

REDES NEURONALES YOLO V2 Y YOLO V4 PARA LA DETECCIÓN TEMPRANA DE LESIONES DE CÁNCER DE MAMA

Luis F. Luque Vega^{1,2}, Salvador Castro Tapia¹, Héctor Alonso Guerrero Osuna³, Luis Octavio Solís Sánchez³, Rafael Reveles Martínez⁴, Jorge Pablo Vega Borrego^{1,3}, Roilhi Frajo Ibarra Hernández⁵

¹Tecnológico Nacional de México, Instituto Tecnológico Superior de Jerez, salvador.ct@jerez.tecnm.mx, jorge.vb@jerez.tecnm.mx.

²Departamento de Procesos Tecnológicos e Industriales, ITESO, luisluque@iteso.mx.

³Posgrado en Ingeniería y Tecnología Aplicada, Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, hectorguerrero@uaz.edu.mx, lsolis@uaz.edu.mx.

⁴Instituto Politécnico Nacional Campus Zacatecas, rvelesm@ipn.mx.

⁵Facultad de Ciencias, Universidad Autónoma de San Luis Potosí, roilhi.ibarra@uaslp.mx.

RESUMEN

El cáncer de mamá sobresa como la principal muerte por tumores malignos en mujeres a nivel mundial con un aproximado de 670,000 defunciones en el año 2022. En México se registraron 7,888 defunciones. La detección temprana al inspeccionar mamografías es la técnica por excelencia para prevenir tal enfermedad y actualmente, el auge de las Redes Neuronales Convolucionales, contribuyen como nuevas técnicas de diagnóstico. En este trabajo se utilizaron un total de 9,344 imágenes de lesiones de cáncer de mama obtenidas de las bases de datos MIAS e INbreast; y se entrenaron las CNN YOLOv2 y YOLOv4; obteniendo 84.43% y 91.21% de precisión promedio al clasificar masas y microcalcificaciones tanto benignas como malignas.

Palabras claves: Cáncer de mama, Diagnóstico temprano, Aprendizaje profundo.

ABSTRACT

Breast cancer stands out as the main death due to malignant tumors in women worldwide, with approximately 670,000 deaths in the year 2022. 7,888 deaths from breast cancer were registered in Mexico. Early detection through mammogram inspection is the technique par excellence to prevent such disease and currently, the rise of Convolutional Neural Networks, contribute as techniques for early diagnosis. In this research, a total of 9,344 images of breast cancer lesions obtained from the MIAS and INbreast databases were used; training the YOLOv2 and YOLOv4 CNNs; obtaining 84.43% and 91.21% of average accuracy when classifying benign and malignant masses and microcalcifications.

Keywords: Breast cancer, Early diagnosis, Deep learning.

1. INTRODUCCIÓN

El cáncer de mama es una enfermedad en la que células anormales en la mama crecen sin control, formando tumores que pueden propagarse y volverse mortales. Acorde a la Organización Naciones Unidas, el cáncer de mama es el más común a nivel mundial y afecta a millones de mujeres. En 2022, se diagnosticaron 2.3 millones de casos con 670,000 muertes por esta enfermedad en todo el mundo y fue el cáncer más común entre las mujeres de 157 de 185 países en 2022 [1] y entre 0.5% y 1% de los casos ocurren en hombres; sin embargo, es prevalente en todos los países, por ejemplo, en el año 2022 en México se registraron 23,790 casos nuevos entre la población de 20 años y más [2].

La Organización Mundial de la Salud, busca reducir la mortalidad por cáncer de mama en un 2.5% anual, para evitar 2.5 millones de muertes entre 2020 y 2040, enfocándose en Promoción de la salud y detección precoz, Diagnóstico oportuno y Tratamiento integral [1].

Actualmente, el Aprendizaje Profundo o Deep Learning (DL) ha logrado un gran avance en técnicas de detección y reconocimiento de objetos, con varias técnicas y topologías de redes neuronales que han sido aplicadas en la detección temprana del cáncer de mama a través del análisis de mamografías. En este trabajo, se presentan los resultados de los entrenamientos de dos arquitecturas de redes neuronales convolucionales (CNN) para la detección temprana de lesiones de cáncer de mama en mamografías al implementar detectores basados en YOLOV2 y YOLOV4 y comparar los resultados de las detecciones realizadas.

2. TEORÍA

La detección de cáncer de mama a través de la aplicación de métodos de inteligencia artificial ha sido ampliamente estudiada. Inicialmente las técnicas de machine learning para segmentación y clasificación se aplicaron en la detección de cáncer de mama con técnicas como árboles de decisión, máquinas de soporte vectorial (SVM), Regresiones lineales y Logísticas, K-Nearest Neighbors (KNN), Bosques Aleatorios y entre otros [3]. Con el desarrollo del aprendizaje profundo o Deep Learning, se introdujo a las CNN para cumplir los propósitos de segmentación, detección y clasificación de imágenes [4].

Posteriormente, nace la familia de CNN rápidas con los métodos R-CNN, Fast R-CNN y Faster R-CNN que operan sobre regiones tratadas como potenciales cajas limitantes y son el preámbulo de la técnica You Only Look Once (YOLO). Debido a su estructura de red simplificada y unificada, YOLO ofrece una alta velocidad que permite predecir a 45 fps y trata a la detección de objetos como un problema de regresión trabajando globalmente sobre los objetos de la imagen; aunque es deficiente al detectar objetos pequeños [5].

Una versión mejorada es YOLO V2, que utiliza cajas de anclaje para detectar objetos en una imagen y predecir 3 atributos para cada referencia, la intersección sobre la unión (IoU) que predice el puntaje de objetividad para cada caja de anclaje, su desplazamiento, y la probabilidad de clase [6]. La arquitectura de YOLOv2 se muestra en la Figura 1.

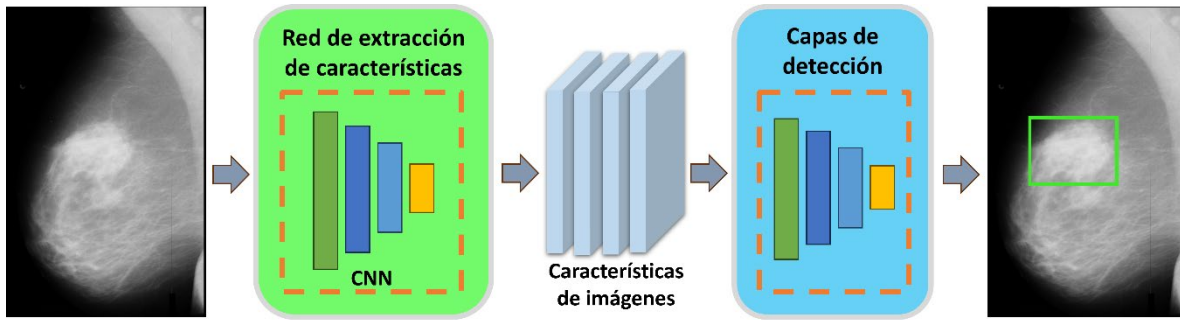


Figura 1. Arquitectura de YOLOv2

Posteriormente surgió YOLOv3, que es un algoritmo diseñado para detectar y clasificar múltiples objetos con una velocidad y precisión excepcionales. Mediante una red piramidal de características puede detectar objetos a tres escalas diferentes tanto pequeños como grandes, usando tres tamaños de núcleos: 13x13, 26x26 y 52x52. Su estructura base Darknet-53, equilibra la precisión y la eficiencia computacional mediante 53 capas convolucionales y en lugar de utilizar softmax para la clasificación [7], YOLOv3 emplea clasificadores logísticos independientes, lo que le permite gestionar etiquetas superpuestas [8].

En la versión YOLOv4, la red de detección de objetos se compone por tres partes. La columna vertebral puede ser una CNN pre entrenada como VGG16, CSPDarkNet53 o ImageNet y actúa como la red de extracción de mapas de características de las imágenes, para pasarlas al Cuello que está compuesto por un módulo de Agrupación Piramidal Espacial (SPP) y una Red de Agregación de Rutas (PAN). En la cabeza se procesan las características agregadas y predice las cajas limitantes, los puntajes de objetividad y los puntajes de clasificación [9]. La arquitectura de YOLOv4 se muestra en la Figura 2.

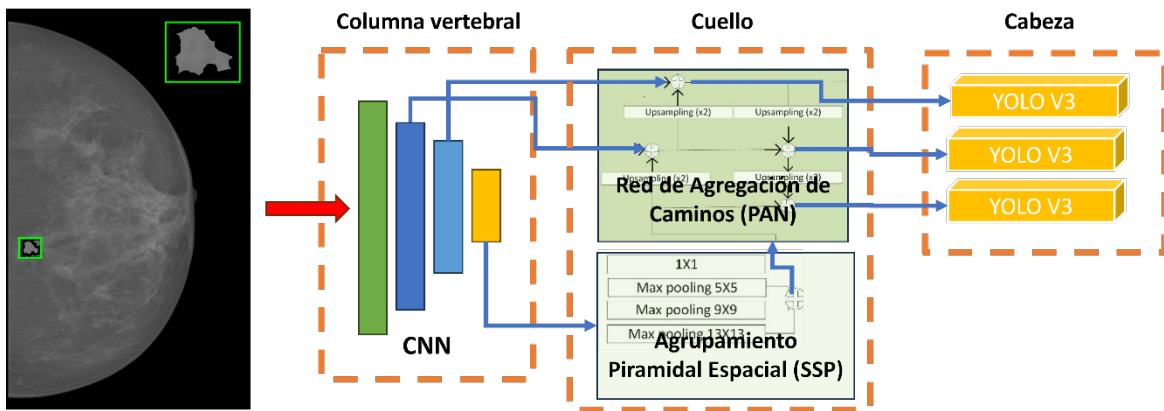


Figura 2. Arquitectura de YOLOv4

En cuanto a la detección temprana de cáncer de mama se han realizados varias investigaciones, dentro de las cuáles, en [10], se propone un sistema Diagnóstico Asistido por Computadora (CAD) que utiliza DL para mejorar la detección y clasificación de lesiones mamarias en mamografías digitales de rayos X mediante YOLO, con puntuaciones F1 del 99,28 % para el conjunto DDSM y 98,02 % para INbreast [11], mejorando debido al

clasificador InceptionResNet-V2 modificado, con altas precisiones del 97,50 % en DDSM y 95,32 % en INbreast con una velocidad de procesamiento de 0,025 segundos por mamografía.

A su vez, en marzo de 2021, un estudio evaluó diferentes versiones del modelo YOLO (v1, v2 y v3) para la detección y clasificación de masas mamarias en mamografías. Usando el conjunto de datos FFDM e INbreast, se determinó que YOLO-v3 ofrecía el mejor rendimiento, detectando el 89,4 % de las masas en INbreast con una precisión promedio del 94,2 % para masas benignas y 84,6 % para malignas y a su vez, se mejoró la precisión al sustituir la red de clasificación de YOLO por ResNet (91,0 %) e InceptionV3 (95,5 %) [12].

Otra investigación presentó un enfoque en tres fases aplicado a la base de datos INbreast para mejorar la detección de masas en cortes superpuestos pequeños, empleando YOLO-V4 para identificar masas tanto en las imágenes completas como en los cortes, obteniendo un 95 % de precisión para clasificar masas como benignas o malignas, superando así a métodos recientes en efectividad diagnóstica [13].

3. PARTE EXPERIMENTAL

3.1 Bases de datos de mamografías y aumentación de datos.

En este trabajo se utilizaron las bases de datos INBreast [11] y Mammographic Image Analysis Society MIAS [14]. INbreast contiene 115 casos con 410 imágenes DICOM y 4 mamografías por caso en sus vistas craneocaudales (CC) y medio lateral oblicuas (MLO) [11]. Originalmente se anotaron 539 lesiones; solo reportan grupos de calcificaciones y no se listan todas las masas; por lo tanto, procedimos a registrar microcalcificaciones por separado y en grupos más pequeños, además de masas más visibles para alimentar las redes neuronales para el entrenamiento y las pruebas.

La base de datos MIAS contiene 161 casos con 322 imágenes digitalizadas en formato PGM con 204 mamografías normales, 118 con alguna lesión y de estas, 66 son benignas y 52 malignas, además de que se provee anotaciones con el tipo de tejido, la clase de la anomalía, su severidad, las coordenadas de la lesión y el radio [14].

Con la finalidad de obtener un mayor número de anotaciones, se procedió a realizar una aumentación de datos, al aplicar la técnica de Ecuilización Adaptiva de Histograma por Limitación de Contraste (CLAHE) sobre las mamografías y sus lesiones o regiones de interés (ROI); en la Figura 3, se observa la basta mejora sobre la mamografía al aplicar la técnica CLAHE [15], siendo que se pueden observar de mejor manera las estructuras de la mama.

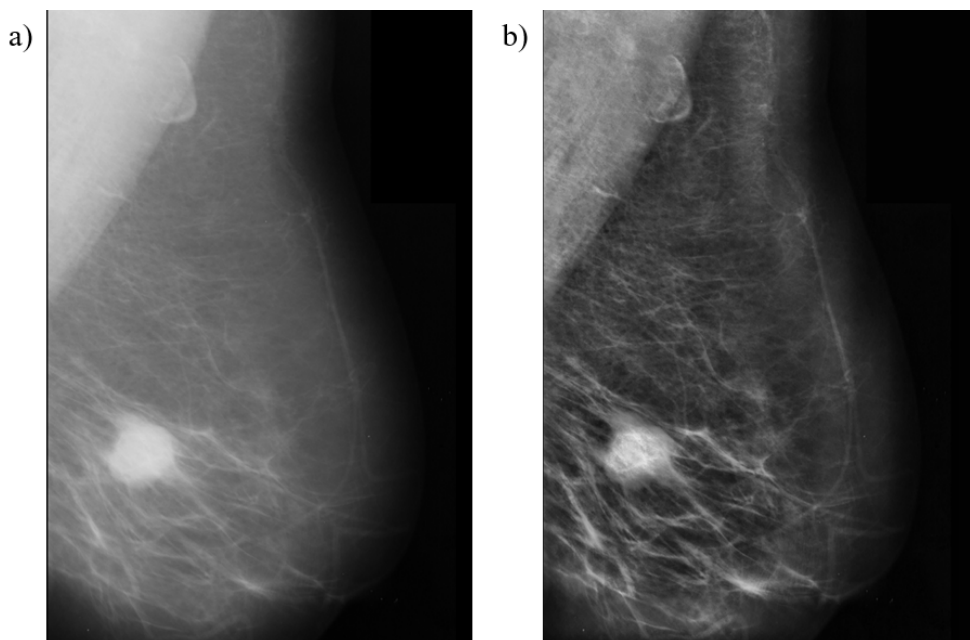


Figura 3. Mejora de la mamografía CLAHE, a) Mamografía original, b) Mamografías tras aplicar CLAHE

Continuando con la aumentación de datos, se realizaron rotaciones de 90° , 180° y 270° de las imágenes, además de quitar el fondo de las de las lesiones y finalmente realizando una rotación horizontal. Tales aumentaciones se pueden observar en las Figura 4.

Con toda esta aumentación de datos, se obtuvieron un total de 9,344 anotaciones de lesiones de mama en las mamografías de ambas bases de datos, y se resumen en la Tabla 1.

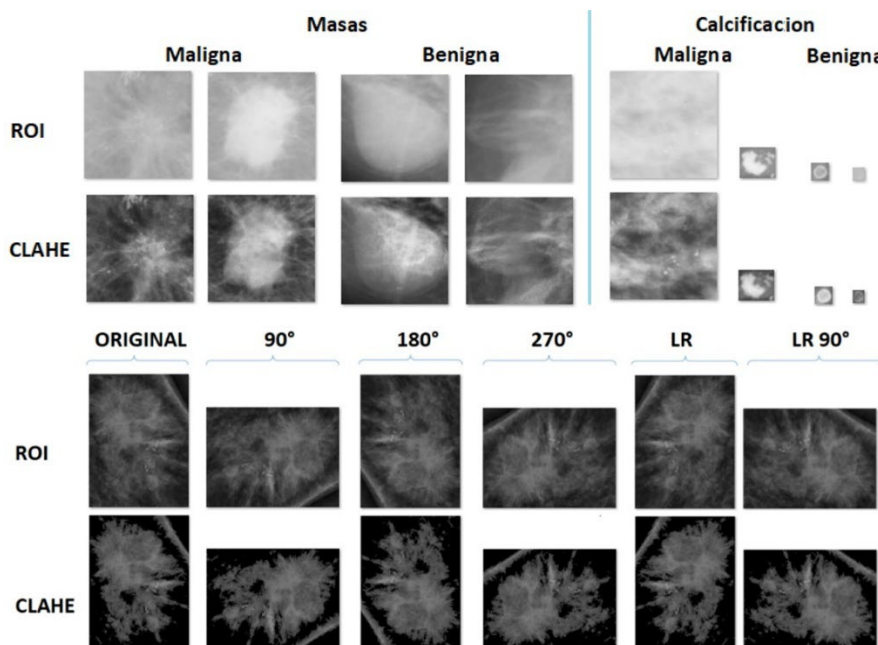


Figura 4. Imagen resultante al aplicar CLAHE y sus rotaciones

Tabla 1. Resultados de la aumentación de datos

ROI	Anotaciones Originales	Anotaciones Extras	CLAHE	90°	180°	270°	Sin fondo	Rotación horizontal	Totales
Calcificación Benigna	240	392	407	803	494	0	0	0	2336
Calcificación Maligna	86	70	73	146	146	146	584	1085	2336
Masa Benigna	91	75	88	217	217	217	868	563	2336
Masa Maligna	122	102	153	306	306	306	0	1041	2336
Totales	539	639	721	1472	1163	669	1452	2689	9344

3.3 Aprendizaje Por Transferencia y Entrenamiento

Utilizando aprendizaje por transferencia, se puede rentrenar una red funcional y probada ampliamente; en este caso, para la detección temprana de lesiones de cáncer de mama. En esta investigación se utilizan las redes YOLOv2 Y YOLOv4, no se utiliza YOLOv3 debido a que la cuarta versión utiliza esta red en su tercer sección o cabeza.

Para la red YOLOv2 se utiliza ResNet-50 como red de extracción de características, y la red de detección se compone de capas de convolución y capas específicas para el modelo V2 que son incrustadas en la capa 141 "activation_40_relu". Como salida se obtiene un mapa de características con un submuestreo por un factor de 16, balanceando la resolución espacial y la intensidad de las características extraídas. Se utilizó optimizador ADAM, un minilote de 4, una tasa inicial de aprendizaje de 1×10^{-3} y un máximo de 10 épocas. Al utilizar la topología de YOLOv4, se establecen las clases, y las cajas de anclaje estimadas, usando tiny-yolov4-coco como red de extracción de características. Se utiliza un optimizador Adam, con un factor de decaimiento del gradiente de 0.9, una tasa de aprendizaje 1×10^{-3} y un tamaño de minilote igual a 4. También se utiliza una pérdida de validación baja para obtener un mejor detector, con una frecuencia de validación sobre los datos igual a 1000, junto con una frecuencia de informe (verbose) igual a 20 para cada iteración; durante 80 épocas. Ambos detectores fueron entrenados en una laptop con procesador AMD Ryzen 5 5500U con 12 GB en RAM y una GPU AMD Radeon (TM), lo cuál explica los tamaños chicos de mini lotes ya que estos recursos son bajos y se emplea mucho tiempo para el entrenamiento.

Al evaluar las métricas de eficiencia de las redes, es importante definir los casos Verdaderos Positivos (TP) y Negativos (TN), los casos Falsos Positivos (FP), y Negativos (FN), que se definen con las ecuaciones (1) a (4), y con ellos podemos obtener la precisión promedio (5) y el recall (6) o recuperación. Con estas métricas se puede construir la matriz de confusión para los detectores basados en YOLOv2 y YOLOv4.

$$TP_i = a_{ii} \quad (1)$$

$$FP_i = \sum_{l=1}^n a_{li} - TP_i \quad (2)$$

$$FN_i = \sum_{l=1}^n a_{il} - TP_i \quad (3)$$

$$TN_i = \sum_{l=1}^n \sum_{k=1}^n a_{lk} - TP_i - FP_i - FN_i \tag{4}$$

$$Precisión = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \tag{5}$$

$$Recall = \frac{TP_i}{TP_i + FN_i} \tag{6}$$

Utilizando las 9344 imágenes de lesiones de cáncer de mama obtenidas de las bases de datos MIAS e INbreast; se entrenaron las redes YOLOV2 y YOLOV4 con el 70% de las imágenes correspondientes a 6540; y el 30% restante se utilizó para pruebas y validación, siendo 2,804 lesiones con sus mamografías. Las métricas de rendimiento se pueden apreciar en las matrices de confusión de la Figura 5, obteniendo un 84.43% de precisión promedio en la red YOLOv2 al clasificar masas y microcalcificaciones tanto benignas como malignas y un 91.21% para YOLOv4; demostrando una gran capacidad de detección para ser implementadas dentro de sistemas CAD.

En la Figura 6, se puede observar la gráfica de la precisión promedio en la red YOLOV4, que mostró mejores resultados; se utilizó un valor de umbral de traslape de cajas de anclaje muy bajo con un valor de 0.01; si este valor se aumenta, la precisión promedio disminuye.

Clase predecida	Clase objetivo				Precisión	Recall
	Calcificacion Benigna	Calcificacion Maligna	Masa Benigna	Masa Maligna		
Calcificacion Benigna	621 22.1%	73 2.6%	5 0.2%	2 0.1%	88.6% 11.4%	
Calcificacion Maligna	71 2.5%	591 21.1%	21 0.7%	18 0.6%	84.3% 15.7%	
Masa Benigna	18 0.6%	10 0.4%	565 20.1%	108 3.9%	80.6% 19.4%	
Masa Maligna	5 0.2%	3 0.1%	103 3.7%	590 21.0%	84.2% 15.8%	
	86.9% 13.1%	87.3% 12.7%	81.4% 18.6%	82.2% 17.8%	84.4% 15.6%	

Clase predecida	Clase objetivo				Precisión	Recall
	Calcificacion Benigna	Calcificacion Maligna	Masa Benigna	Masa Maligna		
Calcificacion Benigna	661 23.6%	32 1.1%	8 0.3%	0 0.0%	94.3% 5.7%	
Calcificacion Maligna	54 1.9%	637 22.7%	9 0.3%	1 0.0%	90.9% 9.1%	
Masa Benigna	1 0.0%	2 0.1%	655 23.4%	43 1.5%	93.4% 6.6%	
Masa Maligna	0 0.0%	3 0.1%	95 3.4%	603 21.5%	86.0% 14.0%	
	92.3% 7.7%	94.5% 5.5%	85.4% 14.6%	93.2% 6.8%	91.2% 8.8%	

Figura 5. Matrices de confusión: a) YOLOv2 con precisión promedio de 84.4% y b) YOLOv4 con precisión promedio de 91.2%

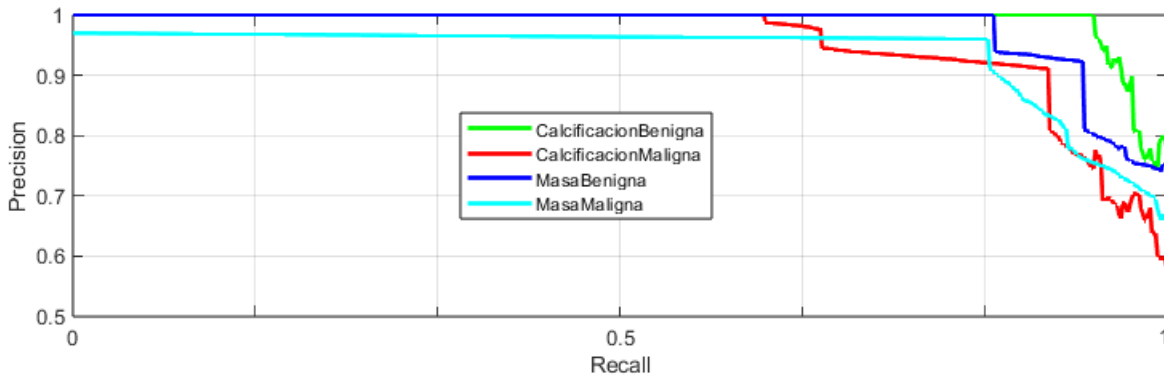


Figura 6. Precisión promedio de la red YOLOV4

4. CONCLUSIONES

La evolución de las CNN ha generado un gran avance en el campo de la imagenología médica y sus altos porcentajes de detección pueden ser utilizados en la detección temprana del cáncer de mama.

Las métricas de desempeño en cuanto a precisión promedio de la red YOLOV4 (91.21%) superan a las de YOLOV2 (84.43%), con una diferencia del 6.78, siendo un algoritmo más robusto para la detección precisa de lesiones de cáncer de mama gracias a su detección multiescala y a las capas de clasificación de YOLOv3.

En la matriz de confusión del detector basado en YOLOv4, se observa que la detección de masas benignas es donde ocurren mayores errores de detección con una precisión del 85.4%, confundiendo con masas malignas, mientras que la mayor precisión del detector es al clasificar las calcificaciones malignas con una precisión del 94.5%.

El valor umbral de traslape de las cajas de anclaje genera una alta precisión promedio en la detección de las lesiones de cáncer de mama y al aumentar este umbral la precisión disminuye.

5. REFERENCIAS

- [1] Organización Mundial de la Salud, “Cáncer de mama,” OMS, 2024. [En línea]. Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/breast-cancer>
- [2] Instituto Nacional de Estadística y Geografía, “Estadística a propósito del Día Internacional de la lucha contra el Cáncer de Mama (19 de octubre),” INEGI, 2023. [En línea]. Disponible en: https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2023/EAP_CMAMA23.pdf.
- [3] A. Khalid, A. Mehmood, A. Alabrah, B. F. Alkhamees, F. Amin, H. AlSalman, y G. S. Choi, “Breast cancer detection and prevention using machine learning,” **Diagnostics**, vol. 13, 3113 (2023) [[doi:10.3390/diagnostics13193113](https://doi.org/10.3390/diagnostics13193113)].
- [4] S. Castro-Tapia, C. L. Castañeda-Miranda, C. A. Olvera-Olvera, H. A. Guerrero-Osuna, J. M. Ortiz-Rodríguez, M. d. R. Martínez-Blanco, G. Díaz-Florez, J. D. Mendiola-Santibañez, y L. O. Solís-Sánchez, “Classification of breast cancer in mammograms with deep learning adding a fifth class,” **Appl. Sci.**, vol. 11, 11398 (2021) [[doi:10.3390/app112311398](https://doi.org/10.3390/app112311398)].
- [5] J. Redmon, S. Divvala, R. Girshick, y A. Farhadi, “You only look once: Unified, real-time object detection,” en **Proc. IEEE Conf. Comput. Vis. Pattern Recognit.** (CVPR), (2016). [En línea]. Disponible en: <http://pjreddie.com/yolo/>.
- [6] J. Redmon y A. Farhadi, “YOLO9000: Better, Faster, Stronger,” University of Washington, Allen Institute for AI, [Online]. Disponible: <http://pjreddie.com/yolo9000/>.
- [7] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York (2006), <https://link.springer.com/book/9780387310732>.

- [8] [1] K. J. Oguine, O. C. Oguine, and H. I. Bisallah, "YOLO v3: Visual and Real-Time Object Detection Model for Smart Surveillance Systems(3s)," Proc. SPIE 56637, Information Technology for Education and Development (ITED), 10051233 (2022). <https://doi.org/10.1109/ITED56637.2022.10051233>
- [9] Bochkovskiy, A., Wang, C. Y., and Liao, H. Y. M., "YOLOv4: Optimal Speed and Accuracy of Object Detection", arXiv, (2020). <https://arxiv.org/pdf/2004.10934>.
- [10] M. A. Al-antari, S.-M. Han, y T.-S. Kim, "Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms," *Comput. Methods Programs Biomed.*, vol. 196, 105584 (2020) [[doi:10.1016/j.cmpb.2020.105584](https://doi.org/10.1016/j.cmpb.2020.105584)].
- [11] Moreira, I. C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M. J., and Cardoso, J. S., "INbreast: Toward a Full-field Digital Mammographic Database", Academic Radiology, 19(2), 93-97, (2024). <https://www.ijarsct.co.in/Paper17515.pdf>
- [12] A. M. Abdel-Zaher y A. M. Eldeib, "Breast cancer classification using deep belief networks," *Comput. Methods Programs Biomed.*, vol. 133, pp. 105–113 (2020) [[doi:10.1016/j.cmpb.2020.105823](https://doi.org/10.1016/j.cmpb.2020.105823)].
- [13] G. Hamed, M. Marey, S. E. Amin, y M. F. Tolba, "Automated breast cancer detection and classification in full field digital mammograms using two full and cropped detection paths approach," *IEEE Access* (2021) [[doi:10.1109/ACCESS.2021.3105924](https://doi.org/10.1109/ACCESS.2021.3105924)].
- [14] Suckling, J., Parker, J., Dance, D., Ashley, S., Hutt, I., Boggis, C., Ricketts, I., Stamatakis E., Cerneaz, N., et al, "Mammographic Image Analysis Society (MIAS)", Database v1.21 [Dataset], (2015). <https://www.repository.cam.ac.uk/items/b6a97f0c-3b9b-40ad-8f183d121eef1459>.
- [15] K. Zuiderveld, "Contrast limited adaptive histogram equalization," en *Graphic Gems IV*, Academic Press Professional, San Diego, CA (1994), pp. 474–485.