

OBJETIVOS Y DESARROLLO

- Segmentar la región de la mama, excluyendo el músculo del brazo. Para ello se probaron los siguientes métodos:
 - Detectar los bordes en la mamografía con Canny, el problema es que el perímetro de la mama está difuminado y el algoritmo no puede detectar el borde que la delimita directamente, y tampoco el del brazo, por lo que antes de hacer Canny se decidió ecualizar el histograma y así amplificar la transición entre el interior y el exterior de la mama, lo cual funcionaba bien con algunas de las imágenes pero se seguía sin obtener el borde del brazo, así que finalmente se descartó este método.
 - Umbralizar la mamografía dos veces, primero con un valor bajo (10/255) para reconocer la forma de la mama incluido el brazo, y segundo con un valor alto (240/255) para solo el brazo y así hacer la resta del primero con el segundo, dejando solo la mama. Los inconvenientes fueron: en la mamografía se encuentran otros elementos no relevantes con tono claro exteriores a la mama que se mantienen en las umbralizaciones; también, en la segunda umbralización, cuando la mama es glandular, se detecta también las glándulas e incluso éstas se unen al brazo, lo que produce que la mama quede vacía y abierta al exterior, por lo que no refleja su correcta forma. Para solucionar estos problemas, se realizó la operación de suavizado morfológico en las imágenes resultantes de las umbralizaciones antes de hacer su resta, y tras hacerlo. De esta manera se consiguió eliminar el elemento que identificaba la radiografía pero no las glándulas ni otros elementos no deseados grandes sin deformar en exceso la forma de la mama, esto podría deberse a los valores escogidos en los umbrales, por lo que se optó por el siguiente método.
 - Formar conjuntos de colores similares con Kmeans, para luego obtener los grupos más claros con una umbralización, en los que se encuentran el músculo del brazo, las glándulas y otros elementos no relevantes todos separados. El siguiente paso es encontrar cada uno de estos contornos con el algoritmo Suzuki85, luego se identifican en base su ángulo y tamaño, calculado con su momento^[1], y ya que el músculo del brazo suele tener un ángulo entre 10° y 60° cuando la mama está a la izquierda, y 150° y 200° a la derecha, y un tamaño mayor a 100, se puede concluir que el contorno que cumpla estas condiciones solo puede ser el músculo del brazo basado en las imágenes de mamografías aportadas. Con esto, solo falta conseguir el contorno de toda la mama incluido el músculo para restárselo, omitiendo los elementos no relevantes, esto se consiguió haciendo otra umbralización con la imagen resultante del Kmeans pero con el color del grupo más oscuro, obteniendo los contornos de todos los grupos externos sin el fondo. Para identificar el contorno de la mama con el músculo entre estos, en base a las imágenes de mamografías aportadas, es el contorno con mayor área. Finalmente se restan el contorno de la mama y el músculo tras hacerle un suavizado morfológico, con el del músculo tras hacerle un suavizado morfológico y dilataciones para eliminar pequeños detalles irrelevantes, resultando en el contorno exclusivo de la mama, el cual se ilustra sobre la mamografía entregada como salida y funciona perfectamente sobre todas las aportadas. Por este motivo, fue el método escogido.
- Clasificar mamografía en, glandular-densa, glandular-grasa, y grasa. Para ello, se investigó^[2] sobre las características que miden la densidad de la mama, concluyendo que la cantidad de glándulas no va en función del tamaño pero sí que es frecuente que un gran tamaño se traduzca en una menor densidad de glándulas y una mayor apreciación de las fibras, lo que se confirma en la tabla 1 al aplicar los siguientes métodos:

- Obtener el tamaño de la mama:
 - Con el contorno de la mama y su momento, se puede calcular su área. Dividiendo ésta con el área de toda la mamografía, resulta en una buena aproximación para comparar su tamaño.
- Obtener la proporción de glándulas en la mama.
 - Con el contorno de la mama a modo de mascara y la mamografía umbralizada a un tono alto (240), se realiza la intersección entre estas, obteniendo una imagen de las partes con mayor tono en el interior de la mama que se interpretan como glándulas. Aunque al probar este método con diferentes mamografías, se puede apreciar que no solo aparecen las glándulas en el resultado, sino que también aparecen fibras próximas al músculo, por lo que se descartó este método.
 - Igual que el anterior método pero en vez de umbralizar la mamografía sin procesar para aplicarle la máscara, se le aplica antes el algoritmo Kmeans y luego umbralizar con el color del grupo con mayor tono, consiguiendo unicamente las glándulas. Finalmente se mide el área que ocupan las glándulas en la mama y se divide por el área total de esta, reflejando una buena estimación de la densidad de glándulas en la mama.
- Obtener la proporción de fibras en la mama.
 - Con el contorno de la mama a modo de mascara y la mamografía, se realiza la intersección entre estas, para a continuación aplicar Canny y así detectar los bordes de las fibras dentro de la mama. La pega es que en muchas mamografías no existe el suficiente contraste en las fibras, por lo que se paso a probar el siguiente método.
 - Lo mismo pero después de aplicar la intersección y antes de Canny, se ecualiza el histograma de la mamografía para aumentar el contraste de las fibras, consiguiendo detectar los bordes de la fibras en todas la mamografías entregadas, el único problema es que también se detecta mucho ruido, por lo que se le realiza un cierre y una apertura morfológica para eliminar este defecto y conectar la fibras adyacentes. Finalmente se mide el área que ocupan las fibras en la mama y se divide por el área total de esta, mostrando una buena estimación de la densidad de fibras en la mama.

Características	Tamaño de la mama		Proporción de fibras		Proporción de glándulas	
Medidas	Media	Desviación típica	Media	Desviación típica	Media	Desviación típica
Glandular-denso	0.3211	0.0862	0.0145	0.0314	0.2868	0.1152
Glandular-graso	0.3181	0.0653	0.0275	0.0303	0.1067	0.0638
Graso	0.3379	0.0896	0.0656	0.0463	0.0242	0.0389

Tabla 1: Resultados con las muestras proporcionadas

CLASIFICACIÓN

Con los atributos de tamaño, proporción de fibras, y proporción de glándulas de la mama, se le pueden entregar al clasificador que se prefiera. Como este punto no es relevante para la práctica, se desarrollo un clasificador simple con pesos en base a la media de cada atributo en las mamografías aportadas, que mostró una precisión del 83,33%

REFERENCIAS

- ^[1]Momento de una imagen: https://en.wikipedia.org/wiki/Image_moment
- ^[2]Características observables en las mamografías para clasificar la densidad de la mama: <https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/mammograms/breast-density-and-your-mammogram-report.html>