

Universidad ORT Uruguay
Facultad de Ingeniería

Identificación de la Región de Interés en imágenes DICOM basado en Inteligencia Artificial

Entregado como requisito para la obtención del
título de Máster en Big Data

Matías Méndez - 140709
Juan Pablo Muñoz - 253250

Tutor: Leonardo Cecilia

2021

Nosotros, Matías Méndez y Juan Pablo Muñoz, declaramos que el trabajo que se presenta en esta obra es de nuestra propia mano. Podemos asegurar que:

- La obra fue producida en su totalidad mientras realizábamos el Proyecto Final de Master en Big Data;

- Cuando hemos consultado el trabajo publicado por otros, lo hemos atribuido con claridad;

- Cuando hemos citado obras de otros, hemos indicado las fuentes. Con excepción de estas citas, la obra es enteramente nuestra;

- En la obra, hemos acusado recibo de las ayudas recibidas;

- Cuando la obra se basa en trabajo realizado conjuntamente con otros, hemos explicado claramente qué fue contribuido por otros, y qué fue contribuido por nosotros;

- Ninguna parte de este trabajo ha sido publicada previamente a su entrega, excepto donde se han realizado las aclaraciones correspondientes.



Matías Méndez



Juan Pablo Muñoz

24-09-2021

Dedicado

A nuestras familias, que nos soportaron durante estos largos meses.

Agradecimientos

- A Leonardo Cecilia, por su guía, disponibilidad, ayuda ante cualquier duda o problema.
- A Diego Espinosa, gran amigo y jefe.
- A Agostina por el apoyo brindado.

Abstract

El presente trabajo aborda la identificación de regiones de interés en imágenes médicas mediante el uso de técnicas de deep learning. Comienza presentando el estado del arte del dominio del problema, analizando diferentes propuestas publicadas para luego realizar una selección de dos arquitecturas que finalmente se implementan. Dichas arquitecturas se prueban con dos datasets diferentes para establecer una comparación entre ellas; por un lado se usan imágenes de mamografías y por otro resonancias magnéticas de cerebros, siendo las regiones de interés posibles lesiones o tumores. Además, en el proceso se investiga y se usa el estándar de imágenes médicas DICOM.

Palabras clave

Inteligencia artificial, deep learning, región de interés, DICOM, imágenes médicas, YOLO, U-Net, redes neuronales, imágenes de resonancia magnética (MRI), mamografía, Python, Pytorch.

Índice general

1	Introducción	8
2	Desarrollo	10
2.1	Estado del arte	10
2.1.1	Región de interés en imágenes	10
2.1.2	Descripción de metodologías	12
2.1.3	Métricas para evaluación	19
2.2	DICOM	20
2.3	Datasets	23
2.3.1	Curated Breast Imaging Subset of DDSM (CBIS-DDSM)	24
2.3.2	Brain MRI segmentation - The Cancer Imaging Archive(TCIA)	25
2.4	Técnica 1: YOLO	28
2.4.1	Arquitectura	28
2.4.2	Implementación 1 Mamografías	32
2.4.3	Implementación 2 MRI	40
2.4.4	Comparación de resultados	44
2.5	Técnica 2: U-Net	45
2.5.1	Arquitectura	46
2.5.2	Implementación 1 MRI	48
2.5.3	Implementación 2 Mamografías	55
2.5.4	Comparación de resultados	59
3	Conclusiones	61
4	Trabajo a futuro	62
5	Bibliografía	64
6	Anexos	69
6.1	Búsqueda de estado del arte	69

6.2 Experimentos Técnica 1	70
6.2.1 Mamografías	70
6.2.2 MRI BRain Segmentation	82

1 Introducción

El presente trabajo se presenta como requisito para obtener el título de Máster en Big Data y describe la implementación de un algoritmo que permita resolver el problema planteado por la empresa ROOTSTRAP de identificación de la Región de interés (ROI, Region Of Interest) en imágenes médicas contenidas bajo el formato de datos DICOM.

El análisis de imágenes médicas se refiere típicamente a la utilización de varios tipos de modalidades y técnicas, para obtener imágenes del cuerpo humano que, a su vez, pueden ser utilizadas por expertos médicos para el diagnóstico y tratamiento de los pacientes[1]. Por ejemplo, la detección temprana del cáncer de pulmón puede ayudar a reducir drásticamente la tasa de mortalidad por este tipo de cáncer [2].

Dado que la Región de Interés (ROI), en imágenes médicas se conoce como aquella región que se considera más importantes que otras, por ejemplo, en resonancias magnéticas cerebrales, el ROI corresponde al tumor[3]. Antes de tomar cualquier decisión de diagnóstico y tratamiento, el médico debe verificar la integridad de la región de interés en la imagen recibida para evitar diagnósticos erróneos[4]. Con el creciente desarrollo de las técnicas de aprendizaje profundo, el análisis de imágenes médicas se ha convertido en un campo activo de investigación[1].

Los estudios médicos que contienen imágenes, típicamente se encuentran bajo el estándar de comunicación Digital Imaging and Communication On Medicine (DICOM)[5], supervisado y respaldado por la National Electrical Manufacturers Association (NEMA)[6], con el fin de garantizar la interoperabilidad de diferentes tipos de estudio y el hardware correspondiente para realizarlo[7].

El trabajo fue realizado en el transcurso de la asignatura “Trabajo final” del

Máster en Big data, la cual comenzó en Abril de 2021 y tuvo una duración de 4 meses. Para gestionar mejor el trabajo, el esfuerzo fue dividido en tres etapas que están descritas a continuación:

- Etapa 1. Investigación de estado del arte
- Etapa 2. Implementación de algoritmos
- Etapa 3. Comparación de resultados

La primera etapa estuvo enfocada en la investigación documental del estado del arte en búsqueda de las bases teóricas que abordaran la temática planteada en el objetivo del trabajo, tales como uso de inteligencia artificial para el análisis de estudio médicos, identificación de región de interés en imágenes, conceptos y lineamientos del estándar DICOM e implementación de algoritmos de inteligencia artificial en imágenes médicas contenidas en dicho estándar.

En la segunda etapa se realizó la búsqueda y selección de dos datasets que contuvieran información de imágenes médicas bajo el estándar DICOM; acto seguido se inició la exploración de la información contenida en ellos y la forma de adaptarlos al uso de redes neuronales. Paralelamente se seleccionaron dos arquitecturas y se realizó su respectiva implementación, cada una de ellas con uno de los datasets seleccionados.

Una vez finalizada la implementación de los dos algoritmos, se inicia la tercera etapa, en donde se procede a adaptar el segundo dataset a cada algoritmo para observar sus resultados y compararlos con los del primer dataset. Finalmente se realiza la comparación de los resultados entre los dos algoritmos implementados con los dos datasets y se procede a extraer las conclusiones.

Todo el desarrollo realizado para este trabajo se encuentra disponible en el siguiente repositorio: https://github.com/Colury/TesisBigData_Mendez-Munoz

2 Desarrollo

2.1 Estado del arte

En esta sección se presenta la investigación realizada del estado del arte en cuanto al análisis de imágenes para obtener la región de interés.

2.1.1 Región de interés en imágenes

Dadas las características y objetivos del proyecto se inicia la exploración del estado del arte con ayuda de los motores de búsqueda Google Académico[8] y FOCO[9], con el fin de encontrar artículos relacionados al problema planteado en el trabajo, usando palabras claves como “región de interés”, “volumen de interés”, “inteligencia artificial”, “visión por computadora”, “DICOM”, etc. Las ecuaciones utilizadas en la búsqueda se encuentran en el anexo 6.1.

2.1.1.1 Artículos seleccionados

Dentro de la búsqueda realizada con los distintos criterios y palabras claves, se realiza una preselección de 33 artículos relacionados al problema y se compilan para su posterior lectura.

A continuación se encuentran listados los artículos seleccionados:

- Sliding Window Based Machine Learning System for the Left Ventricle Localization in MR Cardiac Images[10]
- You Only Look Once: Unified, Real-Time Object Detection[11]
- Automated region-of-interest localization and classification for vision-based visual assesment of civil infrastructure[12]

- Invariants Features from Interest Point Groups[13]
- Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm[14]
- Multi-Task Vehicle Detection With Region-of-Interest Voting[15]
- MRI-PET Correlation in Three Dimensions Using a Volume-of-interest (VOI) Atlas[16]
- The impact of 3D volume of interest definition on accuracy and precision of activity estimation in quantitative SPECT and planar processing methods[17]
- Improving Computer-Aided Detection Using Convolutional Neural Networks and Random View Aggregation[18]
- Volume of Interest Analysis of Spatially Normalized PRESTO Imaging to Differentiate between Parkinson Disease and Atypical Parkinsonian Syndrome[19]
- Discrimination of benign from malignant breast lesions in dense breasts with modelbased analysis of regions-of-interest using directional diffusion-weighted images[20]
- Lung Cancer Detection Using Image Processing and Machine Learning HealthCare[21]
- U-Net: Convolutional Networks for Biomedical Image Segmentation[22]
- Quantitative analysis of medical images: finding relevant regions-of-interest for medical decision support[23]
- Diagnosis and prognosis of cardiovascular diseases by means of texture analysis in magnetic resonance imaging[24]
- Detección jerárquica de grupo de personas[25]
- Template-Based Method for Multiple Volumes of Interest of Human Brain PET Images[26]
- Normalized volume of interest selection and measurement of bone volume in MicroCT scans[27]
- Object detection with sliding window in images including multiple similar objects[28]
- Flexible sliding windows with adaptive pixel strides[29]

- Movie scene segmentation using object detection and set theory[30]
- A Pedestrian Detection Algorithm Based on Score Fusion for Multi-LiDAR Systems[31]
- Visual Object Detection Using Cascades of Binary and One-Class Classifiers[32]
- A Deep Learning-Based Approach for the Detection and Localization of Prostate Cancer in T2 Magnetic Resonance Images[33]
- Automated Breast Cancer Diagnosis Using Deep Learning and Region of Interest Detection (BC-DROID)[34]
- A Method of Segmentation in 3D Medical Image for selection of Region of Interest (ROI)[35]
- Shifting from region of interest (ROI) to voxel-based analysis[36]
- Robust medical image watermarking technique for accurate detection of tamperers inside region of interest and recovering original region of interest[4]
- Reversible Region of Non-Interest (RONI) Watermarking for Authentication of DICOM Images[37]
- Review on Region of Interest Coding Techniques for Medical Image Compression[3]
- Lossless and Lossy Compression of DICOM images With Scalable ROI[38]
- Compression of digital medical images based on multiple regions of interest[39]
- Medical Image Compression based on Region of Interest using Better Portable Graphics (BPG)[40]

2.1.2 Descripción de metodologías

Para la primera revisión de los artículos antes mencionados, se realiza una selección primaria de papers usando criterios como año de publicación y cantidad de revisiones, dado que cada vez existen más desarrollos en arquitecturas basadas en redes neuronales, se prefiere algo a la vanguardia, que contenga un mayor número de revisiones y sea más citado en artículos académicos. Se entiende que el mayor número de revisiones garantizan más fiabilidad para agregar mayor solidez al objetivo del trabajo.

En base a los criterios explicados, los siguientes artículos fueron seleccionados para un análisis más profundo:

- Improving Computer-Aided Detection Using Convolutional Neural Networks and Random View Aggregation[18]
- Automated Breast Cancer Diagnosis Using Deep Learning and Region of Interest Detection (BC-DROID)[34]
- A Deep Learning-Based Approach for the Detection and Localization of Prostate Cancer in T2 Magnetic Resonance Images[33]
- Sliding Window Based Machine Learning System for the Left Ventricle Localization in MR Cardiac Images[10]
- U-Net: Convolutional Networks for Biomedical Image Segmentation[22]

Una vez realizada esta selección se procede a revisarlos para analizar las problemáticas tratadas, las técnicas que fueron usadas y los resultados obtenidos en cada uno de ellos. Con el resultados de este análisis se seleccionan los algoritmos a implementar y posteriormente analizarlos dentro del problema objeto del proyecto.

Con relación al análisis detallado de cada uno de los artículos seleccionados se han elegido dos de ellos para su implementación y comparación con los distintos datasets.

El primer paper seleccionado por los autores corresponde a Automated Breast Cancer Diagnosis Using Deep Learning and Region of Interest Detection (BC-DROID)[34], el cual ha sido implementado en imágenes de formato DICOM e incluye técnicas de clasificación y detección en una sola pasada, basado en la arquitectura You Only Look Once: Unified, Real-Time Object Detection[11]. Sumado a los criterios ya mencionados, se añade que el algoritmo YOLO fue implementado por los autores del presente trabajo en el transcurso del cursado de la materia de Deep Learning del posgrado; si bien la arquitectura a implementar aquí no es exactamente igual, se espera que la experiencia obtenida con dicho algoritmo sea de utilidad con esta nueva variante.

Dentro de la selección del segundo paper se buscó encarar el problema con una técnica distinta a la de clasificación y detección usada en el primer modelo, por ese motivo se eligió segmentación. En primera instancia se seleccionó Sliding Window Based Machine Learning System for the Left Ventricle Localization in MR Cardiac Images[10], pero este se encuentra basado en imágenes en 3D, lo que hacía complejo comparar los resultados entre un paper y otro. Por esta razón se buscó un artículo que incluyera técnicas de segmentación y a su vez estuviese implementado para imágenes en 2D como lo es U-Net: Convolutional Networks for Biomedical Image Segmentation[22].

Los dos artículos seleccionados para implementación se encuentran descritos en mayor detalle en los capítulos 2.4 y 2.5.

A continuación se encuentra un breve resumen de los artículos analizados cuya implementaciones no han sido incluidas en el alcance de este trabajo.

2.1.2.1 Improving Computer-Aided Detection Using Convolutional Neural Networks and Random View Aggregation

Los autores emplean técnicas de inteligencia artificial para ayudar a resolver el problema de detección de falsos positivos generados en la búsqueda con alta sensibilidad de metástasis escleróticas, ganglios linfáticos y pólipos colónicos en imágenes correspondientes a tomografías computarizadas (CT), por medio de sistemas de detección por computadora (CADe: Computer Aided Detection) de vanguardia[18].

El trabajo plantea un diseño de un sistema en cascada que utiliza en primera parte un sistema CADe para generar los inputs tales como ROI y VOI en imágenes 2D y 2.5D, los cuales son pasados por una red neuronal convolucional (CNN) que ayuda a resolver el problema de detección de Falsos Positivos, asignando probabilidades de clase (por ejemplo: lesión, patología, etc) para un conjunto de vistas aleatorias que se promedian para determinar una probabilidad final de clasificación.

Las imágenes generadas por el sistema CADe generan los inputs para alimentar la Red Neuronal Convolucional - ConvNets, como se muestra en la figura 2.1, que consta de una secuencia de dos capas convolucionales con una operación de Max Polling entre ellas. A continuación de la segunda capa convolucional se aplican dos capas Locally Conected similares a las de convolución pero sin compartir los pesos y finalmente una capa Fully Connected para que ejecute las tareas de clasificación. Para evitar el overfitting usan Drop Connect el cual corresponde a una derivación de Dropout y ReLU como función de activación. A la hora de entrenar el modelo se usó Data augmentation para corregir el problema de Overfitting con operaciones de Cropped y Translation.

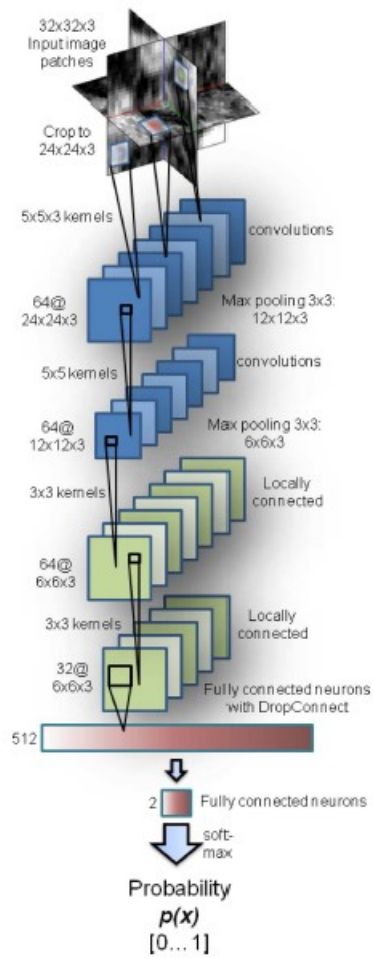


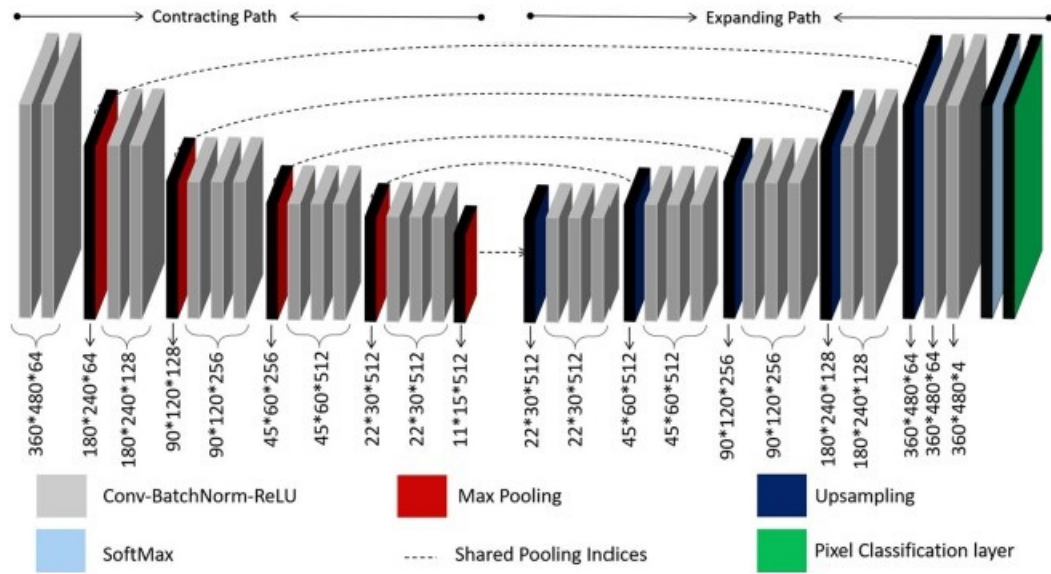
Figura 2.1: Arquitectura Improving Computer-Aided Detection Using Convolutional Neural Networks and Random View Aggregation[18]

Este método fue evaluado en tres conjuntos de datos, correspondiente a un total de 1421 pacientes distribuidos en 59 pacientes para la detección de metástasis escleróticas, 176 pacientes para la detección de ganglios linfáticos y 1.186 pacientes para la detección de pólipos colónicos.

2.1.2.2 A Deep Learning-Based Approach for the Detection and Localization of Prostate Cancer in T2 Magnetic Resonance Images

Los autores plantean el uso de algoritmos de inteligencia artificial para ayudar a resolver problemas de detección, localización y segmentación de lesiones prostáticas en imágenes de resonancias magnéticas T2W MRI 3D, generando una segmentación semántica de la estructura anatómica de la próstata y posibles lesiones malignas[33].

El trabajo realizado plantea una arquitectura de red convolucional Encoder - Decoder para realizar las tareas de segmentación de imágenes 3D, acompañado de un enfoque de ventana deslizante 3D, que conserva la complejidad del dominio 2D mientras que explota la información 3D. la arquitectura de la red está compuesta por dos segmentos comprendidos por una ruta de contracción del lado izquierdo (Encoder) y una ruta de expansión del lado derecho (Decoder), ver figura 2.2. La etapa de contracción está compuesta por 5 capas convolucionales con Batch Normalization y función de activación ReLU; En medio de cada convolución se aplica una capa de Max Polling. La etapa de expansión consta de 5 capas convolucionales del mismo tamaño que las de Contracción, en vista simétrica, dentro de cada una de las capas de convolución se encuentra una capa de Upsampled que recibe una Skip Connection desde las salida de las capas de Max Polling del lado de contracción y las salidas de cada capa convolucional de su lado para ir recuperando el tamaño de la imagen. Al final de la red, se encuentra una función Softmax para clasificar y una capa de Pixel Classification para etiquetar las probabilidades de clasificación resultantes.



[10]

Figura 2.2: Arquitectura A Deep Learning-Based Approach for the Detection and Localization of Prostate Cancer in T2 Magnetic Resonance Images[33]

Para el entrenamiento y evaluación del modelo se extrajeron 2356 cortes multicanal del conjunto de datos, los cuales fueron divididos en 1413 cortes para entrenamiento correspondientes al 60 %, 236 cortes para validación correspondientes al 10 % y el 30 % restante para test.

2.1.2.3 Sliding Window Based Machine Learning System for the Left Ventricle Localization in MR Cardiac Images

Los autores abordan el uso de inteligencia artificial para el desarrollo de la tarea específica de detección de ventrículos izquierdos del corazón en imágenes de Resonancias magnéticas, empleando un denominado sistema inteligente (Backpropagation Neural Network - BPNN) para lograr la tarea de detección. Ajustan el problema de detección del ventrículo izquierdo a tareas de clasificación binaria, asignando una clase a muestras recolectadas de ventrículos izquierdos y otra clase para objetos aleatorios (no ventrículos izquierdos). El trabajo realizado divide el problema en dos fases, una para el reconocimiento de ventrículos izquierdos dentro de la imagen y la otra para la detección de objetos dentro del ventrículo izquierdo[10].

La arquitectura planteada corresponde a una (Backpropagation Neural Network - BPNN) convencional con 1600 capas de entrada, 80 capas ocultas y funciones de activación tipo Log-Sigmoid, entrenada en 1215 Epochs, en la figura 2.3 se observa un ejemplo de una BPNN.

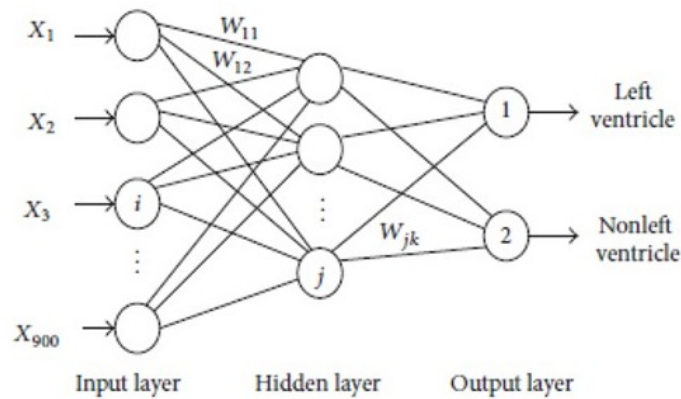


Figura 2.3: Ejemplo de red neuronal de retropropagación, utilizada en Arquitectura de Sliding Window Based Machine Learning System for the Left Ventricle Localization inMR Cardiac Images[10]

Este método fue evaluado con 250 muestras para entrenamiento y 50 muestras usadas para test.

2.1.2.4 Conclusiones de los papers que no fueron implementados

Los artículos seleccionados, si bien están todos centrados en el marco de los objetivos de este trabajo como son el análisis de estudios médicos por medio de Inteligencia Artificial, identificación de Regiones de Interés - ROI, describen distintas técnicas de implementación entre las que están la clasificación, la cual ayuda a identificar lo que está contenido en una imagen, detección que especifica la ubicación de varios objetos en la imagen y segmentación que creará una máscara de píxeles de cada objeto en las imágenes[41] .

A su vez, describen diferentes formas de recibir los datos de entrada a las redes neuronales, como son imágenes provenientes de Sistemas CADE e imágenes en formato 2D, 2.5D y 3D, los autores del presente trabajo consideran que los seleccionados para la implementación son los que pueden llegar a ser comparados dado que manejan distintas técnicas y se alimentan con el mismo formato de datos, en este caso imágenes 2D.

2.1.3 Métricas para evaluación

En la búsqueda del cumplimiento del objetivo principal del trabajo, correspondiente a la identificación de Regiones de interés en imágenes médicas en formato DICOM, se buscó probar 3 técnicas diferentes, correspondientes a Clasificación, Detección y Segmentación, por lo que ha sido necesario identificar métricas que permitieran evaluar y comparar los resultados de cada técnica con los diferentes datasets seleccionados y entre ellas. Por los que se seleccionaron únicamente dos métricas para esto.

- Accuracy
- Intersection over Union (IoU)

2.1.3.1 Accuracy

El Accuracy de un algoritmo de clasificación de aprendizaje automático es una forma en la que se mide la frecuencia con la que un algoritmo clasifica correctamente un punto de datos. El Accuracy es el número de puntos de datos predichos correctamente de todos los puntos de datos. Más formalmente, se define como el número de verdaderos positivos y verdaderos negativos dividido por el número de verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos [42].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.1)$$

Donde: TP corresponde a Verdaderos Positivos, TN Verdaderos Negativos, FP falsos positivos y FN falsos negativos.

Un verdadero positivo o verdadero negativo es un punto de datos que el algoritmo clasificó correctamente como verdadero o falso, respectivamente. Un falso positivo o falso negativo, por otro lado, es un punto de datos que el algoritmo clasificó incorrectamente[42].

2.1.3.2 Intersection over Union - IoU

Intersection over Union - IoU , también conocido como índice Jaccard, es la métrica más utilizada para comparar la similitud entre dos formas arbitrarias. IoU codifica las propiedades de forma de los objetos en comparación, por ejemplo, los

anchos, alturas y ubicaciones de dos cuadros delimitadores, en la propiedad de la región y luego calcula una medida normalizada que se centra en sus áreas (o volúmenes). Esta propiedad hace que IoU sea invariante a la escala del problema en consideración. Debido a esta atractiva propiedad, todas las medidas de rendimiento utilizadas para evaluar la segmentación, la detección de objetos y el seguimiento se basan en esta métrica[43].

Se utiliza para determinar verdaderos positivos y falsos positivos en un conjunto de predicciones. Cuando se utiliza IoU como métrica de evaluación, se debe elegir un umbral de precisión[43].

$$IoU = \frac{A \cap B}{A \cup B} \quad (2.2)$$

Para efectos de este trabajo, se plantea la validación de los resultados de la identificación de la Región de Interés bajo esta métrica.

2.2 DICOM

DICOM, por sus siglas en inglés Digital Imaging and Communications in Medicine[5], es un estándar internacional de comunicaciones para imágenes médicas e información relacionada a dichas imágenes de la NEMA también por sus siglas en inglés, National Electrical Manufacturers Association[6].

El estándar DICOM, define el formato en que están representadas imágenes médicas, la estructura de la metadata y el protocolo de comunicación para intercambiar los datos entre dispositivos. DICOM es un estándar utilizado en la mayoría de dispositivos de radiología, cardiología, radioterapia (Rayos X, CTm MRI, ultrasonido, etc)[7, 44].

Uno de los requisitos del cliente, fue el uso de archivos DICOM. El primer objetivo fue encontrar un dataset que cumpliera con estas características, luego de explorar varios, se opta por usar el de mamografías (CBIS-DDSM) [45]. El uso del mismo implicó varios problemas incluso antes de comenzar la exploración. Los problemas más relevantes se detallan a continuación:

- El tamaño del dataset [45] es de 2 TB, lo que implicó que no se incluyera Google Colab [46] como posible herramienta para utilizar con dicho dataset. Se utilizó el cluster de la Universidad ORT [47]

- Para la lectura de los archivos, fue necesario investigar/probar distintas librerías que permitieran leer la metadata y el contenido de la imagen, para finalmente seleccionar la librería pydicom [48].
- Convertir la información de la imágenes para que permitiera la manipulación en Python y posteriormente por las diferentes redes.

Una vez mitigados estos problemas, se inicia con la exploración de los datos. Antes de ver las metadatos e imágenes por implementación de código, se usó la herramienta ImageJ [49]. La figura 2.4 es un ejemplo de visualización de un archivo DICOM en dicha herramienta.

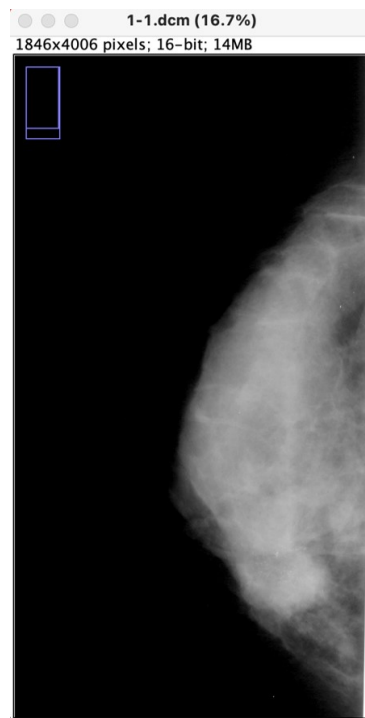


Figura 2.4: Ejemplo de DICOM [5] utilizando ImageJ [49]

Luego desde Python, se utilizó la librería pydicom [48] para extraer la información de la metadata, a modo de ejemplo se puede ver la información obtenida para una imagen en las figuras 2.5 y 2.6.

```

<class 'pydicom.dataset.FileDataset'>
Dataset.file_meta -----
(0002, 0000) File Meta Information Group Length  UL: 194
(0002, 0001) File Meta Information Version       OB: b'\x00\x01'
(0002, 0002) Media Storage SOP Class UID        UI: Secondary Capture Image Storage
(0002, 0003) Media Storage SOP Instance UID     UI: 1.3.6.1.4.1.9590.100.1.2.326970568112801613524204885792264795431
(0002, 0010) Transfer Syntax UID               UI: Implicit VR Little Endian
(0002, 0012) Implementation Class UID          UI: 1.2.40.0.13.1.1.1
(0002, 0013) Implementation Version Name       SH: 'dcm4che-1.4.35'
-----
(0008, 0005) Specific Character Set            CS: 'ISO_IR 100'
(0008, 0016) SOP Class UID                     UI: Secondary Capture Image Storage
(0008, 0018) SOP Instance UID                  UI: 1.3.6.1.4.1.9590.100.1.2.326970568112801613524204885792264795431
(0008, 0020) Study Date                       DA: '20160720'
(0008, 0023) Content Date                     DA: '20160426'
(0008, 0030) Study Time                       TM: '191745'
(0008, 0033) Content Time                     TM: '134255.140000'
(0008, 0050) Accession Number                 SH: ''
(0008, 0060) Modality                         CS: 'MG'
(0008, 0064) Conversion Type                  CS: 'WSD'
(0008, 0090) Referring Physician's Name       PN: ''
(0008, 103e) Series Description                LO: 'full mammogram images'
(0010, 0010) Patient's Name                    PN: 'Mass-Training_P_01431_LEFT_MLO'
(0010, 0020) Patient ID                       LO: 'Mass-Training_P_01431_LEFT_MLO'
(0010, 0030) Patient's Birth Date             DA: ''
(0010, 0040) Patient's Sex                    CS: ''
(0013, 0010) Private Creator                  LO: 'CTP'
(0013, 1010) Private tag data                 UN: b'CBIS-DDSM '
(0013, 1013) Private tag data                 UN: b'43352602'
(0018, 0015) Body Part Examined               CS: 'BREAST'
(0018, 1016) Secondary Capture Device Manufactur LO: 'MathWorks'
(0018, 1018) Secondary Capture Device Manufactur LO: 'MATLAB'
(0020, 000d) Study Instance UID               UI: 1.3.6.1.4.1.9590.100.1.2.338694788812105197809861717921252838478
(0020, 000e) Series Instance UID             UI: 1.3.6.1.4.1.9590.100.1.2.265953197912973660327069359641506533925
(0020, 0010) Study ID                         SH: 'DDSM'
(0020, 0011) Series Number                    IS: "1"
(0020, 0013) Instance Number                  IS: "1"
(0020, 0020) Patient Orientation              CS: 'MLO'
(0020, 0060) Laterality                       CS: 'L'
(0028, 0002) Samples per Pixel                US: 1
(0028, 0004) Photometric Interpretation       CS: 'MONOCHROME2'

```

Figura 2.5: DICOM metadata parte 1

```

(0028, 0010) Rows                             US: 5446
(0028, 0011) Columns                          US: 2581
(0028, 0100) Bits Allocated                   US: 16
(0028, 0101) Bits Stored                      US: 16
(0028, 0102) High Bit                        US: 15
(0028, 0103) Pixel Representation             US: 0
(0028, 0106) Smallest Image Pixel Value       US: 0
(0028, 0107) Largest Image Pixel Value        US: 65535
(7fe0, 0010) Pixel Data                       OW: Array of 28112252 elements

```

Figura 2.6: DICOM metadata parte 2

En la figura 2.6 se puede observar la propiedad pixel data, la cual contiene la información relacionada a la imagen[50]. Utilizando dicha propiedad se desplegó las imágenes, mascarar y crops del dataset de mamografías.

Por medio de la investigación realizada al estándar de DICOM [7] se observó que existe el campo ROI dentro de la metadata[51]. Sin embargo es un módulo que se encuentra en algunos DICOM y en el caso del dataset de las mamografías (CBIS-DDSM) [45] no se encuentra disponible.

2.3 Datasets

La selección de los conjuntos de datos utilizados está enfocada en el cumplimiento del objetivo principal del trabajo correspondiente a la identificación de la Región de interés (ROI) en imágenes médicas en formato DICOM, para esto se recurre a una de las principales fuentes disponibles como es la “The Cancer Imaging Archive - TCIA” [52].

La TCIA es un proyecto financiado por el Programa de Imágenes de Cáncer del Instituto Nacional del Cáncer y utiliza el software NBIA en su núcleo para facilitar las descargas. La TCIA aloja un gran archivo de imágenes médicas de cáncer accesibles para descarga pública. Los datos están organizados como “colecciones”; típicamente imágenes de pacientes relacionadas con una enfermedad común (por ejemplo, cáncer de pulmón), modalidad o tipo de imagen (resonancia magnética, tomografía computarizada, histopatología digital, etc.) o foco de investigación[53].

Los criterios utilizados para la selección de los datasets utilizados en este trabajo, además de haber sido elegidos por ser imágenes provenientes de estudios médicos y contenidas bajo el formato de imágenes DICOM[7] y que incluyeran información de la Región de Interés - ROI, objetivos principales, se buscó que incluyeran suficiente número de datos y que las imágenes contenidas contaran con la calidad suficiente necesaria para el entrenamiento de redes neuronales.

Los datasets seleccionados para el trabajo se encuentran descritos a continuación:

- Curated Breast Imaging Subset of DDSM (CBIS-DDSM) - The Cancer Imaging Archive (TCIA)[53].
- Brain MRI segmentation[54] (TCG – LGG)[55] - The Cancer Imaging Archive (TCIA)[53].

2.3.1 Curated Breast Imaging Subset of DDSM (CBIS-DDSM)

El subconjunto de datos de imágenes de mamografías seleccionadas CBIS - DDSM, corresponde a una actualización y estandarización de la base de datos digital de mamografías de detección DDSM, la cual contiene información patológica de casos normales, benignos y malignos seleccionadas por un mamógrafo capacitado. Las imágenes originales han sido descomprimidas y convertidas a formato DICOM y se ha incluido su respectiva máscara de segmentación y recorte de la región de interés[53].

En la figura 2.7 se observa un ejemplo de visualización de las imágenes contenidas en el dataset DDSM por medio de la herramienta ImageJ[49].



Figura 2.7: Ejemplo de imágenes del dataset DDSM

El subconjunto contiene la información correspondiente a 691 pacientes para un total de 1318 imágenes. Se puede ver en el histograma en la figura 2.8 la cantidad de ejemplos por clase con el que cuenta el dataset.

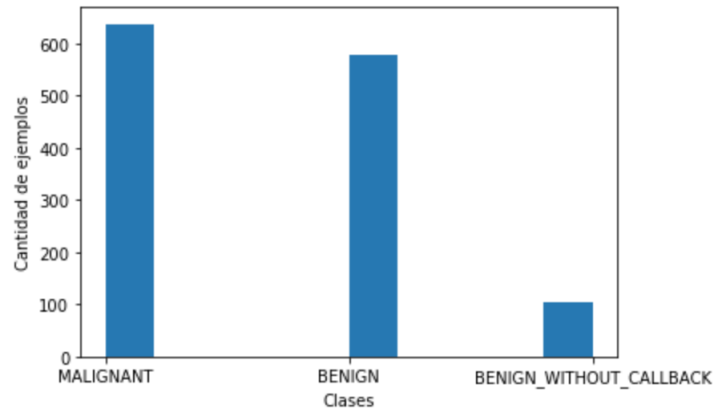


Figura 2.8: Cantidad de ejemplos que contienen clases Malignant, Benign y Benign without callback. Para el dataset de Mamografías

A efectos de este trabajo, la clase “Benign without callback” sera considerada como benigna (Benign). Ya que, no es de interés para los objetivos, la distinción entre estas dos clases.

2.3.2 Brain MRI segmentation - The Cancer Imaging Archive(TCIA)

Basado en la experiencia obtenida en la exploración de los datos del primer dataset conformado por imágenes de mamografías en formato DICOM y entendiendo que con esto ya se cumple en parte con el objetivo del estudio, se procede a realizar una búsqueda de un segundo dataset compuesto de imágenes médicas, pero no limitadas al formato DICOM. De forma que las imágenes contasen con el proceso de transformación a formato convencional ya realizado, permitiendo enfocarse en la implementación de los algoritmos seleccionados para la identificación de la región de interés.

EL dataset seleccionado corresponde a una versión resumida y transformada del conjunto de imágenes de TCGA – LGG de la The Cancer Imaging Archive - TCIA[55]. Y ha sido usado en el estudio “Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm” [56] y extraído de kaggle y renombrado como Brain MRI segmentation[54].

El conjunto de datos de segmentación de imágenes de resonancia magnética del cerebro contiene información de imágenes con sus respectivas máscaras segmentadas de forma manual con la información correspondiente a la ROI, en total

el dataset cuenta con 2090 imágenes con su respectiva máscara procedentes de 111 pacientes. En la Figura 2.9 se puede observar algunos ejemplos en donde al lado izquierdo se encuentra la imagen, la máscara correspondiente a esta imagen y al lado derecho la imagen y la máscara superpuesta sobre la ROI, en este caso tumores.

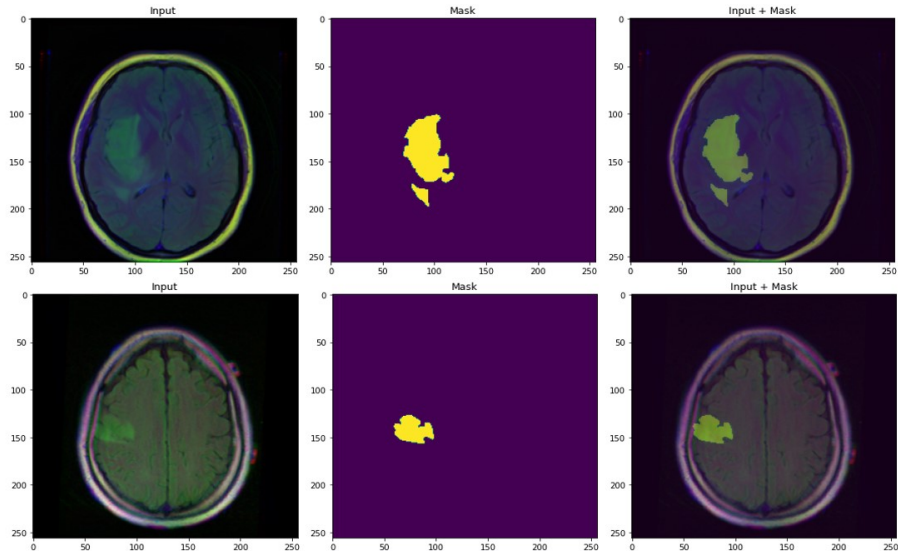


Figura 2.9: Imagen y máscara Brain MRI segmentation[54]

En la figura 2.10 se observa una visualización de las imágenes contenidas en el dataset de TCGA – LGG representadas con la herramienta ImageJ[49].

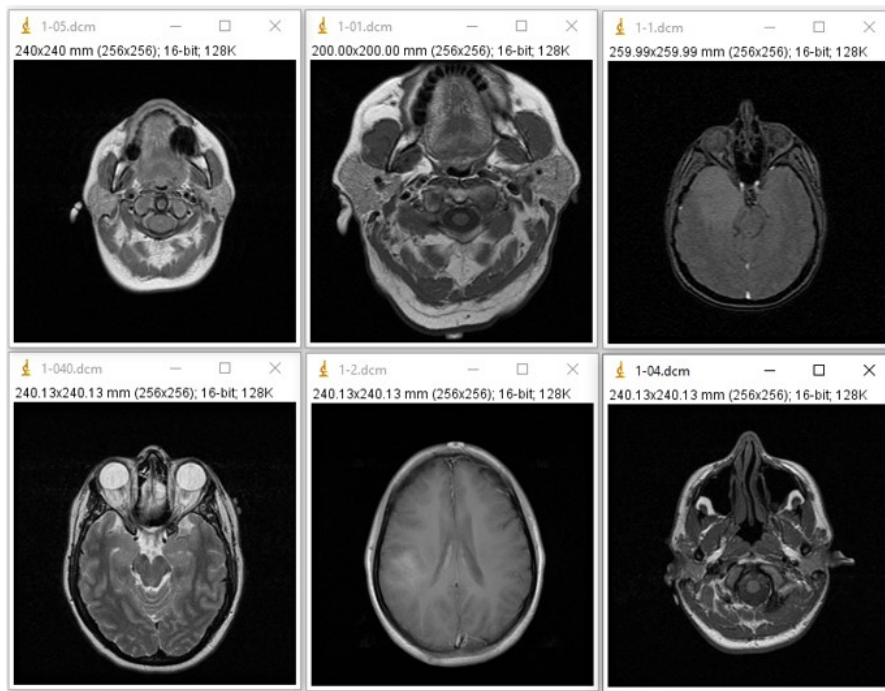


Figura 2.10: Ejemplo de imágenes del dataset TCGA-LGG[55]

Dentro del conjunto de datos seleccionado se encuentran en promedio entre 18 y 20 imágenes por paciente, de las cuales un 34.36 % contienen información referente a la región de interés y un 65.64 % en con región de interés es nula, como se puede ver en el histograma en la figura 2.11.

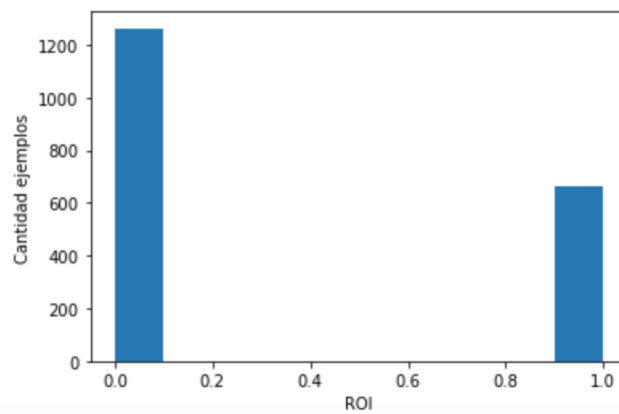


Figura 2.11: Cantidad de ejemplos que contienen ROI (1) y que no (0), para el dataset Brain MRI segmentation

2.4 Técnica 1: YOLO

En el presente trabajo, se denomina técnica 1 a una adaptación de la arquitectura planteada en el paper: “Automated Breast Cancer Diagnosis Using Deep Learning and Region of Interest Detection (BC-DROID)” [34]

Los autores del paper adaptan la arquitectura YOLO [11] pasando por dos procesos de entrenamiento. En el primero llamado “pre-entrenamiento” se encarga de clasificar los recortes de los tumores en dos categorías malignos y benignos. En el segundo denominado “entrenamiento”, se toman los pesos de la etapa de pre-entrenamiento y son utilizados para detectar la ROI en las imágenes.

En el presente trabajo, el código fue escrito completamente por los integrantes del grupo, ya que los autores no publicaron su implementación. Como ayuda en el desarrollo de la elaboración, se utilizó un tutorial de YOLO V1 from scratch [57].

2.4.1 Arquitectura

Como se observa en la figura 2.12, la arquitectura contiene dos redes, cada una de ellas con sus propios objetivos. La red de la izquierda tiene la responsabilidad de clasificar los recortes (crops) de las ROI en maligno y benigno. Mientras que la red de la derecha es responsable de la detección de la ROI en las imágenes.

Como ya fue mencionado, durante la clasificación, la red de la izquierda espera recibir recortes de las ROI de cada imagen. El objetivo de esta red es clasificar los distintos crops en benigno o maligno. Luego de este pre-entrenamiento los pesos de las capas convolucionales son transferidos a las capas convolucionales de la parte derecha.

En la detección de la ROI que se realiza en la red de la derecha, la entrada corresponde a imágenes completas de mamografías y la salida a tensores, siguiendo el formato de definido YOLO: [11].

$$[P_c, x_c, y_c, w, h, P(\text{Maligno}), P(\text{Benigno})]$$

[11]

La cantidad de tensores de salida por imagen esta dada por la siguiente fórmula:

$$S * S * (B * (C + 5))$$

[11]

Donde S es la cantidad de divisiones que se realizan a la imagen, C la cantidad de clases y B la cantidad de bounding boxes.[11]

En este modelo se utilizaron los valores S=7, C=2 y B=1, Es una configuración simple de YOLO, ya que sólo hay una predicción de bounding box. Observando la función de loss 2.3 definida en el paper de Platania[34] el cual omite la sumatoria de los bounding boxes con la función de loss definida en YOLO [11], en el presente trabajo se asumime que es utilizado B=1.

Por lo tanto se espera tener tensores de $[7 * 7 * 7]$ para cada imagen.

A continuación se detallan las capas y los componentes utilizados. El modelo esta constituido por 9 capas convolucionales para los 2 modos, donde las primeras 6 capas están conectadas a max pooling.

Para el pre-entrenamiento se utiliza dos capas fully connected, LeakyReLU con valor 0.1 como función de activación, cross-entropy como función de loss y ADAM como optimizador.

En el entrenamiento no se utilizan las capas fully connected del pre-entrenamiento (ni sus pesos). Entonces se suman a las 9 capas convolucionales 3 nuevas capas fully connected.

Capa	in channels	out channels	Kernel	Stride	Padding
BatchNorm	1	-	-	-	-
Conv2d	1	16	(3x3)	0	1
MaxPool2d	-	-	(2x2)	2	0
Conv2d	16	32	(3x3)	0	1
MaxPool2d	-	-	(2x2)	2	0
Conv2d	32	32	(3x3)	0	1
MaxPool2d	-	-	(2x2)	2	0
Conv2d	32	64	(3x3)	0	1
MaxPool2d	-	-	(2x2)	2	0
Conv2d	64	128	(3x3)	0	1
MaxPool2d	-	-	(2x2)	2	0
Conv2d	128	64	(3x3)	0	1
MaxPool2d	-	-	(2x2)	2	0
Conv2d	64	32	(3x3)	0	1
Conv2d	32	64	(3x3)	0	1
Conv2d	64	128	(3x3)	0	1
Dropout	-	-	-	-	-
Linear (clas)	512	50	-	-	-
Linear (clas)	50	2	-	-	-
Dropout	-	-	-	-	-
Linear (ROI)	1152	2000	-	-	-
Dropout	-	-	-	-	-
Linear (ROI)	2000	1000	-	-	-
Linear (ROI)	1000	$S * S * (C + B * 5)$	-	-	-

Tabla 2.1: Detalle de las capas

Si bien la mayoría de la arquitectura esta basada en el paper (BC-DROID) [34], en este trabajo se realizaron cambios que dieron mejor resultado durante los experimentos. El detalle de las capas es posible consultarlo en la tabla 2.1

En cuanto a la función de activación se vuelve a utilizar LeakyReLU con valor 0.1 y como ADAM como optimizador.

A continuación se define la función de loss utilizada [34, 11] para determinar ROI (parte derecha de la red.)

$$loss = \sum_i^{k^2} \alpha(loss_{coo}^i + loss_{box}^i) + loss_{cls}^i + loss_{cnf}^i \quad (2.3)$$

Cada uno de los distintos componentes se detalla a continuación:

$$loss_{coo}^i = \begin{cases} (x_i - \widehat{x}_i)^2 + (y_i - \widehat{y}_i)^2, & \text{ROI present} \\ 0, & \text{otherwise} \end{cases} \quad (2.4)$$

Representa la función de loss entre las coordenadas del centro de cada celda.

$$loss_{box}^i = \begin{cases} (\sqrt{w_i} - \sqrt{\widehat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\widehat{h}_i})^2, & \text{ROI present} \\ 0, & \text{otherwise} \end{cases} \quad (2.5)$$

Representa la función de loss entre el tamaño de la caja.

$$loss_{cls}^i = \begin{cases} \sum_j (P_i[j] - \widehat{P}_i[j]), & \text{ROI present} \\ 0, & \text{otherwise} \end{cases} \quad (2.6)$$

Representa la función de loss de clasificación.

$$loss_{cnf}^i = \begin{cases} (c_i - \widehat{c}_i)^2, & \text{ROI present} \\ \beta(c_i - \widehat{c}_i)^2, & \text{otherwise} \end{cases} \quad (2.7)$$

Representa la función de loss de si existe o no objeto en una celda.

Es difícil intuir el comportamiento o encontrar los “mejores” parámetros de α o β , los autores seleccionan 5 y 0.4 respectivamente, [34] dado que la función de loss es una composición de otras[11]. Para tener un mejor entendimiento de lo que está sucediendo durante el entrenamiento, se grafica cada uno de los componentes de la función de loss.

Los autores [34] utilizan SGD como optimizador y ReLU como activación. Se selecciona ADAM y no SGD, ya que es casi un estándar dentro de la comunidad por sus resultados. Se utiliza LeakyReLU basado en el paper de YOLO [11]. Además a la arquitectura se agregó una capa de Dropout con 0.5 de probabilidad al igual que en YOLO [11].

2.4.2 Implementación 1 Mamografías

La primera implementación de esta técnica se probó con el dataset de mamografías (descrito en 2.3.1) por dos motivos principales. En primer lugar analizan mamografías en el paper dónde se publica esta arquitectura [34]; permitiendo co-

robustecer que los resultados sean similares incluso cuando el código fue escrito completamente por los integrantes del grupo. En segundo lugar, como ya se comentó antes, este dataset contiene imágenes en formato DICOM [5] que es uno de los objetivos de investigación del trabajo.

Las imágenes de las mamografías, sus máscaras y crops[45] están en formato DICOM[5]. Como consecuencia, no es posible enviar dichos archivos directamente o con técnicas convencionales al modelo. El formato DICOM es un estándar[5] con una estructura particular, fue utilizada la librería pydicom[48] para extraer la metadata y dentro de esta, la propiedad pixel array que contiene la información de la imagen.

Durante el desarrollo se observaron varios problemas: el primero, relacionado con la posibilidad de visualizar los DICOM [5] como imágenes PIL, ya que la imagen se veía en blanco o negro y no en escala de grises, como debería. Se utilizaron varias funcionalidades de la librería Pydicom [48], para lograr visualizarlas correctamente.

El segundo problema fue que la red demoraba mucho en entrenar, luego de investigar el problema, se observó que se trataba por el resize en la etapa de data augmentation. Esto implicó un cambio en la implementación, debiendo extraer la información de la imagen, aplicar el resize y guardarlas en disco como archivos [.npy]. El tamaño seleccionado es (200x200) para las imágenes y las máscaras y (128x128) para los crops.

Como tercer problema, era necesario obtener las coordenadas de las ROI ya que YOLO [11] utiliza esta información y no está presente como dato en el dataset [45], utilizando cada máscara fueron calculadas 4 coordenadas (xmin, ymin, xmax, ymax) con la que fue posible formar un rectángulo conteniendo la ROI.

Como resultado de lo implementado, ver la Figura 2.13 muestra a la izquierda una mamografía y a la derecha su correspondiente máscara, señalando la caja que contiene la ROI.

2.4.2.1 Clasificación

Se puede observar en la figura 2.14 las curvas de training/validation para la loss.

Luego en la figura 2.15 la gráfica precisión correspondiente:

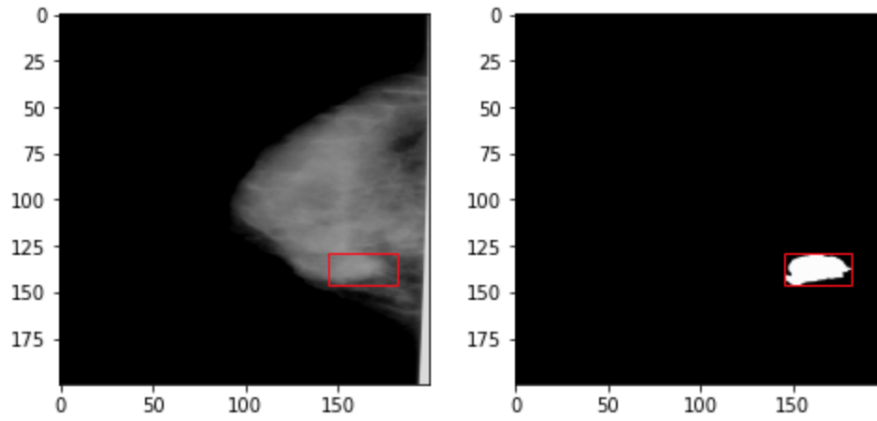


Figura 2.13: Imagen y máscara del dataset de mamografías [45]

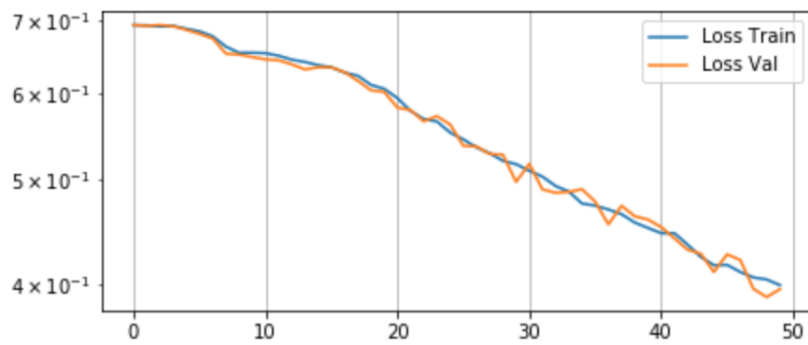


Figura 2.14: Gráfica de train y validation para clasificación

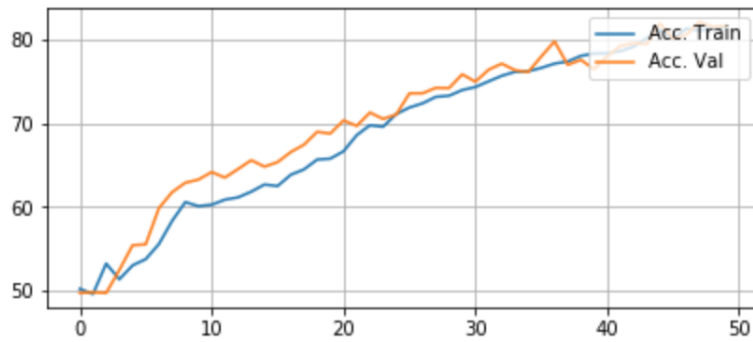


Figura 2.15: Gráfica de precisión train y validation para clasificación

Se puede observar en los ejemplos 2.16 el funcionamiento del clasificador que llegó a una precisión de 71.96.

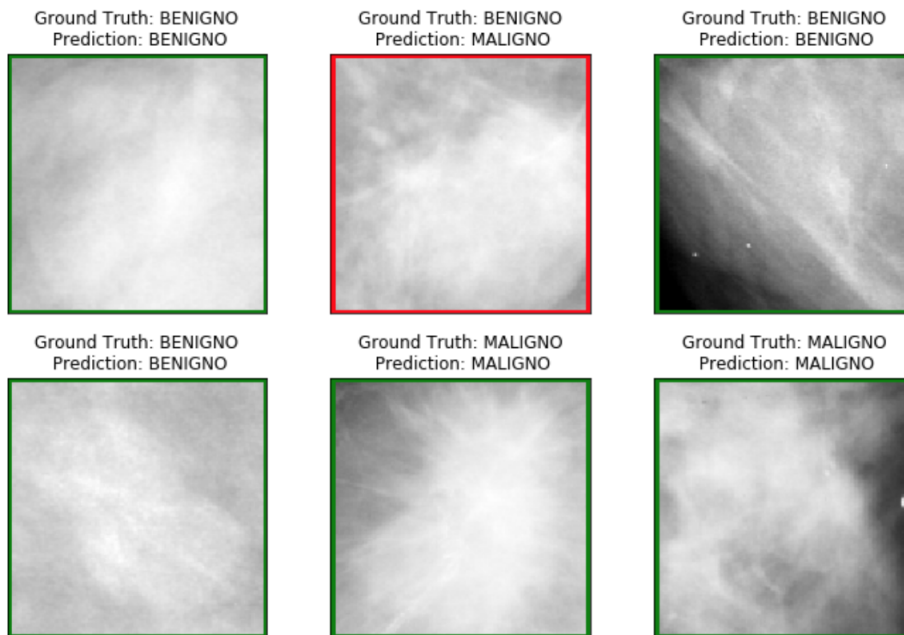


Figura 2.16: Resultado de la clasificación para las mamografías

2.4.2.2 Detección Región de Interés

Luego de entrenada la parte de la izquierda de la red, es necesario utilizar esos pesos para entrenar la parte derecha y comenzar el entrenamiento de las ROI. Durante esta etapa fueron usados distintos hiperparámetros los cuales podemos observar en la tabla 2.2 y sus resultados en el anexo 6.2.1.

Exp	Learning Rate	Epochs	Alfa	Beta	W.Decay	Max IoU (test)
0	1e-5	200	5	0.5	1e-5	0.44
1	1e-5	200	5	0.7	1e-5	0.53
2	1e-5	200	10	0.2	1e-5	0.76
3	1e-3	200	3	0.4	1e-5	0.69
4	1e-4	200	1	0.4	1e-5	0.65
5	2e-4	200	1	0.4	1e-5	0.70
6	2e-5	200	1	0.4	1e-5	0.84
7	8e-10	200	1	0.4	1e-5	0.0035
8	4e-5	200	5	0.4	1e-5	0.59
9	8e-6	200	3	0.3	1e-5	0.65
10	2e-5	200	1	0.4	1e-5	0.63
11	2e-5	200	6	0.4	1e-5	0.71
12	2e-5	200	0.8	0.4	1e-5	0.42

Tabla 2.2: Hiperparámetros ROI dataset mamografías

Después de entrenar los 12 modelos, se observó que el que brindaba mejor resultado correspondía al número 5, del cual se muestran sus correspondientes gráficas 2.17

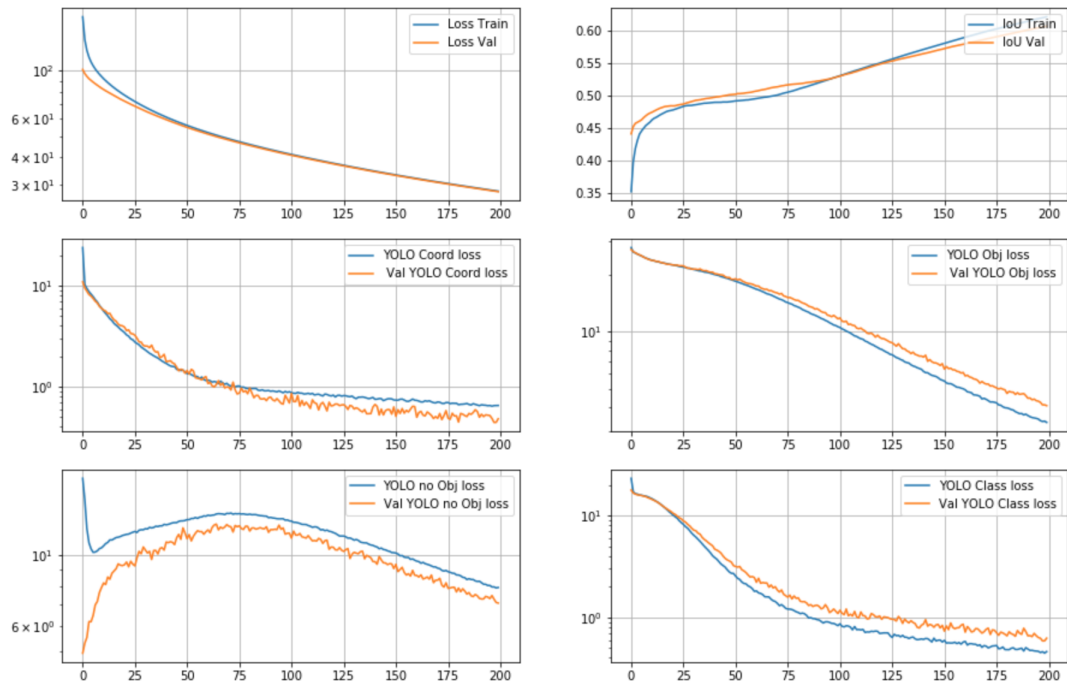


Figura 2.17: ROI Gráficas de entrenamiento modelo 5

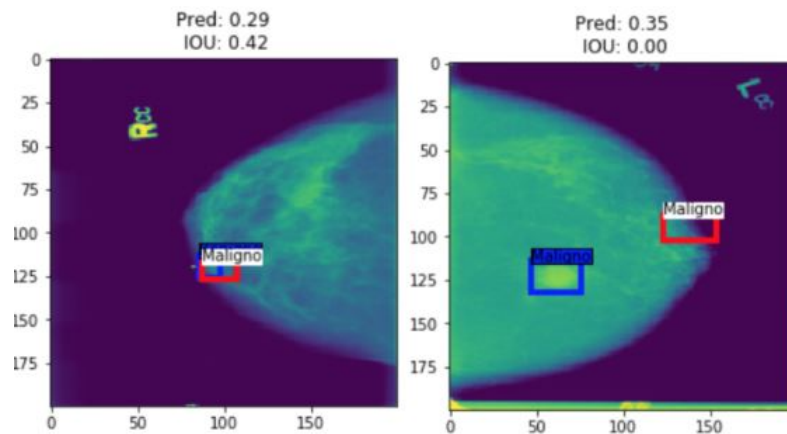


Figura 2.18: Ejemplos de predicciones del modelo con datos de test de mamografías, en rojo se marca la predicción y en azul el target. [45].

A continuación se encuentran los histogramas de train, validation y test:

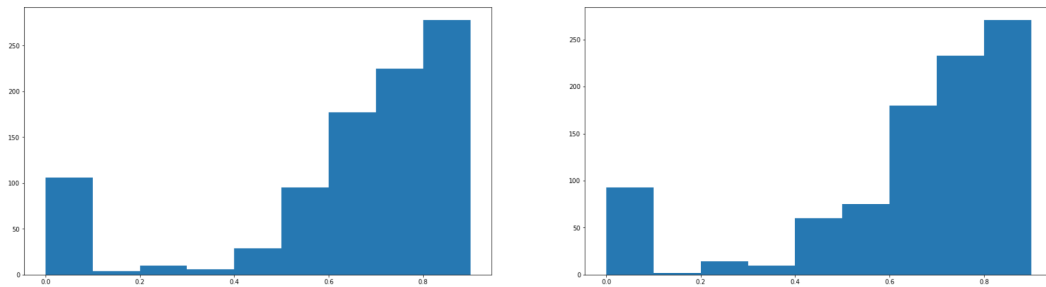


Figura 2.19: Histograma dataset mamografías en train, en el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

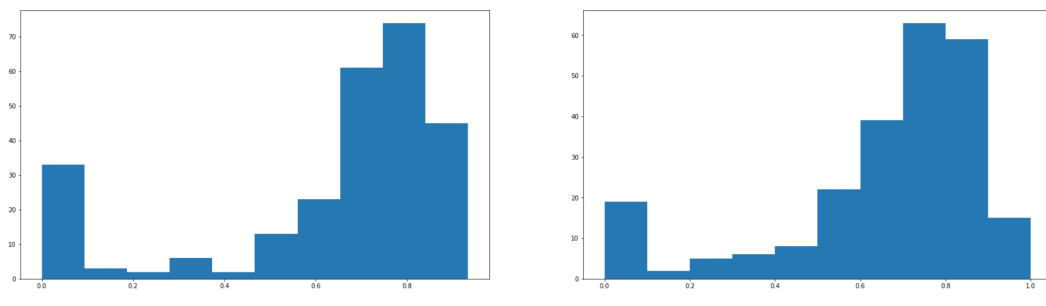


Figura 2.20: Histograma dataset mamografías en validation, en el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

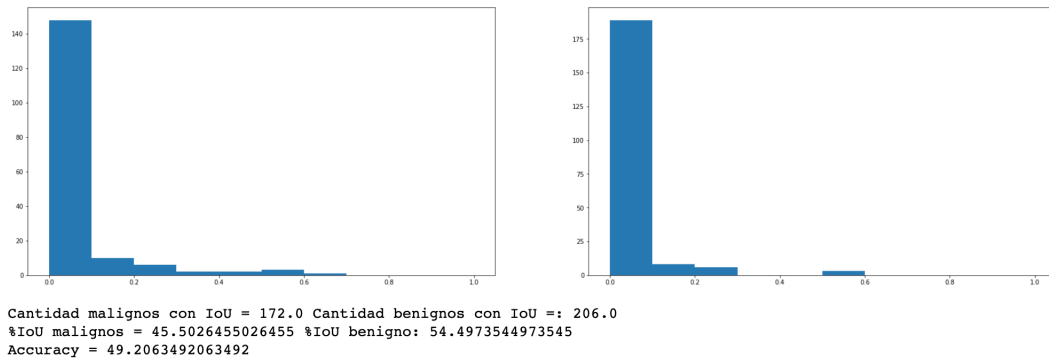


Figura 2.21: Histograma para testing, en el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

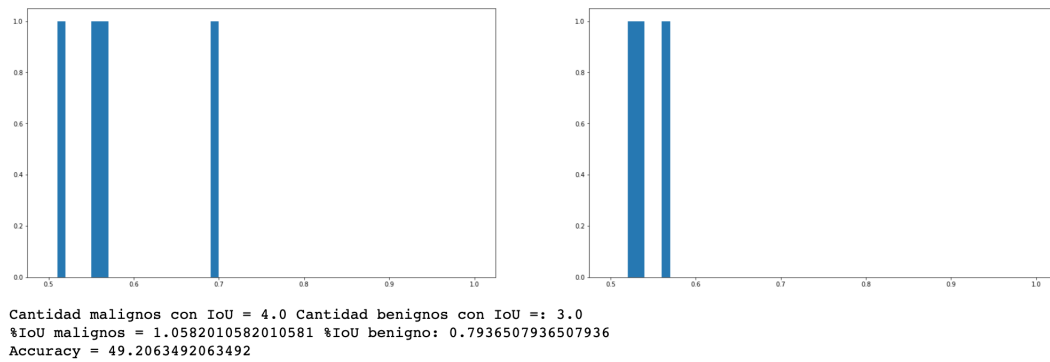


Figura 2.22: Histograma para testing, en el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU. Rango de 0.5 a 1

El total de datos clasificados por la red de la derecha, con etiquetas correspondientes a tumores malignos son 172 (45.50%) y 206 (54.50%) benignos. El accuracy de clasificación es de 49.21%.

Se puede considerar que el modelo tiene un mal rendimiento en el dataset de test con respecto al IOU, con muy pocos valores que superan 0.5. Sin embargo vemos que en train y validation, se obtienen buenos resultados con respecto al IOU. El modelo fue entrenado durante 200 epochs y por lo que se observa en las gráficas de la Figura 2.17, no hay presencia de overfitting.

Debido a los tiempos acotados de este trabajo, no es posible realizar más experimentos para mejorar la performance de la técnica. De todas maneras, se considera importante haber logrado la implementación del modelo y que el mismo entrene.

2.4.3 Implementación 2 MRI

Una vez probada la implementación con el dataset de las mamografías, el siguiente paso fue probar la misma arquitectura con datos diferentes. La idea es comparar el comportamiento del modelo cuando se le presentan imágenes de otra naturaleza, a la vez que se prueba la técnica con un dataset distinto al descrito en el paper donde se presenta esta arquitectura [34].

Para este segundo experimento se utilizó el dataset de imágenes de resonancias magnéticas del cerebro presentado en el capítulo 2.3.2. En esta sección se describen los cambios realizados para trabajar con estas imágenes y los resultados obtenidos, que fueron comparados con los del primer dataset descrito en el capítulo 2.3.1.

A diferencia del dataset de las mamografías [45], el dataset de los MRI Brain Segmentation[54] no brinda los recortes ni las etiquetas. Por lo tanto, fue necesario procesar las imágenes del dataset y realizar los recortes correspondientes de máscara a las imágenes. Para la implementación de la arquitectura con este dataset, se observó que no todas las imágenes contienen ROI, por lo que para la parte derecha y clasificación de las imágenes se determinó que si tiene máscara (píxeles distintos al negro) entonces tiene un tumor maligno en caso contrario, el paciente no tiene tumor. Para mantener las etiquetas y favorecer la comparación, si no tiene un tumor, es etiquetado como benigno.

2.4.3.1 Clasificación

Luego de preparadas las imágenes, máscaras y los crops. Se inicia el entrenamiento de la parte izquierda de la red (clasificación). Una diferencia con las imágenes del dataset de mamografías [45], es que sólo tienen un canal (escala de grises), las imágenes de este dataset [54] tiene 3 canales (RGB).

Al entrenar parte de la izquierda se obtienen las siguientes gráficas:

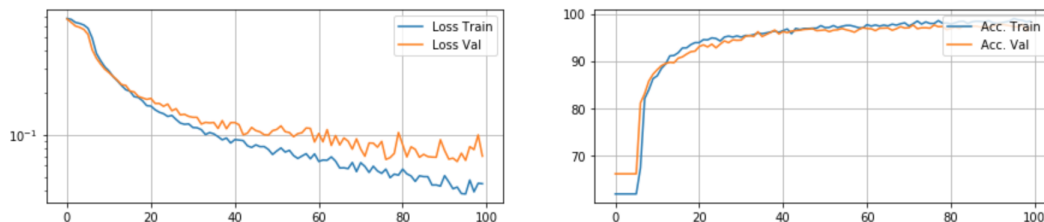


Figura 2.23: Clasificación loss y accuracy

El modelo llega a una accuracy de 97% para validación y loss 0.070316. Se puede observar ejemplos de la clasificación sobre el dataset de test.

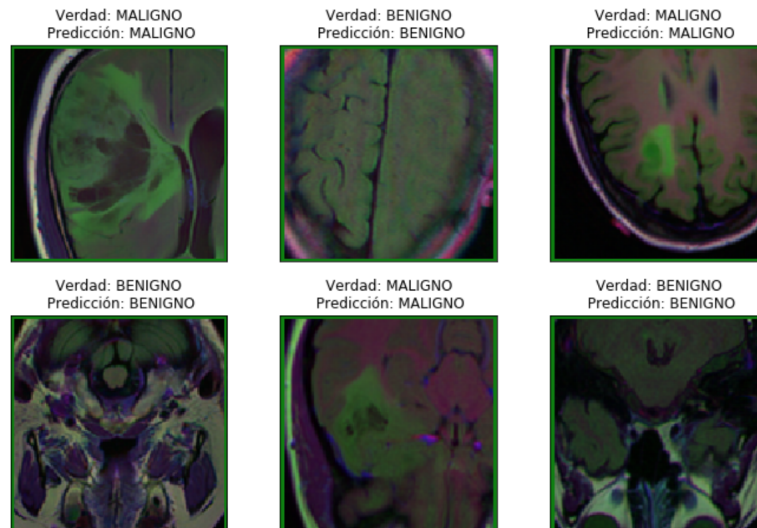


Figura 2.24: Resultado de la clasificación en el dataset de MRI Brain Segmentation

Se puede observar que la parte izquierda de la red genera buenos resultados, y se procede a entrenar la parte de la derecha y comprobar que detecta las ROI correctamente.

2.4.3.2 Detección de Región de Interés

Continuando con el entrenamiento se utilizan los pesos de la clasificación de las capas convoluciones para la detección, al igual que se hizo en el caso de las mamografías. Para este dataset, no todas las imágenes tienen máscara con información, si bien el archivo existe, ese se ve como una imagen negra. Para el presente trabajo se opta por ver los resultados de los modelos, sin excluir a dichas imágenes.

Los experimentos se realizaron con los hiperparámetros que se ven en la tabla 2.3 y sus resultados en el anexo 6.2.2.

Exp	Learning Rate	Epochs	Alfa	Beta	W.Decay	Max IoU (test)
0	1e-2	200	5	0.4	1e-5	1.0
1	1e-3	200	5	0.4	1e-5	1.0
2	1e-4	200	5	0.4	1e-5	1.0

Tabla 2.3: Hiperparámetros ROI dataset MRI Brain Segmentation[54]

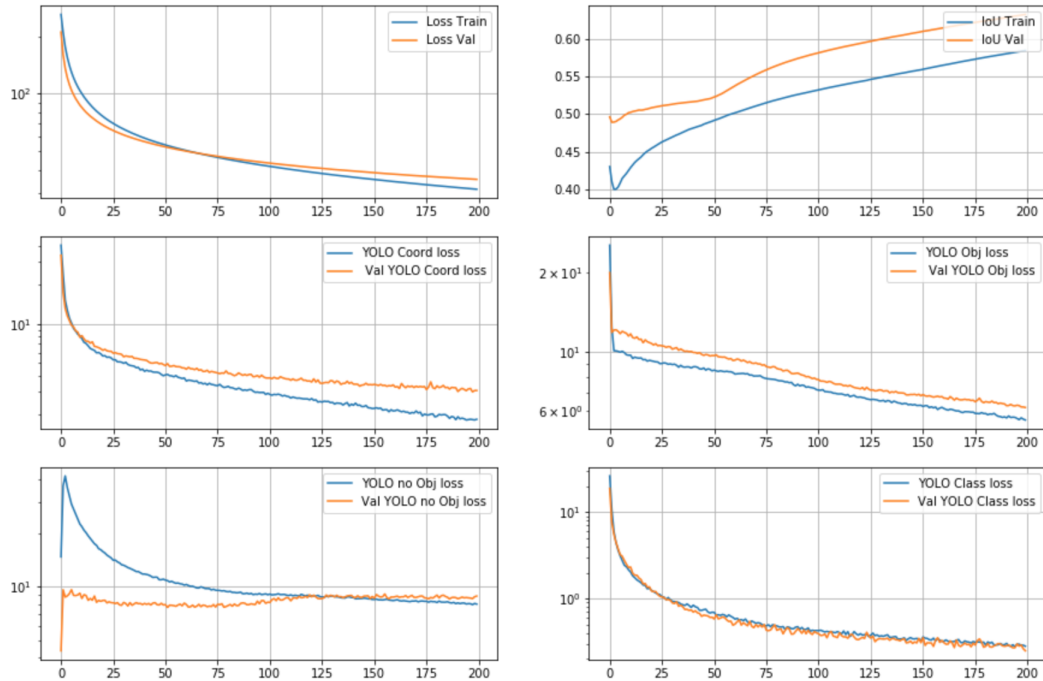


Figura 2.25: Entrenamiento de ROI modelo 0

Observando las gráficas, no hay señales de overfitting en train/validation, el IOU se mantiene creciente y es posible que con un mayor número de Epochs se vean mejoras. (actualmente ha sido entrenado con 200 Epochs).

A continuación, se observan histogramas para determinar la cantidad de ejemplos por valor IoU y clase.

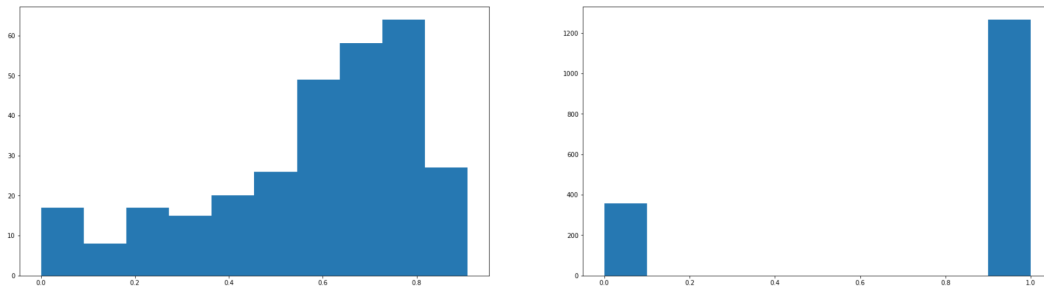


Figura 2.26: Histograma para el entrenamiento, en el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

El total de clasificados en entrenamiento, etiquetados con tumores malignos son 301 (15.64%) y los benignos 1623 (84.36%)

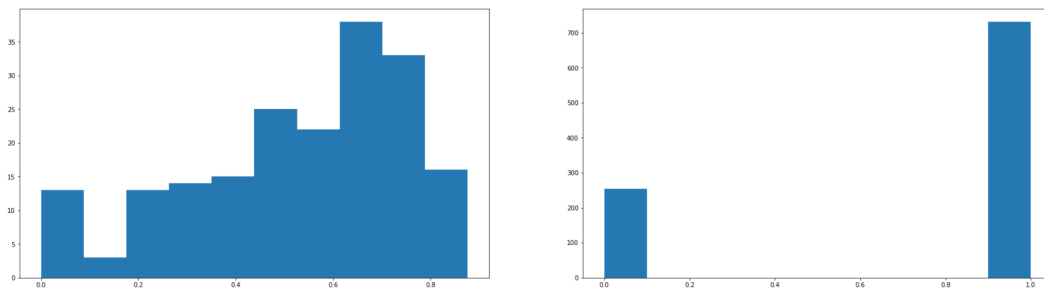


Figura 2.27: Histograma para la validación, en el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

El total de clasificados en validación, etiquetados con tumores malignos son 192 (16.28%) y los benignos 987 (83.72%)

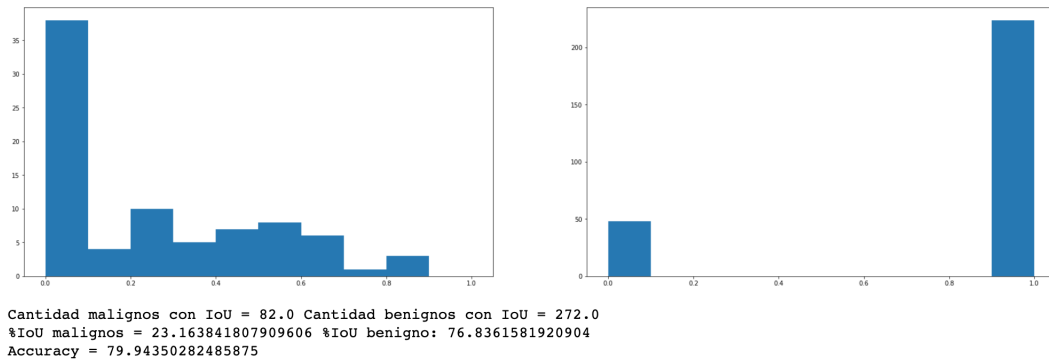


Figura 2.28: Histograma para testing, en el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

El total de clasificados en testing, etiquetados como tumores malignos corresponde a 91 (25.7 %) y los benignos 263 (74.7 %) La precisión de clasificación es de 75.71 %.

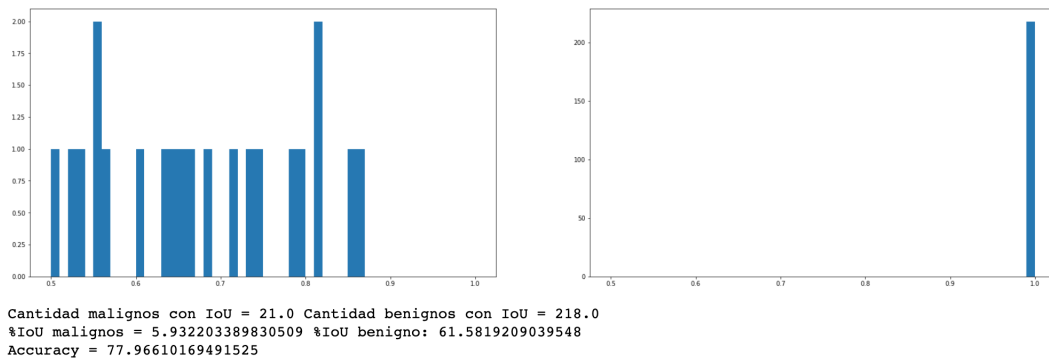


Figura 2.29: Ejemplos de predicciones del modelo con datos de test de MRI Brain Segmentation[54]

Se puede observar en la Figura 2.29 que cuando no existe una ROI, ejemplo 4 y 6 el IOU es igual a 1 debido, a que los valores de las coordenadas son iguales a 0. Para los otros casos, ejemplos 1,2,3 y 5, los valores de IOU indican la relación entre la caja real y la predicha.

2.4.4 Comparación de resultados

En la tabla 2.4 se muestra distintas comparaciones entre los dos datasets aplicando esta arquitectura.

Comparativa	Mamografías	MRI
Precisión Clasificación	71.96 %	97 %
IOU \geq 0.5	1.85 %	67.51 %
IOU \geq 0.5 Malignos	1.06 %	5.93 %
IOU \geq 0.5 Benignos	0.79 %	61.58 %

Tabla 2.4: Comparación de resultados de YOLO entre los dos datasets

Como resultado del análisis de la tabla 2.4 se concluye:

- En cuanto al resultado con la métrica IOU, el dataset de MRI Brain Segmentation[54] ha sido el que generó mejores resultados en testing. Cabe resaltar que las imágenes de este dataset [54] fueron pre-procesadas por humanos, por lo que la gran diferencia puede estar en dicho pre-procesamiento. Otro posible motivo es que las imágenes del dataset de MRI Brain Segmentation, son mas “simples o evidentes.”^{en} cuanto a si hay o no un tumor. Incluso los autores de este trabajo (sin formación en el área de la medicina) son capaces de distinguir la presencia de una posible anomalía en la mayoría de los casos
- Los resultados del DDSM [45] son equilibrados en cuanto a la clasificación por clase, aunque con baja precisión en identificar la ROI. Se supone que esto se debe a que el dataset cuenta con las etiquetas y mascararas con información para todos ejemplos. En cambio el dataset de los MRI Brain Segmentation[54] los benignos no tienen máscaras con información (son pixeles negros). Llevando a que la red pondere a este gran conjunto de imágenes y dar predicciones favorables. Dando una precisión de 97 %.

2.5 Técnica 2: U-Net

La selección de la segunda técnica implementada en el desarrollo de este trabajo con objetivo principal la identificación de regiones de interés en imágenes con formato DICOM, se centró en la búsqueda de trabajos anteriores que se hayan enfrentado a estos desafíos y a su vez, sus resultados permitieran ser comparados con los resultados obtenidos por la técnica 1 usando los mismos conjuntos de datos. Como ha sido descrito en el capítulo 2.4, la arquitectura de la técnica 1 está enfocada a la detección y clasificación de Regiones de interés en imágenes 2D. Para

esto se buscó un trabajo condicionado a recibir datos de entrada en imágenes de 2D y que a su vez utilizara una técnica diferente; conforme al resultado de la búsqueda se decidió implementar un algoritmo de segmentación semántica denominado U-Net.

En el presente trabajo se denomina Técnica 2 a la adaptación del modelo original de la arquitectura U-NET basado en el paper U-Net: Convolutional Networks for Biomedical Image Segmentation desarrollado en mayo de 2015 y ganador del desafío ISBI 2015 - Segmentation of neuronal structures in EM stacks[22, 58].

El trabajo original realizado por los autores intenta resolver problemas en tareas de localización de regiones de interés de forma precisa y limitada por el bajo volumen de datos de muestras para entrenamiento. Si bien como indican los autores, hasta el momento de su publicación, en trabajos anteriores el uso típico de redes convolucionales estaba centrado en tareas de clasificación, donde el resultado que se obtenía eran etiquetas pertenecientes a una clase; aquí se deseaba obtener una localización, es decir generar etiquetas para cada píxel[22].

El resultado, según comentan los autores, es aplicable a distintos problemas de segmentación en imágenes biomédicas, pero en el paper se muestran resultados sobre la segmentación de imágenes de estructuras neuronales obtenidas de un microscopio electrónico (Segmentation of neuronal structures in EM stacks). Con esta arquitectura participaron del desafío ISBI por sus siglas en inglés International Symposium on Biomedical Imaging 2015[58] de la (IEEE - Institute of Electrical and Electronics Engineers)[59], en donde resultaron ganadores[22].

2.5.1 Arquitectura

La arquitectura planteada en el paper corresponde a una arquitectura en forma de U dividida en dos partes, la primera denominada la ruta de contracción, corresponde a una arquitectura típica de red convolucional (lado izquierdo de la red) y una segunda etapa denominada ruta de expansión (lado derecho de la red)[22]; a nivel visual genera una vista simétrica entre las dos.

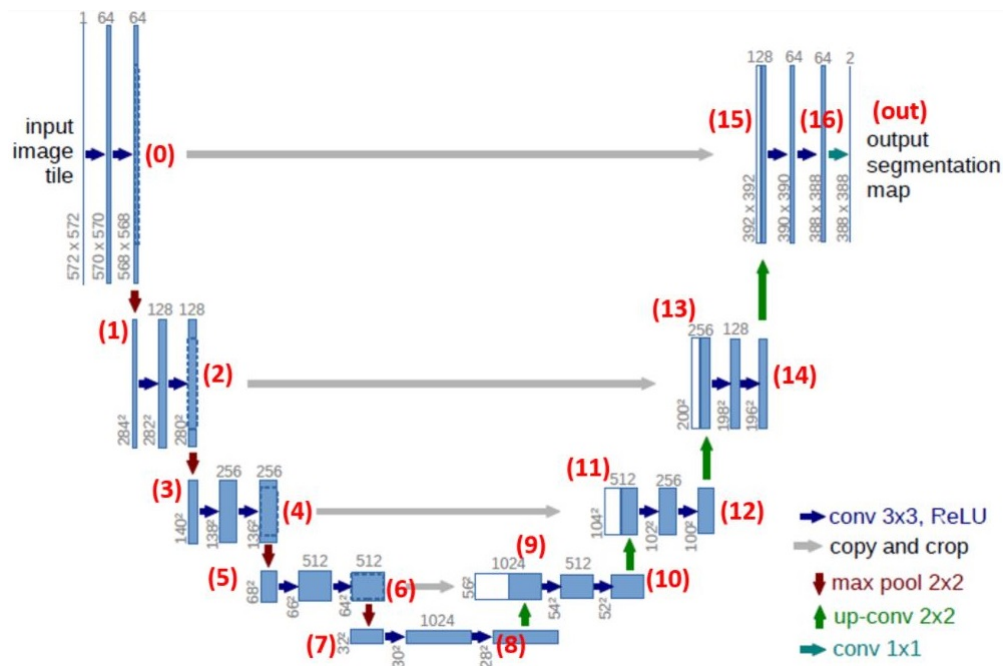


Figura 2.30: Arquitectura U-Net[22]

La red está compuesta por 18 capas convolucionales activadas por ReLU y kernel de tamaño 3x3, 8 para la etapa de contracción y 10 para la etapa de expansión. Las capas de la etapa de convolución están unidas en bloques de 2, representadas en la Figura 2.30 por las salidas (0), (2), (4) y (6) y seguido cada bloque por Max Polling con tamaño de kernel 2x2 y pasos de 2, representadas en la Figura 2.30 por las salidas (1), (3), (5) y (7).

En la etapa de expansión, al igual que en la etapa de contracción se generan bloques de dos capas convolucionales con tamaño de kernel de 3x3, representadas en la Figura 2.30 (8), (10), (12), (14) y (16), seguidas de una convolución transpuesta con tamaño de kernel de 2x2 y pasos de 2, representada en la Figura 2.30 (9), (11), (13) y (15), a excepción del ultimo bloque, representado en la Figura 2.30 (16) que está seguido de una capa convolucional con tamaño de kernel de 1x1, representado en la Figura 2.30 (Out). Finalmente después de cada salida de la etapa de contracción, antes del Max Pooling se crea una conexión directa con a la etapa de expansión, representados en la Figura 2.30, (0) - (15), (2) - (13), (4) - (11) y (6) - (9).

	Capa	in ch.	out ch.	Kernel	Stride
(0)	Conv2d, ReLU	1		(3x3)	-
	Conv2d, ReLU		64	(3x3)	-
(1)	MaxPool2d	-	-	(2x2)	1
(2)	Conv2d, ReLU	64		(3x3)	-
	Conv2d, ReLU		128	(3x3)	-
(3)	MaxPool2d	-	-	(2x2)	1
(4)	Conv2d, ReLU	128		(3x3)	-
	Conv2d, ReLU		256	(3x3)	-
(5)	MaxPool2d	-	-	(2x2)	1
(6)	Conv2d, ReLU	256		(3x3)	-
	Conv2d, ReLU		512	(3x3)	-
(7)	MaxPool2d	-	-	(2x2)	1
(8)	Conv2d, ReLU	512		(3x3)	-
	Conv2d, ReLU		1024	(3x3)	-
(9)	ConvTranspose2d	1024	512	(2x2)	2
(10)	Conv2d, ReLU	1024		(3x3)	-
	Conv2d, ReLU		512	(3x3)	-
(11)	ConvTranspose2d	512	256	(2x2)	2
(12)	Conv2d, ReLU	512		(3x3)	-
	Conv2d, ReLU		256	(3x3)	-
(13)	ConvTranspose2d	256	128	(2x2)	2
(14)	Conv2d, ReLU	256		(3x3)	-
	Conv2d, ReLU		128	(3x3)	-
(15)	ConvTranspose2d	128	64	(2x2)	2
(16)	Conv2d, ReLU	128		(3x3)	-
	Conv2d, ReLU		64	(3x3)	-
(Out)	Conv2d	64	2	(1x1)	0

Tabla 2.5: Arquitectura original paper U-Net

2.5.2 Implementación 1 MRI

La primera implementación de esta técnica ha sido realizada con el conjunto de datos de Brain MRI segmentation[54] descrito en el capítulo 2.3.2, el cual contiene imágenes de Resonancias magnéticas del cerebro. Dado que esta técnica ha sido ampliamente utilizado en la comunidad de imágenes médicas[33] y ha sido referenciado en más de 30.000 publicaciones técnicas[60], existe variedad de implementaciones;aprovechando lo aprendido en la asignatura Deep Learning del Master

en relación al framework Pytorch[61, 62] y con ayuda de tutoriales de implementaciones enfocadas a otros problemas específicos[63, 64], se ha implementado toda la red por los autores de esta tesis, acorde a lo descrito en el paper original en base a este framework , para posteriormente realizar los ajustes para este trabajo específico, los cuales serán detallados a continuación:

Dadas las limitaciones de tiempo y recursos computacionales con que se contaba a la hora de la implementación y con el objetivo de comprobar su desempeño, se inició con una red U-Net reducida de 15 convoluciones similar a la del paper “Convolutional Neural Networks enable efficient, accurate and finegrained segmentation of plant species and communities from high-resolution UAV imagery” [65]. Una vez implementada la primera arquitectura y comprobados sus resultados con los datos, se procede a realizar la adecuación de la red a la arquitectura original con sus respectivas modificaciones agregando las operaciones de convolución restantes, para el enfoque del problema específico. Los primeros cambios en la arquitectura corresponde al número de canales de entrada y de salida de la red. La arquitectura original recibe datos de entrada de 1 canal y genera elementos de salida de dos canales y las imágenes del conjunto de datos cuenta con entrada RGB, por consiguiente 3 canales con valores de píxel en cada canal con valores de entre 0 y 256 y salidas de 1 canal con valores por píxel entre 0 y 256, ver figura 2.31, se procede a cambiar en la arquitectura de la implementación para que reciba 3 canales de entrada y genere uno de salida respectivamente. Además se aplica normalización para que los valores de píxel estén entre 0 y 1, escala de grises.

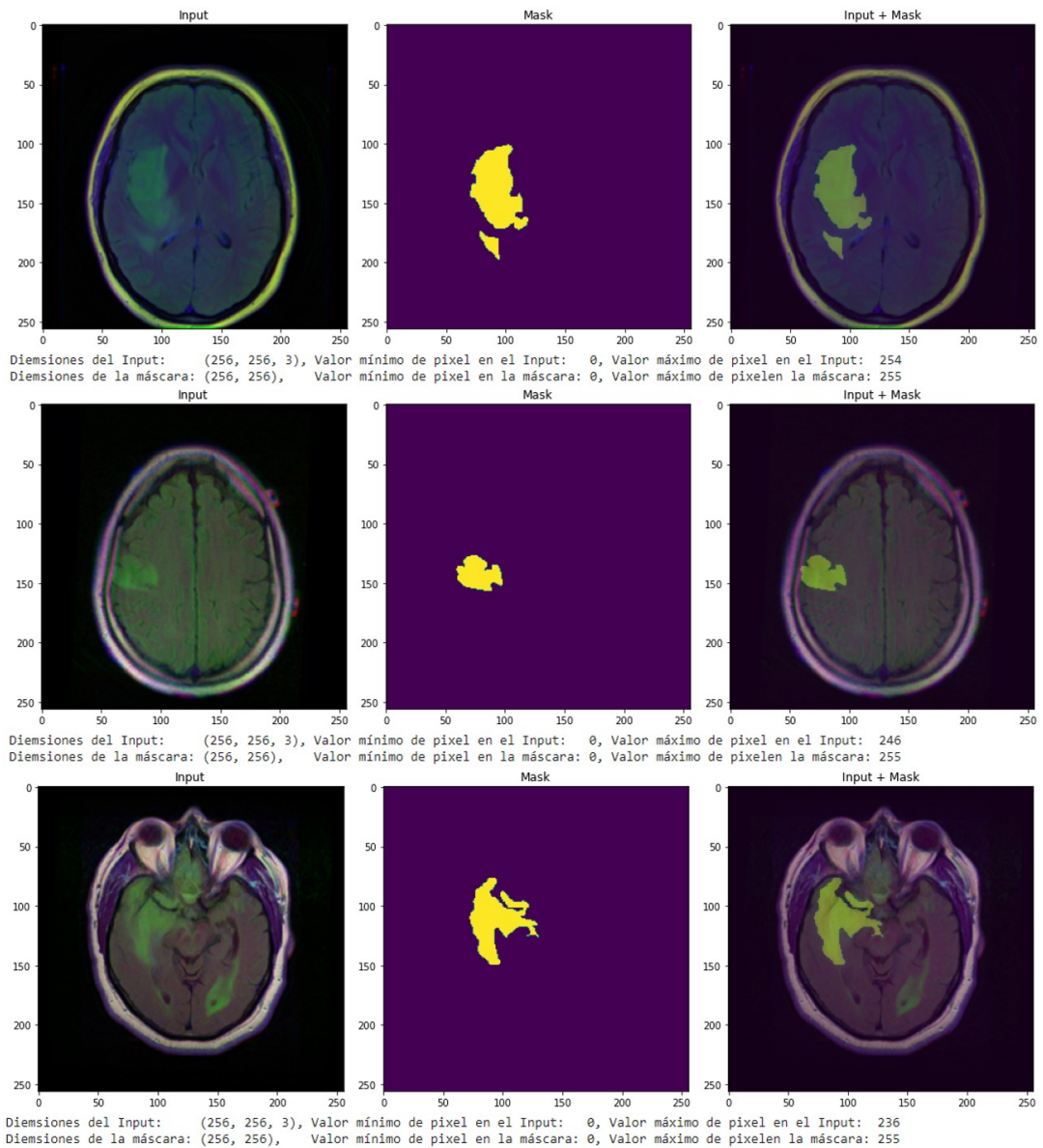


Figura 2.31: Imágenes originales del conjunto de datos MRI Brain Segmentation[54]

Una vez realizada la primera modificación se observa que el tamaño de las imágenes del conjunto de datos utilizado tanto de entrada y salida son de 256 de ancho por 256 de alto, se procede a hacer Resize para que se amolden al tamaño de la entrada de la arquitectura original 572 de ancho por 572 de alto, ver figura 2.32.

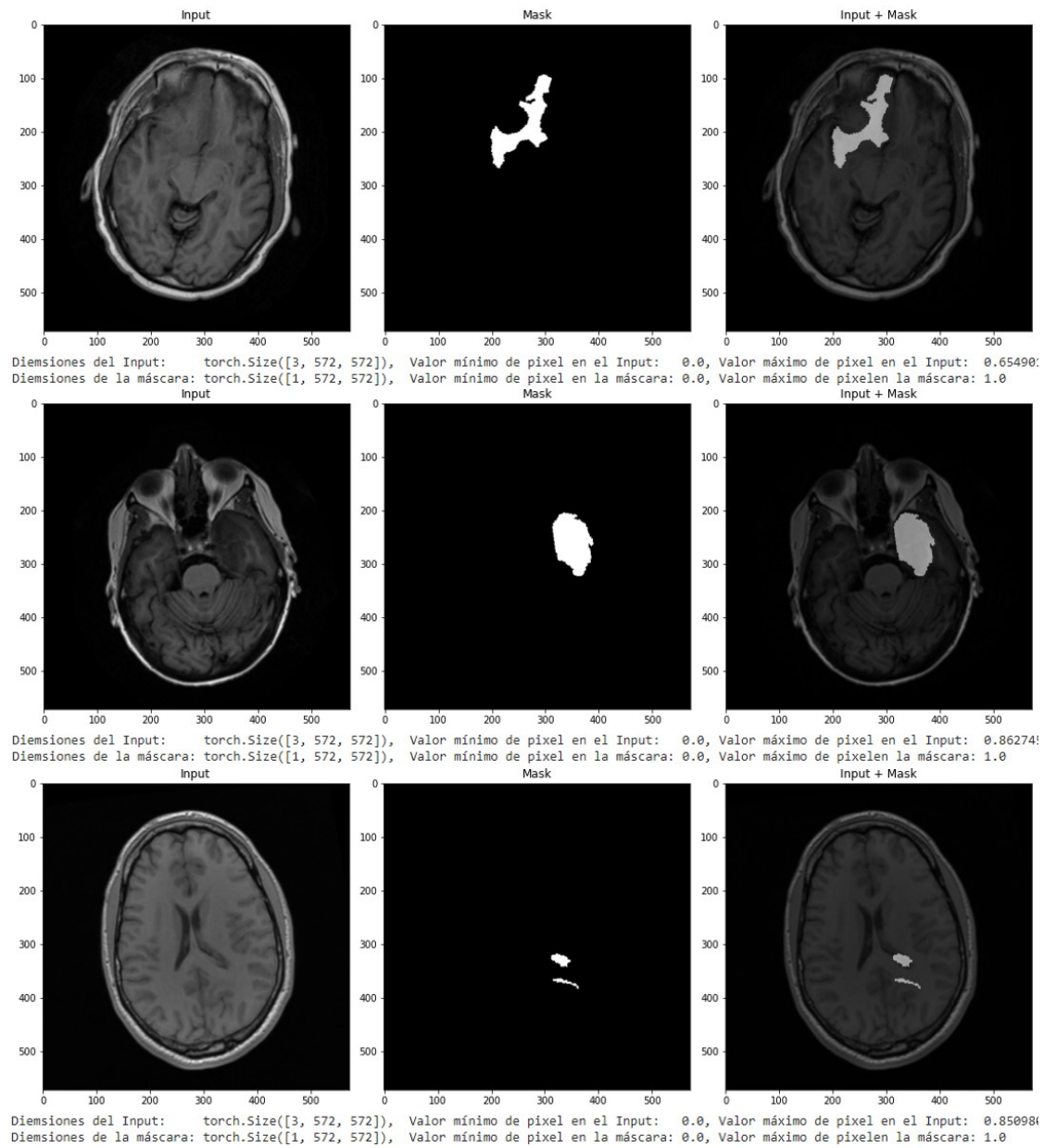


Figura 2.32: Imágenes transformadas

Como se observa en la Figura 2.30, la salida de la arquitectura original genera salidas de 338 pixeles de ancho por 388 pixeles de alto, se agrega Padding de 1 a las salidas de las capas convolucionales (0), (2), (4), (6), (8), (10), (12), (14) y

(16) ver Figura 2.30 para conseguir una salida de la red de tamaño 1 x 572 x 572. Además se agregó Batch Normalization a continuación de cada doble convolución. La arquitectura resultante se muestra en la tabla 2.7.

	Capa	in ch.	out ch.	Kernel	Stride	Padding
(0)	Conv2d, ReLU	3		(3x3)	-	1
	Conv2d, ReLU		64	(3x3)	-	-
(1)	MaxPool2d	-	-	(2x2)	1	-
(2)	Conv2d, ReLU	64		(3x3)	-	1
	Conv2d, ReLU		128	(3x3)	-	-
(3)	MaxPool2d	-	-	(2x2)	1	-
(4)	Conv2d, ReLU	128		(3x3)	-	1
	Conv2d, ReLU		256	(3x3)	-	-
(5)	MaxPool2d	-	-	(2x2)	1	-
(6)	Conv2d, ReLU	256		(3x3)	-	1
	Conv2d, ReLU		512	(3x3)	-	-
(7)	MaxPool2d	-	-	(2x2)	1	-
(8)	Conv2d, ReLU	512		(3x3)	-	1
	Conv2d, ReLU		1024	(3x3)	-	-
(9)	ConvTranspose2d	1024	512	(2x2)	2	-
(10)	Conv2d, ReLU	1024		(3x3)	-	1
	Conv2d, ReLU		512	(3x3)	-	-
(11)	ConvTranspose2d	512	256	(2x2)	2	-
(12)	Conv2d, ReLU	512		(3x3)	-	1
	Conv2d, ReLU		256	(3x3)	-	-
(13)	ConvTranspose2d	256	128	(2x2)	2	-
(14)	Conv2d, ReLU	256		(3x3)	-	1
	Conv2d, ReLU		128	(3x3)	-	-
(15)	ConvTranspose2d	128	64	(2x2)	2	-
(16)	Conv2d, ReLU	128		(3x3)	-	1
	Conv2d, ReLU		64	(3x3)	-	-
(Out)	Conv2d	64	1	(1x1)	0	-

Tabla 2.6: Arquitectura modificada para imágenes de MRI Brain Segmentation [54]

Para el entrenamiento, validación y test, se dividió el dataset en lo máximo permitido por los recursos computacionales, correspondiente a batches de 6 imágenes, para un total de 262 batches para Train con 1572 imágenes, 70 para validación

para 420 imágenes y 11 para test para un total de 108 imágenes. Para evitar el sobre ajuste en el entrenamiento, se agregó Data Augmentation por medio de la herramienta Albumentation[66], la cual permite aplicar a la máscara las mismas transformaciones que se aplicaron a la imagen de entrada, Dropout al final de la red, con probabilidad 0.5 y Weight Decay con valor 0.0001. las transformaciones realizadas consistieron en Rotaciones de 15 grados con probabilidad 1 y Flip horizontal con probabilidad de 0.5. La tasa de aprendizaje utilizada corresponde al valor 0.0001. Adicionalmente se modificaron la función de perdida por Binary Cross Entropy with Logits y el optimizador por Adam.

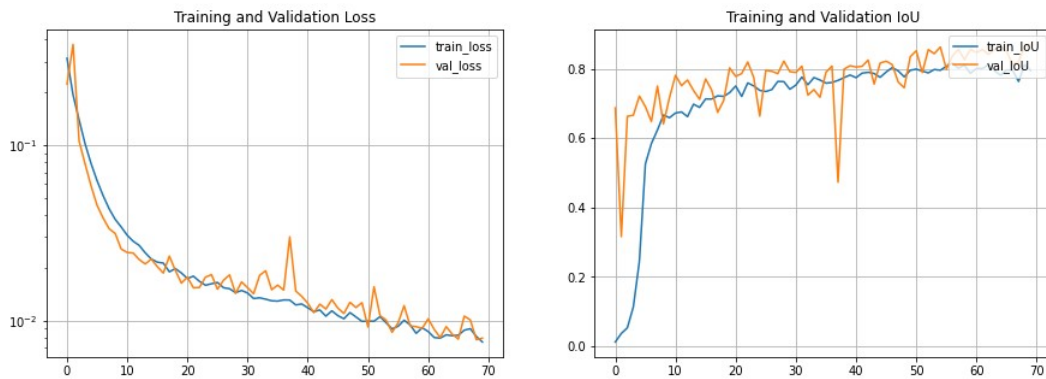


Figura 2.33: Resultados de Loss e IoU en datos de Train y Validation

Como se puede observar en la Figura 2.33, el entrenamiento y validación para esta red fue realizado con 70 Epochs durante 10 horas sobre la plataforma Google Colab[46], obteniendo 0.8 de IoU y Loss de 0.007 promedio por batch, para train y para validation. Además se observa que el IoU en el primer Epoch inicia con valores superiores a los 0.6, esto como resultado de la cantidad de imágenes que no contienen información de ROI, en los que a efectos del cálculo de la métrica les da el valor unitario.

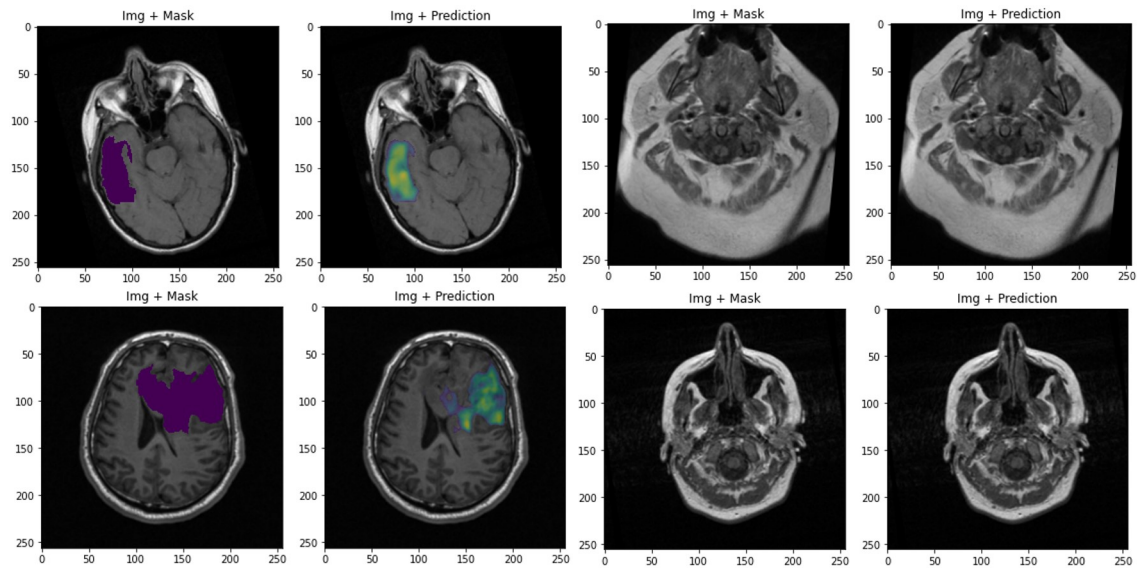


Figura 2.34: Resultados de test de detección de ROI

La Figura 2.34 muestra ejemplos de resultados de detección de la ROI en imágenes de test, se puede observar en los ejemplos que a las imágenes que no contienen información de máscara no les genera datos de ROI en las predicciones y a las imágenes que contienen información de ROI, las logra identificar.

Para la prueba de las detecciones del IoU de las imágenes del dataset de test, se evaluó el modelo con todo los datos, obteniendo resultados superiores a 0.7 y la mayoría alcanzó el valor de 0.85.

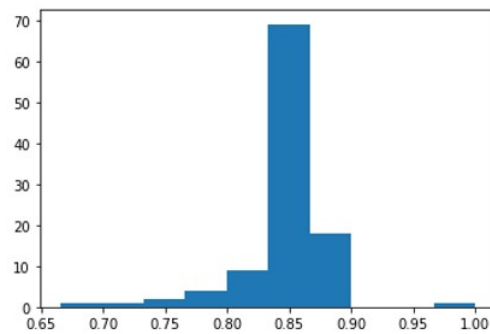


Figura 2.35: Resultados IoU dataset de test

La figura 2.35 muestra los resultados del IoU calculado en 150 imágenes contenidas en el dataset.

2.5.3 Implementación 2 Mamografías

Una vez finalizadas las pruebas con el dataset de MRI Brain Segmentation descrito en el Capítulo 2.3.2, y acorde lo planteado en los objetivos del trabajo, se procede a comprobar la técnica con el segundo conjunto de datos descrito en el Capítulo 2.3.1, el cual ha sido utilizado como primer dataset en la implementación de la técnica 1, con la diferencia que para esta prueba únicamente se tomaron datos de masas etiquetadas como Malignas.

Cómo fue mencionado en capítulos anteriores, las imágenes de mamografías fueron transformadas para ser convertidas a tensores y adaptadas a cada técnica, para este caso al igual que en la implementación 1 en donde se trabajó con el dataset de MRI Brain Segmentation, se buscó aplicar las mismas transformaciones, para poder utilizar la red ya implementada.

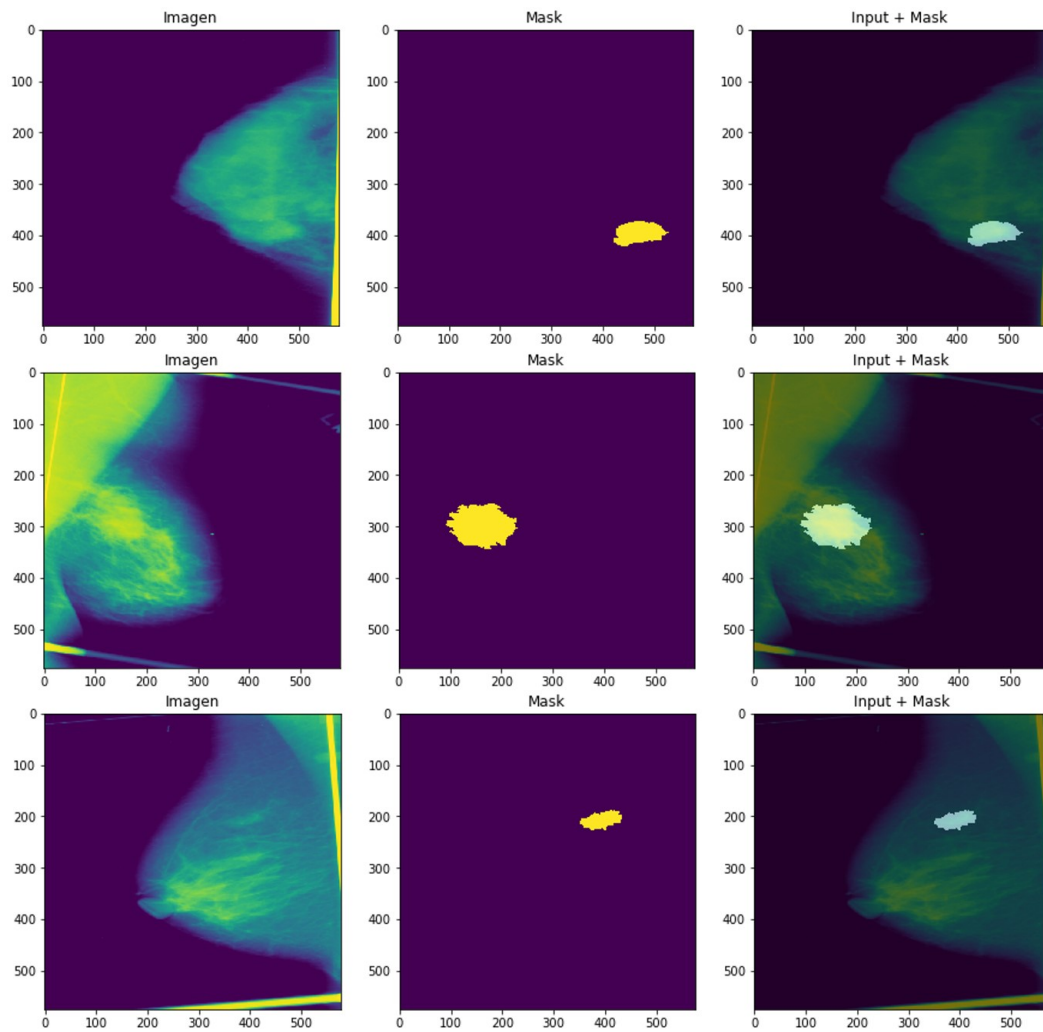


Figura 2.36: Imágenes transformadas del conjunto de datos Mamografías[45]

En cuanto al tamaño y forma de las imágenes, se aplicó normalización a la imagen y a su respectiva máscara, y resize acorde a las dimensiones de las capas de la arquitectura de la red. Finalmente tanto la imagen como la máscara quedaron con valores de $[1 \times 576 \times 576]$.

Una de las diferencias importantes a la hora de realizar las pruebas con esta técnica, corresponde al número de imágenes que contienen información de la ROI, dado que todas las imágenes contienen información con valor unitario en su respectiva máscara.

Para el entrenamiento y validación del modelo se tomaron 2.500 imágenes divididas en batches de 6 imágenes cada uno, 2.000 imágenes para train y 500 para validation.

	Capa	in ch.	out ch.	Kernel	Stride	Padding
(0)	Conv2d, ReLU	1		(3x3)	-	1
	Conv2d, ReLU		64	(3x3)	-	-
(1)	MaxPool2d	-	-	(2x2)	1	-
(2)	Conv2d, ReLU	64		(3x3)	-	1
	Conv2d, ReLU		128	(3x3)	-	-
(3)	MaxPool2d	-	-	(2x2)	1	-
(4)	Conv2d, ReLU	128		(3x3)	-	1
	Conv2d, ReLU		256	(3x3)	-	-
(5)	MaxPool2d	-	-	(2x2)	1	-
(6)	Conv2d, ReLU	256		(3x3)	-	1
	Conv2d, ReLU		512	(3x3)	-	-
(7)	MaxPool2d	-	-	(2x2)	1	-
(8)	Conv2d, ReLU	512		(3x3)	-	1
	Conv2d, ReLU		1024	(3x3)	-	-
(9)	ConvTranspose2d	1024	512	(2x2)	2	-
(10)	Conv2d, ReLU	1024		(3x3)	-	1
	Conv2d, ReLU		512	(3x3)	-	-
(11)	ConvTranspose2d	512	256	(2x2)	2	-
(12)	Conv2d, ReLU	512		(3x3)	-	1
	Conv2d, ReLU		256	(3x3)	-	-
(13)	ConvTranspose2d	256	128	(2x2)	2	-
(14)	Conv2d, ReLU	256		(3x3)	-	1
	Conv2d, ReLU		128	(3x3)	-	-
(15)	ConvTranspose2d	128	64	(2x2)	2	-
(16)	Conv2d, ReLU	128		(3x3)	-	1
	Conv2d, ReLU		64	(3x3)	-	-
(Out)	Conv2d	64	1	(1x1)	0	-

Tabla 2.7: Arquitectura modificada para imágenes de Mamografías [45]

Como se observa en la Tabla 2.7, las modificaciones que se realizaron a la arquitectura utilizada para las pruebas con el dataset de MRI Brain Segmentation, únicamente corresponden al número de canales de la imagen de entrada. Se mantiene el padding y el Batch Normalization agregado a la arquitectura original.

Con el objetivo de comparar los resultados de la arquitectura con los dos datasets, para el entrenamiento con este dataset se mantuvieron los mismos parámetros como $LR = 0.0001$, Adam como optimizador y Binary Cross Entroy with Logits como función de costo. Al igual que se aplicaron a las imágenes las mismas transformaciones de Data Augmentation para evitar el overfitting.

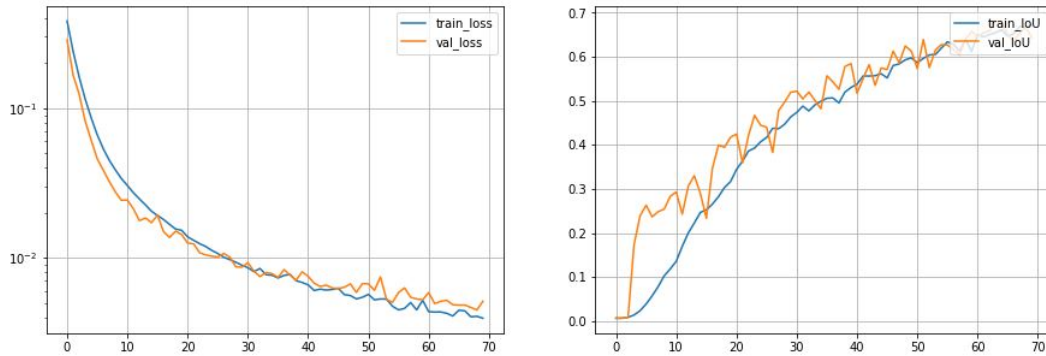


Figura 2.37: Resultados de Loss e IoU en datos de Train y Validation

Como se observa en la Figura 2.37, a igual que con el primer dataset, se entrenó en 70 Epochs durante 12 horas sobre la plataforma Google Colab[46], y se obtuvo como resultado máximo de IoU el valor correspondiente a 0.67, con un valor mínimo de Loss de 0.004, en train y validation.

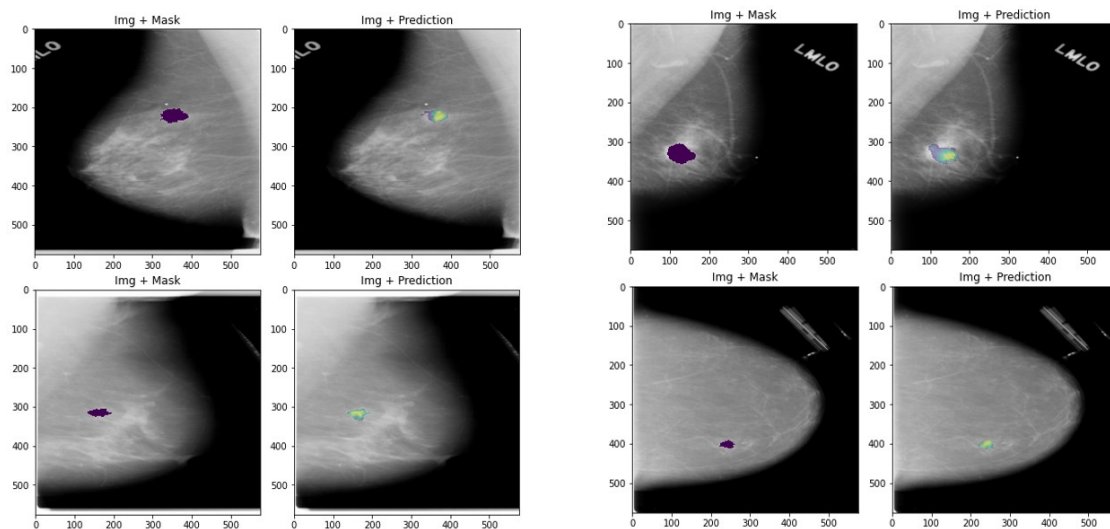


Figura 2.38: Ejemplo de resultados de validación de detección de ROI Acertados

En la Figura 2.38, se observan ejemplos de detecciones correctas obtenidas de los datos de validación comparadas con los datos de la máscara de cada imagen.

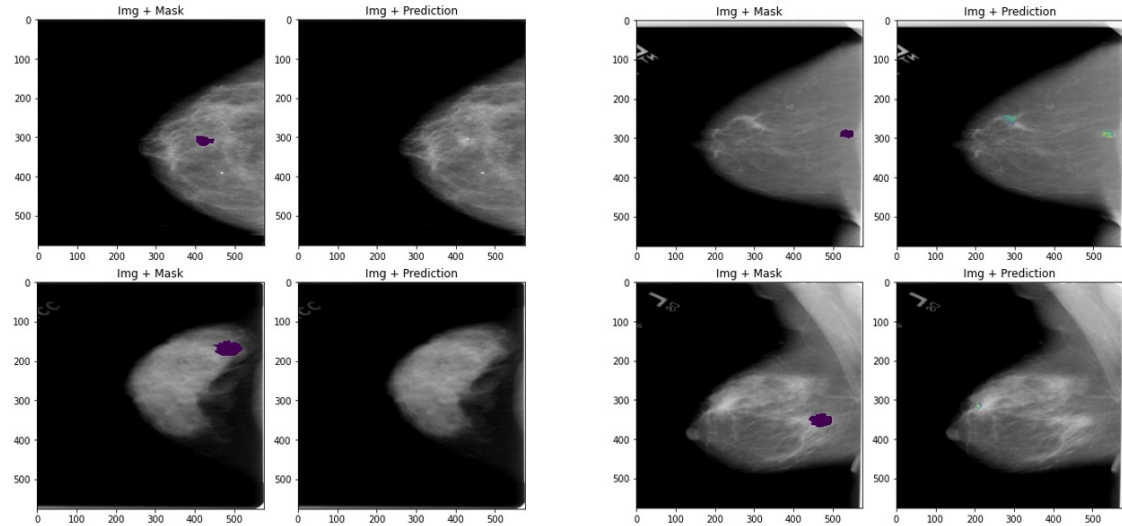


Figura 2.39: Ejemplo de resultados de validación de detección de ROI NO acertados

En la Figura 2.39, se observan ejemplos de identificación correcta de la ROI, obtenidas a partir de los datos de validación comparadas con los datos de la máscara de cada imagen, pero que a su vez incluyen valores de Falsos positivos.

2.5.4 Comparación de resultados

Comparativa	Mamografías	MRI
$\text{IoU} \geq 0.5$	0.67	0.85

Tabla 2.8: Comparación de resultados de U-Net entre los dos datasets

Una vez finalizados los entrenamientos y pruebas con los dos modelos y acorde al análisis de la tabla 2.8 se concluye:

- El modelo U-Net en general logra obtener buenos resultados a la hora de identificar la región de interés en los dos problemas específicos, pero a cuentas

de largas horas de entrenamiento, excediendo las 10 horas para 70 epochs en cada uno usando recursos computacionales disponibles. Por consiguiente, es una red que se puede amoldar a estos tipos de problema pero requiere largas horas para entrenarse en una tarea específica.

- Dado que para el entrenamiento con los dos datasets, se usaron más de 2.000 imágenes y a las mismas se les aplicó Data Augmentation, también es necesario mencionar que se requiere un gran volumen de muestras de entrenamiento para un solo problema y consecuente recursos computacionales.
- Los resultados obtenidos para el dataset de MRI Brain segmentation superaron ampliamente a los de mamografías. Si bien ante el ojo inexperto, en el dataset de MRI, se puede identificar a simple vista las regiones con tumor, en el de mamografías dado el balance de blancos y negros, se hace casi imposible identificarlo. Se presume que esto puede ser un beneficio para la red en su entrenamiento a la hora de observar los tumores.
- Observando los resultados entre las predicciones para los dos datasets y comparándolos con la máscara, se observa que no necesariamente el ROI identificado abarca la zona seleccionada por el experto en el etiquetado de la máscara. Esto puede corresponder a que las etiquetas han sido generadas de forma manual. Por lo que se concluye que sigue siendo necesaria la validación del ojo experto a la hora de generar algún pronóstico.
- Si bien el modelo de U-Net que se usó con los dos dataset se utilizó únicamente para identificar regiones de interés, es posible que, ingresando al modelo las imágenes de las máscaras etiquetadas, sea posible generar clasificaciones de las mismas. Por ejemplo, ingresando el total de datos de mamografías con patologías benignas y malignas, se pueda llegar a clasificarlas. Esto se propone como ampliación a este trabajo.

3 Conclusiones

Dado que la región de interés en los dataset que contienen imágenes en el formato DICOM son generadas de forma manual por personal especializado, no todos los datasets disponibles contienen esta información por lo que se complica a la hora de intentar implementar y comparar con modelos de Inteligencia Artificial.

El entrenamiento de modelos que permiten el análisis de imágenes médicas implica el consumo de grandes recursos computacionales y bastante tiempo de ejecución.

No todas las imágenes DICOM contienen la ROI. Comparación entre modelos o arquitecturas, Aplicabilidad a distintos datasets. Performance – Costo de ejecución (tiempo – recursos).

Luego de las distintas ejecuciones de experimentos y comparaciones de resultados, se concluye:

- Ambas arquitecturas arrojan los mejores resultados con las imágenes del dataset de MRI BRAIN Segmentation. Se entiende que esto se debe a que incluso para el ojo humano las imágenes son más “simples o evidentes“ en cuanto a si hay o no un tumor o al pre procesamiento por humanos realizado.
- Dentro de las arquitecturas, el modelo UNet 2.5 es la que logró mejor IoU en los dos dataset. Sin embargo los tiempos de entrenamiento de UNet son considerablemente mayores y requiere más recursos en comparación a una configuración similar con BC-DROID 2.4.
- Debido a la duración del trabajo y en búsqueda de cumplir el objetivo de comparar dos técnicas y dos datasets, fue necesario dividir los recursos para cada una de las técnicas, lo cual dió como resultado que la cantidad de experimentos realizados y el tiempo de entrenamiento para cada uno fuese acotado.

4 Trabajo a futuro

Se cree que el presente trabajo abre varias posibilidades para profundizar en la identificación de la Región de Interés (ROI) en imágenes médicas contenidas bajo el estándar DICOM. Dado que en el trabajo se abordan las arquitecturas YOLO y U-Net y dos datasets, el lector tendrá acceso a una variedad de problemas y alternativas para solucionarlos.

Dado el tiempo y los recursos dispuestos para la elaboración de este trabajo y con el fin de hacer más extenso el uso de modelos de Inteligencia Artificial en el campo de la medicina y el análisis de imágenes médicas, se plantean algunas posibilidades de mejoras y algunas posibles extensiones a este proyecto:

Posibles extensiones de esta tesis:

- Implementación con otros dataset de archivos DICOM. Observar distintas casos de estudios.
- Implementación de otras arquitecturas para comparar con las dos presentadas aquí.
- Uso de modelos pre-entrenados y comparación con los datos resultantes de este trabajo y los trabajos citados en el mismo.
- Implementación de arquitecturas para identificación de región de interés que no han sido probadas con imágenes médicas.
- Implementación y comparación arquitecturas que permitan la identificación de región y volumen de interés en imágenes 3D.
- Implementación de U-Net con todos los datos de mamografías, incluyendo las etiquetas de patologías (Maligna, Benigna) para obtener clasificación.
- Obtención de otro tipo de métricas para la evaluación de imágenes.

- Investigación con incremento de recursos computacionales para mejorar los tiempos de entrenamiento y posibilitar el uso intensivo de Data Augmentation.
- Puesta en producción de modelos para su uso en casos reales.

5 Bibliografía

- [1] A. Rehman, D. M. A. Butt, and D. M. Zaman, “A survey of medical image analysis using deep learning approaches,” Tech. Rep., 2021.
- [2] D. Kumar, A. Wong, and D. A. Clausi, “Lung nodule classification using deep features in ct images,” Tech. Rep., 2015.
- [3] P. Jangbari and D. Patel, “Review on region of interest coding techniques for medical image compression,” Tech. Rep., 2016.
- [4] R. Eswaraiah and E. S. Reddy, “Robust medical image watermarking technique for accurate detection of tampers inside region of interest and recovering original region of interest,” Tech. Rep., 2014.
- [5] “Dicom standar.” [Online]. Available: <https://www.dicomstandard.org/>
- [6] “Nema.” [Online]. Available: <https://www.nema.org/>
- [7] “Dicom ps3.1 2021d.” [Online]. Available: <http://dicom.nema.org/medical/dicom/current/output/html/part01.html>
- [8] “Google scholar.” [Online]. Available: <https://scholar.google.com/>
- [9] “Foco.” [Online]. Available: <https://foco.timbo.org.uy/home>
- [10] A. Helwan and D. U. Ozsahin, “Sliding window based machine learning system for the left ventricle localization in mr cardiac images,” Tech. Rep., 2017.
- [11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified real-time object detection,” Tech. Rep., 2016.
- [12] C. M. Yeum, J. Choi, and S. J. Dyke, “Automated region-of-interest localization and classification for vision-based visual assessment of civil infrastructure,” Tech. Rep., 2019.

- [13] M. Brown and D. Lowe, “Invariant features from interest point groups,” Tech. Rep., 2002.
- [14] J. Blasco, N. Aleixos, and E. Molto, “Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm,” Tech. Rep., 2006.
- [15] W. Chu, Y. Liu, C. Shen, D. Cai, and X.-S. Hua, “Multi-task vehicle detection with region-of-interest voting,” Tech. Rep., 2018.
- [16] A. C. Evans, S. Marrett, J. Torrescorzo, S. Ku, and L. Collins, “Mri-pet correlation in three dimensions using a volume-of-interest (vol) atlas,” Tech. Rep., 1991.
- [17] B. He and E. C. Frey, “The impact of 3d volume of interest definition on accuracy and precision of activity estimation in quantitative spect and planar processing methods,” Tech. Rep., 2010.
- [18] H. R. Roth, L. Lu, J. Liu, J. Yao, A. Seff, K. Cherry, L. Kim, and R. M. Summers, “Improving computer-aided detection using convolutional neural networks and random view aggregation,” Tech. Rep., 2016.
- [19] K. Sakurai, E. Imabayashi, A. M. Tokumaru, K. Ito, K. Shimoji, M. Nakagawa, Y. Ozawa, M. Shimohira, M. Ogawa, S. Morimoto, I. Aiba, N. Matsukawa, and Y. Shibamoto, “Volume of interest analysis of spatially normalized presto imaging to differentiate between parkinson disease and atypical parkinsonian syndrome,” Tech. Rep., 2016.
- [20] A. I. Penn, M. Medved, V. Dialani, E. D. Pisano, E. B. Cole, D. Brousseau, G. S. Karczmar, G. Gao, B. D. Reich, and H. Abe, “Discrimination of benign from malignant breast lesions in dense breasts with modelbased analysis of regions-of-interest using directional diffusion-weighted images,” Tech. Rep., 2020.
- [21] W. Rahane, H. Dalvi, Y. Magar, and A. Kalane, “Lung cancer detection using image processing and machine learning healthcare,” Tech. Rep., 2018.
- [22] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” Tech. Rep., 2015.
- [23] O. J. del Toro, “Quantitative analysis of medical images: finding relevant regions-of-interest for medical decision support (thesis),” Tech. Rep., 2017.

- [24] A. M. L. Santacruz, “Diagnosis and prognosis of cardiovascular diseases by means of texture analysis in magnetic resonance imaging (thesis),” Tech. Rep., 2017.
- [25] R. S. Matilla, “Detección jerárquica de grupos de personas,” Tech. Rep., 2014.
- [26] F. Yasuno, A. H. Hasnine, T. Suhara, T. Ichimiya, Y. Sudo, M. Inoue, A. Takano, T. Ou, T. Ando, and H. Toyama, “Template-based method for multiple volumes of interest of human brain pet images,” Tech. Rep., 2001.
- [27] T. Snoeks, E. Kaijzel, I. Que, I. Mol, C. Löwik, and J. Dijkstra, “Normalized volume of interest selection and measurement of bone volume in microct scans,” Tech. Rep., 2011.
- [28] J. Lee, J. Bang, and S.-I. Yang, “Object detection with sliding window in images including multiple similar objects,” Tech. Rep., 2017.
- [29] X. Jiang, Y. Pang, J. Pan, and X. Li, “Flexible sliding windows with adaptive pixel strides,” Tech. Rep., 2014.
- [30] I. U. Haq, K. Muhammad, T. Hussain, S. Kwon, M. Sodanil, S. W. Baik, and M. Y. Lee, “Movie scene segmentation using object detection and set theory,” Tech. Rep., 2020.
- [31] T. Wu, J. Hu, L. Ye, and K. Ding, “A pedestrian detection algorithm based on score fusion for multi-lidar systems,” Tech. Rep., 2021.
- [32] H. Cevikalp and B. Triggs, “Visual object detection using cascades of binary and one-class classifiers,” Tech. Rep., 2017.
- [33] R. Alkadi, F. Taher, A. El-baz, and N. Werghi, “A deep learning-based approach for the detection and localization of prostate cancer in t2 magnetic resonance images,” Tech. Rep., 2018.
- [34] R. Platania, S. Shams, S. Yang, J. Zhang, K. Lee, and S.-J. Park, “Automated breast cancer diagnosis using deep learning and region of interest detection (bc-droid),” Tech. Rep., 2016.
- [35] A. Kumar and Pratibha, “A method of segmentation in 3d medical image for selection of region of interest (roi),” Tech. Rep., 2018.
- [36] L. G. Astrakas and M. I. Argyropoulou, “Shifting from region of interest (roi) to voxel-based analysis in human brain mapping,” Tech. Rep., 2009.

- [37] J. M. Zain and M. Clarke, “Reversible region of non-interest (roni) watermarking for authentication of dicom images,” Tech. Rep., 2007.
- [38] L. Y. S. M.V.Latte, “Lossless and lossy compression of dicom images with scalable roi,” Tech. Rep., 2010.
- [39] M. Firoozbakht, J. Dehmeshki, M. Martini, Y. Ebrahimdoost, H. Amin, M. Dehkordi, A. Youannic4, and S. Qanadli, “Compression of digital medical images based on multiple regions of interest,” Tech. Rep., 2010.
- [40] D. Yee, S. Soltaninejad, D. Hazarika, G. Mbuyi, R. Barnwal, and A. Basu, “Medical image compression based on region of interest using better portable graphics (bpg),” Tech. Rep., 2017.
- [41] “mage classification vs. object detection vs. image segmentation.” [Online]. Available: <https://medium.com/analytics-vidhya/image-classification-vs-object-detection-vs-image-segmentation-f36db85fe81>
- [42] “Deepai.” [Online]. Available: <https://deepai.org/machine-learning-glossary-and-terms/accuracy-error-rate>
- [43] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, “Generalized intersection over union: A metric and a loss for bounding box regression,” Tech. Rep., 2020.
- [44] “Dicom, nema.” [Online]. Available: <https://www.nema.org/standards/view/Digital-Imaging-and-Communications-in-Medicine>
- [45] “Cbis-ddsm.” [Online]. Available: <https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM>
- [46] “Google colab.” [Online]. Available: <https://colab.research.google.com/>
- [47] “Universidad ort.” [Online]. Available: <https://www.ort.edu.uy/>
- [48] “Pydicom,” [Mayo 02, 2021]. [Online]. Available: <https://pydicom.github.io/>
- [49] “Imagej.” [Online]. Available: <https://imagej.nih.gov/ij/download.html>
- [50] “Dicom metadata.” [Online]. Available: http://dicom.nema.org/medical/Dicom/2016e/output/chtml/part03/sect_C.7.6.3.html#table_C.7-11c
- [51] “Roi dicom estándar.” [Online]. Available: http://dicom.nema.org/medical/dicom/current/output/chtml/part03/sect_C.8.8.6.html

- [52] “Cancer imaging archive - tcia.” [Online]. Available: <https://www.cancerimagingarchive.net/>
- [53] “About - tcia.” [Online]. Available: <https://www.cancerimagingarchive.net/about-the-cancer-imaging-archive-tcia/>
- [54] “Mri brain segmentation.” [Online]. Available: <https://www.kaggle.com/mateuszbudalgg-mri-segmentation>
- [55] “Tcga-lyg.” [Online]. Available: <https://wiki.cancerimagingarchive.net/display/Public/TCGA-LGG>
- [56] M. Buda, A. Saha, and M. A. Mazurowski, “Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm,” Tech. Rep., 2019.
- [57] “Yolo v1 from scratch de aladdin persson.” [Online]. Available: https://www.youtube.com/watch?v=n9_XyCGr-MI
- [58] “Isbi-2015.” [Online]. Available: <https://biomedicalimaging.org/2015/program/isbi-challenges/>
- [59] “Ieee.” [Online]. Available: <https://www.ieee.org/>
- [60] “Citaciones de u-net.” [Online]. Available: https://scholar.google.com/uy/scholar?cites=10845403114495995712&as_sdt=2005&scioldt=0,5&hl=es
- [61] “Pytorch.” [Online]. Available: <https://pytorch.org/>
- [62] E. Stevens, L. Antiga, and T. Viehmann, “Deep learning with pytorch,” Tech. Rep., 2020.
- [63] “Competición myms-2.” [Online]. Available: https://www.youtube.com/watch?v=iCBU_1xgUg4
- [64] “Pytorch image segmentation tutorial with u-net.” [Online]. Available: <https://www.youtube.com/watch?v=IHq1t7NxS8k>
- [65] T. Kattenborn, J. Eichel, and F. E. Fassnacht, “Convolutional neural networks enable efficient, accurate and finegrained segmentation of plant species and communities from high-resolution uav imagery,” Tech. Rep., 2019.
- [66] “Albumentation.” [Online]. Available: https://albumentations.ai/docs/getting_started/mask_augmentation/

6 Anexos

6.1 Búsqueda de estado del arte

Ecuaciones para búsqueda de datasets:

- computer vision Region of interest volume of interest
- computer vision “Region of interest ROI” volume of interest
- Region of interest ROI volume of interest VOI
- Computer vision Region of interest ROI volume of interest VOI
- Selection of “Region of interest” ROI “volume of interest” VOI in DICOM
- Region of interest machine learning
- Sliding Window Computer vision Region of interest ROI Volume of interes VOI
- Detection techniques for region of interest ROI and volume of interest VOI
- Object detection with sliding window
- neural networks and deep learning healthcare + Region of interest ROI + Volume of interest VOI
- Method for Region of interest
- Methods for region of interest ROI and Volume of interest in DICOM

Los artículos fueron almacenados con la siguiente estructura de carpetas:

“Principal/Motor” – “Ecuación/Artículo”

6.2 Experimentos Técnica 1

6.2.1 Mamografías

Modelo 0

En las figuras 6.1,6.2 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

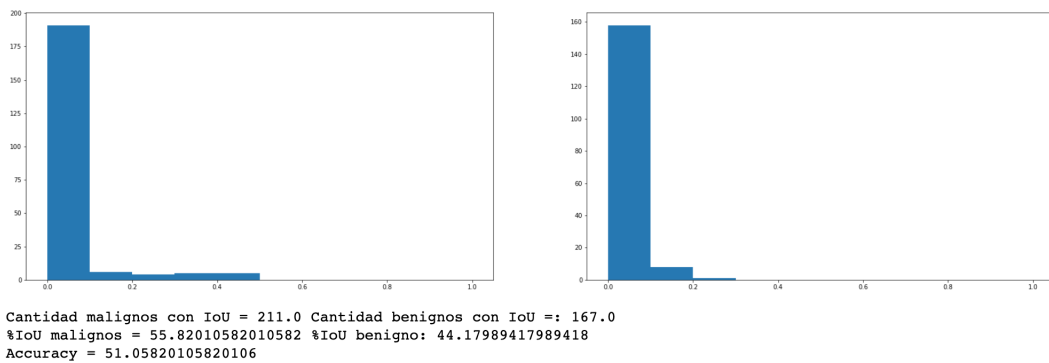


Figura 6.1: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

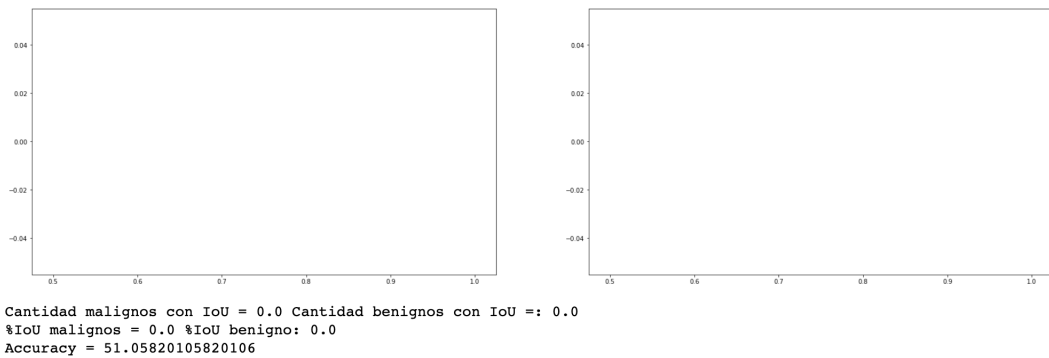


Figura 6.2: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 1

En las figuras 6.3,6.4 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

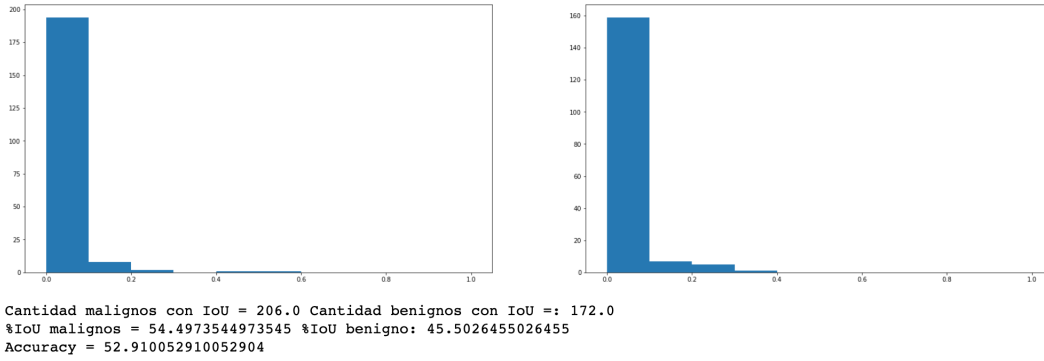


Figura 6.3: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

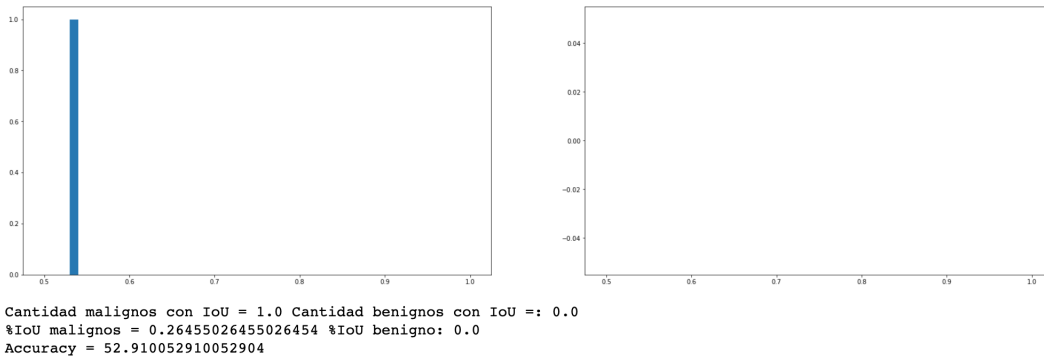


Figura 6.4: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 2

En las figuras 6.5,6.6 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

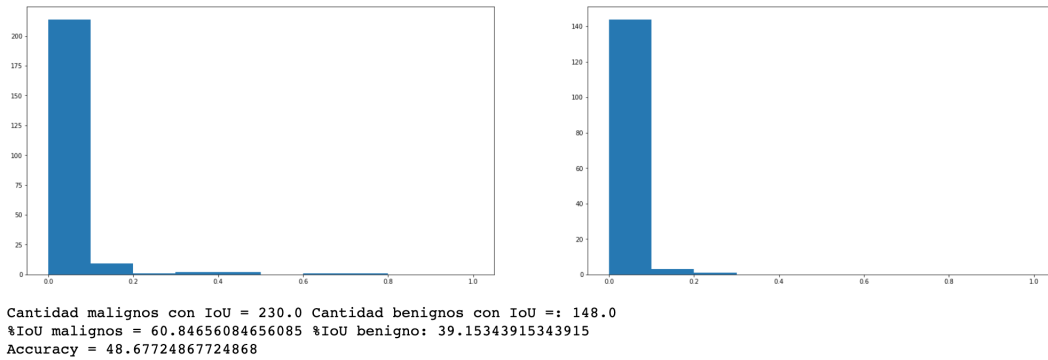


Figura 6.5: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

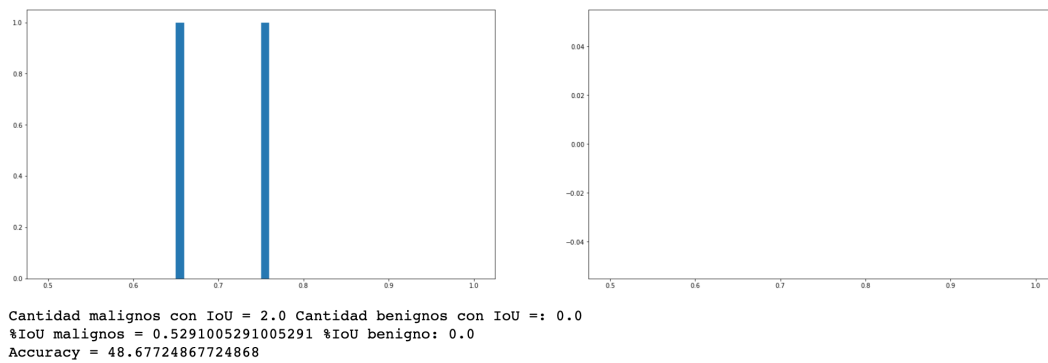


Figura 6.6: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 3

En las figuras 6.7,6.8 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

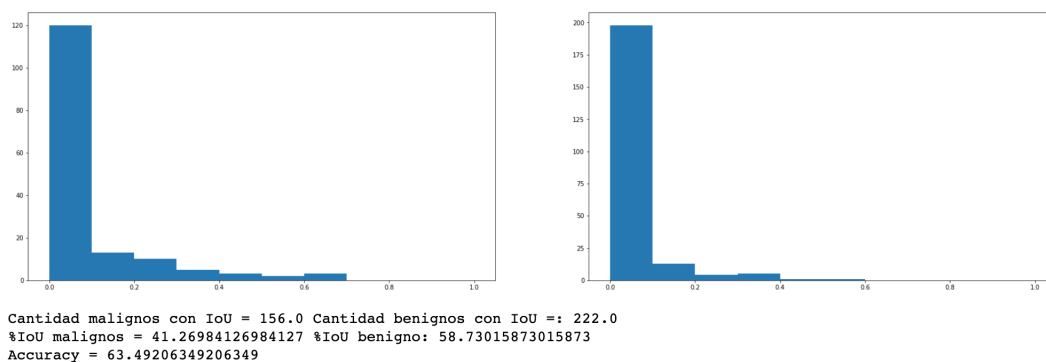


Figura 6.7: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

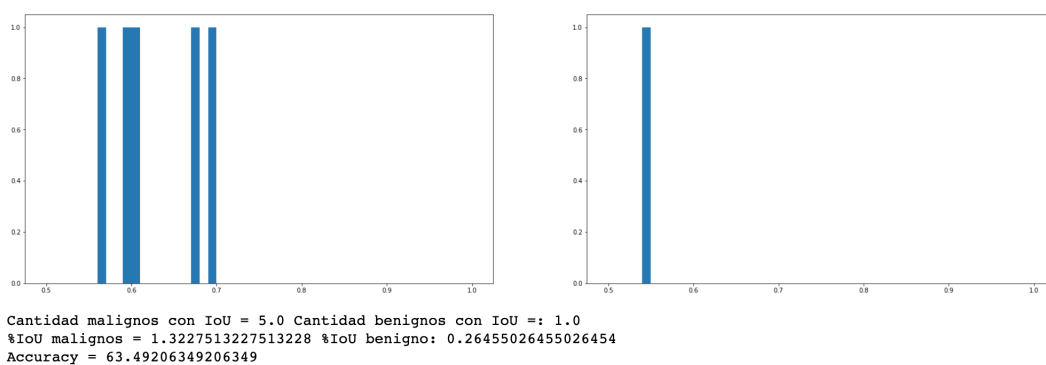


Figura 6.8: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 4

En las figuras 6.9,6.10 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

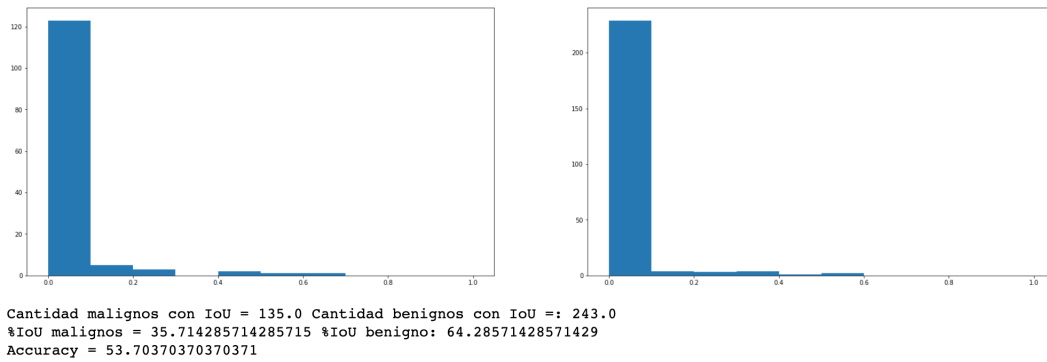


Figura 6.9: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

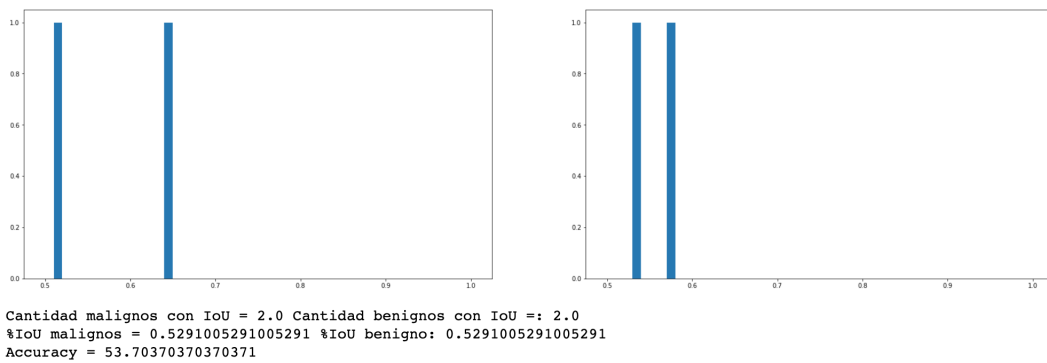


Figura 6.10: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 5

En las figuras 6.11,6.12 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

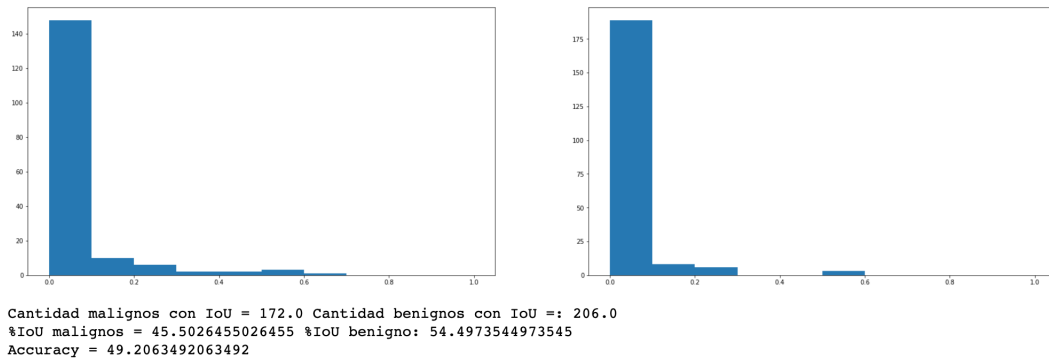


Figura 6.11: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

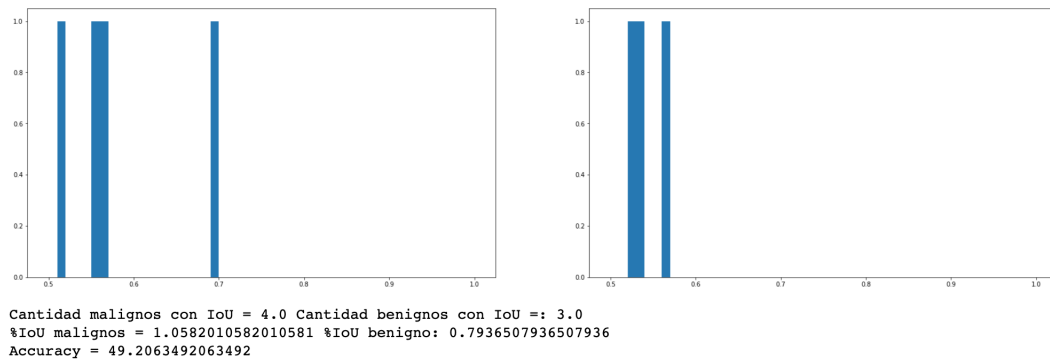


Figura 6.12: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 6

En las figuras 6.13,6.14 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

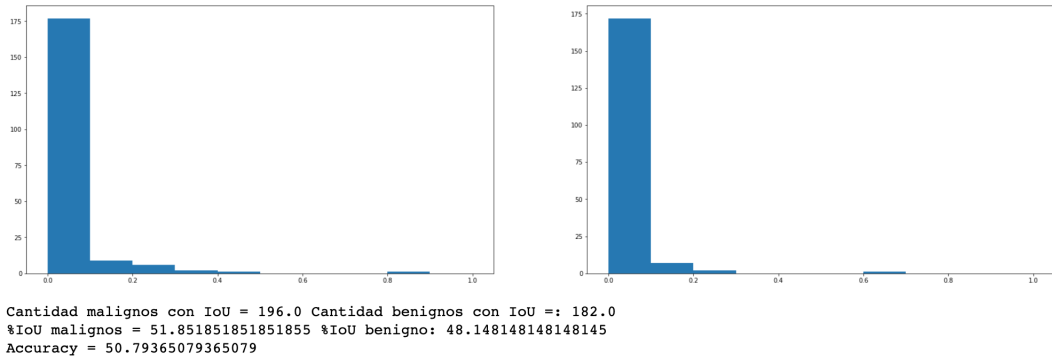


Figura 6.13: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

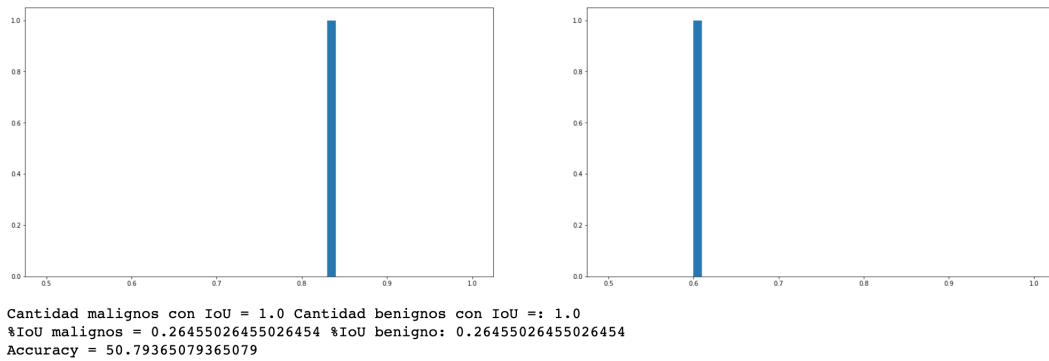


Figura 6.14: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 7

En las figuras 6.15,6.16 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

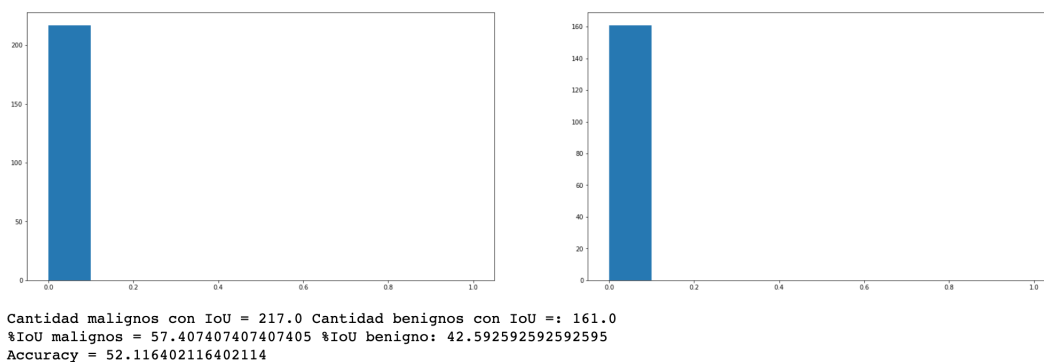


Figura 6.15: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

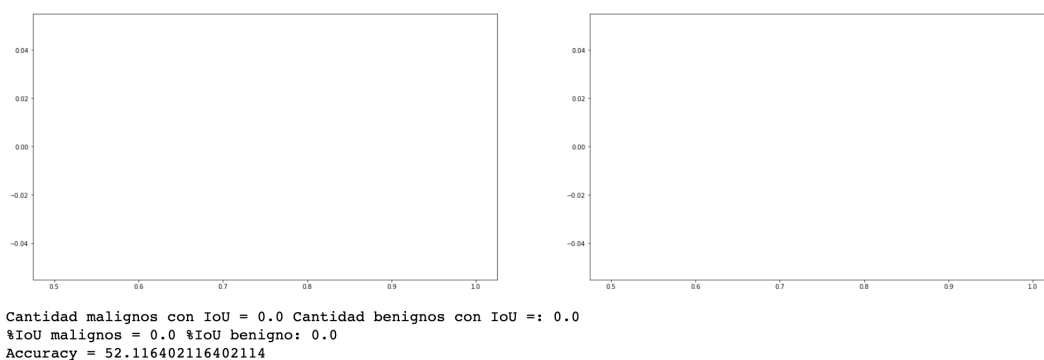


Figura 6.16: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 8

En las figuras 6.17,6.18 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

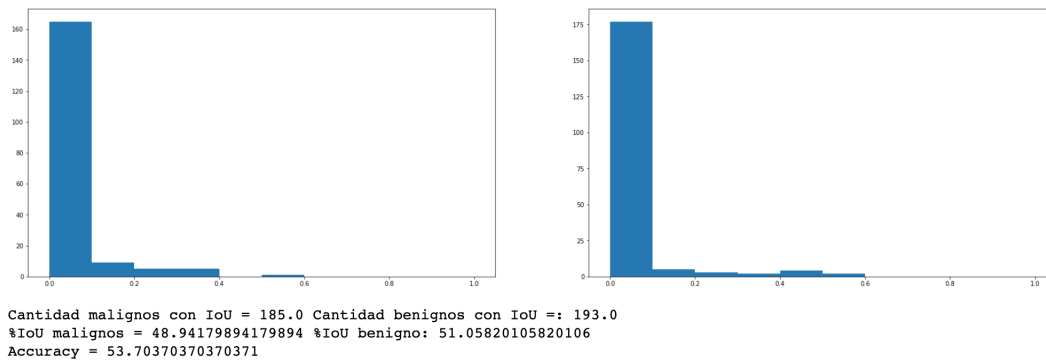


Figura 6.17: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

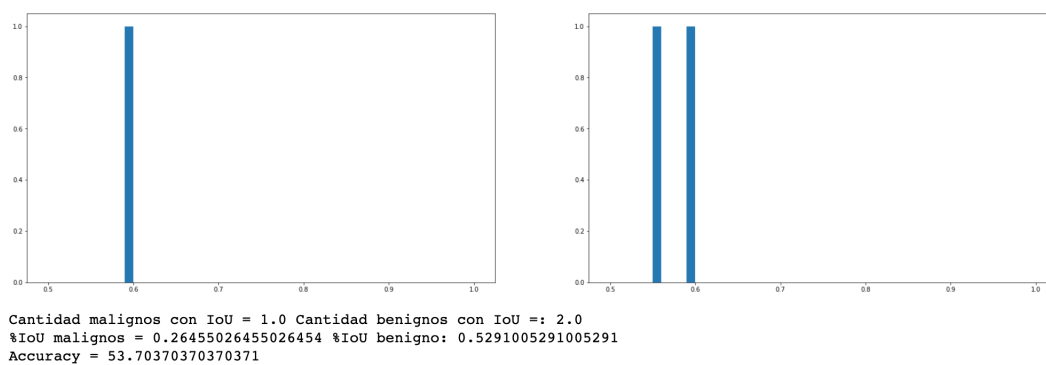


Figura 6.18: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 9

En las figuras 6.19,6.20 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

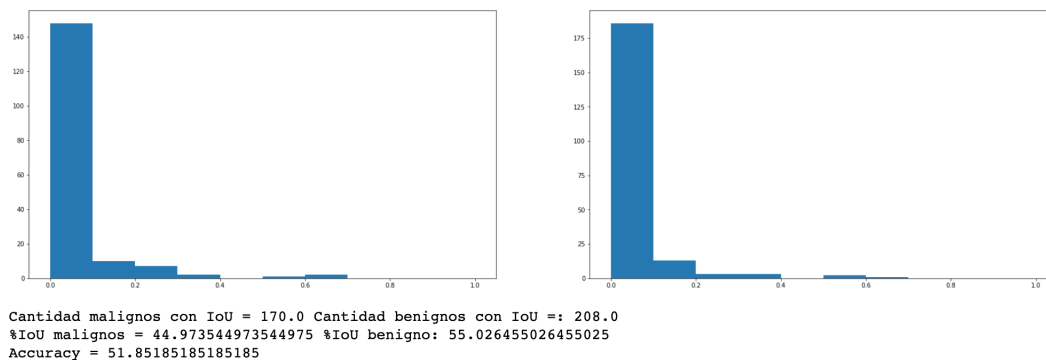


Figura 6.19: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

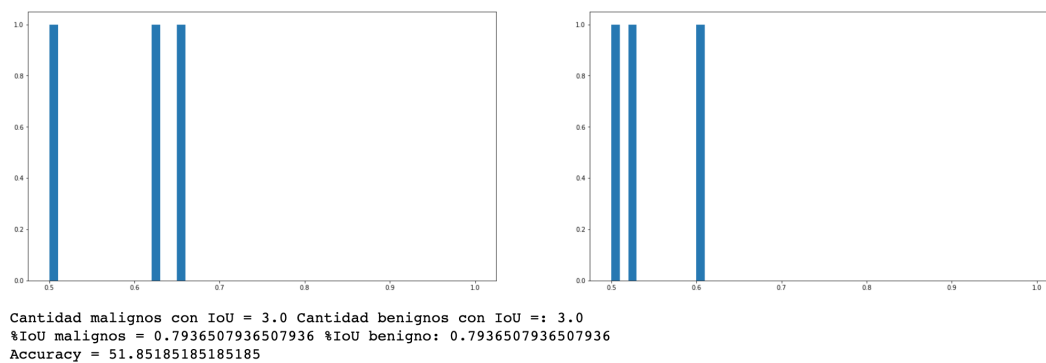


Figura 6.20: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 10

En las figuras 6.21,6.22 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

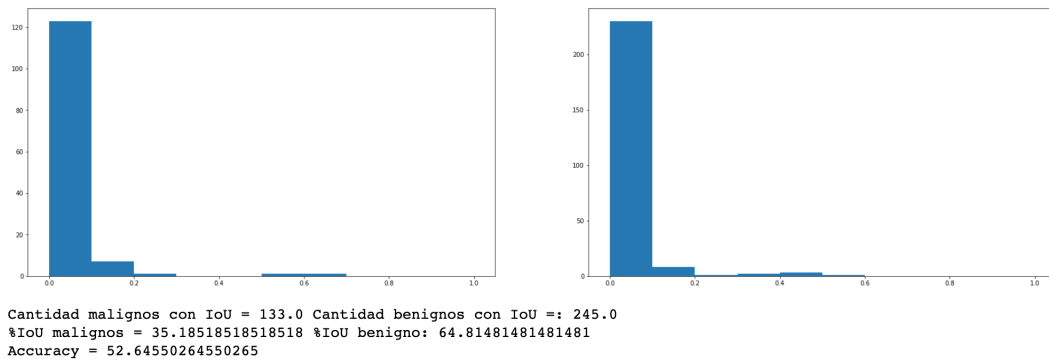


Figura 6.21: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

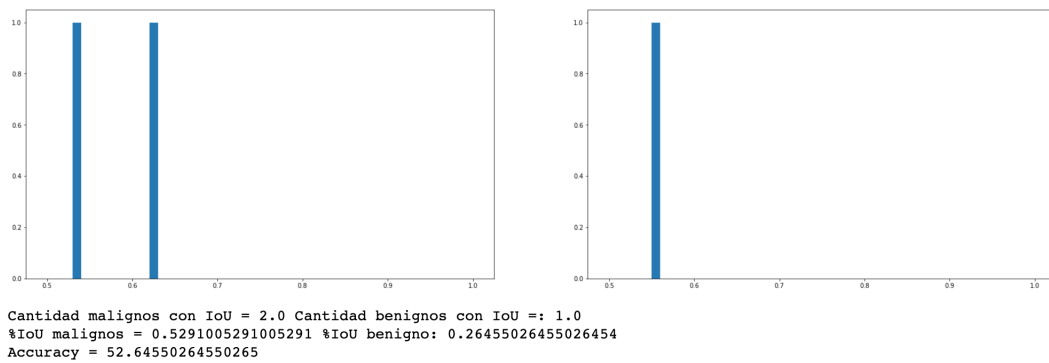


Figura 6.22: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 11

En las figuras 6.23,6.24 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

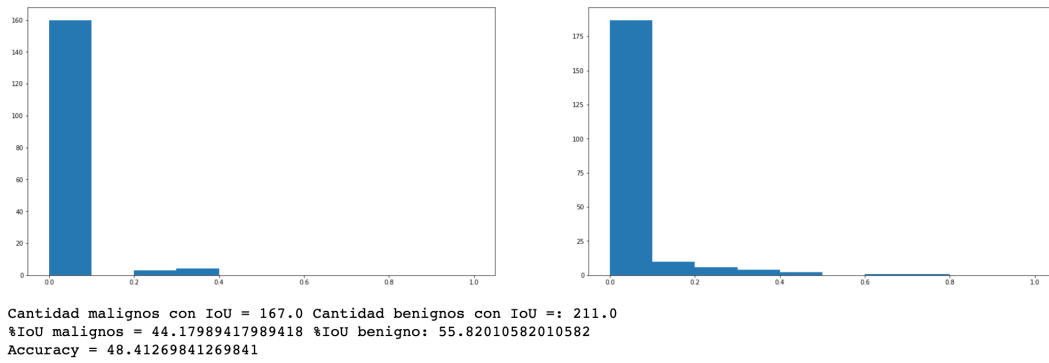


Figura 6.23: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

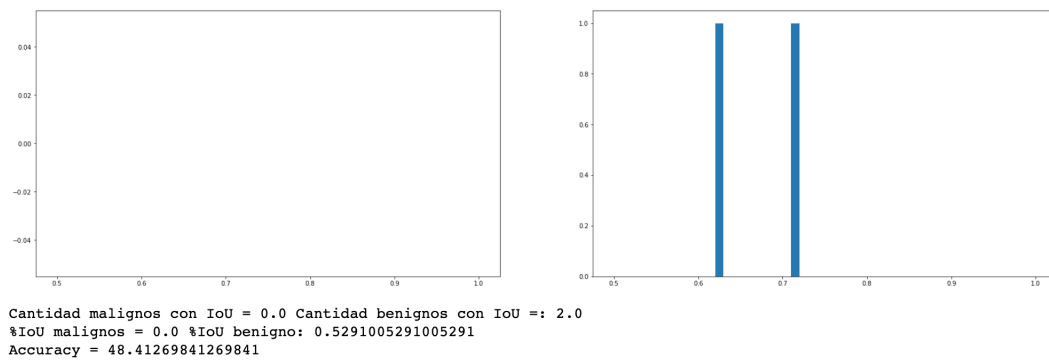


Figura 6.24: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 12

En las figuras 6.23,6.24 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

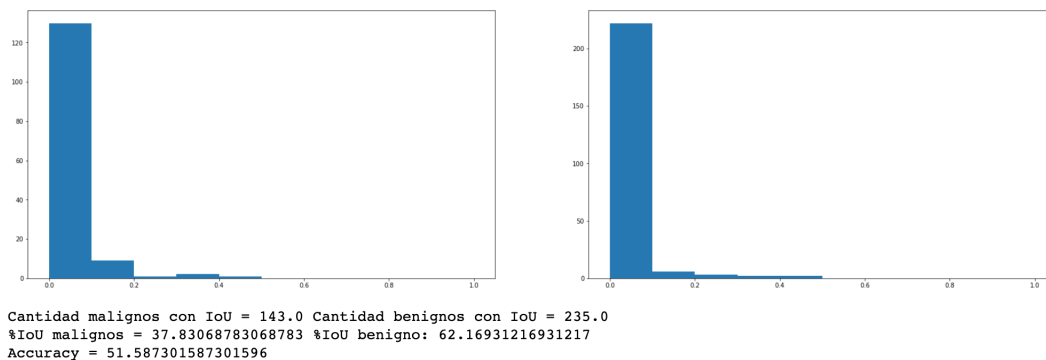


Figura 6.25: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

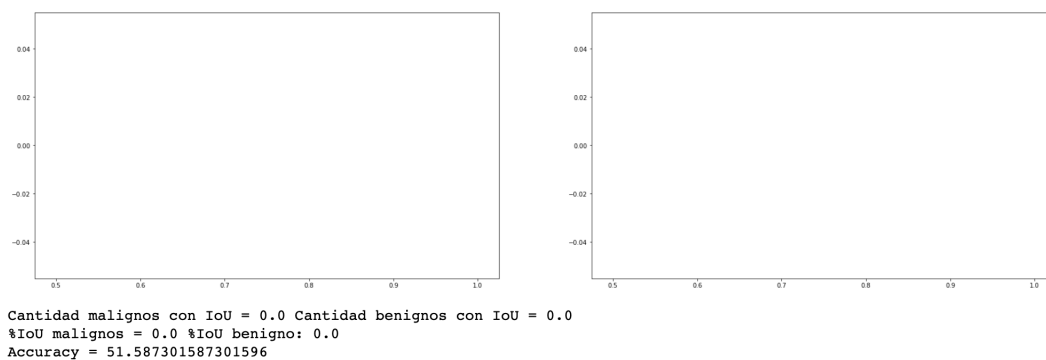


Figura 6.26: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

6.2.2 MRI BRain Segmentation

Modelo 0

En las figuras 6.27,6.28 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

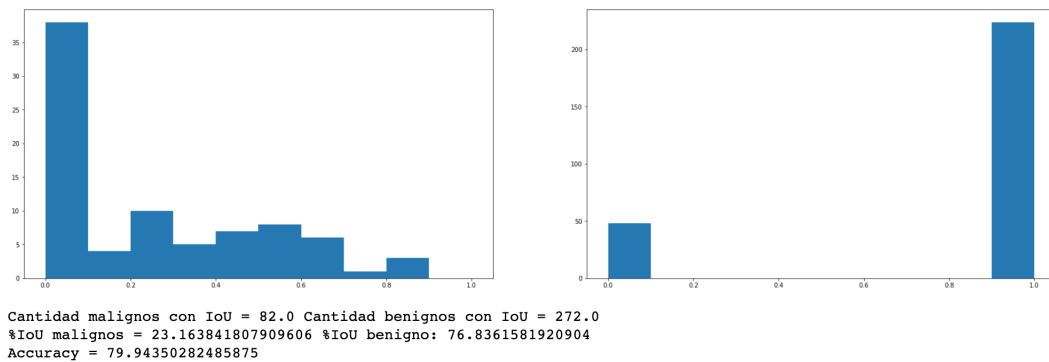


Figura 6.27: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

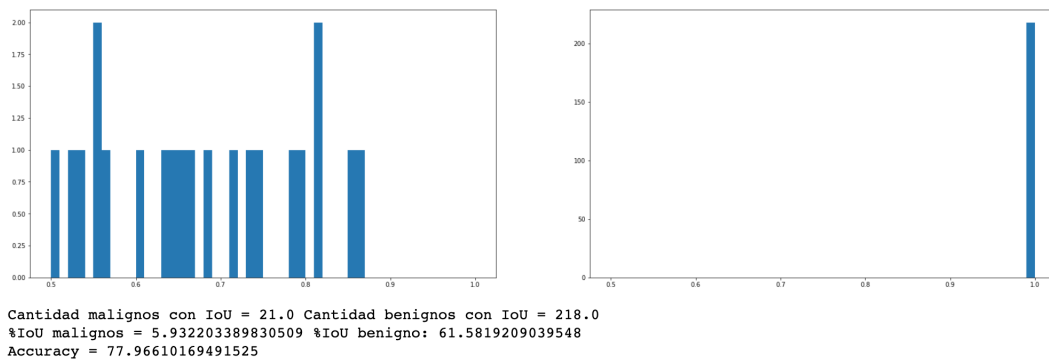


Figura 6.28: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 1

En las figuras 6.29,6.30 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

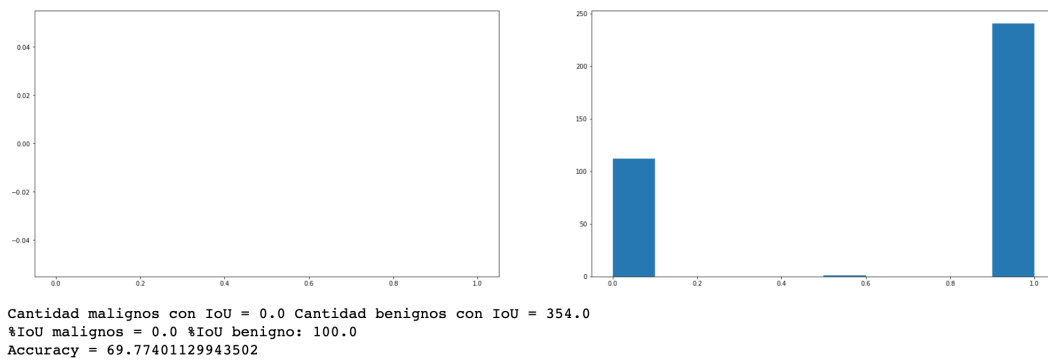


Figura 6.29: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

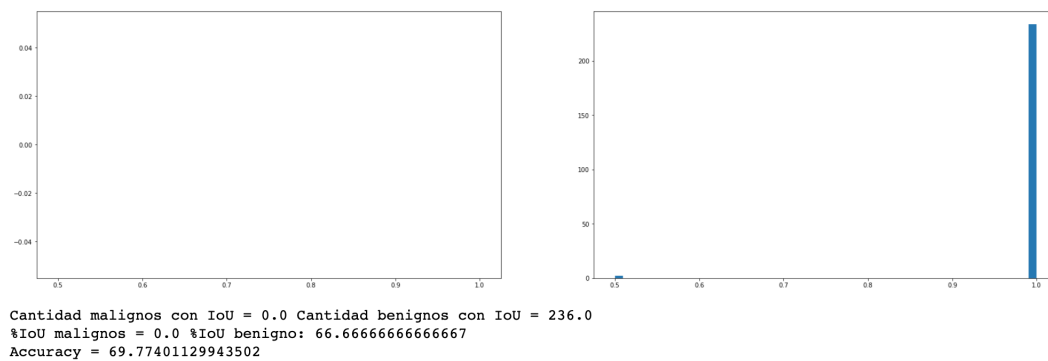


Figura 6.30: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

Modelo 2

En las figuras 6.31,6.32 se puede observar los histogramas el dataset de test. Se utilizando el valor de IoU en el eje de la x y la cantidad de ejemplos en el eje de la y.

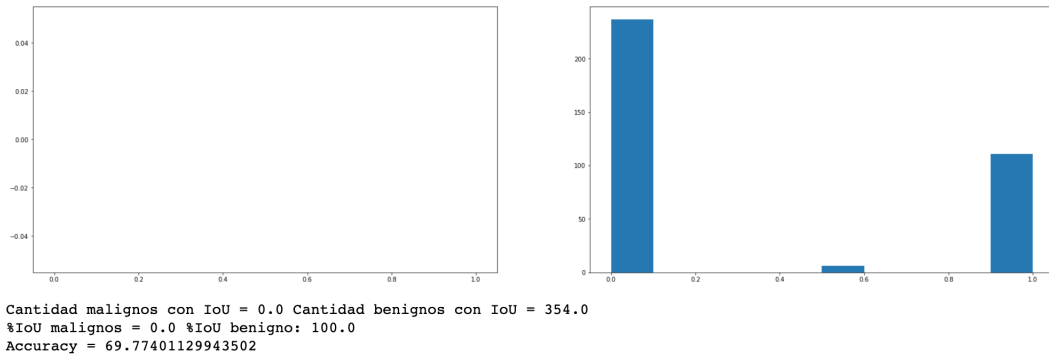


Figura 6.31: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU

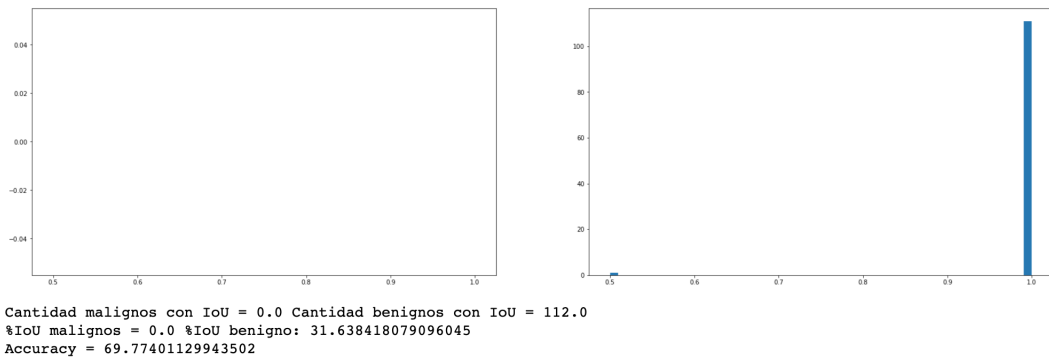


Figura 6.32: Histograma para testing del modelo. En el lado izquierdo figuran los tumores malignos y en la derecha los benignos, cantidad de ejemplos x IOU