

**UNIVERSITÀ
DEGLI STUDI
DI PADOVA**



DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN INGEGNERIA BIOMEDICA

Tesi di Laurea

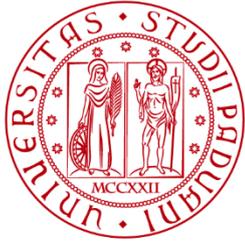
**BRAIN TUMOR SEGMENTATION (BraTS) CHALLENGE:
ASPETTI IMPLEMENTATIVI DEGLI ALGORITMI DI
SEGMENTAZIONE E DI PREDIZIONE**

Relatore: Dott. Ing. Marco Castellaro

Laureando: Lamberto Morandin

Anno Accademico 2022/2023

Data di Laurea 19/07/2023



**UNIVERSITÀ
DEGLI STUDI
DI PADOVA**



DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN INGEGNERIA BIOMEDICA

Tesi di Laurea

**BRAIN TUMOR SEGMENTATION (BraTS) CHALLENGE:
ASPETTI IMPLEMENTATIVI DEGLI ALGORITMI DI
SEGMENTAZIONE E DI PREDIZIONE**

Relatore: Dott. Ing. Marco Castellaro

Laureando: Lamberto Morandin

Anno Accademico 2022/2023

Data di Laurea 19/07/2023

INDICE

1. Introduzione	1
2. La risonanza magnetica	3
2.1. Segmentazione	6
3. Storia del BraTS.....	9
3.1. Modalità di iscrizione e svolgimento.....	9
4. I set di dati.....	11
5. Evoluzione dei task	15
6. Evoluzione degli algoritmi	17
6.1. Come si valutano gli algoritmi	17
6.2. Le reti convoluzionali.....	18
6.2.1 U-Net.....	20
6.3 BraTS 2019: Two-Stage Cascaded U-Net.....	23
6.4 BraTS 2020: nnU-Net for Brain Tumor Segmentation.....	25
6.5 BraTS 2021: Extending nn-UNet for Brain Tumor Segmentation.....	27
7 Aggiornamento della challenge.....	31
8. Risultati.....	33
9. Conclusioni.....	35
10. Bibliografia	37
11. Sitografia.....	39

1. INTRODUZIONE

Il glioma è uno dei tumori al cervello più frequenti negli adulti e coinvolge alcune cellule del sistema nervoso centrale, le cellule gliali. Il tumore può crescere infiltrando i tessuti circostanti. Per valutare gli esiti della strategia terapeutica adottata, vengono utilizzati dei protocolli di neuroimaging basati su Risonanza Magnetica per Immagini (RMI) pre e post terapia. L'accurata identificazione delle differenti parti della neoplasia risulta fondamentale. Il rilevamento e il tracciamento delle diverse sotto-regioni dei tumori solidi, se eseguito manualmente, risulta un lavoro laborioso e, soprattutto, operatore-dipendente. Data la variabile e complessa forma delle masse neoplastiche, la segmentazione delle stesse a partire da immagini multimodali rappresenta uno degli aspetti più complessi e impegnativi dell'analisi delle immagini mediche. Data l'alta ricaduta clinica, questo problema ha attirato l'attenzione durante gli ultimi 30 anni, portando alla creazione di differenti algoritmi (automatici, semiautomatici e interattivi) per la segmentazione delle masse tumorali. In particolare, data l'esigenza di segmentare accuratamente le sotto-regioni dei gliomi, nel 2012 nasce BraTS (Brain Tumor Segmentation), una challenge scientifica proposta in concomitanza della conferenza internazionale MICCAI (Medical Image Computing and Computer Assisted Interventions), con lo scopo di confrontare i vari metodi adottati per la segmentazione automatica delle immagini di risonanza magnetica per i tumori cerebrali. La challenge è stata riproposta annualmente ed arrivata, ad oggi, all'undicesima edizione. In questo elaborato riporterò l'evoluzione delle varie metodiche che, a partire dal 2019, hanno dato i risultati migliori nelle diverse edizioni della challenge.

2. LA RISONANZA MAGNETICA

La RMI è una metodica diagnostica che fornisce immagini dettagliate e ad alto contrasto del corpo umano. A differenza di altre tecniche, essa non si basa su radiazioni ionizzanti ma sull'applicazione di campi magnetici statici e tempo-varianti. Questi elementi rendono la RMI meno invasiva di altre tecniche diagnostiche per immagini, come la tomografia a emissione di positroni (PET) o la Tomografia Computerizzata ai raggi X (TAC). La RMI è in grado di fornire immagini tridimensionali dei tessuti del corpo umano e viene utilizzata per la diagnosi di molteplici condizioni patologiche, poiché permette di visualizzare diversi tipi di strutture (e.g. organi parenchimatosi, organi cavi, raccolte, cavità, articolazioni, ossa). Questo rende la RMI una tecnica versatile, fondamentale in campo medico, e in particolare in neurologia, neurochirurgia, urologia, ortopedia, gastroenterologia, cardiologia e oncologia [1].

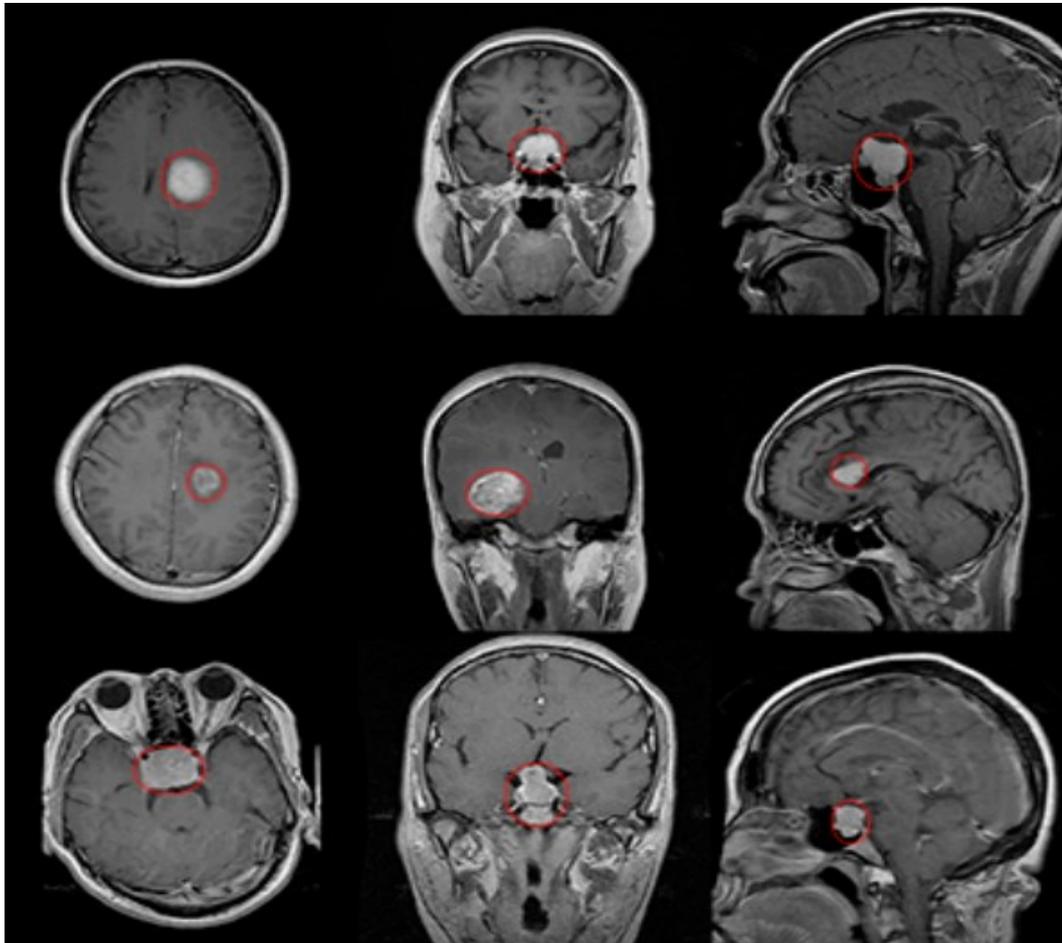


Figura 1: Immagini di risonanza magnetica di un meningioma, un glioma e dell'ipofisi, in proiezioni assiale, coronale e sagittale (1).

Il funzionamento della RMI si basa sulle caratteristiche fondamentali degli isotropi di alcuni elementi i cui nuclei esibiscono proprietà magnetiche (tra i più interessanti i nuclei con numero di massa dispari, quali idrogeno, carbonio, fluoro, sodio, fosforo ecc.), e in particolare su come questi ritornino in equilibrio dopo un'opportuna stimolazione con un segnale a radiofrequenza.

Il tessuto cerebrale abbonda di protoni di idrogeno, dato che il cervello umano è principalmente composto da acqua. In condizione di riposo, i protoni di idrogeno ruotano attorno al proprio asse e, finché sono caricati positivamente, inducono un piccolo campo magnetico, noto come momento magnetico o *spin*. Questo momento magnetico si comporta come un piccolo magnete; tuttavia, dato che ogni momento magnetico è orientato casualmente, non si può osservare un campo magnetico macroscopico, almeno fino a che il campione non viene messo all'interno di un campo magnetico esterno più grande.

Quando il cervello viene messo all'interno di un campo magnetico più forte (chiamato B_0), i momenti magnetici iniziano ad interagire con il campo, orientando ogni spin dei protoni con il campo di B_0 . Al cessare degli impulsi, i momenti magnetici tornano a orientarsi secondo l'asse del campo magnetico, emettendo un debolissimo segnale, detto segnale di risonanza. Il segnale di risonanza viene captato da ricevitori radio e convertito in impulsi digitali, a loro volta elaborati al computer, ottenendo un'immagine la cui scala dei grigi corrisponde alle diverse intensità del segnale di risonanza (2).

La risonanza magnetica strutturale fornisce una caratterizzazione anatomica statica della struttura del cervello, dando informazioni riguardo la forma, la grandezza e l'integrità dei tessuti. Esempi più comuni di immagini strutturali sono T_1 -ponderato (T_1 -W), T_2 -ponderato (T_2 -W), PD-ponderato (PD), T_1 +iniezione di contrasto e la sequenza FLAIR.

- Le immagini T_1 -w forniscono un buon contrasto fra la materia grigia (GM) e la materia bianca (WM), con le parti ad alto contenuto di acqua (come il liquido cerebrospinale [CSF] e le ossa) che appaiono scure. Questo tipo di immagini viene tipicamente usato per la segmentazione dei tessuti e per scopi di registrazione;

- Le immagini T₂-w forniscono un buon contrasto fra il CSF (chiaro) ed il tessuto cerebrale (grigio scuro), mentre il contrasto GM-WM è meno marcato che nelle immagini T₁-w. Questo tipo di immagine viene utilizzato per scopi diagnostici;
- Le immagini PD forniscono un buon contrasto fra WM e GM (anche se non al livello delle immagini T₁-w), ma il contrasto tra GM e CSF è basso. Nella pratica clinica, una fase di preparazione è spesso aggiunta per sopprimere il segnale del CSF e migliorare il contrasto dell'immagine (2).
- Le sequenze FLAIR mostrano un'elevata sensibilità a una vasta gamma di malattie. La sequenza è particolarmente utile per rilevare sottili cambiamenti alla periferia degli emisferi cerebrali, intorno alle cisterne basali, nel tronco cerebrale, nelle interfacce della sostanza bianca grigia e nella regione periventricolare (3).

Ci sono casi in cui viene utilizzato un mezzo di contrasto per evidenziare le strutture vascolari e per aiutare a caratterizzare stati infiammatori e neoplasie. Gli agenti più comunemente usati sono i derivati del gadolinio, che hanno proprietà magnetiche che influenzano i tempi di rilassamento dei protoni. La RM delle strutture intra-articolari può richiedere l'iniezione di un derivato del gadolinio diluito in un'articolazione [II]

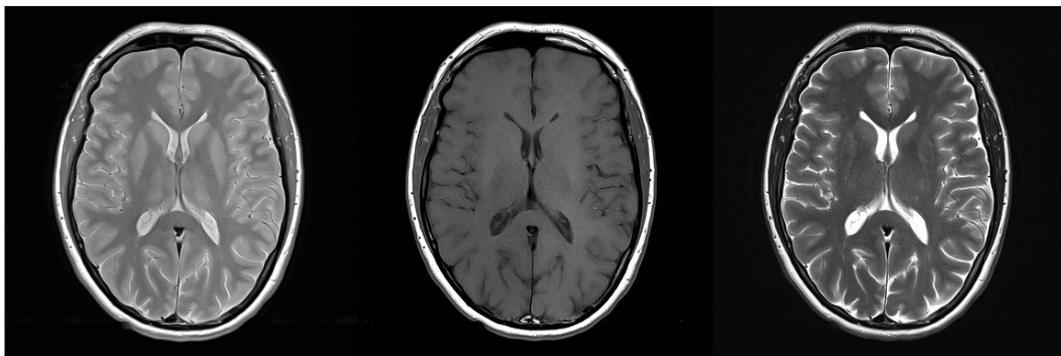


Figura 2: da sinistra a destra, PD, T1-W e T2-W [III]

2.1. Segmentazione

La segmentazione del glioma è un passaggio fondamentale per l'analisi quantitativa dei dati strutturali. Essa consiste nell'identificazione di ciascun voxel (i.e. il singolo valore di una griglia in tre dimensioni) dell'immagine e nell'assegnazione a ognuno di essi un'etichetta, la quale rappresenta uno specifico tessuto: il nucleo tumorale (compresa anche la parte necrotica), l'enhancing del tumore, l'edema circostante la neoplasia e le eventuali infiltrazioni verso i tessuti circostanti.

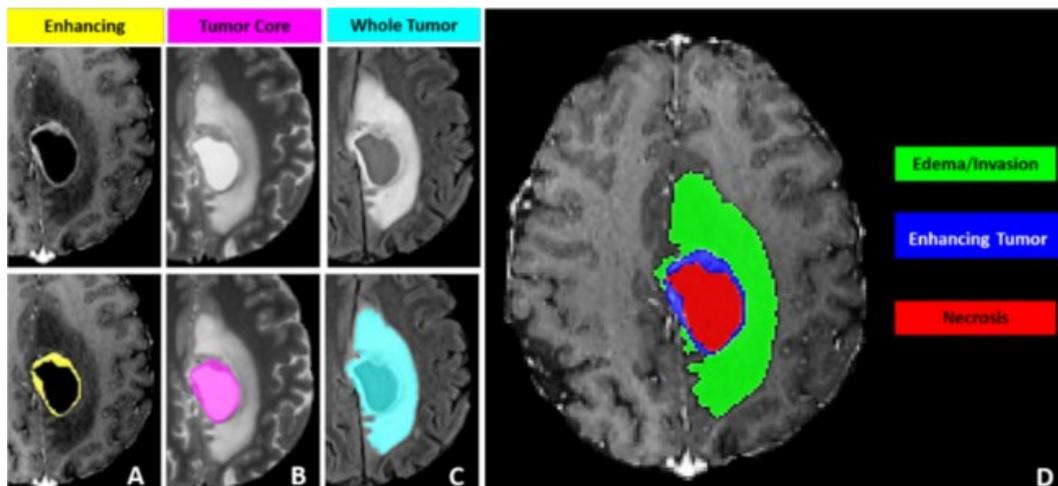


Figura 3: Subregioni del glioma considerate nel RSNA-ASNR-MICCAI BraTS 2021 challenge (4)

L'approccio più semplice nella segmentazione del cervello è il threshold method (metodo di soglia). Con questa metodica viene selezionato un livello di soglia in base all'intensità dell'istogramma dell'immagine; il livello così selezionato viene usato per assegnare ogni voxel al tipo di tessuto. Si tratta di una metodica di facile e rapido utilizzo, soggetta tuttavia ad alcune limitazioni. Una di queste è il fatto che la scelta del valore soglia dipende dalla differenza esistente fra i valori di picco nell'istogramma. Questo metodo non viene quindi utilizzato per la segmentazione di strutture complesse, ma viene tipicamente usato come punto di partenza per metodiche più sofisticate.

Un secondo approccio alla segmentazione di immagini è l'uso dei clustering algorithms (algoritmi di raggruppamento), come ad esempio il k-means, il fuzzy c-means o l'expectation maximisation. L'algoritmo k-means assegna ad ogni voxel una classe. Successivamente esso calcola, iterativamente per ogni classe, la media dell'intensità del segnale ed assegna ogni voxel alla classe con la media più vicina.

Questa procedura iterativa termina quando non vengono più applicati cambiamenti nell'assegnazione dei voxel. L'algoritmo fuzzy c-means generalizza l'algoritmo k-means, calcolando un'assegnazione soft per ogni voxel a ciascuna classe (ossia fornendo una probabilità di affiliazione per ogni classe). Infine, l'expectation maximisation segmenta l'istogramma dell'intensità dell'immagine in modo simile ad un approccio a soglia rigida; tuttavia, invece di fare affidamento a confini rigidi, questo approccio modella l'istogramma come somma di funzioni gaussiane e calcola la probabilità a posteriori di ogni voxel di appartenere ad ogni classe, definita a priori. Similmente al metodo di soglia, questi algoritmi sono sensibili alle intensità disomogenee e non includono presupposti geometrici o anatomici quando si esegue la segmentazione dell'immagine; inoltre, l'accuratezza è limitata dalla variabilità individuale. Queste limitazioni possono essere superate combinando clustering approaches con modelli più complessi.

Un altro approccio frequentemente utilizzato per la classificazione dei tessuti è quello di assegnare le etichette in base ad un modello generato da dati precedentemente raccolti. I dati del modello possono comprendere tessuti etichettati manualmente da esperti o possono essere processati usando approcci più complessi. Con questo approccio, la trasformazione tra l'immagine da segmentare ed il modello è calcolata, ed il modello della mappa delle probabilità è successivamente spostato sull'area da analizzare usando trasformazioni lineari o non lineari. Questo metodo fa affidamento sulla fase di registrazione ed ha delle buone performance solo se la struttura da segmentare nell'immagine iniziale può essere accoppiata propriamente con la stessa struttura del modello. Il metodo della registrazione del modello ottiene segmentazioni non appropriate nelle regioni con un'alta variabilità intra-soggetto o in soggetti con lesioni o patologie anatomiche (2).

3. STORIA DEL BraTS

Per misurare lo stato di avanzamento scientifico nella segmentazione automatica di tumori al cervello e poter comparare fra loro i diversi metodi, è stata organizzata nel 2012 la BraTS (*Multimodal Brain Tumor Image Segmentation Benchmark*), una challenge in unione con il MICCAI¹ (*Medical Image Computing and Computer Assisted Interventions*). Per questo scopo, era stato preparato e reso fruibile un dataset la cui segmentazione era già nota. Ognuno dei venti differenti algoritmi di segmentazione del tumore era stato ottimizzato dai rispettivi programmatori su un sottoinsieme di questo dataset e successivamente utilizzato sulle rimanenti immagini, per valutarne le performance “contro” la delineazione manuale eseguita da esperti nel campo (4).

3.1. Modalità di iscrizione e svolgimento

Il primo benchmark è stato organizzato il 1° ottobre 2012 a Nizza, in Francia, in un seminario tenuto durante la conferenza MICCAI 2012. I partecipanti dovevano registrarsi con uno dei sistemi online ed avevano la possibilità di scaricare un dataset annotato, in modo da testare gli algoritmi. Ad ogni partecipante è stato inoltre chiesto di creare un documento di quattro pagine per riepilogare il loro algoritmo e i problemi riscontrati durante la pratica. Le proposte sono state revisionate dagli organizzatori ed un gruppo finale di dodici partecipanti è stato invitato a contribuire alla challenge. Il training dataset che i partecipanti avevano a disposizione per mettere a punto i propri algoritmi consisteva in 30 immagini di risonanza magnetica multi-contrasto (10 immagini acquisite da pazienti con glioma di basso grado, 20 immagini acquisite da pazienti con glioma di alto grado) e 50 immagini simulate (25 immagini di glioma di basso grado, 25 immagini di glioma di alto grado); le 80 immagini complessive erano state precedentemente annotate con le due etichette del tumore (edema e nucleo) da un esperto. Durante la prova sul posto al workshop del MICCAI, le squadre avevano 12 ore per valutare immagini mai viste prima. Le immagini per il test consistevano in 11 gliomi di alto

¹ La MICCAI society diede inizio, nel 1998 a Boston (Massachusetts, USA), a una serie di conferenze annuali dedicate alla promozione, conservazione e facilitazione della ricerca, dell'istruzione e della pratica nel campo dell'elaborazione delle immagini mediche e degli interventi medici assistiti da computer, inclusi l'imaging biomedico e la robotica medica.

grado e 4 gliomi di basso grado da casi reali, alle quali aggiungere 10 immagini di glioma di alto grado e 5 di gliomi basso grado simulate. Le segmentazioni effettuate da ogni squadra sono state poi caricate su una piattaforma online, che ha calcolato il punteggio delle prestazioni per l'identificazione delle strutture del tumore. Dei dodici gruppi partecipanti, sei gruppi hanno inviato i propri risultati entro la scadenza prevista, un gruppo ha inviato i risultati oltre le 12 ore concesse (5).

Dato il successo ottenuto, la challenge è stata riproposta annualmente. Negli anni, oltre a cambiare la sede della conferenza MICCAI (due edizioni sono state fatte online a causa del covid), si sono aggiornati anche i database e le task richieste, portando ad ulteriori cambiamenti negli algoritmi di segmentazione inviati dai partecipanti (5).

4. I SET DI DATI

I set di dati BraTS sono costituiti da una raccolta retrospettiva di scansioni RMI multi-parametrica (mpMRI) di tumori cerebrali, acquisite da istituzioni diverse nella pratica clinica, con apparecchiature e protocolli di imaging diversi che esitano in una eterogenea qualità delle immagini.

Seguendo il paradigma della valutazione algoritmica nell'apprendimento automatico, i dati inclusi nella BraTS 2021 sono suddivisi in training data (dati di addestramento), validation data (dati di convalida) e testing data (dati che verranno poi usati per valutare gli algoritmi). Ai partecipanti alla sfida vengono fornite le etichette della segmentazione solo per i dati di addestramento. I dati di convalida vengono forniti ai partecipanti senza alcuna annotazione associata e i dati del test vengono tenuti nascosti ai partecipanti per tutto il tempo di svolgimento della challenge.

A differenza del primo anno della challenge, in cui vennero messi a disposizione un dataset di 80 immagini (30 dalla pratica clinica, 50 simulate) come training data e un dataset di 15 immagini per testare gli algoritmi, dal 2013 si decise di rimuovere le immagini simulate, mantenendo un dataset di 30 immagini reali come training data e aumentando a 25 le immagini per i test. Oltre ai due dataset messi a disposizione, dal 2017 ne venne fornito un terzo, contenente 46 immagini per la convalida, in modo da dare la possibilità ai gruppi di analizzare immagini mai viste prima. Ogni anno sono state aggiunte immagini inedite arrivando, nel 2021, a un totale di 2040 immagini, delle quali 1251 di training data, 46 di validation data e 570 di testing data. Questo rappresenta uno dei più grandi set di dati annotati multi-etichetta di scansioni mpMRI per una singola entità patologica (6). I membri delle comunità RSNA (Radiological Society of North America) e ASNR (American Society of Neuroradiology) si sono offerti volontari per perfezionare le annotazioni della sub-regione tumorale i 2.000 casi inclusi nella sfida BraTS 2021, raggiungendo una qualità elevata della valutazione dei dati.

Le scansioni mpMRI incluse nella sfida BraTS 2021 descrivono: (i) volumi nativi (T1); (ii) volumi post-contrasto T1-weighted (T1Gd [Gadolinio]); (iii) volumi T2-weighted (T2); (iv) volumi T2 Fluid Attenuated Inversion Recovery (T2-FLAIR).

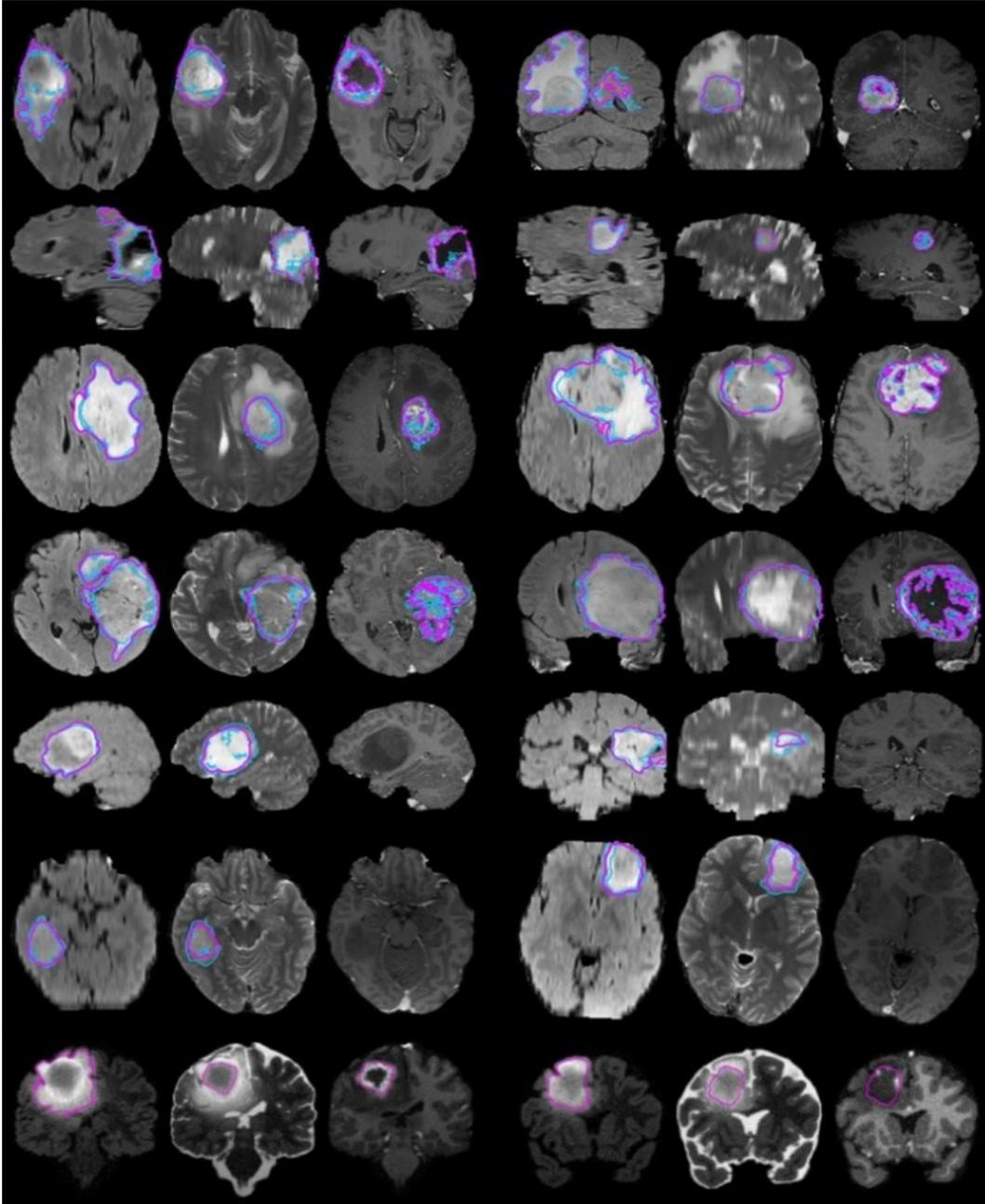


Figura 4: Esempio di immagine preso dal training data del BraTS (5).

In particolare, per la segmentazione delle regioni tumorali, tutti i volumi di imaging sono stati segmentati utilizzando un insieme dei precedenti algoritmi BraTS di alto

livello (i.e. nnU-Net², DeepScan³, DeepMedic⁴). Queste etichette fuse sono state perfezionate manualmente da neuroradiologi volontari, con diversi gradi di esperienza, seguendo un protocollo di annotazione costantemente comunicato. Le annotazioni raffinate manualmente sono state infine approvate da neuroradiologi esperti, con almeno 15 anni di esperienza di lavoro con i gliomi. Le sottoregioni tumorali annotate si basano su osservazioni visibili ai radiologi più addestrati (caratteristiche VASARI⁵) e comprendono il Gd-enhancing tumor (ET), il tessuto edematoso peritumorale /tessuto invaso (ED) e il nucleo tumorale necrotico (4). L'espansione dei set di dati su cui testare gli algoritmi di segmentazione rappresenta uno degli elementi centrali da ottenere nel corso degli anni, al fine di migliorare e validare ulteriormente i risultati.

² Approfondito nel capitolo 6.4.

³ Le architetture DeepSCAN vengono introdotte per la segmentazione dei tumori cerebrali. Ispirata dalla recente architettura Densenet e U-Net, l'architettura DeepSCAN è stata proposta per la segmentazione semantica (6).

⁴ DeepMedic è un software per la segmentazione di immagini 3D, basato su una rete neurale convoluzionale profonda 3D multiscala, del gruppo BioMedIA dell'Imperial College di Londra [IV].

⁵ Il set di funzionalità MRI VASARI (Visually AcceSAbLe Rembrandt Images) è un sistema progettato per consentire una descrizione coerente dei gliomi utilizzando un insieme di caratteristiche visive definite e un vocabolario controllato [V].

5. EVOLUZIONE DEI TASK

Come già accennato nel capitolo dedicato alla storia del BraTS, nella prima edizione della challenge l'unico task richiesto era la segmentazione del tumore nella parte "attiva" e nel "nucleo". Tuttavia, al riepilogo finale è emerso che l'utilizzo di due sole classi di distinzione della neoplasia era insufficiente, dato che l'etichetta "nucleo" conteneva sottostrutture con aspetti molto differenti per ogni modalità. Le immagini del dataset sono state quindi ri-annotate, ridefinendo la classe "nucleo" con le etichette "necrosi", "cisti", "enhancing"(5).

Nel 2014, oltre la task di segmentazione della lesione, sono state aggiunte delle task facoltative: (i) segmentazione longitudinale della lesione (venivano indicati i dataset che formavano una sequenza longitudinale di osservazione dello stesso paziente, in modo da poter usare queste informazioni per identificare il miglior algoritmo di segmentazione longitudinale, clinicamente più rilevante); (ii) classificazione diagnostica delle immagini, chiedendo ai partecipanti di distinguere il grado del glioma visualizzato, informazione che non era fornita; (iii) analisi MRI e classificazione istologica. Nel 2016 è stato inoltre aggiunto il task di quantificazione dei cambiamenti longitudinali (5).

Nel 2017 le task facoltative sono state sostituite con un nuovo quesito: una volta prodotta la segmentazione in etichette nelle immagini preoperatorie, si è chiesto di usare tali etichette per estrarre le caratteristiche delle immagini dai dati RMI disponibili, in modo da predire la sopravvivenza globale del paziente. Oltre ai parametri volumetrici, era possibile considerare l'intensità, la morfologia, l'istogramma, le caratteristiche strutturali, le informazioni spaziali e le proprietà di diffusione del glioma estratte dai modelli di crescita del glioma stesso. Sono state identificate tre classi di sopravvivenza: (i) lunga (>15 mesi); (ii) media (fra i 10 e i 15 mesi); (iii) corta (<10 mesi) [VI].

Nell'ultimo anno, come task da affiancare alla consueta richiesta di segmentazione, è stata richiesta una classificazione radiogenomica. Ai partecipanti venivano forniti dati mpMRI e dello stato di metilazione del promotore MGMT (O⁶-Methylguanine-DNA methyltransferase) associato a ciascun caso (6). MGMT è un gene che codifica per un enzima implicato nella riparazione del DNA e il suo grado di metilazione ha un ruolo importante nella chemio-resistenza agli agenti alchilanti;

quest'ultimo è infatti un problema significativo, che influenza l'efficacia dei farmaci antitumorali. La valutazione dello stato del promotore di MGMT è quindi considerato un elemento importante per il trattamento del tumore (7).

Nel file csv fornito con i dati, i casi con promotore metilato venivano etichettati come "1", quelli con promotore non metilato come "0". I ricercatori hanno proposto un metodo per predire lo stato di metilazione del promotore, con appropriate immagini ed estrazione delle caratteristiche radiologiche, analizzandole con algoritmi di machine learning. Come nel 2017, i partecipanti, oltre ai parametri volumetrici, potevano considerare anche l'intensità, la morfologia, l'istogramma, le caratteristiche strutturali, le informazioni spaziali e le proprietà di diffusione del glioma ricavate dai modelli di crescita del glioma. I partecipanti sono stati valutati per la predizione dello stato dell'MGMT del soggetto indicato nel foglio di calcolo allegato (4).

6. EVOLUZIONE DEGLI ALGORITMI

6.1. Come si valutano gli algoritmi

Per ognuna delle tre regioni si ottiene una mappa binaria con la previsione algoritmica $P \in \{0,1\}$ e la valutazione degli esperti $T \in \{0,1\}$, e calcoliamo il Dice score:

$$Dice(P, T) = \frac{|P_1 \wedge T_1|}{(|P_1| + |T_1|)/2},$$

dove \wedge è l'operatore logico AND, $|\cdot|$ è la grandezza del set (ovvero il numero di voxel che gli appartengono), e P_1 e T_1 rappresentano rispettivamente il set di voxel dove $P=1$ e $T=1$ (fig. 4). Il Dice score normalizza il numero dei veri positivi con la grandezza media delle due aree segmentate.

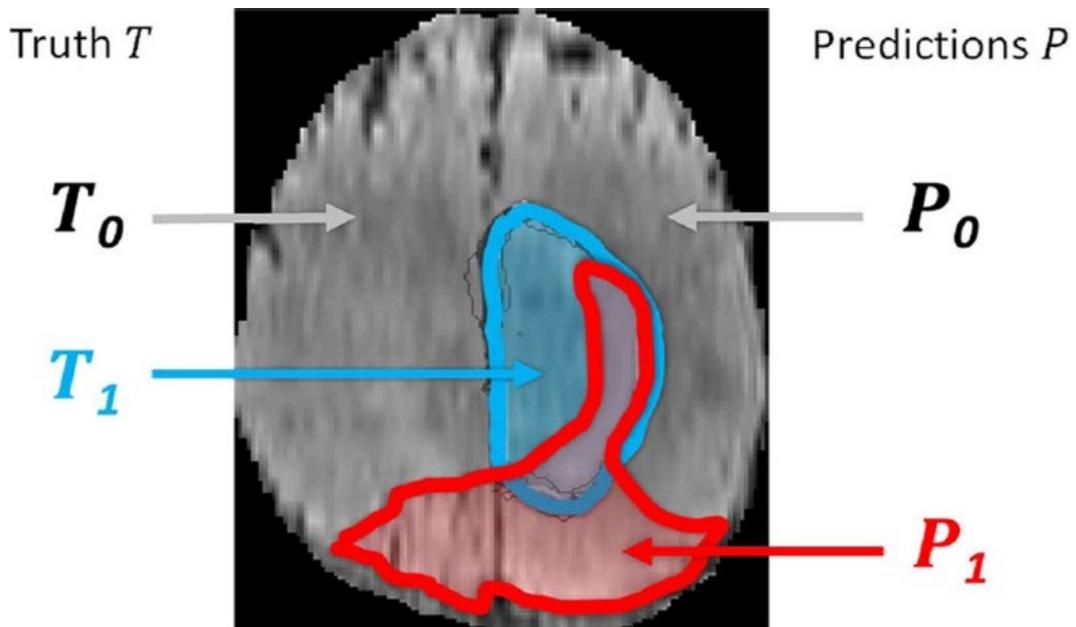


Figura 5: Regioni usate per calcolare il Dice score, sensitivity, specificity e robust Hausdorff score (5).

Viene anche calcolato sensitività (ratio di positività vera) e specificità (ratio di negatività vera):

$$Sens(P, T) = \frac{|P_1 \wedge T_1|}{|T_1|},$$

$$Spec(P, T) = \frac{|P_0 \wedge T_0|}{|T_0|},$$

dove P_0 e T_0 rappresentano rispettivamente i voxel dove $P=0$ e $T=0$.

Dice score, sensitività, e specificità misurano la grandezza dei voxel sovrapposta della regione segmentata. Una differente classe di score valuta invece la distanza fra i contorni delle aree segmentate, ovvero la distanza superficiale. Un esempio è la distanza di Hausdorff, che si ottiene calcolando per tutti i punti p della superficie ∂P_I di un volume dato P_I la distanza più piccola del minimo quadrato $d(p,t)$ dal punto t sulla superficie ∂T_I dell'altro volume dato T_I , e viceversa:

$$Haus(P, T) = \max \left\{ \sup_{p \in \partial P_I} \inf_{t \in \partial T_I} d(p, t), \sup_{t \in \partial T_I} \inf_{p \in \partial P_I} d(t, p) \right\}$$

Restituendo il massimo su tutte le distanze superficiali, tuttavia, rende la misura di Hausdorff molto suscettibile a piccole sotto regioni periferiche in entrambi P_I e T_I . Nella valutazione della challenge della regione del “tumore attivo”, per esempio, entrambi P_I e T_I possono essere formati da più aree di piccola dimensione oppure strutture non convesse con superficie ad area elevata. Nella valutazione dell'intero tumore, le previsioni con regioni false positive (che non incidono sostanzialmente la qualità totale della segmentazione potendole rimuovere in postproduzione) potrebbero avere un impatto drastico sul valore totale dell'Hausdorff score. Per queste motivazioni si usa una versione più robusta dell'Hausdorff, riportando non la massima distanza fra P_I e T_I , ma il 95% quantile di esso (5).

6.2. Le reti convoluzionali

Vi sono vari tipi di reti neurali che vengono utilizzati per casi d'uso e tipi di dati differenti. Ad esempio le reti neurali ricorrenti, comunemente utilizzate per l'elaborazione del linguaggio naturale e il riconoscimento vocale, o le reti neurali convoluzionali (ConvNets o CNN) sono più spesso utilizzate per le attività di classificazione e di computer vision [VII]. Noi approfondiremo il secondo tipo.

È una classe di reti neurali specializzata nell'elaborazione di dati che hanno una topologia a griglia, come un'immagine. Un'immagine digitale è una rappresentazione binaria di dati visivi. Contiene una serie di pixel disposti in modo simile a una griglia che contiene i valori dei pixel per indicare quanto luminoso e di che colore dovrebbe essere ogni pixel.

Il cervello umano elabora un'enorme quantità di informazioni nel momento in cui vede un'immagine. Ogni neurone lavora nel proprio campo recettivo ed è collegato

ad altri neuroni in modo da coprire l'intero campo visivo. Così come ogni neurone risponde agli stimoli solo nella regione ristretta del campo visivo chiamato campo recettivo, anche ogni "neurone" di una CNN elabora i dati solo nel suo campo recettivo. Gli strati sono disposti in modo da rilevare prima i modelli più semplici (linee, curve, ecc.) e più avanti quelli più complessi (volti, oggetti, ecc.). Utilizzando una CNN, è possibile abilitare la vista ai computer [VIII].

Una CNN ha tipicamente tre strati:

- Il livello convoluzionale è l'elemento costitutivo principale di una CNN ed è il punto in cui si verifica la maggior parte dei calcoli. Richiede pochi componenti, ovvero dati di input, un filtro e una mappa delle funzioni. Supponiamo che l'input sia un'immagine a colori, costituita da una matrice di pixel in 3D. Ciò significa che l'input avrà tre dimensioni: un'altezza, una larghezza e una profondità, che corrispondono all'RGB in un'immagine. Inoltre, è presente un rilevatore di funzioni (generalmente di dimensioni 3x3), definito anche kernel o filtro, che si sposterà attraverso i campi recettivi dell'immagine, verificando la presenza della funzione. Questo processo è noto come convoluzione. Quindi, il filtro viene applicato a un'area dell'immagine e viene calcolato un prodotto di punti tra i pixel di input e il filtro. Questo prodotto di punti viene quindi inserito in un array di output. Successivamente, il filtro si sposta di un passo, ripetendo il processo fino a quando il kernel non avrà attraversato l'intera immagine. Dopo ogni operazione di convoluzione, una CNN applica una trasformazione ReLU (Rectified Linear Unit) alla mappa delle funzioni, introducendo la non linearità nel modello. La funzione di attivazione lineare rettificata, o in breve ReLU, è una funzione lineare a tratti che emette direttamente l'input se è positivo, altrimenti emette zero [IX]. È diventata la funzione di attivazione predefinita per molti tipi di reti neurali perché un modello che la utilizza è più facile da addestrare e spesso ottiene prestazioni migliori [VII].
- I livelli di pooling, definiti anche sottocampionamento, eseguono la riduzione della dimensionalità, riducendo il numero di parametri nell'input. In modo simile al livello convoluzionale, l'operazione di pooling applica un

filtro sull'intero input, ma la differenza è che a questo filtro non è associato alcun peso. Invece, il kernel applica una funzione di aggregazione ai valori all'interno del campo ricettivo, popolando l'array di output. Sebbene il livello di pooling comporti la perdita di molte informazioni, offre una serie di vantaggi alla CNN. Contribuisce alla riduzione della complessità, al miglioramento dell'efficienza e alla limitazione del rischio di sovradattamento [VII].

- Il livello completamente connesso esegue l'attività di classificazione in base alle funzioni estratte tramite i livelli precedenti e i loro diversi filtri. Mentre i livelli convoluzionali e di pooling tendono a utilizzare le funzioni ReLu, i livelli completamente connessi (FC) solitamente fanno leva su una funzione di attivazione softmax per classificare gli input in modo appropriato, producendo una probabilità da 0 a 1 [VII].

6.2.1 U-Net

I recenti vincitori della sfida BraTS si basano esclusivamente su reti neurali convoluzionali profonde, più specificamente su architetture encoder-decoder, un modello che è stato introdotto per la prima volta da U-Net.

A partire dalla sua invenzione nel 2015, U-Net è stata l'architettura dominante per la segmentazione biomedica ed è stata altamente ottimizzata, soprattutto per una competizione annuale come BraTS.

U-Net è una rete neurale convoluzionale sviluppata per la segmentazione di immagini biomediche presso il Dipartimento di Informatica dell'Università di Friburgo. La rete è basata sulla rete completamente convoluzionale e la sua architettura è stata modificata ed estesa per lavorare con meno immagini di addestramento e per produrre segmentazioni più precise. La segmentazione di un'immagine 512×512 richiede meno di un secondo su una GPU moderna [X].

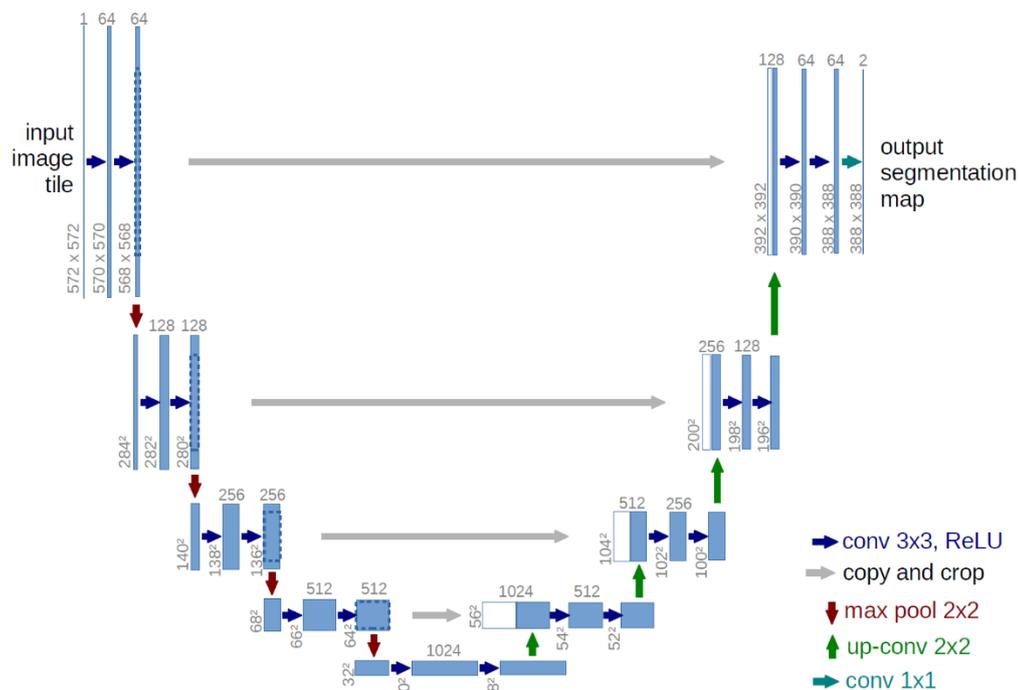


Figura 6: Architettura U-net (esempio per 32×32 pixel nella risoluzione più bassa). Ogni casella blu corrisponde a una mappa delle caratteristiche multicanale. Il numero di canali è indicato nella parte superiore della box. La dimensione x-y è fornita sul bordo inferiore sinistro del box. Le caselle bianche rappresentano le mappe delle feature copiate. Le frecce indicano le diverse operazioni (8)

L'architettura di rete è illustrata nella Figura 6, come si può vedere il nome della rete viene dalla forma a U. Consiste in un percorso di contrazione di encoder (lato sinistro) e un percorso espansivo di decoder (lato destro). Il percorso di contrazione segue l'architettura tipica di una rete convoluzionale. Consiste nell'applicazione ripetuta di due convoluzioni 3x3 (convoluzioni *unpadded*), ciascuna seguita da un'unità lineare rettificata (ReLU) e da un'operazione di pooling max 2x2 con passo 2 per il sottocampionamento. Ad ogni fase di sottocampionamento raddoppiamo il numero di canali di funzionalità. Ogni passo nel percorso espansivo consiste in un sovracampionamento della mappa delle feature seguito da una convoluzione 2x2 ("up-convolution") che dimezza il numero di canali di feature, una concatenazione con la feature map corrispondentemente ritagliata dal percorso di contrazione e due convoluzioni 3x3, ciascuna seguita da un ReLU. Il ritaglio è necessario a causa della perdita di pixel del bordo in ogni convoluzione. Al livello finale viene utilizzata una convoluzione 1x1 per mappare ogni vettore di feature a 64 componenti al numero desiderato di classi. In totale la rete ha 23 livelli convoluzionali. Per consentire una suddivisione uniforme della mappa di segmentazione dell'output (vedere la Figura 2), è importante selezionare la dimensione del riquadro di input in modo tale che tutte le operazioni di pooling max 2x2 vengano applicate a un livello con una dimensione x e y uniforme (8).

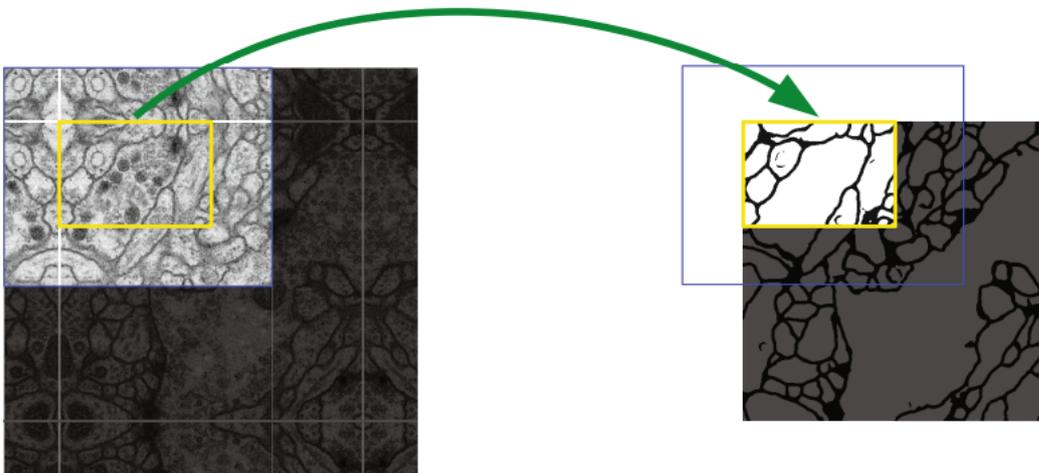


Figura 7: Strategia di sovrapposizione per la segmentazione senza soluzione di continuità di immagini arbitrarie di grandi dimensioni (qui segmentazione di strutture neuronali in stack EM). La previsione della segmentazione nell'area gialla richiede i dati dell'immagine all'interno dell'area blu come input. I dati di input mancanti vengono estrapolati mediante mirroring (8).

6.3 BraTS 2019: Two-Stage Cascaded U-Net

Nel 2019, al 5° workshop BrainLes, tenuto in congiunzione del MICCAI 2019, a Shenzhen, Cina, il vincitore della prima task di segmentazione fu il team “Questionmarks” dell’Università di tecnologia della Cina meridionale. Nel loro paper elaborano una nuova U-Net a cascata a due stadi basata sugli elaborati dei vincitori degli anni precedenti, ispirati soprattutto dalla proposta di Myronenko l’anno precedente di una U-Net asimmetrica con una variazione nel ramo dell’encoder.

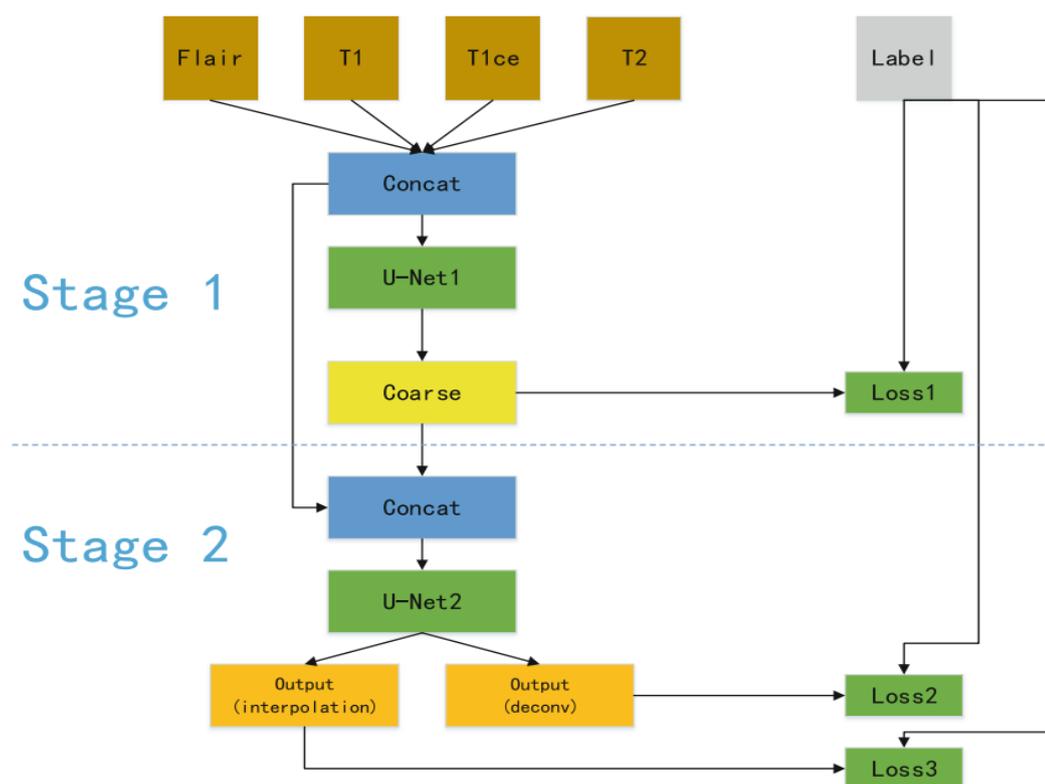


Figura 8: Panoramica della rete a cascata a due fasi [7]

Come si può vedere dalla Figura 8, nel primo stadio, una immagine multimodale di risonanza magnetica ($4 \times 128 \times 128 \times 128$) viene elaborata dal primo stage U-Net e predice una mappa di segmentazione in modo grossolano. Questa mappa di segmentazione, assieme all’immagine grezza, viene passata al secondo stadio U-Net. Questa fase può fornire una mappa di segmentazione più accurata con più parametri di rete. La rete a cascata a due stadi viene addestrata in modo end-to-end, ovvero che il training avviene su tutta la rete in contemporanea, non usando strategie di training che spezzano la rete in sottomoduli. La variabilità di un singolo

modello può essere piuttosto elevata. Per ovviare a questo passaggio vengono utilizzati un totale di cinque reti con un approccio 5-fold cross-validation per fornire una segmentazione più robusta del validation dataset di BraTS 2019.

Vengono riportati i risultati del loro approccio con il validation dataset del BraTS 2019 (che conteneva 125 casi con un glioma di grado sconosciuto e senza segmentazione). Tutti i valori riportati vengono calcolati tramite la piattaforma di valutazione online (<https://ipp.cbica.upenn.edu/>) per la valutazione del punteggio di Dice, della sensibilità, della specificità e della distanza di Hausdorff (95%). I risultati del set di convalida sono disponibili nella Tabella 1. Le prestazioni del miglior modello singolo sono leggermente migliori rispetto all'insieme della 5-fold cross-validation. L'insieme di 12 modelli si traduce in un miglioramento minore rispetto al miglior modello singolo. I risultati dei set di test sono presentati nella tabella 2. Questo algoritmo ha ottenuto il primo posto su oltre 70 squadre partecipanti.

a)

Method	DSC			HD95		
	WT	TC	ET	WT	TC	ET
Validation						
Ensemble of 5-fold	0.90797	0.85888	0.79667	4.35413	5.69195	3.12642
Best single model	0.90819	0.86321	0.80199	4.44375	5.86201	3.20551
Ensemble of 12 models	0.90941	0.86473	0.80211	4.26398	5.43931	3.14581

b)

Method	DSC			HD95		
	WT	TC	ET	WT	TC	ET
Testing						
Ensemble of 12 models	0.88796	0.83697	0.83267	4.61809	4.13071	2.65056

Tabella 1: Misurazioni medie Dice e Hausdorff del metodo di segmentazione proposto su a) set di validazione e b) sul set di test BraTS 2019. DSC - coefficiente di somiglianza al Dice, HD95 - distanza di Hausdorff (95%), WT - tumore intero, TC - nucleo tumorale, ET - enhancing tumorale (9)

I Risultati sperimentali sul set di test dimostra che il metodo proposto ha raggiunto punteggi medi nel Dice di 0,83267, 0,88796 e 0,83697, nonché distanze di Hausdorff (95%) di 2,65056, 4,61809 e 4,13071, rispettivamente per l'enhancing tumor, il tumore intero e il nucleo tumorale. Questo approccio ha vinto il 1° posto nel compito di segmentazione della sfida BraTS 2019, con oltre 70 squadre che hanno partecipato alla sfida (9).

6.4 BraTS 2020: nnU-Net for Brain Tumor Segmentation

Nel 2020, Fabian Isensee ed il suo team hanno proposto nnU-Net, un metodo di segmentazione che configura automaticamente le pipeline di segmentazione per set di dati biomedici arbitrari. nnU-Net ha stabilito nuovi risultati all'avanguardia sulla maggior parte dei 23 set di dati su cui è stato testato, sottolineando l'efficacia di questo approccio.

In primo luogo, viene applicato nnU-Net senza alcuna modifica per fornire una linea di base per le modifiche successive. nnU-Net normalizza i voxel cerebrali di ogni immagine sottraendo la loro media e dividendo per la loro deviazione standard. I voxel non cerebrali rimangono a 0. L'architettura di rete generata da nnU-Net è visualizzata in Figura 9. Segue un modello 3D U-Net che consiste in un encoder e un decodificatore che sono interconnessi da connessioni di salto (*skip connection*).

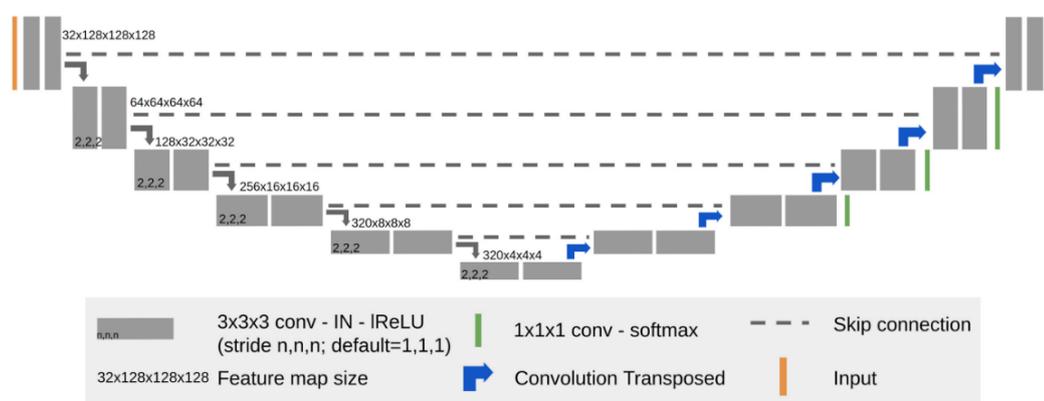


Figura 9: Architettura di rete generata da nnU-Net (10).

Il sottocampionamento viene eseguito con *strided convolutions*⁶ ed il sovracampionamento viene eseguito con la *convolution transposed*. Le uscite di segmentazione ausiliarie, utilizzate per una supervisione profonda, si diramano a tutte tranne le due risoluzioni più basse nel decodificatore. La dimensione della patch di input è selezionata per essere $128 \times 128 \times 128$ con una *batch size*⁷ di 2. Vengono eseguite in totale cinque operazioni di downsampling, con una dimensione della mappa delle caratteristiche di $4 \times 4 \times 4$ nel collo di bottiglia. Il

⁶ È un elemento fondamentale della convoluzione che viene utilizzato nelle reti neurali convoluzionali [XI].

⁷ Si può suddividere il training set in sottogruppi uniformi, chiamati batch. Il numero di esempi contenuti in ogni batch è chiamato batch size [XII].

numero iniziale di *kernel*⁸ convoluzionali è impostato a 32, che viene raddoppiato con ogni sottocampionamento fino ad un massimo di 320. Il numero di kernel nel decodificatore rispecchiano quelli del codificatore.

L'obiettivo dell'addestramento è la somma della perdita di Dice e della perdita di cross-entropia. La perdita agisce sulle tre etichette di classe edema, necrosi e tumore in aumento. nnU-Net utilizza la discesa stocastica del gradiente con un tasso di apprendimento iniziale di 0,01 e un momento di Nesterov di 0,99. L'addestramento viene eseguito per un totale di 1000 epoche, dove un'epoca è definita come 250 iterazioni. Il tasso di apprendimento decade con un programma polinomiale. Le patch di addestramento sono ritagliate da casi di addestramento selezionati in modo casuale. L'aumento dei dati viene applicato al volo durante l'addestramento.

Con un set di dati più ampio, può essere utile aumentare le dimensioni del batch (compromesso della varianza di polarizzazione). A tal fine, modifichiamo la configurazione di nnU-Net per aumentare la batch size da 2 a 5 nel tentativo di migliorare la precisione del modello.

Sulla convalida incrociata del set di allenamento, le prestazioni della baseline nnU-Net sono molto vicine alle configurazioni che sono state specificamente modificate per la sfida BraTS. Sul set di convalida, tuttavia, le modifiche specifiche di BraTS proposte forniscono prestazioni di segmentazione sostanzialmente più elevate. L'insieme finale è stato selezionato in base ai tre migliori modelli del set di convalida (come determinato con la reimplementazione della classifica BraTS). Ha ottenuto punteggi Dice di convalida di 91,24, 85,06 e 79,89 e HD95 di 3,69, 7,82 e 23,5 rispettivamente per il tumore intero, il nucleo del tumore e il tumore potenziante.

Questo approccio ha vinto il concorso BraTS 2020. Sul test set, il modello ha ottenuto punteggi Dice di 88,95, 85,06 e 82,03 e valori HD95 di 8,498, 17,337 e 17,805 rispettivamente per il tumore intero, il nucleo tumorale e il tumore potenziante (10).

⁸ Nell'apprendimento automatico, le macchine kernel sono una classe di algoritmi per l'analisi dei pattern, il cui membro più noto è la macchina a vettori di supporto (SVM). Questi metodi implicano l'utilizzo di classificatori lineari per risolvere problemi non lineari [XIII]

6.5 BraTS 2021: Extending nn-UNet for Brain Tumor Segmentation

Nel 2021 il vincitore fu il team KAIST-MRI-Lab, composto da Huan Minh Luu e Sung-Hong Park. Portando come elaborato un'estensione del framework nnU-Net proposto dal vincitore dello scorso anno aggiungendo diversi componenti. Grazie alla facilità di adattamento di nnU-Net a un nuovo set di dati, nonché ai codici e ai modelli completamente open source, nnU-Net funge da eccellente base per ulteriori sperimentazioni. Le modifiche che sono state fatte per la competizione di quest'anno sono l'utilizzo della normalizzazione di gruppo anziché la normalizzazione batch, l'utilizzo di un encoder asimmetrico per U-Net e l'utilizzo dell'*axial attention* nel decodificatore. La prima modifica che è stata apportata è stata quella dell'aumento delle dimensioni della rete in modo asimmetrico raddoppiando il numero di filtri nell'encoder mantenendo gli stessi filtri nel decoder. Anche il numero massimo di filtri è stato aumentato a 512. Poiché la quantità di dati di addestramento è quadruplicata rispetto all'anno precedente, l'aumento della capacità della rete aiuterà a modellare la più ampia varietà di dati. La struttura della rete modificata è mostrata in Figura 10.

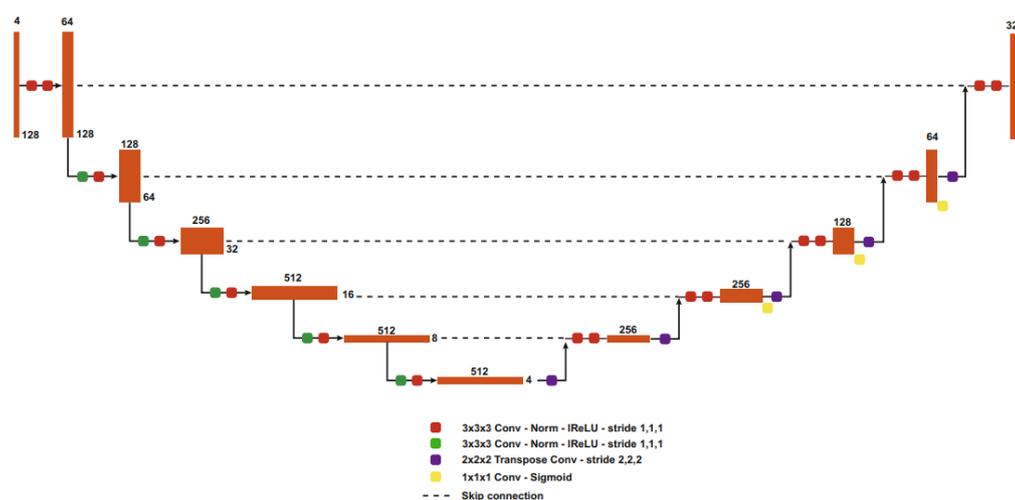


Figura 10: Struttura dell'U-Net più grande con scala asimmetrica sull'encoder. Il numero in alto e in basso indica rispettivamente il numero di canali e la dimensione dei vettori delle feature (11).

La seconda modifica è stata quella di sostituire tutta la normalizzazione batch con la normalizzazione di gruppo. Le reti di convoluzione 3D richiedono memorie GPU significative anche con training a precisione mista, limitando le dimensioni del batch utilizzato durante l'addestramento. La normalizzazione di gruppo ha

dimostrato di funzionare meglio della normalizzazione dei batch per il regime di bassa dimensione dei lotti ed è stata adottata anche dai precedenti vincitori del concorso. Se non diversamente specificato, il numero di gruppi è stato fissato a 32. Il modello nn-UNet di base ha 31,2 milioni di parametri e richiede 8,2 GB di VRAM⁹ con dimensioni del lotto di 2 e addestramento di precisione misto. Il modello più grande proposto ha 87,4 milioni di parametri e richiede 14,4 GB di VRAM con le stesse impostazioni (11).

È stata seguita la metodologia di formazione di nnU-Net per tutte le reti. Ogni rete è stata addestrata con una 5-fold cross-validation. L'aumento dei dati è stato applicato al volo durante l'allenamento per migliorare la generalizzazione. L'aumento dei dati consisteva in rotazione e ridimensionamento casuali, deformazione elastica, aumento della luminosità additiva e ridimensionamento gamma. L'obiettivo per l'ottimizzazione è la somma della perdita binaria cross-entropia e della Dice loss¹⁰, calcolata all'output finale a piena risoluzione e agli output ausiliari con risoluzione inferiore. La versione batch della Dice loss è stata utilizzata al posto della Dice loss campione, calcolando la perdita trattando l'intero lotto come un campione invece di calcolare la media dei dadi di ciascun campione nel minibatch. I dadi batch aiutano a stabilizzare l'addestramento riducendo gli errori dai campioni con pochi campioni annotati

Model	Dice				HD95			
	ET	TC	WT	Average	ET	TC	WT	Average
BL	83.73	87.45	92.63	87.94	22.44	10.56	3.55	12.18
BL + L	83.28	86.53	92.51	87.44	22.50	12.75	3.71	12.99
BL + GN	84.09	87.85	92.77	88.24	22.41	9.20	3.42	11.68
BL + L + GN	84.51	87.81	92.75	88.36	20.73	7.623	3.47	10.61

Tabella 2: Dadi e distanza di Hausdorff del 95% delle reti calcolati dalla piattaforma di valutazione online. ET: migliorare il tumore, TC: nucleo tumorale, WT: tumore intero, HD95: 95% distanza di Hausdorff (11).

⁹ VRAM (RAM video) si riferisce a qualsiasi tipo di memoria ad accesso casuale (RAM) specificamente utilizzata per memorizzare i dati di immagine per il display di un computer [XIV].

¹⁰ La Dice loss deriva dal coefficiente di Sørensen-Dice, che è una statistica sviluppata nel 1940 per valutare la somiglianza tra due campioni. È stato portato alla comunità di computer vision da Milletari et al. nel 2016 per la segmentazione di immagini mediche 3D [XV].

La tabella 2 mostra i risultati con il Dice e HD95 calcolati dagli organizzatori della competizione e visualizzati nella classifica pubblica. Per l'inferenza è stato utilizzato un approccio a finestra scorrevole. La segmentazione è stata prevista per 8 sottovolumi, ponderati con un gaussiano 3D e riassemblati per l'intera segmentazione del volume. Per l'aumento del tempo di prova (TTA), è stato applicato il mirroring dei 3 assi per ciascun sottovolume. Sono state eseguite 64 inferenze ($8 \text{ sottovolumi} \times 8 \text{ mirroring}$) per ciascun volume nel set di convalida e test, con una latenza media di 20 secondi per volume. Tutte le etichette di segmentazione previste sono state quindi raggruppate da 5-fold facendo la media delle probabilità sigmoide¹¹, ad eccezione della configurazione BL + L. I risultati sono stati poi post-elaborati convertendo la classe tumorale potenziante in tumore necrotico se i voxel ET erano inferiori a 200.

I modelli proposti hanno vinto il primo posto nella classifica finale sui dati dei test invisibili, raggiungendo un punteggio di Dice di 88,35%, 88,78%, 93,19% per il tumore potenziante, il nucleo del tumore e l'intero tumore, rispettivamente.

¹¹ L'attivazione sigmoidea è la funzione di attivazione dell'output standard nella classificazione binaria e nella segmentazione con reti neurali. Tuttavia, esiste una varietà di altre potenziali funzioni di attivazione dell'output, che possono portare a risultati migliori nella segmentazione delle immagini mediche (12).

7 AGGIORNAMENTO DELLA CHALLENGE

La Brain Tumor Segmentation (BraTS) Continuous Challenge cerca di identificare gli attuali algoritmi di segmentazione all'avanguardia per i pazienti con glioma diffuso al cervello e le loro sottoregioni. I dati di formazione ed i set di immagini inedite sono resi disponibili tutto l'anno per la creazione e la valutazione di algoritmi di segmentazione. Le metriche di valutazione per le previsioni sui dati di immagini inedite vengono restituite immediatamente ai partecipanti con una classifica continuamente aggiornata. Questa sfida continua culminerà in una valutazione annuale dello stato del campo in cui chiederemo ai partecipanti di presentare versioni containerizzate dei loro migliori modelli che verranno eseguite su un set di dati di test e i risultati presentati alla conferenza annuale MICCAI.

I dati di formazione BraTS e di convalida disponibili per il download e lo sviluppo metodologico da parte dei gruppi partecipanti descrivono un totale di 5.880 scansioni MRI da 1.470 pazienti con glioma diffuso cerebrale e sono identici ai dati usati per la RSNA-ASNR-MICCAI BraTS 2021 Challenge; tuttavia, il set di dati dei test utilizzato nella sfida di quest'anno è stato aggiornato con molte più scansioni mpMRI clinicamente acquisite di routine. Le annotazioni di verità di base delle sottoregioni tumorali sono create e approvate da neuroradiologi esperti per ogni soggetto incluso nei set di dati di addestramento, convalida e test per valutare quantitativamente le segmentazioni tumorali previste [XVI].

8. RISULTATI

I risultati ottenuti dai vincitori degli ultimi tre anni sono esemplificati nella Tabella 3. Dalla tabella si evince che, mentre il valore del Dice sta migliorando progressivamente, arrivando persino ad un valore di 0,9319 sulla segmentazione del tumore intero e migliorando il valore medio del 4,85%, la distanza di Hausdorff (95%) è leggermente peggiorata, passando mediamente da 3,79979 a 9,71333 dimostrando quindi un regresso del 1,58%.



Tabella 3: Valori dei vincitori del BraTS, rispettivamente per l'anno 2019,2020 e 2021

9. CONCLUSIONI

Questo studio si è posto l'obiettivo di analizzare la challenge BraTS, dando priorità ai compiti di segmentazione, approfondendo e confrontando i vincitori degli ultimi tre anni.

È risultato evidente che, la rete U-Net, sviluppata nel 2015, ha rappresentato per gli anni successivi le fondamenta della maggior parte degli algoritmi in gara. Sono stati infatti proposti negli anni seguenti centinaia di algoritmi fondati su U-Net, fino al 2020, anno in cui F. Isensee e il suo team ne hanno sviluppato una variante, la nnU-Net, usata come algoritmo di base e come struttura per lo sviluppo del metodo. nnU-Net ha dimostrato la sua generalizzabilità fornendo un'elevata precisione di segmentazione in modo immediato.

Diversi partecipanti hanno riferito che la strategia di valutazione potrebbe non essere ideale, in quanto il metodo di classificazione utilizzato per determinare gli algoritmi vincenti funziona su un principio diverso da quello utilizzato dagli stessi team per valutare i propri progetti. In effetti, confrontando le classifiche dei modelli dei singoli gruppi sul set di addestramento e sul set di convalida, si osservato disparità tra "aggrega quindi classifica" e "classifica quindi aggrega". Nonostante questo, le tendenze generali erano simili tra i metodi di classificazione, indicando che una selezione del modello basata sull'aggregazione metrica stia ancora andando nella giusta direzione. La tabella 3 ci permette di notare il miglioramento degli algoritmi di segmentazione, ed è fondamentale per la comunità scientifica continuare in questa direzione per rendere sempre più efficiente e valido l'utilizzo di questi algoritmi in campo medico.

Questo elaborato si è concentrato esclusivamente sull'implementazione del deep learning nella segmentazione dei tumori. Tuttavia, le sue applicazioni in radiologia sono molteplici e tendono ad alleggerire il lavoro del medico radiologo, eliminando i compiti semplici che richiedono molto tempo e dando la possibilità allo specialista di concentrarsi solo sulle mansioni più complesse, nelle quali l'intelligenza artificiale non può ancora intervenire.

10. BIBLIOGRAFIA

1. Badža MM, Barjaktarović MČ. Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network. *Appl Sci.* gennaio 2020;10(6):1999.
2. Peter S. Bloomfield, Sabrina Brigadoi, Gaia Rizzo, Mattia Veronese. BASIC NEUROIMAGING A GUIDE TO THE METHODS AND THEIR APPLICATIONS. Second edition. - London (UK): Createspace Independent Publishing Platform;
3. De Coene B, Hajnal JV, Gatehouse P, Longmore DB, White SJ, Oatridge A, et al. MR of the brain using fluid-attenuated inversion recovery (FLAIR) pulse sequences. *AJNR Am J Neuroradiol.* 1992;13(6):1555–64.
4. Baid U, Ghodasara S, Mohan S, Bilello M, Calabrese E, Colak E, et al. The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification [Internet]. arXiv; 2021 [citato 11 luglio 2023]. Disponibile su: <http://arxiv.org/abs/2107.02314>
5. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans Med Imaging.* ottobre 2015;34(10):1993–2024.
6. Zeineldin RA, Karar ME, Burgert O, Mathis-Ullrich F. Multimodal CNN Networks for Brain Tumor Segmentation in MRI: A BraTS 2022 Challenge Solution [Internet]. arXiv; 2022 [citato 11 luglio 2023]. Disponibile su: <http://arxiv.org/abs/2212.09310>
7. Yu W, Zhang L, Wei Q, Shao A. O6-Methylguanine-DNA Methyltransferase (MGMT): Challenges and New Opportunities in Glioma Chemotherapy. *Front Oncol.* 2019;9:1547.

8. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, curatori. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Cham: Springer International Publishing; 2015. p. 234–41. (Lecture Notes in Computer Science).
9. Jiang Z, Ding C, Liu M, Tao D. Two-Stage Cascaded U-Net: 1st Place Solution to BraTS Challenge 2019 Segmentation Task. In 2020. p. 231–41.
10. nnU-Net for Brain Tumor Segmentation - Google Search [Internet]. [citato 12 luglio 2023]. Disponibile su: https://www.google.com/search?q=nnU-Net+for+Brain+Tumor+Segmentation&rlz=1C1VDKB_itIT1017IT1017&oq=nnU-Net+for+Brain+Tumor+Segmentation&gs_lcrp=EgZjaHJvbWUqBggAEEUYOzIGCAAQRRg7MgYIARBFGEAyBggCEEUYPdIBCDEwODBqMGo3qAIAAsAIA&sourceid=chrome&ie=UTF-8
11. Luu HM, Park SH. Extending nn-UNet for brain tumor segmentation [Internet]. arXiv; 2021 [citato 12 luglio 2023]. Disponibile su: <http://arxiv.org/abs/2112.04653>
12. Nieradzik L, Scheuermann G, Saur D, Gillmann C. Effect of the output activation function on the probabilities and errors in medical image segmentation [Internet]. arXiv; 2021 [citato 12 luglio 2023]. Disponibile su: <http://arxiv.org/abs/2109.00903>

11. SITOGRAFIA

- I. “Risonanza Magnetica.” *W*www.airc.it, www.airc.it/cancro/affronta-la-malattia/guida-agli-esami/risonanza-magnetica. Accessed 4 Feb. 2023.

- II. “Imaging a Risonanza Magnetica - Argomenti Speciali.” *Manuali MSD Edizione Professionisti*, www.msmanuals.com/it/professionale/argomenti-speciali/principi-di-imaging-radiologico/imaging-a-risonanza-magnetica. Accessed 10 July 2023.

- III. “Biomedical Imaging.” *Fraunhofer Institute for Integrated Circuits IIS*, www.iis.fraunhofer.de/en/ff/zfp/tech/mr-imaging-in-biomedicine.html. Accessed 26 June 2023.

- IV. “DeepMedic.” *Deepmedic.org*, deepmedic.org/. Accessed 10 July 2023.

- V. Gaillard, Frank. “VASARI MRI Feature Set | Radiology Reference Article | Radiopaedia.org.” *Radiopaedia*, radiopaedia.org/articles/vasari-mri-feature-set. Accessed 10 July 2023.

- VI. “MICCAI BraTS 2017: Tasks | Section for Biomedical Image Analysis (SBIA) | Perelman School of Medicine at the University of Pennsylvania.” *W*www.med.upenn.edu, www.med.upenn.edu/sbia/brats2017/tasks.html. Accessed 8 July 2023.

VII. “Cosa Sono Le Reti Neurali Convoluzionali? | IBM.” *Www.ibm.com*, www.ibm.com/it-it/topics/convolutional-neural-networks.

VIII. Mishra, Mayank. “Convolutional Neural Networks, Explained.” *Medium*, 2 Sept. 2020, towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939.

IX. Jason Brownlee. “A Gentle Introduction to the Rectified Linear Unit (ReLU) for Deep Learning Neural Networks.” *Machine Learning Mastery*, 20 Apr. 2019, machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/.

X. “U-Net.” *Wikipedia*, 29 July 2020, en.wikipedia.org/wiki/U-Net.

XI. datahacker.rs. “#005 CNN Strided Convolution.” *Master Data Science*, 2 Nov. 2018, datahacker.rs/what-is-stride-cnn/#:~:text=A%20strided%20convolution%20is%20another. Accessed 10 July 2023.

XII. Casadei, Cristiano. “IA: Un Po’ Di Nozioni Prima Della Pratica.” *Maggiolidevelopers*, 12 Dec. 2018, www.developersmaggioli.it/blog/ia-un-po-di-nozioni-prima-della-pratica/. Accessed 12 July 2023.

XIII. Wikipedia Contributors. “Kernel Method.” *Wikipedia*, Wikimedia Foundation, 29 Jan. 2019, en.wikipedia.org/wiki/Kernel_method.

XIV. “What Is VRAM?” *SearchStorage*,

www.techtarget.com/searchstorage/definition/video-RAM.

XV. Du, Shuchen. “Understanding Dice Loss for Crisp Boundary Detection.” *AI*

Salon, 30 July 2022, [medium.com/ai-salon/understanding-dice-loss-for-crisp-](https://medium.com/ai-salon/understanding-dice-loss-for-crisp-boundary-detection-bb30c2e5f62b)

[boundary-detection-bb30c2e5f62b](https://medium.com/ai-salon/understanding-dice-loss-for-crisp-boundary-detection-bb30c2e5f62b). Accessed 30 July 2023.

XVI. <https://www.synapse.org/#!/Synapse:syn51156910/wiki/622351>