

Multi-Observer Clinical Validation of AI-Based Probabilistic Breast CTV Maps

Maria Giulia Ubeira-Gabellini¹, Cecilia Riani^{1,2}, Gabriele Palazzo¹, Davide Monticelli¹, Andrei Fodor³, Nadia Gisella Di Muzio^{3,4}, Antonella del Vecchio¹, Alessandra Palma⁵, Anna Balsamo⁶, Angela Coniglio⁶, Claudio Fiorino¹

¹IRCCS San Raffaele Scientific Institute | ²University of Pavia | ³Radiation Oncology, IRCCS San Raffaele | ⁴Vita-Salute San Raffaele University | ⁵Istituto Superiore di Sanità | ⁶Ministry of Health

alessandra.palma@iss.it

BACKGROUND e OBIETTIVI

La segmentazione manuale del *Clinical Target Volume* (CTV) della mammella rappresenta una fase fondamentale della pianificazione radioterapica, è soggetta però a significativa variabilità inter-osservatore (IOV) e richiede un notevole impegno in termini di tempo.

Recentemente è stata sviluppata una pipeline di Deep Learning basata su Transfer Learning per adattare modelli di segmentazione automatica alle caratteristiche di contornamento dei singoli clinici.

L'obiettivo di questo studio è **validare clinicamente mappe probabilistiche del CTV** ottenute mediante l'integrazione di modelli specifici per osservatore e quantificare la variabilità sistematica del contornamento.

MATERIALI E METODI

961

7

CLINICI

40

CASI DI VALIDAZIONE

5

PIANI MAMMILLA SINISTRA

Chirurgia conservativa seguita da radioterapia; TC di pianificazione e contorni clinici del CTV dal 2017 al 2022.

UNET sviluppata internamente, MONAI v1.3; fine-tuning mediante TL per 200 epoche su ≥ 70 casi per clinico.

Sette modelli specifici per clinico combinati in una mappa di isoprobabilità 14%-100%.

Validazione: tre clinici esperti e uno specializzando.

Metriche: ECDF dei residui di sotto/sovra-segmentazione, DSC e ASD.

Confronto dosimetrico mediante piani ottimizzati automaticamente.

WORKFLOW

1

Contorni clinici

2

TL fine-tuning

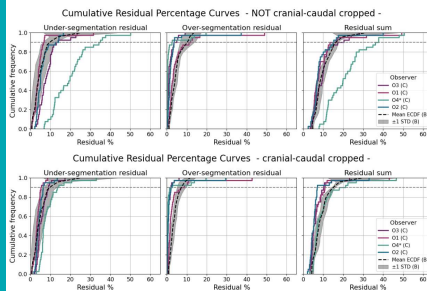
3

Mappa probabilistica

4

Validazione clinica e dosimetrica

RISULTATI 1 | QUALITÀ DELLA MAPPA



9.3% $\pm$ 3.4

residuo di sotto-segmentazione

10.4% $\pm$ 3.5

residuo di sovra-segmentazione

16.5% $\pm$ 1.3

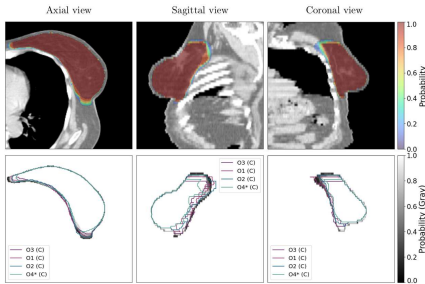
somma dei residui

L'ECDF media ha raggiunto 0,9 ai livelli di residuo riportati sopra.

I contorni degli esperti erano in gran parte compresi nella banda di variabilità prevista.

Il medico in formazione ha mostrato deviazioni maggiori, soprattutto lungo l'asse cranio-caudale.

RISULTATI 2 | CTV PROBABILISTICO

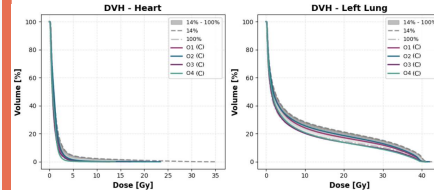


Intervallo di variabilità 14%-100%

Le differenze più rilevanti sono state osservate:
- ai margini cranio-caudali;
- nelle regioni laterali del volume mammario.

Differenza media di volume: 193 $\pm$ 103 cm³

RISULTATI 3 | IMPATTO DOSIMETRICO



DVH di cuore e polmone sinistro

I contorni clinici risultano contenuti all'interno dell'intervallo di variabilità previsto 14%-100%.

I *Dose-Volume Histogram* (DVH) ottenuti dai contorni manuali si collocano prevalentemente vicino al limite inferiore dell'intervallo probabilistico.

I risultati supportano la possibilità di una stima diretta dell'impatto dosimetrico della variabilità inter-osservatore.

RISULTATI PRINCIPALI

Transfer Learning

Miglioramento dell'adattamento dei modelli alle modalità di contornamento dei singoli clinici.

IOV

Quantificazione oggettiva della variabilità sistematica.

Assicurazione e di Qualità

Identificazione automatica di contorni potenzialmente anomali.

Formazione

Supporto all'addestramento di giovani clinici evidenziando le aree di maggiore incertezza.

DVH

Possibilità di stima della variabilità dosimetrica associata alle differenze di contornamento

CONCLUSIONI

Le mappe probabilistiche del CTV basate su Deep Learning hanno dimostrato robustezza e affidabilità in una coorte indipendente multi-osservatore.

L'approccio proposto consente non solo di ottenere una segmentazione automatica, ma anche di quantificare la variabilità inter-osservatore e il relativo impatto dosimetrico.

Possibili applicazioni per questa metodologia:

- controllo qualità dei contorni;
- formazione e tutoraggio;
- armonizzazione multicentrica.

RILEVANZA CLINICA

ASSICURAZIONE DI QUALITÀ

Confronto oggettivo dei contorni clinici con un intervallo di variabilità atteso

FORMAZIONE

Supporto al tutoraggio dei medici in formazione e identificazione delle deviazioni sistematiche di delineazione

TRIAL CLINICAL

Assicurazione di qualità standardizzata dei contorni tra osservatori e centri.

TAKE-HOME MESSAGE

Il transfer learning ha generato mappe probabilistiche del CTV specifiche per clinico, capaci di descrivere la variabilità inter-osservatore sistematica e di supportare la valutazione della qualità dei contorni, la formazione e l'assicurazione di qualità nei trial clinici.

BIBLIOGRAFIA

[1] Ubeira-Gabellini MG, Palazzo G, et al. Development and external multicentric validation of a deep learning-based clinical target volume segmentation model for whole-breast radiotherapy. *Physics and Imaging in Radiation Oncology*. 2025;34:100749. doi:10.1016/j.phro.2025.100749.

FINANZIAMENTI

Il progetto è condotto con il supporto tecnico e finanziario del Ministero della Salute – Programma-CCM 2024: "Segmentazione del CTV basata su IA nella radioterapia della mammella: esplorazione di modelli utilizzabili su larga scala, incluso mappe di probabilità intra ed inter-Centro"