

# Segmentación y Clasificación del Tejido Mamario en Mamografías Laterales

Pablo Torreira Pardo

*Máster en Bioinformática para las Ciencias de la Salud*

*Universidad de A Coruña*

A Coruña, España

pablo.torreiral@udc.es

30/05/2025

**Abstract**—La interpretación de mamografías puede verse limitada por la presencia de estructuras no relevantes como el músculo pectoral y artefactos, dificultando el análisis automatizado del tejido mamario. En este trabajo se propone un sistema automático para segmentar la región mamaria en mamografías laterales y clasificar el tejido como graso, glandular-graso o glandular-denso. Para la segmentación se ha diseñado un pipeline morfológico que elimina artefactos y aísla la región de interés. Posteriormente, se extraen descriptores de textura que alimentan clasificadores entrenados mediante validación cruzada. La segmentación alcanza una IoU media del 91.3%, y la clasificación obtiene un 87.0% de acierto con Random Forest y 82.6% con SVM. Estos resultados demuestran la viabilidad del enfoque propuesto y su potencial como apoyo al diagnóstico radiológico.

**Index Terms**—mamografía, segmentación, clasificación, aprendizaje automático

## I. INTRODUCCIÓN

El cáncer de mama es una de las principales causas de mortalidad entre mujeres a nivel global. La mamografía es una técnica diagnóstica y preventiva fundamental, ya que permite detectar afecciones en etapas tempranas y aumenta la probabilidad de un tratamiento efectivo. No obstante, su interpretación puede ser subjetiva, y la variabilidad entre especialistas motiva la necesidad de desarrollar herramientas automáticas de apoyo al diagnóstico.

En esta práctica se aborda la segmentación y clasificación del tejido mamario en mamografías laterales. El objetivo es doble: por un lado, aislar automáticamente la región de tejido mamario eliminando elementos irrelevantes como el músculo pectoral o artefactos; por otro, clasificar el tipo de tejido presente (graso, glandular-graso o glandular-denso), tarea relevante por su relación con el riesgo de cáncer y la densidad mamaria.

Se utiliza un subconjunto de 23 imágenes de la base de datos mini-MIAS, cada una acompañada de una máscara manual y una etiqueta de clase. Para la segmentación se ha desarrollado un pipeline basado en técnicas morfológicas y umbralización, que extrae la región de interés con alta concordancia respecto a la Ground Truth. Posteriormente, se extraen características de textura sobre la región segmentada y se emplean algoritmos de aprendizaje automático para realizar la clasificación, evaluando su rendimiento mediante validación cruzada estratificada.

## II. PROPUESTA TÉCNICA

La metodología propuesta se divide en dos fases principales: segmentación automática del tejido mamario y clasificación del tipo de tejido. Ambas han sido implementadas en Python, empleando bibliotecas de procesamiento de imagen y aprendizaje automático.

### A. Segmentación del tejido mamario

El objetivo principal de esta etapa es aislar automáticamente la región correspondiente al tejido mamario, excluyendo estructuras irrelevantes como el músculo pectoral, artefactos textuales o zonas externas al cuerpo. El pipeline fue implementado en Python, empleando bibliotecas como OpenCV para procesamiento morfológico, NumPy para operaciones matriciales y matplotlib para visualización de resultados. El proceso completo se compone de los siguientes pasos, ilustrados en la Figura 1:

- 1) **Carga y filtrado:** Se carga la mamografía en escala de grises y se aplica un filtro de mediana de  $11 \times 11$  píxeles para eliminar ruido conservando los bordes relevantes.
- 2) **Umbralización global ( $T=15$ ):** Se realiza una binarización con un umbral bajo para capturar regiones brillantes, como el seno, el músculo y artefactos. Se eliminan píxeles de los márgenes extremos para evitar detecciones espurias.
- 3) **Segmentación por componentes conectados:** Se identifican las regiones mediante `connectedComponentsWithStats` y se selecciona la de mayor área como primera aproximación a la región del seno.
- 4) **Aplicación de máscara primaria:** La máscara obtenida se utiliza para anular el fondo y enfocar el análisis en el tejido relevante.
- 5) **Detección del músculo pectoral:** Se realiza una nueva umbralización ( $T=170$ ), seguida de apertura morfológica ( $7 \times 7$ ), que permite identificar regiones brillantes en la zona superior de la imagen ( $y < 30$ ), característico del músculo pectoral.
- 6) **Filtrado anatómico:** Las regiones candidatas se validan según dos criterios: área mínima de 2000 píxeles y posición anatómica esperada.

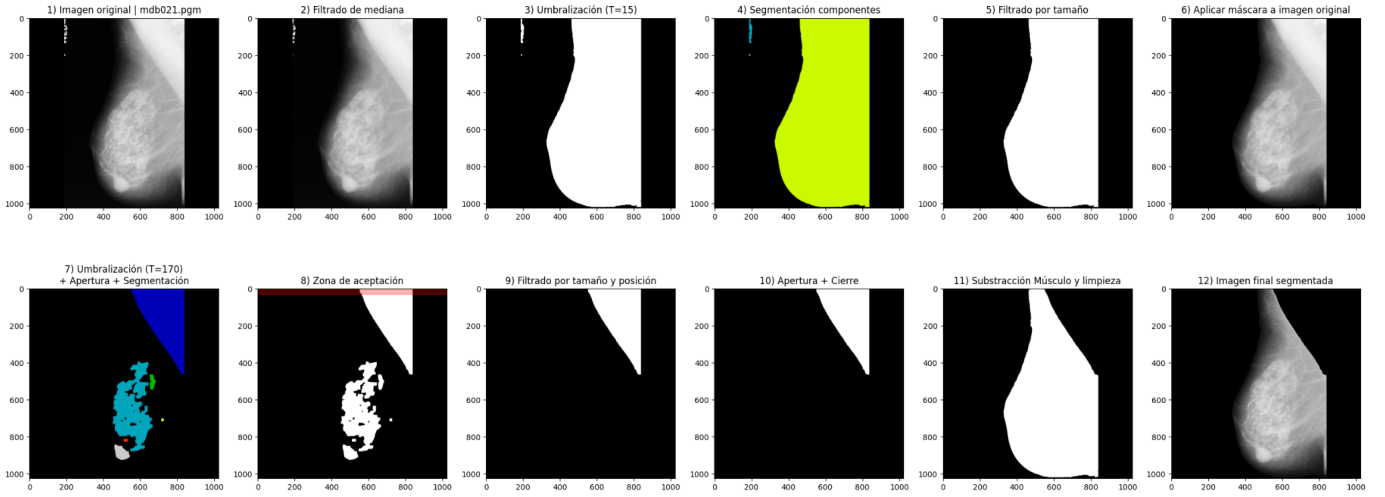


Fig. 1. Pipeline completo de segmentación automática aplicado a la imagen mdb021.pgm. (1) Imagen original, (2) filtrado de mediana, (3) umbralización ( $T=15$ ), (4) segmentación por componentes conectados, (5) filtrado por tamaño, (6) aplicación de máscara, (7) umbralización alta ( $T=170$ ) + apertura, (8) validación anatómica (zona de aceptación), (9) filtrado por posición y tamaño, (10) cierre morfológico, (11) sustracción del músculo y limpieza, (12) imagen final segmentada.

- 7) **Relleno y limpieza morfológica:** Se aplican operaciones de cierre y apertura con un elemento estructurante elíptico de  $21 \times 21$  píxeles para eliminar huecos e irregularidades.
- 8) **Sustracción final:** Se resta la región del músculo pectoral de la máscara general, obteniendo la máscara final del seno, que se aplica sobre la imagen original para obtener la segmentación definitiva.

#### Evaluación de la segmentación

La calidad de la segmentación automática se evaluó utilizando la métrica **Intersection over Union (IoU)**, definida como la proporción entre la intersección y la unión de la máscara predicha y la máscara de referencia, conocida como **Ground Truth (GT)**:

$$\text{IoU} = \frac{|\text{Segmentación} \cap \text{GT}|}{|\text{Segmentación} \cup \text{GT}|} \quad (1)$$

Además, se calcularon las siguientes métricas adicionales:

- **True Positives (TP):** píxeles correctamente segmentados como tejido.
- **False Positives (FP):** píxeles segmentados incorrectamente fuera del área de la GT.
- **False Negatives (FN):** píxeles presentes en la GT pero ausentes en la segmentación.

Estos valores se obtuvieron para cada imagen del conjunto, y se visualizan. Se puede ver un ejemplo en la Figura 2.

#### B. Clasificación del tejido mamario

Tras la segmentación del seno, se procedió a la clasificación del tejido mamario en una de tres clases: graso (clase 0), glandular-graso (clase 1) y glandular-denso (clase 2). Para ello, se extrajeron características de textura a partir de la región segmentada y se entrenaron modelos de aprendizaje automático supervisado para la tarea de clasificación.

#### Vectores de características

Cada imagen fue representada mediante un vector de seis características extraídas exclusivamente sobre la región segmentada (a partir de la máscara binaria). Las características fueron seleccionadas por su capacidad para describir la textura y la intensidad del tejido, que son los principales criterios para diferenciar entre clases. Se usaron:

- **Media de intensidad.**
- **Desviación típica de intensidad.**
- **Entropía local media**, calculada mediante un filtro de entropía sobre una vecindad circular con `skimage.filters.rank.entropy`.
- **Contraste, homogeneidad y energía**, derivados de la matriz de co-ocurrencia de niveles de gris (GLCM) calculada con `skimage.feature.graycomatrix`.

Estas seis variables se almacenaron como un vector por imagen y sirvieron como entrada para los clasificadores.

#### Clasificadores empleados

Se implementaron dos algoritmos de clasificación usando la biblioteca `scikit-learn`:

- **Random Forest (RF):** Se empleó un clasificador `RandomForestClassifier` con 100 árboles (`n_estimators=100`) y profundidad máxima de 100. No se aplicó normalización, ya que los árboles de decisión no son sensibles a la escala de los datos.
- **Máquinas de Vectores de Soporte (SVM):** Se utilizó `SVC` con kernel RBF y parámetros por defecto ( $C = 1.0$ ,  $\gamma = \text{scale}$ ). Se aplicó normalización Z-score con `StandardScaler` dentro de cada fold para evitar fuga de datos entre entrenamiento y validación. La normalización es clave en el uso de SVMs, ya que este clasificador es muy sensible a la diferencia en escala de las variables.

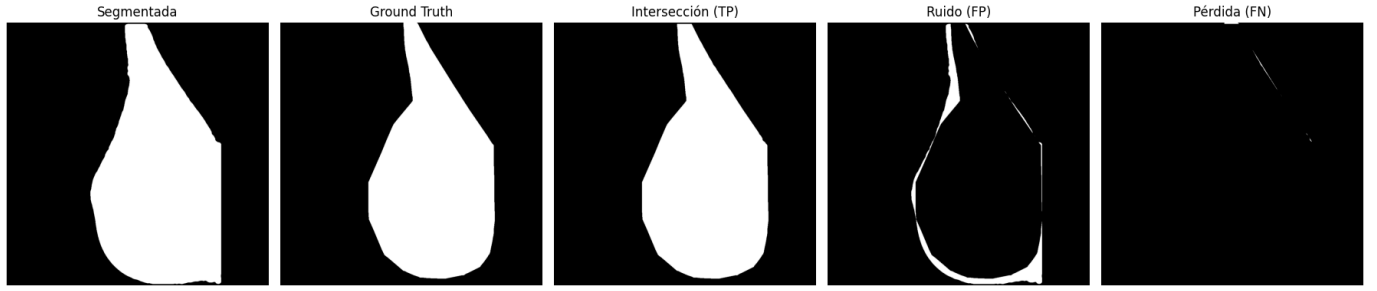


Fig. 2. Evaluación de segmentación: (de izquierda a derecha) máscara segmentada, Ground Truth (GT), intersección (TP), ruido (FP) y pérdida (FN).

### Validación cruzada

Para ambos clasificadores se aplicó validación cruzada estratificada (*StratifiedKfold*,  $k = 4$ ), garantizando que cada fold conservara la proporción original de clases (8-8-7). En cada iteración, los modelos se entrenaron con 3 folds y se evaluaron en el fold restante.

Las métricas obtenidas incluyen precisión (*accuracy*), matriz de confusión, y medidas por clase (precisión, *recall* y F1-score). Los resultados se agregaron para obtener métricas globales.

## III. RESULTADOS

### A. Evaluación de la segmentación

El método propuesto fue evaluado sobre un conjunto de 23 mamografías mediante comparación entre la máscara generada automáticamente y la Ground Truth (GT) proporcionada para cada imagen. La métrica empleada fue la **intersección sobre la unión (IoU)**. Los resultados por imagen oscilaron entre un mínimo del 81.7% y un máximo del 96.7%, con una **IoU media de 91.3%**.

### B. Evaluación de la clasificación

Se entrenaron y validaron dos modelos de clasificación Random Forest (RF) y SVM utilizando validación cruzada estratificada de 4 folds. Las Tablas I y II muestran las métricas por clase (precisión, *recall* y F1-score) en cada uno de los clasificadores.

TABLE I  
MÉTRICAS POR CLASE PARA RANDOM FOREST

Clase	Precisión	Recall	F1-score	Soporte
Graso (0)	0.857	0.750	0.800	8
Glandular-graso (1)	0.778	0.875	0.824	8
Glandular-denso (2)	1.000	1.000	1.000	7

TABLE II  
MÉTRICAS POR CLASE PARA SVM

Clase	Precisión	Recall	F1-score	Soporte
Graso (0)	0.857	0.750	0.800	8
Glandular-graso (1)	0.750	0.750	0.750	8
Glandular-denso (2)	0.875	1.000	0.933	7

Además de las métricas por clase, se calculó la precisión global (*accuracy*) de cada clasificador como proporción de

predicciones correctas sobre el total de muestras. El modelo **Random Forest** alcanzó una *accuracy* media del **87.0%**, mientras que el modelo **SVM** obtuvo un **82.6%**. Estas cifras, junto con los valores elevados de F1-score por clase, indican que ambos clasificadores son capaces de discriminar entre los tres tipos de tejido mamario con alto grado de acierto. No obstante, el modelo basado en árboles mostró una mayor robustez y consistencia, siendo menos sensible a la escala de los datos y presentando menor variabilidad entre clases. Los resultados reflejan una mayor precisión y equilibrio general en el modelo Random Forest, especialmente en la clase 1. Ambos modelos muestran un rendimiento excelente en la clase **glandular-denso**, con F1-score cercano o igual a 1.000, lo que sugiere que la textura densa es fácilmente reconocible por los descriptores empleados.

## IV. CONCLUSIONES Y POSIBLES MEJORAS

En este trabajo se ha desarrollado un sistema automático para el análisis de mamografías laterales, con el objetivo de segmentar la región del tejido mamario y clasificarla en uno de tres tipos: graso, glandular-graso y glandular-denso. La metodología se ha basado en técnicas de procesamiento morfológico, extracción de descriptores de textura y modelos de aprendizaje automático.

Los resultados obtenidos validan la eficacia del enfoque propuesto. La segmentación automática logró una **IoU media del 91.3%**, con alta concordancia respecto a las máscaras manuales (GT). La clasificación posterior, basada en características simples de intensidad y textura, alcanzó una **accuracy del 87.0%** con Random Forest, mostrando un rendimiento robusto en las tres clases. El modelo SVM también ofreció buenos resultados (82.6% de acierto), aunque mostró mayor sensibilidad a la normalización y más errores entre las clases intermedias.

### Posibles mejoras

Aunque los resultados son prometedores, el sistema puede beneficiarse de varias mejoras:

- **Uso de características más avanzadas:** Incorporar descriptores más ricos como LBP podría mejorar la diferenciación entre clases de textura similares.
- **Ampliación del conjunto de datos:** El pequeño tamaño de la muestra (23 imágenes) limita la capacidad de

generalización. Evaluar el sistema sobre un conjunto más amplio y variado permitiría validar su robustez.

- **Clasificadores más complejos:** El uso de redes neuronales convolucionales (CNN) podría permitir una clasificación end-to-end sin necesidad de extracción manual de características.
- **Evaluación clínica:** La validación con datos clínicos reales y en un entorno radiológico podría aportar evidencia sobre la aplicabilidad práctica del sistema como herramienta de apoyo al diagnóstico.

En conjunto, la solución desarrollada demuestra ser funcional y reproducible, constituyendo una base sólida sobre la que construir futuros sistemas automáticos de análisis mamográfico más precisos y escalables.