

  
Universitat de Girona  
**Escola Politècnica Superior**

## **Treball final de grau**

**Estudi: Grau en Tecnologies Industrials**

**Títol: Síntesi d'imatges de RM per a la millora en la segmentació de lesions d'esclerosi múltiple en imatges de ressonància magnètica.**

**Document:** Memòria

**Alumne:** Arnau Cabarrocas Salvador

**Tutor:** Arnau Oliver i Jordi Freixenet

**Departament:** Arquitectura i tecnologia de computadors

**Àrea:** ATC

**Convocatòria (9/2022)**



## **AGRAÏMENTS**

A totes les meves amistats i companys, pel seu interès i alleugeriment del camí que m'han acompanyat de principi a fi.

Als meus tutors, l'Arnau Oliver i en Jordi Freixenet, per l'assessorament al llarg del recorregut, l'orientació, i la seva confiança.

I com no, la família: Els meus pares i a la meva germana per el seu recolzament i estima inqüestionables.

A tots vosaltres: Moltes gràcies.



## ÍNDEX DE CONTINGUTS

Índex de continguts.....	i
1 Introducció .....	1
1.1 Antecedents.....	1
1.2 Objecte .....	1
1.3 Abast .....	1
2 Esclerosi múltiple (EM): .....	3
2.1 Imatges de ressonància magnètica (IRM) per detectar EM: .....	4
2.1.1 Què és l'IRM? .....	4
2.1.2 Perquè fer servir IRM? .....	5
3 Estat del art en la segmentació d'EM.....	7
3.1 Segmentació de lesions d'EM.....	7
3.2 Estratègies per la segmentació de lesions. ....	8
3.2.1 Estratègies supervisades. ....	8
3.2.2 Estratègies no supervisades .....	8
3.3 Mesures d'avaluació. ....	9
3.4 Conclusions.....	10
3.5 Mètodes de síntesi d'imatges.....	11
4 Base teòrica mètode segmentació. ....	13
4.1 Xarxes Neuronals Artificials: .....	13
4.1.1 Arquitectura: .....	13
4.1.2 Entrenament model.....	18
4.1.3 Prediccions model. ....	21
4.2 Xarxa neuronal convolucional .....	22
4.3 U-Net. ....	25
5 Modelització.....	27
5.1 Estructura U-Net. ....	27
5.2 Híper-paràmetres. ....	28
5.3 Bases de Dades. ....	30

6	Skull Stripping.....	31
6.1	Pretractament. ....	31
6.2	Entrenament .....	33
6.3	Resultats.....	34
6.3.1	Resultats mostra.....	34
6.3.2	Mètode binari.....	34
6.3.3	Mètode semi-binari.....	37
6.3.4	Mètode no-binari. ....	38
6.4	Resum resultats mètodes.....	39
6.5	Resultats totals .....	39
7	Segmentació de lesions .....	43
7.1	Pretractament dades. ....	43
7.2	Entrenament. ....	47
7.3	Resultats.....	48
7.3.1	Resultats pretractament T1-w amb rotació i funció Binary Crossentropy. 48	
7.3.2	Resultats pretractament T1-w i funció Binary Crossentropy. ....	50
7.3.3	Resultats pretractament T1-w amb rotació i funció Dice Loss. ....	51
7.3.4	Resultats pretractament T1-w i funció Dice Loss. ....	52
7.3.5	Resultats sense Skull Stripping. ....	53
7.4	Resum resultats. ....	55
8	Adaptabilitat mètodes segmentació.....	57
8.1	Pretractament de dades. ....	57
8.2	Resultats.....	58
8.2.1	Resultats combinació A:.....	60
8.2.2	Resultats combinació B: .....	62
8.2.3	Resultats combinació C: .....	63
8.2.4	Resultats combinació D:.....	64
8.2.5	Resultats combinació E: .....	65
8.3	Resum resultats. ....	66
8.4	Síntesis d'imatges. ....	66
8.4.1	Pretractament.....	67
8.4.2	Entrenament.....	68

## Índex de continguts

8.4.3	Resultats síntesis Vall d'Hebron. ....	69
8.4.4	Resultats síntesis Trueta. ....	70
8.4.5	Conclusions. ....	71
9	Resum del pressupost .....	73
10	Conclusions .....	75
11	Futurs estudis .....	79
12	Referències .....	81





# 1 INTRODUCCIÓ

## 1.1 Antecedents

La segmentació d'imatges consisteix en la separació d'una imatge en diferents parts d'interès, i té un rol important en l'anàlisi d'aquestes. Per exemple, en el cas d'imatges mèdiques, aquests algorismes permeten ajudar als metges a detectar, diagnosticar i fer el seguiment dels pacients. Els últims avenços en Deep Learning han permès la possibilitat d'aplicar mètodes de xarxes neuronals convolucionals (CNN) en imatges mèdiques, obtenint uns excel·lents resultats en diferents aplicacions. Un exemple és l'aplicació de mètodes CNN per la segmentació d'imatges de ressonància magnètica (RM) cerebrals. Tot i això, encara hi ha importants limitacions per a la generalització d'aquestes tècniques, essent el principal problema la pèrdua de sensibilitat dels models entrenats amb un tipus d'escàner quan es proven en imatges provinents d'altres escàners (problema conegut com adaptació de domini).

## 1.2 Objecte

L'objectiu d'aquest Treball Final de Grau, és aconseguir millorar la segmentació de lesions d'esclerosi múltiple en imatges de ressonància magnètica. Per tal d'aconseguir aquest objectiu, el TFG es divideix en tres tasques principals.

## 1.3 Abast

Primerament s'estudiaran les tècniques actuals d'estratègies de Deep Learning del camp de síntesi d'imatges (generació de noves imatges).

El segon apartat consisteix en implementar i avaluar l'aplicabilitat dels mètodes anteriors per resoldre el problema de l'adaptació del domini d'imatges entre diversos escàners de RM.

L'últim apartat consisteix en l'aplicació dels millors mètodes analitzats per a la millora de les tècniques existents de segmentació de lesions cerebrals d'esclerosi múltiple. S'avaluaran els resultats dels algorismes de forma quantitativa i qualitativa.

Les dades que s'utilitzaran en aquest TFG provindran de reptes internacionals (60 casos del WMH Segmentation Challenge) i dels hospitals col·laboradors, Trueta, Vall d'Hebron i Clínica Girona (15 casos per hospital). En ambdós casos tenim imatges

Síntesi d'imatges de RM per la millora en la segmentació de lesions d'esclerosi múltiple

corresponents, respectivament, a , escàners Philips, Siemens i GE. Els entorns de programació que s'utilitzaran seran Matlab i Python

Esclerosi múltiple (EM):

## 2 ESCLEROSI MÚLTIPLE (EM):

L'esclerosi múltiple (EM) és la causa més comuna de discapacitat no traumàtica en afectar als joves. El nombre d'afectats en el món s'estima que es troba entre 1,3 i 2,5 milions de persones. A Catalunya els afectats per la malaltia són més de 6.500. Aquests números estan en creixement, ja que cada any tendeixen en augmentar el nombre de persones afectades. L'esclerosi múltiple afecta entre al doble i triple de dones que d'homes, l'índex d'incidència és molt baix en infants, però augmenta ràpidament passats els 18 anys, arribant així al seu pic entre els 25 i 35 anys a on comença a baixar lentament, trobant així valors molt baixos de noves incidències, passats els 50 anys.

L'EM és una malaltia neuronal crònica autoimmune, degenerativa i inflamatòria-desmielinitzant del sistema nerviós central (SNC). Caracteritzada patològicament per àrees d'inflamació, desmielinització, pèrdua axonal i gliosi repartides per el SNC, causant a vegades impediments motors, sensorials, visuals, cognitius i coordinatius. Respecte a les causes que fan que s'iniciï la malaltia, no estan gaire clares. S'ha observat que es produeix per una resposta autoimmune anòmala en individus amb certa predisposició genètica, sobre el quals certs factors ambientals o virus, com uns nivells baixos de vitamina D, el tabac, un índex de massa corporal elevat durant l'adolescència o el virus d'Epstein-Barr, podrien influir en el desenvolupament i progrés de la malaltia.

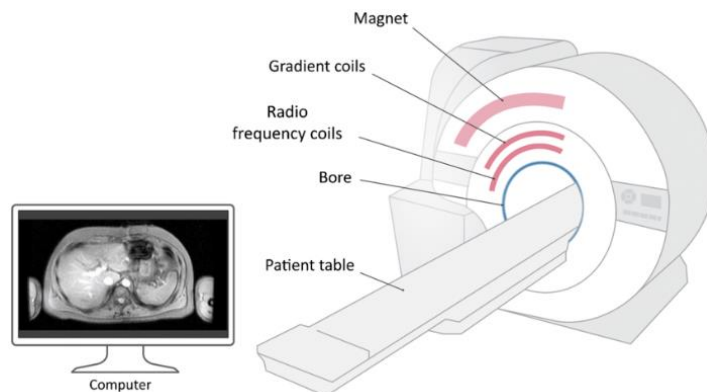
Els pacients poden ser agrupats en quatre categories diferents segons el procés de la malaltia:

- Recurrent-remitent (EMRR): És el tipus més comú, els símptomes apareixen en forma de brots, poden durar des de dies fins a mesos, depenent del cas. Després d'aquest període de temps, els símptomes poden millorar parcial o totalment al produir-se remielinització, baixar la inflamació i altres mecanismes de compensació, els quals van essent menys eficients després d'atacs recurrents.
- Secundària progressiva (EMSP): Alguns casos d'EMRR amb el pas dels anys pateixen un empitjorament progressiu que pot estar o no relacionat amb els brots, la recuperació sol ser incompleta però l'empitjorament pot ser molt lent, cosa que també el fa difícil de detectar.
- Primària progressiva (EMPP): Se sol diagnosticar a edats més avançades, no conté recurrències ni millores, però els símptomes continuen empitjorant des de l'inici.
- Progressiva-recurrent (EMPR): És una variant bastant inusual, on els símptomes són progressius des de l'inici però al llarg del temps es van produint brots que empitjoren els símptomes.

## 2.1 Imatges de ressonància magnètica (IRM) per detectar EM:

### 2.1.1 Què és l'IRM?

La ressonància magnètica és un tipus d'escàner que utilitza forts camps magnètics i ones radio per produir imatges detallades de l'interior del cos.



*Figura 1: Imatge esquemàtica dels principals components d'un escàner de ressonància magnètica.*

Tal com podem veure a la Figura 1 l'escàner consta dels següents components principals:

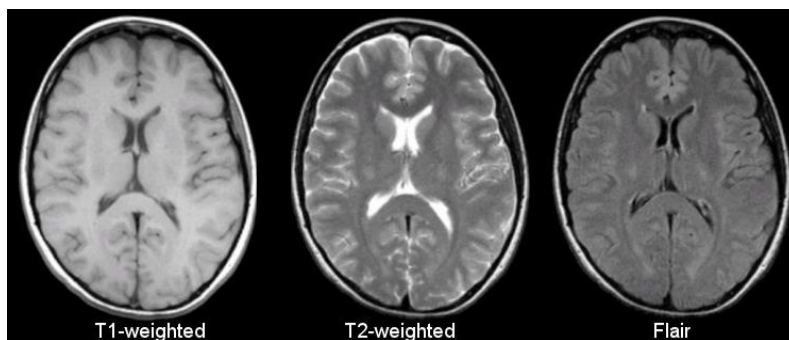
- Un gran imant (Magnet) superconductor per generar un camp magnètic homogeni mitjançant bobines.
- Bobines per generar canvis lineals en el camp magnètic per poder generar senyals detectables en localitzacions específiques del cos, permetent així la creació d'imatges concretes (Gradient coils).
- Unes bobines de radio-freqüència (Radio frequency coils) que transmeten senyals radio a la part del cos que es vol escanejar.
- Un ordinador (Computer) que reconstrueix les senyals radio en una imatge final, normalment mitjançant transformades de Fourier.

Els escàners de ressonància magnètica fan servir les propietats magnètiques intrínseques del cos per produir imatges detallades de qualsevol part del cos. Per fer les imatges s'utilitzen nuclis d'hidrogen (protons), aquests protons d'hidrogen poden vincular-se al planeta terra, rotant sobre el seu propi eix amb un pol nord-sud, com si fos un imant de barra. El seu comportament és normalment totalment aleatori, amb els eixos alineats aleatòriament. Quan el cos es posa sota un gran camp magnètic, tots els eixos s'alineen. Aquest moviment provoca la generació d'un vector magnètic orientat en el sentit del imant (l'escàner). Si en el camp magnètic s'hi afegeix energia addicional que el modifica, el vector es mou i les ones de radio-freqüència ressonen amb els nuclis. Quan s'aturen les senyals radio, els nuclis tornen al estat inicial i això crea una senyal que és la que s'utilitza per crear les imatges de ressonància magnètica.

Esclerosi múltiple (EM):

Hi ha certs paràmetres que fa que obtinguem diferents imatges resultants. Per exemple podem fer un pols d'ones que afecti a menys protons, així rebaixant si ens interessa la densitat de protons detectats. També en certs casos s'utilitzen polsos d'ones en seqüència, per tal de donar èmfasi a certs teixits o anomalies. També es tenen en consideració els temps que tarden en relaxar-se els protons. El temps que tarda el vector en tornar a l'estat de repòs (eixos alineats) s'anomena relaxació T1 i el temps que tarden els eixos en tornar a tenir el spin de l'estat de repòs, s'anomena relaxació T2.

Les seqüències de polsos d'ones més freqüents són la T1-w, T2-w, i FLAIR. Aquestes es defineixen a partir de la variació de dos temps, el temps de repetició entre dos polsos (TR) i el temps entre el pols inicial i el seu eco (TE). La seqüència T1-w fa servir uns TR i TE curts. La T2-w fa servir uns temps TR i TE llargs. La FLAIR és una variació d'aquestes seqüències on aconseguim, canviant el tipus de la senyal, un major contrast entre el teixit i les lesions (Figura 2).



*Figura 2: Diferents tipus d'imatges obtingudes en un mateix cas.*

Moltes malalties es manifesten amb un increment en la quantitat d'aigua, així que l'escàner IRM és un test bastant bo per la detecció de malalties, però li és més difícil determinar la naturalesa de la patologia, per exemple, una infecció i un tumor en alguns casos pot semblar molt similar, així requerint la supervisió d'un radiòleg per obtenir la resposta correcta.

### **2.1.2 Perquè fer servir IRM?**

Les tècniques convencionals de ressonància magnètica són altament sensibles en la detecció de plaques d'EM i poden proporcionar valors quantitius sobre les inflamacions i lesions. Això ha fet que les mètriques derivades de les IRM es considerin les eines para-clínicas més importants a l'hora de diagnosticar, entendre i monitoritzar l'EM, així com l'eficiència dels seus tractaments.

En una imatge de ressonància magnètica, podem visualitzar aquestes lesions produïdes per l'EM en forma de petites taques clares. També ens permeten visualitzar la matèria

blanca (capa externa del cervell), la matèria gris (capa interna del cervell) i el fluid espinal (líquid que envolta el cervell i la espina dorsal).

Per l'estudi de l'anàlisi quantitatiu de les lesions, s'han fet aproximacions manuals o semi-automàtiques de segmentacions de diverses imatges de ressonància magnètica, per tal de computar el numero total de lesions i el volum d'aquestes. El problema és que aquest procés és molt lent, degut a la gran quantitat d'imatges a analitzar per cada cas. També és susceptible a variacions degut a discrepàncies, al segmentar el mateix cas en un altre moment o per algun altre expert.

És per això que l'automatització d'aquest procés que permeti segmentar lesions de grans quantitats de dades permetria estalviar molt de temps i eliminaria la variabilitat, així que s'ha convertit en un tema d'interès i recerca actiu. Desafortunadament els mètodes aconseguits fins ara no donen un resultat equivalent a si s'hagués realitzat de manera manual per un expert. Això és degut en part a que els escàner pot arribar a solapar les lesions amb altres tipus de teixits. També trobem difícil eliminar les variacions produïdes per les discrepàncies entre experts ja que per automatitzar necessitem dades de resultats de segmentacions duts a terme per experts.

Si aconseguíssim crear un mètode automàtic de segmentació de lesions, no només ajudaria en estalvi de temps, sinó que també en millorar els diagnòstics i seguiments dels pacients, tant per estudis de recerca com en pacients del dia a dia.

*Per obtenir més informació es poden consultar les referències [10] [11] [12] [13] [14] [15] [16] [17]*

## **3 ESTAT DEL ART EN LA SEGMENTACIÓ D'EM.**

### **3.1 Segmentació de lesions d'EM.**

Aplicat al camp d'anàlisi d'imatges, definim segmentació com el procés de delimitar una regió d'unes certes característiques o propietats dins de la pròpia imatge. Un exemple seria la delimitació de les regions que formen els diversos teixits en una ressonància magnètica. Aquest procés és altament relacionable amb el concepte de detecció, que tracta de trobar un objecte en concret dins d'una imatge, per exemple, les lesions produïdes per l'esclerosis múltiple. Si només volem detectar cert objecte, només fa falta tenir un aproximat de la regió que aquest ocupa, però podem realitzar simultàniament la segmentació i la detecció per poder delimitar l'objecte que ens interessa.

Un altre concepte important per l'anàlisi d'imatges és la classificació, que és defineix com aquell procés d'assignació d'un objecte a certa categoria. Aquest concepte també és molt important en el camp de la intel·ligència artificial, a on la classificació implica un procés de dos etapes. En la primera etapa el model entrena amb un conjunt de dades ja classificades, i en la segona el model es posa a prova amb dades noves i les classifica. Traduït a imatges, el model pot classificar els diversos píxels així convertint la classificació en un problema de segmentació.

Una de les propietats més utilitzada per la segmentació de lesions a partir d'imatges és la intensitat dels vòxels (píxels volumètrics) d'una imatge, degut a que com hem pogut veure a la Figura 2 els diversos teixits cerebrals tenen diferents intensitats que permeten poder realitzar la segmentació segons la intensitat. A vegades per trobar les lesions podem fer servir combinacions de més d'un tipus d'imatge (T1-w, T2-w i FLAIR), amb aquest mètode provoquem un augment del nombre de característiques així fent més notòria la diferència entre els diversos teixits i facilitant la segmentació.

Altres mètodes que es fan servir, inclouen la informació espacial de les imatges en els seus algorismes. Normalment es fan servir els Markov Random Fields (MRF) per modelar les interaccions amb els veïns de cada vòxel. Altres alternatives són els mètodes de segmentació FC (fuzzy connectedness) o l'aparició d'un atlas probabilístic.

## **3.2 Estratègies per la segmentació de lesions.**

Com hem comentat prèviament, la delimitació manual de les lesions d'EM és difícil i consumeix molt de temps, deguda a la gran quantitat de dades a revisar per cada pacient. A més a més cada estudi pot tenir certa variabilitat segons qui i quant el faci, això fa que l'anàlisi requereixi d'algorismes com el STAPLE que fusionen els diversos anàlisis en un. Aquests problemes han fet que es vulguin buscar estratègies per la segmentació de lesions que no siguin manuals. Aquestes estratègies les podem classificar en dos grans blocs segons el nivell de supervisió: estratègies supervisades i estratègies no supervisades (Figura 3).

### **3.2.1 Estratègies supervisades.**

Les estratègies supervisades són aquelles que fan servir informació o coneixements previs per poder realitzar la segmentació. Aquest grup es pot subdividir en dos subapartats diferenciats. El primer es basa en un atlas d'informació que té en compte la informació espacial. La majoria d'atles són estadístics, però podem arribar a trobar atles topològics. Un atlas estadístic té en compte la probabilitat de cada vòxel de pertànyer a un tipus de teixit i està construït a partir d'un conjunt de segmentacions manuals. Al basar-se en probabilitats, aquest mètode continua funcionant molt bé davant de soroll. Per poder realitzar la segmentació, l'atlas s'ha de registrar amb les imatges de ressonància magnètica, provocant que la dificultat del mètode provingui de la necessitat d'alinear l'atlas amb les imatges i no tant de la segmentació. La segona estratègia supervisada es basa en l'aprenentatge a partir de la segmentació manual de les lesions, a diferència de la segmentació per atlas, aquesta no necessita un procés de registre.

Construir un atlas no és fàcil, i encara que ajudi a reduir el nombre de falsos positius en els resultats, el procés de registre es complica més com més lesions hi hagin. Per altre banda, la segmentació basada en aprenentatge es troba en el problema que la precisió de la segmentació depèn del tipus de dades d'entrada, si la imatge prové a partir de una altre seqüència de polsos o escàner, segurament haurem de tornar a entrenar un model amb dades similars al input o adaptar les dades per tal que s'assemblin a aquelles amb les que hem entrenat.

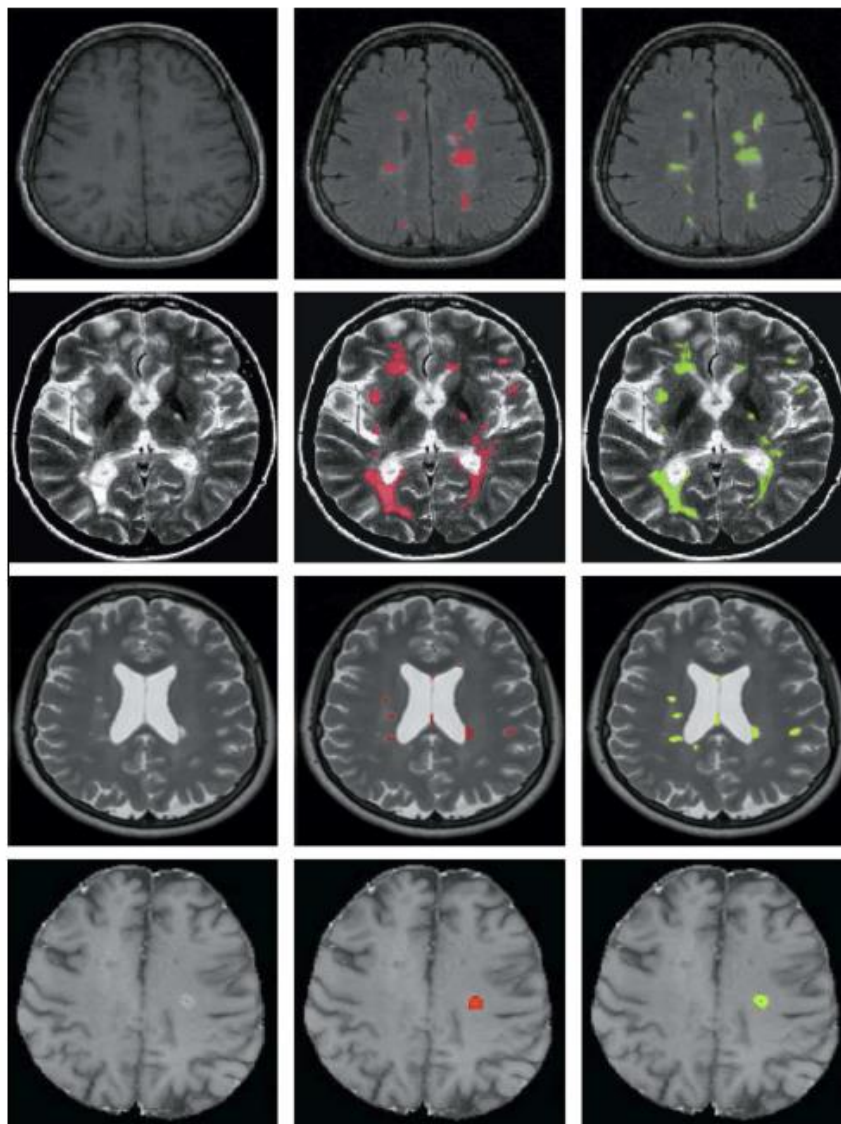
### **3.2.2 Estratègies no supervisades**

Les estratègies no supervisades són aquelles a on no s'hi apliquen coneixements previs, es poden classificar en dos grups segons el què segmenten, en el primer grup primerament es separen els teixits i acte seguit les lesions (o simultàniament). El segon grup directament separen les lesions.

La segmentació es realitza gracies a que és capaç de detectar que les imatges estan composades per diversos teixits. El primer mètode però al separar primer els teixits, la qualitat de la segmentació de les lesions depèn directament de la precisió amb la que els



ha separat. I el segon mètode té dificultats en detectar petites lesions, però assoleix una bona precisió en lesions de mitjanes i grans dimensions.



*Figura 3: Exemples de segmentacions d'esclerosi múltiple. D'esquerra a dreta, les columnes representen: la imatge original, la seva segmentació real i la segmentació que ha trobat el model. La primera fila representa els resultats d'un model supervisat amb atlas. La segona dona els resultats d'un model supervisat d'entrenament. Les dos últimes files representen la segmentació no supervisada basada en teixits i en lesions respectivament.*

### 3.3 Mesures d'avaluació.

Un cop segmentades les lesions, per poder avaluar els models, necessitem unes mètriques de referència per poder comparar-los. Per poder comparar, classifiquem els diversos vòxels en 4 categories: veritables positius (TP), veritables negatius (TN), falsos positius (FP) i falsos negatius (FN). L'objectiu és aconseguir el major nombre possible de valors veritables minimitzant alhora els valors falsos. Però això és idíl·lic, a la realitat el que passa és que a l'intentar augmentar els veritables positius, augmenten els falsos positius també, de la mateixa manera que intentar disminuir els falsos negatius, farà disminuir els veritables positius també. És per això que existeixen diverses

mètriques segons en el què volem centrar-nos. Les principals mètriques són: exactitud (Accuracy), precisió (Precision), sensibilitat (Recall), especificitat (Specificity), el coeficient de Dice Similarity (DSC) i el valor f1 (combinació de precisió i sensibilitat) (Taula 1).

Taula 1: Fórmules mètriques.

Accuracy	$\frac{ TN  +  TP }{ TN  +  TP  +  FP  +  FN }$
Precision	$\frac{ TP }{ TP  +  FP }$
Recall	$\frac{ TP }{ TP  +  FN }$
Specificity	$\frac{ TN }{ TN  +  FP }$
Dice Similarity Coefficient	$\frac{2 \times  TP }{2 \times  TP  +  FP  +  FN }$
F1 score	$\frac{2 \times Precision \times Recall}{Precision + Recall}$

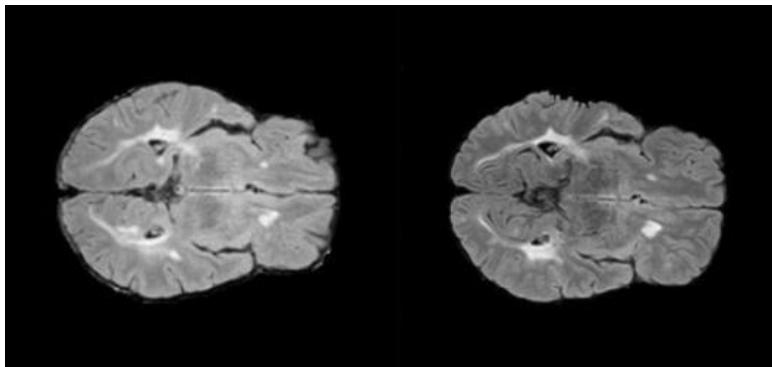
### 3.4 Conclusions.

Un cop definits els mètodes existents per la segmentació de lesions d'esclerosi múltiple, hem pogut veure com una de les dificultats que es comparteix és la falta de adaptabilitat deguda al gran numero diferent d'imatges de ressonància magnètica que podem obtenir si canviem les seqüències de polsos, el tipus d'imatge o l'escàner. Tot hi això, en cas de tenir un input similar a una base de dades resolta per un expert, ens surten molt més a compte i donen resultats millors els mètodes supervisats, ja que al tenir una sèrie de segmentacions similars ajuden a obtenir uns resultats més precisos. Si a més a més tenim primer definides les segmentacions reals dels teixits, podem aplicar una primera segmentació per eliminar tot allò que no ens interessi, i aconseguir tenir unes dades més netes a la hora de fer la segmentació de lesions. Sempre és molt important en els mètodes supervisats el fet de tenir un input net i que s'ajusti bastant al model que tenim, ja sigui un atlas o una base de dades.

Hi haurà molts casos en que degut a les diferències entre les imatges que provenen d'escànners diferents, no coincideixin l'escàner del qual volem segmentar-ne les lesions, i l'escàner del qual tenim les lesions segmentades, fent que l'input no s'assembli el suficient com perquè es pugui ajustar a les dades manualment segmentades, privant-nos de un resultat precís. En aquests casos, si volem fer servir un dels mètodes supervisats per obtenir més precisió, necessitarem trobar una base de dades nova o canviar d'alguna

Estat del art en la segmentació d'EM.

manera les dades d'entrada per ajustar-les a la base de dades que tenim a la nostre disposició. Aquest problema és molt usual trobar-lo quan tenim dades d'un escàner diferent als que hem entrenat (Figura 4), i una manera de solucionar-lo és a partir de la creació d'imatges sintètiques fruits de la transformació de la nostre imatge a una del mateix cas però de característiques similars a les del escàner desitjat.



*Figura 4: Mateix cas escanejat per l'hospital de Santa Catarina (esquerra) i Vall d'Hebron (dreta).*

### **3.5 Mètodes de síntesi d'imatges.**

El camp de la síntesi d'imatges és molt extens i divers, i ens és capaç de donar un mètode per solucionar el problema d'adaptabilitat que ens trobem davant de la utilització d'escàners diferents als manualment segmentats. Per poder acostar les imatges d'entrada a les del mètode supervisat, ens veiem en la necessitat de crear imatges sintètiques a partir d'una d'original. Aquest tipus de síntesis d'imatge, es pot realitzar de dos maneres diferents, classificats de la mateixa manera que els mètodes de segmentació d'imatges. Poden ser estratègies supervisades o no supervisades.

Les estratègies supervisades de síntesi d'imatges estan formades principalment per xarxes neuronals convolucionals. Concretament s'utilitzen les xarxes U-Net, les quals són usualment usades per la segmentació d'imatges, per a obtenir un model a base d'entrenar la xarxa a partir d'un conjunt de dades segmentades manualment. Aquestes xarxes també ens permeten la creació d'un model que transformi les nostres imatges.

Les estratègies no supervisades estan liderades per la utilització de les tècniques d'aprenentatge de les CycleGAN. Els mètodes CycleGAN, són capaços de trobar un mapa que uneixi dos tipus d'imatge, sense la necessitat de que aquestes estiguin aparellades. Per exemple, a partir de una base de dades de cavalls i una de zebres, és capaç de acabar convertint els cavalls perquè semblin zebres. A diferència de les xarxes U-Net, nosaltres no li hauríem indicat com seria un cavall que s'assemblés a una zebra (Figura 5).

## Síntesi d'imatges de RM per la millora en la segmentació de lesions d'esclerosi múltiple

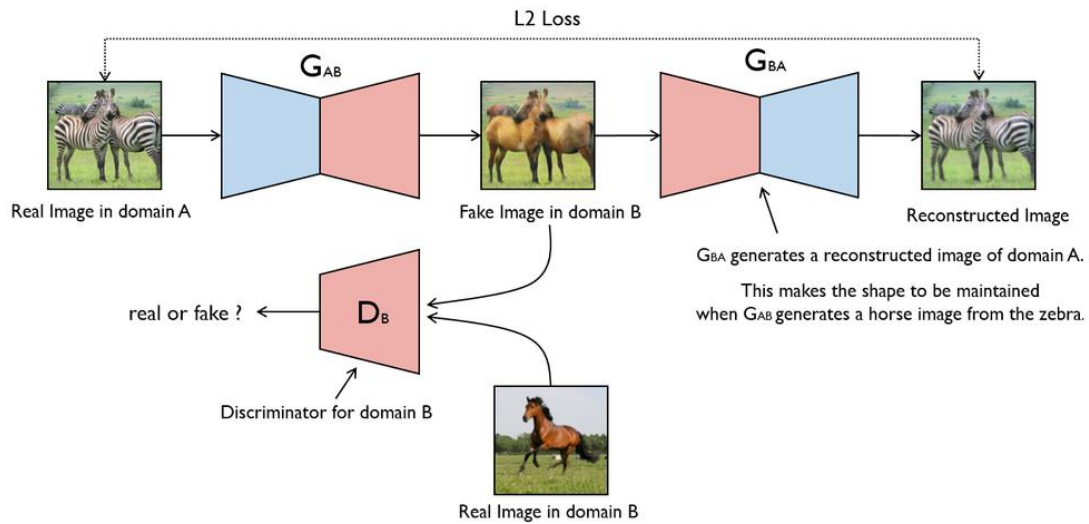


Figura 5: Esquema estructura sistema CycleGAN.

Entre aquestes dues tècniques, si disposem de dades emparellades, la U-Net aconseguirà donar més bons resultats que una CycleGAN, encara que si tenim poques dades, la U-Net tendeix a donar uns resultats més borrosos, igualment però, més encertats.

Per obtenir més informació es poden consultar les referències [3] [4] [5]

## **4 BASE TEÒRICA MÈTODE SEGMENTACIÓ.**

Com hem comentat en el capítol anterior, la segmentació automàtica de lesions d'EM és complexa de realitzar deguda a la vulnerabilitat dels mètodes davant de les diferències entre imatges. Per realitzar aquesta segmentació, si tenim una gran quantitat de segmentacions realitzades per experts que ens podem prendre com a dades verdaderes, hem vist que els mètodes supervisats funcionen millor que els no supervisats. Entre els mètodes supervisats, podem optar per segmentar a partir d'un atlas o mitjançant un model entrenat. En aquest treball, estudiarem i posarem en pràctica la segmentació supervisada mitjançant tècniques basades en l'aprenentatge profund, concretament, utilitzarem l'estructura U-Net. L'U-Net, a diferència d'altres tècniques d'aprenentatge, ens permet obtenir un output en forma d'imatge així permetent obtenir la imatge segmentada com a resultat final. Aquest mètode, com hem vist anteriorment és molt versàtil, ja que també ens permetrà realitzar més endavant la creació d'imatges sintètiques per poder solucionar el problema d'adaptabilitat que ens trobem davant dels diversos tipus d'escàners.

### **4.1 Xarxes Neuronals Artificials:**

L'estructura U-Net és una Xarxa Neuronal Convolucional (CNN) que es basa en el funcionament de codificador – descodificador i en el principi d'una xarxa neuronal artificial per poder donar els resultats.

Les xarxes neuronals artificials tenen aquest nom degut a les semblances que tenen amb una xarxa neuronal biològica. Són sistemes de xarxes composts per nodes anomenats neurones. Estan organitzats per capes que transmeten informació entre si per col·laborar i generar un resultat final a partir del processament de les dades d'entrada.

#### **4.1.1 Arquitectura:**

Les xarxes neuronals artificials estan conformades per dos estructures bàsiques, les capes i les neurones. Aquestes es relacionen a partir de certs paràmetres com els pesos, biaixos i les funcions d'activació (Figura 6).

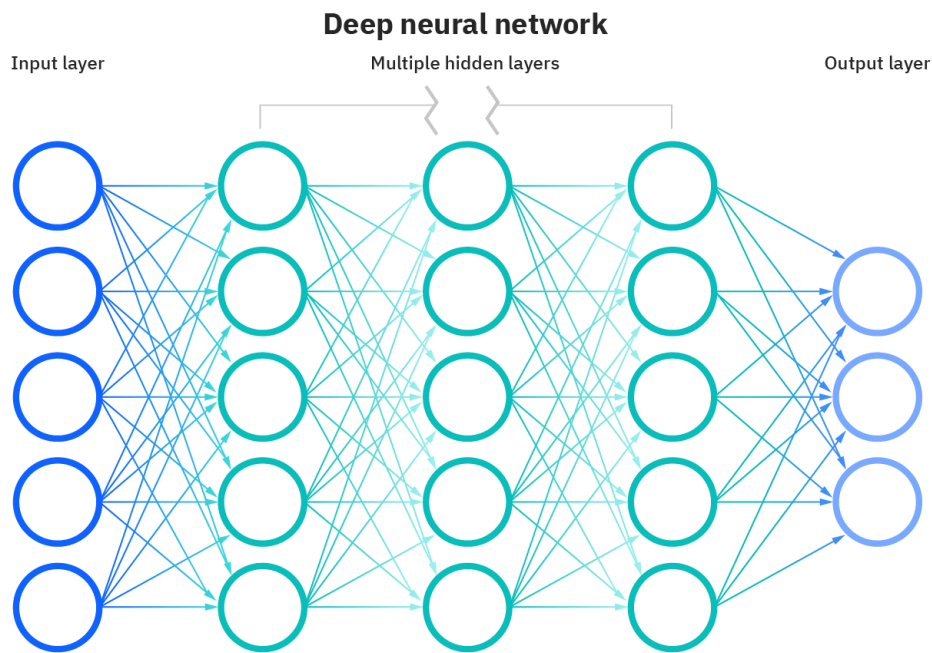


Figura 6: Esquema estructura interna d'una xarxa neuronal artificial.

Una xarxa neuronal està conformada per una sèrie de nodes que s'encarreguen de processar les dades rebudes i enviar-les a altres nodes. Aquests nodes s'organitzen en capes, totes les neurones d'una capa es connecten amb les neurones de la següent capa i cada capa pot tenir un número indefinit de neurones.

Podem distingir tres tipus de capes diferents segons en quin punt de la xarxa estan. La capa d'entrada (Input) està formada per neurones que contenen la informació inicial i s'encarrega de transmetre les dades a la següent capa. La capa de sortida (Output) està formada per tants nodes com classes de sortida que volem que tingui la xarxa, si volguéssim detectar un objecte, podria ser una sola neurona probabilística que ens donaria el valor de la probabilitat de que aquell objecte fos en la imatge. En els casos de segmentació binària tindríem tantes neurones com píxels de la imatge final i cadascun contindria informació sobre si forma part del objecte a segmentar o no, així aconseguint delimitar-lo.

Finalment, trobem les capes ocultes, que estan situades entre la capa inicial i la de sortida i s'encarreguen de processar les dades. Si ens trobem amb més de una capa oculta ja podem considerar la xarxa com una xarxa d'aprenentatge profund (Deep Learning).

Els pesos representen la força de connexió entre dos nodes, com més gran el pes més influència té el node original en el node objectiu. Els biaixos són constants dels nodes

que permeten representar patrons no provinents del origen, permetent que la xarxa tingui més flexibilitat i es pugui aproximar al resultat amb més llibertat.

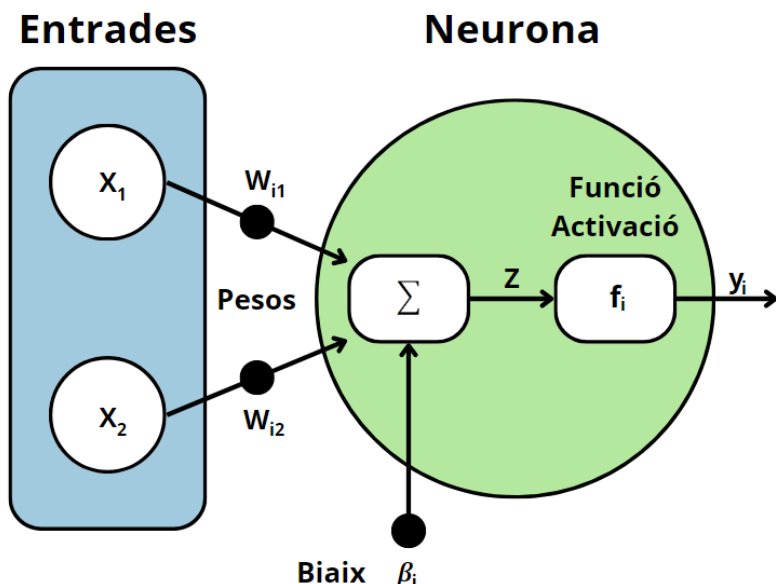


Figura 7: Relació entre neurones d'una xarxa neuronal artificial.

Per determinar les sortides dels nodes, hem de pensar cada node com si fos una regressió lineal, a on els valors de les neurones anteriors, multiplicades per el seu pes associat, es recullen en un sumatori a on també es suma el biaix de la neurona en qüestió. Un cop obtingut el sumatori, el passem per una funció d'activació la qual ens donarà un resultat diferent segons el tipus de funció i el tipus de resultat que hi arribi (Figura 7). Aquest procés queda expressat matemàticament com:

$$y_i = f_i(Z_i) = f_i(\sum_{j=1}^n W_{ij} \times X_j + \beta_i) \quad (1)$$

Les funcions d'activació, ens permeten aplicar una transformació no lineal a el sumatori Z. Hi ha moltes funcions i molt diverses, una de les mes conegudes és la funció esglaó (equació 2), a on segons si es supera un límit establert la funció té un valor o un altre (Figura 8).

$$y_i = \begin{cases} 1, & Z_i < 0 \\ 0, & Z_i \geq 0 \end{cases} \quad (2)$$

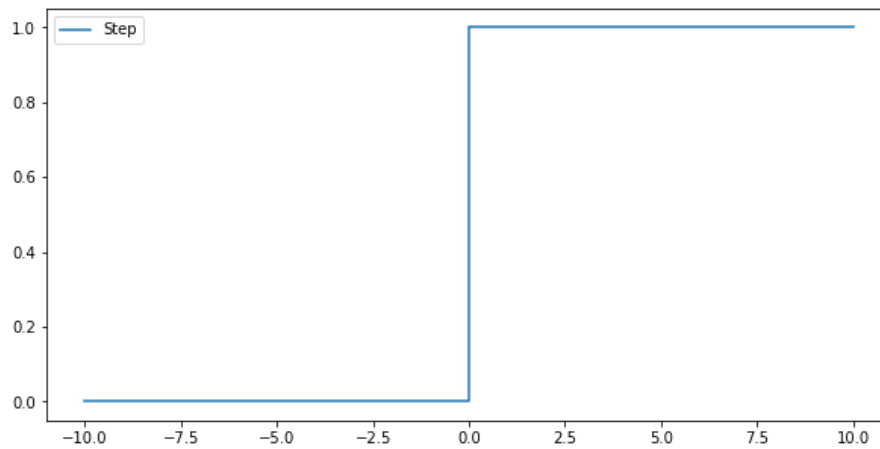


Figura 8: Gràfica funció esglaó unitari.

La funció ReLU (Rectified Linear Unit) (equació 3), és molt similar a una funció esglaó. Funciona de tal manera que tots els valors de  $Z$  positius, conserven el valor de  $Z$ , en canvi si són negatius, queden igualats a 0 (Figura 9).

$$R(Z_i) = \max(0, Z_i) \quad (3)$$

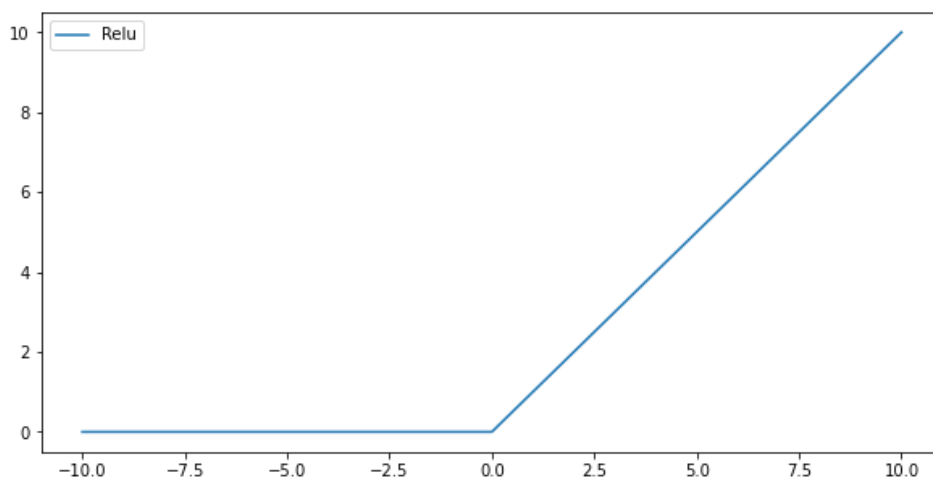


Figura 9: Gràfica funció ReLU.

Una de les altres funcions més important és la funció continua sigmoide (equació 4), que va ser la funció d'activació més popular durant molts anys, a diferència de les funcions anteriors, la funció continua sigmoide és completament derivable, però manté la forma de una funció esglaó (Figura 10).



$$y_i = \frac{1}{1+e^{-z_i}} \quad (4)$$

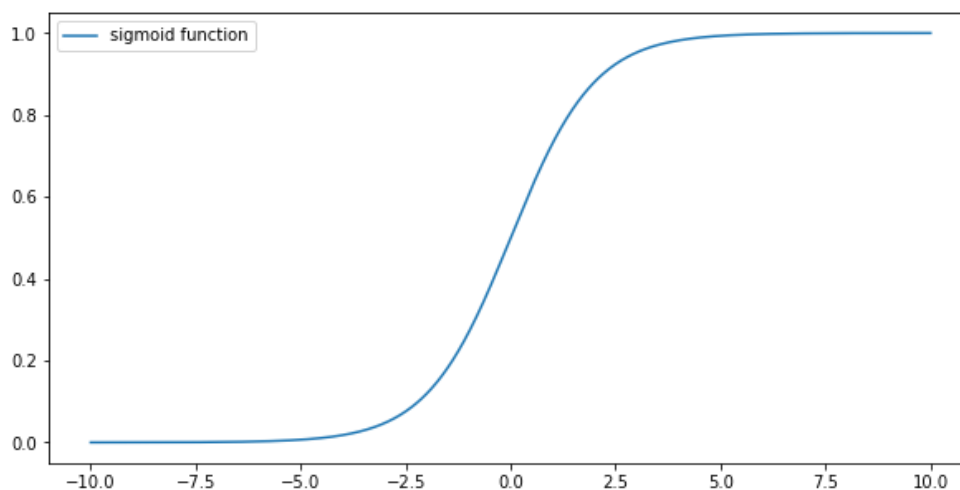


Figura 10: Gràfica funció sigmoide.

Per últim, tenim la funció Softmax (equació 5), aquesta funció ens permet transformar la sortida en una distribució de probabilitat, és per això que s'utilitza en la capa final, per poder donar un resultat probabilístic a les prediccions del model.

$$y_i = \sigma(Z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (5)$$

Per tal que el nostre model sigui capaç d'aprendre, necessitem definir una funció que quantifiqui com d'incertada estava la solució que ens proposava. A partir d'aquesta funció, aleshores, el model intentarà minimitzar la quantitat d'error per cada imatge d'entrenament que vagi processant. Això fa que el problema d'aprenentatge es converteixi en un problema d'optimització a on volem trobar el mínim d'aquesta funció de pèrdues. Hi ha moltes funcions de pèrdua i cadascuna té les seves pròpies aplicacions, en el camp de la segmentació binaria, les funcions de pèrdua més usuals són la Binary Crossentropy Loss, o altres funcions basades en estadístics, com la Dice Loss i la f1 Loss. Si ara mirem el camp de la síntesis d'imatges, les funcions més utilitzades són la del error quadràtic mitjà (MSE), l'error absolut mitjà (MAE), i la Structural Similarity Index (SSIM) la qual és una sèrie de fórmules que tenen en compte la intensitat de la imatge, el contrast, i l'estructura (Taula 2).

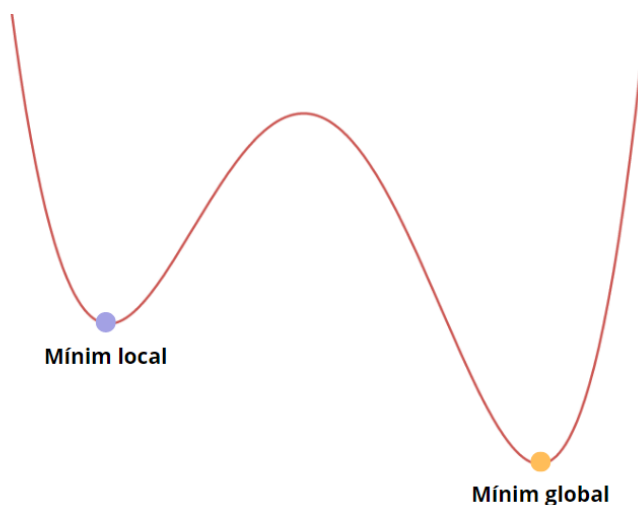
Taula 2: Fórmules funcions pèrdua.

Binary Crossentropy	$l(\hat{y}, y) = -\frac{1}{K} \sum_{i=1}^K y \times \log(\hat{y}) + (1 - y) \times \log(1 - \hat{y})$
Dice	$l(\hat{y}, y) = 1 - \text{Dice Coeficient}$
F1	$l(\hat{y}, y) = 1 - f1 \text{ Coeficient}$
MSE	$l(\hat{y}, y) = \frac{1}{K} \sum_{i=0}^K (y - \hat{y})^2$
MAE	$l(\hat{y}, y) = -\frac{1}{K} \sum_{i=0}^K  y - \hat{y} $
SSIM	$l(\hat{y}, y) = \text{brillo}(\hat{y}, y) \times \text{contrast}(\hat{y}, y) \times \text{estructura}(\hat{y}, y)$

#### 4.1.2 Entrenament model.

El nostre objectiu amb la funció de pèrdua és aconseguir parametritzar correctament la nostre xarxa neuronal per tal d'aconseguir un error mínim. Els paràmetres a entrenar són els pesos i els biaixos, els quals s'inicialitzen aleatòriament i s'han de anar actualitzant fins al seus valors òptims. Per aconseguir aquesta optimització s'utilitzen algorismes d'optimització.

Els algorismes d'optimització han de aconseguir minimitzar correctament la funció de pèrdua, sense convergir en mínims locals (Figura 11). L'optimitzador més utilitzat és l'Adam, és un algorisme que aprofita l'impuls par convergir més ràpidament i a més a més utilitza tècniques adaptatives d'aprenentatge (learning rate) per tal que acceleri a les regions planes i vagi més lent en regions més abruptes per tal de no saltar-se el mínim.



*Figura 11: Diferència entre un mínim local i un mínim global.*

El procés d'entrenament consta de dos grans fases, en la primera fase, el model fa circular endavant les dades d'entrada, passant per totes les capes ocultes fins a obtenir un resultat a la capa final. Un cop hem obtingut la solució, la avaluem segons la nostra funció de pèrdues seleccionada, per poder quantificar l'error. Un cop tenim l'error quantificat, podem començar el segon pas, en el qual mitjançant el mètode de Backpropagation. Aquest mètode ens permet actualitzar els pesos a partir dels gradients de pèrdua que podem obtenir de cada node al fer el recorregut a la inversa a partir de derivades parcials. Les derivades ens permeten saber com de ràpid canvia l'error a mesura que surt/entra dels nodes així com també la velocitat en que canvia el error de cada pes quan entren a un node. En aquesta actualització també entra en joc la taxa d'aprenentatge (Learning Rate) que ens defineix la rapidesa amb la que la xarxa aprendrà. Aquestes dos fases es van repetint cíclicament fins assolir el resultat òptim (Figura 12).

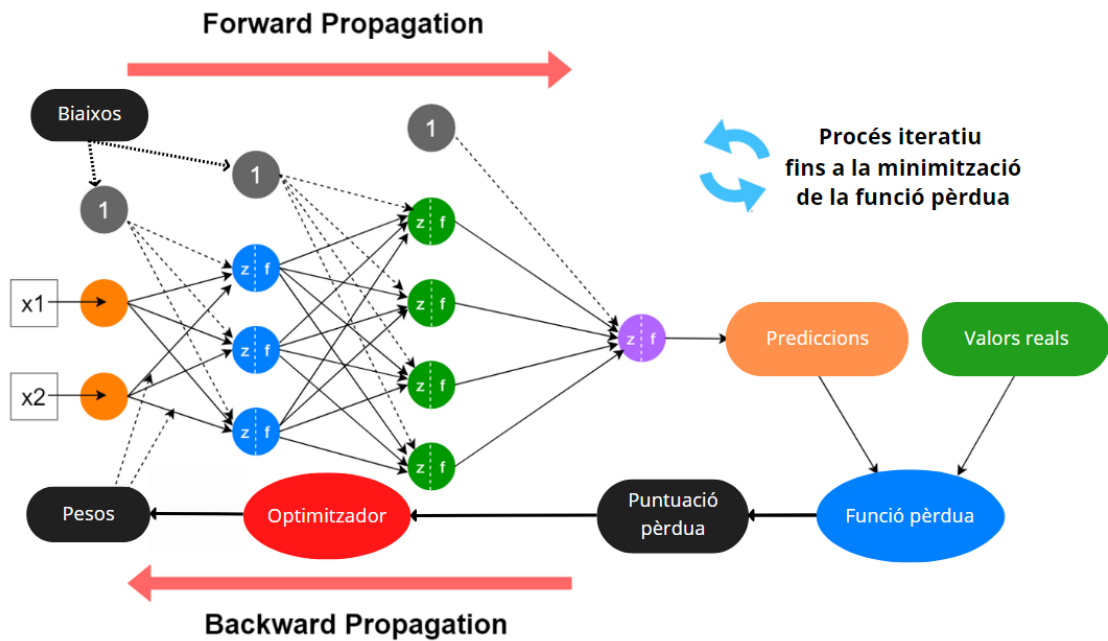


Figura 12: Esquema del procés d'entrenament d'una xarxa neuronal artificial.

A l'hora de crear el model, hi ha certs hiperparàmetres el valor dels quals ha de ser establert abans de començar a entrenar la xarxa. A diferència dels altres paràmetres, aquests no s'actualitzen amb cada iteració, sinó que es mantenen constants en tot moment. Això provoca que una bona modelització del sistema prèvia al entrenament, ens ajudarà molt a tenir uns resultats més bons. El primer paràmetre important és la mida del lot (Batch Size). Per tal de passar les dades a la xarxa pel seu entrenament, dividim aquestes dades en lots i entrem els lots per separat, amb aquesta acció aconseguim que el model convergeixi més ràpid al mínim, ja que processa el conjunt de dades de manera paral·lela. Si la mida del lot és massa petita, el model pot ser susceptible a ser sobre-entrenat, mentre que si és massa gran, pot ser que no pugui entrenar el suficient. Relacionat amb conjunts de dades, també trobem el nombre d'èpoques que entrenem un conjunt de dades. Cada època fa referència a cada cop que totes les dades han sigut processades, tres èpoques significaria que cada dada ha passat 3 cops per la xarxa. A la pràctica, funciona a la inversa que la mida del lot, un numero molt gran d'èpoques comportaria que el model quedés sobre-entrenat, mentre que un numero petit d'èpoques generaria un model sub-entrenat. Per últim trobem que hem de definir prèviament la taxa d'aprenentatge. Com hem comentat anteriorment, és un paràmetre molt important ja que ens determina la velocitat en la que el nostre model aprendrà i si serà capaç de convergir en un mínim global o no

### 4.1.3 Prediccions model.

Un cop el model entrenat, hem de realitzar un procés de test o validació per saber realment com de bé el model aconsegueix predir resultats davant de dades noves. Aquestes dades de test, són del mateix tipus que les d'entrenament. Normalment, de les dades resoltes, en separem la majoria per entrenar, i amb la porció restant avaluem el model per poder saber la precisió real del model. Aquest pas és necessari ja que pot passar que el model no sigui capaç de fer prediccions correctes sobre les noves dades. Aquest problema pot sorgir de dos situacions diferents: underfitting (sub-entrenament) o overfitting (sobre-entrenament) (Figura 13).

Un model amb underfitting és aquell a on tant amb les dades d'entrenament com amb les de test, el model no aconsegueix donar-nos un bon resultat de precisió, si passa això pot solucionar-se augmentant el nombre d'èpoques, incrementant la complexitat del model, o com a últim recurs, netejant, canviant i/o augmentant la base de dades. Ja que pot ser que amb les dades que disposem, el model no ens és capaç d'arribar a uns bons resultats de precisió.

En el cas d'overfitting, el model seria capaç d'assolir molts bons resultats en les dades d'entrenament, però al estar massa acostumat a aquestes dades, al introduir dades noves no ens genera unes prediccions amb bona precisió. Per solucionar aquest problema, podem monitoritzar simultàniament les dades de test quan entrenem, i així podem fer que les iteracions parin quan comenci a baixar la precisió en les dades de test. Un altre mètode seria reduir la complexitat del model o afegint més dades diverses d'entrenament. En aquest últim cas, s'utilitzen tècniques de per augmentar el nombre de dades, com per exemple la creació d'imatges noves a partir de la rotació de les imatges que disposem.

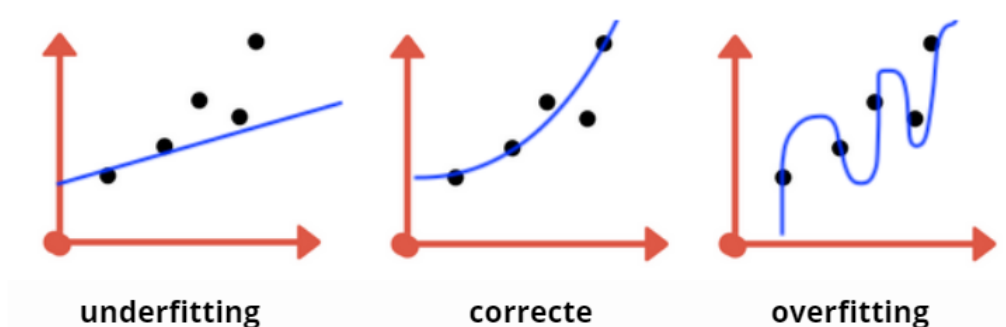


Figura 13: Gràfiques representatives de les diferents corbes de resultats d'entrenaments.

## 4.2 Xarxa neuronal convolucional

A l'intentar introduir una imatge com a input en la nostra xarxa neuronal, veiem que cada píxel correspon a una entrada. Això comporta que per una imatge RGB de 1 MP, tenim 3.000.000 d'inputs, que davant d'una capa de 100 neurones, obtindríem 300.000.000 pesos en la primera capa. Per intentar evitar aquest problema, es van crear les xarxes neuronals convolucionals. Aquestes xarxes continuen obligant a la imatge a descompondre's per píxels, però aconseguim evitar els valors altíssims de paràmetres, gràcies a la utilització de les capes convolucionals també anomenades filtres.

Cada capa convolucional està formada per una sèrie de filtres convolucionals, que tenen com a funció la detecció de formes i figures en les imatges. Aquests filtres consisteixen en matrius petites de valors, que realitzen el producte escalar amb la imatge, per tal d'obtenir un resultat. Poden ser des de molt simples com la detecció de vores verticals, o més complexes com per exemple la detecció d'ulls en una imatge.

Una convolució consisteix en el recorregut d'aquests filtres per la imatge, realitzant el producte escalar i emmagatzemant els resultats en una nova matriu de sortida.

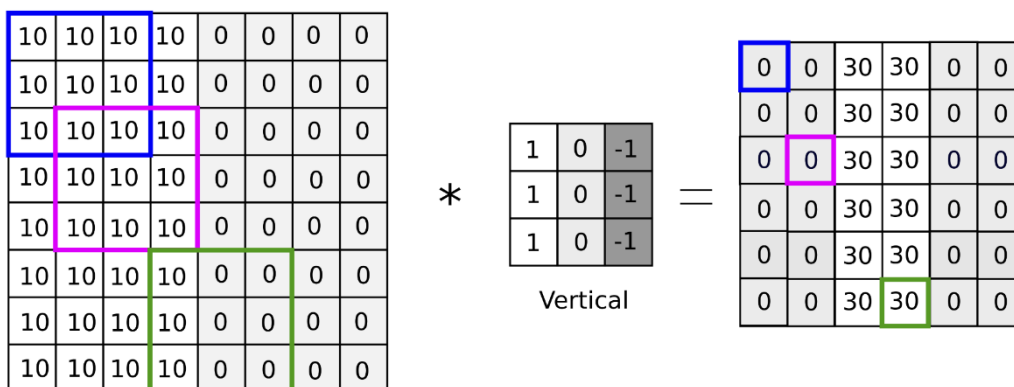


Figura 14: Esquema resultats produïts per la convolució d'un filtre vertical en una xarxa neuronal convolucional.

En la Figura 14 podem veure si els valors representen píxels, la imatge original representa una imatge de dos colors separats verticalment. A aquesta imatge, li apliquem un filtre de detecció de vores verticals. Aquest filtre veiem com al anar movent-se per la imatge va emmagatzemant els valors del seu producte escalar amb els píxels de la imatge corresponents en una matriu resultat. La operació que realitza per trobar el resultat a emmagatzemar és:

$$\begin{bmatrix} 10 \times 1 = 10 & 0 \times 0 = 0 & 0 \times -1 = 0 \\ 10 \times 1 = 10 & 0 \times 0 = 0 & 0 \times -1 = 0 \\ 10 \times 1 = 10 & 0 \times 0 = 0 & 0 \times -1 = 0 \end{bmatrix}$$

$$Resultat = 10 + 10 + 10 + 0 + 0 + 0 + 0 + 0 + 0$$

El requadre verd, al haver-se mogut 4 píxels a la dreta i 6 píxels respecte la part superior esquerra, el seu resultat es veu emmagatzemat en la fila 6 i columna 4 de la matriu resultant. En la matriu resultant veiem com ha pogut detectar a la perfecció la separació vertical de la imatge original, diferenciant-la del resta de la imatge.

Del funcionament de les convolucions en podem extreure bastanta informació. Primerament, podem observar com la mida de la matriu resultant, és menor a la matriu original. Això és degut a que les dimensions venen determinades per els moviments que pugui realitzar el filtre dins de la matriu original, per tant, si augmentem la mida del filtre, la matriu resultant serà més reduïda. També veiem com aquells píxels dels extrems de la imatge no es tenen en compte en tantes operacions com el píxels centrals de la imatge. Per exemple els píxels de les 4 cantonades, només es tenen en compte en la operació per donar resultat als píxels de les 4 cantonades de les imatges resultants. És per això que si també ens interessin aquelles parts de les imatges que estan properes als extrems, hem de tenir mètodes que ens solucionin el problema i donin més importància a aquests píxels. El padding dona solució a aquest problema creant un marc de valor 0 al voltant de la imatge així fent que els nous píxels dels extrems siguin zeros i no els reals de la imatge (Figura 15).

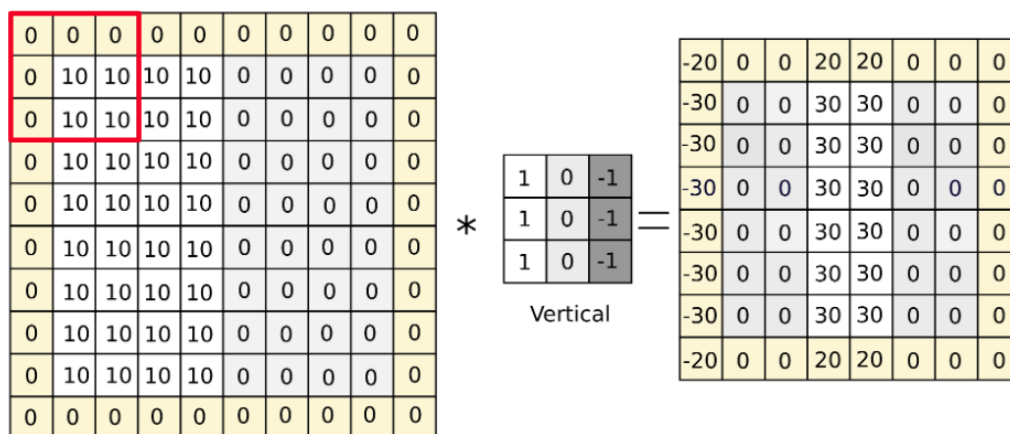


Figura 15: Aplicació i resultat d'un padding de 1.

D'altra banda, també ens interessa en certs casos, reduir la mida de la matriu resultant sense haver de augmentar la mida del filtre. Al fer aquesta reducció, aconseguim reduir el cost computacional de l'operació i també ens ajuda a evitar el overfitting. Aquesta reducció la podem fer de varies maneres diferents, una d'elles és l'aplicació d'un stride o pas. L'stride simplement és el nombre de píxels que et mous en una direcció perquè es consideri que has fet un pas. Un dels altres mètodes és el pooling, en aquest mètode, el filtre que es recorre la imatge original no realitza cap producte escalar. Si el pooling és segons el màxim (max pooling), el valor que posaríem en la matriu resultant seria simplement el màxim valor dels píxels en els que està intervenint. També hi ha altres

mètodes de pooling com l'average pooling, a on el resultat final és la mitjana dels píxels de la submatriu (Figura 16). Cal notar la clara diferència amb un filtre normal, els filtres de pooling, no tenen cap valor associat, per tant a la hora d'entrenar el model, no tenen cap paràmetre. És per això que hi ha diversitat d'opinions sobre si considerar el pooling com una capa, i en molts articles veiem que ho consideren un procés extra però no al nivell de una capa convolucional.

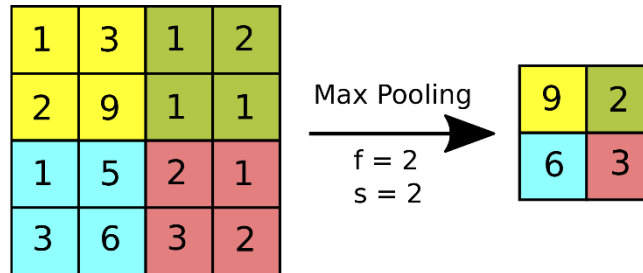


Figura 16: Aplicació i resultat d'una capa de max pooling amb un stride de 2.

Una capa convolucional pot tenir múltiples filtres, passant a ser el nombre de filtres, el nombre de sortides que tindrà aquella capa. Cal considerar però que les imatges RGB al tenir tres canals diferents, cada filtre també tindrà tres capes i cada capa multiplicarà el seu canal corresponent, la matriu de cada capa no té perquè ser la mateixa per cada canal.

En l'aprenentatge automàtic, cada filtre està parametritzat i són els valors d'aquests filtres els que s'han de actualitzar per tal de optimitzar el resultat. Si observem tot el procés que s'aplica a cada input, podem veure com el mètode és similar al estudiat en les xarxes neuronals automàtiques. L'input és la imatge, els filtres simulen els pesos, a la matriu resultant es concatenen tot el conjunt de multiplicacions input-filtre, i finalment, al resultat li sumem un biaix i li apliquem una funció d'activació per obtenir el resultat final. Un cop aquest procés també podem fer canvis a posteriori com per exemple un max pooling (Figura 17).



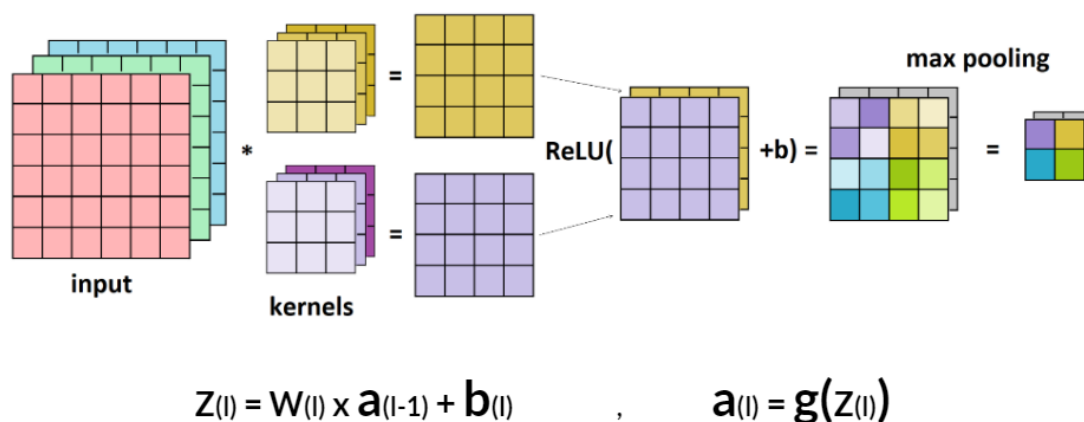


Figura 17: Esquema de la equivalència de les fórmules d'una xarxa neuronal artificial amb l'estructura d'una xarxa neuronal convolucional.

Aquest procés de transformació del input és el procés de transferència d'informació que es realitzarà a la resta de la xarxa neuronal, convertint el output d'un procés amb el input del següent. Podem veure com si anem repetint aquest procés, cada cop obtindrem imatges més reduïdes però amb més capes representant cadascuna d'aquestes les diverses característiques de la imatge. Arribarem a un punt on podrem representar la imatge com un vector columna de valors, que és el mateix tipus d'output que obteníem en les xarxes neuronals no-convolucionals. És gràcies que a aquest vector li podem aplicar una funció d'activació per obtenir resultats probabilístics, que aquest mètode es pot utilitzar per exemple per la detecció d'objectes en la imatge (Figura 18). Si volem obtenir una segmentació, necessitarem tornar a reconstruir la imatge per obtenir un output en forma de imatge segmentada, és aquí on apareix la xarxa U-Net.

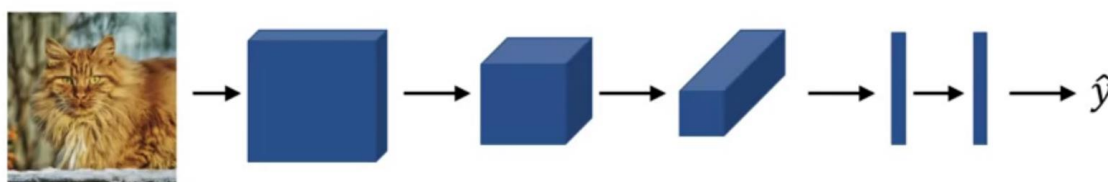


Figura 18: Procés de detecció d'animals via una xarxa neuronal convolucional.

### 4.3 U-Net.

Es va descobrir que en les xarxes convolucionals, en aquelles capes més comprimides que tenen informació més concreta, però han perdut la informació original de la imatge, si també li comunicàvem informació més detallada, ajudava al seu entrenament i obteníem millors resultats. Aquestes xarxes neuronals s'anomenen Res-Net (Residual Networks) i fan que la informació d'una capa sigui transmesa dos capes més endavant. Les xarxes neuronals convolucionals U-Net, es basen en aquest principi. Per tal de

descomprimir la imatge, reben informació de quan la xarxa estava descomprimida anteriorment. Això actua com a guia i permet crear una bona descompressió generant així un resultat més precís.

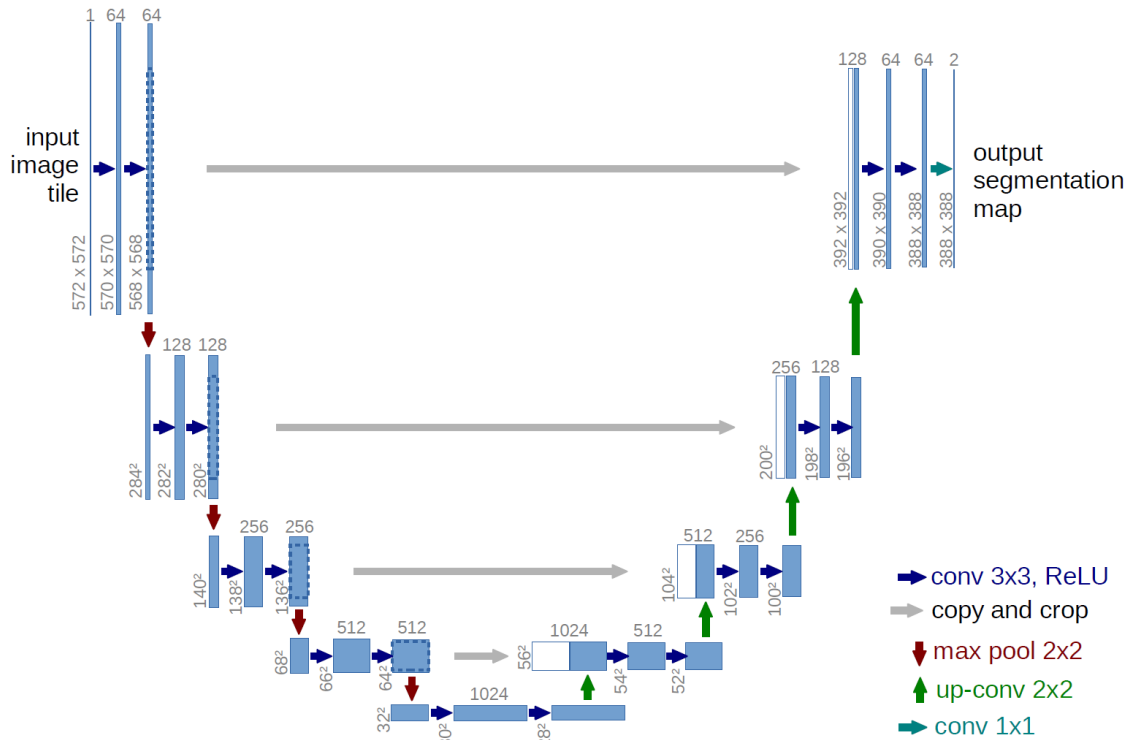


Figura 19: Esquema del procés de compressió i descompressió d'una xarxa U-Net.

La xarxa s'anomena U-Net degut a la forma de "U" que té el seu diagrama. Podem observar en la Figura 19 com anem comprimint la imatge inicial a base de poolings. També veiem com el nombre d'imatges de cada capa augmenta gràcies a les convolucions, fins que arribem a un punt on fem la transposada de les convolucions i juntament amb la informació de les imatges pre-comprimides de la mateixa mida, aconseguim revertir el procés i acabar amb una imatge segmentada.

Per obtenir més informació es poden consultar les referències [6] [7] [8] [9]

## 5 MODELITZACIÓ.

En aquest projecte, l'objectiu és realitzar la segmentació de les lesions i intentar solucionar el problema d'adaptabilitat. Per aconseguir-lo, es treballarà amb la xarxa neuronal convolucional U-Net. Com hem vist anteriorment, les U-Net ens permeten generar bons resultats de segmentació i alhora també la podem utilitzar per solucionar el problema d'adaptabilitat mitjançant la síntesis d'imatges. El procés de segmentació es pot diferenciar en tres grans apartats, en el primer, crearem un model de segmentació que separi els teixits cerebrals de la resta del cos. Aquest primer model, ens permet reduir informació que podria confondre a l'hora de detectar lesions, ja que altres components cranials tenen la mateixa intensitat de vòxels que les lesions a detectar. Aquest mètode s'anomena Skull Stripping i és molt comú en casos de segmentacions de lesions cerebrals ja que en millora la precisió.

Un cop realitzat el Skull Stripping, es farà la segmentació de lesions d'esclerosis múltiple a partir dels casos ja segmentats amb el model de Skull Stripping. Quan s'obtinguin uns bons models de segmentació de lesions, es posaran a prova amb els casos reals de diversos escàners. És aquí a on podrem observar les diferències de les segmentacions segons el escàner. A partir d'aquestes diferències, podrem realitzar la síntesis d'imatges d'aquells escàners més diferents i observar si ens ajuden a millorar els resultats o no.

### 5.1 Estructura U-Net.

En tots els tres processos mencionats, l'estructura de la U-Net serà la mateixa, només variaran els inputs, alguns híper-paràmetres i de les funcions de pèrdua. Les U-Net es poden diferenciar en 3D U-Net i 2D U-Net, segons si contenen informació espacial o no. Per fer el nostre treball, ens hem basat en l'estructura d'una U-Net 2D que no manté aquesta informació degut a la seva simplicitat que permet reduir el cost computacional i és capaç de donar molts bons resultats igualment.

La U-Net modelada està composta per 4 capes de compressió i 4 capes de descompressió, a on en cada capa s'hi realitzen dos convolucions de filtres 3x3 amb un stride unitari i un padding adaptable que fa que es conservi la mida entre la imatge input i la del output. En cada canvi de capa s'hi realitza un max pooling de 2x2 (Figura 19). El nombre de filtres de les convolucions es duplica en cada primera convolució de cada capa, començant amb 32 filtres i acabant a la última capa de compressió amb 256 filtres.

Aquesta estructura, al tenir 4 capes de compressió i 4 de descompressió, obliga a les dimensions inicials de la imatge a ser múltiples de 8, ja que sinó les dimensions podrien sortir decimals i ens ocasionarien problemes.

## 5.2 Híper-paràmetres.

Per l'entrenament de la nostre xarxa U-Net, es necessiten definir els híper-paràmetres. Es treballarà amb l'optimitzador Adam el qual per defecte té una taxa d'aprenentatge de 0,1 que és un valor alt, però podria no ser el més adequat per el nostre problema de segmentació. Per saber com de bé s'ajusta a la segmentació, hem realitzat una prova amb dades de Skull Stripping a on per cada època, la taxa d'aprenentatge anava augmentant, fins que ens arribava al valor original.

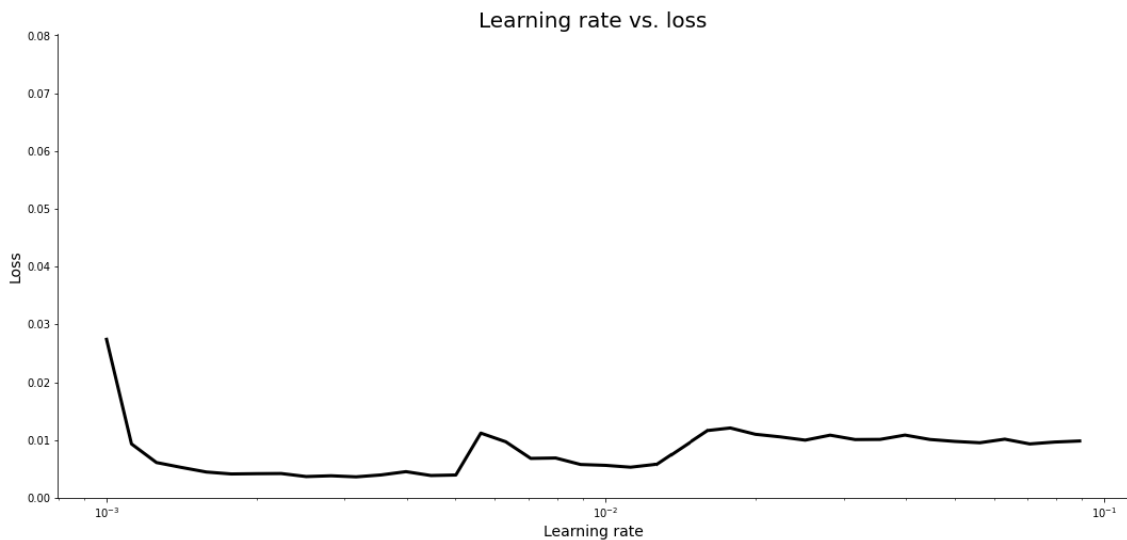


Figura 20: Gràfica de la quantitat d'error segons el learning rate (taxa aprenentatge).

Veiem en la Figura 20 com la pèrdua al inici, amb valors més baixos és degut al error inicial de la primera època, i per la resta de valors de Learning Rate, la quantitat de pèrdua és la mateixa independentment del valor de la taxa. Això ens demostra que l'optimitzador Adam funciona bé amb qualsevol valor del Learning Rate. Per tant, per poder fer els models, s'utilitzarà un Learning Rate de 0,1.

Un dels altres paràmetres a ajustar, és la mida de lot (Batch Size). Aquest paràmetre ens ajuda a evitar underfitting i overfitting i a estalviar temps d'execució si ha estat parametritzat correctament. Valors massa alts provocarien underfitting, i valors massa petits podrien ocasionar overfitting. Per quantificar els resultats, hem realitzat proves amb les dades de Skull Stripping i hem observat com variava el temps d'execució, i les precisions d'entrenament i test. Aquests valors els hem recopilat en la Taula 3:

## Modelització.

Taula 3: Temps d'execució segons mida de lot.

Mida del lot	Precisió entrenament	Precisió validació	Temps execució
256	0.9960	0.9931	77064 s (13 èpoques)
128	0.9958	0.9931	40397 s (7 èpoques)
64	0.9958	0.9933	29640 s (5 èpoques)
32	0.9960	0.9933	23712 s (4 èpoques)
16	0.9970	0.9942	5754 s (14 èpoques)

En la Taula 3 podem observar com a més mida de lot, per més que el temps per època sigui menor, al haver de fer més èpoques per convergir a un resultat, tarda més que si rebaixem la mida de lot. Totes aquestes proves s'han fet amb el mètode de Early Stopping, que consisteix en parar de fer èpoques quan el model no millora. En aquest cas, s'ha indicat que el model pari de entrenar quan la precisió de validació porta 3 èpoques sense millorar, i que agafi el model de 3 èpoques anteriors a la parada. Gràcies a això, no ens hem de preocupar per overfitting, i ens permet veure quant tarda en convergir segons el nombre de imatges en el lot. L'Early Stopping també fa que els valors de precisió siguin similars, ja que els models no paren fins que no arriben a un punt a on s'estanquin.

També veiem com amb una mida de lot de 16 imatges, els resultats són completament superiors a els altres. El dispositiu en els que s'han realitzats els experiments és un ordinador amb una GPU Geforce GTX 1050 de 2 GB i una CPU Intel Core i7. És important entendre el context de programació, ja que la intel·ligència artificial funciona especialment bé gràcies a la velocitat de càlculs que aporten les targetes gràfiques (GPU) respecte als processadors (CPU). La targeta gràfica que disposem, no admet una mida de lot superior a 16 imatges, ja que es queda sense memòria per poder fer tots els càlculs corresponents. És per això que hi ha una enorme diferència entre els temps d'execució dels lots superiors a 16 i els de 16. També veiem com les precisions han augmentat respecte a les altres mides, això és així gràcies a que al ser lots petits, el resultat oscil·la molt segons la època i aconseguix trobar un mínim més baix que en altres lots. En les proves de mida 16, en la tercera època ja arribava a una precisió de validació de 0.9934, però al anar oscil·lant tant, no s'estanca al cap de 3 èpoques i continua entrenant. Això no vol dir que els altres valors de mida de lot no podrien arribar a aquesta precisió millorada, sinó que hauríem de augmentar la paciència (el nombre de valors inferiors abans de indicar a el model que deixi d'entrenar) per tal de poder obtenir-los. Fent així que encara tardi més en convergir a un resultat.

Si tenim en compte que aquestes proves s'han realitzat amb 1/3 de les dades que disposem, veiem que una mida de lot de 16 és la millor opció. No només arriba a molts bons resultats, sense caure en underfitting, sinó que aconseguim obtenir resultats molt més ràpidament.

### **5.3 Bases de Dades.**

Per poder entrenar els nostres models, necessitem unes bases de dades amb informació sobre les segmentacions que es pugui acceptar com a certa. A més a més, per mirar els problemes d'adaptabilitat, necessitem varies dades dels mateixos pacients en diversos escàners d'on extreure lesions, per poder veure en un mateix cas, com varien els resultats segons l'escàner amb els que s'ha realitzat la ressonància magnètica.

En el cas del Skull Stripping, s'ha utilitzat la base de dades de Calgary-Campinas. Conté informació de tres escàners diferents (GE, Philips i Siemens) sotmesos a camps magnètics diferents (1,5 i 3 T) d'on tenim 359 imatges de ressonància magnètica tant d'homes com de dones, en un rang d'edat de 29-80 anys. Aquestes imatges les tenim en format T1-w.

Les dades per la segmentació de lesions d'esclerosi múltiple, hem utilitzat les dades del WHM Segmentation Challenge 2017 a on tenim un total de 60 casos, de 3 escàners diferents (20/escàner). Tots aquests casos els tenim en format T1-w i FLAIR. Per detectar les lesions, utilitzarem les imatges en format FLAIR.

Les dades per la síntesi d'imatges provenen de 15 pacients no lesionats que s'han realitzat ressonàncies magnètiques en 3 hospitals diferents: Vall d'Hebron, Trueta i Santa Catarina. Aquests pacients s'han escanejat 3 cops per hospitals, una d'elles en un any de diferència.

Per poder veure l'efecte que tenen els mètodes en casos reals de lesions, disposem de 10 pacients lesionats que s'han escanejat en els mateixos 3 hospitals que les dades de síntesis d'imatges. Dels escàners del hospital Trueta, en disposem tant de dades actuals com de dades antigues del mateix pacient.

*Per obtenir més informació es poden consultar les referències [1] [2]*

## 6 SKULL STRIPPING.

### 6.1 Pretractament.

Les imatges provinents de les ressonàncies de la base de dades, requereixen un pretractament per tal de poder segmentar-les amb precisió. A més a més segons com volem entrenar les dades, necessitem que l'input sigui d'un tipus o d'un altre. Per fer la segmentació, realitzarem diversos entrenaments a on compararem els resultats de la minimització de diverses funcions de pèrdua segons el tipus d'imatge inicial. Arrel d'aquesta comparació, podrem seleccionar el millor mètode de segmentació entre els analitzats.

Davant de les imatges de ressonància magnètica, podem entrenar el model de tres maneres diferents segons com són els inputs de les imatges completes i la seva segmentació, també anomenada màscara. Com que la segmentació només indica quines parts formen part dels teixits cerebrals i quines no, podem entrenar el model de tal manera que ens doni una predicció binària en format de 0 i 1. Si després ens interessa recuperar els valors dels píxels de les imatges, podem multiplicar la predicció amb la imatge original. Aquesta multiplicació farà que tots aquells píxels no segmentats al tenir valor 0, no recuperaran el píxel original. Mentre que aquells píxels segmentats, al tenir valor 1, la multiplicació fa que adoptin el valor original, així aconseguint obtenir la segmentació de la imatge original amb el valor dels píxels originals.

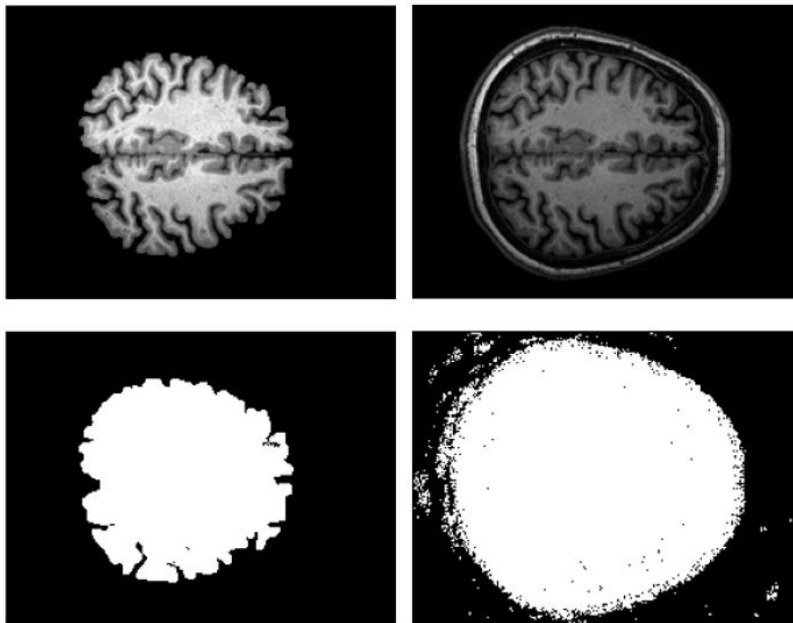
Gràcies a que ens és possible aquesta recuperació, el model el podem entrenar amb un input natural (directament la imatge de la ressonància magnètica, sense pretractament previ) i una màscara binaritzada. També podem tenir l'input binaritzat, gràcies a que els teixits cerebrals tenen una coloració més clara que el seu entorn, es poden veure separats en la binarització de les imatges de T1-w, per tant ens és possible realitzar-ne la segmentació ja que el model en pot extreure les seves característiques. Finalment, podem no realitzar cap binarització ni de l'input ni de la màscara, i que el model ens retorni com a output la segmentació ja en escala de gris.

Les imatges de ressonància magnètica, tenen el valors dels seus píxels entre 0 i 4500. Aquests valors provoquen un augment en el cost computacional del entrenament, per tant convé normalitzar-los. La normalització és l'operació de transformar tots els valors a una escala de 0 a 1 (equació 6). Aquest procés el realitzarem amb totes les imatges de ressonància magnètica de la nostre base de dades, indiferentment del mètode que volem implementar. Per fer aquesta operació, el que es fa és dividir tots els píxels de la imatge per el valor del píxel més alt d'aquesta, aconseguint que es mantinguin les diferències

entre cada píxel. Un cop ho tenim en escala de 0 a 1, podem llavors multiplicar la imatge per el nombre que desitgem per tal de canviar la escala.

$$\text{Imatge Normalitzada} = \frac{\text{Imatge Original}}{\max(\text{Imatge Original})} \quad (6)$$

El procés de binarització d'imatges consisteix en la conversió d'aquells píxels que superin cert valor límit a color blanc (valor=1), i aquells que no el superin, convertir-los a color negre (valor=0). El problema recau en determinar el valor límit per el qual realitzar la binarització. Les imatges al tenir el fons de color negre, posar el propi límit a 0 sembla una bona opció. Aquest cas és cert en el cas de les màscares, ja que al haver estat segmentades prèviament per experts, només tenen valor aquells píxels que formin part de la zona segmentada. Si apliquem el mateix principi en la imatge no segmentada, podrem observar com no binaritza correctament l'estructura cerebral degut a que en les parts fosques, els valors no són 0 sinó valors molt propers a aquest (Figura 21).



*Figura 21: Resultats de les binaritzacions de la màscara (columna esquerra) i el cap (columna esquerra) amb un valor límit de 0.*

Així doncs quin valor límit prenem? La resposta a aquesta solució no és fàcil, el problema que tenim és que per diverses parts de la ressonància la línia convé traçar-la a partir de cert valors, i aquests valors no tenen perquè ser constants per la resta de les zones del mateix escàner. A més a més ens trobem amb un problema d'adaptabilitat si fem servir el mateix límit amb altres imatges d'altres escàners. Com que en la nostre base de dades, tractem amb tres escàners diferents, ens trobem en la necessitat d'utilitzar un valor límit adaptatiu segons la imatge que es vol segmentar.



## Skull Stripping.

El mètode d'Otsu, estableix un valor límit variable segons la diferència entre els valors del píxel de la imatge i el nombre de píxels de cada rang de valors. Això permet agrupar els píxels segons valors i quantitat, i traçar el límit a on els diversos blocs de valors es separen (Veure Figura 22 i Figura 23 ).

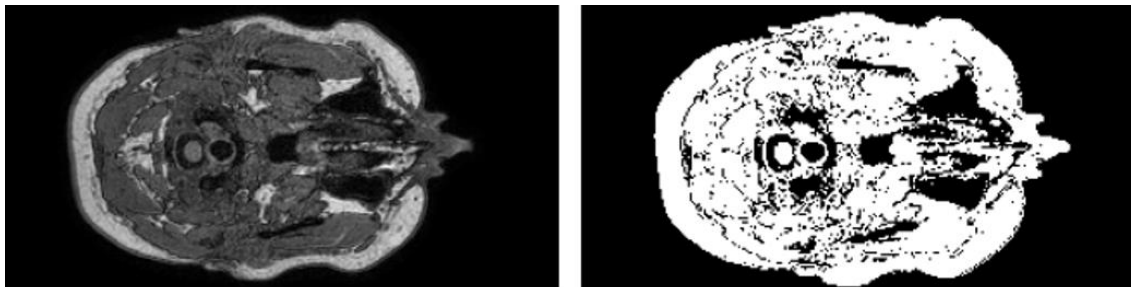


Figura 22: Resultat de la binarització mitjançant el mètode d'Otsu.

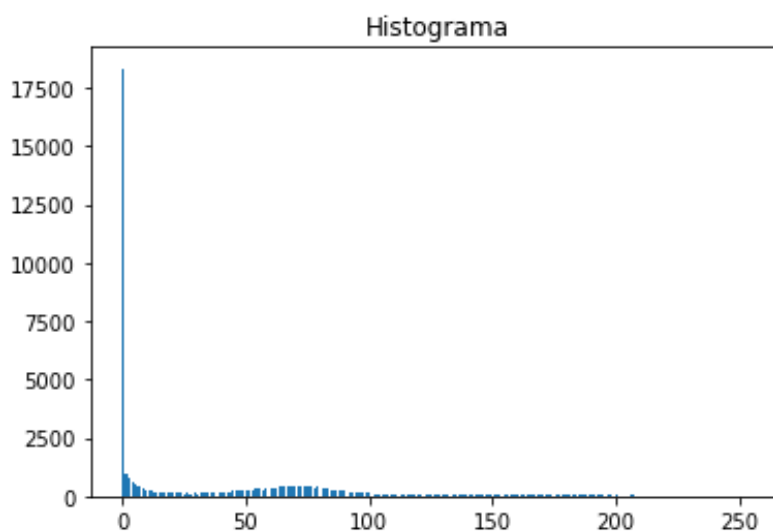


Figura 23: Histograma corresponent a la Figura 22.

## 6.2 Entrenament

Per entrenar les dades pretractades, ens falta decidir quines funcions pèrdua utilitzar. Depenent del mètode que fem servir, utilitzarem certes funcions o altres. Per comparar diverses funcions i veure els resultats, hem creat i definit dos funcions pèrdua, basades en penalitzar els falsos positius (FP) i els falsos negatius (FN), en comptes de la funció de pèrdua Dice o F1, cada funció només penalitza un d'ells, no una combinació d'aquests paràmetres. També hem utilitzat la funció de pèrdua de Poisson com a pèrdua logarítmica, per poder obtenir un resultat en color en el mètode no binari (equació 7).

$$l(\hat{y}, y) = \frac{1}{K} \sum_{l=0}^K (\hat{y} - y \times \log(\hat{y})) \quad (7)$$

Les funcions segons els mètodes són les següents:

- Mètode Binari (input i màscara binaris): Binary Crossentropy, Dice Loss, Penalització FP, Penalització FN.
- Mètode Semi Binari (input no-binari, màscara binària): Binary Crossentropy, Dice Loss, Penalització FP, Penalització FN.
- Mètode No Binari (input i màscara no binaris): Poisson Loss, Dice Loss, Penalització FP, Penalització FN.

En tots el mètodes, farem parar la iteració de les èpoques quan els valors d'exactitud de les dades de validació baixin durant tres èpoques seguides. El model que es guardarà per fer les prediccions serà el model de la època a partir de la qual han baixat consecutivament els valors de l'exactitud.

## **6.3 Resultats**

### **6.3.1 Resultats mostra**

Al tenir 3 mètodes amb 4 funcions per mètode, tenim 12 combinacions diferents a comparar de segmentacions. Per poder observar quina combinació de mètode-funció ens funciona millor per realitzar les segmentacions, ho podem mirar a partir d'una mostra de les dades. Al fer això, ens estalviem temps d'execució. Un cop seleccionats les combinacions més efectives, les tornarem a modelar amb la totalitat de les dades per obtenir la precisió desitjada.

Els resultats de les combinacions els avaluarem segons els valors d'exactitud (Accuracy), precisió (Precision) i sensibilitat (Recall) de les dades de validació.

### **6.3.2 Mètode binari.**

Si apliquem el mètode binari, obtenim els següents resultats:

## Skull Stripping.

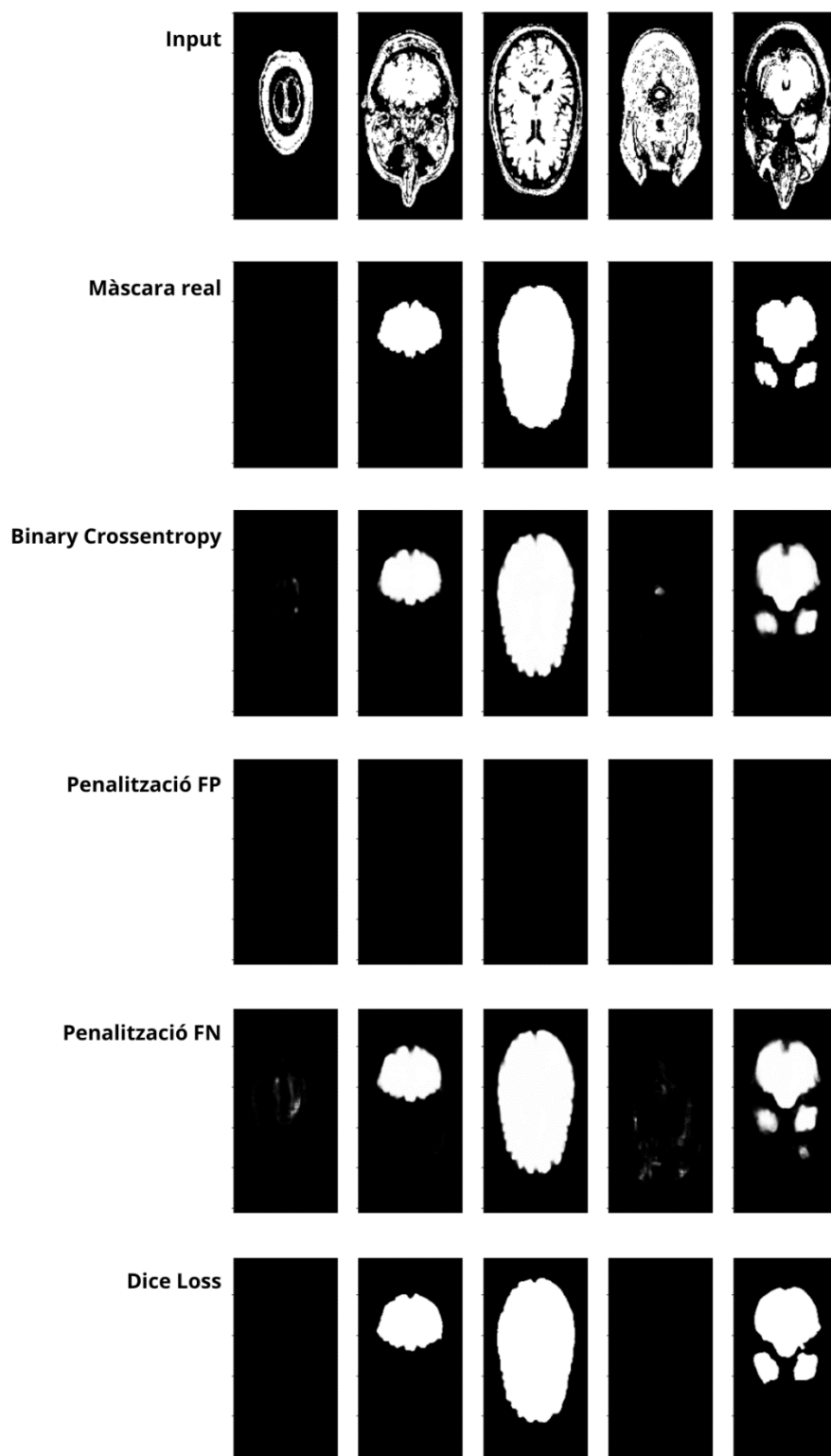


Figura 24: Resultats segons funció de pèrdua del mètode binari.

Taula 4: Resultats estadístics segons funció

Funció	Accuracy	Precision	Recall
Binary Crossentropy	0.9942	0.9777	0.9759
Penalització FP	0.8752	0	0
Penalització FN	0.9934	0.9666	0.9815
Dice Loss	0.9902	0.9331	0.9918

Com podem veure en la Taula 4 El mètode de penalització no genera cap predicció, per això ens surten nuls els valors de Precision i Recall, però ens manté un 0.8752 d'accuracy degut a que gran part de les imatges és de valor 0 (negre). La funció Binary Crossentropy, veiem que ens proporciona el valor d'Accuracy més elevat, i uns valors de Precision i Recall equilibrats, semblant així, la millor fórmula del mètode. Als resultats de la funció Dice Loss, obtenim uns bons valors d'Accuracy, però el Recall és notablement superior a la precisió. Finalment la funció de penalització FP ens obté uns resultats molt bons d'Accuracy amb uns resultats de Precision i Recall que si bé tendeixen més a augmentar el Recall, estan ben equilibrats.

Com afecten aquests diversos estadístics en les prediccions realitzades, ho podem veure en la Figura 24. Podem observar, com efectivament, la funció de penalització FP no ens genera cap predicció. També veiem l'efecte que fa el tenir un valor alt de Recall i baix de Precision, al tenir aquesta descompensació, el model genera unes prediccions més grans del que hauria, segmentant així més teixit que el desitjat. Això no significa però, que les prediccions siguin errònies, com podem veure en la imatge, la funció Dice Loss que és la que té els estadístics més desequilibrats, no ens genera prediccions en aquelles capes que realment no ha de predir res. És per això que per avaluar els mètodes de segmentació, necessitem visualitzar els resultats i no només mirar la informació que ens aporten els estadístics. Cal veure també que la funció Dice Loss ens dona els resultats ja binaritzat, mentre que tant els resultats de les funcions de penalització com la Binary Crossentropy, els resultats són probabilístics, permetent-nos agafar com a correctes els valors que estiguin per sobre de cert valor límit (binaritzar). Dit això, en aquest mètode, veiem que la funció Dice i la Binary Crossentropy són les que funcionen millor. Els resultats obtinguts per la funció Dice, poden sotmetre's a Postractament, reduint el gruix de la part segmentada, per obtenir uns resultats més ajustats, millorant així el resultat. De la mateixa manera, podríem millorar els resultats de la funció Binary Crossentropy si realitzem una binarització, ja que en cap moment sobredimensionem segmentacions detectades, sinó que el problema que tenim és que genera petites prediccions probabilístiques de zones a les que no hi ha segmentacions. Els resultats de la funció de

## Skull Stripping.

penalització de falsos negatius, veiem com són inferiors en Accuracy i Precision, però superiors en Recall. Veiem que això és degut que també detecta petites probabilitats de teixits a on no n'hi ha, però a major quantitat, ajudant així a millorar el Recall sense perdre molta Precision.

### 6.3.3 Mètode semi-binari.

Si apliquem el mètode semi-binari, obtenim els següents resultats:

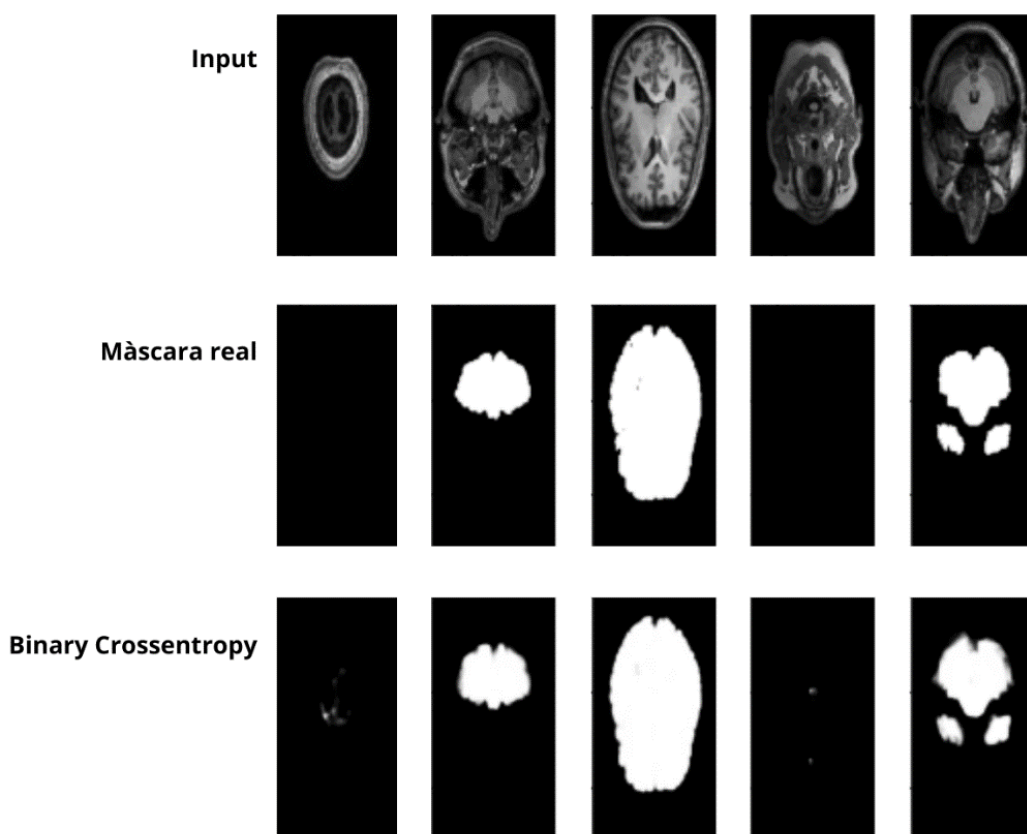


Figura 25: Resultats de la funció de pèrdua Binary Crossentropy en el mètode semi-binari

Taula 5: Resultats estadístic segons funció.

Funció	Accuracy	Precision	Recall
Binary Crossentropy	0.9950	0.9807	0.9795
Penalització FP	0.8752	0	0
Penalització FN	0.8752	0	0
Dice Loss	0.8752	0	0

Podem observar en la Taula 5 i la Figura 25 com l'única funció de pèrdua que ens funciona en aquest mètode és la Binary Crossentropy. Podem veure com respecte els resultats obtinguts en el anterior mètode, els resultats són casi idèntics, només es diferencien per una lleugera millora genera en els estadístics. De la mateixa manera que en el mètode anterior, aquestes deteccions falses en altres capes, es poden corregir aplicant una binarització a la predicció. Aquesta operació fa que aquesta combinació de semi-binari i la funció Binary Crossentropy, ens proporcionin molts bons resultats.

#### 6.3.4 Mètode no-binari.

Si apliquem el mètode no-binari, obtenim els següents resultats:

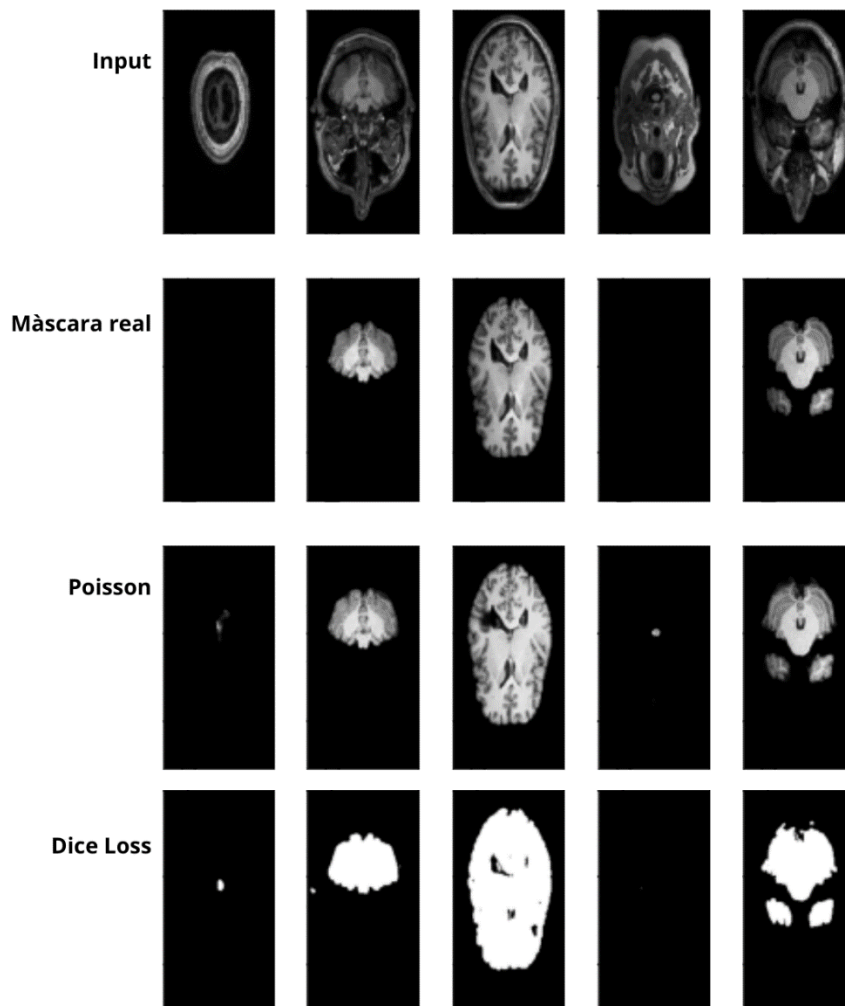


Figura 26: Resultats de les diverses funcions de pèrdua del mètode no-binari.

Taula 6: Resultats estadístics segons funció.

Funció	Accuracy	Precision	Recall
Poisson	0.8757	0.9945	0.2642
Penalització FP	0.8752	0	0
Penalització FN	0.8752	0	0
Dice Loss	0.8717	0.9722	0.9621

En el mètode no binari, hem volgut comprovar l'efecte de la Dice Loss encara que ens donés un resultat binari, veiem en la Figura 26 i Taula 6 com ens dona uns resultats pitjors que a la resta de mètodes. La funció Dice Loss, que com hem vist anteriorment, prioritza els valors veritables positius, no ens segmenta la totalitat dels teixits. A més a més, ens detecta segmentacions en capes que no n'hi ha, problema amb el qual no ens hi hem trobat en els altres mètodes. En la Taula 6 veiem també com els valors de l'Accuracy han baixat dràsticament respecte els altres mètodes, això és degut a que es té en compte els valors de cada píxel. Podem observar en la Figura 26 com els resultats de la funció de Poisson tenen una coloració diferent, pateixen alguna deformació en els teixits interns de la predicció. També veiem com apareixen falses segmentacions en capes que no n'hi haurien de haver, problema similar al que ens passava amb la funció de pèrdua de Binary Crossentropy. A diferència d'aquesta, però, la funció de Poisson ens dona els resultats en escala de gris, imitant la imatge original, no obtenim uns valors probabilístics als que poder binaritzar per corregir aquest error. Per aquests motius, en comparativa amb els altres mètodes, aquest no ens proporciona uns bons resultats de segmentació.

## 6.4 Resum resultats mètodes

Com hem pogut observar en les Figura 24 Figura 25 i Figura 26, i la Taula 4, Taula 5 i Taula 6 .No només les combinacions de mètode binari amb les funcions Binary Crossentropy i Dice, i la combinació del mètode semi-binari amb la funció Binary Crossentropy, ens obtenen uns resultats molt bons, sinó que també podem corregir els errors que trobem en les prediccions generades per aquestes. Per aquests motius, entrenarem el nostre model amb les 359 imatges, amb cadascuna de les tres combinacions, per tenir un model més robust davant de inputs diferents.

## 6.5 Resultats totals

No podem perdre de vista que el nostre model de Skull Stripping, ha de servir per segmentar correctament les dades de pacients lesionats tant de els casos reals, com dels casos que farem servir per entrenar el model i segmentar lesions. Amb això en ment, necessitem definir una mida d'input a la nostre U-Net, que no deformi excessivament

cap de les imatges de les tres bases de dades, la d'entrenament, la de segmentació de lesions i la de casos reals. Com que totes aquestes bases de dades tenen informació d'escàners diferents dintre les pròpies bases de dades, la mida de les imatges de ressonància magnètica d'aquestes, no és constant. La majoria d'escàners tenen la imatge allargada, amb una relació d'aspecte de 3:5, però també ens trobem en molts escàners que donen la imatge en un format quadrat de relació d'aspecte 1:1. Per entrenar el nostre model complet, utilitzarem una relació de píxels de 192 x 256 que correspon a una relació d'aspecte de 3:4, per entrenar un punt entremig i evitar la deformació excessiva de les imatges.

Si realitzem l'entrenament de les tres combinacions amb la mida d'imatge adequada, obtenim els següents resultats:

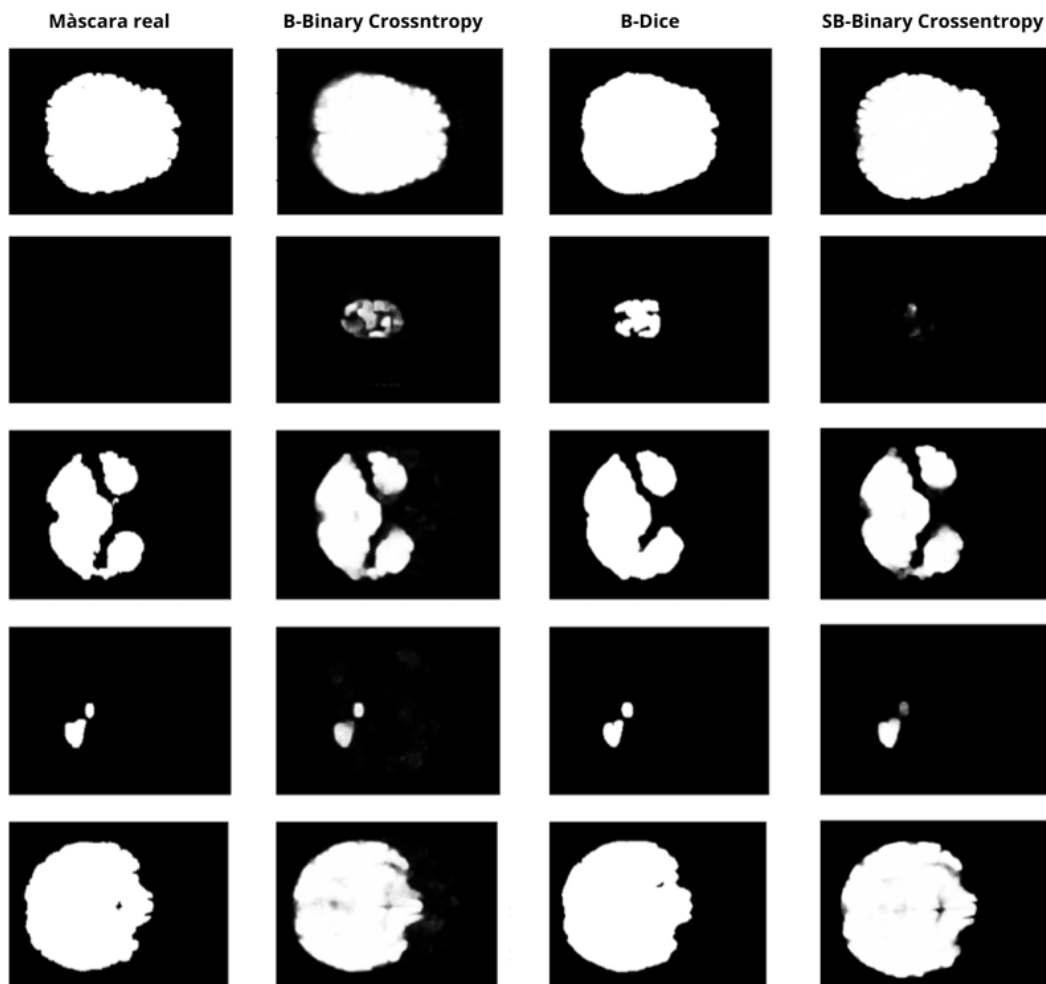


Figura 27: Màscara i resultats obtinguts amb diversos mètodes de segmentació.



## Skull Stripping.

Taula 7: Resultats estadístics segons funció.

Funció	Accuracy	Precision	Recall
B-Binary Crossentropy	0.9947	0.9753	0.9835
B-Dice Loss	0.9911	0.9294	0.9942
SB-Binary Crossentropy	0.9949	0.9873	0.9725

Podem veure en la Figura 27 i la Taula 7 com al entrenar amb més dades, aconseguim augmentar els estadístics dels models. També podem observar com en alguns casos, la segmentació no es produeix del tot bé, això pot ser degut en alguns casos a les variabilitats que tenim en les dades resoltes o bé perquè les imatges són una mica diferents a les esperades per el model. Observem com tots tres mètodes segmenten bé les imatges, veiem com la funció de Binary Crossentropy del mètode semi-binari aconseguix els millors resultats d'Accuracy i Precision. La combinació del mètode binari amb la funció Dice Loss ens proporciona el millor resultat de Recall, però sacrifica molta precisió. Per últim el mètode binari amb la funció de Binary Crossentropy ens dona bons resultats en tots els estadístics, actua de la mateixa manera que el del mètode semi-binari, però es nota la falta de precisió ja que segmenta més del que hauria.

Tots tres models entrenats, treballen bé amb les dades, i depenent del tipus d'input, podrien funcionar més o menys bé. La funció Dice Loss ens pot ajudar en aquelles segmentacions a on el model entrenat per la funció Binary Crossentropy no segmenta el suficient teixit, ja que la funció Dice es basa més en l'obtenció d'un bon valor de Recall. També pot ser que encara que haguem entrenat el model amb una base de dades que conté escàners diferents, si intentem segmentar amb el model generat per el mètode semi-binari, ens podem trobar que la diferent intensitat de píxels que pot tenir la nova imatge ens produeixi problemes en la segmentació. Per últim ens podem trobar en que el model amb la funció de Binary Crossentropy segmenta bé i els resultats de la funció Dice fan que les segmentacions quedin sobredimensionades. Com que no podem saber en quina situació ens podem trobar al segmentar altres casos fora de la base de dades, abans de entrenar models per la segmentació de lesions, mirarem de quina manera podem segmentar millor els casos d'entrenament.



## 7 SEGMENTACIÓ DE LESIONS

### 7.1 Pretractament dades.

Per obtenir uns bons resultats en la segmentació de lesions d'esclerosis múltiple, necessitem que les imatges que entrem al model estiguin netes i aptes per ser segmentades. Aquest pretractament l'haurem de realitzar sobretot per obtenir unes bones imatges de segmentació de Skull Stripping, ja que al formar part de bases de dades diferents, ens podem trobar problemes a la hora de fer les segmentacions.

Trobem que no avaluar directament quin dels nostres models de Skull stripping segmenta millor les imatges de la base de dades de lesions, al no tenir la màscara dels teixits cerebrals.

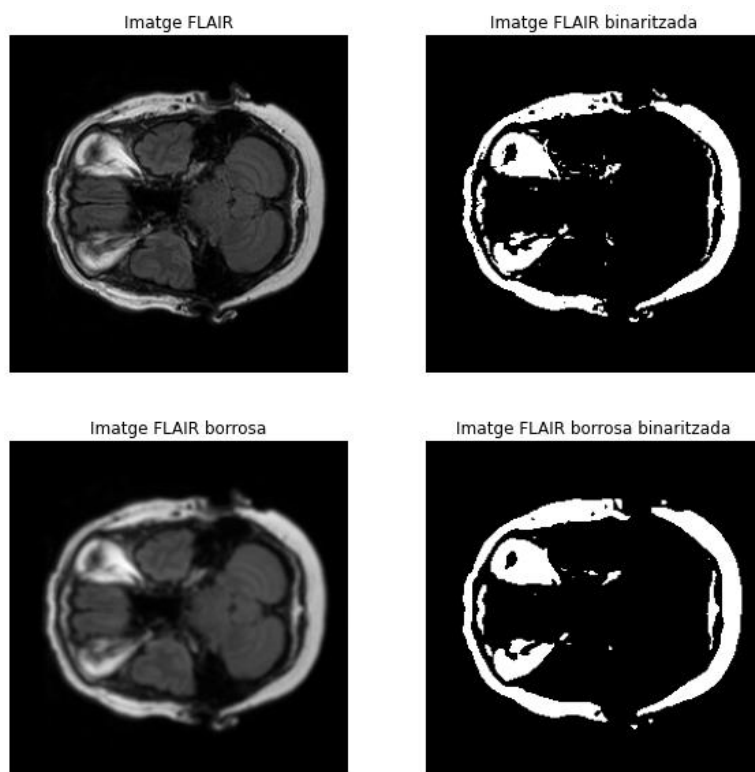
Per solucionar aquest problema, el que podem fer és realitzar les prediccions de les imatges i comprovar que la totalitat de les lesions estan a dins de les regions segmentades. El nostre model, al fer la segmentació, ens dona un output binari (o binaritzable). Aquest output el podem multiplicar directament amb la màscara de les lesions, convertint així tots els píxels en comú en valor 1 i aquells que no apareguin en la predicció o en la màscara com a valor 0. Fent aquesta operació, obtenim una imatge resultant que si la predicció funciona correctament, hauria de ser igual a la màscara. També podem trobar-nos que les segmentacions marcades en la màscara s'hagin marcat excessivament i sobrepassin la zona de teixit cerebral, així fent que la segmentació no ens agafi la totalitat de les lesions.

Els models que hem obtingut de Skull Stripping, han sigut entrenats amb imatges T1-w, però per entrenar les lesions, utilitzarem les imatges FLAIR. Aquesta diferència afecta directament a aquells mètodes a on la imatge d'entrada al sistema no estava binaritzada deguda a la diferència entre les intensitats dels píxels de les imatges, però també pot arribar a afectar a els mètodes binaris si la imatge té una gran quantitat de píxels d'alta intensitat que no formin part dels teixits cerebrals. Aquesta situació es pot solucionar si primerament realitzem l'Skull Stripping en la imatge T1-w de la base de dades, i el resultat de la predicció el multipliquem per la imatge FLAIR per obtenir el resultat de la segmentació amb els píxels corresponents a les imatges que necessitem per la segmentació de lesions.

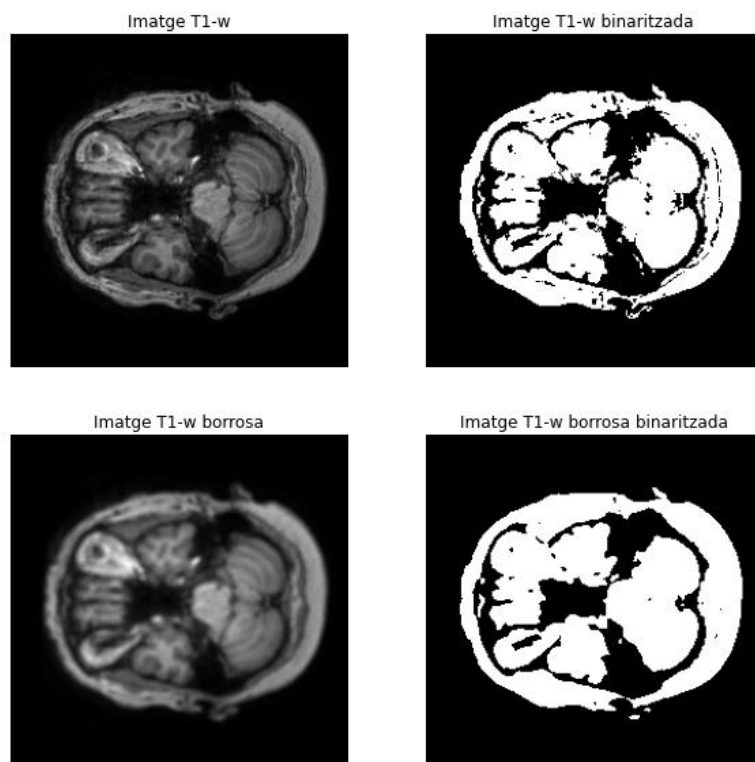
Com hem comentat anteriorment, la prioritat a la hora de avaluar els resultats dels models de Skull Stripping, és seleccionar aquells models capaços de contenir més

## Síntesi d'imatges de RM per la millora en la segmentació de lesions d'esclerosi múltiple

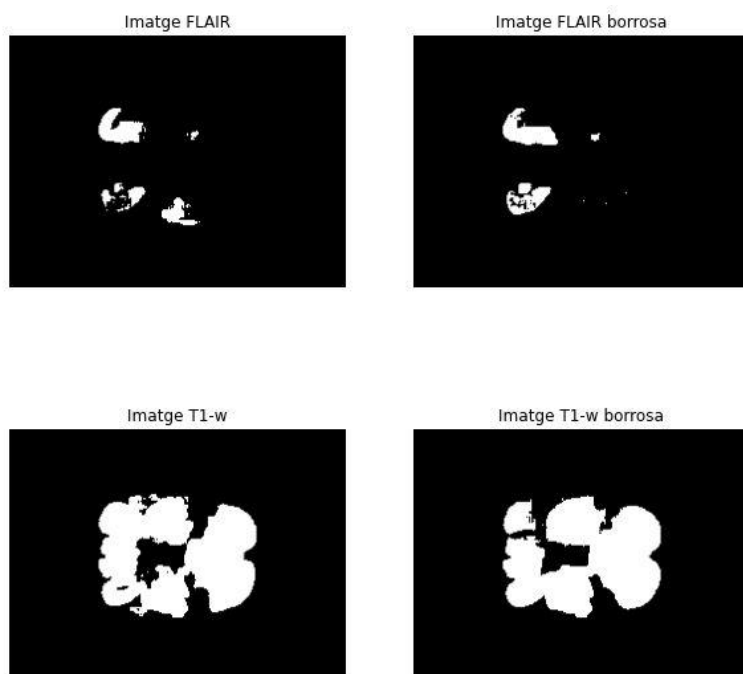
lesions en la seva solució. Per ajudar a crear resultats més aptes, en el pretractament de l'Skull Stripping, apliquem una operació que difongui una mica les imatges, així aconseguint uns valors més estables i ajudant a la binarització. Aquesta operació només es realitza si el mètode de segmentació és binari.



## Segmentació de lesions



*Figura 28: Imatges resultats binaritzacions segons pretractament.*



*Figura 29: Resultats segmentacions segons pretractament.*

En la Figura 28 i Figura 29 veiem com ens podem trobar en que en certes capes no podem binaritzar bé les imatges FLAIR, veient-nos obligats a realitzar les binaritzacions en les imatges T1-w per obtenir resultats. També ens serveixen per observar com

gràcies a l'addició de borrositat a les imatges, aconseguim netejar el soroll i obtenir uns bons resultats de segmentació. Aquest procés de reducció de soroll és important en aquelles imatges molt sorolloses, que ens creen una binarització poc uniforme. Per últim, cal observar que en aquests nous casos, les imatges estan rotades 180° respecte les dades d'entrenament de Skull Stripping. Això provoca que encara que obtinguem un bon resultat amb les imatges rotades, haguem de estudiar com varien els resultats si les orientem de la manera correcta.

Finalment, per acabar d'ajudar a que els models siguin capaços de contenir més lesions en la seva solució, també realitzem un post-tractament en les prediccions que generin els nostres models. A aquest post-tractament, hi afegim dos operacions. La primera d'elles dilata les prediccions, expandint els valors positius de les vores de les segmentacions. La segona operació, busca els contorns de la segmentació i omple tots els forats que tingui amb valors positius (Figura 30).



Figura 30: Imatge original amb la predicció sense i amb post-tractament respectivament.

A la hora d'avaluar els diversos models de Skull Stripping, hem agafat una mostra de 10 casos, i hem comprovat si la segmentació contenia la totalitat de les lesions. Aquestes comprovacions les hem fet amb diversos tipus de pretractaments. Hem realitzat comprovacions de 8 pretractaments diferents amb cadascun dels 3 models de segmentació i hem quantificat quants casos de la mostra tenien la totalitat de les lesions.

Taula 8: Numero de casos que contenen la totalitat de les lesions segons pretractament.

Pretractament\Models	B-BinCross	B-Dice Loss	SB-BinCross
FLAIR	5/10	6/10	0/10
FLAIR + rotació	4/10	5/10	0/10
FLAIR + complements	7/10	6/10	0/10
FLAIR + rotació + complements	7/10	6/10	0/10
T1-w	7/10	6/10	1/10

T1-w + rotació	7/10	7/10	1/10
T1-w + complements	9/10	7/10	2/10
T1-w + rotació + complements	9/10	8/10	3/10

En la Taula 8 podem observar com l'adició del les operacions de post-tractament (complements), ens ajuden a abastar més lesions en les imatges. També podem veure com el procés de rotació ajuda lleugerament al resultat, però en els dos millors casos (9/10), veiem com hi és present en un i en l'altre no. Aquest resultat ha estat agafat a partir de una mostra i no és completament representatiu del conjunt de la base de dades. És per això que la taula ens serveix per tenir una comparació orientativa dels diversos mètodes de segmentació, i poder avaluar els mètodes de segmentació de lesions a partir d'aquests.

## 7.2 Entrenament.

En la segmentació de lesions, necessitem conservar els colors, i ja hem vist que els mètodes completament no binaris, no ens obtenen una segmentació precisa del color. Per aquests motius, tots els mètodes d'entrenament seran semi-binari. Per entrenar el model utilitzarem 4 mètodes diferents corresponents a les segmentacions de les fórmules de Binary Crossentropy i Dice Loss a partir de les imatges T1-w amb post-tractament (tant girades com no).

A la hora de determinar les funcions de pèrdua, hem de tenir en compte que les segmentacions són molt petites degudes a la mida que tenen les lesions en comparació a la mida de la imatge, això fa que sigui molt fàcil que les funcions caiguin en un mínim local a on no ens detectin res perquè la pèrdua serà poca. Per corregir aquest problema, hem de seleccionar funcions que tinguin en compte paràmetres com la sensibilitat i la precisió, o mètodes similars. Les funcions que hem seleccionat per els nostres casos són la Binary Crossentropy, la Dice Loss i la f1 Loss. Aquestes fórmules s'aplicaran a tots 4 mètodes de pretractament ja que el tipus d'input sempre és el mateix.

En el procés d'entrenament, per evitar el overfitting, parem d'iterar èpoques quan un valor estadístic no millora durant un numero determinat d'èpoques. Normalment aquest estadístic sol ser l'exactitud de les dades de validació, però en el cas de la segmentació de lesions, el valor de l'exactitud sempre serà molt alt per culpa de les mides de les lesions. És per aquest motiu, que per la funció Binary Crossentropy, ens fixarem en el valor del estadístic f1 que ens determina la qualitat dels valors de precisió i especificitat. Per les funcions pèrdua Dice i f1, tendeixen a sobredimensionar les lesions, descompensant així els valors de precisió i especificitat, essent aquest últim significativament superior. Per tant, per agafar la època que faci el model més equilibrat, monitoritzarem l'estadístic de la precisió.

### 7.3 Resultats.

En tots els entrenaments, treballarem amb la totalitat de les dades, ja que la base de dades consta de 60 casos, els quals queden separats en 50 casos d'entrenament i 10 de validació. Aquests 50 d'entrenament, ens permeten poder realitzar proves amb ells sense tenir molt de temps d'execució.

#### 7.3.1 Resultats pretractament T1-w amb rotació i funció Binary Crossentropy.

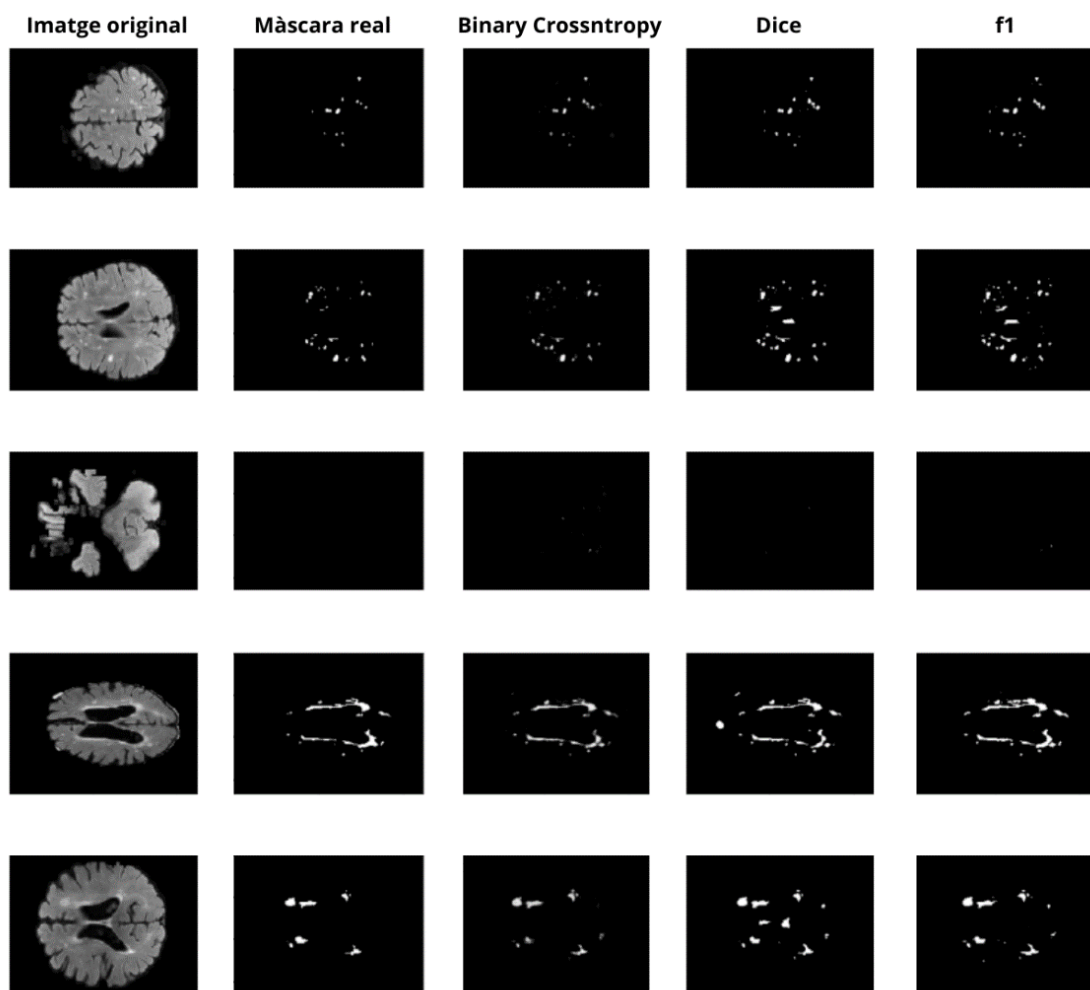


Figura 31: Resultats segons funció de pèrdua.



## Segmentació de lesions

*Taula 9: Resultats estadístics segons mètode.*

Funció	Accuracy	Precision	Recall	F1
Binary Crossentropy	0.9979	0.7481	0.8007	0.7735
Dice Loss	0.9980	0.6427	0.7581	0.6956
F1 Loss	0.9992	0.7041	0.7767	0.7386

Si observem la Taula 9, podem veure com totes les funcions assoleixen un valor molt alt d'Accuracy, tal com havíem suposat que passaria degut a la petita mida de les lesions. També veiem com totes tres funcions ens donen un millor valor de Recall que no de Precision, però la funció Binary Crossentropy ens dona el millor valor f1. Per poder entendre si els valors de la segmentació són bons, hem agafat de referència els valors obtinguts per els equips de recerca al treballar la segmentació d'aquesta base de dades en el WMH Challenge 2017, i veiem que els valors de f1 solen oscil·lar de 0.71 a 0.79. Amb aquesta informació podem observar com les nostres funcions ens obtenen uns valors molt decents de segmentació de lesions. Podem veure a la Figura 31 com les funcions Dice i f1 ens sobredimensionen les lesions, mentre que les lesions detectades per la funció Binary Crossentropy, s'ajusten bastant a les reals.

### 7.3.2 Resultats pretractament T1-w i funció Binary Crossentropy.

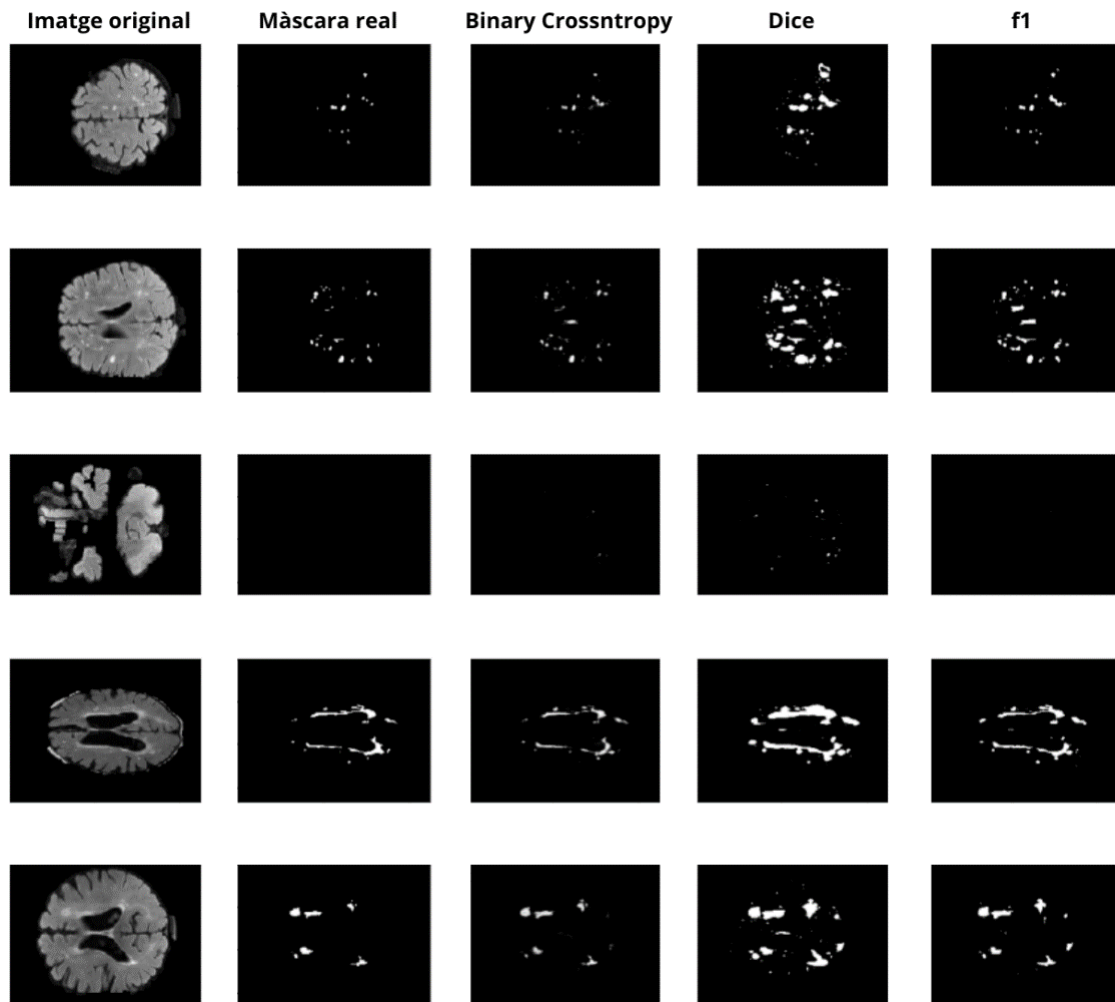


Figura 32: Resultats segmentacions segons funció de pèrdua.

Taula 10: Resultats estadístics segons funció.

Funció	Accuracy	Precision	Recall	F1
Binary Crossentropy	0.9943	0.7421	0.7749	0.7581
Dice Loss	0.9930	0.2788	0.9353	0.4296
F1 Loss	0.9975	0.5492	0.7959	0.6499

Si observem la Taula 10 podem veure com totes les funcions continuen assolint un valor molt alt d'Accuracy. En aquest cas, queda representat, la gran diferència entre els valors

de Precision i Recall de les funcions Dice i f1. Aquesta diferència ens indica que hi ha molt de sobredimensionament de les lesions segmentades, ja que casi no hi ha falsos negatius però trobem molts falsos positius. Podem veure a la Figura 32 com les funcions Dice i f1, en aquest cas, plasmen la notòria diferència en el sobredimensionament que ens realitza la funció Dice amb el de la funció f1, on el d'aquesta última és considerablement menor.

### 7.3.3 Resultats pretractament T1-w amb rotació i funció Dice Loss.

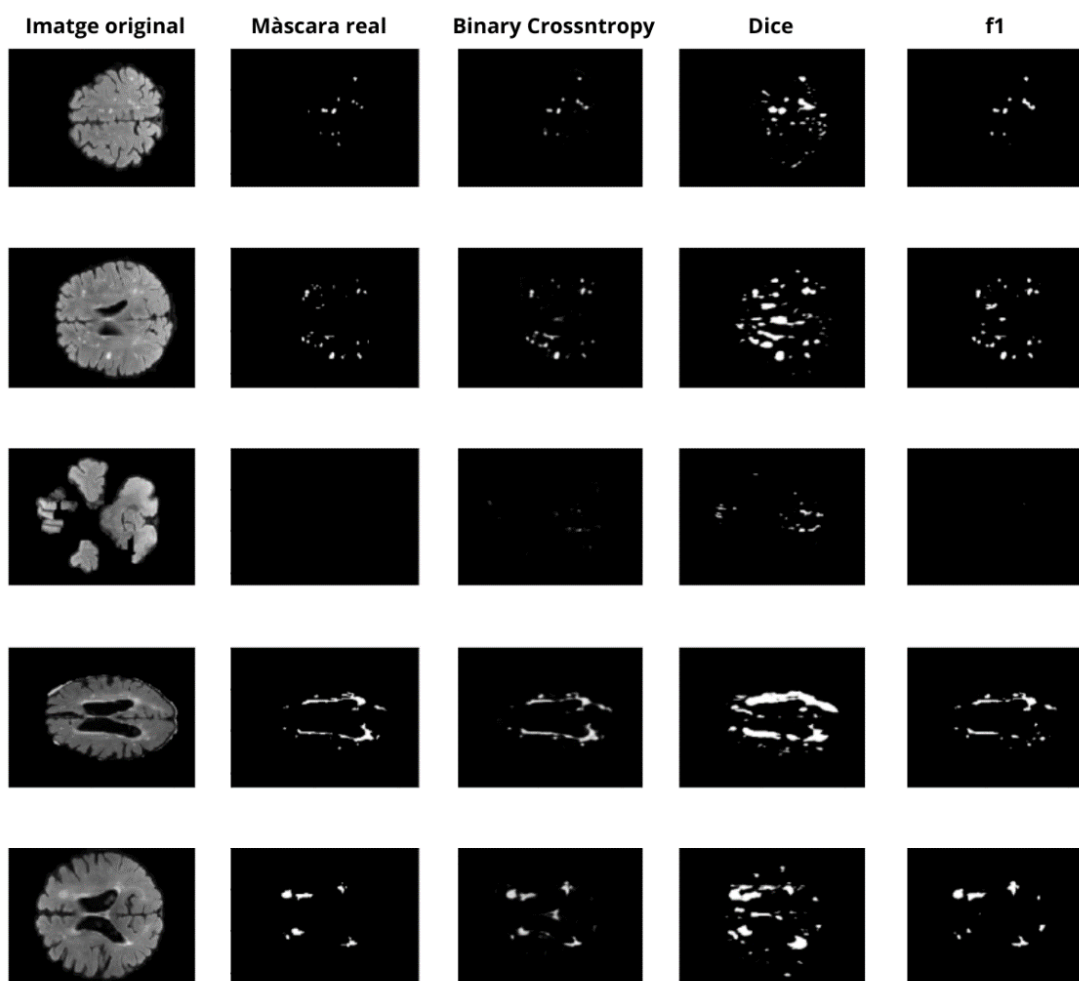


Figura 33: Resultats segons funció de pèrdua.

Taula 11: Resultats estadístics segons funció.

Funció	Accuracy	Precision	Recall	F1
Binary Crossentropy	0.9914	0.6845	0.8357	0.7526
Dice Loss	0.9863	0.0916	0.8213	0.1648
F1 Loss	0.9984	0.5331	0.7488	0.6228

Si observem la Taula 11, podem veure com només els valors del estadístic f1 de la funció Binary Crossentropy s'ha mantingut en un bon valor, els altres valors de f1 han baixat dràsticament degut als valors extremadament baixos de precisió que indiquen el sobredimensionament mencionat, i que podem observar en la Figura 33.

### 7.3.4 Resultats pretractament T1-w i funció Dice Loss.

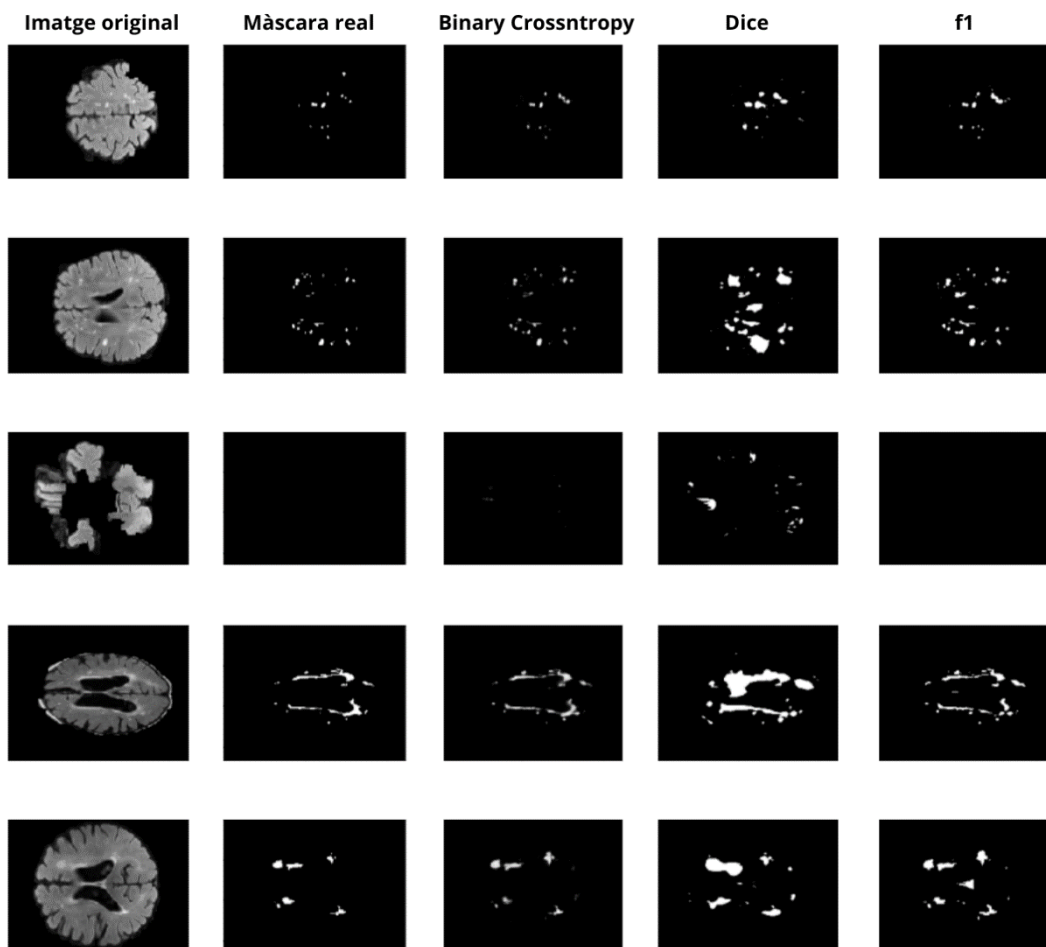


Figura 34: Resultats segons funció de pèrdua.

*Taula 12: Resultats estadístics segons funció.*

Funció	Accuracy	Precision	Recall	F1
Binary Crossentropy	0.9991	0.6323	0.7643	0.6921
Dice Loss	0.9939	0.1674	0.6899	0.2694
F1 Loss	0.9986	0.5571	0.8068	0.6591

Si observem la Taula 12, podem veure en aquest cas com no s'ha conservat el valor de f1 de la funció Binary Crossentropy i s'ha quedat bastant reduït. També veiem com el valor del Recall de la funció Dice que abans era molt elevat, s'ha reduït dràsticament. Podem observar en la Figura 34 com ens continua sobredimensionant molt el mètode per funció Dice per més que hagi baixat molt el valor del seu Recall. Això ens està indicant que no és capaç de detectar bé les lesions ja que ni sobredimensionant aconsegueix tenir valors de falsos negatius baixos.

### **7.3.5 Resultats sense Skull Stripping.**

Com a últims resultats, hem volgut comparar els resultats obtinguts amb els mètode d'Skull Stripping amb els resultats que obtindríem si haguéssim realitzat la segmentació directament de la imatge del escàner, per saber si realment ens ajuda el procés d'Skull Stripping en la segmentació de lesions d'esclerosis múltiple.

## Síntesi d'imatges de RM per la millora en la segmentació de lesions d'esclerosi múltiple

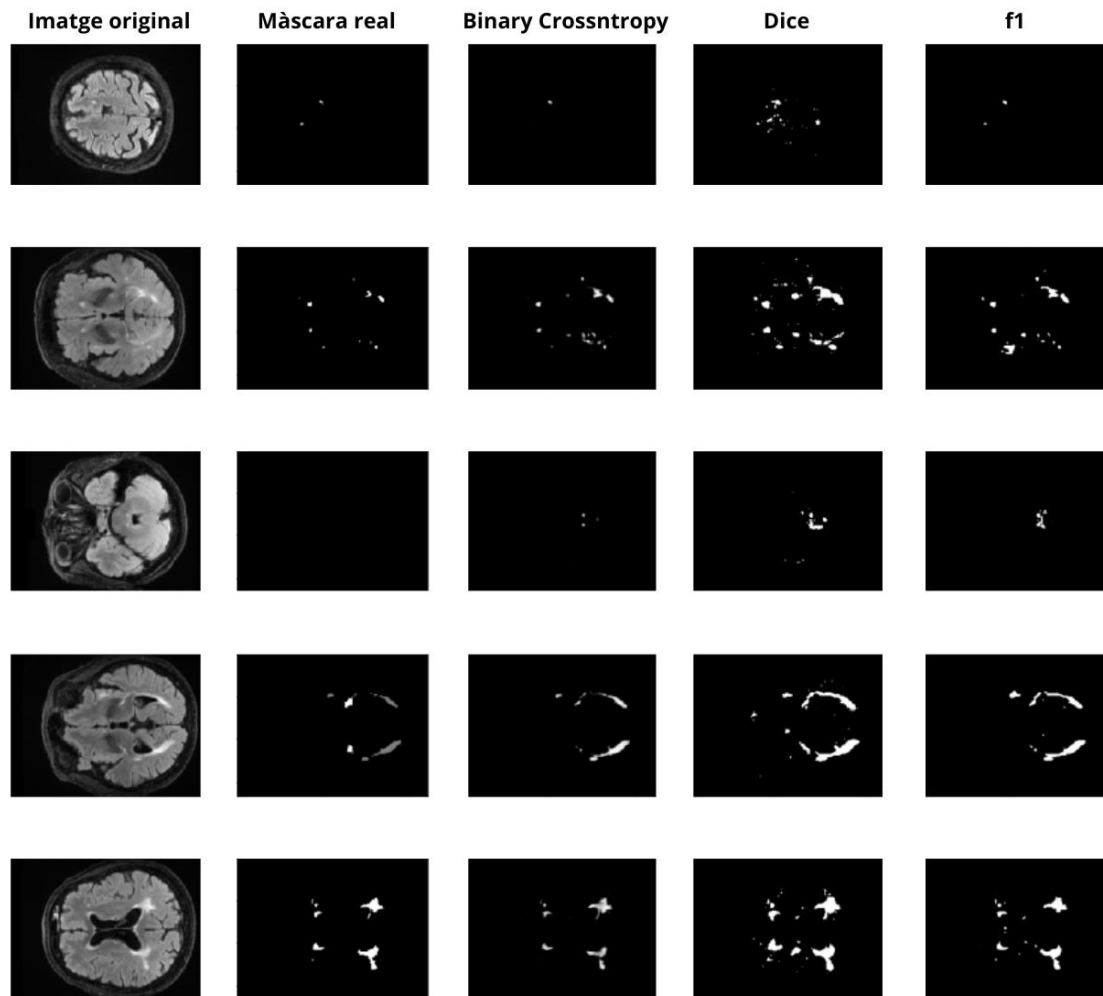


Figura 35: Resultats segons funció de pèrdua.

Taula 13: Resultats estadístics segons funció.

Funció	Accuracy	Precision	Recall	F1
Binary Crossentropy	0.9998	0.7010	0.8350	0.7622
Dice Loss	0.9976	0.3380	0.9199	0.4944
F1 Loss	0.9985	0.4460	0.8488	0.5847

Com podem veure en la Taula 13 aconseguim obtenir un molt bon valor de f1 mitjançant la fórmula de Binary Crossentropy, però els resultats baixen dràsticament en les altres dos fórmules, trobant en la funció f1 el pitjor valor de tots els models. Igualment, aconseguim trobar uns resultats similars a els altres mètodes d'Skull

Stripping. Si ens fixem en la Figura 35, podem observar que el sobredimensionament que ens produeixen les funcions, a diferència dels altres mètodes, es produeix sobretot a la detecció de lesions a on no n'hi ha i no tant a el engrandiment de les lesions existent, cosa que és un problema més gran ja que no podem realitzar un tractament posterior per corregir el error.

### **7.4 Resum resultats.**

Al analitzar els resultats de cada mètode, veiem com tots tenen en comú el sobredimensionament que generen les funcions Dice i f1. També veiem com el primer mètode emprat que es basava en utilitzar una imatge T1-w pretractada, amb rotació i amb l'Skull Stripping segmentat amb la funció Binary Crossentropy, és el mètode que ens dona millors estadístics en totes les funcions. Tot i així, podem observar un sobredimensionament clar de les funcions Dice i f1, que per més que sigui més suau que el que ens prediuen altres mètodes, no és negligible.

Per últim, podem veure la gran importància que té el utilitzar un bon mètode d'Skull Stripping, ja que si bé hem vist que ben utilitzat millora els resultats de la nostre segmentació, si l'apliquem de manera errònia, podem obtenir uns resultats pitjors que els que obtindríem si no l'haguéssim realitzat.

La nostre base de dades de segmentació de lesions, contenia imatges de tres escàners diferents, creant així un model amb certa flexibilitat, i podent agafar com a vàlids els mètodes emprats per obtenir uns bons resultats amb aquesta base de dades.





Adaptabilitat mètodes segmentació.

## **8 ADAPTABILITAT MÈTODES SEGMENTACIÓ.**

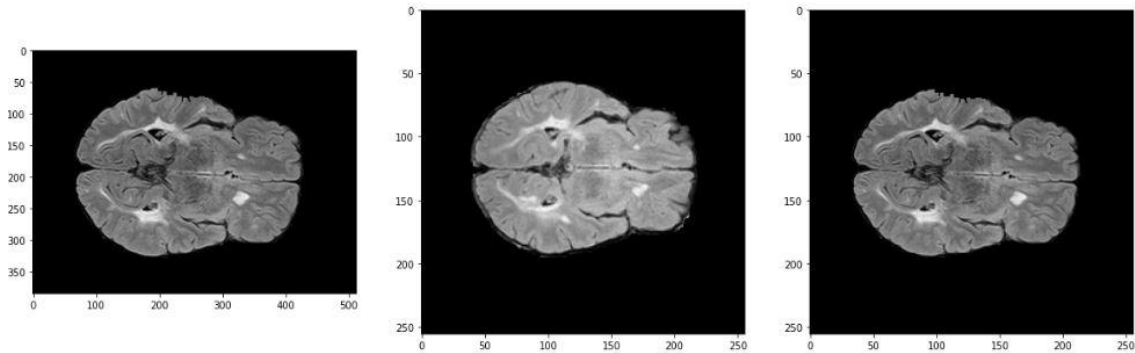
Per comprovar l'eficàcia de la segmentació de lesions d'esclerosi múltiple, no en tenim prou amb els resultats obtinguts de les dades de validació dels models de segmentació. Si bé els resultats obtinguts ens serveixen per determinar els millors mètodes de segmentació, necessitem comprovar les prediccions que ens realitzen els models si els utilitzem amb altres escàners. Al realitzar aquesta prova, podrem veure representada l'adaptabilitat del model davant d'inputs provinents d'imatges de diferents escàners.

Per visualitzar l'adaptabilitat dels models, utilitzarem informació de 10 pacients amb lesions d'esclerosi múltiple que han passat per tres hospitals diferents a realitzar-se una ressonància magnètica (Santa Catarina, Trueta (disposem de dades de fa un any i dades actuals) i Vall d'Hebron) .Això ens permetrà observar com el model ens prediu les mateixes lesions, visualitzades de manera diferent.

### **8.1 Pretractament de dades.**

Les imatges de la nostre base de dades, provenen de tres hospitals diferents, que tenen escàners diferents. Per poder comparar els mètodes, necessitem que les imatges de tots tres hospitals siguin com més semblants possibles. Al fer que els inputs s'assemblin entre els diferents escàners, aconseguim eliminar les diferències en els resultats que provindrien del format del input introduït, aconseguint així comparar millor la capacitat d'adaptabilitat del model. Aquests pretractaments de semblança, no inclouen canvis en la coloració ni síntesis d'imatges per fer que els escàners, sinó que consta de processos de canvis de mida i relacions d'aspecte, per tal que totes les imatges es deformin de la mateixa manera al ser introduïdes al model.

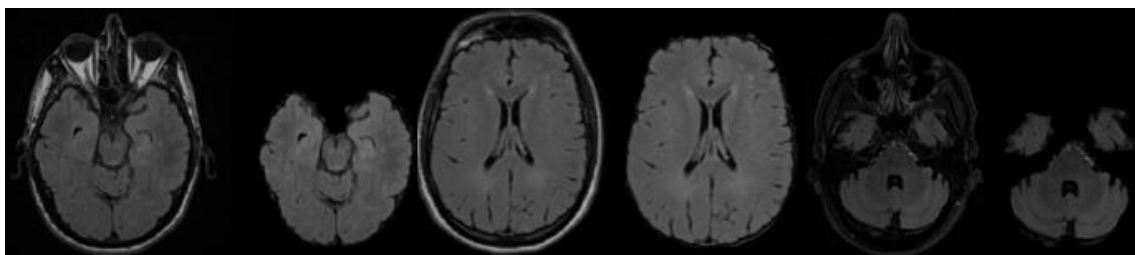
Per saber en quins casos fa falta aquest pretractament, analitzem les mides de les imatges provinents dels escàners dels tres hospitals. Tant els escàners de l'Hospital Santa Catarina com els del Trueta, ens generen unes imatges de relació d'aspecte 1:1. Els escàners de la Vall d'Hebron però, ens generen imatges 3:4.



*Figura 36: D'esquerra a dreta: imatge original Vall d'Hebron, imatge original Santa Catarina, imatge retocada Vall d'Hebron*

En la Figura 36, podem veure com tot hi que les relacions d'aspecte són diferents, el que varia respecte les imatges d'altres escàners és el fons de la imatge. Per solucionar aquest problema, hem realitzat un padding (Addició de píxels negres en els extrems de la imatge) fins que hem assolit la relació d'aspecte 1:1. Podem apreciar com hem aconseguit així la semblança que volíem tenir entre les diverses imatges.

Un cop tenim les imatges en el format correcte, hem de realitzar el procés que hem determinat al apartat anterior. Primerament realitzarem el Skull Stripping amb la funció de pèrdua Binary Crossentropy. A diferència dels casos de la base de dades de la segmentació de lesions, no disposem d'imatges T1-w de les ressonàncies. Així que realitzarem la segmentació directament des de la imatge FLAIR.



*Figura 37: Exemples segmentació a partir d'imatge FLAIR*

En la Figura 37, podem veure com tot hi no haver pogut segmentar les imatges T1-w, les segmentacions es poden realitzar bé en les imatges FLAIR provinents dels diversos hospitals.

## 8.2 Resultats.

En la segmentació de lesions, hem obtingut diversos models segons els diversos mètodes de Skull Stripping. Això no significa però, que cada model està vinculat a la

## Adaptabilitat mètodes segmentació.

manera en la que hem realitzat el seu pretractament. Els diversos mètodes de Skull Stripping, ens permetien entrenar amb diverses dades i comprovar l'eficàcia de diversos models de segmentació de lesions per tal de determinar el millor model d'entre aquests. Els models obtinguts, poden funcionar correctament independentment del mètode de Skull Stripping que fem per obtenir les dades d'input.

És per aquest motiu, que per analitzar l'eficàcia de segmentació en els diferents casos dels hospitals, comprovarem l'eficàcia amb models diferents per comparar els resultats d'aquests. Per fer el anàlisi de l'eficàcia, utilitzarem el model obtingut de les combinacions de les funcions Binary Crossentropy i f1 Loss amb l'Skull Stripping a partir de T1-w rotat i segmentat amb la funció Binary Crossentropy. Dels altres models, utilitzarem les combinacions amb la funció Binary Crossentropy.

Per futures referències, abreviarem els noms de les 5 combinacions que comprovarem (Taula 14).

*Taula 14: Nous noms combinacions*

Nom abreviat	Pretractament Skull Stripping	Funció Skull Stripping	Funció Segmentació Lesions
Combinació A	T1-w rotat	Binary Crossentropy	Binary Crossentropy
Combinació B	T1-w rotat	Binary Crossentropy	F1 Loss
Combinació C	T1-w	Binary Crossentropy	Binary Crossentropy
Combinació D	T1-w rotat	Dice Loss	Binary Crossentropy
Combinació E	T1-w	Dice Loss	Binary Crossentropy

Al no ser casos resolts, no podem realitzar cap càlcul de mètriques. Totes les observacions i valoracions es realitzaran a partir de les imatges obtingudes de les prediccions.

### 8.2.1 Resultats combinació A:

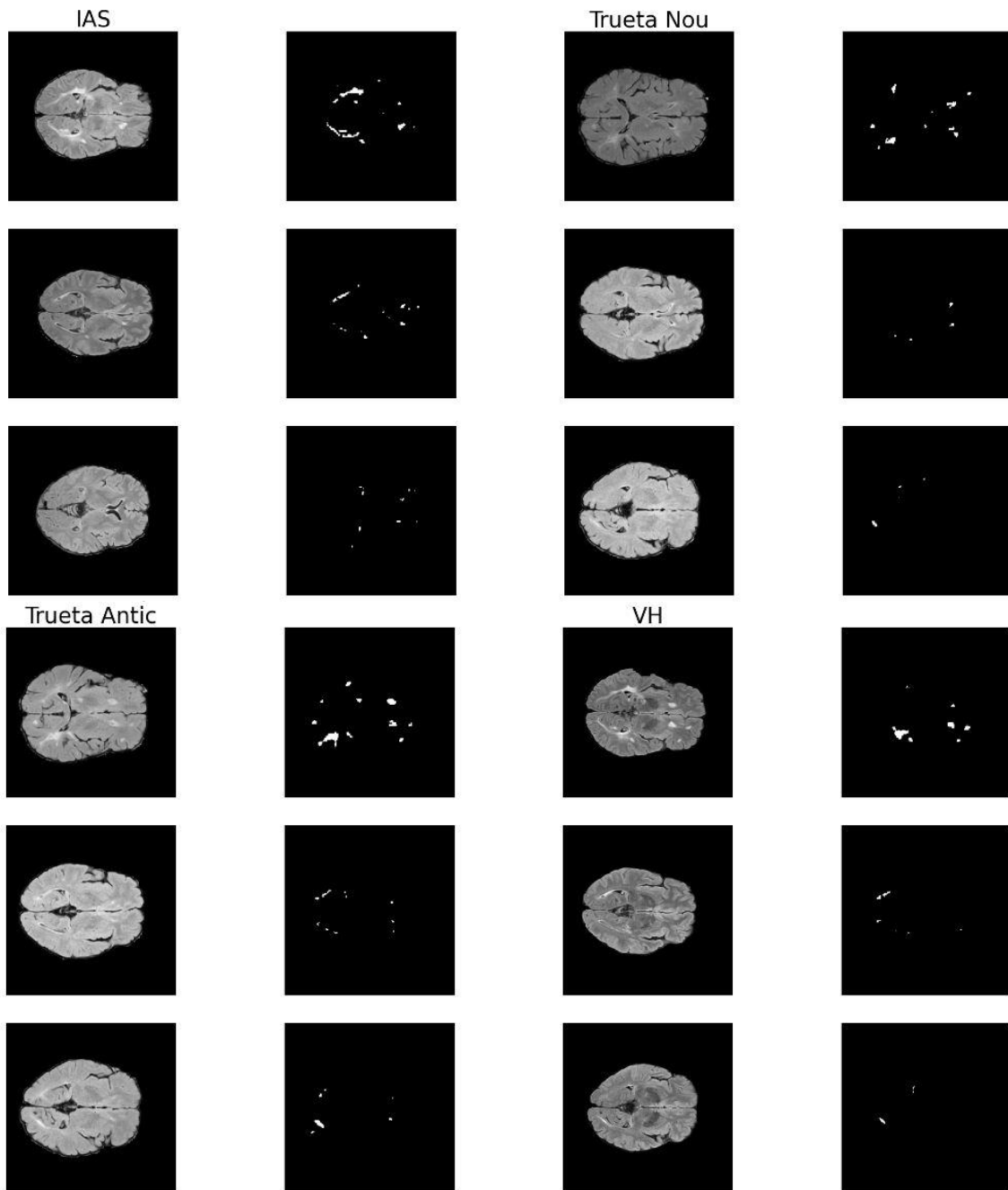


Figura 38: Resultats combinació A segons 3 casos de cada hospital.

Com podem observar en la Figura 38, aconseguim segmentar bé les lesions provinents dels escàners del hospital Santa Catarina, però observem certa tendència al sobre-dimensionament, ja que en alguns casos ens detecta lesions a on no n'hi ha. Les prediccions que realitza al les imatges provinents del Trueta (tant les noves com les

Adaptabilitat mètodes segmentació.

antigues), també estan ben segmentades, però en comptes de sobredimensionament, trobem que en alguns casos detecta menys del que hi ha. Finalment, els escàners de la Vall d'Hebron, veiem com no aconsegueix detectar molt bé les lesions, deixant-se de segmentar lesions de una mida considerablement gran.

### 8.2.2 Resultats combinació B:

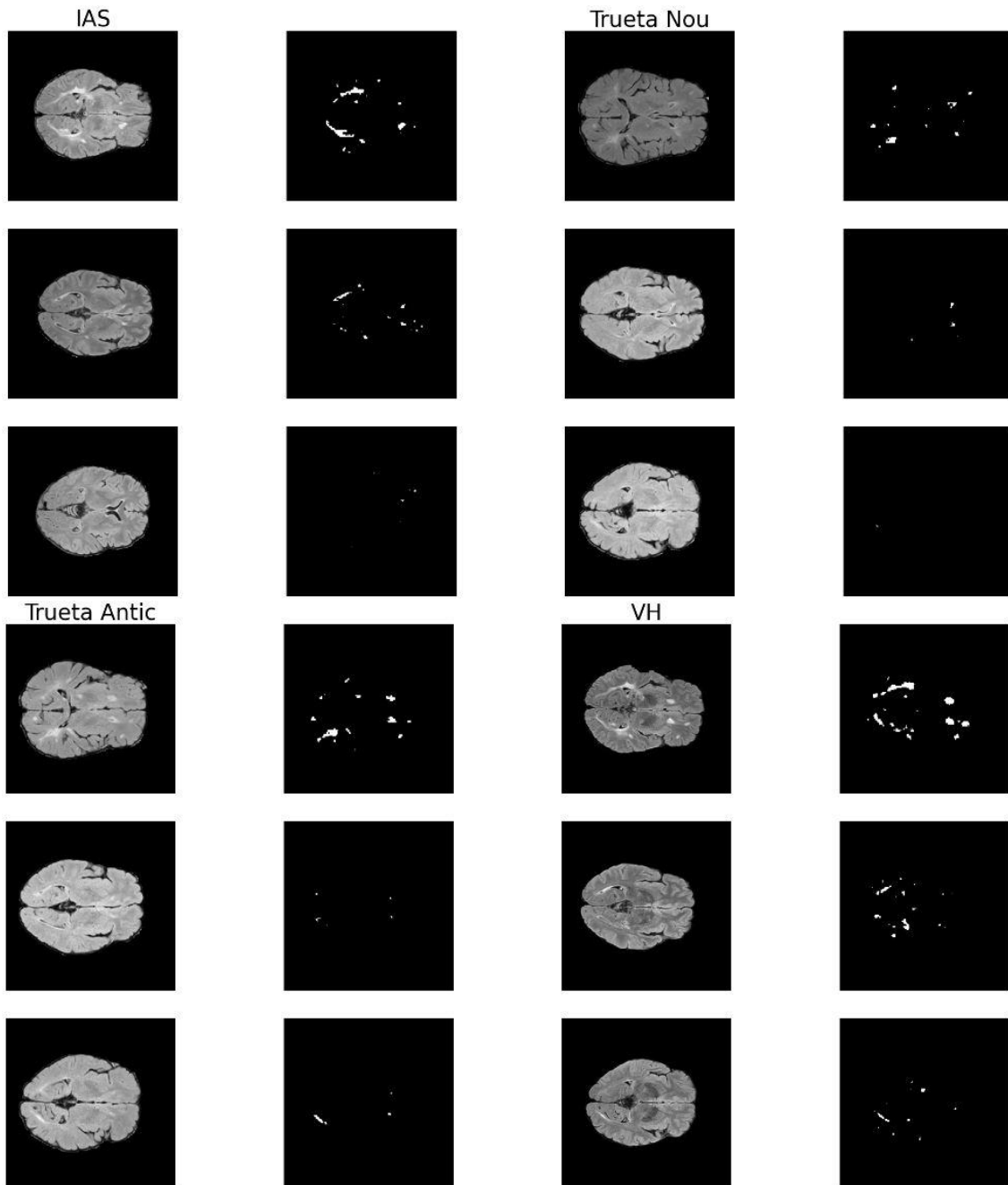


Figura 39: Resultats combinació B segons 3 casos de cada hospital.

Al utilitzar la funció f1 per realitzar la segmentació, esperàvem uns resultats bastant sobre-dimensionats. Es pot observar en la Figura 39 com a diferència del cas anterior, ara el model ens segmenta les lesions dels escàners de la Vall d'Hebron fins al punt de segmentar lesions que no hi són realment, i sobredimensionant les existents. Per la resta d'escàners, podem observar com no solem trobar resultats sobredimensionats i veiem com no ens segmenta tampoc lesions falses, obtenint així uns bons resultats.

Adaptabilitat mètodes segmentació.

### 8.2.3 Resultats combinació C:

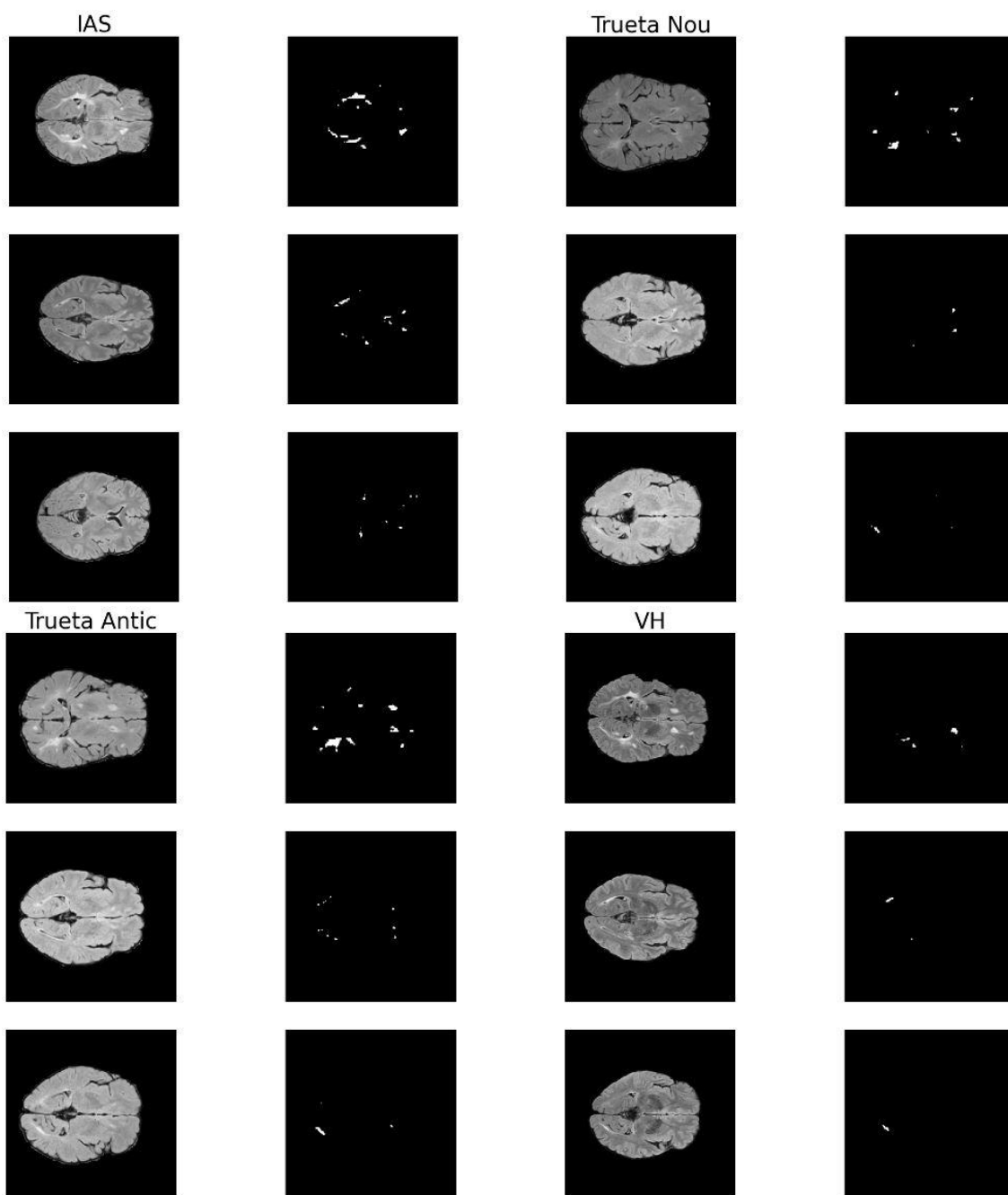
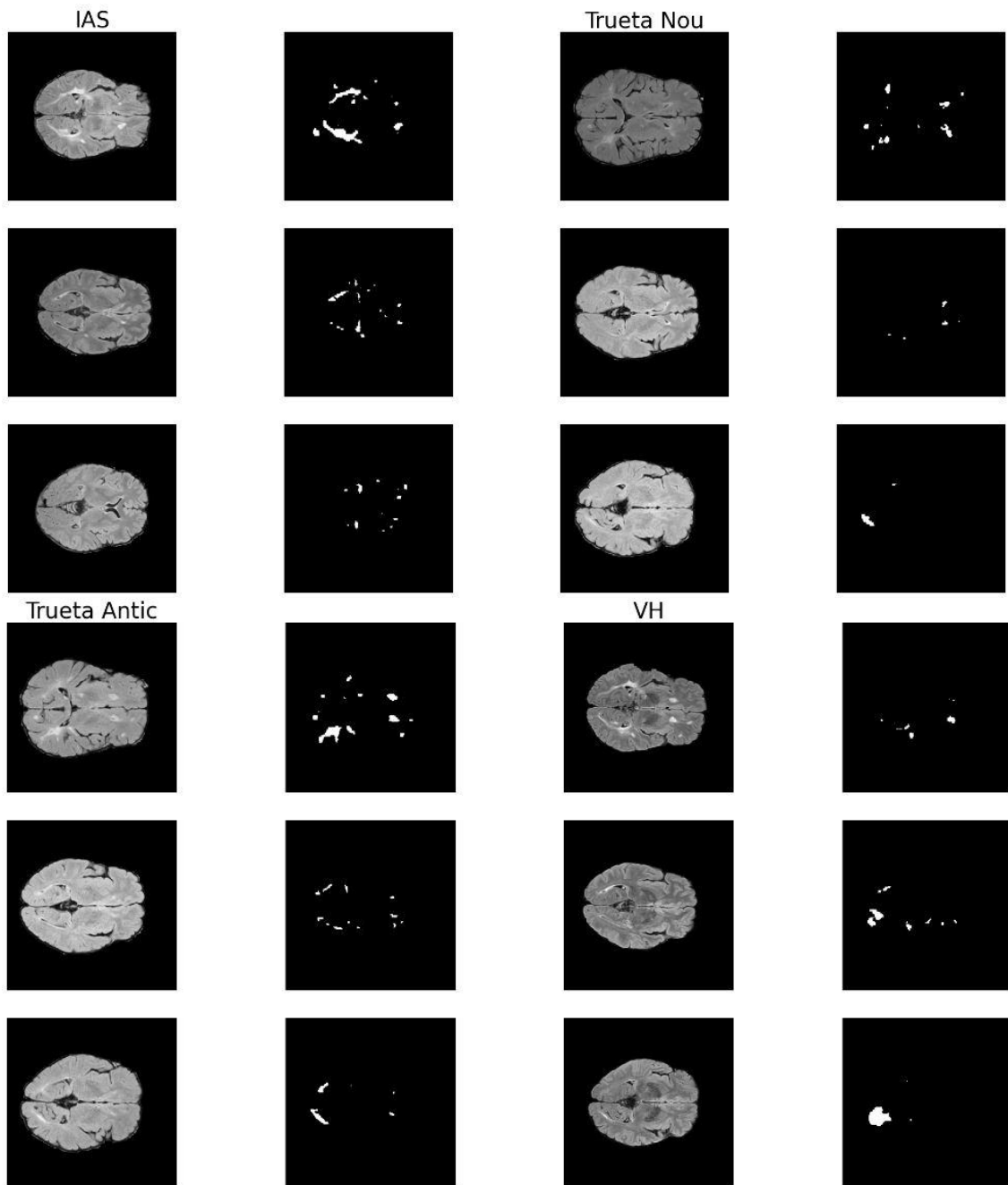


Figura 40: Resultats combinació C segons 3 casos de cada hospital.

Els resultats de la combinació C, com podem veure en la Figura 40, són molt similars als obtinguts en la combinació A, però sense tant sobre-dimensionament. Podem apreciar com les lesions predites dels escàners són més petites, fins al punt a on en el cas de les imatges de la Vall d'Hebron, casi que no detecta lesions algunes.

### 8.2.4 Resultats combinació D:



*Figura 41: Resultats combinació D segons 3 casos de cada hospital.*

Com podem veure en la Figura 41, aquesta combinació ens segmenta amb una sobredimensionament molt més marcat, juntament amb la segmentació de lesions que no són certes. Però a diferència de la combinació B, aquest sobredimensionament no fa que ens predigui correctament les imatges dels escàners de la Vall d'Hebron.



Adaptabilitat mètodes segmentació.

### 8.2.5 Resultats combinació E:

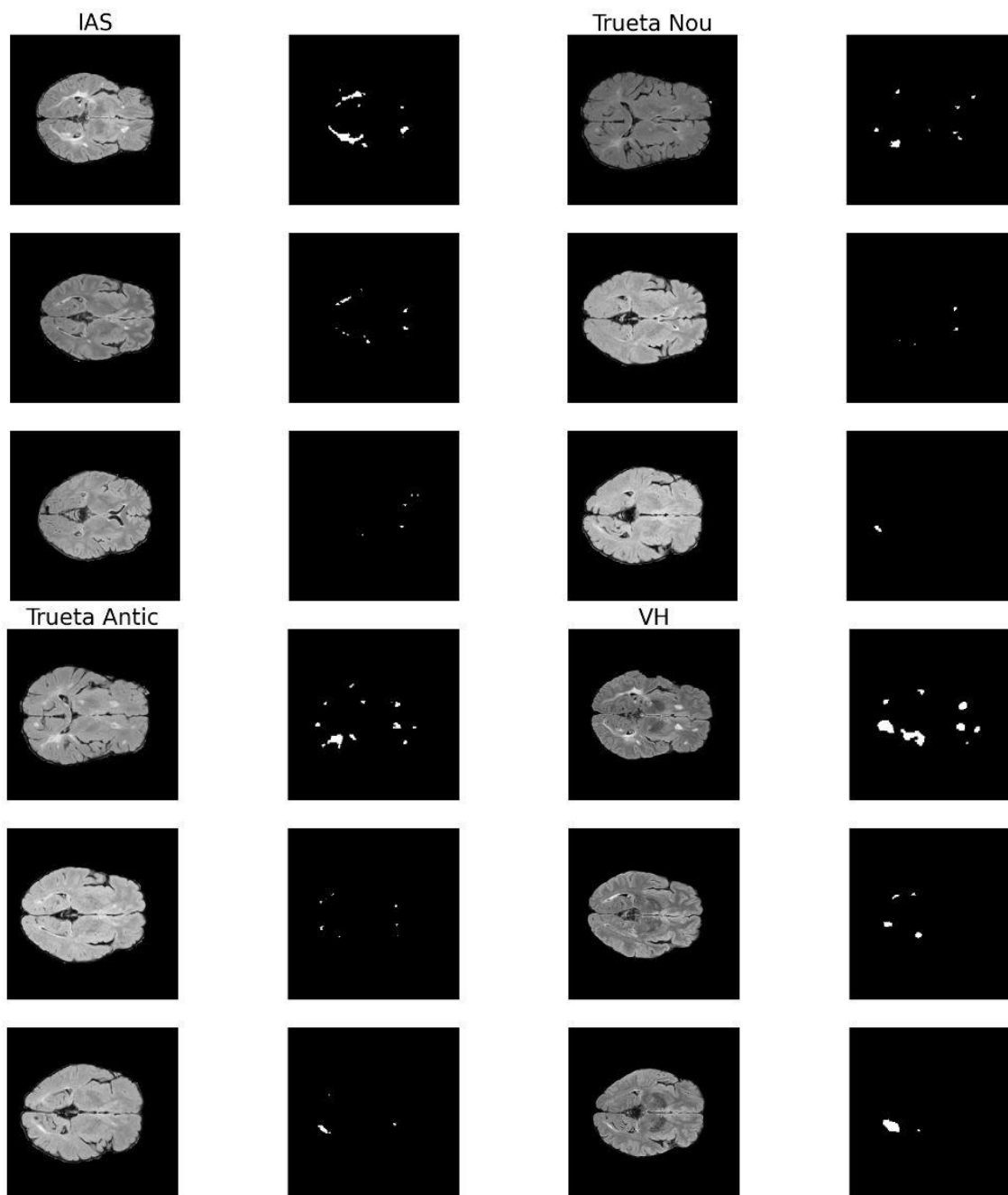


Figura 42: Resultats combinació E segons 3 casos de cada hospital.

Finalment tenim els resultats provinents de la combinació E (Figura 42). Podem veure com de la mateixa manera que la combinació D, tenim un sobredimensionament més marcat. La diferència és que amb aquesta combinació, no segmentem lesions falses, a

més a més aconseguim detectar millor les lesions dels escàners de la Vall d'Hebron. Tot hi això, no aconseguim obtenir uns bons resultats d'aquests.

### **8.3 Resum resultats.**

Al comparar els casos, hem pogut observar com els models tenen problemes en la detecció de lesions d'escàners provinents de l'Hospital de Vall d'Hebron. També veiem com en tots els casos, aconseguim detectar amb precisió els casos del hospital Santa Catarina. Els casos del Trueta també els aconseguim segmentar bé, però solem trobar errors més de manera més recurrent que amb els escàners de Santa Catarina. Pel que fa a les combinacions, hem pogut comprovar com aquelles provinents dels models de segmentació de lesions obtinguts a partir de un Skull Stripping fet a partir de la funció Dice Loss, ens obtenen uns resultats més sobredimensionats, amb bastants falsos positius però pocs falsos negatius. El sobredimensionament de les lesions, pot corregir-se en una etapa de post-tractament, si és constant i no molt accentuat, amb un simple procés d'erosió. El problema que tenim si hi ha massa sobredimensionament i no és constant és que per corregir-ho hauríem de fer una erosió més gran i podríem acabar eliminant petites lesions segmentades en el procés. Per aquests motius, podem determinar que les primeres tres combinacions són més adequades per la segmentació de lesions.

Amb la comparació també hem pogut confirmar el problema d'adaptabilitat que ens trobem si les dades d'entrenament del model de segmentació de lesions, varien molt de les dades que volem segmentar. Això s'ha vist clar en el cas d'aquelles imatges provinents de la Vall d'Hebron, les quals només la combinació B ens les ha pogut detectar. Per solucionar aquest problema, ens plantejem buscar un mètode que ens permeti transformar aquestes imatges de la Vall d'Hebron en unes imatges que siguin més compatibles amb el model de segmentació.

### **8.4 Síntesis d'imatges.**

Com hem vist als resultats obtinguts al aplicar el models de segmentació en casos reals, les dades més compatibles amb els models són les del Hospital Santa Catarina. Per realitzar la síntesis d'imatges tenim una base de dades de 15 persones sense lesions, que es van escanejar 3 cops per hospital (en la Vall d'Hebron només 2). Al poder etiquetar per cada imatge provinent de la Vall d'Hebron, al seu equivalent al Hospital Santa Catarina, ens permet realitzar la síntesis d'imatges a partir d'una xarxa neuronal convolucional U-Net a on l'input seran les imatges de la Vall d'Hebron i la màscara seran les que vinguin del hospital Santa Catarina.

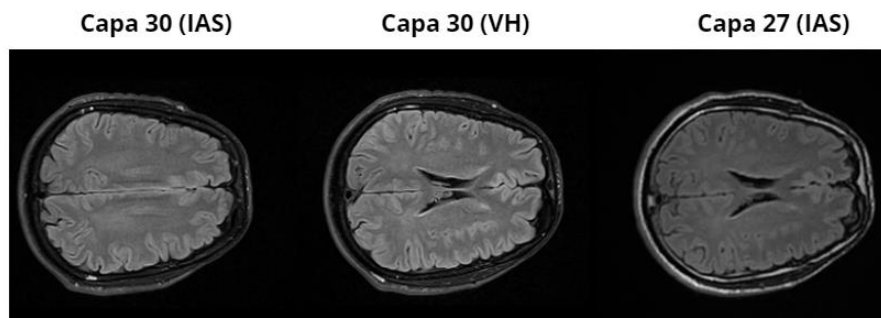
Adaptabilitat mètodes segmentació.

#### 8.4.1 Pretractament.

A la hora d'obtenir un bon model, necessitem aportar a la nostre xarxa neuronal amb la major quantitat de dades possibles. En els altres casos, disposàvem de suficients imatges per entrenar els nostres models, però ara trobem que disposem d'una quantitat limitada de dades per poder entrenar el nostre model. Si bé en el cas de la segmentació de lesions, en la base de dades teníem 60 casos i ara en tenim 30, trobem que hi ha una diferència molt gran en el nombre de capes que té cada cas. Hem de recordar que cada imatge és volumètrica, i les imatges que nosaltres predim son en 2 dimensions, per tant la quantitat de capes que tingui cada imatge volumètrica, afectarà al nombre d'imatges amb les que podrem entrenar el model. Si ara ens fixem en les bases de dades, en la base de dades de la segmentació de lesions, disposem de 60 casos d'una mitjana de 260 capes. En canvi, en la nostre base de dades actual, disposem de 30 casos de 44 capes. Això fa que el nombre d'imatges que podrem donar al model perquè entreni, és significativament inferior.

Per donar solució a aquest problema, hem realitzat un augment en la nostre base de dades a base de girar la imatge a partir del eix horitzontal, permetent-nos duplicar el nombre de casos. Tot hi així la diferència continua sent molt gran, per poder veure com afecten el nombre de dades que li donarem al nostre model, també farem un model de síntesi d'imatges per l'hospital Trueta, al qual també li aplicarem un augment de dades.

A part de realitzar un augment en el nombre de dades, al necessitar comparar les imatges dels diferents escàners capa a capa, necessitarem dos processos molt importants. Primerament, el nombre de capes ha de ser el mateix entre les dos imatges. Això es compleix per tots els casos excepte als de la Vall d'Hebron que tenen 46 capes (respecte les 44 capes de la resta de casos). Per fer que el nombre de capes sigui igual, simplement canviem la mida de la imatge volumètrica perquè estigui més comprimida (mateix procés que canviar la resolució de la imatge, però en la dimensió z). L'altre procés, i el més important dels 2, és l'anivellament entre capes. Quan es realitza una ressonància magnètica, els punts de partida i finalització són constants ni entre casos ni entre hospitals. En cada relació cas hospital a cas hospital, necessitem que els punts de partida i final coincideixin, per tal de poder afirmar que les capes d'un cas corresponen a les capes de l'altre. Per aconseguir aquest anivellament, hi ha diversos mètodes, en el nostre cas al tenir una base de dades reduïda, hem realitzat l'anivellament manualment, eliminant capes inicials o finals per tal d'anivellar-les (Figura 43).

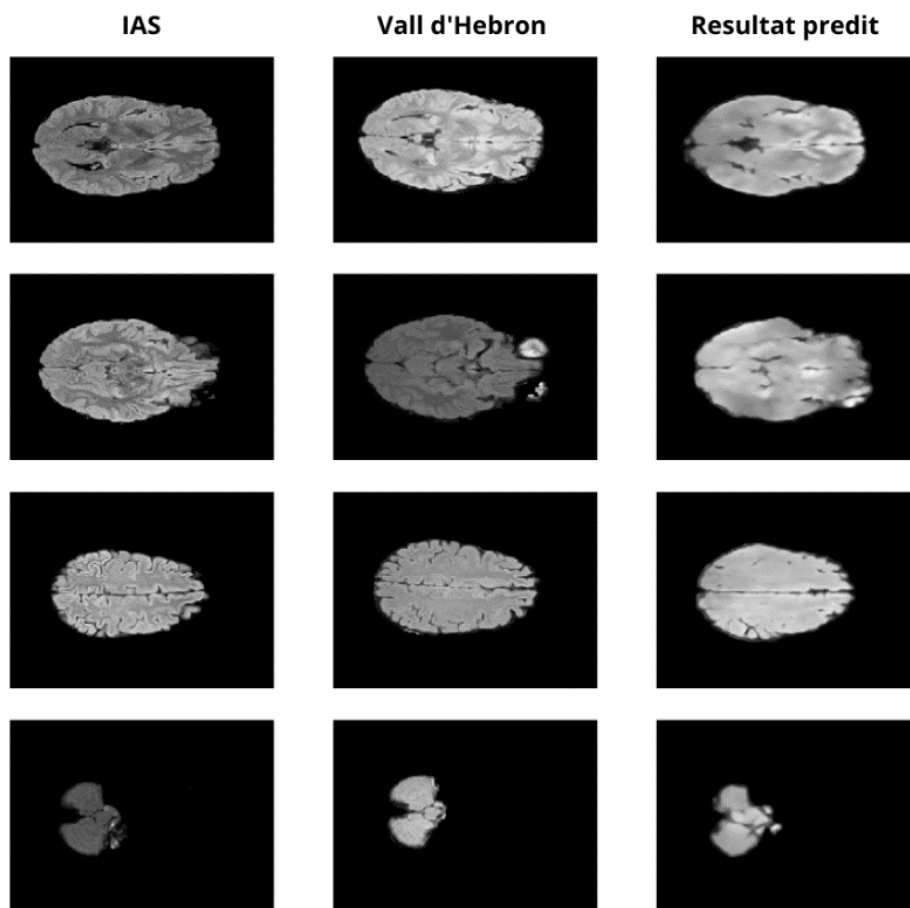


*Figura 43: Diferència entre les imatges associades a cada capa en diferents hospitals (Vall d'Hebron i Santa Catalina).*

#### **8.4.2 Entrenament.**

Per crear el model, hem treballat amb diverses funcions de pèrdua com la Poisson, MSE i MAE, però cap d'aquestes ens aconseguia convergir en cap resultat. La funció que ens ha permès obtenir un resultat ha sigut la SSIM Loss.

### 8.4.3 Resultats síntesis Vall d'Hebron.



*Figura 44: Resultats prediccions del model de síntesis d'imatges per l'hospital de la Vall d'Hebron*

Tal com podem veure en la Figura 44, si bé el model ens arriba a un resultat, no aconsegueix delimitar-lo prou bé i crea una imatge molt borrosa de la qual n'és difícil extreure'n característiques. Per intentar solucionar aquesta borrositat, abans de comparar els casos amb la transformació de la síntesis, hem aplicat un filtre que afila la imatge i treu borrositat. Si ara aquest model l'apliquem als casos que teníem, podem veure com encara que li haguem aplicat el filtre, no obtenim una bona imatge final, i no aconseguim millorar el resultat (Figura 45).

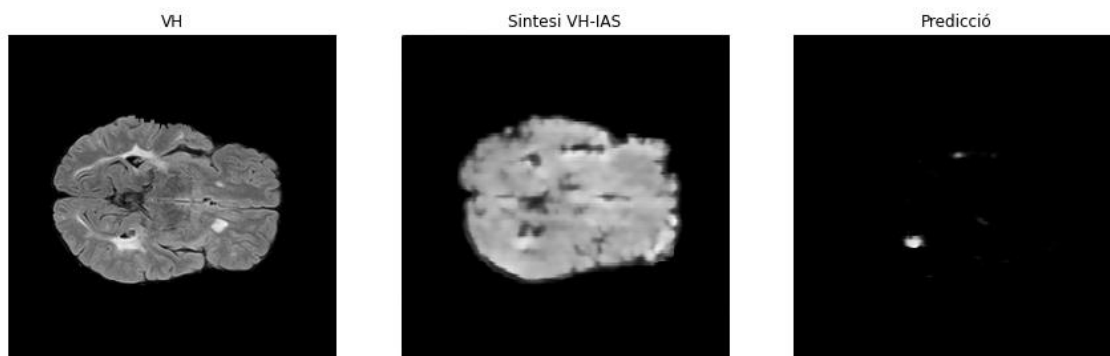


Figura 45: D'esquerra a dreta: Imatge original, imatge transformada, predicció feta.

#### 8.4.4 Resultats síntesis Trueta.

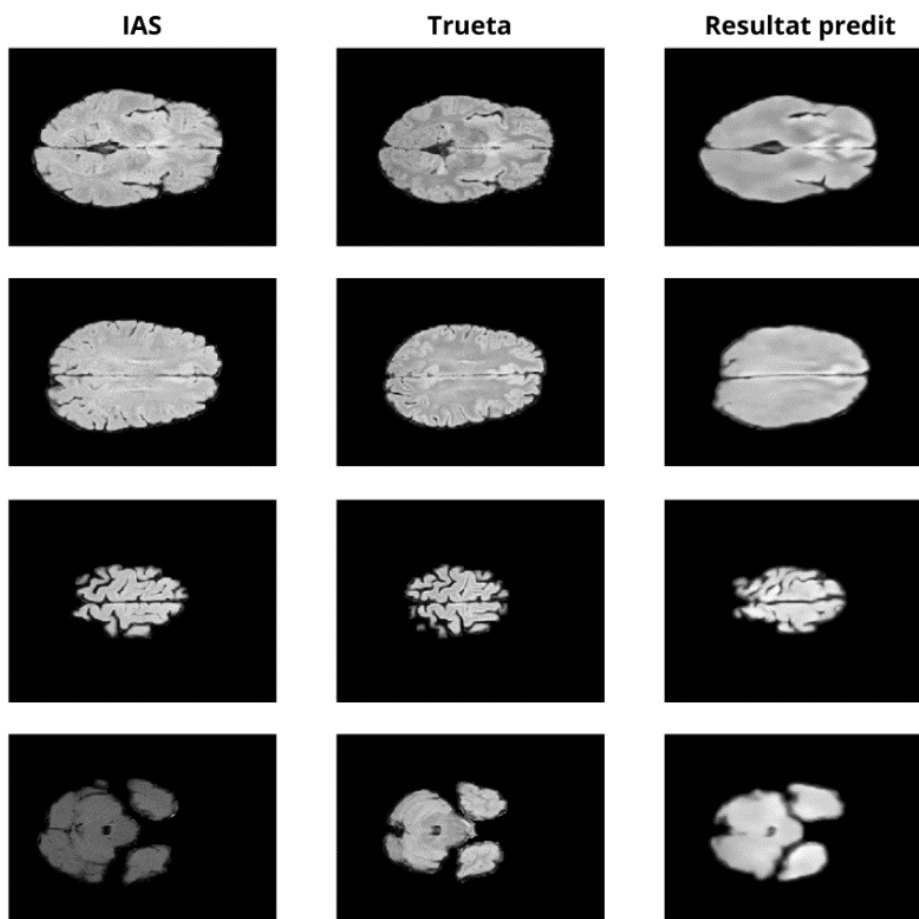


Figura 46: Resultats prediccions del model de síntesis d'imatges de l'hospital Trueta

Ara hem provat d'obtenir resultats amb una base de dades més gran, veiem en la Figura 46, com els resultats han millorat una mica, però continuem trobant molta borrositat en

Adaptabilitat mètodes segmentació.

les imatges, cosa que perjudica a la segmentació de lesions. Veiem com encara que utilitzem el mètode per transformar les imatges del Trueta en imatges del hospital Santa Catarina i ho passem per un filtre que ajudi a treure borrositat, no aconseguim obtenir bon resultat (**Error! Reference source not found.**).

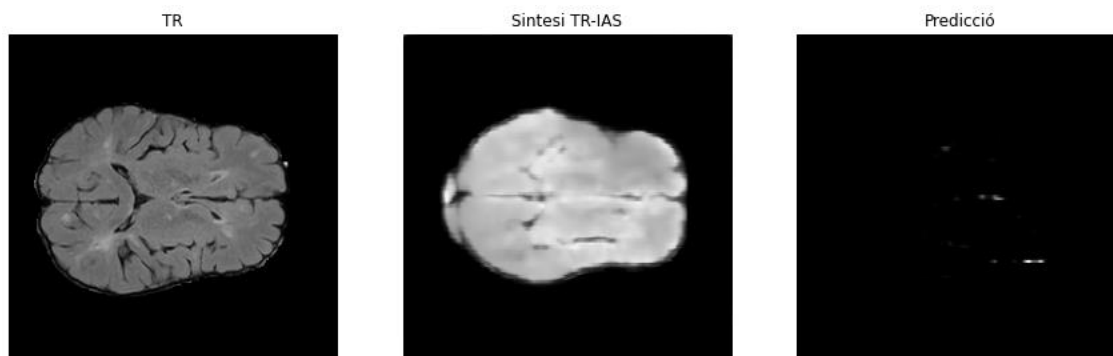


Figura 47: D'esquerra a dreta: Imatge original, imatge transformada, predicció feta.

#### 8.4.5 Conclusions.

Hem pogut comprovar com al aplicar els models en casos diferents als entrenats, ens hem trobat amb un problema d'adaptabilitat. Aquest problema, no hi hem pogut donar solució. Aquesta falta de solució podria ser degut a que el model no funciona o no tenim suficients dades per poder convergir en una bona solució. Per sortir de dubtes, vam realitzar la síntesi d'imatges per convertir una imatge FLAIR en T1-w, aquest model ens podria servir per realitzar el Skull Stripping a partir de imatges T1-w en aquells casos en que només disposem d'imatges FLAIR. Per fer això, hem utilitzat la base de dades de segmentació de lesions, que com hem comentat anteriorment, és molt més extensa que la que hem disposat per fer la síntesi entre hospitals. Si observem els resultats obtinguts, podem veure com encara trobem borrositat en les imatges, però també veiem com el model ha sigut capaç de predir amb molta més exactitud la imatge que volíem obtenir (**Error! Reference source not found.**).

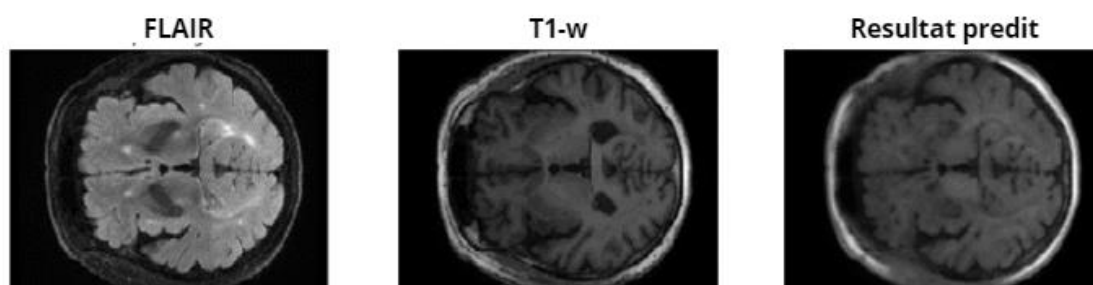
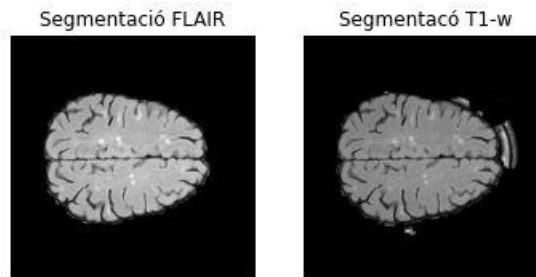


Figura 48: D'esquerra a dreta: Imatge original, imatge transformada, predicció feta.

Això ens ha permès aplicar aquest model per realitzar el Skull Stripping a partir de imatges FLAIR sintetitzades per semblar T1-w, i si bé no ens dona uns resultats tan

bons com els que obtindríem si no féssim la síntesis, si que veiem que són prou encertats (**Error! Reference source not found.**).



*Figura 49: Comparativa segmentació via imatge FLAIR i la segmentació via la transformada a T1-w d'aquesta*

Això ens reafirma, que si aconseguíssim augmentar la nostre base de dades, podríem aconseguir un bon model de síntesis d'imatges. Per altre banda, també veiem com per tal d'obtenir bons resultats de síntesis d'imatges a partir d'una xarxa U-Net, necessitem més dades que les que necessitaríem per obtenir un bon resultat de segmentació.



## **9 RESUM DEL PRESSUPOST**

TOTAL PRESSUPOST PER CONTRACTE € 11250

Aquest pressupost d'execució per contracte puja a  
(ONZE MIL DOS-CENTS CINQUANTA EUROS)

Arnau Cabarrocas Salvador  
Autor del projecte

Girona, 6 de setembre de 2022



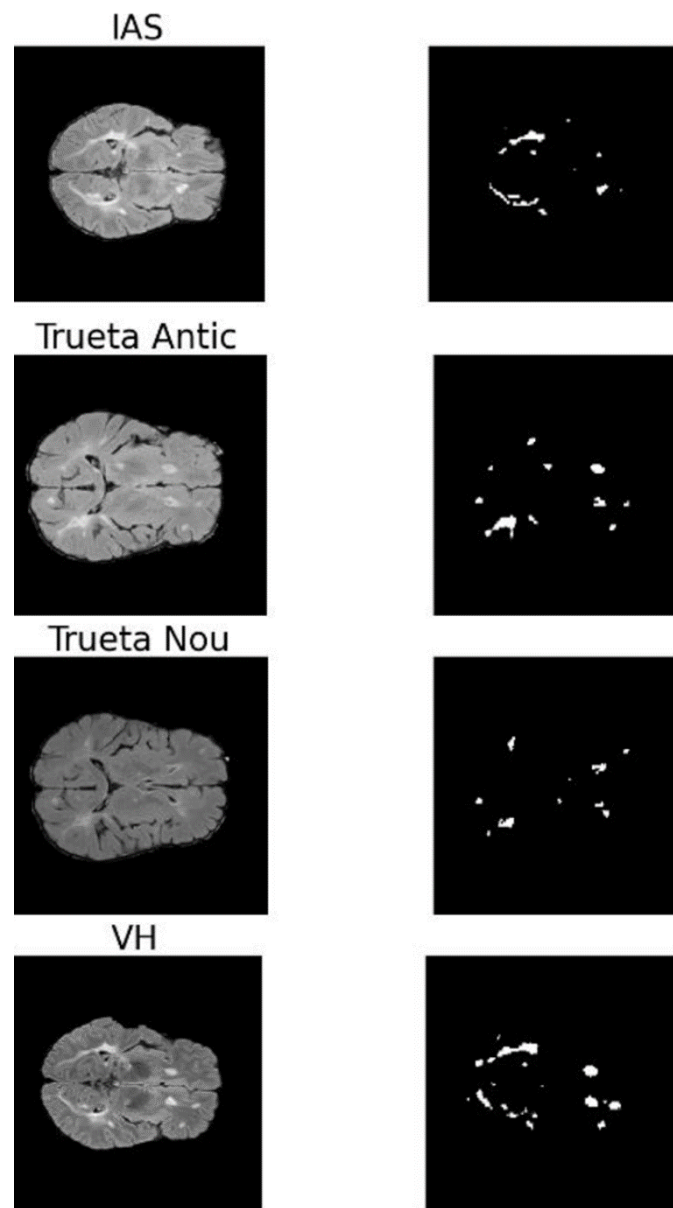
## 10 CONCLUSIONS

Al iniciar el treball hem pogut veure la gran presència que té en el món la esclerosi múltiple i la importància que té la intel·ligència artificial en el camp de la medicina, essent un dels camps d'estudi de més interès per els investigadors, al automatitzar processos que als experts els hi comportarien temps i esforç. És per això que un dels camps de més atenció de la aplicació mèdica de la intel·ligència artificial és la segmentació d'imatges mèdiques per l'extracció d'àrees d'interès.

Durant l'estudi de les diferents aproximacions a la segmentació de lesions d'esclerosi múltiple, hem pogut veure com les xarxes neuronals convolucionals, són molt útils per la segmentació automàtica de lesions, essent un dels millors mètodes de segmentació si es disposen de les suficients dades. A més a més, hem teoritzat l'aparició d'un problema d'adaptabilitat en que es troben els models de segmentació de lesions, al treballar amb dades diferents a les que han estat entrenats els models.

En el transcurs de la creació dels models necessaris per segmentar les lesions podem observar l'efecte que causen els diferents hiperparàmetres del model i els pretractaments de les imatges de ressonància magnètica en els resultats. Hem vist la gran importància de la selecció de les funcions de pèrdua a la hora de dissenyar models, juntament amb les repercussions que comporten no pretractar les dades ja que si bé sabíem que processos com el Skull Stripping ajuden a obtenir millors resultats en la segmentació de lesions d'esclerosi múltiple, en aquest treball hem vist com al no modelar bé la xarxa neuronal i no pretractar les dades, obtenim pitjors resultats en la segmentació de lesions que els que haguéssim obtingut si haguéssim segmentat les lesions directament de les imatges originals.

Finalment, al aplicar els models en casos reals, hem comprovat la nostre hipòtesis del problema d'adaptabilitat que tenim, degut a les diferències entre les imatges que obtenim dels hospitals analitzats amb les dades que hem fet servir per trobar els models de segmentació de lesions d'esclerosi múltiple. Al buscar solució al problema, hem descobert la gran importància que té la mida de la base de dades en l'ús de les xarxes U-Net en el àmbit de la síntesis d'imatges.



*Figura 50: Resultats de la segmentació de lesions obtinguts a través de diverses combinacions en imatges dels diferents hospitals.*

En la Figura 50 podem veure com hem aconseguit segmentar lesions en tot tipus de casos, gràcies a la utilització de mètodes de pretractament diferents a la hora de generar models de segmentació.

Malgrat la falta d'un major número de dades que ens ha privat d'obtenir un bon model de síntesi, hem pogut veure com amb una mateixa base de dades d'entrenament de segmentació de lesions, si bé no podem trobar un model que funcioni per tots els hospitals alhora, si que hem pogut realitzar la segmentació de les lesions correctament

## Conclusions

en tots els casos dels hospitals, aconseguint superar el problema d'adaptabilitat, sense poder assegurar que funcioni davant d'altres dades. Però si que ens ha permès veure la importància que té el disseny del model, no només és posar-hi dades i obtenir resultats.

Arnau Cabarrocas Salvador

Autor del projecte

Girona, 6 de setembre de 2022



## 11 FUTURS ESTUDIS

A partir dels resultats obtinguts en aquest treball, es poder realitzar continuacions del estudi, que permetin profunditzar més sobre la segmentació de lesions d'esclerosi múltiple. Aquests estudis es poden basar en la millora dels models de síntesi d'imatges gràcies a l'augment de la base de dades, ja sigui gràcies a l'obtenció de noves dades, o la segmentació realitzada per experts de conjunts de dades ja existents. El camp de la segmentació d'imatges, com hem pogut veure, no té una sola solució, hi ha diversos mètodes i alternatives, que permetrien realitzar un estudi comparatiu entre aquests, per tal de poder determinar les millors estratègies de segmentació.

Com hem pogut veure en el treball, hem aconseguit solucionar el problema de l'adaptabilitat dels mètodes gràcies als diversos models obtinguts a partir de una sola base de dades. Això obre portes a la realització d'estudis de l'efecte que realitzen diversos pretractaments en la solució dels problema de segmentació, i comprovar si realment és una alternativa a la falta de dades viable o no.





## 12 REFERÈNCIES

- 1- Home - WMH Segmentation Challenge. (n.d.). WMH Segmentation Challenge. Retrieved September 6, 2022, from <https://wmh.isi.uu.nl/>
- 2- Valverde, S., Coll, L., Valencia, L., Clèrigues, A., Oliver, A., Vilanova, J. C., Ramió-Torrentà, L., Rovira, À., & Lladó, X. (2021). Assessing the Accuracy and Reproducibility of PARIETAL : A Deep Learning Brain Extraction Algorithm. *Journal of Magnetic Resonance Imaging*. <https://doi.org/10.1002/jmri.27776>
- 3- Li, Y., Li, W., Xiong, J., Xia, J., & Xie, Y. (2020). Comparison of Supervised and Unsupervised Deep Learning Methods for Medical Image Synthesis between Computed Tomography and Magnetic Resonance Images. *BioMed Research International, 2020*, 1–9. <https://doi.org/10.1155/2020/5193707>
- 4- Lladó, X., Oliver, A., Cabezas, M., Freixenet, J., Vilanova, J. C., Quiles, A., Valls, L., Ramió-Torrentà, L., & Rovira, À. (2012). Segmentation of multiple sclerosis lesions in brain MRI: A review of automated approaches. *Information Sciences, 186*(1), 164–185. <https://doi.org/10.1016/j.ins.2011.10.011>
- 5- Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L., & Erickson, B. J. (2017). Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *Journal of Digital Imaging, 30*(4), 449–459. <https://doi.org/10.1007/s10278-017-9983-4>
- 6- *What are Neural Networks?* (n.d.). [www.ibm.com](https://www.ibm.com/cloud/learn/neural-networks#toc-what-are-n-2oQ5Vepe). <https://www.ibm.com/cloud/learn/neural-networks#toc-what-are-n-2oQ5Vepe>
- 7- Malhotra, P., Gupta, S., Koundal, D., Zaguia, A., & Enbeyle, W. (2022). Deep Neural Networks for Medical Image Segmentation. *Journal of Healthcare Engineering, 2022*, 1–15. <https://doi.org/10.1155/2022/9580991>
- 8- Ronneberger, O., Fischer, P., & Brox, T. (n.d.). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. <https://arxiv.org/pdf/1505.04597v1.pdf>

- 9- He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*. ArXiv.org. <https://arxiv.org/abs/1512.03385>
- 10- *Esclerosi múltiple*: (n.d.). Fundación Esclerosi Múltiple. - <https://www.fem.es/es/esclerosis-multiple/>
- 11- Tarun Dua, Rompani, P., Organización Mundial De La Salud, & Multiple Sclerosis International Federation. (2008). *Atlas : multiple sclerosis resources in the world, 2008*. Who.
- 12- Goldenberg, M. M. (2012). Multiple sclerosis review. *P & T : A Peer-Reviewed Journal for Formulary Management*, 37(3), 175–184. [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3351877/#\\_\\_sec2title](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3351877/#__sec2title)
- 13- Berger, A. (2002). Magnetic resonance imaging. *BMJ (Clinical Research Ed.)*, 324(7328), 35. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1121941/>
- 14- McDonald, W. I., Compston, A., Edan, G., Goodkin, D., Hartung, H.-P., Lublin, F. D., McFarland, H. F., Paty, D. W., Polman, C. H., Reingold, S. C., Sandberg-Wollheim, M., Sibley, W., Thompson, A., Van Den Noort, S., Weinshenker, B. Y., & Wolinsky, J. S. (2001). Recommended diagnostic criteria for multiple sclerosis: Guidelines from the international panel on the diagnosis of multiple sclerosis. *Annals of Neurology*, 50(1), 121–127. <https://doi.org/10.1002/ana.1032>
- 15- Dobson, R., & Giovannoni, G. (2018). Multiple sclerosis - a review. *European Journal of Neurology*, 26(1), 27–40. <https://doi.org/10.1111/ene.13819>
- 16- *Esclerosi múltiple*. (n.d.). Canal Salut. Retrieved September 6, 2022, from [https://canalsalut.gencat.cat/ca/detalls/article/Esclerosi\\_multiple](https://canalsalut.gencat.cat/ca/detalls/article/Esclerosi_multiple)
- 17- Chapter 2 - Principles of Magnetic Resonance Imaging. (n.d.). Users.fmrib.ox.ac.uk. [https://users.fmrib.ox.ac.uk/~stuart/thesis/chapter\\_2/section2\\_6.html](https://users.fmrib.ox.ac.uk/~stuart/thesis/chapter_2/section2_6.html)