



Universidad Autónoma de Querétaro

Facultad de Ingeniería
M.C. en Inteligencia Artificial
Querétaro, Qro.
Mayo 2022
Proyecto II

Detección de anomalías en mamografías

Flores Muñoz Jorge de Jesus
Rodríguez Flores Cecilia Gabriela

Dr. Sebastián Salazar Colores

Optativa II- Deep Learning

ÍNDICE

I. Introducción	3
II. Justificación	4
III. Marco Teórico	5
Composición de la mama	5
Mamografía	5
Detección de objetos	7
Etiquetado	7
Modelos	9
Métricas de evaluación	12
IV. Metodología	13
Base de datos	13
Conjunto de Datos de Microcalcificación de mama	13
Base de datos pública MIAS	14
Desarrollo	15
V. Resultados	18
VI. Análisis de resultados	22
VII. Conclusión	25
VIII. Referencias	26

I. Introducción

El cáncer de mama es una de las principales causas de mortalidad en el género femenino a nivel mundial. Esta enfermedad se suele detectar por medio de diversas técnicas de imagen dependiendo de la gravedad de la situación, pero aun así la mamografía es la técnica más utilizada para la detección de dicha enfermedad.

Así bien, el diagnóstico y la prevención temprana es lo más importante para un buen tratamiento y cura de la posible enfermedad. Las estadísticas muestran un descenso en la mortalidad por cáncer de pecho en las mujeres que se realizan estudios mamográficos frecuentes.

Las microcalcificaciones en la mamografía han sido principalmente un signo precoz fiable del cáncer de mama y su detección es de vital importancia para un mejor pronóstico. Lamentablemente la dificultad radica en su detección debido a la muy pequeña diferencia de contraste en las zonas con lesiones respecto a las zonas sanas de la mama, por lo que las lesiones pudieran pasar desapercibidas cuando están en etapas iniciales de la enfermedad [1].

Actualmente se cuentan con técnicas de inteligencia artificial capaces de analizar imágenes y generar la detección de objetos de interés. La detección de objetos es una tarea importante en muchos campos, como el diagnóstico médico, la navegación robótica, la conducción automática, la realidad aumentada, etc. En este trabajo se hace uso de la red YOLOv4 (por sus siglas en inglés "You Only Look Once"), el modelo Efficient y el modelo Detectron2, las cuales se entrenarán con una base de datos con imágenes de anomalías mamarias, con el fin de realizar una comparación de los resultados obtenidos.

II. Justificación

El presente trabajo es pertinente a la materia de Deep Learning dentro del contexto del procesamiento y análisis de imágenes médicas, permitiendo la detección de anomalías precursoras del cáncer en el tejido mamario en las mamografías, con lo cual puede contribuir a conformar en un futuro una herramienta indispensable en el ámbito médico.

El cáncer de mama es el tipo de cáncer más común a nivel mundial, con alrededor de 685,000 fallecimientos de mujeres como consecuencia de esta enfermedad. Por lo cual, es la principal causa de mortalidad en las mujeres [11].

En México, según el comunicado de prensa preliminar por parte del Instituto Nacional Estadístico y Geografía [6], se han registrado 7880 defunciones por cáncer de mama, entre los cuales 58 casos corresponden al género masculino, por consiguiente, hubo un total de 7821 defunciones correspondientes al género femenino, en donde se concentra un mayor número de fallecimientos en el rango de edad entre los 65 años y más.

El cáncer de mama es una enfermedad en la cual las células mamarias se duplican sin control. Generalmente, la detección en etapas tempranas de esta enfermedad, ayuda en gran medida en su tratamiento y control.

En una etapa preliminar, en la detección del cáncer de mama, se intentan localizar pequeños depósitos de calcio, conocidos como microcalcificaciones, en las arterias de la mama. En la mayoría de los casos, esta búsqueda se hace analizando visualmente las mamografías, en ocasiones se cuenta con una herramienta que permite filtrar las imágenes, permitiendo su detección.

Los algoritmos basados en Inteligencia Artificial representan una vía prometedora para mejorar la precisión de la mamografía, así bien pueden ser utilizados para proveer referencia al radiólogo en la evaluación con el fin de comparar sus propios resultados o bien atraiga la atención del médico especialista a áreas específicas que puede estar pasando por alto.

III. Marco Teórico

Composición de la mama

Biológicamente, el seno es una glándula sudorípara, cuya misión primordial es la producción de leche, cada seno contiene un conjunto de lóbulos que se ramifican desde el pezón, los cuales se encuentran unidos por conductos. Así bien, cuenta también con vasos sanguíneos y vasos linfáticos, estos últimos conducen a los nodos linfáticos, los cuales son racimos que se localizan en varias partes del cuerpo, como son en la axila, clavícula, pecho, entre otros [10].

El seno está compuesto principalmente por tejido adiposo y la glándula mamaria. En general, las mujeres que se encuentran en sus ciclos hormonales, en la etapa del embarazo, lactancia suelen contar con una mayor cantidad de tejido glandular.

A diferencia de la etapa de la menopausia, en donde la glándula se atrofia y el volumen de la mama depende básicamente del tejido adiposo conocido como grasa, por consiguiente, la mama será poco densa.

El sistema linfático es una red de ganglios linfáticos y conductos linfáticos que ayuda a combatir las infecciones, por medio de este sistema se capturan las sustancias perjudiciales para posteriormente expulsarlas del organismo. Este punto tiene especial relevancia sobre todo en los tumores malignos, los cuales usan el drenaje linfático para dispersarse y por consiguiente producir metástasis [18].

Mamografía

La mamografía es una modalidad de imagen de rayos X de baja dosis. Los usos principales de la mamografía son en la detección temprana del cáncer de mama cuando aún no se presentan ningún síntoma ni problema en los senos, así como en la revisión evolutiva del posible cáncer mamario. El tamizaje mamográfico de rutina consiste en tomar una proyección oblicua mediolateral (MLO) y una craneocaudal (CC) de cada mama [5].

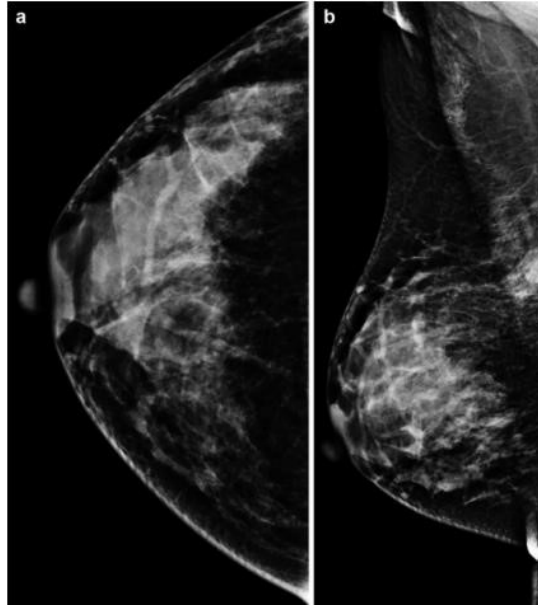


Figura 1. Proyecciones de rutina a) Oblicua Mediolateral (MLO) y b) Craneocaudal (CC).

Esta modalidad es comúnmente utilizada por los médicos especialistas como herramienta para obtener una mejor visualización de cada seno. Así bien, cuenta con un sistema de categorización estandarizada de lesiones mamarias conocido como BI-RADS, el cual fue creado por el Colegio Estadounidense de Radiología (ACR).

Este sistema es comúnmente implementado para la interpretación de los resultados obtenidos en la mamografía, ya que brinda a los médicos especialistas una guía para definir el grado de sospecha de alguna anomalía reduciendo las discrepancias en la interpretación.

El sistema BI-RADS se encuentra dividido en seis categorías, las cuales se basan en las características de las lesiones observables en la mamografía. En la tabla siguiente se visualiza cada una de las categorías con su correspondiente descripción, características e intervención [1].

Tabla 1. Categoría del Sistema BI-RADS.

Categoría	Descripción	Características
0	Mastografía incompleta. Evaluación de imagen adicional necesaria y/o mamografías anteriores para su comparación	No valorable
1	Estudio negativo	Normal. Mamas simétricas, sin nódulos, distorsiones ni calcificaciones
2	Estudio negativo con hallazgos benignos	Fibroadenomas, calcificaciones de la piel o vasculares, quistes, galactoceles, hamartomas, ganglios linfático
3	Hallazgo probablemente benignos	Calcificaciones amorfas ya agrupadas, Nódulo solitario normal, Densidad, Asimetría Focal
4	Hallazgos sospechosos 4A: Baja sospecha de malignidad 4B: Moderada sospecha de malignidad 4C: Alta sospecha de malignidad	Nódulos irregulares Densidad similar a parénquima Microcalcificaciones heterogéneas Perdida arquitectura de mama Ganglios aumentados de tamaño
5	Alta probabilidad de malignidad	Nódulo irregular, espiculado Densidad mayor a parénquima Microlobulado Microcalcificaciones morfología lineal
6	Malignidad comprobada	Confirmado por biopsia

Nota. Adaptado del libro Sistema de informes y registro de datos de estudios por imágenes de la mama.

Detección de objetos

La detección de objetos se basa en redes neuronales convolucionales que han sido entrenadas previamente con una gran cantidad de diferentes tipos de objetos, es decir, imágenes, con las cuales las redes van a aprender y a diferir entre ellas. Así bien, al pasar una imagen de un objeto de la misma clase la red será capaz de detectar el objeto incluso cuando presente diferencias de tamaño, rotación, etc.

A lo largo del tiempo han surgido varias soluciones que han buscado obtener el resultado más preciso y eficiente, además de contar con un coste computacional bajo.

Etiquetado

La anotación de imágenes es un tipo de etiquetado de datos, el cual se aplica tanto a la anotación de imagen como de video. Al igual que un conjunto de imágenes, los videos también se pueden anotar continuamente, como una fuente de imagen o cuadro por cuadro.

Esta técnica se utiliza principalmente para identificar objetos y límites, para posteriormente categorizarlos, por ejemplo, el significado o la comprensión de la imagen completa. Se necesita una cantidad considerable de tiempo para entrenar, validar y evaluar un modelo para obtener el resultado esperado para todas estas aplicaciones.

- Anotación de imagen simple: se enfoca en el proceso para etiquetar una imagen con las palabras que mejor representan los objetos mencionados en ella. Por ejemplo, puede anotar la imagen de un perro con la etiqueta "perro doméstico".
- Anotación de imagen compleja: se define como un tipo de anotación de imagen que se puede usar para reconocer, contar o rastrear más de un objeto o región en una imagen. La complejidad de su anotación diferirá según la complejidad del proyecto en cuestión. Por ejemplo, el etiquetado de las diferentes razas de perros.

Software de etiquetado

Existe una variedad de herramientas especializadas para la realización del proceso de etiquetado como son Labellmg, Roboflow, VGG Image Annotator, Supervise, MakeSense, RectLabel, etc. [13].

Labellmg

Labellmg es un programa de software gratuito para etiquetar imágenes. El software fue lanzado por Tzutalin en 2015, el cual está escrito en Python; utiliza QT para su GI (interfaz gráfica). Así bien, es una herramienta sencilla y básica para etiquetar unos pocos cientos de imágenes para crear un conjunto de datos para el entrenamiento de algún modelo de visión por computadora.

Labellmg admite el etiquetado en formato de archivo de texto YOLO o VOC XML. Se recomienda que los usuarios utilicen el formato XML VOC predeterminado para producir etiquetas. Debido a que es el formato de ImageNet, es un estándar universal ya que está asociado con la detección de objetos, mientras que las diferentes implementaciones de YOLO tienen diferentes formatos de línea de texto. Además, puede convertir fácilmente de VOC XML a

cualquier otro formato utilizando herramientas. Por ejemplo, puede convertir de VOC XML a COCO JSON.

Roboflow

Roboflow permite a los desarrolladores crear sus propias aplicaciones de visión por computadora, sin importar su conjunto de habilidades o experiencia. Proporciona las herramientas para convertir imágenes sin procesar en un modelo de visión por computadora capacitado personalizado y lo implementa para su uso en aplicaciones. Roboflow admite modelos de detección y clasificación de objetos [15].

Roboflow se utiliza para:

- Imágenes o cargar anotaciones existentes
- Las anotaciones VOC XML existentes en anotaciones COCO JSON
- Genera formatos de anotación como TFRecords, CreateML y Turi Create, e implementaciones personalizadas de YOLOv3 (archivos de texto sin formato o Darknet).

Modelos

Red You Only Look Once (YOLO)

La red YOLO como sugiere su nombre, una sola "mirada" es suficiente para encontrar todos los objetos en una imagen e identificarlos, su arquitectura se puede apreciar en la figura 2. En términos de aprendizaje automático, podemos decir que todos los objetos se detectan a través de una única ejecución de algoritmo. Se hace dividiendo una imagen en una cuadrícula y prediciendo cuadros delimitadores y probabilidades de clase para cada celda en una cuadrícula.

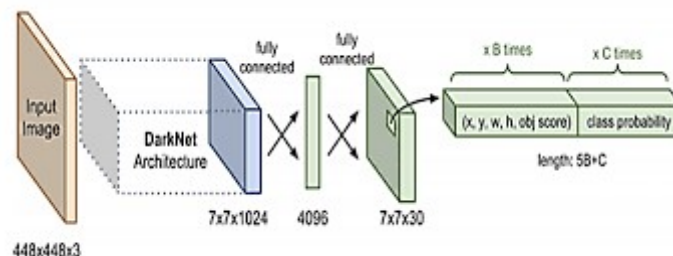


Figura 2. Arquitectura de una red YOLO.

Vale la pena señalar que la salida sin procesar de YOLO contiene muchos cuadros delimitadores para el mismo objeto, estas cajas difieren en forma y tamaño, algunos cuadros son mejores para capturar el objeto de destino, mientras que otros ofrecidos por un algoritmo funcionan mal.

Para seleccionar el mejor cuadro delimitador para un objeto determinado, se aplica un algoritmo de supresión no máxima (NMS), el cual es una técnica para filtrar las predicciones de los detectores de objetos, la canalización típica de detección de objetos tiene un componente para generar propuestas de clasificación. Las propuestas no son más que las regiones candidatas para el objeto de interés.



Figura 3. Supresión no máxima.

Todas las casillas que predice YOLO tienen un nivel de confianza asociado. NMS utiliza estos valores de confianza para eliminar las casillas que se predijeron con poca certeza. Por lo general, estos son todos los cuadros que se predicen con confianza por debajo de 0,5.

Cuando se eliminan todos los cuadros delimitadores inciertos, solo quedan los cuadros con el nivel de confianza alto. Para seleccionar el mejor entre los candidatos de mayor rendimiento, NMS selecciona la casilla con el nivel de confianza más alto y calcula cómo se cruza con las otras casillas alrededor. Si una intersección es más alta que un nivel de umbral particular, se elimina el cuadro delimitador con menor confianza. En caso de que NMS compare dos casillas que tienen una intersección por debajo de un umbral seleccionado, ambas casillas se mantienen en las predicciones finales [19].

EfficientDet

EfficientDet es una familia de modelos que expresan la misma arquitectura en diferentes escalas de tamaño de modelo. La arquitectura EfficientDet fue escrita por Google Brain, la cual está construida sobre EfficientNet, una red neuronal convolucional que está preentrenada en la base de datos de imágenes de ImageNet para la clasificación. EfficientDet agrupa y mezcla porciones de la imagen en granularidades dadas y forma características que pasan a través de una capa de fusión de características NAS-FPN, la cual combina varias funciones en diferentes granularidades y las pasa al cabezal de detección, donde se predicen los cuadros delimitadores y las etiquetas de clase [16].

El modelo EfficientDet logró una mayor precisión con la menor cantidad de épocas de entrenamiento con tareas de detección de objetos. Por lo cual, esta arquitectura logró superar al modelo RetinaNet, Mask R-CNN y YOLOv3 [8].

Detectron2

Detectron2 es un modelo de visión por computadora propio escrito en PyTorch por el grupo de investigación de IA de FAIR Facebook, el cual utiliza conjuntos de datos en formato COCO JSON. Este modelo incluye todos los modelos que estaban disponibles en el Detectron original, como son los modelos Faster R-CNN, Mask R-CNN, RetinaNet y DensePose, así como algunos modelos más nuevos, incluidos Cascade R-CNN, Panoptic FPN y TensorMask. Así bien, Detectron2 puede ser utilizado para realizar la detección de puntos clave, la detección de objetos y la segmentación semántica [17].

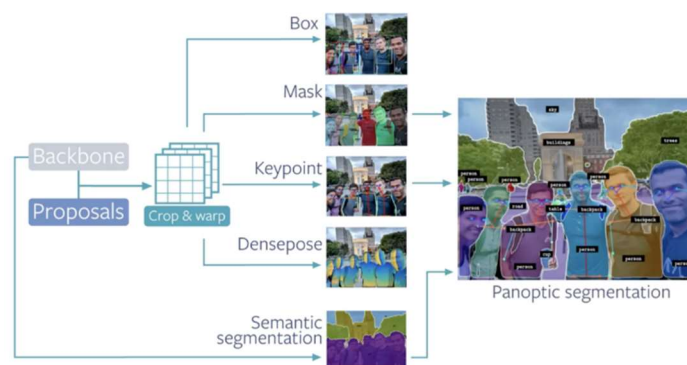


Figura 4. Marco flexible del modelo Detectron2.

Métricas de evaluación

Para la tarea de detección de objetos, es necesario el uso de métricas que se adapten al problema e indiquen que tan precisas son las predicciones. Como solución, existen varias métricas frecuentemente usadas, como Intersección sobre Unión (IoU), precisión promedio (AP) y precisión promedio media (mAP).

Intersección sobre Unión (IoU)

Es una métrica de evaluación basada en el índice Jaccard, el cual evalúa la superposición entre dos cuadros delimitadores, es decir, requiere un cuadro delimitador de verdad y un cuadro delimitador previsto. Usualmente esta métrica de evaluación es utilizada en retos de evaluación de objetos como es Pascal VOC challenge. Para poder aplicar este método es necesario tener las etiquetas que especifican las coordenadas de los objetos en las imágenes y las coordenadas predichas por el detector [19].

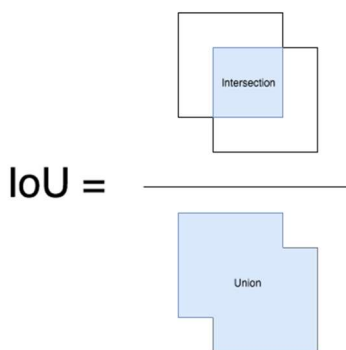


Figura 5. Intersección sobre Unión.

Precisión promedio (AP)

Otra forma de comparar el rendimiento de los detectores de objetos como Faster R-CNN, SSD, etc. es por medio de la precisión promedio (AP), la cual es una métrica popular que resume una curva de recuperación de precisión como la media ponderada de las precisiones logradas en cada umbral, con el aumento en la recuperación desde el umbral anterior utilizado como peso [20]:

$$AP = \sum_n (R_n - R_{n-1})P_n$$

Donde:

P_n y R_n son la precisión y la recuperación en el umbral n .

Precisión promedio media (mAP)

La precisión promedio media es el promedio de AP de cada clase. Sin embargo, la interpretación de AP y mAP varía en diferentes contextos. El mAP se calcula encontrando la precisión promedio (AP) para cada clase. Por otra parte, el código de los modelos utilizados tiene métodos implementados para computar estas métricas automáticamente; y luego promediando varias clases.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

El mAP incorpora el compromiso entre precisión y recuperación, considera tanto los falsos positivos (FP) como los falsos negativos (FN). Esta propiedad hace que mAP sea una métrica adecuada para la mayoría de las aplicaciones de detección[14].

IV. Metodología

Base de datos

Para la realización de este trabajo se decidió utilizar una base de datos, la cual es una fusión de dos bases de datos, en donde únicamente se seleccionan las mamografías que contengan alguna anomalía, por lo cual el total de imágenes a implementar es de 272 imágenes, en donde 225 imágenes son del conjunto de datos de Microcalcificación de mama y 47 imágenes son de la base de datos MIAS. A continuación, se dará una breve descripción de cada una de las bases de datos a utilizar.

Conjunto de Datos de Microcalcificación de mama

El conjunto de datos consta de 100 pares de mamografías, las cuales contienen vistas CC (cráneo caudal) y MLO (medio lateral oblicuo) de cada paciente. Se trata de una completa base de datos para la detección y clasificación BI-RADS de microcalcificaciones mamarias, mediante mamografías digitales [7].

Tabla 2. Características sobre la base de datos.

Formato de la imagen	Total de imágenes	Vista de la Imagen	Información contenida en la base de datos
.dcm .jpg	400 imágenes	MLO CC	<p>Casos normales contiene 50 pacientes Casos sospechosos, contiene 50 pacientes.</p> <p>Cada una de las carpetas de los pacientes contiene:</p> <ul style="list-style-type: none"> - 4 archivos en formato.dcm - 4 archivos en formato.jpg

Base de datos pública MIAS

El conjunto de datos MIAS se utiliza ampliamente para fines académicos debido a que no proporciona información personal de las pacientes en cuestión. En la siguiente tabla se pueden visualizar las características sobre esta base de datos [9].

Tabla 3. Características sobre la base de datos MIAS.

Formato de la imagen	Total de imágenes digitalizadas	Vista de la Imagen	Información contenida en la base de datos
.pgm	322 imágenes, de las cuales 299 – Normales 51 – Malignas 62 - Benignas	MLO	<p>Número de referencia de cada mamografía</p> <p>Característica del tejido:</p> <ul style="list-style-type: none"> - F Grasa - G Grasa-glandular - D Denso-glandular <p>Clase de anomalía presente:</p> <ul style="list-style-type: none"> - CALC Calcificación - CIRC Masas bien definidas/ circunscritas - SPIC Masas Espiculadas - MISC Masas mal definidas, Otras - ARCH Distorsión de la Arquitectura - ASYM Asimetría - NORM Normal <p>Gravedad de la anomalía</p> <ul style="list-style-type: none"> - B Benigno - M Maligno <p>Coordenadas (x, y) del centro de la anomalía en la imagen</p> <p>Radio aproximado (en pixeles) de un círculo que encierra la anomalía</p>

Desarrollo

Para la elaboración de este trabajo, se hizo uso de las redes Yolov4, EfficientDet D0 y Detectron2. La metodología utilizada para la implementación de los modelos mencionados anteriormente fue muy similar entre sí. Antes de detallar cada uno de los modelos propuestos, se procede a mencionar el proceso de acondicionamiento de la base de datos.

Primeramente, fue necesario realizar una modificación en el formato en el que se encuentran las mamografías, debido a que en el formato en el que se encuentran no es compatible con el software Labellmg. Por lo cual, se decidió pasar del formato.dcm al formato .jpg y del formato.pgm al formato .jpg.

Posteriormente, una vez realizado el paso anterior se prosiguió con la modificación de los nombres de las imágenes con el fin de no tener dificultades en los siguientes pasos. Así bien, se prosiguió con el etiquetado de las imágenes por medio del software Labellmg, en donde se decidió implementar únicamente dos clases: Anomalía y Anomalía sospechosa. Los archivos de las anotaciones se generaron en formato.xml, formato.json y formato.txt.

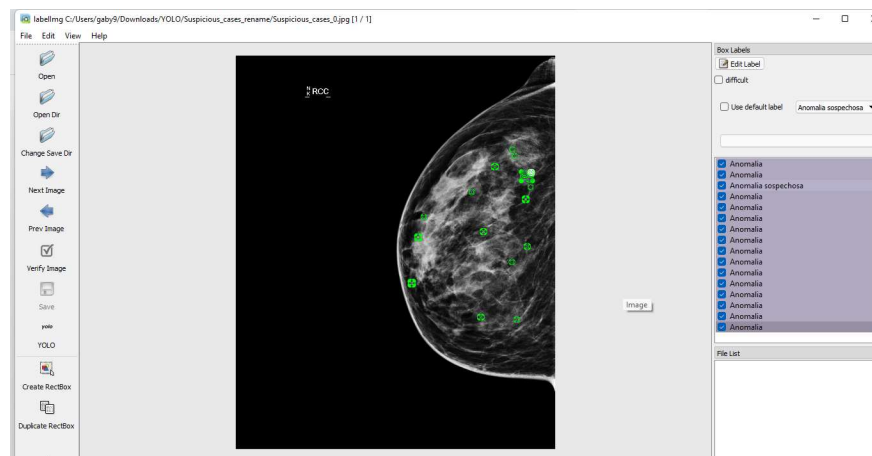
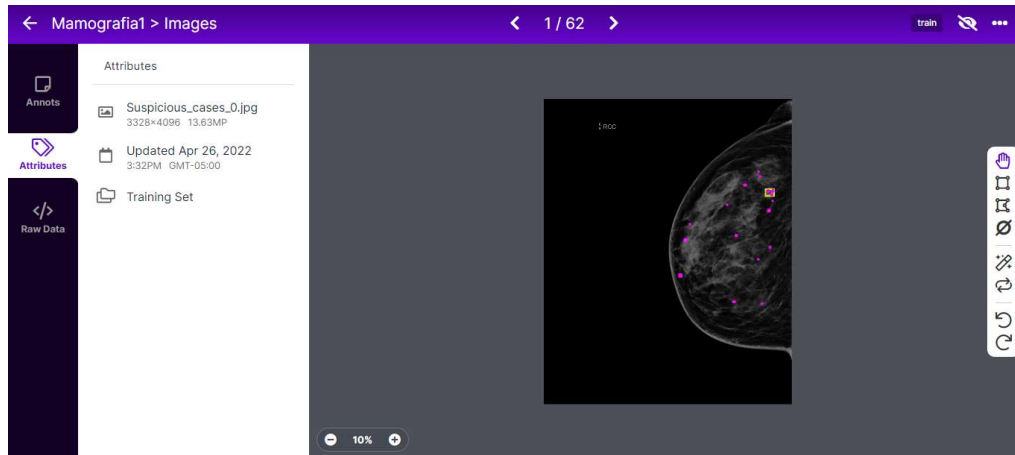


Figura 6. Etiquetado de las mamografías en Labellmg.

Así mismo, para la implementación de los modelos EfficientDet D0 y Detectron2 fue necesario ingresar la base de datos junto con las etiquetas en formato.json a la página de Roboflow, en donde se pudieron obtener las etiquetas en el formato TFrecord y formato COCO.



Antes de realizar la implementación de los modelos fue necesario separar la base de datos en un conjunto de entrenamiento, el cual cuenta con un 60% de la base de datos; un conjunto de validación, el cual únicamente contiene un 20% de la base de datos; y un conjunto de prueba, en donde se cuenta con un 20% de la base de datos.

Para la separación de la base de datos para la implementación del modelo YOLOv4, se realizó la división de la base de datos por medio de un código de separación en Google Colab, en donde fue necesario poner los porcentajes para cada conjunto de datos.



Figura 8. División de la base de datos 60x20x20.

Así mismo, la separación del conjunto de datos para la implementación de los modelos de Detectron2 y EfficientDet D0 se realizó por medio de la página de Roboflow.

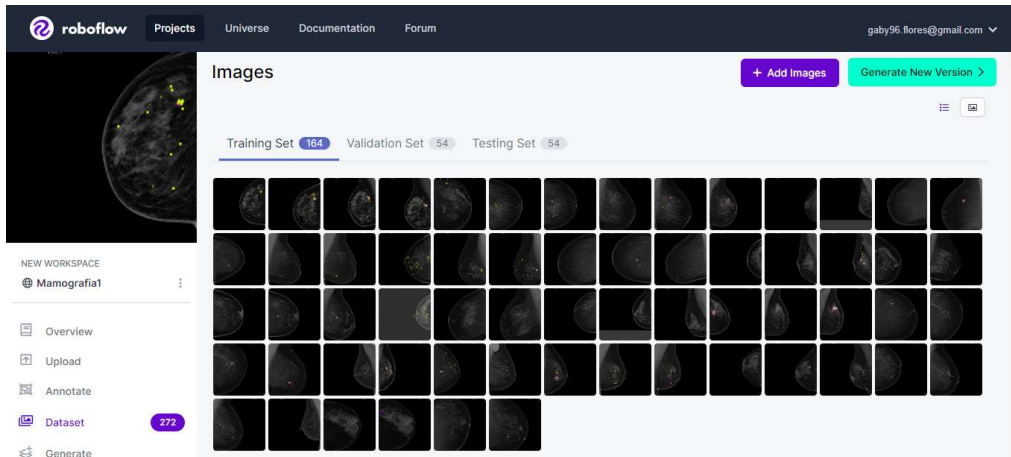


Figura 9. División de la base de datos en Roboflow.

Para cada uno de los modelos se utilizaron redes previamente entrenadas. Así como también la utilización de diferentes tutoriales, en donde se debieron realizar ciertas modificaciones para la utilización de la base de datos propuesta como son el link de Roboflow en donde se encontraba la base de datos con sus correspondientes etiquetas, la ruta de las carpetas de los conjuntos de entrenamiento, validación y evaluación en Drive; número de clases, actualización de ciertas versiones de librerías, etc.

Tabla 4. Tutoriales para la implementación de los modelos.

Modelo	Tutorial
YOLOv4	https://www.youtube.com/watch?v=SBNlIdAcq5k
EfficientDet D0	https://colab.research.google.com/drive/1MJSm8U_ctQ8yl2G7fdNG4mIDw5fWgvQI
Detectron2	https://colab.research.google.com/drive/1-TNOcPm3Jr3fOJG8rnGT9gh60mHUsvaW#scrollTo=kc8MmgZugZWR

Así bien, por medio de la página de Roboflow fue posible realizar un entrenamiento rápido con un modelo propuesto por la misma página con el propósito de saber de una manera anticipada los posibles resultados que se pueden obtener al implementar la base de datos propuesta con sus correspondientes etiquetas.

V. Resultados

Después de la etapa de entrenamiento y evaluación fue posible obtener diversos resultados para cada uno de los modelos propuestos, por lo cual a continuación se mostrarán los diferentes parámetros obtenidos, así como las predicciones del conjunto de evaluación de cada uno de los modelos.

Aunque primeramente se visualizará el resultado obtenido por parte del modelo rápido propuesto por Roboflow, esto con el fin de tener un posible pronóstico de los resultados conforme a la base de datos propuesta.

Tabla 5. Parámetros obtenidos del modelo Roboflow 2.0 Object Detection(Fast).

mAP	Precisión	Recall
2.1%	32.4%	4.4%



Figura 10. Gráficos del conjunto de entrenamiento.

YOLOv4

Tabla 6. Parámetros obtenidos con la red YOLOv4.

Iteración	<u>mAP@0.5</u>	Mejor mAP	Avg loss	Learning Rate	Tiempo de entrenamiento
1605	10.29%	10.38%	1.679972	0.001000	4 hrs

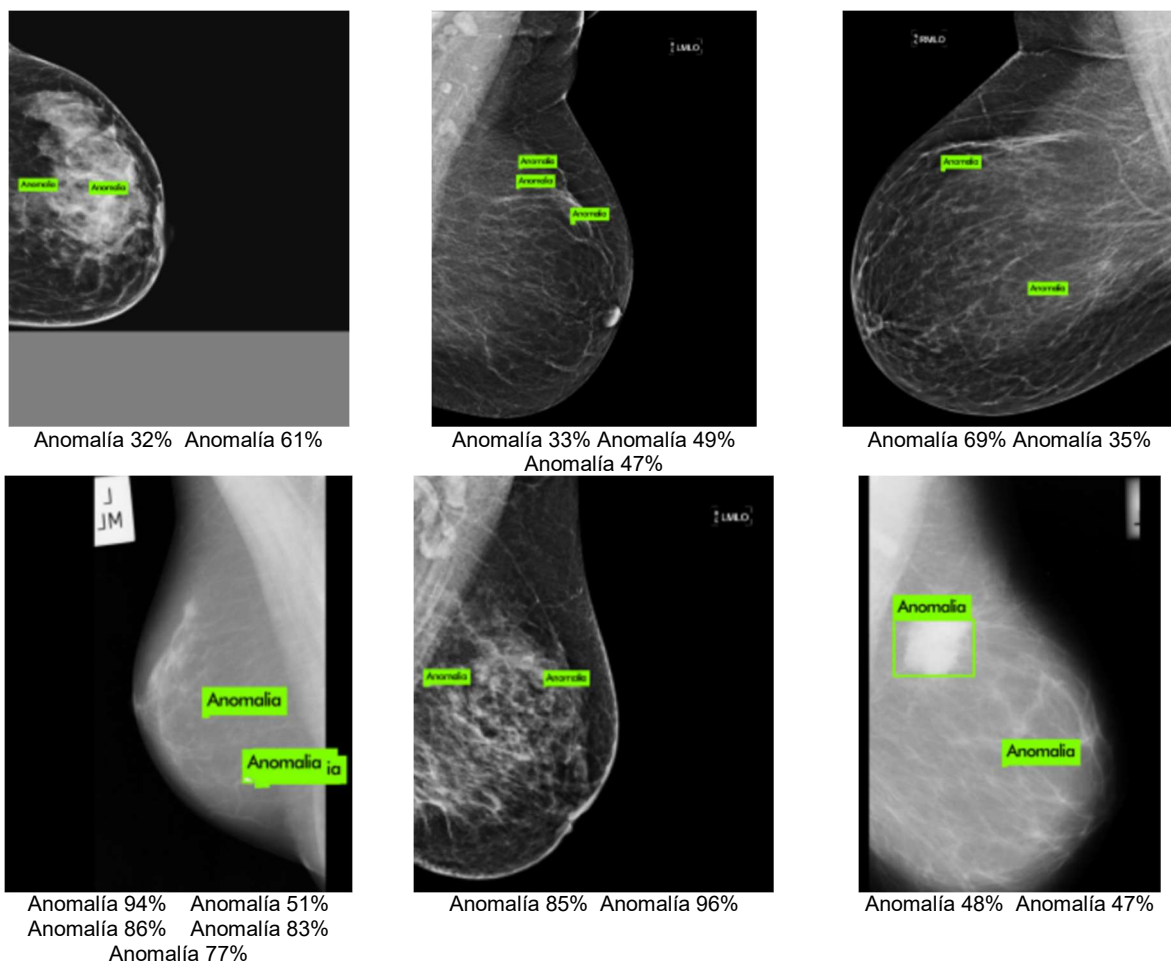


Figura 11. Predicciones del modelo YOLOv4.

EfficientDet D0

Tabla 7. Parámetros obtenidos con el modelo EfficientDet D0.

Pérdida de clasificación	Pérdida de localización	Pérdida de regularización	Pérdida Total	Tasa de Aprendizaje	Tiempo de entrenamiento
0.35131317	0.5757987	0.06644203	0.9935539	0.0799058	2 hrs

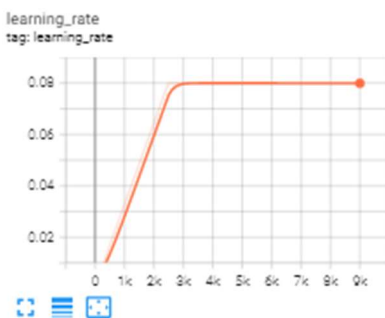


Figura 12. Gráfica de la Tasa de Aprendizaje en TensorBoard.

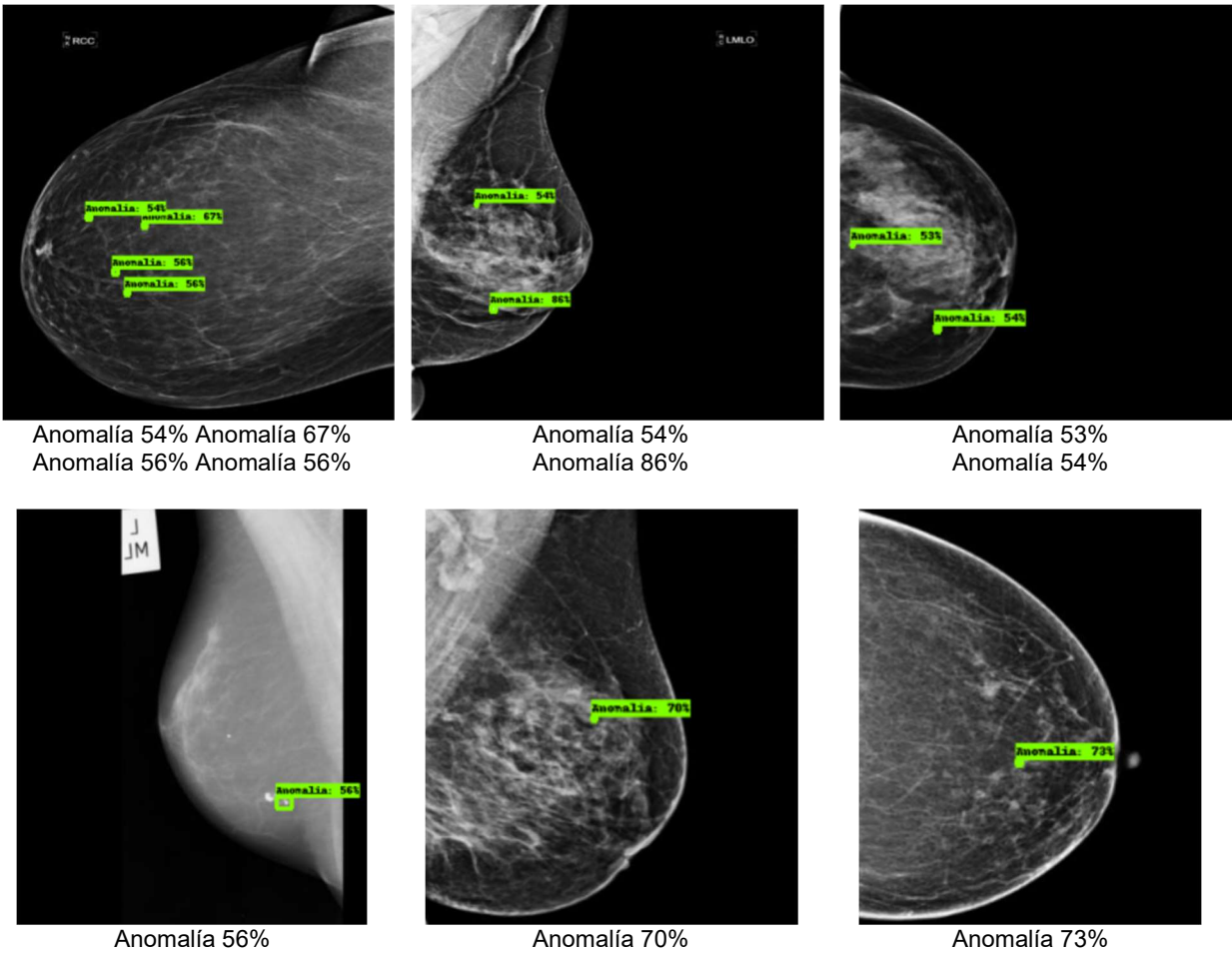


Figura 13. Resultados del modelo EfficientDet D0.

Detectron2

Etapas de entrenamiento

Tabla 8. Valores obtenidos en la etapa de entrenamiento.

AP	AP50	AP75	APs	APm
2.997	5.565	2.698	3.168	0.000

Tabla 9. Valores obtenidos para cada categoría en la etapa de entrenamiento.

Categoría	AP	Categoría	AP
Anomalía	2.312	Anomalía sospechosa	3.682

Tabla 10. Parámetros obtenidos en la etapa de entrenamiento.

Total de iteraciones	Total de pérdida	Tasa de Aprendizaje	Tiempo de entrenamiento
1499	0.949	0.000050	30 min

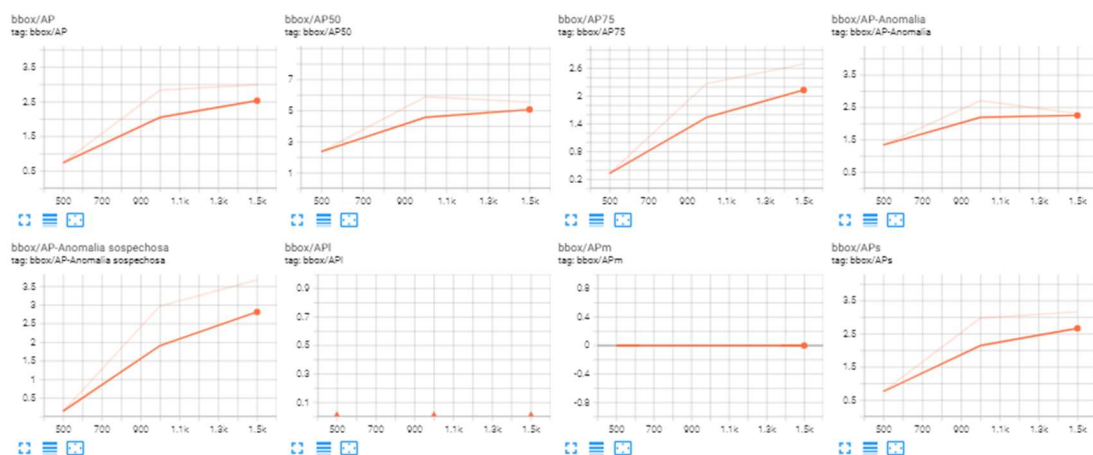


Figura 14. Gráficas de los resultados obtenidos en la etapa de entrenamiento en TensorBoard.

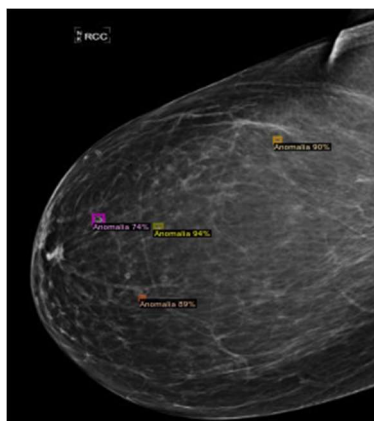
Etapa de evaluación

Tabla 11. Valores obtenidos en la etapa de evaluación.

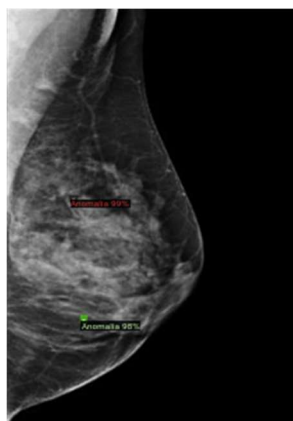
AP	AP50	AP75	APs	APm
3.679	11.022	1.633	3.689	1.087

Tabla 12. Valores obtenidos para cada categoría en la etapa de evaluación.

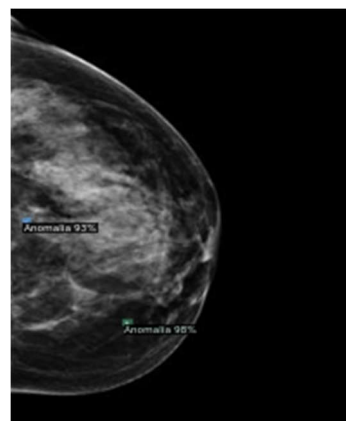
Categoría	AP	Categoría	AP
Anomalia	4.203	Anomalia sospechosa	3.155



Anomalia 90% Anomalia 74%
Anomalia 94% Anomalia 89%



Anomalia 99%
Anomalia 98%



Anomalia 93%
Anomalia 98%

Figura 15. Resultados del modelo Detectron2.

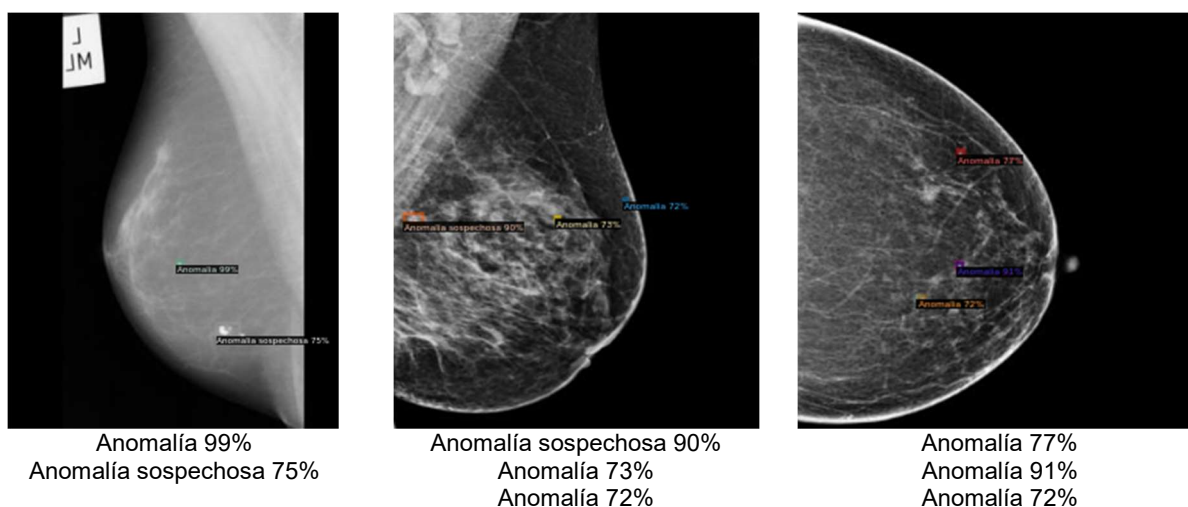


Figura 15. Continuación Resultados del modelo Detectron2.

VI. Análisis de resultados

En esta sección se realizará una comparación de los resultados obtenidos en la implementación de cada uno de los modelos propuestos, por lo cual a continuación se podrán visualizar las diferentes tablas en donde se encuentran los diferentes parámetros como son el tiempo de ejecución en la etapa de entrenamiento, Tasa de Aprendizaje, el valor del total de pérdidas, así como sus correspondientes métricas.

Tabla 13. Comparación de los tiempos de entrenamiento.

Modelos	Tiempo de entrenamiento
YOLOv4	4 hrs
EfficientDet D0	2 hrs
Detectron2	30 min

Tabla 14. Comparación de los valores de la Tasa de Aprendizaje de los modelos propuestos.

Modelos	Tasa de Aprendizaje
YOLOv4	0.001000
EfficientDet D0	0.0799058
Detectron2	0.000050

Tabla 15. Comparación del resultado del total de pérdidas.

Modelos	Total de pérdidas
EfficientDet D0	0.9935539
Detectron2	0.949

Tabla 16. Comparación del resultado de mAP.

Modelos	mAP
YOLOv4	10.29%
Roboflow 2.0 Object Detection (Fast)	2.1%

Tabla 17. Comparación de los valores obtenidos modelo Detectron2 en la etapa de entrenamiento y evaluación.

	AP	AP50	AP75
Etapas de Entrenamiento	2.997	5.565	2.698
Etapas de Evaluación	3.679	11.022	1.633

Tabla 18. Valores obtenidos para cada categoría en la etapa de entrenamiento del modelo Detectron2.

	Categoría	AP	Categoría	AP
Etapas de Entrenamiento	Anomalía	2.312	Anomalía sospechosa	3.682
Etapas de Evaluación	Anomalía	4.203	Anomalía sospechosa	3.155

Así bien, se realizó una comparación entre los resultados obtenidos de cada uno de los modelos con respecto al etiquetado de la mamografía del conjunto de evaluación.

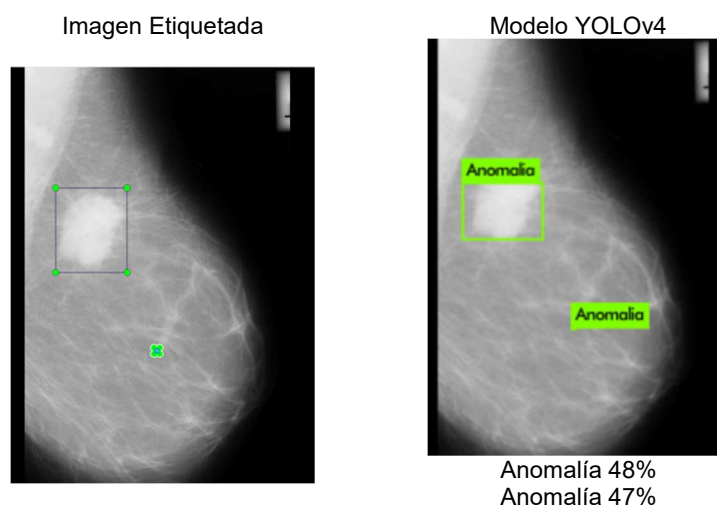


Figura 16. Comparación de la mamografía con el resultado obtenido en el Modelo YOLOv4.

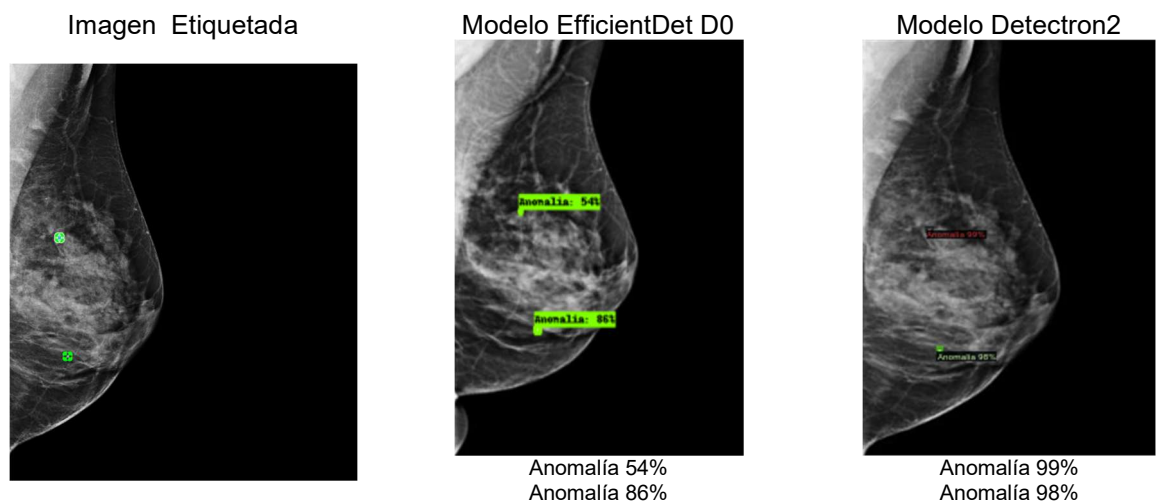


Figura 17. Comparación de la mamografía con el resultado obtenido del Modelo EfficientDet D0 y Modelo Detectron2.

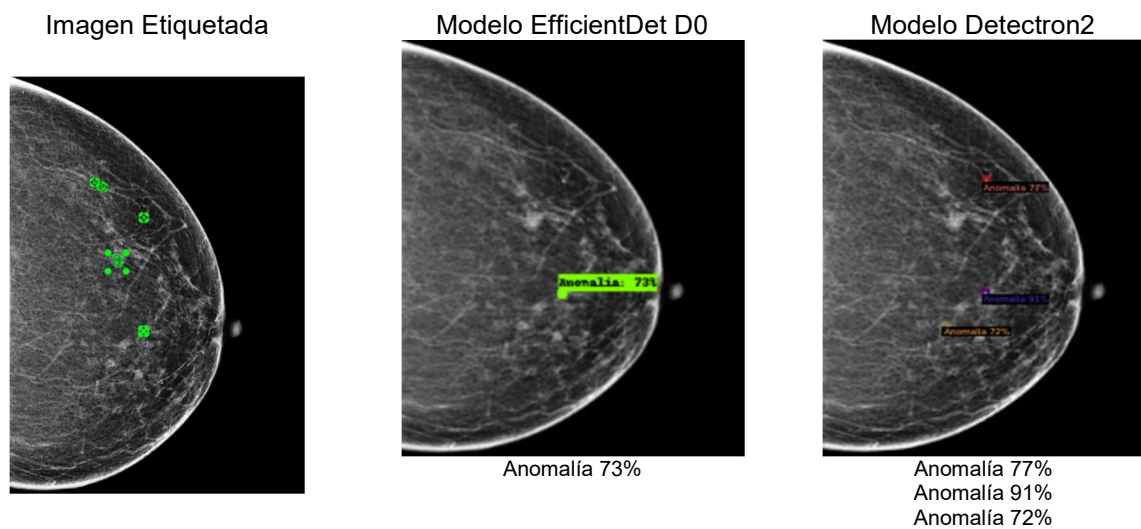


Figura 18. Comparación de la mamografía con el resultado obtenido del Modelo EfficientDet D0 y Modelo Detectron2.

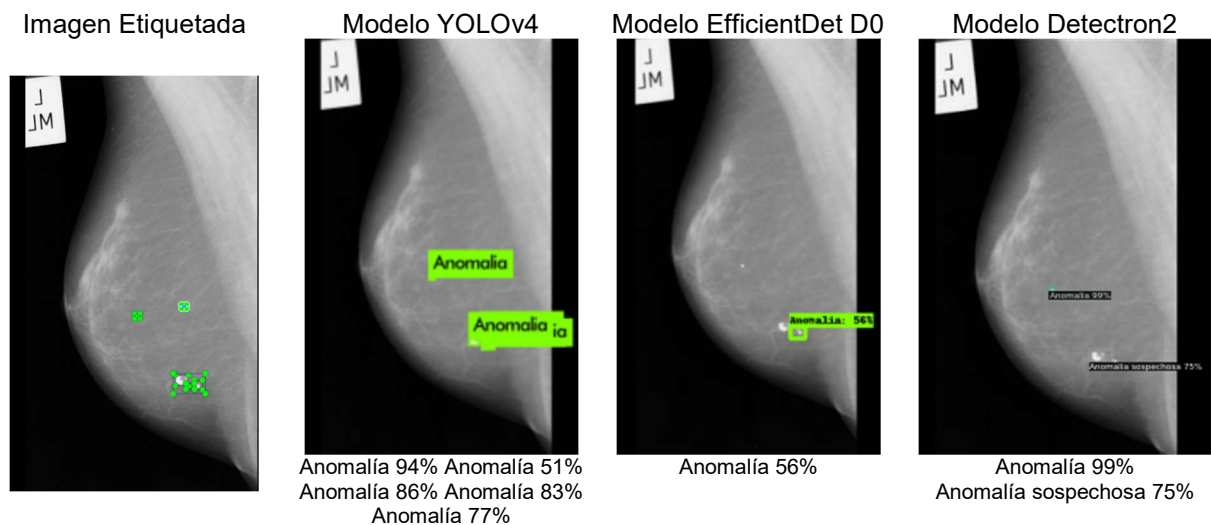


Figura 19. Comparación de la mamografía con el resultado obtenido en el Modelo YOLOv4, Modelo EfficientDet D0 y Modelo Detectron2.

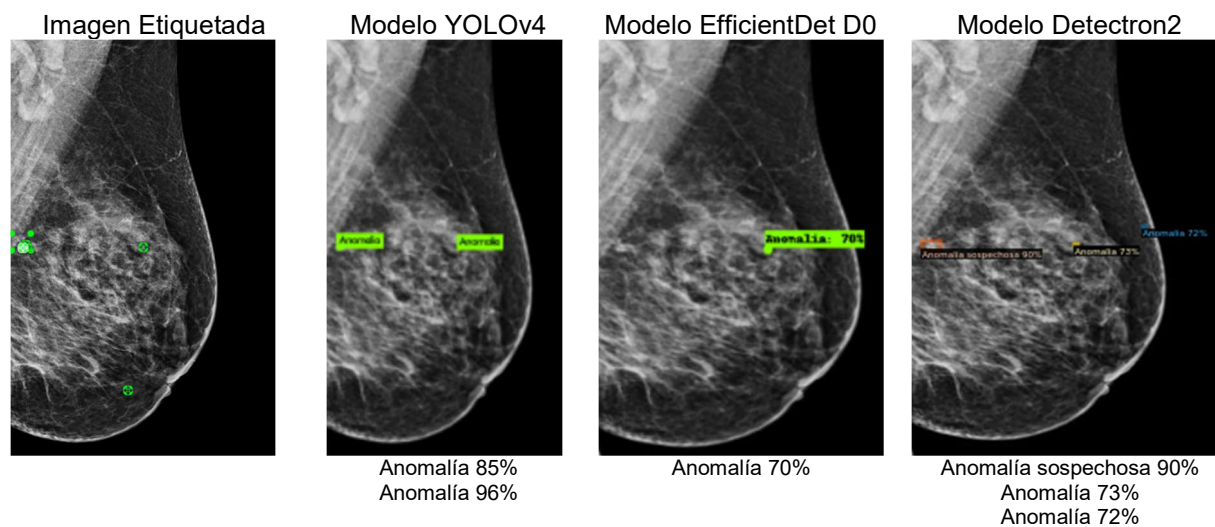


Figura 20. Comparación de la mamografía con el resultado obtenido en el Modelo YOLOv4, Modelo EfficientDet D0 y Modelo Detectron2.

VII. Conclusión

El Aprendizaje profundo ha logrado grandes avances en el campo de la investigación en detección, en donde como se pudo observar existe una gran variedad de modelos para diferentes aplicaciones. En este trabajo se han comparado tres arquitecturas enfocadas a la detección de objetos (YOLOv4, EfficientDet D0, Detectron2), en donde cada una de ellas propuso como resultado un porcentaje de confianza al detectar alguna anomalía en las mamografías del conjunto de evaluación.

Esta comparación tiene como fin el determinar cuál algoritmo podría ser una opción viable para la base de datos propuesta, en donde se pudo observar que los valores obtenidos en las métricas (mAP y AP) fueron considerablemente pequeñas. De igual manera, se pudo observar que el modelo Detectron2 tuvo un mayor porcentaje de confianza en la detección de anomalías, aunque en algunas de sus detecciones no fueron del todo certeras. Por otra parte, los modelos YOLOv4 y EfficientDet D0 tuvieron un porcentaje menor de confianza, así como también estos dos modelos pasaron por alto algunas anomalías en la etapa de evaluación.

Es posible que la razón de estos resultados se deba a que las anomalías son demasiado pequeñas e irregulares. Además de eso, si se cuenta con tejido denso en la mamografía, esto puede provocar que haya una mayor probabilidad de obtener falsos positivos en la detección de anomalías, como se pudo observar anteriormente. Por tal motivo para futuros trabajos será interesante probar otras estrategias de preprocesamiento con el fin de poder visualizar mejores resultados en la detección de las anomalías.

VIII. Referencias

- [1] BI-RADS (2016). Sistema de informes y registro de datos de estudios por imágenes de la mama. Journal. <http://booksmedicos.org>
- [2] Bosch Gaudenz. LabelImg para anotación de imágenes. Recuperado el 28 de Abril de 2022 de <https://viso.ai/computer-vision/labelimg-for-image-annotation/>
- [3] Camacho, E. (2021) Personalizando el detector YOLO (Detección de objetos con YOLO y Google-Colab). Recuperado el día 13 de Abril de 2022 de <https://www.youtube.com/watch?v=SBNlIdAcq5k>
- [4] Chevalier, M., Torres, R., 2010. Mamografía digital. Rev. Física Médica 11
- [5] Heywang-Koebrunner, S., Schreer, I., y Barter, S. (2014). Diagnostic Breast Imaging. Mammography, Sonography, Magnetic Resonance Imaging, and Interventional Procedures. Thieme <http://www.medlibrary.info/>
- [6] INEGI. (2021). Características de las Defunciones Registradas en México Durante 2020, PRELIMINAR. Comunicado de Prensa Num. 402/21, 1–5. https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2021/EstSociodemo/DefuncionesRegistadas2020_Pre_07.pdf
- [7] Loizidou, K., Skouroumouni, G., Pitris, C. et al. (2021) Breast Micro-Calcification Dataset. Recuperado el día 10 de Abril de 2022 de <https://zenodo.org/record/5036062#.Ym2HpOjMLre>
- [8] Maithani, M (2020) EfficientDet: Guía del modelo de detección de objetos de última generación. Recuperado el día 30 de Abril de 2022 de <https://analyticsindiamag.com/efficientdet/>
- [9] MIAS (2021). Mammographic Image Analysis Homepage. Recuperado el 31 de octubre del 2021 de <https://www.mammoimage.org/databases/>
- [10] Muñoz Sánchez, M. J. (2007). Los senos. Anatomía, evolución y cuidados dermatocósméticos. Revista Offarm. Elsevier. 26(11), 64-68.
- [11] Organización Mundial de la Salud (2021). Cáncer de mama. Recuperado el día 31 de octubre del 2021 de <https://www.who.int/es/news-room/fact-sheets/detail/breast-cancer>.
- [12] Palacios Bedia, M. (2021) Desarrollo de un sistema de posicionamiento 3D de personas mediante inteligencia artificial empleando imágenes 360°. Recuperado el 28 de Abril de 2022 de https://oa.upm.es/68669/1/TESIS_MASTER_MIGUEL_BEDIA_PALACIOS_2021.pdf
- [13] Pokhrel, S. (2020) Image Data Labelling and Annotation- Everything you need to know. Recuperado el día 30 de Abril de 2022 de <https://towardsdatascience.com/image-data-labelling-and-annotation-everything-you-need-to-know-86ede6c684b1>

- [14] Rizoli, A. (2022) Explicación de la precisión promedio (mAP): todo lo que necesita saber. Recuperado el día 28 de Abril de 2022 de [https://www.v7labs.com/blog/mean-average-precision#:~:text=Mean%20Average%20Precision\(mAP\)%20is%20a%20metric%20used%20to%20evaluate,values%20from%200%20to%201.](https://www.v7labs.com/blog/mean-average-precision#:~:text=Mean%20Average%20Precision(mAP)%20is%20a%20metric%20used%20to%20evaluate,values%20from%200%20to%201.)
- [15] Roboflow(2022) Roboflow. Recuperado el día 10 de Abril de 2022 de <https://docs.roboflow.com/>
- [16] Roboflow. EfficientDet-D0-D7. Recuperado el día 29 de Abril de 2022 de <https://models.roboflow.com/object-detection/efficientdet-d0-d7>
- [17] Roboflow. Detector2. Recuperado el día 29 de Abril de 2022 de <https://models.roboflow.com/object-detection/detectron2>
- [18] Selvi, R. (2014). Breast Diseases. Imaging and Clinical Management. Springer
<https://doi.org/10.1007/978-81-322-2077-0>
- [19] Sambasivarao, K.(2019) Non-maximum suppression(NMS). Recuperado el día 10 de Abril de <https://towardsdatascience.com/non-maximum-suppression-nms-93ce178e177c>
- [20] scikitlearn.Sklearn.metrics.average_precision_score.
Recuperado el día 30 de Abril de 2022 de https://scikit-learn.org/stable/modules/generated/sklearn.metrics.average_precision_score.html
- [21] TFC (2021) cómo funciona la detección de objetos de TensorFlow y dónde usarla. Recuperado el día 28 de Abril de 2022 de <https://www.tfcertification.com/blog/how-tensorflow-object-detection-works-where-to-use-it>