



UNIVERSITAT DE  
BARCELONA

**Treball de Fi de Grau**

**GRAU D'ENGINYERIA INFORMÀTICA**

**Facultat de Matemàtiques i Informàtica**

**Universitat de Barcelona**

---

**SEGMENTACIÓ DE MAMOGRAFIES  
UTILITZANT TÈCNIQUES  
D'APRENENTATGE PROFUND**

---

**Daniel Juárez Gutiérrez**

Directora: Laura Igual Muñoz  
Realitzat a: Departament de  
Matemàtiques i Informàtica  
Barcelona, 12 de juny de 2023

## Resum

Els sistemes CADe i CADx (computer-aided detection i computer-aided diagnosis) són sistemes que ajuden als professionals mèdics a analitzar i avaluar en poc temps informació obtinguda a través de rajos X, ressonàncies magnètiques o ultrasons, entre d'altres. Aquests sistemes combinen elements de visió per ordinador i intel·ligència artificial amb tècniques d'imatge mèdica.

Un camp de treball important és l'anàlisi de mamografies per tal d'ajudar en el diagnòstic del càncer de mama.

Aquest treball té com a objectiu desenvolupar un sistema de segmentació de mamografies utilitzant l'aprenentatge profund, específicament l'arquitectura de xarxa neuronal U-Net. Per fer-ho, s'utilitza el conjunt de dades públiques CBIS-DDSM, que és un dels conjunts de dades més grans i àmpliament emprats en el camp de la mamografia per poder validar nous mètodes automàtics de segmentació.

## Resumen

Los sistemas CADe y CADx (computer-aided detection y computer-aided diagnosis) son sistemas que ayudan a los profesionales médicos a analizar y evaluar en poco tiempo información obtenida a través de rayos X, resonancias magnéticas o ultrasonidos, entre otros. Estos sistemas combinan elementos de visión por ordenador e inteligencia artificial con técnicas de imagen médica.

Un campo de trabajo importante es el análisis de mamografías para ayudar en el diagnóstico del cáncer de mama.

Este trabajo tiene como objetivo desarrollar un sistema de segmentación de mamografías utilizando el aprendizaje profundo, específicamente la arquitectura de red neuronal U-Net. Para hacerlo, se utiliza el conjunto de datos públicos CBIS-DDSM, que es uno de los conjuntos de datos más grandes y ampliamente empleados en el campo de la mamografía para poder validar nuevos métodos automáticos de segmentación.

## Abstract

CADe and CADx (computer-aided detection and computer-aided diagnosis) systems are designed to assist medical professionals in quickly analyzing and evaluating information obtained through X-rays, magnetic resonance imaging (MRI), ultrasounds, among others. These systems combine elements of computer vision and artificial intelligence with medical imaging techniques.

An important field of work for these systems is the analysis of mammograms to aid in the diagnosis of breast cancer.

The objective of this work is to develop a mammogram segmentation system using deep learning, specifically the U-Net neural network architecture. To accomplish this, the publicly available CBIS-DDSM dataset is utilized, which is one of the largest and widely employed datasets in the field of mammography to validate new automatic segmentation methods.

## ÍNDIX DE CONTINGUTS

<b>1. Introducció.....</b>	<b>7</b>
1.1. Contextualització del problema.....	7
1.2. Objectius del treball.....	8
1.3. Estructura del treball.....	9
<b>2. Antecedents i background tècnic sobre la segmentació de mamografies.....</b>	<b>10</b>
2.1. Introducció a la segmentació d'imatges i aplicacions en mamografies.....	10
2.1.1. Què és una mamografia? Quines són les anomalies en una mamografia?.....	10
2.1.2. Què és la segmentació d'imatges?.....	11
2.1.3. Segmentació de mamografies.....	12
2.1.4. Mètodes de segmentació tradicionals.....	12
2.1.5. Avantatges de les xarxes neuronals.....	14
2.1.6. La U-Net: breu introducció.....	14
2.2. Xarxes neuronals i tècniques d'aprenentatge profund.....	14
2.2.1. Què són les xarxes neuronals?.....	15
2.2.2. Algorismes d'aprenentatge profund.....	16
2.2.3. Tipologies de xarxes neuronals.....	17
2.2.4. Tècniques d'optimització.....	17
2.3. Xarxes neuronals convolucionals i la seva aplicació en la segmentació d'imatges....	18
2.3.1. Funcionament de les CNN en la segmentació d'imatges.....	19
2.4. Model U-Net i altres models per a segmentació de mamografies.....	20
2.4.1. FCN: Fully Convolutional Networks.....	20
2.4.2. ResNet.....	21
2.4.3. U-Net.....	21
2.5. Augment de dades per millorar el rendiment dels models.....	22
<b>3. Materials i mètodes.....</b>	<b>24</b>
3.1. Llibreries i frameworks.....	24
3.2. Descripció dels conjunts de dades disponibles.....	26
3.2.1. MIAS.....	27
3.2.2. DDSM.....	27
3.2.3. CBIS-DDSM.....	28
3.2.4. Anàlisi exploratori de les dades CBIS-DDSM.....	29
3.3. Preprocessament de les imatges.....	36
3.3.1. Eliminació de vores.....	36
3.3.2. Eliminació d'artefactes.....	38
3.3.3. Ajust al pit.....	39
3.3.4. CLAHE (Contrast Limited Adaptive Histogram Equalization).....	40
3.3.5. Normalització.....	41

3.4. Descripció de l'arquitectura de la xarxa U-Net.....	41
3.5. Funcions de pèrdua i algorismes d'optimització.....	42
3.6. Descripció de les diferents estratègies d'entrenament.....	44
3.6.1. Entrenament sense augment de dades.....	45
3.6.2. Entrenament amb augment de dades.....	45
<b>4. Experiments i resultats.....</b>	<b>48</b>
4.1. Metodologia d'avaluació dels resultats.....	48
4.2. Resultats obtinguts en els diferents experiments.....	48
4.2.1. Resultats quantitius.....	49
4.2.2. Resultats qualitius.....	51
<b>5. Anàlisi de resultats.....</b>	<b>55</b>
5.1. Comparació de resultats amb treballs relacionats.....	55
5.2. Discussió dels resultats obtinguts i interpretació dels mateixos.....	56
<b>6. Conclusions i treball futur.....</b>	<b>64</b>
6.1. Conclusions generals del treball.....	64
6.2. Limitacions del treball i possibles direccions futures.....	65
<b>7. Referències bibliogràfiques.....</b>	<b>67</b>
<b>8. Annex.....</b>	<b>72</b>

# 1. Introducció

L'augment de la prevalença del càncer de mama a escala mundial ha portat a una major necessitat de desenvolupar mètodes d'avaluació i diagnòstic més precisos i eficaços. La mamografia, un tipus d'imatge mèdica, és una eina crucial en la detecció precoç d'aquest tipus de càncer, ja que permet visualitzar anomalies en les imatges dels teixits mamaris.<sup>1</sup>

La segmentació de mamografies, que consisteix en la separació digital (divisió de la imatge en parts) del teixit mamari d'altres estructures i fons, és una tasca important per a l'anàlisi d'aquestes imatges. En els últims anys, les xarxes neuronals han demostrat ser un mètode efectiu per a la segmentació de mamografies. Una de les arquitectures més utilitzades en aquesta tasca és la U-Net, que ha estat aplicada amb èxit en diferents estudis.<sup>2</sup>

Tanmateix, malgrat els beneficis que ofereixen les xarxes neuronals, encara hi ha reptes a superar en la segmentació de mamografies, com ara la necessitat de dades de qualitat i la dificultat per a la segmentació de lesions de petita mida.<sup>3</sup>

En aquest treball, es proposa la utilització de l'arquitectura U-Net per a la segmentació de mamografies del conjunt de dades públic CBIS-DDSM<sup>4</sup>. Amb l'objectiu d'avaluar l'efectivitat del mètode proposat, es realitzaran diferents experiments utilitzant mètriques d'avaluació com el coeficient de Dice.

A més a més, es discutiran les limitacions dels mètodes tradicionals emprats en la segmentació de mamografies i com la utilització de xarxes neuronals pot ajudar a superar aquestes limitacions.

## 1.1. Contextualització del problema

La detecció precoç del càncer de mama és crucial per a la seva cura i supervivència del pacient, i les mamografies són una eina efectiva per a la seva detecció. No obstant això, la interpretació precisa de les imatges mamogràfiques és un repte, ja que la superposició del teixit mamari pot ocultar tumors o generar falsos positius, el que fa de la mateixa una tasca important que pot millorar la precisió de la detecció de càncer de mama.

---

<sup>1</sup> "Screening for breast cancer with mammography - Cochrane." 4 jun. 2013, [https://www.cochrane.org/CD001877/BREASTCA\\_screening-for-breast-cancer-with-mammography](https://www.cochrane.org/CD001877/BREASTCA_screening-for-breast-cancer-with-mammography). Se consultó el 5 jun. 2023.

<sup>2</sup> "U-Net: Convolutional Networks for Biomedical Image Segmentation." 18 may. 2015, <https://arxiv.org/abs/1505.04597>. Se consultó el 5 jun. 2023.

<sup>3</sup> "Curated Breast Imaging Subset of Digital Database for Screening ...." 21 jun. 2016, <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629>. Se consultó el 5 jun. 2023.

<sup>4</sup> "Curated Breast Imaging Subset of Digital Database for Screening ...." 21 jun. 2016, <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629>. Se consultó el 5 jun. 2023.

La segmentació de mamografies és una tasca crucial en aquesta detecció, ja que ajuda als radiòlegs a localitzar possibles anomalies. Els mètodes tradicionals de segmentació, basats en la segmentació manual (divisió de la imatge per un professional), són laboriosos, subjectius i poden portar a errors humans, cosa que pot comprometre la precisió del diagnòstic.

Els mètodes tradicionals per a la segmentació automàtica de mamografies es basen en algorismes de processament d'imatges que busquen diferències en la intensitat dels píxels per separar el teixit mamari dels possibles tumors.<sup>5</sup> Aquestes tècniques presenten limitacions, ja que les imatges mamogràfiques són molt variables en termes de contrast i resolució, i la segmentació precisa del teixit mamari és una tasca complexa.

Les xarxes neuronals s'han convertit en una eina valuosa per a la segmentació de mamografies, perquè poden aprendre a reconèixer patrons i característiques del teixit mamari a partir de grans quantitats de dades. En particular, l'arquitectura U-Net ha demostrat ser efectiva per a la segmentació de mamografies, perquè permet la integració de característiques d'alta i baixa resolució per a la segmentació precisa del teixit mamari.

Per a l'avaluació de la segmentació, s'utilitzen conjunts de dades rellevants, com el Digital Database for Screening Mammography (DDSM<sup>6</sup>) i el Curated Breast Imaging Subset of DDSM (CBIS-DDSM).<sup>7</sup> Aquests conjunts de dades contenen imatges mamogràfiques amb anotacions de la segmentació realitzats per part dels experts, que es poden utilitzar per a l'entrenament i avaluació de models.

## 1.2. Objectius del treball

El desenvolupament d'un mètode eficaç de segmentació automàtica de mamografies té el potencial d'ajudar als professionals mèdics a detectar de manera més precisa i eficient el càncer de mama, millorant així les taxes de supervivència i la qualitat de vida de les pacients.

L'objectiu principal d'aquest projecte és explorar diferents models de xarxes neuronals profundes (basats en l'arquitectura U-Net) per a la segmentació automàtica de mamografies per millorar els procediments emprats en la detecció precoç del càncer de mama, una de les malalties més comunes i mortals en dones de tot el món.

---

<sup>5</sup> "Effect of image binarization thresholds on breast cancer ...." <https://www.sciencedirect.com/science/article/pii/S2667305321000351>. Se consultó el 5 jun. 2023.

<sup>6</sup> "University of South Florida Digital Mammography Home Page DDSM." <http://www.eng.usf.edu/cvprg/mammography/database.html>. Se consultó el 11 jun. 2023.

<sup>7</sup> "Curated Breast Imaging Subset of Digital Database for Screening ...." 21 jun. 2016, <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629>. Se consultó el 5 jun. 2023.



### 1.3. Estructura del treball

Aquest treball sobre segmentació de mamografies mitjançant tècniques d'aprenentatge profund ha estat estructurat de la següent manera:

En el capítol número dos, *Antecedents i background tècnic*, s'exposen els antecedents i background tècnic del món de la segmentació, i en concret de la segmentació de mamografies. Veurem els conceptes bàsics, introduïrem les xarxes neuronals, les CNN i es parlarà sobre la U-Net, l'arquitectura número u en segmentació mèdica.<sup>8 9</sup>

A continuació trobarem l'apartat número tres, *Materials i Mètodes*. L'objectiu d'aquesta secció és la d'introduir al lector en les eines específiques que s'han utilitzat per a la realització del treball. Aquí es mostraran les tècniques, llibreries o arquitectures que s'han fet servir i les seves característiques.

En la secció número quatre es mostren els resultats obtinguts utilitzant els materials i mètodes comentats anteriorment. Introduïrem els criteris per a l'avaluació dels resultats i els mostrarem separant-los en resultats quantitius i qualitius. Un cop mostrats els resultats obtinguts, farem un anàlisi dels mateixos a la secció cinc. Comentarem quines característiques s'han assolit i quines no, quin és l'origen d'un resultat o d'un altre i quines alternatives o solucions existeixen per a allò que no funciona.

En últim lloc, al capítol número sis, extraurem les conclusions pertinents de l'anàlisi dels resultats, en aquest apartat també comentarem les limitacions trobades en els resultats i en quina direcció es podria portar el treball futur.

---

<sup>8</sup> "U-Net: Convolutional Networks for Biomedical Image Segmentation." 18 may. 2015, <https://arxiv.org/abs/1505.04597>. Se consultó el 5 jun. 2023.

<sup>9</sup> "Grand Challenge: Home." <https://isbi-aida.grand-challenge.org/>. Se consultó el 5 jun. 2023.

## 2. Antecedents i background tècnic sobre la segmentació de mamografies.

La segmentació de mamografies és una tasca important en la detecció precoç del càncer de mama, ja que ajuda els radiòlegs a detectar anomalies en els teixits mamaris. Diferents mètodes han estat desenvolupats per a la segmentació de mamografies en el passat, que inclouen l'ús de mètodes tradicionals com la umbralització i la morfologia matemàtica.<sup>10</sup> Tanmateix, aquests mètodes presenten limitacions, especialment en la segmentació de lesions i estructures anatòmiques detallades.

Amb l'aparició de les xarxes neuronals, la segmentació de mamografies s'ha beneficiat de la seva capacitat per a l'extracció de característiques i la segmentació precisa de les regions d'interès. Entre les arquitectures de xarxes neuronals utilitzades en la segmentació de mamografies, la U-Net és una de les més populars a causa de la seva eficàcia en la segmentació d'estructures detallades i de les lesions.

Per a l'avaluació de la segmentació, s'utilitzen diferents bases de dades públiques, com ara el Digital Database for Screening Mammography (DDSM) i el Curated Breast Imaging Subset of DDSM (CBIS-DDSM). A més, l'ús de dades de qualitat és crucial per a la validesa dels resultats. En aquesta secció també s'exploraran altres mètodes com l'augment de dades i les funcions de pèrdua.

### 2.1. Introducció a la segmentació d'imatges i aplicacions en mamografies

#### 2.1.1. *Què és una mamografia? Quines són les anomalies en una mamografia?*

En aquest apartat es proporciona una breu introducció sobre què és una mamografia i quines són algunes de les anomalies que es poden trobar en aquest tipus d'imatge.

La mamografia és un tipus d'imatge mèdica que utilitza baixes dosis de raigs X amb la finalitat d'obtenir imatges detallades dels pits (veure un exemple a la Figura 1). És una eina clau en la detecció del càncer de mama i d'altres possibles malalties. Les mamografies poden mostrar als experts la presència de masses, calcificacions o altres possibles anomalies en la zona del pit:

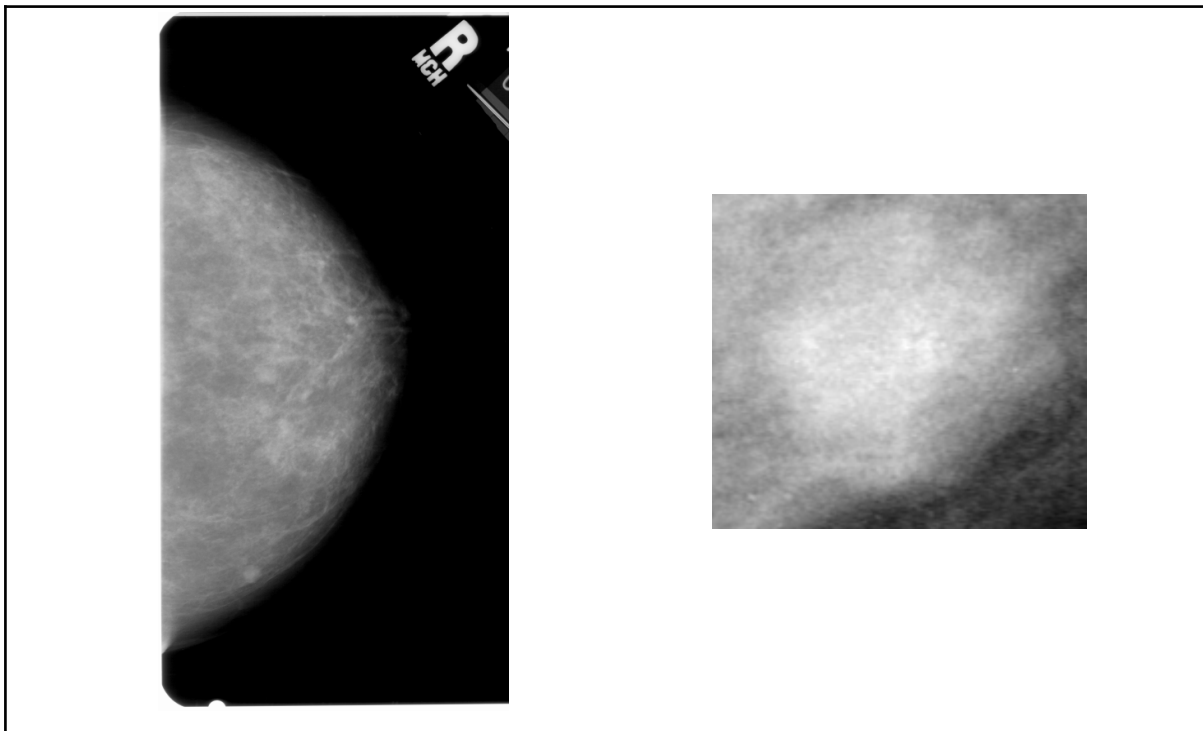
- **Masses:** les masses són zones del teixit de la mama amb una major densitat o aparença diferent de la resta del teixit (vegeu Figura 1). Les masses poden ser

---

<sup>10</sup> "Breast Cancer Segmentation Methods: Current Status and Future ...."  
<https://www.hindawi.com/journals/bmri/2021/9962109/>. Se consultó el 5 jun. 2023.

benignes (no canceroses) o malignes (canceroses). La detecció i la segmentació d'aquestes zones és de gran rellevància a l'hora de realitzar diagnòstic i posterior tractament.

- **Calcificacions:** les calcificacions són petites acumulacions de calci que es formen als pits. Algunes calcificacions poden ser benignes, però d'altres poden ser un indicatiu de la presència de càncer. Aquestes anomalies es poden trobar amb diferents formes, però acostumen a ser nòduls (agrupacions denses de cel·lules) o grups de nòduls, així com entramats (estructura de teixits) complexos.
- **Asimetries i distorsions del teixit:** les asimetries són diferències en l'aparença i densitat del teixit que poden provocar diferents distorsions del pit. Algunes asimetries i distorsions són normals, d'altres poden requerir avaluació addicional per poder ser indicatius de certes malalties.



*Figura 1: Exemple de mamografia, amb la corresponent massa augmentada.*

### *2.1.2. Què és la segmentació d'imatges?*

La segmentació d'imatges és una tasca important en el processament d'imatges, que consisteix a dividir una imatge en diferents regions o segments. Aquests segments són grups de píxels que tenen característiques visuals similars, com ara color, textura o forma. Es tracta d'una tasca complexa i important en molts camps, com ara la visió per computador, la robòtica, la medicina i la indústria del cinema i dels videojocs. En la visió per computador, la segmentació d'imatges es pot utilitzar per a la detecció d'objectes, en la segmentació de la

imatge mèdica per a la identificació de lesions o tumors i en la segmentació de la imatge en la indústria del cinema per a l'edició i la postproducció.

Hi ha diferents mètodes utilitzats a l'hora d'aplicar la segmentació, des de l'ús de filtres i algorismes matemàtics fins a l'aprenentatge automàtic i la intel·ligència artificial. Els mètodes més senzills poden utilitzar la detecció de canvis en la lluminositat o el color, mentre que els mètodes més complexos poden involucrar xarxes neuronals per a la detecció de contorns i formes.

### 2.1.3. Segmentació de mamografies

En una mamografia, la segmentació pot ajudar a detectar les zones anòmales que poden indicar la presència de càncer de mama, com ara les microcalcificacions o les masses. La segmentació també pot ajudar a diferenciar les estructures normals de les anormals en la imatge de la mama, cosa que facilita la tasca dels radiòlegs.

Per tant, la segmentació d'imatges és una tasca important en la detecció precoç del càncer de mama, que pot ajudar a identificar els possibles casos de càncer en una primera fase i millorar les possibilitats de tractament i curació. A més a més, també pot ajudar a reduir el nombre de falsos positius o falsos negatius en les mamografies, cosa que pot millorar la precisió i la confiança en aquesta tècnica de diagnòstic.

### 2.1.4. Mètodes de segmentació tradicionals

Els mètodes tradicionals en la segmentació de mamografies han estat àmpliament utilitzats en el passat i es basen en tècniques matemàtiques com la umbralització, les transformades morfològiques i la segmentació basada en regions. A continuació, descriurem breument cadascun d'aquests mètodes i les seves limitacions:

- **Umbralització**<sup>11</sup>: La umbralització és un mètode senzill i popular en la segmentació d'imatges, incloent-hi les mamografies. Aquesta tècnica consisteix en establir un llindar o umbral que diferenciï les zones de la imatge que es volen segmentar de les que no (Figura 2). Tot i que la umbralització és una tècnica senzilla i ràpida, pot ser sensible a les variacions de lluminositat i contrast, i a vegades no és prou precisa per a la detecció de petites masses o microcalcificacions en la mamografia.
- **Transformades morfològiques**<sup>12</sup>: Les transformades morfològiques són una família de tècniques de processament d'imatges que es basen en l'ús d'operacions

---

<sup>11</sup> "Dynamic multiple thresholding breast boundary detection algorithm ...."

<https://pubmed.ncbi.nlm.nih.gov/20175501/>. Se consultó el 9 abr. 2023.

<sup>12</sup> "Amorphous calcifications (breast) | Radiology Reference Article." 25 sept. 2021,

<https://radiopaedia.org/articles/amorphous-calcifications-breast?lang=us>. Se consultó el 9 abr. 2023.

matemàtiques per a la segmentació. Aquestes operacions s'utilitzen per a la detecció de contorns i per a l'eliminació de soroll en la imatge. Tanmateix, les transformades morfològiques poden ser sensibles a les petites variacions en la textura de la mamografia i poden no ser prou precisos per a la detecció de microcalcificacions.

- **Segmentació basada en regions**<sup>13 14</sup>: La segmentació basada en regions és un mètode que es basa en l'ús d'algoritmes per a la identificació de regions o segments significatius en la imatge. Aquesta tècnica pot ser útil per a la detecció de masses o altres anomalies en la mamografia, però pot ser complicada de configurar i no sempre és prou precisa per a la detecció de microcalcificacions.
- **Aprenentatge automàtic**<sup>15 16</sup>: L'aprenentatge automàtic també ha estat utilitzat amb èxit per a la segmentació d'imatges, utilitzant per exemple les *Support Vector Machines* o SVM. Aquest mètode de segmentació consisteix en tractar cada pixel i els seus veïns com una instància d'entrenament, buscant etiquetar-la com a pertanyent a la regió d'interès o no i maximitzant el marge entre elles.

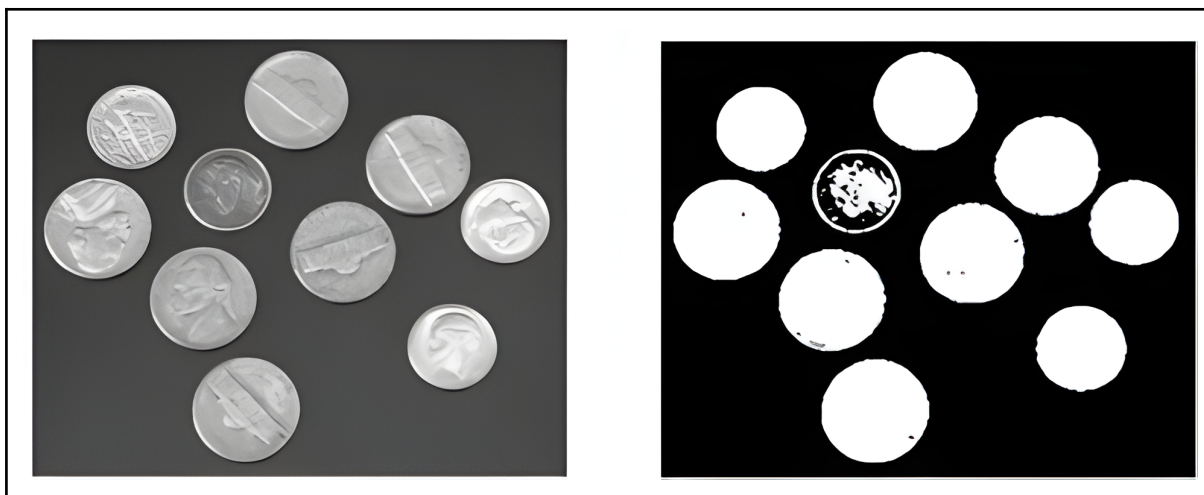


Figura 2: Segmentació realitzada mitjançant umbralització.<sup>17</sup>

Els mètodes tradicionals en la segmentació de mamografies han estat útils en el passat, però poden tenir limitacions en la seva capacitat per a la detecció precisa de petites

<sup>13</sup> "Region-Based Semantic Segmentation with End-to-End Training." 17 sept. 2016, [https://link.springer.com/chapter/10.1007/978-3-319-46448-0\\_23](https://link.springer.com/chapter/10.1007/978-3-319-46448-0_23). Se consultó el 5 jun. 2023.

<sup>14</sup> "A new conditional region growing approach for microcalcification ...." 24 jul. 2021, <https://link.springer.com/article/10.1007/s11517-021-02379-x>. Se consultó el 5 jun. 2023.

<sup>15</sup> "Image segmentation based on Support Vector Machine - IEEE Xplore." <http://ieeexplore.ieee.org/document/6525921/>. Se consultó el 12 jun. 2023.

<sup>16</sup> "SVM for Image Segmentation? - Cross Validated - Stack Exchange." 25 abr. 2012, <https://stats.stackexchange.com/questions/27085/svm-for-image-segmentation>. Se consultó el 12 jun. 2023.

<sup>17</sup> "Thresholding in Image Processing - YouTube." 14 jun. 2020, <https://www.youtube.com/watch?v=vtbdqq7yAcc>. Se consultó el 5 jun. 2023.

anomalies. Actualment, les tècniques més avançades basades en l'aprenentatge automàtic i la intel·ligència artificial són més precises i efectives.

### *2.1.5. Avantatges de les xarxes neuronals*

Les xarxes neuronals són un tipus d'algorisme de l'aprenentatge automàtic que imiten el funcionament del cervell humà per a la resolució de problemes complexos. A diferència dels mètodes tradicionals de segmentació d'imatges, les xarxes neuronals poden adaptar-se de forma automàtica a les característiques de les imatges i aprendre a partir de grans quantitats de dades.

En la segmentació de mamografies, les xarxes neuronals poden ajudar a detectar anomalies com microcalcificacions i masses amb major precisió que els mètodes tradicionals. Això es deu a la seva capacitat per a l'aprenentatge profund (vegeu apartat 2.2), que els permet aprendre característiques ocultes en les imatges que els mètodes tradicionals no poden detectar.

Els avantatges de les xarxes neuronals també inclouen la seva capacitat per a la segmentació d'imatges en temps real i la seva adaptabilitat a diferents tipus d'imatges i situacions clíniques, fent-les ideals per a la detecció de les anomalies amb major precisió.

### *2.1.6. La U-Net: breu introducció*

La U-Net<sup>18</sup> és una arquitectura de xarxa neuronal convolucional que s'ha utilitzat amb èxit en la segmentació d'imatges, especialment en el camp de la detecció precoç del càncer de mama. La seva estructura simètrica en forma de U és el que li dona el nom. La U-Net utilitza una tècnica anomenada "skip connections" per a la detecció d'estructures més petites i proporciona una segmentació més precisa. Això permet la detecció de detalls fins i tot en regions amb pocs píxels. Es tracta d'una arquitectura molt eficaç en la segmentació d'imatges mèdiques i ha estat utilitzada amb èxit en molts estudis anteriors.

Informació corresponent a l'arquitectura específica utilitzada en aquest treball la trobarem a l'apartat 3.3, "Descripció de l'arquitectura de la xarxa U-Net".

## **2.2. Xarxes neuronals i tècniques d'aprenentatge profund**

L'aprenentatge profund ha revolucionat el camp de la visió per computador en els últims anys. Les xarxes neuronals són capaces d'aprendre de manera automàtica patrons a

---

<sup>18</sup> "U-Net: Convolutional Networks for Biomedical Image Segmentation." 18 may. 2015, <https://arxiv.org/abs/1505.04597>. Se consultó el 5 jun. 2023.

partir de dades i d'utilitzar aquest coneixement per a resoldre tasques complexes, com ara la segmentació d'imatges. En aquest apartat es descriuen les bases teòriques de les xarxes neuronals i les tècniques d'aprenentatge profund més utilitzades en la segmentació d'imatges mèdiques. També s'aborden les seves limitacions i les tècniques per a mitigar-les, així com les diferents arquitectures de xarxes neuronals que s'han utilitzat en la segmentació de mamografies, entre elles la U-Net.

### 2.2.1. Què són les xarxes neuronals?

Les xarxes neuronals són models matemàtics complexos que s'utilitzen en moltes aplicacions d'aprenentatge automàtic i visió per computador. Una xarxa neuronal està composta per capes de neurones interconnectades, cadascuna de les quals processa la informació rebuda de la capa anterior per produir una sortida (vegeu Figura 3). Els tipus de capes més comuns en les xarxes neuronals són les capes convolucionals, de rectificació lineal (ReLU), de submostreig i totalment connectades:

- Les **capes convolucionals** s'utilitzen per a la detecció de característiques d'una imatge. Utilitzen filtres que apliquen una operació de convolució a través de la imatge per a detectar diferents característiques en diferents regions de la imatge.
- Les **capes ReLU** s'utilitzen per a introduir la no-linealitat en la xarxa neuronal, cosa que és important per a aprendre patrons complexos a partir de les dades.
- Les **capes de submostreig** s'utilitzen per a reduir la dimensionalitat de les dades i simplificar el model.
- Les **capes totalment connectades** s'utilitzen per a realitzar la classificació final.

En general, una xarxa neuronal es configura connectant diferents capes entre si de manera que la sortida d'una capa sigui l'entrada de la capa següent, fins a la capa final que produeix la sortida desitjada.<sup>19</sup>

---

<sup>19</sup> "Understanding Neural Networks - Prince Canuma." 13 feb. 2019, <https://prince-canuma.medium.com/understanding-neural-networks-22b29755abd9>. Se consultó el 5 jun. 2023.

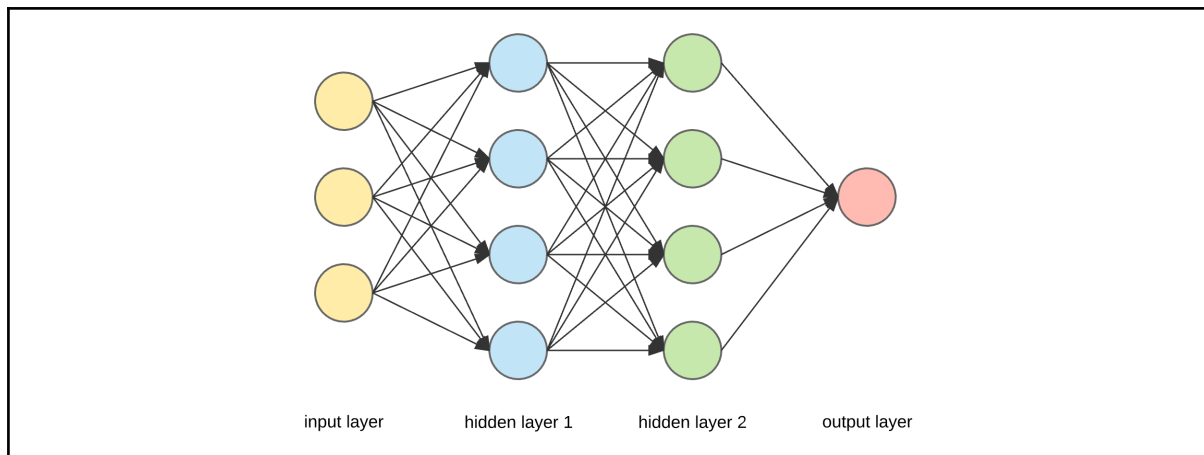


Figura 3: Xarxa neuronal amb la capa d'entrada, dues capes ocultes i una capa de sortida.<sup>20</sup>

### 2.2.2. Algorismes d'aprenentatge profund

Els algorismes d'aprenentatge profund són una classe d'algorismes d'aprenentatge automàtic que són capaços d'aprendre representacions de dades de gran dimensionalitat a través de la utilització de xarxes neuronals profundes.

Aquests algorismes són capaços de resoldre problemes molt complexos, com ara la detecció d'objectes en imatges, el reconeixement de veu i la traducció automàtica de llenguatges naturals. Per entendre el funcionament dels algorismes d'aprenentatge profund, és important primer entendre els conceptes de funció d'activació i propagació de l'error:

- La **funció d'activació** és una funció matemàtica que s'aplica a la sortida d'una capa de neurones per determinar la seva activació, que a continuació es transmet a la capa següent.
- La **propagació de l'error** és el procés pel qual l'error de la sortida de la xarxa es propaga cap enrere a través de les capes per ajustar els pesos de les connexions i minimitzar l'error.

La xarxa neuronal es compon de diferents capes, cada una de les quals té un nombre determinat de neurones. La capa d'entrada de la xarxa rep les dades d'entrada i les transmet a la següent capa, anomenada capa oculta. Aquestes capes ocultes són capaces d'aprendre característiques cada vegada més abstractes de les dades, a través de l'aplicació d'una funció d'activació i la propagació de l'error. Finalment, la capa de sortida de la xarxa emet una sortida que es compara amb la sortida desitjada per calcular l'error i ajustar els pesos de les connexions.

<sup>20</sup> "Understanding Neural Networks - Prince Canuma." 13 feb. 2019, <https://prince-canuma.medium.com/understanding-neural-networks-22b29755abd9>. Se consultó el 5 jun. 2023.



### 2.2.3. Tipologies de xarxes neuronals

Hi ha diferents tipologies de xarxes neuronals, i cada tipus té una estructura i un funcionament únic:

- Les **xarxes feedforward** són un dels tipus més comuns de xarxes neuronals. En aquest tipus de xarxes, les dades es mouen en una direcció, passant per capes successives de neurones que són totalment connectades. Cada neurona en una capa dona com a resultat una sortida que és passada a totes les neurones de la capa següent. Això es repeteix fins que es produeix una sortida final. Les xarxes feedforward són excel·lents per a tasques de classificació, com la identificació d'objectes en imatges.
- Les **xarxes recurrents** són xarxes neuronals que contenen bucles que permeten als neurones comunicar-se amb capes anteriors. Això significa que les xarxes recurrents poden utilitzar la informació de les dades anteriors per fer prediccions més precises. Aquest tipus de xarxes es poden utilitzar en tasques de processament de llenguatge natural, com la traducció automàtica, on les paraules anteriors d'una frase influeixen en el significat de les paraules següents.
- Les **xarxes convolucional**s són una tipologia especial de xarxes neuronals que són específicament dissenyades per a la tasca de processament d'imatges. Aquest tipus de xarxes estan formades per capes que detecten característiques en una imatge i les utilitzen per fer prediccions. Les xarxes convolucional són molt útils en la segmentació d'imatges, una tasca important en la detecció de càncer de mama en mamografies.

Cada tipologia té una estructura i un funcionament diferents. Les xarxes feedforward són excel·lents per a tasques de classificació, les xarxes recurrents són excel·lents per a tasques de processament de llenguatge natural, i les xarxes convolucional són excel·lents per a tasques de processament d'imatges com la segmentació, l'objectiu d'aquest treball.

### 2.2.4. Tècniques d'optimització

Per aconseguir que les xarxes neuronals puguin aprendre de forma efectiva, són necessàries tècniques d'optimització adequades. Hi ha diverses tècniques d'optimització que es poden utilitzar per millorar el rendiment dels models de xarxes neuronals.

L'algorisme de retropropagació és una tècnica d'optimització comuna en les xarxes neuronals. Consisteix a passar la informació de la sortida de la xarxa cap a la seva entrada, propagant l'error i ajustant els pesos de les connexions de la xarxa. Això permet que la xarxa aprengui a partir de les dades d'entrada i ajusti els seus paràmetres per aconseguir una sortida òptima.

Els mètodes d'optimització basats en gradient també són molt comuns en les xarxes neuronals. Aquest tipus de mètodes utilitzen la derivada de la funció d'error per ajustar els pesos de la xarxa. La idea és que es van actualitzant els pesos en la direcció del gradient negatiu de la funció d'error, de manera que es minimitza aquesta funció. Algunes de les variants més populars són el mètode del gradient descendent estocàstic (SGD, veure Figura 4), el mètode d'Adam i el mètode d'Adagrad.

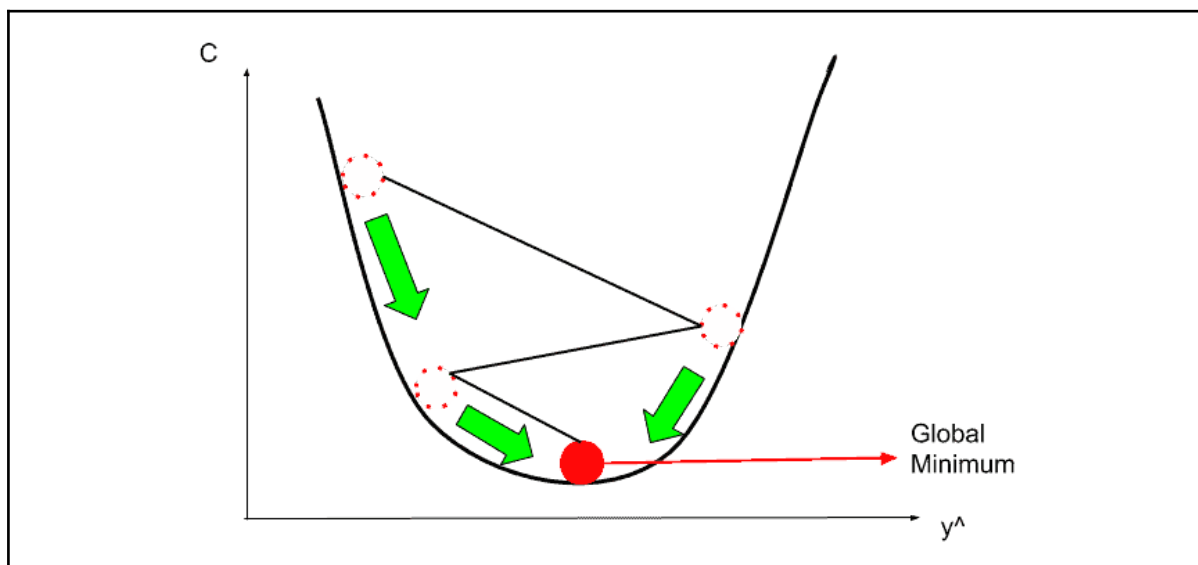


Figura 4: Visualització del procés seguit pel SGD per trobar un mínim.

Altres tècniques d'optimització comunes inclouen l'ús de tècniques de regularització, com la regularització L1 i L2, que ajuden a prevenir l'overfitting del model. També es poden utilitzar tècniques d'aprenentatge per transferència per transferir el coneixement d'una xarxa entrenada prèviament en una tasca relacionada a la tasca actual. L'elecció de la tècnica d'optimització adequada pot ser crítica per obtenir un model de xarxa amb un bon rendiment en la tasca que es vol resoldre.

### 2.3. Xarxes neuronals convolucionals i la seva aplicació en la segmentació d'imatges

Les xarxes neuronals convolucionals (CNN<sup>21</sup>) són un tipus de xarxa neuronal que ha demostrat ser molt efectiva en la tasca de segmentació d'imatges. Les CNN es basen en la idea que les característiques visuals importants es poden identificar i aprendre a través de la

<sup>21</sup> "Convolutional neural networks for medical image segmentation - arXiv." 17 nov. 2022, <https://arxiv.org/abs/2211.09562>. Se consultó el 12 jun. 2023.

convolució (Figura 5), que és una operació matemàtica que aplica un filtre a una imatge per obtenir una nova imatge.

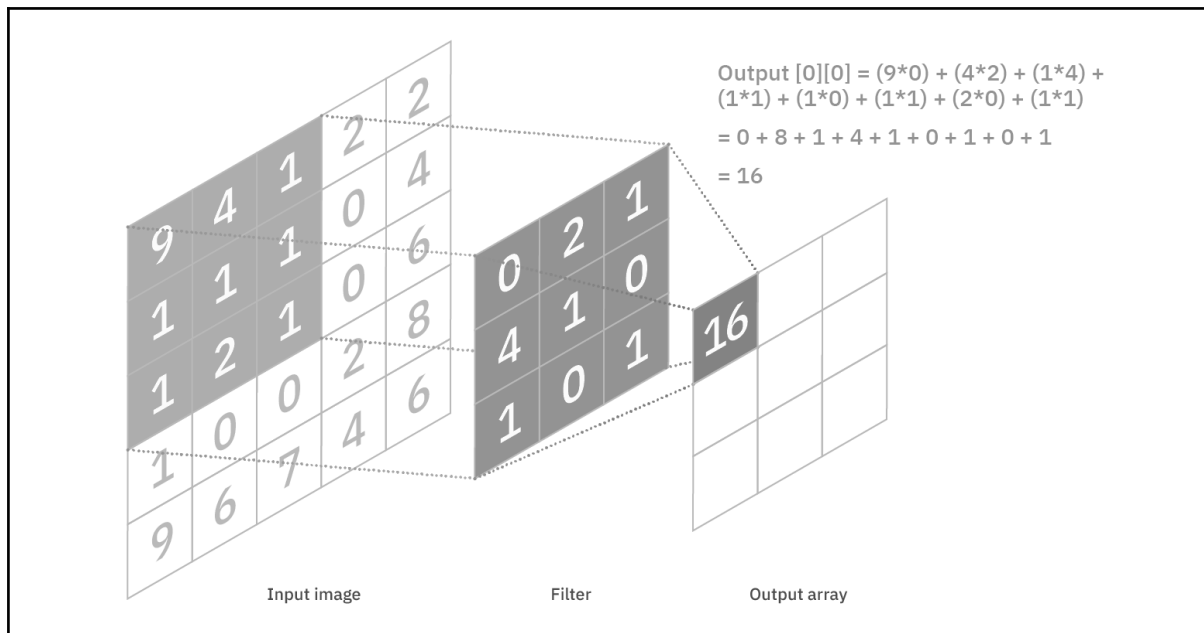


Figura 5: Exemple de convolució.

### 2.3.1. Funcionament de les CNN en la segmentació d'imatges

Les xarxes neuronals convolucionals (CNN) són una classe d'algorismes d'aprenentatge profund que són especialment útils en tasques de segmentació d'imatges i són capaços de processar la informació de forma localitzada. Les CNN processen les imatges en diferents capes, que són capaces d'aprendre característiques cada vegada més complexes a mesura que es profunditza en la xarxa. Això es fa mitjançant capes convolucionals, que consisteixen en una sèrie de filtres que s'apliquen a l'entrada per obtenir un mapa de característiques.

El primer pas en la segmentació d'imatges amb CNN és la preparació de les dades d'entrada. Això pot incloure la normalització de les imatges per garantir que estiguin en un rang de valors específic i la segmentació manual de les imatges d'entrenament per crear les màscares de referència.

A continuació, es construeix la xarxa, que consisteix en diferents capes de convolució, activació i agrupació. A mesura que la informació passa per la xarxa, els filtres s'apliquen a l'entrada per detectar característiques específiques de la imatge. Les activacions s'apliquen després de les capes de convolució per introduir una no-linealitat a la xarxa. Les capes de max-pooling o de "reducció" s'utilitzen per reduir la dimensionalitat de la imatge a mesura que es profunditza en la xarxa. Això permet aprendre característiques cada vegada més abstractes i complexes a conforme es processa la informació.

Finalment, la xarxa produeix una sortida en forma de màscara que indica la segmentació de la imatge d'entrada. Això es fa mitjançant capes de convolució amb un únic canal de sortida, que produeixen una matriu de puntuacions que s'interpreta com una màscara.

Les CNN són capaces de segmentar imatges amb una gran precisió i són capaces d'aprendre característiques de forma automàtica i adaptativa. Això fa que siguin una eina potent per a la segmentació d'imatges en molts àmbits, inclosa la medicina.

## **2.4. Model U-Net i altres models per a segmentació de mamografies**

L'arquitectura U-Net és una xarxa neuronal convolucional (CNN) que es va desenvolupar en primera instància amb la finalitat de resoldre problemes de segmentació d'imatges mèdiques. Aquest model va ser creat per Olaf Ronneberger, Philipp Fischer i Thomas Brox l'any 2015, i va ser publicat per primer cop al paper "U-Net: Convolutional Networks for Biomedical Image Segmentation".<sup>22</sup> L'origen i motivació de la seva creació va ser la de millorar les FCN, descrites anteriorment per Evan Shelhamer, Jonathan Long i Trevor Darrell<sup>23</sup>.

### *2.4.1. FCN: Fully Convolutional Networks*

Quan es van presentar per primer lloc les "Fully Convolutional Networks" (FCN) es va poder demostrar que sobrepassen a les tècniques anteriors en la segmentació semàntica. L'objectiu principal dels seus desenvolupadors era construir xarxes capaces de prendre entrades de mida arbitrària i retornar una sortida de la mateixa mida amb un procés d'aprenentatge eficient, tant en inferència com en nivell d'aprenentatge. A diferència de les xarxes neuronals convencionals que estan dissenyades per a tasques de classificació, les FCN són capaces de produir mapes de píxels segmentats de manera eficient.

Les FCN (com indica el seu nom), estan compostes només per capes convolucionals i de pooling. És precisament això el que permet l'entrada d'imatges de mida arbitrària, no com en altres xarxes que requereixen dimensions d'entrada fixes. La segona característica important és l'ús de convolucions transposades o interpolació bilineal en les operacions d'augment, el que permet una segmentació fina i que conserva els detalls i informació.

Pel que fa al seu ús en la segmentació de mamografies, les FCN van suposar un gran avanç en aquesta tasca, ja que, com en la resta de casos d'aplicació de segmentació, permeten

---

<sup>22</sup> "U-Net: Convolutional Networks for Biomedical Image Segmentation." 18 may. 2015, <https://arxiv.org/abs/1505.04597>. Se consultó el 22 may. 2023.

<sup>23</sup> "Fully Convolutional Networks for Semantic Segmentation - arXiv." 14 nov. 2014, <https://arxiv.org/abs/1411.4038>. Se consultó el 22 may. 2023.

detectar àrees sospitoses de ser tumors o d'altres anormalitats de forma més precisa que models i arquitectures anteriors.

#### 2.4.2. ResNet

ResNet és l'abreviatura de "Residual Network", un tipus de xarxa CNN coneguda per poder ser molt profunda sense patir problemes amb el descens del gradient. Va ser creada l'any 2015 per Kaiming He, Xiangyu Zhang, Shaoqing Ren, i Jian Sun<sup>24</sup> i la idea principal és l'ús de blocs residuals, que permeten fer salts de connexions per tal de saltar capes. Aquestes connexions son conegudes com a connexions *skip*.

Aquestes connexions consisteixen en salts entre capes, en comptes de propagar de forma típica l'entrada a través de les mateixes. El seu nom, connexió *skip* o connexió *residual*, ve de què es transporta la informació original de l'entrada fins a la sortida del bloc residual (capa de sortida). El resultat d'això és que en comptes d'aprendre's la transformació completa, només s'aprèn la diferència entre l'entrada i la sortida de les capes desitjades.

El resultat és una millor optimització, i permeten la creació de xarxes més profundes, on es poden obtenir millors resultats, més precisos. Els experiments realitzats en tasques de visió artificial amb xarxes ResNet per a classificació o segmentació han demostrat un rendiment alt, sobretot amb conjunts de dades petits o de característiques complexes.

#### 2.4.3. U-Net

Com ja s'ha comentat anteriorment, la U-Net<sup>25</sup> es una CNN utilitzada per a la segmentació d'imatges, i el seu origen es la cerca de la millora de les FCN. Des de la seva aparició es el marc de referència en models de segmentació d'imatges mèdiques.

El seu nom deriva de la seva arquitectura en forma de U. Aquesta arquitectura es caracteritza per tenir dues parts principals: un encoder i un decoder. L'encoder consta de múltiples capes convolucionals i de pooling, la funció de les quals és la d'extreure les característiques i reduir la dimensió dels nivells anteriors. El decoder està format per capes d'upsampling (ampliació de la resolució de la imatge) i concatenació. A la Figura 6 es pot veure l'estructura proposada originalment.

---

<sup>24</sup> "[1512.03385] Deep Residual Learning for Image Recognition - arXiv." 10 dic. 2015, <https://arxiv.org/abs/1512.03385>. Se consultó el 22 may. 2023.

<sup>25</sup> "U-Net: Convolutional Networks for Biomedical Image Segmentation." 18 may. 2015, <https://paperswithcode.com/paper/u-net-convolutional-networks-for-biomedical>. Se consultó el 5 jun. 2023.

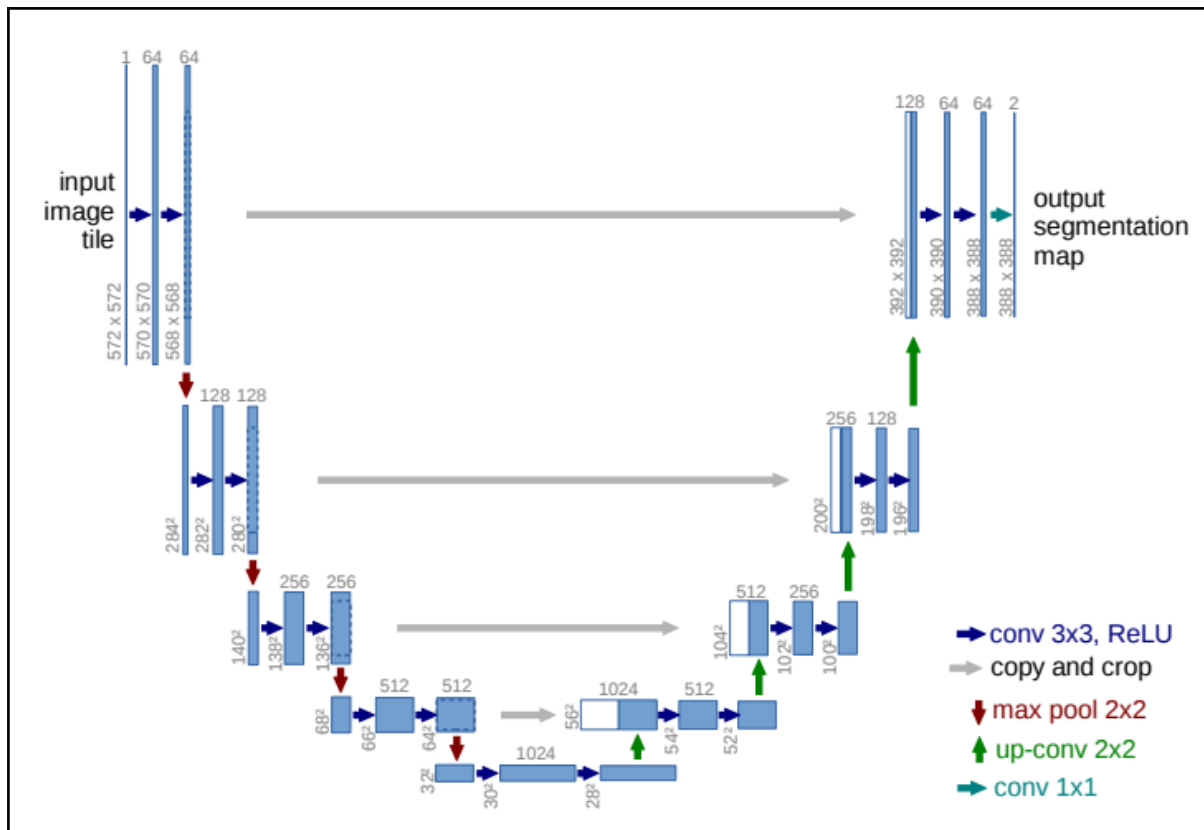


Figura 6: Arquitectura U-Net com va ser presentada al paper original.<sup>26</sup>

La característica principal de la U-Net és l'ús de connexions *skip*<sup>27</sup> de forma similar a la ResNet, que permeten recuperar informació de capes prèvies de encoding a la part del decoder. Això permet alterar el camí del gradient quan entrenem la xarxa, i ha demostrat ser efectiva a l'hora de millorar la segmentació de les imatges.

## 2.5. Augment de dades per millorar el rendiment dels models

L'augment de dades es una part fonamental en l'entrenament de models de segmentació. El seu us permet augmentar la varietat i el nombre de mostres d'exemple que podem mostrar i usar per a l'entrenament de la xarxa. El seu objectiu és el de millorar la capacitat de generalització del model presentant nous casos fent ús de diverses manipulacions de les dades originals. Algunes d'aquestes manipulacions poden ser rotacions, desplaçaments, deformacions i canvis en el contrast.<sup>28</sup>

<sup>26</sup> "U-Net: Convolutional Networks for Biomedical Image Segmentation." 18 may. 2015, <https://arxiv.org/abs/1505.04597>. Se consultó el 5 jun. 2023.

<sup>27</sup> "Intuitive Explanation of Skip Connections in Deep Learning." 23 mar. 2020, <https://theaisummer.com/skip-connections/>. Se consultó el 22 may. 2023.

<sup>28</sup> "Data Augmentation | Papers With Code." <https://paperswithcode.com/task/data-augmentation>. Se consultó el 12 jun. 2023.

En aquest apartat s'exposen certes tècniques d'augment emprades normalment en la segmentació. El seu ús no és exclusiu, normalment s'apliquen diverses tècniques de forma conjunta. La seva selecció dependrà del context, tipus de problema, característiques de les dades i criteri del programador:

### ***Desplaçaments***

Aquesta tècnica consisteix a desplaçar una imatge original en diverses direccions, variant la posició de les regions d'interès i modificant la seva localització espacial.

### ***Augments de mida***

Els augments de mida permeten mostrar a la xarxa diferents canvis d'escala a partir de dades originals, ajudant al model a comprendre millor les variacions de mida.

### ***Mirall i volteig***

Aplicant un efecte de mirall o de volteig, podem generar noves dades reflectides de forma horitzontal o vertical, també aplicant rotacions de diversos angles. D'aquesta forma augmentem el nombre de variacions geomètriques.

### ***Deformacions elàstiques***

La tècnica de deformacions elàstiques permet aplicar deformacions de forma local a les imatges. Mitjançant aquest procés es poden simular deformacions pel que fa a estructura, no només de perspectiva.

### ***Canvis de brillantor i contrast***

Aplicant canvis de contrast i brillantor a les imatges podem augmentar la robustesa del model davant de diferents situacions d'il·luminació, així com la robustesa a la representació de textures depenent del contrast existent.

### 3. Materials i mètodes

En aquest apartat es proporciona una descripció detallada de les eines, recursos i estratègies utilitzades en aquest treball.

En primer lloc, veurem quin llenguatge i llibreries s'han fet servir. Les llibreries són eines fonamentals per al desenvolupament del treball, ja que proporcionen funcions, components i estructures que faciliten la creació dels models de xarxes neuronals així com faciliten el tractament de les dades.

En segon lloc, es descriuen els conjunts de dades implicats en l'entrenament i avaluació del model de segmentació. En concret es parla de dos coneguts conjunts de dades de mamografies: el DDSM<sup>29</sup> i el CBIS-DDSM<sup>30</sup>. Es detallen les seves característiques, les dades que contenen i el seu origen.

Seguidament, es presenta el procés de preprocessament realitzat sobre les imatges utilitzades. S'inclouen les transformacions i processos seguits de forma prèvia a l'entrenament: cropping, eliminació d'artefactes, escalats, normalitzacions i ajust del contrast. Aquest preprocessament té com a objectiu augmentar la qualitat i la consistència de les imatges.

A continuació es detallen els components i estructura de la xarxa U-Net emprada en el treball. En concret es parla de les capes, filtres, connexions i altres característiques rellevants, com les funcions de pèrdua o optimitzadors.

Per últim, es parla sobre l'entrenament amb i sense augment de dades, amb l'objectiu de preveure els resultats que obtindrem amb els experiments.

#### 3.1. Llibreries i frameworks

Els models implementats, així com la resta de codi han sigut implementats utilitzant Python (versió 3.9), ja que existeixen diferents llibreries i frameworks per a la implementació de models de xarxes neuronals. En aquest text, es detallen algunes de les llibreries i frameworks més populars, com TensorFlow, PyTorch i Keras, incloent-hi les seves característiques i diferències:

---

<sup>29</sup> "University of South Florida Digital Mammography Home Page DDSM." <http://www.eng.usf.edu/cvprg/mammography/database.html>. Se consultó el 22 may. 2023.

<sup>30</sup> "Curated Breast Imaging Subset of Digital Database for Screening ...." 16 may. 2023, <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629>. Se consultó el 22 may. 2023.



- **TensorFlow**<sup>31</sup> (versió 2.11.0)

És una llibreria d'aprenentatge automàtic desenvolupada per Google Brain Team que es va llançar el 2015. És una llibreria de codi obert i suporta diferents llenguatges de programació, com ara Python, C++ i Java. TensorFlow té la capacitat de crear xarxes neuronals complexes i és amplament utilitzada en aplicacions d'aprenentatge profund, visió per computador i processament de llenguatge natural. TensorFlow ofereix una API molt àmplia que permet als desenvolupadors personalitzar els seus models i ajustar-los a les seves necessitats.

- **Keras**<sup>32</sup> (versió 2.11.0)

És una API d'aprenentatge profund de codi obert que s'executa sobre TensorFlow. Els models de Keras es defineixen en capes que es poden connectar de forma seqüencial o no seqüencial per crear xarxes neuronals. Keras ofereix una sintaxi simple i fàcil d'utilitzar, el que fa que sigui una eina popular per als desenvolupadors que volen crear prototips ràpids de models de xarxes neuronals. Keras també ofereix la possibilitat d'integrar-se amb altres llibreries i frameworks d'aprenentatge profund.

- **Keras-Unet**<sup>33</sup> (versió 0.1.2) i **visualker**<sup>34</sup> (versió 0.0.2)

Són dues llibreries que han resultat vitals en la realització del present treball. La primera és una encapsulació de diverses implementacions de U-Net, oferint també altres funcions d'utilitat per a la generació de noves imatges i de visualització. Visualker ha sigut utilitzada per a la visualització de l'arquitectura del model de segmentació creat, altament flexible i amb bons resultats.

- **Pandas**<sup>35</sup> (versió 1.5.3)

És una llibreria de codi obert per a manipular i analitzar dades. Aquest ha sigut el seu ús, ja que ha resultat de gran utilitat a l'hora d'analitzar els resultats obtinguts així com el contingut del conjunt de dades total.

---

<sup>31</sup> "TensorFlow.org." <https://www.tensorflow.org/>. Se consultó el 5 jun. 2023.

<sup>32</sup> "Keras: Deep Learning for humans." <https://keras.io/>. Se consultó el 5 jun. 2023.

<sup>33</sup> "karolzak/keras-unet: Helper package with multiple U-Net ... - GitHub." <https://github.com/karolzak/keras-unet>. Se consultó el 5 jun. 2023.

<sup>34</sup> "visualker for Keras / TensorFlow - GitHub." <https://github.com/paulgavrikov/visualker>. Se consultó el 5 jun. 2023.

<sup>35</sup> "pandas - Python Data Analysis Library." <https://pandas.pydata.org/>. Se consultó el 13 jun. 2023.

L'execució del codi s'ha realitzat tant localment com al núvol. La manca dels recursos necessaris a la màquina local ha fet necessari l'ús de recursos com Kaggle<sup>36</sup> per tal de poder fer ús de més memòria i GPU més potent. Kaggle es una plataforma en línia que permet accedir a conjunts de dades públics i participar en la resolució de problemes d'aprenentatge automàtic. A través de Kaggle es pot accedir a recursos computacionals de gran capacitat, el que permet executar certs algorismes que els requereixen sense necessitat de tenir els recursos de forma local.

### 3.2. Descripció dels conjunts de dades disponibles

Per tal d'obtenir resultats òptims en la tasca de segmentació, com en gairebé qualsevol problema d'aprenentatge profund, és fonamental la tria del conjunt de dades a partir del qual treballarem. De la seva selecció depèn el rendiment de l'entrenament, així com de la validació dels models creats.

En el món de les mamografies trobem diversos conjunts de dades que ja han sigut utilitzats en un passat per a la tasca de segmentació, però també per a problemes de classificació o altres.

Els sistemes de detecció assistida per ordinador (CADE) i diagnòstic (CADx) són sistemes l'objectiu dels quals és assistir als radiòlegs professionals en la interpretació de mamografies i altres imatges. Un problema comú en els CADE ha sigut la seva limitació per l'alta proporció de falsos positius que generen. No podem obviar el fet que bona part d'aquesta limitació sigui causada per la tria de les dades d'avaluació i, en el cas de l'entrenament de models, de la tria de dades d'entrenament.<sup>37</sup>

Existeixen un gran nombre de bancs d'imatges d'ús obert en el món no mèdic, per mencionar només un parell d'exemples: MNIST (70k imatges de números escrits a mà) i ImageNet (més de 14M d'imatges d'animals, menjar, cotxes...). Aquest caràcter obert ha permès que els conjunts de dades siguin altament accessibles, i han esdevingut una mena d'estàndard que tothom pot utilitzar per a comparar tècniques i/o models amb les mateixes dades.

Per contra, això no ha sigut sempre el cas en l'àmbit de la imatge mèdica. L'alt cost, en temps i recursos, que requereix l'obtenció de certes imatges, així com el limitat accés a pacients i casos diversos, han fet que l'existència d'un conjunt de dades estandarditzat hagi arribat més tard.

---

<sup>36</sup> "Kaggle: Your Machine Learning and Data Science Community." <https://www.kaggle.com/>. Se consultó el 12 jun. 2023.

<sup>37</sup> "A curated mammography data set for use in computer ... - Nature." 19 dic. 2017, <https://www.nature.com/articles/sdata2017177>. Se consultó el 27 may. 2023.

En els següents apartats s'introdueixen un parell dels conjunts de dades més coneguts, MIAS i DDSM, i per últim parlarem sobre CBIS-DDSM, un conjunt de dades creat amb l'objectiu de trencar les limitacions anteriorment mencionades.

### 3.2.1. MIAS

El conjunt MIAS<sup>38</sup> (Mammographic Image Analysis Society) es una de les grans referents en l'àmbit de les mamografies. Conté un total de 322 imatges obtingudes de 161 pacients. Es troben en escala de grisos i tenen una resolució de 1024 x 1024 píxels. Cada imatge compta amb una màscara que ens proporciona la informació necessària sobre les possibles regions d'interès presents en la mamografia. MIAS també proporciona dades sobre l'edat del pacient, historial i altra informació clínica que pot ser rellevant segons la tasca a realitzar.

MIAS va ser el primer conjunt de dades que es va considerar en la realització d'aquest treball, però la seva limitada quantitat d'imatges va motivar la cerca d'altres opcions. Conseqüència de la seva mida, és la falta de diversitat en els casos que conté, el que podria motivar una pèrdua de generalització a l'hora d'entrenar un model de segmentació. També cal dir que MIAS és un conjunt de dades antic, i per tant les imatges que conté poden no reflectir la millora tecnològica en l'adquisició de les imatges.

En la cerca d'un conjunt de dades més complet, amb més imatges i més modern, és on trobem CBIS i DBIS-DDSM. Tots dos són usats comunment en el món de la investigació contra el càncer de mama i, per tant, han esdevingut una mena d'estàndard a l'hora d'avaluar i comparar treballs d'investigació.

### 3.2.2. DDSM

El conjunt DDSM (Curated Breast Imaging Subset) és un conjunt de dades públic creat pel NCI<sup>39</sup> i el NIH<sup>40</sup> als Estats Units d'Amèrica. Conté imatges de càncer de mama de diferents tipus, com mamografies digitals i d'ultrasons. Aquest fet permet per exemple comparar l'efectivitat de diferents tècniques de detecció depenent del tipus de metodologia de presa d'imatges.

El conjunt conté mamografies de caràcter benigne i maligne. Per tal de poder saber de quin tipus, s'inclouen anotacions sobre les característiques de cada imatge i pacient. Aquestes anotacions inclouen l'existència de màscares a escala de píxel sobre les diferents masses o

---

<sup>38</sup> "Mammographic Image Analysis Society (MIAS) database v1.21." 28 ago. 2015, <https://www.repository.cam.ac.uk/handle/1810/250394>. Se consultó el 26 may. 2023.

<sup>39</sup> "Comprehensive Cancer Information - NCI." <https://www.cancer.gov/>. Se consultó el 26 may. 2023.

<sup>40</sup> "National Institutes of Health (NIH) | Turning Discovery Into Health." <https://www.nih.gov/>. Se consultó el 26 may. 2023.

calcificacions que puguin existir en les imatges. Aquest nivell d'informació permet desenvolupar algoritmes de detecció i classificació, com per exemple un model de segmentació.

La mida del conjunt DDSM és de 2620 casos, amb les seves imatges i anotacions. Es pot obtenir lliurement des de les pàgines de descàrrega de NIH i NCI.

### 3.2.3. *CBIS-DDSM*

El conjunt CBIS-DDSM<sup>41</sup> (Curated Breast Imaging Subset of DDSM) és un altre conjunt de dades disponible públicament i creat pel NIH i el NCI. Aquest conjunt de dades es centra en la mamografia digital, i conté informació com la màscara de regió d'interès i informació clínica del pacient (de forma similar a DDSM).

Com indica el seu nom, és una versió actualitzada i estandarditzada del conjunt DDSM. Això significa que les dades presents en CBIS-DDSM han sigut seleccionades i curades per un expert en mamografies.

Conté mamografies digitalitzades, un gran nombre d'estudis/casos (2620) i descripcions detallades sobre el pacient, el seu historial i característiques específiques de les anormalitats. Conté casos normals (sense anormalitats), benignes i malignes, amb informació patològica confirmada. La seva mida conjuntament amb les màscares el converteix en una eina de gran utilitat per a desenvolupament de noves tècniques de diagnòstic, així com a estàndard per a contrastar resultats entre diferents proves<sup>42</sup>.

Precisament va ser l'estandardització el detonant per a la creació de CBIS-DDSM, ja que resultava complex replicar experiments per part de tercers utilitzant conjunts de dades que en molts casos eren privats. MIAS i DDSM són dels pocs conjunts públics que existien prèviament a la creació del CBIS-DDSM, però com hem pogut veure els principals problemes que tenen és la limitada mida d'aquests. CBIS-DDSM resol aquest problema i és per aquest motiu que ha sigut el conjunt de dades seleccionat per a la realització d'aquest treball.

---

<sup>41</sup> "A curated mammography data set for use in computer-aided ...." 19 dic. 2017, <https://www.nature.com/articles/sdata2017177>. Se consultó el 26 may. 2023.

<sup>42</sup> "CBIS-DDSM Dataset | Papers With Code." <https://paperswithcode.com/dataset/cbis-ddsm>. Se consultó el 26 may. 2023.

### 3.2.4. Anàlisi exploratori de les dades CBIS-DDSM

Ja hem pogut veure algunes de les característiques principals de CBIS-DDSM, com el seu origen, mida i informació que conté. Però quines dades específicament construeixen aquest conjunt és el tema d'aquest apartat.

El conjunt DDSM<sup>43</sup> conté força informació sobre tots els seus casos, però la informació d'alguns camps específics pot ser força limitada i altra és complexa en el seu accés. La millora entre DDSM i CBIS-DDSM es va centrar principalment en aquests punts: la descompressió de les imatges, l'annotació, l'extracció de metadades i l'aplicació d'un nou format.

El resultat d'aquesta millora és un conjunt de dades que conté mamografies, regions d'interès (màscara) i seccions (crop) d'interès. A la Figura 7 trobem el flux seguit pels creadors de CBIS-DDSM per aconseguir les millores esmentades en el conjunt de dades.

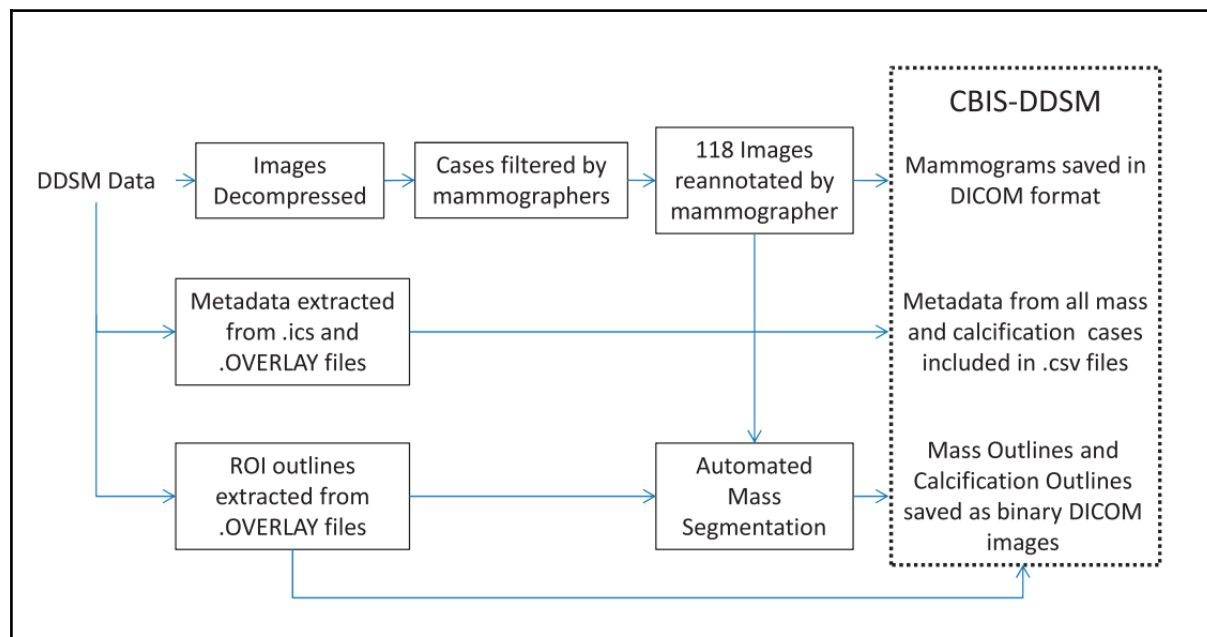


Figura 7: Diagrama de flux<sup>44</sup> de la preparació de CBIS-DDSM.

La descàrrega de les dades es fa a través d'una plataforma del NBIA<sup>45</sup>, el NBIA Data Retriever<sup>46</sup>. També ens cal el fitxer de manifest on s'indica aquelles dades que volem

<sup>43</sup> "A curated mammography data set for use in computer-aided ...." 19 dic. 2017, <https://www.nature.com/articles/sdata2017177>. Se consultó el 26 may. 2023.

<sup>44</sup> "A curated mammography data set for use in computer-aided ...." 19 dic. 2017, <https://www.nature.com/articles/sdata2017177>. Se consultó el 26 may. 2023.

<sup>45</sup> "National Biomedical Imaging Archive - NBIA - NCI Wiki." <https://wiki.nci.nih.gov/display/NBIA>. Se consultó el 27 may. 2023.

<sup>46</sup> "NBIA Data Retriever FAQ - TCIA Online Help." 17 oct. 2022, <https://wiki.cancerimagingarchive.net/display/NBIA/NBIA+Data+Retriever+FAQ>. Se consultó el 27 may. 2023.

descarregar, en format *.tcia*<sup>47</sup>. Un cop descarregat el Data Retriever i el fitxer de manifest, podrem descarregar les dades i les tindrem disponibles localment. El conjunt de dades conté mamografies de calcificacions, masses, o ambdues.

Una de les primeres característiques que crida l'atenció en navegar pels continguts és l'estructura de carpetes que utilitza, on trobarem carpetes dins carpetes, amb noms de carpeta i de fitxer poc descriptius en general.

Pel que fa a fitxers, trobarem dos tipus: fitxers de tipus DICOM<sup>48</sup> i fitxers CSV (*.dcm* i *.csv* respectivament). Cal destacar que un mateix pacient pot tenir dos tipus d'escaneig per al mateix pit, cadascun realitzat des d'un angle diferent: Craneo-Cabal i Mig-Lateral-Oblicua (CC i MLO). Els fitxers *.dcm* poden contenir tres tipus d'informació, la mamografia completa, un retall de la regió d'interès (zoom) i la regió d'interès en forma de màscara. Les mamografies completes es troben en una branca de carpetes i les regions d'interès es troben juntes en una altra branca de carpetes. Aquestes imatges (mamografia, retall i màscara, mostrades a la Figura 8), es troben a fitxers DICOM separats, per tant, un pacient amb només una massa, tindrà associat 3 fitxers independents. Els fitxers DICOM no només contenen la informació d'imatge, també porten incrustats altra informació relacionada amb el pacient, el tipus d'imatge, característiques sobre l'orientació i màquina emprada, entre d'altres.

En el cas de trobar un pit amb més d'una massa, podrem veure que trobarem un retall i una màscara per cada una de les masses.

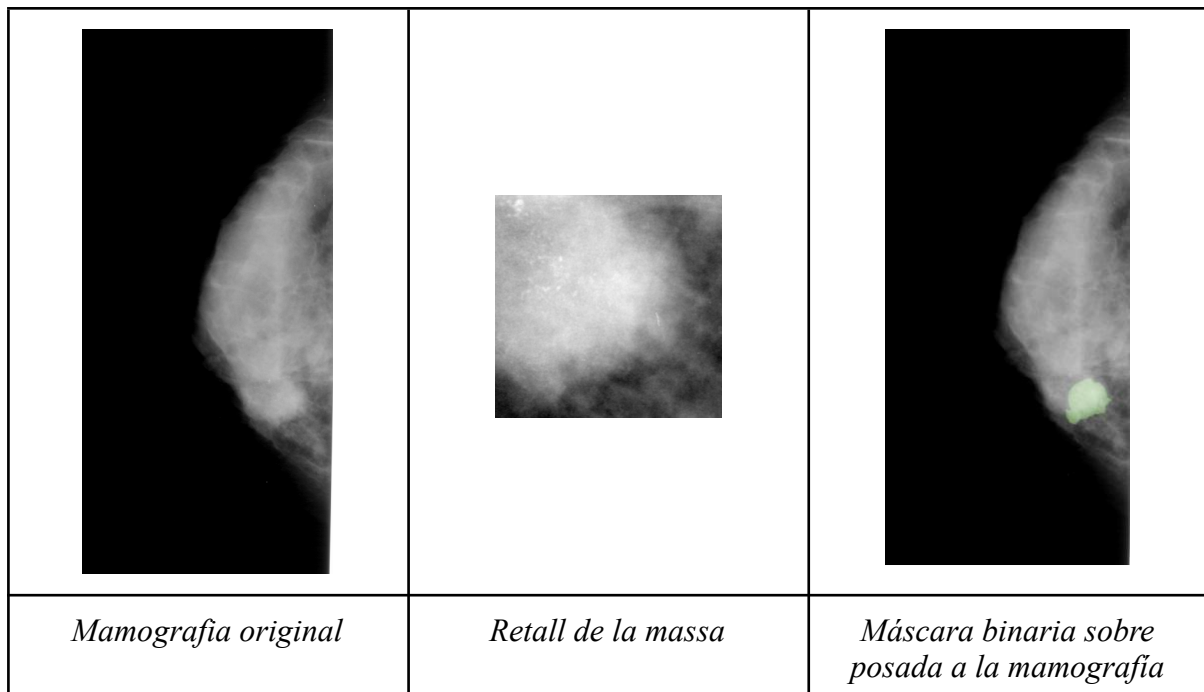


Figura 8: Mamografia original, retall de la massa i màscara sobre la mamografia.

<sup>47</sup> "Curated Breast Imaging Subset of Digital Database for Screening ...." 21 jun. 2016, <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629>. Se consultó el 27 may. 2023.

<sup>48</sup> "DICOM Standard." <https://www.dicomstandard.org/>. Se consultó el 31 may. 2023.

Els fitxers CSV contenen informació relativa tant als fitxers (localització dins les carpetes), com dels pacients. Utilitzant la llibreria *pandas* de Python, podem carregar i manipular fàcilment la informació que contenen aquests fitxers amb l'objectiu, per exemple, d'analitzar més en profunditat les característiques úniques de cada cas.

En la detecció del càncer de mama s'ha demostrat múltiples vegades la importància que té la forma i estructura de les masses<sup>49</sup>, així com l'estat dels seus marges i la densitat del pit. Trobem aquesta informació fàcilment accessible als fitxers CSV.

patient_id	breast_density	image view	mass shape	mass margins
P_00016	4	CC	IRREGULAR	SPICULATED
P_00016	4	MLO	IRREGULAR	SPICULATED
P_00017	2	CC	ROUND	CIRCUMSCRIBED
P_00017	2	MLO	ROUND	ILL_DEFINED
P_00032	3	CC	ROUND	OBSCURED

Figura 9: Dades contingudes al fitxer CSV "mass\_case\_description\_test\_set.csv".

Hem seleccionat cinc característiques principals (un exemple dels continguts a la Figura 9). A continuació es mostra una explicació dels valors més comuns i una taula de distribució per tipus i nombre de casos en el subconjunt de dades de masses:

- **patient\_id**: identificador únic del pacient. Podem veure a la Figura 9 que aquest es repeteix, la causa és que un mateix pacient pot tenir més d'una mamografia on cadascuna és realitzada amb una tècnica diferent.
- **breast\_density**: la densitat del pit. En aquest cas la trobem representada amb un rang de sencers d'un a quatre:
  - Densitat baixa (1): baixa proporció de teixit fibroglandular en comparació amb teixit gras.
  - Densitat dispersa (2): Densitat mamària intermèdia.
  - Densitat heterogènia (3): proporcions de densitat mixta, amb zones més i menys denses.
  - Densitat alta (4): representa una alta proporció de teixit fibroglandular en comparació amb el teixit gras.

<sup>49</sup> "Specimen mammography for intraoperative margin assessment in ...." 2 nov. 2022, <https://www.nature.com/articles/s41598-022-23234-5>. Se consultó el 31 may. 2023.

En la Taula 1 es mostra la quantitat de casos per cadascun dels tipus de densitat del pit.

<b>Tipus de densitat</b>	<b>Nombre de casos</b>
<i>Densitat 1</i>	757
<i>Densitat 2</i>	449
<i>Densitat 3</i>	337
<i>Densitat 4</i>	153

*Taula 1: Densitats i número d'imatges per densitat.*

- **image view**: ens indica la perspectiva utilitzada per a l'obtenció d'aquesta imatge, CC o MLO:
  - CC: Craneo-Caudal, projecció de dalt a baix.
  - MLO: Mig-Lateral Oblicua, projecció obtinguda de forma obliqua.

En la Taula 2 es mostra el nombre de casos per tipus de vista.

<b>Tipus de vista</b>	<b>Nombre de casos</b>
<i>MLO</i>	912
<i>CC</i>	784

*Taula 2: Tipus de vista i número d'imatges per vista.*

- **mass shape**: el tipus de forma de la massa present a la mamografia. Algunes de les més presents es mostren a continuació (vegeu Taula 3):
  - IRREGULAR: forma de la massa irregular.
  - OVAL: forma de la massa ovalada.
  - ROUND: forma de la massa rodona.
  - LOBULAR: forma lobular de la massa.

Existeixen altres formes, i també combinacions com LOBULATED-IRREGULAR, per exemple.



Tipus de forma	Nombre de casos	Tipus de forma	Nombre de casos
<i>IRREGULAR</i>	464	<i>FOCAL_ASYMMETRIC_DENSITY</i>	25
<i>OVAL</i>	412	<i>ASYMMETRIC_BREAST_TISSUE</i>	25
<i>LOBULATED</i>	384	<i>OVAL-LYMPH_NODE</i>	6
<i>ROUND</i>	164	<i>LOBULATED-IRREGULAR</i>	6
<i>ARCHITECTURAL_DISTORSION</i>	103	<i>LOBULATED-LYMPH_NODE</i>	4
<i>IRREGULAR-ARCHITECTURAL_DISTORSION</i>	52	<i>ROUND-OVAL</i>	3
<i>LYMPH_NODE</i>	35	<i>LOBULATED-ARCHITECTURAL_DISTORTION</i>	2

Taula 3: Nombre de casos per forma de massa.

- **mass margins:** aquest camp descriu els marges de la massa detectada. Les més presents són les següents (es mostra la seva distribució en la Taula 4):
  - *CIRCUMSCRIBED*: marges ben delimitats, definits.
  - *MICROLOBULATED*: marges amb petites protuberàncies.
  - *OBSCURED*: marges ocults, no clarament visibles.
  - *ILL\_DEFINED*: marges mal definits, no nítids.

Marges	Nombre de casos	Marges	Nombre de casos
<i>CIRCUMSCRIBED</i>	392	<i>MICROLOBULATED</i>	129
<i>ILL_DEFINED</i>	370	<i>ILL_DEFINED-SPICULATED</i>	30
<i>SPICULATED</i>	363	<i>CIRCUMSCRIBED-ILL_DEFINED</i>	29
<i>OBSCURED</i>	247	<i>OBSCURED-ILL_DEFINED</i>	24

Taula 4: Nombre de casos per tipus de marges.

La selecció d'aquests camps no és arbitrària, com ja s'ha esmentat prèviament, ha estat demostrat que la densitat del pit, la forma i els marges de la massa són vitals per a la

detecció d'aquestes. Hem seleccionat també l'identificador del pacient (*patient\_id*) i el tipus de vista per tal de poder distingir de forma individual cada cas.

L'estructura de carpetes de CBIS-DDSM és altament complexa si aprofundim en els seus continguts. En un primer nivell ens trobarem les dades dividides en un conjunt de test i un d'entrenament. En un segon nivell trobem una divisió en *calc* (calcificacions) i *mass* (masses). I en un tercer nivell trobarem dues branques més, una de mamografies completes i una altra per als retalls i màscares. El següent nivell conteé una carpeta per cas d'estudi, el nom de la carpeta inclou el tipus (calcificació o massa) el "nom" del pacient, el pit (RIGHT o LEFT) i el tipus d'escaneig (CC o MLO). I per últim trobem dos nivells de carpetes més amb noms no descriptius abans d'arribar a la imatge, que pot tenir de nom 1-1.dcm o 1-2.dcm. A la Figura 10 podem veure el camí necessari per arribar fins a una de les imatges.



Figura 10: L'estructura de carpetes fins a arribar al fitxer final.

Una de les solucions a aquesta complexitat de navegació entre carpetes i la falta de descriptibilitat dels noms dels fitxers és reestructurar el sistema de carpetes, eliminant subcarpetes innecessàries i canviant el nom dels fitxers DICOM.<sup>50</sup>

Tot i l'existència d'aquesta opció, en la realització d'aquest treball s'ha optat per una alternativa més simple: utilitzar la informació present als fitxers DICOM per obtenir el nom del cas i el tipus d'imatge. A la Figura 11 podem veure part dels continguts del fitxer DICOM.

<sup>50</sup> "Segmenting Abnormalities in Mammograms (Part 1 of 3) | by Cleon W." 25 feb. 2021, <https://towardsdatascience.com/can-you-find-the-breast-tumours-part-1-of-3-1473ba685036>. Se consultó el 27 may. 2023.

Modality	MG
Conversion Type	WSD
Referring Physician Name	
Series Description	ROI mask images
Patient Name	Mass-Test_P_00016_LEFT_CC_1
Patient ID	Mass-Test_P_00016_LEFT_CC_1
Patient Birth Date	
Patient Sex	
Body Part Examined	BREAST

Figura 11: Part de la informació contenida al fitxer DICOM, podem veure com el camp Series Description conte el tipus d'imatge que conté, en aquest cas una mascara.

Utilitzant aquesta informació podem carregar les dades a memòria sense necessitat de modificar el sistema de carpetes originals, cosa que permet a qualsevol utilitzar l'estructura estàndard sense haver de crear o descarregar una de nova.

Si analitzem les característiques del conjunt de dades, trobarem que tenim 2620 mamografies en total. D'aquestes, 1592 contenen masses, la resta només contenen calcificacions. En aquest treball es va decidir treballar amb masses, per tant, el conjunt de dades emprades seran aquestes 1592 mamografies. És possible que alguna mamografia tingui més d'una massa, en concret trobem 71 mamografies d'aquest tipus (Figura 12).<sup>51</sup>

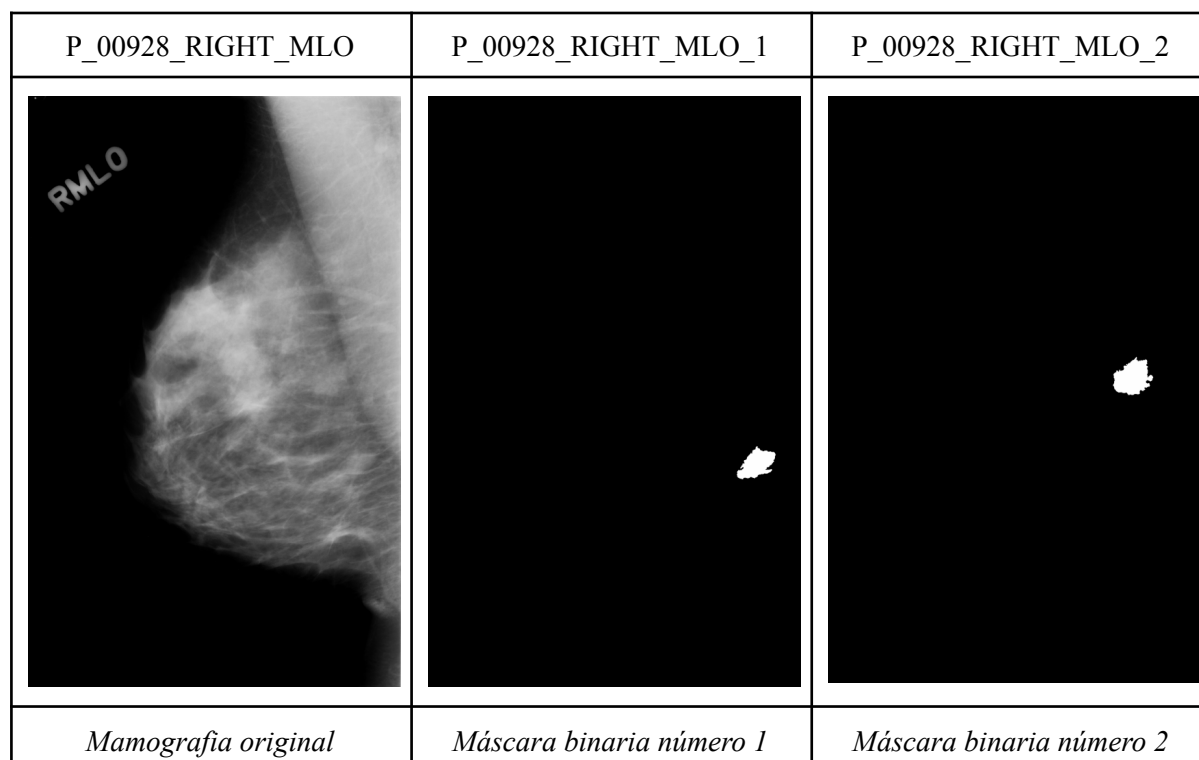


Figura 12: Mamografia i les màscares corresponents a les dues masses.

<sup>51</sup> "Segmenting Abnormalities in Mammograms (Part 1 of 3) | by Cleon W." 25 feb. 2021, <https://towardsdatascience.com/can-you-find-the-breast-tumours-part-1-of-3-1473ba685036>. Se consultó el 27 may. 2023.

És important considerar els casos en els quals trobem més d'una massa, ja que no podem fer servir per a l'entrenament del nostre model les màscares individuals: és necessari unir les diverses màscares d'un mateix pit amb més d'una massa abans de passar-les al model.

Si mostrem una mateixa mamografia al nostre model però mostrem dues o més màscares diferents perdrem coherència i precisió, ja que en realitat existeixen casos amb més d'una massa per mamografia.

### **3.3. Preprocessament de les imatges**

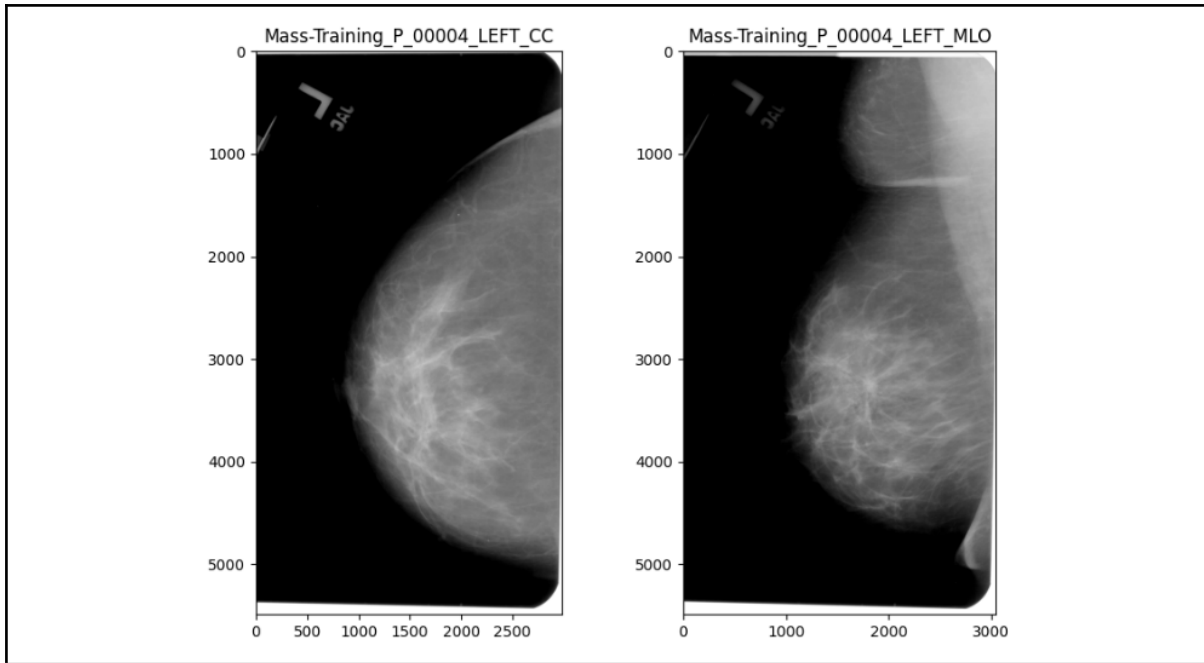
Com ja hem pogut veure en l'apartat anterior, les mamografies presents al conjunt de dades CBIS-DDSM han estat seleccionades i revisades, el que ens garanteix una certa qualitat d'aquestes. Tot i així existeixen una sèrie de passos a seguir per tal de garantir bons resultats. En aquest apartat ens centrarem en aquests passos, el preprocessat, una etapa vital en el procés de segmentació i altres tècniques.

L'objectiu d'aquests passos és millorar la qualitat de les mamografies, ja sigui eliminant artefactes, millorant el contrast o qualsevol altra tècnica que permet obtenir una representació òptima de les característiques que es considerin d'interès en el problema en qüestió. Aquesta millor representació ens permetrà facilitar la posterior segmentació de les masses.

El preprocessat implica un seguit de tècniques o algorismes, com per exemple normalització, eliminació d'artefactes, millora del contrast, retall de seccions no desitjades, reducció de soroll o redimensionat entre d'altres. Veurem quines tècniques s'han implementat en la realització d'aquest treball i per què. Es comentaran també possibles variacions realitzades o que es poden realitzar, i quin impacte en la tasca de segmentació poden tenir l'ús d'uns algorismes o d'altres.

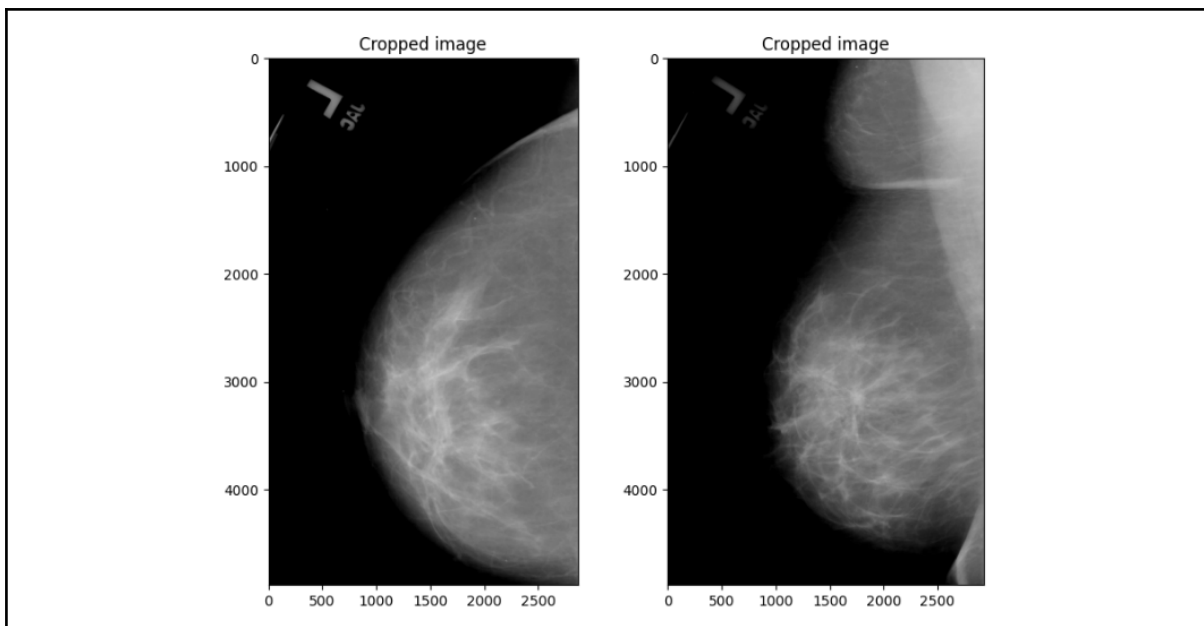
#### *3.3.1. Eliminació de vores*

Si obrim una imatge qualsevol del conjunt (mamografia, no màscara), podem veure que en la gran majoria de casos les vores de la mateixa no són rectes o presenten algun tipus d'artefacte (vegeu exemple en la Figura 13).



*Figura 13: Mamografies originals, sense cap tipus de processat. Observem com les vores presenten marges blancs i les cantonades no són angulars.*

La causa és que les mamografies digitalitzades són escanejos d'una placa, sobre la qual s'ha realitzat la mamografia original. La presència d'aquestes zones pot provocar una pèrdua de rendiment en el model de segmentació, ja que les zones d'un blanc intens poden ser considerades com una massa per un model entrenat sobre imatges no tractades. La solució més simple és retallar cada imatge de forma que eliminem les vores irregulars (Figura 14).



*Figura 14: Mamografies amb les vores irregulars eliminades.*

### 3.3.2. Eliminació d'artefactes

A les imatges mostrades en l'apartat anterior podem veure a la part superior dreta unes lletres. Aquestes lletres (es troben invertides) fan referència al tipus de vista en què s'ha realitzat la mamografia (MLO o CC) i el pit al qual correspon (L per a esquerra, R per a dreta). Aquestes marques les fan els metges per tal de saber ràpidament quin tipus de mamografia estan veient, però a l'hora d'entrenar un model de segmentació, és important eliminar-les pels mateixos motius que els artefactes que veiem a les vores.

En aquest cas s'ha decidit eliminar aquests artefactes utilitzant la llibreria *cv2*<sup>52</sup> i els contorns. El procediment bàsic és el següent:

1. Apliquem *OTSU Thresholding*<sup>53</sup> a la imatge, la qual cosa ens permet separar el primer pla del fons de la imatge.
2. Apliquem una sèrie d'operacions morfològiques amb l'objectiu d'eliminar imperfeccions en els contorns.
3. Apliquem una operació de dilatació als contorns, per compensar la possibilitat de no haver detectat alguna regió.
4. Se seleccionen tots els contorns generats, es guarda el més gran (pit).
5. La resta de contorns es transformen en màscares i s'utilitzen per eliminar els contorns restants de la imatge.

El resultat d'aquest procediment és una imatge amb un fons net, on l'únic objecte present és el pit (vegeu Figura 15).

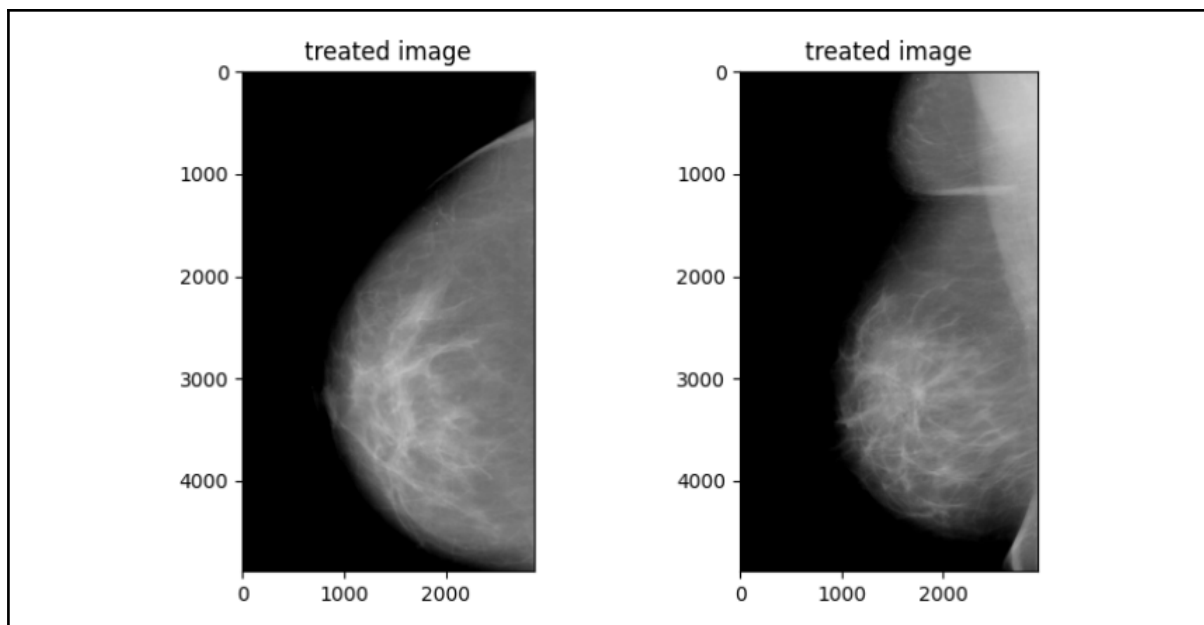


Figura 15: Mamografies amb artefactes eliminats.

<sup>52</sup> "OpenCV modules." <https://docs.opencv.org/4.x/>. Se consultó el 30 may. 2023.

<sup>53</sup> "Otsu's Thresholding Technique - LearnOpenCV." 5 ago. 2020, <https://learnopencv.com/otsu-thresholding-with-opencv/>. Se consultó el 31 may. 2023.

### 3.3.3. Ajust al pit

Si ara mirem les imatges obtingudes després de l'eliminació d'artefactes, podrem veure que ens queda una regió negra sense cap mena d'informació rellevant. Pot ser interessant aplicar algun tipus d'algorisme, que ens permet ajustar la mida de les imatges als límits del pit, almenys en l'eix horitzontal. El motiu és que en l'eix horitzontal trobarem una millor alineació amb l'anatomia del pit, el que pot disminuir la pèrdua de detall en el processat posterior.

En primera instància es va implementar una funció d'ajust utilitzant contorns, de forma similar a com s'eliminen els artefactes en l'apartat anterior. Aquest mètode produeix discrepàncies entre les imatges i les màscares, resultant en mides diferents. També la qualitat del resultat obtingut depèn en gran manera dels paràmetres de múltiples mètodes de detecció de contorns, i obtenir un resultat correcte per a totes les imatges va resultar ser complicat. Es va decidir finalment implementar un algorisme ad hoc alternatiu que no utilitza contorns i es beneficia de l'ús de la llibreria de *NumPy* per a fer càlculs matricials de forma ràpida.

Aquest algorisme realitza un “escaneig” de la mamografia amb l'objectiu de detectar el marge del pit a la imatge, aprofitant la flexibilitat dels slices de Python juntament amb *NumPy*. El procediment a seguir es mostra a continuació:

1. Calculem a quina banda de la imatge es troba el pit.
2. Obtenim la mitjana de valors de la imatge.
3. Creem un rectangle, el qual és la zona que escanegem.
4. Realitzem una sèrie d'iteracions, comparant a cada pas la mitjana de valors del rectangle amb la mitjana de tota la imatge.
5. Si la mitjana de la zona d'escaneig és superior, hem trobat el pit, retrocedim i retallem. La imatge resultant es mostra a la Figura 16.

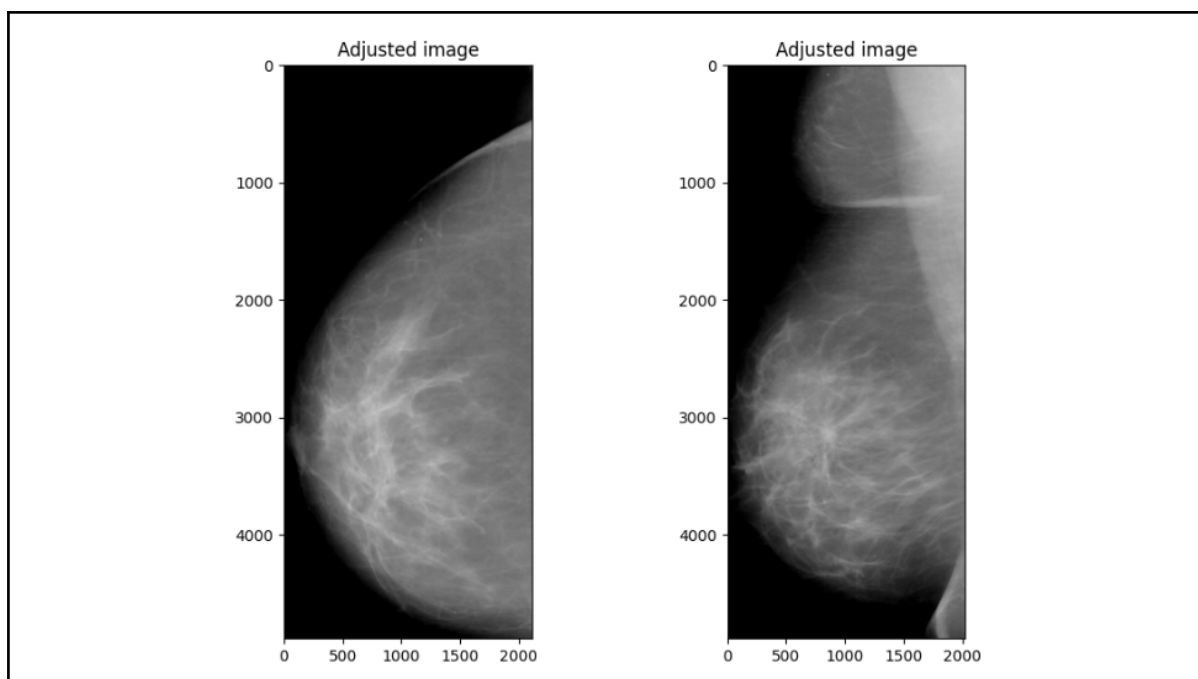


Figura 16: Imatges ajustades a la zona del pit.

L'eliminació de les zones de la imatge on només tenim un fons negre ens permet obtenir diversos beneficis, com la reducció de soroll, estalviar recursos computacionals, l'enfocament en la regió d'interès i, per tant, la millora en la segmentació. Per descomptat és important fer aquesta eliminació de forma correcta, ja que eliminar regions on existeix informació rellevant podria tenir un efecte contrari al desitjat.

### 3.3.4. CLAHE (Contrast Limited Adaptive Histogram Equalization)

CLAHE<sup>54</sup> és un algorisme de processament d'imatges que s'utilitza de forma regular amb l'objectiu de millorar el contrast local d'una imatge donada. Normalment, l'ajust del contrast (equalització de l'histograma) pot introduir soroll no desitjat en la imatge, artefactes, i pot generar una pèrdua de detall en zones on el contrast sigui exagerat. L'algorisme CLAHE té com a objectiu combatre aquestes deficiències.

El seu funcionament es basa a realitzar l'equalització de l'histograma de forma local, per regions o blocs d'imatge més petits, i no aplicant un únic ajust a tota la imatge. D'aquesta forma ens permet millorar el contrast de forma local, adaptant-se a les característiques del bloc en qüestió. Podem veure el resultat en la Figura 17.

Aquest procés és de gran utilitat en imatges mèdiques, sobretot en imatges com les emprades: imatges amb un fons negre del qual no volem exagerar-ne el soroll. Alhora, permet millorar el contrast en les estructures fines, millorant en la representació dels detalls. De forma similar a l'equalització de l'histograma convencional cal controlar els paràmetres de l'algorisme, ja que podem afegir artefactes en zones de forta transició, així com generar soroll.

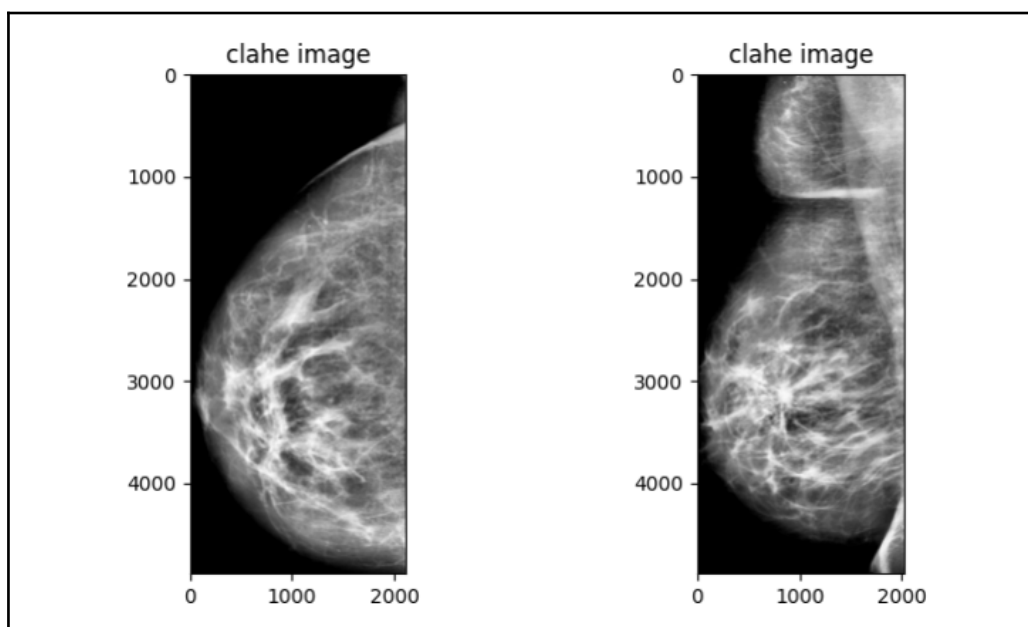


Figura 17: Imatges amb CLAHE aplicat (originals a la Figura 13). L'estructura i detall del pit es veu incrementat, el fons es conserva negre.

<sup>54</sup> "Image Contrast Enhancement Using CLAHE - Analytics Vidhya." 17 ago. 2022, <https://www.analyticsvidhya.com/blog/2022/08/image-contrast-enhancement-using-clahe/>. Se consultó el 12 jun. 2023.



### 3.3.5. Normalització

L'objectiu de la normalització és la de portar els valors presents en les nostres dades a una escala comuna. En l'entrenament de xarxes neuronals l'escala típica és tenir els valors d'entrada de 0 a 1. El motiu és que la majoria de funcions d'activació i algorismes d'optimització treballen en aquest rang.

En el cas de les imatges, portarem els valors de cada píxel al rang de 0 a 1. Això ens proporciona estabilitat en l'entrenament, evitant l'explosió del gradient i i, per tant, millorant la convergència.

## 3.4. Descripció de l'arquitectura de la xarxa U-Net

Com es va discutir en l'apartat 2, la xarxa U-Net és una arquitectura de xarxa convolucional, especialment utilitzada per a segmentació i en concret en segmentació d'imatges mèdiques. La seva característica principal és la forma en U, on trobem un encoder (downsampling) i un decoder (upsampling).

En primer lloc, es va implementar un model d'U-Net utilitzant Keras. En ser l'arquitectura U-Net una de les arquitectures més utilitzades en segmentació, i la número u en segmentació d'imatges mèdiques, existeixen múltiples fonts i exemples de models ja construïts. Gràcies a l'API de Keras construir un model U-Net és relativament simple, utilitzant les capes proporcionades per la llibreria.

Utilitzarem capes de convolució (*Conv2D*), de submostreig (*MaxPooling2D*), de convolució transposada (*Conv2DTranspose*) i capes de concatenació (*skip connections*). El model comença amb la ruta d'encoder, que consta de capes de convolució i capes de pooling. A continuació trobem la ruta de decoder, que consta de capes de transposició i concatenació.

Per tal de facilitar la lectura del codi (encapsulant la implementació de la xarxa) i amb la finalitat de treballar amb un model estàndard, finalment s'ha decidit treballar amb la llibreria *keras-unet*. Aquesta llibreria proporciona la implementació mencionada anteriorment, utilitza *keras* i les seves capes, però les implementa sota una llibreria que proporciona mètodes per tal de construir el model amb només una crida a la mateixa (també implementa altres funcions d'utilitat per a la visualització i creació d'iteradors). A la Figura 18 es mostra el model real utilitzat en el treball.

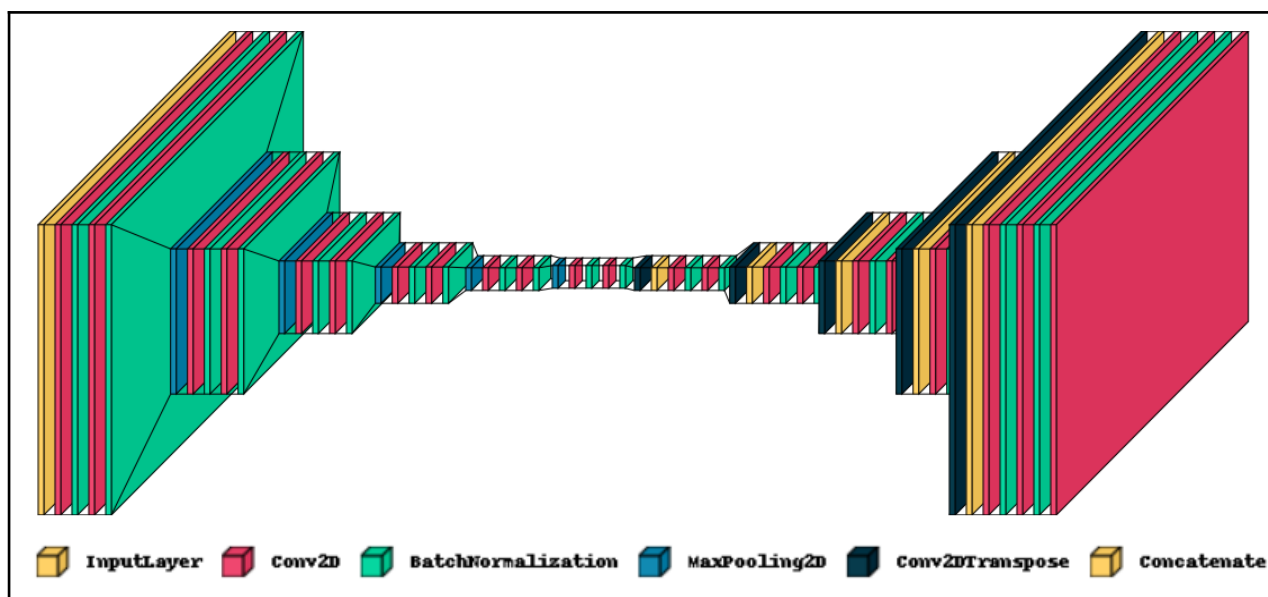


Figura 18: Model U-Net utilitzat en aquest treball. Visualització realitzada amb la llibreria *visualker*<sup>55</sup>.

Observem també la presència de capes de normalització per lots (BatchNormalization). Aquestes capes s'utilitzen per normalitzar les activacions de capes anteriors i va ser proposada per Sergey Ioffe i Christian Szegedy<sup>56</sup> l'any 2015 com a tècnica per a accelerar la convergència en l'entrenament de xarxes neuronals.

En la creació del model s'ha experimentat amb diversos valors de nombre de capes, nombre de filtres i diversos valors de dropout. Cal considerar el fet que variar aquests paràmetres no només afecta el rendiment del model, sinó que també té un cost en recursos computacionals.

S'ha experimentat amb diversos valors en el transcurs de la realització del treball, i com es tracta a l'apartat *Limitacions del treball i possibles direccions futures*, no sempre ha sigut possible executar la segmentació.

### 3.5. Funcions de pèrdua i algorismes d'optimització

La selecció de la funció de pèrdua i dels optimitzadors és fonamental en l'entrenament de models d'aprenentatge automàtic. El seu paper és optimitzar els models, de forma que aquests mateixos aprenguin de forma precisa i efectiva.

Les funcions de pèrdua o cost s'utilitzen per a quantificar la discrepància entre la predicció del model i els valors reals presents al conjunt de dades. La seva funció és minimitzar aquesta discrepància, i la seva elecció depèn completament del tipus de problema

<sup>55</sup> "visualker for Keras / TensorFlow - GitHub." <https://github.com/paulgavrikov/visualker>. Se consultó el 1 jun. 2023.

<sup>56</sup> "Accelerating deep network training by reducing internal covariate shift.." 2 mar. 2015, <https://arxiv.org/pdf/1502.03167>. Se consultó el 31 may. 2023.

a tractar. En el cas de la segmentació volem una funció que compari la màscara predita amb la màscara del dataset. S'han utilitzat dues funcions de pèrdua diferents per a la realització del treball, ambdues típiques en tasques de segmentació:

1. **Coefficient de Dice**<sup>57</sup>: el coeficient de Dice o de Sørensen-Dice, s'utilitza per a mesurar la superposició entre dues segmentacions.

$$Dice = \frac{2 \cdot |A \cap B|}{|A| + |B|}$$

- On A i B son les segmentacions.
- $|A \cap B|$  és la intersecció entre A i B.
- $|A|$  i  $|B|$  són les cardinalitats, és a dir, el nombre d'elements en A i en B.

Dice varia entre 0 i 1, on 0 indica la falta de superposició i 1 indica la superposició o similitud exacta. Cal destacar que Dice penalitza menys els falsos negatius en comparació amb altres índexs.

2. **IOU**<sup>58</sup> (Intersection Over Union): també conegut com a Índex de Jaccard, és un coeficient de superposició, de forma similar que Dice. Es calcula de la següent forma:

$$IoU = \frac{|A \cap B|}{|A \cup B|}$$

- On A i B son les segmentacions.
- $|A \cap B|$  és la intersecció entre A i B.
- $|A \cup B|$  és la unió entre A i B.

De la mateixa forma que Dice, l'IOU varia entre 0 i 1, indicant 0 la manca d'intersecció i 1 la màxima intersecció possible.

Pel que fa als algorismes d'optimització, el seu objectiu és el d'ajustar els paràmetres del model durant l'entrenament, amb l'objectiu de minimitzar la funció de pèrdua determinant els pesos i esbiaixos que s'utilitzen a cada pas de l'entrenament. En la realització del treball s'han utilitzat dos algorismes d'optimització o optimitzadors: SGD (Stochastic Gradient Descent) i Adam:

---

<sup>57</sup> "Metrics to Evaluate your Semantic Segmentation Model | by Ekin Tiu." 9 ago. 2019, <https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2>. Se consultó el 5 jun. 2023.

<sup>58</sup> "Metrics to Evaluate your Semantic Segmentation Model | by Ekin Tiu." 9 ago. 2019, <https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2>. Se consultó el 5 jun. 2023.

- **SGD**<sup>59</sup>: es tracta d'un optimitzador molt utilitzat en l'entrenament de models d'aprenentatge automàtic. El seu objectiu és minimitzar la funció de pèrdua ajustant els paràmetres del model. SGD realitza aquest ajust utilitzant una única mostra o conjunt de mostres, fet que el converteix en l'algorisme més eficient i vàlid per a grans quantitats de dades. Consisteix en les etapes d'Inicialització, Mostreig, Càlcul del Gradient, Actualització de paràmetres i Repetició. La seva component estocàstica permet evitar límits locals. Els paràmetres que podem ajustar al SGD són el Learning Rate i el Moment.<sup>60</sup>
- **Adam**<sup>61</sup>: Adam (Adaptative Moment Estimation) és una millora sobre el SGD, combinant el descens del gradient amb l'adaptació del Learning Rate. La seva característica principal és que adapta el Learning Rate de forma automàtica en funció de la magnitud dels gradients calculats en el moment i dels anteriors. De forma similar, consta de diversos passos: Inicialització, Càlcul del Gradient, Actualització del Momentum, Correccions, Actualització dels Paràmetres, Repetició.<sup>62</sup>

Els algorismes d'optimització utilitzats es troben ja implementats en la API de keras, i per tant no és necessari crear noves funcions o classes.

### 3.6. Descripció de les diferents estratègies d'entrenament

Un cop s'han seleccionat les dades, s'han preprocessat, hem definit l'arquitectura del model de segmentació i hem establert els seus paràmetres inicials, el pas següent és aplicar tècniques d'augment de dades per tal de garantir una millor generalització del problema.

Tot i això pot ser interessant també realitzar proves sense aplicar augment de dades, amb l'objectiu de comparar els resultats i les mètriques obtingudes (errors, coeficients...).

---

<sup>59</sup> "Stochastic Gradient Descent — Clearly Explained !!" 6 sept. 2019, <https://towardsdatascience.com/stochastic-gradient-descent-clearly-explained-53d239905d31>. Se consultó el 5 jun. 2023.

<sup>60</sup> "Stochastic Gradient Descent — Clearly Explained !!" 6 sept. 2019, <https://towardsdatascience.com/stochastic-gradient-descent-clearly-explained-53d239905d31>. Se consultó el 31 may. 2023.

<sup>61</sup> "Gentle Introduction to the Adam Optimization Algorithm for Deep ...." 3 jul. 2017, <https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/>. Se consultó el 5 jun. 2023.

<sup>62</sup> "Adam — latest trends in deep learning optimization.." 22 oct. 2018, <https://towardsdatascience.com/adam-latest-trends-in-deep-learning-optimization-6be9a291375c>. Se consultó el 31 may. 2023.

### *3.6.1. Entrenament sense augment de dades*

Aquest entrenament es caracteritza per realitzar l'entrenament del model de segmentació sense augmentar les dades disponibles. En aquesta estratègia no s'aplica cap modificació sintètica al conjunt d'entrenament i és possible que en certes situacions sigui una tècnica vàlida o preferible.

Un cas pot ser on les dades originals són prou diverses i representatives del problema, i per tant ja inclou variacions suficients i situacions que proporcionen una mostra completa i adequada de l'entrada. Tot i això és important considerar que aquesta estratègia pot tenir limitacions i ens pot portar a altres problemes, ja que si ens ajustem només a les mostres específiques del conjunt de dades original podem limitar la capacitat del model de segmentació per a generalitzar.

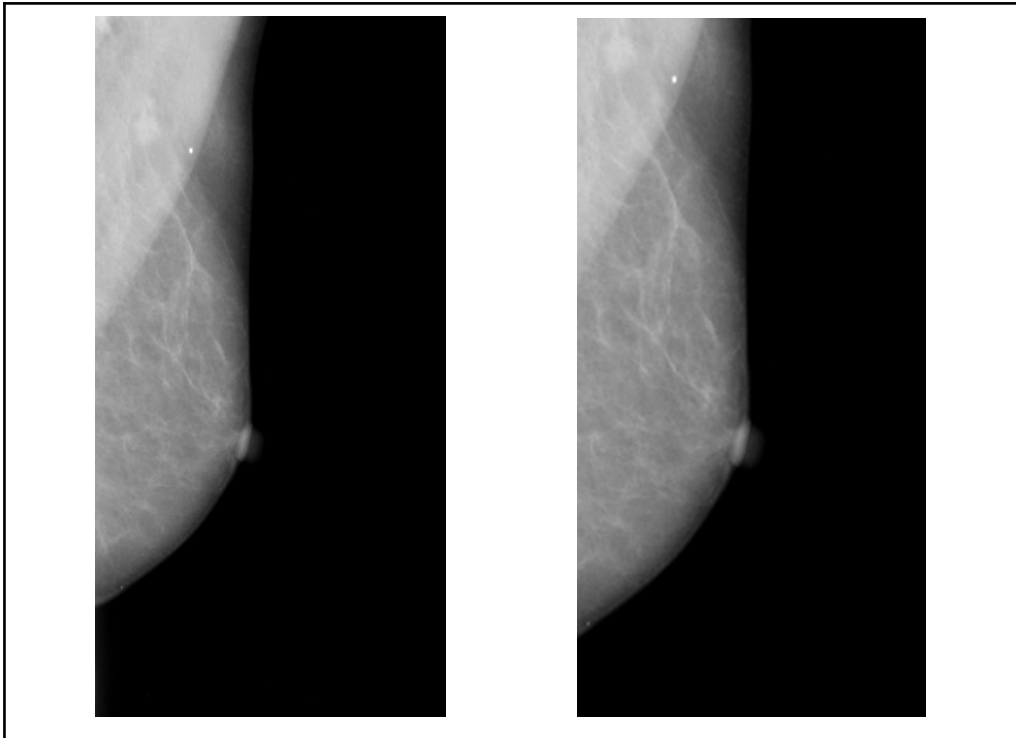
En el transcurs del treball s'han fet entrenaments sense augment de dades. Però com veurem en l'apartat de Resultats els resultats milloren substancialment quan apliquem augment de dades.

### *3.6.2. Entrenament amb augment de dades*

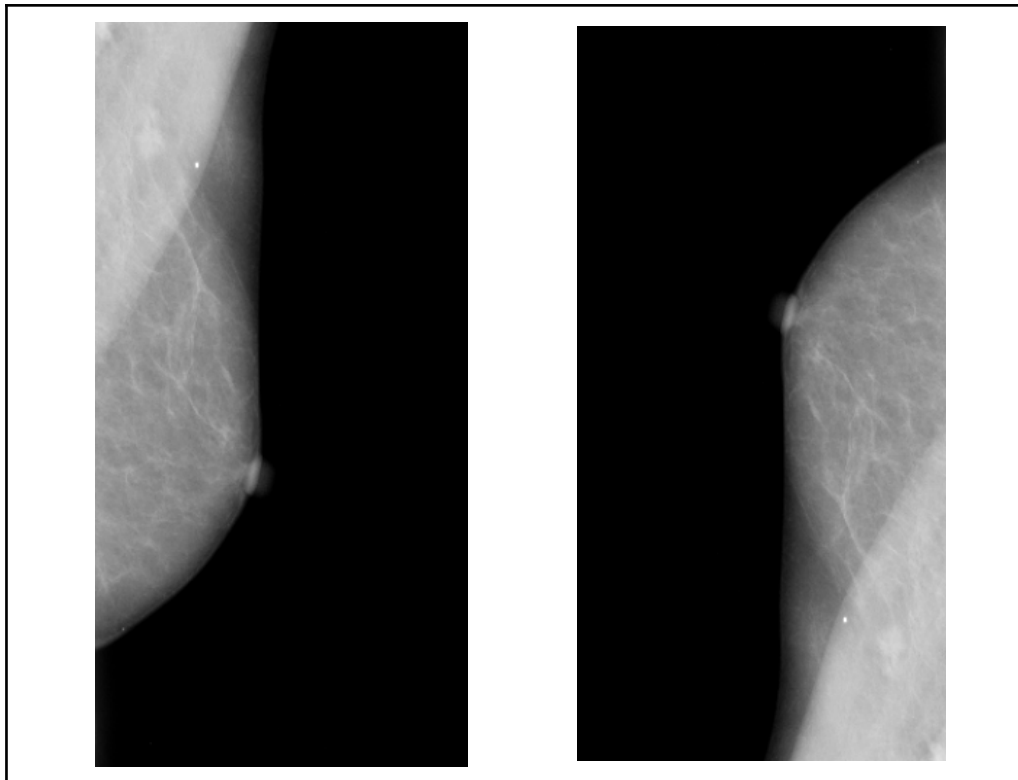
En casos on el conjunt de dades és limitat en quantitat o diversitat, l'entrenament sense augment de dades ens pot portar a problemes d'overfitting per la falta de generalització. L'overfitting es dona quan el model s'ajusta en excés als detalls proporcionats a les dades d'entrenament, i en introduir noves dades de diferents característiques, no és capaç de generalitzar allò après a la nova situació. Això implica una baixa capacitat de resposta del model entrenat a entrades on les variacions respecte al conjunt d'entrenament és diferent.

La solució a aquests problemes és l'augment de dades; ja descrit a l'apartat 2.5. En aquest treball s'han explorat diverses tècniques, algunes ja descrites en l'apartat mencionat. La seva implementació s'ha fet utilitzant la llibreria Keras-unet, que implementa una funció que permet l'entrada de diversos paràmetres per tal d'obtenir un iterador que contingui les dades augmentades. Les tècniques explorades han sigut:

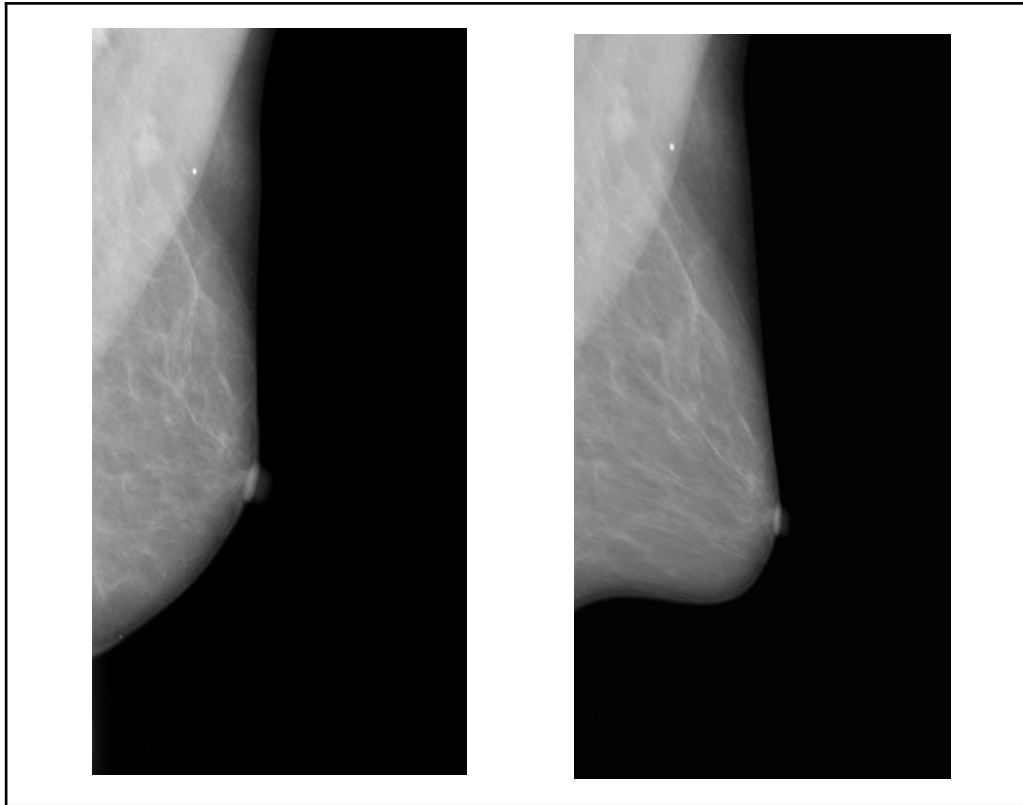
- Augments de mida (vegeu Figura 19)
- Mirall i volteig (vegeu Figura 20)
- Deformacions elàstiques (vegeu Figura 21)
- Canvis de brillantor i contrast (vegeu Figura 22)



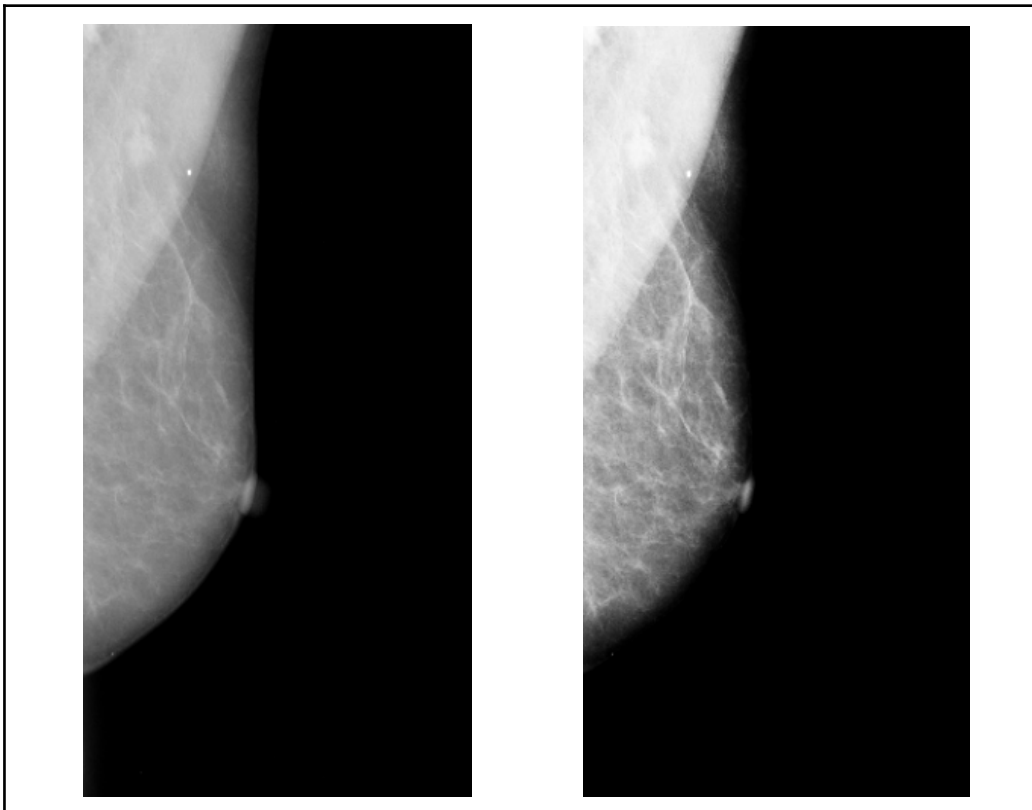
*Figura 19: Exemple de lleuger augment de mida (zoom positiu).*



*Figura 20: Exemple d'imatge voltejada.*



*Figura 21: Exemple de transformació elàstica.*



*Figura 22: Exemple de canvi en brillantor i contrast.*

## **4. Experiments i resultats**

En aquest apartat es presenten els experiments realitzats utilitzant els recursos, mètodes i tècniques que s'han vist fins ara. També es presenten els resultats d'aquests, obtinguts del preprocessat i augment de les dades, i de l'entrenament i avaluació dels models. L'objectiu és analitzar els diversos procediments seguits, avaluant-ne l'efectivitat.

En primer lloc, però, presentarem la metodologia emprada per avaluar els resultats, quines mètriques s'han utilitzat per mesurar la qualitat de la segmentació i el procés d'avaluació. Un cop completada aquesta part passarem a mostrar els resultats quantitius i qualitius.

### **4.1. Metodologia d'avaluació dels resultats**

Per a l'avaluació dels resultats primer cal recordar que s'han fet servir les següents funcions de pèrdua: Coeficient de Dice i IOU; que ja van ser introduïts a l'apartat 3.4. Aquestes mètriques són àmpliament emprades en l'avaluació de qualitat de segmentacions i donen una mesura quantitativa del nivell de coincidència entre la màscara original i la predicció feta.

S'ha utilitzat, com també es va introduir anteriorment a l'apartat 3.1.3, el conjunt de dades CBIS-DDSM, que ja inclou un split de test per tal de poder avaluar les segmentacions fetes. D'aquest conjunt s'han utilitzat 361 imatges de test, i de les 1231 imatges d'entrenament, s'han reservat 100 per a validació. Això ens deixa un split (partició) de train-validation-test de 1131-100-361.

Durant l'entrenament s'han monitorat les funcions de pèrdua mencionades, s'han calculat per al conjunt de dades de test, i per últim s'han generat mostres de segmentació per poder comparar-les amb les originals.

### **4.2. Resultats obtinguts en els diferents experiments**

A continuació es presenten els resultats obtinguts en els diferents experiments amb la finalitat d'avaluar el rendiment dels models de segmentació entrenats i testeats sobre el conjunt de dades CBIS-DDSM. Els resultats s'han dividit en dos subapartats: resultats quantitius i resultats qualitius.



#### 4.2.1. Resultats quantitativus

En total s'han comptabilitzat 10 experiments, tot i que s'han realitzat molts més. El motiu per fer aquesta selecció és que no tots han obtingut resultats vàlids o no s'han pogut executar per manca de recursos. En total s'han realitzat uns 100 experiments, dels quals s'han seleccionat els 10 esmentats com els més representatius del desenvolupament del treball.

	Filters	Layers	Steps	Epochs	Augmented	Dice	IOU	Loss
Experiment #1	32	5	600	75	Yes	0.43	0.31	0.073
Experiment #2	32	5	600	100	Yes	0.50	0.39	0.062
Experiment #3	32	5	600	150	Yes	0.73	0.59	0.010
Experiment #4 *	32	5	600	150	Yes	0.78	0.64	0.012
Experiment #5 **	32	5	600	150	Yes	<b>0.83</b>	<b>0.72</b>	<b>0.009</b>

Taula 5: Resultats obtinguts amb augment de dades.

	Filters	Layers	Steps	Epochs	Augmented	Dice	IOU	Loss
Experiment #6	32	5	600	75	No	0.31	0.24	0.162
Experiment #7	32	5	600	100	No	0.46	0.32	0.107
Experiment #8	32	5	600	150	No	0.56	0.47	0.062
Experiment #9 *	32	5	600	150	No	0.65	0.49	0.013
Experiment #10**	32	5	600	150	No	0.77	0.63	0.019

Taula 6: Resultats dels experiments realitzats amb diferents configuracions. Els experiments amb \* indiquen que es van ajustar els valors de l'augment de dades i amb \*\* indica que es va aplicar l'algorisme d'ajust al pit explicat anteriorment a l'apartat 3.2.3.

A les taules 5 i 6 observem els resultats obtinguts en els diferents experiments realitzats. Podem veure que les característiques d'arquitectura del model s'ha mantingut (Filters i Layers), de la mateixa forma s'ha mantingut el paràmetre Steps. La diferència la trobem en el número d'epochs i si s'ha utilitzat estratègia d'augment de dades o no. A més, a

l'experiment #4 i #9 es van ajustar els valors de l'augment de dades i als experiments #5 i #10 es va aplicar l'algorisme d'ajust al pit explicat anteriorment a l'apartat 3.2.3. .

En experiments previs als mostrats a les taules, i comparant els resultats amb els mostrats, es va poder veure que models amb un nombre d'Epochs inferior a 75 i steps inferiors a 500 no produeixen bons resultats. En molts casos aquests models seleccionaven gairebé tota la zona del pit, o totes les zones denses. Aquest fet donava com a resultat uns valors de Dice i de IOU molt baixos, dins l'interval [0.2, 0.3]. De forma similar, però en l'altre extrem, experiments amb més de 150 epochs o 600 steps, donen bons resultats, però no superiors o iguals als mostrats.

En els experiments #1, #2, #3, #6, #7 i #8 podem veure l'evolució de les mètriques en funció de les epoch establertes, la qual té una evolució creixent respecte d'aquest paràmetre.

A l'experiment #4 i #9 es van ajustar els valors de l'augment de dades, variant els valors de rotació i deformació de la imatge. En concret es va afegir rotació sobre l'eix horitzontal a més del vertical (*flip* de 180° en els dos eixos) i es va incrementar el *shear range* de 5 a 15 graus.

Als experiments #5 i #10 es va aplicar l'algorisme d'ajust al pit mencionat anteriorment a l'apartat 3.3.3. L'experiment #5 és el que millor resultat ha obtingut amb un coeficient de Dice de 0.83 (83%) i un IOU de 0.72 (72%) i serà l'analitzat a continuació.

Al gràfic de la Figura 23 podem veure quina ha sigut la progressió de les mètriques a mesura que avança l'entrenament en l'experiment #5.

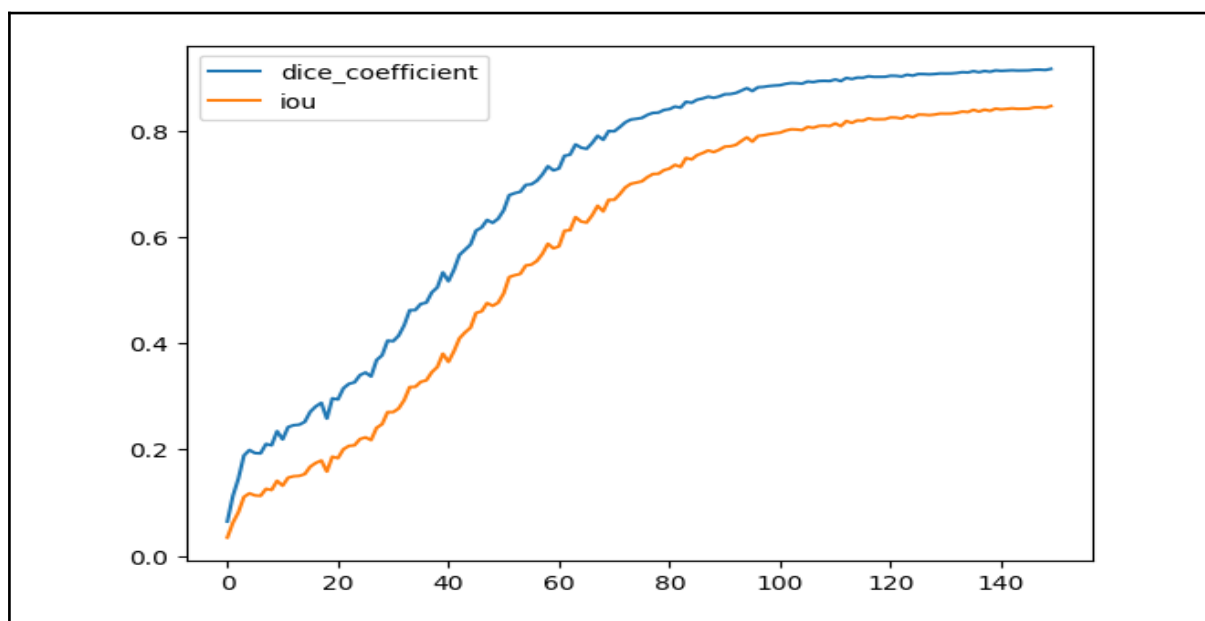


Figura 23: Gràfic corresponent a l'evolució de les mètriques d'avaluació. Eix X correspon a epochs, eix Y correspon al tant per u.

Ens trobem amb una ràpida pujada, de forma contínua, en la similitud mesurada pel coeficient de Dice i l'IOU des del principi de l'entrenament fins a aproximadament l'epoch 80. Es pot apreciar l'inici d'un estancament situat aproximadament sobre l'epoch 100 en endavant, on el rendiment obtingut és cada cop menor.

Fent un petit anàlisi estadístic dels valors de Dice obtingut per a cada imatge de test, obtenim les estadístiques presentades a la Taula 7.

Mitjana	Mediana	Desviació Estàndard	Màxim	Mínim
0.83	0.905	0.291	0.958	0.0

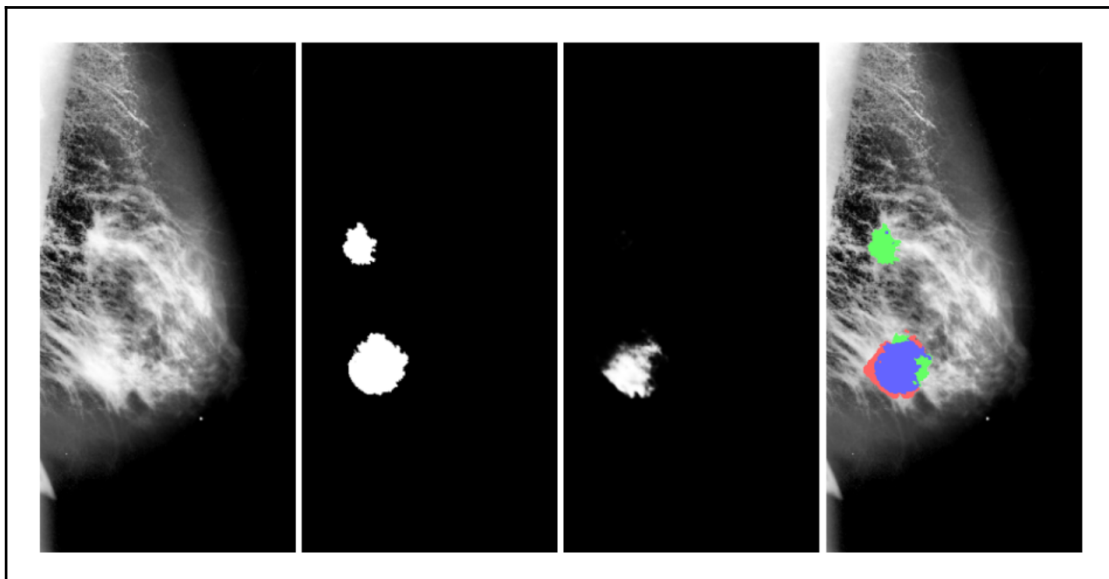
*Taula 7: Estadístiques de les prediccions sobre el conjunt de test del millor model.*

- **Mitjana:** valor promig del coeficient de Dice en el conjunt de test. En aquest cas ens indica que tenim una superposició d'un 83% en promig entre les regions segmentades i de referència.
- **Mediana:** la mediana és el valor trobat al punt mitjà quan els coeficients són ordenats de menor a major. En aquest cas ens indica que el 50% dels coeficients es troben per sota del 90.5% i l'altre 50% per sobre.
- **Desviació estàndard:** la desviació estàndard és una mesura de dispersió que indica la variació del valor del coeficient de Dice respecte a la mitjana.
- **Màxim:** el valor màxim obtingut, en aquest cas del 95.8%.
- **Mínim:** el valor mínim obtingut, en aquest cas del 0%, és a dir, un mínim d'una segmentació té una intersecció del 0% o molt pròxima.

#### 4.2.2 Resultats qualitius

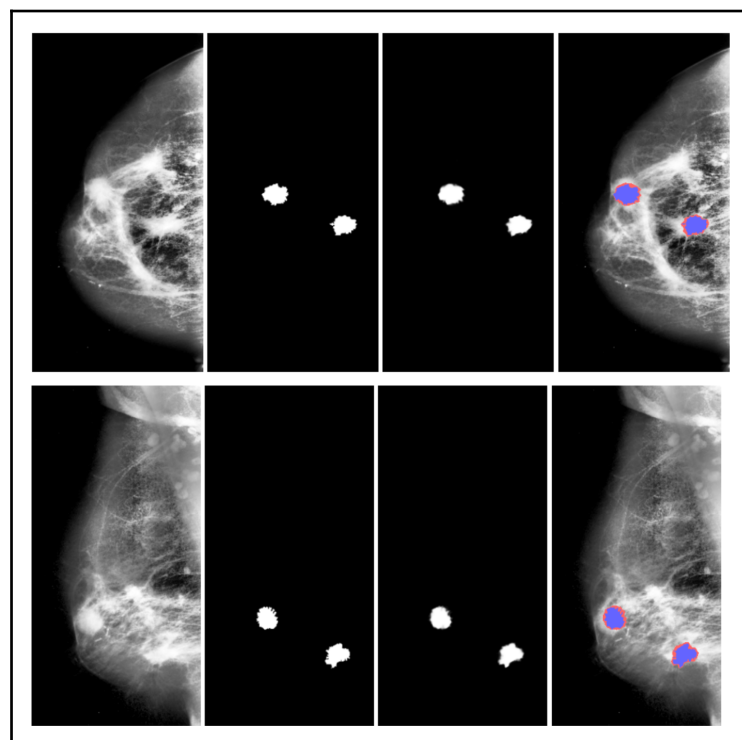
A continuació es presenten exemples visuals dels resultats obtinguts pels models. En concret ens centrarem en el millor model obtingut. En primer lloc, es mostra un exemple de segmentació on el resultat no ha estat tan bo, però resulta útil per tal de mostrar de forma visual l'estructura amb la qual es presenten més endavant la resta de resultats.

A la Figura 24 es comparen les prediccions amb les màscares originals per a mostrar com es representen ambdues a sobre de la mamografia original. Aquesta segmentació ha obtingut un resultat força baix, amb un coeficient de Dice de 0.57 o del 57%.



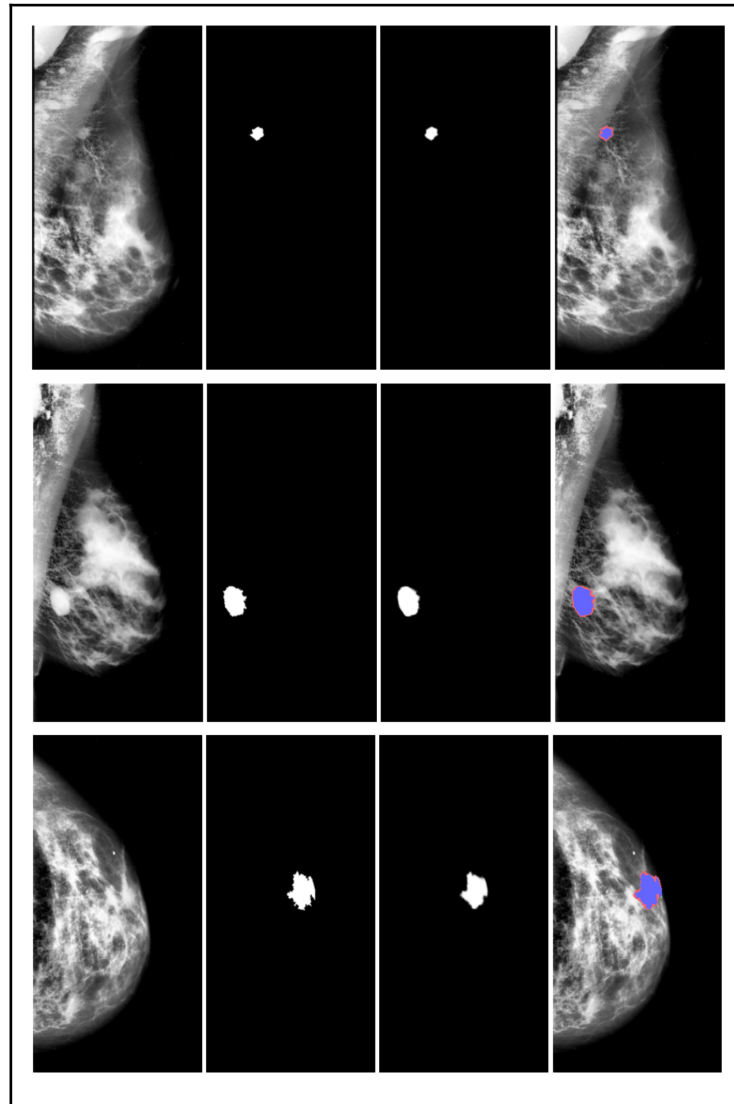
*Figura 24: D'esquerra a dreta: mamografia, màscara original, predicció, sobreposició. Blau = True Positive, Vermell = False Positive, Verd = False Negative. Dice de 57%.*

A la Figura 25, es mostren dos casos, en els quals el nombre de masses presents a les mamografies i màscares originals és igual a 2. Amb un bon nivell d'intersecció.



*Imatge 25: Dues segmentacions amb un nivell d'intersecció elevat on el nombre de masses és igual a 2.*

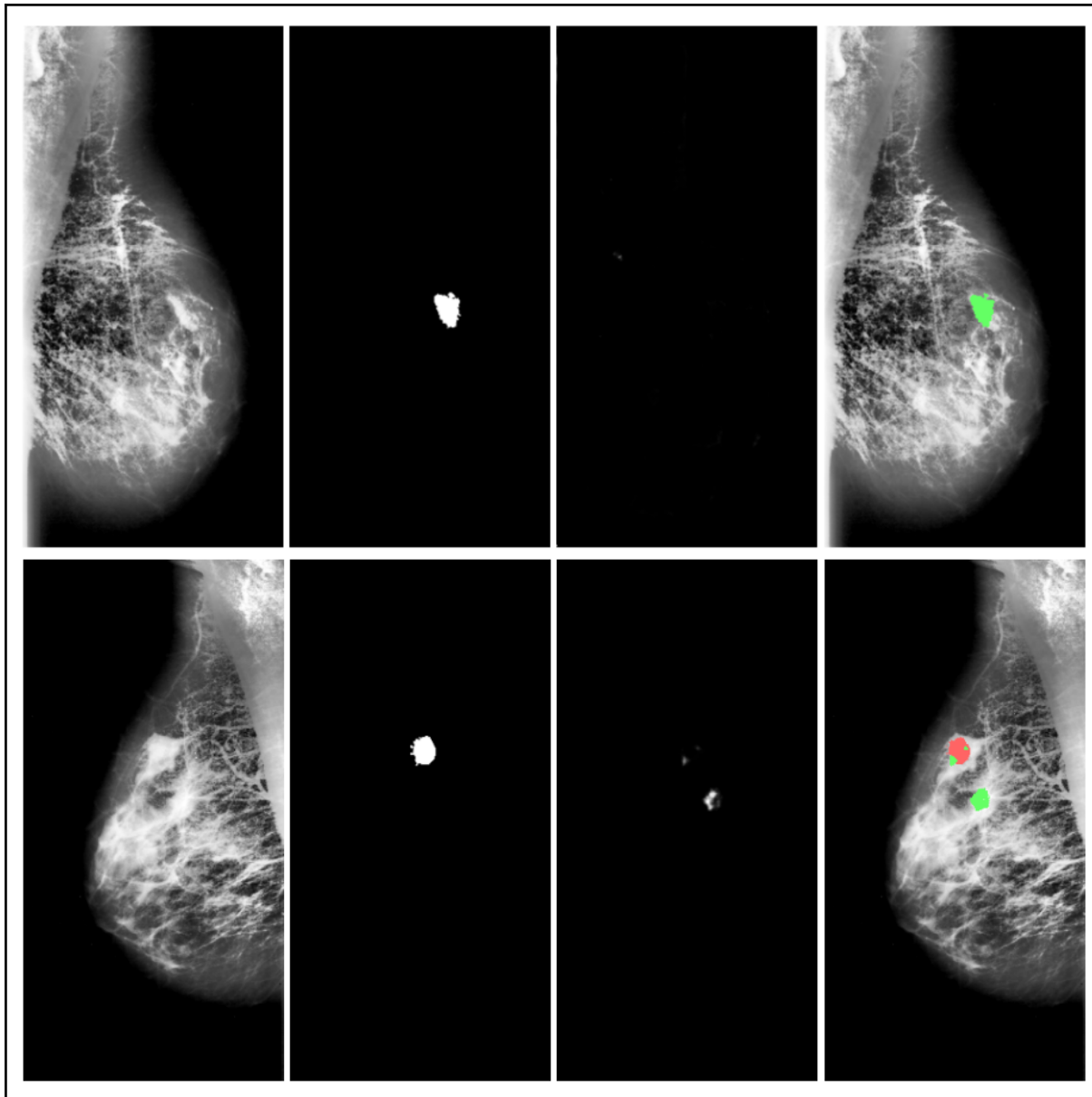
A la Figura 26, es mostren resultats per a diferents segmentacions on el nombre de masses presents a la mamografia original és igual a 1, on la intersecció, de nou, és elevada.



*Figura 26: Diferents segmentacions realitzades pel model on el nombre de masses es igual a 1 i amb una bona intersecció.*

Amb un anàlisi visual ràpid, podem veure com les segmentacions obtingudes pel model són força precises. El model és capaç de segmentar de forma correcta gairebé totes les anormalitats (masses), incloent-hi aquells casos en què trobem més d'una anormalitat per imatge.

En alguns casos, com els mostrats a la figura 27, podem trobar que el model ha segmentat regions que no es troben en les màscares binàries originals. Visualment, aquests falsos positius acostumen a tenir el perfil d'anormalitats, inclòs forma i brillantor. També trobem algun cas en què no ha realitzat cap classe de segmentació.



*Figura 27: Mostra de segmentació. En el primer cas (primera fila) no s'ha detectat cap massa, en el segon cas (segona fila) la regió detectada és diferent de la màscara original.*

## 5. Anàlisi de resultats

En aquest apartat es fa un anàlisi exhaustiu dels resultats obtinguts en els experiments de segmentació realitzats. En concret compararem aquests resultats amb altres treballs de característiques similars, i ens centrarem a comprendre les causes darrere del que s'ha obtingut.

En primer lloc, al subapartat “Comparació dels resultats amb treball relacionats”, compararem la feina feta i els resultats amb altres treballs relacionats amb la segmentació de mamografies. Destaquem les similituds, però també les diferències en conjunts de dades, mètriques i altres tècniques o algorismes implementats. L'objectiu és explorar si els resultats són consistents amb feina prèvia, i mirarem de veure si s'ha pogut superar o igualat l'estat de l'art en la segmentació mamogràfica.

En segon lloc, al subapartat “Discussió dels resultats obtinguts i interpretació dels mateixos”, realitzarem una discussió i interpretació en detall del que s'ha obtingut en els experiments. Parlarem sobre rendiment, interpretació, punts forts i limitacions. En un apartat posterior lligarem amb oportunitats de millora que s'obren en analitzar aquests punts, discutint els punts a millorar.

### 5.1. Comparació de resultats amb treballs relacionats

En general els resultats obtinguts en aquest treball són prometedors, i es troben pròxims a altres resultats obtinguts en treballs anteriors.

Al treball fet per Asma Baccouche et al.<sup>63</sup>, els autors van realitzar una sèrie de comparacions de rendiment entre diversos tipus d'arquitectures, models i conjunts de dades. Podem utilitzar aquesta informació per a comparar els resultats obtinguts en aquest model.

Els resultats obtinguts per exemple en el treball de Al-Antari et al.<sup>64</sup> amb una variació d'una FCN (la FrCN), va obtenir una puntuació de Dice d'un 92.69% en el conjunt de dades INbreast. La feina feta per Zhu et al.<sup>65</sup>, per la seva banda, va obtenir resultats del 90.97 i a INbreast i del 91.30 al DDSM-BCRP<sup>66</sup> utilitzant una altra variació de FCN (multi-scale FCN).

---

<sup>63</sup> "Connected-UNets: a deep learning architecture for breast mass ...." 2 dic. 2021, <https://pubmed.ncbi.nlm.nih.gov/34857755/>. Se consultó el 3 jun. 2023.

<sup>64</sup> "A fully integrated computer-aided diagnosis system for digital X-ray ...." <https://www.sciencedirect.com/science/article/abs/pii/S1386505618302880>. Se consultó el 3 jun. 2023.

<sup>65</sup> "Adversarial Deep Structured Nets for Mass Segmentation ... - arXiv." 24 oct. 2017, <https://arxiv.org/abs/1710.09288>. Se consultó el 3 jun. 2023.

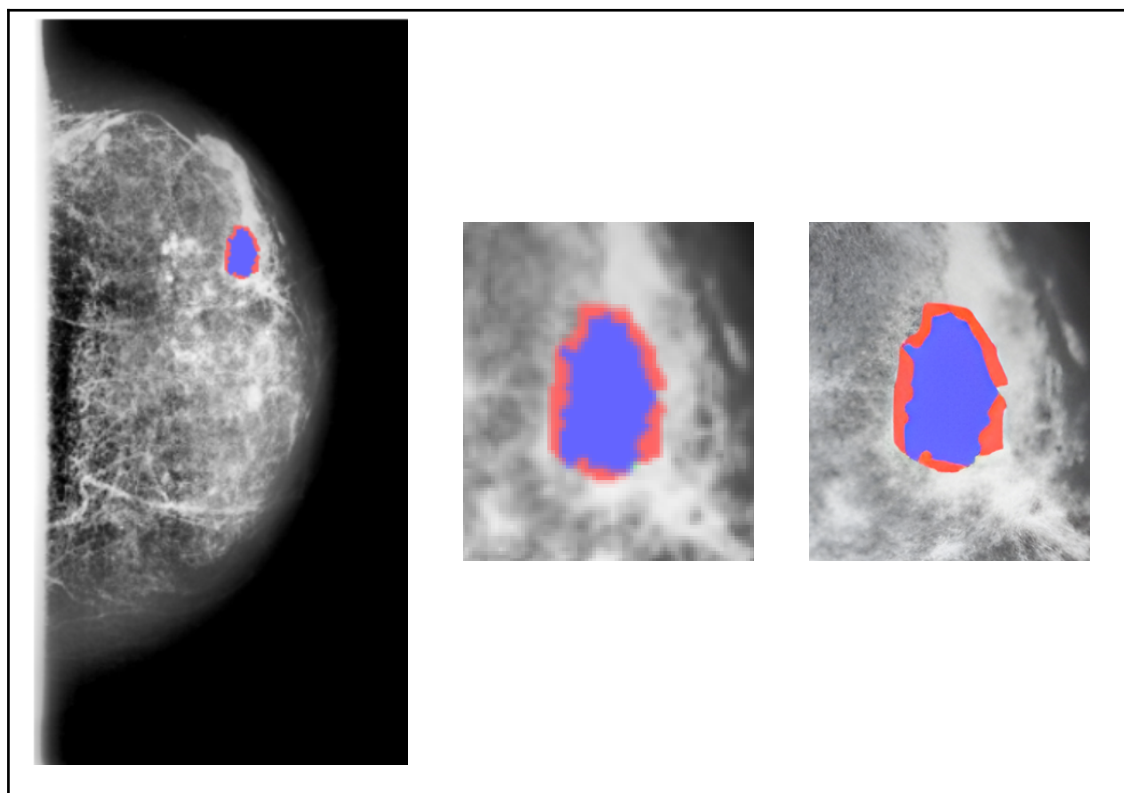
<sup>66</sup> "DoD BCRP Mammography Datasets at USF." <http://www.eng.usf.edu/cvprg/mammography/DDSMB/BCRP/bcrp.html>. Se consultó el 3 jun. 2023.

A la feina de Asma Baccouche et al. trobem que finalment es van aconseguir precisions de fins al 95.39% en la segmentació de les imatges utilitzant una variació de la U-Net, la Connected-ResUNet.

## 5.2. Discussió dels resultats obtinguts i interpretació dels mateixos

Si realitzem un anàlisi visual de les segmentacions realitzades pel millor model, podem arribar a la conclusió que els resultats produïts són generalment bons, amb algunes excepcions.

En les segmentacions on la localització i forma de la massa han sigut segmentades de forma precisa, observem que la regió segmentada és lleugerament superior a la segmentació original (ground truth). Veure Figura 28.



*Figura 28: Mamografia amb superposició de predicció i màscara original, esquerra. Secció segmentada, centre. Secció segmentada augmentada artificialment, dreta.*

Existeixen possibles raons que explicarien aquesta discrepància entre la predicció i la màscara original:

1. **Variabilitat en les estructures fines:** les masses poden tenir una gran variabilitat en la seva estructura, sobretot en els marges. Aquesta variabilitat pot resultar en una generalització massa “general” del que és el marge de la massa. Això pot portar al fet



que el model interpreti una certa regió al voltant de la massa com a tumor que un expert no interpretaria com a tal.

2. **Diferències en la definició de contorns:** les màscares originals que hagin sigut creades per experts, poden tenir marges molt definits, amb contorns suaus i precisos. Els marges de la predicció poden ser més sensibles a certs detalls, resultant en una major superfície considerada com a tumor.
3. **Limitacions del model:** és possible que (com s'ha comentat en el punt 1), el model no pugui captar amb màxima precisió totes les variacions i detalls de cada massa. Això pot introduir desviacions respecte de les màscares originals.
4. **Preprocessament i augment de dades:** les variacions de resolució, mida i/o forma realitzades en el preprocessament i augment de dades poden ser una causa d'aquesta discrepància. Si el procés no és òptim, és possible que de nou, certes discrepàncies s'acumulin durant la segmentació.

La variació en superfície entre les prediccions i les màscares originals no és intrínsecament dolent, és possible que en certs casos sigui preferible seleccionar una zona de risc lleugerament més gran que la massa en si. Tampoc implica un baix rendiment quan considerem els resultats generals obtinguts.

L'existència de falsos positius (segmentació de regions considerades com masses però que no ho són) també és un dels punts que s'han de tractar. Algunes de les potencials causes dels falsos positius:

1. **Variabilitat en l'aparença de les masses:** les masses poden tenir gran variabilitat en forma i mida. Algunes masses poden tenir característiques similars a altres estructures presents al teixit del pit visible a les mamografies. Alguns exemples poden ser el teixit dens o estructures fibroglandulars.
2. **Artefactes:** l'existència d'artefactes invisibles a simple vista però visibles per al model, poden interferir amb la detecció i segmentació de les masses.
3. **Sensibilitat a patrons:** és possible que el model se sensibilitzi a certs patrons específics a masses, però que en realitzar les segmentacions generalitzi aquests a regions que no ho són, generant falsos positius.
4. **Limitacions del model:** és possible que el model no sigui capaç d'aprendre certes característiques de les masses, ja sigui per limitacions del conjunt de dades, la seva arquitectura o els seus paràmetres (profunditat, filtres...).

També s'han trobat casos on tenim falsos negatius. On el model no ha donat cap sortida quan en realitat si existeix alguna massa a la mamografia. Les causes poden ser similars a les citades anteriorment, però també afegim dues possibilitats més:

1. **Masses subtils:** certes masses poden no tenir el contrast o l'estructura necessària perquè el model les consideri com a masses, sent omeses pel mateix.
2. **Mida i resolució:** depenent de la mida i resolució de les masses, és possible trobar-se amb limitacions a l'hora de segmentar la zona per part del model.

Anteriorment, es van mostrar una sèrie de valors estadístics pertanyents a l'avaluació del model amb les dades de test. A la Taula 8 es mostren els valors estadístics de les prediccions sobre aquest conjunt, els quals mostren un rendiment promig sòlid en la tasca de segmentació.

Mitjana	Mediana	Desviació Estàndard	Màxim	Mínim
0.83	0.905	0.291	0.958	0.0

*Taula 8: Estadístiques de les prediccions sobre el conjunt de test.*

El valor promig per al coeficient de Dice és del 83%, indicant una superposició important entre les màscares originals i les prediccions. De forma general això suggereix que el model és capaç de capturar correctament la forma i posició a la mamografia de les masses.

La mediana del coeficient de Dice és del 90.5%. El que indica que un 50% dels resultats es troben per sobre i un 50% es troben per sota. Això ens indica que els coeficients obtinguts es troben desplaçats cap a la dreta, cap a valors de Dice elevats.

La desviació estàndard de 0.291 ens indica una certa variabilitat que hem de considerar. Aquesta variabilitat juntament amb el fet que les dades es troben desplaçades cap a valors alts, ens indica que trobem certa dispersió en els valors baixos.

El valor màxim de 95.8% en el coeficient de Dice ens indica un molt bon rendiment en almenys un cas, però el valor mínim de 0% ens indica que almenys un cas té una intersecció nul·la. Aquest últim cas fa interessant analitzar en detall que succeeix en els casos on el coeficient de Dice és baix, i que trobem en aquesta regió que produeix una dispersió elevada.

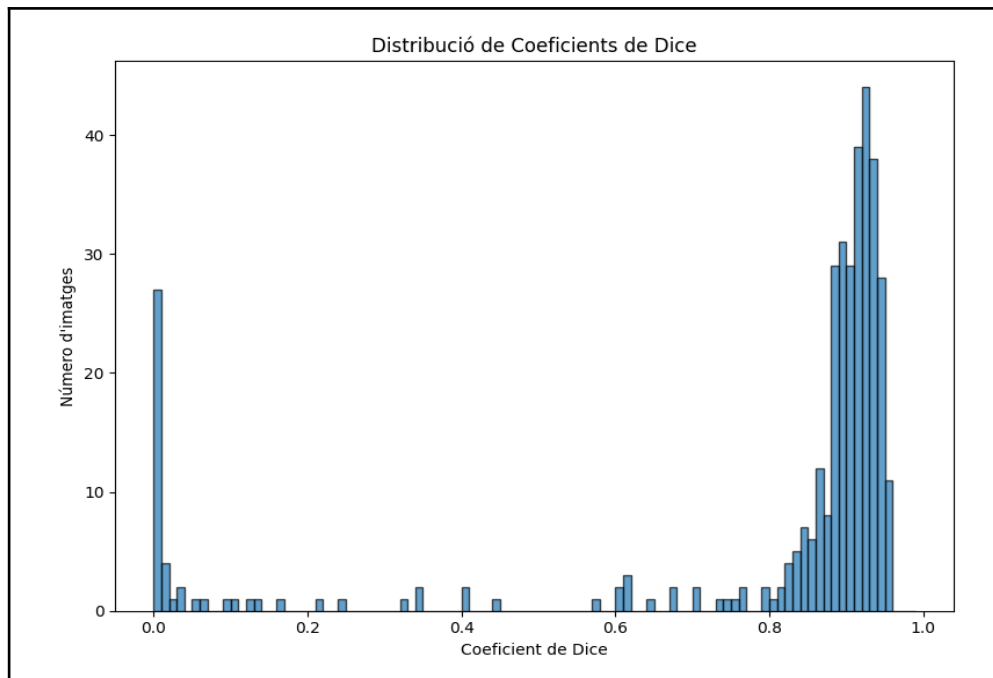


Figura 29: Nombre d'imatges trobades en intervals de 0.1 per coeficient de Dice.

A la Figura 29, es mostra de forma gràfica la distribució del nombre d'imatges en relació al coeficient de Dice. S'observa com una gran majoria de les imatges es troben en la part dreta del gràfic, dins l'interval [0.8, 1.0]. Però també trobem un pic important d'imatges on la intersecció obtinguda és zero. L'observació d'aquest fenomen va motivar la investigació d'aquest, amb l'objectiu d'entendre quins són els casos on el rendiment és tan baix i poder analitzar-ne les característiques per obtenir una resposta.

Del conjunt d'imatges de test, s'han extret aquelles amb un coeficient de Dice inferior al 10%. En total s'han extret 37 segmentacions amb aquest baix rendiment. Podem observar-ne un exemple a la Figura 30.

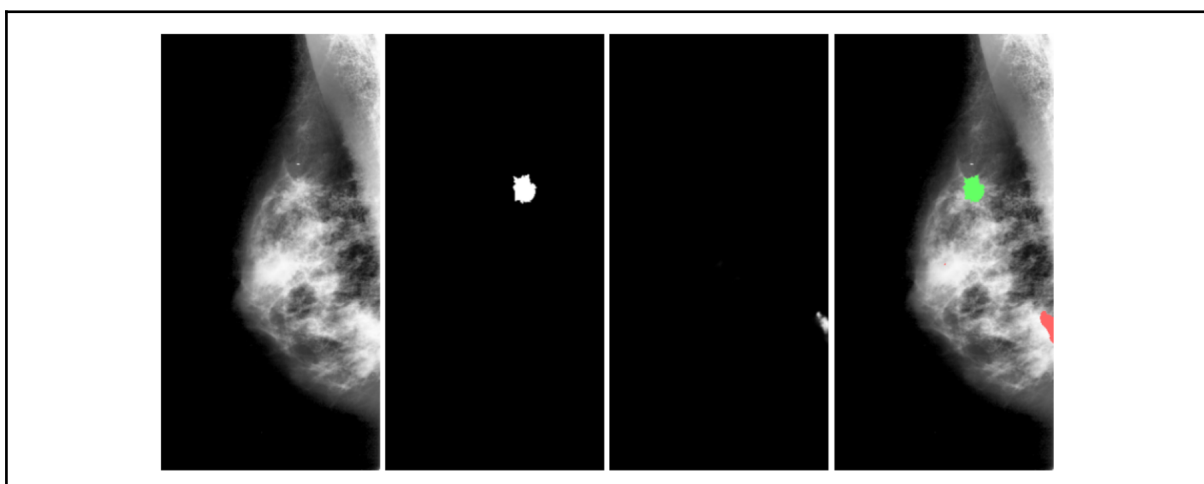


Figura 30: Exemple de segmentació amb coeficient de Dice inferior al 10%.

A continuació, utilitzant la llibreria pandas, s'han llegit els fitxers CSV de descripció del casos, seleccionant-ne els atributs d'identificador de pacient, de densitat del pit, de tipus de vista, forma de la massa i marges de la massa.

Utilitzant pandas s'ha fet un recompte de cada valor existent per a cada atribut per a tot el conjunt de dades. D'aquesta forma obtenim una perspectiva general de les característiques del conjunt de dades. De la mateixa forma s'ha realitzat el mateix procediment per a les 37 segmentacions esmentades anteriorment. Cal recordar que el nombre de casos no és igual al nombre de pacients, ja que cada pacient pot tenir més d'una mamografia, algunes amb més d'una massa. En total tenim 1696 mamografies al conjunt complet i 64 mamografies al subconjunt de baix rendiment.

A les Figures 31 i 32 observem les característiques del conjunt total de casos i del subconjunt de baix rendiment.

<pre> #_____All set data_____# Recuento para la columna 'breast_density': 2   757 3   449 1   337 4   153 Name: breast_density, dtype: int64  Recuento para la columna 'image view': MLO   912 CC    784 Name: image view, dtype: int64  Recuento para la columna 'mass shape': IRREGULAR                464 OVAL                     412 LOBULATED                384 ROUND                    164 ARCHITECTURAL_DISTORTION 103 IRREGULAR-ARCHITECTURAL_DISTORTION 52 LYMPH_NODE                35 FOCAL_ASYMMETRIC_DENSITY 25 ASYMMETRIC_BREAST_TISSUE 25 OVAL-LYMPH_NODE           6 LOBULATED-IRREGULAR       6 LOBULATED-LYMPH_NODE      4 ROUND-OVAL                 3 LOBULATED-ARCHITECTURAL_DISTORTION 2 IRREGULAR-FOCAL_ASYMMETRIC_DENSITY 2 IRREGULAR-ASYMMETRIC_BREAST_TISSUE 1 OVAL-LOBULATED            1 LOBULATED-OVAL            1 ROUND-IRREGULAR-ARCHITECTURAL_DISTORTION 1 ROUND-LOBULATED           1 Name: mass shape, dtype: int64  Recuento para la columna 'mass margins': CIRCUMSCRIBED            392 ILL_DEFINED               370 SPICULATED                363 OBSCURED                  247 MICROLOBULATED           129 ILL_DEFINED-SPICULATED   30 CIRCUMSCRIBED-ILL_DEFINED 29 OBSCURED-ILL_DEFINED     24 CIRCUMSCRIBED-OBSCURED   22 MICROLOBULATED-ILL_DEFINED 5 OBSCURED-ILL_DEFINED-SPICULATED 5 CIRCUMSCRIBED-OBSCURED-ILL_DEFINED 4 OBSCURED-SPICULATED      4 CIRCUMSCRIBED-MICROLOBULATED-ILL_DEFINED 3 CIRCUMSCRIBED-MICROLOBULATED 2 OBSCURED-CIRCUMSCRIBED   2 MICROLOBULATED-ILL_DEFINED-SPICULATED 2 MICROLOBULATED-SPICULATED 2 CIRCUMSCRIBED-SPICULATED 1 Name: mass margins, dtype: int64 </pre>	<pre> #_____Low Dice coefficient data_____# Recuento para la columna 'breast_density': 2   29 3   15 1   13 4    7 Name: breast_density, dtype: int64  Recuento para la columna 'image view': MLO   35 CC    29 Name: image view, dtype: int64  Recuento para la columna 'mass shape': OVAL                20 LOBULATED           19 IRREGULAR           15 ROUND                5 IRREGULAR-ARCHITECTURAL_DISTORTION 3 ARCHITECTURAL_DISTORTION 1 FOCAL_ASYMMETRIC_DENSITY 1 Name: mass shape, dtype: int64  Recuento para la columna 'mass margins': OBSCURED            20 CIRCUMSCRIBED       13 ILL_DEFINED          12 SPICULATED           9 MICROLOBULATED       5 OBSCURED-ILL_DEFINED 2 OBSCURED-ILL_DEFINED-SPICULATED 2 ILL_DEFINED-SPICULATED 1 Name: mass margins, dtype: int64 </pre>
<p><i>Figura 31: Recompte de casos per característiques de totes les mamografies.</i></p>	<p><i>Figura 32: Recompte de casos per característiques de les mamografies amb poca o nul·la intersecció.</i></p>

D'aquestes taules podem extreure la següent informació, calculant en tant per u la proporció de les característiques presents al conjunt complet i al subconjunt amb Dice baix.

	<b>Dataset complet</b>	<b>Subconjunt Dice baix</b>
Density 2	$757/1696 = 0.44$	$29/64 = 0.45$
Density 3	$449/1696 = 0.26$	$15/64 = 0.23$
Density 1	$337/1696 = 0.20$	$13/64 = 0.20$
Density 4	$153/1696 = 0.09$	$7/64 = 0.10$

*Taula 9: Proporció de la característica "breast\_density" al conjunt de dades total vs. mamografies amb baix rendiment de segmentació.*

	<b>Dataset complet</b>	<b>Subconjunt Dice baix</b>
CIRCUMSCRIBED	$392/1696 = 0.23$	$13/64 = 0.20$
ILL_DEFINED	$370/1696 = 0.21$	$12/64 = 0.18$
SPICULATED	$363/1696 = 0.21$	$9/64 = 0.14$
OBSCURED	$247/1696 = 0.14$	$20/64 = 0.31$
MICROLOBULATED	$129/1696 = 0.07$	$5/64 = 0.07$

*Taula 10: Proporció de la característica "mass\_margins" al conjunt de dades total vs. mamografies amb baix rendiment de segmentació.*

Les taules 9 i 10 ens mostren informació sobre diferents categories, analitzem l'última:

- Es divideix en les categories "CIRCUMSCRIBED", "ILL\_DEFINED", "SPICULATED", "OBSCURED" i "MICROLOBULATED".
- Per a cada categoria es mostren dues relacions, el nombre d'elements de la categoria sobre el nombre d'elements del conjunt complet, i el nombre d'elements de la categoria al subconjunt sobre la mida del subconjunt.

Aquests valors representen la proporció d'elements a cada categoria en relació amb el conjunt complet i al subconjunt.

Podem observar com la categoria “OBSCURED” representa el 14% dels elements al conjunt de dades complet, i sobre el 31% en el subconjunt de baix rendiment. Podem extreure dues conclusions:

1. La proporció molt més gran de casos “OBSCURED” al conjunt de rendiment baix pot indicar que els casos d’aquest subconjunt són més propensos a tenir característiques obstruïdes o amb manca de claredat en comparació a totes les dades. Això indicaria que el model mostra problemes a l’hora de segmentar casos on les masses es troben en condicions de baixa visibilitat a causa de la interferència d’altres teixits.
2. La importància de la categoria “OBSCURED” fa que sigui important donar un extra d’anàlisi a aquelles masses etiquetades amb aquesta categoria, ja que es tracta del punt feble del model desenvolupat. Això pot incloure l’aplicació d’altres tècniques per tractar els casos de forma especial.

## 6. Conclusions i treball futur

En aquest treball s'han obtingut bons resultats de segmentació utilitzant les mètriques d'avaluació de Dice i IOU juntament amb un model U-Net. Un cop obtinguts i explicats els resultats en els apartats anteriors, cal extreure les conclusions de la feina feta.

### 6.1. Conclusions generals del treball

L'objectiu principal d'aquest projecte ha estat explorar diferents models de xarxes neuronals profundes (utilitzant l'arquitectura U-Net) per a la segmentació automàtica de mamografies per millorar els procediments emprats en la detecció precoç del càncer de mama, una de les malalties més comunes i mortals en dones de tot el món. En la seva realització s'han pogut veure i experimentar amb els procediments i eines típics d'aquest camp de la imatge mèdica, el que ha permès tenir contacte directe amb el que podria ser un cas real de resolució d'un problema fent ús d'aprenentatge profund.

S'ha investigat sobre els antecedents de la segmentació, com les metodologies han evolucionat i quines són més destacades avui en dia. S'ha arribat a la conclusió que les tècniques d'aprenentatge profund són les preferides per les seves característiques, capacitat de generalització i nivell de precisió.

També s'han pogut analitzar les diverses dades disponibles en el món de la segmentació de mamografies, enfocant-se principalment en el conjunt CBIS-DDSM. Aquest conjunt de dades resulta de gran importància, ja que permet a qualsevol investigador fer ús de dades obertes, poder desenvolupar noves estratègies, tant per a la segmentació com per a altres tasques, i comparar-se amb altres mètodes que les utilitzen.

L'ús de tècniques d'aprenentatge profund i en concret de models de xarxes neuronals convolucionals és de vital importància en la investigació en tècniques de detecció i prevenció del càncer de mama. En concret la U-Net, arquitectura central del treball, demostra les seves capacitats de segmentació.

Respecte als resultats obtinguts, podem afirmar que les estratègies seleccionades per a la segmentació de les masses en mamografies han sigut correctes. Els resultats mostren una alta superposició entre les màscares originals i les prediccions fetes pel model, com ens indiquen els valors de Dice i IOU obtinguts. En caràcter general el model és capaç d'extreure les característiques de les masses i separar-les de la resta de teixit de les mamografies. Tot i això, també podem observar l'existència de falsos positius i falsos negatius. El que això indica és que existeixen reptes encara per tractar, i que poden ser l'objectiu de treballs futurs. A l'anàlisi dels resultats hem pogut veure algunes de les possibles causes d'aquestes inconsistències, i aquests poden ser el punt de focus principal més endavant.



## 6.2. Limitacions del treball i possibles direccions futures

En aquest apartat discutirem les limitacions del treball fet i es proposen possibles propostes de millora o idees que pot ser interessant explorar en el futur. Cap mètode és perfecte i és important investigar noves possibilitats.

Una de les principals limitacions està relacionada amb l'ús de recursos computacionals, com memòria RAM, VRAM i GPU, dels quals el mètode proposat pot requerir una quantitat significativa . Això pot arribar a limitar la seva aplicació en entorns de recursos limitats o situacions on l'accés als mateixos no és possible o consistent amb la seva necessitat.

La dependència d'un únic conjunt de dades específic es pot veure com una limitació. En concret hem utilitzat el conjunt de dades CBIS-DDSM tant per a l'entrenament com per a l'avaluació. Aquest conjunt de dades podria contenir certes característiques o esbiaixos que poden fer que el model no generalitzés bé en situacions on les entrades pertanyin a altres dades de característiques clíniques diferents (altres centres, màquines o protocols d'adquisició).

Les propostes de millora són diverses. En primer lloc, podem considerar el que hem comentat sobre l'ús amb pocs recursos i l'ús d'altres conjunts de dades per tal de millorar la generalització. Aquestes millores ja serien avenços importants, però no són les úniques millores que es poden aplicar.

Per tal de poder tractar les limitacions de memòria, s'haurien d'explorar diferents tipus de tècniques d'optimització i compressió, així com tècniques d'entrenament alternatives amb un enfocament en plataformes de recursos limitats. També seria interessant treballar en una plataforma amb més recursos, el que permetria per exemple utilitzar imatges de major resolució o desenvolupar i validar models més grans (nombre de capes i/o filtres).

Es poden explorar arquitectures avançades, com ara variacions de l'arquitectura U-Net. Tot i que l'estructura bàsica segueix sent la mateixa, existeixen variants que permeten tractar diferents problemes (com la detecció de patrons o detalls de menor resolució) de forma diferent i millorar el seu rendiment. Alguns exemples són U-Net++<sup>67</sup>, Attention

---

<sup>67</sup> "UNet++: A Nested U-Net Architecture for Medical Image Segmentation." 18 jul. 2018, <https://arxiv.org/abs/1807.10165>. Se consultó el 12 jun. 2023.

U-Net<sup>68</sup> i V-Net<sup>69</sup>. Aquesta última seria interessant sobretot si també es pogués tractar amb conjunts de mamografies 3D.

Altres idees que es poden explorar serien diferents tipus de tècniques d'augment de dades o preprocessat. En concret podria ser interessant l'aplicació d'un model previ a la tasca per tal d'eliminar la zona de múscul pectoral de la mamografia. També es pot investigar l'aplicació de la transferència d'aprenentatge (*transfer learning*), cosa que permetria utilitzar models prèviament entrenats per a adaptar-los al problema de la segmentació de mamografies. Per últim, també es podria experimentar amb imatges obtingudes amb altres tècniques d'imatge, com ara ressonància i ecografies.

---

---

<sup>68</sup> "Attention U-Net: Learning Where to Look for the Pancreas - arXiv." 11 abr. 2018, <https://arxiv.org/abs/1804.03999>. Se consultó el 12 jun. 2023.

<sup>69</sup> "V-Net: Fully Convolutional Neural Networks for Volumetric Medical ...." 15 jun. 2016, <https://arxiv.org/abs/1606.04797>. Se consultó el 12 jun. 2023.

## 7. Referències bibliogràfiques

- [1] "Screening for breast cancer with mammography - Cochrane." 4 jun. 2013, [https://www.cochrane.org/CD001877/BREASTCA\\_screening-for-breast-cancer-with-mammography](https://www.cochrane.org/CD001877/BREASTCA_screening-for-breast-cancer-with-mammography). Se consultó el 5 jun. 2023.
- [2] Ronneberger, O., Fischer, P., & Brox, T. (2015a, May 18). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. ArXiv.org. <https://arxiv.org/abs/1505.04597>
- [3] "Curated Breast Imaging Subset of Digital Database for Screening ...." 21 jun. 2016, <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629>. Se consultó el 5 jun. 2023.
- [4] "Curated Breast Imaging Subset of Digital Database for Screening ...." 21 jun. 2016, <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629>. Se consultó el 5 jun. 2023.
- [5] "Effect of image binarization thresholds on breast cancer ...." <https://www.sciencedirect.com/science/article/pii/S2667305321000351>. Se consultó el 5 jun. 2023.
- [6] "University of South Florida Digital Mammography Home Page DDSM." <http://www.eng.usf.edu/cvprg/mammography/database.html>. Se consultó el 11 jun. 2023.
- [7] "Curated Breast Imaging Subset of Digital Database for Screening ...." 21 jun. 2016, <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629>. Se consultó el 5 jun. 2023.
- [8] Ronneberger, O., Fischer, P., & Brox, T. (2015, May 18). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. ArXiv.org. <https://arxiv.org/abs/1505.04597>
- [9] "Grand Challenge: Home." <https://isbi-aida.grand-challenge.org/>. Se consultó el 5 jun. 2023.
- [10] Michael, E., Ma, H., Li, H., Kulwa, F., & Li, J. (2021). Breast Cancer Segmentation Methods: Current Status and Future Potentials. *BioMed Research International*, 2021, 1–29. <https://doi.org/10.1155/2021/9962109> <https://www.hindawi.com/journals/bmri/2021/9962109/>
- [11] Wu, Y.-T., Zhou, C., Chan, H.-P., Paramagul, C., Hadjiiski, L. M., Daly, C. P., Douglas, J. A., Zhang, Y., Sahiner, B., Shi, J., & Wei, J. (2009). Dynamic multiple thresholding breast boundary detection algorithm for mammograms. *Medical Physics*, 37(1), 391–401. <https://doi.org/10.1118/1.3273062>
- [12] "Amorphous calcifications (breast) | Radiology Reference Article." 25 sept. 2021, <https://radiopaedia.org/articles/amorphous-calcifications-breast?lang=us>. Se consultó el 9 abr. 2023.
- [13] "Region-Based Semantic Segmentation with End-to-End Training." 17 sept. 2016, [https://link.springer.com/chapter/10.1007/978-3-319-46448-0\\_23](https://link.springer.com/chapter/10.1007/978-3-319-46448-0_23). Se consultó el 5 jun. 2023.

- [14] "A new conditional region growing approach for microcalcification ...." 24 jul. 2021, <https://link.springer.com/article/10.1007/s11517-021-02379-x>. Se consultó el 5 jun. 2023.
- [15] Wang, X., Wang, S., Zhu, Y., & Meng, X. (2012). Image segmentation based on Support Vector Machine. *International Conference on Computer Science and Network Technology*. <https://doi.org/10.1109/iccst.2012.6525921>  
<http://ieeexplore.ieee.org/document/6525921/>
- [16] "SVM for Image Segmentation? - Cross Validated - Stack Exchange." 25 abr. 2012, <https://stats.stackexchange.com/questions/27085/svm-for-image-segmentation>. Se consultó el 12 jun. 2023.
- [17] "Thresholding in Image Processing - YouTube." 14 jun. 2020, <https://www.youtube.com/watch?v=vtbdqq7yAcc>. Se consultó el 5 jun. 2023.
- [18] Ronneberger, O., Fischer, P., & Brox, T. (2015, May 18). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. ArXiv.org. <https://arxiv.org/abs/1505.04597>
- [19] "Understanding Neural Networks - Prince Canuma." 13 feb. 2019, <https://prince-canuma.medium.com/understanding-neural-networks-22b29755abd9>. Se consultó el 5 jun. 2023.
- [20] "Understanding Neural Networks - Prince Canuma." 13 feb. 2019, <https://prince-canuma.medium.com/understanding-neural-networks-22b29755abd9>. Se consultó el 5 jun. 2023.
- [21] Bertels, J., Robben, D., Lemmens, R., & Vandermeulen, D. (2022, November 17). *Convolutional neural networks for medical image segmentation*. ArXiv.org. <https://doi.org/10.48550/arXiv.2211.09562>
- [22] Ronneberger, O., Fischer, P., & Brox, T. (2015, May 18). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. ArXiv.org. <https://arxiv.org/abs/1505.04597>
- [23] Long, J., Shelhamer, E., & Darrell, T. (2014). *Fully Convolutional Networks for Semantic Segmentation*. ArXiv.org. <https://arxiv.org/abs/1411.4038>
- [24] He, K., Zhang, X., Ren, S., & Sun, J. (2015, December 10). *Deep Residual Learning for Image Recognition*. ArXiv.org. <https://arxiv.org/abs/1512.03385>
- [25] Ronneberger, O., Fischer, P., & Brox, T. (2015a, May 18). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. ArXiv.org. <https://arxiv.org/abs/1505.04597>
- [26] Ronneberger, O., Fischer, P., & Brox, T. (2015a, May 18). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. ArXiv.org. <https://arxiv.org/abs/1505.04597>
- [27] "Intuitive Explanation of Skip Connections in Deep Learning." 23 mar. 2020, <https://theaisummer.com/skip-connections/>. Se consultó el 22 may. 2023.

- [28] "Data Augmentation | Papers With Code." <https://paperswithcode.com/task/data-augmentation>. Se consultó el 12 jun. 2023.
- [29] "University of South Florida Digital Mammography Home Page DDSM." <http://www.eng.usf.edu/cvprg/mammography/database.html>. Se consultó el 22 may. 2023.
- [30] "Curated Breast Imaging Subset of Digital Database for Screening ...." 16 may. 2023, <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629>. Se consultó el 22 may. 2023.
- [31] "TensorFlow.org." <https://www.tensorflow.org/>. Se consultó el 5 jun. 2023.
- [32] "Keras: Deep Learning for humans." <https://keras.io/>. Se consultó el 5 jun. 2023.
- [33] "karolzak/keras-unet: Helper package with multiple U-Net ... - GitHub." <https://github.com/karolzak/keras-unet>. Se consultó el 5 jun. 2023.
- [34] "visualkeras for Keras / TensorFlow - GitHub." <https://github.com/paulgavrikov/visualkeras>. Se consultó el 5 jun. 2023.
- [35] "pandas - Python Data Analysis Library." <https://pandas.pydata.org/>. Se consultó el 13 jun. 2023.
- [36] "Kaggle: Your Machine Learning and Data Science Community." <https://www.kaggle.com/>. Se consultó el 12 jun. 2023.
- [37] Lee, R. S., Gimenez, F., Hoogi, A., Miyake, K. K., Gorovoy, M., & Rubin, D. L. (2017). A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data*, 4(1). <https://doi.org/10.1038/sdata.2017.177>
- [38] "Mammographic Image Analysis Society (MIAS) database v1.21." 28 ago. 2015, <https://www.repository.cam.ac.uk/handle/1810/250394>. Se consultó el 26 may. 2023.
- [39] "Comprehensive Cancer Information - NCI." <https://www.cancer.gov/>. Se consultó el 26 may. 2023.
- [40] "National Institutes of Health (NIH) | Turning Discovery Into Health." <https://www.nih.gov/>. Se consultó el 26 may. 2023.
- [41] Lee, R. S., Gimenez, F., Hoogi, A., Miyake, K. K., Gorovoy, M., & Rubin, D. L. (2017). A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data*, 4(1). <https://doi.org/10.1038/sdata.2017.177>
- [42] "CBIS-DDSM Dataset | Papers With Code." <https://paperswithcode.com/dataset/cbis-ddsm>. Se consultó el 26 may. 2023.
- [43] Lee, R. S., Gimenez, F., Hoogi, A., Miyake, K. K., Gorovoy, M., & Rubin, D. L. (2017). A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data*, 4(1). <https://doi.org/10.1038/sdata.2017.177>

- [44] Lee, R. S., Gimenez, F., Hoogi, A., Miyake, K. K., Gorovoy, M., & Rubin, D. L. (2017). A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data*, 4(1). <https://doi.org/10.1038/sdata.2017.177>
- [45] "National Biomedical Imaging Archive - NBIA - NCI Wiki." <https://wiki.nci.nih.gov/display/NBIA>. Se consultó el 27 may. 2023.
- [46] "NBIA Data Retriever FAQ - TCIA Online Help." 17 oct. 2022, <https://wiki.cancerimagingarchive.net/display/NBIA/NBIA+Data+Retriever+FAQ>. Se consultó el 27 may. 2023.
- [47] "Curated Breast Imaging Subset of Digital Database for Screening ...." 21 jun. 2016, <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629>. Se consultó el 27 may. 2023.
- [48] "DICOM Standard." <https://www.dicomstandard.org/>. Se consultó el 31 may. 2023.
- [49] Lin, C., Wang, K., Chen, H., Xu, Y., Pan, T., & Chen, Y. (2022). Specimen mammography for intraoperative margin assessment in breast conserving surgery: a meta-analysis. *Scientific Reports*, 12(1), 18440. <https://doi.org/10.1038/s41598-022-23234-5>
- [50] "Segmenting Abnormalities in Mammograms (Part 1 of 3) | by Cleon W." 25 feb. 2021, <https://towardsdatascience.com/can-you-find-the-breast-tumours-part-1-of-3-1473ba685036>. Se consultó el 27 may. 2023.
- [51] "Segmenting Abnormalities in Mammograms (Part 1 of 3) | by Cleon W." 25 feb. 2021, <https://towardsdatascience.com/can-you-find-the-breast-tumours-part-1-of-3-1473ba685036>. Se consultó el 27 may. 2023.
- [52] "OpenCV modules." <https://docs.opencv.org/4.x/>. Se consultó el 30 may. 2023.
- [53] "Otsu's Thresholding Technique - LearnOpenCV." 5 ago. 2020, <https://learnopencv.com/otsu-thresholding-with-opencv/>. Se consultó el 31 may. 2023.
- [54] "Image Contrast Enhancement Using CLAHE - Analytics Vidhya." 17 ago. 2022, <https://www.analyticsvidhya.com/blog/2022/08/image-contrast-enhancement-using-clahe/>. Se consultó el 12 jun. 2023.
- [55] "visualkeras for Keras / TensorFlow - GitHub." <https://github.com/paulgavrikov/visualkeras>. Se consultó el 1 jun. 2023.
- [56] Ioffe, S. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. <https://arxiv.org/pdf/1502.03167.pdf>
- [57] "Metrics to Evaluate your Semantic Segmentation Model | by Ekin Tiu." 9 ago. 2019, <https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2>. Se consultó el 5 jun. 2023.

- [58] "Metrics to Evaluate your Semantic Segmentation Model | by Ekin Tiu." 9 ago. 2019, <https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2>. Se consultó el 5 jun. 2023.
- [59] "Stochastic Gradient Descent — Clearly Explained !!" 6 sept. 2019, <https://towardsdatascience.com/stochastic-gradient-descent-clearly-explained-53d239905d31>. Se consultó el 5 jun. 2023.
- [60] "Stochastic Gradient Descent — Clearly Explained !!" 6 sept. 2019, <https://towardsdatascience.com/stochastic-gradient-descent-clearly-explained-53d239905d31>. Se consultó el 31 may. 2023.
- [61] "Gentle Introduction to the Adam Optimization Algorithm for Deep ...." 3 jul. 2017, <https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/>. Se consultó el 5 jun. 2023.
- [62] "Adam — latest trends in deep learning optimization.." 22 oct. 2018, <https://towardsdatascience.com/adam-latest-trends-in-deep-learning-optimization-6be9a291375c>. Se consultó el 31 may. 2023.
- [63] Baccouche, A., Garcia-Zapirain, B., Castillo Olea, C., & Elmaghraby, A. S. (2021). Connected-UNets: a deep learning architecture for breast mass segmentation. *Npj Breast Cancer*, 7(1). <https://doi.org/10.1038/s41523-021-00358-x>
- [64] Al-antari, M. A., Al-masni, M. A., Choi, M.-T., Han, S.-M., & Kim, T.-S. (2018). A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. *International Journal of Medical Informatics*, 117, 44–54. <https://doi.org/10.1016/j.ijmedinf.2018.06.003>
- [65] Zhu, W., Xiang, X., Tran, T. D., Hager, G. D., & Xie, X. (2017, December 25). *Adversarial Deep Structured Nets for Mass Segmentation from Mammograms*. ArXiv.org. <https://doi.org/10.48550/arXiv.1710.09288>
- [66] "DoD BCRP Mammography Datasets at USF." <http://www.eng.usf.edu/cvprg/mammography/DDSM/BCRP/bcrp.html>. Se consultó el 3 jun. 2023.
- [67] Zhou, Z., Rahman, M., Tajbakhsh, N., & Liang, J. (2018). *UNet++: A Nested U-Net Architecture for Medical Image Segmentation*. ArXiv.org. <https://arxiv.org/abs/1807.10165>
- [68] Oktay, O., Schlemper, J., Folgoc, Loic Le, Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, Nils Y, Kainz, B., Glocker, B., & Rueckert, D. (2018). *Attention U-Net: Learning Where to Look for the Pancreas*. ArXiv.org. <https://arxiv.org/abs/1804.03999>
- [69] Milletari, F., Navab, N., & Ahmadi, S.-A. (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *ArXiv:1606.04797 [Cs]*. <https://arxiv.org/abs/1606.04797>
-

## 8. Annex

Es pot consultar el codi utilitzat en el treball al repositori següent:

<https://github.com/computingcreeper/TFG-Segmentacio-de-mamografies-utilitzant-tecniques-daprenentatge-profund.git>