

DOI: <https://doi.org/10.47300/actasidi-unicyt-2024-14>

BRYSTIA: DETECCIÓN Y CLASIFICACIÓN DE CÁNCER DE MAMA EN ULTRASONIDOS MEDIANTE PROCESAMIENTO DIGITAL DE IMÁGENES E INTELIGENCIA ARTIFICIAL

Velasco Urbano, Jose David

Universidad Del Cauca

Popayán, Colombia

jvelascou@unicauca.edu.co

<https://orcid.org/0009-0003-1082-8151>

Guerrero Erazo, Sebastian Camilo

Universidad Del Cauca

Popayán, Colombia

secguerreo@unicauca.edu.co

<https://orcid.org/0009-0007-8412-9467>

Ramirez Mina, Fabián Andres

Universidad Del Cauca

Popayán, Colombia

framiezm@unicauca.edu.co

<https://orcid.org/0009-0009-5082-1898>

RESUMEN

En este artículo, se presenta un enfoque novedoso para la detección y clasificación del cáncer de mama en ecografías mediante el uso de procesamiento digital de imágenes e inteligencia artificial. La herramienta desarrollada, denominada BrystIA, utiliza algoritmos avanzados para la mejora y segmentación de ecografías mamarias, junto con máquinas de soporte vectorial (SVM), bosques aleatorios y redes neuronales entrenadas para identificar y clasificar masas anómalas, benignas o malignas, con alta precisión. Los resultados de las pruebas indican una mejora significativa en el diagnóstico en comparación con los métodos tradicionales. La interfaz, diseñada de forma intuitiva para los profesionales de la salud, facilita la visualización rápida y el análisis de los resultados, permitiendo también la supervisión y validación profesional. Este estudio destaca la eficacia de la IA en imagenología mamaria y subraya la importancia de la colaboración interdisciplinaria para mejorar la detección temprana del cáncer de mama, prometiendo una integración efectiva de tecnologías inteligentes en la práctica clínica para mejorar la precisión diagnóstica y la atención al paciente.

Palabras clave: cáncer de mama, ecografías, inteligencia artificial, procesamiento de imágenes.

ABSTRACT

In this article, a novel approach is presented for the detection and classification of breast cancer in ultrasound images using digital image processing and artificial intelligence. The developed tool, named BrystIA, utilizes advanced algorithms for the enhancement and segmentation of breast ultrasounds, along with support vector machines (SVM), random forests, and neural networks trained to identify and classify abnormal, benign, or malignant masses with high accuracy. Test results indicate a significant improvement in diagnostics compared to traditional methods. The interface, intuitively designed for healthcare professionals, facilitates rapid visualization and

analysis of results, also allowing for professional supervision and validation. This study highlights the effectiveness of AI in breast imaging and underscores the importance of interdisciplinary collaboration to improve early breast cancer detection, promising an effective integration of intelligent technologies into clinical practice to enhance diagnostic precision and patient care.

Keywords: breast cancer, ultrasound, artificial intelligence, image processing.

1. INTRODUCCIÓN

El cáncer de mama es una de las principales causas de mortalidad entre las mujeres a nivel mundial, con millones de nuevos casos diagnosticados cada año. A pesar de los avances en tecnologías de imagen médica, como las mamografías y las ecografías, la detección temprana sigue siendo un desafío, especialmente en mujeres con tejido mamario denso, donde las imágenes pueden ser difíciles de interpretar. Las ecografías no utilizan radiación ionizante y tienen la ventaja de ser más efectivas en la detección de masas en tejidos densos; sin embargo, su interpretación es altamente dependiente de la habilidad y experiencia del profesional médico, lo que introduce un componente subjetivo en el diagnóstico.

Ante estos desafíos, ha surgido el uso de técnicas de procesamiento digital de imágenes y, recientemente, de inteligencia artificial (IA) para mejorar la precisión en la detección y clasificación de masas mamarias. La IA ha demostrado ser una herramienta valiosa en el análisis de grandes cantidades de imágenes, ayudando a los profesionales de la salud a identificar patrones que podrían pasar desapercibidos para el ojo humano. Estas tecnologías permiten mejorar la precisión y agilizar el tiempo de análisis en el diagnóstico.

BrystIA es una interfaz innovadora que combina procesamiento digital de imágenes e IA con el fin de asistir en la detección y clasificación de cáncer de mama utilizando algoritmos de mejora y segmentación de imágenes, junto con métodos de clasificación como máquinas de soporte vectorial (SVM), bosques aleatorios y redes neuronales convolucionales para diferenciar entre tumores benignos y malignos.

1.1 Marco Conceptual

La capacidad de la IA para analizar grandes volúmenes de datos y aprender patrones complejos, ha permitido su aplicación en diversas áreas de la medicina, particularmente en el análisis de imágenes. Según estudios recientes, el uso de técnicas de aprendizaje automático, como las SVM y los bosques aleatorios, ha demostrado ser altamente efectivo en la identificación de enfermedades a partir de imágenes médicas, destacando su aplicación en el diagnóstico del cáncer de mama.

1.1.1. Procesamiento Digital de Imágenes en Radiología

El procesamiento digital de imágenes es una técnica clave para mejorar la calidad de las imágenes médicas antes de aplicar algoritmos de IA. Este proceso incluye la mejora del contraste, la eliminación de ruido y la segmentación de las regiones de interés. En el caso de BrystIA, se emplean técnicas avanzadas como "Histogram Stretching" para resaltar las áreas sospechosas dentro de las ecografías mamarias, lo que facilita la posterior clasificación de las masas.

La segmentación juega un rol crucial en la detección de tumores, ya que permite delimitar con precisión las áreas que serán analizadas por los algoritmos de clasificación. Diferentes estudios han mostrado que el uso de métodos morfológicos para la segmentación, como los umbrales adaptativos y las operaciones morfológicas, mejora significativamente la detección de anomalías en ecografías.

1.1.2. Inteligencia Artificial y Clasificación de Tumores

La inteligencia artificial, especialmente el aprendizaje automático, ha impulsado avances clave en el diagnóstico médico. Técnicas como máquinas de soporte vectorial (SVM) y bosques aleatorios se utilizan ampliamente para clasificar imágenes médicas. Las SVM destacan en espacios de alta dimensionalidad, identificando el hiperplano óptimo que separa las clases. Los bosques aleatorios, por su parte, son ideales para conjuntos de datos con múltiples características, ya que combinan varios árboles de decisión, mejorando la precisión y minimizando el sobreajuste; además, las redes neuronales convolucionales (CNN), una de las técnicas más avanzadas en deep learning para procesamiento de imágenes, han mostrado una capacidad notable para aprender características complejas, esenciales en tareas de clasificación. En BrystIA, estas técnicas se aplican para distinguir entre tumores benignos y malignos a partir de características extraídas de ecografías, logrando un análisis preciso y profundo de las imágenes.

1.1.3. Revisión Bibliográfica

Diversos estudios respaldan la efectividad del uso de IA en la detección de cáncer de mama. Investigaciones como las de Villavicencio et al. (2021) y Ramos Sánchez (2022) muestran cómo las redes neuronales y los bosques aleatorios han mejorado significativamente la precisión diagnóstica en la radiología mamaria. Estas investigaciones subrayan la importancia de integrar la IA en el flujo de trabajo clínico, facilitando la detección temprana y la personalización de los tratamientos, lo cual es esencial en la atención a pacientes con cáncer.

2. MATERIALES Y MÉTODOS

La metodología de este estudio se diseñó para implementar y evaluar el rendimiento de BrystIA, un sistema basado en inteligencia artificial para la detección y clasificación de tumores mamarios en ecografías. El proceso de desarrollo y evaluación del sistema incluyó varias fases:

2.1. Recopilación de Datos

Para el desarrollo y validación de BrystIA, se utilizó el conjunto de datos público *Breast Ultrasound Images Dataset (BUSI)*, disponible en la plataforma Kaggle (Breast Ultrasound Images Dataset, 2023). Este conjunto de datos contiene más de 780 imágenes de ecografías de mama, etiquetadas en tres categorías: benignas, malignas y normales. Las imágenes fueron tomadas con diversas condiciones de calidad, lo que ofreció una amplia variedad de ejemplos para entrenar los algoritmos de detección y clasificación. Para fines de entrenamiento y validación del modelo, las imágenes se dividieron en tres conjuntos: entrenamiento (70%), validación (15%), y prueba (15%).

2.2. Preprocesamiento de Imágenes

El preprocesamiento de las imágenes fue fundamental para asegurar que las ecografías mamarias estén en condiciones óptimas para su análisis y clasificación. En primer lugar, las imágenes en formato RGB fueron convertidas al espacio de color HSV, donde se trabajó principalmente con el plano “V” (valor) que contiene la información de intensidad. La elección del plano V fue clave, ya que la información de intensidad es crucial para resaltar las características estructurales de los tumores mamarios.

Posteriormente, se aplicó la técnica de “Histogram Stretching” en el plano V, con el objetivo de mejorar el contraste y destacar las áreas sospechosas, lo que permitió expandir el rango de valores de intensidad en la imagen, facilitando la identificación visual de los tumores. Después, se eliminaron ruidos no deseados utilizando un filtro de media.

2.3. Segmentación de Tumores

La segmentación fue realizada mediante umbrales adaptativos y técnicas morfológicas. Dado que los tumores pueden tener formas irregulares y variadas, fue necesario implementar un proceso de segmentación robusto que pudiera adaptarse a estas condiciones.

Se calculó un umbral adaptativo en el plano V de cada imagen, ajustándose automáticamente según las características de luminosidad. Este método permitió generar una imagen binaria donde las áreas con potencial presencia de tumor aparecían resaltadas; se llevaron a cabo operaciones morfológicas de apertura y cierre para eliminar el ruido de fondo y definir con precisión los bordes de los tumores.

Para mejorar la precisión de la segmentación, se empleó una forma estructural de disco en las operaciones morfológicas para modelar la geometría de los tumores mamarios, suavizando los bordes irregulares y permitiendo una segmentación coherente. Finalmente, la imagen resultante de la segmentación fue presentada en la interfaz de usuario, permitiendo a los profesionales revisar manualmente las áreas seleccionadas para su validación.

2.4. Extracción de Características

Una vez segmentadas las masas, se procedió a la extracción de características utilizando la función "regionprops" de MATLAB. Las características seleccionadas fueron: área, circularidad, excentricidad, longitud del eje mayor, longitud del eje menor. Estas características morfológicas fueron seleccionadas por su relevancia en la diferenciación entre tumores benignos y malignos.

2.5. Clasificación

BrystIA empleó tres técnicas de clasificación:

- **Máquinas de soporte vectorial (SVM):** este clasificador es útil para problemas de alta dimensionalidad. Se empleó un kernel de tipo Radial Basis Function (RBF).
- **Bosques aleatorios:** este clasificador combinó 50 árboles de decisión para mejorar la precisión y reducir el riesgo de sobreajuste.
- **Redes neuronales convolucionales (CNN):** se utilizaron para la detección de características complejas en las ecografías. La arquitectura de la CNN consiste de múltiples capas para su creación.

2.6. Validación del Modelo

El rendimiento de los modelos fue evaluado mediante las métricas de precisión, exactitud, sensibilidad y especificidad. Estas métricas se calcularon utilizando las predicciones de los modelos en el conjunto de prueba; además, se generaron matrices de confusión para cada modelo con el fin de evaluar su capacidad para clasificar correctamente los tumores benignos y malignos.

3. RESULTADOS Y DISCUSIÓN

3.1. Resultados Obtenidos con Bosques Aleatorios

Este clasificador mostró el mejor rendimiento entre los tres métodos evaluados. Este clasificador alcanzó una **precisión del 95%**, una **exactitud del 95%**, una **sensibilidad de 0.95** y una **especificidad de 0.95**. Estas métricas indican que el modelo es altamente efectivo tanto para detectar tumores malignos como para identificar correctamente los tumores benignos, minimizando los errores de clasificación.

3.2. Resultados Obtenidos con Máquinas de Soporte Vectorial (SVM)

Se alcanzó una **precisión del 90%**, una **exactitud del 90%**, una **sensibilidad de 0.85** y una **especificidad de 0.95**. Aunque las SVM presentaron una alta especificidad, su sensibilidad fue ligeramente inferior en comparación con los bosques aleatorios, lo que sugiere que este modelo tuvo dificultades para detectar algunos tumores malignos, resultando en un mayor número de falsos negativos.

3.3. Resultados Obtenidos con Redes Neuronales Convolucionales (CNN)

El modelo tuvo una **precisión del 86%**, una **exactitud del 88%**, una **sensibilidad de 0.90** y una **especificidad de 0.85**. A pesar de que la sensibilidad del modelo fue relativamente alta, lo que indica que las CNN fueron capaces de detectar el 90% de los tumores malignos, su especificidad fue inferior en comparación con los otros métodos, lo que llevó a un mayor número de falsos positivos.

3.4. Discusión

Tabla 1.

Resultados de los métodos de clasificación.

Método	Precisión (%)	Exactitud (%)	Sensibilidad	Especificidad
Bosques Aleatorios	95	95	0,95	0,95
SVM	90	90	0,85	0,95
CNN	86	88	0,90	0,85

Los resultados obtenidos reflejan la efectividad de BrystIA para la detección y clasificación de tumores mamarios, en particular cuando se utiliza el clasificador de bosques aleatorios. Este modelo sobresale en todas las métricas evaluadas, lo que lo convierte en una opción confiable para su implementación en entornos clínicos. Las SVM también ofrecen un rendimiento competitivo, con una alta especificidad que reduce la probabilidad de falsos positivos, aunque su sensibilidad podría mejorar.

Por otro lado, las CNN demostraron ser prometedoras, aunque requieren una mayor optimización para reducir los falsos positivos, su capacidad para procesar imágenes complejas es un punto fuerte.

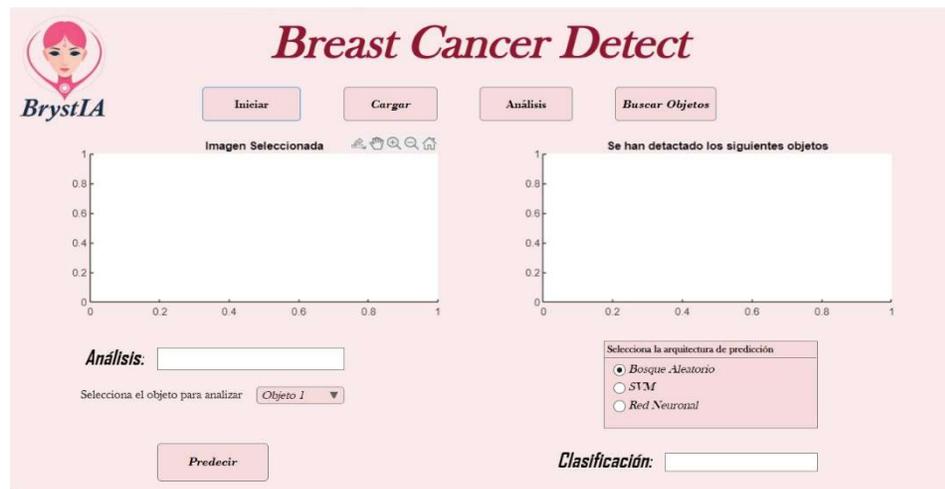
Es importante destacar que el éxito de BrystIA también depende de su capacidad para integrarse en el flujo de trabajo clínico. La interfaz diseñada permite que los profesionales de la salud validen los resultados automáticamente generados por el sistema, manteniendo un equilibrio entre la automatización proporcionada por la IA y la experiencia del especialista.

4. CONCLUSIONES Y REFERENCIAS

Los resultados sugieren que BrystIA es una herramienta poderosa para la detección y clasificación de cáncer de mama en ecografías, especialmente al utilizar bosques aleatorios, demostrando que la implementación de inteligencia artificial en la imagenología mamaria no solo mejora la precisión diagnóstica, sino que también proporciona una herramienta fiable para los profesionales de la salud.

La incorporación de tres métodos de clasificación incrementa la confiabilidad del diagnóstico, facilitando un trabajo colaborativo eficaz entre la interfaz y el profesional de salud. Esto optimiza los procesos de diagnóstico, permitiendo ayudar a más pacientes en menos tiempo.

Figura 1.
Interfaz BrystIA



REFERENCIAS

- Ramos Sánchez, E., & López de Ramos, A. (2022). Advances in deep learning for breast cancer detection. *International Journal of Radiology and Medical Imaging*, 5(3), 45-52. <https://doi.org/10.33412/ijrmi.2022.3078>
- Villavicencio, M., García, P., & Hernández, J. (2021). Machine learning in medical image analysis: A review. *Journal of Medical Imaging and Health Informatics*, 11(2), 204-210. <https://doi.org/10.1016/j.jmih.2021.08.023>
- Breast Ultrasound Images Dataset (BUSI). (2023, 1 de marzo). Kaggle. <https://www.kaggle.com/datasets/sabahesaraki/breast-ultrasound-images-dataset>

Los autores del trabajo autorizan a la Universidad Internacional de Ciencia y Tecnología (UNICYT) a publicar este resumen en extenso en las Actas del Congreso IDI-UNICYT 2024 en Acceso Abierto (Open Access) en formato digital (PDF) e integrarlos en diversas plataformas online bajo la licencia CC: Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

La Universidad Internacional de Ciencia y Tecnología y los miembros del Comité Organizador del Congreso IDI-UNICYT 2024 no son responsables del contenido ni de las implicaciones de lo expresado en este artículo.