CLASIFICACIÓN DE NORMALIDAD EN TEJIDO DE MAMAS A PARTIR DE IMÁGENES DIGITALES DE RAYOS X

DIEGO FERNANDO VELASCO GIRALDO

ESCUELA DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

FACULTAD DE INGENIERÍA

UNIVERSIDAD DEL VALLE

SANTIAGO DE CALI

OCTUBRE DE 2022

CLASIFICACIÓN DE NORMALIDAD EN TEJIDO DE MAMAS A PARTIR DE IMÁGENES DIGITALES DE RAYOS X

DIEGO FERNANDO VELASCO GIRALDO

Proyecto de grado para optar al título de

MAGISTER EN INGENIERÍA CON ÉNFASIS EN INGENIERÍA ELECTRÓNICA

En modalidad de

PROFUNDIZACIÓN

Director

HUMBERTO LOAIZA CORREA, Ph.D.

Escuela de Ingeniería Eléctrica y Electrónica

Codirectora

MARIA FERNANDA TOBAR BLANDÓN, M.Sc.

Escuela de Salud Pública

GRUPO DE INVESTIGACIÓN PERCEPCIÓN Y SISTEMAS INTELIGENTES

ESCUELA DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

FACULTAD DE INGENIERÍA

UNIVERSIDAD DEL VALLE

SANTIAGO DE CALI

OCTUBRE DE 2022

TABLA DE CONTENIDO

LISTA DE FIGURAS

LISTA DE TABLAS

Nota de aceptación

_____

_____

Director

_____

Jurado

_____

Jurado

Resumen

El cáncer de mama es uno de los tipos de cáncer más comunes en el mundo, con mayor incidencia y mortalidad en mujeres. Sin embargo, se ha demostrado que la mayoría de mujeres que obtienen un diagnóstico temprano de la enfermedad y que reciben tratamiento, tienen muy alta probabilidad de supervivencia. Entre los métodos iniciales más utilizados para estudiar y diagnosticar el cáncer de mama están las mamografías digitales (imágenes digitales del interior del seno tomadas con rayos X), en las cuales se pueden observar nódulos, calcificaciones u otras anormalidades, e incluso, detectar la naturaleza de las mismas (benigna o maligna).

Los radiólogos, médicos especialistas en interpretar imágenes de rayos X, deben leer y diagnosticar cantidades considerables de mamografías por día. Además de que no son fáciles de interpretar, la experticia del radiólogo y factores adicionales como la calidad de la imagen y la fatiga del profesional juegan un papel importante en la sensibilidad y especificidad del diagnóstico. Por lo que, en países desarrollados, sugieren un doble diagnóstico por distintos radiólogos, y si no coinciden, consultan a un tercero. Como herramienta tecnológica de apoyo para que estos especialistas realicen diagnósticos con mayor certeza o cuenten con una opinión adicional, en el presente proyecto se implementó un sistema de visión artificial que permite extraer y relacionar características de imágenes mamográficas digitales para clasificar su tejido como anormal o normal.

Para lo anterior, se utilizaron técnicas de procesamiento de imágenes y reconocimiento de patrones. Específicamente, se implementaron dos métodos distintos para la segmentación de la mama, una que se basa en operaciones morfológicas, entropía, etc., y otra en la arquitectura de red neuronal convolucional U-Net. Para la clasificación y detección de anormalidades, se aplicaron técnicas basadas en máquinas de aprendizaje profundas (arquitecturas de redes neuronales convolucionales DenseNet121 + U-Net) y en máquinas de aprendizaje superficiales

(perceptrón multicapa MLP y máquina de soporte vectorial SVM). Las entradas de estas dos últimas, son características estadísticas extraídas del histograma de las imágenes, junto con descriptores de textura a partir de la matriz de co-ocurrencia en niveles de gris, de las imágenes originales y de los patrones binarios locales de las imágenes (LBP).

Se obtuvieron resultados sobresalientes. En las dos técnicas de segmentación de la mama desarrolladas se obtuvo un índice de similitud de Jaccard de 0,72 con el primer método y de 0,95 con el segundo, respecto a los *ground truth*.

La clasificación de mamografías como normales o anormales con la arquitectura de red profunda DenseNet121 presentó una exactitud del 94% para el conjunto de *test* y, la detección de anormalidades con la arquitectura U-Net en mamografías de *test* clasificadas como anormales por DenseNet121, presentó un promedio de intersección con los *ground truth* de 0,94.

Con las máquinas de aprendizaje superficiales, se clasificaron regiones normales y anormales de las imágenes para detectarlas posteriormente con un algoritmo de ventana deslizante. La clasificación de estas regiones tuvo una exactitud de 96,2% con el perceptrón multicapa (MLP) y de 97,1% con la máquina de soporte vectorial (SVM). Con estas dos máquinas de aprendizaje, la ventana deslizante detectó algunas anormalidades en las imágenes, pero se puede mejorar con técnicas complementarias en trabajos futuros.

Se espera que este sistema contribuya a la elaboración de herramientas de soporte para los médicos radiólogos y, por ende, reduzca el margen de error en el diagnóstico de mamografías.

Palabras clave: detección de anormalidades en mama, cáncer de mama, mamografía, inteligencia artificial, visión artificial, reconocimiento de patrones.

Abstract

Breast cancer is one of the most common types of cancer worldwide, with a higher incidence and mortality in women. However, it has been shown that most women who get an early diagnosis of the disease and receive treatment have a very high chance of survival. Among the initial methods used to study and diagnose breast cancer are digital mammograms (digital images of the inside of the breast taken with X-rays), in which nodules, calcifications, or other abnormalities can be observed, and even detect the nature of the same (benign or malignant).

Radiologists, doctors who specialize in interpreting X-ray images, must read and diagnose significant numbers of mammograms each day. In addition to not being easy to interpret, the radiologist's expertise and additional factors such as image quality and operator fatigue play an essential role in the sensitivity and specificity of the diagnosis. Therefore, in developed countries, they suggest a dual diagnosis by different radiologists, and if they do not coincide, they consult a third party. As a technological support tool for these specialists to make diagnoses with greater certainty or have a different opinion, in this project, a computer vision system is implemented, it allows extracting and relating features of digital mammographic images to classify their tissue as abnormal or normal.

For this, image processing and pattern recognition techniques were used. Specifically, two different methods for breast segmentation were implemented, one based on morphological operations, entropy, etc., and the other based on the U-Net convolutional neural network architecture. For the classification and detection of abnormalities, techniques based on deep learning machines (DenseNet121 + U-Net convolutional neural network architectures) and surface learning machines (multilayer perceptron MLP and support vector machine SVM) were applied. The inputs of these last two are statistical characteristics extracted from the histogram of the images, together with texture descriptors from the gray-level co-occurrence

matrix (GLCM), from the original images, and from the local binary patterns of the images (LBP).

Outstanding results were obtained. In the two breast segmentation techniques developed, a Jaccard similarity index of 0.72 was obtained with the first method and 0.95 with the second regarding ground truth.

The classification of mammograms as normal or abnormal with the DenseNet121 deep network architecture gave an accuracy of 94% for the test set, and the detection of abnormalities with the U-Net architecture in test mammograms classified as abnormal by DenseNet121 gave an average intersection with the ground truth of 0.94.

Using surface learning machines, normal and abnormal regions of the images were classified for later detection using a sliding window algorithm. The classification of these regions had an accuracy of 96.2% with the multilayer perceptron (MLP) and 97.1% with the support vector machine (SVM). With these two trained learning machines,, the sliding window detected some abnormalities in the images, but it can be improved with complementary techniques in future works.

This system is expected to contribute to the development of support tools for radiologists and thus reduce the margin of error in diagnosing mammograms.

Keywords: detection of breast abnormalities, breast cancer, mammography, artificial intelligence, computer vision, pattern recognition.

1. Introducción

Entre los tipos de cáncer conocidos en la actualidad, el cáncer de mama o de seno es el más común a nivel mundial y uno de los que más muertes de mujeres genera. A pesar de ser una enfermedad tratable con gran posibilidad de éxito de recuperación si es detectada a tiempo, frecuentemente es diagnosticada en estadios avanzados, por lo que los métodos y herramientas de diagnóstico juegan un papel fundamental en el tratamiento adecuado del cáncer de mama.

Las mamografías son una de las primeras etapas para la detección de anormalidades después del tacto mamario. Los encargados de interpretar las mamografías son los radiólogos, esta interpretación está sujeta a errores humanos relacionados con la fatiga y experticia del especialista debido a la complejidad de la lectura de muchas de ellas.

Para contribuir a los radiólogos a aumentar el número de diagnósticos de mamografías digitales y a reducir su margen de error, en el presente proyecto de grado de maestría se estableció la siguiente pregunta problematizadora:

¿Cómo se puede establecer la condición de normalidad en tejido mamario a partir del procesamiento digital de mamografías?

Para dar solución al problema planteado, se desarrolló un sistema para determinar la normalidad y anormalidad del tejido mamario a partir del análisis de mamografías, utilizando técnicas de procesamiento digital de imágenes y de reconocimiento de patrones. El sistema se entrenó y probó con mamografías de mujeres disponibles en bases de datos públicas. Este trabajo no se enfocó en determinar el tipo de anormalidad presente en el tejido.

Específicamente, se implementaron dos técnicas de segmentación de imágenes y

tres técnicas de reconocimiento de patrones para clasificar el tejido mamario de mujeres como normal o anormal a partir de mamografías digitales.

Finalmente, se evaluaron las rutas de solución implementadas y se eligió la óptima para resolver el problema planteado.

Se pretende que lo desarrollado se utilice en trabajos futuros para crear una herramienta de apoyo que sea utilizada por los especialistas en el campo en cuestión.

1.1. Justificación

El cáncer de mama es el cáncer con mayor incidencia a nivel mundial y es la principal causa de muerte en las mujeres según los últimos indicadores publicados por la Organización Mundial de la Salud. En el año 2020 se registraron más de 2,2 millones de casos en el mundo y aproximadamente 685.000 mujeres fallecieron a causa de esta enfermedad durante el mismo año (Figura 1). Según esta entidad, una de cada 12 mujeres desarrollará cáncer de mama en su vida. La mayoría de los casos y muertes por cáncer de mama ocurren en países de ingresos bajos y medios.

La enfermedad se relaciona generalmente con las mujeres, pero los hombres también la pueden padecer, aunque solo se da entre el 0,5 y 1% de los casos respecto a los presentados en mujeres [1].

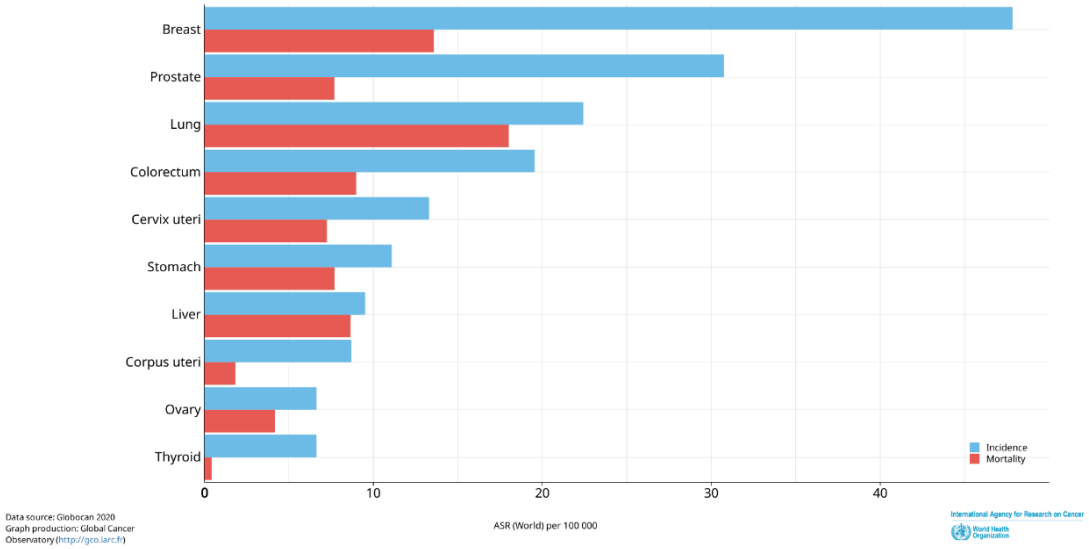


Figura 1. Tasas estimadas de incidencia y mortalidad estandarizadas por edad en el mundo, ambos sexos, todas las edades (2020).
[2]

En Colombia, el cáncer de mama presenta el nivel de incidencia más alto en comparación con los demás tipos de cáncer y ocupa el cuarto lugar en número de muertes ocasionadas, por debajo del cáncer de estómago, de pulmón y colorrectal. En nuestro país, se registraron más de 15.500 casos de cáncer de mama en 2020 y debido a esta enfermedad aproximadamente 4.400 mujeres fallecieron [2].

Aunque las cifras relacionadas con la presencia de la enfermedad son altas, las muertes ocasionadas por cáncer de mama pueden reducir notablemente si se realiza un diagnóstico temprano de la enfermedad. Cuando el cáncer de mama se detecta y trata a tiempo, las posibilidades de supervivencia son muy altas [3].

Entre los métodos de diagnóstico para detectar patologías en mama, los basados en imágenes diagnósticas son los más utilizados en etapas iniciales antes de realizar pruebas de patología del tejido. El diagnóstico del cáncer de mama basado en imágenes se orienta a la detección de anomalías en el tejido de los senos empleando técnicas como ecografías, imágenes de resonancia magnética, imágenes a partir de sensores infrarrojos (termografías) y la más utilizada; imágenes

de rayos X (mamografías). Actualmente la mayoría de los programas de detección de cáncer de seno establecen que las imágenes diagnósticas sean leídas de manera independiente por dos especialistas distintos, como estrategia para reducir el riesgo de interpretaciones erradas [4]. Sin embargo, la doble lectura incurre en una carga de trabajo, retardos y costos adicionales para las instituciones de salud, el especialista y el paciente.

Algunas de las anormalidades en las mamografías que pueden asociarse con una posible presencia de cáncer son las calcificaciones, nódulos o masas, asimetrías en densidad, y/o distorsión de la arquitectura del seno. La cantidad, ubicación y características de las anormalidades encontradas, son utilizadas por los radiólogos para interpretar y realizar los diagnósticos de las mamografías. Debido a que la inspección de las mamografías se efectúa regularmente de forma visual, la precisión del diagnóstico está fuertemente condicionada por la experticia, agudeza visual y fatiga del especialista médico.

Los sistemas de visión artificial pueden contribuir a la caracterización y detección del tejido presente en una mamografía, favorecer la disminución de diagnósticos errados por el especialista y podrían aportar la doble lectura mencionada anteriormente, debido a que siempre no se cuenta con dos especialistas para diagnosticar una misma mamografía, sobre todo en países de recursos económicos limitados, como el nuestro. Sin embargo, el análisis automático a partir de imágenes de rayos X presenta un reto debido a las diferencias en contraste, ubicación, traslape, tamaño y forma de los tejidos sanos y anormales.

En la actualidad existen algunos sistemas comerciales para detección asistida por computador (CAD por sus siglas en inglés) en mamografías digitales y algunos en resonancias magnéticas (MRI). Los más utilizados por clínicas especializadas son los sistemas “R2 ImageChecker CAD” de la empresa “Hologic Inc.” y el sistema “SecondLook” de “iCad Inc.” [5]. Ambos cuentan con patentes y certificaciones de

la FDA [6][7]. Clínicas reconocidas como el “Washington Hospital Healthcare System” [8], el “GWU Medical Center” [9] y el “Baptist Health Louisville” [10] en Estados Unidos los utilizan.

Actualmente, el desarrollo de sistemas automáticos de apoyo al diagnóstico de imágenes médicas es un campo abierto de investigación [11] [12] [13]. Inclusive, diversas instituciones organizan retos abiertos a la comunidad científica sobre análisis de imágenes biomédicas[1]. En la Universidad del Valle, el laboratorio de Visión Artificial del grupo PSI cuenta con experiencia en trabajos de pregrado y proyectos de investigación de posgrado orientados al análisis del tejido mamario con imágenes de rayos X[2] e imágenes termográficas[3].

1.2. Objetivos

1.2.1. Objetivo General

Implementar un sistema de visión artificial para establecer automáticamente la normalidad del tejido mamario a partir del análisis de imágenes digitales de rayos X.

1.2.2. Objetivos Específicos

1. Identificar las principales características presentes en mamografías digitales que permitan clasificar la normalidad del tejido.

2. Evaluar dos técnicas de segmentación para localizar regiones de interés en

[1] https://grand-challenge.org/challenges/

[2] Herramienta de Ayuda en la Localización de Microcalcificaciones en Mamografías Mediante Técnicas de Procesamiento Digital de Imágenes, trabajo de Grado, Mónica Millán Giraldo, Universidad del Valle, 2005.

[3] Sistema de Procesamiento de Imágenes Termográficas Como Apoyo a la Detección de Nódulos en Mamas de Mujeres, Trabajo de Grado de Maestría, Steve Rodríguez Guerrero, Universidad del Valle, 2017.

las mamografías.

3. Evaluar tres técnicas de reconocimiento de patrones para establecer la normalidad en el tejido mamario sobre las regiones segmentadas.

4. Establecer alcances y limitaciones de las técnicas de segmentación y reconocimiento de patrones implementadas.

2. Antecedentes

Existen trabajos importantes sobre procesamiento de mamografías digitales o digitalizadas, orientados a segmentar e identificar características para detectar la presencia de anormalidades, cáncer de mama o establecer su estadio. A continuación, se presentan algunos trabajos desarrollados a nivel nacional e internacional relacionados con esta temática.

2.1. Antecedentes nacionales

2.1.1. Representation learning for mammography mass lesion classification with convolutional neural networks

Este es un trabajo realizado por investigadores de la Universidad Nacional de Colombia y la Universidad Industrial de Santander, tiene como objetivo la clasificación automática de lesiones en imágenes mamarias, donde integran técnicas de aprendizaje profundo para aprender automáticamente las características discriminatorias, evitando el diseño de detectores de características específicos basados en imágenes hechos a mano.

Construyeron un conjunto de datos a partir de 344 casos de pacientes con cáncer de mama que contenían un total de 736 vistas de mamografías (medio lateral oblicuo y craneocaudal), representativas de lesiones segmentadas manualmente asociadas a masas: 426 lesiones benignas y 310 lesiones malignas. Estas imágenes fueron tomadas del “Repositorio Digital de Cáncer de Mama” (BCDR por sus siglas en inglés), el cual es un amplio repositorio público compuesto por casos de pacientes con cáncer de mama en la región norte de Portugal. La creación de BCDR fue apoyada por el proyecto IMED (Para el Desarrollo de Algoritmos para Análisis de Imágenes Médica: traducción del significado de sus siglas en inglés) destinado a crear repositorios de imágenes médicas y exploración masiva de métodos de Diagnóstico Asistido por Computadora (CADx por sus siglas en inglés).

El método desarrollado para lograr el objetivo de la investigación comprende dos etapas principales: preprocesamiento para mejorar los detalles de la imagen y entrenamiento supervisado para inferir tanto las características como el ajuste de los parámetros del clasificador de lesiones de imágenes mamarias.

Adoptan un enfoque híbrido en el que se utilizan redes neuronales convolucionales para aprender la representación de forma supervisada en lugar de diseñar descriptores particulares para explicar el contenido de las imágenes de mamografía.

Los resultados obtenidos luego de aplicadas distintas técnicas de procesamiento y clasificación mencionadas obtuvieron una precisión máxima del 82% con una arquitectura de red convolucional propia, que llamaron CNN3, la cual fue entrenada con la Unidas de Procesamiento Gráfico: Tesla K40 GPU y tomó 1.4 horas en culminar [13].

2.1.2. Detección de microcalcificaciones mamarias agrupadas

Este es otro trabajo realizado en Colombia relacionado con el presente proyecto de grado de maestría, no se enfoca en la anormalidad del tejido mamario, sino específicamente en la detección de microcalcificaciones mamarias, pero se tiene en cuenta porque el procesamiento de las imágenes para detectar estas microcalcificaciones debe ser muy bueno, porque identificar las microcalcificaciones en mamografías digitales es una tarea compleja debido a su diminuto tamaño.

Para el desarrollo de este trabajo se empleó el *dataset* de imágenes mamográficas que dispone la Universidad del Sur de la Florida, denominado “Base de Datos Digital para Mamografía de Detección” (DDSM por sus siglas en inglés). El cual es constituido por más de por más de 2620 casos, donde cada uno posee cuatro imágenes correspondientes a dos proyecciones de cada seno: medio lateral oblicua

(MLO) y cráneo caudal (CC).

En el algoritmo propuesto, aplican inicialmente una umbralización global a las imágenes para eliminar la información que no se requiere en las mamografías, como etiquetas y demás, luego implementan técnicas de segmentación como detección de mayor área y métodos de realce donde hacen uso de la transformada discreta Wavelet.

Finalmente, con las microcalcificaciones resaltadas en las mamografías digitales, proponen realizar un conteo manual de las mismas para que el especialista pueda brindar un diagnóstico con mayor facilidad [14].

Las técnicas de procesamiento utilizadas en este trabajo se tuvieron en cuenta en este trabajo, pues, aunque no aplican técnicas de clasificación, un buen procesamiento de las imágenes es de gran importancia, ya que imágenes bien procesadas, que son la entrada a las máquinas de aprendizaje, permitirán una mejor clasificación automática.

2.2. Antecedentes internacionales

2.2.1. Feature subset selection for classification of malignant and benign breast masses in digital mammography

El objetivo principal de la investigación mostrada en este artículo es el de seleccionar un conjunto de características óptimo para mejorar el rendimiento de la clasificación de masas benignas y malignas en mamografías digitales utilizando técnicas de visión artificial y máquinas de aprendizaje automático.

No indican el *dataset* utilizado para entrenamiento y pruebas de las máquinas de aprendizaje implementadas.

Las etapas principales del sistema son:

- Extracción de características.
- Reducción de dimensiones de las características.
- Evaluación de comportamiento.
- Selección de características.

Realizaron una extracción de distintas características estadísticas, características morfológicas utilizando la Transformada Wavelet, hallaron y definieron las áreas de interés teniendo en cuenta el área, perímetro, orientación, etc. Luego de obtener gran cantidad de características aplicaron una reducción de la dimensionalidad para seleccionar solamente las características más discriminantes.

Las máquinas de aprendizaje utilizadas fueron:

- Redes neuronales multicapa.
- Máquinas de soporte vectorial.
- K vecinos más cercanos.

Durante las pruebas de clasificación la máquina de aprendizaje que mejor desempeño tuvo fue la máquina de soporte vectorial, con una precisión de 90,9% [15].

2.2.2. Mammogram classification using convolutional neural networks

En este artículo, se plantea principalmente el desarrollo de un sistema que permita determinar la existencia de anormalidades en mamografías digitales utilizando Redes Neuronales Convolucionales (CNNs), donde posteriormente se muestra el análisis de su rendimiento.

El *dataset* utilizado para entrenamiento y pruebas de la CNN es el MIAS *dataset*.

El proceso de procesamiento y clasificación de las imágenes mamográficas digitales seguido se muestra a continuación:

- Preprocesamiento de imágenes recortando las anormalidades en las coordenadas indicadas en el *dataset* (submuestreo), también implementaron técnicas de aumento de datos aplicando rotaciones a las imágenes mamográficas.

- Clasificación con arquitectura de red neuronal convolucional propia utilizando TensorFlow.

Máquina de aprendizaje utilizada:

- Red Neuronal Convolucional con arquitectura propia.

Durante las pruebas de clasificación, la precisión que se obtuvo con la máquina de aprendizaje implementada (CNN), presentó una precisión de 75% [16].

2.2.3. Convolutional neural network improvement for breast cancer classification

Este artículo propone un sistema similar al que se pretende desarrollar en el presente proyecto, con la diferencia que no clasifica la normalidad del tejido mamario presente en mamografías digitales, sino que clasifica la benignidad o malignidad del tejido mamario presente en mamografías digitales.

El *dataset* utilizado para entrenamiento y pruebas de la CNN es el mini-MIAS *dataset*.

Las fases principales del sistema desarrollado son:

- Redimensionamiento de las imágenes del *dataset*.
- Entrenamiento de las CNNs utilizadas (distintas arquitecturas)
- Pruebas de clasificación con las diferentes arquitecturas.
- Porcentajes obtenidos de mamografías digitales con tejido normal, masas benignas y malignas.

En el procesamiento de las imágenes implementaron aumento de datos, submuestreo y cambio de resolución, teniendo en cuenta ubicación de la anormalidad conocida en las mamografías catalogadas como anormales en el *dataset*.

Las arquitecturas de las Redes Neuronales Convolucionales (CNNs) implementadas y probadas fueron:

- Adaboost.
- Neuroph MLP
- BCDCNN (versiones 1, 2, 3).
- CNNI-BCC.

Durante las pruebas de clasificación con las distintas arquitecturas, se obtuvo una precisión de clasificación máxima de 90,5% [17].

2.3. Discusión

En la Tabla 1, se presentan de manera resumida las principales características de los sistemas de procesamiento y clasificación de tejido mamario expuestos anteriormente. Se puede apreciar que se utilizan diferentes *dataset*s, técnicas de

procesamiento y de clasificación. De igual manera, se destaca la diversidad de arquitecturas en los clasificadores implementados.

Tabla 1. Resumen discusión antecedentes.

Referencia	***Dataset* Utilizado**	**Técnicas de Procesamiento**	**Técnicas de Reconocimiento de Patrones**	**Precisión**
[13]	BCDR	Recorte, Aumento de Datos, Normalización Global del Contraste, Normalización Local del Contraste	Redes Neuronales Convolucionales	82%
[14]	DDSM	Umbralización Global, Segmentación, Método de realce, Método de realce con Transformada Wavelet	No aplica	No aplica
[15]	No indicado	Extracción de diversas características estadísticas - Características Morfológicas con Transformada Wavelet - Pixeles en área de interés – Perímetro – Orientación, etc. - Reducción de dimensiones - Selección de características más relevantes	Redes Neuronales Multicapa, Máquinas de Soporte Vectorial, K Vecinos más Cercanos	90,9%
[16]	MIAS *Dataset* (Mini)	Recorte de anormalidades en coordenadas indicadas en el *dataset* (submuestreo) - Técnicas de aumento de datos aplicando rotaciones a las imágenes mamográficas.	Redes Neuronales Convolucionales utilizando SciKit Flow	75%
[17]	MIAS *Dataset* (Mini)	Aumento de Datos – Submuestreo y cambio de resolución (teniendo en cuenta ubicación de la malignidad conocida)	Redes Neuronales Convolucionales (Adaboost, Neuroph MLP, BCDCNN, CNNI-BCC)	90,5%

3. Marco Teórico

A continuación, se definen algunos conceptos sobre la glándula mamaria, el cáncer de mama y su diagnóstico. También se muestran los conceptos principales de visión artificial, reconocimiento de patrones e inteligencia artificial aplicados en el desarrollo del presente proyecto.

3.1. Mama, cáncer de mama y diagnóstico

3.1.1. Anatomía y fisiología de la mama

Las mamas son dos estructuras ubicadas en la pared torácica anterior del cuerpo humano. Se extienden desde la clavícula hasta unos centímetros más abajo de la axila y hasta la mitad de la caja torácica, en cada lado. Están presentes en hombres y mujeres, pero en las mujeres se siguen desarrollando después de la pubertad hasta hacerse más prominentes. Esto porque en las mujeres, los senos contienen glándulas mamarias, las cuales cumplen una función nutricia durante la lactancia materna. Estas glándulas también segregan algunas hormonas.

La glándula mamaria se compone de varias secciones llamadas lóbulos (15 a 20 lóbulos). Cada uno de estos lóbulos está conformado por muchos lóbulos más pequeños denominados lobulillos, estos son los que producen la leche en mujeres lactantes. Tanto los lóbulos como los lobulillos están conectados por unos ductos o tubos conocidos como conductos galactóforos, los cuales cumplen la función de llevar la leche al pezón. Estos ductos son las estructuras mamarias donde generalmente comienza a formarse el cáncer.

La mama también está compuesta por vasos sanguíneos y linfáticos, estos últimos traen linfa, en su mayoría, desde los ganglios linfáticos ubicados en las axilas.

Se compone principalmente de tejido adiposo (graso). A medida que una mujer envejece, especialmente una vez que llega a la menopausia, el tejido mamario glandular se comienza a atrofiar y es reemplazado por más tejido adiposo. El grupo de tejidos en la mama conformados por el tejido conectivo que brinda sostén a los lóbulos y lobulillos, junto con el tejido adiposo, se le conoce como estroma y el conformado por las unidades secretoras y el sistema de conductos, se denomina parénquima [18] (Figura 2).

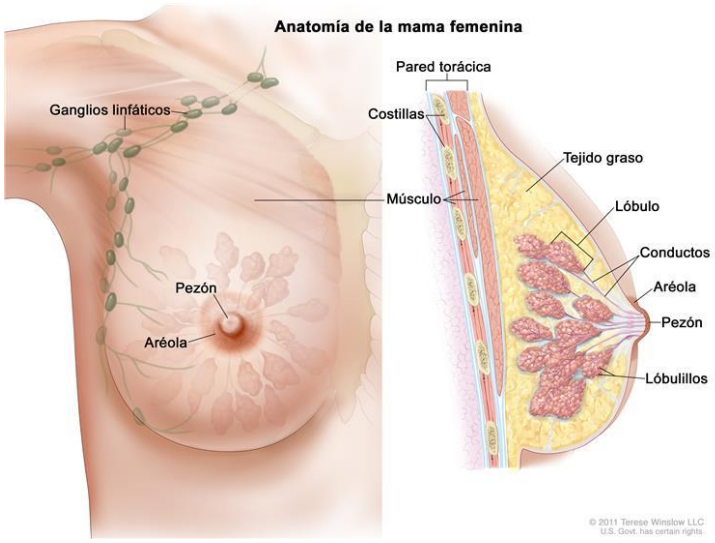


Figura 2. Anatomía de la mama femenina.
[19]

3.1.2. Cáncer de mama

Las células del cuerpo humano se dividen de forma controlada con el objetivo de reemplazar a las células envejecidas o muertas y cuando se presentan células defectuosas, el organismo normal induce en ellas la apoptosis o muerte celular programada. Así se mantiene el correcto funcionamiento de los distintos órganos. Este proceso está regulado por una serie de mecanismos que indican a la célula cuándo comenzar a dividirse, cuándo permanecer estable o cuándo morir (comunicación intercelular a través de distintas proteínas).

Cuando estos mecanismos se alteran, las células pierden su comunicación normal e inician una proliferación acelerada y descontrolada que posteriormente dará lugar a un tumor o nódulo. Si estas células dejan de depender del oxígeno para continuar vivas y proliferándose, adquieren la capacidad de infiltrar otros órganos aledaños y de viajar a través del torrente sanguíneo para luego iniciar un proceso de proliferación similar (metástasis), se dice que hay presencia de cáncer, si no, puede tratarse de un tumor benigno [20] (Figura 3).

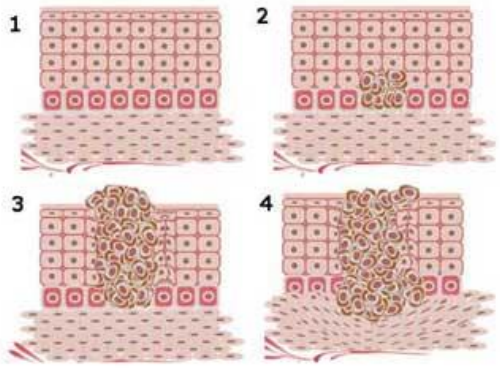


Figura 3. Proliferación e invasión de células cancerígenas.
[20]

El cáncer de mama, comienza cuando las células en el seno presentan los comportamientos anteriormente mencionados, las células del parénquima mamario y en algunos casos del estroma mamario forman un tumor que se puede sentir al tacto y generalmente se puede ver en una imagen de rayos X (mamografía). Esos tumores pueden ser benignos o malignos según sus características morfológicas y patológicas [21].

El tipo de cáncer de mama más común es el carcinoma ductal, que empieza en los conductos galactóforos. Otro tipo de cáncer de mama es el carcinoma lobulillar, que empieza en los lobulillos de la mama. El cáncer de mama invasivo es el que se disemina desde el sitio donde empezó, en los conductos mamarios o lobulillos, hasta el tejido normal circundante [22].

En la Figura 4 se ilustran algunos de los tipos de cáncer de mama:

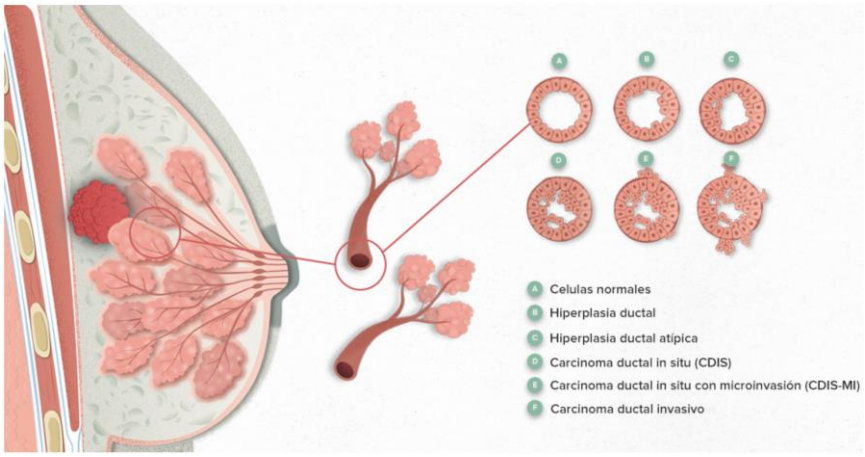


Figura 4. Tipos de cáncer de mama comunes**.**
[23]

3.1.3. Mamografías digitales

La mamografía digital es un procedimiento diagnóstico que utiliza rayos X de baja intensidad para el estudio de patologías de la mama, donde se puede ver el interior de la mama como en una mamografía tradicional, pero la imagen obtenida, al ser digital, puede ser procesada en sistemas computarizados con mayor facilidad y precisión, con todas las ventajas del proceso de imagen digital, que permite aumentar el contraste en las áreas de interés. Algunas mamografías convencionales son digitalizadas en alta calidad, también podrían considerarse mamografías digitales poque pueden ser procesadas en sistemas computacionales.

Las mamografías digitales (Figura 5), permiten la detección a tiempo del cáncer de mama siempre y cuando sea bien tomada e interpretada correctamente. Permiten un resultado preciso en poco tiempo. Según su finalidad, pueden ser preventivas o diagnósticas después de notar síntomas.

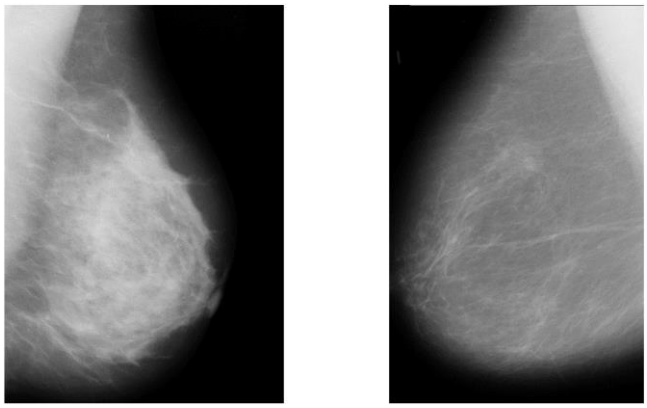


Figura 5. Ejemplos de mamografías digitales.
[24]

3.1.4. Técnicas de toma de mamografías

La exploración se realiza principalmente con dos proyecciones en cada seno, una superior llamada vista o proyección cráneo-caudal (CC) y una oblicua llamada vista o proyección medio-lateral oblicua (MLO). Existen otras técnicas que el radiólogo puede solicitar, pero son estas dos las más utilizadas. En la Figura 6 se ilustran las técnicas de exploración mencionadas y también la de vista lateral, que, aunque no es muy utilizada, vale la pena mencionarla [25].

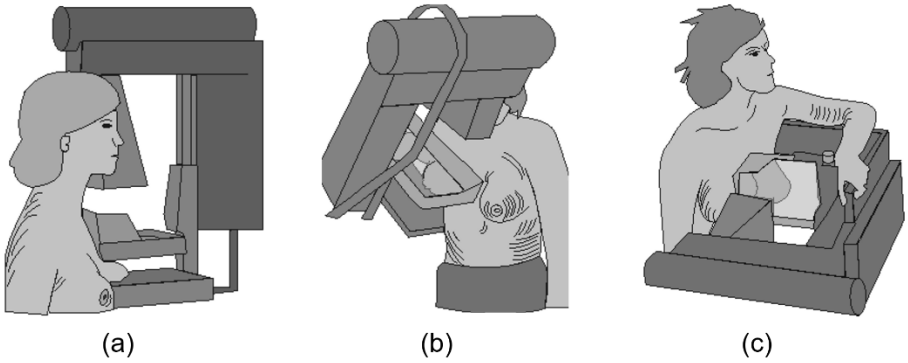


Figura 6. Proyecciones: (a) cráneo-caudal, (b) medio-lateral oblicua, (c) lateral.
[25]

3.1.5. Anormalidades en mamografías digitales

Existen diferentes tipos de anormalidades, benignas y malignas, que se pueden observar en las mamografías. Las principales anormalidades se muestran en la Figura 7 y se describen posteriormente.

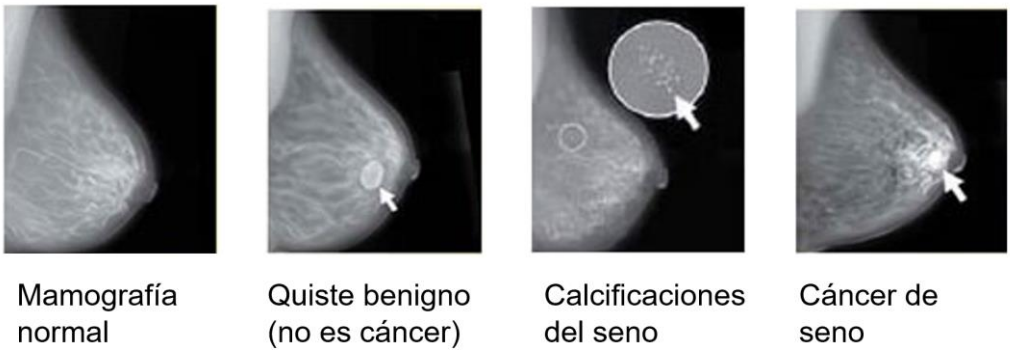


Figura 7. Algunas anormalidades en mamografías. [26]

3.1.6. Tumores, masas o bultos

Pueden aparecer tumores de distintas formas y tamaños. Los que tienen líquido en su interior, son lisos, redondos y tienen bordes bien definidos, estos no son cancerosos. Los que tienen forma irregular tienen mayor posibilidad de ser malignos [26].

3.1.7. Calcificaciones

Las calcificaciones grandes y redondas (macrocalcificaciones) son comunes en mujeres mayores de 50 años. Se ven como pequeños puntos blancos en una mamografía. Es poco probable que estén relacionados con el cáncer. Rara vez necesita más pruebas. Las microcalcificaciones más pequeñas son pequeños parches de calcio que se pueden ver en una mamografía. La mayoría de las veces, estos no son signos de cáncer. Sin embargo, estas áreas pueden necesitar más investigación dependiendo de cómo aparezcan en la mamografía.

3.1.8. Tejido denso de seno

Un seno denso tiene relativamente menos grasa y más tejido glandular y conectivo. Este resultado de una mamografía es tan común como normal, especialmente entre las mujeres jóvenes y mujeres que usan terapia hormonal para la menopausia. El tejido denso de seno puede dificultar más la interpretación de una mamografía porque tanto el tejido denso de seno como los tumores de seno aparecen como zonas sólidas blancas en la imagen [26].

3.2. Banco de datos (*dataset*)

Lo primero que se requiere es un banco de datos de imágenes digitales de rayos X de mama (mamografías digitales o mamografías digitalizadas) para preprocesarlas, entrenar los sistemas de clasificación propuestos y verificar su funcionamiento después de entrenados.

Lo ideal sería contar con un banco de datos que cuente con una gran cantidad de imágenes etiquetadas por expertos, es decir, con las anormalidades marcadas correctamente por radiólogos. Al no contar con un banco de datos con esas características suministrado por una clínica o radiólogo, se realiza la búsqueda de *dataset*s públicos disponibles en la Web.

En la Tabla 2, se presenta la comparación de algunos de los *dataset*s más utilizados en el campo de visión artificial [27].

Tabla 2. Comparación de algunos de los *datasets* de mamografías más utilizados.

Dataset	Dimensiones	Vista	Formato	Bits por pixel	Normales	Benignas	Malignas	Ventajas	Desventajas
DDSM	3.118x5.001	CC y MLO	LJPEG	12	914	870	695	Ampliamente usado.	Formato no estándar. Posición no precisa de las lesiones.
IRMA	Varias	CC y MLO	PNG	12	1.108	1.284	1.284	Posición precisa de las lesiones. Alta resolución.	Formato no estándar.
INbreast	Varias	CC y MLO	DICOM	16	67	220	49	Posición precisa de las lesiones. Formato estándar de archivo.	Tamaño limitado. Variaciones limitadas de las formas de las masas. *Dataset* antiguo.
mini-MIAS	1.024x1.024	MLO	PGM	8	209	67	54	Muy utilizado en la actualidad. Buena calidad con bajo peso. Permite manipulación con bajo costo computacional. Información detallada en metadatos suministrados.	Tamaño limitado. Imágenes de baja resolución. Solo vista MLO.

Para el desarrollo de este proyecto se eligió trabajar con el banco de datos mini-MIAS debido a que contiene imágenes que no son de gran peso, lo cual es de utilidad para los efectos de este proyecto porque en las redes neuronales convolucionales que se utilizan, incluso se reducen un poco más de dimensión.

Las imágenes que contiene este *dataset* provienen de la base de datos MIAS original (digitalizada con un borde de pixel de 50 micras), esta se ha reducido a un borde de pixel de 200 micras, se ha recortado y rellenado para que cada imagen tenga 1.024 × 1.024 pixeles. Cuenta con 322 archivos de imagen en formato PGM y con una tabla de metadatos que aportan información detallada de las mamografías, como el tipo de tejido dominante visible, las coordenadas de ubicación de las anormalidades presentes, en su mayoría muy precisas, etc. En la Figura 8, se muestran algunas de las imágenes presentes en el mini-MIAS *dataset.*

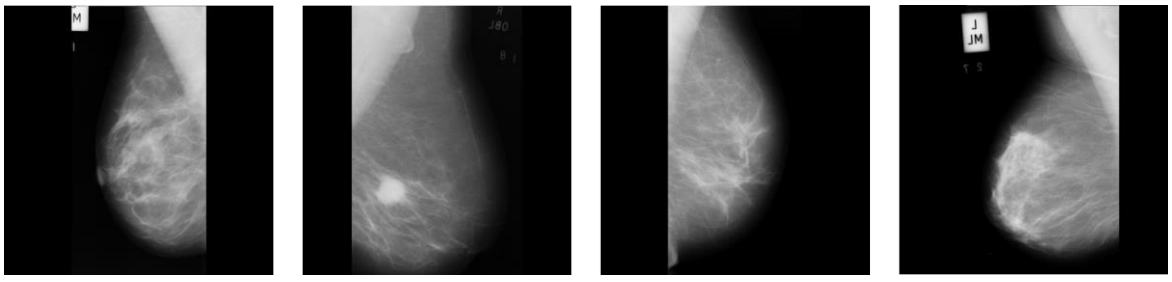


Figura 8. Ejemplos de mamografías presentes en el mini-MIAS *dataset*

Se observa que todas las mamografías presentes en el mini-MIAS *dataset* se encuentran en orientación vertical.

Nota importante: este *dataset* se tomó inicialmente de la Web www.kaggle.com, y durante el desarrollo de este proyecto, se encontró que la ubicación de algunas anormalidades estaba incorrecta en los metadatos. Por lo cual se sugiere a quien esté interesado en trabajar con este banco de datos, tomarlo de la página oficial http://peipa.essex.ac.uk/info/mias.html [24].

3.2.1. Información detallada del mini-MIAS *dataset* [24]

A continuación, se muestran las cinco primeras instancias de los metadatos suministrados en el mini-MIAS *dataset* para ilustrar su contenido:

```
mdb001 G CIRC B 535 425 197
mdb002 G CIRC B 522 280 69
mdb003 D NORM
mdb004 D NORM
mdb005 F CIRC B 477 133 30
```

La siguiente lista proporciona los detalles de los metadatos del mini-MIAS *dataset*:

- **Columna 1:** número de referencia de la base de datos MIAS.

- **Columna 2:** carácter del tejido de fondo:
  F: graso
  G: graso-glandular
  D: denso-glandular

- **Columna 3:** clase de anormalidad presente:
  CALC: calcificación
  CIRC: masas bien definidas/circunscritas

SPIC: masas espiculadas
MISC: otras masas mal definidas
ARCH: distorsión arquitectónica
ASYM: asimetría
NORM: normal

- **Columna 4:** gravedad de la anormalidad:
  B: benigna
  M: maligna

Columnas 5 y 6: coordenadas x, y del centro de la anormalidad.

Columna 7: radio aproximado en pixeles de uno o más círculos que encierran la anormalidad o las anormalidades.

La lista está organizada en pares de imágenes, donde las mamografías de la mama izquierda son los archivos con número impar y las de la mama derecha, los archivos con número par, de un mismo paciente.

El tamaño de todas las imágenes es de 1.024 pixeles x 1.024 pixeles. Las imágenes se han centrado en esa matriz de 1.024 x 1.024.

Cuando hay calcificaciones presentes, las ubicaciones de los centros y los radios se aplican a grupos en lugar de calcificaciones individuales. El origen del sistema de coordenadas es la esquina inferior izquierda.

En algunos casos, las calcificaciones se distribuyen ampliamente por toda la imagen en lugar de concentrarse en un solo sitio. En estos casos, las ubicaciones de los centros y los radios son inapropiados y se han omitido.

3.3. Preprocesamiento de datos (*Data Wrangling*)

En el ámbito de la inteligencia artificial, para referirse a la etapa de preprocesamiento de datos, comúnmente se utiliza el término “*Data Wrangling*”, que en español significa disputa o enfrentamiento de datos. Este consiste en limpiar, reorganizar y transformar los datos de un conjunto de datos "sin procesar" para hacerlos más útiles, que permitan aportar más información. El proceso exacto se modifica de un proyecto a otro según sus datos y lo que se está tratando de lograr.

En este proyecto, por ejemplo, a los metadatos del *dataset* seleccionado se les realiza reemplazo o llenado de valores vacíos o nulos, cambios de tipos de dato, agrupación y eliminación por condiciones establecidas, separación y visualización de variables según su tipo (variables categóricas, numéricas y sus subgrupos) y balanceo de datos [28].

3.4. Visión artificial

A continuación, se presentan los conceptos principales relacionados con visión artificial aplicados en el desarrollo de este proyecto.

3.4.1. Operaciones morfológicas

Las operaciones morfológicas son operaciones entre pixeles que permiten transformar la forma de los objetos presentes en imágenes a través de elementos estructurantes sin modificar el tamaño de la imagen original. Se opera el valor de los pixeles de interés con los más cercanos. Las operaciones morfológicas principales y más básicas son erosión (*erosion*) y dilatación (*dilation*).

La erosión busca eliminar pixeles de los bordes de los elementos de una imagen con el objetivo de remover pixeles ruidosos aislados, suavizar bordes del objeto o remover la capa exterior de pixeles del mismo, esto hace que el objeto se haga un

poco más pequeño. La dilatación, por el contrario, agrega pixeles para llenar huecos o espacios y también para el suavizado de bordes de objetos, esto expande levemente las formas.

En la erosión de una máscara binaria, el valor del pixel de salida se establece en 0 si alguno de los pixeles vecinos tiene el valor 0 y en la dilatación, el pixel de salida es 1 si alguno de los pixeles vecinos tiene el valor 1.

El elemento estructurante o de estructuración (SE por las siglas en inglés de *Structuring Element*) es una imagen secundaria, que podría considerarse como un filtro que recorrerá la imagen, este se utiliza para evaluar la imagen en busca de las propiedades de interés. La cantidad de pixeles modificados depende del tamaño y la forma de los elementos de estructuración seleccionados.

El elemento estructurante puede ser de cualquier forma o tamaño, representado digitalmente, con un origen definido, usualmente en el centro del elemento.
Las operaciones de erosión y dilatación se utilizan para implementar las operaciones de apertura (*opening*) y cierre (*closing*), respectivamente, que consisten en una combinación de las dos primeras.

La operación de apertura consiste en aplicar primero la operación de erosión, seguida de la operación de dilatación, se utiliza el mismo elemento estructurante para ambas operaciones. La operación de cierre consiste en aplicar primero la operación de dilatación y luego la operación de erosión, también utilizando el mismo elemento estructurante para ambas operaciones. Después de aplicadas estas operaciones, el tamaño del objeto se conserva [29].

Las operaciones morfológicas mencionadas, se definen de la siguiente manera [30]:

- **Erosión de N por M:** $N \ominus M = Y$
- **Dilatación de N por M:** $N \oplus M = Y$

- **Apertura de N por M:** $N \circ M = (N \ominus M) \oplus M$
- **Cierre de N por M:** $N \bullet M = (N \oplus M) \ominus M$

En la Figura 9 se observan los resultados de las diferentes operaciones morfológicas aplicadas a una imagen de entrada (*Original*).

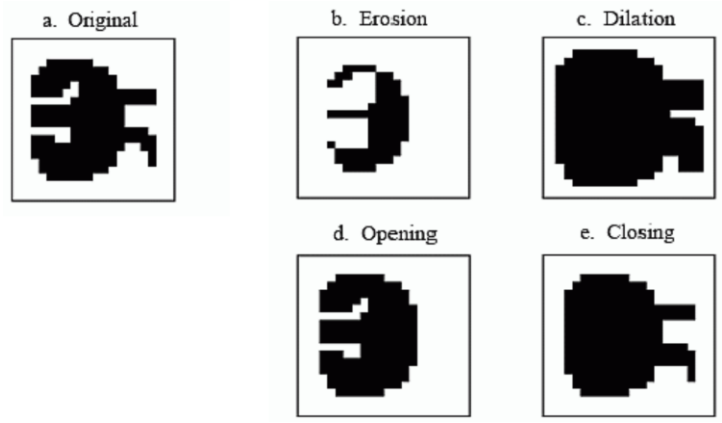


Figura 9. Ejemplo ilustrativo de las operaciones morfológicas. [30]

3.4.2. Envolvente convexa

"Se dice que un conjunto A es convexo si el segmento de línea recta que une dos puntos cualesquiera de A se encuentra completamente dentro de A. La envolvente convexa H de un conjunto arbitrario S es el conjunto convexo más pequeño que contiene a S" [29] (Figura 10).

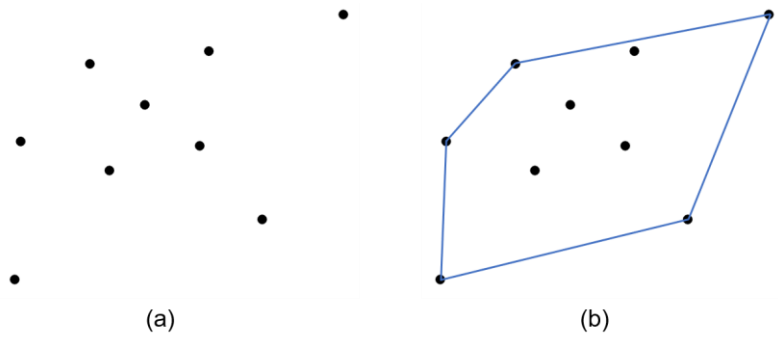


Figura 10. (a) conjunto de puntos *S*, (b) envolvente convexa de *S.*

3.4.3. Entropía

La entropía es una función de las frecuencias de los valores de un conjunto dado. Mide la variación o desorden de los datos, por ejemplo, si las variables de un conjunto de datos tienen el mismo valor, entonces la entropía es cero. Por el contrario, si la variable asume todos los valores con casi la misma frecuencia en el conjunto, entonces la entropía es máxima. La entropía de Shannon es la más popular.

$$Entropía = -\sum_{i=1}^{d} p_i \log p_i$$

Donde p_i es la probabilidad del valor i y d es el tamaño del dominio de la variable considerada [28].

3.4.4. Ecualización del histograma

El histograma de una imagen consiste en la estimación de la distribución de probabilidad de los valores de intensidad de los pixeles presentes en la misma. En una imagen en escala de gris, la probabilidad de ocurrencia de un pixel con valor de gris $\mathrm{r_k}$ se define de la siguiente manera:

$$\mathrm{p(r_k)} = \frac{\mathrm{n_k}}{\mathrm{n}}, \qquad 0 \leq \mathrm{r_k} \leq 1, \qquad \mathrm{k} = 0,1,2,\dots,\mathrm{L}-1$$

Donde L representa el número niveles de gris, $\mathrm{n_k}$ la cantidad de pixeles con valor de gris k y n, la cantidad total de pixeles en la imagen.

Se especifica una cantidad de contenedores (*bins*), cada *bin* equivale a la probabilidad del número de pixeles dentro del rango establecido para el *bin* correspondiente, esto depende del número de *bins* en el que se divida el

histograma. Se cuenta con un histograma para cada canal, en imágenes en escala de grises, se tiene un solo histograma, como se ilustra en la Figura 11.

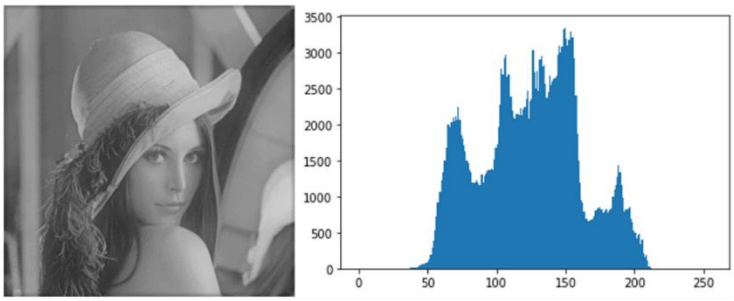


Figura 11. (a) Imagen original, (b) Histograma correspondiente.

El histograma está relacionado directamente con el contraste de la imagen, cuando los valores posibles de intensidad de pixeles se encuentran acumulados en un rango de valores cercano, se observa un bajo contraste, para mejorarlo, se implementa la ecualización del histograma, que consiste en distribuir los valores de intensidad de pixeles de una imagen de una manera más uniforme y menos estrecha dentro del histograma para mejorar su contraste global. Se obtiene a partir de la función acumulativa de la intensidad o histograma acumulado, es decir, de la suma de los valores del histograma, como se muestra a continuación [29]:

$$s_k = T(r_k) = \sum_{j=0}^{k} p_r(r_j) = \sum_{j=0}^{k} \frac{n_j}{n}$$

En la Figura 12, se observa la imagen ilustrativa con mejor contraste después de aplicar la ecualización del histograma global:

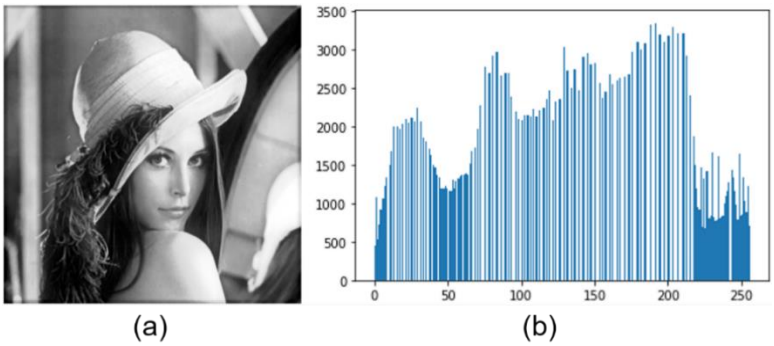


Figura 12. (a) Imagen con histograma ecualizado, (b) Histograma correspondiente.

3.4.5. Ecualización del histograma adaptativo limitado por contraste (CLAHE)

En ocasiones la ecualización del histograma convencional presenta un problema de saturación de brillo en la imagen después de aplicarlo, esto se debido a que no se aplica a una región específica de la imagen, sino globalmente (Figura 12). La ecualización del histograma adaptativo por contraste resuelve este problema al tomar sectores o “mosaicos” de la imagen (por ejemplo, de 8 x 8) y a estos les aplica la ecualización del histograma convencional, mencionada en el apartado anterior.

El histograma se limita a cada mosaico, con el inconveniente de que, si hay ruido presente, lo amplificará. Para evitar esto, se aplica la limitación de contraste.

Si algún *bin* de histograma está por encima del límite de contraste especificado, esos pixeles se recortan y distribuyen uniformemente a otros *bins* antes de aplicar la ecualización de histograma. Después de la ecualización, se aplica interpolación bilineal para eliminar los artefactos ruidosos. El resultado, se muestra a continuación (Figura 13).

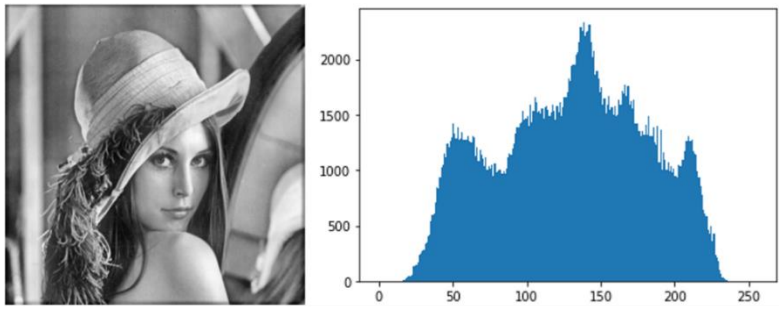


Figura 13. (a) Imagen con CLAHE, (b) Histograma correspondiente.

3.4.6. Matriz de coocurrencia de niveles de gris (GLCM)

La matriz de coocurrencia de niveles de gris (GLCM) fue propuesta por Robert M. Haralick en la década de 1970 [31] para obtener información de textura de imágenes en escala de gris. Es un método de análisis estadístico de segundo orden.

Como su nombre lo indica, la matriz de coocurrencia de niveles de gris cuantifica la coocurrencia simultánea de determinados valores de gris en determinadas orientaciones. Concretamente, describe la frecuencia de aparición de un par de valores de gris en una orientación θ determinada (0°, 45°, 90° o 135°) [32].

Para obtener la matriz de coocurrencias, se selecciona la orientación (ángulo θ) a la que se quiere hacer el análisis, se considera cada combinación posible de nivel de gris. Se cuenta la frecuencia a la que aparece la combinación determinada en la orientación seleccionada. Es decir, para un desplazamiento θ dado, la matriz de coocurrencia de nivel de gris $P_{i,j}$ describe la frecuencia de aparición del nivel de gris $i(ésimo)$ y $j(ésimo)$ en la orientación θ.

3.4.7. Características extraídas de la matriz de coocurrencia de niveles de gris

La GLCM de una imagen o región determinada no da como resultado una única medida, se obtiene una matriz cuadrada, cuyo tamaño depende del número de niveles de gris de la imagen. Para obtener información específica de la matriz, se

suelen calcular ciertas medidas que son las que constituyen los descriptores de coocurrencias.

Algunos de estos descriptores se expresan a continuación:

- **Contraste:** $\sum_{i,j=0}^{L-1} P_{i,j}(i-j)^2$
- **Disimilitud:** $\sum_{i,j=0}^{L-1} P_{i,j}|i-j|$
- **Homogeneidad:** $\sum_{i,j=0}^{L-1} \frac{Pi,j}{1+(i-j)^2}$
- **Segundo momento angular (ASM):** $\sum_{i,j=0}^{L-1} P_{i,j}^2$
- **Energía:** $\sqrt{ASM}$
- **Correlación:**

$$\sum_{i,j=0}^{L-1} P_{i,j}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right]$$

La matriz de entrada es el histograma de coocurrencia de niveles de gris para el cual se calcula la propiedad especificada. El valor escrito como $P_{i,j}$ en realidad tiene en cuenta a $P[i,j,d,\ \theta]$, que es el número de veces que el nivel de gris j aparece a una distancia d y en un ángulo θ desde el nivel de gris i [33].

3.4.8. Patrones binarios locales (LBP)

Los patrones binarios locales se utilizan para detectar información de textura de una imagen. Este filtra los pixeles y obtiene un valor binario representativo, que usualmente se da en decimal.

Se comparan los pixeles o puntos que rodean un punto central. Si los puntos circundantes son mayores o menores que el punto central, se obtiene un resultado binario.

Se hace de la siguiente manera:

- Si valor de pixel vecino >= valor de pixel central, se agrega un 1 al número binario descriptor LBP.

- Si valor de pixel vecino < valor de pixel central, se agrega un 0 al número binario descriptor LBP.

- Después, el número binario resultante, se le asigna al pixel central

El histograma del resultado LBP también es una buena medida para clasificar texturas. [34], [35]

En la Figura 14 se observan ejemplos de texturas dependiendo los patrones binarios locales (de 8x8). Se dice que cuando los pixeles circundantes son todos negros o todos blancos, la región de la imagen es plana. Los grupos de pixeles continuos en blanco o negro se consideran patrones que pueden interpretarse como esquinas o bordes. Si los pixeles alternan entre pixeles blancos y negros, el patrón se considera "no uniforme":

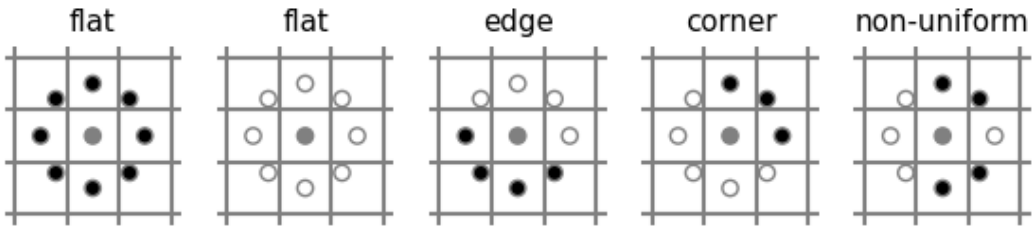


Figura 14. Ejemplos de texturas con LBP [35]

En la Figura 15 se presenta el LBP (negativo) de una imagen mamográfica segmentada del conjunto de datos utilizado en el presente proyecto:

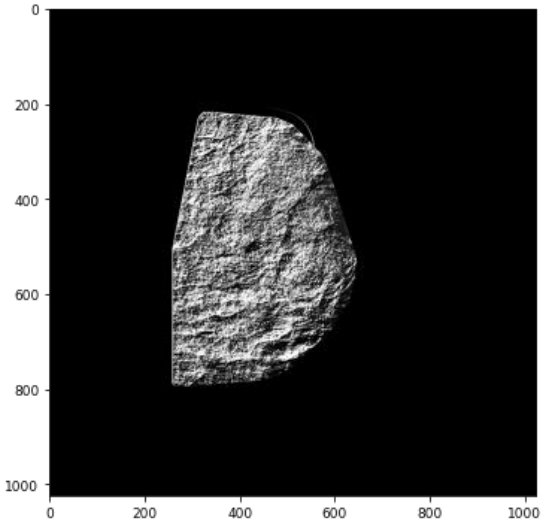


Figura 15. Negativo de imagen LBP de mamografía segmentada.

3.4.9. Índice de similitud de Jaccard

El índice de Jaccard o coeficiente de Jaccard cuantifica el nivel de similitud entre dos conjuntos de datos, consiste en calcular la intersección sobre la unión de ambos conjuntos [28]. A continuación, se muestra la expresión para calcularlo:

.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

3.5. Máquinas de aprendizaje superficial

3.5.1. Redes neuronales artificiales

Las redes neuronales artificiales se basan en el funcionamiento de los sistemas neuronales biológicos. En una red neuronal artificial, las salidas están determinadas normalmente por una combinación lineal de las entradas, su valor final se ve modificado por una función de activación específica.

De este modo, cada entrada se multiplica por un peso correspondiente, el cual se ajusta durante el entrenamiento de la red a través del cálculo del gradiente

descendente o alguna de sus variaciones para determinar el mínimo de la función de pérdida (*backpropagation* [36]).

Los resultados de dichas multiplicaciones se suman, en conjunto con un término de desplazamiento (*bias*), que permite lograr un desplazamiento en la salida de la red, para generar la entrada a la función de activación.

Este modelo se expresa de la siguiente manera:

$$n_{out} = f_{act}(w_1 * i_1 + w_2 * i_2 + \cdots + w_b * b)$$

n_{out} equivale a la salida de una neurona, f_{act} es la función de activación e i_k, w_k representan una entrada a la neurona con su respectivo peso, b y w_b la entrada del término de desplazamiento y el peso asignado, respectivamente.

La función de activación se utiliza para determinar si la neurona debe dispararse o no. Existen diferentes funciones de activación tales como: sigmoide, lineal, ReLU, etc., que se utilizan dependiendo la necesidad del problema abordado.

El tipo de red neuronal artificial más utilizado es un perceptrón multicapa completamente conectado, donde cada neurona de una capa determinada toma como entradas las salidas de todas las neuronas de la capa anterior, como se muestra a continuación, en la Figura 16 [36].

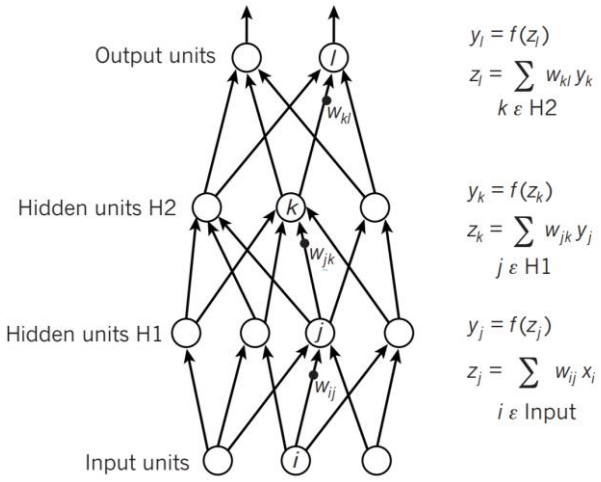


Figura 16. Estructura de un perceptrón multicapa.
[36]

3.5.2. Máquinas de soporte vectorial

Una máquina de soporte vectorial o máquina de vectores de soporte (SVM) es un clasificador que separa dos clases mediante un hiperplano. Este hiperplano puede ser la representación la N-dimensional de una línea. A partir de un problema de clasificación binaria con los datos de entrenamiento etiquetados, la máquina de soporte vectorial encuentra el hiperplano óptimo que separa los datos de entrenamiento entre las dos clases. Esto puede extenderse a un problema con más clases.

A manera de ejemplo, en la Figura 17 se ilustra un problema bidimensional, es decir, con dos parámetros de entrada (x, y) y se representan dos clases. Esto para facilidad de comprensión y visualización. Se debe tener en cuenta que en la mayoría de casos se aplican hiperplanos de mayor dimensionalidad para separar clases con muchos parámetros de entrada y estas no se pueden graficar.

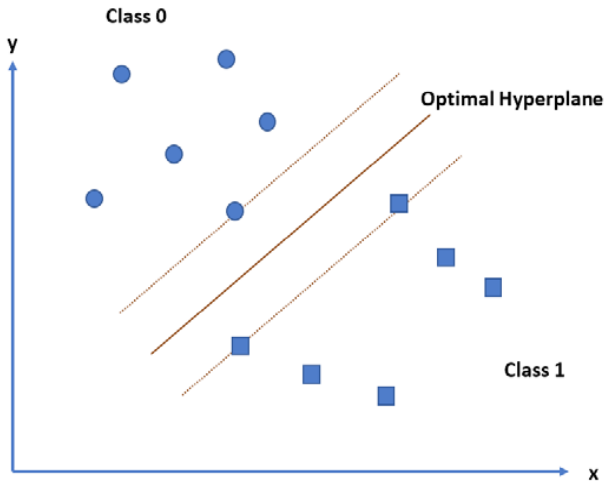


Figura 17. Separación de clases con un hiperplano. [37]

Se pueden observar las dos clases (puntos y cuadros) separadas por una línea sólida que representa el mejor hiperplano para separar las dos clases. Podrían dibujarse infinitas líneas diferentes para separar las dos clases, pero esta es la óptima porque maximiza la distancia de cada punto a la línea de separación. Los puntos sobre las líneas punteadas se denominan vectores de soporte. La distancia entre las dos líneas punteadas se denomina margen máximo [28], [37].

3.6. Máquinas de aprendizaje profundo

3.6.1. Redes neuronales convolucionales

Las redes neuronales convolucionales (CNNs) se basan en el funcionamiento de la corteza visual del cerebro, son un tipo de redes neuronales artificiales profundas muy utilizadas en la actualidad en el ámbito de la visión artificial.

Las capas convolucionales consisten en neuronas, al igual que en una red neuronal clásica, pero los pesos están controlados por conexiones en conjuntos llamados filtros, que son responsables de detectar diferentes características de entrada.

Cada capa convolucional realiza una operación de convolución entre la entrada de esa capa y un conjunto de filtros definidos para ella, produciendo resultados para cada filtro (mapas de características); estos mapas son una representación de la entrada que resalta su propiedad e intenta extraerla con el filtro apropiado.

A las salidas de cada capa de convolución se le realiza una operación de "submuestreo" (*pooling*) para obtener un mapa más pequeño que contiene solo la información más importante para pasar a la siguiente capa. Una capa totalmente conectada conserva la estructura básica de un perceptrón multicapa. La entrada al conjunto de capas completamente conectado es la muestra de salida de la última capa convolucional.

Esto permite filtrar continuamente los parámetros de entrada según su relevancia para la decisión hasta llegar a la capa de salida, que devuelve el valor de probabilidad de que la entrada pertenezca a cada clase distinguida por la función. CNN. El marco anterior puede tomar una entrada inicial (como una imagen) y usar capas convolucionales para extraer diferentes tipos de características y luego enviarlas a una red neuronal (última capa) completamente conectada (densa), responsable de evaluar la posible correspondencia de la entrada [36]. En la Figura 18 se ilustra la estructura de la arquitectura de red neuronal convolucional LeNet-5.

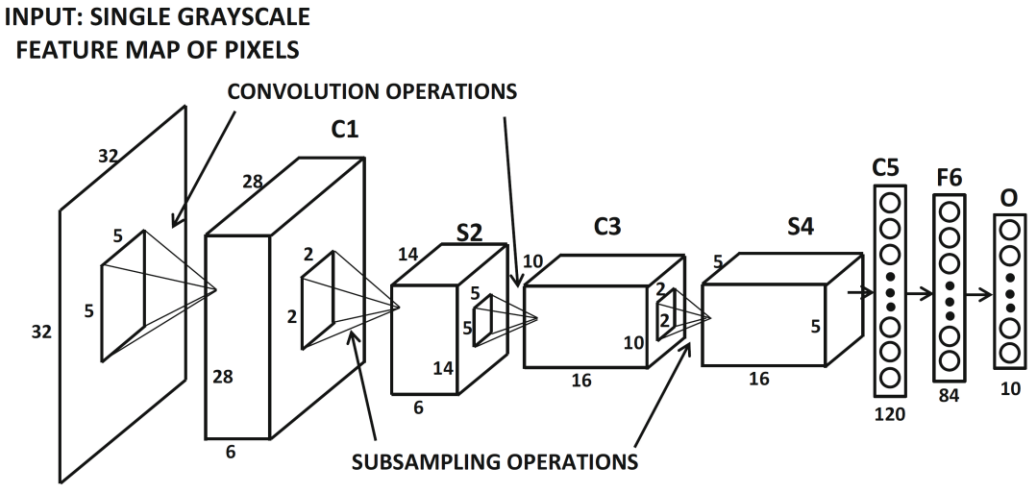


Figura 18. LeNet-5: Una de las primeras redes neuronales convolucionales [38].

3.6.2. Transferencia de aprendizaje

El aprendizaje por transferencia es una técnica utilizada para reducir las exigencias computacionales y de tiempo necesarias para entrenar una máquina de aprendizaje. Esto implica tomar características similares de un modelo ya entrenado para resolver el problema que se está atacando y adaptarlo al problema en cuestión de manera efectiva, aprovechando así el conocimiento previo del modelo y reduciendo la cantidad de datos de entrenamiento necesarios para lograr buenos resultados. Para hacer esto, algunas capas de procesamiento (generalmente las últimas capas) se modifican y vuelven a entrenar, y algunas capas cambian sus pesos; este reentrenamiento requiere menos conjuntos de datos de los que se necesitarían para entrenar un modelo desde cero y lograr resultados similares.

3.6.3. Arquitectura DenseNet121

DenseNet121 pertenece a una familia de arquitecturas de red conocidas como DenseNet [39]. En las que cada capa obtiene entradas adicionales de todas las capas anteriores y pasa sus propios mapas de características a todas las capas posteriores. Es decir, cada capa recibe el "conocimiento" de todas las capas anteriores.

Mientras que las redes convolucionales tradicionales con L capas tienen L conexiones (una entre cada capa y su capa subsiguiente). DenseNet tiene $\frac{L(L+1)}{2}$ conexiones directas.

En la Figura 19, se ilustra la arquitectura DenseNet121. Se puede observar que, para cada capa, los mapas de características de todas las capas anteriores se utilizan como entradas, y sus propios mapas de características se utilizan como entradas en todas las capas posteriores.

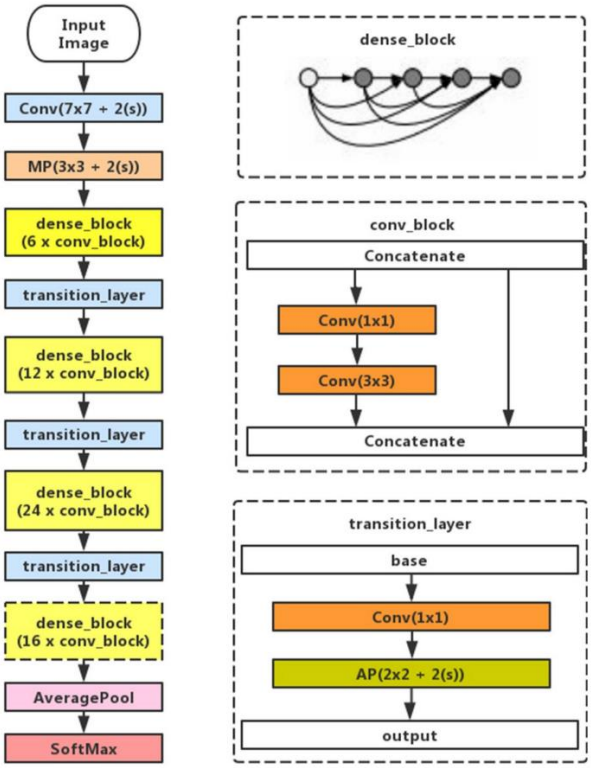


Figura 19. Arquitectura DenseNet121 [40].

Las capas entre dos bloques adyacentes se denominan capas de transición y cambian tamaños de mapas de características a través de convolución y agrupación [39].

La familia de DenseNet tiene varias ventajas: alivian el problema de la fuga del gradiente, fortalecen la propagación de funciones, fomentan la reutilización de funciones y reducen sustancialmente la cantidad de parámetros.

3.6.4. Arquitectura U-Net

U-Net es una arquitectura de red profunda netamente convolucional, que permite realizar segmentación semántica en imágenes, es decir, asignar una clase o etiqueta a cada pixel presente en la imagen, es muy utilizada en segmentación de imágenes médicas.

Una de las ventajas principales de la arquitectura U-Net es que, aunque para el entrenamiento exitoso de redes profundas se requieren grandes cantidades de imágenes para su entrenamiento, incluso millones de muestras etiquetadas correctamente, esta se basa en el uso de aumento de datos para usar las muestras disponibles, así no sean muchas, de manera más eficiente.

La arquitectura consta de una ruta de contracción (*contracting path*) para capturar el contexto de la imagen de entrada y una ruta de expansión (*expanding path*) simétrica que permite una localización precisa (ver Figura 20).

Se puede entrenar de extremo a extremo a partir de muy pocas imágenes. Esta superó al mejor método anterior (una red neuronal convolucional de ventana deslizante).

Además, es rápida. La segmentación de una imagen de 512 x 512 toma menos de un segundo en una GPU moderna. [41]

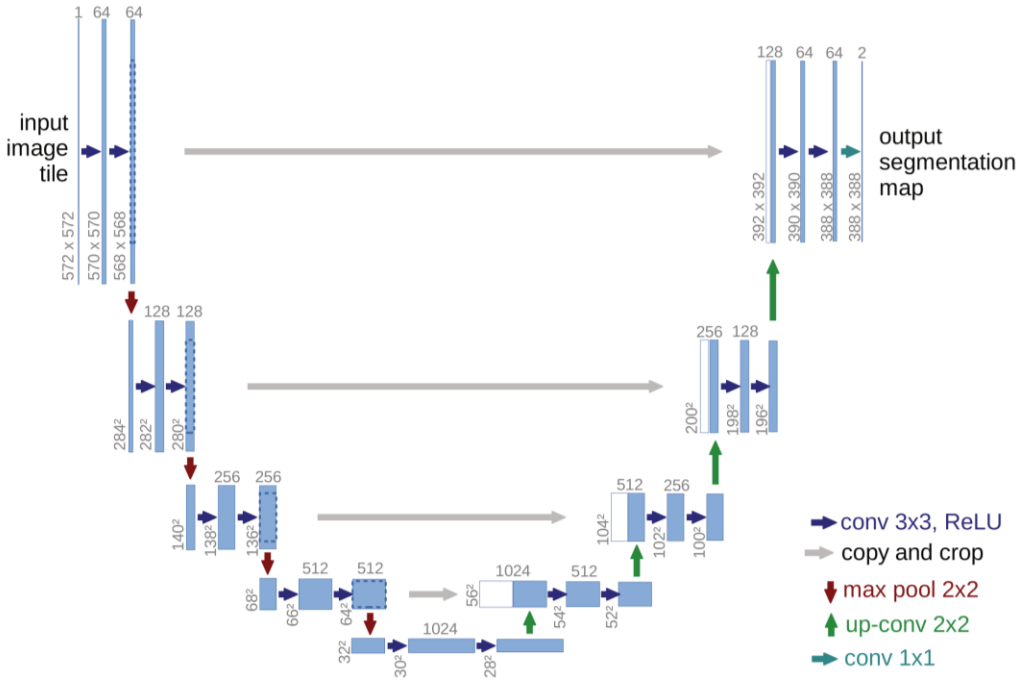


Figura 20. Arquitectura U-Net [41].

3.7. Medición de desempeño de los clasificadores

3.7.1. Métricas

En este apartado se detallan algunas métricas para la evaluación de los sistemas de clasificación desarrollados, para esto se deben conocer los siguientes términos [42]:

- **Verdaderos positivos (VP o TP):** clasificación de entradas como pertenecientes a una clase, cuando realmente pertenecen a esa clase.
- **Falsos positivos (FP):** clasificación de entradas como pertenecientes a una clase, cuando no pertenecen a esa clase.
- **Verdaderos negativos (VN o TN):** clasificación entradas como no pertenecientes a una clase, cuando realmente no pertenecen a esa clase.
- **Falsos negativos (FN):** clasificación de entradas como no pertenecientes a una clase, cuando sí pertenecen a la misma.
- **Matriz de confusión:** consiste en una matriz o tabla que permite visualizar fácilmente el número de verdaderos positivos, falsos positivos, falsos negativos y verdaderos negativos del modelo, para evaluar su rendimiento [42]. En la Figura 21, que se expone a continuación, se ilustra una matriz de confusión con la disposición de filas y columnas utilizada en este proyecto:

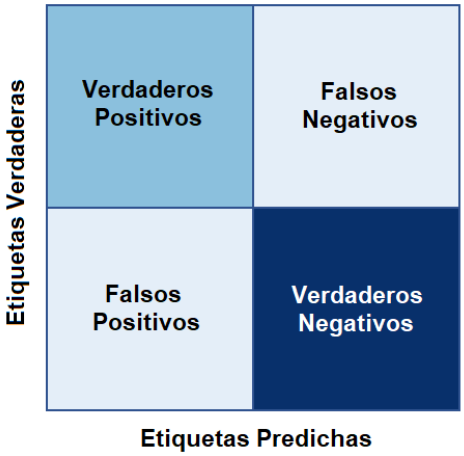


Figura 21. Matriz de confusión y disposición utilizada.

Las métricas de desempeño basadas utilizadas son las siguientes [42]:

- **Exactitud:** *accuracy* en inglés, es la tasa de predicciones correctas respecto al total de observaciones del conjunto de *test*.

$$Exactitud = \frac{VP + VN}{VP + FP + VN + FN}$$

- **Precisión:** *precision* en inglés, es el porcentaje de verdaderos positivos respecto al total de positivos, entre falsos y verdaderos.

$$Precisión = \frac{VP}{VP + FP}$$

- **Exhaustividad o sensibilidad:** *recall* en inglés, mide la cantidad de etiquetas que el modelo es capaz de identificar correctamente. Se evalúan los verdaderos positivos respecto a todos los que realmente debieron ser positivos.

$$Exhaustividad = \frac{VP}{VP + FN}$$

- **Valor F1:** combina las medidas de precisión y exhaustividad en una sola métrica para indicar el desempeño general del sistema respecto a las predicciones positivas. Se halla calculando la media armónica entre la precisión y la exhaustividad.

$$F1 = 2 * \frac{Precisión * Exhaustividad}{Precisión + Exhaustividad}$$

- **Área bajo la curva característica operativa del receptor (AUC):** consiste en el área bajo la curva característica operativa del receptor, ROC por sus siglas en inglés. Esta curva brinda el valor de la exhaustividad (tasa de

verdaderos positivos) contra la tasa de falsos positivos, que equivale a:

$$TFP = \frac{FP}{FP + VN}$$

El AUC es calculada entre las coordenadas (0,0) y (1,1). Un AUC igual a 0 indica que todas las predicciones del modelo son incorrectas, mientras que un AUC igual a 1 indica que todas las predicciones fueron correctas. Se puede definir como un resumen de la calidad del modelo [43].

- **Función de pérdida de entropía cruzada**

La pérdida de entropía cruzada se utiliza para medir el rendimiento de un modelo de clasificación. La pérdida, costo o error, se mide como una probabilidad entre 0 y 1, siendo 0 el caso ideal, sin errores de clasificación.

Para clasificación binaria, la entropía cruzada se calcula de la siguiente manera (entropía cruzada binaria):

$$Pérdida = -\frac{1}{n}\sum_{i=1}^{n} y_i * \log \hat{y}_i + (1 - y_i) * \log (1 - \hat{y}_i)$$

En clasificación multiclase, se calcula una pérdida separada para cada etiqueta de clase y se suman los resultados (entropía cruzada categórica):

$$Pérdida = -\sum_{i=1}^{n} y_i * \log \hat{y}_i$$

n equivale al número de clases, y al valor verdadero (etiqueta) y $\hat{y}_i$ a la probabilidad predicha de pertenencia a la clase.

4. Solución

En este capítulo se muestra la solución de software implementada. Esta se desarrolla principalmente en Python 3.7.13 [44] en el entorno de ejecución online Google Colab Pro [45], acelerado por una GPU NVIDIA Tesla P100-PCIE-16GB. Se desarrollan las funciones propias necesarias y se utilizan librerías que permiten implementar los conceptos descritos en el marco teórico, principalmente las siguientes:

- *Pandas*, para gestión de estructuras de datos [46].
- *NumPy*, para manejo de arreglos y matrices [47].
- *Matplotlib* y *Seaborn* para visualización de datos [48], [49].
- *OpenCV*, para procesamiento de imágenes [50].
- *Scikit-learn*, para aplicar técnicas de *Machine Learning*, como redes neuronales artificiales y máquinas de soporte vectorial, entre muchas otras [51].
- *TensorFlow y Keras*, para la implementación de redes neuronales convolucionales y *Deep Learning* [52], [53].

Una de las técnicas de segmentación se implementa en *MATLAB R2019a* [54].

A continuación, se realiza la descripción detallada de la solución desarrollada.

4.1. Introducción

Con base en los objetivos, se implementaron dos técnicas distintas de segmentación de imágenes para segmentar la región de interés, el seno. Posteriormente, se utilizaron tres técnicas distintas de reconocimiento de patrones para la clasificación y detección de anormalidades en las mamografías del *dataset* seleccionado. A continuación, se presentan los diagramas de bloques generales de las distintas etapas de las soluciones implementadas.

4.1.1. Descripción general del preprocesamiento de metadatos y los métodos de segmentación de la mama

Como se menciona en el marco teórico, el Banco de datos (*dataset*) seleccionado fue el mini-MIAS *dataset*.

En la Figura 22, se muestra que la primera etapa del desarrollo de la solución inicia con el preprocesamiento de los metadatos del *dataset* (*Data Wrangling*) para comprenderlo y seleccionar los datos definitivos, en este caso, imágenes con las que se entrenan los sistemas de clasificación subsiguientes.

Una vez se cuenta con el banco de datos construido, se procede a aplicar dos técnicas de segmentación distintas a las mamografías originales para segmentar el área de interés, en este caso, la mama, sin tejido pectoral o artefactos extraños que no pertenezcan a ella. Se implementa el método 1 de segmentación de la mama: a partir de operaciones morfológicas, entropía y otras, y el método 2 de segmentación de la mama: con la arquitectura de red neuronal convolucional U-Net.

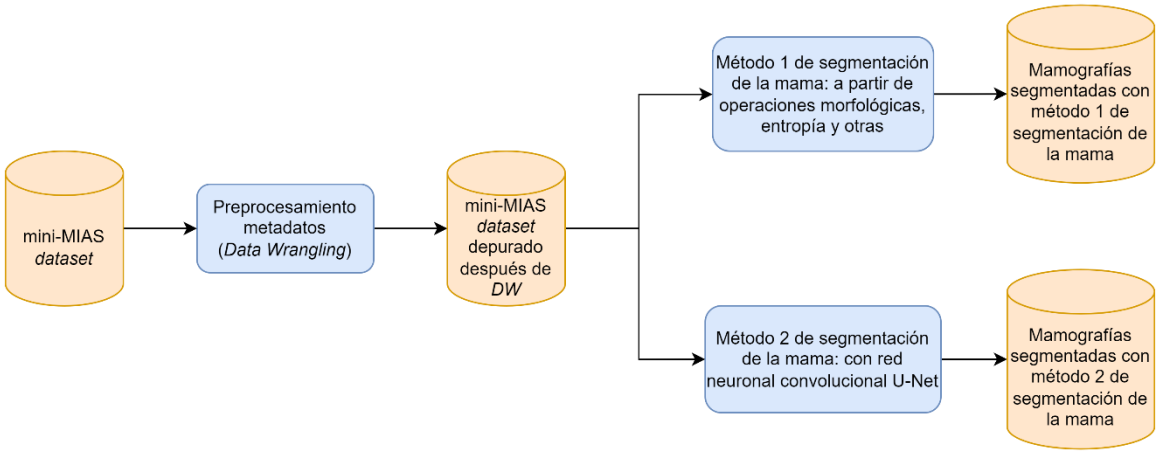


Figura 22. Diagrama de bloques métodos de segmentación de la mama.

4.1.2. Descripción general de método de clasificación y detección de anormalidades 1: con redes neuronales convolucionales DenseNet121 + U-Net

Esta solución consiste en la implementación de máquinas de aprendizaje profundas, que infieren las características discriminantes por sí mismas. Se prueba con las imágenes resultado de ambos procesos de segmentación. Se ajusta y se realiza transferencia de aprendizaje con la arquitectura DenseNet121 para clasificar las mamografías como normales o anormales. Posteriormente, se detecta el tejido anormal (de las mamografías marcadas como anormales) utilizando la arquitectura de red neuronal convolucional U-Net, esta se entrena con etiquetas que son máscaras binarias con las regiones anormales, creadas a partir de la información de ubicación y radio de la región presentes en los metadatos del *mias_dataframe.* Para entrenamiento y pruebas (*test*) de este modelo, se utilizan las imágenes anormales (ya detectadas con DenseNet121) originales del mini-MIAS *dataset*, solo con el histograma ecualizado adaptativamente (ver Figura 23).

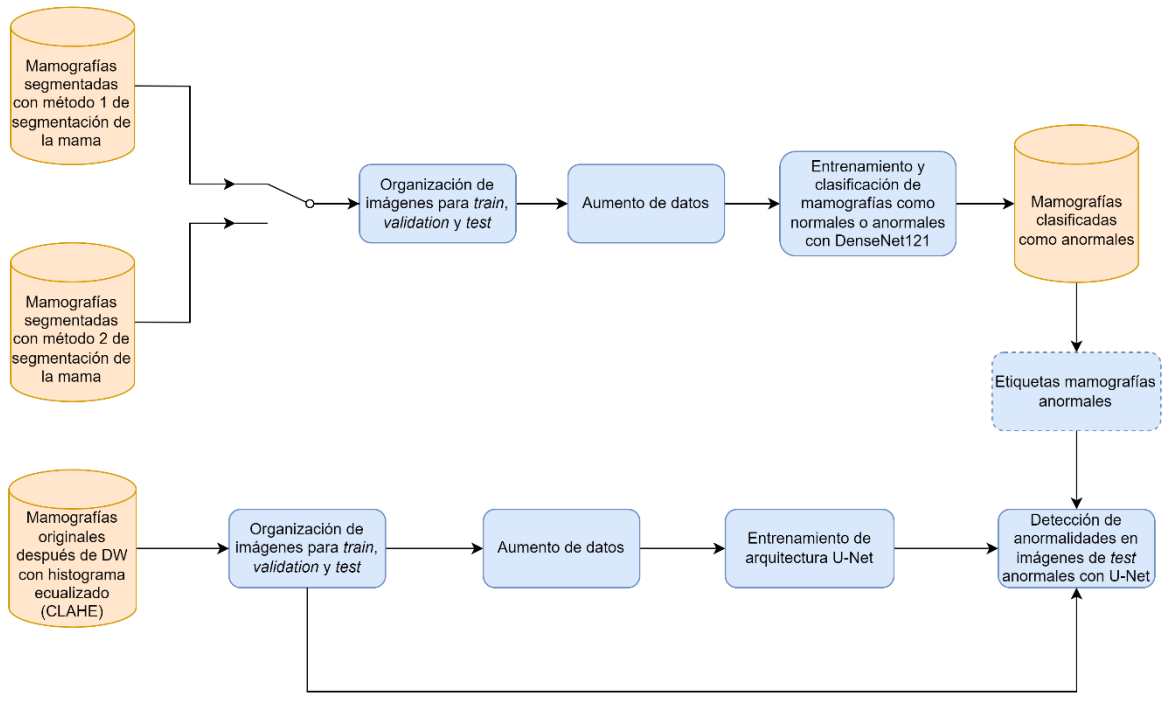


Figura 23. Diagrama de bloques de método de clasificación y detección de anormalidades 1: con redes neuronales convolucionales DenseNet121 + U-Net.

4.1.3. Descripción general de método de clasificación y detección de anormalidades 2: a partir de extracción de características, ventana deslizante, MLP y SVM.

Este método consiste en la extracción de características estadísticas del histograma y de textura que permitan clasificar el tejido de la mama con una red neuronal artificial, específicamente un perceptrón multicapa (MLP) y una máquina de soporte vectorial (SVM). Estos modelos se entrenan con las características mencionadas obtenidas de las regiones o *patches* de las anormalidades extraídas con la información de ubicación y radio de las mismas, también con regiones de tejido normal extraídas aleatoriamente de las mamografías normales. Posteriormente, se evalúa cuál de los dos clasificadores tuvo mejor rendimiento para clasificar las regiones o *patches* de las imágenes de *test* implementando un algoritmo de ventana deslizante (ver Figura 24).

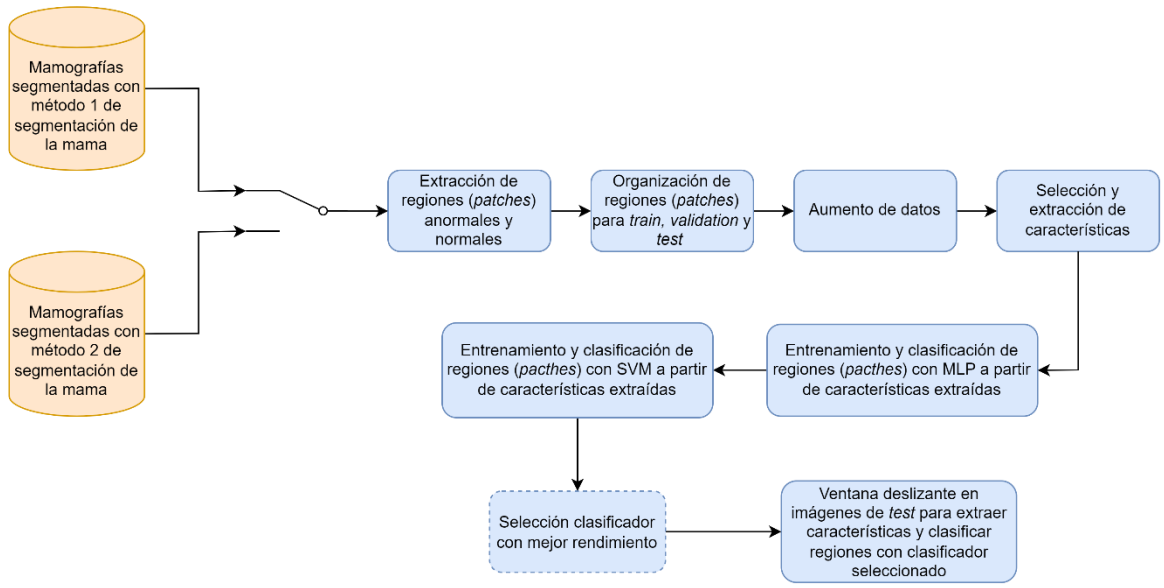


Figura 24. Diagrama de bloques método de clasificación y detección de anormalidades 2: a partir de extracción de características, ventana deslizante, MLP y SVM.

Finalmente, se prueba el funcionamiento de los métodos utilizados y se analizan los resultados.

Cada una de las etapas desarrolladas, se detallan a continuación.

4.2. Preprocesamiento

En este capítulo se muestra el proceso de preprocesamiento de los metadatos del mini-MIAS *dataset* y los dos métodos de segmentación del seno implementados.

4.2.1. Preprocesamiento de metadatos del *dataset* (*Data Wrangling*)

Inicialmente, se copian los metadatos de la página oficial del mini-MIAS *dataset* y se guardan en un documento de texto (.TXT). Este se convierte en un *dataframe* de Pandas para facilitar su manipulación de aquí en adelante. Para fácil entendimiento, en este documento se denomina *mias_dataframe* y para mejor visualización, en la Tabla 3 se muestran solo las 3 primeras instancias del mismo.

Tabla 3. Instancias del *mias_dataframe* (330, 7)

	REFNUM	BG	CLASS	SEVERITY	X	Y	RADIUS
0	mdb001	G	CIRC	B	535.0	425.0	197.0
1	mdb002	G	CIRC	B	522.0	280.0	69.0
2	mdb003	D	NORM	NaN	NaN	NaN	NaN

Posteriormente, se grafican los datos de las variables categóricas nominales presentes en el *mias_dataframe* para comenzar a interpretarlo (Figura 25).

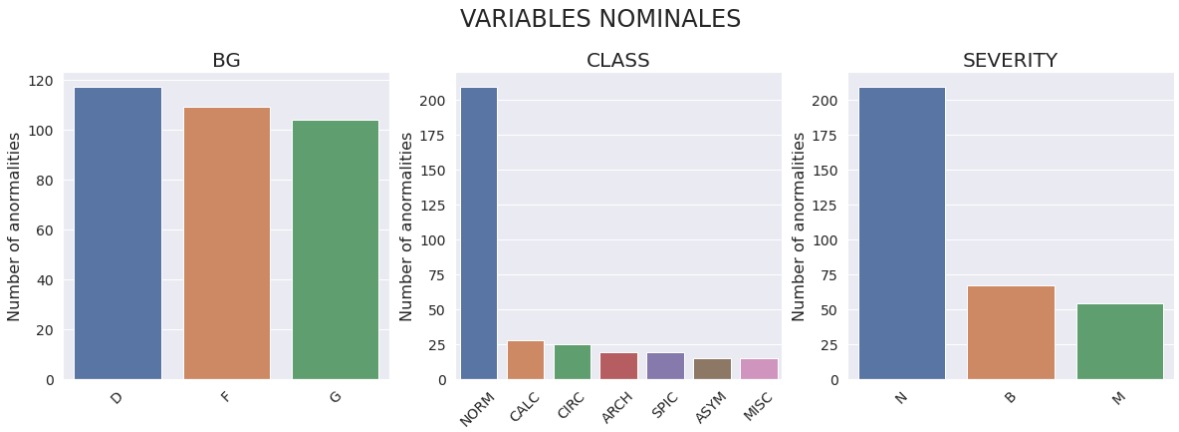


Figura 25. Variables categóricas nominales del *mias_dataframe.*

En el gráfico de barras BG de la Figura 25, se observa la cantidad de instancias presentes en el *mias_dataframe* según la característica dominante del tejido de fondo (BACKGROUND), *Dense*, *Fatty* o *Glandular* (ver 3.2.1).

En el gráfico CLASS, se contempla la cantidad de mamografías normales y el número de tipos de anomalías presentes en las mamografías anormales (CLASS) (ver 3.2.1).

El gráfico de SEVERITY muestra la cantidad de mamografías normales (N) y la cantidad de anormalidades benignas (B) y malignas (M) presentes en las mamografías anormales (SEVERITY).

Algunas mamografías etiquetadas como anormales presentan más de una anomalía. Por esta razón, en las afirmaciones anteriores se mencionan cantidades de anormalidades y no de imágenes (presentes en la carpeta all-mias). En el caso de las mamografías normales, las etiquetas coinciden con el número de imágenes.

Las mamografías cuyo nombre (REFNUM) se repite en el *mias_dataframe*, se detectan y se muestran a continuación, en la Tabla 4:

Tabla 4. Referencias repetidas en el *mias_dataframe.*

REFNUM	Número de apariciones	SEVERITY
mdb005	2	B
mdb132	2	B
mdb144	2	B y M
mdb223	2	B
mdb226	3	B
mdb239	2	M
mdb249	2	M

En la Tabla 4 se observa que hay 8 referencias repetidas, pues la imagen mdb144 tiene anomalías pertenecientes a las dos clases (*benign* y *malignant*).

Teniendo en cuenta que el número de etiquetas difiere del número de imágenes, en la Tabla 5 se sintetizan estas cantidades.

Tabla 5. Cantidad de etiquetas vs. cantidad de imágenes.

Etiquetas anormalidades benignas: **67**	Imágenes con anormalidades benignas: **62**
Etiquetas anormalidades malignas: **54**	Imágenes con anormalidades malignas: **51**
Etiquetas mamografías normales: **209**	Imágenes normales: **209**
Total etiquetas: **330**	Total imágenes: **322**

Dado que el objetivo del presente proyecto es clasificar las mamografías y el tejido como normal o anormal, y no clasificar las anormalidades como benignas o malignas, en la columna SEVERITY, se unifican las anormalidades benignas y malignas como una sola clase anormal. Por lo tanto, se pasa de tener 3 tipos de SEVERITY, a tener 2: anormal (A) y normal (N). También se llenan los valores nulos con 0,0 (Tabla 6):

Tabla 6. Instancias del *mias_dataframe* con dos clases de SEVERITY (330, 7).

	REFNUM	BG	CLASS	SEVERITY	X	Y	RADIUS
0	mdb001	G	CIRC	A	535.0	425.0	197.0
1	mdb002	G	CIRC	A	522.0	280.0	69.0
2	mdb003	D	NORM	N	0.0	0.0	0.0

Después, los datos numéricos se convierten a enteros debido a la naturaleza de las variables X, Y y RADIO. También, se realiza una validación de los REFNUM etiquetados como anormales que no contengan las coordenadas X, Y y RADIO de la anormalidad. En la Tabla 7, se muestran las imágenes anormales que no se encuentran etiquetadas correctamente. Estas imágenes se eliminan del *mias_dataframe* porque, al no contener la ubicación y el radio de las anormalidades, no permitirían entrenar un sistema para detectar la localización de las mismas en nuevas imágenes.

Tabla 7. Imágenes anormales sin etiquetar correctamente.

Abnormal with no X, Y labeled:

	REFNUM	BG	CLASS	SEVERITY	X	Y	RADIUS
59	mdb059	F	CIRC	A	0	0	0
218	mdb216	D	CALC	A	0	0	0
238	mdb233	G	CALC	A	0	0	0
251	mdb245	F	CALC	A	0	0	0

Al finalizar este procesamiento de los datos, el *mias_dataframe* pasa de tener 330 etiquetas a tener 326, como se observa en la Tabla 8 y en la Figura 26.

Tabla 8. *mias_dataframe* después del *Data Wrangling* (326, 7).

	REFNUM	BG	CLASS	SEVERITY	X	Y	RADIUS
0	mdb001	G	CIRC	A	535	425	197
1	mdb002	G	CIRC	A	522	280	69
2	mdb003	D	NORM	N	0	0	0

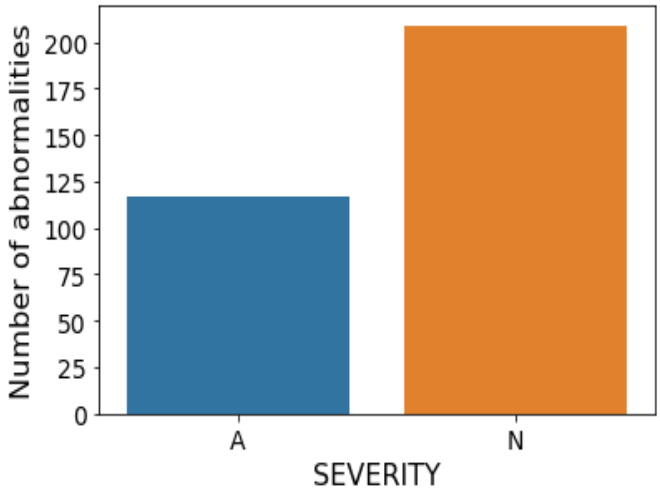


Figura 26. Número de mamografías por clase después del *DW.*

Después del *Data Wrangling* de los datos, el *dataset* queda distribuido de la siguiente manera (Tabla 9):

Tabla 9. Cantidad de etiquetas vs. cantidad de imágenes después del *DW.*

Etiquetas anormalidades: **117**	Imágenes con anormalidades: **109**
Etiquetas mamografías normales: **209**	Imágenes normales: **209**
Total etiquetas: **326**	Total imágenes: **318**

Estas son las etiquetas e imágenes definitivas que se utilizan en el entrenamiento de las máquinas de aprendizaje en la etapa de clasificación y detección de anormalidades. Para los dos métodos de segmentación de la mama que se presentan a continuación, se procesarán las 322 imágenes, con el fin de evaluar los resultados de la segmentación en una muestra mayor.

4.2.2. Método 1 de segmentación de la mama: a partir de operaciones morfológicas, entropía y otras

Este método de segmentación del seno se implementa en MATLAB R2019a (todo lo demás se desarrolla en Python, como se mencionó anteriormente). El objetivo es

eliminar los artefactos extraños y el tejido pectoral presentes en las mamografías para segmentar solo la región de interés, la mama. Se basa en el método propuesto por Derouiche, Chaima *et al.* [55], específicamente, en el orden de algunas de las operaciones morfológicas aplicadas y en el procedimiento para la detección y cambio de orientación de las mamografías para la localización del tejido pectoral.

Las operaciones y procesamiento de las imágenes se ajustan a las necesidades específicas de este proyecto y los demás procedimientos, como la eliminación del pectoral, son propuestos en este trabajo.

A continuación, se muestra el diagrama de bloques resumido del método 1 de segmentación de la mama: a partir de operaciones morfológicas, entropía y otras (Figura 27).

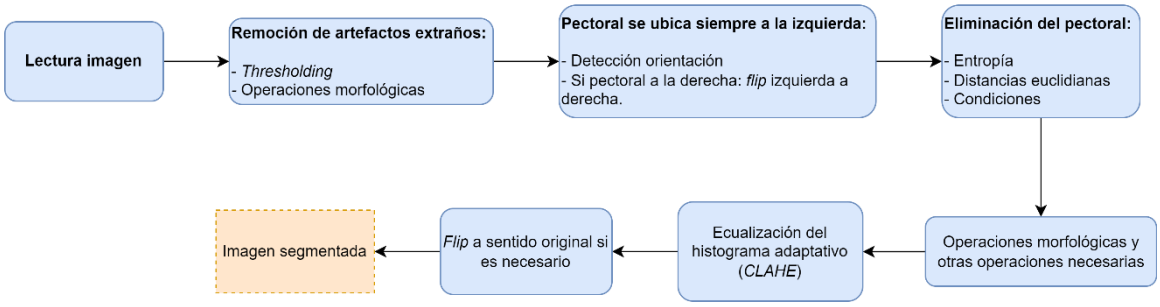


Figura 27. Diagrama general método 1 de segmentación de la mama: a partir de operaciones morfológicas, entropía y otras.

La descripción detallada de este proceso de segmentación se encuentra en el Anexo 1: Descripción detallada del método 1 de segmentación de la mama: a partir de operaciones morfológicas, entropía y otras.

4.2.3. Método 2 de segmentación de la mama: con red neuronal convolucional U-Net

Este método de segmentación de la mama se implementó con la arquitectura de red neuronal convolucional U-Net, recomendada para la segmentación de imágenes

médicas (ver marco teórico: Arquitectura U-Net).

Para ser entrenada, se requerían las máscaras binarias de la región de interés (mama) de cada mamografía (*ground truth*). Estas son las clases a las que la red busca llegar ajustando los filtros durante su entrenamiento y siguiendo su proceso de *contracting* y *expanding path*.

El mini-MIAS *dataset* no suministra estas máscaras binarias. Por lo que se optó por realizarlas “manualmente” con un software para Linux llamado LabelMe, desarrollado por el MIT [56]. Este permite crear máscaras a través de la marcación manual de puntos para formar polígonos y delimitar las regiones de interés, como se muestra en la Figura 28.

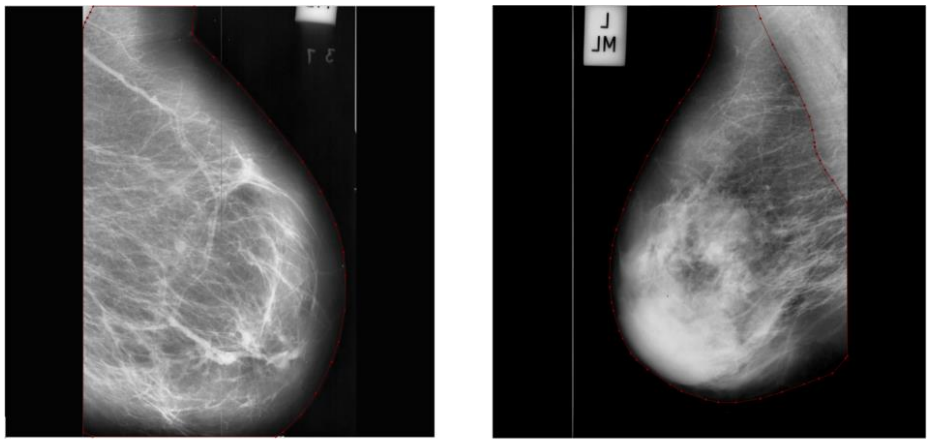


Figura 28. Creación de polígonos con LabelMe: (a) mdb314, (b) mdb257.

Normalmente, las máscaras de imágenes médicas deben ser definidas por un experto, en este caso, un radiólogo. Pero lo que se requiere en este apartado, es segmentar la mama como tal y esta se puede distinguir fácilmente de los artefactos extraños y del tejido pectoral sin necesidad de ser un experto, por lo que se procedió a crear las 322 máscaras “manualmente”.

Este software genera un archivo JSON por cada máscara con las coordenadas de

los polígonos creados. Se implementó una función para leer estos archivos y convertirlos en imágenes binarias (máscaras), las cuales se almacenan en formato PNG.

En la Figura 29, se ilustran algunas imágenes del *dataset* y las máscaras binarias creadas (*ground truth*).

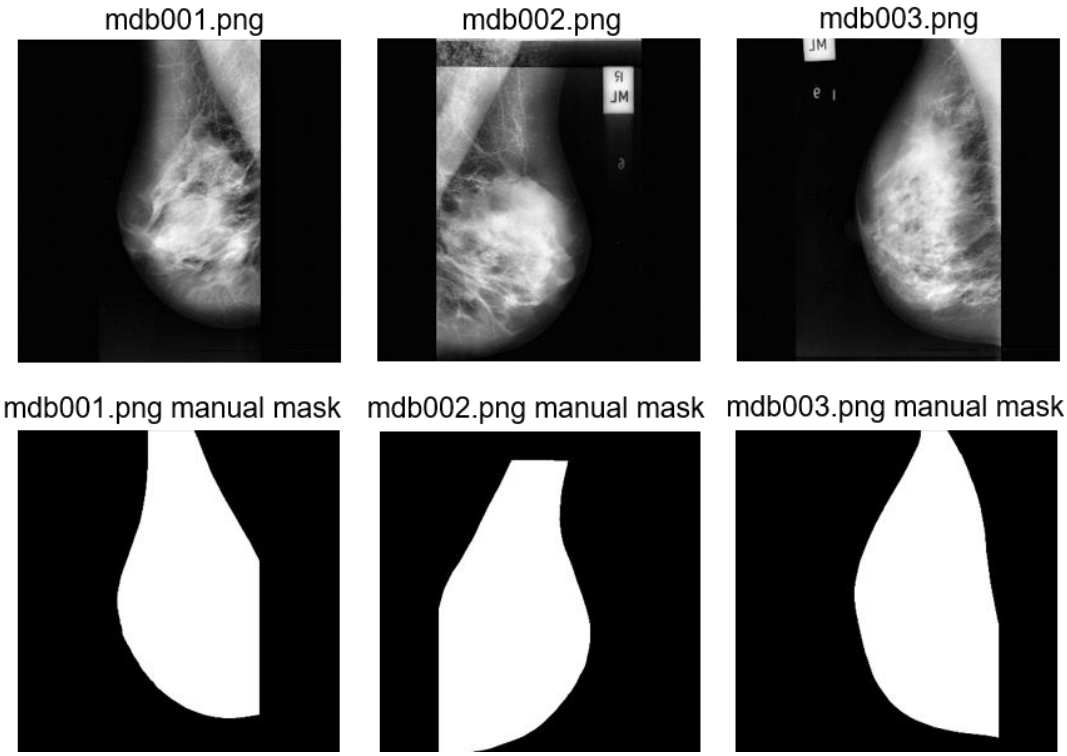


Figura 29. Imágenes y sus máscaras creadas manualmente con LabelMe.

Después de crear todas las máscaras, se procedió a modelar la red neuronal convolucional U-Net. Siguiendo su arquitectura, se realiza el apilamiento (*stack*) de cada capa.

Las **imágenes de entrada** a la red son imágenes con histograma ecualizado que originalmente eran de 1.024 x 1.024 pixeles, se redimensionaron a 128 x 128 y se normalizaron.

Las **clases o *labels*** a los que la red U-Net busca llegar durante el entrenamiento,

son las máscaras creadas con LabelMe (*ground truth images*), las cuales también se redimensionaron a 128 x 128. Los parámetros definitivos para el entrenamiento fueron los siguientes:

- **Imágenes para entrenamiento:** 274
- **Porcentaje de datos para validación:** 10% de conjunto de *train.*
- **Imágenes para *test*:** 48
- **Optimizador:** estimación de momento adaptativo (Adam), que combina dos metodologías de descenso de gradiente.
- **Tasa de aprendizaje:** 0,0001.
- **Función de pérdida**: entropía cruzada binaria.
- **Métrica:** exactitud.
- **Tamaño de lotes (*batch size*)** = 20 (para que sean 13 pasos por época, porque son 274 imágenes de *train* = 274/20 = 13,7).
- **Épocas:** 120

A continuación, en las Figuras 30 y 31, se muestran las gráficas de exactitud y pérdida, respectivamente, para el conjunto de validación durante el entrenamiento:

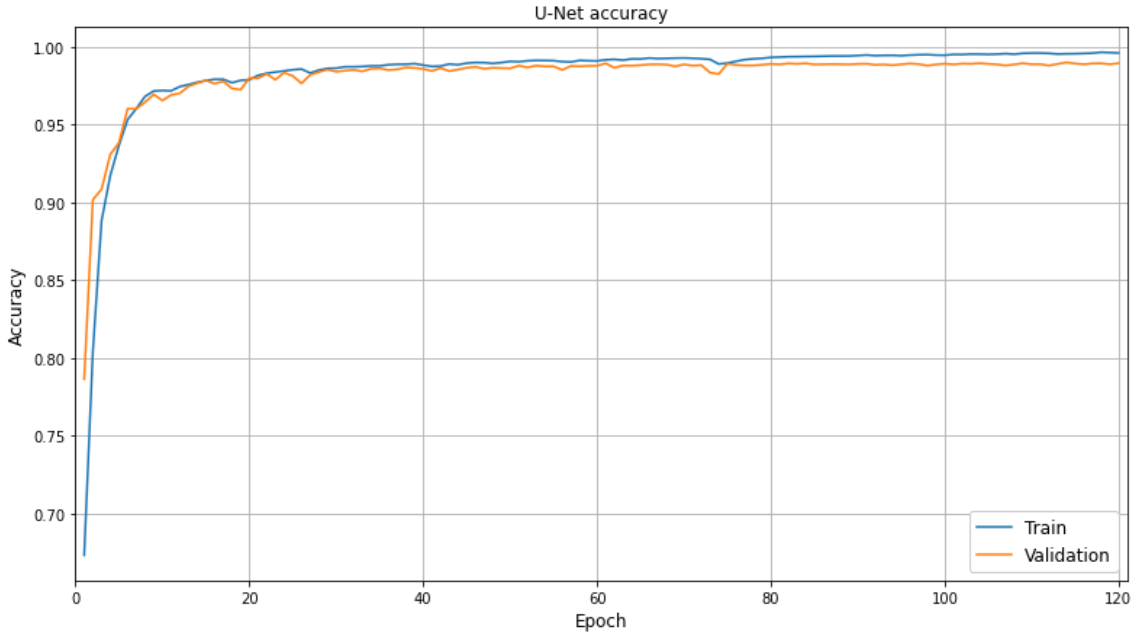


Figura 30. Exactitud durante entrenamiento de U-Net.

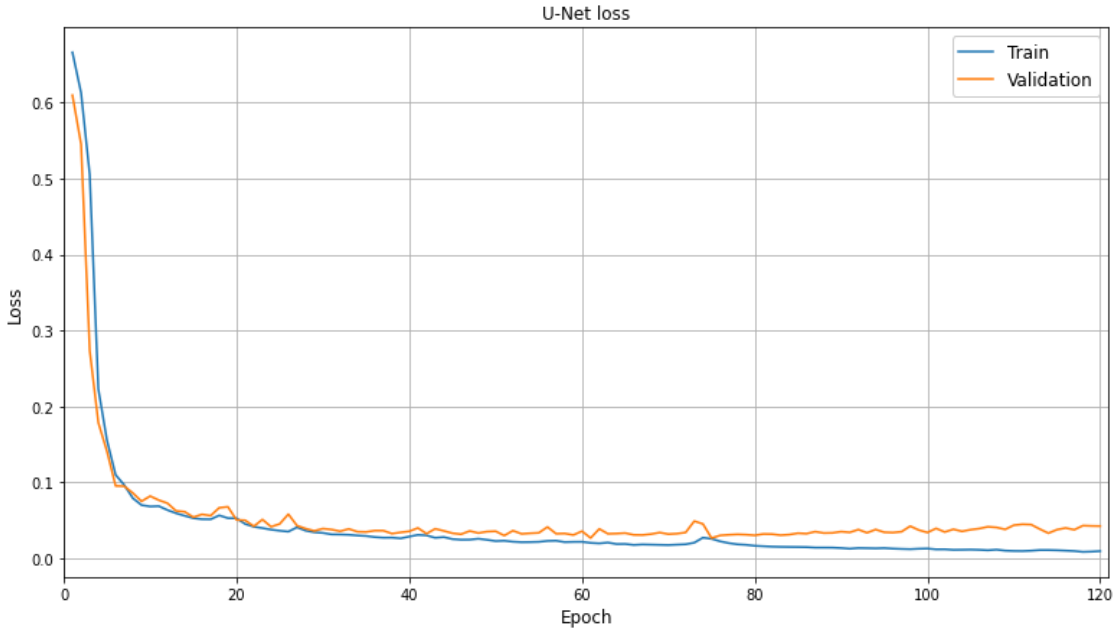


Figura 31. Pérdida o error durante entrenamiento de U-Net.

Se observa que aproximadamente a partir de la época 40 se alcanzó casi el valor de máximo de exactitud y mínimo de error en el conjunto de validación.

Se guardó el modelo con menor error en el conjunto de validación:

- **Exactitud en conjunto de validación (*validation accuracy*):** 0,9899
- **Error en conjunto de validación (*validation loss*):** 0,0263

Posteriormente, las imágenes resultantes segmentadas, se redimensionaron a su tamaño original.

Como se mencionó en el marco teórico, U-Net es una arquitectura de red profunda netamente convolucional, que permite realizar segmentación semántica en imágenes, es decir, asignar una clase o etiqueta a cada pixel presente en la imagen, es muy utilizada en segmentación de imágenes médicas. Es por esta razón que los valores de exactitud y precisión son tan altos y tan bajos, respectivamente. Porque la mayor parte de los pixeles de la imagen son clasificados correctamente.

4.3. Clasificación y detección de anormalidades

4.3.1. Introducción

Se implementaron dos soluciones para la clasificación y detección de anormalidades en las mamografías del mini-MIAS *dataset*, la primera con la combinación de las arquitecturas de redes neuronales convolucionales profundas DenseNet121 + U-Net. La segunda, a partir de extracción “manual” de características, perceptrón multicapa, máquina de soporte vectorial y ventana deslizante.

4.3.2. Método de clasificación y detección de anormalidades 1: con redes neuronales convolucionales DenseNet121 + U-Net

4.3.2.1. Introducción

Este método consiste en la utilización de redes neuronales convolucionales profundas, específicamente, las arquitecturas DenseNet121 para clasificar las mamografías como normales o anormales y U-Net para la detección de las anomalías en las mamografías clasificadas como anormales.

4.3.2.2. Clasificación de mamografías como normales o anormales con arquitectura DenseNet121

4.3.2.2.1. Organización de imágenes para entrenamiento, validación y *test*

Después del *Data Wrangling* del *mias_dataframe*, se conoce que las imágenes que cuentan con etiquetas correctas de ubicación y radio de las anormalidades son 318 imágenes y 326 etiquetas en total. También que las clases en las que se dividió el *dataset*, pasaron de ser tres (normal, benigno o maligno) a dos, “anormal” (*abnormal*) y “normal” (*normal*). De las 318 imágenes, 109 pertenecen a la clase anormal y 209 a la clase normal.

Por lo anterior, se creó una función que toma del *mias_dataframe* los nombres de los archivos con su clase correspondiente y genera dos carpetas, de la siguiente manera (Figura 32):

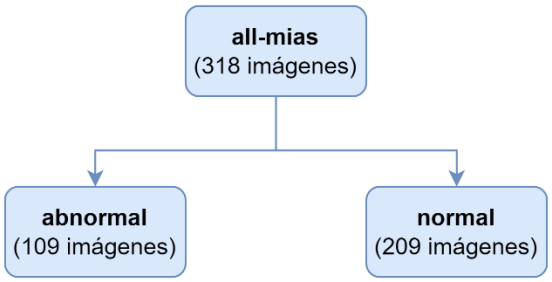


Figura 32. Carpetas iniciales

Luego se implementaron funciones para automatizar el proceso de división del *dataset* en conjuntos de entrenamiento (70%), validación (15%) y *test* (15%) y la creación de sus respectivas carpetas. Con lo que se obtiene el siguiente organigrama (Figura 33):

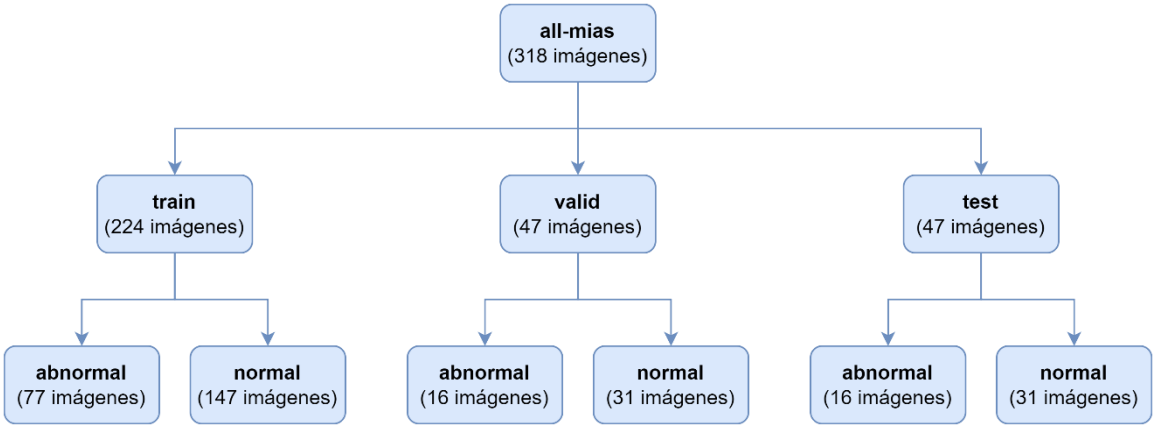


Figura 33. Organigrama inicial de conjuntos de entrenamiento, validación y *test*.

4.3.2.2.2. Aumento de datos

Debido al evidente desbalance entre clases que se puede observar en la Figura 33, y a la necesidad de un banco de datos con mayor cantidad de imágenes para el entrenamiento de los modelos, se aplicaron técnicas de aumento de datos para generar nuevas muestras a partir de las imágenes existentes. Esto con el objetivo

de tener mejores resultados en la clasificación y evitar efectos como el sobreajuste (*overfitting*) y el subajuste (*underfitting*) de los modelos entrenados.

Por lo cual, se desarrolló una función que aplica iterativamente distintas transformaciones a las imágenes de entrenamiento, es decir, a las pertenecientes a las carpetas *abnormal* y *normal* dentro de la carpeta de entrenamiento (*train*) (Figura 33). Específicamente, se generaron 3 imágenes por cada imagen original.

Se aplicaron combinaciones aleatorias de las siguientes transformaciones:

- Rotación con un valor aleatorio en un rango de $\pm 12°$.

- Desplazamiento aleatorio horizontal respecto al origen de la imagen en un rango de $\pm 15\%$ del ancho total de la misma.

- Desplazamiento aleatorio vertical respecto al origen de la imagen en un rango del $\pm 15\%$ del alto total de la misma.

- Giro (*flip*) horizontal aleatorio.

- Al aplicar las transformaciones mencionadas, se deben rellenar los espacios “vacíos” que quedan en la imagen, para lo cual se evalúa el valor de intensidad del pixel [0 0], con este, se completan estos espacios. Aunque los bordes de las mamografías del mini-MIAS *dataset* son aparentemente negros, se rellenan esos espacios con el nivel de gris del fondo para conservar los mismos niveles.

- En este caso, no se aplicaron modificaciones de brillo y contraste porque las imágenes de entrada fueron preprocesadas y segmentadas previamente.

A continuación, en la Figura 34 se muestran algunos ejemplos de las transformaciones aplicadas:

Figura 34. Ejemplos de transformaciones aplicadas para aumento de datos

Después del aumento de datos, se crea una función para eliminar las imágenes sobrantes generadas e igualar las cantidades de los conjuntos de entrenamiento anormal y normal. Con lo que se obtiene el organigrama de carpetas definitivo, que se ilustra a continuación, en la Figura 35:

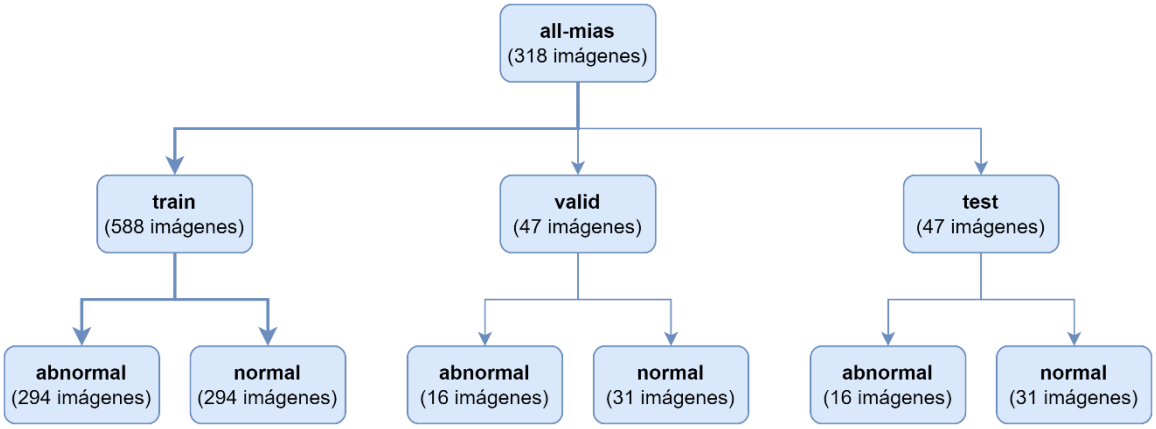


Figura 35. Organigrama de conjuntos de entrenamiento, validación y *test* después del aumento de datos de entrenamiento.

4.3.2.2.3. Configuración de arquitectura DenseNet121

Para aprovechar las bondades de la transferencia de aprendizaje (ver marco teórico: Transferencia de aprendizaje), se cargó el modelo de la arquitectura de red neuronal convolucional profunda DenseNet121 desde la librería Keras, junto con los pesos

pre-entrenados con la base de datos de imágenes ImageNet [57].

Se configuró la capa de entrada con dimensiones 224, 224, 3. Tiene 3 canales porque la arquitectura original de esta red trabaja con imágenes RGB, por lo cual, en este caso cada canal se entrenó en paralelo con la misma imagen mamográfica en escala de grises.

Con las anteriores dimensiones de entrada, los parámetros totales de la arquitectura DenseNet121 pre-entrenada con la base de datos ImageNet para 1.000 clases son los siguientes:

- Parámetros totales: 8.062.504
- Parámetros entrenables: 7.978.856
- Parámetros no entrenables: 83.648

4.3.2.2.4. Ajuste fino (*fine tuning)* de la arquitectura DenseNet121

La capa de salida de la arquitectura DenseNet121, es una capa densa (red neuronal artificial completamente conectada (Dense)) cuenta con 1.000 clases y 1.025.000 parámetros. Por lo cual, se elimina la última capa (de la red neuronal artificial Dense) que tenía 1.000 neuronas de salida y se conecta una nueva con 2 neuronas de salida con función de activación *softmax*, correspondientes a las dos clases que nos interesan clasificar (*abnormal* (‘A’: 0) y *normal* (‘N’: 1)).

Los parámetros totales de la arquitectura DenseNet121 modificada para 2 clases, son los siguientes:

- Parámetros totales: 7.039.554
- **Parámetros entrenables: 6.955.906**
- Parámetros no entrenables: 83.648

Y la nueva capa densa de salida queda con 2 clases y 2.050 parámetros.

Posteriormente, solo se habilitan las últimas 20 capas para ser entrenadas y se congelan los pesos de las capas anteriores. Reduciendo significativamente los parámetros a entrenar:

- Parámetros totales: 7.039.554
- **Parámetros entrenables: 368.962**
- Parámetros no entrenables: 6.670.592

4.3.2.2.5. Entrenamiento de la red neuronal convolucional DenseNet121

A continuación, se sintetizan los parámetros de configuración y entrenamiento que se definieron para la red neuronal convolucional profunda DenseNet121:

Las **imágenes de entrada** a la red son imágenes con el histograma ecualizado adaptativamente (resultado del método 1 de segmentación de la mama: a partir de operaciones morfológicas y otras) que originalmente eran de 1.024 x 1.024 pixeles, se redimensionaron a 224 x 224 y se normalizaron.

Las **clases o *labels*** anormal y normal se extrajeron del *mias_dataframe* y se relacionan con las imágenes de entrada.

Los conjuntos de entrenamiento, validación y *test,* se ilustran en la Figura 35.

Los parámetros para el entrenamiento fueron los siguientes:

- **Optimizador:** estimación de momento adaptativo (Adam), que combina dos metodologías de descenso de gradiente.
- **Tasa de aprendizaje:** 0,0001
- **Función de pérdida**: entropía cruzada categórica.

- **Métrica:** exactitud.
- **Tamaño de lotes (*batch size*)** = 49
- **Épocas:** 50

En las figuras subsiguientes (Figuras 36 y 37), se muestran respectivamente las gráficas de exactitud y pérdida de los conjuntos de entrenamiento y validación durante el entrenamiento del modelo:

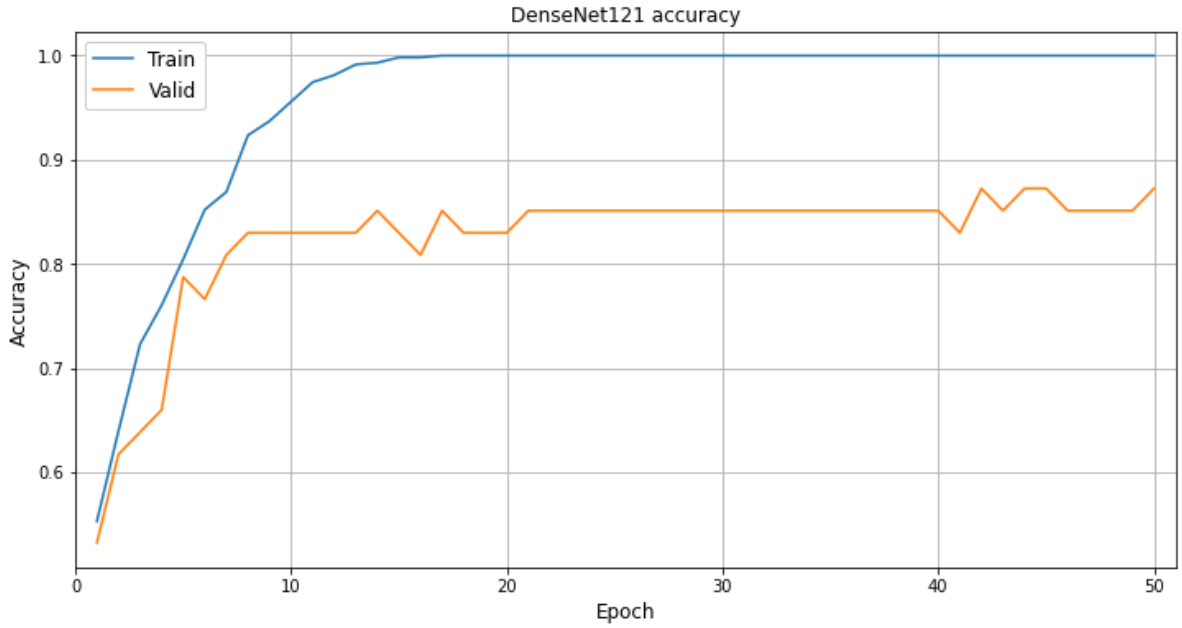


Figura 36. Exactitud durante entrenamiento de DenseNet121.

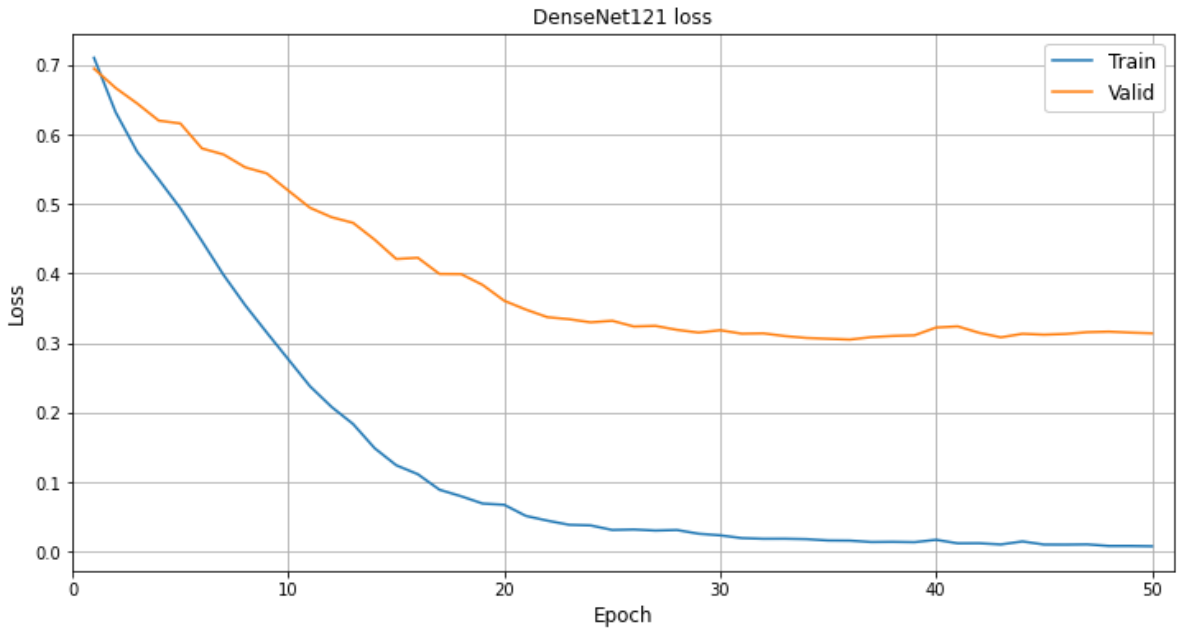


Figura 37. Pérdida o error durante entrenamiento de DenseNet121.

En estas gráficas de entrenamiento de exactitud y pérdida, Figuras 36 y 37, se puede observar que aproximadamente a partir de la época 30 se llegó casi al valor de máximo de exactitud (*accuracy*) y mínimo de pérdida (*loss*) en el conjunto de validación.

Se guardó el modelo con menor error en el conjunto de validación:

- **Exactitud en conjunto de validación (*validation accuracy*):** 0,8951
- **Error en conjunto de validación (*validation loss*):** 0,3134

4.3.2.3. Detección de anormalidades en imágenes anormales con arquitectura U-Net

4.3.2.3.1. Organización de imágenes para entrenamiento, validación y *test*

Para implementar este método de detección y entrenar esta red convolucional, se utilizaron las imágenes originales del mini_MIAS *dataset,* con artefactos extraños y pectoral incluidos. De esta carpeta, se tomaron las 318 imágenes correspondientes a las etiquetas de imágenes definitivas después del *Data Wrangling*, donde 109 pertenecen a la clase anormal y 209 a la clase normal. A estas imágenes solo se les ecualizó el histograma adaptativamente (CLAHE).

Posteriormente, se implementó una función para crear máscaras binarias de las anormalidades a partir de las coordenadas X, Y y del radio suministrados en el *mias_dataframe.* Se transformaron las coordenadas de Y porque las suministradas ubican la coordenada 0 en la parte inferior izquierda de la imagen y OpenCV, lo hace desde la esquina superior izquierda. También se entrenó con imágenes y máscaras de la clase normal (imágenes de ceros), para reforzar el aprendizaje de las características de tejido normal.

En la Figura 38 se exponen algunas imágenes anormales del mini-MIAS *dataset* con la ubicación de sus anomalías y las máscaras binarias creadas correspondientes (*ground truth*).

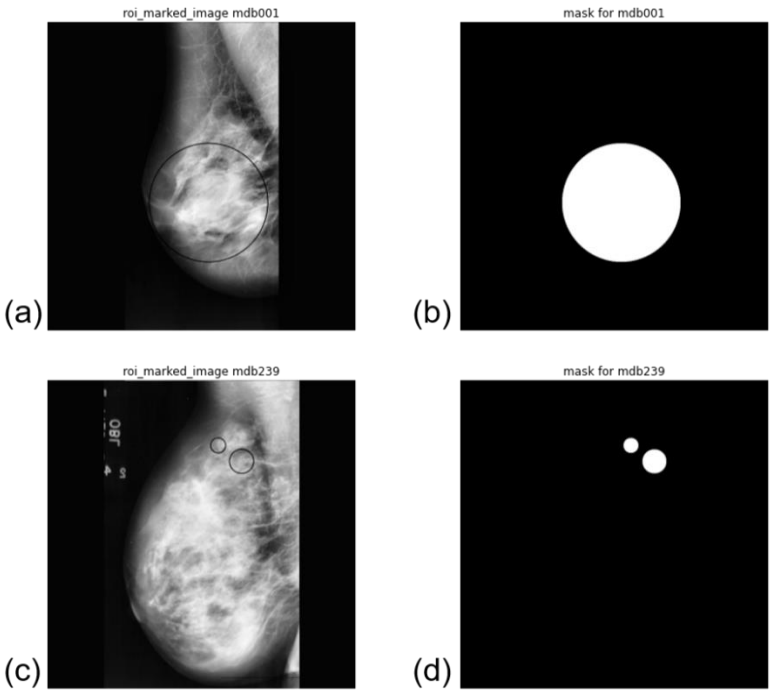


Figura 38. Máscaras extraídas a partir de información suministrada en *mias_dataframe.*

Partiendo de lo anterior, se implementaron funciones para generar carpetas con las imágenes y máscaras (clases o etiquetas a los que la red busca llegar) de la siguiente manera (Figura 39):

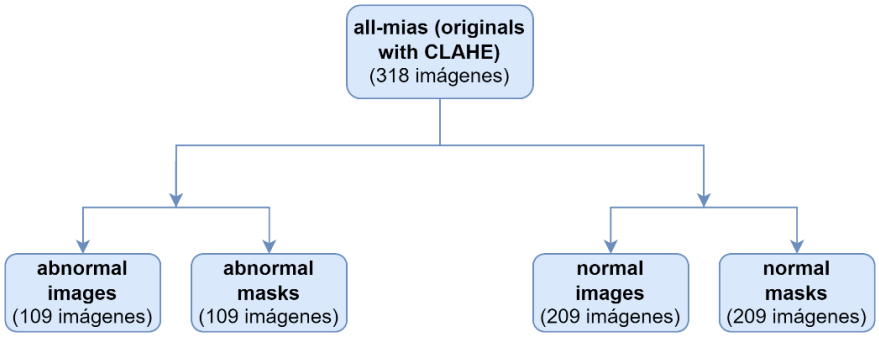


Figura 39. Organigrama original de carpetas de imágenes para U-Net.

Para elaborar los conjuntos de entrenamiento, validación y *test*, se separaron las imágenes en carpetas de entrenamiento y *test* porque durante el entrenamiento se especifica cuánto porcentaje del conjunto de entrenamiento se toma para validación. Los porcentajes definidos fueron 70% para datos de entrenamiento, 15%

para validación y 15% para *test*.

Después de realizar la separación mencionada, la organización de los conjuntos de datos (imágenes y máscaras) se muestra en la Figura 40.

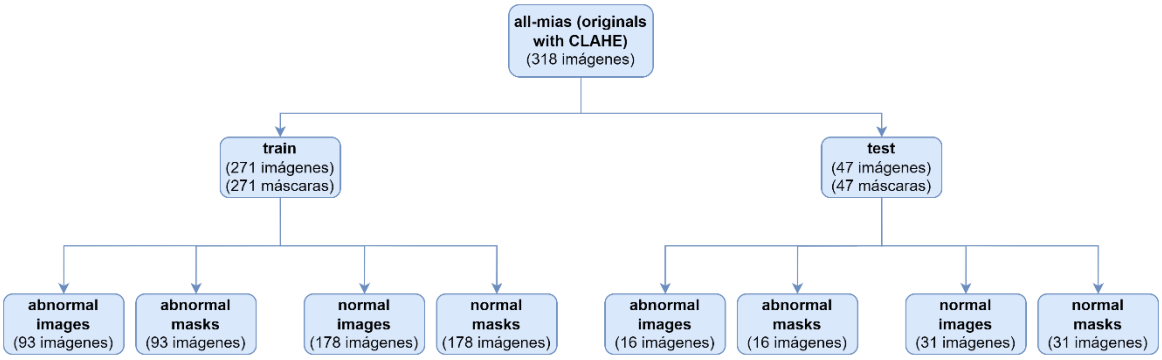


Figura 40. Organigrama de carpetas de imágenes y máscaras dividido en conjuntos para entrenamiento y *test.*

4.3.2.3.2. Aumento de datos

Para aumentar el número de muestras y balancear las clases para evitar los efectos de sobreajuste y subajuste, se aplicaron a las imágenes de entrenamiento y sus máscaras correspondientes, las mismas transformaciones realizadas a las imágenes de entrenamiento del apartado anterior (DenseNet121).

Después de aplicadas las transformaciones de aumento de datos a los conjuntos de entrenamiento, los cantidad y distribución de los conjuntos de imágenes definitivos para las distintas etapas de este método de detección de anormalidades con la arquitectura de red neuronal convolucional U-Net, se muestran a continuación, en la Figura 41:

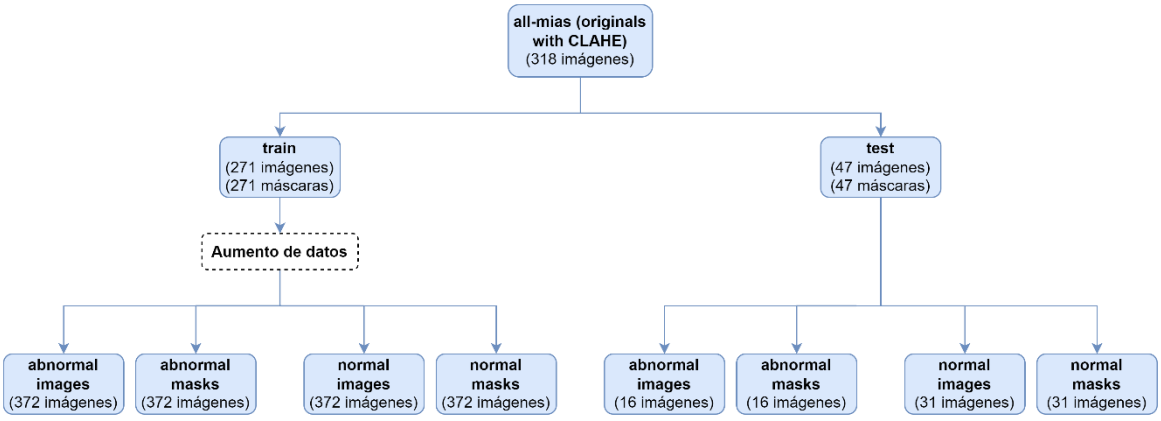


Figura 41. Organigrama de conjuntos de entrenamiento y *test* definitivos después de aumento de datos.

4.3.2.3.3. Entrenamiento de la red neuronal convolucional U-Net

Se implementó el modelo de la red neuronal convolucional U-Net. Siguiendo su arquitectura, se realizó el apilamiento (*stack*) de cada capa. Seguidamente, se exponen los parámetros de configuración y entrenamiento que se definieron para la red neuronal convolucional U-Net:

Las **imágenes de entrada** a la red, se redimensionaron a 256 x 256 y se normalizaron (originalmente son de 1.024 x 1.024).

Para el entrenamiento, se cargaron las imágenes y máscaras en arreglos para reducir la velocidad de carga de los datos durante el entrenamiento y en la predicción posterior del conjunto de *test*.

- **Porcentaje de datos para validación:** 15% del conjunto de entrenamiento.
- **Optimizador:** estimación de momento adaptativo (Adam), que combina dos metodologías de descenso de gradiente.
- **Tasa de aprendizaje:** 0,0001
- **Función de pérdida**: entropía cruzada binaria.
- **Monitor:** error o pérdida.

- **Tamaño de lotes (*batch size*)**: 20
- **Épocas:** 200

A continuación, en la Figura 42 se presenta la gráfica de exactitud y en la Figura 43 la gráfica de pérdida, de los conjuntos de entrenamiento y validación durante el entrenamiento del modelo:

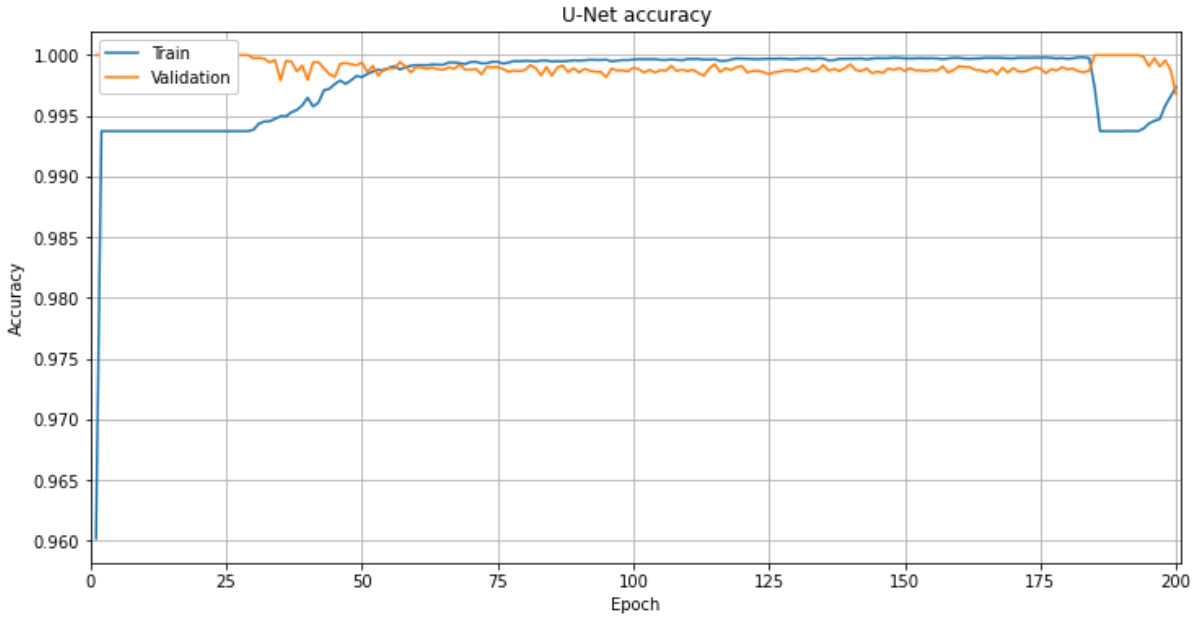


Figura 42. Exactitud en los conjuntos de entrenamiento y validación durante el entrenamiento de U-Net.

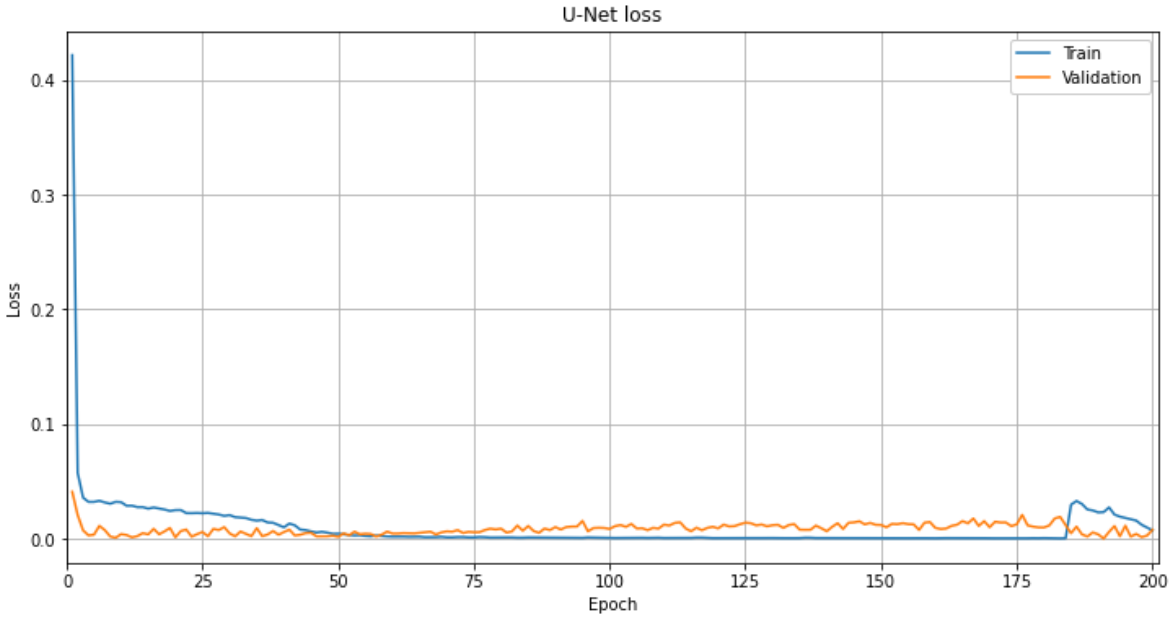


Figura 43. Pérdida o error en los conjuntos de entrenamiento y validación durante el entrenamiento de U-Net.

Se guardó el modelo que menor error presentó en el conjunto de validación.

En las gráficas de entrenamiento de exactitud y pérdida, Figuras 42 y 43, respectivamente, se observan resultados aparentemente muy positivos desde la primera época, presentando visualmente variaciones muy pequeñas durante las épocas siguientes. Pero se debe tener en cuenta que esta red realiza segmentación semántica y clasifica los pixeles de la imagen, que siendo imágenes de 256 x 256, el total de pixeles es 65.536, la mayoría de pixeles presentes en la imagen serán fácilmente clasificados por la red y los de interés, que son pocos respecto a toda la imagen, representan un porcentaje muy pequeño, por lo cual, variaciones muy mínimas en el error de validación, darán como resultado cambios muy significativos en la detección final de las anormalidades. Debido a este análisis, la red se entrenó por bastantes épocas para encontrar el menor valor de error de validación, aunque la mejora fuera por algunos decimales.

Se guardó el modelo con menor error en el conjunto de validación:

- **Exactitud en conjunto de validación (*validation accuracy*):** 0,9996
- **Error en conjunto de validación (*validation loss*):** 9,7412e-04

4.3.2.4. Conclusiones

- Se entrenó la red neuronal profunda DenseNet121 con el método 1 de segmentación de la mama: a partir de operaciones morfológicas y otras, porque se obtuvieron mejores resultados, esto debido a que es más cercana al tejido glandular de las mamas y contienen menos información “ruidosa” para los sistemas de clasificación.

- Se probó con la arquitectura de red convolucional MobileNet, por tener pocos parámetros comparada con otras arquitecturas, pero no permitió clasificar correctamente mamografías como normales o anormales, como sí lo hizo la arquitectura DenseNet121. También se tuvo en cuenta la arquitectura Inception v3 [58], que es utilizada y recomendada en el ámbito de imágenes médicas, pero por su gran cantidad de parámetros y costo computacional, no se utilizó.

- Después de haber sido entrenada, la red neuronal convolucional U-Net permitió detectar muy bien las anomalías presentes en mamografías anormales, pero en varias mamografías normales también encontraba anormalidades que obviamente no existían. Se entrenó la red con las imágenes anormales, normales y sus respectivas máscaras, (para las normales las máscaras son imágenes con los pixeles en cero). También se probó entrenando solo con las imágenes anormales y sus máscaras. En ambas opciones se realizó aumento de datos. Pero siempre, en muchos de los casos de *test*, U-Net detectó anormalidades en las imágenes normales.

- Al clasificar las mamografías como normales o anormales con la red profunda DenseNet121 con muy buena exactitud, se puede pensar que esta

podría permitir aportar información para la detección de las anormalidades sin necesidad de utilizar U-Net posteriormente. Pero se realizaron pruebas con Grad-Cam, que se define como mapeo de activación de clase ponderado por gradiente, el cual utiliza la información de gradiente específica de la clase de la capa convolucional final para producir un mapa de localización aproximado de las regiones importantes (discriminantes) en la imagen. Estas regiones se observaron en un mapa de calor y aunque este mapa abarcaba pixeles de zonas cercanas a la anormalidad, cubría un área extensa que no permitía deducir cuál era la ubicación real o bien aproximada de la anormalidad.

- Debido a que el aprendizaje profundo requiere grandes cantidades de muestras para inferir características y poder clasificar entre clases, se optó por realizar *transfer learning* debido a que, aunque se hizo aumento de datos, sigue siendo una cantidad baja comparada con las cantidades que normalmente se utilizan para entrenar estos modelos desde cero.

4.3.3. Método de clasificación y detección de anormalidades 2: a partir de extracción de características, ventana deslizante, MLP y SVM

4.3.3.1. Introducción

Para el desarrollo de este método de clasificación y detección, se tomaron como entrada las carpetas de imágenes del organigrama que se presenta en la Figura 32, que contiene 109 mamografías anormales y 209 mamografías normales.

Se entrenaron dos máquinas de aprendizaje: un perceptrón multicapa y una máquina de soporte vectorial para comparar el rendimiento de ambas. Se entrenaron con características extraídas a regiones (*patches*) de las imágenes mamográficas preprocesadas y segmentadas.

La construcción del vector de características definitivo se basó en características

estadísticas del histograma y descriptores a partir de la matriz de co-ocurrencia de niveles de gris (GLCM). Estas características y descriptores se calcularon de las regiones extraídas de las imágenes originales y de las imágenes resultado de la localización de patrones binarios (LBP) de las regiones extraídas.

Posteriormente, se implementó un algoritmo de ventana deslizante para clasificar cada región y marcarla si se clasifica como parte de una anormalidad.

Los pasos se detallan a continuación.

4.3.3.2. Extracción de las regiones anormales

De las mamografías anormales, se extrajeron las regiones (*patches*) de las anomalías a partir de las coordenadas X, Y y del radio suministrados en los metadatos del *mias_dataframe.* Las anormalidades tienen dimensiones distintas, por lo que se redimensionaron a 64 x 64. A continuación, en la Figura 44, se muestran algunas de las regiones anormales extraídas redimensionadas:

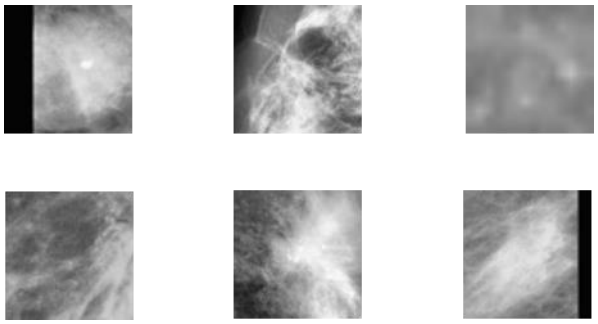


Figura 44. Regiones anormales.

4.3.3.3. Extracción de regiones normales

Se extrajeron regiones aleatorias de las imágenes normales de entrada de la misma dimensión de los anteriores (64 x 64), dentro de un rango de ubicación entre 128 y 770, tanto en X, como como en Y (recordar que las imágenes son de 1.024 x 1.024), para que no predominaran regiones oscuras pertenecientes a las franjas izquierda

y derecha de la imagen. En la Figura 45, que se presenta seguidamente, se exponen algunas de las regiones normales extraídas de dimensiones 64 x 64:

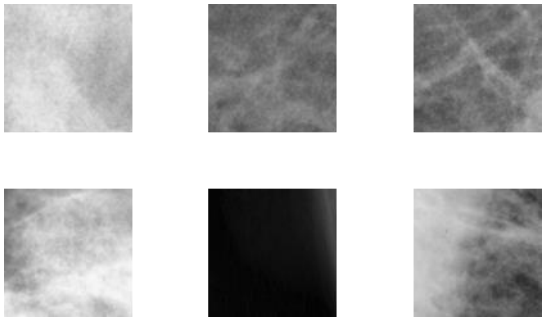


Figura 45. Regiones de tejido normal.

4.3.3.4. Organización de regiones (*patches*) para entrenamiento y *test*

El total de regiones anormales fue 117 (número de etiquetas de anormalidades del *mias_dataframe*), pocas para el entrenamiento, por lo cual, se aplicaron técnicas de aumento de datos. Las regiones de tejido normal se extrajeron después de realizar el aumento de datos de las regiones anormales, ya que se tienen más imágenes normales y debido a que sus regiones se extraen aleatoriamente como se mencionó en el apartado anterior (Extracción de regiones normales), se puede decidir el número de regiones necesario para balancear las clases.

A continuación, se muestra el proceso de aumento de datos de las regiones anormales.

4.3.3.5. Aumento de datos

A cada una de las 117 regiones de tejido anormal, se le aplicaron combinaciones aleatorias de las siguientes transformaciones:

- Rotación con un valor aleatorio en un rango de ±5°.

- Desplazamiento aleatorio horizontal respecto al origen de la imagen en un

rango de ±10% del ancho total de la misma.

- Desplazamiento aleatorio vertical respecto al origen de la imagen en un rango del ±10% del alto total de la misma.

- Giro (*flip*) horizontal aleatorio.

- Al aplicar las transformaciones mencionadas, se deben rellenar los espacios "vacíos" que quedan en la imagen, para lo cual se evalúa el valor de intensidad del pixel [0 0], que corresponde al fondo de una de las imágenes mamográficas de entrada, con este valor, se completan los espacios vacíos luego de las transformaciones de los *patches* para conservar los mismos niveles de gris de fondo.

En la Figura 46, se expone un ejemplo de una región original y dos de las transformaciones de aumento aplicadas a la misma:

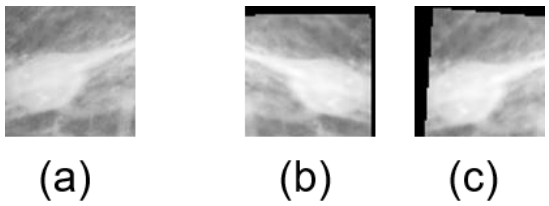


Figura 46. (a) mdb132, (b) y (c) transformaciones de mdb132 para aumento de datos.

Después del aumento de datos, se crea una función para eliminar las imágenes sobrantes generadas e igualar las cantidades de entrenamiento anormal y normal.

Por lo cual, el nuevo organigrama de carpetas de imágenes con las divisiones respectivas de entrenamiento, validación, *test* y sus respectivas clases, se ilustra a continuación, en la Figura 47.

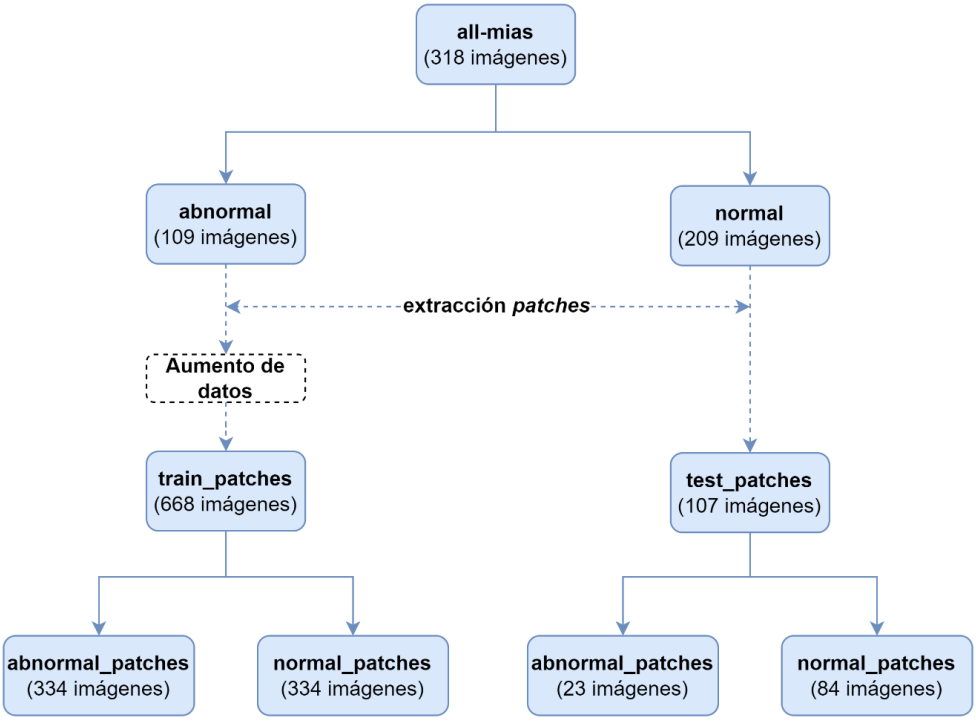


Figura 47. Organigrama de imágenes de regiones extraídas (*patches*) después del aumento de datos.

4.3.3.6. Extracción de características

Con las imágenes (regiones) de entrenamiento balanceadas, se procedió a crear las respectivas funciones para la extracción de características de cada una para entrenar los sistemas de clasificación y detección subsiguientes.

Cada uno de los pasos siguientes se aplicó iterativamente a cada una de las regiones.

4.3.3.6.1. Cálculo de patrones binarios locales (LBP)

1. Se calculan los patrones binarios locales de cada imagen (región), con los siguientes parámetros:

    - **Radio** = 1

- **Número de pixeles vecinos** = 8

2. Se calcula el negativo de la imagen LBP.

4.3.3.6.2. Extracción de características del histograma

1. Se calcula la media, la mediana y la entropía del histograma de la imagen.

2. Se calcula la media, la mediana y la entropía del histograma del negativo de la imagen LBP.

4.3.3.6.3. Extracción de características desde la matriz de co-ocurrencia de niveles de gris (GLCM)

Se calcula la matriz de co-ocurrencia de niveles de gris (GLCM) con los siguientes parámetros:

- **Ángulo** = 0°
- **Distancia** = 1
- **Niveles** = 256
- **Simétrica**
- **Normalizada**

De la matriz de co-ocurrencia de niveles de gris se calculan los siguientes descriptores:

- **Disimilitud**
- **Correlación**
- **Homogeneidad**
- **Energía**

4.3.3.6.4. Construcción de vector de características

Sintetizando lo anterior, se obtiene el siguiente vector de características para el entrenamiento (Tabla 10):

Tabla 10. Vector de características de entrenamiento.

	HIST_MEAN	HIST_MEDIAN	HIST_ENTROPY	LBP_HIST_MEAN	LBP_HIST_MEDIAN	LBP_HIST_ENTROPY	DISSIMILARITY	CORRELATION	HOMOGENEITY	ENERGY	LBP_DISSIMILARITY	LBP_CORRELATION	LBP_HOMOGENEITY	LBP_ENERGY
0	14.294821	15.0	5.342287	14.227092	1.0	3.967462	9.608135	0.967523	0.194715	0.095016	47.645585	0.637626	0.291910	0.084982
1	14.063745	13.0	5.371941	14.067729	1.0	3.903732	9.554067	0.978658	0.200286	0.099953	51.865823	0.569270	0.284130	0.096970
2	13.892430	13.0	5.369847	13.916335	1.0	3.939059	9.818452	0.970737	0.202119	0.109656	51.602927	0.587127	0.283992	0.095923
3	14.876494	7.0	4.796910	14.039841	1.0	3.884106	7.677579	0.899846	0.280410	0.069794	53.727431	0.562964	0.269229	0.091570
4	14.617530	10.0	4.919968	13.737052	1.0	3.915002	5.337302	0.971540	0.297890	0.092968	51.294147	0.605999	0.289460	0.110300
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
663	16.318726	0.0	4.113790	13.637450	3.0	4.448380	5.885913	0.877822	0.195261	0.036470	88.930804	0.263185	0.074188	0.036014
664	16.318726	2.0	4.602547	14.051793	3.0	4.255864	2.965774	0.991036	0.306059	0.036998	81.689980	0.209136	0.080067	0.037270
665	6.629482	0.0	2.010128	4.693227	0.0	3.546996	0.932788	0.963721	0.788668	0.596011	44.839286	0.368538	0.641825	0.609798
666	0.000000	0.0	0.000000	1.418327	0.0	1.525773	0.024554	0.955178	0.987723	0.691752	9.133433	0.582290	0.920520	0.891314
667	0.000000	0.0	0.000000	1.250996	0.0	1.424882	0.015873	0.900395	0.992063	0.908306	7.020337	0.633801	0.936591	0.908130

El vector de características de entrada para entrenamiento definitivo cuenta con 14 características y 668 instancias, como se muestra en la Tabla 11.

Tabla 11. Vector de características de *test.*

	HIST_MEAN	HIST_MEDIAN	HIST_ENTROPY	LBP_HIST_MEAN	LBP_HIST_MEDIAN	LBP_HIST_ENTROPY	DISSIMILARITY	CORRELATION	HOMOGENEITY	ENERGY	LBP_DISSIMILARITY	LBP_CORRELATION	LBP_HOMOGENEITY	LBP_ENERGY
0	16.318726	8.0	4.755613	14.689243	2.0	4.248358	4.843998	0.981495	0.205358	0.025572	70.840030	0.438076	0.126826	0.039288
1	16.318726	1.0	4.567416	14.928287	1.0	3.991448	2.530258	0.993764	0.358594	0.043424	52.235615	0.620786	0.200081	0.064649
2	16.318726	16.0	4.955676	14.689243	3.0	4.209592	5.341022	0.985209	0.207376	0.023714	61.931300	0.530802	0.153138	0.053970
3	16.318726	16.0	4.926003	15.505976	1.0	3.590327	2.712302	0.996399	0.358420	0.034241	39.391865	0.694103	0.328330	0.095695
4	16.318726	11.0	4.830153	15.223107	1.0	3.817388	2.791915	0.995786	0.337639	0.035334	46.748512	0.664556	0.236513	0.085286
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
102	16.318726	0.0	4.389631	12.557769	3.0	4.171442	2.377232	0.989136	0.342620	0.041545	102.487351	0.009841	0.056384	0.047925
103	16.318726	0.0	3.829626	13.525896	3.0	4.414030	5.477183	0.811023	0.198330	0.041792	87.199901	0.304045	0.073167	0.038103
104	16.318726	1.0	4.601772	13.792829	3.0	4.415136	5.324901	0.962600	0.193433	0.026920	86.119048	0.274895	0.079697	0.032879
105	16.318726	7.0	4.705695	14.318725	3.0	4.077912	3.043899	0.991123	0.278448	0.032967	75.873016	0.238547	0.104491	0.050419
106	0.000000	0.0	0.000000	1.003984	0.0	1.467805	0.000000	1.000000	1.000000	1.000000	2.926091	0.827894	0.968285	0.938366

El vector salida para el entrenamiento (*labels* o clases), correspondientes a cada instancia (SEVERITY), originalmente estaba dispuesto de manera binaria, es decir, 0 para anormal y 1 para normal, pero se convirtieron a dos categorías distintas (*dummies*), para tener dos neuronas a la salida, como se expone a continuación, en la Tabla 12:

Tabla 12. Vector de etiquetas o *labels* de entrenamiento y *test.*

```
     A  N
0    1  0
1    1  0
2    1  0
3    1  0
4    1  0
..  .. ..
663  0  1
664  0  1
665  0  1
666  0  1
667  0  1

[668 rows x 2 columns]
     A  N
0    1  0
1    1  0
2    1  0
3    1  0
4    1  0
..  .. ..
102  0  1
103  0  1
104  0  1
105  0  1
106  0  1

[107 rows x 2 columns]
```

4.3.3.6.5. Entrenamiento de perceptrón multicapa (MLP)

Se entrenó un perceptrón multicapa con 14 neuronas de entrada, 19 neuronas en la capa oculta, a la cual se le configuró un abandono o *dropout* del 15% (omisión aleatoria de neuronas durante el proceso de entrenamiento para evitar el sobreajuste) y finalmente, dos neuronas de salida. La función de activación de las neuronas de la capa oculta fue ReLU y la función de activación de las neuronas de salida fue la función Softmax. El esquema de la red se muestra a continuación, en la Figura 48:

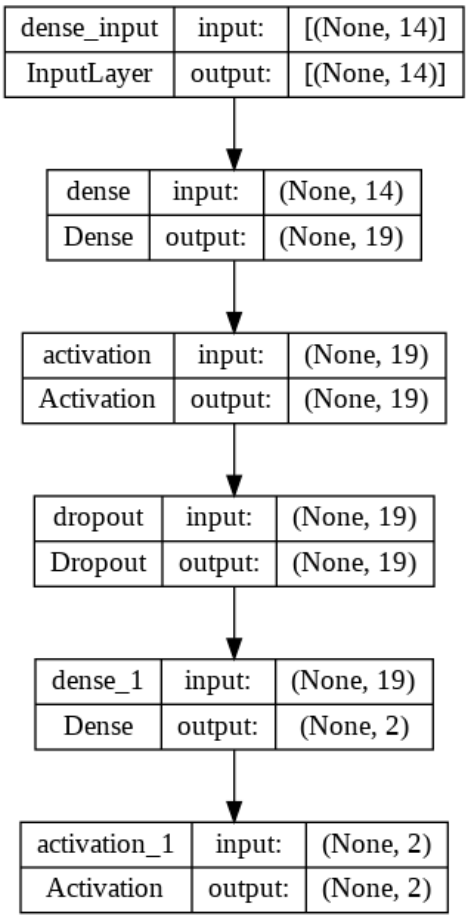


Figura 48. Modelo del perceptrón multicapa implementado.

- **Porcentaje de datos para validación:** 15% del conjunto de entrenamiento.
- **Optimizador:** estimación de momento adaptativo (Adam), que combina dos metodologías de descenso de gradiente.
- **Tasa de aprendizaje:** 0,0001
- **Función de pérdida:** entropía cruzada categórica
- **Métrica:** exactitud
- **Tamaño de lotes (*batch size*):** 2
- **Épocas:** 40

En las figuras que se presentan a continuación, Figuras 49 y 50, se muestran las gráficas de exactitud y pérdida, respectivamente, de los conjuntos de entrenamiento y validación durante el entrenamiento del perceptrón multicapa. Se puede observar

que aproximadamente desde la época 12, se llega a los valores de exactitud máxima y error mínimo para el conjunto de validación.

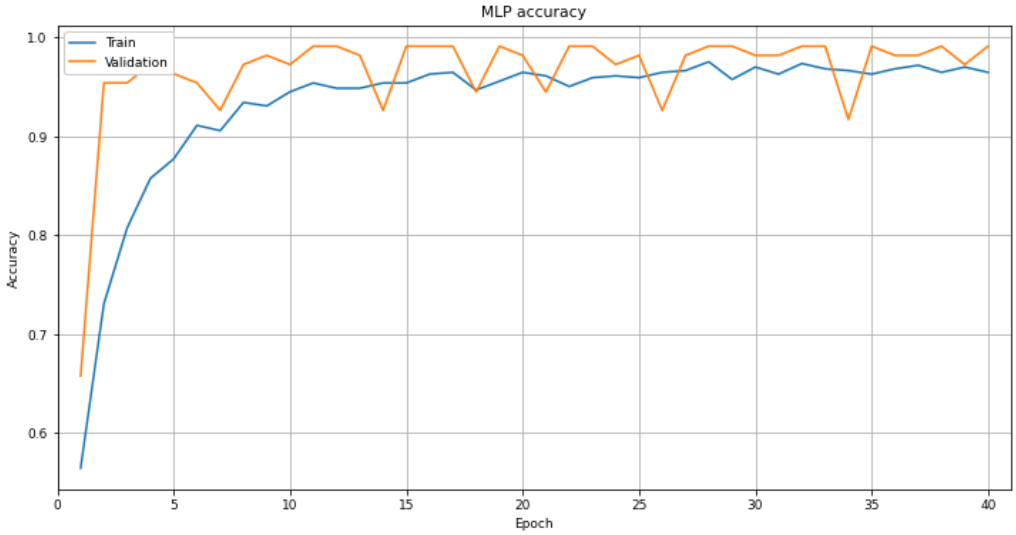


Figura 49. Exactitud con las características del conjunto de entrenamiento y validación durante el entrenamiento del MLP.

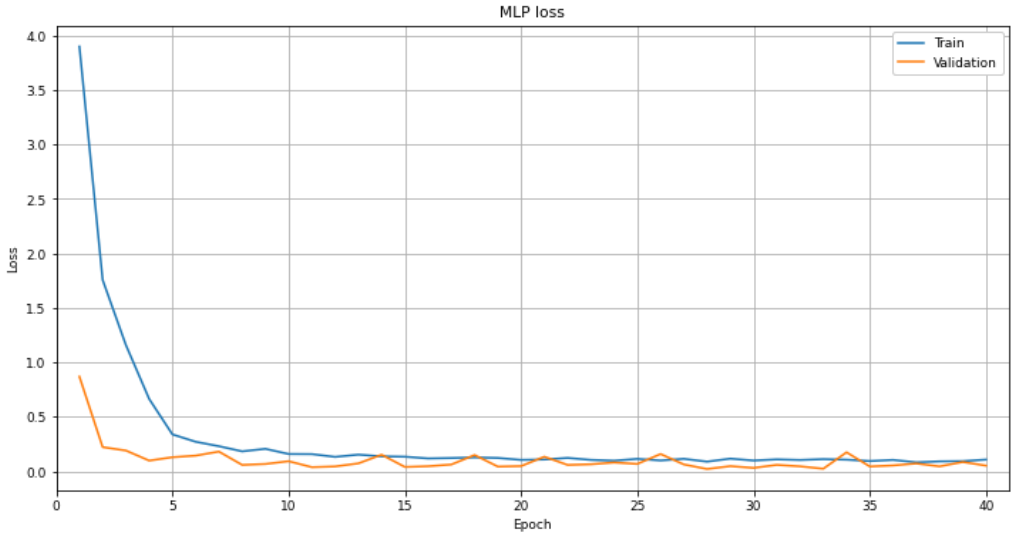


Figura 50. Error o pérdida con las características del conjunto de entrenamiento y validación durante el entrenamiento del MLP.

4.3.3.6.6. Clasificación de regiones con SVM

Se entrenó la máquina de soporte vectorial (SVM) con el vector de características

definitivo y sus respectivas etiquetas o *labels* (SEVERITY) para clasificar las regiones como anormales o normales.

Se configuró con un **kernel lineal** (*kernel='linear'*).

4.3.3.6.7. Ventana deslizante

Se desarrolló un algoritmo de ventana deslizante (*sliding window*) para recorrer las imágenes de *test* en pasos de 8 pixeles, horizontal y verticalmente, como se ilustra en la Figura 51, para extraer iterativamente regiones de 64 x 64 de las imágenes mamográficas preprocesadas y segmentadas con el método 1 de segmentación de la mama: a partir de operaciones morfológicas y otras.

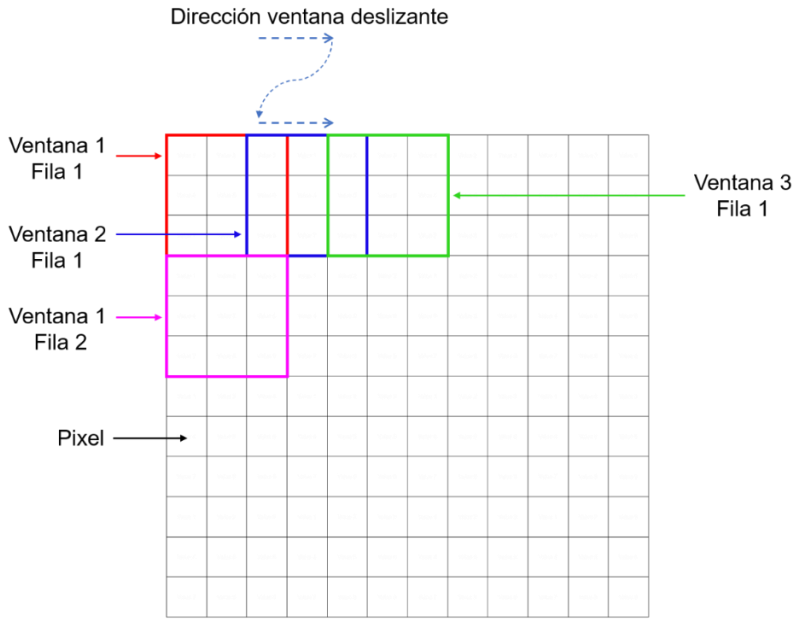


Figura 51. Ilustración ventana deslizante.

Los pasos realizados para la clasificación de las regiones extraídas, se detallan a continuación:

1. A cada región se le extraen las características del histograma y de textura mencionadas anteriormente (vector de características definitivo).

2. Con base en las características extraídas a la región, esta se clasifica como anormal o normal con la máquina de soporte vectorial entrenada. Esto porque la SVM presentó un mejor rendimiento que el perceptrón multicapa (MLP), además de tener menos costo computacional.

3. Si la región es clasificada como anormal, se marca con un recuadro rojo, de lo contrario, no se marca (ver capítulo de resultados).

4.3.3.7. Conclusiones

- Para la clasificación de las regiones, también se utilizaron las características estadísticas del histograma: desviación estándar, oblicuidad (*skewness*) y curtosis, además de características extraídas del histograma de gradientes orientados HOG. Se realizaron distintas combinaciones de características y se observó que estas no aportaban positivamente a la clasificación, por lo que se eliminaron del vector de características definitivo.

- A las características iniciales y a las extraídas se les aplicó análisis en componentes principales (PCA), con el fin de reducir el vector de características y mejorar su rendimiento, pero no se obtuvo mejora, por lo que no se utilizaron en el algoritmo definitivo.

- Con las máquinas de aprendizaje superficial se intentaron clasificar las imágenes completas inicialmente, pero no se obtuvieron buenos resultados al clasificar las mamografías como normales o anormales por lo que se optó por extraer regiones y clasificar los segmentos de la imagen con el algoritmo de ventana deslizante.

- La matriz de co-ocurrencia de niveles de gris (GLCM), se creó y probó con distintos ángulos de inclinación, pero se utilizó el ángulo cero porque fue el que mejores resultados arrojó.

5. Pruebas y resultados

5.1. Introducción

A continuación, se dan a conocer las diferentes pruebas realizadas a cada uno de los métodos desarrollados, con sus respectivos resultados.

5.2. Prueba método 1 de segmentación de la mama: a partir de operaciones morfológicas y otras

5.2.1. Descripción

Para evaluar la segmentación realizada, se calculó el índice de similitud de Jaccard entre las máscaras resultantes del método 1 de segmentación de la mama: a partir de operaciones morfológicas y otras, y las máscaras extraídas manualmente con el software LabelMe (*ground truth*).

5.2.2. Procedimiento

1. Se implementó una función para calcular iterativamente el índice de Jaccard (ver marco teórico) entre todas las máscaras resultantes de la segmentación de la mama con este método y sus respectivos *ground truth* (mamas segmentadas manualmente con software LabelMe).

2. Se implementó una función para graficar las imágenes originales y el resultado del traslape entre las máscaras correspondientes resultado del método 1 de segmentación de la mama: a partir de operaciones morfológicas y otras, y sus respectivos *ground truth* (ver Figura 52).

3. Se calculó la media de los índices de Jaccard de todas las máscaras segmentadas respecto a las máscaras *ground truth* (ver Tabla 13).

5.2.3. Resultados y análisis

En la Figura 52, se presentan visualmente los resultados del método 1 de segmentación de la mama: a partir de operaciones morfológicas y otras, de las 4 primeras imágenes del mini-MIAS *dataset.*

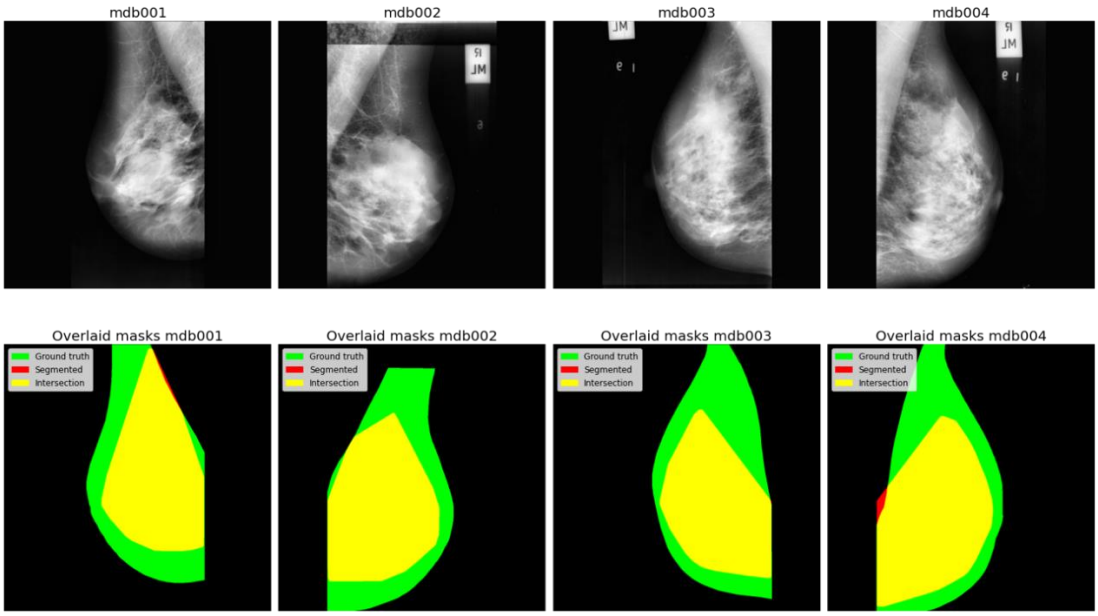


Figura 52. Resultados método segmentación 1

En la Tabla 13, se observa el índice de similitud de Jaccard entre las 5 primeras y 5 últimas máscaras segmentadas con este método y sus respectivos *ground truth*. En esta tabla también se puede observar la media del índice de similitud de Jaccard de todas las máscaras segmentadas con este método de segmentación de la mama.

Tabla 13. Índices Jaccard método 1 de segmentación de la mama: a partir de operaciones morfológicas y otras

	GT_MASK	SEG_MASK	JACCARD_INDEX
0	mdb001_gt	mdb001_seg	0.647013
1	mdb002_gt	mdb002_seg	0.629524
2	mdb003_gt	mdb003_seg	0.569048
3	mdb004_gt	mdb004_seg	0.643100
4	mdb005_gt	mdb005_seg	0.736155
...	...	...	...
317	mdb318_gt	mdb318_seg	0.832201
318	mdb319_gt	mdb319_seg	0.737878
319	mdb320_gt	mdb320_seg	0.778285
320	mdb321_gt	mdb321_seg	0.728492
321	mdb322_gt	mdb322_seg	0.681400

322 rows × 3 columns

Average Jaccard similarity: 0.7224

Media del índice de similitud de Jaccard de todas las máscaras segmentadas con el método 1 de segmentación de la mama: a partir de operaciones morfológicas y otras: 0,7224

Los resultados de este método de segmentación permiten observar que el resultado no es muy exacto en comparación a las máscaras de *ground truth*.

A pesar de esto, en las imágenes de la Figura 52 se puede analizar que la segmentación se acerca más al tejido glandular de la mama, que es donde regularmente se presentan las anormalidades en mama.

Es decir, permite anular información “ruidosa” para los clasificadores implementados.

5.3. Prueba método 2 de segmentación de la mama: con red neuronal convolucional U-Net

5.3.1. Descripción

Para evaluar este método, también se calculó el índice de similitud de Jaccard entre las máscaras resultantes y las máscaras extraídas manualmente con el software LabelMe (*ground truth*).

5.3.2. Procedimiento

1. Se implementó una función que permite graficar las imágenes originales y el resultado del traslape entre las máscaras segmentadas resultantes de este método y sus respectivos *ground truth* (ver Figura 53).

2. Se implementó una función que calcula iterativamente el índice de Jaccard entre todas las máscaras segmentadas resultantes de este método y sus respectivos *ground truth* (mamas segmentadas manualmente con software LabelMe).

3. Se calculó la media de los índices de Jaccard de todas las máscaras segmentadas respecto a las máscaras *ground truth* (ver Tabla 14).

5.3.3. Resultados y análisis

En la Figura 53, se presentan visualmente los resultados de este método de segmentación (método 2 de segmentación de la mama: con red neuronal convolucional U-Net) de las 4 primeras imágenes del mini-MIAS *dataset.*

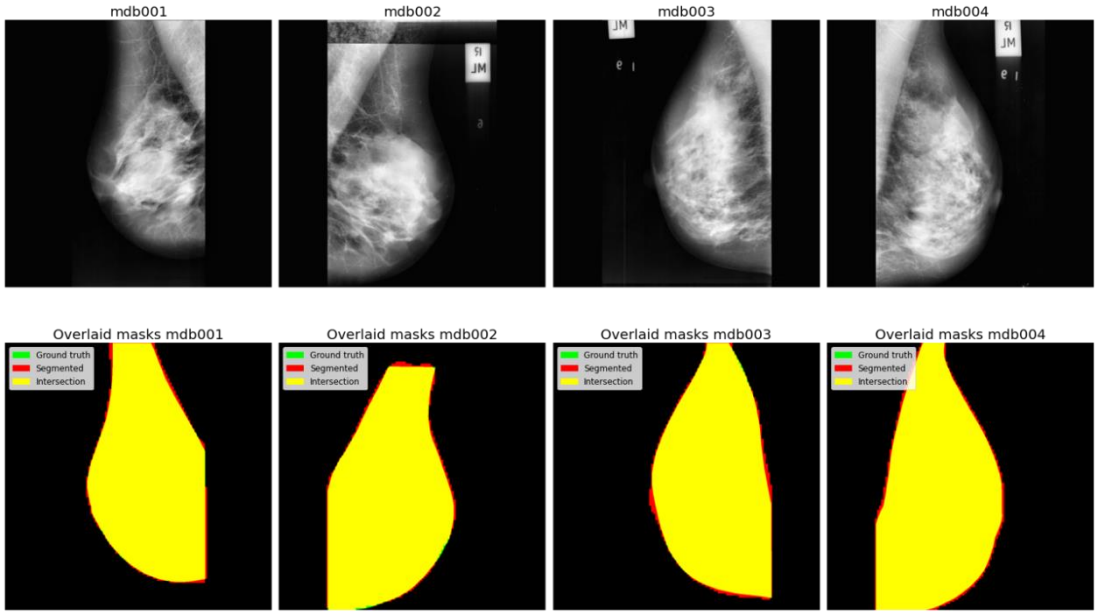


Figura 53. Resultados Método de segmentación de la mama con U-Net.

En la Tabla 14, se observa el índice de similitud de Jaccard entre las máscaras segmentadas con este método de segmentación de la mama y sus respectivos *ground truth*, de las 5 primeras y 5 últimas máscaras. En esta tabla también se puede observar la media del índice de similitud de Jaccard de todas las máscaras segmentadas con este método.

Tabla 14. Índices Jaccard Método de segmentación de la mama con U-Net.

	GT_MASK	SEG_MASK	JACCARD_INDEX
0	mdb001_gt	mdb001_seg	0.961804
1	mdb002_gt	mdb002_seg	0.962702
2	mdb003_gt	mdb003_seg	0.960331
3	mdb004_gt	mdb004_seg	0.956987
4	mdb005_gt	mdb005_seg	0.971610
...	...	...	...
317	mdb318_gt	mdb318_seg	0.890565
318	mdb319_gt	mdb319_seg	0.969557
319	mdb320_gt	mdb320_seg	0.933151
320	mdb321_gt	mdb321_seg	0.956306
321	mdb322_gt	mdb322_seg	0.907736

322 rows × 3 columns

Average Jaccard similarity: 0.9563

Media del índice de similitud de Jaccard de todas las máscaras segmentadas con el método 2 de segmentación de la mama: con red neuronal convolucional U-Net: 0,9563

Los resultados de este método se segmentación de la mama (método 2 de segmentación de la mama: con red neuronal convolucional U-Net), son significativamente buenos, la red permitió segmentar con gran similitud el área de interés, la mama.

Aunque este resultado de segmentación es mucho mejor y más preciso que el primero, a través de las pruebas con las máquinas de aprendizaje para la clasificación y detección de anormalidades, se observa que el rendimiento es menor que con el primer método, esto debido a que este introduce más “ruido” y el primero se aproxima más al tejido netamente glandular de la mama, que es donde se presentan las masas o calcificaciones.

5.4. Prueba método de clasificación y detección de anormalidades 1: con redes neuronales convolucionales DenseNet121 + U-Net

5.4.1. Prueba de clasificación de mamografías como normales o anormales con DenseNet121

5.4.1.1. Descripción

Para evaluar el rendimiento de la arquitectura de red neuronal convolucional DenseNet121, se clasificó el conjunto de *test* con el modelo previamente entrenado.

5.4.1.2. Procedimiento

Con el modelo DenseNet121 ajustado y entrenado como se mencionó en el capítulo de la implementación de la solución, se procederá a predecir las imágenes del conjunto de *test* (16 imágenes normales y 31 anormales) y a calcular las métricas de desempeño mencionadas en el marco teórico, a continuación, en el apartado de resultados y análisis.

5.4.1.3. Resultados y análisis

En la matriz de confusión que se presenta a continuación, en la Figura 54, se puede observar que, de las predicciones del conjunto de *test,* se presentaron 14 verdaderos positivos, 30 verdaderos negativos, 1 falso positivo y 2 falsos negativos.

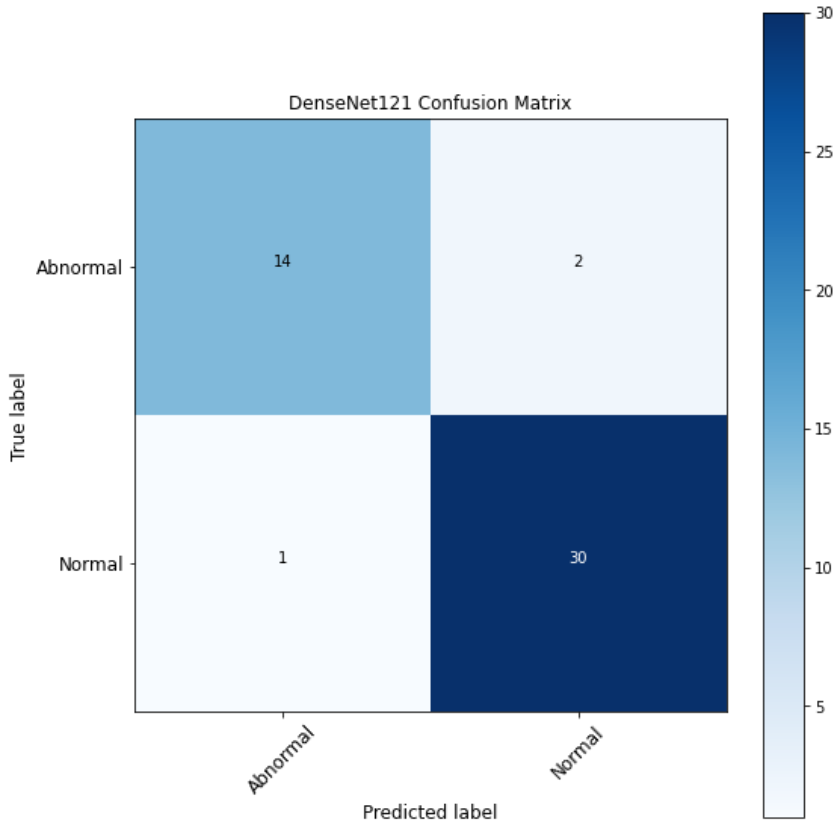


Figura 54. Matriz de confusión de clasificación de imágenes de *test* con DenseNet121

A continuación, se muestran las métricas de desempeño de clasificación de mamografías como anormales o normales con la arquitectura DenseNet121:

- **Exactitud DenseNet121 (*Test*):** 0,94
- **Precisión DenseNet121 (*Test*):** 0,93
- **Exhaustividad DenseNet121 (*Test*):** 0,88
- **Valor F1 DenseNet121 (*Test*):** 0,91

Se observa que se obtuvo una exactitud del 94% en la clasificación de imágenes mamográficas de *test* como normales o anormales con este método.

Aunque lo ideal sería que la exhaustividad fuera más alta que la precisión en este caso, porque es preferible tener más falsos positivos (falsos anormales) que falsos negativos (falsos normales), el error es mínimo, se considera que es un resultado

sobresaliente en la clasificación de mamografías como normales o anormales.

En la Figura 55, se expone la curva ROC y el AUC de la misma, resultante de la clasificación del conjunto de *test*.

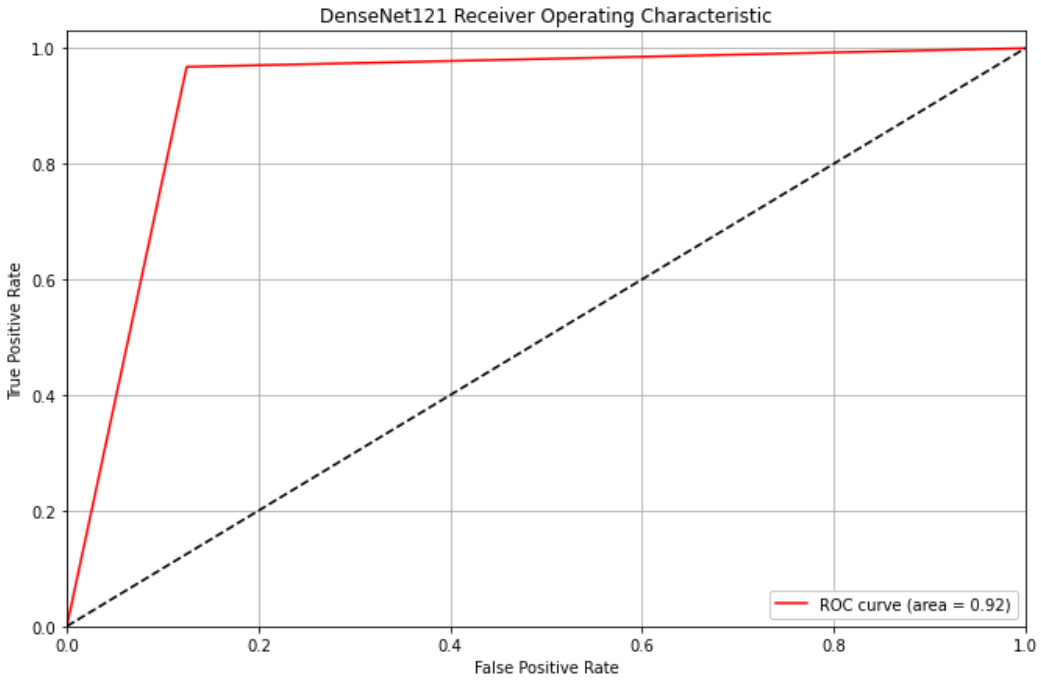


Figura 55. Curva ROC y AUC de clasificación de conjunto de *test* con arquitectura DenseNet121.

Se obtuvo un área bajo la curva ROC de 0,92 con el conjunto de *test.*

Con los resultados mostrados, se afirma que la red neuronal convolucional profunda DenseNet121 pudo inferir las características discriminantes entre mamografías normales y anormales a partir de las imágenes de entrenamiento, porque clasificó correctamente la mayoría de las imágenes de *test*.

5.4.2. Prueba de detección de anormalidades en imágenes anormales con arquitectura U-Net

5.4.2.1. Descripción

Para evaluar el desempeño de la detección de anormalidades de la red convolucional U-Net entrenada, se introdujeron como entrada al modelo entrenado, el conjunto de imágenes de *test* anormales, que son 16 imágenes, con sus respectivos *ground truth* (ver organización de los conjuntos de datos para entrenamiento, validación y *test* en el apartado correspondiente de la solución).

5.4.2.2. Procedimiento

Se predijeron iterativamente las 16 imágenes anormales del conjunto de *test.* Pues, como se mencionó anteriormente, la red U-Net se utilizó para detectar anormalidades en mamografías ya clasificadas como anormales por DenseNet121.

Para mostrar el desempeño de la detección de anormalidades del modelo U-Net entrenado, se mostrarán las imágenes originales, las máscaras *ground truth,* las máscaras predichas por el modelo y finalmente el traslape de las máscaras sobre la imagen de entrada, donde el color verde indicará el *ground truth*, el rojo la detección del modelo U-Net y el amarillo, la intersección entre ambos.

También, se calculará la intersección entre las máscaras detectadas y las máscaras *ground truth.* Se define esta métrica y no Jaccard en este caso, porque la red fue entrenada con máscaras que contienen círculos que encierran las anormalidades, pero no tienen su forma exacta y la red posiblemente haya encontrado características de las anormalidades como tal, no de círculos que las encierren. Por esto, si la anormalidad detectada se cruza con la anormalidad de la máscara *ground truth,* esta se tomará como correcta, y se le dará el valor de 1, si no se presenta intersección, se le dará el valor de 0, para finalmente promediar la intersección de las 16 imágenes de *test* y así contar con una métrica de desempeño.

5.4.2.3. Resultados y análisis

En las Figuras 56, 57 y 58, que se presentan a continuación, se ilustra el resultado de las 3 primeras imágenes del conjunto de *test* (mdb012, mdb021 y mdb023), las restantes se pueden observar en el Anexo 2 del presente documento.

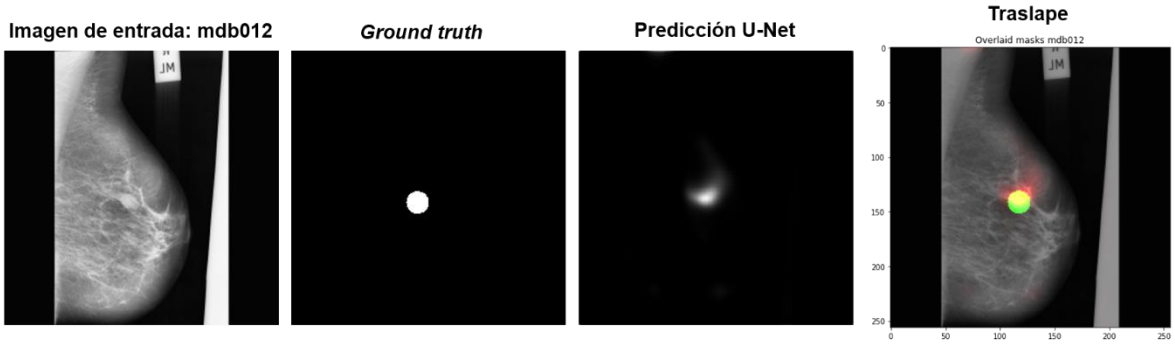


Figura 56. Resultado de detección con U-Net para la mamografía mdb012.

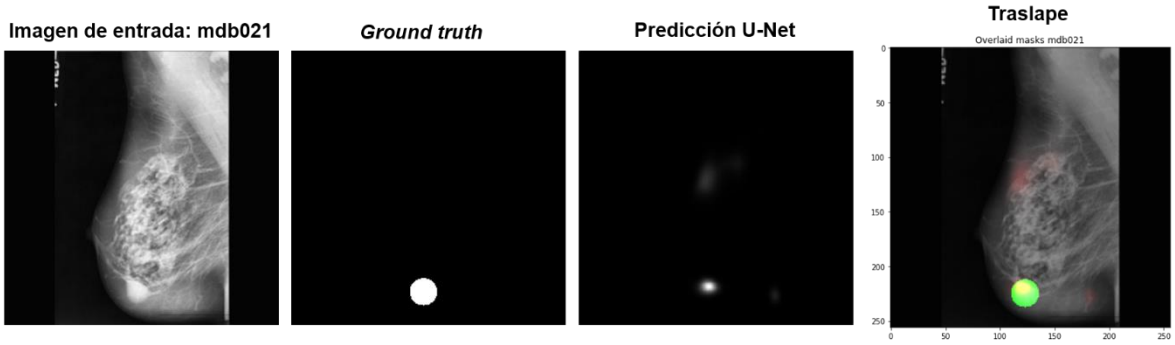


Figura 57. Resultado de detección con U-Net para la mamografía mdb021.

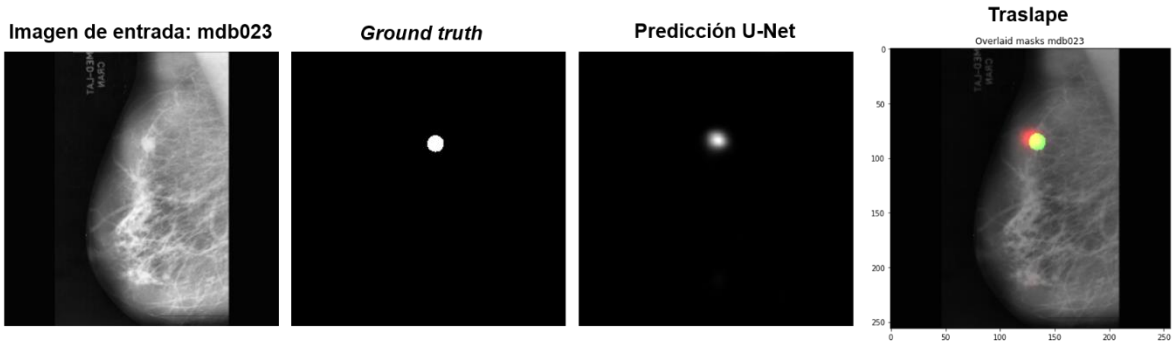


Figura 58. Resultado de detección con U-Net para la mamografía mdb023.

Al analizar todas las predicciones del conjunto de *test*, se observó que el modelo

detecta correctamente la mayoría de a normalidades en mamografías anormales.

Se realizaron pruebas con mamografías normales y en varias detectaba anormalidades, razón por la cual, dentro de este método de clasificación y detección, primero se clasifican como normales o anormales con U-Net.

Se recomienda observar los demás resultados en el Anexo 2, ya que algunas detecciones llevan a pensar que siguen la forma de la anormalidad como tal, e incluso que detectan otras que podrían ser y que no fueron etiquetadas.

Tabla 15. Intersección entre máscaras *ground truth* y detectadas por la red U-Net.

	GT_MASK	PRED_MASK	INTERSECTION
0	mdb012_gt	mdb012_predicted	1.0
1	mdb021_gt	mdb021_predicted	1.0
2	mdb023_gt	mdb023_predicted	1.0
3	mdb058_gt	mdb058_predicted	1.0
4	mdb081_gt	mdb081_predicted	1.0
5	mdb102_gt	mdb102_predicted	1.0
6	mdb148_gt	mdb148_predicted	1.0
7	mdb165_gt	mdb165_predicted	1.0
8	mdb184_gt	mdb184_predicted	1.0
9	mdb188_gt	mdb188_predicted	1.0
10	mdb193_gt	mdb193_predicted	1.0
11	mdb206_gt	mdb206_predicted	1.0
12	mdb218_gt	mdb218_predicted	1.0
13	mdb267_gt	mdb267_predicted	1.0
14	mdb290_gt	mdb290_predicted	1.0
15	mdb314_gt	mdb314_predicted	0.0

Average intersection: 0.9375

En la Tabla 15, se observa que en la última imagen del conjunto de *test* hubo intersección entre las máscaras, es decir, no se detectó la anomalía correctamente.

Pero se destaca que un solo error entre 16 imágenes de *test*, para una media de éxito de 0,9375 es un resultado sobresaliente.

5.5. Prueba método de clasificación y detección de anormalidades a partir de extracción de características, MLP, SVM y ventana deslizante

5.5.1. Descripción

Se ejecutaron pruebas de clasificación con las regiones extraídas establecidas para *test.* También con las imágenes completas del conjunto de *test* inicial (ver organigramas de carpetas en el apartado de la solución correspondiente) para clasificar y marcar las regiones anormales con el algoritmo de ventana deslizante.

5.5.2. Procedimiento

Inicialmente, se evaluará el resultado de los dos clasificadores (MLP y SVM) clasificando las regiones del conjunto de *test* a partir de las características mencionadas del histograma y de textura.

Con el clasificador que mejor rendimiento muestre entre el MLP y la SVM, se correrá el algoritmo de ventana deslizante, extrayendo y clasificando cada región de 64 x 64 en pasos de 8 durante su recorrido, marcando las regiones anormales con cuadros rojos (de 64 x 64).

5.5.3. Resultados y análisis

- **Resultados método de clasificación con perceptrón multicapa (con extracción de características de regiones de 64 x 64)**

  A continuación, en las Figura 59 y 60, se muestran la matriz de confusión y la curva ROC, respectivamente, resultantes de la clasificación de regiones

del conjunto de *test* con el perceptrón multicapa entrenado, a partir de las características extraídas:

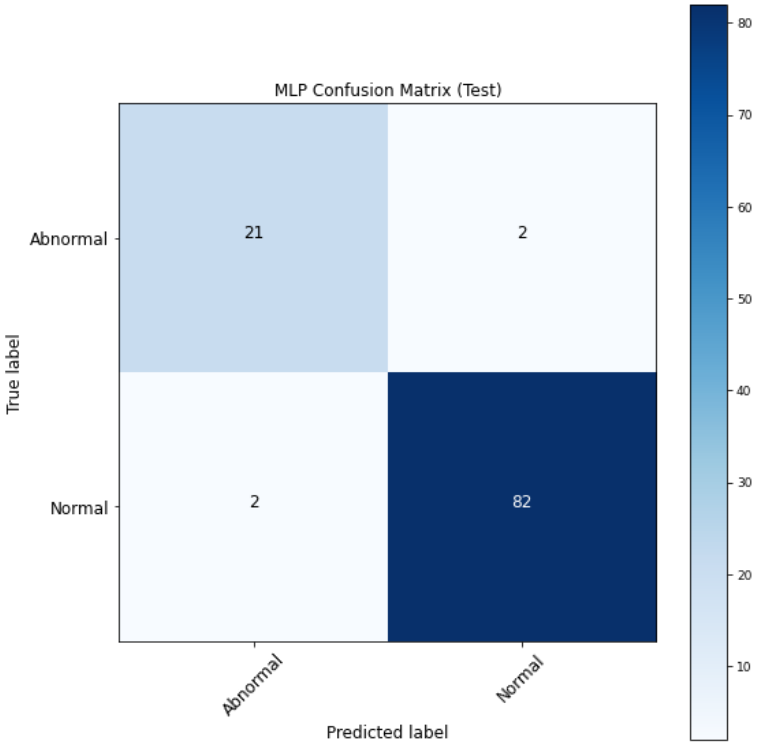


Figura 59. Matriz de confusión de clasificación de regiones de *test* con MLP.

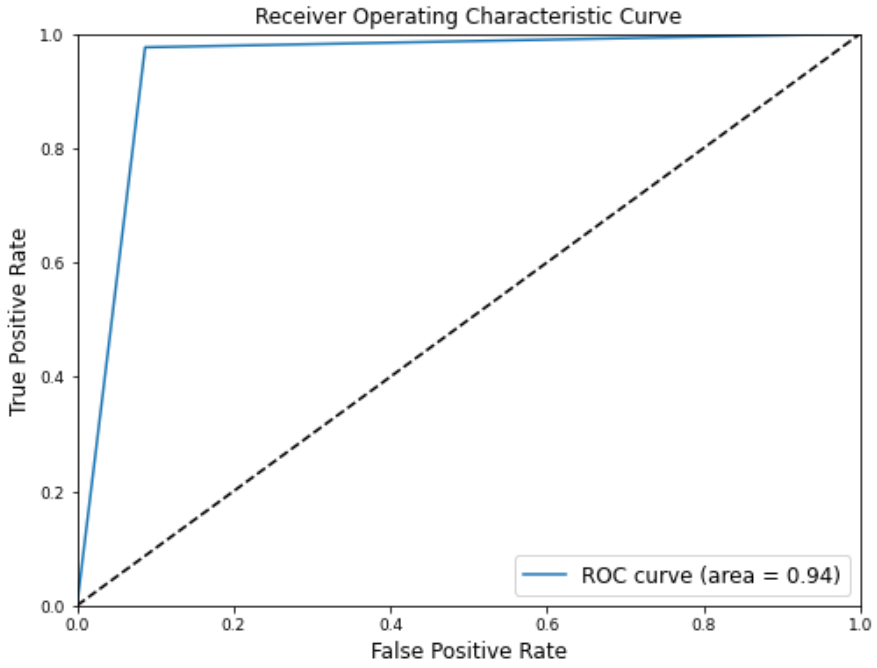


Figura 60. Curva ROC y AUC de clasificación de regiones de *test* con MLP.

- **Exactitud MLP (*Test*):** 0,962
- **Precisión MLP (*Test*):** 0,976
- **Exhaustividad MLP (*Test*):** 0,976
- **Valor F1 MLP (*Test*):** 0,976

Se obtuvo una exactitud de 96,2% y un AUC de la curva ROC de 0,94. Solo se observan 2 falsos positivos y 2 falsos negativos. Un resultado satisfactorio para el conjunto de *test*.

- **Resultados método de clasificación con SVM (con extracción de características manuales a partir de patches)**

En las Figuras 61 y 62, que se exponen seguidamente, se muestran la matriz de confusión y la curva ROC, respectivamente, de la clasificación de regiones del conjunto de *test* con la máquina de soporte vectorial entrenada:

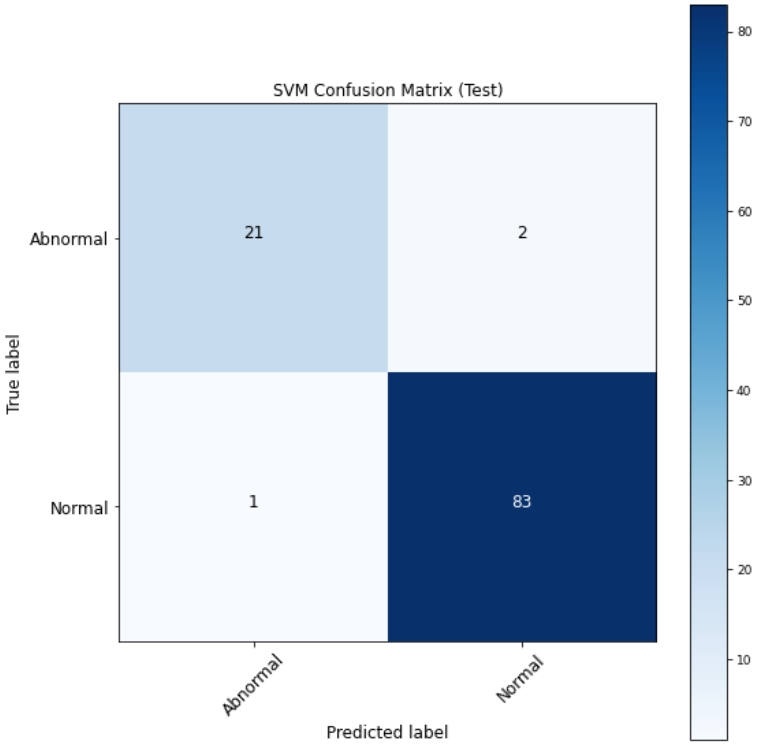


Figura 61. Matriz de confusión clasificación patches de *test* con SVM.

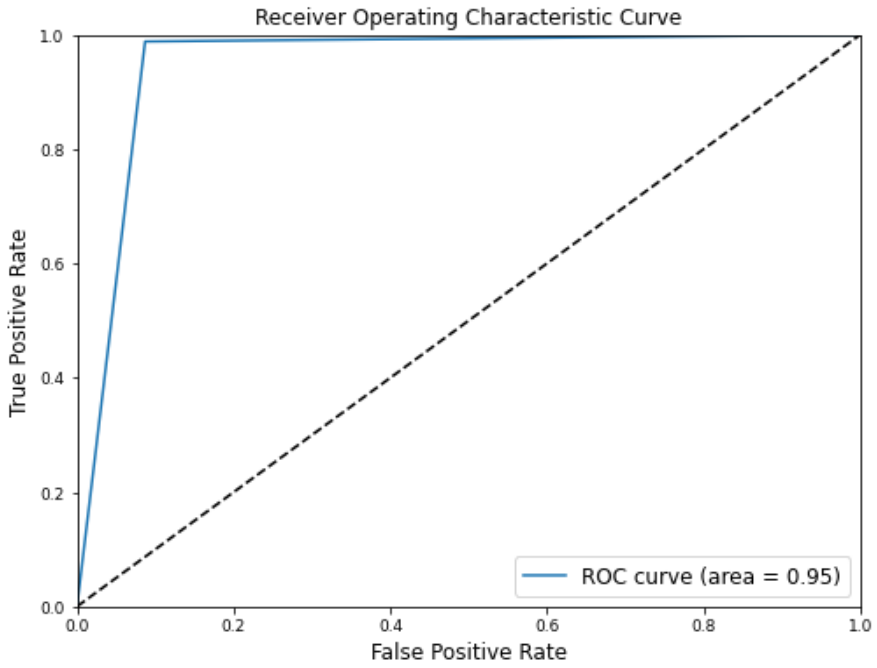


Figura 62. Curva ROC y AUC de clasificación de *patches* de *test* con SVM.

- **Exactitud SVM (*Test*):** 0,971
- **Precisión SVM (*Test*):** 0,976
- **Exhaustividad SVM (*Test*):** 0,988
- **Valor F1 SVM (*Test*):** 0,982

Se obtuvo una exactitud de 97,1% y un AUC de la curva ROC de 0,95. Solo se observan 1 falso positivo y 2 falsos negativos. Lo cual representa un resultado sobresaliente en la clasificación del conjunto de *test*. Se observa que la máquina de soporte vectorial presentó mejor rendimiento que el perceptrón multicapa.

- **Detección con ventana deslizante:**

Las entradas en este caso, fueron imágenes completas del conjunto de *test* segmentadas con el método 1 de segmentación de la mama: a partir de operaciones morfológicas y otras. Por la manera como fue implementado el algoritmo de ventana

deslizante y por su lógica de funcionamiento, su costo computacional es alto, debido a esto, se realizaron pruebas con pocas imágenes del conjunto de *test* elegidas aleatoriamente.

Para mejorar la exactitud en la detección de anormalidades con este método, se establecieron distintos umbrales de rechazo, específicamente, se evaluó si la región presente en la ventana deslizante pertenece a la clase anormal y además, si sobrepasa el umbral de rechazo establecido, para estas pruebas se establecieron dos umbrales de rechazo (0,5 y 0,6), que se comparan con el resultado de pertenencia a la clase por parte del clasificador SVM que, como se mencionó anteriormente, fue el escogido para la realización de estas pruebas de detección con ventana deslizante. A continuación, en las Figuras 63 y 64, se presenta la detección de anormalidades en dos imágenes del conjunto de *test*:

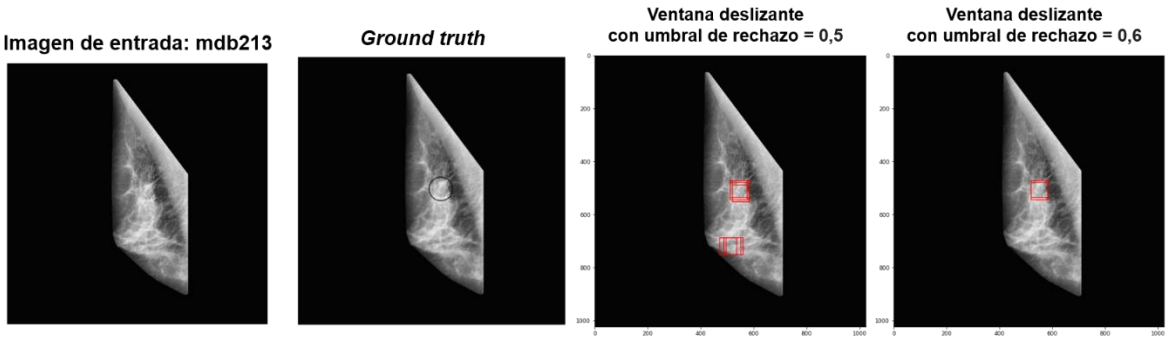


Figura 63. Detección con ventana deslizante a distintos umbrales de rechazo (imagen mdb213).

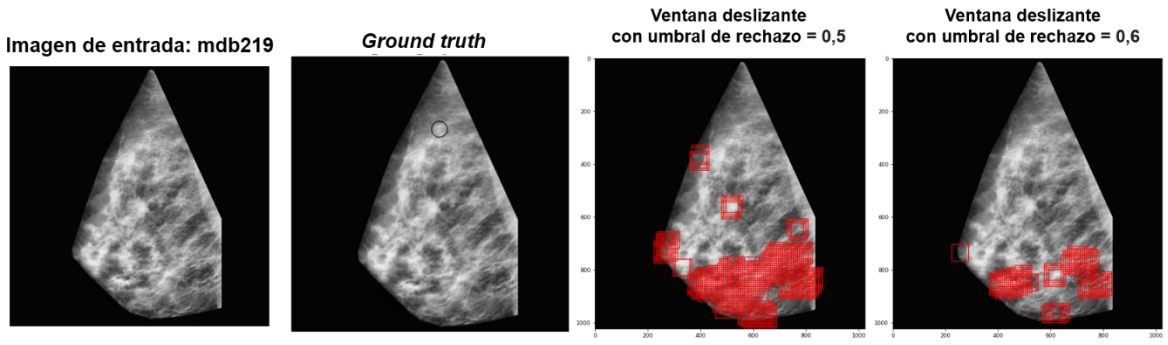


Figura 64. Detección con ventana deslizante a distintos umbrales de rechazo (imagen mdb219).

En la Figura 62 se observa una correcta detección de la anormalidad, sin presencia de falsos positivos, cuando se establece un umbral de rechazo de 0,6. Al disminuir el umbral de rechazo a 0,5 se observa la presencia de falsos positivos.

En la Figura 63, se observan falsos positivos para ambos umbrales de rechazo, sin detectar la anormalidad verdadera.

Se ilustró un resultado correcto (Figura 63) y otro incorrecto (Figura 64). Este patrón de resultados se repite en otras imágenes del conjunto de *test*, por lo que solo se muestran estas dos pruebas en el presente documento y se sugieren mejoras complementarias a este procedimiento en el apartado de trabajos futuros.

- **Información adicional**

Este método también se implementó creando regiones sin redimensionar, con un tamaño de ventana mínimo de 64 x 64. Para las anormalidades con dimensiones de 64 x 64 o menos, se extrajeron regiones de 64 x 64 con centro en la ubicación de la anormalidad. Para las anormalidades que superaban esas dimensiones, se extrajeron subregiones de 64 x 64 con ventana deslizante en pasos de 8 (horizontal y verticalmente).

Con esto se generó una gran cantidad de regiones. Después de separar las regiones obtenidas, para entrenamiento se establecieron 10.041 regiones de la clase anormal y 10.041 regiones de la clase normal. Para pruebas, 1.771 regiones pertenecientes a la clase anormal y 1.771 regiones pertenecientes a la clase normal.

Por lo anterior, no hubo necesidad de realizar aumento de datos, pues se crearon las regiones normales aleatorias necesarias para equilibrar las clases. Este método podría trabajarse más a fondo en el futuro, no se plasmó en el presente documento porque los resultados no fueron superiores a los expuestos (redimensionando las regiones).

6. Conclusiones generales

- Se seleccionaron y extrajeron características de las mamografías digitales presentes en el mini-MIAS *dataset*, que permitieron clasificar el tejido mamario entre normal o anormal.

- Se utilizaron dos métodos distintos de segmentación para segmentar el área de interés, que en este caso era la mama, porque no se contaba con etiquetado de las anormalidades con su forma, para la segmentación de las anormalidades propiamente dichas, se extrajo la información de los metadatos del *mias_dataframe*, los cuales brindaban las coordenadas X, Y y el radio de las mismas, pero no su contorno exacto.

- Se implementó una técnica de aprendizaje profundo (arquitectura DenseNet121) para clasificar las mamografías digitales como normales o anormales.

- Se utilizó una técnica de aprendizaje profundo netamente convolucional (U-Net) para detectar las anormalidades presentes en mamografías clasificadas previamente como anormales con DenseNet121.

- Se implementaron técnicas de visión artificial para extraer características estadísticas del histograma y características de textura de regiones pequeñas de las mamografías digitales.

- Se utilizaron dos técnicas de aprendizaje superficial (perceptrón multicapa y máquina de soporte vectorial) para clasificar las pequeñas regiones extraídas de las imágenes mamográficas, como normales o anormales.

- Se desarrolló un algoritmo de ventana deslizante para recorrer las mamografías digitales y clasificar región por región de 64 x 64, como normal

o anormal. Esto con la máquina de soporte vectorial, que tuvo mejor rendimiento que el perceptrón multicapa. Las regiones o *patches* clasificados como anormales, se marcan con un recuadro rojo. Este método no trabajó bien para todas las imágenes

- La segmentación a partir de operaciones morfológicas requiere implementar diversas pruebas de ensayo y error para lograr un resultado óptimo para la mayoría de las imágenes. De igual manera, se recomienda utilizar porque su costo computacional no es muy alto.

- La segmentación a partir de la red convolucional U-Net tiene un costo computacional mayor pero los resultados de segmentación fueron mejores que con la técnica de aprendizaje superficial.

- Aunque la segmentación con la red convolucional U-Net fue mejor y arrojó un índice de Jaccard mejor, se obtuvieron mejores resultados de la predicción de mamografías normales y anormales a partir de la segmentación 1, esto porque eliminaba información que no aportaba para la clasificación. Quiere decir que no siempre la mejor segmentación y más compleja, permite obtener los mejores resultados de clasificación.

- Para la ubicación de las anormalidades con redes convolucionales se requirió primero clasificar las mamografías como normales o anormales con la arquitectura DenseNet121 y luego localizarlas con la arquitectura U-Net, esto debido a que la arquitectura U-Net siempre busca detectar anormalidades partiendo de las máscaras con que fue entrenada, incluso, en mamografías normales.

- La combinación de técnicas y arquitecturas de máquinas de aprendizaje permite contar con distintas opciones para el cumplimiento del objetivo principal planteado.

7. Trabajos futuros

Con base en el trabajo desarrollado y los resultados obtenidos, se plantean las siguientes perspectivas de trabajo:

1. A partir del sistema desarrollado, se propone implementar una interfaz de usuario para la clasificación de mamografías digitales.

2. Para poder utilizar este sistema en el campo real de aplicación, se requiere la construcción de un *dataset* con una cantidad de imágenes mucho mayor, etiquetadas correctamente por expertos.

3. Con un banco de datos más robusto, se podría entrenar el sistema para predecir el tipo de anormalidad presente y, además, clasificarla como benigna o maligna.

4. Adicional a la detección y clasificación de mamografías digitales, se podría entrenar un sistema con una base de datos de tratamientos para cada anormalidad, de manera que se muestre al usuario la información correspondiente a esta, y a su tratamiento.

5. El algoritmo de ventana deslizante para clasificar regiones de las imágenes, podría mejorarse aplicando técnicas como “comité” de pesos de las regiones y/o pirámide de imágenes para deslizar la ventana en la misma imagen en distintas resoluciones. Esto para mejorar su rendimiento, ya que este no trabajó de manera óptima para varias imágenes de *test*.

6. Existe la posibilidad de implementar funciones de análisis de datos sobre los resultados de diagnóstico del sistema. Estas funciones podrían relacionar el diagnóstico con diferentes factores, como grupo étnico, edad, ubicación geográfica, peso, altura, etc. y obtener información muy valiosa

para los médicos especialistas.

7. Si se implementa un sistema como este, se podría hacer uso de la transferencia de aprendizaje para realizar un mejoramiento continuo del modelo de clasificación, para mejorar cada vez más su rendimiento a partir de las nuevas muestras que se vayan introduciendo.

Referencias

[1] “Breast cancer.” https://www.who.int/news-room/fact-sheets/detail/breast-cancer (accessed May 20, 2022).

[2] Ferlay J *et al.*, “Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer.,” 2020. https://gco.iarc.fr/today (accessed May 22, 2022).

[3] O. Ginsburg *et al.*, “Breast cancer early detection: A phased approach to implementation,” *Cancer*, vol. 126, pp. 2379–2393, May 2020, doi: 10.1002/CNCR.32887.

[4] C. E. Ventura-Alfaro, “Errores de medición en la interpretación mamográfica por radiólogos,” *Revista de Salud Pública*, vol. 20, no. 4, pp. 518–522, Jul. 2018, doi: 10.15446/rsap.v20n4.52035.

[5] S. Leon, L. Brateman, J. Honeyman-Buck, and J. Marshall, “Comparison of two commercial CAD systems for digital mammography.,” *J Digit Imaging*, vol. 22, no. 4, pp. 421–3, Aug. 2009, doi: 10.1007/s10278-008-9144-x.

[6] “Image Analytics | Breast Image Analysis | Hologic.” https://www.hologic.com/hologic-products/breast-skeletal/image-analytics (accessed Mar. 06, 2019).

[7] “SecondLook for 2D Mammography.” https://icadmed.com/secondlook.html (accessed Mar. 06, 2019).

[8] “Digital Mammography Services | Imaging Center System.” https://www.whhs.com/Services/Specialized-Programs/Outpatient-Imaging-Center/Digital-Mammography-Services.aspx (accessed Mar. 06, 2019).

[9] “CAD Gives a Boost to Breast Imaging | Imaging Technology News.” https://www.itnonline.com/article/cad-gives-boost-breast-imaging (accessed Mar. 06, 2019).

[10] “R2 Image Checker | Baptist Health Louisville, KY.” https://www.baptisthealth.com/louisville/pages/services/imaging-and-diagnostics/screenings-and-radiology-services/mammography/r2-image-checker.aspx (accessed Mar. 06, 2019).

[11] W. Xie, Y. Li, and Y. Ma, “Breast mass classification in digital mammography based on extreme learning machine,” *Neurocomputing*, vol. 173, pp. 930–941, Jan. 2016, doi: 10.1016/J.NEUCOM.2015.08.048.

[12] F. Pak, H. R. Kanan, and A. Alikhassi, “Breast cancer detection and classification in digital mammography based on Non-Subsampled Contourlet Transform (NSCT) and Super Resolution,” *Comput Methods Programs Biomed*, vol. 122, no. 2, pp. 89–107, Nov. 2015, doi: 10.1016/J.CMPB.2015.06.009.

[13] J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, and M. A. Guevara Lopez, “Representation learning for mammography mass lesion classification with convolutional neural networks,” 2016, doi: 10.1016/j.cmpb.2015.12.014.

[14] D. por Ing Rodrigo Javier Herrera García, “Detección de microcalcificaciones mamarias agrupadas.” Accessed: Jul. 30, 2019. [Online]. Available: https://pdfs.semanticscholar.org/b470/bff7a49a639988b48f610638f1585da79d19.pdf

[15] R. Chaieb and K. Kalti, “Feature subset selection for classification of malignant and benign breast masses in digital mammography,” *Pattern Analysis and Applications*, pp. 1–27, Nov. 2018, doi: 10.1007/s10044-018-0760-x.

[16] H. Zhou, Y. Zaninovich, and C. Gregory, “Mammogram Classification Using Convolutional Neural Networks.” Accessed: Apr. 24, 2019. [Online]. Available: https://ehnree.github.io/documents/papers/mammogram_conv_net.pdf

[17] F. F. Ting, Y. J. Tan, and K. S. Sim, “Convolutional neural network improvement for breast cancer classification,” *Expert Syst Appl*, vol. 120, pp. 103–115, Apr. 2019, doi: 10.1016/J.ESWA.2018.11.008.

[18] K. P. McGuire, “Breast Anatomy and Physiology,” *Breast Disease: Diagnosis and Pathology*, vol. 1, pp. 1–14, Jan. 2015, doi: 10.1007/978-3-319-22843-3_1.

[19] Instituto Nacional del Cáncer de los Institutos Nacionales de la Salud de EE.UU., “Definición de mama - Diccionario de cáncer - National Cancer Institute.” https://www.cancer.gov/espanol/publicaciones/diccionario/def/mama

(accessed Jul. 04, 2019).

[20] Asociación Española Contra el Cáncer, “Anatomía de la mama: ¿Qué es el Cáncer de Mama? | AECC.” https://www.aecc.es/es/todo-sobre-cancer/tipos-cancer/cancer-mama/que-es-cancer-mama (accessed Jul. 04, 2019).

[21] American Cancer Society, “What Is Breast Cancer? | Breast Cancer Definition.” https://www.cancer.org/cancer/breast-cancer/about/what-is-breast-cancer.html (accessed Jul. 04, 2019).

[22] EE. UU. Instituto Nacional del Cáncer, “Diccionario de cáncer - National Cancer Institute.” https://www.cancer.gov/espanol/publicaciones/diccionario/buscar?contains=false&q=cancer+de+mama (accessed Jul. 04, 2019).

[23] A. Prat, B. Ádamo, A. Rodríguez, and M. Muñoz, “¿Qué es el Cáncer de Mama? | PortalCLÍNIC.org.” Accessed: Jul. 04, 2019. [Online]. Available: https://portal.hospitalclinic.org/enfermedades/cancer-de-mama/definicion

[24] “The mini-MIAS database of mammograms.” http://peipa.essex.ac.uk/info/mias.html (accessed May 23, 2022).

[25] Université de Caen - Oncoprof, “Mammography.” https://oncoprof.net/Generale2000/g04_Diagnostic/Mammographie/gb04_mm01.html (accessed Jul. 04, 2019).

[26] EE. UU. Instituto Nacional de Cáncer, “Significado de cambios y afecciones de los senos - National Cancer Institute.” https://www.cancer.gov/espanol/tipos/seno/significado-cambios-en-los-senos (accessed Jul. 04, 2019).

[27] D. Abdelhafiz, C. Yang, R. Ammar, and S. Nabavi, “Deep convolutional neural networks for mammography: Advances, challenges and applications,” *BMC Bioinformatics*, vol. 20, Jun. 2019, doi: 10.1186/S12859-019-2823-4.

[28] M. N. Murty and V. S. Devi, “Introduction to Pattern Recognition and Machine Learning,” vol. 5, Jun. 2015, doi: 10.1142/8037.

[29] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, vol. Third Edition.

[30] S. Kassir, I. Hassoon, and D. Riyadh, “Brain tumor localization and extraction algorithm in MRI images,” 2019.

[31] R. M. Haralick, I. Dinstein, and K. Shanmugam, “Textural Features for Image Classification,” *IEEE Trans Syst Man Cybern*, vol. SMC-3, no. 6, pp. 610–621, 1973, doi: 10.1109/TSMC.1973.4309314.

[32] “Cooccurrence Matrix - an overview | ScienceDirect Topics.” https://www.sciencedirect.com/topics/engineering/cooccurrence-matrix (accessed Oct. 23, 2022).

[33] “Module: feature — skimage v0.19.2 docs.” https://scikit-image.org/docs/stable/api/skimage.feature.html#skimage.feature.graycoprops (accessed Oct. 23, 2022).

[34] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Trans Pattern Anal Mach Intell*, vol. 24, no. 7, pp. 971–987, Jul. 2002, doi: 10.1109/TPAMI.2002.1017623.

[35] “Local Binary Pattern for texture classification — skimage v0.19.2 docs.” https://scikit-image.org/docs/stable/auto_examples/features_detection/plot_local_binary_pattern.html (accessed Oct. 30, 2022).

[36] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[37] A. Artasanchez and P. Joshi, *Artificial Intelligence with Python - Second Edition*.

[38] C. C. Aggarwal, “Neural Networks and Deep Learning.”

[39] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Nov. 2017, vol. 2017-January, pp. 2261–2269. doi: 10.1109/CVPR.2017.243.

[40] Q. Ji, J. Huang, W. He, and Y. Sun, “Optimized deep convolutional neural networks for identification of macular diseases from optical coherence tomography images,” *Algorithms*, vol. 12, no. 3, 2019, doi: 10.3390/a12030051.

[41] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for

biomedical image segmentation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015, vol. 9351, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[42] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. Elsevier Inc., 2012. doi: 10.1016/C2009-0-61819-5.

[43] "Clasificación: Curva ROC y AUC | Machine Learning | Google Developers." https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc (accessed Nov. 17, 2022).

[44] "3.7.13 Documentation." https://docs.python.org/release/3.7.13/ (accessed May 25, 2022).

[45] "Te damos la bienvenida a Colaboratory - Colaboratory." https://colab.research.google.com/ (accessed May 25, 2022).

[46] "pandas - Python Data Analysis Library." https://pandas.pydata.org/about/ (accessed May 25, 2022).

[47] "NumPy documentation — NumPy v1.22 Manual." https://numpy.org/doc/stable/ (accessed May 25, 2022).

[48] "Matplotlib — Visualization with Python." https://matplotlib.org/ (accessed May 25, 2022).

[49] "seaborn: statistical data visualization — seaborn 0.11.2 documentation." https://seaborn.pydata.org/ (accessed May 25, 2022).

[50] "OpenCV: OpenCV-Python Tutorials." https://docs.opencv.org/4.1.2/d6/d00/tutorial_py_root.html (accessed May 25, 2022).

[51] "scikit-learn: machine learning in Python — scikit-learn 1.1.1 documentation." https://scikit-learn.org/stable/ (accessed May 25, 2022).

[52] "TensorFlow." https://www.tensorflow.org/ (accessed May 25, 2022).

[53] "About Keras." https://keras.io/about/ (accessed May 25, 2022).

[54] "R2019a - Actualizaciones de las familias de productos MATLAB y Simulink - MATLAB & Simulink." https://la.mathworks.com/products/new_products/release2019a.html

(accessed May 25, 2022).

[55] Chaima Derouiche and Akram Boukhamla, “Pectoral Muscle Boundary detection using digital mammograms,” 2017, Accessed: Oct. 06, 2019. [Online]. Available: https://www.researchgate.net/publication/321600175_Pectoral_Muscle_Boundary_detection_using_digital_mammograms

[56] “LabelMe. The Open annotation tool.” http://labelme.csail.mit.edu/Release3.0/ (accessed Oct. 30, 2022).

[57] “ImageNet.” https://www.image-net.org/ (accessed Oct. 30, 2022).

[58] C. Szegedy *et al.*, “Going deeper with convolutions.”

8. Anexos

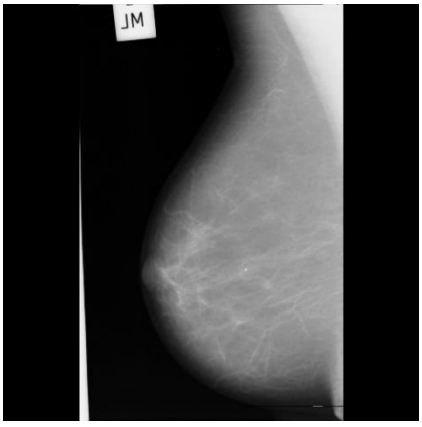
8.1. Anexo 1: Descripción detallada del método 1 de segmentación de la mama: a partir de operaciones morfológicas, entropía y otras

Se desarrolló una función que se aplica iterativamente a todas las imágenes del mini-MIAS *dataset*. Para efectos ilustrativos, seguidamente se muestra el procedimiento para una imagen aleatoria del *dataset* (mdb009).

Este proceso se realizó heurísticamente, es decir, se analizaron los resultados de las operaciones en la mayoría de las imágenes, con base en estos, se definió el orden de las operaciones, elementos estructurantes y demás. Esto se aclara porque en la imagen ejemplo algunas operaciones podrían parecer no surtir efecto alguno.

A continuación, se exponen los pasos desarrollados para este método de segmentación:

1. Se lee el archivo de imagen PGM para convertirlo en un arreglo (1.024 x 1.024) y poder manipularlo:



2. Se crea una máscara binaria de la imagen (*thresholding*):

3. Por las características de los artefactos extraños que se quieren eliminar inicialmente (etiquetas con letras), se crea un elemento estructurante morfológico de forma line de longitud 20 y ángulo 135° y se aplica la operación morfológica de cierre (*closing*):

15×15 **logical** array

1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

4. Se crea un elemento estructurante morfológico de forma *disk* de radio 1 y se aplica la operación morfológica de erosión (*erosion*):

```
3×3 logical array

     0     1     0
     1     1     1
     0     1     0
```

5. Se crea un elemento estructurante morfológico de forma line de longitud 5 y ángulo 45° y se aplica nuevamente la operación morfológica de erosión (*erosion*):

```
3×3 logical array

     0     0     1
     0     1     0
     1     0     0
```

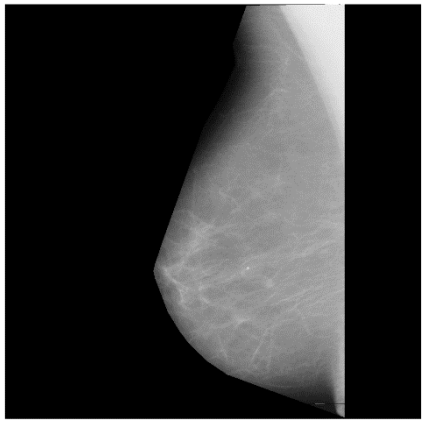
6. Se rellenan los huecos de la imagen binaria con la función *imfill* de MATLAB. “En su sintaxis, un hueco es un conjunto de pixeles de fondo que no se puede alcanzar rellenando el fondo desde el borde de la imagen”:

7. Se detecta el área mayor entre los objetos separados presentes en la imagen binaria. No se aplica desde el principio porque las operaciones morfológicas anteriores, buscan, entre otras cosas, separar la mama de los artefactos extraños porque en algunas mamografías se juntan:

8. Se calcula la envolvente convexa de los puntos en la imagen binaria:

9. Se calcula la envolvente convexa de los puntos en la imagen binaria:

10. Se aplica la máscara binaria a la imagen original:



11. Se encuentran los índices y valores de elementos distintos a cero. devuelve un vector que contiene los índices lineales de cada elemento distinto de cero.

12. Se hallan el mínimo y máximo de filas arrojadas en el paso anterior. Y mínimo y máximo de columnas arrojadas en el paso anterior.

13. Se crea un nuevo arreglo que vaya de los límites mínimo a máximo de filas y mínimo a máximo de columnas, es decir, se recorta la imagen para luego proceder a detectar la orientación de la mama:

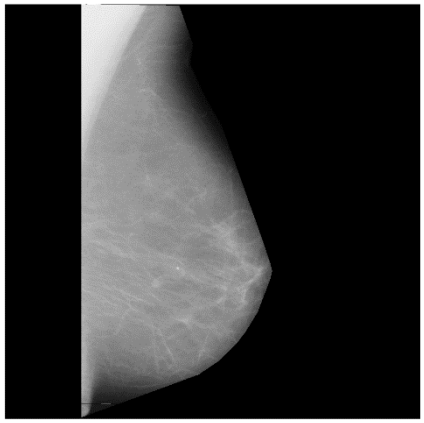
14. Se hallan las dimensiones del arreglo obtenido. En este caso:

- Número de filas = 1021
- Número de columnas = 471

15. Para la localización del tejido pectoral, se busca orientar las mamografías hacia la derecha, es decir, que el tejido pectoral siempre quede situado a la izquierda de la imagen, por lo cual, se suman las todas filas de las primeras 20 columnas y se suman las todas filas de las últimas 20 columnas y se comparan esos dos valores.

16. Si la suma de los valores de intensidad de las primeras 20 columnas es mayor que la suma de intensidad de las últimas 20 columnas, es porque la imagen está orientada hacia la derecha. Del modo contrario, la mamografía en la imagen estará orientada hacia la izquierda. Esta posición se almacena en una variable.

17. Si la mama está orientada hacia la izquierda, se rota la imagen del paso 10 para orientarla hacia la derecha. Como en el caso de la imagen ejemplo (mdb009):



18. A este arreglo resultante, se le calculan los índices lineales de cada valor de intensidad distinto de cero y se ubican en un vector columna.

19. Se crea un nuevo vector columna con los valores presentes en los índices entregados por el vector columna anterior.

20. Se calcula la entropía del arreglo.

21. Se establece un umbral. Si la entropía la entropía es alta (imagen con más variaciones) se establece umbral más bajo, si es más plana, se establece umbral más alto. La idea es que donde esté más desordenado es porque no hay pectoral, o sea que las intensidades deben ser algo bajas, por eso se pone el umbral más bajo.

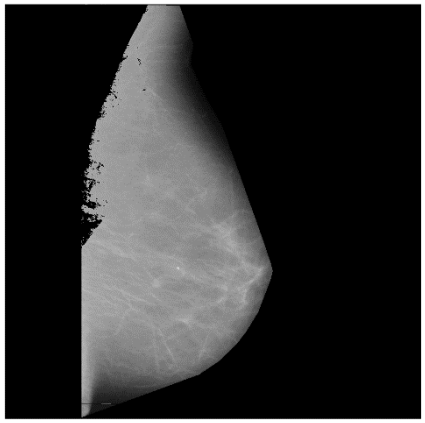
22. Se recorre cada valor del arreglo (pixel de la imagen) y se calcula:
    a. Su distancia euclidiana hasta el origen (0, 0)
    b. Y su distancia euclidiana hasta un punto ubicado el 60% de las filas y el 80% de las columnas, partiendo desde el origen (0,0).

23. Se evalúa si el valor de intensidad del pixel es mayor al del umbral de intensidad establecido gracias a la entropía de la imagen y menor al valor máximo de intensidad (255 para este caso de imágenes con profundidad de

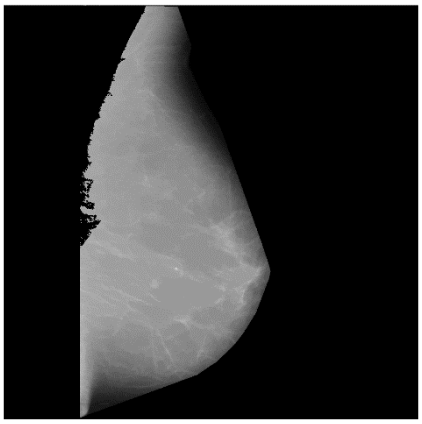
8 bits).

24. Si lo anterior se cumple, se evalúa si la distancia euclidiana a. es menor que la distancia euclidiana b., es decir, que el pixel en cuestión no ha sobrepasado la diagonal trazada por el 60% de filas y 80% de columnas

25. Si ambas condiciones se cumplen, se considera que el pixel correspondiente hace parte del tejido pectoral y se iguala a cero, es decir, deja de ser visible en la imagen.



26. Se rellenan los huecos de la imagen binaria con la función *imfill* de MATLAB. “En su sintaxis, un hueco es un conjunto de pixeles de fondo que no se puede alcanzar rellenando el fondo desde el borde de la imagen”:



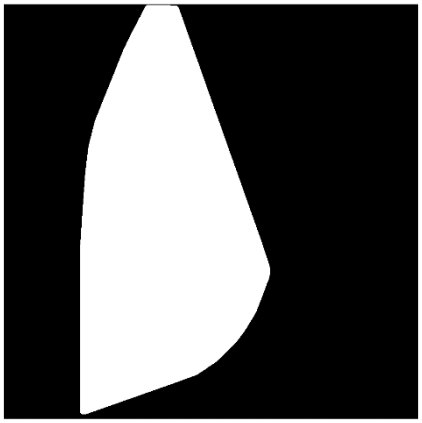
27. Se crea una máscara binaria de la imagen (*thresholding*):

28. Se crea un elemento estructurante morfológico de forma disk de radio 10 y se aplica la operación morfológica de apertura (*opening*):

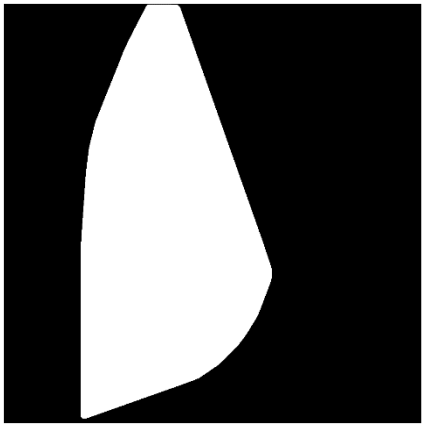
```
19×19 logical array

   0   0   0   0   1   1   1   1   1   1   1   1   1   1   1   0   0   0   0
   0   0   0   1   1   1   1   1   1   1   1   1   1   1   1   1   0   0   0
   0   0   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   0   0
   0   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   0
   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
   0   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   0
   0   0   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   0   0
   0   0   0   1   1   1   1   1   1   1   1   1   1   1   1   1   0   0   0
   0   0   0   0   1   1   1   1   1   1   1   1   1   1   1   0   0   0   0
```

Con esta operación se logra terminar de separar algunos artefactos extraños y/o etiquetas que continuaban muy cercanos a la región de interés de la imagen. En la actual, no se logra apreciar.

29. Se calcula la envolvente convexa:

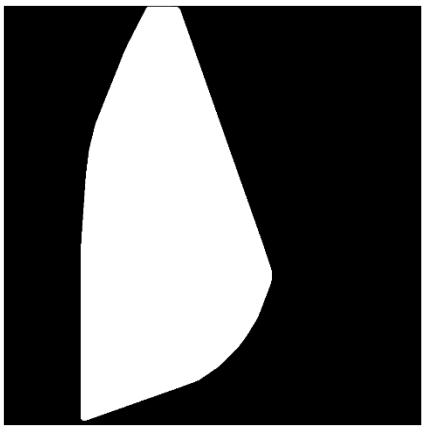


30. Se aplica la operación morfológica de cierre (*closing*) con el mismo elemento estructurante de la operación morfológica anterior (paso 28):

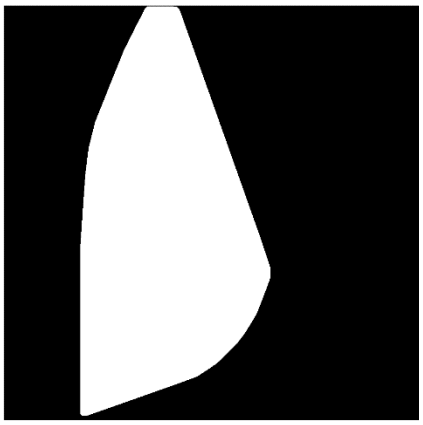


En la imagen ejemplo (mdb009) no se aprecia ningún cambio con esta operación, pero se incluye porque en otras logra cerrar correctamente la máscara luego de la operación *opening* anterior (paso 28).

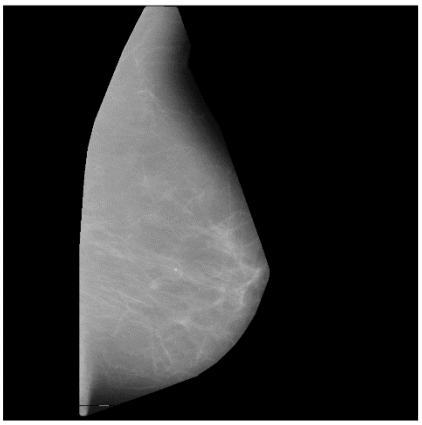
31. Se detecta el objeto de mayor área entre los objetos separados presentes en la imagen binaria:



32. Se calcula la envolvente convexa:

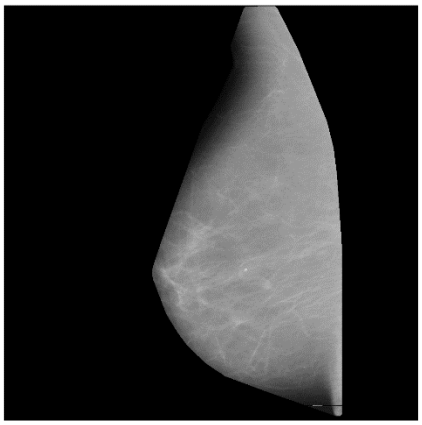


33. Se aplica la máscara binaria a la imagen:

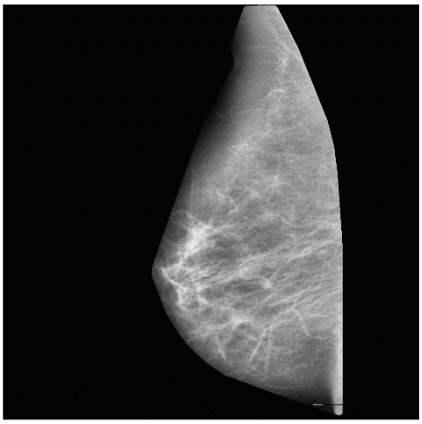


34. Se verifica la variable que indica la orientación original de la imagen (paso 16), si originalmente la imagen correspondiente estaba orientada hacia la

izquierda, se hace *flip* horizontalmente. Como en el caso de esta imagen ejemplo (mdb009):



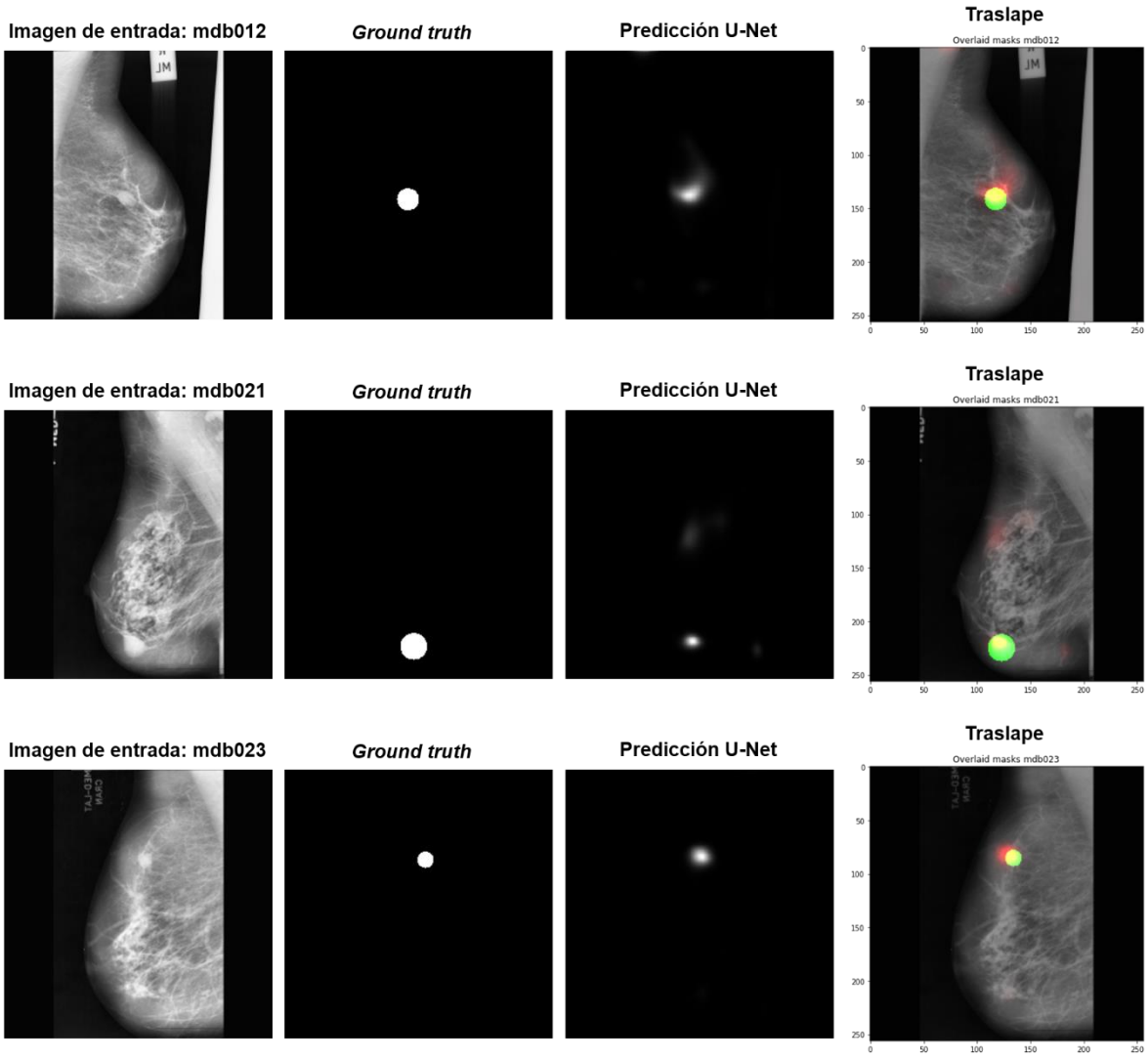
35. Se aplica la ecualización del histograma adaptativo (CLAHE):



Como se mencionó anteriormente, las imágenes de entrada estaban en formato PGM originalmente, después de aplicar este proceso de segmentación a todas las imágenes, estas se almacenan en formato PNG. Al igual que las imágenes generadas en el aumento de datos. En otras palabras, el formato PNG es el formato con el cual se trabaja en todas las etapas de este proyecto.

8.2. Anexo 2: Resultados detección de anormalidades con U-Net para el conjunto de *test* (16 imágenes):

A continuación, se presentan los resultados de la detección de anormalidades con U-Net para el conjunto de prueba, donde se observa para cada una: la imagen de entrada, la salida que se debería obtener (*ground truth*), la predicción realizada por U-Net y el traslape entre las 3 imágenes para observar la detección sobre la mamografía y su intersección con el *ground truth* correspondiente.



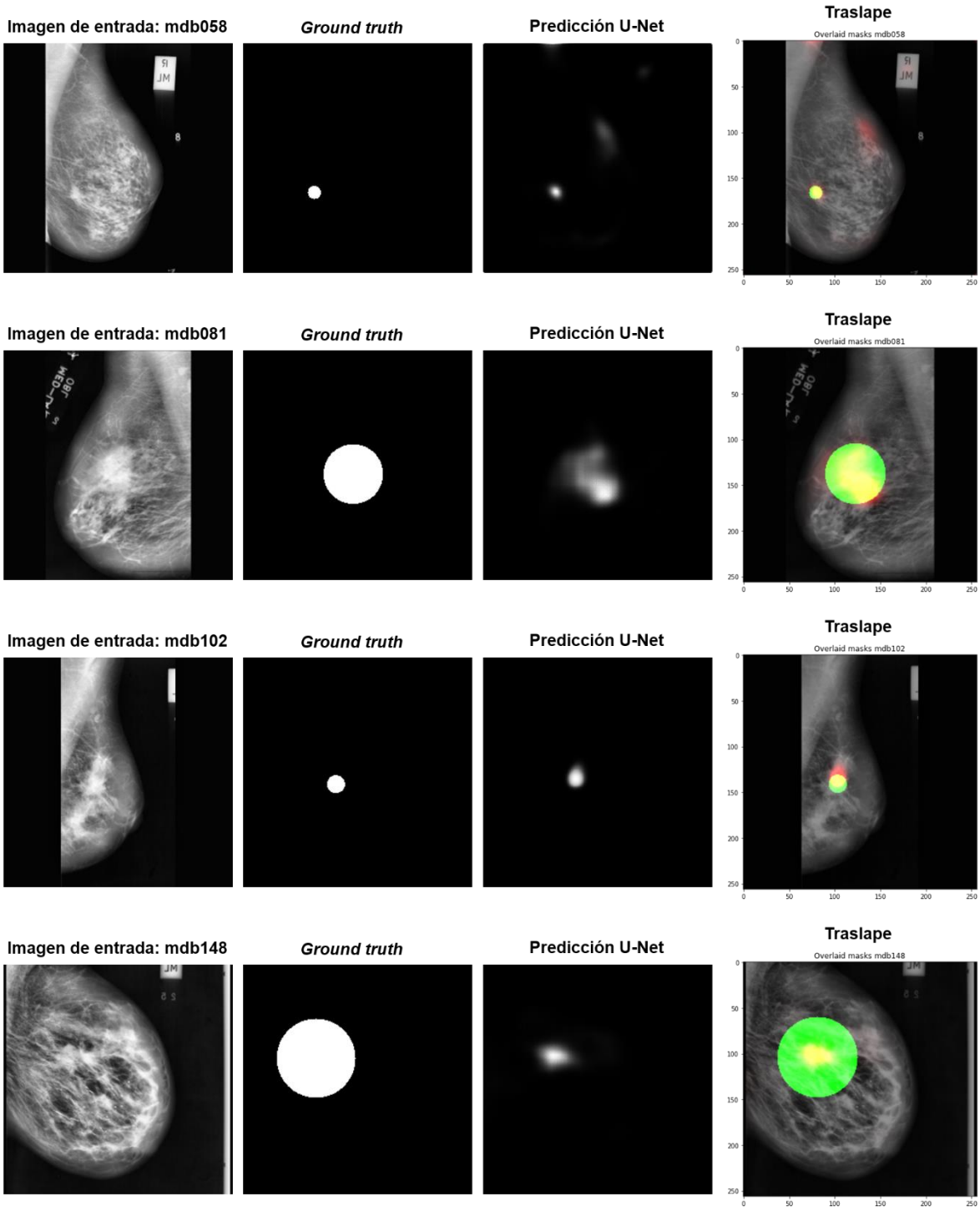


Imagen de entrada: mdb058
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb058
Imagen de entrada: mdb081
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb081
Imagen de entrada: mdb102
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb102
Imagen de entrada: mdb148
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb148

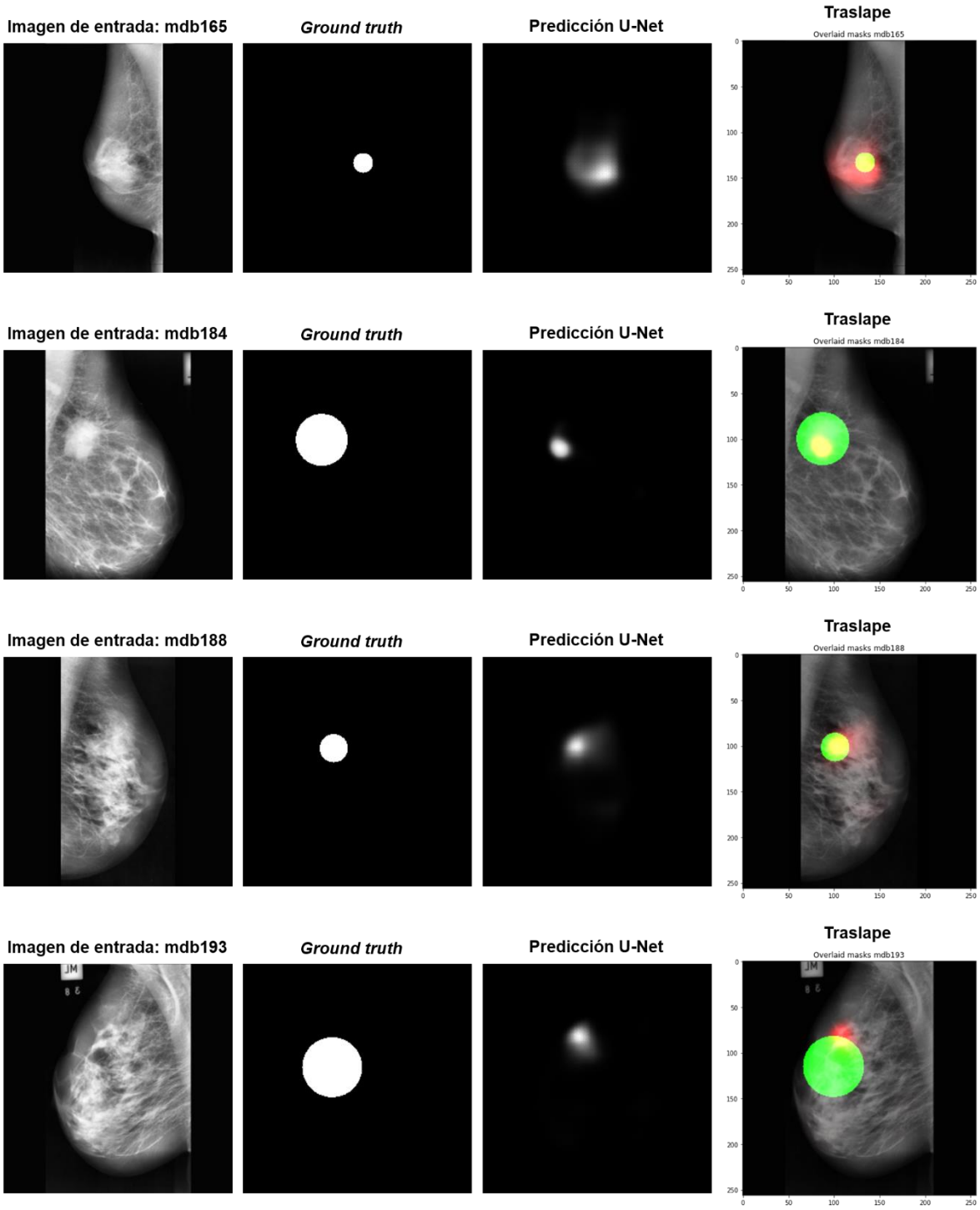


Imagen de entrada: mdb165
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb165
Imagen de entrada: mdb184
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb184
Imagen de entrada: mdb188
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb188
Imagen de entrada: mdb193
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb193

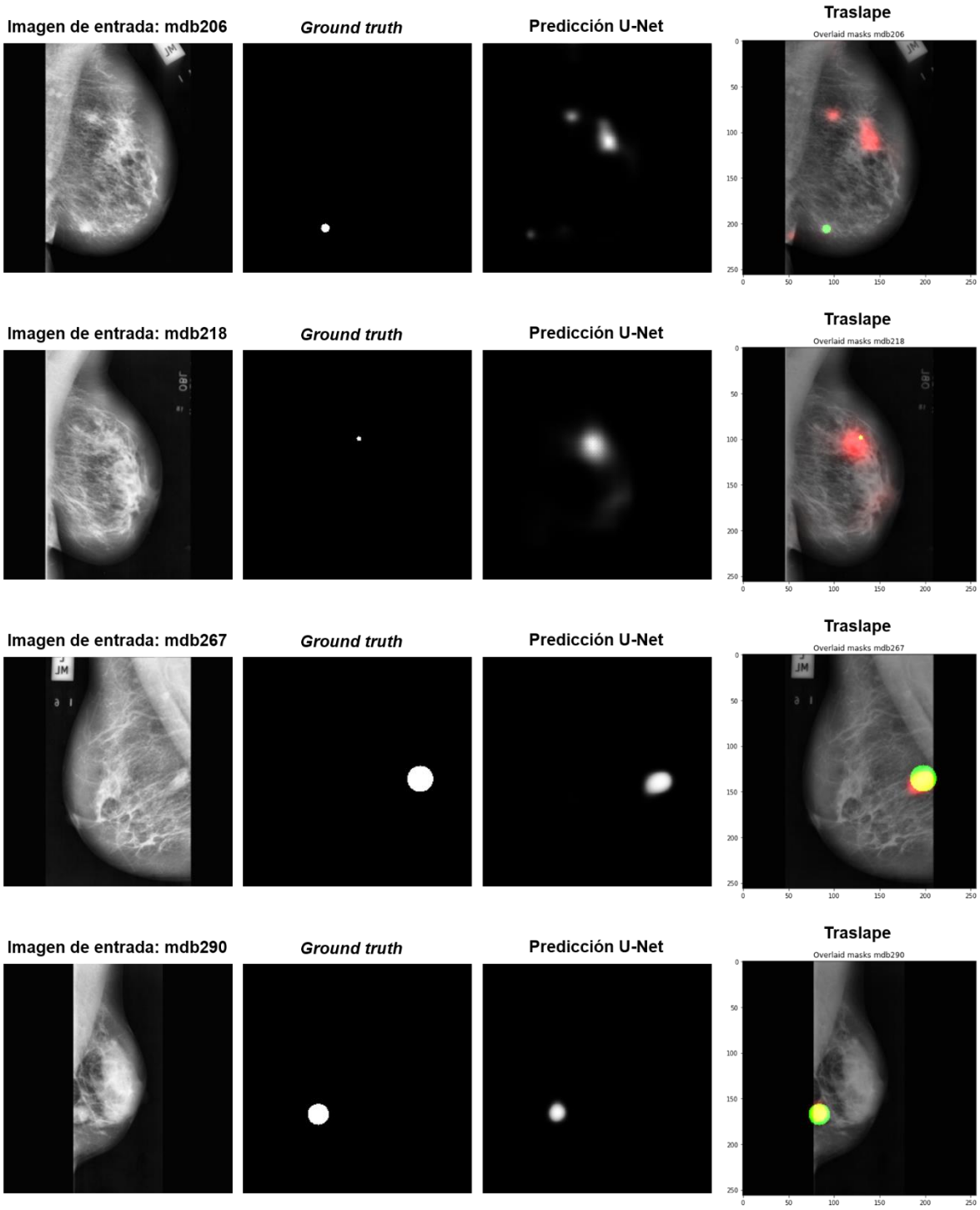


Imagen de entrada: mdb206
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb206
Imagen de entrada: mdb218
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb218
Imagen de entrada: mdb267
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb267
Imagen de entrada: mdb290
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb290

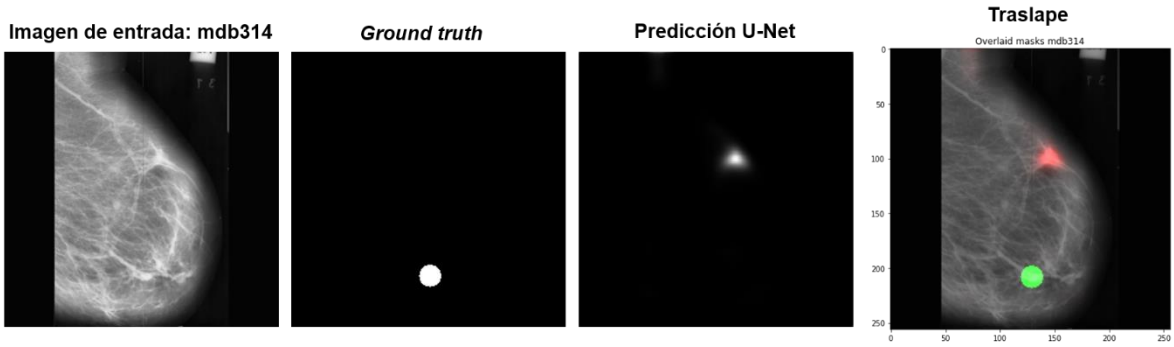


Imagen de entrada: mdb314
Ground truth
Predicción U-Net
Traslape
Overlaid masks mdb314
0
50
100
150
200
250