



Detección de cáncer en imágenes de mamografía usando aprendizaje profundo

Mario Andrés Mejía Cortés

**Facultad de Ingeniería
Programa de Ingeniería Electrónica
Ibagué, 2023**



Detección de masas en imágenes de mamografía por medio de técnicas de aprendizaje profundo

Mario Andrés Mejía Cortés

Trabajo de grado que se presenta como requisito parcial para optar al título de:
Ingeniero electrónico

Director (a):

PhD. Manuel Guillermo Forero Vargas

Universidad de Ibagué

Facultad de Ingeniería

Programa de Ingeniería Electrónica

Ibagué, 2023



A mi familia y amigos, por su apoyo y paciencia durante todo el proceso. Gracias por estar ahí he impulsarme a culminar este proceso.

Agradecimientos

Un especial agradecimiento al director de tesis Doctor Ing. Manuel Guillermo Forero Vargas y al semillero de investigación LUN



Resumen

El cáncer de mama es una enfermedad que afecta a millones de mujeres en todo el mundo y se ha convertido en una de las principales causas de muerte en esta población. Detectar a tiempo esta enfermedad es crucial para mejorar las tasas de supervivencia; sin embargo, los médicos se enfrentan a desafíos significativos en su evaluación precisa, ya que las imágenes mamográficas pueden presentar características sutiles y variables que dificultan la detección manual de masas malignas o benignas. En este contexto, se presenta este proyecto, con el objetivo de implementar y evaluar la eficacia de las redes neuronales en la segmentación y clasificación precisa de masas mamográficas, lo que ha arrojado resultados con una precisión del 85 por ciento en la segmentación de masas y un 75 por ciento en la segmentación de estas según su severidad (maligno o benigno).

Palabras clave: Cáncer de mama, imágenes mamográficas, redes neuronales, segmentación

Abstract

Breast cancer is an illness that affects millions of women around the world and has become one of the main causes of death in this population. Detecting this disease in a timely manner is crucial to improve survival rates; however, doctors face significant challenges in its accurate assessment, as mammographic images can exhibit subtle and variable characteristics that make manual detection of malignant or benign masses difficult. In this context, this project is presented, with the aim of implementing and evaluating the effectiveness of neural networks in the segmentation and accurate classification of mammographic masses, which has yielded results with an 80 percent accuracy in mass segmentation and a 70 percent accuracy in the segmentation of these according to their severity (malignant or benign).

Keywords: Breast cancer, mammographic images, neural networks, segmentation

Contenido

Introducción	1
Capítulo 1: Objeto de estudio.....	2
1.1 Planteamiento del problema.....	2
1.2 Objetivos	3
1.2.1 Objetivo general.....	3
1.2.2 Objetivos específicos	3
1.3 Estado del arte	4
Capítulo 2: Marco Teórico.....	5
2.1 Aprendizaje de maquina.....	5
2.2 Aprendizaje supervisado	5
2.3 Perceptrón.....	6
2.4 Red neural.....	6
2.4.1 Capa de entrada	7
2.4.2 Capa oculta	7
2.4.3 Capa de salida.....	7
2.5 Aprendizaje profundo	8
2.6 Redes neurales convolucionales	8
2.6.1 Capas de convolución.....	9
2.6.2 Capas de agrupación (pooling)	9
2.6.3 Capas completamente conectadas	9
2.6.4 Funciones de activación	9
2.7 Arquitectura de aprendizaje.....	10
2.8 UNET	10
2.9 Maxpooling.....	11
2.10 Sobreajuste	11
2.11 Regularización	11
2.12 Dropout	11
2.13 Batch Normalization	12
2.14 Data Augmentation.....	12
2.15 Exactitud	13
2.16 Intersección sobre la unión (IoU).....	13
Capítulo 3: Materiales	14
3.1 Base de datos CBIS-DDSM	14
3.2 Software y Hardware	15
Capítulo 4: Desarrollo	17
3.3 Preparación y muestreo de la base de datos.....	17
3.4 Ajuste de la red	19
3.5 Clasificación	29
Capítulo 5: Resultados.....	31
6. Conclusiones y recomendaciones.....	33
6.1 Conclusiones.....	33
6.2 Recomendaciones.....	34
Referencias bibliográficas	35



Lista de figuras

Figura 1. Estructura de un perceptron	6
Figura 2. Estructura red neuronal	7
Figura 3. Estructura de una red neural convolucional	8
Figura 4. Estructura de red UNET	10
Figura 5. Ejemplo de data augmentation en imágenes	12
Figura 6. Información estadística base de daros CBIS-DDMS.....	14
Figura 7. Caso base de datos a) mamografía. b) máscaras de segmentación individuales para cada lesión	17
Figura 8. Caso base de datos. a) mamografía. b) máscara de segmentación con lesiones unificadas	17
Figura 9. Cuadro de selección alrededor de la lesión	18
Figura 10. Muestra obtenida	18
Figura 11. Modelo red UNET clásica	20
Figura 12. Curvas de aprendizaje entrenamiento UNET clásica. a) exactitud, b) precisión	21
Figura 13. Modelo red UNET con capas de batch normalization	22
Figura 14. Curvas de aprendizaje con batch normalization a) exactitud, b) perdida	23
Figura 15. Modelo red UNET con batch normalization y dropout en codificador	24
Figura 16. Curvas de aprendizaje con batch normalization y dropout en codificador a) exactitud, b) perdida	25
Figura 17. Modelo red UNET con batch normalization y dropout en decodificador	25
Figura 18. Curvas de aprendizaje con batch normalization y dropout en decodificador a) exactitud, b) perdida	26
Figura 19. Modelo red UNET con batch normalization y dropout en codificador y decodificador	27
Figura 20. Curvas de aprendizaje con batch normalization, dropout en codificador y decodificador a) exactitud, b) perdida	28
Figura 21. Curvas de aprendizaje clasificación a) exactitud, b) perdida	30

Lista de tablas

Tabla 1. Tabla comparativa de resultados	31
--	----



Introducción

El cáncer de mama representa una carga significativa para la salud global, como lo señaló la Organización Mundial de la Salud[1] en el año 2020, con 2.3 millones de diagnósticos y 685,000 muertes a nivel mundial. Un factor que resaltar es que la mitad de estos casos se presentan en mujeres sin factores de riesgo específicos, a excepción de su edad y género. Además, el cáncer de mama tiende a ser asintomático en sus etapas tempranas, lo que complica enormemente su detección precoz.

Una de las herramientas principales para la prevención y detección temprana del cáncer de mama son las mamografías, exámenes que deben realizarse periódicamente. No obstante, los médicos se enfrentan a desafíos significativos en la evaluación precisa de estas imágenes, ya que pueden mostrar características sutiles y variables que dificultan la detección manual de esta afectación [2].

Es en este contexto que surge la necesidad de desarrollar proyectos que estudien técnicas que permitan asistir a los médicos en su tarea de detectar esta enfermedad. En este proyecto, se explorarán técnicas de detección computarizada de masas en imágenes de mamografía mediante el uso de redes neuronales convolucionales. Estas redes tienen el potencial de mejorar la precisión y eficiencia en la identificación de masas mamarias, lo que podría llevar a una detección temprana más efectiva y, en última instancia, a un mejor pronóstico para las pacientes con cáncer de mama.

Capítulo 1: Objeto de estudio

1.1 Planteamiento del problema.

El cáncer de mama, una enfermedad que afecta a millones de mujeres en todo el mundo, tiene un impacto significativo en la salud y la calidad de vida. Según la Organización Mundial de la Salud (OMS)[1], esta enfermedad se origina en las células del revestimiento de los conductos o lóbulos del tejido glandular mamario. En sus primeras etapas, suele ser asintomático y localizado en el conducto o lóbulo, con un potencial de diseminación mínimo. Sin embargo, con el tiempo, puede progresar a una etapa invasiva, donde invade el tejido mamario circundante y se propaga a ganglios linfáticos cercanos u otros órganos, proceso conocido como metástasis, lo que lo vuelve altamente mortal.

El tratamiento del cáncer de mama puede ser efectivo, especialmente cuando se detecta tempranamente. La combinación de cirugía, radioterapia y medicación, incluyendo terapias hormonales, quimioterapia y terapias biológicas dirigidas, puede ser crucial para frenar la progresión y diseminación del cáncer, salvando vidas en el proceso[3].

Sin embargo, el mayor desafío en el tratamiento del cáncer de mama reside en la detección temprana y precisa[4]. Dado que en sus primeras etapas suele ser asintomático y difícil de identificar mediante métodos tradicionales como la autoexploración o evaluación clínica, se necesitan herramientas de diagnóstico más avanzadas.

Las mamografías, aunque fundamentales para la prevención, pueden resultar complejas de interpretar debido a la sutileza y variabilidad en las imágenes, lo que dificulta la detección manual de masas malignas o benignas[2]. La imperiosa necesidad de una detección temprana y precisa del cáncer de mama es evidente, ya que puede marcar la diferencia en el pronóstico y tratamiento.

Por lo tanto, surge un desafío crucial: desarrollar un enfoque más efectivo y preciso para la detección temprana del cáncer de mama, especialmente en sus etapas iniciales, cuando las posibilidades de tratamiento exitoso son más altas. Este proyecto se propone abordar esta problemática mediante el uso de técnicas de detección computarizada de masas en imágenes de mamografía, aprovechando el potencial de las redes neuronales convolucionales. La pregunta central que guio esta investigación es cómo estas técnicas pueden ser empleadas para mejorar la detección temprana y precisión en la identificación de masas mamarias, lo que podría conducir a un diagnóstico y tratamiento más efectivos, y en última instancia, salvar vidas en la lucha contra el cáncer de mama.

1.2 Objetivos

1.2.1 Objetivo general

Implementar y comparar técnicas de aprendizaje profundo para la detección de cáncer en imágenes de mamografía.

1.2.2 Objetivos específicos

- Recopilar el estado de arte sobre el desarrollo de métodos basados en aprendizaje profundo para la detección de cáncer en imágenes de mamografía.
- Analizar las muestras con el fin de explorar la posibilidad de mejorar la calidad y determinar si tienen características particulares que permitan aumentar los índices de la identificación del cáncer.
- Implementar arquitecturas neuronales para la detección de cáncer de mama en imágenes de mamografía usando la base de datos MINIS-DSM.
- Comparar los resultados obtenidos para identificar el mejor resultado posible en cuanto a carga computacional y precisión de las redes aplicadas.

1.3 Estado del arte

El cáncer de mama, de acuerdo con la Organización Mundial de la Salud (OMS), se ha establecido como la forma más prevalente de cáncer a nivel mundial, con más de 2.2 millones de casos registrados en el año 2020. Esta cifra es motivo de grave preocupación, ya que indica que aproximadamente una de cada 12 mujeres enfrentará esta devastadora enfermedad en algún punto de sus vidas. Sin embargo, la situación más alarmante reside en que el cáncer de mama es una de las principales causas de mortalidad entre las mujeres, cobrándose la vida de alrededor de 685,000 mujeres en el mismo año [1].

La imperante necesidad de detectar esta enfermedad en sus primeras etapas, cuando las posibilidades de tratamiento exitoso son más altas, ha impulsado la investigación y desarrollo de técnicas avanzadas[5]. En este contexto, las redes neuronales, en particular las redes convolucionales, han emergido como herramientas prometedoras para auxiliar en el diagnóstico y detección temprana del cáncer de mama.

Un estudio relevante en este campo, realizado por Cantone *et al*[6]., evaluó el rendimiento de 33 modelos utilizando la base de datos OMI-DB. Entre estos modelos, la red EfficientNet-B3 se destacó, logrando una precisión del 85.2%.

Además, el trabajo de Shen *et al*[7]. merece atención, ya que empleó redes completamente convolucionales para analizar la base de datos CBIS-DDSM. En este caso, lograron una precisión de 89% en una red Resnet50 pre entrenada.

En el estudio de Singh *et al*[8]., se propuso una red adversativa inversa condicional. Esta red alcanzó una intersección sobre la unión del 87.03%, siendo un resultado similar a los estudios anteriores.

Capítulo 2: Marco Teórico

El capítulo actual se sumerge en el marco teórico requerido para una comprensión exhaustiva de los principios, técnicas y aplicaciones del aprendizaje automático en el contexto de la tesis. A medida que se avanza en este capítulo, se exploran las bases conceptuales y teóricas que respaldan la utilización de algoritmos de aprendizaje automático.

2.1 Aprendizaje de máquina

El aprendizaje de máquina (machine learning en inglés) es un subcampo de la inteligencia artificial que se enfoca en el desarrollo de algoritmos y modelos computacionales que permiten a las computadoras aprender y mejorar su rendimiento en tareas específicas a través de la experiencia y sin una programación explícita. En lugar de seguir un conjunto de reglas fijas, las máquinas de aprendizaje pueden analizar datos, identificar patrones y tomar decisiones basadas en la información que se les proporciona. El aprendizaje de máquina se divide principalmente en tres clases principales que son el aprendizaje supervisado, no supervisado y por refuerzo[9].

2.2 Aprendizaje supervisado

El aprendizaje supervisado es una de las categorías fundamentales de técnicas dentro del campo del aprendizaje de máquina. En el aprendizaje supervisado, un algoritmo se entrena utilizando un conjunto de datos etiquetado, lo que significa que cada ejemplo de entrada en el conjunto de datos se asocia con una etiqueta o resultado conocido. El objetivo principal del aprendizaje supervisado es aprender una relación funcional entre las entradas y las etiquetas para poder hacer predicciones precisas sobre nuevos datos no vistos[9].

2.3 Perceptrón

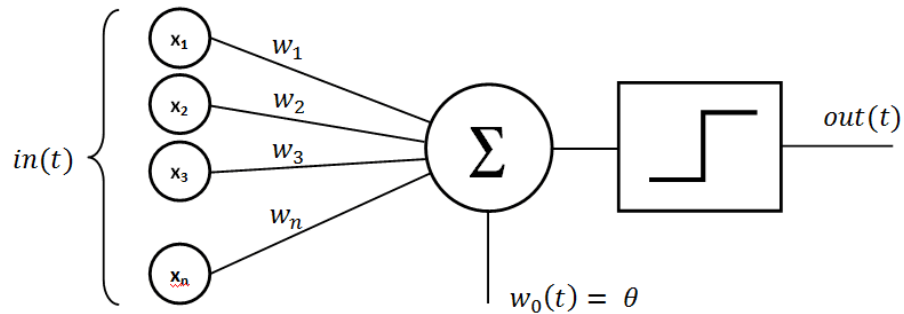


Figura 1. Estructura de un perceptron
Fuente: datascientest [10]

Un perceptrón es un tipo de unidad básica en las redes neuronales artificiales y es un concepto fundamental en el campo del aprendizaje de máquina y la inteligencia artificial. Fue desarrollado por Frank Rosenblatt en 1957 y se considera una de las arquitecturas neuronales más simples. El perceptrón se utiliza principalmente para tareas de clasificación binaria, donde se determina si una entrada pertenece a una de dos clases (por ejemplo, sí/no o spam/no spam).

El perceptrón toma múltiples entradas (características) y asigna un peso a cada una de ellas. Luego, suma las entradas ponderadas y aplica una función de activación para producir una salida. La salida se compara con un umbral (también conocido como umbral de activación), y dependiendo de si la salida supera o no este umbral, se clasifica en una de las dos clases[10].

2.4 Red neural

Una red neuronal es un tipo de proceso de aprendizaje de máquina que utiliza nodos o neuronas interconectados en una estructura de capas que se asemeja al cerebro humano. Está formada por diferentes nodos que funcionan como neuronas y que transmiten señales e información entre sí[11].

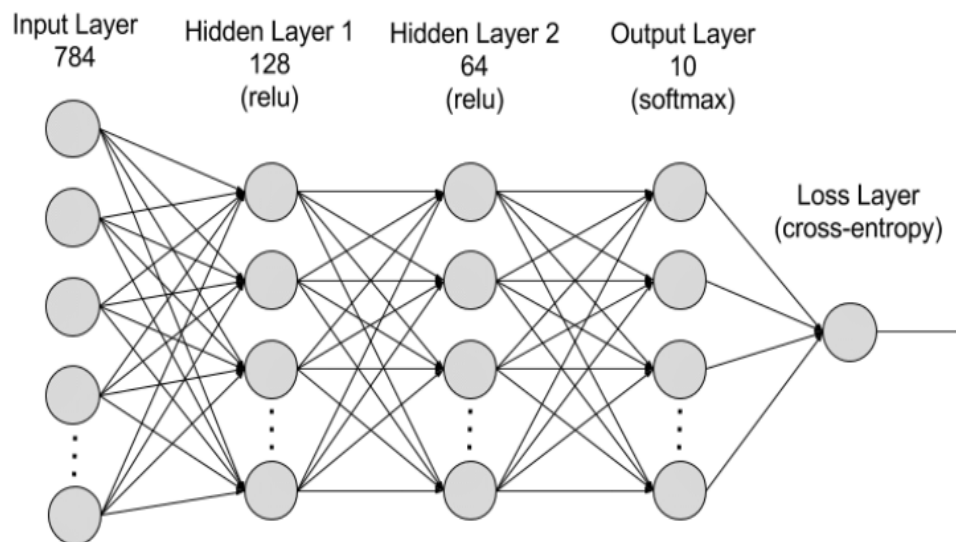


Figura 2. Estructura red neuronal
Fuente: Amazon[11]

Estas redes están compuestas de tres capas principales:

2.4.1 Capa de entrada

Esta capa está compuesta por neuronas que reciben los datos o señales procedentes del entorno. Es la primera capa de la red neuronal y se encarga de recibir los datos de entrada y enviarlos a la siguiente capa.

2.4.2 Capa oculta

Esta capa no tiene conexión directa con el entorno y puede ser precedida por otras capas ocultas o por la capa de entrada. Esta capa es la encargada de procesar los datos de entrada y transformarlos en una forma que pueda ser utilizada por la capa de salida.

2.4.3 Capa de salida

Esta capa está compuesta por neuronas que proporcionan la solución de la red neuronal. Es la última capa de la red neuronal y se encarga de producir la salida de la red neuronal.

2.5 Aprendizaje profundo

El aprendizaje profundo, también conocido como deep learning en inglés, es una rama especializada del campo del aprendizaje de máquina que se enfoca en entrenar redes neuronales con múltiples capas ocultas para abordar tareas complejas de procesamiento y análisis de datos[12].

2.6 Redes neurales convolucionales

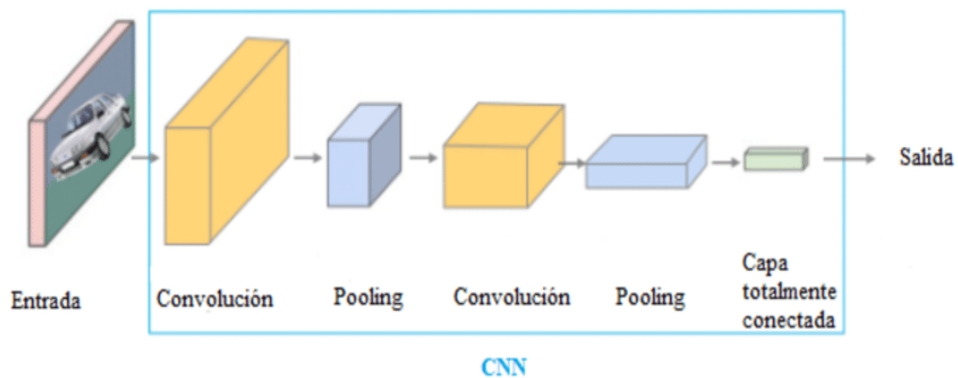


Figura 3. Estructura de una red neural convolucional
Fuente: ResearchGate[13]

La IBM define una Red Neuronal Convolucional, o Convolutional Neural Network (CNN) en inglés, como un tipo de arquitectura de red neuronal diseñada específicamente para el procesamiento de datos bidimensionales, como imágenes y vídeos. Las CNN son ampliamente utilizadas en tareas de visión por computadora, reconocimiento de patrones y procesamiento de imágenes debido a su capacidad para capturar características locales y jerárquicas en los datos [14]. Los elementos principales de estas redes son:

2.6.1 Capas de convolución

Estas capas aplican filtros convolucionales a las regiones locales de la imagen de entrada para extraer características relevantes. Los filtros se deslizan a lo largo de la imagen, calculando productos escalares en cada posición y produciendo mapas de características. Esto permite detectar bordes, texturas y patrones simples en la imagen.

2.6.2 Capas de agrupación (pooling)

Después de la convolución, se aplican capas de agrupación para reducir la dimensionalidad de las características y preservar las características más importantes. La operación de agrupación toma un conjunto de valores cercanos y los reduce a un solo valor representativo, como el máximo o el promedio, para reducir el costo computacional y mejorar la invarianza a pequeñas deformaciones en la imagen.

2.6.3 Capas completamente conectadas

Estas capas son similares a las de las redes neuronales tradicionales y se utilizan para realizar tareas de clasificación o regresión. Toman las características extraídas por las capas anteriores y las utilizan para tomar decisiones finales, como etiquetar una imagen con una clase específica.

2.6.4 Funciones de activación

Al igual que en otras redes neuronales, se aplican funciones de activación (como ReLU) después de las operaciones de convolución y agrupación para introducir no linealidad en la red y permitir la captura de relaciones más complejas en los datos.

2.7 Arquitectura de aprendizaje

La arquitectura de aprendizaje se refiere a la estructura de una red neuronal o modelo de aprendizaje de máquinas. Esta arquitectura se compone de diferentes capas y nodos que procesan los datos de entrada y producen una salida, y puede ser personalizada y optimizada para mejorar la precisión y eficiencia del modelo.

2.8 UNET

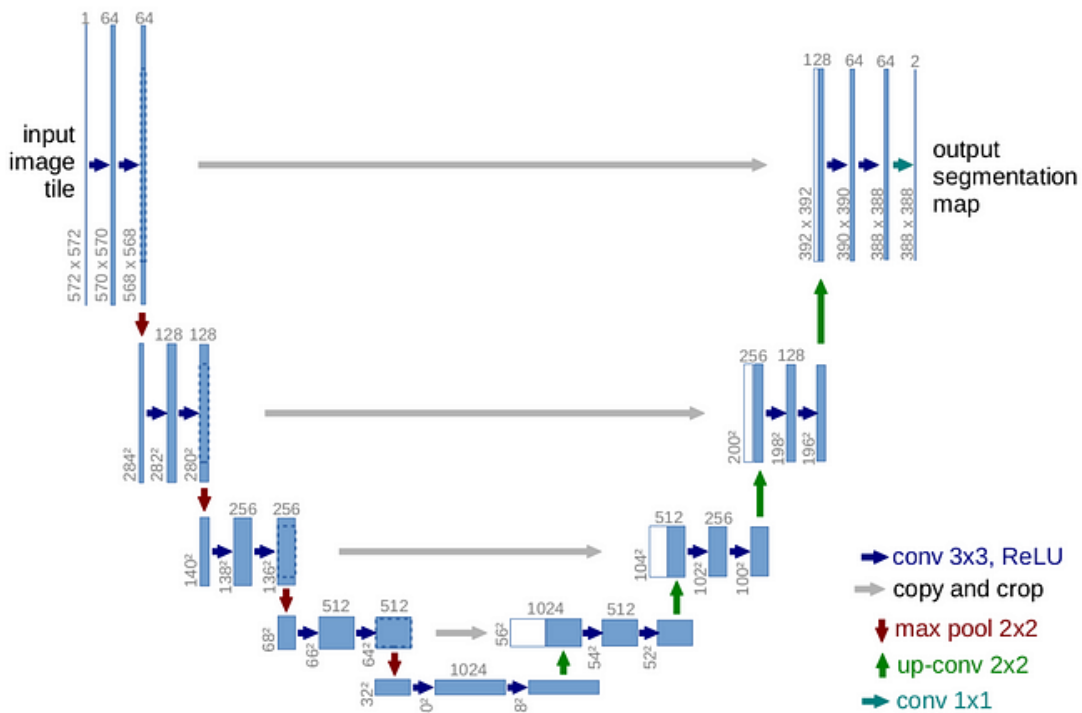


Figura 4. Estructura de red UNET
Fuente: towardsdatascience[9]

La red UNET, propuesta por primera vez por O. Ronneberger, P. Fischer, and T. Brox [15], es una arquitectura de red neuronal convolucional utilizada para la segmentación semántica de imágenes médicas. Esta arquitectura se compone de una estructura en forma de U con capas de convolución y deconvolución, y utiliza la técnica de "skip connections" para mejorar la precisión de la segmentación.

2.9 Maxpooling

Maxpooling es una técnica utilizada en las redes neuronales convolucionales para reducir la dimensionalidad de los datos de entrada y extraer características importantes de las imágenes[16] Esta técnica se realiza aplicando un filtro máximo a regiones no superpuestas de la representación inicial y ayuda a reducir la cantidad de parámetros y cálculos en la red neuronal.

2.10 Sobreajuste

El sobreajuste (overfitting) es un comportamiento no deseado de un modelo de aprendizaje de máquinas utilizado para el modelado predictivo. El sobreajuste ocurre cuando un modelo se ajusta demasiado bien a los datos de entrenamiento y no generaliza bien a nuevos datos[17] Esto puede suceder cuando el modelo es demasiado complejo o cuando se entrena con datos anómalos. El sobreajuste puede ser identificado mediante la comparación del rendimiento del modelo en los datos de entrenamiento y en los datos de prueba. Si el rendimiento del modelo es significativamente mejor en los datos de entrenamiento que en los datos de prueba, es probable que el modelo esté sobre ajustado.

2.11 Regularización

La regularización en el aprendizaje de máquinas es una técnica utilizada para evitar el sobreajuste de un modelo de aprendizaje de máquinas. Esta técnica consiste en añadir un término de penalización a la función de coste del modelo y puede mejorar la capacidad del modelo para generalizar a nuevos datos y reducir la complejidad del modelo[18].

2.12 Dropout

Dropout es una técnica de regularización, propuesta por Srivastava et al[19], utilizada en el aprendizaje de máquinas para evitar el sobreajuste de un modelo. Esta técnica consiste en desactivar aleatoriamente algunas neuronas durante el

entrenamiento de la red neuronal. De esta manera, se evita que las neuronas se especialicen en patrones específicos y se obliga a la red neuronal a aprender características más generales y robustas. El dropout se implementa por capa en una red neuronal y puede ser utilizado con la mayoría de los tipos de capas, como capas densas completamente conectadas, capas convolucionales y capas recurrentes. El dropout no se utiliza en la capa de salida y se puede ajustar mediante un hiperparámetro que indica la probabilidad de desactivar una neurona en una capa determinada. La técnica de dropout ha demostrado ser efectiva en la reducción del sobreajuste y en la mejora del rendimiento de la red neuronal.

2.13 Batch Normalization

Batch Normalization es una técnica de regularización utilizada en el aprendizaje de máquinas para evitar el sobreajuste de un modelo. Esta técnica consiste en normalizar las salidas entre capas de una red neuronal en un formato estándar, lo que ayuda a mejorar la eficiencia y la velocidad de aprendizaje de la red neuronal. Batch Normalization se realiza entre las capas de una red neuronal y se aplica a mini-batches en lugar del conjunto completo de datos. Batch Normalization fue propuesto por Sergey Ioffe y Christian Szegedy [20] en 2015.

2.14 Data Augmentation

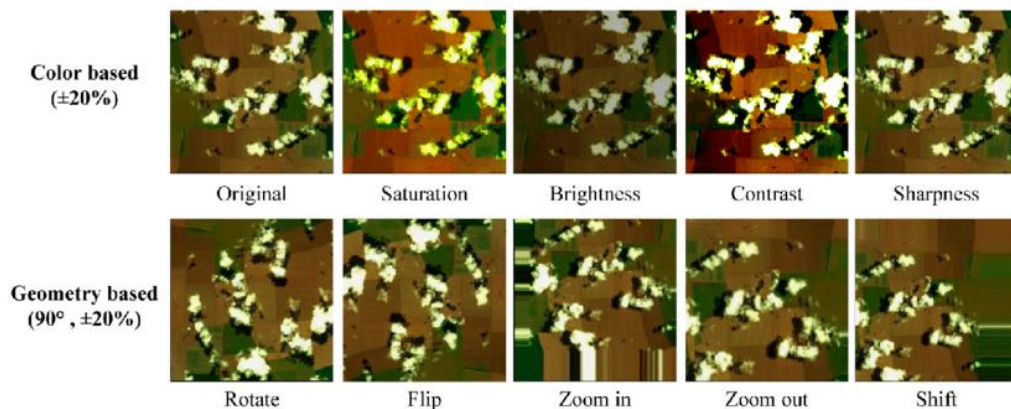


Figura 5. Ejemplo de data augmentation en imágenes
Fuente: researchgate [21]

Data augmentation es una técnica utilizada en el aprendizaje de máquinas para aumentar artificialmente la cantidad de datos de entrenamiento mediante la generación de nuevos puntos de datos a partir de los datos existentes. Esta técnica se utiliza para mejorar la precisión y la eficiencia de los modelos de aprendizaje de máquinas, especialmente cuando el conjunto de datos de entrenamiento es limitado[22].

2.15 Exactitud

La exactitud se define como la capacidad de un modelo para predecir con de manera correcta las etiquetas o valores de un conjunto de datos de prueba. Se puede calcular como el cociente entre el número de predicciones correctas realizadas por el modelo y el número total de registros en el conjunto de prueba. La precisión se encuentra en un rango que va desde 0, en caso de que el modelo no realice ninguna predicción correcta, hasta 1, cuando el modelo acierta en todas las muestras del conjunto de prueba[23].

2.16 Intersección sobre la unión (IoU)

La Intersección sobre la Unión (IoU) se utiliza comúnmente en la evaluación del rendimiento de algoritmos de detección de objetos, especialmente en el contexto de aplicaciones de visión por computadora. Su definición se fundamenta en la relación entre la intersección de dos cajas, una correspondiente a la predicción y otra a la verdad, y la unión de ambas cajas. La IoU proporciona un valor que varía en un rango de 0, cuando no hay intersección entre las dos cajas, hasta 1, cuando ambas cajas son idénticas[23].

Capítulo 3: Materiales

Una de las tareas primordiales en el marco del proyecto consistió en la identificación de una base de datos de acceso público que albergara imágenes de mamografía, acompañadas de sus correspondientes anotaciones. Entre las opciones evaluadas, se identificaron tres candidatas viables: CBIS-DDSM, MIAS e INbreast.

La elección recayó en la base de datos CBIS-DDSM debido a su notoria amplitud, que comprende un total de 1566 casos, en comparación con MIAS, que consta de 161 casos, y INbreast, que alberga 101. Cabe destacar que la base de datos CBIS-DDSM también se destaca por proporcionar anotaciones en forma de máscaras de segmentación, una característica que facilita de manera significativa la manipulación y la alimentación de la información a la red neuronal durante el desarrollo del proyecto.

3.1 Base de datos CBIS-DDSM

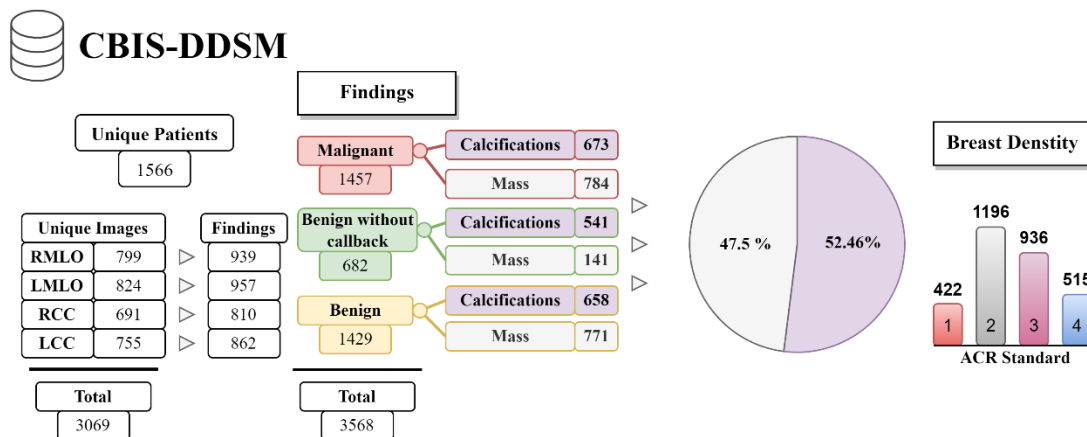


Figura 6. Información estadística base de datos CBIS-DDSM
Fuente: MračkoA et al.[24]

Aunque originalmente se planteó el uso de la base de datos MINI-DDSM, durante la investigación se encontró que es un subconjunto de la base de datos DDSM, la cual cuenta con una versión revisada por un mayor número de expertos, aunque excluye la edad del paciente, información que no fue necesaria en esta investigación.

La base de datos CBIS-DDSM[25] (Curated Breast Imaging Subset of DDSM) representa una versión avanzada y estandarizada de la base de datos Digital Database for Screening Mammography (DDSM). Esta fuente de datos valiosa comprende estudios de mamografía en película que han sido cuidadosamente digitalizados. En su interior, alberga una amplia variedad de casos que incluyen muestras normales, benignas y malignas, todas ellas respaldadas por información de patología verificada.

La compilación de CBIS-DDSM se caracteriza por la selección y curación minuciosa realizada por un radiólogo experimentado. Este proceso garantiza la calidad y coherencia de los datos. Además, las imágenes originales se han transformado al formato DICOM, formato estándar para la visualización de imágenes biomédicas.

Esta base de datos se compone de un conjunto considerable de recursos: 1566 pacientes y 3069 imágenes en las que se encuentran 3568 anotaciones. Para simplificar el enfoque en la detección y diagnóstico de diferentes tipos de lesiones, se ha dividido en dos grupos principales: calcificaciones y masas. Asimismo, cada grupo de imágenes se ha subdividido en conjuntos de entrenamiento y prueba, lo que facilita la evaluación y validación de algoritmos de aprendizaje automático.

En última instancia, CBIS-DDSM no solo proporciona una amplia gama de imágenes médicas, sino también información clínica relevante, como la densidad del seno, la sutileza de las lesiones y la gravedad de las patologías. Para masas, se detallan aspectos como la forma y los márgenes de la lesión, mientras que, para las calcificaciones, se incluye información sobre su tipo y distribución dentro del seno.

3.2 Software y Hardware

El proyecto se llevó a cabo en un sistema operativo Ubuntu 20.04 como plataforma base. El equipo empleado en la ejecución del proyecto constaba de un procesador Intel Xenon E5-2650 v4, una tarjeta gráfica NVIDIA GM204GL (Quadro M4000),

una capacidad de memoria RAM de 16GB y una unidad de almacenamiento con una capacidad de 1.5TB.

La implementación del proyecto se realizó utilizando el entorno de desarrollo Visual Studio Code, con el lenguaje de programación Python como núcleo principal. La construcción y ejecución de las redes neuronales se basó en el uso de la biblioteca TensorFlow. Además, se hicieron uso de librerías adicionales como NumPy, Sci-learn y OpenCV para llevar a cabo el desarrollo y la ejecución de las tareas de manipulación de datos e imágenes.

Capítulo 4: Desarrollo

En el presente capítulo, se presenta un detallado análisis del procedimiento seguido durante el proceso de entrenamiento. En primer término, se aborda el proceso de muestreo de la base de datos, que comprende la extracción de muestras de las lesiones presentes en las imágenes mamográficas. En segundo lugar, se describe el proceso de ajuste y entrenamiento de la red seleccionada. Finalmente, se profundiza en un proceso de entrenamiento destinado a la segmentación multiclase.

3.3 Preparación y muestreo de la base de datos.

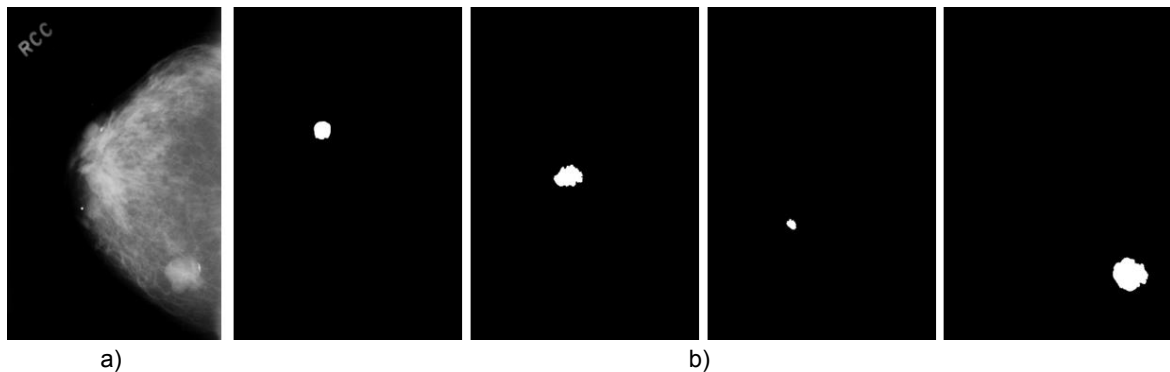


Figura 7. Caso base de datos a) mamografía. b) máscaras de segmentación individuales para cada lesión
Fuente: Autor

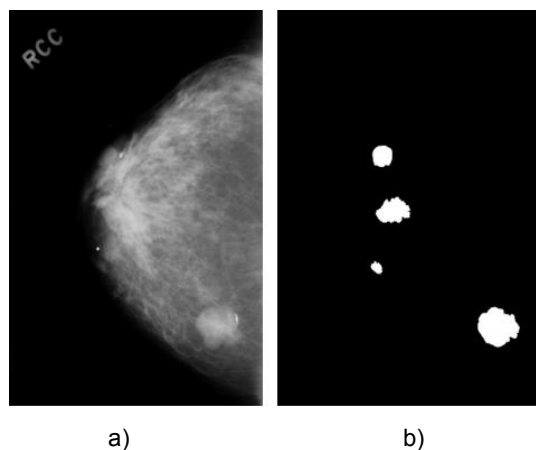


Figura 8. Caso base de datos. a) mamografía. b) máscara de segmentación con lesiones unificadas
Fuente: Autor

En el contexto de la base de datos CBIS-DDSM, que contiene imágenes de mamografía junto con máscaras correspondientes para cada lesión mamaria, surgió la necesidad de consolidar las máscaras relacionadas con un mismo paciente en una sola máscara que englobe todas las lesiones detectadas. Este enfoque nos permitió simplificar y organizar de manera más efectiva la información asociada con cada caso.

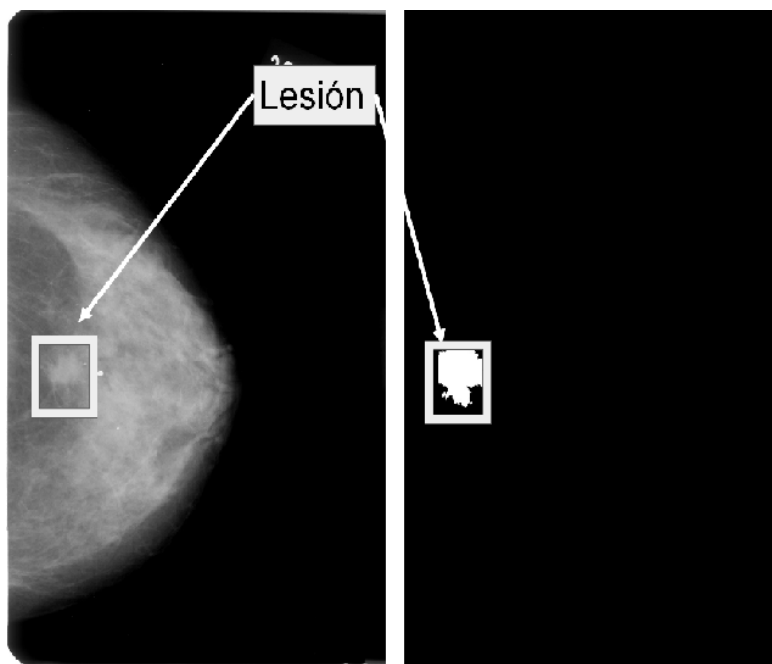


Figura 9. Cuadro de selección alrededor de la lesión
Fuente: Autor

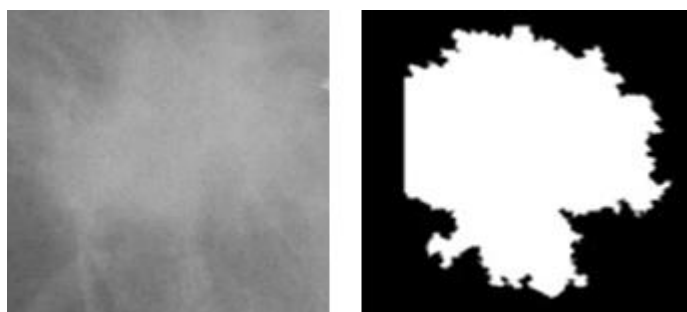


Figura 10. Muestra obtenida
Fuente: Autor

Para el muestreo de esta base de datos, se implementó el siguiente procedimiento. Se extrajo el centro de cada lesión y se generó un cuadro de dimensiones

estandarizadas, específicamente, un cuadro de 320 pixeles de altura por 320 pixeles de ancho que rodeaba a la lesión.

Con el objetivo de aumentar significativamente el número de muestras disponibles y enriquecer la información que se proporciona a la red neuronal, se tomó la decisión de realizar un proceso de data augmentation o aumento de datos. Este proceso involucró tomar el cuadro previamente generado y aplicar desplazamientos de al menos un 30 % en cada dirección, lo que resultó en múltiples muestras que abarcaban diferentes áreas del tejido circundante a la lesión.

Adicionalmente, se llevaron a cabo transformaciones geométricas para enriquecer aún más el conjunto de datos. Estas transformaciones incluyeron transformación por espejo X e Y, así como giros en ángulos de 90, 180 y 270 grados. Se experimentó con transformación por zoom no obstante estas generaban valores de interpolados en las máscaras lo que afectaba su integridad. Al finalizar este proceso se obtuvieron un total de 56937 imágenes.

3.4 Ajuste de la red

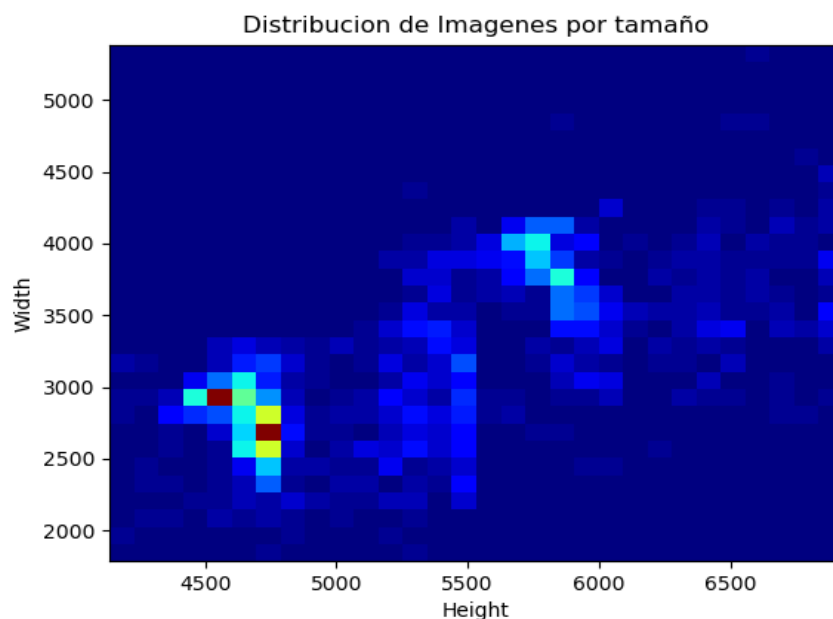


Figura 11. Distribución de imágenes por tamaño
Fuente: autor

Para abordar la tarea de la segmentación de masas, se inició utilizando una arquitectura de red UNET clásica como punto de partida, esta red fue elegida por su amplio uso en segmentación de imágenes biomédicas como el trabajo realizado por el semillero de investigación LUN en detección de nódulos pulmonares. Sin embargo, para agilizar el proceso y reducir la carga computacional, se optó por entrenar inicialmente la red en un conjunto de imágenes reducido representativo de 19534 muestras. Para lograr una representación óptima, se procedió a la generación de un mapa que visualiza la distribución de las imágenes en función de sus dimensiones. Se realizó la selección de una región caracterizada por una densidad notable de imágenes, como se evidencia en la Figura 11. Esta área específica comprende imágenes con alturas situadas entre 2300 y 3500 píxeles y anchos que oscilan entre 4250 y 4700 píxeles.

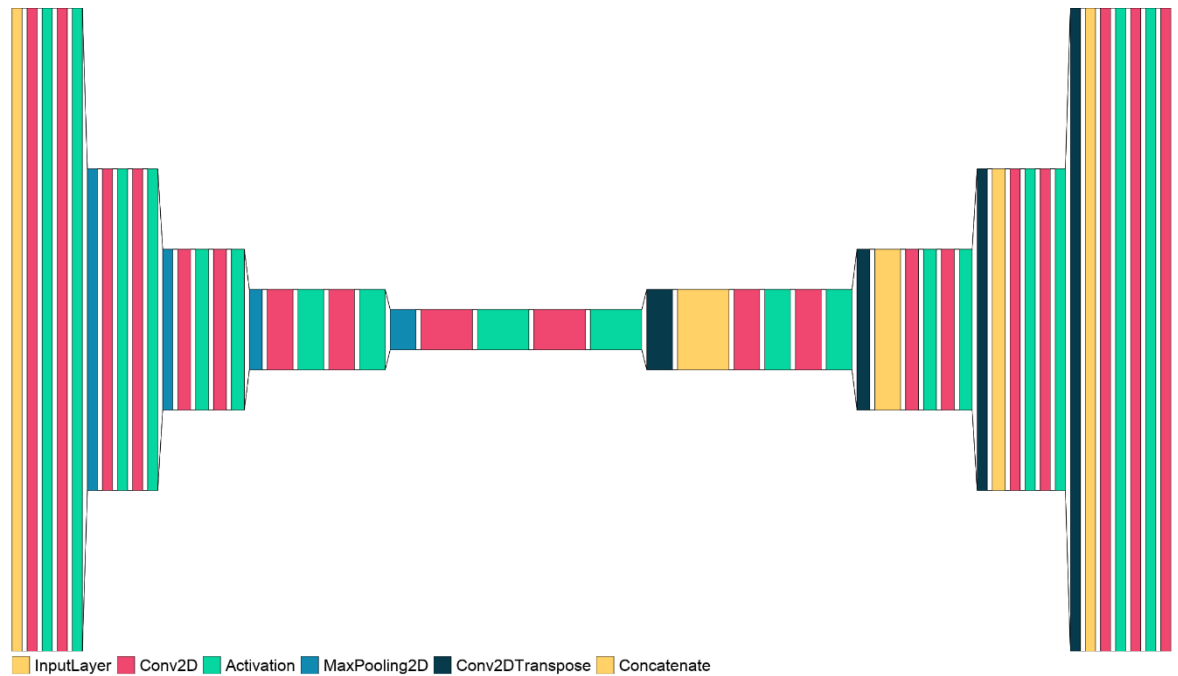
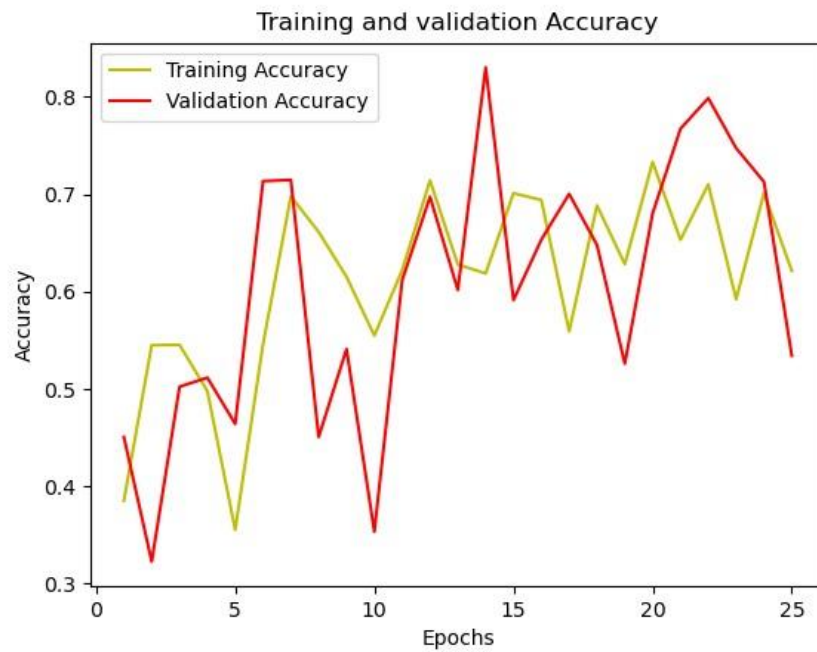


Figura 12. Modelo red UNET clásica

Fuente: autor



a)



b)

Figura 13. Curvas de aprendizaje entrenamiento UNET clásica. a) exactitud, b) precisión
Fuente: autor

El entrenamiento se llevó a cabo a lo largo de 25 épocas, durante las cuales se aplicó un enfoque de tasa de aprendizaje decreciente lo que va a permitir al modelo hacer ajustes más precisos al ritmo que va avanzando el proceso. Inicialmente, los resultados no fueron satisfactorios, ya que el aprendizaje no convergía de manera óptima como se puede apreciar en la figura 13

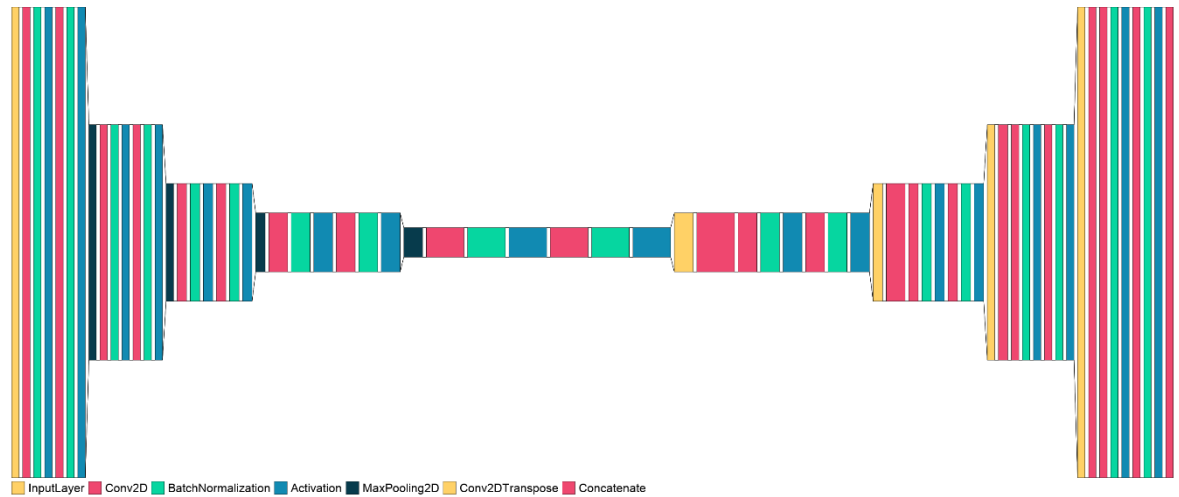
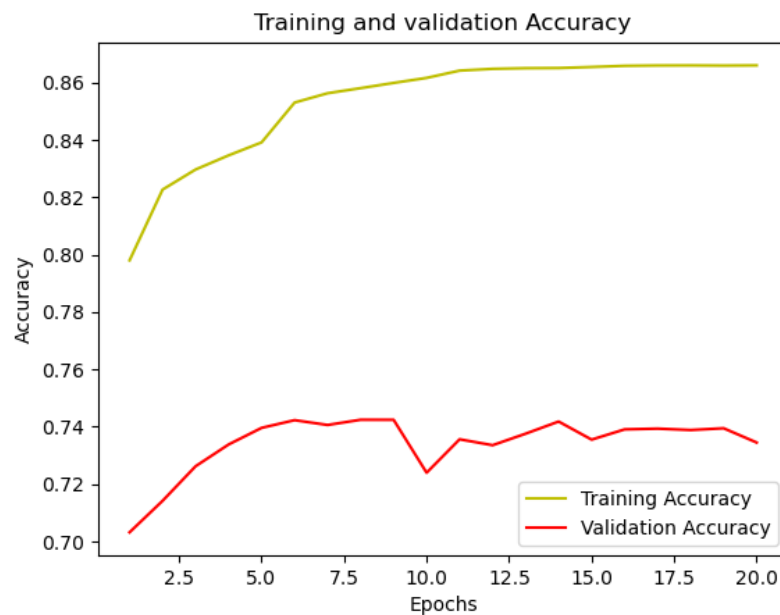


Figura 14. Modelo red UNET con capas de batch normalization
Fuente: autor



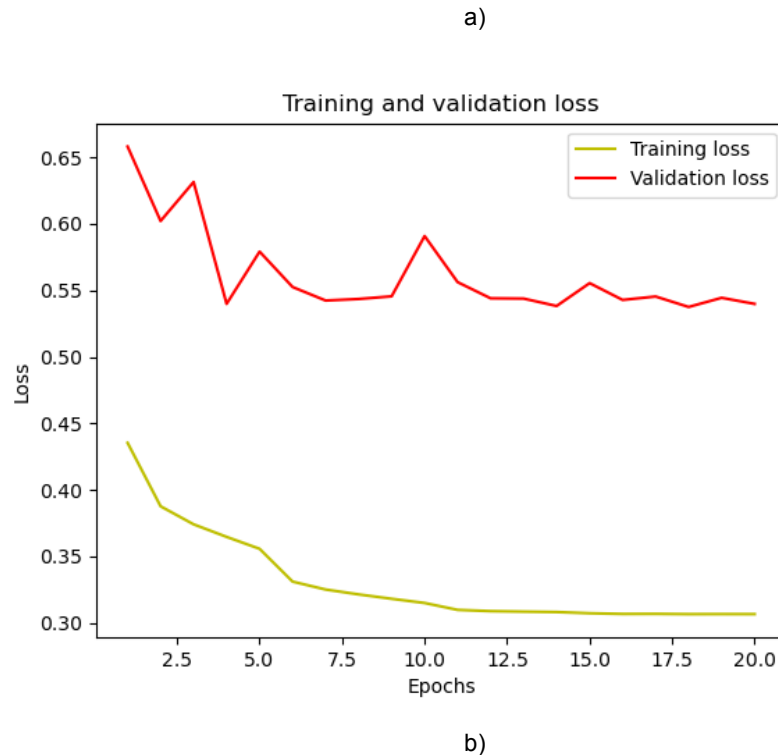


Figura 15. Curvas de aprendizaje con batch normalization a) exactitud, b) pérdida
Fuente: autor

Ante esta situación, se decidió tomar medidas para acelerar el proceso de convergencia. Se introdujeron capas de batch normalization antes de las capas de convolución de la red. Esto permitió una mejor normalización de los datos y, como resultado, una mayor eficiencia durante el entrenamiento.

Sin embargo, a medida que el modelo se desarrollaba, se identificó un problema de sobreajuste. Esta discrepancia se hizo evidente al comparar las curvas de aprendizaje en la figura 15, donde se observó un mejor rendimiento en el conjunto de entrenamiento en comparación con el conjunto de validación.

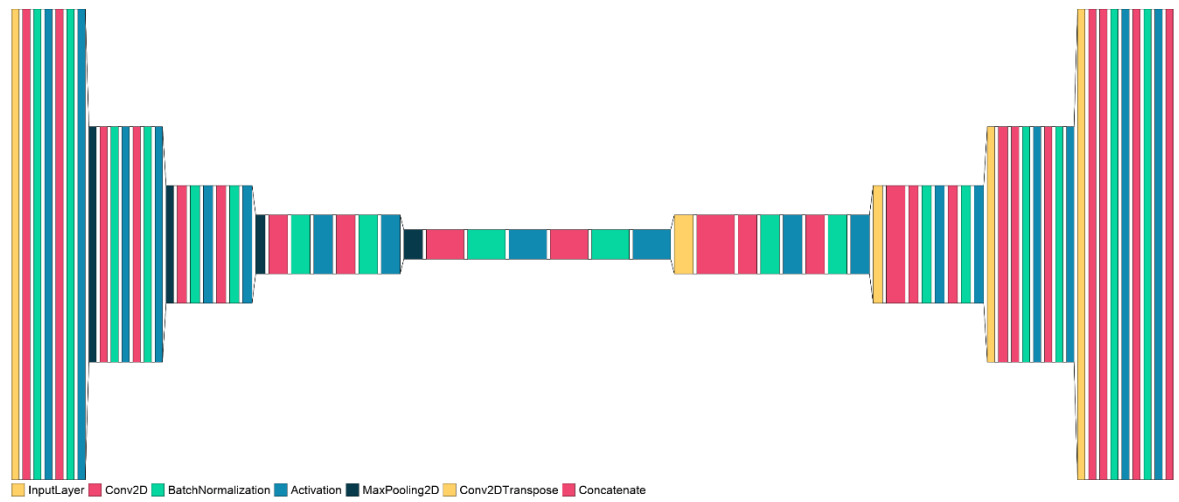
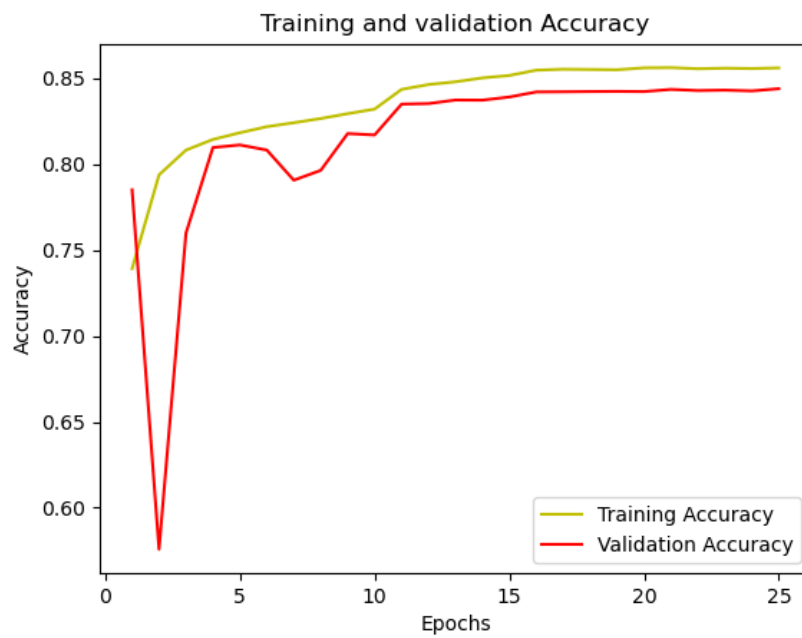
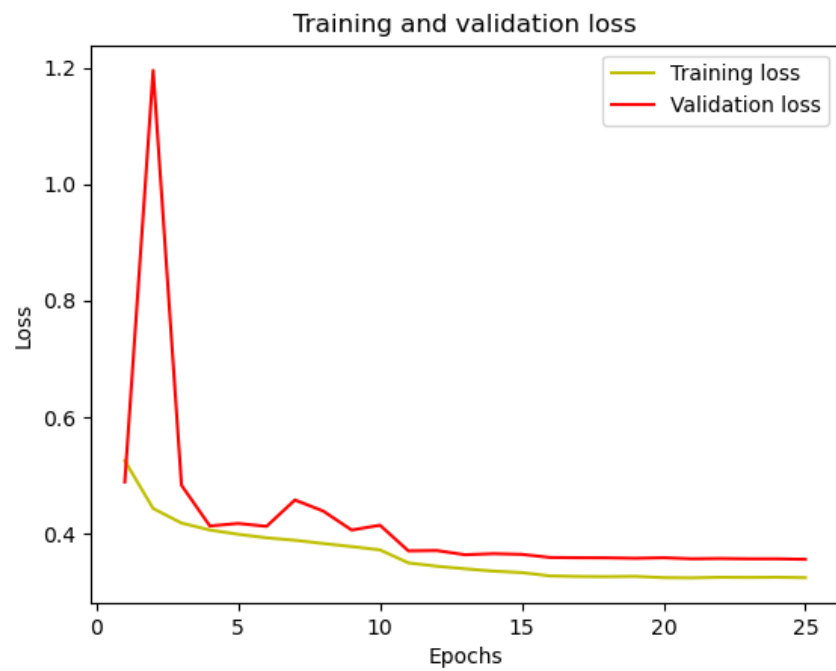


Figura 16. Modelo red UNET con batch normalization y dropout en codificador
Fuente: autor



a)



b)

Figura 17. Curvas de aprendizaje con batch normalization y dropout en codificador a) exactitud, b) perdida
Fuente: autor

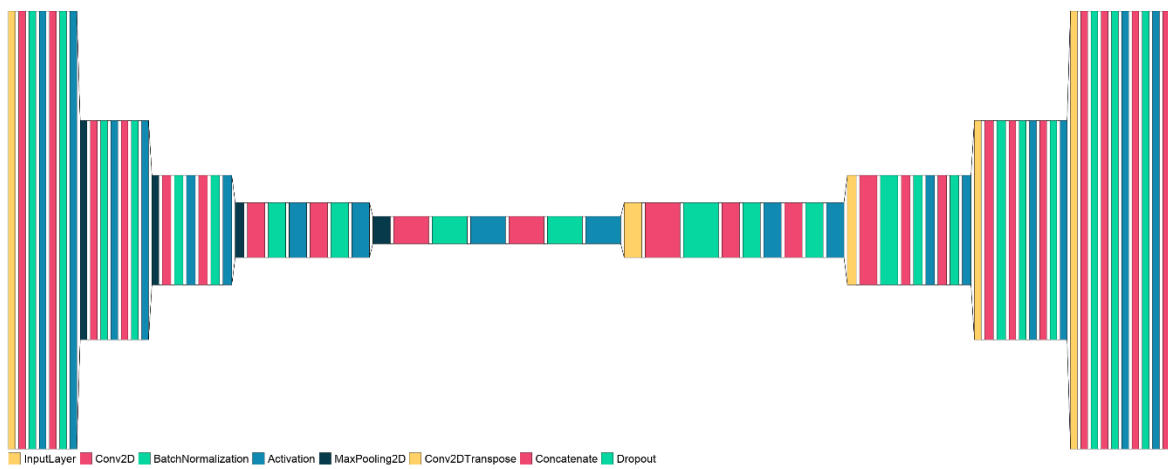
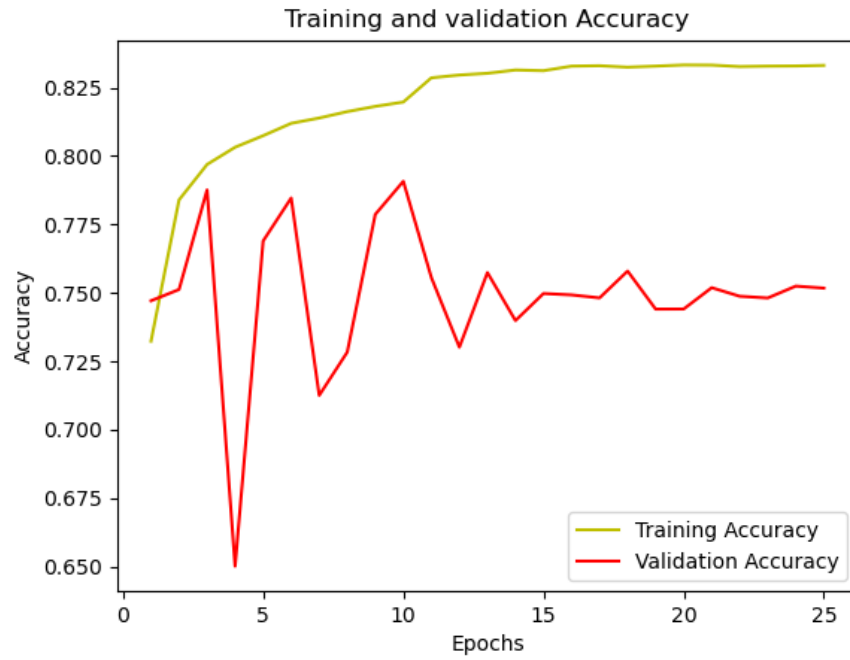
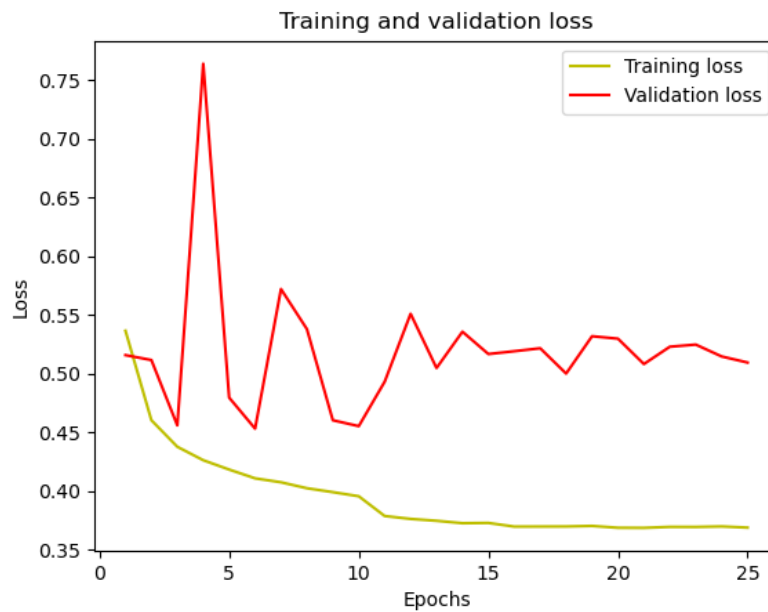


Figura 18. Modelo red UNET con batch normalization y dropout en decodificador
Fuente: autor



a)



b)

Figura 19. Curvas de aprendizaje con batch normalization y dropout en decodificador a) exactitud, b) perdida
Fuente: autor

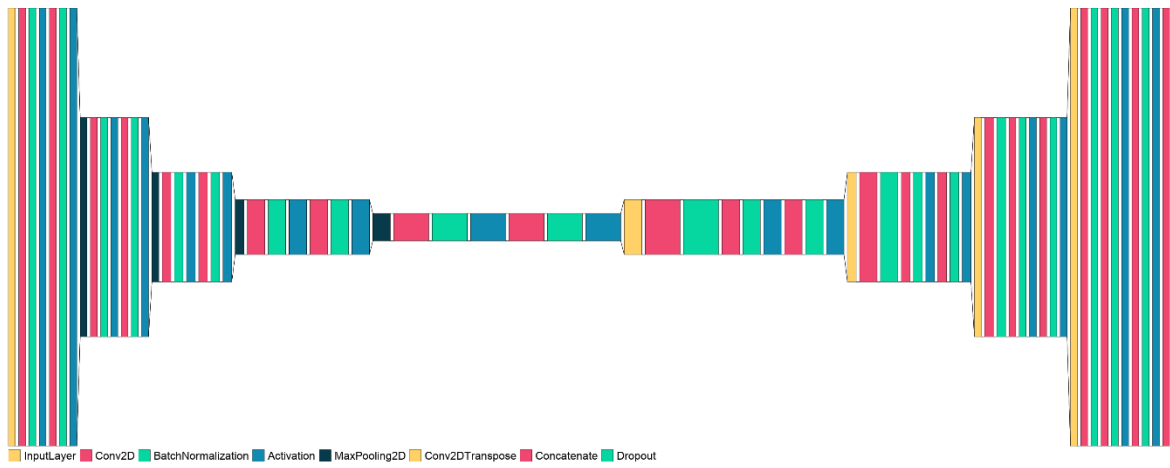
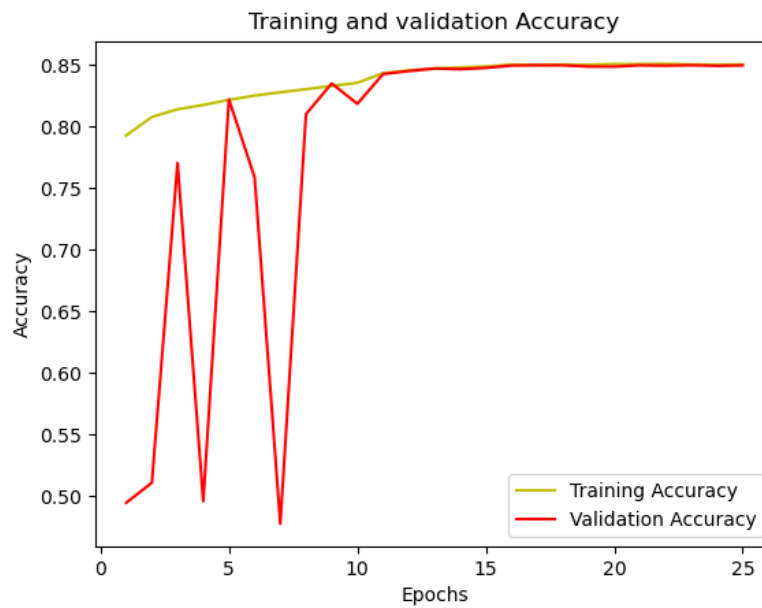
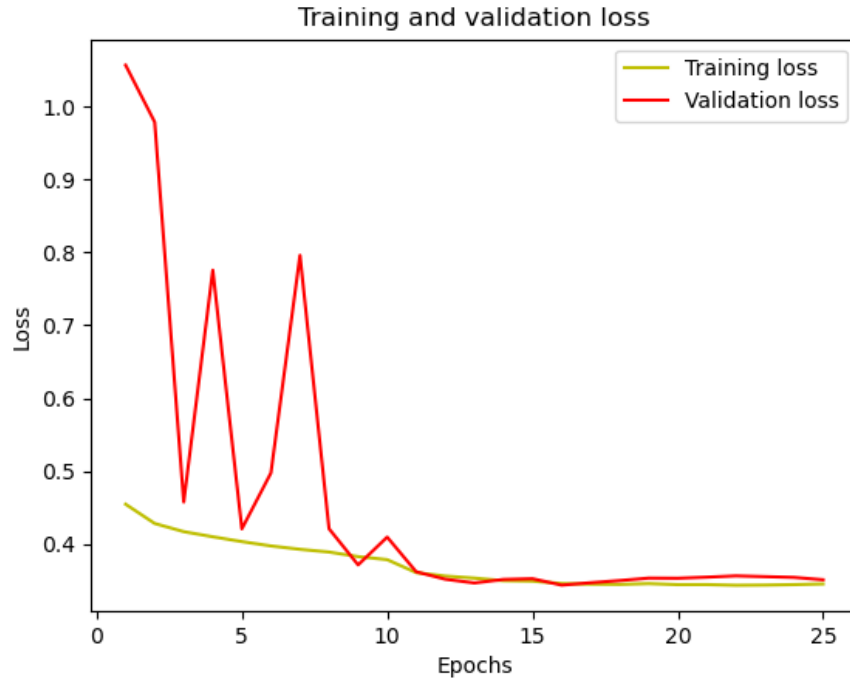


Figura 20. Modelo red UNET con batch normalization y dropout en codificador y decodificador
Fuente: autor



a)



b)

Figura 21. Curvas de aprendizaje con batch normalization, dropout en codificador y decodificador a) exactitud, b) pérdida

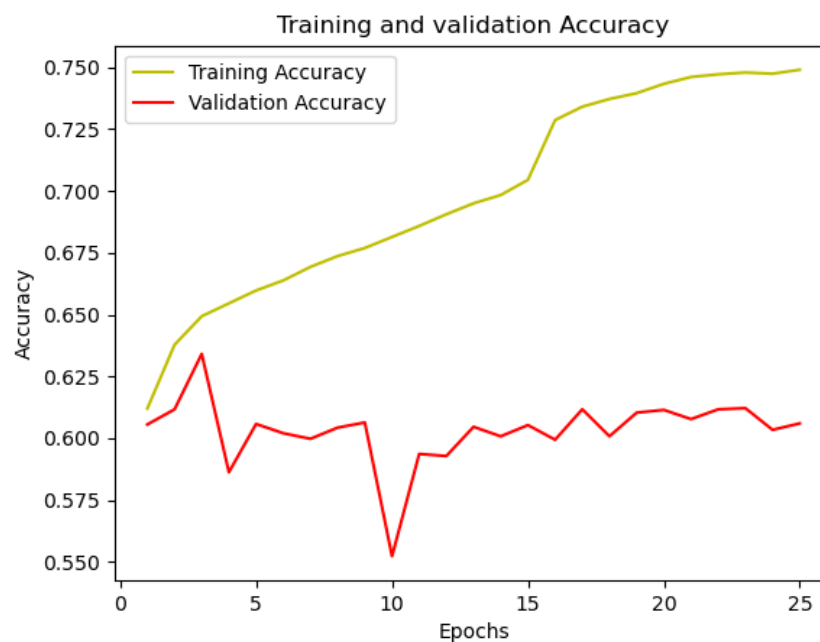
Fuente: autor

Para abordar este desafío, se optó por implementar técnicas de regularización. Se eligió la técnica de dropout, ya que esta impide al modelo depender de los pesos de una única neurona para obtener el resultado, y se realizaron experimentos con tres variantes del modelo. En el primer enfoque, se aplicó la capa de dropout después de cada operación de pooling en el codificador. En el segundo enfoque, el dropout se aplicó únicamente después de cada operación de deconvolución en el decodificador. Finalmente, en la tercera variante, se implementó dropout tanto en el codificador como en el decodificador.

3.5 Clasificación

Finalmente se procedió con el proceso de clasificación, para lo que se optó por entrenar con la red UNET con dropout en decodificador y codificador, se procedió con el proceso de clasificar las lesiones según su estado de benigno o maligno

Este nuevo desafío requería una preparación previa esencial. Para lograrlo, primero se llevó a cabo un proceso de ajuste de las máscaras de las lesiones. Esto debido a que era necesario codificar estas máscaras en un formato denominado "ONE-HOT" formato requerido por la herramienta para tareas de segmentación multiclase.



a)



b)
Figura 22. Curvas de aprendizaje clasificación a) exactitud, b) perdida
Fuente: autor

Una vez concluido el proceso de entrenamiento, se constató la presencia nuevamente de inconvenientes relacionados con el sobreajuste, tal como se visualiza en la Figura 22. En esta representación gráfica, se destaca una notoria disparidad de rendimiento entre los conjuntos de entrenamiento y validación.

Capítulo 5: Resultados

En la fase final del proyecto, se procedió a la realización de pruebas mediante un conjunto de datos de imágenes previamente no empleadas en el proceso de entrenamiento. Con el objetivo de efectuar predicciones sobre la totalidad de las imágenes, se adoptó un enfoque basado en la técnica de ventana deslizante. Este método implica la subdivisión de la imagen en secciones más pequeñas, las cuales son sometidas a predicciones individuales, para posteriormente reconstruir un resultado final mediante la combinación de estas predicciones. A continuación, se procedió a calcular métricas fundamentales, tales como la exactitud y la intersección sobre la unión (IoU).

	Imágenes de entrenamiento	Tiempo(h)	Exactitud	IoU
UNET	19534	25.15	0.55	0.42
UNET + batch normalization	19534	25.10	0.69	0.62
UNET + batch normalization+ dropout en codificador	19534	24.90	0.73	0.68
UNET + batch normalization+ dropout en decodificador	19534	24.90	0.71	0.69
UNET + batch normalization+ dropout en decodificador y codificador	19534	25.10	0.74	0.70
Clasificación	56937	127	0.57	0.39

Tabla 1. Tabla comparativa de resultados
Fuente: autor

En el ámbito de la segmentación binaria, se pudo constatar una mejora sustancial en el rendimiento al implementar técnicas como batch normalization y la aplicación de dropout tanto en el codificador como en el decodificador de la red, llegando a una exactitud del 74%.

Por otra parte, en el contexto de la segmentación multiclase, se identificó un notable incremento en el tiempo de entrenamiento. Este fenómeno se atribuye a la mayor complejidad inherente a esta tarea la cual requiere considerar y distinguir entre múltiples clases y al incrementado número de datos.

6. Conclusiones y recomendaciones

6.1 Conclusiones

Se implementó y estudió la red UNET, la cual fue diseñada originalmente para la clasificación y segmentación de imágenes biomédicas obteniendo una precisión del 74% segmentación de masas, y un 57% en la clasificación.

El modelo demostró éxito en mitigar el sobreajuste en tareas de segmentación de masas, aunque se enfrentó a dificultades en la transición hacia tareas de clasificación multiclase, como la segmentación por severidad de lesión. A pesar de no superar los resultados obtenidos por investigaciones previas, este proyecto representa un paso significativo en la exploración del tema dentro del ámbito del semillero de investigación, contribuyendo al avance en el estudio de la materia.

Al analizar las muestras se vio la necesidad de realizar tareas de data augmentation, y de generar nuevas muestras por medio de desplazamientos para que la red pueda aprender información del contexto alrededor a la lesión.

Se observó que el modelo que contiene batch normalization y dropout en decodificador y codificador obtuvo la mejor exactitud con del 69% y una intersección sobre la unión del 62%, no obstante, este modelo no reflejó este desempeño en la tarea de clasificación donde no solo experimentó un sobreajuste, sino que también recibió bajos valores de exactitud con 57% y una intersección sobre la unión del 39%.

El modelo encontró desafíos particulares en la segmentación de imágenes donde la masa no era claramente discernible del tejido circundante y en casos donde la lesión era de tamaño relativamente reducido. Dichas situaciones presentaron obstáculos considerables en la obtención de resultados precisos.

Un factor determinante durante el proyecto fue la limitación de tiempo, ya que el proceso de entrenamiento del modelo requería aproximadamente de 27 a 28 horas con un conjunto de muestras reducido, y alrededor de 5 días con un conjunto completo de muestras. Estos tiempos prolongados representaron uno de los principales desafíos en la ejecución del proyecto.

6.2 Recomendaciones

Dada la particular naturaleza del tejido mamario y la importancia de distinguir las masas dentro del tejido mamario con más claridad, se recomienda la exploración y aplicación de técnicas de mejora de contraste. Estas técnicas tienen el potencial de realzar de manera significativa las características de las masas en las mamografías.

Además, se sugiere como una valiosa estrategia la búsqueda de acceso a la base de datos OMI-DB, una fuente de datos de origen inglés que se encuentra disponible mediante solicitud. Esta base de datos alberga exámenes mamográficos más modernos y actualizados, lo que puede ser de gran utilidad para investigaciones y estudios en el campo de la radiología mamaria.

En el contexto de la utilización de modelos de aprendizaje automático, se recomienda explorar el uso de redes totalmente convolucionales (Fully Convolutional Networks). Estas redes permiten llevar a cabo el entrenamiento utilizando imágenes de mamografías en su resolución original, eliminando la necesidad de redimensionarlas. Esta herramienta puede resultar beneficiosa al preservar los detalles y la información relevante de las imágenes.

Además, se recomienda el continuar pruebas con los métodos de regularización ya que como se pudo observar el modelo de comportamiento de manera adecuada en tareas de segmentación binaria, no obstante, todavía encuentra problemas de sobreajuste en tareas de clasificación multiclase.

Referencias bibliográficas

- [1] "Breast cancer." Accessed: Sep. 04, 2023. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>
- [2] I. Díaz-Yuñez, M. Porras, G. Parra, S. Serrano Montes, and E. De Nubila, "Artículo Original ACCURACY OF A DIGITAL MAMMOGRAPHY SERVICE IN THE DIAGNOSIS OF BREAST CÁNCER ABSTRACT," 2012.
- [3] "Tratamiento del cáncer de seno | American Cancer Society." Accessed: Nov. 02, 2023. [Online]. Available: <https://www.cancer.org/es/cancer/tipos/cancer-de-seno/tratamiento.html>
- [4] "Detección temprana y diagnóstico del cáncer de seno | American Cancer Society." Accessed: Nov. 02, 2023. [Online]. Available: <https://www.cancer.org/es/cancer/tipos/cancer-de-seno/pruebas-de-deteccion-y-deteccion-temprana-del-cancer-de-seno.html>
- [5] J. Armando Ortiz Murillo *et al.*, "Detección de cáncer de mama usando técnicas avanzadas de minería de datos con redes neuronales," 2016.
- [6] M. Cantone, C. Marrocco, F. Tortorella, and A. Bria, "Convolutional Networks and Transformers for Mammography Classification: An Experimental Study," *Sensors*, vol. 23, no. 3, Feb. 2023, doi: 10.3390/S23031229.
- [7] L. Shen, L. R. Margolies, J. H. Rothstein, E. Fluder, R. McBride, and W. Sieh, "Deep Learning to Improve Breast Cancer Detection on Screening Mammography," *Sci Rep*, vol. 9, no. 1, Dec. 2019, doi: 10.1038/S41598-019-48995-4.
- [8] V. K. Singh *et al.*, "Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network," *Expert Syst Appl*, vol. 139, p. 112855, Jan. 2020, doi: 10.1016/j.eswa.2019.112855.
- [9] "¿Qué es el machine learning? - Explicación sobre el machine learning empresarial - AWS." Accessed: Nov. 02, 2023. [Online]. Available: <https://aws.amazon.com/es/what-is/machine-learning/>
- [10] "Perceptron : qu'est-ce que c'est et à quoi ça sert ?" Accessed: Sep. 05, 2023. [Online]. Available: <https://datascientest.com/es/perceptron-que-es-y-para-que-sirve>

- [11] “¿Qué es una red neuronal? - Explicación de las redes neuronales artificiales - AWS.” Accessed: Sep. 06, 2023. [Online]. Available: <https://aws.amazon.com/es/what-is/neural-network/>
- [12] “¿Qué es el aprendizaje profundo? - Explicación del aprendizaje profundo - AWS.” Accessed: Nov. 02, 2023. [Online]. Available: <https://aws.amazon.com/es/what-is/deep-learning/>
- [13] “Descripción del funcionamiento de una red neuronal convolucional (CNN)... | Download Scientific Diagram.” Accessed: Sep. 06, 2023. [Online]. Available: https://www.researchgate.net/figure/Figura-1-Descripcion-del-funcionamiento-de-una-red-neuronal-convolucional-CNN-10_fig1_348825166
- [14] “¿Qué son las redes neuronales convolucionales? | IBM.” Accessed: Sep. 06, 2023. [Online]. Available: <https://www.ibm.com/es-es/topics/convolutional-neural-networks>
- [15] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” May 2015, [Online]. Available: <http://arxiv.org/abs/1505.04597>
- [16] “Max Pooling Definition | DeepAI.” Accessed: Nov. 02, 2023. [Online]. Available: <https://deepai.org/machine-learning-glossary-and-terms/max-pooling>
- [17] “¿Qué es el sobreajuste? - Explicación del sobreajuste en machine learning - AWS.” Accessed: Nov. 02, 2023. [Online]. Available: <https://aws.amazon.com/es/what-is/overfitting/>
- [18] “Regularization in Machine Learning - GeeksforGeeks.” Accessed: Nov. 02, 2023. [Online]. Available: <https://www.geeksforgeeks.org/regularization-in-machine-learning/>
- [19] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” 2014.
- [20] S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” *32nd International Conference on Machine Learning, ICML 2015*, vol. 1, pp. 448–456, Feb. 2015, Accessed: Sep. 06, 2023. [Online]. Available: <https://arxiv.org/abs/1502.03167v3>
- [21] “Data augmentation-based color and geometry transformation. | Download Scientific Diagram.” Accessed: Sep. 06, 2023. [Online]. Available: https://www.researchgate.net/figure/Data-augmentation-based-color-and-geometry-transformation_fig4_342625631

- [22] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, p. 100258, Dec. 2022, doi: 10.1016/J.ARRAY.2022.100258.
- [23] "Evaluating Image Segmentation Models | by Frank Liang | Towards Data Science." Accessed: Nov. 02, 2023. [Online]. Available: <https://towardsdatascience.com/evaluating-image-segmentation-models-1e9bb89a001b>
- [24] A. Mračko, L. Vanovčanová, and I. Cimrák, "Mammography Datasets for Neural Networks—Survey," *Journal of Imaging*, vol. 9, no. 5. MDPI, May 01, 2023. doi: 10.3390/jimaging9050095.
- [25] R. S. Lee, F. Gimenez, A. Hoogi, K. K. Miyake, M. Gorovoy, and D. L. Rubin, "A curated mammography data set for use in computer-aided detection and diagnosis research," *Sci Data*, vol. 4, Dec. 2017, doi: 10.1038/SDATA.2017.177.