

SEGMENTACIÓN AUTOMATIZADA DE CÁNCER DE MAMA EN IMÁGENES DE ULTRASONIDOS MEDIANTE REDES NEURONALES CONVOLUCIONALES

AUTOMATED BREAST CANCER SEGMENTATION IN ULTRASOUND IMAGES USING CONVOLUTIONAL NEURAL NETWORKS



¹Erick Fabián Donado Fonseca

Universidad Simón Bolívar

Cómo citar: Donado Fonseca, E. F. (2026). Segmentación Automatizada de Cáncer de Mama en Imágenes de Ultrasonidos Mediante Redes Neuronales Convolucionales. *Publicaciones E Investigación*, 20(1).
<https://doi.org/10.22490/25394088.10453>

Recibido 16 Septiembre 2025, Aprobado 13 Diciembre 2025

RESUMEN

Esta investigación se centra en el desarrollo de un sistema automático para la detección y segmentación de tumores en imágenes de ultrasonido mamario mediante técnicas avanzadas de aprendizaje profundo. Se diseñó un programa autoaprendiente basado en redes neuronales convolucionales (CNN) con arquitectura tipo U-Net, capaz de entrenarse con conjuntos de datos médicos etiquetados para identificar y resaltar posibles regiones cancerosas. La metodología incluyó el preprocesamiento de las imágenes de ultrasonido mediante escalado a un tamaño uniforme de 256×256 píxeles, la división del conjunto de datos en entrenamiento y prueba (80/20), y la normalización de los datos. El entrenamiento del modelo incorporó funciones de activación, regularización mediante *dropout* y normalización por lotes para mejorar el rendimiento. El sistema mejora su precisión de manera autónoma al adaptarse a nuevos patrones y, al mismo tiempo, ofrece retroalimentación visual mediante la generación de máscaras de segmentación que marcan áreas sospechosas en las imágenes de ultrasonido. El desempeño se evaluó a través de la visualización de predicciones sobre imágenes del conjunto de datos usado para el entrenamiento. Los resultados demuestran la capacidad del modelo para detectar y segmentar regiones tumorales con alta precisión, lo que reduce falsos positivos y apoya el diagnóstico oportuno. Este enfoque evidencia el potencial clínico de las CNN para mejorar la exactitud diagnóstica, la eficiencia y, en última instancia, los resultados en pacientes con cáncer de mama.

Palabras clave: segmentación de imágenes; cáncer de mama; ultrasonido; aprendizaje profundo; U-Net; redes neuronales convolucionales.

¹ erick.donado@unisimon.edu.co / ORCID: 0009-0008-5752-3255

ABSTRACT

This research focuses on the development of an automatic system for tumor detection and segmentation in breast ultrasound images using advanced deep learning techniques. A self-learning program was designed based on convolutional neural networks with a U-Net architecture, capable of training on labeled medical datasets to identify and highlight potential cancerous regions. The methodology included preprocessing ultrasound images by resizing them to a uniform 256×256 pixels, splitting the dataset into training and testing sets (80/20), and normalizing the data. Model training incorporated activation functions, dropout regularization, and batch normalization to enhance performance. The system autonomously improves its accuracy by adapting to new patterns while providing visual feedback through segmentation masks that highlight suspicious areas in ultrasound images. Performance was evaluated by visualizing predictions on images from the training dataset. The results demonstrate the model's ability to detect and segment tumor regions with high accuracy, reducing false positives and supporting timely diagnosis. This approach highlights the clinical potential of convolutional neural networks to improve diagnostic accuracy, efficiency, and ultimately patient outcomes in breast cancer care.

Keywords: image segmentation; breast cancer; ultrasound; deep learning; U-Net; convolutional neural networks.



1. INTRODUCCIÓN

El cáncer de mama representa una de las principales causas de mortalidad en mujeres en el mundo, y la detección temprana y precisa es fundamental (Minaee et al., 2021) para mejorar el pronóstico y reducir la mortalidad. Entre las técnicas de diagnóstico más utilizadas se encuentra el ultrasonido, debido a su carácter no invasivo, de bajo costo y amplia disponibilidad. Sin embargo, la interpretación de estas imágenes suele ser compleja, la calidad puede variar sustancialmente, los contornos de los tumores pueden estar mal definidos y el resultado depende en gran medida de la experiencia del radiólogo.

En este contexto, los avances en inteligencia artificial (IA), visión por computador y aprendizaje profundo ofrecen nuevas oportunidades para desarrollar herramientas que apoyan el diagnóstico. En particular, las CNN y arquitecturas de segmentación como U-Net han demostrado un gran potencial (Ronneberger

et al., 2015; Çiçek et al., 2016; Zhou et al., 2018) para identificar automáticamente regiones de interés en imágenes médicas (Isensee et al., 2021). Estas pueden entrenarse con grandes volúmenes de imágenes de ultrasonido etiquetadas, con ellas aprenden patrones asociados con lesiones malignas y así destacan visualmente áreas sospechosas.

La segmentación automática no solo permite reducir la carga de trabajo manual y mejorar la consistencia en los diagnósticos, sino que también facilita una cuantificación más precisa de características relevantes, como el tamaño y la forma del tumor. Asimismo, la incorporación de salidas visuales que resalten regiones sospechosas proporciona un apoyo directo al especialista, disminuye la carga cognitiva, minimiza el error humano y aumenta la precisión diagnóstica.

Este artículo describe la implementación y evaluación de un modelo basado en U-Net para la segmentación automática de tumores en imágenes de ultrasonido mamario. Se detalla el proceso de desarrollo, entrenamiento y validación del sistema, así como un análisis de las fortalezas, limitaciones, implicaciones clínicas y posibles mejoras futuras para la aplicación en entornos reales de diagnóstico médico.

2. MATERIALES Y MÉTODO

La investigación se desarrolla mediante un enfoque cuasi experimental que evalúa los conocimientos de geometría básica, comparando los resultados de aprendizaje de dos grupos de estudiantes:

2.1. Herramientas y *frameworks*

El objetivo es desarrollar una aplicación capaz de detectar tumores en imágenes de ultrasonido mamario en mujeres entre los 25 y 75 años, utilizando Python y diversas bibliotecas y frameworks relacionados con redes neuronales y ciencia de datos, como Keras, NumPy, Pandas, Scikit-learn y OpenCV. La metodología consistió en la recolección y preparación de un conjunto de datos, que fue procesado y etiquetado para entrenar un modelo de aprendizaje convolucional. Posteriormente, se evaluó la precisión y eficiencia del modelo mediante métricas cuantitativas y pruebas visuales, comparando los resultados de entrenamiento y validación con las máscaras de segmentación de las imágenes originales.

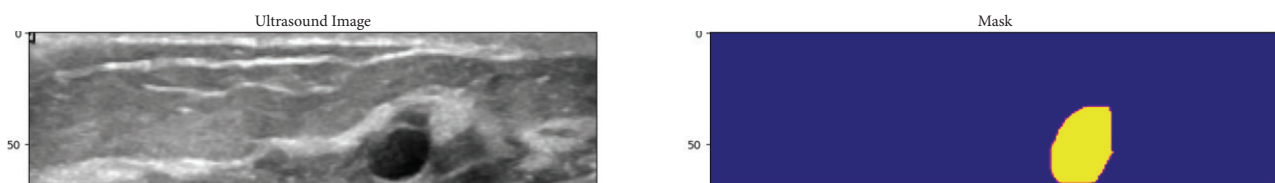
Los requerimientos del proyecto se centraron en dos aspectos fundamentales: la lógica de

programación implementada en Python como herramienta principal de desarrollo y la disponibilidad de datos para el entrenamiento del modelo. Python fue seleccionado por su versatilidad, diseño orientado a objetos y extenso ecosistema de bibliotecas para IA y aprendizaje automático. Keras facilitó la construcción y el entrenamiento de modelos de aprendizaje profundo mediante una interfaz sencilla y eficiente, mientras que Pandas y NumPy permitieron estructuras poderosas para la manipulación de datos y el cómputo científico. OpenCV se utilizó para tareas de visión por computador. Kaggle sirvió tanto como fuente del conjunto de datos de ultrasonido mamario como plataforma de colaboración y validación de modelos.

2.2. Conjunto de datos y preprocesamiento de imágenes

El conjunto de datos, recolectado de Kaggle en 2018 (Al-Dhabyani et al., 2020), estuvo compuesto por 780 imágenes de ultrasonido de 600 pacientes femeninas, categorizadas en normales, benignas y malignas. Las imágenes, con un tamaño promedio de 500×500 píxeles y en formato PNG, se acompañaban de máscaras de segmentación de referencia. Para el entrenamiento del modelo se redimensionaron tanto las imágenes como las máscaras a 256×256 píxeles; luego, se normalizaron y dividieron en conjuntos de entrenamiento y validación en una proporción 80/20. Durante el preprocesamiento, se aplicaron las máscaras para definir las salidas de segmentación esperadas, que sirvieron como referencia para la evaluación de la precisión de las predicciones.

Figura 1. Ejemplo de imagen del conjunto de datos y la implementación de máscaras.



Fuente: elaboración propia

3. DESARROLLO

3.1. Arquitectura y entrenamiento del modelo

La arquitectura utilizada fue una red neuronal convolucional U-Net estándar (Ronneberger et al., 2015), compuesta por bloques de convoluciones y capas de *max pooling* para el proceso de codificación, así como capas de *upsampling* o convolución transpuesta para la decodificación. Para mejorar la capacidad de generalización y reducir el sobreajuste, se incorporaron *dropout* y normalización por lotes en capas seleccionadas. El entrenamiento empleó la función de pérdida *binary cross-entropy*, junto con métricas como precisión, *recall*, intersección sobre unión (IoU) y coeficiente de Dice (Ming y Xiao, 2024). Se aplicaron mecanismos de *early stopping* para evitar el sobreajuste, deteniendo el entrenamiento cuando la pérdida de validación dejaba de mejorar.

El análisis del conjunto de datos puso de manifiesto la necesidad de abordar el desbalance de clases, analizar la distribución de valores de píxeles y aplicar técnicas de aumento de datos (Balaji, 2023), como volteo, recorte y rotación, para incrementar la diversidad y robustez del modelo. Asimismo, a lo largo del entrenamiento, se evidenció la importancia de optimizar hiperparámetros, ajustando tamaños de filtros, tasas de *dropout*, configuraciones de kernel y tasas de aprendizaje. También se emplearon técnicas adicionales de regularización, como programaciones de la tasa de aprendizaje y *checkpoints*, para reforzar la fiabilidad del modelo.

En esta etapa se consideraron propuestas de arquitecturas alternativas surgidas en los últimos años. Ejemplos de ello son UNet++ (Zhou et al., 2018), 3D U-Net para volúmenes médicos (Çiçek et al., 2016) y enfoques híbridos con transformadores, como Swin-Unet (Cao et al., 2023) y *Vision Transformers* aplicados al cáncer de mama (Borah et al., 2022), que ofrecen puntos de comparación para justificar la elección de la U-Net estándar como punto de partida.

La sinergia entre Python como base de programación, Keras como *framework* de modelado y Kaggle como fuente de datos permitió un flujo de trabajo integral que abarcó desde la adquisición y el preprocesamiento hasta el entrenamiento y la validación del modelo.

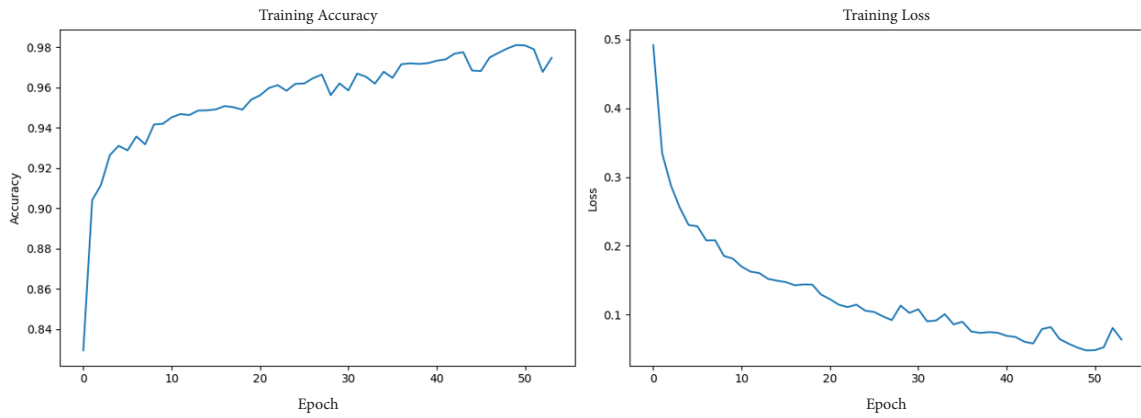
3.2. Evaluación y análisis de resultados

El análisis del proceso de entrenamiento evidenció un alto grado de exactitud al comparar las predicciones con las máscaras de referencia. El modelo alcanzó una precisión de entrenamiento de aproximadamente 98.15 %, con una pérdida final de 0.0466, lo que constituye un resultado notable para este tipo de algoritmos. La curva de precisión mostró una mejora constante a lo largo de las épocas, con fluctuaciones mínimas y alcanzando una meseta alrededor de la época 40. Este valor debe interpretarse junto con métricas específicas como IoU y Dice, que reflejan mejor el desempeño en regiones tumorales, ya que solo se considera la exactitud global (*pixel accuracy*). Este comportamiento sugirió riesgo de sobreajuste, de modo que el entrenamiento se detuvo en el punto óptimo para conservar un equilibrio adecuado entre aprendizaje y generalización.

El proceso de evaluación combinó análisis cuantitativos y visuales. Por un lado, las métricas estadísticas confirmaron que el modelo alcanzó una alta precisión, *recall* y exactitud general. Por otro lado, las comparaciones visuales entre las máscaras de referencia y las predicciones demostraron la capacidad del modelo para detectar regiones tumorales incluso en condiciones complejas, como bordes difusos, bajo contraste o formas irregulares.

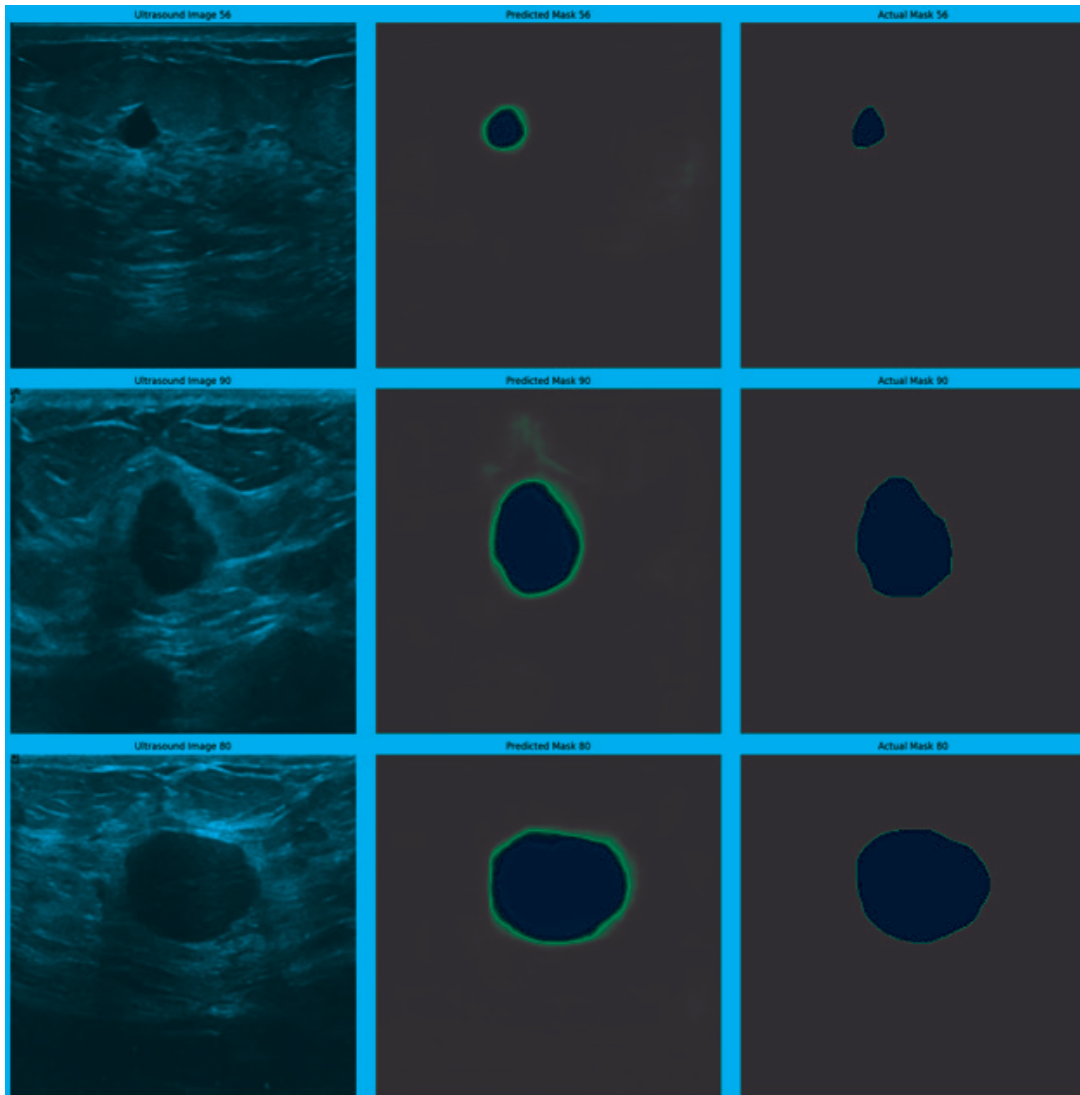
Los resultados mostraron una correspondencia adecuada entre las máscaras predichas y las reales, con visualizaciones que confirman la efectividad del modelo en los casos estudiados. En términos generales, las áreas tumorales se identificaron con bordes precisos, aunque en algunos casos se evidenció sobresegmentación o subsegmentación, lo que refleja las fortalezas y las limitaciones del enfoque.

Figura 2. Resultados de la precisión y pérdida en función del entrenamiento según las épocas



Fuente: elaboración propia

Figura 3. Comparación visual: imagen de ultrasonido, máscara predicha, máscara real



Fuente: Elaboración propia.

Si bien el modelo mostró resultados sólidos, se identificaron áreas de mejora, entre ellas la expansión del conjunto de datos, la implementación de técnicas más sofisticadas de aumento de datos, la experimentación con funciones de pérdida alternativas como *cross-entropy* ponderada o Dice loss, y la integración de métodos de cuantificación de incertidumbre para evaluar la confianza en las predicciones a nivel de píxel.

En conjunto, los resultados experimentales aportan evidencia convincente sobre el potencial de aplicación clínica de los modelos de aprendizaje profundo y, en particular, de las arquitecturas U-Net como herramientas complementarias para los radiólogos. Al permitir la segmentación precisa de regiones tumorales en imágenes de ultrasonido mamario, estos sistemas pueden apoyar significativamente la toma de decisiones médicas, mejorar la exactitud diagnóstica y contribuir a una atención más eficaz a los pacientes.

4. DISCUSIÓN

Los resultados obtenidos en este estudio evidencian la eficacia de las arquitecturas basadas en U-Net para la segmentación automática de tumores en imágenes de ultrasonido mamario, en línea con investigaciones recientes que destacan la aplicabilidad de las CNN en el diagnóstico asistido por IA. Los hallazgos confirman que U-Net constituye una arquitectura base sólida (Ronneberger et al., 2015; Isensee et al., 2021) y robusta, capaz de identificar patrones relevantes incluso en escenarios de bajo contraste o bordes difusos, ofreciendo un apoyo complementario al que realizan los radiólogos. En este caso, la U-Net estándar alcanzó un desempeño competitivo en precisión global, pero no superó el 0.9 en Dice como lo reportan algunos estudios que aplican mecanismos de atención (Byra et al., 2020; Chen et al., 2023).

Esto confirma tanto la validez como las limitaciones de la arquitectura básica respecto a variantes más modernas. Diversos estudios de revisión sistemática (Minaee et al., 2021) destacan la evolución de las técnicas de segmentación profunda, incluyendo variantes

con mecanismos de atención (Sulaiman et al., 2024; Anari et al., 2025; Ejiyi et al., 2024) y arquitecturas híbridas (Cao et al., 2023; Guo et al., 2021; Zarbakhsh, 2023; Sun et al., 2024; Liu et al., 2024). Asimismo, conjuntos de datos complementarios, como BreastDM para imágenes de resonancia dinámica (Zhao et al., 2023), han impulsado nuevas métricas de validación.

De igual modo, modelos híbridos que combinan CNN con módulos de atención (CBAM, spatial attention) o incluso con transformadores han alcanzado métricas superiores (Guo et al., 2021; Zarbakhsh, 2023; Sun et al., 2024; Liu et al., 2024), logrando valores de Dice mayores al 0.9 en conjuntos de datos de referencia (Chen et al., 2023; Guo et al., 2021). En comparación con métodos tradicionales de análisis de imágenes, la segmentación automática ofrece mayor consistencia y reduce la dependencia de la experiencia individual, consolidándose como una alternativa prometedora.

La relevancia de este trabajo radica en la posibilidad de integrar este tipo de modelos en entornos clínicos reales, en particular en contextos de recursos limitados o alta demanda diagnóstica, donde pueden contribuir a mejorar la precisión y rapidez en la detección temprana del cáncer de mama. A futuro, la expansión del conjunto de datos, la aplicación de técnicas más avanzadas de aumento de datos, la experimentación con funciones de pérdida adaptativas y la incorporación de arquitecturas híbridas con mecanismos de atención se presentan como líneas prioritarias para robustecer la capacidad de generalización y la aplicabilidad clínica del modelo.

5. CONCLUSIONES

El desarrollo de un modelo autoaprendiente basado en aprendizaje profundo para la detección de cáncer de mama en imágenes de ultrasonido representa un avance significativo en la imagenología médica y en la tecnología de diagnóstico. Mediante el uso de CNN, el sistema es capaz de aprender de forma autónoma de extensos conjuntos de datos, lo que incrementa su

precisión en la identificación de estructuras potencialmente cancerosas.

Durante el entrenamiento, se observó una mejora constante en el desempeño hasta la época 40, luego de la cual comenzaron a presentarse fluctuaciones similares a una curva logarítmica. Estas variaciones sugirieron el inicio del sobreajuste, lo que motivó detener el entrenamiento en la época con mayor precisión y mejor rendimiento. Esta decisión permitió mantener un equilibrio óptimo entre aprendizaje y capacidad de generalización.

El trabajo demuestra la viabilidad de aplicar arquitecturas de aprendizaje profundo, en particular U-Net, para la segmentación automática de tumores en ultrasonido mamario. El modelo propuesto logró segmentaciones precisas en la mayoría de los casos, consolidándose como una herramienta prometedora de apoyo en el diagnóstico médico.

No obstante, se identificaron limitaciones en situaciones con bordes poco definidos o tumores de morfología compleja, lo que resalta la necesidad de continuar optimizando este enfoque. La integración de soluciones impulsadas por IA en la práctica clínica resulta viable y tiene el potencial de mejorar las tasas de detección temprana.

Los resultados sugieren una reducción de falsos positivos en comparación con los enfoques manuales, aunque se precisa validación clínica para confirmarlo, y una disminución de la carga cognitiva de los profesionales de la salud. Futuras investigaciones deberán enfocarse en la expansión y diversificación de los conjuntos de datos, la comparación con otros enfoques de segmentación, la incorporación de arquitecturas con mecanismos de atención, validaciones externas y la optimización de funciones de pérdida como Dice loss o IoU loss (Ming y Xiao, 2024). Estos avances permitirán incrementar la robustez del modelo en poblaciones diversas y facilitar su integración en los flujos de trabajo clínico, maximizando su impacto en el diagnóstico y tratamiento oportuno del cáncer de mama.

REFERENCIAS

- Al-Dhabyani, W., Gomaa, M., Khaled, H., & Fahmy, A. (2020). Dataset of breast ultrasound images. *Data in Brief*, 28, 104863. <https://doi.org/10.1016/j.dib.2019.104863>
- Anari, S., Sadeghi, S., Sheikhi, G., Ranjbarzadeh, R., & Bende-chache, M. (2025). Explainable attention based breast tumor segmentation using a combination of UNet, ResNet, DenseNet, and EfficientNet models. *Scientific Reports*, 15(1), 1027. <https://doi.org/10.1038/s41598-024-84504-y>
- Balaji, K. (2023). Image Augmentation based on Variational Autoencoder for Breast Tumor Segmentation. *Academic Radiology*, 30, S172-S183. <https://doi.org/10.1016/j.acra.2022.12.035>
- Borah, N., Varma, P. S. P., Datta, A., Kumar, A., Baruah, U., & Ghosal, P. (2022). Performance Analysis of Breast Cancer Classification from Mammogram Images Using Vision Transformer. *2022 IEEE Calcutta Conference (CALCON)*, 238-243. <https://doi.org/10.1109/CALCON56258.2022.10060315>
- Byra, M., Jarosik, P., Dobruch-Sobczak, K., Klimonda, Z., Piotrkowska-Wroblewska, H., Litniewski, J., & Nowicki, A. (2020). *Breast mass segmentation based on ultrasonic entropy maps and attention gated U-Net* (No. arXiv:2001.10061). arXiv. <https://doi.org/10.48550/arXiv.2001.10061>
- Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2023). Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation. En L. Karlinsky, T. Michaeli, & K. Nishino (Eds.), *Computer Vision – ECCV 2022 Workshops* (Vol. 13803, pp. 205-218). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-25066-8_9
- Chen, G., Li, L., Dai, Y., Zhang, J., & Yap, M. H. (2023). AAU-Net: An Adaptive Attention U-Net for Breast Lesions Segmentation in Ultrasound Images. *IEEE Transactions on Medical Imaging*, 42(5), 1289-1300. <https://doi.org/10.1109/TMI.2022.3226268>
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). *3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation* (Versión 1). arXiv. <https://doi.org/10.48550/ARXIV.1606.06650>
- Ejiyi, C. J., Qin, Z., Agbesi, V. K., Ejiyi, M. B., Chikwendu, I. A., Bamisile, O. F., Onyekwere, F. E., & Bamisile, O. O. (2024). Attention-enriched deeper UNet (ADU-NET) for disease diagnosis in breast ultrasound and retina fundus images. *Progress in Artificial Intelligence*, 13(4), 351-366. <https://doi.org/10.1007/s13748-024-00340-1>
- Guo, Y., Duan, X., Wang, C., & Guo, H. (2021). Segmentation and recognition of breast ultrasound images based on an expanded U-Net. *PLOS ONE*, 16(6), e0253202. <https://doi.org/10.1371/journal.pone.0253202>

- Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2), 203-211. <https://doi.org/10.1038/s41592-020-01008-z>
- Liu, S., Liu, D., & Lin, Y. (2024). DPNNet: A dual-attention patching network for breast tumor segmentation in an ultrasound image. *Multimedia Systems*, 30(6), 355. <https://doi.org/10.1007/s00530-024-01562-y>
- Minaee, S., Boykov, Y. Y., Porikli, F., Plaza, A. J., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image Segmentation Using Deep Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1-1. <https://doi.org/10.1109/TPAMI.2021.3059968>
- Ming, Q., & Xiao, X. (2024). Towards Accurate Medical Image Segmentation with Gradient-Optimized Dice Loss. *IEEE Signal Processing Letters*, 31, 191-195. <https://doi.org/10.1109/LSP.2023.3329437>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. En N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (Vol. 9351, pp. 234-241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28
- Sulaiman, A., Anand, V., Gupta, S., Rajab, A., Alshahrani, H., Al Reshan, M. S., Shaikh, A., & Hamdi, M. (2024). Attention based UNet model for breast cancer segmentation using BUSI dataset. *Scientific Reports*, 14(1), 22422. <https://doi.org/10.1038/s41598-024-72712-5>
- Sun, J., Yuan, B., Tian, J., Yuan, Z., Wu, Y., He, X., Yang, C., & Tao, Z. (2024). DDRA-Net: Dual-Channel Deep Residual Attention UPerNet for Breast Lesions Segmentation in Ultrasound Images. *IEEE Access*, 12, 43691-43703. <https://doi.org/10.1109/ACCESS.2024.3373551>
- Zarbakhsh, P. (2023). Spatial Attention Mechanism and Cascade Feature Extraction in a U-Net Model for Enhancing Breast Tumor Segmentation. *Applied Sciences*, 13(15), 8758. <https://doi.org/10.3390/app13158758>
- Zhao, X., Liao, Y., Xie, J., He, X., Zhang, S., Wang, G., Fang, J., Lu, H., & Yu, J. (2023). BreastDM: A DCE-MRI dataset for breast tumor image segmentation and classification. *Computers in Biology and Medicine*, 164, 107255. <https://doi.org/10.1016/j.compbiomed.2023.107255>
- Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation (Versión 1). arXiv. <https://doi.org/10.48550/ARXIV.1807.10165>