.8834	0.8698	0.8766	0.9493	0.9042	0.9031	0.9033	0.9032	0.9278

Los resultados de esta investigación que evaluó el rendimiento de la arquitectura VGG16 en comparación con otras técnicas de imagen sin preprocesamiento se presentan en la Tabla 9. La tabla muestra los resultados de cada técnica en términos de varias métricas. Los resultados indican que la arquitectura VGG16 se desempeñó mejor cuando se utilizó el histograma de ecuilización para el conjunto de datos MIAS en el escenario multiclase. Las imágenes sin preprocesamiento se comportaron mal en comparación con las imágenes con preprocesamiento. Para el escenario multiclase, la técnica de enmascaramiento no nítida produjo mejores resultados para el conjunto de datos CBIS-DDSM en comparación con la clasificación binaria. Estos resultados demuestran el impacto de las técnicas de preprocesamiento en el rendimiento de la arquitectura VGG16 en diferentes escenarios de clasificación.

Tabla 9 El modelo VGGNet16 aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
VGG16	NT	0.9951	0.9546	0.9747	0.9646	0.9945	0.9617	0.9617	0.9617	0.9617	0.9716
VGG16	NO	0.9992	0.9677	0.9768	0.9722	0.9959	0.9658	0.9659	0.9658	0.9658	0.9742
VGG16	CLAHE	0.9984	0.9442	0.98	0.9618	0.9918	0.9515	0.9516	0.9515	0.9516	0.9636
VGG16	HE	0.9995	0.9704	0.9674	0.9689	0.9951	0.9704	0.9706	0.9704	0.9704	0.9774
VGG16	UM	0.9985	0.9586	0.9747	0.9666	0.9965	0.9627	0.9628	0.9627	0.9627	0.9722
VGG16	CS	0.999	0.9726	0.9716	0.9721	0.996	0.9597	0.96	0.9597	0.9597	0.9697
Model	Technique	CBIS-DDSM									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
VGG16	NT	0.9035	0.892	0.916	0.9038	0.9642	0.9184	0.917	0.9175	0.9171	0.9384
VGG16	NO	0.9045	0.9079	0.898	0.9029	0.9627	0.9246	0.9235	0.9238	0.9236	0.9431
VGG16	CLAHE	0.8937	0.9203	0.8596	0.8889	0.9563	0.9166	0.9159	0.9156	0.9156	0.937
VGG16	HE	0.8947	0.9018	0.8833	0.8924	0.95	0.9104	0.9103	0.9124	0.9112	0.9337
VGG16	UM	0.8997	0.9049	0.891	0.8979	0.9616	0.9277	0.9267	0.9269	0.9268	0.9455
VGG16	CS	0.8836	0.9027	0.857	0.8793	0.9526	0.9184	0.9174	0.9175	0.9174	0.9384

Tabla 10 muestra los resultados de un experimento que comparó el rendimiento de la arquitectura VGG19 con otras técnicas de imagen sin preprocesamiento. Los resultados se muestran en términos de varias métricas. Los resultados sugieren que la arquitectura VGG19 funcionó mejor cuando se utilizó el histograma de ecuilización para el conjunto de datos MIAS en el escenario multiclase. La arquitectura VGG16 obtuvo resultados similares a los de la VGG19. Sin embargo, en este experimento, las imágenes sin preprocesamiento superaron a la arquitectura VGG16. Para el escenario multiclase, la técnica de

enmascaramiento de enfoque produjo mejores resultados para el conjunto de datos CBIS-DDSM. Estos resultados destacan la importancia de las técnicas de preprocesamiento para mejorar el rendimiento de las arquitecturas VGG19 y VGG16 en diferentes escenarios de clasificación.

Tabla 10 El modelo VGGNet19 aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	FI-Score	AUC	Accuracy	Precision	Recall	FI-Score	AUC
VGG19	NT	0.9582	0.9588	0.9547	0.9567	0.9909	0.9434	0.9438	0.9434	0.9435	0.9577
VGG19	NO	0.9709	0.9705	0.9695	0.97	0.9953	0.9576	0.9577	0.9576	0.9576	0.9681
VGG19	CLAHE	0.9669	0.9702	0.961	0.9656	0.9932	0.9418	0.9418	0.9418	0.9417	0.9568
VGG19	HE	0.9684	0.9664	0.9684	0.9674	0.9939	0.9566	0.9566	0.9566	0.9566	0.9674
VGG19	UM	0.9725	0.9696	0.9737	0.9716	0.9951	0.9561	0.9562	0.9561	0.9561	0.9673
VGG19	CS	0.973	0.9756	0.9684	0.972	0.9961	0.9546	0.9546	0.9546	0.9546	0.9661
Model	Technique	CBIS-DDSM									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	FI-Score	AUC	Accuracy	Precision	Recall	FI-Score	AUC
VGG19	NT	0.8791	0.8587	0.9044	0.881	0.9489	0.9176	0.9166	0.9168	0.9167	0.9379
VGG19	NO	0.8915	0.9182	0.857	0.8866	0.9591	0.9156	0.9149	0.9148	0.9149	0.9364
VGG19	CLAHE	0.8769	0.8825	0.8666	0.8745	0.949	0.9102	0.9088	0.9092	0.9089	0.9322
VGG19	HE	0.8775	0.8659	0.8903	0.878	0.9501	0.9006	0.9043	0.8996	0.901	0.9243
VGG19	UM	0.8953	0.8819	0.9102	0.8958	0.9588	0.9174	0.9174	0.9162	0.9164	0.9375
VGG19	CS	0.8896	0.8965	0.8782	0.8873	0.9545	0.9143	0.9138	0.9133	0.9134	0.9353

La Tabla 11 presenta los resultados de un experimento que comparó el rendimiento de la arquitectura AlexNet12 con otras técnicas de imagen sin preprocesamiento. Los resultados se muestran en términos de varias métricas. Los resultados indican que AlexNet12 funcionaron mejor cuando se usó el histograma de ecualización para el conjunto de datos MIAS, tanto en clasificación binaria como en escenarios multiclase. A pesar de los malos resultados obtenidos con datos preprocesados, AlexNet12 lograron excelentes resultados con datos sin procesar al utilizar todas las técnicas. Vale la pena mencionar que esta arquitectura fue hecha a mano usando TensorFlow y no usa hiperparámetros de ImageNet. Estos resultados destacan el potencial de AlexNet12, especialmente cuando se trabaja con datos sin procesar, y demuestran el impacto de las técnicas de preprocesamiento en el rendimiento de diferentes arquitecturas en tareas de clasificación de imágenes.

Tabla 11 El modelo AlexNet12 aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
AN12	NT	0.7491	0.7313	0.7613	0.746	0.8491	0.7181	0.7645	0.7181	0.7115	0.7838
AN12	NO	0.9083	0.973	0.8338	0.898	0.9665	0.5927	0.6838	0.5927	0.5654	0.6869
AN12	CLAHE	0.9552	0.9886	0.9179	0.952	0.9951	0.8603	0.8634	0.8603	0.8593	0.8942
AN12	HE	0.9643	0.9383	0.9915	0.9642	0.9958	0.9245	0.9263	0.9245	0.9245	0.9436
AN12	UM	0.9389	0.9235	0.9526	0.9378	0.9869	0.7329	0.7583	0.7329	0.7195	0.7933
AN12	CS	0.9521	0.9246	0.981	0.952	0.9933	0.5823	0.5799	0.5423	0.491	0.6466
Model	Technique	CBIS-DDSM									
		Bi-class					Multi-class				
		Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
AN12	NT	0.8731	0.8629	0.8839	0.8733	0.9392	0.8614	0.866	0.8606	0.8573	0.8956
AN12	NO	0.8496	0.8683	0.8205	0.8437	0.9274	0.4979	0.6624	0.5056	0.3969	0.6281
AN12	CLAHE	0.8309	0.7854	0.9057	0.8413	0.9186	0.3913	0.7386	0.3834	0.2644	0.5378
AN12	HE	0.8363	0.7972	0.8974	0.8443	0.9094	0.3269	0.647	0.3462	0.1864	0.5096
AN12	UM	0.7678	0.6947	0.9467	0.8014	0.9015	0.3968	0.7088	0.389	0.2762	0.5419
AN12	CS	0.8087	0.744	0.9352	0.8287	0.9078	0.7856	0.8131	0.7824	0.7825	0.8376

5.3 Clasificación usando algoritmos de aprendizaje automático.

En primer lugar, se compararon varios conjuntos de imágenes preprocesadas utilizando diferentes algoritmos de aprendizaje automático, como Support Vector Machine, Multi Support Vector Machine, Multi Layer Perceptron, Decision Tree, K Neighbors, Random Forest y Gradient Boosting Classifier, para clasificar estas imágenes. Los resultados de este experimento se presentan en las Tablas 12 a 17, donde se evaluó el rendimiento general de los diferentes algoritmos y técnicas de aprendizaje profundo utilizando perceptrones multicapa (MLP). En general, los resultados indicaron que el rendimiento de estos algoritmos fue deficiente. Sin embargo, el algoritmo de Support Vector Machine se desempeñó excepcionalmente bien en la clasificación binaria utilizando el conjunto de datos CBIS-DDSM y superó a todos los demás modelos y técnicas. En particular, la matriz de longitud de ejecución de nivel de gris (GLLRM) obtuvo buenos resultados en la clasificación binaria con el conjunto de datos CBIS-DDSM. El enfoque de clasificación binaria y GLRLM generó buenos resultados para algunos de los modelos utilizados. Estos resultados destacan la importancia de seleccionar cuidadosamente el algoritmo de aprendizaje automático y la técnica de preprocesamiento adecuados para una tarea de clasificación de imágenes y un conjunto de datos específicos. Tabla 12, Tabla 13, Tabla 14, Tabla 15, Tabla 16, Tabla 17.

Tabla 12 El modelo SVM aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS							
		Bi-class FOF				Multi-class FOF			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
SVM	NT	0.7631	0.7197	0.8667	0.7864	0.6476	0.6597	0.6476	0.6474
SVM	NO	0.7589	0.7163	0.8625	0.7826	0.6476	0.6579	0.6476	0.6478
SVM	CLAHE	0.7526	0.7089	0.8625	0.7782	0.649	0.663	0.649	0.6487
SVM	HE	0.761	0.7172	0.8667	0.7849	0.6517	0.6679	0.6517	0.652
SVM	UM	0.7589	0.7178	0.8583	0.7818	0.6462	0.6567	0.6462	0.6458
SVM	CS	0.7568	0.7168	0.8542	0.7795	0.6448	0.6589	0.6448	0.6449
Model	Technique	Mini-MIAS							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
SVM	NT	0.6583	0.6442	0.7167	0.6785	0.6364	0.6328	0.6364	0.6298
SVM	NO	0.6625	0.6491	0.7167	0.6812	0.6308	0.6268	0.6308	0.6249
SVM	CLAHE	0.6625	0.6479	0.7208	0.6824	0.6322	0.6273	0.6322	0.6255
SVM	HE	0.6646	0.6504	0.7208	0.6838	0.628	0.6245	0.628	0.6221
SVM	UM	0.6646	0.6504	0.7208	0.6838	0.6378	0.6346	0.6378	0.6319
SVM	CS	0.6625	0.6479	0.7208	0.6824	0.6322	0.6279	0.6322	0.6261
Model	Technique	Mini-MIAS							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
SVM	NT	0.4063	0.6066	0.4063	0.4731	0.4839	0.3778	0.8095	0.5152
SVM	NO	0.625	0.6874	0.625	0.6495	0.5	0.3846	0.6522	0.4839
SVM	CLAHE	0.5156	0.578	0.5156	0.5389	0.5333	0.3871	0.5714	0.4615
SVM	HE	0.4063	0.5964	0.4063	0.4311	0.625	0.48	0.5217	0.5
SVM	UM	0.5313	0.5885	0.5313	0.5487	0.5313	0.4146	0.7391	0.5313
SVM	CS	0.4688	0.5751	0.4688	0.5092	0.5968	0.45	0.8571	0.5902
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
SVM	NT	0.7048	0.6666	0.7259	0.6950	0.7584	0.755	0.7584	0.7529
SVM	NO	0.7040	0.6658	0.7250	0.6942	0.7578	0.7545	0.7578	0.7522
SVM	CLAHE	0.7064	0.6666	0.7326	0.6980	0.7576	0.7546	0.7576	0.7518
SVM	HE	0.7033	0.6651	0.7242	0.6934	0.755	0.7516	0.755	0.7492
SVM	UM	0.7052	0.6669	0.7267	0.6955	0.7574	0.7538	0.7574	0.7518
SVM	CS	0.7071	0.6694	0.7267	0.6969	0.7565	0.7532	0.7565	0.7508
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
SVM	NT	0.6213	0.5654	0.7896	0.6589	0.6287	0.6397	0.6287	0.5989
SVM	NO	0.6190	0.5633	0.7904	0.6578	0.6289	0.6403	0.6289	0.5994
SVM	CLAHE	0.6213	0.5650	0.7937	0.6601	0.6294	0.6404	0.6294	0.6
SVM	HE	0.6194	0.5634	0.7929	0.6587	0.6289	0.64	0.6289	0.5993
SVM	UM	0.6194	0.5636	0.7904	0.6580	0.6283	0.6394	0.6283	0.5984
SVM	CS	0.6201	0.5641	0.7921	0.6589	0.63	0.6414	0.63	0.6006
Model	Technique	CBIS-DDSM							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
SVM	NT	1	1	1	1	0.5676	0.6052	0.5676	0.5674
SVM	NO	1	1	1	1	0.4879	0.519	0.4879	0.4914
SVM	CLAHE	1	1	1	1	0.471	0.5138	0.471	0.4753
SVM	HE	1	1	1	1	0.4831	0.5478	0.4831	0.4852
SVM	UM	1	1	1	1	0.4952	0.5885	0.4952	0.500
SVM	CS	1	1	1	1	0.5	0.5245	0.5	0.4988

Tabla 13 El modelo MLP aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS							
		Bi-class FOF				Multi-class FOF			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
MLP	NT	.8868	0.8843	0.8917	0.888	0.8958	0.8978	0.8958	0.8953
MLP	NO	0.8574	0.8772	0.8333	0.8547	0.9129	0.9191	0.9129	0.9122
MLP	CLAHE	0.8931	0.8987	0.8875	0.8931	0.8773	0.8787	0.8773	0.8769
MLP	HE	0.8658	0.8826	0.8458	0.8638	0.9156	0.9154	0.9156	0.9155
MLP	UM	0.891	0.895	0.8875	0.8912	0.8786	0.8791	0.8786	0.8786
MLP	CS	0.9203	0.9244	0.9167	0.9205	0.8668	0.8669	0.8668	0.86689
Model	Technique	Mini-MIAS							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
MLP	NT	0.6583	0.6442	0.7167	0.6785	0.6364	0.6328	0.6364	0.6298
MLP	NO	0.6625	0.6491	0.7167	0.6812	0.6308	0.6268	0.6308	0.6249
MLP	CLAHE	0.6625	0.6479	0.7208	0.6824	0.6322	0.6273	0.6322	0.6255
MLP	HE	0.6646	0.6504	0.7208	0.6838	0.628	0.6245	0.628	0.6221
MLP	UM	0.6646	0.6504	0.7208	0.6838	0.6378	0.6346	0.6378	0.6319
MLP	CS	0.6625	0.6479	0.7208	0.6824	0.6322	0.6279	0.6322	0.6261
Model	Technique	Mini-MIAS							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
MLP	NT	0.5156	0.6542	0.5156	0.5565	0.5806	0.3684	0.3333	0.35
MLP	NO	0.4375	0.5457	0.4375	0.4709	0.5313	0.3793	0.4783	0.4231
MLP	CLAHE	0.4844	0.5759	0.4844	0.5173	0.6667	0.5294	0.4286	0.4737
MLP	HE	0.4219	0.6135	0.4219	0.4513	0.6406	0.5	0.5217	0.5106
MLP	UM	0.4531	0.61	0.4531	0.5061	0.5781	0.4286	0.5217	0.4706
MLP	CS	0.4531	0.5933	0.4531	0.4908	0.4516	0.303	0.4762	0.3704
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
MLP	NT	0.7048	0.6666	0.7259	0.6950	0.7584	0.755	0.7584	0.7529
MLP	NO	0.7040	0.6658	0.7250	0.6942	0.7578	0.7545	0.7578	0.7522
MLP	CLAHE	0.7064	0.6666	0.7326	0.6980	0.7576	0.7546	0.7576	0.7518
MLP	HE	0.7033	0.6651	0.7242	0.6934	0.755	0.7516	0.755	0.7492
MLP	UM	0.7052	0.6669	0.7267	0.6955	0.7574	0.7538	0.7574	0.7518
MLP	CS	0.7071	0.6694	0.7267	0.6969	0.7565	0.7532	0.7565	0.7508
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
MLP	NT	0.6213	0.5654	0.7896	0.6589	0.6287	0.6397	0.6287	0.5989
MLP	NO	0.6190	0.5633	0.7904	0.6578	0.6289	0.6403	0.6289	0.5994
MLP	CLAHE	0.6213	0.5650	0.7937	0.6601	0.6294	0.6404	0.6294	0.6
MLP	HE	0.6194	0.5634	0.7929	0.6587	0.6289	0.64	0.6289	0.5993
MLP	UM	0.6194	0.5636	0.7904	0.6580	0.6283	0.6394	0.6283	0.5984
MLP	CS	0.6201	0.5641	0.7921	0.6589	0.63	0.6414	0.63	0.6006
Model	Technique	CBIS-DDSM							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
MLP	NT	0.9708	0.6136	0.9643	0.75	0.6014	0.6191	0.6014	0.6037
MLP	NO	0.9627	0.5581	0.8571	0.6761	0.5097	0.5288	0.5097	0.5125
MLP	CLAHE	0.9692	0.6	0.9643	0.7397	0.5797	0.5996	0.5797	0.5813
MLP	HE	0.9579	0.5192	0.9643	0.675	0.5386	0.5733	0.5386	0.5412
MLP	UM	0.9611	0.54	0.9643	0.6923	0.4952	0.6015	0.4952	0.492
MLP	CS	0.9708	0.6136	0.9643	0.75	0.5821	0.5867	0.5821	0.5799

Tabla 14 El modelo KNN aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS							
		Bi-class FOF				Multi-class FOF			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
KNN	NT	0.7631	0.7197	0.8667	0.7864	0.6476	0.6597	0.6476	0.6474
KNN	NO	0.7589	0.7163	0.8625	0.7826	0.6476	0.6579	0.6476	0.6478
KNN	CLAHE	0.7526	0.7089	0.8625	0.7782	0.649	0.663	0.649	0.6487
KNN	HE	0.761	0.7172	0.8667	0.7849	0.6517	0.6679	0.6517	0.652
KNN	UM	0.7589	0.7178	0.8583	0.7818	0.6462	0.6567	0.6462	0.6458
KNN	CS	0.7568	0.7168	0.8542	0.7795	0.6448	0.6589	0.6448	0.6449
Model	Technique	Mini-MIAS							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
KNN	NT	0.6583	0.6442	0.7167	0.6785	0.6364	0.6328	0.6364	0.6298
KNN	NO	0.6625	0.6491	0.7167	0.6812	0.6308	0.6268	0.6308	0.6249
KNN	CLAHE	0.6625	0.6479	0.7208	0.6824	0.6322	0.6273	0.6322	0.6255
KNN	HE	0.6646	0.6504	0.7208	0.6838	0.628	0.6245	0.628	0.6221
KNN	UM	0.6646	0.6504	0.7208	0.6838	0.6378	0.6346	0.6378	0.6319
KNN	CS	0.6625	0.6479	0.7208	0.6824	0.6322	0.6279	0.6322	0.6261
Model	Technique	Mini-MIAS							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
KNN	NT	0.4667	0.672	0.4667	0.5378	0.5645	0.3636	0.381	0.3721
KNN	NO	0.4333	0.5461	0.4333	0.4798	0.5938	0.4545	0.6522	0.5357
KNN	CLAHE	0.5333	0.6306	0.5333	0.5721	0.65	0.5	0.6667	0.5714
KNN	HE	0.45	0.6357	0.45	0.5087	0.5938	0.44	0.4783	0.4583
KNN	UM	0.5	0.582	0.5	0.5306	0.4844	0.3438	0.4783	0.4
KNN	CS	0.5667	0.6854	0.5667	0.6059	0.4355	0.25	0.3333	0.2857
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
KNN	NT	0.7048	0.6666	0.7259	0.6950	0.7584	0.755	0.7584	0.7529
KNN	NO	0.7040	0.6658	0.7250	0.6942	0.7578	0.7545	0.7578	0.7522
KNN	CLAHE	0.7064	0.6666	0.7326	0.6980	0.7576	0.7546	0.7576	0.7518
KNN	HE	0.7033	0.6651	0.7242	0.6934	0.755	0.7516	0.755	0.7492
KNN	UM	0.7052	0.6669	0.7267	0.6955	0.7574	0.7538	0.7574	0.7518
KNN	CS	0.7071	0.6694	0.7267	0.6969	0.7565	0.7532	0.7565	0.7508
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
KNN	NT	0.6213	0.5654	0.7896	0.6589	0.6287	0.6397	0.6287	0.5989
KNN	NO	0.6190	0.5633	0.7904	0.6578	0.6289	0.6403	0.6289	0.5994
KNN	CLAHE	0.6213	0.5650	0.7937	0.6601	0.6294	0.6404	0.6294	0.6
KNN	HE	0.6194	0.5634	0.7929	0.6587	0.6289	0.64	0.6289	0.5993
KNN	UM	0.6194	0.5636	0.7904	0.6580	0.6283	0.6394	0.6283	0.5984
KNN	CS	0.6201	0.5641	0.7921	0.6589	0.63	0.6414	0.63	0.6006
Model	Technique	CBIS-DDSM							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
KNN	NT	0.9724	0.6341	0.9286	0.7536	0.401	0.5424	0.401	0.3437
KNN	NO	0.9611	0.55	0.7857	0.6471	0.6135	0.6415	0.6135	0.6166
KNN	CLAHE	0.9643	0.5682	0.8929	0.6944	0.5604	0.5874	0.5604	0.563
KNN	HE	0.9579	0.5192	0.9643	0.675	0.529	0.5707	0.529	0.5315
KNN	UM	0.9562	0.5091	1	0.6747	0.401	0.5424	0.401	0.3437
KNN	CS	0.9692	0.6047	0.9286	0.7324	0.6014	0.6103	0.6014	0.6008

Tabla 15 El modelo RF aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS							
		Bi-class FOF				Multi-class FOF			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
RF	NT	0.7631	0.7197	0.8667	0.7864	0.6476	0.6597	0.6476	0.6474
RF	NO	0.7589	0.7163	0.8625	0.7826	0.6476	0.6579	0.6476	0.6478
RF	CLAHE	0.7526	0.7089	0.8625	0.7782	0.649	0.663	0.649	0.6487
RF	HE	0.761	0.7172	0.8667	0.7849	0.6517	0.6679	0.6517	0.652
RF	UM	0.7589	0.7178	0.8583	0.7818	0.6462	0.6567	0.6462	0.6458
RF	CS	0.7568	0.7168	0.8542	0.7795	0.6448	0.6589	0.6448	0.6449
Model	Technique	Mini-MIAS							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
RF	NT	0.6583	0.6442	0.7167	0.6785	0.6364	0.6328	0.6364	0.6298
RF	NO	0.6625	0.6491	0.7167	0.6812	0.6308	0.6268	0.6308	0.6249
RF	CLAHE	0.6625	0.6479	0.7208	0.6824	0.6322	0.6273	0.6322	0.6255
RF	HE	0.6646	0.6504	0.7208	0.6838	0.628	0.6245	0.628	0.6221
RF	UM	0.6646	0.6504	0.7208	0.6838	0.6378	0.6346	0.6378	0.6319
RF	CS	0.6625	0.6479	0.7208	0.6824	0.6322	0.6279	0.6322	0.6261
Model	Technique	Mini-MIAS							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
RF	NT	0.5313	0.6382	0.5313	0.571	0.5484	0.3939	0.619	0.4815
RF	NO	0.5313	0.6522	0.5313	0.5714	0.6094	0.4706	0.6957	0.5614
RF	CLAHE	0.4844	0.5762	0.4844	0.5175	0.6167	0.4643	0.619	0.5306
RF	HE	0.4531	0.6086	0.4531	0.4762	0.5469	0.3929	0.4783	0.4314
RF	UM	0.5	0.5625	0.5	0.5133	0.5	0.3636	0.5217	0.4286
RF	CS	0.375	0.5185	0.375	0.4198	0.5161	0.3636	0.5714	0.4444
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
RF	NT	0.7048	0.6666	0.7259	0.6950	0.7584	0.755	0.7584	0.7529
RF	NO	0.7040	0.6658	0.7250	0.6942	0.7578	0.7545	0.7578	0.7522
RF	CLAHE	0.7064	0.6666	0.7326	0.6980	0.7576	0.7546	0.7576	0.7518
RF	HE	0.7033	0.6651	0.7242	0.6934	0.755	0.7516	0.755	0.7492
RF	UM	0.7052	0.6669	0.7267	0.6955	0.7574	0.7538	0.7574	0.7518
RF	CS	0.7071	0.6694	0.7267	0.6969	0.7565	0.7532	0.7565	0.7508
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
RF	NT	0.6213	0.5654	0.7896	0.6589	0.6287	0.6397	0.6287	0.5989
RF	NO	0.6190	0.5633	0.7904	0.6578	0.6289	0.6403	0.6289	0.5994
RF	CLAHE	0.6213	0.5650	0.7937	0.6601	0.6294	0.6404	0.6294	0.6
RF	HE	0.6194	0.5634	0.7929	0.6587	0.6289	0.64	0.6289	0.5993
RF	UM	0.6194	0.5636	0.7904	0.6580	0.6283	0.6394	0.6283	0.5984
RF	CS	0.6201	0.5641	0.7921	0.6589	0.63	0.6414	0.63	0.6006
Model	Technique	CBIS-DDSM							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
RF	NT	0.9319	0.3939	0.9286	0.5532	0.5604	0.6065	0.5604	0.5661
RF	NO	0.953	0.4902	0.8929	0.6329	0.5072	0.5334	0.5072	0.5104
RF	CLAHE	0.9433	0.4364	0.8571	0.5783	0.4614	0.5046	0.4614	0.4647
RF	HE	0.953	0.4902	0.8929	0.6329	0.9254	0.3714	0.9286	0.5306
RF	UM	0.919	0.3553	0.9643	0.5192	0.4879	0.5867	0.4879	0.4901
RF	CS	0.9465	0.4545	0.8929	0.6024	0.5411	0.5739	0.5411	0.5444

Tabla 16 El modelo DT aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS							
		Bi-class FOF				Multi-class FOF			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
DT	NT	0.7631	0.7197	0.8667	0.7864	0.6476	0.6597	0.6476	0.6474
DT	NO	0.7589	0.7163	0.8625	0.7826	0.6476	0.6579	0.6476	0.6478
DT	CLAHE	0.7526	0.7089	0.8625	0.7782	0.649	0.663	0.649	0.6487
DT	HE	0.761	0.7172	0.8667	0.7849	0.6517	0.6679	0.6517	0.652
DT	UM	0.7589	0.7178	0.8583	0.7818	0.6462	0.6567	0.6462	0.6458
DT	CS	0.7568	0.7168	0.8542	0.7795	0.6448	0.6589	0.6448	0.6449
Model	Technique	Mini-MIAS							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
DT	NT	0.6583	0.6442	0.7167	0.6785	0.6364	0.6328	0.6364	0.6298
DT	NO	0.6625	0.6491	0.7167	0.6812	0.6308	0.6268	0.6308	0.6249
DT	CLAHE	0.6625	0.6479	0.7208	0.6824	0.6322	0.6273	0.6322	0.6255
DT	HE	0.6646	0.6504	0.7208	0.6838	0.628	0.6245	0.628	0.6221
DT	UM	0.6646	0.6504	0.7208	0.6838	0.6378	0.6346	0.6378	0.6319
DT	CS	0.6625	0.6479	0.7208	0.6824	0.6322	0.6279	0.6322	0.6261
Model	Technique	Mini-MIAS							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
DT	NT	0.5156	0.6085	0.5156	0.5488	0.5161	0.3043	0.3333	0.318
DT	NO	0.4531	0.5066	0.4531	0.4751	0.5313	0.3704	0.4348	0.4
DT	CLAHE	0.4375	0.5161	0.4375	0.4681	0.65	0.5	0.4286	0.4615
DT	HE	0.4688	0.5911	0.4688	0.4915	0.6094	0.4667	0.6087	0.5283
DT	UM	0.4219	0.5615	0.4219	0.4669	0.5781	0.4286	0.5217	0.4706
DT	CS	0.375	0.475	0.375	0.4079	0.4355	0.25	0.3333	0.2857
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
DT	NT	0.7048	0.6666	0.7259	0.6950	0.7584	0.755	0.7584	0.7529
DT	NO	0.7040	0.6658	0.7250	0.6942	0.7578	0.7545	0.7578	0.7522
DT	CLAHE	0.7064	0.6666	0.7326	0.6980	0.7576	0.7546	0.7576	0.7518
DT	HE	0.7033	0.6651	0.7242	0.6934	0.755	0.7516	0.755	0.7492
DT	UM	0.7052	0.6669	0.7267	0.6955	0.7574	0.7538	0.7574	0.7518
DT	CS	0.7071	0.6694	0.7267	0.6969	0.7565	0.7532	0.7565	0.7508
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
DT	NT	0.6213	0.5654	0.7896	0.6589	0.6287	0.6397	0.6287	0.5989
DT	NO	0.6190	0.5633	0.7904	0.6578	0.6289	0.6403	0.6289	0.5994
DT	CLAHE	0.6213	0.5650	0.7937	0.6601	0.6294	0.6404	0.6294	0.6
DT	HE	0.6194	0.5634	0.7929	0.6587	0.6289	0.64	0.6289	0.5993
DT	UM	0.6194	0.5636	0.7904	0.6580	0.6283	0.6394	0.6283	0.5984
DT	CS	0.6201	0.5641	0.7921	0.6589	0.63	0.6414	0.63	0.6006
Model	Technique	CBIS-DDSM							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
DT	NT	0.9643	0.5625	0.9643	0.7105	0.5531	0.5945	0.5531	0.5555
DT	NO	0.9692	0.6154	0.8571	0.7164	0.5652	0.5854	0.5652	0.5699
DT	CLAHE	0.9627	0.5556	0.8929	0.6849	0.5121	0.5371	0.5121	0.5165
DT	HE	0.9514	0.4828	1	0.6512	0.4758	0.5226	0.4758	0.4776
DT	UM	0.9417	0.4375	1	0.6087	0.3937	0.5159	0.3937	0.35
DT	CS	0.9643	0.5652	0.9286	0.7027	0.529	0.5393	0.529	0.529

Tabla 17 El modelo GBC aplicado en cada base de datos y técnica.

Model	Technique	Mini-MIAS							
		Bi-class FOF				Multi-class FOF			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
GBC	NT	0.7631	0.7197	0.8667	0.7864	0.6476	0.6597	0.6476	0.6474
GBC	NO	0.7589	0.7163	0.8625	0.7826	0.6476	0.6579	0.6476	0.6478
GBC	CLAHE	0.7526	0.7089	0.8625	0.7782	0.649	0.663	0.649	0.6487
GBC	HE	0.761	0.7172	0.8667	0.7849	0.6517	0.6679	0.6517	0.652
GBC	UM	0.7589	0.7178	0.8583	0.7818	0.6462	0.6567	0.6462	0.6458
GBC	CS	0.7568	0.7168	0.8542	0.7795	0.6448	0.6589	0.6448	0.6449
Model	Technique	Mini-MIAS							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
GBC	NT	0.6583	0.6442	0.7167	0.6785	0.6364	0.6328	0.6364	0.6298
GBC	NO	0.6625	0.6491	0.7167	0.6812	0.6308	0.6268	0.6308	0.6249
GBC	CLAHE	0.6625	0.6479	0.7208	0.6824	0.6322	0.6273	0.6322	0.6255
GBC	HE	0.6646	0.6504	0.7208	0.6838	0.628	0.6245	0.628	0.6221
GBC	UM	0.6646	0.6504	0.7208	0.6838	0.6378	0.6346	0.6378	0.6319
GBC	CS	0.6625	0.6479	0.7208	0.6824	0.6322	0.6279	0.6322	0.6261
Model	Technique	Mini-MIAS							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
GBC	NT	0.6094	0.4667	0.6087	0.5283	0.4839	0.4543	0.4839	0.4682
GBC	NO	0.6129	0.4118	0.3333	0.3684	0.5	0.5323	0.5	0.5033
GBC	CLAHE	0.5469	0.4063	0.5652	0.4727	0.4032	0.4784	0.4032	0.4327
GBC	HE	0.5625	0.4138	0.5217	0.4615	0.4677	0.5903	0.4677	0.4987
GBC	UM	0.3871	0.1852	0.2381	0.2083	0.4194	0.4886	0.4194	0.4431
GBC	CS	0.6833	0.5556	0.4762	0.5128	0.4677	0.5321	0.4677	0.4926
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
GBC	NT	0.7048	0.6666	0.7259	0.6950	0.7584	0.755	0.7584	0.7529
GBC	NO	0.7040	0.6658	0.7250	0.6942	0.7578	0.7545	0.7578	0.7522
GBC	CLAHE	0.7064	0.6666	0.7326	0.6980	0.7576	0.7546	0.7576	0.7518
GBC	HE	0.7033	0.6651	0.7242	0.6934	0.755	0.7516	0.755	0.7492
GBC	UM	0.7052	0.6669	0.7267	0.6955	0.7574	0.7538	0.7574	0.7518
GBC	CS	0.7071	0.6694	0.7267	0.6969	0.7565	0.7532	0.7565	0.7508
Model	Technique	CBIS-DDSM							
		Bi-class GLCM				Multi-class GLCM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
GBC	NT	0.6213	0.5654	0.7896	0.6589	0.6287	0.6397	0.6287	0.5989
GBC	NO	0.6190	0.5633	0.7904	0.6578	0.6289	0.6403	0.6289	0.5994
GBC	CLAHE	0.6213	0.5650	0.7937	0.6601	0.6294	0.6404	0.6294	0.6
GBC	HE	0.6194	0.5634	0.7929	0.6587	0.6289	0.64	0.6289	0.5993
GBC	UM	0.6194	0.5636	0.7904	0.6580	0.6283	0.6394	0.6283	0.5984
GBC	CS	0.6201	0.5641	0.7921	0.6589	0.63	0.6414	0.63	0.6006
Model	Technique	CBIS-DDSM							
		Bi-class GLRLM				Multi-class GLRLM			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
GBC	NT	0.9643	0.56	1	0.7179	0.57	0.5866	0.57	0.5692
GBC	NO	0.9643	0.5652	0.9286	0.7027	0.5242	0.5311	0.5242	0.5242
GBC	CLAHE	0.9643	0.56	1	0.7179	0.5483	0.579	0.5483	0.5522
GBC	HE	0.9481	0.4667	1	0.6364	0.5266	0.5616	0.5266	0.5303
GBC	UM	0.9514	0.4828	1	0.651	0.4348	0.5534	0.4348	0.412
GBC	CS	0.9627	0.551	0.9643	0.7013	0.5556	0.5719	0.5556	0.5561

5.4 Resultados Metodología 2

Se dividió el conjunto de datos preprocesados 80% en datos de entrenamiento, 10% en datos de validación y 10% en datos de prueba. Usamos los siguientes parámetros para las redes neuronales VGG19, ResNet50, ResNet151 y EfficientNetB7 para la clasificación: usamos el optimizador Adam con una tasa de aprendizaje de 0.0001, establecimos el tamaño del lote en 32 y usamos entropía cruzada binaria y categórica como la pérdida funciones. El proceso de entrenamiento duró 20 épocas.

Así mismo como en la metodología 1 se cambió la red neuronal de clasificación para adaptar esta investigación.

Para el conjunto de entrenamiento, se usó aumento de datos. El aumento de datos introduce variaciones en el modelo al aplicar cambios a los conjuntos de datos. El aumento de datos aumenta la solidez del aprendizaje automático y reduce los gastos operativos.

La Tabla 18 muestra el aumento de datos para las imágenes de la etapa 1. Esta ampliación se realizó para imágenes normales y tumorales en el conjunto de datos Mini-MIAS.

Tabla 18 Usando la técnica de aumento de datos en el conjunto de datos de entrenamiento de la etapa 1.

Sin aumento de datos			
Categoría	Conjunto de entrenamiento	Conjunto de validación.	Conjunto de prueba.
Normal	165	22	22
Tumor	82	11	11
Con aumento de datos			
Categoría	Conjunto de entrenamiento	Conjunto de validación.	Conjunto de prueba.
Normal	4648	720	22
Tumor	4648	560	11

La Tabla 19 muestra el aumento de datos para las imágenes de la etapa 2. Este aumento en el número de imágenes se hace de forma sintética, y en las imágenes sanas se tuvo que agregar más transformaciones de estas imágenes para igualar las imágenes anormales. En este caso se utilizaron las dos bases de datos Mini-MIAS y CBIS-DDSM.

La etapa 2 es notoria por la falta de información. Se usaron métricas adicionales para evaluar el rendimiento real de estos dos modelos en la investigación porque hay menos imágenes normales.

Tabla 19 Usando la técnica de aumento de datos en el conjunto de datos de entrenamiento de la etapa 2.

Sin aumento de datos			
Categoría	Conjunto de entrenamiento	Conjunto de validación.	Conjunto de prueba.
Normal	209	20	22
Abnormal	2,308	288	290
Con aumento de datos			
Categoría	Conjunto de entrenamiento	Conjunto de validación.	Conjunto de prueba.
Normal	18,464	20	22
Abnormal	18,464	288	290

En la Tabla 20 muestra el aumento de datos para las imágenes del estadio 3. Para este estadio, las imágenes que presentaban lesiones fueron aumentadas y etiquetadas como "masas y calcificaciones". Solo para esta etapa se utilizó el conjunto de datos CBIS-DDSM.

En la Tabla 21 se muestra el aumento de datos para las imágenes del estadio 4. Las imágenes de tejido normal fueron aumentadas y rotuladas para este estadio, con lesiones como masas y calcificaciones. Para este problema multiclase se

utilizaron los conjuntos de datos Mini-MIAS para imágenes normales y el CBIS-DDSM para imágenes con lesiones mamarias.

Tabla 20 Usando la técnica de aumento de datos en el conjunto de datos de entrenamiento de la etapa 3.

Sin aumento de datos			
Categoría	Conjunto de entrenamiento	Conjunto de validación.	Conjunto de prueba.
Masses	1,244	155	156
Calcifications	1,064	133	134
Con aumento de datos			
Categoría	Conjunto de entrenamiento	Conjunto de validación.	Conjunto de prueba.
Masses	24,880	155	156
Calcifications	24,880	133	134

Antes de que DCNN pudiera utilizar el conjunto de datos, la cantidad de imágenes fue distribuida equitativamente para mantener el conjunto de datos equilibrado. Las Tablas 2, 4 y 5 requerían mayores aumentos para las imágenes normales debido a que había menos de ellas en comparación con las otras imágenes de las otras clases, lo que afectaría la capacidad de aprendizaje de DCNN. A excepción de las imágenes de validación de la tabla. Mesa. 1, que eran diferentes porque carecían de información, las imágenes de prueba y validación fueron similares para las demás etapas. La Tabla 22 Los valores obtenidos como resultado del entrenamiento se muestran en esta tabla. El final resultado de cada etapa, así como de cada DCNN, solo se observa con imágenes normalizadas (NO).muestra los resultados utilizando el conjunto de datos normalizados y las métricas de rendimiento especificadas de exactitud, precisión, recuperación, puntuación F1 y AUC. La Tabla 23 Los valores obtenidos como resultado del entrenamiento se muestran en esta tabla. Cada resultado final de la etapa se observa junto con cada DCNN utilizando la técnica CLAHE. mostró los resultados utilizando el conjunto de datos con la técnica de preprocesamiento CLAHE. Para comprobar el rendimiento, se utilizaron

las siguientes métricas: exactitud, precisión, recuperación, puntuación F1 y AUC. Se usaron diferentes modelos pre-entrenados y se compararon cada etapa. Se mostraron los resultados generales de prueba y los resultados de precisión, recuperación, puntuación F1 y AUC de cada clase. Todo lo que se mostró para la prueba de las dos clases fue la precisión promedio y los valores de AUC. Estos valores fueron los mismos, pero los valores de AUC individuales para cada clase se presentaron en pruebas con múltiples clases.

Los resultados obtenidos cuando las imágenes solo se normalizaron fueron buenos en comparación con las imágenes con la técnica CLAHE, como se mostró en la Tabla 22 y Tabla 23. Así lo demostró VGG19, que en promedio produjo mejores resultados cuando las imágenes solo se enviaron para la normalización de píxeles con un rango de 0-255. Sin embargo, en otras etapas de esta investigación se utilizó CLAHE para lograr desempeños notablemente altos, por ejemplo, en la Tabla 6. En comparación con otros modelos, los modelos ResNet50 y EfficientNetB7 lograron excelentes resultados. Esto mostró que se logró una precisión de hasta el 99 % utilizando los dos modelos indicados anteriormente junto con el ajuste fino que se realizó para este problema en la etapa 3.

Tabla 21 Usando la técnica de aumento de datos en el conjunto de datos de entrenamiento de la etapa 4.

Sin aumento de datos			
Categoría	Conjunto de entrenamiento	Conjunto de validación.	Conjunto de prueba.
Normal	209	220	22
Masses	1,244	155	156
Calcifications	1,064	133	134
Con aumento de datos			
Categoría	Conjunto de entrenamiento	Conjunto de validación.	Conjunto de prueba.
Normal	19,904	20	22
Masses	19,904	155	156
Calcifications	19,904	133	134

Tabla 22 Los valores obtenidos como resultado del entrenamiento se muestran en esta tabla. El final resultado de cada etapa, así como de cada DCNN, solo se observa con imágenes normalizadas (NO).

DCNN	Class	DCNN Performance using normalise				
		Accuracy	Precision	Recall	F1-score	AUC
Mini-MIAS						
VGG19	Normal		0.9545	0.9545	0.9545	
	Tumor	0.93939	0.9091	0.9091	0.9091	0.9318
	Overall		0.9091	0.9091	0.9091	
ResNet50	Normal		0.9545	0.9545	0.9545	
	Tumor	0.9393	0.9091	0.9091	0.9091	0.9318
	Overall		0.9091	0.9091	0.9091	
ResNet152	Normal		0.9130	0.9545	0.9333	
	Tumor	0.9091	0.9	0.8181	0.8571	0.8863
	Overall		0.9	0.8181	0.8571	
EfficientNetB7	Normal		0.9565	1	0.9777	
	Tumor	0.9696	1	0.9091	0.9523	0.9545
	Overall		1	0.9091	0.9523	
CBIS-DDSM						
VGG19	Calcification		0.8875	0.8656	0.8787	
	Masses	0.8896	0.8875	0.9102	0.8987	0.8879
	Overall		0.8897	0.8896	0.8895	
ResNet50	Calcification		0.9147	0.8805	0.8973	
	Masses	0.9068	0.9006	0.9294	0.9148	0.9050
	Overall		0.9006	0.9294	0.9148	
ResNet152	Calcification		0.9076	0.8805	0.8939	
	Masses	0.9034	0.9	0.9230	0.9113	0.9018
	Overall		0.9	0.9230	0.9113	
EfficientNetB7	Calcification		0.9166	0.9029	0.9097	
	Masses	0.9172	0.9177	0.9294	0.9235	0.9162
	Overall		0.9177	0.9294	0.9235	
VGG19	Normal		0.9829	0.9829	0.9965	
	Abnormal	0.9807	0.9444	0.7727	0.85	0.8846
	Overall		0.9802	0.9807	0.9798	
ResNet50	Normal		0.9965	0.9965	0.9965	
	Abnormal	0.9935	0.9545	0.9545	0.9545	0.9755
	Overall		0.9545	0.9545	0.9545	
ResNet152	Normal		0.9931	1	0.9965	
	Abnormal	0.9935	0.9989	0.9090	0.9523	0.9545
	Overall		0.9090	0.9523	0.9931	
EfficientNetB7	Normal		0.9931	1	0.9965	
	Abnormal	0.9935	0.9989	0.9090	0.9523	0.9545
	Overall		0.9989	0.9090	0.9523	
Mini-MIAS + CBIS-DDSM						
VGG19	Normal		0.8636	0.8636	0.92664	0.9266
	Calcification	0.9253	0.8731	0.8731	0.8731	0.8888
	Masses		0.8974	0.8974	0.8974	0.8974
	Overall		0.8846	0.8846	0.884	0.9042
ResNet50	Normal		0.9987	0.9545	0.9767	0.9772
	Calcification	0.9006	0.9055	0.8582	0.8812	0.8953
	Masses		0.8841	0.9294	0.9062	0.9038
	Overall		0.9014	0.9006	0.9004	0.9255
ResNet152	Normal		0.9988	0.8636	0.9268	0.9318
	Calcification	0.9006	0.8776	0.9104	0.8937	0.9074
	Masses		0.9090	0.8974	0.9032	0.9038
	Overall		0.9020	0.9006	0.9008	0.9143
EfficientNetB7	Normal		1	0.7722	0.8717	0.8863
	Calcification	0.9102	0.875	0.9402	0.9064	0.9195
	Masses		0.9337	0.9038	0.9185	0.9198
	Overall		0.9132	0.9102	0.91007	0.9086

Tabla 23 Los valores obtenidos como resultado del entrenamiento se muestran en esta tabla. Cada resultado final de la etapa se observa junto con cada DCNN utilizando la técnica CLAHE.

DCNN	Class	DCNN Performance using CLAHE				
		Accuracy	Precision	Recall	F1-score	AUC
Mini-MIAS						
VGG19	Normal		0.9545	1	0.9767	
	Tumor	0.9696	1	0.9166	0.9565	0.9583
	Overall		0.9166	0.9565	0.9545	
ResNet50	Normal		0.8333	0.8695	0.9091	
	Tumor	0.9091	0.9091	0.9523	0.9302	0.8928
	Overall		0.9091	0.8333	0.8695	
ResNet152	Normal		0.8333	0.9091	0.9130	
	Tumor	0.9393	0.913	1	0.9545	0.9166
	Overall		1	0.8333	0.9091	
EfficientNetB7	Normal		0.9166	0.9166	0.9523	
	Tumor	0.9393	0.9523	0.9523	0.9523	0.9345
	Overall		0.9166	0.9166	0.9166	
CBIS-DDSM						
VGG19	Calcification		0.8617	0.7910	0.8249	
	Masses	0.8448	0.8323	0.8910	0.8606	0.8410
	Overall		0.8459	0.8448	0.8441	
ResNet50	Calcification		0.8759	0.8955	0.8856	
	Masses	0.8931	0.908	0.8910	0.8996	0.8932
	Overall		0.90844	0.8910	0.8996	
ResNet152	Calcification		0.9076	0.8805	0.8939	
	Masses	0.9034	0.9	0.9230	0.9113	0.9018
	Overall		0.9	0.9230	0.9113	
EfficientNetB7	Calcification		0.8978	0.9179	0.9077	
	Masses	0.9137	0.9281	0.9102	0.9190	0.9140
	Overall		0.9281	0.9102	0.9190	
VGG19	Normal		0.9830	1	0.9914	
	Abnormal	0.9839	1	0.7727	0.8717	0.8863
	Overall		0.7727	0.8717	0.9830	
ResNet50	Normal		0.9989	0.9989	1	
	Abnormal	1	0.9989	0.9989	1	1
	Overall		0.9989	1	0.9989	
ResNet152	Normal		0.9931	1	0.9965	
	Abnormal	0.9935	1	0.9090	0.9523	0.9545
	Overall		0.9090	0.9523	0.9931	
EfficientNetB7	Normal		1	0.9989	1	
	Abnormal	1	0.9989	0.9989	1	1
	Overall		1	0.9989	1	
Mini-MIAS + CBIS-DDSM						
VGG19	Normal		1	0.9090	0.9523	0.95454
	Calcification	0.8557	0.856	0.7985	0.8262	0.8486
	Masses		0.8383	0.8974	0.8668	0.8621
	Overall		0.8573	0.8557	0.8554	0.8884
ResNet50	Normal		1	0.9090	0.9523	0.9545
	Calcification	0.9166	0.9029	0.9029	0.9029	0.9149
	Masses		0.9177	0.9294	0.9235	0.9230
	Overall		0.9171	0.9166	0.9167	0.9308
ResNet152	Normal		0.9090	0.9090	0.9090	0.9510
	Calcification	0.8974	0.8931	0.8731	0.8830	0.8972
	Masses		0.8993	0.9166	0.9079	0.9070
	Overall		0.8973	0.8974	0.8973	0.9184
EfficientNetB7	Normal		1	0.9090	0.9523	0.9545
	Calcification	0.8974	0.9126	0.8582	0.8846	0.8982
	Masses		0.8734	0.9294	0.9006	0.8974
	Overall		0.8992	0.8974	0.8973	0.9167

TESIS TESIS TESIS TESIS TESIS

CAPITULO 6

En el capítulo 6 se mostrarán las conclusiones y trabajos futuros. Este estudio presentó una comparación usando once redes neuronales convolucionales preentrenadas en mamografías usando un ajuste fino en su clasificación. Para ambos casos se tuvo que modificar la parte de clasificación que fue un ajuste de las capas ocultas y la capa de salida que podía ser diferente entre un problema binario o un problema multiclase.

CONCLUSIONES Y TRABAJO FUTURO

6.1 Conclusión

El cáncer de mama es el cáncer más común entre las mujeres de todo el mundo. En este trabajo se demuestra la eficacia de un modelo DCNN para la clasificación y detección del cáncer de mama.

En resumen, el texto describe una metodología de preprocesamiento de imágenes mamográficas y la utilización de modelos de redes neuronales convolucionales para la detección de anomalías. La base de datos utilizada incluye imágenes de Mini-MIAS y CBIS-DDSM. El preprocesamiento incluye la eliminación de manchas negras y la recopilación de imágenes de interés (ROI) mediante un algoritmo. Se utilizó el algoritmo K-means y el método GLCM para excluir las imágenes con manchas negras. La clasificación se realizó mediante la extracción de características de textura estadísticas de segundo orden. La principal limitación para la detección de anomalías en mamografías es que las anomalías están presentes solo en un área pequeña de la imagen completa. Los modelos de redes neuronales convolucionales profundos se han diseñado con un tamaño de entrada estándar de 224x224 o 299x299 píxeles. Cambiar el tamaño de una mamografía completa puede resultar en la pérdida de información de la imagen.

Este estudio presentó una comparación usando once redes neuronales convolucionales preentrenadas en mamografías usando un ajuste fino en su

TESIS TESIS TESIS TESIS TESIS

clasificación. Para la primera metodología se hizo una comparación entre 10 de estos modelos y se comparó con varias técnicas de preprocesamiento para ver la efectividad de cada una de ellas. Los conjuntos de datos utilizados fueron Mini-MIAS y CBIS-DDSM.

Para la segunda metodología solamente se probaron cuatro estas arquitecturas que resultaron ser de las mejores y solamente se realizó el preprocesamiento con la técnica de CLAHE para aumentar la calidad de las imágenes para este caso se probaron cuatro etapas utilizando imágenes de Mini-MIAS y CBIS-DDSM. Se demostró que la tercera etapa que fue la comparación binaria entre las imágenes con anomalías de CBIS-DDSM con las imágenes sanas de Mini-MIAS. Utilizando ResNet50 y EfficientB7 se obtuvo los mejores resultados.

Para ambos casos se tuvo que modificar la parte de clasificación que fue un ajuste de las capas ocultas y la capa de salida que podía ser diferente entre un problema binario o un problema multiclase.

El equipo que utilizamos en este estudio nos impuso algunas limitaciones. Pero para obtener resultados más rápidos y mejores comparaciones. En un futuro, estas tareas de clasificación se puedan realizar con equipos más capaces y especializados.

El objetivo de desarrollar un algoritmo de detección automática de lesiones en mamografías mediante el uso de inteligencia artificial se cumplió exitosamente. Se utilizaron redes neuronales convolucionales, junto con TensorFlow, para identificar y localizar posibles anomalías en las imágenes mamográficas con alta precisión y eficiencia. La implementación de este algoritmo puede mejorar el diagnóstico temprano del cáncer de mama, lo que puede ser fundamental para el tratamiento y la recuperación de las pacientes.

La metodología para la detección de lesiones en mamografías se basó en una combinación de algoritmos que permitió la extracción de características y la clasificación de las imágenes médicas. Se utilizaron múltiples algoritmos, incluyendo aquellos para aprendizaje automático que usaron estadísticas de primer y segundo

orden, así como redes neuronales convolucionales. Todo el proceso se implementó en python y se utilizaron varios algoritmos adicionales para funciones misceláneas.

En cuanto a la evaluación de desempeño, se probaron varios algoritmos de aprendizaje automático, incluyendo aprendizaje profundo, y las redes neuronales convolucionales obtuvieron mejores resultados según la arquitectura seleccionada. Además, se utilizó una combinación de dos bases de datos, Mini-MIAS y CBIS-DDSM, para separar las imágenes dependiendo de las lesiones o la falta de ellas.

Finalmente, se integraron todos los módulos del sistema para su validación, desde la recolección de datos hasta la evaluación del resultado. Se realizaron pruebas exhaustivas y se ajustó el sistema según sea necesario para garantizar su precisión y eficiencia en la detección automática de lesiones en mamografías.

6.2 Trabajo futuro.

Lo que se planea en un futuro es lo siguiente:

- Utilizar este sistema en mamografías de una base datos de México.
- Mejorar el código y cambiarlo para que sea mucho más sencillo al utilizarlo en campo.
- Utilizar otros métodos de clasificación y no solamente redes neuronales artificiales.
- Agregar al CAD una parte utilizando el sistema YOLOv8 para un sistema más robusto y confiable de clasificación.
- Buscar modelos más modernos de clasificación de imágenes. Arquitecturas más actuales y probarlas.

6.3 Preguntas de investigación

¿Qué tipo de técnicas de preprocesamiento son útiles para alimentar el modelo automático de detección?

Las técnicas de preprocesamiento CLAHE (Contrast Limited Adaptive Histogram Equalization) y Ecuador de histograma son útiles para mejorar la calidad de las imágenes mamográficas y prepararlas para ser alimentadas al modelo automático de detección. El CLAHE ajusta el contraste localmente para mejorar la visibilidad de estructuras anatómicas relevantes, mientras que el Ecuador de histograma ajusta la distribución de intensidad de píxeles en la imagen para mejorar la visualización de detalles sutiles. Al aplicar estas técnicas de preprocesamiento, se puede aumentar la eficacia del modelo de detección automática de lesiones en mamografías y reducir la tasa de falsos positivos y falsos negativos.

CLAHE tiene la capacidad de limitar la amplificación de contraste en áreas de la imagen con grandes gradientes de intensidad, lo que puede evitar la sobreexposición y la subexposición en la imagen procesada. Esto significa que CLAHE puede proporcionar una imagen con un contraste mejorado y detalles más claros en áreas específicas, lo que puede ser especialmente útil en imágenes médicas donde se necesita una precisión y detalle máximo en la detección de anomalías.

CLAHE fue más efectivo que otras técnicas de preprocesamiento utilizadas en la primera y segunda metodología, para mejorar la precisión de la detección automática de lesiones en mamografías. Además, el hecho de que CLAHE haya permitido obtener un resultado de detección de anomalías superior al 98% indica que esta técnica de preprocesamiento puede ser altamente efectiva en la detección temprana del cáncer de mama.

¿En qué podría ayudar el aprendizaje automático o el aprendizaje profundo en la detección temprana de cáncer de mama?

Existen varias formas en las que el aprendizaje automático y profundo podrían contribuir a la detección temprana de cáncer de mama.

- **Detección automática de anomalías en mamografías:** Los modelos de aprendizaje automático y profundo pueden ser entrenados para detectar patrones sutiles en las mamografías que puedan indicar la presencia de una anomalía o lesión sospechosa. Esto podría ayudar a detectar tumores en las etapas iniciales, lo que aumentaría las posibilidades de un tratamiento exitoso.
- **Identificación de patrones de crecimiento del tumor:** Los modelos de aprendizaje automático y profundo también pueden ser entrenados para analizar las imágenes de seguimiento de mamografías a lo largo del tiempo y detectar patrones en el crecimiento del tumor. Esto podría ayudar a los médicos a determinar la gravedad del cáncer y desarrollar un plan de tratamiento adecuado.
- **Selección de pacientes para pruebas adicionales:** El aprendizaje automático y el aprendizaje profundo también pueden ayudar a seleccionar a las pacientes que pueden necesitar pruebas adicionales, como una biopsia, para determinar si tienen cáncer de mama. Esto podría reducir la cantidad de pruebas innecesarias y ayudar a los médicos a enfocar su atención en las pacientes que realmente necesitan pruebas adicionales.

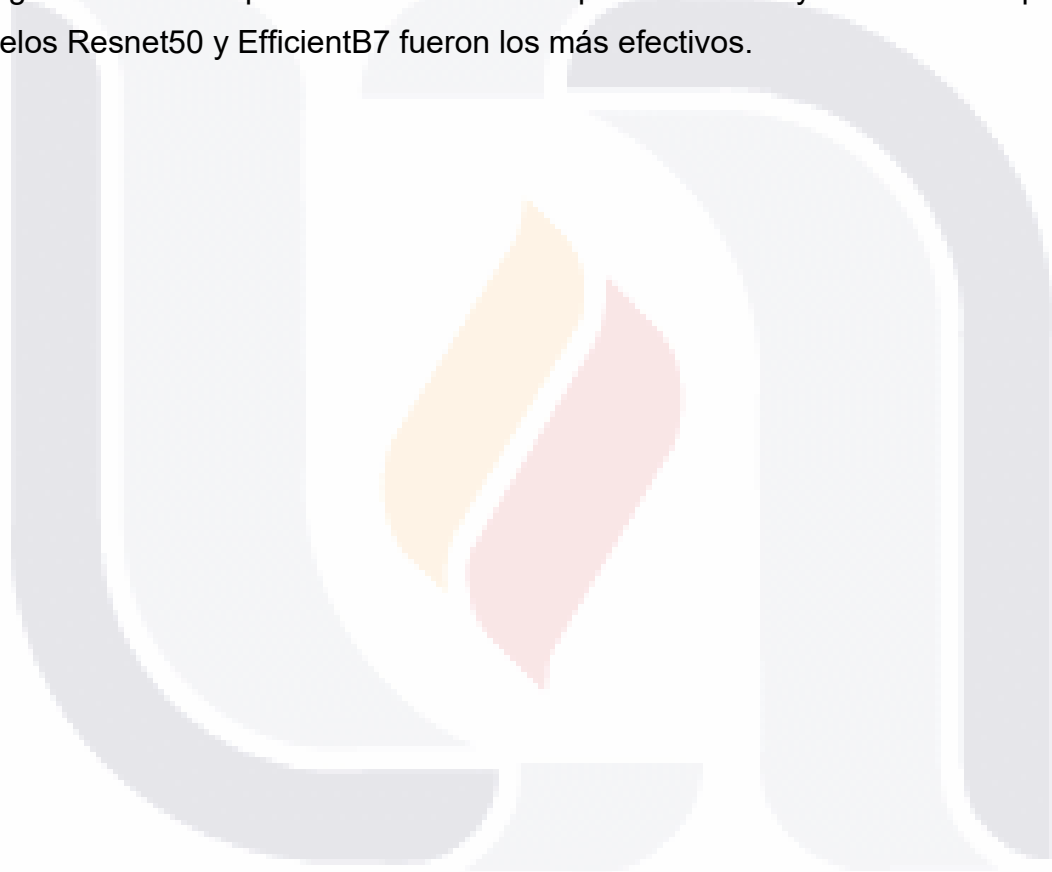
Si existen varios tipos de modelos que permiten detectar el cáncer, ¿Cuál sería el más eficiente?

No hay un modelo único que sea el más eficiente para detectar el cáncer de mama, ya que cada modelo tiene sus propias fortalezas y debilidades y su eficiencia dependerá del conjunto de datos y de las características específicas de cada caso de detección de cáncer. Además, la eficiencia de un modelo también dependerá de la calidad y cantidad de los datos que se le proporcionen para su entrenamiento.

En general, los modelos de aprendizaje profundo, como DenseNet, VGG, ResNet, InceptionV3, MobileNet y EfficientNetB7, son modelos muy populares y utilizados en la detección de cáncer de mama, debido a su capacidad para detectar patrones

sutiles y complejos en las imágenes médicas. Sin embargo, cada modelo tiene su propia arquitectura y características específicas, lo que significa que pueden ser más o menos eficientes en diferentes situaciones y conjuntos de datos.

En este estudio, se evaluó la eficacia de un modelo de red neuronal convolucional profundo para la detección de anomalías en mamografías de Mini-MIAS y CBIS-DDSM. Se aplicó una metodología de preprocesamiento de imágenes que incluyó la eliminación de manchas negras y la selección de imágenes de interés mediante un algoritmo. Se compararon once modelos preentrenados y se demostró que los modelos Resnet50 y EfficientB7 fueron los más efectivos.



BIBLIOGRAFÍA

- [1] C. Villarreal-Garza *et al.*, “Breast Cancer in Young Women in Latin America: An Unmet, Growing Burden,” *Oncologist*, vol. 18, no. 12, pp. 1298–1306, Dec. 2013, doi: 10.1634/theoncologist.2013-0321.
- [2] C. Villarreal-Garza, F. Mesa-Chavez, A. Plata de la Mora, M. Miaja-Avila, M. Garcia-Garcia, A. Fonseca, S. de la Rosa-Pacheco, M. Cruz-Ramos, M. R. García Garza, A. Mohar, and E. Bargallo-Rocha, “Prospective Study of Fertility Preservation in Young Women With Breast Cancer in Mexico,” *Journal of the National Comprehensive Cancer Network*, pp. 1–8, Oct. 2021, doi: 10.6004/jnccn.2020.7692.
- [3] Y. Chávarri-Guerra, C. Villarreal-Garza, P. E. R. Liedke, F. Knaul, A. Mohar, D. M. Finkelstein, and P. E. Goss, “Breast cancer in Mexico: A growing challenge to health and the health system,” *The Lancet Oncology*, vol. 13, no. 8. Aug. 2012. doi: 10.1016/S1470-2045(12)70246-2.
- [4] E. H. Houssein, M. M. Emam, A. A. Ali, and P. N. Suganthan, “Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review,” *Expert Systems with Applications*, vol. 167. Elsevier Ltd, Apr. 01, 2021. doi: 10.1016/j.eswa.2020.114161.
- [5] E. D. Pisano and M. J. Yaffe, “Digital mammography,” *Radiology*, vol. 234, no. 2. pp. 353–362, Feb. 2005. doi: 10.1148/radiol.2342030897.
- [6] J. Wang, H. Zhu, S. H. Wang, and Y. D. Zhang, “A Review of Deep Learning on Medical Image Analysis,” *Mobile Networks and Applications*, vol. 26, no. 1, pp. 351–380, Feb. 2021, doi: 10.1007/s11036-020-01672-7.

- TESIS TESIS TESIS TESIS TESIS
- [7] S. Dong, P. Wang, and K. Abbas, "A survey on deep learning and its applications," *Computer Science Review*, vol. 40. Elsevier Ireland Ltd, May 01, 2021. doi: 10.1016/j.cosrev.2021.100379.
- [8] X. Zhang and W. Dahu, "Application of artificial intelligence algorithms in image processing," *J Vis Commun Image Represent*, vol. 61, pp. 42–49, May 2019, doi: 10.1016/j.jvcir.2019.03.004.
- [9] D. Hassabis, D. Kumaran, C. Summerfield, and M. Botvinick, "Neuroscience-Inspired Artificial Intelligence," *Neuron*, vol. 95, no. 2. Cell Press, pp. 245–258, Jul. 19, 2017. doi: 10.1016/j.neuron.2017.06.011.
- [10] S. Bagchi and A. Huong, "Signal Processing Techniques and Computer-Aided Detection Systems for Diagnosis of Breast Cancer – A Review Paper," *Indian J Sci Technol*, vol. 10, no. 3, Jan. 2017, doi: 10.17485/ijst/2017/v10i3/110640.
- [11] S. Batchu, F. Liu, A. Amireh, J. Waller, and M. Umair, "A Review of Applications of Machine Learning in Mammography and Future Challenges," *Oncology (Switzerland)*, vol. 99, no. 8. S. Karger AG, pp. 483–490, Jul. 01, 2021. doi: 10.1159/000515698.
- [12] A. K. Mohanty, M. R. Senapati, S. Beberta, and S. K. Lenka, "Texture-based features for classification of mammograms using decision tree," *Neural Comput Appl*, vol. 23, no. 3–4, pp. 1011–1017, Sep. 2013, doi: 10.1007/s00521-012-1025-z.
- [13] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27–48, Apr. 2016, doi: 10.1016/j.neucom.2015.09.116.
- [14] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electronic Markets*, vol. 31, pp. 685–695, 2021, doi: 10.1007/s12525-021-00475-2/Published.

- TESIS TESIS TESIS TESIS TESIS
- [15] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, May 27, 2015. doi: 10.1038/nature14539.
- [16] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going Deeper with Convolutions," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Sep. 2014, [Online]. Available: <http://arxiv.org/abs/1409.4842>
- [17] A. Fourcade and R. H. Khonsari, "Deep learning in medical image analysis: A third eye for doctors," *J Stomatol Oral Maxillofac Surg*, vol. 120, no. 4, pp. 279–288, Sep. 2019, doi: 10.1016/j.jormas.2019.06.002.
- [18] P. Mamoshina, A. Vieira, E. Putin, and A. Zhavoronkov, "Applications of Deep Learning in Biomedicine," *Molecular Pharmaceutics*, vol. 13, no. 5. American Chemical Society, pp. 1445–1454, May 02, 2016. doi: 10.1021/acs.molpharmaceut.5b00982.
- [19] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, and J. Dean, "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, no. 1. Nature Publishing Group, pp. 24–29, Jan. 01, 2019. doi: 10.1038/s41591-018-0316-z.
- [20] R. S. Lee, F. Gimenez, A. Hoogi, K. K. Miyake, M. Gorovoy, and D. L. Rubin, "A curated mammography data set for use in computer-aided detection and diagnosis research," *Sci Data*, vol. 4, Dec. 2017, doi: 10.1038/sdata.2017.177.
- [21] J. Suckling *et al.*, "The mammographic image analysis society digital mammogram database," 2015.
- [22] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in*

- Bioinformatics*), vol. 11141 LNCS, pp. 270–279, 2018, doi: 10.1007/978-3-030-01424-7_27.
- [23] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, S. Member, H. Xiong, and Q. He, “A Comprehensive Survey on Transfer Learning,” *Proceedings of the IEEE*, vol. 109, pp. 43–76, 2021.
- [24] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Trans Knowl Data Eng*, vol. 22, no. 10, pp. 1345–1359, 2010, doi: 10.1109/TKDE.2009.191.
- [25] L. G. Falconi, M. Perez, W. G. Aguilar, and A. Conci, “Transfer learning and fine tuning in breast mammogram abnormalities classification on CBIS-DDSM database,” *Advances in Science, Technology and Engineering Systems*, vol. 5, no. 2, pp. 154–165, 2020, doi: 10.25046/aj050220.
- [26] A. Howard *et al.*, “Searching for MobileNetV3,” 2019.
- [27] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, and M. Andreetto, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” 2017.
- [28] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016, [Online]. Available: <http://image-net.org/challenges/LSVRC/2015/>
- [29] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks For Large-Scale Image recognition,” *3rd International Conference on Learning Representations*, 2015, [Online]. Available: <http://www.robots.ox.ac.uk/>
- [30] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely Connected Convolutional Networks,” 2018. [Online]. Available: <https://github.com/liuzhuang13/DenseNet>.

- TESIS TESIS TESIS TESIS TESIS
- [31] C. Szegedy, V. Vanhoucke, S. Ioffe, and J. Shlens, "Rethinking the Inception Architecture for Computer Vision," 2015.
- [32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," 2012. [Online]. Available: <http://code.google.com/p/cuda-convnet/>
- [33] M. Tan and Q. v Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *Proceedings of the 36th International Conference on Machine Learning*, pp. 6105–6114, 2019.
- [34] K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization," in *Graphics Gems*, Elsevier, 1994, pp. 474–485. doi: 10.1016/b978-0-12-336156-1.50061-6.
- [35] J. Sharma, J. K. Rai, and R. P. Tewari, "Identification of Pre-processing Technique for Enhancement of Mammogram Images," *International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom)*, 2014.
- [36] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, "Albumentations: Fast and flexible image augmentations," *Information (Switzerland)*, vol. 11, no. 2, Feb. 2020, doi: 10.3390/info11020125.
- [37] The American Cancer Society medical and editorial content team, "What Is Cancer," *cancer.org*.
- [38] Carlos Bolívar Fiallos Quinteros, "Detección automática de lesiones en mamografías usando técnicas de análisis de imágenes," 2015.
- [39] İ. Işıklı Esener, S. Ergin, and T. Yüksel, "A Genuine GLCM-based Feature Extraction for Breast Tissue Classification on Mammograms," *International Journal of Intelligent Systems and Applications in Engineering*, no. 4, pp. 124–129, 2016, [Online]. Available: www.atscience.org/IJISAE

- TESIS TESIS TESIS TESIS TESIS
- [40] K. Preetha and S. K. Jayanthi, "GLCM and GLRLM based Feature Extraction Technique in Mammogram Images," 2018. [Online]. Available: www.sciencepubco.com/index.php/IJET
- [41] A. Kumar Mohanty, S. Beberta, and S. K. Lenka, "Classifying Benign and Malignant Mass using GLCM and GLRLM based Texture Features from Mammogram," *International Journal of Engineering Research and Applications (IJERA)* www.ijera.com, vol. 1, pp. 687–693, [Online]. Available: www.ijera.com
- [42] B. Dai, R. C. Chen, S. Z. Zhu, and W. W. Zhang, "Using random forest algorithm for breast cancer diagnosis," in *Proceedings - 2018 International Symposium on Computer, Consumer and Control, IS3C 2018*, Feb. 2019, pp. 449–452. doi: 10.1109/IS3C.2018.00119.
- [43] A. Kolchev, D. Pasyukov, I. Egoshin, I. Kliouchkin, O. Pasyukova, and D. Tumakov, "YOLOv4-Based CNN Model versus Nested Contours Algorithm in the Suspicious Lesion Detection on the Mammography Image: A Direct Comparison in the Real Clinical Settings," *J Imaging*, vol. 8, no. 4, Apr. 2022, doi: 10.3390/jimaging8040088.
- [44] S. Li, M. Dong, G. Du, and X. Mu, "Attention Dense-U-Net for Automatic Breast Mass Segmentation in Digital Mammogram," *IEEE Access*, vol. 7, pp. 59037–59047, 2019, doi: 10.1109/ACCESS.2019.2914873.
- [45] N. Ravitha Rajalakshmi, R. Vidhyapriya, N. Elango, and N. Ramesh, "Deeply supervised U-Net for mass segmentation in digital mammograms," *Int J Imaging Syst Technol*, vol. 31, no. 1, pp. 59–71, Mar. 2021, doi: 10.1002/ima.22516.
- [46] R. Patgiri, S. Nayak, T. Akutota, and B. Paul, "Machine Learning: A Dark Side of Cancer Computing," Mar. 2019, [Online]. Available: <http://arxiv.org/abs/1903.07167>
- [47] A. Rampun, K. López-Linares, P. J. Morrow, B. W. Scotney, H. Wang, I. G. Ocaña, G. Maclair, R. Zwiggelaar, M. A. González Ballester, and I.

- Macía, "Breast pectoral muscle segmentation in mammograms using a modified holistically-nested edge detection network," *Med Image Anal*, vol. 57, pp. 1–17, Oct. 2019, doi: 10.1016/j.media.2019.06.007.
- [48] A. Altameem, C. Mahanty, R. C. Poonia, A. K. J. Saudagar, and R. Kumar, "Breast Cancer Detection in Mammography Images Using Deep Convolutional Neural Networks and Fuzzy Ensemble Modeling Techniques," *Diagnostics*, vol. 12, no. 8, p. 1812, Jul. 2022, doi: 10.3390/diagnostics12081812.
- [49] B. Wei, Z. Han, X. He, and Y. Yin, "Deep Learning Model Based Breast Cancer Histopathological Image Classification," *2017 the 2nd IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA 2017)*, 2017.
- [50] W. Auccahuasi, C. Delrieux, F. Sernaqué, E. Flores, and N. Moggiano, "Detection of Microcalcifications in Digital Mammography Images, using Deep Learning Techniques, based on Peruvian Casuistry," *E-Health and Bioengineering Conference (EHB)*, pp. 1–4, 2019.
- [51] M. Alkhaleefah, S. C. Ma, Y. L. Chang, B. Huang, P. K. Chittem, and V. P. Achhannagari, "Double-shot transfer learning for breast cancer classification from x-ray images," *Applied Sciences (Switzerland)*, vol. 10, no. 11, Jun. 2020, doi: 10.3390/app10113999.
- [52] K. R. Qasim and A. J. Ouda, "An Accurate Breast Cancer Detection System Based on Deep Learning CNN," *Medico-legal Update*, vol. 20, no. 1, pp. 984–990, 2020, doi: 10.37506/v20/i1/2020/mlu/194726.
- [53] S. Charan, M. J. Khan, and K. Khurshid, "Breast cancer detection in mammograms using convolutional neural network," *2018 International Conference on Computing, Mathematics and Engineering Technologies: Invent, Innovate and Integrate for Socioeconomic Development, iCoMET 2018 - Proceedings*, vol. 2018-January, pp. 1–5, Apr. 2018, doi: 10.1109/ICOMET.2018.8346384.

- TESIS TESIS TESIS TESIS TESIS
- [54] A. Saber, M. Sakr, O. M. Abo-Seida, and A. Keshk, "Tumor Detection and Classification in Breast Mammography-based on Fine-tuned Convolutional Neural Networks," *International Journal of Computers and Information(IJCI)*, 2022, [Online]. Available: <https://ijci.journals.ekb.eg/>
- [55] S. Montaha, S. Azam, A. K. M. Rakibul Haque Rafid, P. Ghosh, M. Z. Hasan, M. Jonkman, and F. de Boer, "BreastNet18: A High Accuracy Fine-Tuned VGG16 Model Evaluated Using Ablation Study for Diagnosing Breast Cancer from Enhanced Mammography Images," *Biology (Basel)*, vol. 10, no. 12, Dec. 2021, doi: 10.3390/biology10121347.
- [56] V. Reddy Allugunti, "Breast cancer detection based on thermographic images using machine learning and deep learning algorithms Healthcare View project Breast cancer detection based on thermographic images using machine learning and deep learning algorithms," *International Journal of Engineering in Computer Science* 2022, vol. 4, no. 1, pp. 49–56, 2022, [Online]. Available: <https://www.researchgate.net/publication/361228083>
- [57] A. Khamparia, S. Bharati, P. Podder, D. Gupta, A. Khanna, T. K. Phung, and D. N. H. Thanh, "Diagnosis of Breast Cancer Based on Modern Mammography using Hybrid Transfer Learning," *Multidimens Syst Signal Process*, 2021.
- [58] Z. Hameed, S. Zahia, B. Garcia-Zapirain, J. J. Aguirre, and A. M. Vanegas, "Breast cancer histopathology image classification using an ensemble of deep learning models," *Sensors (Switzerland)*, vol. 20, no. 16, pp. 1–17, Aug. 2020, doi: 10.3390/s20164373.
- [59] L. Shen, L. R. Margolies, J. H. Rothstein, E. Fluder, R. McBride, and W. Sieh, "Deep Learning to Improve Breast Cancer Detection on Screening Mammography," *Sci Rep*, vol. 9, no. 1, Dec. 2019, doi: 10.1038/s41598-019-48995-4.

- [60] H. Nasir Khan, A. R. Shahid, B. Raza, A. H. Dar, and H. Alquhayz, "Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network," *IEEE Access*, vol. 7, pp. 165724–165733, 2019, doi: 10.1109/ACCESS.2019.2953318.
- [61] D. A. Ragab, M. Sharkas, S. Marshall, and J. Ren, "Breast cancer detection using deep convolutional neural networks and support vector machines," *PeerJ*, vol. 2019, no. 1, 2019, doi: 10.7717/peerj.6201.
- [62] P. Xi, C. Shu, and R. Goubran, "Abnormality Detection in Mammography using Deep Convolutional Neural Networks," *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Mar. 2018, [Online]. Available: <http://arxiv.org/abs/1803.01906>
- [63] R. Agarwal, O. Díaz, M. H. Yap, X. Lladó, and R. Martí, "Deep learning for mass detection in Full Field Digital Mammograms," *Comput Biol Med*, vol. 121, Jun. 2020, doi: 10.1016/j.combiomed.2020.103774.
- [64] P. U. Hepsa, S. A. Özel, and A. Yazc, "Using Deep Learning for Mammography Classification," *2017 International Conference on Computer Science and Engineering (UBMK)*, pp. 418–423, 2017, doi: 10.5061/dryad.jm6k3.
- [65] A. G. Waks and E. P. Winer, "Breast Cancer Treatment: A Review," in *JAMA - Journal of the American Medical Association*, vol. 321, no. 3, American Medical Association, 2019, pp. 288–300. doi: 10.1001/jama.2018.19323.
- [66] C. para el C. y la P. de E. División de Prevención y Control del Cáncer, "¿Qué es el cáncer de mama?," 2022.
- [67] Sociedad Española de Protección Radiológica, "APLICACIONES MÉDICAS (MAMOGRAFÍA) ," <https://www.sepr.es/archivo-doc/recursos/otros/1857-3-3-aplicaciones-medicas-mamografia-pdf#:~:text=La%20mamograf%C3%ADa%20puede%20utilizarse%20para,la%20mama%2C%20especialmente%20el%20c%C3%A1ncer.>

- [68] Instituto nacional del cancer., “Mamografías,” <https://www.cancer.gov/espanol/tipos/seno/hoja-informativa-mamografias#:~:text=Para%20las%20mamograf%C3%ADas%20se%20usan,aumenta%20el%20riesgo%20de%20c%C3%A1ncer.>, 2021.
- [69] A. Arian, K. Dinas, G. C. Pratilas, and S. Alipour, “The Breast Imaging-Reporting and Data System (BI-RADS) Made Easy,” *Iranian Journal of Radiology*, vol. 19, no. 1, Jan. 2022, doi: 10.5812/iranjradiol-121155.
- [70] C. C. Aggarwal, “Artificial Intelligence A Textbook,” 2021.
- [71] C. Lexcellent, “SPRINGER BRIEFS IN APPLIED SCIENCES AND TECHNOLOGY Are Humans Going to Be Hacked?,” 2019. [Online]. Available: <http://www.springer.com/series/8884>
- [72] Huawei Technologies Co., Ltd., *Artificial Intelligence Technology*. Springer Nature Singapore, 2023. doi: 10.1007/978-981-19-2879-6.
- [73] Y. Ouali, C. Hudelot, and M. Tami, “An Overview of Deep Semi-Supervised Learning,” Jun. 2020, [Online]. Available: <http://arxiv.org/abs/2006.05278>
- [74] P. Cunningham, M. Cord, and S. J. Delany, “Supervised Learning.”
- [75] J. Chaki and N. Dey, “Texture Feature Extraction Techniques for Image Recognition,” *Springer*, 2020, [Online]. Available: <http://www.springer.com/series/10618>
- [76] “Appendix A: Definition of textural features First-order gray-level statistics.”
- [77] Z. Khyioon, H. Jehlol, H. B. Jehlol, A. Subhi, and A. Oleiwi, “Classification of Mammography Image Using Machine Learning Classifiersand Texture Features,” 2015. [Online]. Available: www.ijirae.com

- [78] Flach P, "MACHINE LEARNING The Art and Science of Algorithms that Make Sense of Data," 2012.
- [79] S. Batchu, F. Liu, A. Amireh, J. Waller, and M. Umair, "A Review of Applications of Machine Learning in Mammography and Future Challenges," *Oncology (Switzerland)*, vol. 99, no. 8. S. Karger AG, pp. 483–490, Jul. 01, 2021. doi: 10.1159/000515698.
- [80] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Computer Science*, vol. 2, no. 3. Springer, May 01, 2021. doi: 10.1007/s42979-021-00592-x.
- [81] Venkatesan E and Velmurugan T, "Performance Analysis of Decision Tree Algorithms for Breast Cancer Classification," 2015.
- [82] S. Farhad Khorshid and A. Mohsin Abdulazeez, "BREAST CANCER DIAGNOSIS BASED ON K-NEAREST NEIGHBORS: A REVIEW," *Journal Of Archaeology Of Egypt/Egyptology*, vol. 18, no. 4, pp. 1927–1951, 2021.
- [83] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, "A comparative analysis of gradient boosting algorithms," *Artif Intell Rev*, vol. 54, no. 3, pp. 1937–1967, Mar. 2021, doi: 10.1007/s10462-020-09896-5.
- [84] H. Liang, J. Li, H. Wu, L. Li, X. Zhou, and X. Jiang, "Mammographic Classification of Breast Cancer Microcalcifications through Extreme Gradient Boosting," *Electronics (Switzerland)*, vol. 11, no. 15, Aug. 2022, doi: 10.3390/electronics11152435.
- [85] I. Goodfellow, Y. Bengio, and A. Courville, "Deep Learning."
- [86] A. F. Gad, *Practical Computer Vision Applications Using Deep Learning with CNNs*. Apress, 2018. doi: 10.1007/978-1-4842-4167-7.
- [87] H. Demuth and B. de Jesús, "Neural Network Design 2nd Edition."

- TESIS TESIS TESIS TESIS TESIS
- [88] V. E. Balas, S. Sekhar, R. Dharmendra, S. Pijush, and S. Editors, "Handbook of Deep Learning Applications." [Online]. Available: <http://www.springer.com/series/8767>
- [89] A. F. Agarap, "Deep Learning using Rectified Linear Units (ReLU)," Mar. 2018, [Online]. Available: <http://arxiv.org/abs/1803.08375>
- [90] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," 2014.
- [91] A. Gulli, *TensorFlow 1.x deep learning cookbook*. 2017.
- [92] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation Functions: Comparison of trends in Practice and Research for Deep Learning," Nov. 2018, [Online]. Available: <http://arxiv.org/abs/1811.03378>
- [93] J. S. Bridle, "Training Stochastic Model Recognition Algorithms Training Stochastic Model Recognition Algorithms as Networks can lead to Maximum Mutual Information Estimation of Parameters."
- [94] A. Y. Ng, "Feature selection, L 1 vs. L 2 regularization, and rotational invariance."
- [95] Y. Bai, E. Yang, B. Han, Y. Yang, J. Li, Y. Mao, G. Niu, and T. Liu, "Understanding and Improving Early Stopping for Learning with Noisy Labels," Jun. 2021, [Online]. Available: <http://arxiv.org/abs/2106.15853>
- [96] R. Sun, "Optimization for deep learning: theory and algorithms," Dec. 2019, [Online]. Available: <http://arxiv.org/abs/1912.08957>
- [97] S. Ruder, "An overview of gradient descent optimization algorithms," Sep. 2016, [Online]. Available: <http://arxiv.org/abs/1609.04747>
- [98] D. P. Kingma and J. Lei Ba, "ADAM: A Method for stochastic optimization," *International Conference on Learning Representations*, 2015.

- [99] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A Comprehensive Survey on Transfer Learning," Nov. 2019, [Online]. Available: <http://arxiv.org/abs/1911.02685>
- [100] M. Iman, K. Rasheed, and H. R. Arabnia, "A Review of Deep Transfer Learning and Recent Advancements," Jan. 2022, [Online]. Available: <http://arxiv.org/abs/2201.09679>
- [101] L. G. Falconi, M. Perez, W. G. Aguilar, and A. Conci, "Transfer learning and fine tuning in breast mammogram abnormalities classification on CBIS-DDSM database," *Advances in Science, Technology and Engineering Systems*, vol. 5, no. 2, pp. 154–165, 2020, doi: 10.25046/aj050220.
- [102] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, and J. Dean, "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, no. 1. Nature Publishing Group, pp. 24–29, Jan. 01, 2019. doi: 10.1038/s41591-018-0316-z.
- [103] U. Hahn and M. Oleynik, "Medical Information Extraction in the Age of Deep Learning," *Yearbook of medical informatics*, vol. 29, no. 1. NLM (Medline), pp. 208–220, Aug. 01, 2020. doi: 10.1055/s-0040-1702001.
- [104] H. M and S. M.N, "A Review on Evaluation Metrics for Data Classification Evaluations," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, no. 2, pp. 01–11, Mar. 2015, doi: 10.5121/ijdkp.2015.5201.
- [105] Guido. van Rossum and F. L. Drake, *An introduction to Python : release 2.2.2*. Network Theory Ltd, 2003.
- [106] USENIX Association., ACM SIGMOBILE., ACM Special Interest Group in Operating Systems., and ACM Digital Library., *TensorFlow: A System for Large-Scale Machine Learning*. USENIX Association, 2015.

- [107] Chollet F., "Keras," <https://keras.io/>, 2015.
- [108] J. Hugunin and T. Oliphant, "NumPy," <https://numpy.org/>, 1995.
- [109] W. McKinney, "Pandas," <https://pandas.pydata.org/>, Jan. 11, 2008.
- [110] "Pillow," <https://pillow.readthedocs.io/en/stable/>, 2011.
- [111] "OpenCV," <https://opencv.org/>, 2000.
- [112] "Scikit-learn," <https://scikit-learn.org/stable/>, 2007.
- [113] "Matplotlib," <https://matplotlib.org/>, 2003.
- [114] "Albumentations," <https://albumentations.ai/>.
- [115] "Pydicom," <https://pydicom.github.io/>, 2017.
- [116] E. M. F. el Houbay and N. I. R. Yassin, "Malignant and nonmalignant classification of breast lesions in mammograms using convolutional neural networks," *Biomed Signal Process Control*, vol. 70, Sep. 2021, doi: 10.1016/j.bspc.2021.102954.
- [117] Xi P, Shu C, and Goubran R, *Abnormality Detection in Mammography using Deep Convolutional Neural Networks*. 2018.
- [118] K. K. Singh, S. Kumar, M. Antonakakis, K. Moirogiorgou, A. Deep, K. L. Kashyap, M. K. Bajpai, and M. Zervakis, "Deep Learning Capabilities for the Categorization of Microcalcification," *Int J Environ Res Public Health*, vol. 19, no. 4, Feb. 2022, doi: 10.3390/ijerph19042159.
- [119] Y. J. Tan, K. S. Sim, and F. F. Ting, "Breast Cancer detection Using Convolutional Neural Networks for Mammogram Imaging System."
- [120] Y. Li and H. Wu, "A Clustering Method Based on K-Means Algorithm," *Phys Procedia*, vol. 25, pp. 1104–1109, 2012, doi: 10.1016/j.phpro.2012.03.206.
- [121] M.L. Huang and T. Y. Lin, "Dataset of breast mammography images with masses," *ELSEVIER*, pp. 1–6, 2020, doi: 10.17632/x7bvzv6cvr.1.

TESIS TESIS TESIS TESIS TESIS

[122] J. Tan, J. Yang, S. Wu, G. Chen, and J. Zhao, "A critical look at the current train/test split in machine learning," Jun. 2021, [Online]. Available: <http://arxiv.org/abs/2106.04525>



ANEXO

8.1 Clasificación usando aprendizaje profundo en la primera metodología.

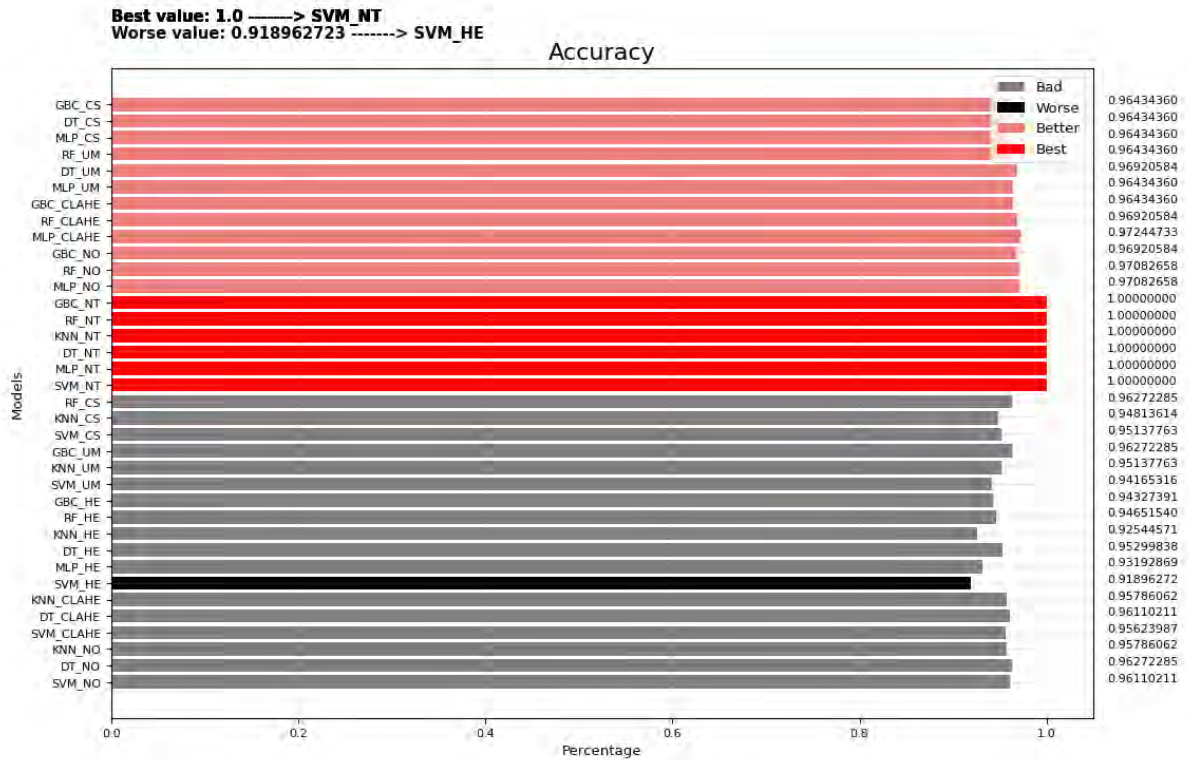


Ilustración 51 Grafica mostrando la métrica de exactitud para la etapa de entrenamiento, comparando los modelos de aprendizaje automatico en una clasificación binaria utilizando el conjunto de datos CBIS-DDSM.

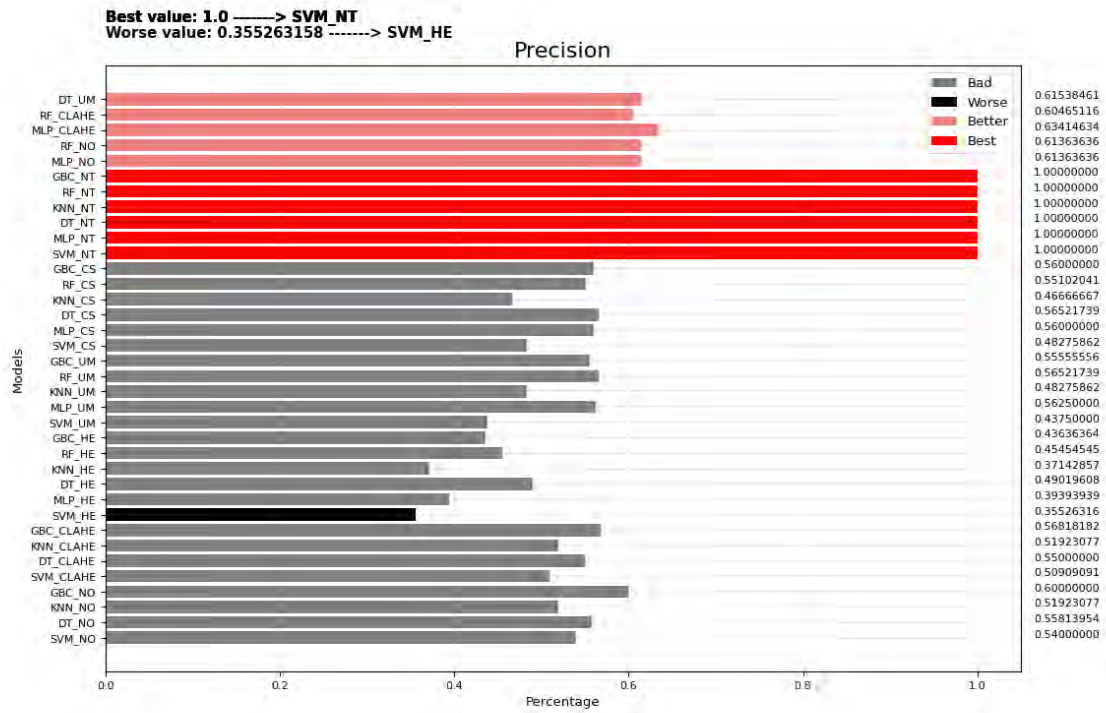


Ilustración 52 Grafica mostrando la métrica de precisión para la etapa de entrenamiento, comparando los modelos de aprendizaje automático en una clasificación binaria utilizando el conjunto de datos CBIS-DDSM.

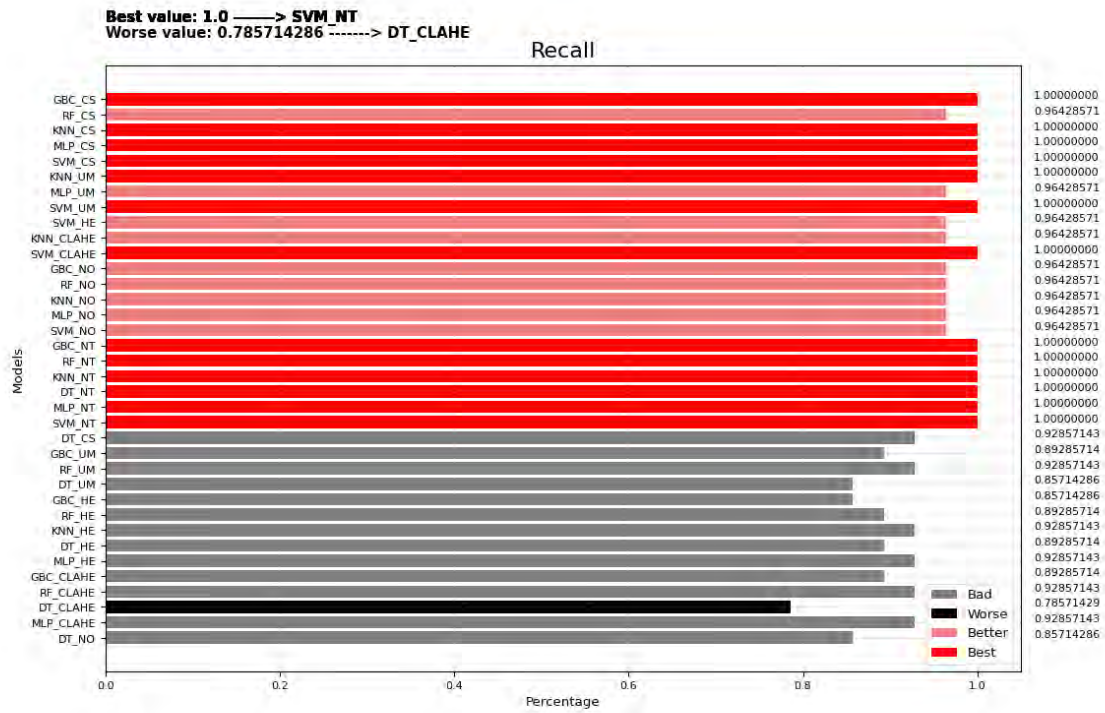


Ilustración 53 Grafica mostrando la métrica de recall para la etapa de entrenamiento, comparando los modelos de aprendizaje automático en una clasificación binaria utilizando el conjunto de datos CBIS-DDSM.

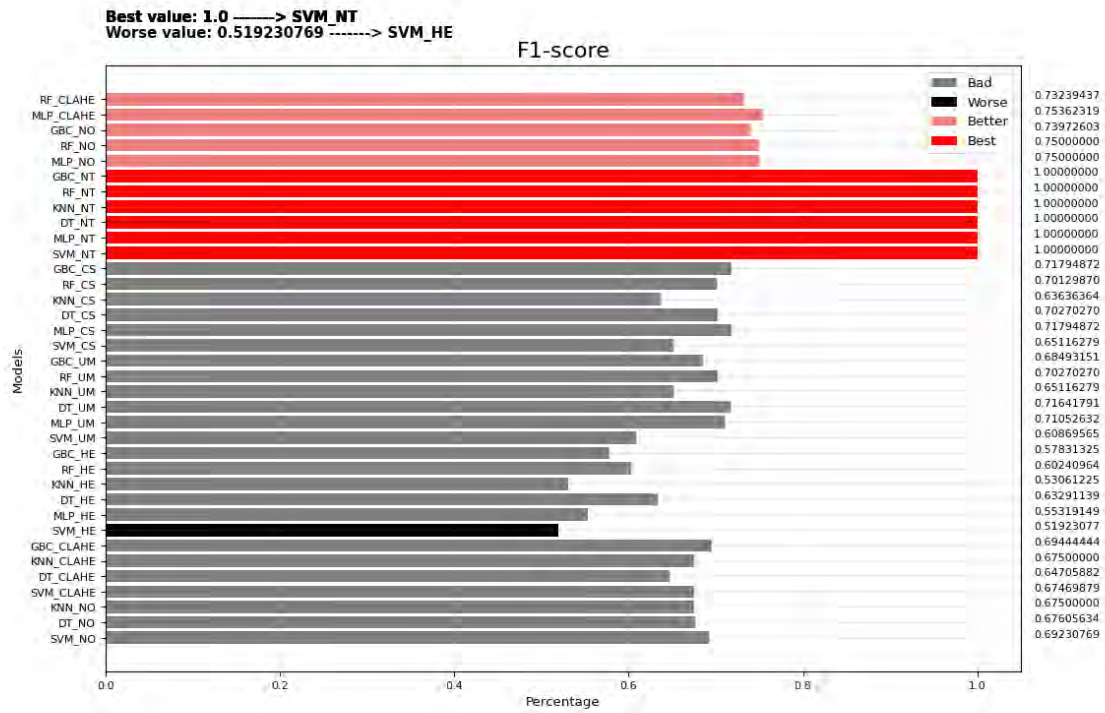


Ilustración 54 Grafica mostrando la métrica de F1-score para la etapa de entrenamiento, comparando los modelos de aprendizaje automático en una clasificación binaria utilizando el conjunto de datos CBIS-DDSM.

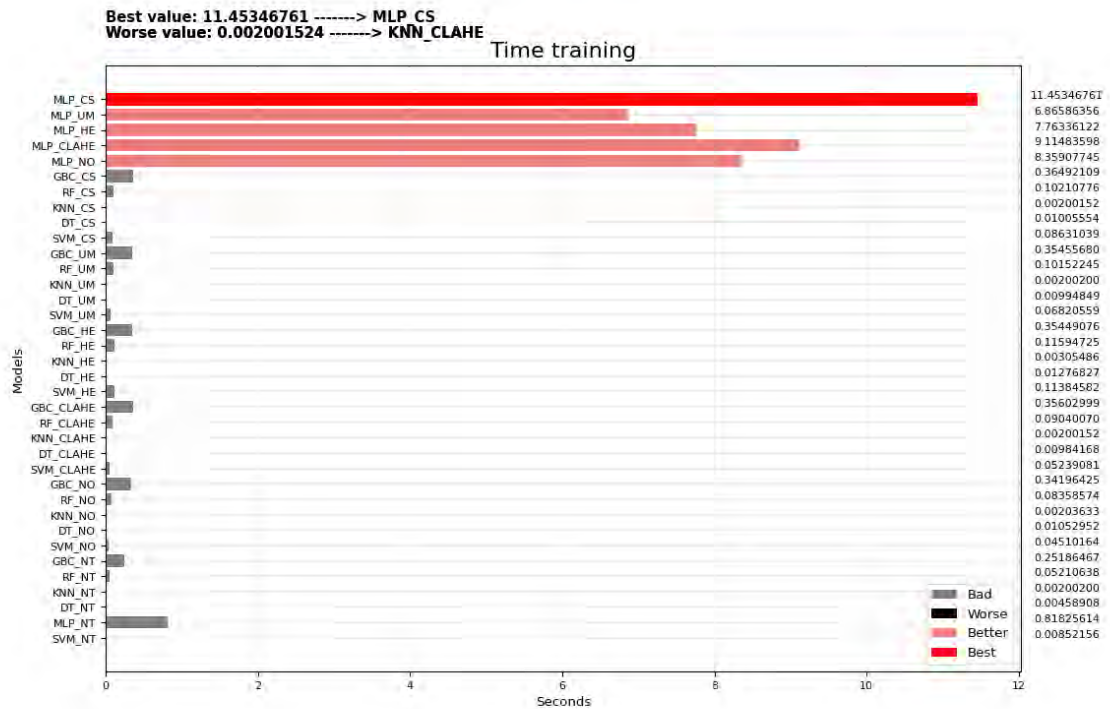


Ilustración 55 Grafica mostrando el tiempo de entrenamiento para la etapa de entrenamiento, comparando los modelos de aprendizaje automático en una clasificación binaria utilizando el conjunto de datos CBIS-DDSM.

Best value: 1.0 -----> RN152_NO
 Worse value: 0.670624852 -----> MN_NT

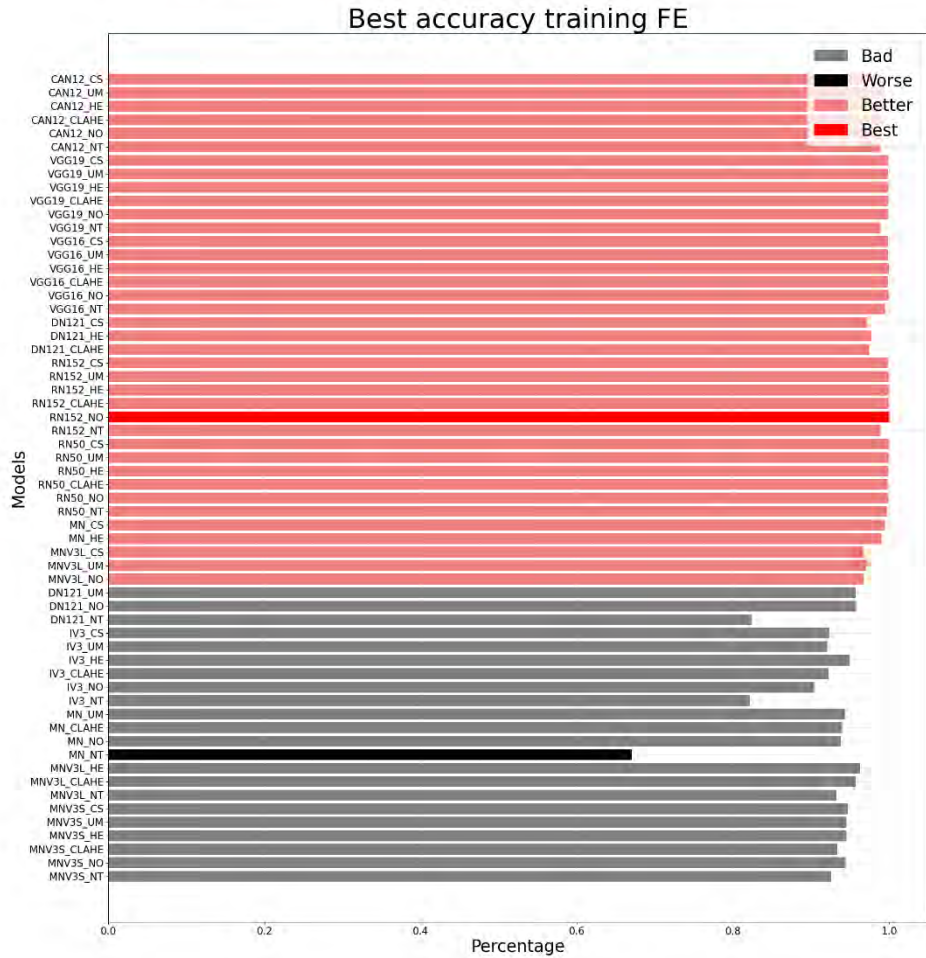


Ilustración 56 Grafica mostrando la métrica de exactitud para la etapa de entrenamiento, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS.

Best value: 0.98625952 -----> RN50_NO
 Worse value: 0.721628487 -----> MN_NT

Best accuracy testing



Ilustración 57 Grafica mostrando la métrica de exactitud para la etapa de prueba, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS.

Best value: 0.985796949 -----> RN50_NO
 Worse value: 0.746007213 -----> CAN12_NT

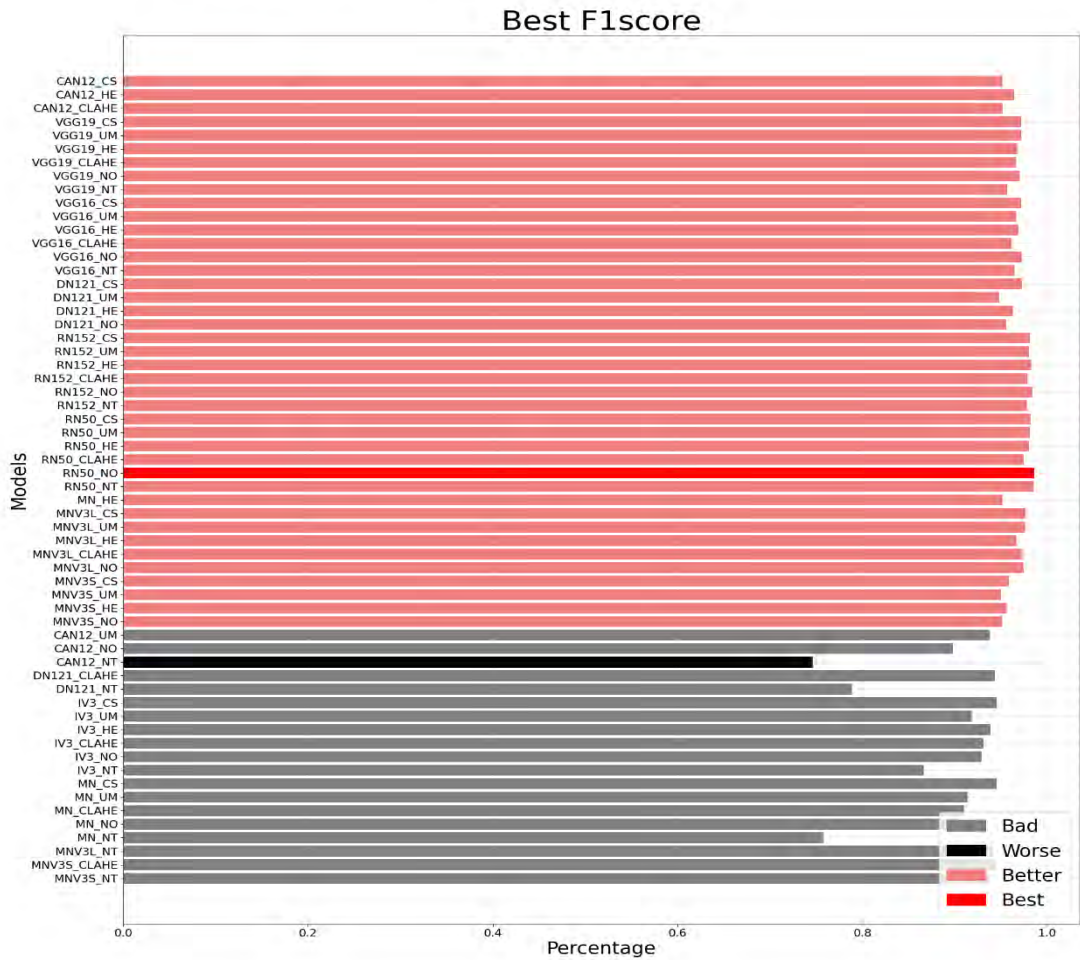


Ilustración 21 Comparación de los modelos en una clasificación binaria en la métrica de F1-score en la etapa de prueba usando el conjunto de datos de Mini-MIAS.

Best value: 0.995332555 -----> DN121_CLAHE
 Worse value: 0.65467075 -----> MN_NT

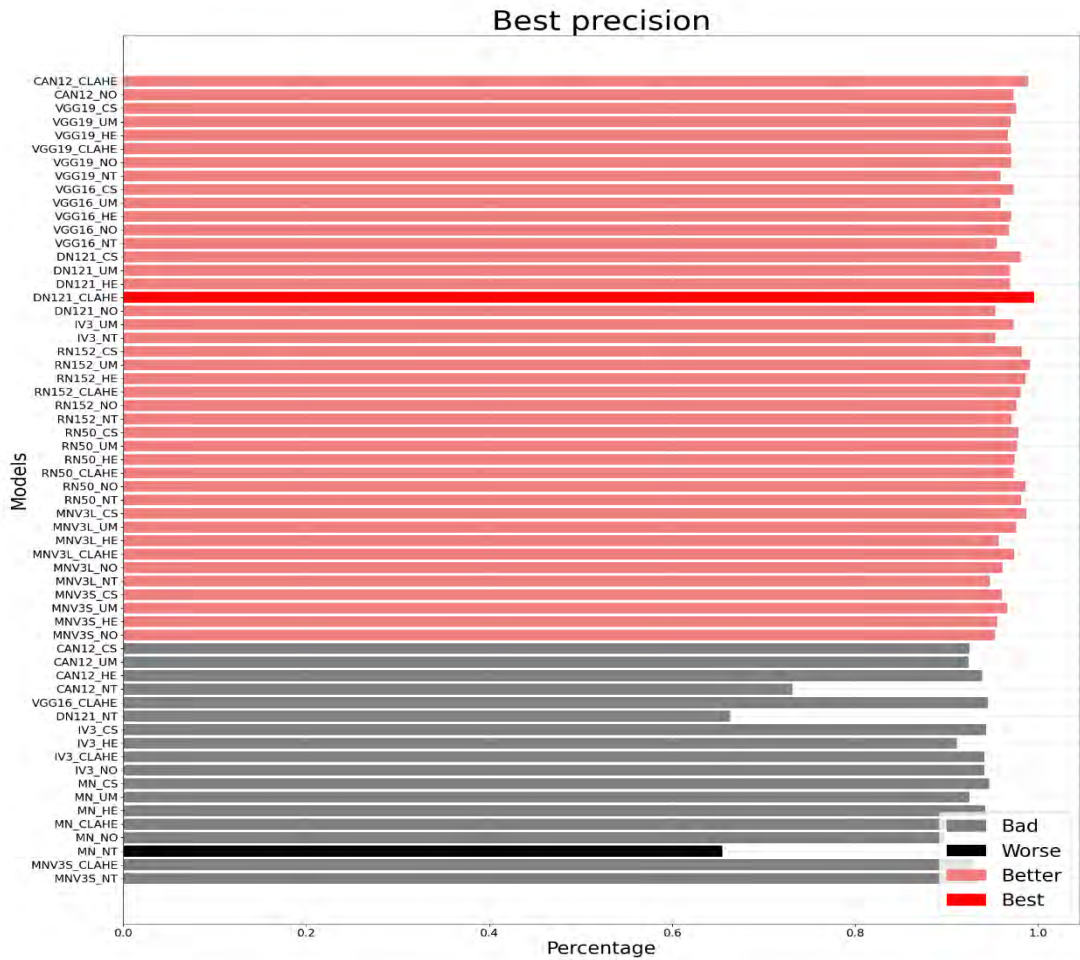


Ilustración 22 Comparación de los modelos en una clasificación binaria en la métrica de precisión en la etapa de prueba usando el conjunto de datos de Mini-MIAS.

Best value: 0.991587802 -----> RN152_NO
 Worse value: 0.761303891 -----> CAN12_NT

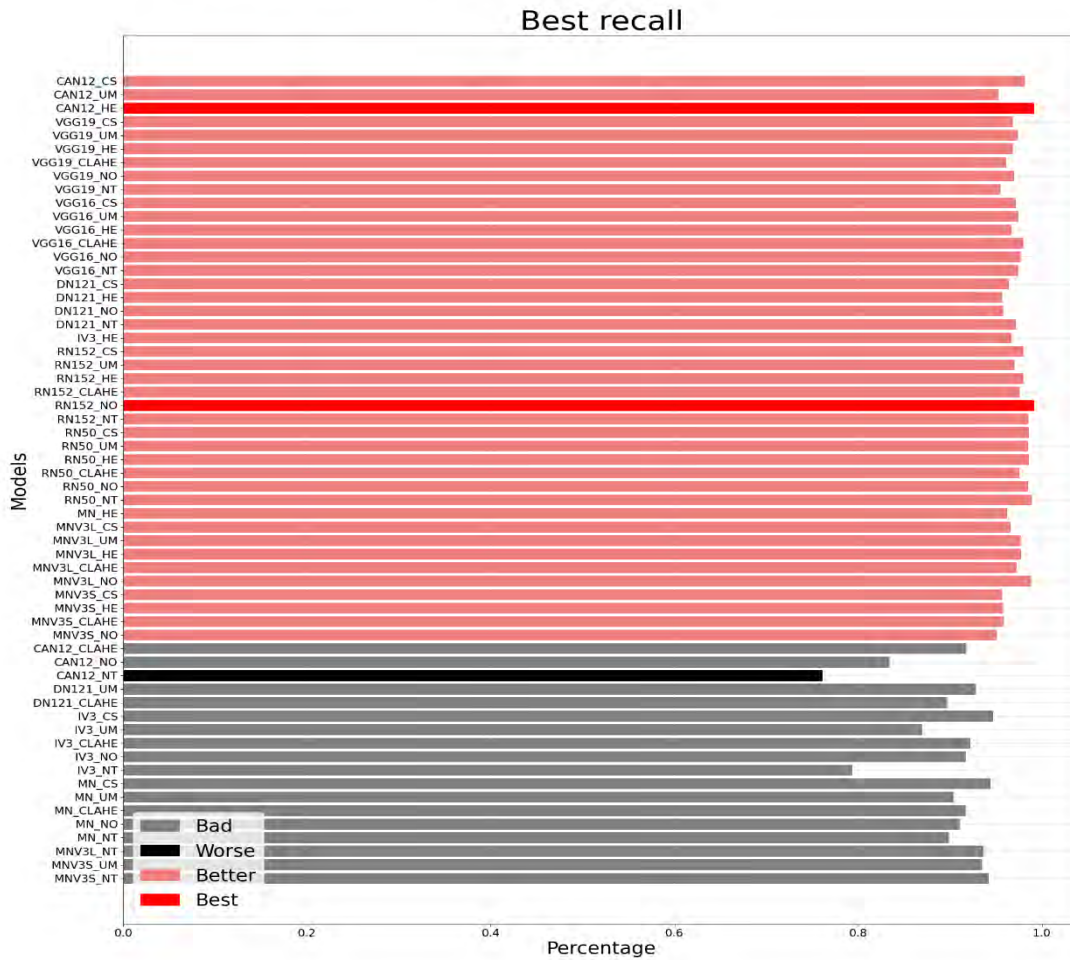


Ilustración 23 Comparación de los modelos en una clasificación binaria en la métrica de exhaustividad en la etapa de prueba usando el conjunto de datos de Mini-MIAS.

Best value: 0.998442352 -----> VGG16_HE
 Worse value: 0.434296221 -----> MN_NT

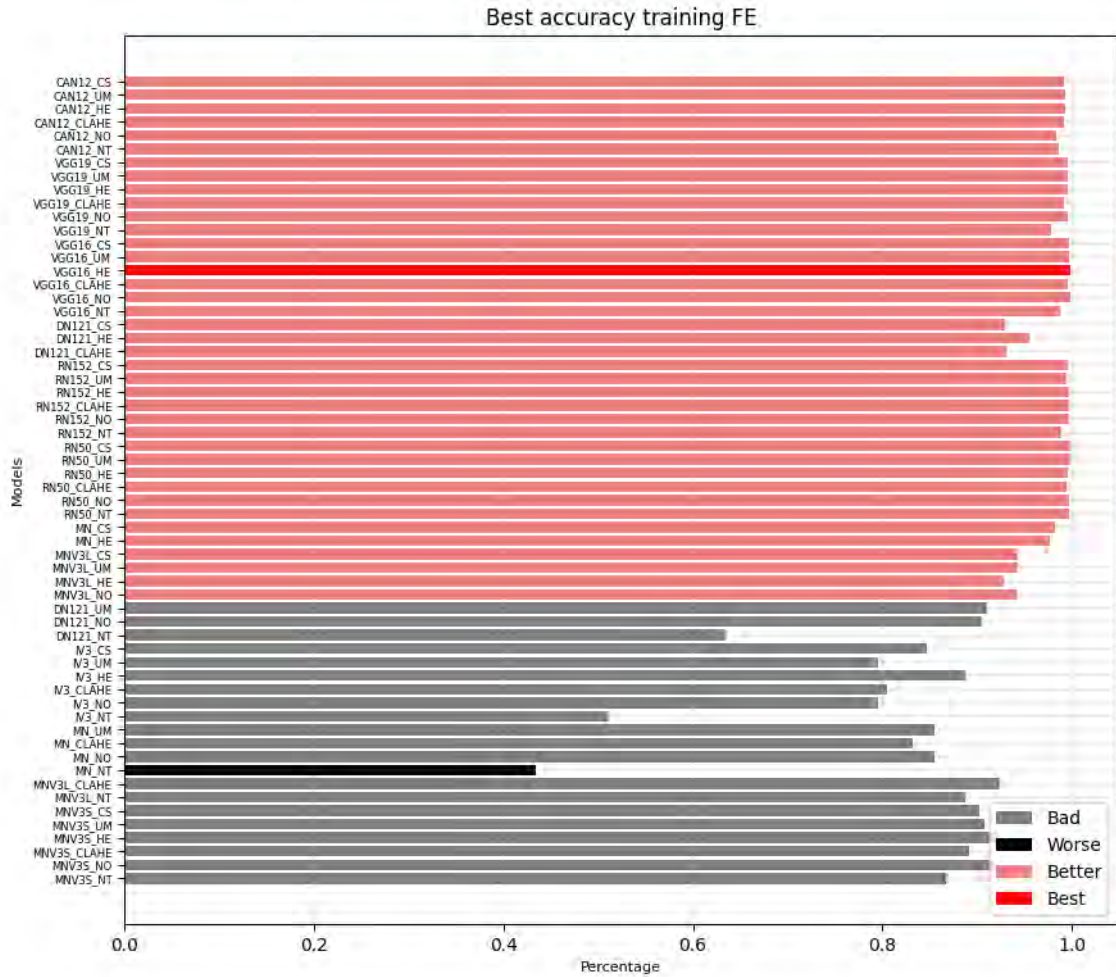


Ilustración 28 Comparación de los modelos en una clasificación multiclase final en la métrica de exactitud en la etapa de entrenamiento usando el conjunto de datos de Mini-MIAS.

Best value: 0.976554513 -----> RN152_CLAHE
 Worse value: 0.465341479 -----> MN_NT

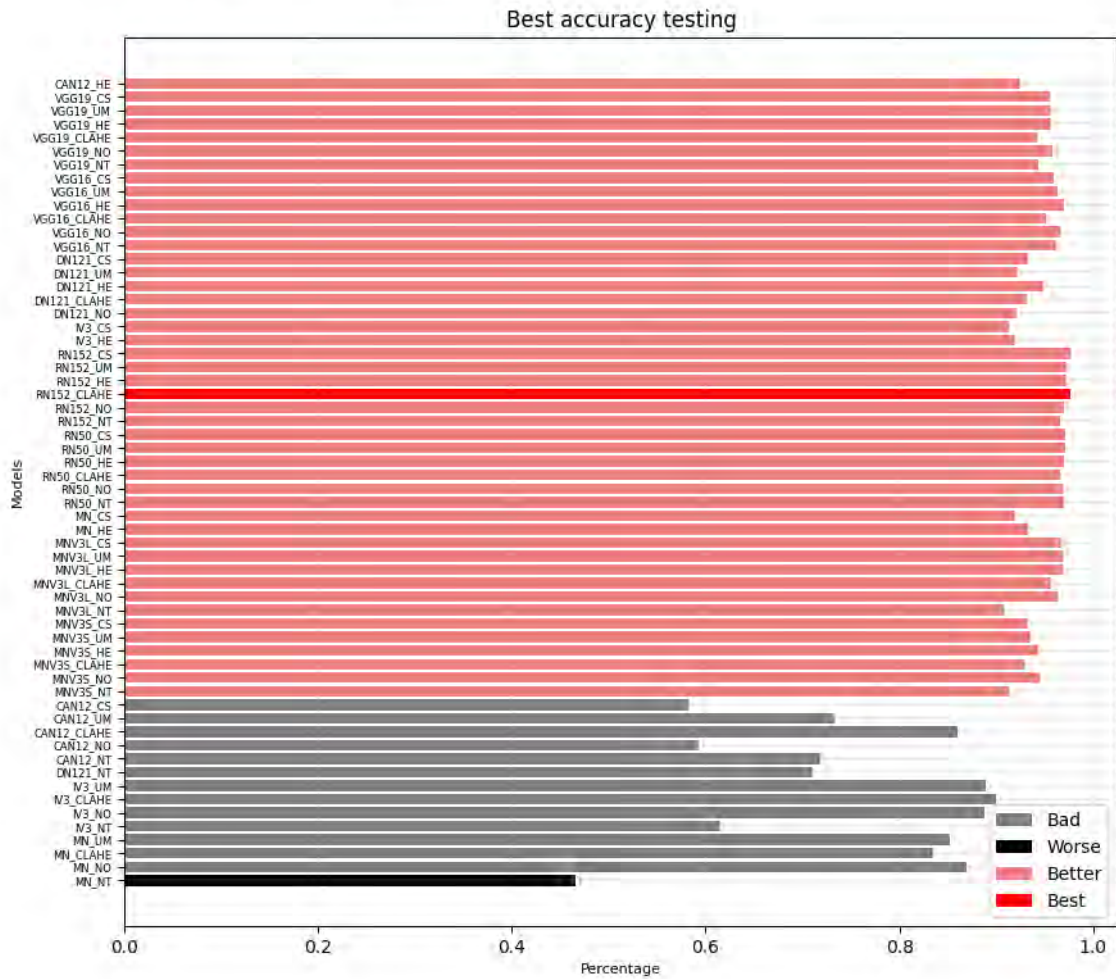


Ilustración 30 Comparación de los modelos en una clasificación multiclase en la métrica de exactitud en la etapa de prueba usando el conjunto de datos de Mini-MIAS.

Best value: 0.976518393 -----> RN152_CLAHE
 Worse value: 0.382319679 -----> MN_NT

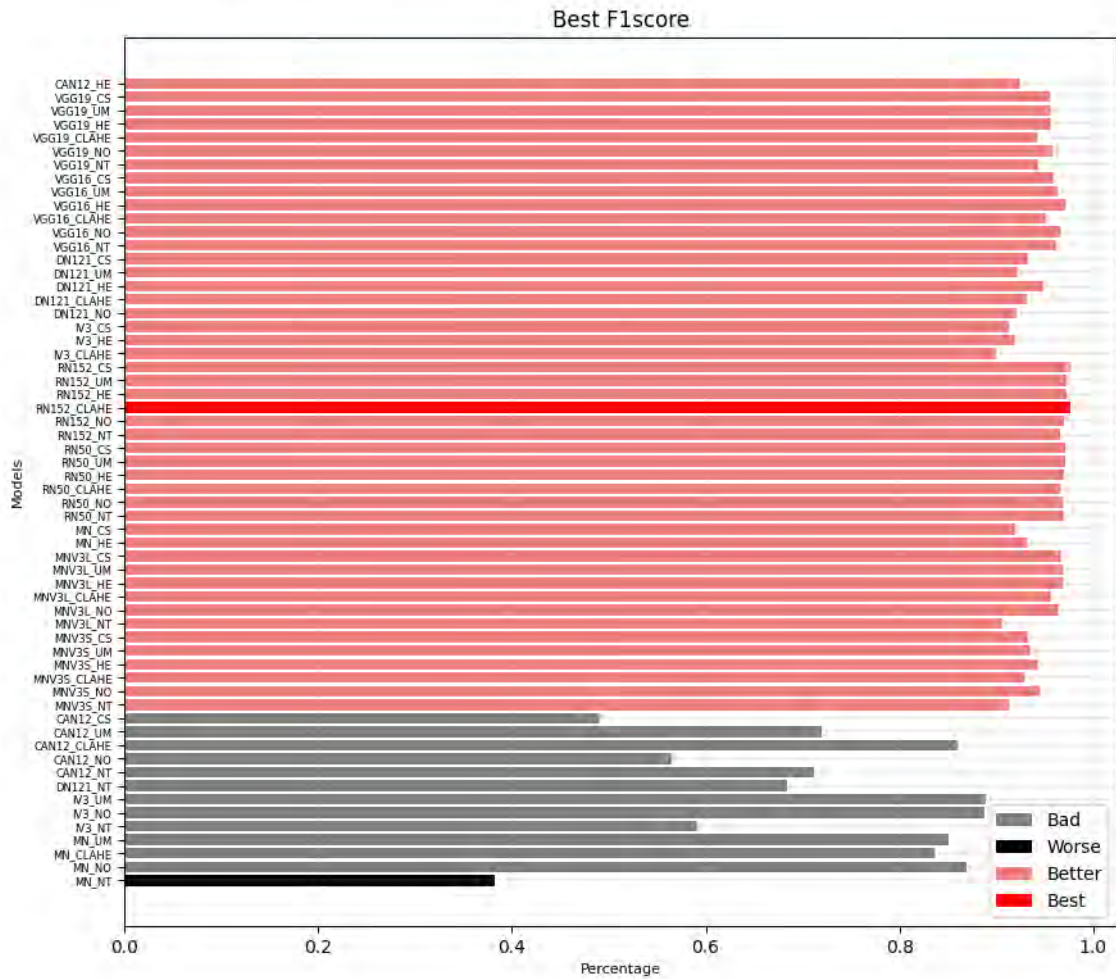


Ilustración 32 Comparación de los modelos en una clasificación multiclase en la métrica de F1-score en la etapa de prueba usando el conjunto de datos de Mini-MIAS.

Best value: 0.976577793 -----> RN152_CLAHE
 Worse value: 0.579976205 -----> CAN12_CS

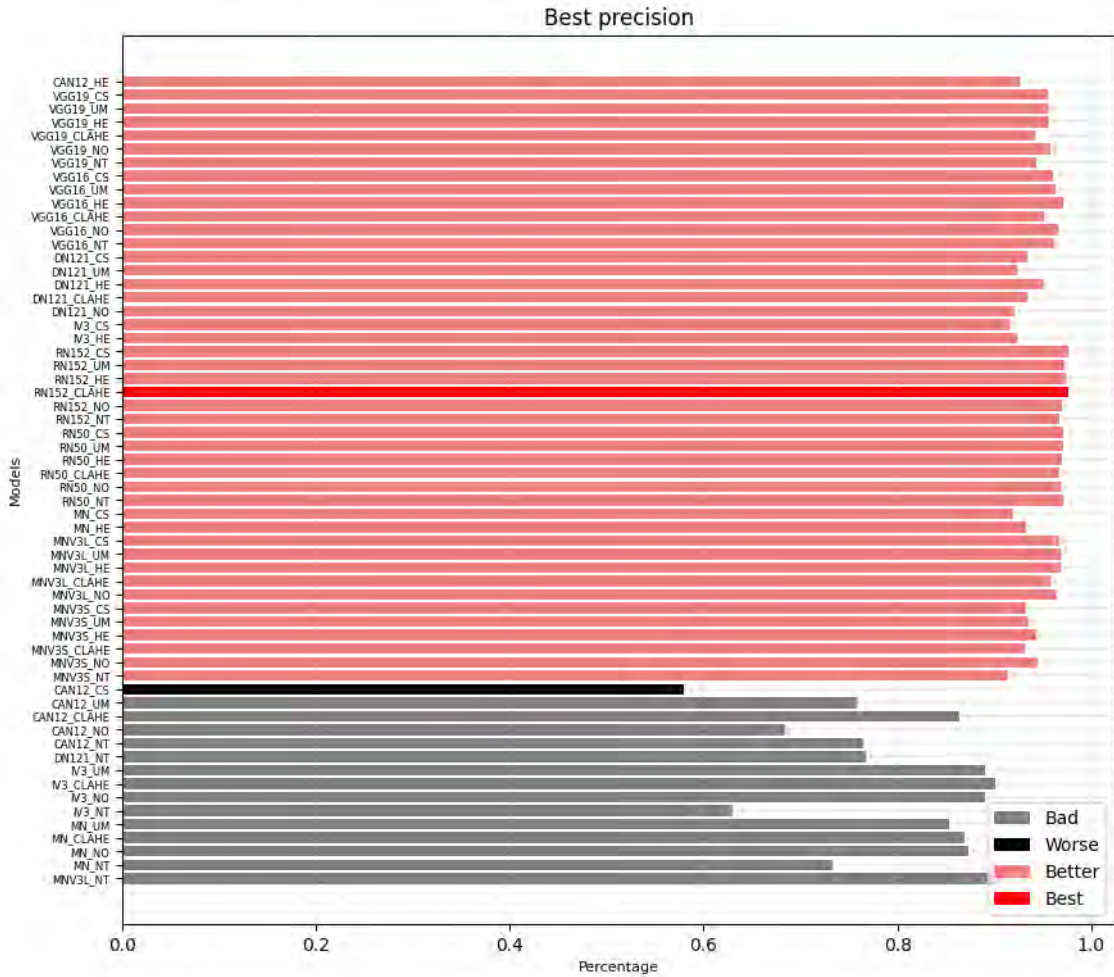


Ilustración 33 Comparación de los modelos en una clasificación multiclase en la métrica de precisión en la etapa de prueba usando el conjunto de datos de Mini-MIAS.

Best value: 0.976554536 -----> RN152_CLAHE
 Worse value: 0.465341488 -----> MN_NT

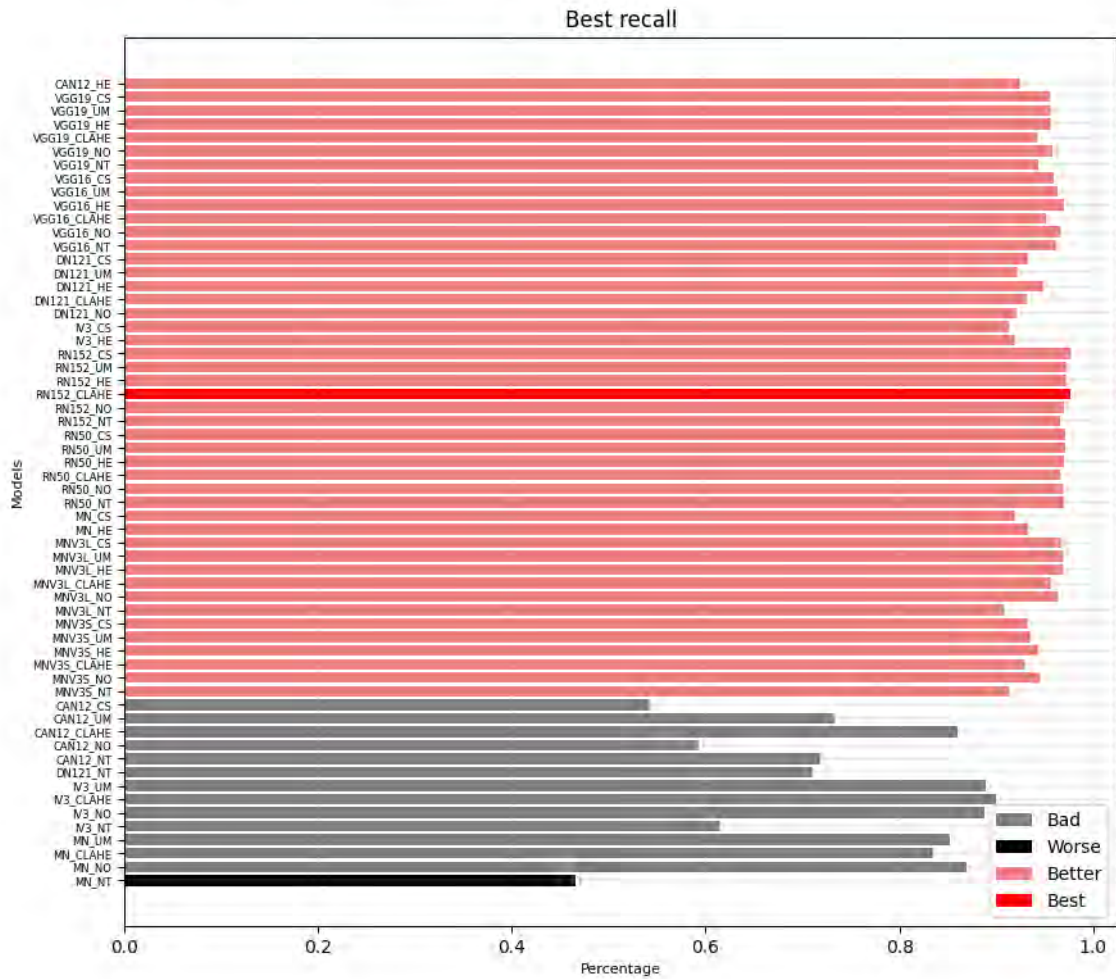


Ilustración 34 Comparación de los modelos en una clasificación multiclase en la métrica de exhaustividad en la etapa de prueba usando el conjunto de datos de Mini-MIAS.

Best value: 0.995329142 -----> RN152_UM
 Worse value: 0.782938242 -----> MN_NT

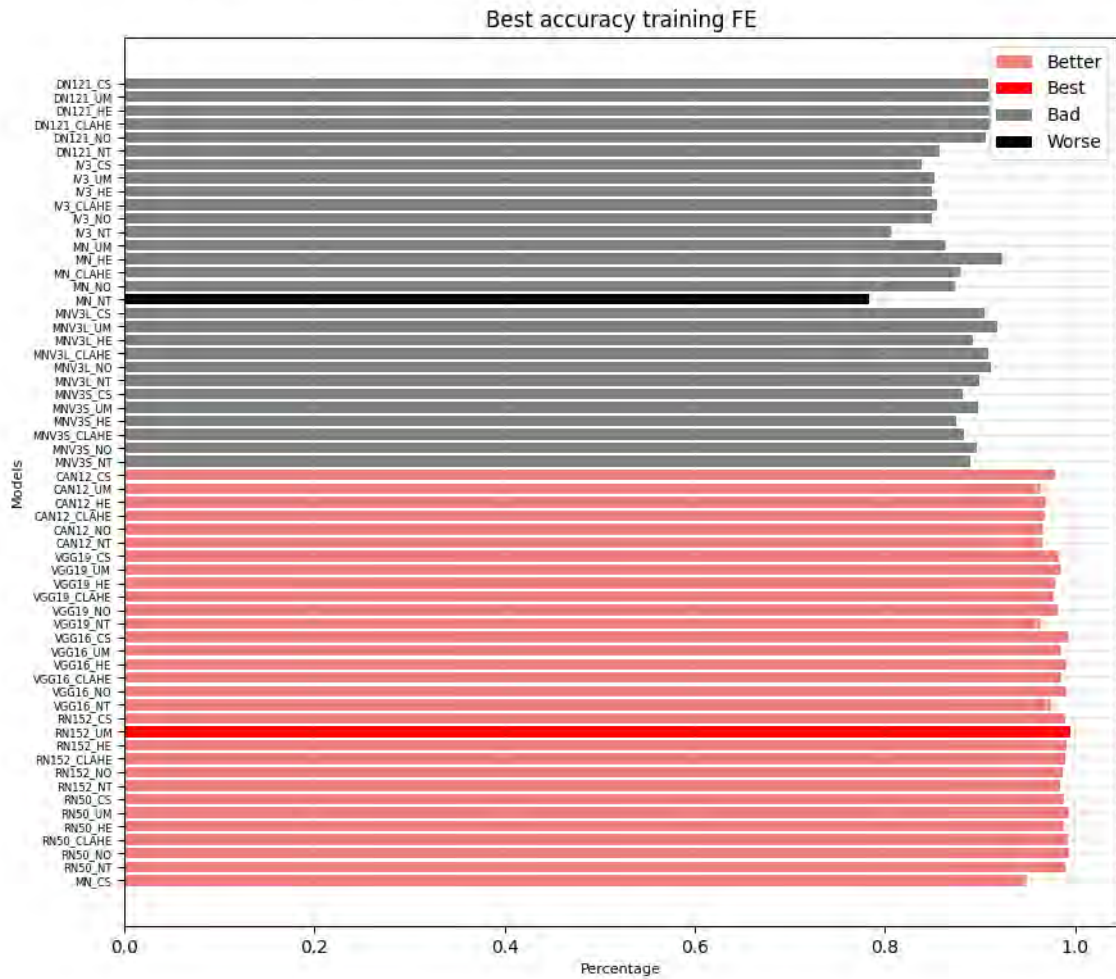


Ilustración 39 Comparación de los modelos en una clasificación binaria final en la métrica de exactitud en la etapa de entrenamiento usando el conjunto de datos de CBIS-DDSM.

Best value: 0.924199164 -----> RN152_UM
 Worse value: 0.767840147 -----> CAN12_UM

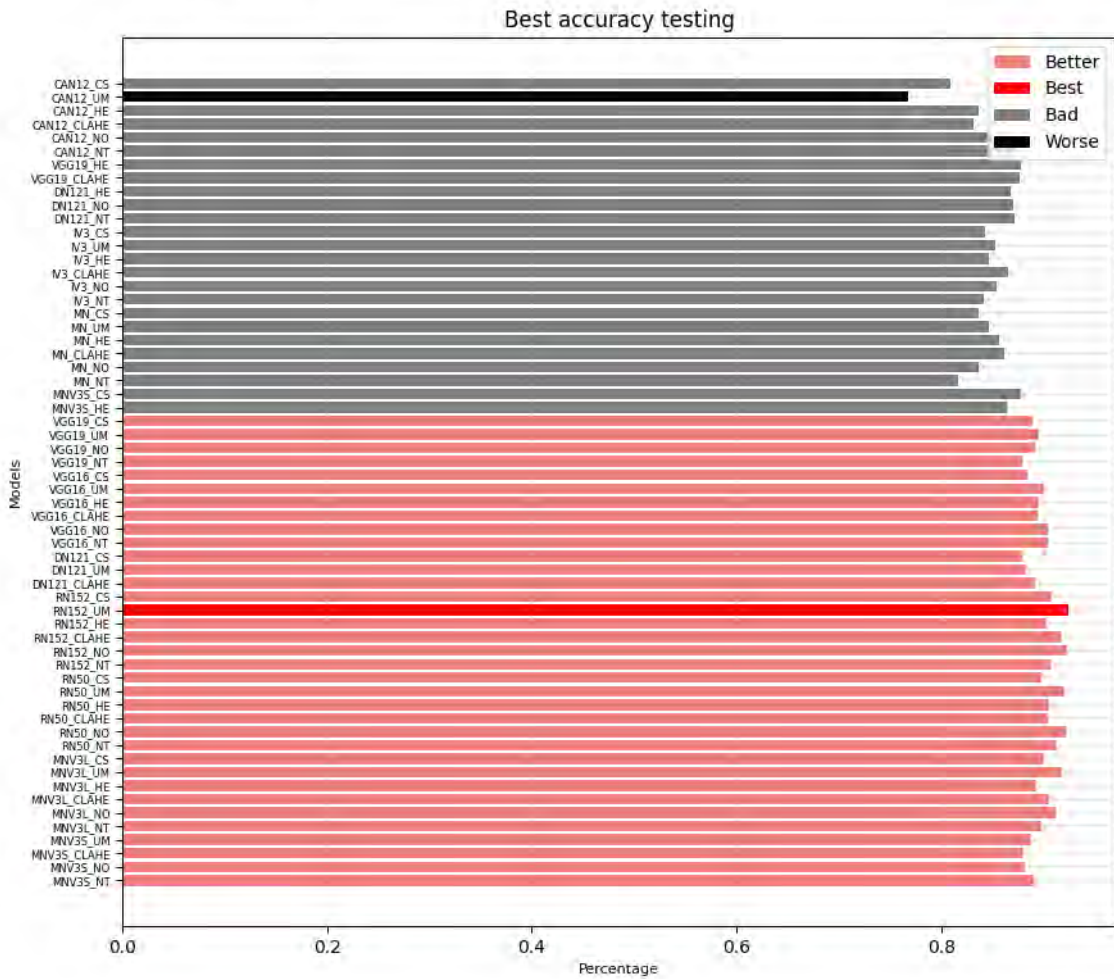


Ilustración 41 Comparación de los modelos en una clasificación binaria en la métrica de exactitud en la etapa de prueba usando el conjunto de datos de CBIS-DDSM.

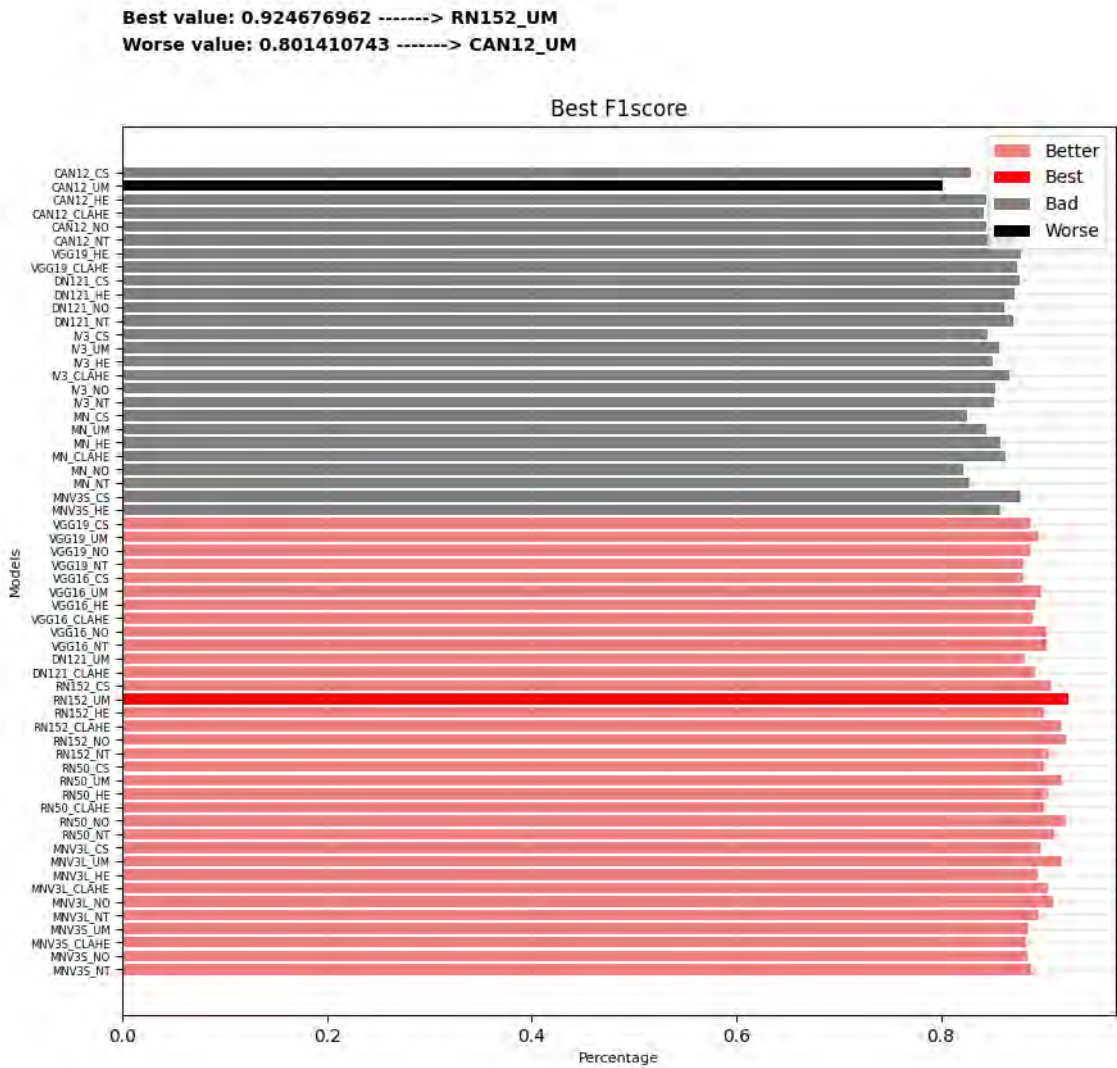


Ilustración 43 Comparación de los modelos en una clasificación binaria en la métrica de F1-score en la etapa de prueba usando el conjunto de datos de CBIS-DDSM.

Best value: 0.930280246 -----> RN50_CLAHE
 Worse value: 0.694731891 -----> CAN12_UM

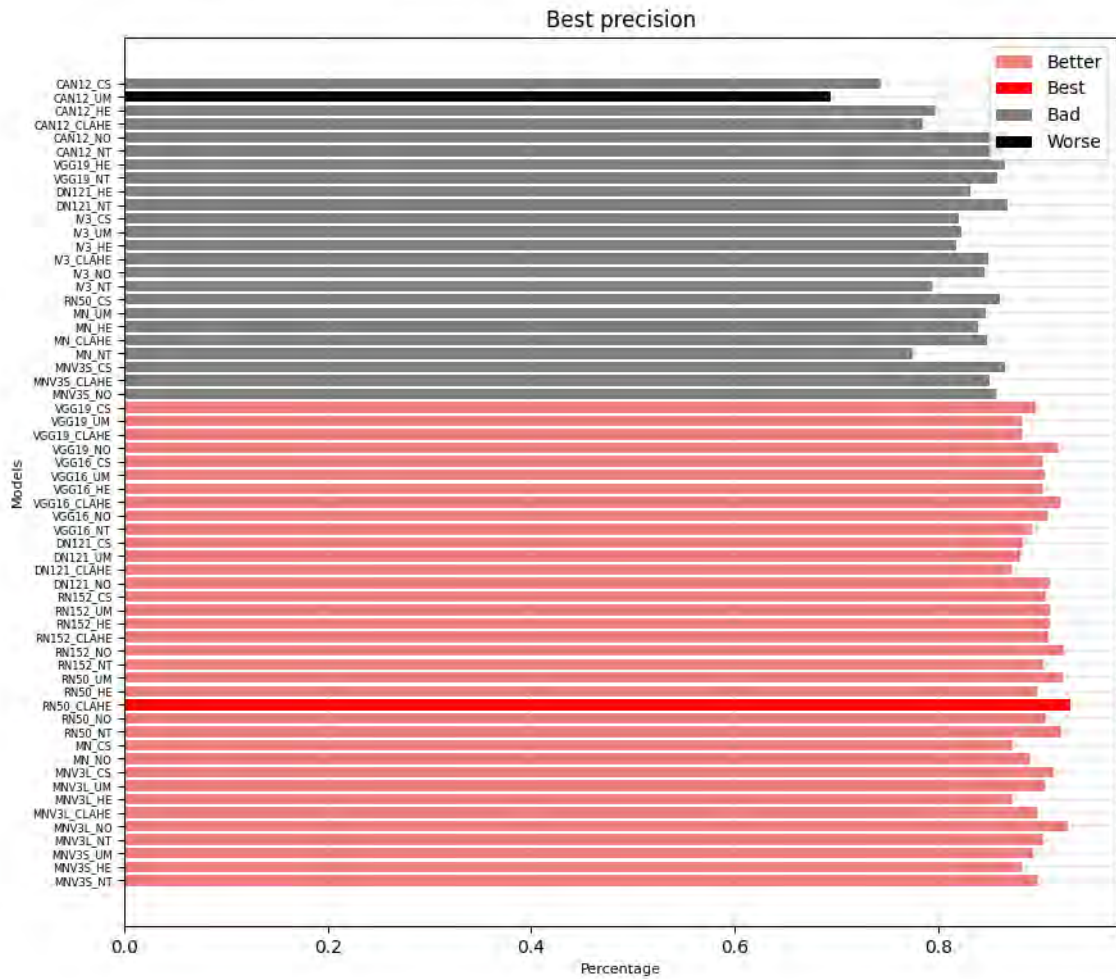


Ilustración 44 Comparación de los modelos en una clasificación binaria en la métrica de precisión en la etapa de prueba usando el conjunto de datos de CBIS-DDSM.

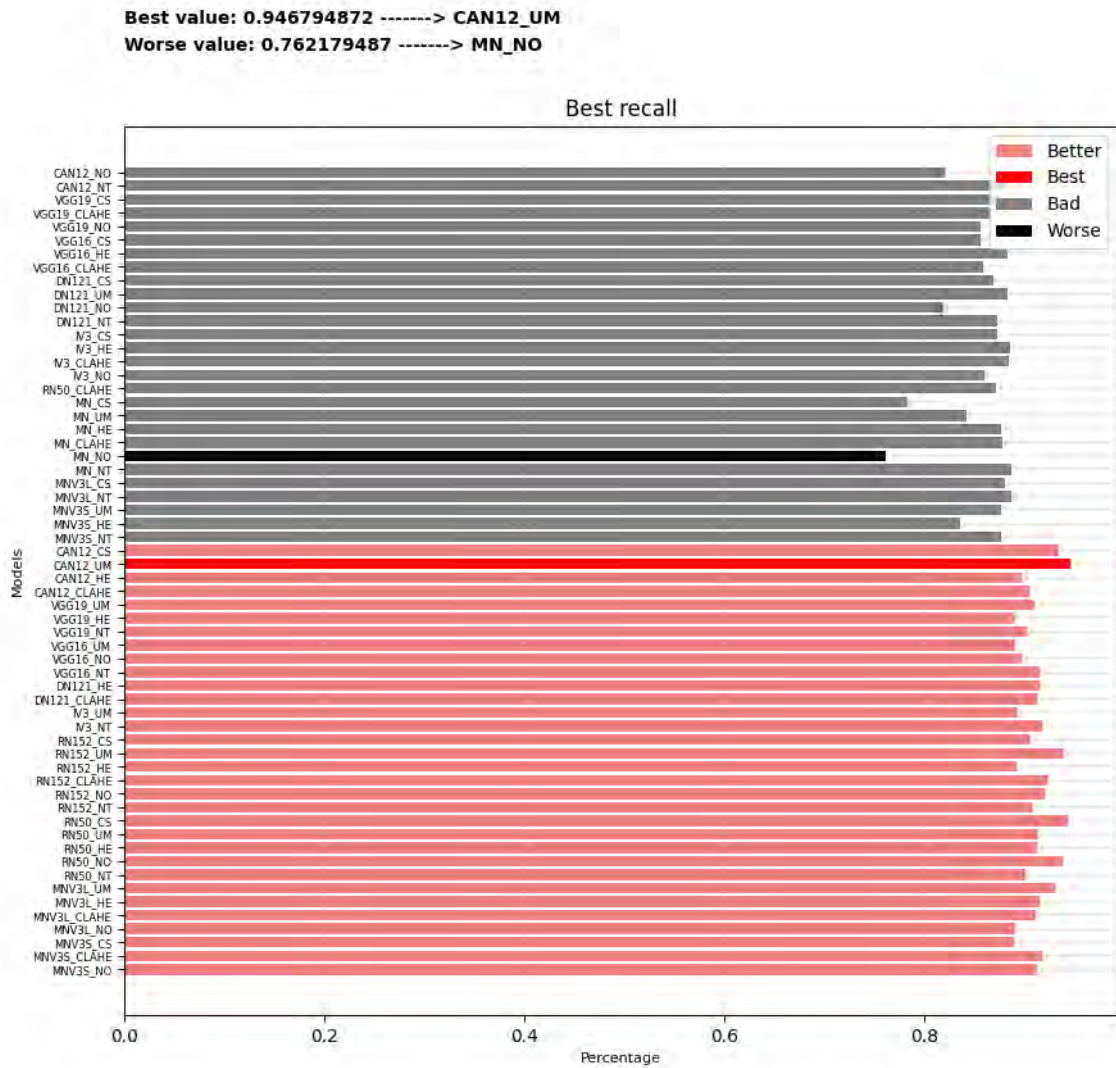


Ilustración 45 Comparación de los modelos en una clasificación binaria en la métrica de exhaustividad en la etapa de prueba usando el conjunto de datos de CBIS-DDSM.

Best value: 0.991684616 -----> RN50_UM
 Worse value: 0.778207898 -----> IV3_NO

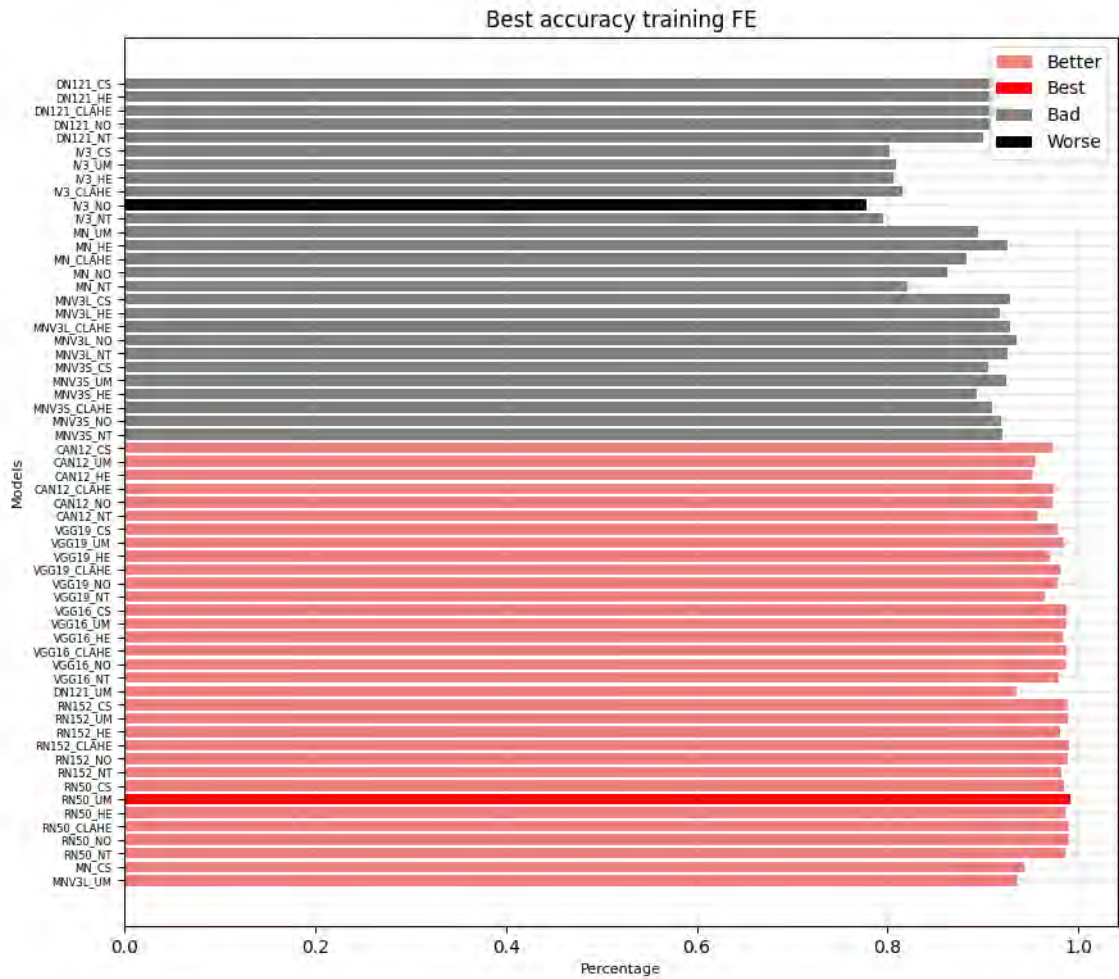


Ilustración 50 Comparación de los modelos en una clasificación multiclase final en la métrica de exactitud en la etapa de entrenamiento usando el conjunto de datos de CBIS-DDSM.

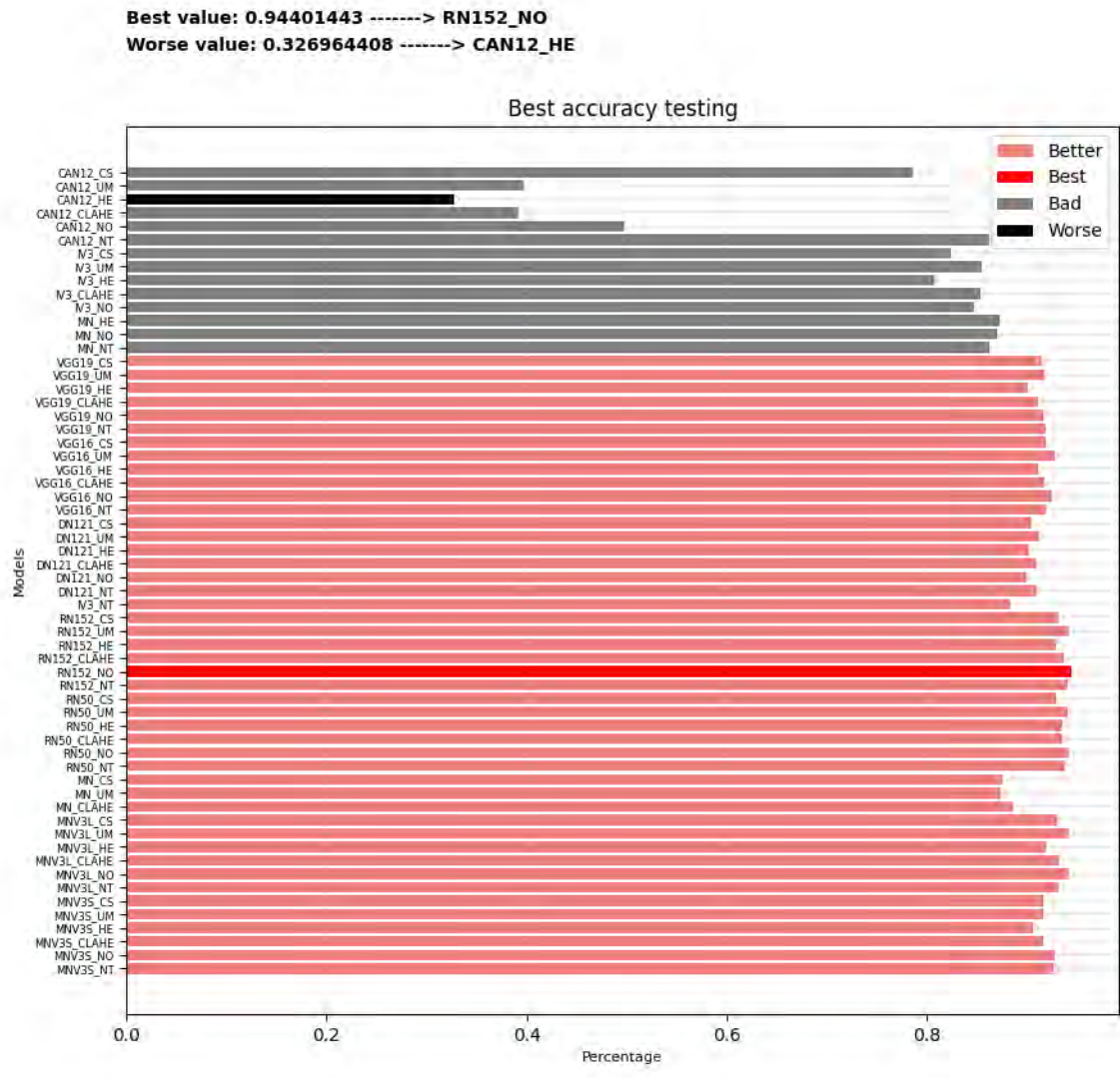


Ilustración 52 Comparación de los modelos en una clasificación multiclase en la métrica de exactitud en la etapa de prueba usando el conjunto de datos de CBIS-DDSM.

Best value: 0.94337144 -----> RN152_NO
 Worse value: 0.186404027 -----> CAN12_HE

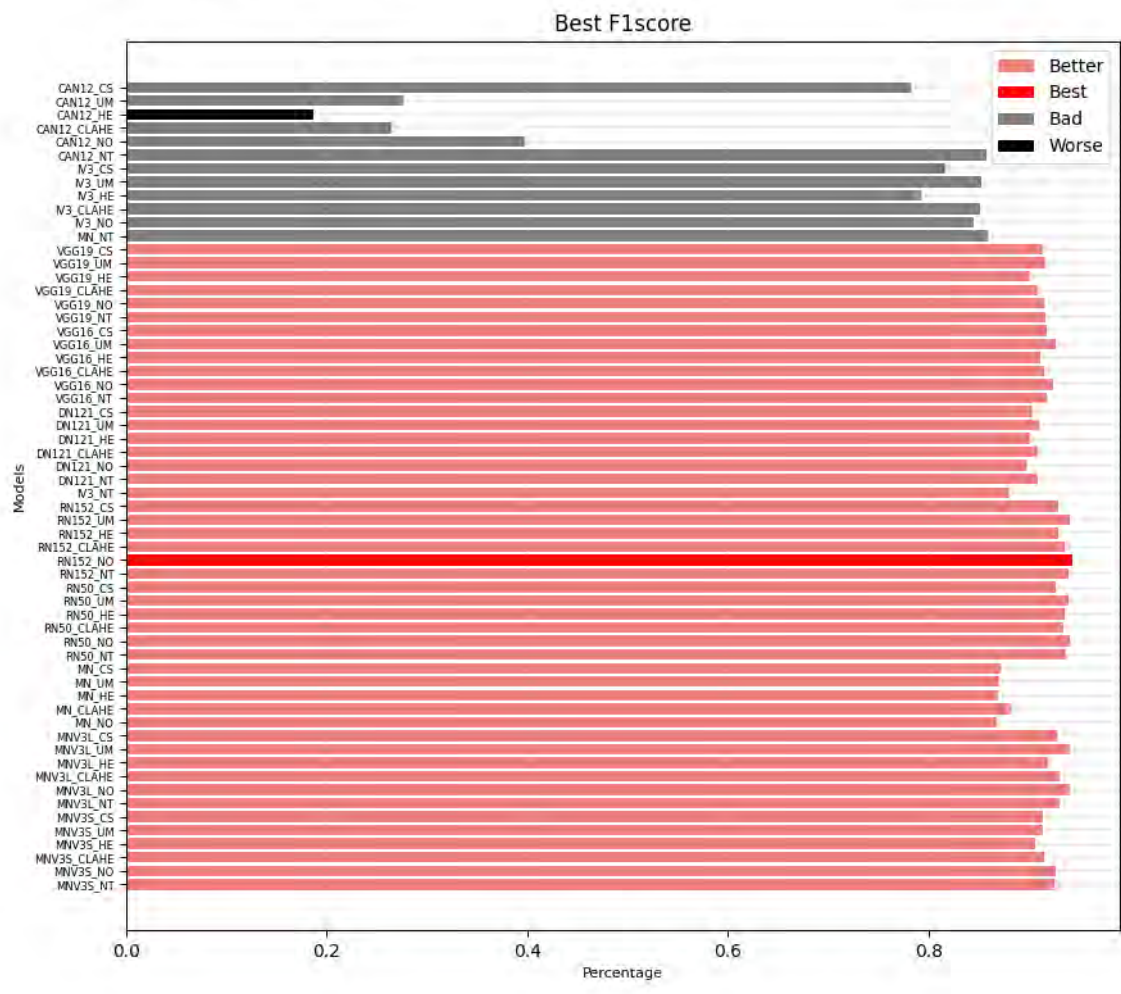


Ilustración 54 Comparación de los modelos en una clasificación multiclase en la métrica de F1-score en la etapa de prueba usando el conjunto de datos de CBIS-DDSM.

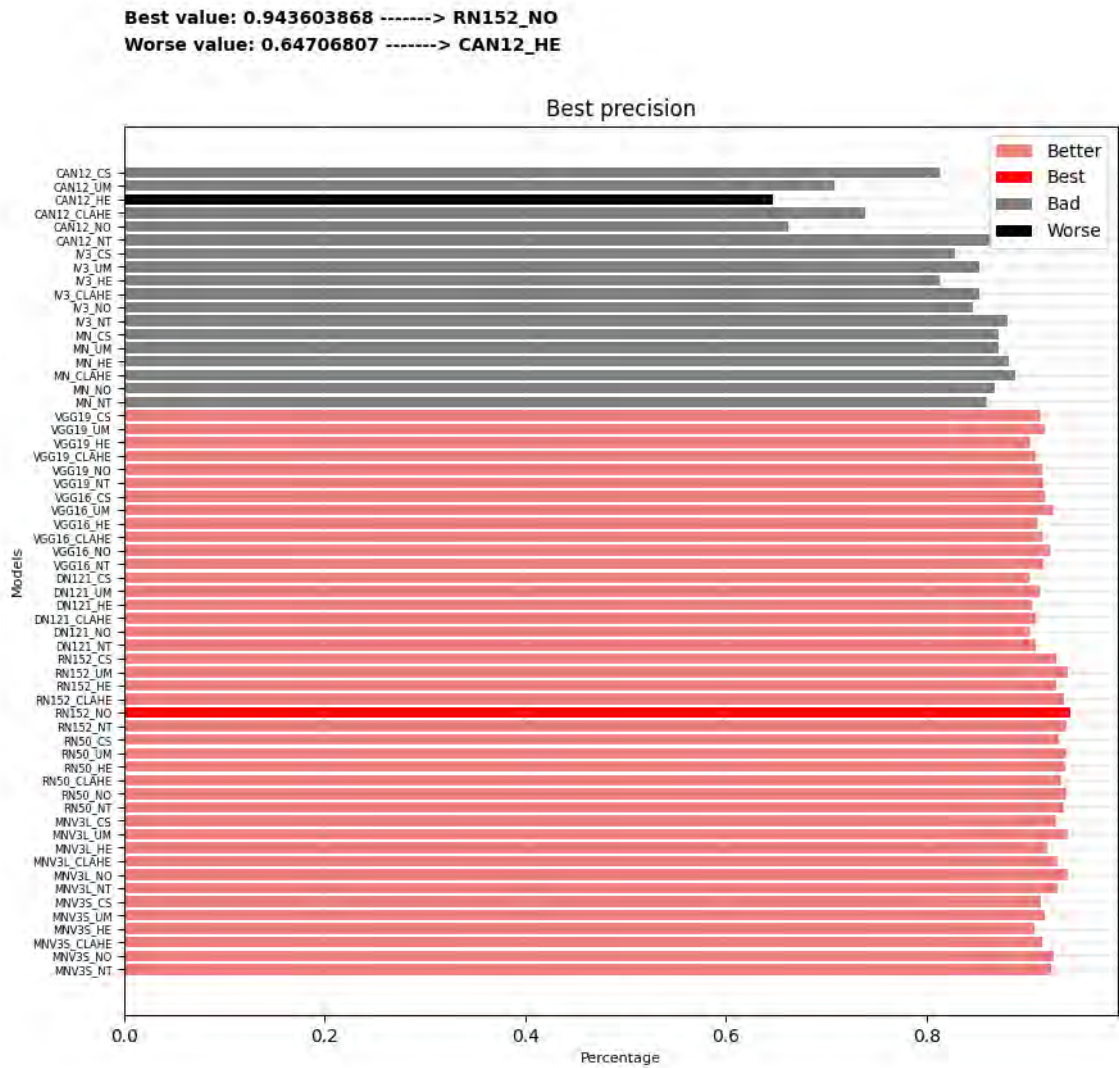


Ilustración 55 Comparación de los modelos en una clasificación multiclase en la métrica de precisión en la etapa de prueba usando el conjunto de datos de CBIS-DDSM.

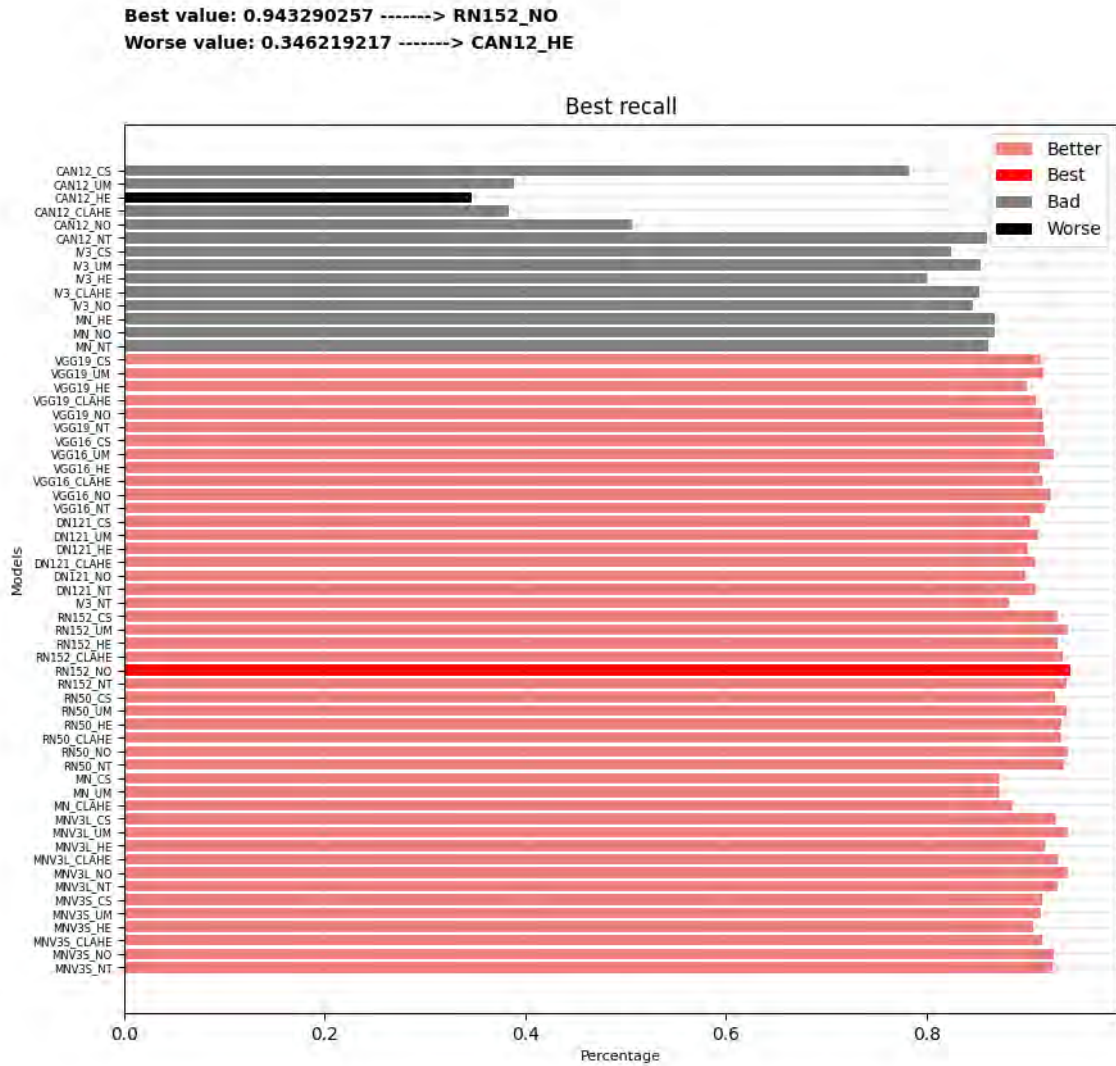


Ilustración 56 Comparación de los modelos en una clasificación multiclase en la métrica de exhaustividad en la etapa de prueba usando el conjunto de datos de CBIS-DDSM.

8.2 Clasificación usando aprendizaje profundo en la segunda metodología.

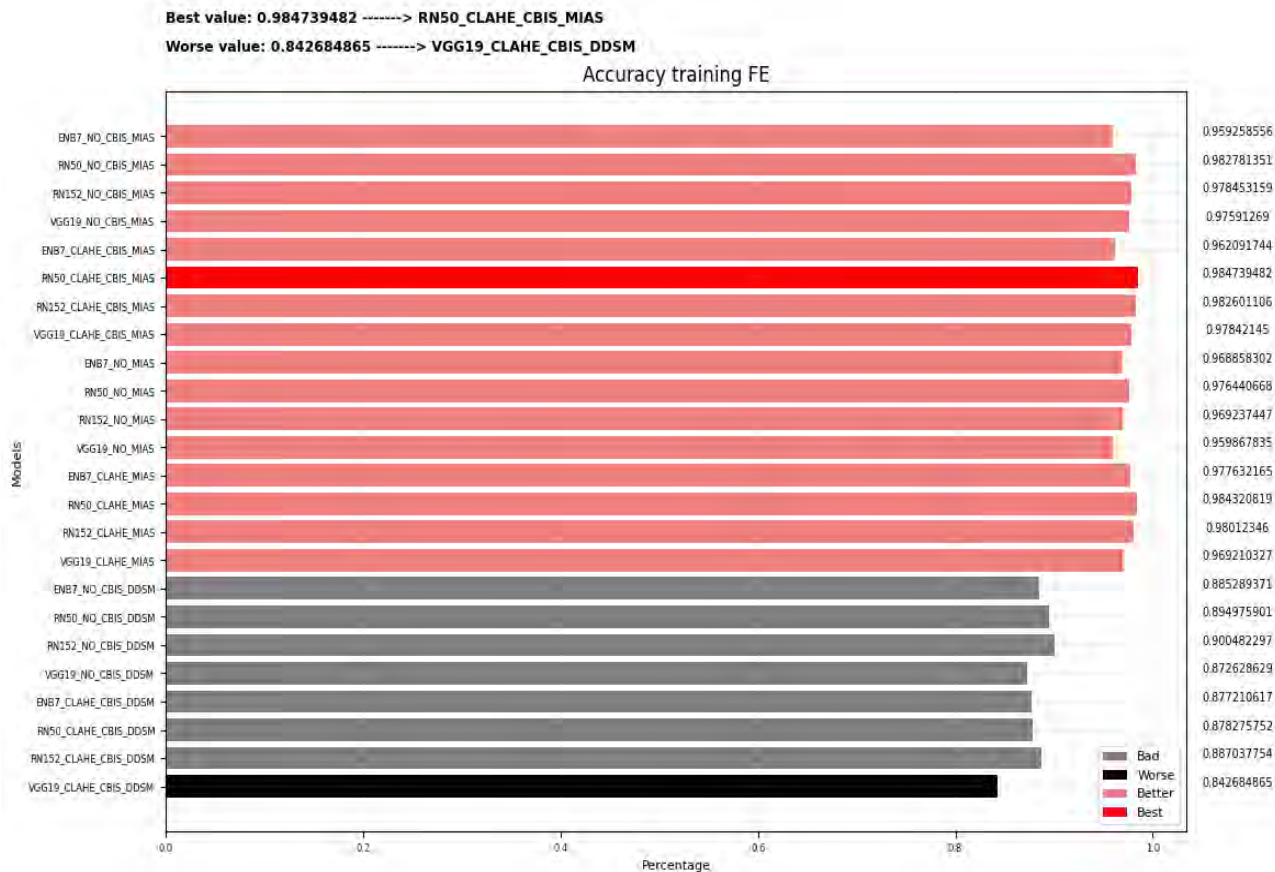


Ilustración 58 Grafica mostrando la métrica de exactitud para la etapa de entrenamiento en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS y CBIS-DDSM. (Las cuatro etapas).

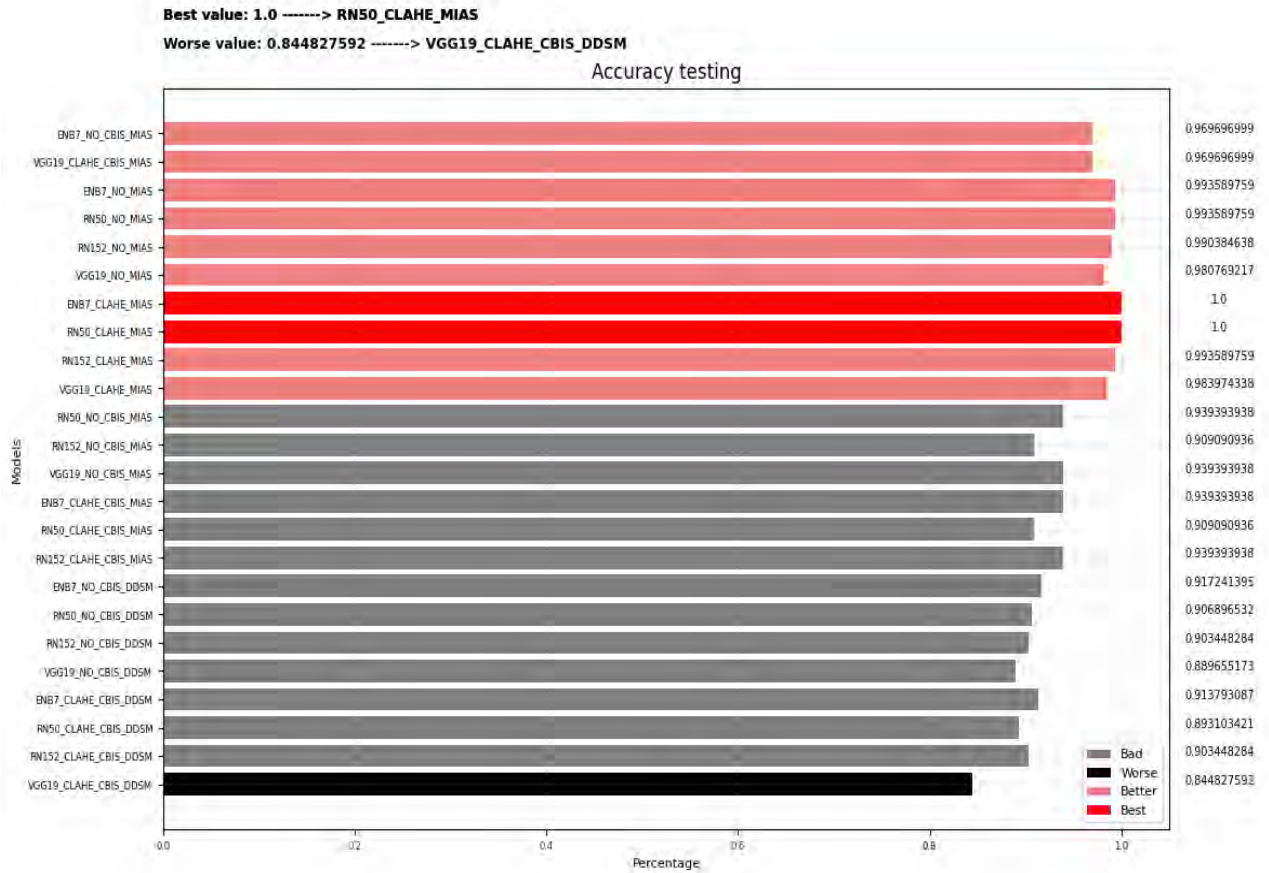


Ilustración 59 Grafica mostrando la métrica de exactitud para la etapa de prueba en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS y CBIS-DDSM. (Las cuatro etapas).

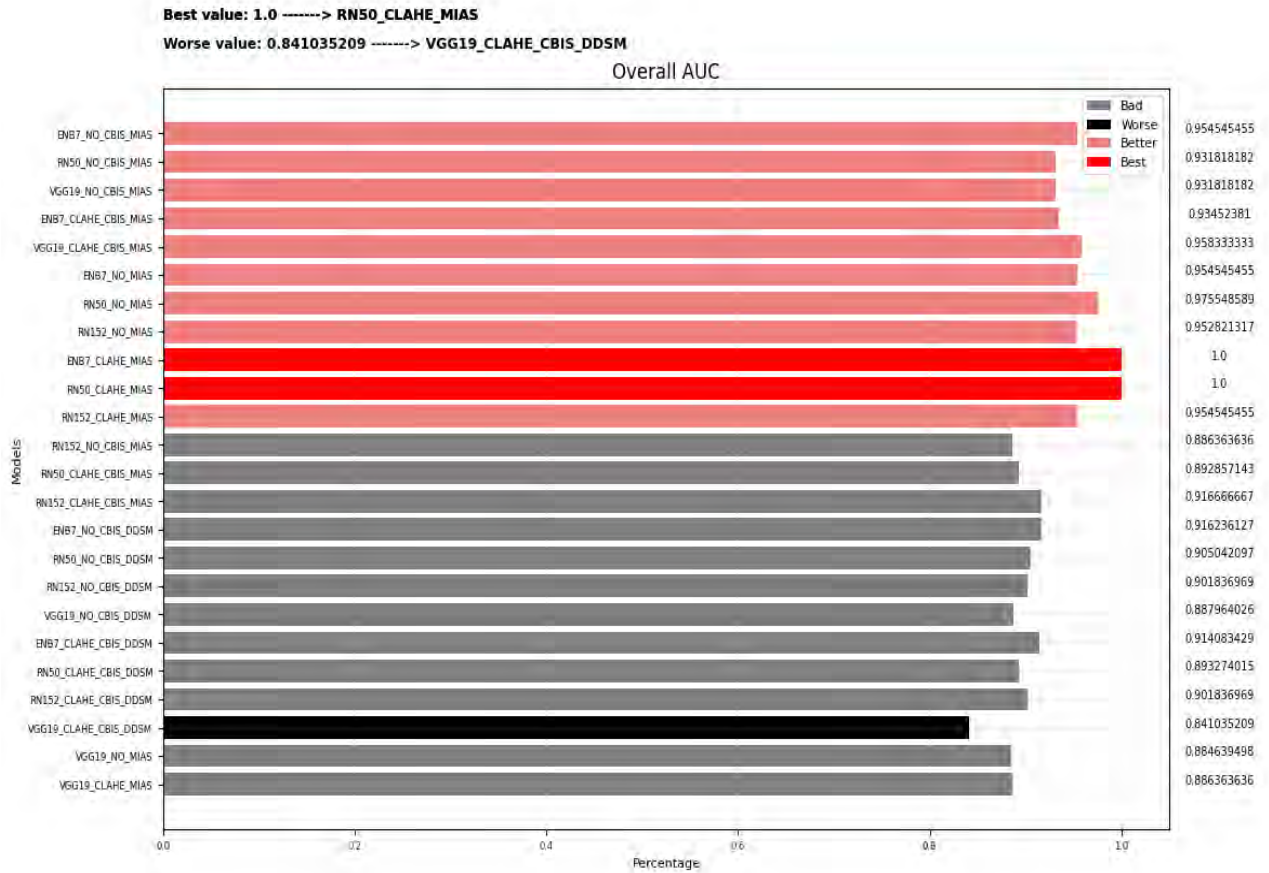


Ilustración 60 Grafica mostrando la métrica de AUC promedio para la etapa de entrenamiento en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS y CBIS-DDSM. (Las cuatro etapas).

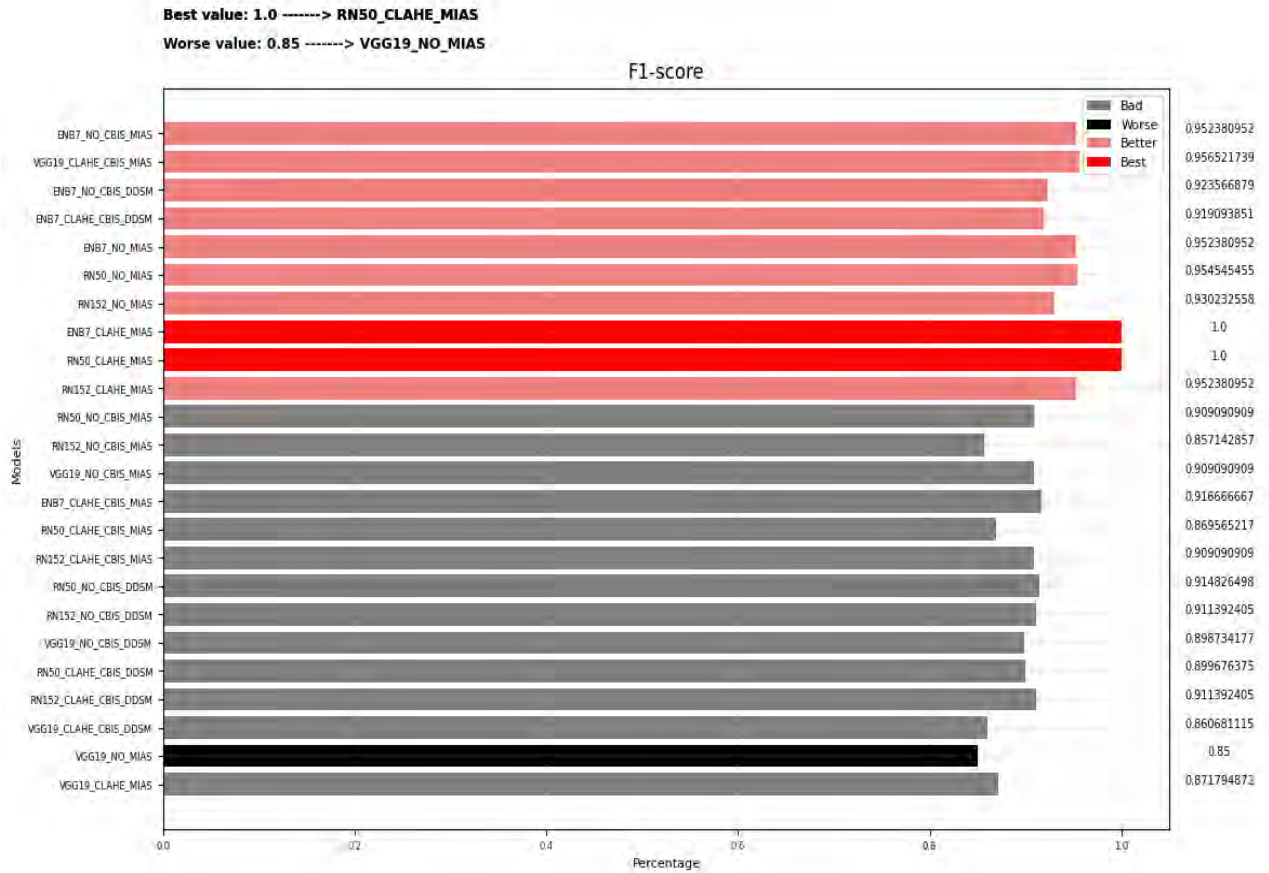


Ilustración 61 Grafica mostrando la métrica de f1-score para la etapa de entrenamiento en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS y CBIS-DDSM. (Las cuatro etapas).

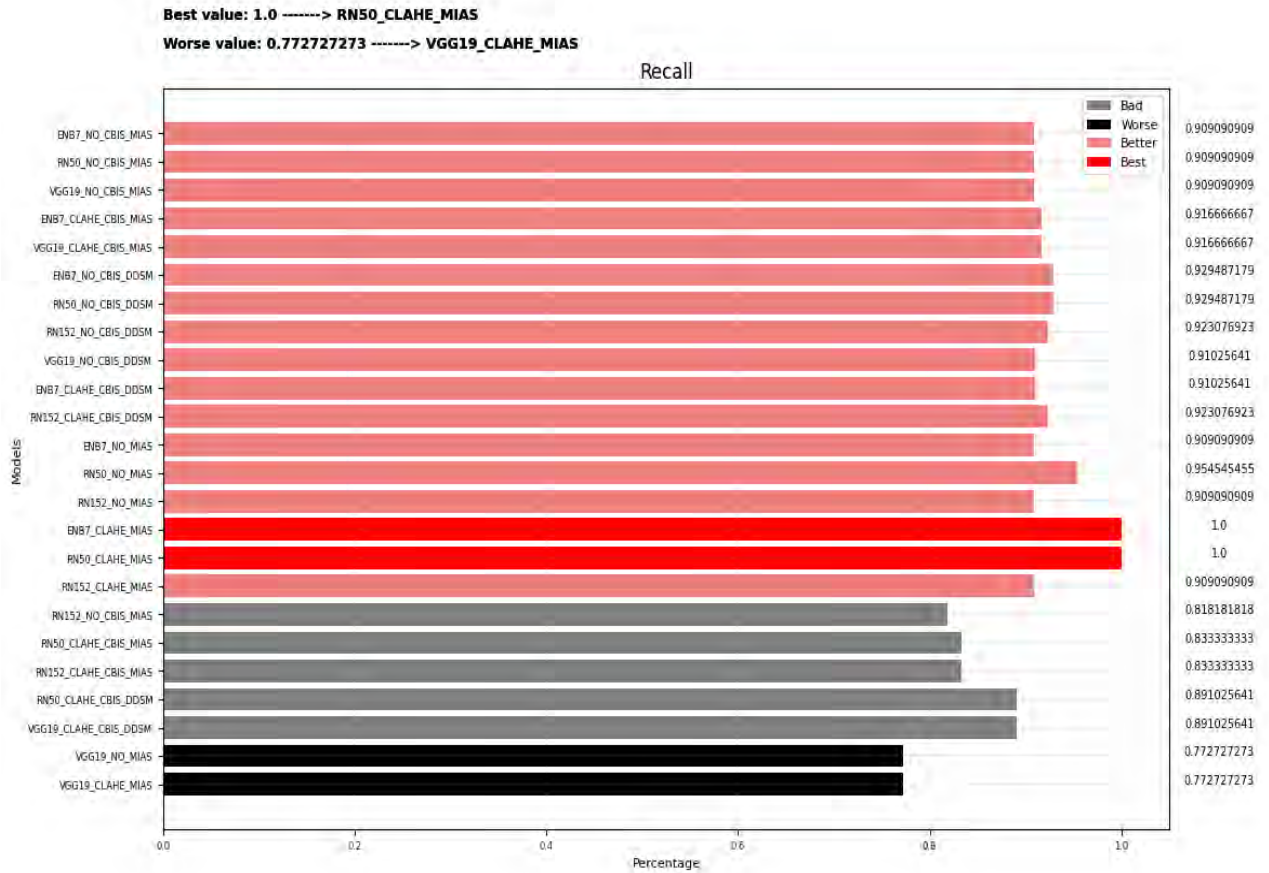


Ilustración 62 Grafica mostrando la métrica de recall para la etapa de entrenamiento en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto de datos Mini-MIAS y CBIS-DDSM. (Las cuatro etapas).

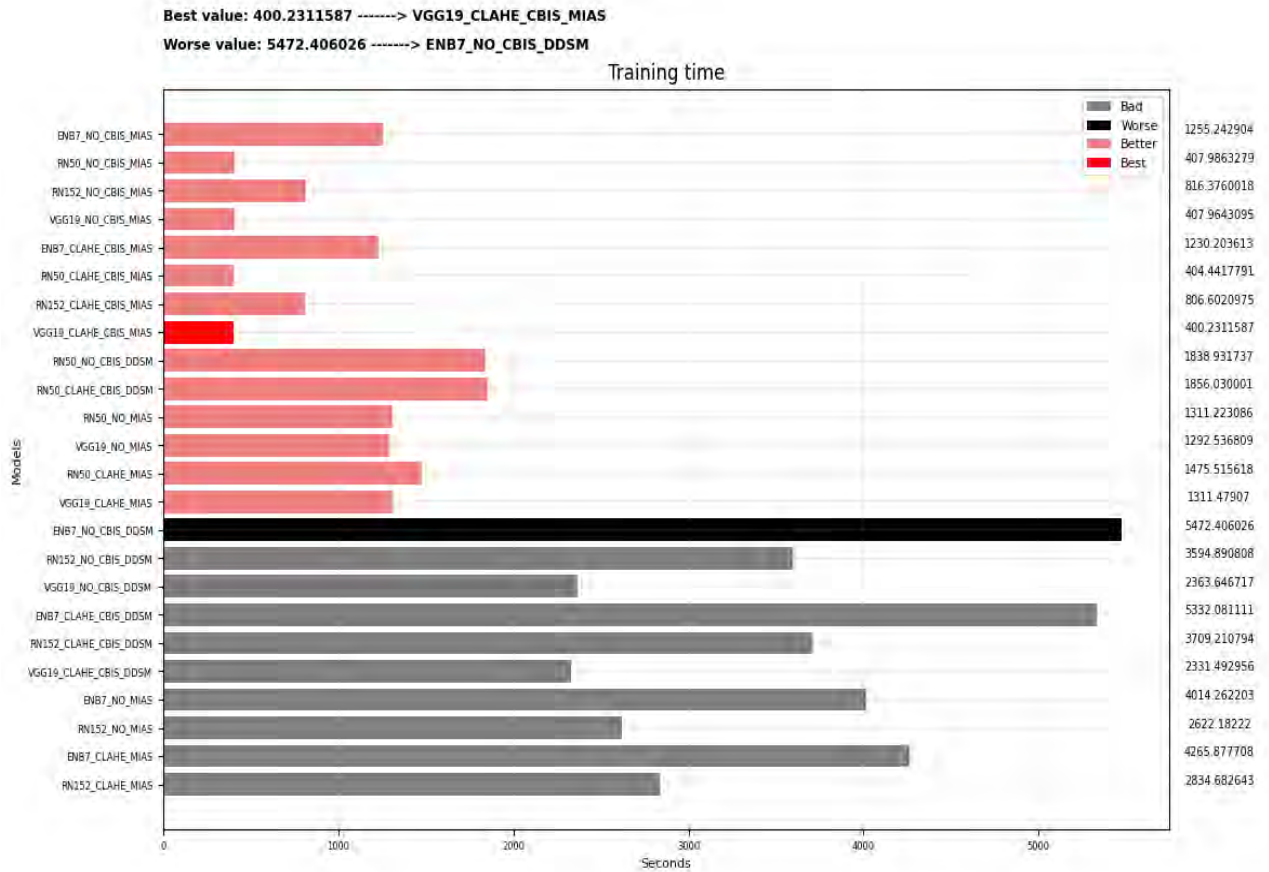


Ilustración 63 Grafica mostrando el tiempo de procesamiento de la etapa de entrenamiento en la segunda metodología, comparando los modelos de CNN pre-entrenadas en una clasificación binaria utilizando el conjunto.