

1. Introducció

Aquest projecte final de carrera s'ha dut a terme amb el grup de Visió per Computador del departament d'Arquitectura i Tecnologia de Computadors de la Universitat de Girona. Esta emmarcat en l'àmbit de recerca en imatge mèdica i el que pretén és oferir una eina d'ajuda pels radiòlegs a l'hora de realitzar el diagnòstic de lesions al teixit mamari.

El càncer de mama consisteix en el creixement anormal de les cèl·lules del teixit mamari, provocant la formació de tumors o masses. Si el tumor no envaeix els teixits i òrgans propers, se l'anomena benigne, en cas contrari és un tumor maligne.

Els objectius del projecte es divideixen en tres blocs: Primerament, realitzar una segmentació automàtica del contorn d'una imatge on hi ha una massa central. Tot seguit, a partir del contorn trobat, caracteritzar la massa. I finalment, utilitzant les característiques anteriors classificar la massa en benigne o maligne.

En el projecte s'utilitza el Matlab com a eina de programació. Concretament les funcions enfocades al processat de imatges del toolbox de *Image processing* (propi de Matlab) i els classificadors de la PRTools de la *Delft University of Technology*.

2. Segmentació de masses

Segmentar consisteix en extreure informació concreta de la imatge, dividint-la en regions. En aquest projecte hem segmentat les imatges mamogràfiques amb els mètodes *Region Growing* (basat en regions) i *Fast Marching* (basat en contorns).

El **Region Growing** o creixement de regions consisteix en escollir una sèrie de píxels inicials (llavor), als que se li aniran afegint píxels amb propietats o característiques semblants (segons un criteri d'agregació o similitud).

S'han implementat sis variants de l'algorisme, sorgides de la combinació dels següents criteris d'agregació i del punt on comença a créixer la regió (llavor):

- La llavor pot trobar-se al centre de la imatge, a una cantonada de la imatge o al píxel amb nivell de gris més alt.
- I el criteri d'agregació pot ser per desviació o per intensitat. En el primer cas es calcula la mitjana i la desviació estàndard d'un quadrat $N \times N$ del centre de la regió d'interès i s'obté el següent criteri: $(mitjana - 3 * desviació) \leq f(x, y) \leq (mitjana + 3 * desviació)$. En el segon, el d'intensitat, es filtraran els píxels que s'agreguen a la regió segons un llindar Δ : $|f(x, y) - \mu_{Ri}| \leq \Delta$

El mètode **Level Set** (LS) és una tècnica numèrica per l'evolució d'interfícies i formes. Nosaltres hem implementat el *Fast Marching* una variant en què l'evolució de les corbes només és en una direcció.

Aquest mètode comença la seva evolució al centre de la imatge i va creixent cap a l'exterior d'aquesta segons el seu gradient. En zones planes l'evolució és ràpida i a mesura que es va acostant al contorn (més gradient) la velocitat de creixement disminueix. Un exemple d'evolució del *Fast Marching* és el que es mostra a la figura:

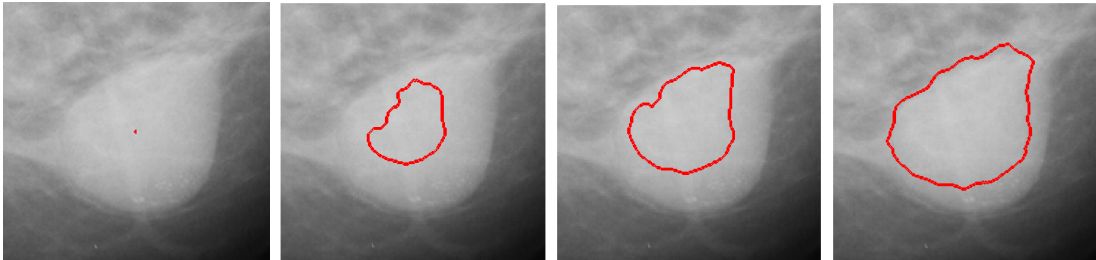


Figura 1: Evolució de la segmentació per la imatge 'rdb021ll.tif'

3. Caracterització i classificació de masses

Una vegada s'han segmentat les 40 masses, n'extraiem característiques que ens ajudaran posteriorment a realitzar el diagnòstic de cadascuna d'elles. Les característiques utilitzades són les següents:

Característiques	Nombre
Intensitat	3
Forma	10
Textura	11
Contorn	27
TOTAL	51

Taula 1: Resum de les característiques extretes

Creem un vector amb 50 d'aquestes característiques (discriminem l'àrea) i un altre vector de 40 posicions amb el diagnòstic del radiòleg.

Classificar consisteix en agrupar conjunts d'elements amb característiques comunes. En el nostre cas disposem de imatges mamogràfiques amb vectors de característiques que les defineixen. Tenim definits dos grups de imatges, les que contenen masses benignes i les malignes.

Per portar a terme la classificació hem utilitzat l'aprenentatge automàtic, que extreu coneixements a partir d'observacions i els utilitza per realitzar una classificació. També hem fet servir l'algorisme de selecció SFFS per escollir les característiques que ens donen millors resultats. I finalment amb els classificadors *fisherc*, *knn*, *treec* i *svc* del PRTools s'ha obtingut un diagnòstic per les imatges mitjançant l'estratègia *leave-one-out* (una imatge de test i les altres 39 d'entrenament i repetir-ho fins que totes les imatges de la base de dades han sigut testejades).

4. Resultats

Una vegada implementats els dos mètodes de segmentació, el mètode per extreure característiques i el classificador s'ha testejat el sistema amb 39 imatges (40 masses)

de la base de dades MIAS, de les quals disposem d'un diagnòstic i una segmentació manual d'un radiòleg.

A l'apartat de segmentació veiem que el mètode *Region Growing* obté millors resultats amb la llavor al centre i criteri d'agregació de desviació, encara que té problemes amb imatges amb poca diferència de contrast entre la massa i la resta de teixit mamari. Per altra banda l'algorisme *Fast Marching* té l'inconvenient de ser semi-automàtic, necessita algú que l'aturi en el moment adequat. La segmentació obtinguda s'acosta més a la del radiòleg, però té tendència a arrodonir la forma de la massa.

Comparant les segmentacions hem vist que el *Fast Marching* ens dona millors resultats, degut possiblement a què ha d'estar supervisat i per tant el creixement de la regió s'atura quan nosaltres ho creiem oportú.

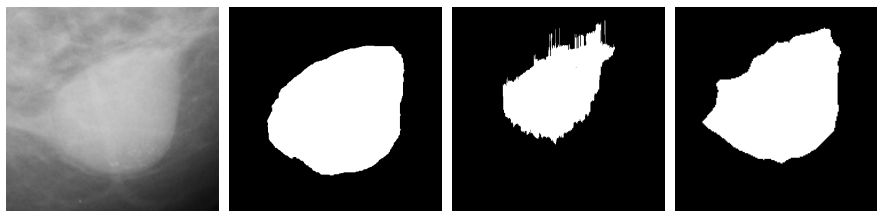


Figura 2: Resultats de la segmentació. Imatge original, segmentació metge, segmentació *Region Growing* i segmentació *Fast Marching*

Per altra banda, els resultats de classificació obtinguts amb la segmentació mèdica es mostren a la següent taula:

<i>Classificadors</i>	<i>Característiques</i>				
	Intensitat	Forma	Textura	Contorn	Totes
fisherc	62,5%	67,5%	70%	55%	70%
knn	57,5%	67,5%	75%	65%	55%
treec	62,5%	62,5%	52,5%	50%	57,5%
svc	70%	70%	60%	67,5%	50%

Taula 2: Resum de tant per cent d'encert a la classificació amb la segmentació del metge i utilitzant les característiques extrems

On s'observa que el classificador *fisherc* és el que millors resultats ens dona utilitzant les 50 característiques calculades. Però si prèviament realitzem una selecció de característiques, el millor classificador passa a ser el *knn*:

Classificadors	Encert
fisherc	62,5%
knn	85%
treec	72,5%
svc	62,5%

Escollir la millor classificació és molt complicat degut a que les diferències entre les diferents segmentacions són poc significatives.

5. Conclusions

Al finalitzar aquest projecte s'han assolit els objectius inicialment proposats. Per realitzar una segmentació automàtica del contorn s'han implementat els algorismes *Region Growing* i *Fast Marching*. S'han extret característiques de forma, intensitat, textura i contorn de la massa, per posteriorment realitzar un diagnòstic. A més a més dels objectius inicials, al llarg del projecte s'han adquirit diversos coneixements sobre les diferents eines utilitzades com el Matlab i l'editor de textos \LaTeX . S'ha estudiat la teoria dels mètodes de segmentació i classificació de masses per entendre què fan en cada moment i posteriorment implementar-los en Matlab. També s'ha après a analitzar documentació tècnica i així extreure'n informació útil pel projecte.

En resum, hem analitzat, dissenyat i implementat una eina d'ajuda al radiòleg per tal de diagnosticar, com a benigne o maligne, una regió d'interès que conté una massa. També hem après a realitzar un projecte informàtic seguint una planificació i objectius inicials. Personalment crec que els resultats finals han estat satisfactoris, encara que com en tots els projectes de recerca queden molts punts per seguir investigant.