



**UNIVERSITÀ
DEGLI STUDI
DI PADOVA**

UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO di MEDICINA - DIMED

**CORSO DI LAUREA TRIENNALE IN “TECNICHE DI RADIOLOGIA MEDICA,
PER IMMAGINI E RADIOTERAPIA”**

Sede di Treviso

Presidente: Prof.ssa Raffaella Motta

Tesi di Laurea:

**Segmentazione automatica delle immagini di risonanza magnetica mediante Intelligenza
Artificiale: una revisione della letteratura con focus tecnico-percettivo**

Relatore: Eleonora Stefani

Correlatore: Giorgia Mel

Laureando: Nicolò Missaglia

Anno Accademico 2024 - 2025

ABSTRACT

Introduzione e scopo

L'avvento dei nuovi metodi di segmentazione automatica sta portando ad importanti evoluzioni nel campo dell'imaging diagnostico di Risonanza Magnetica (RM). L'aumento dei dati delle immagini in Risonanza Magnetica ed i tempi di segmentazione elevati e soggetti a variabilità inter e intra operatore rendono l'esame più complesso. L'obiettivo di questa revisione è andare ad analizzare lo stato dell'arte e i principi di funzionamento delle metodologie più diffuse nel campo della segmentazione automatica in immagini di Risonanza Magnetica, valutando il loro impatto nel workflow tecnico, evidenziando lo sviluppo che questi nuovi metodi introducono nel ruolo del tecnico di radiologia.

Materiali e metodi

È stata svolta una revisione della letteratura narrativa tematica, andando a selezionare articoli sulla segmentazione automatica relativi al suo funzionamento e ai cambiamenti apportati nel workflow di Risonanza Magnetica. Sono state analizzate quattro aree tematiche per la ricerca: metodologie e architetture di segmentazione, qualità dell'immagine e requisiti per la segmentazione con IA, integrazione degli algoritmi IA nel workflow RM, validazione percettiva dell'imaging da parte del TSRM e analisi di caso studio sperimentale con focus sul comportamento degli algoritmi di IA su immagini di RM in presenza di artefatti. Sono stati utilizzati articoli del database di Pubmed, Elsevier, Wiley, LiberKey e IEEE, utilizzando come motore di ricerca Pubmed e Consensus e prendendo anche articoli non open access (visibili con account istituzionale Unipd).

Sono stati analizzati i parametri di acquisizione dell'imaging di RM in uso presso l'AULSS 2 Marca trevigiana per ottenere un corretto funzionamento del software di segmentazione automatica con Deep Learning Icobrain.

Risultati

La letteratura evidenzia che gli algoritmi di Deep Learning, in particolare le architetture con varianti di modelli U-Net, migliorano l'efficienza dei processi di segmentazione e permettono la riduzione dei tempi richiesti rispetto ad una segmentazione manuale. Numerosi studi riportano prestazioni elevate nella delimitazione di strutture patologiche con alti livelli di accuratezza, raggiungendo valori di Dice Similarity Coefficient molto alti, di circa 0.96. Tuttavia in presenza di immagini in input di bassa qualità o artefatte, i valori di Dice Similarity Coefficient calano del 20%, andando a ridurre l'affidabilità della segmentazione automatica.

Conclusioni

Il TSRM assume un ruolo centrale nella supervisione del workflow radiologico specie nel processo di acquisizione delle immagini radiologiche e nella loro validazione percettiva, fondamentale per evitare dati in input di bassa qualità e garantire una segmentazione affidabile.

Parole chiave: segmentazione automatica, Deep Learning, tecnico di radiologia, Risonanza Magnetica, validazione percettiva, Icobrain, workflow

ABSTRACT INGLESE

Background and objective

The advent of new automatic segmentation methods is leading to significant advances in the field of Magnetic Resonance Imaging (MRI). The increase in MRI image data and the long segmentation times, subject to inter- and intra-operator variability, make the examination more complex. The objective of this review is to analyze the performance of the most widely used automatic segmentation methodologies in MRI images, assessing their impact on the technical workflow, and highlighting the developments these new methods introduce in the role of the radiographer.

Materials and methods

A thematic narrative literature review was conducted, selecting articles on automatic segmentation related to its operation and the changes implemented in the MRI workflow. Four research topics were analyzed: segmentation methodologies and architectures, image quality and requirements for AI segmentation, integration of AI algorithms into the MRI workflow, perceptual validation of imaging by the MRI technician, and experimental case study analysis focusing on the behavior of AI algorithms on MRI images in the presence of artifacts. Articles from the PubMed, Elsevier, Wiley, LiberKey, and IEEE databases were used, using PubMed and Consensus as search engines. Non-open access articles (visible with a Unipd institutional account) were also included.

The MRI imaging acquisition parameters used at the AULSS 2 Marca trevigiana were analyzed to ensure proper functioning of the Icobrain Deep Learning automatic segmentation software.

Results

The literature highlights that Deep Learning algorithms, particularly architectures with variants of U-Net models, improve the efficiency of segmentation processes and reduce the time required compared to manual segmentation. Numerous studies report high performance in the delineation of pathological structures with high levels of accuracy, achieving very high Dice Similarity Coefficient values, around 0.96. However, in the presence of low-quality or artifactual input images, Dice Similarity Coefficient values drop by 20%, reducing the reliability of automatic segmentation.

Conclusion

The radiographer plays a central role in supervising the radiological workflow, especially in the radiological image acquisition process and their perceptual validation, which is essential for avoiding low-quality input data and ensuring reliable segmentation.

Keywords: automatic segmentation, Deep Learning, radiology technician, magnetic resonance imaging, perceptual validation, Icobrain, workflow

INDICE

INTRODUZIONE	1
SCOPO TESI	3
MATERIALI E METODI	5
Metodi Area Zero	6
Metodi Area Uno	7
Metodi Area Due	8
Metodi Area Tre	9
RISULTATI E DISCUSSIONE	11
Area Zero - Metodologie e architetture di segmentazione	11
Area Uno – Qualità dell’immagine e requisiti per la segmentazione con IA	27
Area Due – Integrazione degli algoritmi IA nel workflow RM	37
Area Tre – Validazione percettiva delle immagini di RM da parte del TSRM e analisi di caso studio sperimentale	48
CONCLUSIONI	55
LIMITI DELLO STUDIO E SVILUPPI FUTURI	57
BIBLIOGRAFIA E SITOGRAFIA	59

INTRODUZIONE

L'imaging in Risonanza Magnetica (RM) rappresenta una delle metodiche più avanzate e versatili della diagnostica per immagini, grazie alla sua natura non invasiva, all'assenza di radiazioni ionizzanti e all'elevato contrasto intrinseco dei tessuti molli. Oltre alla rappresentazione morfologica delle strutture anatomiche, esso consente di ottenere informazioni funzionali e quantitative attraverso sequenze dedicate, come diffusione e perfusione, permettendo una caratterizzazione più approfondita dei tessuti e dei processi patologici.

Negli ultimi anni, l'evoluzione dell'imaging RM si è progressivamente orientata verso un approccio sempre quantitativo. L'analisi non si limita più alla semplice identificazione visiva della lesione, ma integra parametri numerici, mappe parametriche e misure volumetriche, fondamentali per il monitoraggio terapeutico, la stratificazione prognostica e lo sviluppo di modelli di medicina personalizzata. Questo cambiamento ha determinato un incremento significativo della quantità e della complessità dei dati prodotti, rendendo necessario l'impiego di strumenti avanzati di elaborazione e analisi.

In tale contesto si inserisce la rapida diffusione dell'Intelligenza Artificiale (IA), in particolare degli approcci basati sul Deep Learning. Le reti neurali profonde hanno dimostrato elevate prestazioni in numerose applicazioni, tra cui ricostruzione accelerata delle immagini, riduzione del rumore, classificazione di pattern anatomici e, soprattutto, segmentazione automatica. Quest'ultima, intesa come il processo di suddivisione dell'immagine in regioni corrispondenti a specifiche strutture anatomiche o volumi di interesse, rappresenta un passaggio cruciale per l'analisi quantitativa e per applicazioni successive quali radiomica, pianificazione terapeutica e follow-up.

Tradizionalmente, la segmentazione in RM è eseguita manualmente o con strumenti semi-automatici da operatori esperti. Tuttavia, tale approccio è oneroso in termini di tempo, dipendente dall'esperienza individuale e soggetto a variabilità inter- e intra-osservatore, come evidenziato in letteratura. L'aumento del volume di esami prodotti quotidianamente e la crescente richiesta di analisi quantitative rendono sempre meno sostenibile una gestione esclusivamente manuale.

Le architetture di Deep Learning, in particolare i modelli di tipo U-Net e le loro varianti tridimensionali, rappresentano attualmente lo standard di riferimento per la segmentazione automatica in RM. Numerosi studi riportano performance elevate, spesso valutate tramite metriche quantitative come il Dice Similarity Coefficient¹(DSC), che misura la sovrapposizione tra

¹ Dice Similarity Coefficient(DSC)/DSC: è un indice statistico che misura quanto una segmentazione sia simile ad un'altra di riferimento. Il valore varia da 0 a 1, più alto è il valore più la segmentazione è simile a quella di riferimento fino ad arrivare al valore di 1 in cui le segmentazioni risultano identiche. (Moore C, Bell D, Dice similarity coefficient)

segmentazione automatica e riferimento manuale, e la Hausdorff Distance (HD), che quantifica la massima distanza tra i contorni segmentati automaticamente alle segmentazioni di riferimento. In diversi ambiti applicativi, tali modelli hanno raggiunto livelli di accuratezza comparabili a quelli di operatori esperti, con una drastica riduzione dei tempi di elaborazione.

Nonostante questi risultati promettenti, l'efficacia dei sistemi di segmentazione automatica rimane fortemente condizionata dalla qualità delle immagini di input al sistema di elaborazione. Artefatti da movimento, rumore, disomogeneità del campo magnetico, ridotto contrasto o errori di ricostruzione possono influenzare significativamente le prestazioni degli algoritmi. I modelli di Deep Learning, infatti, apprendono pattern statistici dai dati di addestramento e possono mostrare ridotta robustezza in presenza di variabilità non rappresentata nel dataset. Ne deriva che la qualità tecnica dell'acquisizione e l'ottimizzazione dei protocolli RM costituiscono elementi determinanti per l'affidabilità delle pipeline automatiche.

In questo scenario, il ruolo del Tecnico Sanitario di Radiologia Medica (TSRM) assume una rilevanza strategica. La produzione di immagini tecnicamente adeguate, la gestione degli artefatti, la corretta scelta delle sequenze e dei parametri di acquisizione e l'attenzione ai processi di ricostruzione e pre-processing rappresentano condizioni imprescindibili affinché i sistemi di IA possano operare in modo efficace e sicuro. L'integrazione dell'IA nel workflow non elimina la responsabilità tecnica, ma la ridefinisce, spostando l'attenzione verso la qualità del dato come prerequisito per l'automazione.

Un ulteriore elemento critico riguarda la validazione dei risultati. Le metriche quantitative comunemente utilizzate per valutare le prestazioni algoritmiche non sempre riflettono pienamente la qualità percettiva e la plausibilità anatomica delle segmentazioni. In letteratura è stata evidenziata una possibile discrepanza tra performance numeriche elevate e accettabilità visiva dell'output. Segmentazioni con valori di Dice elevati possono presentare contorni irregolari, eccessivamente semplificati o non coerenti con l'anatomia reale. Per questo motivo, accanto alla valutazione quantitativa, risulta fondamentale una validazione qualitativa e percettiva da parte dell'operatore esperto, finalizzata a garantire coerenza morfologica e affidabilità clinica.

La segmentazione automatica in RM non rappresenta quindi esclusivamente uno strumento di supporto alla diagnosi, ma un elemento di trasformazione del workflow tecnico. Essa consente di standardizzare i processi, ridurre la variabilità operatore-dipendente e ottimizzare i tempi di post-processing, ma richiede al contempo nuove competenze, maggiore consapevolezza tecnologica e capacità critica nella supervisione dei sistemi automatizzati.

SCOPO TESI

Partendo dall'ipotesi di fondo è che l'efficacia dei sistemi di IA in RM non dipenda esclusivamente dalle prestazioni algoritmiche, ma in modo sostanziale dalla qualità tecnica delle immagini prodotte e dalla capacità dell'operatore di integrare valutazioni quantitative e percettive in un'ottica di responsabilità professionale e di sicurezza clinica, il presente lavoro di tesi vuole effettuare una revisione della letteratura sulla segmentazione automatica in RM tramite Deep Learning, con un focus specifico sul ruolo tecnico-percettivo del tecnico di radiologia.

In particolare, la tesi intende:

1. analizzare le architetture e le metodologie più diffuse nella segmentazione automatica in RM;
2. valutare l'impatto di tali sistemi sull'efficienza del workflow tecnico, sui tempi di elaborazione delle immagini e sulla standardizzazione dei dati, con particolare riferimento ai sistemi in uso presso l'azienda ULSS 2 Marca trevigiana;
3. evidenziare il ruolo emergente del TSRM nella supervisione percettiva dei dati di input, nella validazione e nel controllo della qualità delle immagini;
4. fornire indicazioni su prospettive future e possibili sviluppi per l'integrazione dei sistemi di IA nel workflow tecnico.

MATERIALI E METODI

Dopo una prima fase di esplorazione, è stata condotta una revisione tematica della letteratura, strutturata in aree tematiche, che riflettono le principali dimensioni attraverso cui la segmentazione automatica in RM impatta il workflow tecnico-professionale del tecnico di radiologia.

La fase preliminare di esplorazione della letteratura è stata supportata dall'utilizzo di strumenti di ricerca basati su Intelligenza Artificiale quali Consensus², al fine di identificare i principali temi emergenti e orientare la definizione delle aree tematiche della revisione. Successivamente, la ricerca sistematica è stata condotta sui principali database scientifici (fase di scoping).

La letteratura è stata successivamente analizzata e organizzata secondo quattro principali aree tematiche, rilevanti dal punto di vista del workflow tecnico e del ruolo del tecnico di radiologia:

- area zero - metodologie e architetture di segmentazione;
- area 1 – qualità dell'immagine e requisiti per la segmentazione con IA;
- area 2 – integrazione degli algoritmi IA nel workflow RM;
- area 3 – percezione e caso studio sperimentale.

È stata condotta una revisione narrativa della letteratura scientifica utilizzando i seguenti database biomedici, PubMed e Consensus.

La ricerca è stata effettuata nel periodo dal 20 dicembre 2025 al 9 marzo 2026.

Per l'inquadramento dello scopo e lo sviluppo dell'introduzione è stata effettuata una ricerca su PubMed con le seguenti parole chiave:

“Artificial Intelligence”, “machine learning”, “Deep Learning”, “Magnetic Resonance Imaging”, “scan planning”, “slice positioning”, “automatic planning”, “protocol optimization”, “radiographer”, “radiologic technologist”, “workflow”

Composte in questa stringa su PubMed:

((Artificial Intelligence[MeSH] OR machine learning OR Deep Learning)

AND

(Magnetic Resonance Imaging[MeSH])

AND

(scan planning OR slice positioning OR automatic planning OR protocol optimization)

AND

(radiographer OR radiologic technologist OR workflow)

² Consensus: è un motore di ricerca accademico basato su Intelligenza Artificiale che sintetizza risultati provenienti da letteratura scientifica peer-reviewed, prevalentemente indicizzata su PubMed. L'efficacia della ricerca dipende dalla qualità della formulazione del prompt. come descritto in questo articolo: (Faix, A. et al., 2025)

Il risultato della ricerca su PubMed con filtro negli ultimi 5 anni ha dato 52 risultati, i quali sono stati selezionati includendo articoli che parlassero maggiormente di segmentazione relativa alla Risonanza Magnetica e meno di Tomografia Computerizzata, che parlassero anche di workflow e sono state scartate ricerche in ambito veterinario, che non facessero riferimento alla segmentazione, che non includessero il trasferimento di immagini segmentate da RM a TC.

Gli articoli selezionati nelle varie aree sono stati presi non solo open access, ma anche da tutte le librerie e database online che permettevano di visualizzare le revisioni e gli studi tramite l'accesso con un account dell'Università di Padova e nello specifico sono stati utilizzati i database Elsevier, Wiley, LiberKey e IEEE.

Metodi Area Zero

Oltre ai database biomedici tradizionali come Pubmed è stato utilizzato anche Consensus, a cui è stato inserito il seguente prompt³ con l'impostazione iniziale di ricerca in versione "pro" e in modalità "medical":

"Fornisci una panoramica dello stato dell'arte sulle tecniche di segmentazione delle immagini di Risonanza Magnetica, con particolare attenzione agli approcci basati sul Deep Learning.

- Utilizzo delle reti neurali convoluzionali (CNN) per la segmentazione di immagini RM.
- Architettura U-Net e principali varianti (ad esempio U-Net 2D, U-Net 3D, Attention U-Net, Residual U-Net).
- Confronto tra strategie di segmentazione 2D, 2.5D e 3D nelle immagini di Risonanza Magnetica.
- Vantaggi, limiti e principali contesti applicativi di ciascun approccio Focalizza la risposta su studi scientifici peer-reviewed e articoli di revisione relativi alla Risonanza Magnetica (ad esempio imaging cerebrale, muscoloscheletrico o oncologico)"

La scrittura di questo prompt è stata guidata da Chat GPT⁴.

È stato impostato un filtro per la qualità degli articoli Q1-Q3 e negli ultimi 5 anni. La ricerca, nella sezione Risultati, ha dato tre risultati principali per ogni argomento indicato nel Prompt, derivanti completamente dall'archivio PubMed.

³ Prompt design: Frase immessa all'interno di Consensus che influenza la pertinenza, specificità e riproducibilità dei risultati ottenuti. La struttura sintattica, il livello di specificità, la lingua utilizzata (es. "studi peer-reviewed", "ultimi 5 anni", "review articles") possono modificare sensibilmente i risultati ottenuti. Ciò è ampiamente discusso nella letteratura, l'utilizzo dei modelli linguistici nella ricerca scientifica. (Van Dis et al., 2023)

⁴ Chat GPT: utilizzato per la scrittura di prompt e seguendo le indicazioni per l'impegno dell'università Unipd, presenti nel documento "IndicazioniImpiegoAI_Unipd_ITA_rev20251128.pdf" presente nel sito unipd (https://www.unipd.it/sites/unipd.it/files/2025/IndicazioniImpiegoAI_Unipd_ITA_rev20251128.pdf, 2025)

Successivamente sono state fatte delle ricerche specifiche per gli argomenti U-Net e BraTS utilizzando l'archivio Pubmed ed effettuando una ricerca che andasse a selezionare le revisioni degli ultimi 5 anni.

Le stringhe utilizzate all'interno del database di Pubmed per trovare gli articoli specializzati di approfondimento sugli argomenti U-Net e BraTS sono rispettivamente:

U-Net: (U-Net AND brain MRI AND segmentation)

BraTS: (BraTS AND brain tumor segmentation AND Deep Learning)

Sono quindi state selezionate le revisioni in cui si parla di U-Net in Risonanza Magnetica e delle sue varianti mostrando anche le loro prestazioni.

Metodi Area Uno

L'obiettivo dell'Area 1 è quello di fare una revisione della letteratura riguardo la qualità delle immagini e i requisiti per la segmentazione automatica, andando ad analizzare in particolare artefatti e parametri di acquisizione e cosa essi comportano nell'immagine.

Inoltre sono state esposte le caratteristiche delle immagini RM che ne permettono il confronto.

In quest'area sono stati utilizzati i database Pubmed, il manuale "Elementi di Risonanza Magnetica: Dal protone alle sequenze per le principali applicazioni diagnostiche" (Coriasco, M., Rampado, O., & Bradac, G. B., 2014) ed il manuale "The essential physics of medical imaging" (Bushberg J. T. et al., 2012, 3rd ed.) per la ricerca delle informazioni riguardo gli artefatti ed i parametri di acquisizione.

E' stato inoltre utilizzato Consensus nel quale è stato inserito il seguente prompt con l'impostazione iniziale di ricerca in versione "pro" e in modalità "medical":

"Fai una ricerca sulla qualità ideale o migliore attualmente delle immagini diagnostiche in Risonanza Magnetica e cosa cambia dai requisiti delle immagini radiologiche di Risonanza Magnetica per un utilizzo nella segmentazione automatica svolta con intelligenze artificiali."

Nella ricerca sono stati aggiunti filtri di qualità degli articoli Q1-Q3, e la ricerca è stata concentrata per articoli negli ultimi 5 anni.

Per la ricerca all'interno di Pubmed invece è stata inserita la seguente stringa per la revisione dei tipi di artefatti visibili in un'immagine di Risonanza Magnetica:

(MRI artifact*[Title/Abstract])

AND

(radiologist)

In questa ricerca sono state prese solo le revisioni e non sono stati messi filtri temporali perché questa parte dello studio vuole mostrare i diversi tipi artefatti, indipendentemente dall'introduzione della segmentazione automatica e metodi di Deep Learning.

Lo stringa ha quindi ha prodotto 22 risultati dei quali è stato selezionato lo studio che non andasse a parlare degli artefatti nella pratica clinica e non si concentrasse nella loro correzione o riconoscimento da parte di software ed IA, andando anche ad escludere studi che parlassero di nuove possibili cause di artefatto.

Metodi Area Due

Per l'Area 2 è stata condotta una revisione della letteratura focalizzata sull'integrazione degli algoritmi di IA implementati nel workflow della Risonanza Magnetica, con focus nella pianificazione, acquisizione, organizzazione e qualità dell'immagine.

Come per le altre, i database utilizzati sono stati PubMed con l'aiuto di Consensus.

Il prompt utilizzato all'interno di Consensus è "What is the current clinical workflow in magnetic resonance imaging and how is it optimized in radiology practice?" le cui impostazioni iniziali, di ricerca sono state la versione "pro" e la modalità "medical".

In questo caso è stato utilizzato un prompt in lingua inglese, la lingua predominante nell'ambito scientifico internazionale, per migliorare l'efficacia e l'accuratezza dei risultati dal momento che l'utilizzo di un approccio critico è comunque importante nella selezione degli articoli. (Van Dis et al. 2023)

Sono inoltre stati utilizzati filtri di qualità dell'articolo "Q1" e studi degli ultimi 5 anni e dai risultati ho selezionato l'articolo di (Lin D.J. et al., 2024, Designing Clinical MRI for Enhanced Workflow and Value) in quanto l'unico focalizzato su organizzazione e descrizione strutturata delle fasi del workflow in RM in generale senza riferimenti al campo oncologico. Per la parte inerente all'introduzione dell'IA all'interno del workflow radiologico invece è stato utilizzato il database Pubmed in cui è stata inserita la seguente stringa:

(artificial intelligence OR machine learning)

AND

(radiology)

AND

(workflow)

AND

(medical imaging OR magnetic resonance OR MRI)

con un filtro temporale negli ultimi 5 anni. Oltre alla stringa sono state utilizzate altre combinazioni di parole chiave: “artificial intelligence”, “machine learning”, “radiology”, “workflow” e “medical imaging”, al fine di individuare studi relativi all’integrazione dell’intelligenza artificiale nel workflow radiologico.

È stato inoltre utilizzato il sito ufficiale Icobrain che riporta gli articoli validi inerenti al funzionamento ed ai cambiamenti che ha portato Icobrain e non solo, anche in generale sistemi di Deep Learning per la segmentazione e l’introduzione all’interno del workflow radiologico di questi nuovi software.

In particolare sono stati presi articoli riguardanti principalmente l’introduzione di Icobrain o del Deep Learning all’interno del workflow radiologico e i suoi vantaggi.

I parametri della tabella n.4 sono stati raccolti durante un’esperienza di tirocinio nel servizio di Neuroradiologia dell’Ospedale Ca’Foncello di Treviso dell’Azienda ULSS2 Marca Trevigiana.

È stata utilizzata anche la consultazione del giornale “European Radiology” sul sito <https://link.springer.com/journal/330> per poi cercare e valutare, nella sezione articles, quale articolo fosse più inerente all’integrazione di IA nel workflow radiologico, selezionando l’articolo “Leadership in radiology in the era of technological advancements and artificial intelligence”(Wichtmann et al. 2024).

Metodi Area Tre

Quest’Area ha come obiettivo quello di andare a mostrare quali sono gli effetti, quindi come si comportano i diversi software di segmentazione automatica, con la presenza di artefatti all’interno delle immagini RM.

Sono quindi stati ripresi articoli utilizzati nelle altre aree per introdurre e delineare un quadro più completo, andando inoltre a sottolineare la difficoltà di segmentazione automatica in presenza di artefatti e quindi la conseguente importanza di una validazione percettiva tradizionale da parte di una figura specializzata, quale il tecnico di radiologia.

La ricerca è quindi stata fatta basandosi su uno studio “ROXSI: Robust Cross-Sequence Semantic Interaction for Brain Tumor Segmentation on Multi-Sequence MR Images”, per l’impossibilità di andare a fare uno studio autonomo per le ragioni illustrate nel capitolo “LIMITI DELLO STUDIO E SVILUPPI FUTURI”.

Si è scelto di andare a selezionare un singolo studio ritenuto metodologicamente valido che andasse a vedere il comportamento di vari metodi per la segmentazione automatica in presenza di artefatti e

ciò ha permesso di non incontrare ambiguità ed introdurre ulteriori variabili derivanti dal confronto tra studi con metodologie diverse.

Lo studio è stato trovato tramite Pubmed con la seguente stringa:

("brain tumor"[Title/Abstract] OR "brain tumour"[Title/Abstract])

AND

(segmentation[Title/Abstract])

AND

(MRI[Title/Abstract])

NOT

(review)

AND

(artifacts)

Sono inoltre stati utilizzati i filtri temporali che comprendessero gli ultimi 5 anni così da trovare informazioni con l'utilizzo dei più recenti metodi di segmentazione automatica e per avere una visione migliore del progresso in questo campo.

È stato quindi selezionato lo studio in base alla data di pubblicazione più recente e che facesse riferimento ad un confronto dei diversi tipi di artefatto.

Per la stesura dell'elaborato di tesi è stato fatto uso di IA generativa in forma strumentale (riformulazione nella parte di Introduzione e Discussione e risultati), cognitiva (suggerimenti di struttura) e produttiva (generazione di prompt per ricerca su strumenti di IA generativa).

RISULTATI E DISCUSSIONE

Negli ultimi anni, lo sviluppo di metodi di intelligenza artificiale per la segmentazione automatica di immagini di Risonanza Magnetica ha rappresentato un importante avanzamento nell'analisi quantitativa di dati volumetrici e cambiamenti nel lavoro del tecnico di radiologia.

In questa revisione è quindi stata analizzata la letteratura riguardo la segmentazione in Risonanza Magnetica con l'utilizzo di metodiche IA rispetto alle metodiche manuali con poi una parte di focus tecnico-percettivo per descrivere lo stato dell'arte della segmentazione in Risonanza Magnetica.

Area Zero - Metodologie e architetture di segmentazione

La segmentazione delle immagini medicali identifica delimita e classifica i voxel o i pixel di una specifica regione di interesse (ROI), che può includere organi, strutture anatomiche o lesioni patologiche. Questo passo è fondamentale nell'elaborazione delle immagini perché fornisce la base per misurazioni, valutazioni e analisi.

Nel contesto oncologico e radiologico, la segmentazione trasforma l'immagine in un insieme di dati strutturati, consentendo la quantificazione del volume, della forma, dell'eterogeneità e delle relazioni spaziali delle lesioni. Questi parametri non possono essere pienamente valutati con la sola osservazione qualitativa. (Veiga-Canuto et al., 2023)

Segmentazione come strumento diagnostico

Dal punto di vista clinico e diagnostico, la segmentazione si collega direttamente all'interpretazione dell'immagine da parte del radiologo e al supporto decisionale. L'obiettivo non è solo la precisione geometrica della delimitazione, ma anche la rappresentazione accurata dell'estensione della patologia, del coinvolgimento delle strutture adiacenti e delle variazioni temporali della malattia. (Tom Boeken et al., 2023)

La segmentazione diagnostica viene utilizzata principalmente nella valutazione della risposta terapeutica, nella stadiazione oncologica e nel follow-up. In questi contesti, una delimitazione coerente e riproducibile della lesione consente confronti affidabili con l'imaging esistente. Tuttavia, l'approccio dipende fortemente dall'esperienza dell'operatore e dalla variabilità intra- e inter-osservatore, soprattutto quando i margini sono poco definiti o le strutture anatomiche sono complesse. Questa variabilità può portare a interpretazioni discordanti tra diversi clinici o in diverse fasi dello stesso operatore, influenzando il percorso gestionale del paziente, dalla scelta del trattamento alla valutazione della risposta clinica. In ambito di ricerca, tale variabilità potrebbe compromettere l'affidabilità dei risultati, influenzando negativamente la coerenza e la comparabilità

tra studi diversi, incidendo in maniera negativa sulla riproducibilità delle misurazioni e sulla robustezza delle analisi quantitative successive. (Veiga-Canuto et al. , 2022)

Segmentazione manuale

La segmentazione manuale è tuttora considerata il gold standard: viene eseguita da esperti clinici che integrano conoscenze anatomiche, fisiopatologiche e contestuali. È fondamentale in diagnostica, dove l'esperienza del medico radiologo consente di adattare la segmentazione ai casi complessi o atipici. (Veiga-Canuto et al. 2023)

Numerosi studi hanno dimostrato che la segmentazione manuale richiede molto tempo. È soggetta a notevole variabilità tra osservatori diversi e, talvolta, anche per lo stesso osservatore in momenti distinti. Questa variabilità può compromettere la riproducibilità delle misurazioni e la robustezza delle analisi quantitative successive. (Jennifer Monereo-Sánchez et al., 2021) (Lerch et al., 2025) (Veiga-Canuto et al., 2022)

La segmentazione manuale è un processo molto dispendioso in termini di tempo, soprattutto quando si tratta di immagini volumetriche, in cui bisogna segmentare una per una le immagini 2D che compongono il volume.

Come indicato dagli studi di Till D. Lerch, il tempo di segmentazione per un volume 3D è in media di 3 ore ed è il passaggio più dispendioso in termini di tempo nel workflow delle immagini in Risonanza Magnetica. Questo costituisce quindi un elemento limitante, in quanto la segmentazione è un prerequisito per l'analisi quantitativa e la pianificazione chirurgica. (Lerch T. D. et al., 2025)

Il secondo problema deriva dalla variabilità intra- e interosservatore, analizzata nel lavoro di Monereo-Sánchez, in cui si dimostra che la delimitazione delle regioni di interesse dipende in modo importante dall'esperienza del radiologo, dalla complessità anatomica della struttura, dalla qualità dell'immagine e dalla ripetibilità della segmentazione. Questa variabilità, soprattutto nell'ambito oncologico, può indurre a un errore sistematico e quindi a ridurre l'affidabilità delle misure che, a loro volta, sono fondamentali nella pianificazione del trattamento.

Le immagini RM migliorano, ma aumenta anche la complessità dei dati e, di conseguenza, il rischio di perdita di informazioni durante la segmentazione manuale.

(Jennifer Monereo-Sánchez et al., 2021)

Segmentazione automatica

La segmentazione automatica classica, prima del Deep Learning, si basava su approcci deterministici e statici fondati sulla continuità spaziale e su modelli matematici che definivano le

regole di segmentazione. I principali modelli sono il thresholding, il region growing e i modelli statici basati sulla forma.

Il thresholding presuppone che le strutture di interesse abbiano densità diverse. I voxel o i pixel vengono suddivisi in base alla densità e assegnati a una classe; certe classi vengono poi escluse secondo soglie prefissate.

Tuttavia, questo metodo ignora gli artefatti e le variazioni di contrasto, soprattutto in RM, dove le densità non sono standardizzate. Di conseguenza, questo metodo non funziona bene con confini sfumati.

Il region growing seleziona un voxel iniziale che espande la selezione in base a criteri di intensità, in modo automatico o manuale. Questo metodo è soggetto a rumore, a scarsa generalizzabilità e dipende molto dall'impostazione iniziale dei criteri.

I modelli statici usano lo studio geometrico probabilistico di immagini segmentate manualmente, ricavando regole matematiche di probabilità da applicare ad altre immagini.

Tuttavia, questo modello più robusto manca di flessibilità nei confronti di patologie di forma atipica e presenta una fase di progettazione costosa in termini di tempo e denaro, poiché si basa su segmentazioni manuali.

(Wang et al., 2022)

La segmentazione automatica moderna si basa su algoritmi di intelligenza artificiale e di Deep Learning. Utilizza in particolare reti neurali convoluzionali, che apprendono dai dati come identificare e delimitare le strutture di interesse, contribuendo a ridurre l'influenza soggettiva dell'operatore umano. (Aigerim Dautkulova et al., 2025)

Valutazione degli algoritmi e benchmark nella segmentazione

La letteratura scientifica evidenzia come il confronto tra algoritmi di segmentazione automatica sia metodologicamente significativo solo se effettuato su dataset condivisi, annotazioni standardizzate e metriche di valutazione comuni. In assenza di tali condizioni, le differenze di performance osservate tra modelli possono dipendere non tanto dalle caratteristiche degli algoritmi quanto dalle specificità dei dati utilizzati per l'addestramento e la validazione.

Per rispondere a questa criticità, sono stati introdotti benchmark di riferimento, con l'obiettivo di standardizzare il problema clinico e consentire una comparazione oggettiva tra approcci diversi. In questo contesto, il benchmark BRATS (Brain Tumor Segmentation Challenge) rappresenta uno dei principali punti di riferimento per la segmentazione automatica dei tumori cerebrali su immagini di Risonanza Magnetica.

Il dataset BRATS include acquisizioni multimodali (sequenze T1, T1 con mezzo di contrasto, T2 e FLAIR) e segmentazioni manuali effettuate da esperti, consentendo la delimitazione di diverse componenti tumorali, quali tumore intero, core tumorale e regione con enhancement. Oltre ai dati, BRATS definisce anche metriche standard di valutazione, tra cui il Dice Similarity Coefficient e la distanza di Hausdorff, che permettono di confrontare quantitativamente le prestazioni dei modelli. L'introduzione di benchmark condivisi consente quindi di attribuire le differenze di performance prevalentemente agli algoritmi e non alle variazioni intrinseche dei dati clinici, rendendo possibile una reale validazione scientifica delle architetture di intelligenza artificiale applicate alla segmentazione. (Menze B. H., et al., 2015)

Deep Learning e modelli di segmentazione

L'analisi automatica delle immagini mediche è diventata centrale nella diagnostica e nel supporto alle decisioni cliniche. Questo cambiamento è dovuto sia ai progressi nelle tecnologie di acquisizione, che ora producono dati ad alta risoluzione e multimodali, sia allo sviluppo di metodi computazionali avanzati per estrarre informazioni utili. In questo scenario, il Deep Learning si è affermato come uno degli approcci più efficaci e promettenti per elaborare e interpretare immagini mediche complesse.

Il Deep Learning è una branca del machine learning, che fa parte dell'intelligenza artificiale, e si basa su modelli chiamati reti neurali artificiali profonde. Questi modelli hanno una struttura a più livelli che permette di imparare rappresentazioni dei dati, passando da descrizioni semplici a concetti più astratti e significativi. (Renard et al., 2020)

A differenza dei metodi tradizionali di machine learning, in cui l'estrazione delle caratteristiche si basa sull'esperienza dell'operatore e sulla progettazione manuale, con il Deep Learning le caratteristiche utili vengono apprese automaticamente dai dati precedentemente acquisiti. Questo è molto utile nelle immagini mediche, che spesso presentano grande variabilità, confini poco chiari e dati complessi da valutare con un occhio critico capace di cogliere la totalità delle informazioni. (Wang et al., 2022)

Dal punto di vista teorico, una rete neurale profonda è una funzione matematica complessa formata da una serie di trasformazioni non lineari. Ogni livello applica una trasformazione ai dati in ingresso, di solito una combinazione lineare seguita da una funzione di attivazione non lineare. Ripetendo questo schema su più livelli, il modello costruisce rappresentazioni sempre più dettagliate. Il risultato finale dipende dall'insieme di tutte queste trasformazioni.

Un aspetto chiave del Deep Learning è il processo di addestramento, in cui i parametri del modello, come pesi e bias, vengono ottimizzati per ridurre la differenza tra le previsioni e i dati di riferimento (ground truth). L'ottimizzazione avviene con algoritmi che usano il gradiente e il feedback negativo. Questo consente al modello di adattarsi ai dati di addestramento e di migliorare le sue previsioni. (Renard et al., 2020).

Il Deep Learning è strettamente legato alla capacità di apprendere grazie a una visione completa dell'immagine. Questa caratteristica lo rende particolarmente adatto a descrivere fenomeni complessi e di grandi dimensioni, come quelli nelle immagini di Risonanza Magnetica, in cui le informazioni diagnostiche sono spesso distribuite su ampie aree e in più modalità di acquisizione.

Nel campo della segmentazione delle immagini mediche, il Deep Learning ha dimostrato di superare molte delle principali limitazioni dei metodi tradizionali. Tecniche come le soglie di intensità, il region growing o i modelli statistici classici spesso non funzionano bene in presenza di rumore, variazioni di contrasto o strutture patologiche irregolari. Al contrario, i modelli di Deep Learning combinano informazioni locali e globali, sfruttano il contesto spaziale e imparano schemi complessi direttamente dai dati, senza dover fare ipotesi semplificate sulla distribuzione delle intensità o sulla forma delle strutture.

Il successo del Deep Learning non dipende solo dall'aumento della profondità delle reti, ma anche da altri fattori come la disponibilità di dataset annotati e la maggiore potenza di calcolo grazie alle schede grafiche e lo sviluppo di architetture e strategie di addestramento più stabili ed efficienti. Tuttavia, nel campo medico, la scarsità di dati annotati resta una sfida importante, spingendo la ricerca verso soluzioni in grado di ottenere buoni risultati anche con pochi dati. (Wang et al., 2022)

CNN

Le reti neurali convoluzionali (Convolutional Neural Networks) sono il riferimento per l'analisi automatica delle immagini mediche. Sono anche fondamentali per la segmentazione voxel-wise in immagini di Risonanza Magnetica. Il sistema CNN è ampiamente utilizzato nella segmentazione dei tumori cerebrali. Ha inoltre partecipato alla BraTS (Brain Tumor Segmentation), cioè la competizione scientifica internazionale per l'affinamento di algoritmi di segmentazione automatica dei tumori cerebrali in immagini di Risonanza Magnetica.

La CNN è composta da strati convoluzionali che si dividono in strati di sovra- o sottocampionamento, o di sovracampionamento. Questa organizzazione consente alla rete di avere una gerarchia dei dati delle immagini che va da informazioni piccole e locali a informazioni più semanticamente rilevanti, andando sempre più in profondità negli strati della rete neurale.

L'operazione principale è quindi la convoluzione, cioè un insieme di filtri o kernel che vengono posizionati sull'intera immagine, generando così una mappa di attivazione che evidenzia specifici pattern spaziali.

Uno dei problemi delle reti neurali progettate per immagini medicali è il pooling. Questo processo consiste nel ridurre la qualità delle immagini, procedendo sempre più in profondità negli strati. In questo modo si ottiene un'immagine in output a minore risoluzione spaziale. Le CNN per ambito medico-radiologico compensano il pooling mediante un sovracampionamento dell'immagine. In questo modo la risoluzione spaziale diventa comparabile a quella dell'immagine in input.

Ciò che ha reso però migliori le CNN rispetto ai modelli di machine learning è la condivisione dei pesi, cioè la condivisione di filtri locali, che possono quindi essere riutilizzati, riducendo di molto i parametri necessari e evitando di dover creare un filtro nuovo per ogni immagine, che sarebbe un dispendio computazionale elevato.

I vantaggi che hanno portato le CNN rispetto ai metodi tradizionali sono una maggiore robustezza inter-soggetto e inter-scanner, una riduzione della dipendenza dall'operatore, l'eliminazione della necessità di segmentazioni manuali ripetitive e soggettive e dell'accuratezza, come dimostrato dal progressivo miglioramento del dice score riportato nelle edizioni recenti di BraTS, nelle quali, negli ultimi 10 anni, i modelli vincitori erano tutti basati sul modello CNN.

Le architetture basate su reti neurali convoluzionali costituiscono i fondamenti dei moderni modelli di segmentazione delle immagini, che attualmente mantengono l'apprendimento gerarchico e le mappe di segmentazione. A partire da tale modello è stata introdotta la U-Net, una struttura dedicata alle immagini principalmente in Risonanza Magnetica.

Limiti delle CNN nella segmentazione medica

Le reti neurali convoluzionali (CNN) rappresentano oggi uno dei metodi più utilizzati per la segmentazione automatica di immagini mediche, grazie alla loro capacità di estrarre caratteristiche locali e apprendere schemi complessi dai dati (Litjens et al., 2017). Nonostante le elevate prestazioni, queste architetture presentano però alcuni limiti strutturali che ne condizionano l'applicabilità clinica.

Un primo limite è legato all'uso del pooling e del downsampling, che comportano una perdita progressiva di informazione spaziale. Questa riduzione di dettaglio può compromettere la precisione nella localizzazione dei margini delle strutture, un aspetto particolarmente critico in presenza di confini patologici sfumati, come quelli che si riscontrano nelle lesioni cerebrali (Ronneberger et al. 2015; Milletari et al. 2016).

Un secondo limite riguarda la difficoltà delle CNN nel catturare il contesto globale dell'immagine. Le convoluzioni operano su campi recettivi locali e la comprensione delle relazioni tra regioni spazialmente distanti avviene solo indirettamente attraverso la profondità della rete. Questo può risultare meno efficace per strutture estese o distribuite in modo irregolare, motivando lo sviluppo di meccanismi di attention o di architetture multi-scala per integrare informazioni globali (Oktay et al., 2018; Chen et al. 2017).

Infine, le CNN richiedono grandi quantità di dati annotati per ottenere prestazioni elevate. In ambito medico, l'annotazione voxel-wise è costosa e soggetta a variabilità intra- e inter-operatore, aumentando il rischio di overfitting e riducendo la generalizzabilità dei modelli (Shen, Wu, & Suk, 2017; Cheplygina et al. 2019).

Questi limiti hanno stimolato lo sviluppo di architetture avanzate, come U-Net, Attention U-Net, Residual U-Net, nnU-Net e modelli ibridi CNN-Transformer, progettate per preservare l'informazione spaziale, integrare il contesto globale e ridurre la dipendenza da grandi dataset annotati. Tali evoluzioni rappresentano la base dello stato dell'arte nella segmentazione automatica di immagini mediche, in particolare in ambito neuro-oncologico, dove accuratezza, robustezza e riproducibilità dei risultati sono fondamentali per applicazioni cliniche e ricerca.

Questa esigenza di standardizzazione e confronto oggettivo è alla base di lavori come, che hanno proposto il benchmark BRATS per valutare in modo scientifico le prestazioni dei modelli di segmentazione sulle immagini di Risonanza Magnetica.

U-Net

L'architettura U-Net è tra le più importanti per la segmentazione delle immagini mediche ed è basata sul concetto di Deep Learning. È nata per superare i limiti delle CNN (vedi supra), consentendo di ottenere predizioni precise anche con dataset ristretti. (Morandin.L., 2023).

Nell'imaging biomedico, la segmentazione serve al riconoscimento di strutture di interesse e alla localizzazione dei loro margini. Con l'architettura U-Net non abbiamo più il problema della struttura CNN classica, ossia la perdita di informazioni spaziali dovuta al pooling e al downsampling, perché U-Net funziona unendo dettagli semantici a quelli spaziali ad alta risoluzione.

L'U-Net ha una struttura simmetrica a U, composta da una prima fase encoder-decoder.

La fase di encoding, cioè la codifica, utilizza una serie di blocchi convoluzionali e di downsampling, modificando sempre di più l'immagine e quindi distinguendo i diversi tipi di tessuto presenti.

La fase di decodifica, invece, fa l'opposto: partendo da rappresentazioni compatte, ricostruisce l'immagine alla risoluzione originale utilizzando meccanismi di upsampling e di convoluzione, localizzando con precisione le strutture anatomiche e patologiche.

Il passaggio principale che permette alle u-net di non perdere informazioni spaziali è quello delle skip connections, cioè delle connessioni dirette tra un livello di decodifica e il corrispondente livello di codifica, compensando così la perdita di informazioni che altrimenti sarebbe dovuta al downsampling⁵. (Morandin.L., 2023)

L'uso delle skip connections è particolarmente efficace nella Risonanza Magnetica, dove i margini patologici non sono omogenei e risultano sfumati.

Il campo in cui l'architettura U-net ha giovato di più è quello della segmentazione tumorale a livello cerebrale perché, in particolare, gliomi hanno un'alta eterogeneità intratumorale, che viene individuata e definita dai sistemi con architettura U-net in modo preciso grazie alle altre informazioni semantiche e spaziali. (Yousef R. et al., 2023).

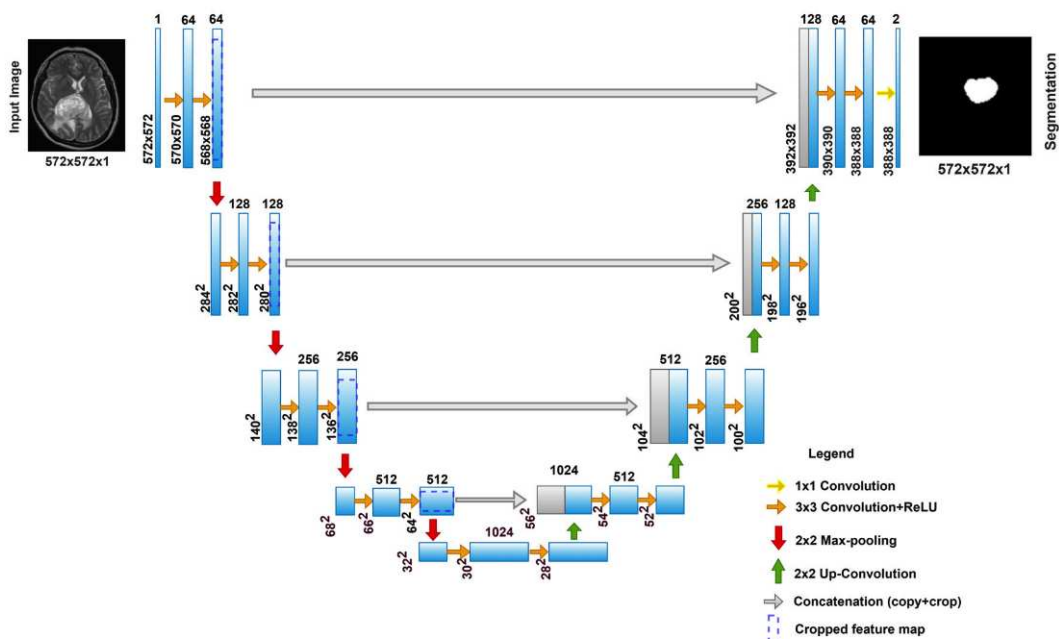


Immagine n.1 (Yousef R. et al., *U-Net-based models towards optimal MR brain image segmentation*. 2023)

L'U-Net ha poi avuto un grande successo, che ha permesso di sviluppare varianti dell'architettura che soppressero alle limitazioni della versione standard o si adattassero a un compito specifico. La presenza di questo sviluppo attesta la validità e l'affidabilità dell'architettura U-Net.

⁵ Operazione anche chiamata *subsampling*, consiste nella riduzione dei pixel che compongono l'immagine.

Tecniche di imaging 2D e 3D

Una delle principali distinzioni tra le immagini mediche è tra immagini 2D e 3D. Questo influisce sulla dimensione dei dati in input e richiede la modifica dell'architettura U-net di partenza, per adattarla meglio al compito.

Le due architetture, U-Net 2D ed U-Net 3D, rappresentano due approcci differenti che rispondono a problemi con caratteristiche diverse, come la dimensione dei dati, la capacità computazionale e l'obiettivo di analisi clinica.

Segmentazione 2D

Il sistema 2D, partendo da immagini volumetriche, come quelle della Risonanza Magnetica, va a scomporre il volume in più unità singole, analizzandole ed elaborandole singolarmente, così da creare una segmentazione per ogni slice dell'immagine.

Questo meccanismo riduce la necessità di strumenti con elevata capacità computazionale e la spesa hardware. L'addestramento risulta semplificato e può essere ampliato, grazie alla minore grandezza delle singole slice rispetto a un volume completo. (Wang et al., 2022)

Questo approccio a singola slice causa però una perdita di informazioni spaziali lungo l'asse tra le slice. Ciò limita la possibilità di avere una visione completa del volume.

Perciò le strutture patologiche, come i tumori che hanno un volume tridimensionale, possono avere una perdita parziale di informazioni creando problemi a causa della potenziale incoerenza di margine e volume tra slice adiacenti. (Yousef R. et al., 2023)

Segmentazione 3D

L'approccio 3D, elaborando direttamente il volume completo, consente di analizzare informazioni in tutti gli assi spaziali, fornendo un'informazione più completa e complessa.

La U-Net, in questo caso, utilizza convoluzioni tridimensionali; ciò consente una continuità volumetrica fondamentale nelle rappresentazioni volumetriche, migliorando nettamente la qualità della segmentazione. (Philip Novosad et al., 2020)

La segmentazione 3D è nettamente superiore a quella 2D in certi ambiti, come in ambito di Risonanza Magnetica cerebrale, in quanto l'approccio 3D permette di mantenere una coerenza anatomica migliore rispetto all'approccio 2D oltre che una migliore distinzione in caso di strutture molto piccole. (Aigerim Dautkulova et al., 2025)

Un'altro campo in cui il modello 3D è nettamente superiore si trova nelle Risonanza Magnetica che viene utilizzata su pazienti molto eterogenei, da neonati ad adulti e ciò permette nel modello 3D di avere un'elevata riproducibilità e robustezza, superiore a quella 2D.

(Simarro J. et al., 2024)

Tuttavia, la segmentazione 3D presenta una complessità nettamente superiore, richiedendo un supporto hardware costoso e un addestramento molto più lento, con conseguente minore numero di campioni nel dataset.

Il problema principale è l'overfitting: si estraggono dal database di addestramento così tanti dati specifici che, a volte, vengono selezionati anche dati che peggiorano la segmentazione, ad esempio memorizzando dati inutili o sfavorevoli come il rumore. (Yousef R. et al., 2023)

Attention U-Net

Una delle principali evoluzioni del modello U-Net è la variante Attention. Questa versione risolve un problema dell'architettura base ovvero che le skip connections trasferiscono indiscriminatamente tutte le informazioni dall'encoder al decoder, ma così possono giungere anche rumore o dati ridondanti e ciò comporta una diminuzione del contrasto, soprattutto nelle strutture piccole, come nel campo della segmentazione cerebrale. (Wang et al., 2022)

L'attenzione U-Net è stata introdotta per contrastare questo limite. Utilizza moduli di attenzione tra l'encoder e il decoder per effettuare un'analisi selettiva dei dati utili. (Yousef et al., 2023)

Questi moduli sono inseriti lungo le skip connections che collegano l'encoding e il decoding. Analizzano due tipi di dati: i dati dell'encoder ad alta risoluzione e quelli dal segnale di gating del decoder, che contengono informazioni semantiche. Si calcola il peso di ogni dato e si trasferisce proporzionalmente all'immagine di output. Così, il filtro enfatizza i dati importanti ed elimina o sopprime quelli meno utili, riducendo il rumore o le strutture non pertinenti. (Yousef et al., 2023)

Questa tecnica riduce drasticamente il problema della segmentazione con U-Net dovuto a confini sfocati o all'eterogeneità intralesionale. È importante in presenza di edema peritumorale o di componenti del nucleo simili ai tessuti circostanti, che rendono difficile una separazione netta. Con questo meccanismo, il modello individua meglio le strutture piccole e poco contrastate, migliorando la precisione del metodo. (Yousef et al., 2023)

La generalizzazione di questo modello funziona meglio rispetto a quella del modello di partenza. I meccanismi di attenzione consentono di concentrare la rete su informazioni significative con maggiore precisione. In questo modo si elimina il rischio che il modello apprenda pattern irrilevanti, in grado di compromettere la segmentazione.

A livello computazionale, l'attention U-Net aumenta i calcoli in modo contenuto, mantenendo un buon equilibrio tra complessità ed efficienza (Wang et al., 2022).

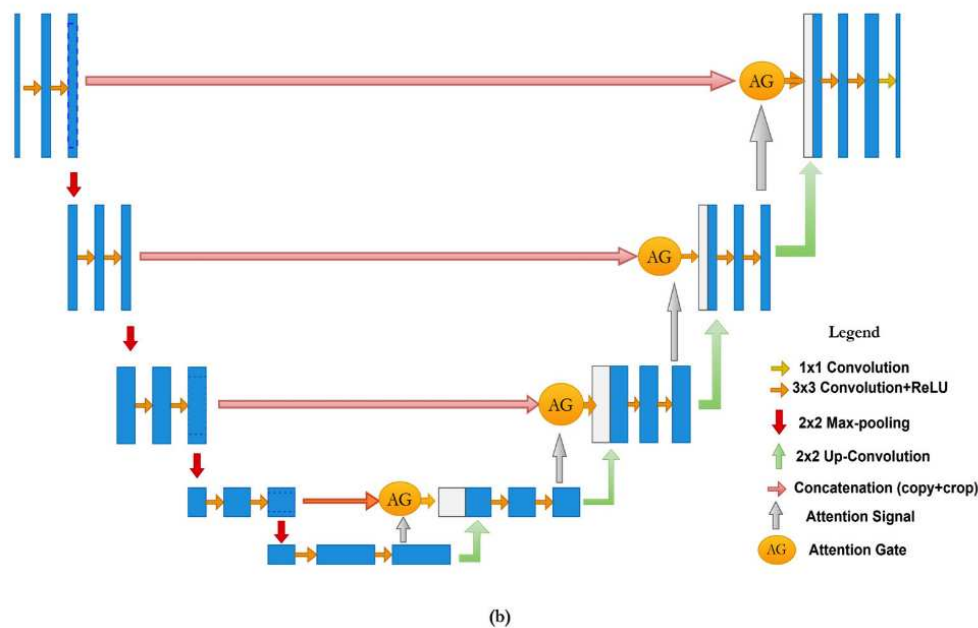


Immagine n.2 (Yousef R. et al., *U-Net-based models towards optimal MR brain image segmentation*. 2023)

Residual U-Net

Con l'aumento della complessità dei compiti di segmentazione in Risonanza Magnetica, è emersa la necessità di architetture sempre più profonde, in grado di catturare pattern complessi nelle immagini medicali. Tuttavia, l'aumento della profondità delle reti neurali convoluzionali è spesso associato a difficoltà di addestramento, tra cui la degradazione delle prestazioni all'aumentare del numero di strati.

Il problema riguarda l'addestramento e la comunicazione tra gli strati. Strati sempre più profondi sono più difficili da controllare e addestrare. Questo rende il sistema instabile a grandi profondità.

La residual U-Net, oltre alla struttura di base con encoder, decoder e skip connection, introduce un quarto elemento: le shortcut che permettono la comunicazione interstrato, andando così a migliorare la comunicazione.

L'input entra in un blocco della rete, il blocco delle convoluzioni, ma, a differenza del modello U-Net base, va a sommare il risultato con l'input originale, andando così a mantenere tutte le informazioni all'interno di tutti gli strati.

Questo concorre quindi ad una maggior stabilità dell'addestramento senza problemi di overfitting, andando ad aumentare significativamente anche la qualità della segmentazione in output, permettendo di arrivare a livelli molto profondi, utile quindi specialmente nelle immagini RM

cerebrali dove c'è una grande variabilità morfologica e confini poco definiti.(Wang et al., 2022; Yousef et al., 2023)

Le migliori informazioni vanno anche ad aumentare la robustezza e riproducibilità dei risultati, data la miglior affidabilità dei dati in output causata dalla diminuzione della sensibilità del modello ad elementi inutili, che possono sviare il focus del processo di segmentazione.

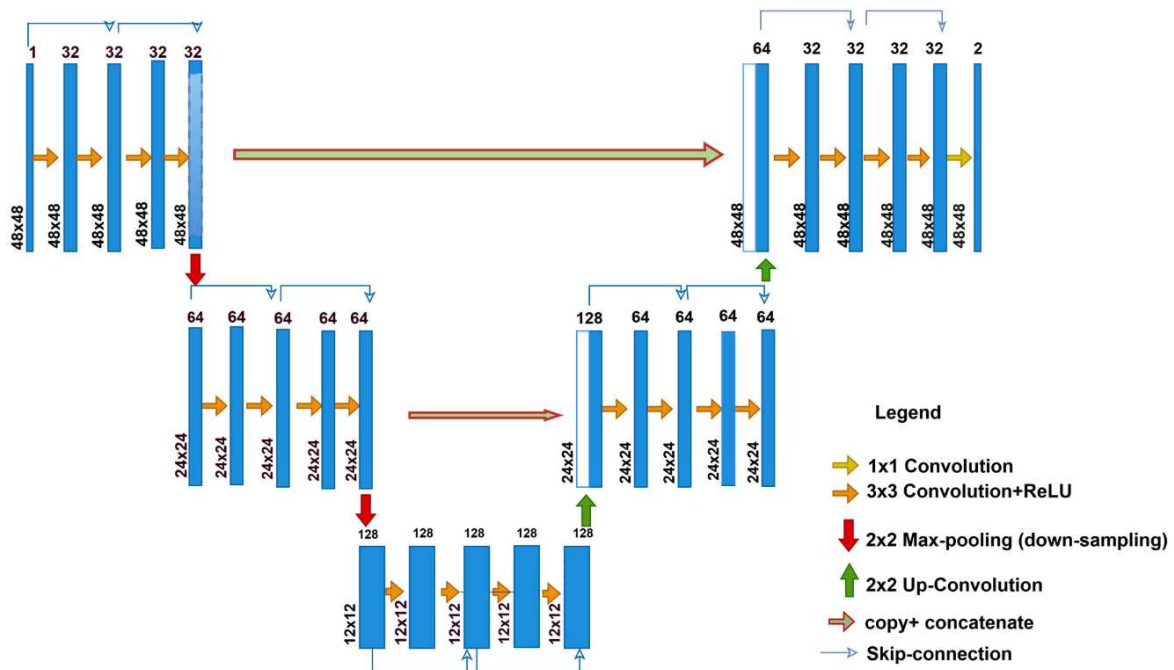


Immagine n.3 (Yousef R. et al., *U-Net-based models towards optimal MR brain image segmentation*. 2023)

Algoritmo nnU-Net

Il riferimento principale per questo sottocapitolo è il lavoro di Fabian Isensee. (Isensee et al., 2021) Nonostante il successo dell'architettura U-Net con lo sviluppo delle sue varianti specializzate, nell'applicazione clinica rimane una forte componente decisionale umana che può inficiare il corretto o appropriato funzionamento dell'architettura.

La scelta della dimensione, della profondità di rete, delle dimensioni dei patch e dei parametri di addestramento richiede l'intervento manuale dell'operatore. Questi fattori possono cambiare significativamente i risultati e rendere la segmentazione difficilmente ripetibile.

Per questo è stata introdotta la nnU-Net, uno strumento che rende automatiche queste decisioni. L'obiettivo di nnU-Net è ottimizzare le prestazioni e rendere ripetibile la segmentazione. Analizza e decide qual è la scelta migliore per ottenere la massima precisione possibile.

nnU-Net utilizza una serie di parametri e regole. Questi definiscono quale sia l'approccio migliore in base ai dati di input forniti. Le regole riguardano la dimensione dei dati, la dimensione delle patch in input, la profondità della rete e i parametri di addestramento e inferenza. In questo modo,

vengono eliminate le scelte manuali e l'architettura è semplice e veloce da impostare. (Veiga-Canuto, 2023)

La scelta automatica della dimensionalità della rete, ovvero la decisione tra analisi 2d e 3D, è una criticità, perché trovare il giusto compromesso, cercando di ottenere il risultato migliore possibile compatibilmente con l'hardware utilizzato e con la qualità del risultato, è molto complesso, soprattutto in Risonanza Magnetica, dove la grandezza dei voxel non è sempre isotropica.

Il sistema decide quale opzione scegliere analizzando le proprietà geometriche del dataset, in particolare la grandezza del voxel sull'asse z. Questo parametro determina lo spessore della fetta e il numero di fette nel volume da analizzare. In questo modo, si sceglie il miglior compromesso tra facilità di calcolo, ampiezza del database, computabilità più difficile, più informazioni, possibili artefatti, rumore e spazio richiesto.

La nnU-Net non si limita a una scelta singola. Può addestrare più configurazioni, che vengono poi analizzate con metriche precise, tra cui la principale è il Dice Similarity Coefficient (DSC).

Questa tecnica automatizzata ha dimostrato elevata robustezza rispetto alla variabilità dei dati. Mantiene prestazioni elevate nonostante l'eterogeneità di acquisizione, protocolli, risoluzioni e sedi anatomiche.

In sintesi, la gestione automatica permette un equilibrio ottimale tra accuratezza, robustezza e sostenibilità computazionale.

Ulteriori studi hanno confermato l'affidabilità dell'architettura, confrontando i risultati ottenuti da medici radiologi esperti con quelli dell'intelligenza artificiale e ottenendo un'elevata concordanza.

Model	DSC			HD95%			Parameters	Time
	ET	WT	TC	ET	WT	TC		
3D U-Net [96]	0.779	0.881	0.827	27.23	7.788	8.278	23 M	6 h (1.2 s/sample)
Modified U-Net	0.781	0.905	0.807	26.607	5.785	18.545	26 M	10 h (3.8 s/sample)
Attention U-Net [44]	0.778	0.878	0.827	26.662	7.794	8.305	23.2 M	6.2 h (1.7 s/sample)
R2 Attention U-Net [97]	0.7426	0.8784	0.7993	36.653	9.228	9.95	22 M	5.8 h (0.8 s/sample)

Tabella n.1 (Yousef R. et al., *U-Net-based models towards optimal MR brain image segmentation*, 2023)

La tabella 1 fornisce un'analisi comparativa delle principali varianti dell'architettura U-Net applicate alla segmentazione dei tumori cerebrali in immagini di Risonanza Magnetica.

L'obiettivo della tabella è mettere a confronto le loro prestazioni quantitative in termini di accuratezza di segmentazione, utilizzando come metrica principale il Dice Similarity Coefficient (DSC, metrica utilizzata per valutare quanto una segmentazione automatica coincide con la segmentazione di riferimento).

Per ciascun modello riportato in letteratura, la tabella presenta i valori di DSC ottenuti nelle tre sottoregioni tumorali: Enhancing Tumor (ET), Whole Tumor (WT) e Tumor Core (TC). Questa

suddivisione, dal momento che le tre regioni presentano caratteristiche morfologiche e di contrasto differenti, pongono sfide diverse ai sistemi di segmentazione automatica.

Dall'analisi dei dati emerge che numerose varianti della U-Net, anche strutturalmente molto differenti tra loro, raggiungano prestazioni comparabili, in particolare per la regione del tumore intero (WT), dove i valori di DSC risultano generalmente più elevati e stabili.

La regione dell'Enhancing Tumor (ET) presenta generalmente una maggiore variabilità nei valori di DSC, risultando quindi più complessa da segmentare rispetto al Whole Tumor (WT) e al Tumor Core (TC), che mostrano valori più elevati e stabili.

La tabella fornisce quindi una visione d'insieme dello stato dell'arte, evidenziando come la U-Net, nelle sue diverse declinazioni, continui a rappresentare una solida architettura di riferimento per la segmentazione di immagini RM in ambito neuro-oncologico.

(Yousef R. et al., 2023)

Transformer e modelli ibridi

Tali conclusioni suggeriscono che la nnU-Net possa essere un valido strumento di supporto alla pratica clinica, riducendo il lavoro manuale a carico del medico radiologo.

La U-Net è la baseline di riferimento per valutare nuove architetture di segmentazione medica. Tra queste, ci sono le architetture basate su meccanismi di azione, che si focalizzano in modo efficace su determinate aree ed eliminano l'influenza di aree non rilevanti o rumorose. Un'ulteriore linea di sviluppo è rappresentata dai modelli ibridi che partono da un modello CNN e ci aggiungono moduli specifici per aumentarne la specializzazione, la robustezza e la capacità di adattarsi a ogni possibile scenario, anche molto eterogeneo.

Tutte queste nuove architetture vengono quindi confrontate con il modello di architettura nnU-Net per la loro valutazione.

Negli ultimi anni, l'interesse verso l'uso dei Transformer e dei modelli ibridi CNN-Transformer nella segmentazione di immagini mediche è aumentato. Questo perché le CNN hanno difficoltà a cogliere il contesto globale nelle immagini complesse di Risonanza Magnetica.

Le reti neurali convoluzionali sono molto efficaci nell'estrarre informazioni locali grazie ai filtri con campi recettivi limitati. Tuttavia, le CNN modellano il contesto globale solo indirettamente, modificando le informazioni tra gli strati. Questo è un limite, soprattutto quando si tratta di una visione d'insieme. È il caso di strutture anatomiche grandi o di lesioni distribuite in modo irregolare.

I Transformer si basano su meccanismi di self-attention che consentono di modellare direttamente le dipendenze globali tra gli elementi dell'input. Questi meccanismi collegano regioni spazialmente distanti. Ogni parte di un'immagine si collega con tutte le altre, permettendo così una valutazione della segmentazione più accurata che tiene conto di tutte le parti dell'immagine.

Nei modelli Transformer, l'immagine viene suddivisa in piccoli patch. Ogni patch valuta le relazioni con le altre e riceve un peso che guida l'attenzione del modello. Tuttavia, la self-attention richiede molta potenza di calcolo per confrontare tutte le patch tra loro. Sono necessari molti dati per addestrare efficacemente il modello.

Per sopperire alle limitazioni e ai problemi dei modelli transformer e CNN, la letteratura propone modelli ibridi. Questi modelli combinano convoluzioni e attention. Le CNN vengono utilizzate per estrarre informazioni locali. I moduli di self-attention sono inseriti per unire le varie patch, ottenendo una visione globale e relazioni tra patch distanti.

Ciò prende quindi le caratteristiche di precisione spaziale dalla CNN e l'accuratezza nella segmentazione e nell'individuazione di confini, anche non ben definiti, fornita dai transformer. I modelli ibridi sono in grado di affrontare compiti complessi e sono molto efficaci nella segmentazione di tumori cerebrali, caratterizzati da elevata eterogeneità e variabilità morfologica. (Wang et al., 2022).

Nonostante i risultati promettenti, l'introduzione dei modelli ibridi nella pratica clinica presenta ancora problemi. L'elevato costo computazionale e la necessità di dataset ampi e ben annotati lo rendono inadatto alla pratica clinica. Inoltre, la struttura complessa riduce la riproducibilità dei risultati. Un singolo cambiamento può causare variazioni negli strati successivi; l'attuazione in ambito medico resta distante.

Per questi motivi, i modelli ibridi restano un elemento in sviluppo, valido per il futuro.

(Renard et al., 2020)

Limiti intrinseci alle tecniche di segmentazione con Deep Learning

Il principale limite dei modelli di Deep Learning per la segmentazione è la scarsa disponibilità di dati etichettati di alta qualità. Il gold standard odierno per la segmentazione di immagini radiologiche si basa sull'annotazione voxel-wise, le cui etichette vengono date dai medici radiologi e clinici specializzati.

Inoltre, la segmentazione manuale su cui si basa presenta problemi di variabilità inter- e intra-operatore, il che incide sulla riproducibilità e quindi richiede una maggiore quantità di dati etichettati. (Renard et al., 2020)

Questa problematica è ulteriormente accentuata in RM, dove la varietà di sequenze e protocolli impiegate rende difficile creare dataset omogenei e annotati.

Di conseguenza i modelli di Deep Learning tradizionali richiedono grandi quantità di dati per non avere il fenomeno dell'overfitting, cioè l'adattarsi ad elementi specifici non pertinenti come rumore o artefatti.

Per questo si sono cercate delle soluzioni che non necessitino di database completamente etichettati rendendo più efficaci i dati disponibili, tali approcci si sono rivelati validi in ambito clinico, consentendo l'uso di dati non annotati o parzialmente annotati. I metodi che utilizzano questa tipologia di approccio sono il semi-supervised e il self-supervised.

(Liu Z. et al., 2024)

Approcci semi-supervised

Il problema della limitata quantità di dati nella segmentazione di immagini in Risonanza Magnetica è discusso nella letteratura recente ed evidenzia come l'annotazione ideale voxel-wise, a ogni singolo voxel, richieda un impegno quasi insostenibile da parte di esperti e clinici, risultando così costosa in termini di denaro e tempo, con variabilità inter- e intra-osservatore.

Quindi un approccio semi-supervised, cioè che combina una quantità ristretta di immagini annotate con immagini non annotate, facendo sì che il sistema crei pseudo-annotazioni (etichette generate automaticamente) e pattern che permettono di utilizzare anche i dati delle immagini non annotate, riducendo così costi e tempi.

Questa tipologia di approccio si rivela molto valida, soprattutto in ambito clinico, per la sua minima quantità di annotazioni necessarie e per le sue prestazioni comunque buone, compatibili con quelle dei modelli completamente annotati. (Liu Z. et al., 2024)

Approcci self-supervised

Un'ulteriore evoluzione dell'approccio semi-supervised è l'approccio self-supervised, che consiste nell'avere un dataset di immagini completamente non annotate che, mediante un pre-addestramento basato sul Deep Learning, apprenda solo da immagini senza annotazioni, permettendo così di utilizzare moltissime immagini che, dopo un passaggio di fine-tuning con poche immagini medicali annotate, offrono una migliore robustezza e generalizzabilità con dataset molto piccoli.

(Jiang J. et al., 2022)

Area Uno – Qualità dell'immagine e requisiti per la segmentazione con IA

Dipendenza degli algoritmi di Deep Learning dalla qualità del dato

La qualità dell'immagine è considerata un prerequisito fondamentale per il corretto funzionamento degli algoritmi di IA in Radiologia.

I sistemi di segmentazione automatica basati su Deep Learning sono modelli data-driven⁶ e sono fortemente dipendenti dai dati di input: in letteratura è ampiamente documentato come rumore, artefatti, variazioni di contrasto e differenze nei parametri di acquisizione possano ridurre significativamente le performance dei modelli di segmentazione, con decrementi misurabili di metriche quali Dice Similarity Coefficient e aumento della Hausdorff Distance. Tutti questi elementi possono infatti andare a modificare la percezione del bordo degli oggetti da parte degli algoritmi, rendendo più difficoltoso per l'IA il riconoscimento della figura. (Najjar, 2023)

La qualità dell'immagine RM non rappresenta quindi solo un requisito diagnostico, ma un prerequisito tecnico per l'affidabilità degli algoritmi di Intelligenza Artificiale.

Ne deriva che la produzione di immagini standardizzate, prive di artefatti rilevanti e con adeguato rapporto segnale-rumore, costituisce una condizione essenziale per la corretta integrazione dell'IA nel workflow tecnico.

Artefatti RM

Un artefatto è una caratteristica dell'immagine che non rappresenta fedelmente l'anatomia o la fisiologia reale del paziente.

Gli artefatti in RM vengono suddivisi in grandi categorie: quelli legati alla non uniformità del campo magnetico statico (B0), quelli legati ai gradienti, quelli legati al sistema a radiofrequenza e quelli legati alla fase di codifica del segnale MR. (Graves & Mitchell, 2013)

Dal punto di vista della segmentazione automatica, gli artefatti possono:

- alterare i contorni anatomici;
- modificare l'intensità dei voxel;
- generare strutture duplicate o distorte;
- introdurre pattern non anatomici che l'algoritmo può erroneamente interpretare come tessuto reale.

⁶ Data-driven: modello che implica l'utilizzo di mezzi computazionali come ad esempio l'IA per andare ad analizzare dati e guidare alla risposta.

Artefatti legati al campo magnetico statico (B0)

La Risonanza Magnetica richiede un campo magnetico statico uniforme. Tuttavia, questa uniformità non è mai perfetta. Le cause possono essere costruttive, ambientali o dipendere dalla presenza del paziente, poiché i tessuti hanno differenti suscettibilità magnetiche.

La non uniformità del campo B0 diventa evidente nella saturazione selettiva del grasso. Questa tecnica funziona solo se la frequenza dell'acqua e quella del grasso sono ben separate e costanti nel volume studiato. Se il campo non è uniforme, la soppressione del grasso può essere incompleta o, in alcuni casi, sopprimere anche parte del segnale dell'acqua. In regioni toraciche o addominali, soprattutto vicino alle interfacce aria-tessuto, questo problema è frequente.

La non uniformità del campo influisce su:

- soppressione del grasso incompleta o non selettiva;
- tecniche di separazione acqua/grasso (Dixon), che possono fallire in presenza di forti variazioni locali del campo e generare errori di classificazione dei voxel;
- distorsioni locali del campo e perdita di segnale in prossimità di metallo. (Graves & Mitchell, 2013)



Immagine n.4 (Graves, M.J. and Mitchell, D.G., Body MRI artifacts in clinical practice: A physicist's and radiologist's perspective, 2013)

L'immagine n.4 rappresenta il distretto addominale visto in una sezione coronale di una sequenza DIXON only water (Graves & Mitchell, 2013) che mostra solamente i tessuti ricchi di acqua come muscoli organi parenchimatosi liquidi o lesioni con alta concentrazione di acqua o strutture vascolari, questa sequenza è quindi utile per la visione di edema o infiammazione dei tessuti molli. L'immagine presenta un artefatto negli angoli dovuto ad una non uniformità del campo magnetico B₀, che si traduce in una deformazione dei contorni anatomici e quindi un errore nei volumi segmentati, inoltre provoca una presenza di bande nere che vanno ad alterare completamente la classificazione dei tessuti.

L'immagine infatti è una DIXON only water che mostra solamente i tessuti ricchi di acqua come muscoli organi parenchimatosi liquidi o lesioni con alta concentrazione di acqua o strutture vascolari, questa sequenza è quindi utile per la visione di edema o infiammazione dei tessuti molli.

Artefatti legati ai gradienti

I gradienti sono responsabili della codifica spaziale del segnale. Quando vengono commutati rapidamente ed inducono correnti parassite del magnete. Queste correnti generano campi magnetici secondari che distorcono le forme d'onda dei gradienti stessi. Il risultato può essere instabilità del segnale, distorsioni o ghosting, specialmente in sequenze sensibili come la diffusione o la phase-contrast.

Un altro limite strutturale è la non linearità dei gradienti. Idealmente il gradiente dovrebbe variare in modo lineare nello spazio, ma ai bordi del campo di vista questa linearità si perde. Ciò comporta distorsioni geometriche dell'immagine, con possibile impatto sulle misurazioni quantitative o sulle procedure guidate. (Graves & Mitchell, 2013)

Per la segmentazione automatica ciò implica:

- deformazione della morfologia reale delle strutture;
- errore nelle misurazioni volumetriche;
- ridotta affidabilità nelle analisi quantitative longitudinali.

In applicazioni di follow-up o confronto inter-esame, tali distorsioni possono tradursi in variazioni apparenti di volume non correlate a modificazioni cliniche reali.

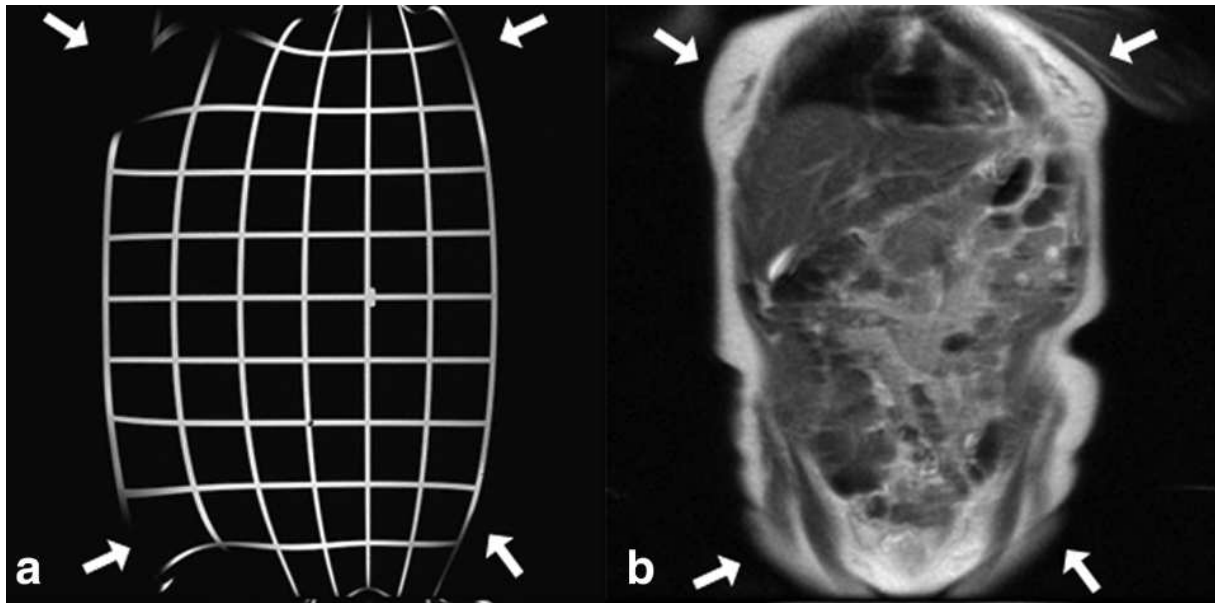


Immagine n.5 (Graves, M.J. and Mitchell, D.G., Body MRI artifacts in clinical practice: A physicist's and radiologist's perspective, 2013)

L'immagine n.5 rappresenta nell'immagine a. una phantom grid, cioè una griglia utilizzata nei controlli di qualità per la verifica di artefatti da deformità di campo e nell'immagine b. una slice di un'esame entero RM acquisito in coronale, con pesatura T2, soggetta a deformazioni magnetiche più frequentemente data la necessità di avere un FOV grande, in questo caso è di 48 cm, che aumenta la probabilità di distorsioni d'immagine avendo il campo magnetico della RM una diminuzione di linearità agli estremi.

L'immagine b. mostra una deformità di campo dovuta alla non linearità dei gradienti che può dare deformazione della morfologia reale degli organi, errori su misurazioni volumetriche che quindi portano ad una ridotta affidabilità nelle analisi quantitative.

Artefatti legati al sistema a radiofrequenza (B1)

Il sistema RF ha il compito di eccitare i protoni e di ricevere il segnale. Anche qui possono verificarsi non uniformità.

A campi elevati, soprattutto a 3T, il campo di trasmissione (B1+) può diventare non uniforme a causa delle interferenze tra le onde RF. Si possono creare regioni di interferenza costruttiva o distruttiva, con aree più intense e più scure (shading). Questo fenomeno è legato alla lunghezza d'onda della RF nel corpo umano, che diventa paragonabile alle dimensioni corporee.

Il campo di ricezione (B1-) può anch'esso essere non uniforme, soprattutto quando si utilizzano coil array. In questi casi l'intensità del segnale varia con la profondità e possono comparire aree di ipo-intensità, ad esempio se un elemento della coil è guasto.

Esistono poi artefatti dovuti a segnali provenienti da fuori dal volume di interesse, come il cusp artifact nelle sequenze fast spin echo, e artefatti dovuti a interferenze RF esterne, come lo zipper artifact o i fenomeni da spike noise, che producono pattern periodici o a bande nell'immagine. (Graves & Mitchell, 2013)

Queste disomogeneità di intensità possono influenzare i modelli di Deep Learning che non siano stati addestrati su dataset con analoga variabilità, determinando segmentazioni incoerenti o errori di classificazione tessutale.

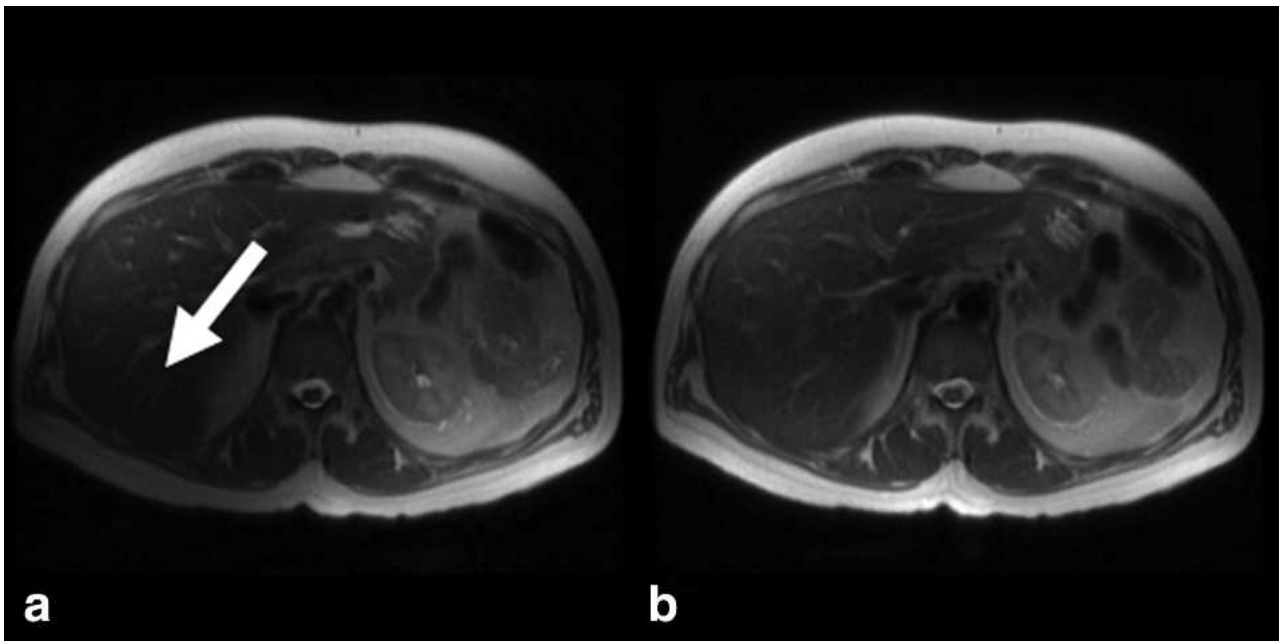


Immagine n.6 (Graves, M.J. and Mitchell, D.G., Body MRI artifacts in clinical practice: A physicist's and radiologist's perspective, 2013)

L'immagine n.6 rappresenta dati di un'esame RM addome superiore in particolare sezioni assiali con pesatura T1, a livello delle vertebre T10-11, in cui è presente un artefatto che deriva dal malfunzionamento di una o più coil di una bobina⁷ che va a ricevere il segnale in modo anomalo o parziale, dando quindi un calo di segnale nella zona difettata, potendo dare così in certi casi delle false aree di ipodensità.

⁷ Bobina/coil: sono componenti fondamentali della Risonanza Magnetica, composti da fili di rame avvolti in più canali ed agiscono come antenne per la trasmissione e ricezione di impulsi a radiofrequenze. (Gruber B., et al., 2018)

Artefatti legati alla codifica del segnale MR

Durante la formazione dell'immagine, il segnale viene codificato tramite gradienti di frequenza e di fase. Se il campo di vista nella direzione di fase è troppo piccolo, strutture al di fuori del FOV vengono aliasate, cioè "ripiegate" nell'immagine.

Le tecniche di parallel imaging, come SENSE o GRAPPA, riducono il tempo di acquisizione, ma possono generare artefatti se la calibrazione delle coil non è corretta o se il FOV è inadeguato, con aliasing centrale o ghosting.

Il chemical shift è un effetto intrinseco dovuto alla differenza di frequenza di precessione tra l'acqua e il grasso. Questo produce disallineamento spaziale lungo la direzione di frequenza e fenomeni di cancellazione di segnale nelle immagini out-of-phase.

Nelle sequenze fast spin echo, il decadimento T2 durante l'eco train genera una sfocatura (echo train blurring) lungo la direzione di fase.

Infine, il movimento è una delle cause più comuni di artefatti. La respirazione, il battito cardiaco, il flusso sanguigno o il movimento del paziente determinano modulazioni del segnale che si manifestano come ghosting nella direzione di fase, sfocatura o perdita di segnale. Anche le immagini ottenute per sottrazione possono essere alterate da piccoli disallineamenti tra le acquisizioni. (Graves & Mitchell, 2013)

Per un algoritmo di segmentazione, tali alterazioni possono generare:

- segmentazione di strutture duplicate
- frammentazione dei contorni
- errori nella localizzazione spaziale

La letteratura evidenzia come la presenza di artefatti da movimento sia associata a riduzione significativa delle performance di modelli di segmentazione automatica. (Najjar, 2023)

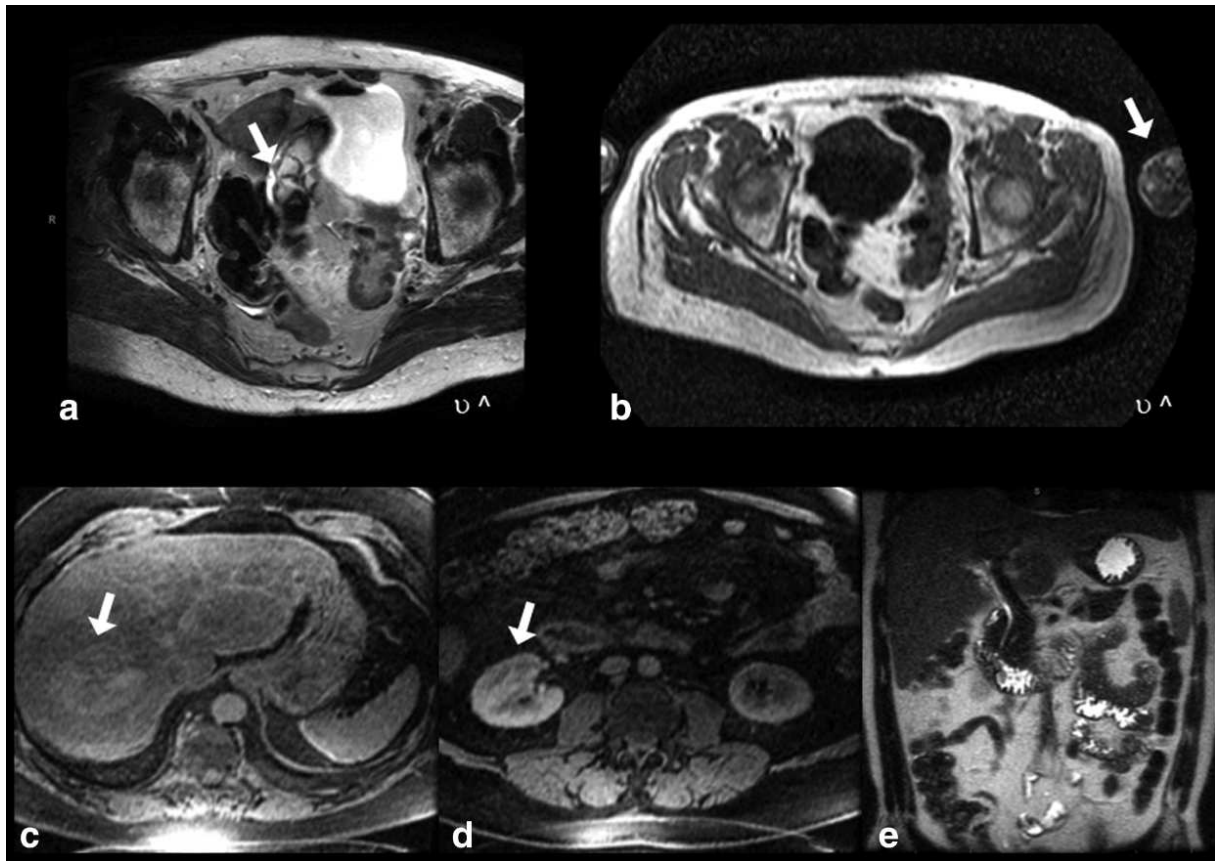


Immagine n.7 (Graves M.J. and Mitchell D.G., Body MRI artifacts in clinical practice: A physicist's and radiologist's perspective, 2013)

Nell'immagine n.7 si possono vedere esempi di artefatti da aliasing di fase 2D e 3D, nell'immagine a. il campo visivo nella direzione della codifica di fase è troppo piccolo, quindi il polso sinistro risulta soggetto ad aliasing nel bacino. Nell'immagine b. invece si ha la stessa acquisizione di a ma con campo più ampio in cui il polso sinistro nella sua posizione corretta.

L'immagine c. mostra l'artefatto da aliasing di fase della selezione della slab: l'immagine mostra una struttura che sembra trovarsi all'interno del fegato. L'immagine d. della parte inferiore dell'acquisizione mostra che la struttura sovrastante il fegato era in aliasing rispetto al rene destro.

Vi è quindi una segmentazione di strutture in realtà duplicate portando così confusione tra anatomia reale ed aliasata. Vi è dunque un errore nella localizzazione spaziale.

Parametri quantitativi di qualità dell'immagine

Vi sono dei parametri che influenzano direttamente la qualità delle immagini RM, che sono il rapporto segnale-rumore, la risoluzione spaziale, il contrasto e il rapporto contrasto-rumore. (Coriasco M., Rampado O., Bradac G.B., "Elementi di Risonanza Magnetica")

Presenza di rumore

Il rumore rappresenta la componente casuale e non informativa del segnale che si sovrappone al segnale utile proveniente dai tessuti, determinando una riduzione del rapporto segnale-rumore (SNR) e della qualità complessiva dell'immagine. (Bushberg et al., 2012)

Nei sistemi di Deep Learning applicati all'imaging medico, la presenza di rumore può compromettere la corretta identificazione delle strutture anatomiche, poiché il modello può apprendere pattern non rilevanti o non generalizzabili. Quando elementi estranei alla morfologia del paziente vengono comunque presi in considerazione dagli algoritmi IA si ha riduzione delle performance di segmentazione. (Litjens et al., 2017)

Rapporto segnale-rumore

Il Signal-to-Noise Ratio (SNR) rappresenta il rapporto tra il segnale utile proveniente dai tessuti e la componente casuale e non informativa che si sovrappone ad esso. Un basso SNR si traduce in un aumento della granulosità dell'immagine e in una maggiore variabilità dell'intensità di segnale anche tra pixel appartenenti a tessuti omogenei, con conseguente riduzione della qualità complessiva dell'immagine. (Coriasco et al., 2014)

Dal punto di vista della segmentazione automatica, un elevato livello di rumore può compromettere la stabilità dei contorni generati dal modello, aumentando la variabilità inter-esame e riducendo la riproducibilità delle misurazioni.

Il Contrast-to-Noise Ratio (CNR) esprime invece la capacità di distinguere, in presenza di rumore, una struttura dallo sfondo circostante. Anche in condizioni di contrasto teoricamente adeguato, un'elevata componente di rumore può ridurre la discriminabilità tra tessuti, poiché le differenze di segnale legate alla morfologia possono sovrapporsi alle fluttuazioni casuali del rumore. (Coriasco et al., 2014)

Nei modelli di Deep Learning applicati all'imaging medico, un basso CNR può ridurre la separabilità statistica tra classi tessutali, con impatto negativo sulle metriche di accuratezza e sulla capacità di generalizzazione del modello. (Tajbakhsh et al., 2020)

Risoluzione spaziale e risoluzione di contrasto

Esistono due tipi di risoluzione, quella spaziale e quella di contrasto.

La risoluzione spaziale è la capacità di distinguere due strutture vicine come separate ed è direttamente influenzata dalla dimensione dei pixel e dallo spessore di strato; la risoluzione di

contrasto è la capacità di distinguere la differenza di intensità di grigi fra due tessuti simili. (Coriasco M., Rampado O., Bradac G.B., “Elementi di Risonanza Magnetica”)

Una bassa risoluzione spaziale significa che i dettagli fini non sono chiaramente distinguibili, poiché i pixel o voxel rappresentano una porzione più grande di tessuto, quindi ne risulta in margini dei tessuti sfumati e non chiaramente distinguibili soprattutto nel caso dell’IA.

Invece in presenza di una bassa risoluzione di contrasto tessuti con caratteristiche simili possono apparire uguali e quindi difficili da distinguere, l’IA però in questo caso riesce meglio a definire le minime differenze come descritto nell’Area zero.

Variabilità dei parametri di acquisizione e standardizzazione

L’ottenimento di un’immagine di qualità ottimale deve confrontarsi con tempi di acquisizione accettabili e questo dipende fortemente da parametri di acquisizione quali:

- tempo di eco (TE): influenza direttamente la pesatura T2 delle immagini e in maniera inversamente proporzionale il livello di segnale, per cui un aumento del TE si traduce in una maggior pesatura T2 e peggior SNR, una sua diminuzione invece comporta una riduzione della pesatura T2 ma un aumento del SNR;
- tempo di ripetizione (TR): un TR breve aumenta la pesatura in T1 e quindi il contrasto ma si traduce in un minor tempo di recupero della magnetizzazione longitudinale e un SNR minore;
- flip angle: influenza in maniera direttamente proporzionale il SNR;
- spessore di fetta: uno spessore maggiore comporta più segnale emesso e quindi un maggior SNR;
- distanza tra due fette: due fette adiacenti possono influenzare il segnale emesso a vicenda, con il cosiddetto effetto cross-talk;
- FOV (field of view): a matrice costante, un maggior campo di vista si traduce in dimensione dei pixel più elevata e quindi minor risoluzione spaziale ma maggior SNR;
- numero di medie: è il numero di volte in cui viene acquisita ogni linea del k-spazio, per cui un maggior numero di medie comporta un aumento del SNR;
- dimensioni della matrice: è data dalla dimensione dei pixel lungo la direzione della codifica di frequenza e di fase, una matrice più grande significa voxel più piccoli e quindi minor SNR ma maggior risoluzione spaziale e viceversa.

(Coriasco M., Rampado O., Bradac G.B., “Elementi di Risonanza Magnetica”)

La variazione dei parametri di acquisizione è l'operazione compiuta dal TSRM prima dell'acquisizione in cui si vanno ad impostare tutti i parametri utili per avere la migliore acquisizione possibile. Ciò per l'IA rappresenta un problema perché va ad aggiungere una variabilità enorme nelle immagini che quindi sono più difficili da analizzare dall'algoritmo soprattutto nel fondamentale database e confronto di immagini. Si parla di domain shift, cioè di discrepanza tra le caratteristiche dei dati di addestramento e quelle dei dati clinici reali. I modelli possono mostrare ridotta generalizzabilità quando applicati a immagini acquisite con protocolli differenti rispetto al dataset di training.

Il tecnico quindi con l'utilizzo di questi software deve cercare di seguire protocolli preimpostati per garantire standardizzazione e uniformità e ridurre la variabilità tecnica, andandoli ad adattare in base alle dimensioni e caratteristiche del paziente, cercando quindi mantenere immagini che siano il più comparabili possibile con i dati di addestramento e garantendo quindi una buona qualità senza artefatti. La standardizzazione dei protocolli di acquisizione risulta fondamentale non solo per finalità diagnostiche bensì anche per garantire robustezza e riproducibilità dei sistemi automatici. (Fortin et al., 2017)

I modelli possono mostrare ridotta generalizzabilità quando applicati a immagini acquisite con protocolli differenti rispetto al dataset di training. (Litjens, et al., 2017)

Caratteristiche dell'immagine che ne garantiscono la confrontabilità

Perché un'immagine sia confrontabile con altre e processabile correttamente da algoritmi di segmentazione automatica, devono essere garantite alcune caratteristiche fondamentali:

1. segnale sufficientemente alto e rumore controllato (SNR e CNR elevati): permettono di distinguere i tessuti tra loro, riducendo la confusione nelle mappe di segmentazione; (Coriasco et al., 2014; Tajbakhsh et al., 2020)
2. assenza di artefatti significativi: gli artefatti possono indurre errori morfologici nei modelli di IA, compromettendo la plausibilità anatomica; (Renard et al., 2020; Kuang et al., 2025)
3. uniformità dei parametri di acquisizione: la stabilità di TR, TE, FOV, spessore di strato e di ricostruzione garantisce che le immagini siano comparabili longitudinalmente e tra pazienti; (Litjens et al., 2017; Tajbakhsh et al., 2020)
4. coerenza inter-sequenza e inter-scanner: la ridotta variabilità tecnica limita il fenomeno di domain shift, aumentando la generalizzabilità del modello; (Tajbakhsh et al., 2020)

5. adeguata qualità percettiva dell'immagine: margini nitidi, uniformità di segnale e leggibilità visiva facilitano la validazione percettiva da parte del tecnico di radiologia.

(Monereo-Sánchez et al., 2021; Veiga-Canuto et al., 2022)

La presenza o l'assenza di queste caratteristiche definisce la confrontabilità dell'immagine: senza di esse, gli algoritmi possono produrre segmentazioni errate o incoerenti, come evidenziato negli studi di Kuang et al., 2025

Implicazioni tecnico-percettive per il tecnico di radiologia

L'analisi dei fattori di qualità dell'immagine evidenzia come la segmentazione automatica non possa prescindere dalla competenza tecnica nella fase di acquisizione. La letteratura sottolinea infatti come l'affidabilità dei sistemi di IA sia strettamente dipendente dalla qualità e coerenza dei dati di input. (Tajbakhsh et al., 2020)

In tale contesto, l'integrazione dell'Intelligenza Artificiale nel workflow dell'imaging RM non riduce la responsabilità professionale, ma la ridefinisce, richiedendo un controllo umano continuo della qualità del dato e dei risultati prodotti. (European Society of Radiology, 2019)

Il TSRM assume pertanto un ruolo determinante nella prevenzione e gestione degli artefatti, nell'ottimizzazione dei parametri di acquisizione, nel mantenimento della coerenza tra protocolli e nella produzione di immagini comparabili nel tempo, configurandosi come primo garante dell'idoneità tecnica dell'immagine all'elaborazione automatizzata.

Area Due – Integrazione degli algoritmi IA nel workflow RM

Introduzione al concetto di workflow in RM

Nell'ambito dell'imaging di Risonanza Magnetica, il workflow radiologico è un sistema organizzato in fasi interdipendenti, in cui ogni segmento incide sull'efficienza complessiva, sulla qualità dell'immagine e sul valore clinico ottenuto. (Lin et al., 2024)

Secondo il modello di Lin, il tempo complessivo dell'esame RM può essere suddiviso in tre categorie fondamentali:

1. value-added time: tempo direttamente produttivo ai fini diagnostici, cioè l'acquisizione dei dati definita "gradient time";
2. business value-added time: tempo necessario per il posizionamento del paziente e delle bobine, le regolazioni dello scanner e la preparazione del paziente;

3. non-value-added time: tempo che non apporta valore alla metodica, cioè non migliora qualità, sicurezza o outcome clinico, come ad esempio tempi di attesa, ripetizioni inutili di passaggi, attività manuali ridondanti, trasferimenti di dati non necessari. All'interno della sala RM definito room time, si può dividere in due tempi:

Un concetto importante è anche quello di room-time (tempo che passa il paziente all'interno della sala radiologica), ed è diviso in:

- gradient time: tempo di acquisizione pura, che va dalla selezione del pacchetto da parte del TSRM alla fine dell'acquisizione; si può affermare quindi che il gradiente-time coincide con il value added time, poiché contribuisce direttamente alla generazione di immagini e quindi alla produzione di informazioni cliniche;
- non-gradient time: tempo in cui non è attiva l'acquisizione, ma il paziente è all'interno della sala di Risonanza Magnetica per la preparazione e posizionamento.

L'articolo scientifico di Lin afferma che, con l'evoluzione delle tecniche, il gradient time è progressivamente ridotto, spostando quindi il focus sulla gestione del non gradient time come obiettivo principale per ottimizzare il workflow. (Lin et al., 2024)

Richiesta clinica

Il workflow radiologico in Risonanza Magnetica ha origine dalla fase di richiesta clinica, momento che ne determina il workflow clinico del paziente. La corretta indicazione clinica è un fattore importante del processo perché permette un'ottimizzazione economica e di tempo andando ad evitare esami non consoni alla sospetta patologia. Inoltre, la definizione precisa del quesito clinico è molto importante per la scelta dell'indagine radiologica più appropriata e, in caso di Risonanza Magnetica, è fondamentale per la scelta del protocollo di studio da adottare e dell'uso o meno del mezzo di contrasto, informazioni importanti che, se non precise, possono influire sul non gradient-time. (Lin et al., 2024)

Organizzazione del tempo e workflow in RM

L'organizzazione temporale, con la costante riduzione del tempo nell'acquisizione, grazie al Deep Learning reconstruction⁸, è concentrata sull'ottimizzazione del turnaround time, cioè il tempo che intercorre tra il completamento dell'ultima sequenza di un paziente e l'inizio della prima sequenza

⁸ Deep Learning Reconstruction: è una tecnica di ricostruzione delle immagini mediche che utilizza il Deep Learning e quindi va a ricostruire le immagini apprendendo come si può ricostruire l'immagine in modo ottimale, grazie allo studio di dataset in cui si confrontano immagini ideali, sotto-campionate o rumorose. (Yousef R. et al., 2023)

del paziente successivo. Questo intervallo include l'uscita del paziente dalla sala magnetica, la pulizia della diagnostica, la preparazione del paziente successivo, il suo posizionamento e i settaggi iniziali. Il turnaround time rappresenta oggi una porzione importante dell'organizzazione, in esami con room time di circa 25 minuti, in cui può arrivare fino a un terzo del tempo totale. (Lin, A. et al., 2023)

Un elemento che influisce sul turnaround time è la programmazione dell'esame e la registrazione del paziente, che rientra nella prima fase del workflow di Risonanza Magnetica. Analisi mostrano infatti che proprio questa fase incide la media di turnover time durante un turno: è la corretta pianificazione degli esami, senza avere slot di prenotazione vuoti e sicuri, senza cancellazioni all'ultimo momento dovute ad errori di programmazione. In questo, il ruolo del Tecnico è coinvolto nella standardizzazione delle tempistiche di svolgimento dei vari e differenti protocolli. (Streit, et al., 2021)

In una prima fase, con il paziente si effettua un controllo di tutti i documenti e moduli di consenso all'esame firmati con dati importanti come la creatinina in caso di un esame con mezzo di contrasto. Poi si procede con la preparazione del paziente che avviene in stanze adiacenti alla sala RM che prevede il cambio abiti, un'ulteriore verifica delle condizioni di sicurezza ed eventuale posizionamento dell'accesso venoso. (Lin A. et al., 2023; Streit et al., 2021)

Acquisizione dell'esame e post-processing

L'acquisizione è il momento in cui il TSRM va a impostare il protocollo e tutti i parametri di acquisizione per ottenere un'immagine diagnostica. Con l'introduzione del Deep Learning Reconstruction non si accelera direttamente l'acquisizione, ma si consente di acquisire dati più rumorosi o sotto campionati e di ricostruire la parte di dati mancante ad una qualità elevata e diagnostica, ottenendo di conseguenza acquisizioni più rapide. (Lin A. et al., 2023)

Refertazione

Il passaggio successivo all'acquisizione e al post-processing è la segmentazione, discussa ampiamente nell'Area zero, con anche, in molti esami, un'analisi volumetrica e funzionale e un'analisi avanzata eseguita dal medico radiologo per la valutazione delle caratteristiche di eterogeneità e il riconoscimento di pattern compatibili con una o più patologie.

Quindi, in questi passaggi, si effettuano una descrizione anatomica, una valutazione quantitativa, misurazioni di diametri e volumi con indici funzionali e un confronto con esami precedenti, per poi arrivare a una conclusione diagnostica, rappresentata dal referto clinico.

Fase	sotto-processo	attività principali	tempo medio
Scheduling e Registration	Prenotazione	Assegnazione Slot, Scelta Protocollo	Variabile
	Arrivo Paziente	Check-in, Verifica Dati	-
	Trasporto Ricoverati	Richiesta e Arrivo Trasporto	22 Min
	Avvio Primo Esame Della Giornata	Preparazione Inizio Giornata	32 Min
Organization & Preparation	Preparazione Paziente	Cambio Abiti, Consenso, Sicurezza	
	Accesso Venoso	Posizionamento Accesso Venoso	Variabile con Pz., 3 Min
	Chiarimenti Clinici	Verifica Protocollo/ Laboratori	Variabile
	Patient Preparation Time	Tempo Totale Prima Ingresso in Sala	19 Min
MR Scan	Scan Preparation Time	Posizionamento Paziente e Bobine	6 Min
	Effective Scan Time	Acquisizione Sequenze	40 Min
	Clean-Up Time	Uscita Paziente, Riassetto Sala	4 Min
	Scan Room Occupation Time	Tempo Totale in Sala RM	50 Min

Tabella n.2 delle fasi del workflow con rispettivi tempi tecnici. (informazioni prese da Streit, U., et al., Analysis of core processes of the MRI workflow for improved capacity utilization, 2021)

Turnover	Intervallo tra Pazienti	Fine Esame - Inizio Successivo	13 Min
-----------------	-------------------------	--------------------------------	--------

Tabella n.3 delle fasi del workflow con rispettivi tempi tecnici. (informazioni prese da Streit, U., et al., Analysis of core processes of the MRI workflow for improved capacity utilization, 2021)

Il turnover è un elemento molto variabile, anche di 20 min in base all'esame o al paziente, ed è un dato importante per riuscire ad inquadrare come e perché il ruolo del tecnico sia cambiato in questi anni.

L'analisi fatta nell'articolo di Streit, (Streit et al., 2021) mostra analizzando i dati, come il focus, organizzativo e di ottimizzazione, si debba spostare al non-gradient-time rispetto al gradient-time già più ottimizzato grazie all'utilizzo del Deep Learning (vedi supra).

Integrazione di algoritmi IA nell'imaging di RM

L'intelligenza artificiale viene utilizzata nel workflow radiologico in diverse fasi del processo.

Tuttavia, tali benefici risultano dipendenti dall'integrazione strutturata dell'IA nel processo organizzativo e dalla qualità dei dati prodotti nelle fasi preliminari (Wichtmann et al., 2026).

Di seguito vengono illustrate le funzioni dell'IA nel contesto della Risonanza Magnetica.

Acquisizione e post-processing.

L'IA in questa fase può andare a migliorare la qualità dell'immagine e quindi va a ridurre i tempi riducendo la necessità di un'acquisizione più lunga.

Come abbiamo già visto nell'Area zero, il post-processing rappresenta uno dei principali ambiti di applicazione dell'IA all'interno del workflow radiologico, consentendo l'analisi automatizzata delle immagini.

Lo studio di (Wichtmann et al., 2026) evidenzia come l'IA possa ottimizzare le fasi di imaging, analisi e reporting attraverso sistemi automatizzati data-driven.

In questa fase l'IA agisce direttamente sull'elaborazione dell'immagine, consentendo analisi qualitative e riducendo le attività manuali. (Wichtmann et al., 2026)

Caso studio sull'impatto dell'uso di sistemi di IA nel workflow tecnico nell'imaging RM

Icobrain è un ottimo esempio di software di intelligenza artificiale utilizzato anche attualmente da vari servizi ospedalieri e tra questi è in uso presso la Neuroradiologia dell'Ospedale Ca' Foncello dell'Azienda ULSS2 Marca Trevigiana.

Il workflow di Icobrain è pensato per integrarsi automaticamente nei processi radiologici, riducendo gli interventi manuali e accelerando l'acquisizione delle immagini.

Partendo dalla Risonanza Magnetica eseguita dal paziente che produce immagini volumetriche, il workflow continua con dati volumetrici in formato DICOM⁹ che vengono inviati automaticamente al RIS¹⁰ per poi essere intercettate da Icobrain.

Icobrain intercetta quindi le immagini RM e le crittografa per garantire maggiore sicurezza; successivamente le invia a Icometrix, la parte di Icobrain che elabora l'immagine.

⁹ DICOM: (Digital Imaging and Communication in Medicine) è uno standard che definisce i criteri per la comunicazione, la visualizzazione, l'archiviazione di immagini radiologiche.

¹⁰ RIS: (Radiology Information System) è un sistema informativo radiologico, è utilizzato per gestire il flusso di dati alfanumerici relativi ai pazienti.

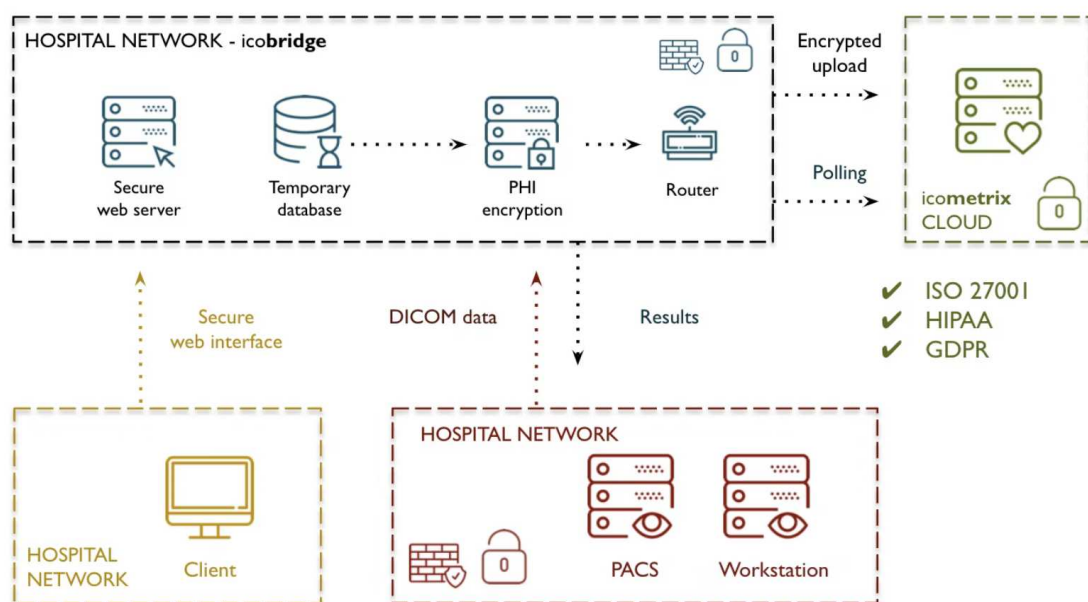


Immagine n.8 (Immagine presa da Icobrain 2026, Clinical Workflow Integration, sito web)

Una volta entrate nel software Icobrain, le scansioni vengono elaborate da algoritmi di IA di Icobrain, che effettuano un'analisi quantitativa automatica, ad esempio dei volumi delle strutture, delle lesioni o di anomalie specifiche; inoltre, valutano la qualità dei risultati con parametri interni di Icobrain.

Il risultato è un report quantitativo, che riporta dati e misurazioni oggettive dei volumi delle strutture cerebrali, e qualitativo visivo, che dà segmentazioni che mostrano visivamente le aree di interesse nelle immagini (vedi area zero).

I report sono utili al medico radiologo, che potrà utilizzare quelle informazioni per scrivere il referto. In questa fase, quindi, l'IA va solo ad affiancare e supportare il medico nelle valutazioni diagnostiche di imaging.

Icobrain in particolare va a fornire un modello di referto già sviluppato che il medico radiologo può utilizzare. (Icobrain 2026, Clinical Workflow Integration, Sito web)

Nella pratica clinica Icobrain è già in uso presso le strutture ospedaliere e nello specifico presso il servizio di Neuroradiologia dell'Ospedale Ca' Foncello di Treviso è stato adottato per lo studio di sclerosi multipla e di decadimento cognitivo. Ho potuto verificare direttamente sul campo, durante la mia esperienza da tirocinante presso tale sede, come avviene l'esecuzione di un'indagine RM che prevede l'utilizzo di Icobrain per la segmentazione automatica delle immagini e come questo induca dei cambiamenti per il TSRM e all'interno del workflow. Qui sono presenti due scanner RM della

ditta Siemens, che sono Siemens Magnetom Vida con intensità di campo magnetico pari a 3 tesla e Siemens Magnetom Sola con intensità di campo magnetico pari a 1.5 tesla.

Come riportato nell'area 1, Icobrain necessita, per un corretto e ottimale funzionamento, di immagini di qualità elevata e di parametri di acquisizione standardizzati al fine di ottenere uniformità nei risultati. Quelli che seguono in tabella sono i parametri di acquisizione delle sequenze T1 e Flair adottati dal servizio di Neuroradiologia dell'Ospedale Ca' Foncello che consentono la riproducibilità e la standardizzazione dei risultati di output per il corretto funzionamento del software di analisi di Icobrain.

	FLAIR in RM 1.5T	FLAIR in RM 3T	T1 mprage 1.5T	T1 mprage 3T
Slab group	1	1	1	1
Slabs	1	1	1	1
Slices per slab	176	176	176	176
Distance Factor			50	50
Position	isocenter	isocenter	isocenter	isocenter
Orientation	Sagittal	Sagittal	sagittal	sagittal
Phase encoding direction	Antero-posteriore	Antero-posteriore	Antero-posteriore	Antero-posteriore
Phase oversampling	0	0	0	0
Slice oversampling	0%	12,5%	0	0
FOV read	245 mm	250 mm	256 mm	256 mm
FOV phase	100	100	100%	100%
Slice thickness	1 mm	1 mm	1 mm	1 mm
TR	5500 ms	5500 ms	2300 ms	2300 ms
TE	415 ms	377 ms	2,98 ms	2,98 ms
Averages	2,0	1,8	1	1
Concatenations	1	1	1	1
Tempo	5:47 min	5:32 min	4:18 min	5:21 min
Voxel	1.0x1.0x1.0 mm ³	1.0x1.0x0.6 mm ³	1.0x1.0x1.0 mm ³	1.0x1.0x1.0 mm ³

Tabella n.4, mostra i parametri di acquisizione utilizzati presso la Neuroradiologia dell'ospedale Ca' Foncello di Treviso per imaging di RM con Icobrain.

Il TSRM deve conoscere i parametri necessari al corretto funzionamento di Icobrain e pertanto deve attenersi ai protocolli di acquisizione impostati e deve soprattutto mantenere lo spessore di acquisizione pari a 1 mm, pertanto deve essere consapevole di non poter effettuare modifiche eccessive agli altri parametri di scansione quali FOV read e FOV phase. Si può notare inoltre come

i tempi di acquisizione delle sequenze siano lunghi, per cui la preparazione del paziente all'esame diventa fondamentale.

Di seguito sono descritte le fasi del workflow per l'esecuzione di una RM cerebrale per studio di sclerosi multipla e di decadimento cognitivo che ho osservato durante il tirocinio presso il servizio di Neuroradiologia di Treviso:

1. programmazione e registrazione: al momento della prenotazione dell'esame, viene riservato uno slot di tempo di 40 minuti. I controlli di sclerosi multipla sono ripetuti nello stesso scanner per ridurre al minimo la variabilità dei risultati;
2. preparazione del paziente: il paziente arriva in diagnostica e avviene il controllo della documentazione in sala di preparazione RM che prevede la verifica del quesito clinico dell'esame, del questionario per l'idoneità all'esecuzione dell'esame, all'accesso in sala magnete e all'eventuale somministrazione del mezzo di contrasto, della documentazione radiologica precedente, e la spiegazione del funzionamento dell'esame. Solo dopo tali verifiche, il paziente indossa un camice fornito dal servizio per evitare la presenza di corpi metallici nascosti negli indumenti personali, rimuove gli oggetti metallici e avviene il posizionamento dell'accesso venoso se necessario;
3. acquisizione esame RM: il paziente viene posizionato in sala magnete in modo da garantirne il massimo confort (nel caso di esame nella RM Magnetom Vida viene scelta la bobina "Head Coil" più adeguata alla fisionomia del paziente, essendo il sistema provvisto di una bobina da 64 canali fissa utilizzata maggiormente per la miglior qualità di immagini, e di una da 20 canali che consente invece il sollevamento della testa del paziente nel caso esso sia particolarmente cifotico o con difficoltà respiratorie). Questi accorgimenti permettono una minor probabilità di movimento durante l'esame e quindi un miglior output di risultato. Icobrain si utilizza in studi di sclerosi multipla o di demenza e si tratta di pazienti non sempre collaboranti per cui vengono acquisite, compatibilmente con il protocollo di studio, prima le sequenze fondamentali per il funzionamento dell'IA (T1 Mprage e Flair). Il TSRM procede poi al controllo delle immagini, momento di massima importanza in quanto per evitare errori di analisi di Icobrain non ci devono essere artefatti da movimento ed altri che possono inficiare sulla qualità del dato in uscita. Ad esame terminato il TSRM procede all'archiviazione dell'esame a PACS ed invio solo delle sequenze utili al software Icobrain.

Le fasi del workflow rispecchiano quelle viste precedentemente e riportate in letteratura da Streit nello studio del 2021. Si evince che è fondamentale la corretta preparazione del paziente all'esame rm per evitare perdite di tempo una volta iniziato lo studio e come sia il TSRM la figura con cui si

interfaccia principalmente il paziente e da cui dipende in gran misura la buona riuscita dell'esame e il corretto funzionamento dell'IA.

Ruolo del Tecnico Sanitario di Radiologia: cambiamenti in seguito all'introduzione della segmentazione automatica.

L'integrazione dell'IA nella Risonanza Magnetica ha trasformato il workflow e, con esso, ha portato cambiamenti diretti e indiretti nel lavoro del TSRM che, sempre di più, si trova a doversi adattare a essi.

Verifica del corretto invio delle immagini

In un contesto senza IA, il tecnico conclude l'esame con la corretta archiviazione nel PACS delle immagini radiologiche; con la segmentazione automatica, invece, si aggiunge un passaggio tra l'acquisizione e la refertazione, che consiste nell'invio corretto delle immagini prima al motore IA per l'elaborazione e poi al PACS per l'archiviazione.

Il ruolo del tecnico in questo campo è quello definito orchestration engine, cioè colui che ha il compito di instradare le immagini all'IA, selezionare le serie corrette che devono e non devono essere modificate dall'IA e quindi attivare l'algoritmo. (Korfiatis et al., 2025)

Da un punto di vista operativo, questo fa sì che non sia più sufficiente che l'esame sia solo svolto correttamente, ma che risponda anche a quelle caratteristiche dell'immagine necessarie al corretto funzionamento dell'IA (vedi Area 1 - Qualità dell'immagine ideale per IA).

Occorre quindi verificare, da parte del tecnico, che il sistema abbia inviato ed elaborato correttamente le immagini dell'esame, aggiungendo così un compito fondamentale all'interno del workflow di segmentazione automatica.

Per riconoscere eventuali anomalie nel flusso di dati durante il controllo, il tecnico deve anche conoscere il workflow delle immagini, che integra la segmentazione automatica, consentendogli di acquisire una maggiore padronanza dei dati e di ridurre gli errori.

In pratica, nel processo di acquisizione, questo punto non apporta grandi cambiamenti, ma aggiunge responsabilità al tecnico di radiologia. (Wichtmann et al., 2026)

La selezione corretta delle serie da inviare all'algoritmo presuppone la consapevolezza delle caratteristiche tecniche necessarie al corretto funzionamento del sistema (vedi Area 1)

Controllo dell'output

Si aggiunge anche una fase di post-production monitoring, che si concentra sulla necessità continua di un controllo delle performance, degli errori sistematici, della sicurezza, della manutenzione degli algoritmi e di una supervisione costante. (Korfiatis et al. 2025)

Quindi, dal punto di vista tecnico, ciò comporta la necessità di una supervisione non solo del flusso dei dati, ma anche del controllo della qualità delle immagini stesse e, soprattutto, delle prestazioni dell'IA nel tempo, con la conseguente funzione di segnalazione di problemi o cali di prestazioni.

Ciò influisce anche sul rapporto tra il TSRM e il medico radiologo, aumentando la comunicazione tra le due figure e favorendo il confronto sul corretto funzionamento del flusso di dati tra algoritmo e IA. (Wichtmann et al. 2026)

Impatti nell'assetto organizzativo.

La segmentazione automatica va ad aggiungere come abbiamo già visto la verifica dell'output. Ciò ricade sul TSRM che ora avrà un elemento in più da controllare e che potrebbe andare ad influenzare le tempistiche dell'interno della procedura d'esame, che quindi esige un'attenzione e organizzazione più solida da parte della singola figura del TSRM ed inoltre una maggior attività del coordinatore tecnico che ora dovrà anche implementare nel modello organizzativo un passaggio aggiuntivo e tenere conto del coordinamento tra sistemi e professionisti sanitari che diventa sempre più complesso. (Wenderott et al. 2024, Korfiatis et al. 2025)

Inoltre ciò comporta che il personale per essere efficiente ed efficace abbia bisogno di una formazione continua per poter lavorare al meglio, garantendo così, il funzionamento del workflow. (Lin, A., Et al., 2023)

Acquisizione data-driven.

La segmentazione automatica influisce anche sulla parte di acquisizione perché, come già visto nell'Area 1, l'immagine per essere processata correttamente dall'algoritmo deve rispettare determinate caratteristiche.

Ciò implica che le prestazioni degli algoritmi di Deep Learning siano strettamente dipendenti dalla coerenza e dalla qualità dei dati di input. (Litjens et al., 2017; Tajbakhsh et al., 2020)

In particolare, elementi come il SNR e CNR, l'assenza di artefatti significativi, la coerenza dei parametri di acquisizione (TR, TE, FOV, spessore di strato) e l'uniformità di ricostruzione, elementi

(già discussi nell'Area 1) rappresentano condizioni essenziali per garantire stabilità e affidabilità dei dati di output. (Coriasco et al., 2014; Litjens et al., 2017; Tajbakhsh et al., 2020)

In questo contesto, il ruolo del Tecnico Sanitario di Radiologia Medica assume una dimensione ulteriormente strategica: la qualità tecnica dell'immagine non è più finalizzata esclusivamente alla valutazione diagnostica del medico radiologo, ma diventa determinante anche per l'elaborazione automatizzata. L'acquisizione può quindi essere definita *data-driven*, poiché orientata non solo alla produzione di immagini diagnostiche, ma alla generazione di dati strutturati, coerenti e processabili dagli algoritmi. Nel lavoro del tecnico, la qualità tecnica dell'immagine e la rilevazione e l'accorgersi di eventuali artefatti nell'immagine che possano influenzare direttamente l'output diventano quindi ancora più importanti.

Inoltre, la gestione dei dataset annotati e la consapevolezza dei possibili errori sistematici o della variabilità interpretativa dei sistemi di IA richiedono una supervisione continua e una validazione percettiva da parte del professionista, al fine di mantenere elevata l'affidabilità complessiva del sistema. (Wichtmann et al., 2026; Korfiatis et al., 2025)

L'integrazione dell'IA non riduce dunque la responsabilità tecnica nella fase di acquisizione, ma ne amplia la portata, trasformando il TSRM in garante della qualità del dato destinato all'analisi automatizzata.

L'analisi dell'integrazione degli algoritmi di Intelligenza Artificiale nel workflow della Risonanza Magnetica evidenzia come l'innovazione tecnologica non si limiti all'introduzione di strumenti di supporto diagnostico, ma comporti una riorganizzazione strutturale del processo operativo.

Dalla richiesta clinica alla refertazione, l'IA interviene in più fasi del workflow, contribuendo alla riduzione dei tempi di acquisizione, all'ottimizzazione del turnaround time e all'automatizzazione del post-processing. Tuttavia, tali benefici risultano strettamente dipendenti dalla qualità e dalla standardizzazione dei dati prodotti nelle fasi preliminari.

Il focus organizzativo, già orientato alla gestione del non-gradient time (Lin et al., 2024; Streit et al., 2021), si estende oggi al controllo del flusso dei dati, alla supervisione degli output algoritmici e alla validazione delle performance nel tempo. (Korfiatis et al., 2025; Wichtmann et al., 2026)

In questo scenario, il Tecnico Sanitario di Radiologia Medica non assume un ruolo marginale rispetto all'automazione, ma diventa figura centrale nell'orchestrazione del sistema: responsabile della qualità del dato di input, della corretta integrazione dell'algoritmo nel workflow e della verifica dell'affidabilità dell'output.

L'Area 3 analizzerà quindi in modo approfondito come questa evoluzione tecnologica stia ridefinendo competenze, responsabilità e identità professionale del TSRM nell'era dell'imaging data-driven.

Area Tre – Validazione percettiva delle immagini di RM da parte del TSRM e analisi di caso studio sperimentale

La letteratura evidenzia come modelli con performance numericamente elevate possano comunque generare segmentazioni morfologicamente incoerenti o scarsamente plausibili dal punto di vista visivo, in particolare in presenza di rumore, artefatti o variazioni dei parametri di acquisizione. (Renard et al., 2020)

Nel contesto della Risonanza Magnetica, caratterizzato da margini sfumati, eterogeneità del segnale e variabilità inter sequenziale, la componente percettiva assume un ruolo centrale. Studi sulla variabilità inter- e intra-osservatore dimostrano che anche la segmentazione manuale è influenzata dalla qualità dell'immagine e dall'interpretazione visiva dell'operatore. (Monereo-Sánchez et al., 2021; Veiga-Canuto et al., 2022)

La validazione percettiva dell'output dell'intelligenza artificiale si configura pertanto come un passaggio necessario nel workflow tecnico: non come atto diagnostico, ma come controllo qualitativo volto a verificare coerenza anatomica e plausibilità morfologica della segmentazione rispetto all'immagine RM originale.

La percezione visiva del Tecnico Sanitario di Radiologia Medica rappresenta un elemento integrativo alle metriche quantitative, contribuendo a garantire l'affidabilità complessiva del processo di segmentazione automatica.

Comportamento dei software di segmentazione automatica su immagini con artefatti

Come detto nel sottocapitolo 1.5, per poter funzionare correttamente gli algoritmi di IA necessitano di immagini di ottima qualità. L'obiettivo di questo paragrafo è quello di studiare il comportamento di software di segmentazione automatica in presenza di immagini non standard e di valutare che effetti possa avere la mancanza di validazione da parte di una figura esperta come quella del TSRM medica sui dati di imaging (che quindi porterebbe ad un dataset di dati di input con artefatti).

Lo studio proposto è basato principalmente sullo studio di Z. Kuang et al., 2025, che mostra come cambia il comportamento degli algoritmi di segmentazione automatica in base a imaging di RM con diversi tipi di artefatti e pesature, utilizzando come metro di misura il DSC.

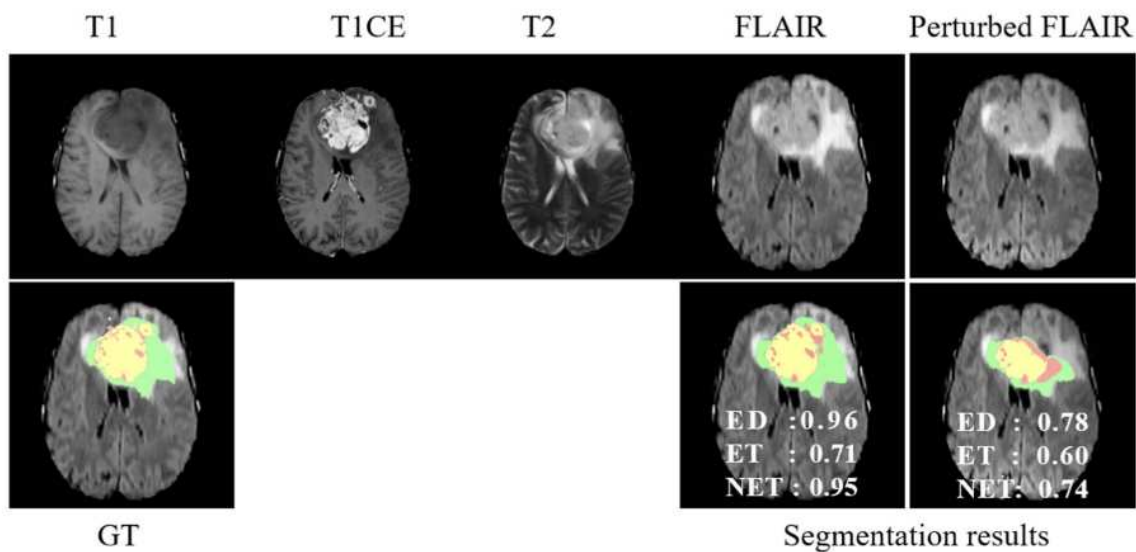


Immagine n.9 (Z. Kuang, Z. Yan, A. Abayazeed, F. Wagner, L. Yu and M. Reyes Robust Cross-Sequence Semantic Interaction for Brain Tumor Segmentation on Multi-Sequence MR Images 2025)

Lo studio è stato fatto andando ad acquisire un RM multi-sequenza del distretto cerebrale, con diverse pesature, per poi confrontare l'immagine segmentata GT (Ground Truth) e le immagini "segmentation result" in cui è stata svolta una segmentazione tramite un modello U-Net, con la presenza di artefatti e senza.

Sono state scelte immagini di RM cerebrale perché particolarmente sensibile ad artefatti.

L'immagine n.9 mostra come in presenza di perturbazioni/artefatti, con il software basato sul modello U-Net si vada ad avere una grande differenza nella segmentazione che in questo caso non riconosce il lobo frontale sinistro come patologico.

Nell'immagine si possono vedere le diverse pesature della stessa sezione quali T1, T1CE (Contrast Enhancement), T2 e FLAIR. L'immagine segmentata in basso a sinistra nell'immagine 9 (GT) è l'immagine ground-truth, che è stata segmentata manualmente ed utilizzata come base dello studio e del confronto tra le immagini.

Le sequenze utilizzate sono state T1, T1CE, T2 e FLAIR, che sono sequenze standard per un esame RM cerebrale, perché evidenziano diverse componenti della lesione e nello specifico:

- T1: fornisce informazioni anatomiche e consente di visualizzare la struttura generale del cervello e della lesione. Il tumore appare generalmente ipo-intenso rispetto al parenchima sano;
- T1CE: evidenzia la porzione di tumore con enhancement, va quindi a mostrare la parte attiva e vascolarizzata della lesione in cui arriva il mezzo di contrasto;

- T2: principalmente per la visione di edema peritumorale;
- FLAIR: sopprime il segnale del liquido cerebrospinale e consente di visualizzare in modo più chiaro l'edema peritumorale (ED) e l'infiltrazione tumorale nel parenchima circostante.

Possiamo quindi vedere un buon risultato nella prima immagine dei segmentation results fatta sull'immagine con pesatura FLAIR, ottenendo valori alti di DSC.

Nelle immagini segmentate sia in modo automatico sia nell'immagine ground-truth la segmentazione ha colori diversi per indicare diverse componenti del tumore: il giallo indica la zona tumorale con la presenza di mezzo di contrasto, la zona verde indica l'edema peritumorale e in rosso il core necrotico o non captante.

Possiamo notare una notevole differenza tra i due risultati della segmentazione automatica: il primo senza artefatti presenta un'area segmentata nettamente maggiore rispetto all'immagine con l'artefatto, e se poi confrontate con il ground truth possiamo concludere che la porzione tumorale più anteriore non è stata riconosciuta come patologica nelle immagini artefatte. Si può quindi dire che la segmentazione con artefatti non è affatto valida andando anche a mostrare zone di necrosi più ampie rispetto a altre due. Ciò è anche confermato dall'analisi matematica del DSC delle diverse zone tumorali, le quali indicano un netto calo di performance nell'immagine artefatta passando da valori alti nell'immagine senza artefatto:

$ED(\text{Edema}) = 0.96,$

$ET(\text{Enhancing Tumors}) = 0.71,$

$NET(\text{Necrotic and Non-Enhancing Tumor Cores}) = 0.95,$

a valori molto più bassi:

$ED(\text{Edema}) = 0.78,$

$ET(\text{Enhancing Tumors}) = 0.60,$

$NET(\text{Necrotic and Non-Enhancing Tumor Cores}) = 0.74$

Una riduzione di prestazione media del 19%

(Z. Kuang et al., 2025)

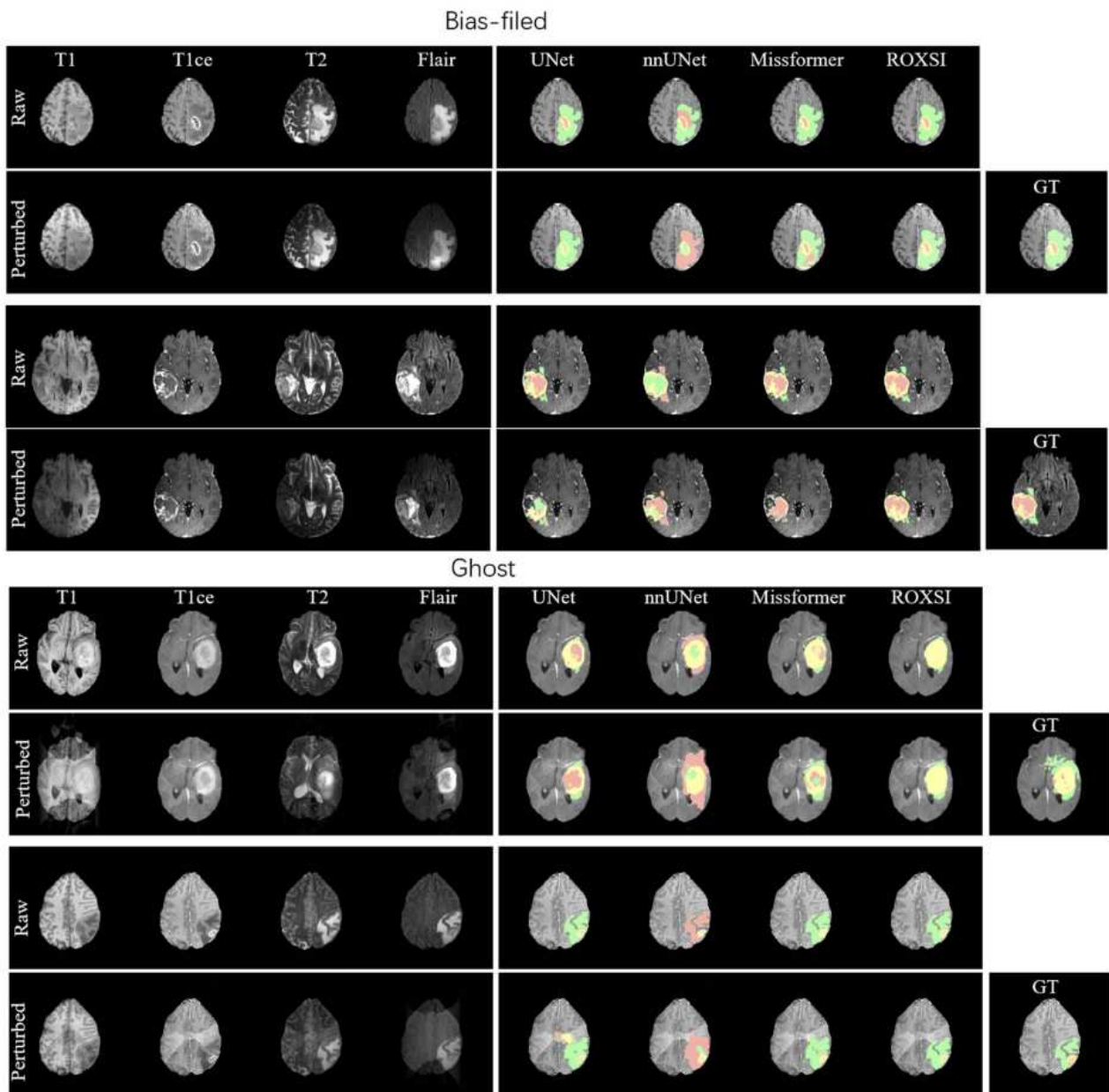


Immagine n.10 (Z. Kuang, et al. Robust Cross-Sequence Semantic Interaction for Brain Tumor Segmentation on Multi-Sequence MR Images, 2025)

Nell'immagine n.10 sono presenti quattro blocchi principali che rappresentano diversi tipi di artefatto nelle immagini RM, i quali sono bias-field, ghost, spike e motion.

Quindi si sono simulate le tipologie di artefatti realistici per valutare come si comportino le segmentazioni automatiche di diversi modelli di segmentazione automatica.

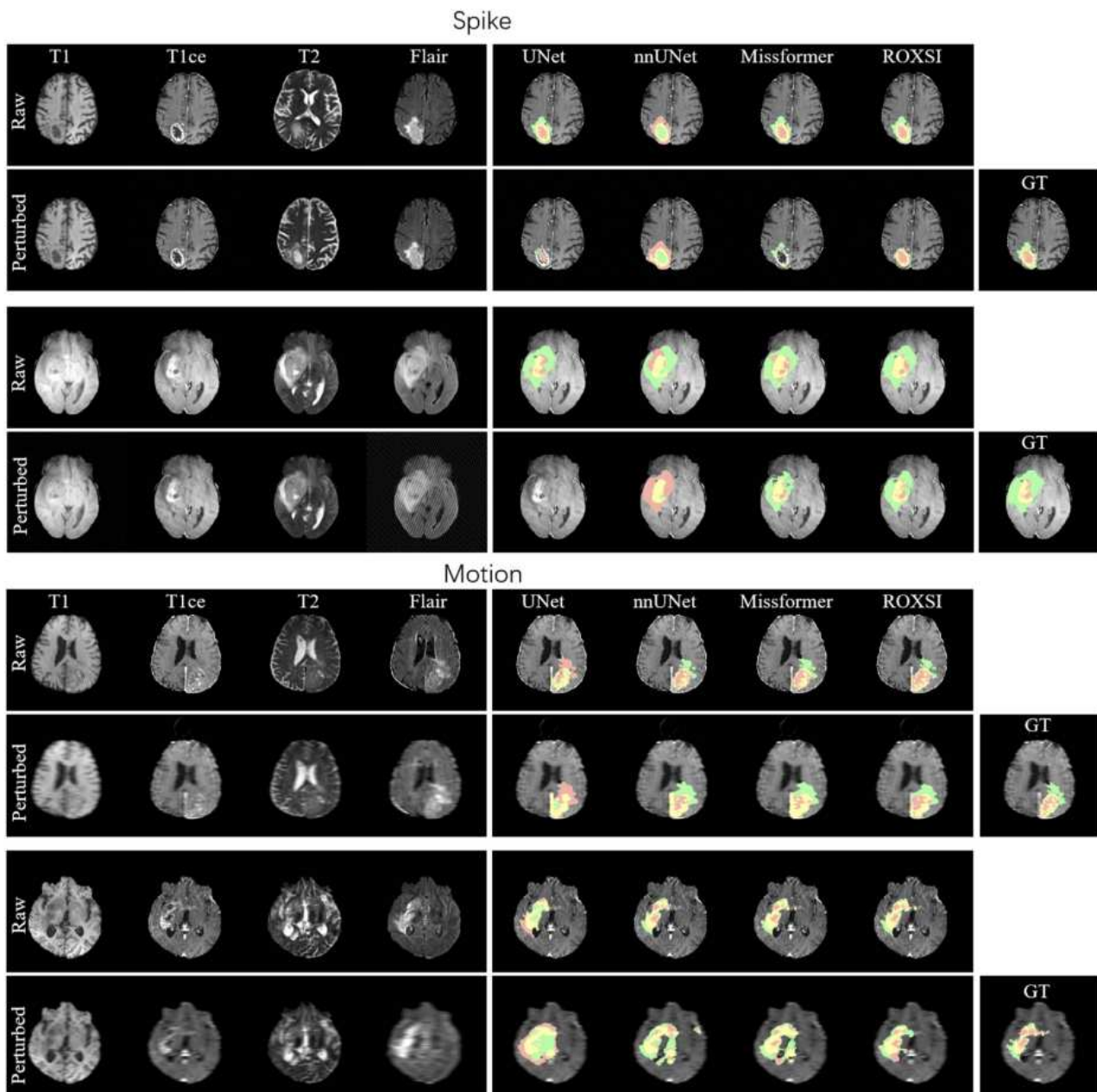


Immagine n.11 (Z. Kuang, et al. Robust Cross-Sequence Semantic Interaction for Brain Tumor Segmentation on Multi-Sequence MR Images, 2025)

Nelle immagini n.10 e n.11, ogni blocco contiene due righe, la prima “Raw” ha le immagini RM nelle diverse pesature originali e la seconda riga “Perturbed” ha le immagini con l’artefatto.

Nelle colonne U-Net, nnU-Net, Missformer, ROXSI sono rappresentate rispettivamente le immagini segmentate da un tipo di segmentazione automatica, quindi questa tabella va a comparare diversi modelli di segmentazione automatica mettendoli a confronto con l’immagine GT.

Le immagini segmentate con i modelli U-Net e nnU-Net hanno aree segmentate complessivamente simili al ground truth, ma con diverse imprecisioni per quanto riguarda la caratterizzazione dei

diversi tessuti tumorali, ciò è accentuato dalla presenza di artefatti come possiamo vedere comparando le immagini “raw” e “perturbed” con l’utilizzo del modello nnU-Net in tutti i tipi di artefatto.

Con l’utilizzo dei modelli Missformer e ROXSI si possono vedere dei netti miglioramenti nella segmentazione e nella caratterizzazione dei diversi tessuti tumorali, garantendo, anche con la presenza di artefatti, un risultato molto più simile all’immagine ground truth.

Nel caso del modello Missformer l’architettura è basata su un modello transformer (vedi paragrafo Transformer e modelli ibridi), mentre ROXSI è un software che è stato progettato specificatamente per le segmentazioni in RM con artefatti. La differenza è quindi che va ad analizzare ogni sequenza con pesatura diversa separatamente, una per una, andando così a compensare il problema degli artefatti, che se presenti in anche una sola immagine potevano andare ad inficiare l’intera segmentazione.

Questa analisi ci fa capire che il ruolo del TSRM rimane fondamentale nel processo di acquisizione dell’immagine e validazione percettiva.

Gli algoritmi di Deep Learning hanno mostrato che quando applicati ad un contesto clinico reale dove possono esserci artefatti le prestazioni dei modelli possono diminuire in modo significativo, dimostrando inoltre la sensibilità di questi modelli con degradazione dell’immagine.

In questo contesto quindi la figura del TSRM è essenziale in quanto operatore qualificato in grado di riconoscere l’artefatto e se necessario andare a ripetere la sequenza per avere un risultato ottimale e senza artefatti.

Il principale vantaggio dell’osservazione umana rispetto agli algoritmi di segmentazione è quello di avere una visione globale dell’immagine. Il TSRM non analizza solamente singole strutture, ma interpreta l’intero contesto anatomico e la coerenza complessiva dell’immagine.

Molti algoritmi di segmentazione si concentrano su pattern appresi da database locali ristretti che possono quindi essere meno solidi quando l’immagine presenta alterazioni non previste, come artefatti.

CONCLUSIONI

Da questa revisione si è concluso che l'utilizzo della segmentazione automatica è una delle innovazioni più importanti nell'ambito radiologico in RM.

Essa offre vantaggi principalmente in termini di tempo, non solo nella parte di acquisizione, ma anche in quella di post-elaborazione, andando a fornire un supporto ai medici radiologi, i quali possono già usufruire di software, come Icobrain, per avere supporto alla refertazione.

È emerso inoltre che anche il ruolo del TSRM cambia molto con l'introduzione dell'IA, focalizzandosi maggiormente su controllo, sicurezza e validazione del processo, andando ad aumentare la responsabilità di questa professione.

Si può concludere affermando che la segmentazione automatica non sostituisce la valutazione umana, dal momento che l'IA modifica i parametri di acquisizione, ottenendo un miglior bilanciamento tra velocità e qualità, ma richiede una collaborazione attiva tra tecnologia e competenze del personale tecnico, al fine di massimizzare l'accuratezza diagnostica e l'efficienza del processo clinico. Inoltre l'integrazione di sistemi come Icobrain nel workflow di imaging deve essere accompagnata da formazione specifica e linee guida per la standardizzazione delle acquisizioni, garantendo che l'IA operi su dati adeguati e produca risultati riproducibili e clinicamente utili.

La segmentazione automatica quindi rappresenta non solo un'innovazione tecnologica, ma un elemento di evoluzione del workflow radiologico, che ha però bisogno tempo per avere uno studio accurato per una corretta introduzione nell'ambiente clinico.

LIMITI DELLO STUDIO E SVILUPPI FUTURI

Lo studio di revisione presenta alcune limitazioni, le quali potranno essere sviluppate in futuro per migliorare o ridefinire la tesi.

In particolare per lo studio della percezione ed il confronto tra immagini sarebbe stato utile non basarsi su studi già fatti, ma elaborarlo separatamente creando un database di immagini open source¹¹ o meno, con il quale poi, tramite i software IA di Deep Learning e segmentazione, andare a segmentare delle immagini con la presenza o meno di artefatti.

Questo studio non è stato possibile per più ragioni, partendo dalla forza computazionale necessaria per il processo di creazione ed etichettatura del database che indipendentemente dalle dimensioni e quantità delle immagini, avrebbe richiesto molte ore, se non giorni di elaborazione da parte del computer che deve essere potente ed equipaggiato con la tecnologia cuda core, presente solo nelle schede video Nvidia, della quale il mio computer non è provvisto, inoltre eliminando la possibilità di utilizzo di computer equipaggiati con sistemi operativi diversi da linux o windows.

Un'altro limite trovato è la conoscenza informatica necessaria per la creazione del database e l'installazione tramite prompt di tutte le varie applicazioni utili per l'elaborazione e comparazione delle immagini.

La soluzione potrebbe essere quella di utilizzare un database pre-addestrato presente anche in varie versioni open source, ad esempio il dataset di BraTs, ma in qualsiasi caso si ha bisogno delle caratteristiche hardware sopra elencate.

Lo studio può inoltre essere ampliato con l'integrazione di scale percettive che vadano ad analizzare in modo sistematico la percezione della qualità d'immagine oppure con il coinvolgimento sul campo di un TSRM per la validazione ed analisi delle immagini.

¹¹ La scelta dell'utilizzo di un database open source sarebbe dettata dalla difficoltà di reperimento delle immagini dai singoli pazienti, che secondo il GDPR 2016/679 devono firmare il modulo per la delibera all'utilizzo delle immagini. La gestione dei moduli e la quantità di dati che dovrebbero essere archiviati, per 10 anni secondo il codice civile art.2220, sarebbe quindi molto complessa per la creazione del database.

BIBLIOGRAFIA E SITOGRAFIA

- Yousef, R., Khan, S., Gupta, G., Siddiqui, T., Albahlal, B. M., Alajlan, S. A., & Haq, M. A. (2023). U-Net-based models towards optimal MR brain image segmentation. *Sensors*, 23(9), 4275. <https://doi.org/10.3390/s23094275>
- Novosad P, Fonov V, Collins DL; Alzheimer's Disease Neuroimaging Initiative†. Accurate and robust segmentation of neuroanatomy in T1-weighted MRI by combining spatial priors with deep convolutional neural networks. *Hum Brain Mapp*. 2020 Feb 1;41(2):309-327. doi: 10.1002/hbm.24803. Epub 2019 Oct 21. PMID: 31633863; PMCID: PMC7267949
- Wang R., T. Lei, R. Cui, B. Zhang, H. Meng, and A. K. Nandi, “Medical image segmentation using Deep Learning: A survey,” *IET Image Processing*, Jan. 2022, doi: <https://doi.org/10.1049/ipr2.12419>
- Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M.-A., Prastawa, M., Reyes, M., & Van Leemput, K. (2015). The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 34(10), 1993–2024. <https://doi.org/10.1109/TMI.2014.2377694>
- Yao, A. D., Cheng, D. L., Pan, I., & Kitamura, F. (2020). Deep Learning in neuroradiology: A systematic review of current algorithms and approaches for the new wave of imaging technology. *Radiology: Artificial Intelligence*, 2(2), e190026. <https://doi.org/10.1148/ryai.2020190026>
- Gruber B, Froeling M, Leiner T, Klomp DWJ. RF coils: A practical guide for nonphysicists. *J Magn Reson Imaging*. 2018 Jun 13;48(3):590–604. doi: 10.1002/jmri.26187. Epub ahead of print. PMID: 29897651; PMCID: PMC6175221
- Moore C, Bell D, Dice similarity coefficient. Reference article, *Radiopaedia.org* <https://doi.org/10.53347/rID-75056>
- Faix, A., 2025, Consensus: Using AI to Analyze Scientific Literature. *Library Trends*, 344-354. <https://dx.doi.org/10.1353/lib.2025.a961198>
- Van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L., (2023). Five priorities for research. *Nature*, 614, 224–226. <https://doi.org/10.1038/d41586-023-00288-7>

Area zero

- Najjar R. (2023). Redefining Radiology: A Review of Artificial Intelligence Integration in Medical Imaging. *Diagnostics* (Basel, Switzerland), 13(17), 2760. <https://doi.org/10.3390/diagnostics13172760>
- Veiga-Canuto, D., et al. (2023). Artificial intelligence–based segmentation in oncologic imaging: Current status and future perspectives. *Cancers*, 15(10), 2573. <https://doi.org/10.3390/cancers15102573>
- Boeken, T., Feydy, J., Lecler, A., Soyer, P., Feydy, A., Barat, M., & Duron, L. (2023). Artificial intelligence in diagnostic and interventional radiology: Where are we now?. *Diagnostic and interventional imaging*, 104(1), 1–5. <https://doi.org/10.1016/j.diii.2022.11.004>
- Veiga-Canuto, D., et al. (2022). Comparative multicentric evaluation of interobserver variability in manual versus Deep Learning–based tumor segmentation. *Cancers*, 14(15), 3648. <https://doi.org/10.3390/cancers14153648>
- Monereo-Sánchez, J., et al. (2021). Quality control strategies for brain MRI segmentation: Impact on variability and reproducibility. *Journal of Magnetic Resonance Imaging*, 61(5), 1353–1365. <https://doi.org/10.1002/jmri.1353>
- Dautkulova, A., Aider, O. A., Teulière, C., Coste, J., Chaix, R., Ouachik, O., Pereira, B., & Lemaire, J. J. (2025). Automated segmentation of deep brain structures from Inversion-Recovery MRI. *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, 120, 102488. <https://doi.org/10.1016/j.compmedimag.2024.102488>
- Morandin L., BRAIN TUMOR SEGMENTATION (BraTS) CHALLENGE: ASPETTI IMPLEMENTATIVI DEGLI ALGORITMI DI SEGMENTAZIONE E DI PREDIZIONE, 2023, <https://hdl.handle.net/20.500.12608/48847>
- Novosad, P., Fonov, V., Collins, D. L., & Alzheimer's Disease Neuroimaging Initiative† (2020). Accurate and robust segmentation of neuroanatomy in T1-weighted MRI by combining spatial priors with deep convolutional neural networks. *Human brain mapping*, 41(2), 309–327. <https://doi.org/10.1002/hbm.24803>
- Lerch, T. D., Zeng, G., Boschung, A., Meier, M. K., Steppacher, S. D., Tannast, M., Heverhagen, J., & Schmaranzer, F. (2025). Magnetic resonance imaging–based three-dimensional models of the pelvis and hip using machine learning for automatic bone

- segmentation in a dynamic hip impingement simulation. *Orthopaedic Journal of Sports Medicine*. <https://doi.org/10.1177/23259671251334138>
- Wang, S., Zhou, M., Liu, Z., Liu, Z., Gu, D., Zang, Y., Dong, D., Gevaert, O., & Tian, J. (2022). Medical image segmentation using Deep Learning: A survey. <https://doi.org/10.1049/ipr2.12419>
 - Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., ... Van Leemput, K. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BraTS). <https://doi.org/10.1109/TMI.2014.2377694>
 - Renard, F., Guedria, S., De Palma, N., & Vuillerme, N. (2020). Variability and reproducibility in Deep Learning for medical image segmentation. *Scientific Reports*, 10. <https://doi.org/10.1038/s41598-020-69920-0>
 - Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on Deep Learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
 - Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Proceedings of MICCAI* (pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28
 - Milletari, F., Navab, N., & Ahmadi, S. A. (2016). V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)* (pp. 565–571). IEEE. <https://doi.org/10.1109/3DV.2016.79>
 - Yousef, R., Khan, S., Gupta, G., Siddiqui, T., Albahlal, B. M., Alajlan, S. A., & Haq, M. A. (2023). U-Net-based models towards optimal MR brain image segmentation. <https://doi.org/10.3390/diagnostics13091624>
 - Simarro, J., Meyer, M. I., Van Eyndhoven, S., Phan, T. V., Billiet, T., Sima, D. M., & Ortibus, E. (2024). A Deep Learning model for brain segmentation across pediatric and adult populations. *Scientific Reports*, 14, 11735
 - Isensee F., Jaeger P., Kohl S., Petersen J., Maier-Hein K. (2021). nnU-Net: a self-configuring method for Deep Learning-based biomedical image segmentation. *Nature Methods*
 - Liu, Z., Kainth, K., Zhou, A., Deyer, T. W., Fayad, Z. A., Greenspan, H., & Mei, X. (2024). A review of self-supervised, generative, and few-shot Deep Learning methods for

- data-limited magnetic resonance imaging segmentation. *NMR in Biomedicine*.
<https://doi.org/10.1002/nbm.5143>
- Jiang, J., Tyagi, N., Tringale, K., Crane, C., & Veeraraghavan, H. (2022). Self-supervised 3D anatomy segmentation using self-distilled masked image transformer. <https://arxiv.org/abs/2205.10342>
 - Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE TPAMI*, 40(4), 834–848. <https://doi.org/10.1109/TPAMI.2017.2699184>
 - Cheplygina, V., de Bruijne, M., & Pluim, J. P. W. (2019). Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. *Medical Image Analysis*, 54, 280–296. <https://doi.org/10.1016/j.media.2019.03.009>
 - Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B., & Rueckert, D. (2018). Attention U-Net: Learning where to look for the pancreas. *Medical Image Analysis*, 53, 193–203. <https://doi.org/10.1016/j.media.2019.01.012>
 - Shen, D., Wu, G., & Suk, H.-I. (2017). Deep Learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>

Area 1

- Graves MJ, Mitchell DG. Body MRI artifacts in clinical practice: a physicist's and radiologist's perspective. *J Magn Reson Imaging*. 2013 Aug;38(2):269-87. doi: 10.1002/jmri.24288. PMID: 23960007
- Coriasco, M., Rampado, O., & Bradac, G. B. (2014). *Elementi di Risonanza Magnetica: Dal protone alle sequenze per le principali applicazioni diagnostiche*. Springer Italia. <https://doi.org/10.1007/978-88-470-5641-1>
- Bushberg, J. T., Seibert, J. A., Leidholdt, E. M., & Boone, J. M. (2012). *The essential physics of medical imaging (3rd ed.)*. Wiley-Blackwell
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciampi, F., Ghahfarooian, M., van der Laak, J. A., van Ginneken, B., & Sánchez, C. I. (2017). A survey on Deep Learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>

- Tajbakhsh N, Jeyaseelan L, Li Q, Chiang JN, Wu Z, Ding X. Embracing imperfect datasets: A review of Deep Learning solutions for medical image segmentation. *Med Image Anal.* 2020;63:101693
- Fortin JP, et al. Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage.* 2017;167:104–120
- Renard, F., Guedria, S., De Palma, N., & Vuillerme, N. (2020). Variability and reproducibility in Deep Learning for medical image segmentation. *Scientific Reports.* <https://doi.org/10.1038/s41598-020-69920-0>
- Z. Kuang, Z. Yan, A. Abayazeed, F. Wagner, L. Yu and M. Reyes, ROXSI: Robust Cross-Sequence Semantic Interaction for Brain Tumor Segmentation on Multi-Sequence MR Images, *IEEE Journal of Biomedical and Health Informatics*, vol. 29, no. 4, April 2025. doi: 10.1109/JBHI.2024.3513479
- Monereo-Sánchez, J., et al. (2021). Quality control strategies for brain MRI segmentation: Impact on variability and reproducibility. *Journal of Magnetic Resonance Imaging*, 61(5), 1353–1365. <https://doi.org/10.1002/jmri.1353>
- Veiga-Canuto, D., et al. (2022). Comparative multicentric evaluation of interobserver variability in manual versus Deep Learning–based tumor segmentation. *Cancers*, 14(15), 3648. <https://doi.org/10.3390/cancers14153648>
- European Society of Radiology (ESR) What the radiologist should know about artificial intelligence – an ESR white paper. *Insights Imaging.* 2019;10:44

Area 2

- Streit, U., Uhlig, J., Lotz, J., Panahi, B., & Seif Amir Hosseini, A. (2021). Analysis of core processes of the MRI workflow for improved capacity utilization. *European journal of radiology*, 138, 109648. <https://doi.org/10.1016/j.ejrad.2021.109648>
- Wichtmann, B.D., Paech, D., Pianykh, O.S. et al. Leadership in radiology in the era of technological advancements and artificial intelligence. *Eur Radiol* 36, 548–552 (2026). <https://doi.org/10.1007/s00330-025-11745-4>
- Icobrain 27 febbraio 2026, Clinical Workflow Integration, Sito web, <https://www.icometrix.com/clinical-workflow>.
- Korfiatis, P., Kline, T. L., Meyer, H. M., Khalid, S., Leiner, T., Loufek, B. T., Blezek, D., Vidal, D. E., Hartman, R. P., Joppa, L. J., Missert, A. D., Potretzke, T. A., Taubel, J. P., Tjelta, J. A., Callstrom, M. R., & Williamson, E. E. (2024). Implementing Artificial Intelligence Algorithms in the Radiology Workflow: Challenges and Considerations. *Mayo*

Clinic proceedings. Digital health, 3(1), 100188.
<https://doi.org/10.1016/j.mcpdig.2024.100188>

- Wenderott, K., Krups, J., Zaruchas, F. et al. Effects of artificial intelligence implementation on efficiency in medical imaging—a systematic literature review and meta-analysis. *npj Digit. Med.* 7, 265 (2024). <https://doi.org/10.1038/s41746-024-01248-9>
- Monereo-Sánchez, J., et al. (2021). Quality control strategies for brain MRI segmentation: Impact on variability and reproducibility. *Journal of Magnetic Resonance Imaging*, 61(5), 1353–1365. <https://doi.org/10.1002/jmri.1353>

Area 3

- Z. Kuang, Z. Yan, A. Abayazeed, F. Wagner, L. Yu and M. Reyes, ROXSI: Robust Cross-Sequence Semantic Interaction for Brain Tumor Segmentation on Multi-Sequence MR Images, *IEEE Journal of Biomedical and Health Informatics*, vol. 29, no. 4, April 2025, doi: 10.1109/JBHI.2024.3513479
- Renard, F., Guedria, S., De Palma, N., & Vuillerme, N. (2020). Variability and reproducibility in Deep Learning for medical image segmentation. *Scientific Reports*. <https://doi.org/10.1038/s41598-020-69920-0>

