

Análisis de imágenes biomédicas – Práctica 3

Máster en Bioinformática para las Ciencias de la Salud

Curso 2024/2025

Clasificación del tejido mamario en mamografía lateral

Las mamografías son una modalidad de imagen de los senos utilizando rayos X de baja dosis. Estas imágenes son ampliamente utilizadas en la detección temprana de cáncer de mama, lo que hace que su análisis automático sea de especial importancia. En este trabajo nos centraremos en la detección del área mamaria en tomas laterales y en la categorización del tejido en tres tipos, graso, glandular-graso y glandular-denso (Figura 1)

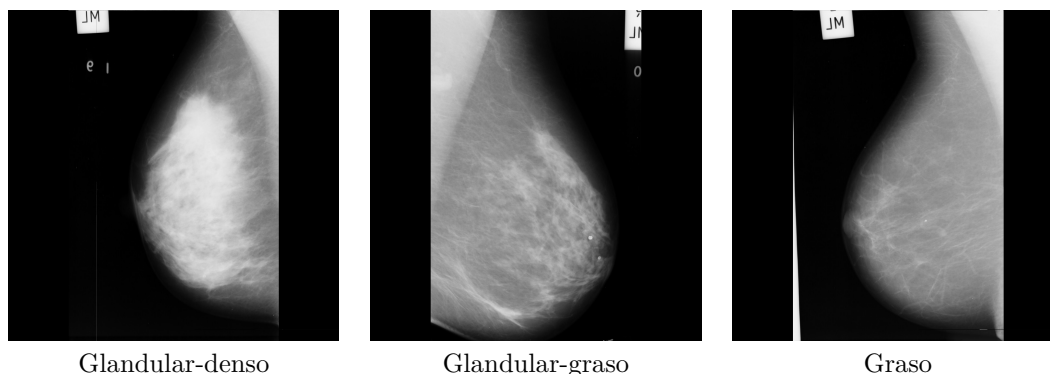


Figura 1: Tipos de tejido mamario

Materiales

Para este trabajo se utilizará un subconjunto de 24 imágenes provenientes del mini-MIAS database of mammograms¹ que se encuentra disponible en Moodle. Este conjunto contiene 8 imágenes por cada tipo de tejido (glandular-denso, glandular-graso y graso) con sus correspondientes máscaras de segmentación.

Se proporciona además un fichero con la clasificación de todas las imágenes de la base de datos (`info.csv`) así como un programa en python y un notebook para cargar el conjunto de datos (`read_files.py` y `p3.ipynb`).

Tarea 1: Segmentación

La primera tarea consistirá en el desarrollo de un método que permita extraer de forma automática la región de tejido mamario de la imagen. Como se puede observar en la Figura 2, para extraer dicha región es necesario descartar otros elementos presentes en la imagen: la región del músculo pectoral y otros artefactos (recuadro blanco con texto, letras, zonas blancas, ...). Para validar los resultados de la segmentación se proporciona un conjunto de máscaras manuales.

Notas

- Los niveles de intensidad de los distintos elementos son similares en todas las imágenes.

¹<http://peipa.essex.ac.uk/info/mias.html>

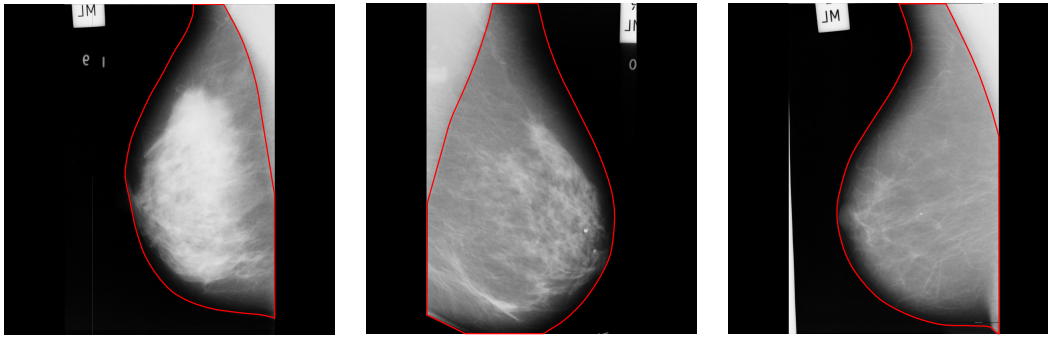


Figura 2: Área de tejido mamario

- El fondo de la imagen es prácticamente negro y los artefactos tienen un color muy claro
- La región de músculo es uniforme, de un color más claro que el tejido mamario, presenta una forma casi triangular y se encuentra en la parte superior, izquierda o derecha, de la imagen.
- Utilizad alguna métrica para evaluar el ajuste de vuestra propuesta.

Tarea 2: Clasificación

Una vez obtenida la región de interés, la segunda tarea consiste en la clasificación del tejido mamario que aparece en esta región. Para ello es necesario extraer una o varias características de la región que describan el tipo de tejido y permitan distinguirlo de los otros tipos.

Notas

- Las principales diferencias entre los tres tipos de tejidos son de color y textura. Tened en cuenta que sólo es relevante la zona de tejido mamario.
- El fichero info.csv contiene las clases de tejido mamario correspondientes a cada imagen.
- Para la clasificación podéis usar heurísticas o algoritmos de aprendizaje máquina. En este último caso, utilizad validación cruzada para obtener el porcentaje de acierto en el conjunto de validación.

Entrega

La fecha de entrega de esta práctica es el 30 de Mayo de 2025. La realización de la práctica es individual. Cada estudiante deberá entregar antes de la fecha límite un fichero comprimido con el siguiente contenido:

- El código escrito en `python`
- Un informe describiendo y justificando la propuesta en formato conferencia IEEE ² de 4 páginas como máximo. La estructura sugerida para el informe es la siguiente:
 - Título, autores, fecha
 - Abstract: breve resumen del problema, la propuesta, los resultados y las conclusiones.
 - Introducción al problema, a los materiales y al método propuesto
 - Propuesta técnica: descripción de los pasos seguidos para resolver el problema
 - Resultados: evaluación de los resultados obtenidos en ambas tareas
 - Conclusiones y posibles mejoras

Cada estudiante deberá defender su propuesta en una reunión de Teams con el profesor tras la entrega de la práctica.

²<https://www.ieee.org/conferences/publishing/templates.html>

Requisitos/Rúbrica³

- Implementar un método para extraer de forma satisfactoria la región de tejido mamario (hasta 1.25 puntos)
- Desarrollar un algoritmo que permita clasificar correctamente el tejido mamario (hasta 1.25 puntos)
- Informe (hasta 0.5 puntos)

³Puntuación condicionada a la defensa de la práctica