

POLITECNICO DI TORINO

**Corso Laurea Magistrale
in Ingegneria Biomedica**



Segmentazione automatica di metastasi tumorali epatiche a partire da immagini TC tramite convolutional neural networks

Relatori

prof. Gabriella Balestra

ing. Valentina Giannini

firma dei relatori

Candidato

Lorenzo Beldi

firma candidato

A.A. 2019/2020

INDICE

INDICE FIGURE	4
INDICE TABELLE	6
INDICE EQUAZIONI E FORMULE.....	7
SOMMARIO	8
INTRODUZIONE.....	9
MATERIALI E METODI	12
Pazienti	12
Individuazione regione epatica	12
Ricerca limiti addome piano trasverso	12
Soglie piano frontale	13
Clustering.....	14
Segmentazione	17
Criticità.....	20
U-Net	23
Preparazione Dataset.....	23
Divisione Train Set, Test Set e Validation Set.....	23
Pre-processing immagini	25
Architettura rete	25
Post-Processing.....	27
RISULTATI	28
Confronto architettura reti.....	28
Analisi per pool di provenienza.....	36
Analisi intensità dei pixel.....	40
Analisi per grandezza della lesione.....	42
Analisi per tipologia di lesione	44
Esempi di segmentazione.....	46
Analisi lesioni critiche.....	49
DISCUSSIONE	51
BIBLIOGRAFIA.....	52

INDICE FIGURE

Figura 1: Incidenza e mortalità nel mondo dei tumori più frequenti (1).....	9
Figura 2: Rappresentazione vena porta.....	9
Figura 3: Flowchart lavoro Vorontsov e Tang (4).....	10
Figura 4: Flowchart lavoro Sun e Guo (5).....	11
Figura 5: Binarizzazione addome piano trasverso prima dell'analisi delle componenti connesse (A) e dopo (B).....	13
Figura 6: Binarizzazione piano trasverso prima dell'analisi delle componenti connesse (A), dopo l'analisi delle componenti connesse (B) e immagine dopo il riempimento dei buchi all'interno dell'oggetto (C).....	14
Figura 7: Rappresentazione dell'individuazione del fegato su piano frontale.....	14
Figura 8: Rappresentazione degli step di segmentazione del fegato.....	18
Figura 9: Esempio segmentazione fegato paziente 66.....	18
Figura 10: Esempio segmentazione fegato paziente C1007.....	18
Figura 11: Esempio segmentazione fegato paziente C1017.....	19
Figura 12: Esempio segmentazione fegato paziente C2005.....	19
Figura 13: Esempio segmentazione fegato paziente N1027.....	19
Figura 14: Esempio segmentazione fegato paziente N1039.....	19
Figura 15: Esempio criticità lesione under-segmentata.....	20
Figura 16: Esempio criticità centroid tracking.....	22
Figura 17: Esempio criticità centroid tracking.....	22
Figura 18: Architettura Unet.....	26
Figura 19: Confronto architettura reti sul Train set.....	30
Figura 20: Confronto architettura reti sul Test set.....	31
Figura 21: Percentuale di pazienti del Train test appartenenti ai range indicanti le percentuali di lesioni individuate per paziente.....	35
Figura 22: Percentuale di pazienti del Test set appartenenti ai range indicanti le percentuali di lesioni individuate per paziente.....	36
Figura 23: Percentuali di lesioni individuate per pool di provenienza.....	38
Figura 24: Mediana di Dice, Precision e Recall delle lesioni del Train set per pool di provenienza.....	39
Figura 25: Mediana di Dice, Precision e Recall delle lesioni del Test set per pool di provenienza.....	39
Figura 26: Rappresentazione della distribuzione dei valori di intensità dei pixel delle metastasi appartenenti al Train set.....	41
Figura 27: Rappresentazione della distribuzione dei valori di intensità dei pixel delle metastasi appartenenti al Test set.....	42
Figura 28: Percentuale di lesioni individuate per grandezza appartenenti rispettivamente del Train set e del Test set.....	43
Figura 29: Percentuale di lesioni individuate per profondità appartenenti rispettivamente del Train set e del Test set.....	44
Figura 30: Percentuale di lesioni individuate per tipologia appartenenti rispettivamente del Train set e del Test set.....	45

Figura 31: Paziente 1, in rosso la segmentazione manuale e in blu la segmentazione automatica.....	46
Figura 32: Paziente N2004, in rosso la segmentazione manuale e in blu la segmentazione automatica.....	47
Figura 33: Paziente C1005, in rosso la segmentazione manuale e in blu la segmentazione automatica.....	48
Figura 34: Percentuale di metastasi apparenti alla diverse criticità.....	49
Figura 35: Criticità segmentazione del fegato errata.....	50
Figura 36: Esempio di segmentazione fallita.....	50
Figura 37: Criticità lesione sul bordo del fegato.....	50
Figura 38: Lesioni con intensità e distribuzione pixel simile al fegato.....	50

INDICE TABELLE

Tabella 1: Valori delle soglie di double thresholding per ogni paziente.....	16
Tabella 2: Elenco di pazienti esclusi, di pazienti modificati e di pazienti usati solo per il testing.....	25
Tabella 3: Confronto architetture sul numero di lesioni considerate non individuate e individuate.....	29
Tabella 4: Confronto architetture sulla media di Dice, Precision e Recall.....	29
Tabella 5: Confronto architetture sulla deviazione standard di Dice, Precision e Recall.....	29
Tabella 6: Confronto architetture sulla mediana di Dice, Precision e Recall.....	30
Tabella 7: Percentuale di lesioni individuate di ogni paziente appartenente al Test set, derivanti dalla rete con 5 layer.....	32
Tabella 8: Percentuale di lesioni individuate di ogni paziente appartenente al Test set, derivanti dalla rete con 4 layer.....	32
Tabella 9: Percentuale di lesioni individuate di ogni paziente appartenente al Train set, derivanti dalla rete con 5 layer.....	33
Tabella 10: Percentuale di lesioni individuate di ogni paziente appartenente al Train set, derivanti dalla rete con 4 layer.....	34
Tabella 11: Numero di metastasi considerate individuate e non individuate per pool di provenienza.....	37
Tabella 12: Media di Dice, Precision e Recall delle lesioni per pool di provenienza.....	37
Tabella 13: Mediana di Dice, Precision e Recall delle lesioni per pool di provenienza.....	38
Tabella 14: Intensità media dei pixel.....	41
Tabella 15: Numero e relativa percentuale di metastasi piccole e grandi individuate, rispettivamente per Train set e Test set.....	43
Tabella 16: Numero e relativa percentuale di metastasi individuate per ogni range, rispettivamente per Train set e Test set.....	44
Tabella 17: Numero e relativa percentuale di metastasi individuate per ogni tipo, rispettivamente per Train set e Test set.....	45
Tabella 18: Cause di criticità con relativo numero e percentuale di metastasi appartenenti ad ogni gruppo.....	49

INDICE EQUAZIONI E FORMULE

Equazione 1: Formula per il calcolo del range dell'addome su piano trasverso.....	13
Equazione 2: Formula per il calcolo della soglia minore e maggiore per il double thresholding.....	15
Equazione 3: Z-score normalization.....	24
Equazione 4: Dice Loss Function.....	27
Equazione 5: Dice, Precision e Recall.....	28
Equazione 6: Intervallo di valori assunto dal 95% di una distribuzione Normale.....	40

SOMMARIO

La medicina personalizzata è una delle nuove direzioni della medicina moderna, la quale prevede misure diagnostiche e terapeutiche specifiche per ogni singolo paziente. Nel caso di tumori con metastasi epatiche, diventerebbe importante, per l'ottimizzazione della terapia, conoscere a priori se una determinata metastasi possa reagire positivamente o meno ad un ciclo di terapia, prima di intervenire con un intervento chirurgico.

Attualmente si stanno sviluppando studi di caratterizzazione automatica per prevenire la risposta alla terapia. Questi metodi però richiedono la segmentazione manuale delle metastasi.

L'implementazione di un sistema di segmentazione automatico diventa imperativo per diminuire il tempo di diagnosi.

Lo stato dell'arte attuale presenta dei metodi di segmentazione semiautomatici, quindi con supervisione dell'operatore necessaria, funzionanti.

Scopo di questa Tesi è la segmentazione automatica delle metastasi tramite Deep Learning, nello specifico usando la U-Net: quest'ultima è una *fully-convoluted neural network* pensata specificatamente per la segmentazione di immagini mediche.

Per raggiungere questo obiettivo le CT sono state pre-processate per rendere l'input della rete neurale più specifico possibile. Il pre-processing consiste nell'individuazione automatica della regione anatomica del fegato. La segmentazione del fegato è stata fatta tramite thresholding a doppia soglia, individuate tramite utilizzo di un fuzzy c-means clustering.

Dopo aver segmentato il fegato, quella regione è stata croppata per diminuire i pixel di input ed è stato creato il Dataset per la rete.

Il Dataset è stato diviso in Train set e Test set secondo il rapporto 70-30%.

La rete implementata, presenta un input, una parte di down-sampling, una parte di up-sampling e un output.

Passo successivo è stato quello di lavorare sui parametri e i risultati della rete, per cercare di migliorarne le prestazioni. Le prove sono state fatte sul Test set, che contiene delle immagini che non sono state utilizzate in fase di allenamento.

INTRODUZIONE

Il tumore al colon-retto è il quarto tumore per incidenza (6,1%) e il quinto per mortalità (5,8%) tra tutti i tipi tumori(1).

Nel tempo la possibilità di sopravvivenza a questo tumore è largamente aumentata grazie a screening diagnostici e a trattamenti migliori e più specifici; comunque circa la metà dei pazienti con questa malattia incorrono nella morte.

I pazienti con tumore al colon-retto sviluppano spesso metastasi al fegato a causa della vena porta che drena quasi tutto il sangue dell'apparato digerente nel fegato (Figura 2). Più nello specifico il 35% dei pazienti sottoposti ad una TC di controllo prima di un intervento chirurgico per la rimozione di un tumore al colon-retto, presentano metastasi epatiche (2).

CANCER SITE	NO. OF NEW CASES (% OF ALL SITES)	NO. OF DEATHS (% OF ALL SITES)
Lung	2,093,876 (11.6)	1,761,007 (18.4)
Breast	2,088,849 (11.6)	626,679 (6.6)
Prostate	1,276,106 (7.1)	358,989 (3.8)
Colon	1,096,601 (6.1)	551,269 (5.8)
Nonmelanoma of skin	1,042,056 (5.8)	65,155 (0.7)
Stomach	1,033,701 (5.7)	782,685 (8.2)
Liver	841,080 (4.7)	781,631 (8.2)
Rectum	704,376 (3.9)	310,394 (3.2)
Esophagus	572,034 (3.2)	508,585 (5.3)
Cervix uteri	569,847 (3.2)	311,365 (3.3)
Thyroid	567,233 (3.1)	41,071 (0.4)

Figura 1. Incidenza e mortalità nel mondo dei tumori più frequenti

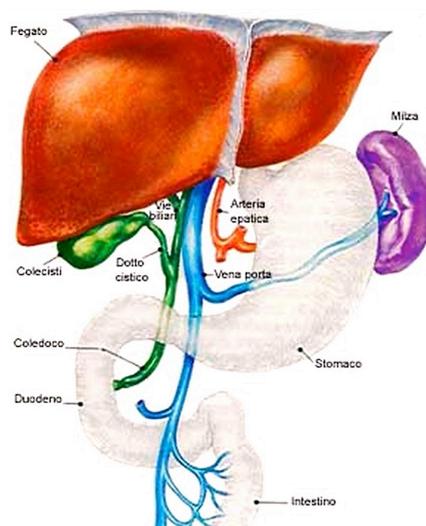


Figura 2. Rappresentazione vena porta

La diagnostica delle metastasi tumorali al fegato prevede la segmentazione manuale da parte dell'operatore, processo che comporta un grande quantitativo di tempo ed è fortemente dipendente dall'esperienza dell'operatore.

Un sistema di segmentazione automatico potrebbe far risparmiare tempo prezioso all'operatore, fornendogli contemporaneamente un confronto attendibile in fase diagnostica.

In letteratura si trovano diversi approcci per la segmentazione di metastasi epatiche.

Qi1 e Xiong (3) hanno implementato un sistema di segmentazione semi-automatico dove utilizzano un region growing sul volume 3D. Il contro di questo approccio sono la richiesta di inizializzazione dei seed da parte dell'utente e la tendenza del region growing ad uscire dai contorni dell'oggetto da segmentare, propagando l'errore alle slice adiacenti.

Altro approccio, sempre semi-automatico, è stato tentato da Vorontsov e Tang (4), che hanno proposto una segmentazione tramite modello deformabile, supportata da una rete neurale. Nello specifico, hanno allentato una rete MLP per disegnare una mappa di probabilità della metastasi tumorale. Nella fase di testing viene generata suddetta mappa di probabilità, vengono scelti manualmente dei pixel appartenenti alla metastasi per l'inizializzazione del modello deformabile e quest'ultimo si aggiusta seguendo le informazioni ottenute dalla mappa di probabilità

Questo metodo produce dei buoni risultati ma ha bisogno di un ampio set di immagini per il training e inoltre necessita dell'inizializzazione manuale da parte dell'utente.

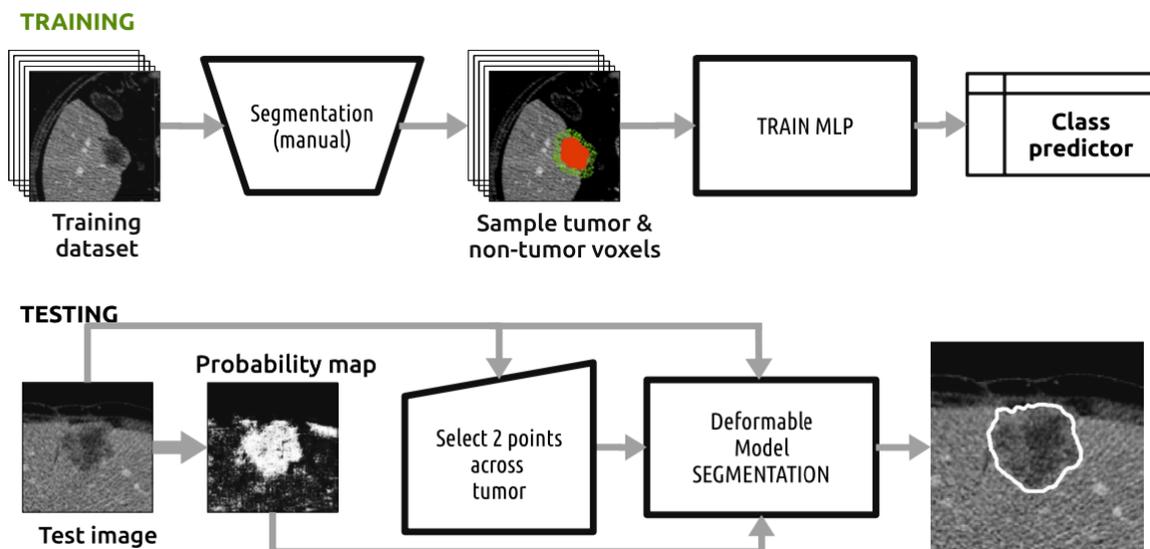


Figura 3. Flowchart lavoro Vorontsov e Tang

Metodi automatici sono stati invece proposti da Sun e Guo (5), o da Vorontsov e Cerny (6) che esplorano le potenzialità delle fully-convolutional neural networks applicate alla segmentazione delle metastasi epatiche.

In particolare il primo studio propone una multy-channel fully-convolutional neural network, dove ogni fase di contrasto della CT concorre all'allenamento della rete neurale in maniera separata. Questo metodo ha ottenuto buoni risultati.

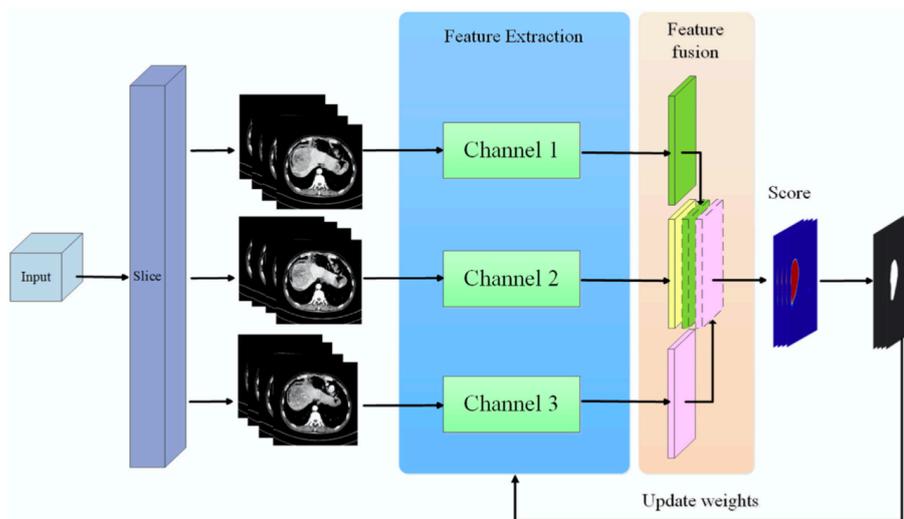


Figura 4. Flowchart lavoro Sun e Guo..

Anche Vorontsov e Cerny (6) trovano nelle fully-convolutional neural networks un potenziale metodo per una segmentazione automatica efficace e più veloce rispetto ad altri metodi in letteratura. Anche questo approccio ottiene buoni risultati.

Lo scopo di questa Tesi è quello di implementare un modello di segmentazione automatica delle metastasi tumorali epatiche tramite l'utilizzo della U-Net, una fully-convolutional neural network, da affiancare ad una caratterizzazione automatica della risposta della lesione alla terapia.

MATERIALI E METODI

Pazienti

Per questa Tesi sono state utilizzate le CT addominali di 130 pazienti di prima e terza linea. Nello specifico i pazienti di prima linea sono 92, di cui 31 derivanti dall'Istituto Di Candiolo e 61 dall'Ospedale Niguarda. Invece i pazienti di terza linea sono stati forniti da più centri diversi e in tutto sono 38.

Inoltre sono state fornite anche le segmentazioni manuali delle metastasi tumorali epatiche relative alle CT addominali.

In tutto sono state fornite dunque le segmentazioni di 540 lesioni, di cui 172 provenienti da pazienti di terza linea, 125 da pazienti di prima linea dell'Istituto Di Candiolo e 240 da pazienti di prima linea dell'Ospedale Niguarda.

Le immagini fornite sono in formato NIfTI, 512x512 int16.

Queste immagini sono state pre-processate e poi si è proceduto con l'individuazione automatica della regione anatomica del fegato al fine di ridurre, e rendere più specifico, l'input della rete Unet per la segmentazione delle metastasi.

Individuazione regione epatica

L'algoritmo per l'individuazione del fegato è stato sviluppato in ambiente MATLAB®. Le criticità della segmentazione automatica del fegato sono date dalla estrema variabilità della sua morfologia e dell'intensità dei pixel che compongono la CT tra un paziente e l'altro.

Per ovviare a questa criticità si è implementato un sistema di thresholding che utilizza come soglie i valori ricavati da un fuzzy c-means clustering.

Di seguito sono riportati i passi fondamentali dell'algoritmo di segmentazione.

Ricerca limiti addome piano trasverso

In alcune CT la schiena del lettino è della stessa intensità del fegato, quindi per evitare che l'algoritmo segmenti delle regioni sbagliate si è proceduto a definire delle coordinate che individuino la parte di interesse.

Per fare ciò, si è presa la slice centrale della CT su piano trasverso e si è creata una maschera binaria in cui tutti i pixel con valore maggiore di -600 sono stati messi ad 1. Successivamente si è fatta un'analisi delle componenti connesse (Figura 5), eliminando tutti gli oggetti tranne il più grande, che risulta essere l'addome.

Si è poi proceduto a determinare i limiti dell'addome che verranno utilizzati per cercare le soglie sul piano frontale.

Il range dell'addome è calcolato usando le seguenti formule:

$$val_min = y_{min} + 0,2 * (y_{max} - y_{min})$$

$$val_max = y_{min} + 0,8 * (y_{max} - y_{min})$$

Equazione 1. Formula per il calcolo del range dell'addome su piano trasverso.

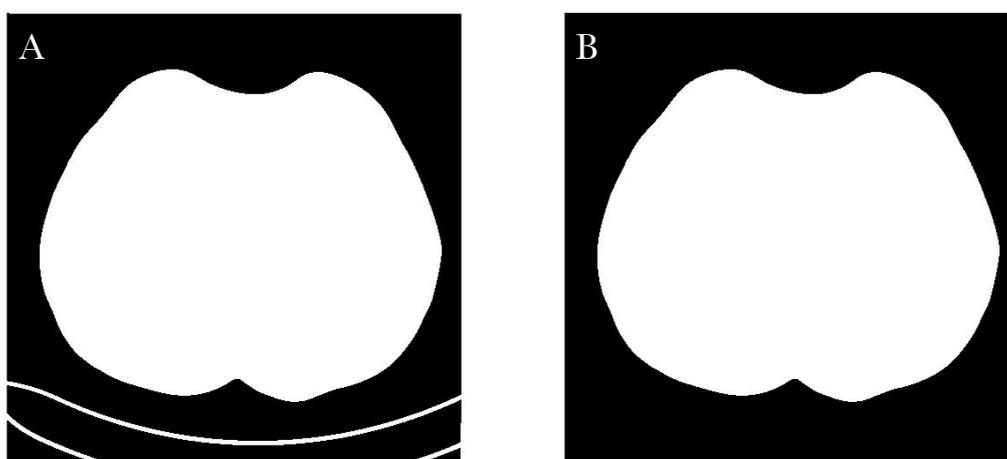


Figura 5. Binarizzazione addome piano trasverso prima dell'analisi delle componenti connesse (A) e dopo (B).

Soglie piano frontale

Per passare sul piano frontale si è applicata una rotazione 3D di 90° sull'asse frontale. Si considerano solo le slice nell'intervallo individuato precedentemente (*Equazione 1*) e si procede a filtrare l'immagine, slice per slice, con un filtro mediano con un kernel di dimensioni 3x3 per uniformare i pixel.

Dopo il filtraggio calcolo il 65 percentile delle intensità dei pixel, sempre slice per slice. Questo perchè si utilizza la conoscenza a priori sulla grandezza del fegato sul piano frontale della CT, e si fa dunque l'assunzione che il 65% dell'immagine sia composta da fegato.

Per trovare le soglie del thresholding frontale si prende il massimo dei percentili. Le soglie sono calcolate sommando e sottraendo al massimo percentile la deviazione standard di tutti i percentili calcolati, questo per allargare il range e includere eventuali metastasi che hanno intensità caratteristiche generalmente più basse.

Si procede dunque alla binarizzazione dell'immagine sul piano frontale usando le due soglie individuate.

Viene poi svolta un'analisi delle componenti connesse per eliminare tutti gli oggetti, eccetto il più grosso, che rappresenta il fegato. Successivamente vengono riempiti i buchi all'interno dell'oggetto.



Figura 6. Binarizzazione piano trasverso prima dell'analisi delle componenti connesse (A), dopo l'analisi delle componenti connesse (B) e immagine dopo il riempimento dei buchi all'interno dell'oggetto (C).

Viene infine salvata la coordinata superiore e inferiore del fegato sul piano frontale, in modo da limitare le slice di segmentazione sul piano trasverso.

Clustering

La slice centrale del piano frontale dell'immagine originale viene moltiplicata per la maschera ottenuta al punto precedente, ottenendo il fegato sul piano frontale (Figura 7) rappresentato dall'intensità dei propri pixel.



Figura 7. Rappresentazione dell'individuazione del fegato su piano frontale.

Si è effettuato poi, su tutti i pixel diversi da 0, un fuzzy c-means clustering tramite la funzione “*fc*m” di Matlab. Questo tipo di clustering assegna ogni pixel a tutti i cluster, specificando il grado di appartenenza a quest’ultimi tramite una membership fuzzy.

Si è deciso di utilizzare 3 cluster per differenziare meglio le varie regioni della CT: sfondo, fegato e altre zone non appartenenti al fegato.

Dopo che i pixel sono stati assegnati ai loro cluster secondo il loro massimo grado di appartenenza, si è individuato il cluster maggiore, rappresentante i valori di intensità del fegato.

Le soglie per il clustering sono state scelte secondo le seguenti formule:

$$th_{\min} = \min(\text{cluster}) - 2 * \text{std}(\text{cluster})$$

$$th_{\max} = \max(\text{cluster}) + \text{std}(\text{cluster})$$

Equazione 2. Formula per il calcolo della soglia minore e maggiore per il double thresholding.

La scelta di sbilanciare la soglia minore rispetto alla maggiore (Equazione 2) è data dal fatto che le metastasi rappresentano valori di intensità dei pixel minori rispetto alla media di quelli del fegato.

Nella *Tabella 1*, sono riportati i valori delle soglie di segmentazione inferiore e superiore per ogni paziente.

Paziente	Th_min	Th_max	Paziente	Th_min	Th_max	Paziente	Th_min	Th_max
1	56,30	112,60	C1006	103,01	129,00	N1018	97,38	133,06
2	90,75	219,62	C1007	61,92	107,54	N1019	36,97	190,77
4	47,40	116,05	C1008	17,73	219,64	N1020	52,60	132,20
5	25,58	128,46	C1009	80,42	120,04	N1021	61,94	142,28
6	101,15	161,43	C1010	76,72	139,89	N1022	75,59	187,46
10	88,06	213,97	C1011	58,49	540,01	N1023	45,95	248,27
12	57,66	108,17	C1012	62,23	122,14	N1024	67,11	113,45
17	71,74	148,88	C1013	109,63	152,44	N1025	59,24	275,13
18	100,21	150,15	C1014	99,25	142,37	N1026	86,61	164,95
19	75,76	166,62	C1015	70,51	128,99	N1027	99,13	151,94
20	-22,64	67,07	C1016	62,64	126,18	N1029	50,89	131,05
22	75,99	126,26	C1017	73,14	123,93	N1030	50,84	121,83
23	-104,14	149,82	C1018	73,89	124,31	N1031	68,52	142,49
25	72,78	135,36	C1019	63,17	136,17	N1032	71,12	115,19
26	87,79	155,36	C1020	99,67	123,67	N1033	86,61	147,70
27	42,49	164,50	C1021	54,95	120,77	N1034	-26,67	125,59
28	27,83	76,34	C1023	75,35	162,57	N1035	80,50	149,50
29	64,37	140,07	C1024	82,25	129,12	N1036	44,73	233,89
30	83,31	154,10	C1025	62,21	273,65	N1037	51,29	120,36
31	104,99	303,01	C1026	66,66	151,67	N1038	79,26	124,87
32	81,03	137,49	C1027	74,19	138,16	N1039	51,30	124,60
33	74,22	131,14	C2001	63,05	130,47	N1040	34,93	106,53
36	84,18	141,91	C2002	62,95	114,52	N1041	88,77	560,87
37	75,52	131,49	C2003	70,04	122,48	N1042	66,83	161,59
38	12,37	102,32	C2004	-1,21	122,85	N1043	90,78	148,36
41	72,05	112,23	C2005	58,39	124,31	N1044	96,33	147,83
45	61,07	122,22	N1002	62,10	113,20	N1045	46,87	140,82
46	47,55	118,98	N1003	73,70	143,90	N1046	48,28	128,36
49	68,80	164,35	N1004	66,42	131,29	N1047	44,39	470,31
50	81,07	122,22	N1005	12,58	150,46	N1048	60,63	129,94
52	94,85	130,33	N1006	76,64	130,18	N1049	61,49	154,26
54	31,22	117,14	N1007	89,45	160,28	N1050	46,31	111,10
55	65,39	159,80	N1008	68,94	168,03	N1051	53,30	140,35
57	62,41	129,55	N1009	54,33	153,09	N1053	68,17	129,16
62	67,05	150,47	N1010	82,33	157,34	N1054	45,52	692,74
63	43,44	137,53	N1011	49,52	159,24	N1055	54,14	143,43
66	71,21	146,65	N1012	-16,12	120,81	N1056	54,18	136,41
C1001	56,71	171,15	N1013	65,33	126,58	N1057	26,44	109,78
C1002	97,13	129,43	N1014	48,47	158,51	N2001	35,05	111,47
C1003	96,73	170,13	N1015	82,97	352,27	N2002	71,21	140,65
C1004	98,55	129,73	N1016	82,85	585,33	N2003	60,20	119,65
C1005	100,22	135,89	N1017	63,14	110,18	N2004	91,20	155,65

Tabella 1. Valori delle soglie di double thresholding per ogni paziente.

Segmentazione

Ottenute le soglie per il thresholding, si ritorna sul piano trasverso per la segmentazione.

Sono state prese solo le slice nell'intervallo di azione individuato sul piano frontale (*Equazione 2*) e si è proceduto alla segmentazione slice per slice.

I passi eseguiti per la segmentazione sono:

- filtraggio mediano con kernel 5x5 dell'immagine per unificare i pixel (*Figura 8.B*);
- applicazione delle soglie di thresholding (*Figura 8.C*);
- analisi delle componenti connesse per eliminare tutti gli oggetti tranne il più grosso (*Figura 8.C*). Nelle slice in cui il fegato è piccolo, questo passaggio può portare alla segmentazione dell'oggetto sbagliato. Questa criticità è stata risolta facendo un controllo sul baricentro degli oggetti. Utilizzando la conoscenza a priori sull'ubicazione del fegato, si è individuato un quadrante in cui deve risiedere il baricentro dell'oggetto, affinché sia considerato fegato;
- dilatazione della maschera ottenuta con un elemento strutturale a disco di raggio 8 e chiusura dei buchi all'interno dell'oggetto (*Figura 8.D*);

Una volta segmentate tutte le slice nell'intervallo di interesse, si è proceduto con una parte di post-processing (*Figura 8.E*):

- si sono portate le maschere ottenute sul piano frontale e si è eseguita un'apertura con elemento strutturale a disco di raggio 3 per eliminare gli errori di segmentazione ed avere una continuità tra una slice e l'altra;
- l'algoritmo si sposta di nuovo sul piano trasverso dove avviene una dilatazione con elemento strutturale a disco di raggio 10. Questo perché l'obiettivo ultimo non è la segmentazione del fegato, ma individuare la regione in cui si trova. Viene prediletta la certezza di evitare esclusioni di parti di fegato a scapito di una maggiore precisione sui bordi della segmentazione;

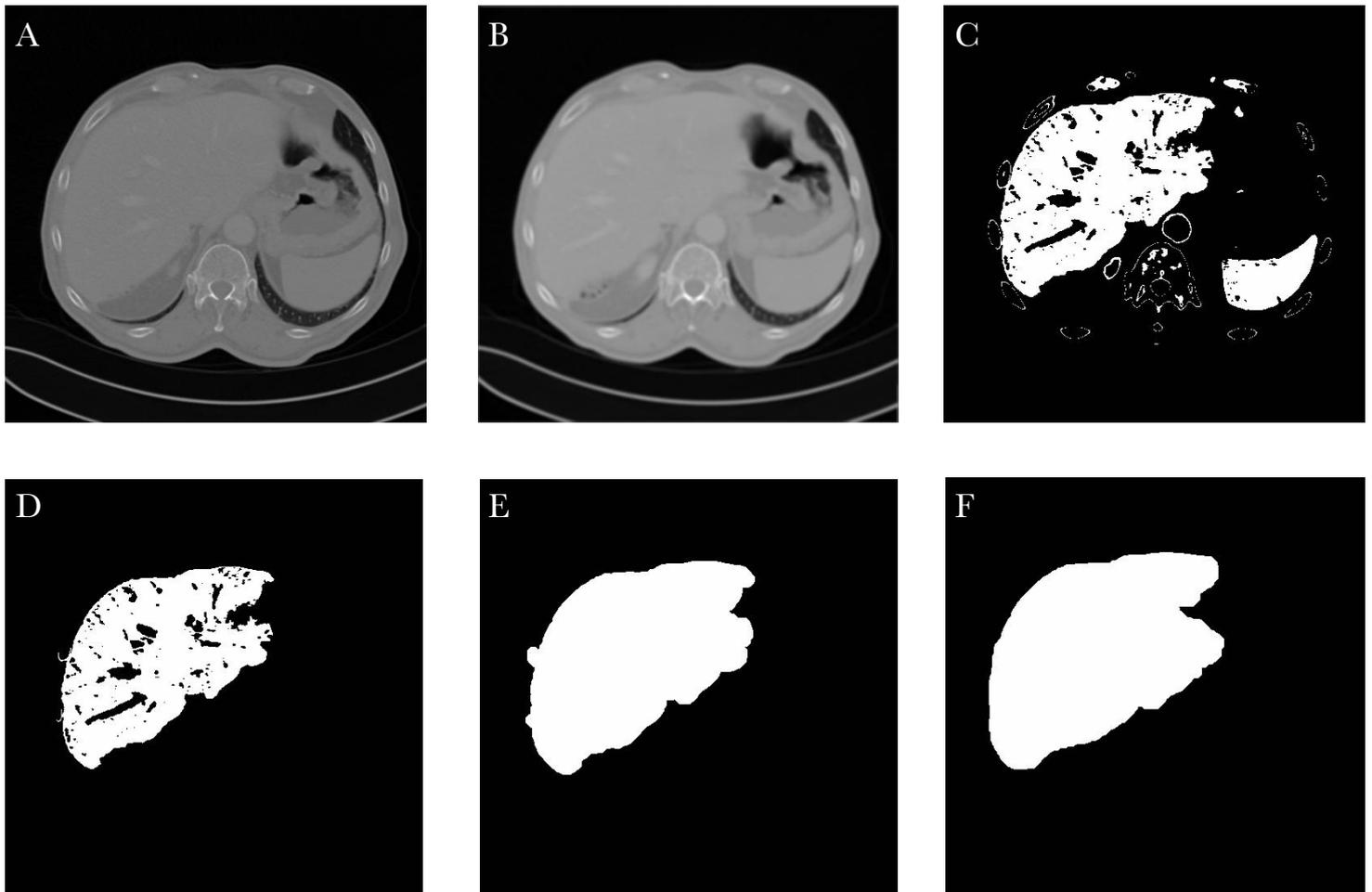


Figura 8. Rappresentazione degli step di segmentazione del fegato: TC originale (A); applicazione filtraggio mediano (B); applicazione delle soglie di thresholding (C); analisi componenti connesse (D); dilatazione e chiusura buchi all'interno dell'oggetto (E); post-processing (F).

Una volta ottenute le segmentazioni si moltiplica la CT per le maschere e si ottiene la rappresentazione del solo fegato. Di seguito sono riportati alcuni esempi.

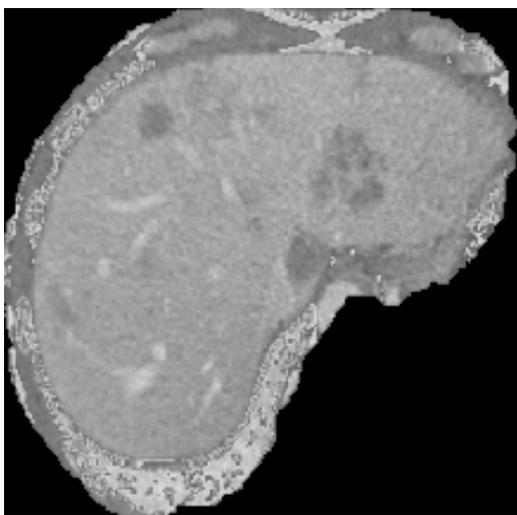


Figura 9. Esempio segmentazione fegato paziente 66.

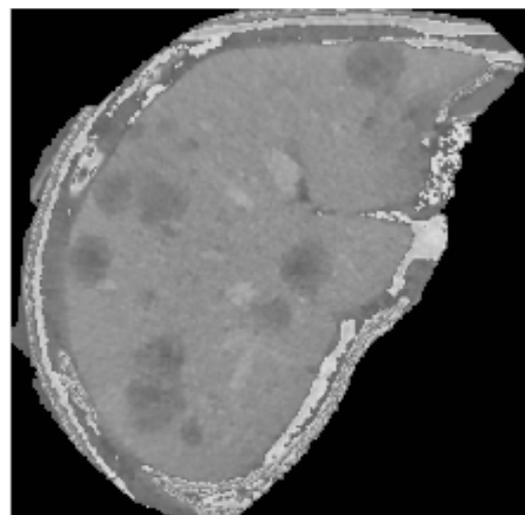


Figura 10. Esempio segmentazione fegato paziente C1007.

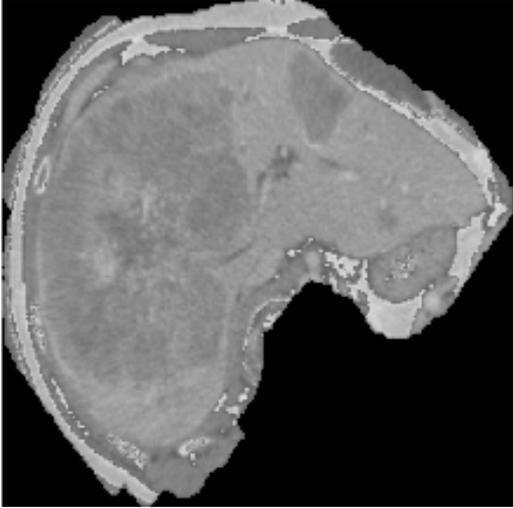


Figura 11. Esempio segmentazione fegato paziente C1017.

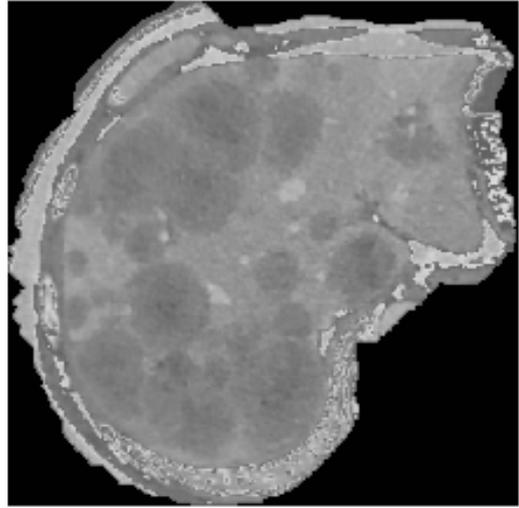


Figura 12. Esempio segmentazione fegato paziente C2005.

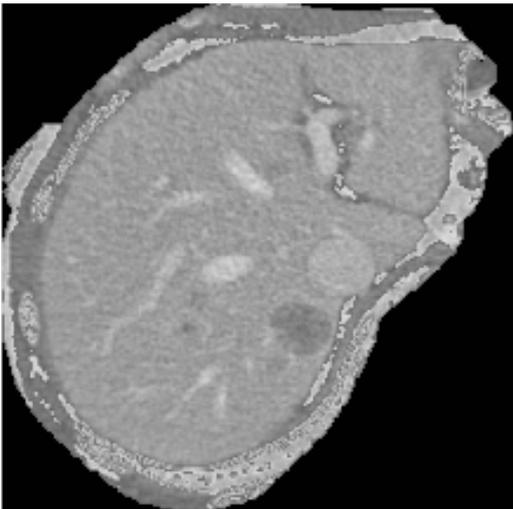


Figura 13. . Esempio segmentazione fegato paziente N1027.

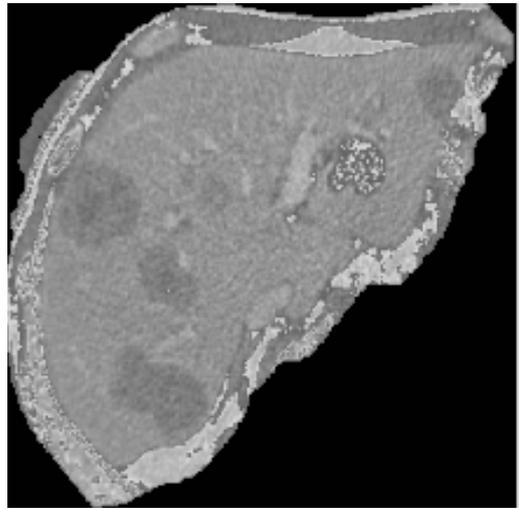


Figura 14. Esempio segmentazione fegato paziente N1039.

Criticità

Questo algoritmo presenta due criticità principale:

1. Under-segmentazione metastasi sui bordi del fegato;
2. Segmentazione di oggetti sbagliati quando il fegato è piccolo o ci sono altri oggetti più grossi di quest'ultimo;

Sono state attuate le seguenti strategie per cercare di ovviare ai problemi presentati.

1. L'under-segmentazione è rappresentata da concavità che si creano dove l'algoritmo ha fallito nell'individuare le metastasi come facenti parti del fegato. Per riempire queste concavità si sono aggiunti pixel 1 sui bordi della maschera tramite la funzione "bwmorph" di MATLAB® con l'operazione "majority" e limite a 15 cicli. Questo permette di mantenere inalterato il numero di oggetti presenti nell'immagine. Successivamente si è considerato l'intorno 8 connesso dei pixel a contatto con il bordo della maschera: se nell'intorno sono presenti 5 o più pixel con valore 1, il pixel viene messo ad 1. Così facendo la concavità si riduce notevolmente. In alcuni casi questo non è sufficiente e si perde una parte di lesione;

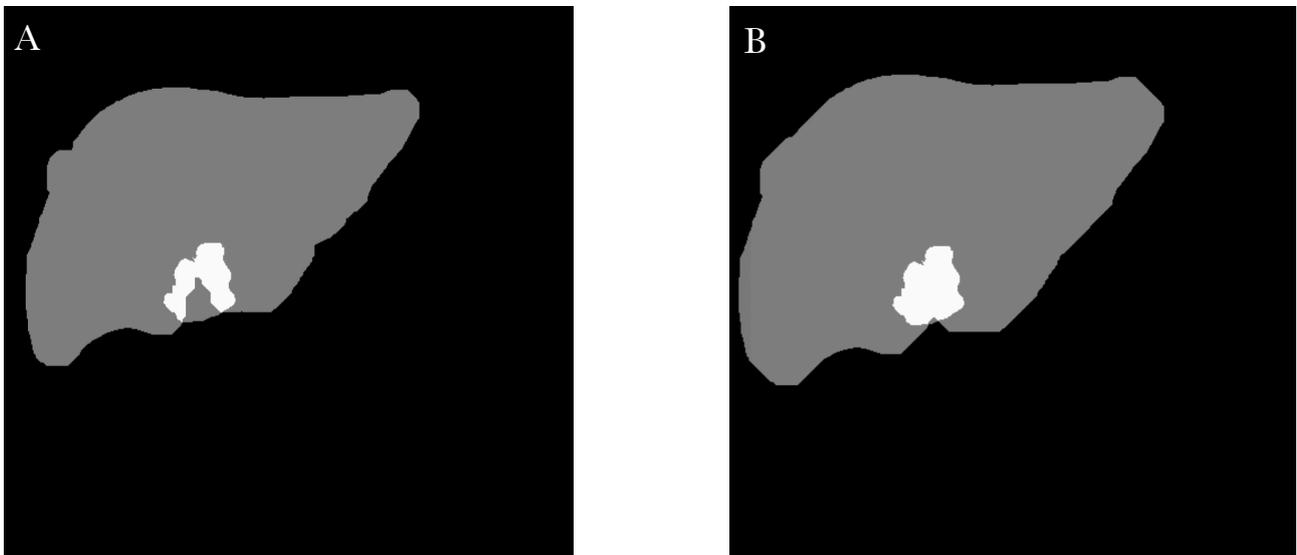


Figura 15. Esempio di maschera del fegato con lesione under-segmentata; maschera in uscita dall'algoritmo di segmentazione (A); maschera dopo l'applicazione della funzione "bwmorph" (B);

2. Quando è presente un oggetto più grosso del fegato con analoghi valori di intensità dei pixel, l'algoritmo lo seleziona come fegato. Questi oggetti però si trovano sempre in una posizione non corretta, quindi sono facili da individuare. Per ovviare a questo problema si è implementato un sistema di centroid tracking, che sfrutta la continuità della posizione del fegato nella terza dimensione della CT. Dopo aver ultimato la segmentazione di tutte le slice del fegato, si calcolano le coordinate del centroide della segmentazione, slice per slice. Sperimentalmente si è fissata una soglia, normalizzata a seconda del valore di risoluzione della CT (in mm) sulla terza dimensione. Se la distanza tra i centroidi di due segmentazioni di slice contigue è maggiore o uguale a questa soglia, la seconda segmentazione viene sostituita dalla media tra la segmentazione della slice che precede, con la segmentazione della slice che segue. Ad ogni modifica si compie nuovamente il calcolo dei centroidi delle slice modificate e si procede fin quando non sono state modificate tutte le segmentazioni che presentano distanze, tra i centroidi delle slice contigue, superiori alla soglia scelta;

Viene riportato un esempio di questa criticità.

Nella *Figura 16* è rappresentata una slice della CT del paziente C1015 e le rispettive maschere di segmentazione senza (*Figura 16.B*) e con centroid tracking (*Figura 16.C*). Si può notare che in questa slice la segmentazione è corretta in quanto il fegato è ancora l'oggetto più grosso nel range di intensità selezionato.

Invece nella *Figura 17*, è rappresentata la slice successiva, sempre dello stesso paziente, ancora con le rispettive maschere di segmentazione senza (*Figura 17.B*) e con centroid tracking (*Figura 17.C*). In questo caso il fegato non è più l'oggetto più grosso presente nel range di intensità e quindi viene eliminato dall'analisi delle componenti connesse descritte nel paragrafo Segmentazione, portando ad una segmentazione errata (*Figura 17.B*). Invece implementando il centroid tracking si può vedere come la maschera errata venga sostituita, risultando in una segmentazione corretta (*Figura 17.C*).

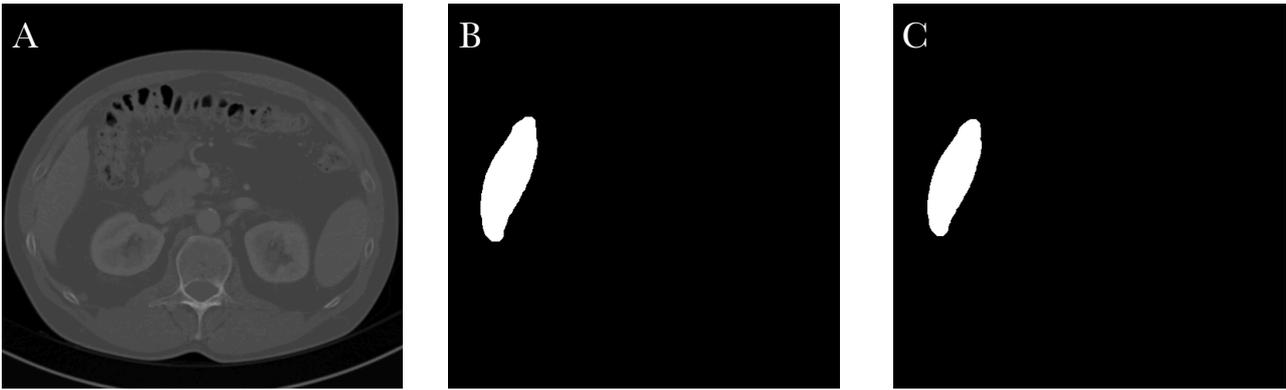


Figura 16. Slice 12 della CT del paziente C1015 (A), maschera del fegato della slice 12 della CT del paziente C1015 senza centroid tracking (B) e con centroid tracking (C)

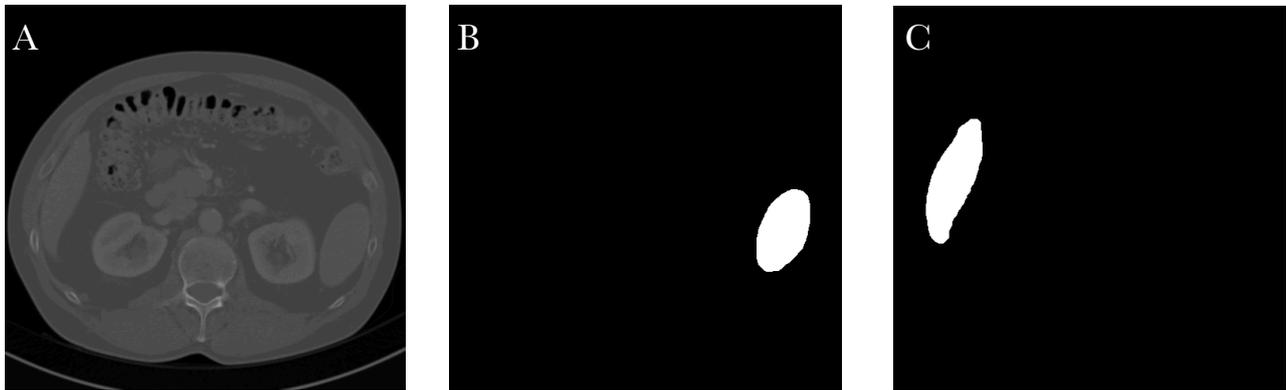


Figura 17. Slice 11 della CT del paziente C1015 (A), maschera del fegato della slice 11 della CT del paziente C1015 senza centroid tracking (B) e con centroid tracking (C)

U-Net

Di seguito vengono presentati gli step per la preparazione delle immagini e l'architettura della rete utilizzata. I seguenti codici sono stati implementati in ambiente Spyder in Python usando le seguenti librerie:

- Keras: libreria per reti neurali e deep learning;
- Skimage: libreria per l'immagine processing;
- Nibabel: libreria per la gestione di immagini mediche;
- Numpy: libreria per la gestione di array multidimensionali;
- Scipy: libreria per la gestione di algoritmi e strumenti matematici;
- Os: libreria per l'interfaccia con i sistemi operativi;
- Pickle: libreria per la serializzazione e de-serializzazione di strutture di oggetti;
- Json: libreria per l'utilizzo del data format JSON;
- Pandas: libreria per la manipolazione e l'analisi di dati;

Preparazione Dataset

Per rendere le immagini compatibili con l'input della U-Net si è proceduto a fare un crop delle immagini nell'intorno del fegato e delle rispettive segmentazioni manuali delle maschere, convertendo le dimensioni delle immagini ad un formato 256x256.

L'algoritmo di segmentazione del fegato ha fallito nella segmentazione su un totale di 4 pazienti (*Tabella 2*), portando il totale dei pazienti inclusi nel Dataset a 126.

Divisione Train Set, Test Set e Validation Set

Si è proceduto ad un'analisi dei risultati prodotti dall'algoritmo di segmentazione del fegato, per individuare eventuali segmentazioni poco precise, da escludere nella fase di training, ma utilizzabili nella fase di testing.

Alcune segmentazioni presentavano lievi imprecisioni che sono state corrette manualmente tramite MIPAV, un'applicazione che permette l'analisi e la visualizzazione di immagini mediche.

Dall'analisi si evince che dei 126 pazienti, 12 presentavano segmentazioni poco precise, che potrebbero corrompere l'allenamento della rete, per questa ragione sono escluse.

Di seguito sono riportati i pazienti esclusi dalla fase di allenamento e i pazienti che sono stati modificati manualmente (*Tabella 2*).

	Esclusi	Modificati	Solo Testing
	48	4	23
	N1001	6	29
	N1028	17	33
	N1052	28	C1001
		30	C1015
		57	N1014
		52	N1017
		66	N1022
		C1005	N1049
		C1006	N2002
		C1009	N2003
		C1016	N2004
		C1019	
		C2023	
		C2026	
		C2001	
		N1002	
		N1020	
		N1021	
		N1031	
		N1032	
		N1036	
		N1043	
		N1054	
		N1057	
TOTALE	4	25	12

Tabella 2. Elenco di pazienti esclusi, di pazienti modificati e di pazienti usati solo per il testing

Tenendo conto di quei pazienti assegnati a priori al testing, i pazienti sono stati assegnati in modo randomico al Train set e Test set con un rapporto rispettivamente del 70-30%, per un totale di 88 pazienti assegnati al Train set e di 38 pazienti assegnati al Test set.

Questa divisione porta ad avere 365 lesioni nel Train set e 175 lesioni nel Test set.

Degli 88 pazienti destinati al Train set, il 10% è tenuto da parte per comporre il Validation set: questi pazienti non saranno coinvolti nella fase di apprendimento, ma serviranno per monitorare l'andamento di quest'ultimo durante la fase di training.

Pre-processing immagini

Tutte le immagini 256x256 sono state standardizzate tramite *Z-score normalization* (Equazione 3). Invece, per quanto riguarda le maschere di segmentazione manuale, si sono binarizzate le maschere, mettendo ad 1 i pixel appartenenti a metastasi e a 0 i pixel appartenenti allo sfondo.

$$img_{std} = \frac{img - mean(img)}{std(img)}$$

Equazione 3. *Z-score normalization.*

Architettura rete

L'architettura scelta è quella proposta da Ronneberger e Fischer(7). Si tratta di una *full-convolutional neural network* ideata specificatamente per la segmentazioni di immagine mediche.

L'architettura (Figura 18) è composta da un *contracting path* seguito da un *expansive path*.

L'input è costituito dalle immagini di Training 256x256 precedentemente processate e dalle relative maschere di segmentazione.

Di seguito sono riportate in dettaglio le parti componenti l'architettura:

- Il *contracting path* è diviso in layer e prevede l'applicazione di:
 - Due step di convoluzioni 3x3, seguite da un rettificatore lineare (ReLU). Ottengo così delle feature map;
 - La feature map viene copiata, croppata e messa da parte per futuro utilizzo; il crop dell'immagine è necessario per la perdita sui bordi dovuta alle convoluizioni (*unpadded convolution*);
 - Successivamente si opera un *max pooling* 2x2 con passo 2. Ad ogni layer si raddoppiano le feature appartenenti alla feature map. Questo viene ripetuto per un numero di layer scelti dallo sviluppatore;

- L' *expansive path* è costituito invece da:
 - up-sampling della feature map seguito da una convoluzione 2x2 che dimezza il numero delle feature;
 - concatenazione della feature map corrente, con la feature map corrispondente salvata e croppata in precedenza durante il *contracting path*;
 - susseguirsi di due convoluzioni 3x3, seguite a loro volta da un rettificare lineare(ReLU);
- Il *layer di output* è composto da una convoluzione 1x1 che ha il compito di mappare i 64 canali di feature nel numero scelto di classi;

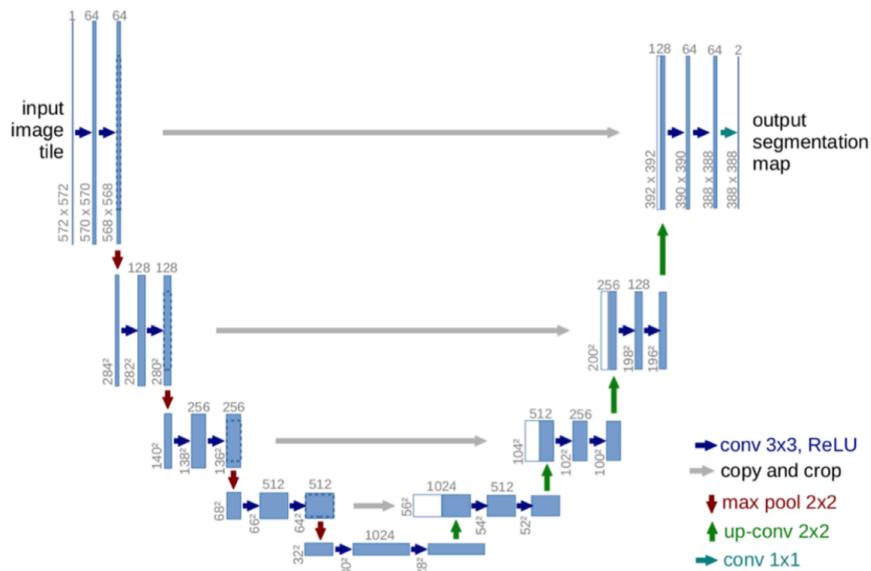


Figura 18. Architettura U-Net.

Per migliorare le prestazioni della rete, e renderla quindi più specifica a seconda del task richiesto, ci sono una serie di iperparametri che possono essere modificati. Di seguito sono elencati e descritti brevemente gli iperparametri principali.

L'*ottimizzatore* è l'iperparametro che gestisce la fase di apprendimento, in quanto valuta il gradiente di errore nella fase di *backpropagation* della rete. Una di quelle più usate ed efficienti con grandi dataset è l'*Adam* (Adaptive Moment Estimation) (8). La funzione di ottimizzazione è stato implementata con un learning rate di 1e-4. Il learning rate rappresenta il passo con cui vengono aggiornati i pesi nella fase di *backpropagation*.

Per quanto riguarda le funzioni di attivazione dei neuroni di convoluzione e del layer di output sono stati utilizzati rispettivamente la funzione *ReLU* e la funzione *Sigmoid*.

La *loss function* è l'iperparametro che valuta le performance della rete durante l'allenamento, influenzando quindi il processo di apprendimento. In questo lavoro è stato scelto il *Dice Loss Function* (Equazione 4).

$$\text{Dice Loss Function} = - \frac{2 \cdot |A \cap B|}{|A| + |B|}$$

Equazione 4. *Dice Loss Function*

Il *batch size* invece indica il numero di campioni che vengono analizzati prima che si effettui l'aggiornamento dei pesi interni della rete. Il *batch size* è stato impostato a 100.

Il numero di feature iniziali dell'immagine di input è stato impostato a 64.

Dei parametri elencati si è provato a modificare il numero di layer della rete, mantenendo fissi gli altri parametri. Sono state provate le architetture con 5 e 4 layer. Il numero di epoche è stato impostato a 30, imponendo però un controllo sul Validation set: nel caso in cui non dovesse esserci un incremento della *loss function* sul Validation set per 5 epoche successive, l'allenamento viene interrotto per evitare overfitting.

Post-Processing

L'immagine in uscita dalla rete è stata filtrata attraverso un filtro Gaussiano, e poi binarizzata usando il Metodo Otsu. Quest'ultimo presume che ci siano solamente due oggetti nell'immagine, in questo caso sfondo e metastasi, e cerca la migliore soglia per separarli.

RISULTATI

Per l'analisi dei risultati sono stati presi in considerazione i seguenti parametri:

$$DICE = \frac{2TP}{2TP + FP + FN}$$

$$PRECISION = \frac{TP}{TP + FP}$$

$$RECALL = \frac{TP}{TP + FN}$$

Equazione 5. Formula rispettivamente del Dice Coefficient, Precision e Recall.

Dove TP, FP, TN e FN sono rispettivamente veri positivi, falsi positivi, veri negativi e falsi negativi.

Una volta ottenuta la maschera di segmentazione predetta dalla rete, si è compiuta un'analisi 3D per metastasi, confrontando su tutto il volume, pixel per pixel, i valori assegnati dalla predizione rispetto a quelli assegnati dalla segmentazione manuale di riferimento.

Dei parametri elencati si è calcolata media, deviazione standard e mediana.

Importante sottolineare che per l'analisi sono state considerate solo le lesioni di cui si possiede la segmentazione manuale. Eventuali altre metastasi individuate dalla rete, che non avevano una corrispondenza nelle segmentazioni manuali, non sono state prese in considerazione.

Confronto architettura reti

Per confrontare le performance delle due architetture si è compiuta un'analisi per lesione e una per paziente.

Nell'analisi per lesione si è preso come riferimento il parametro di Dice, e si è considerata una lesione come individuata dalla rete nel caso presenti un Dice > 0,2 .

Quindi si è contato il numero di lesioni considerate individuate e il numero di quelle considerate non individuate, calcolando la percentuale di successo della rete nell'individuare le metastasi (*Tabella 3*).

Si è calcolato poi media (*Tabella 4*), deviazione standard (*Tabella 5*) e mediana (*Tabella 6*) dei parametri di controllo solo per le metastasi considerate individuate.

Questa procedura è stata fatta per le due architetture sia sul Train set che sul Test set.

LESIONI INDIVIDUATE E NON INDIVIDUATE

		NON INDIVIDUATE(Dice<0,2)	INDIVIDUATE (Dice ≥ 0,2)
5 LAYER	TRAIN	82	282
	TEST	50	125
4 LAYER	TRAIN	114	250
	TEST	59	116

Tabella 3. Numero di lesioni considerate non individuate e individuate, rispettivamente del train set e del test set delle due architetture.

MEDIA

		DICE	PRECISION	RECALL
5 LAYER	TRAIN	0,7703	0,8329	0,7716
	TEST	0,6637	0,8480	0,6122
4 LAYER	TRAIN	0,7158	0,8925	0,6449
	TEST	0,6253	0,8873	0,5407

Tabella 4. Media di Dice, Precision e Recall di train set e test set rispettivamente della architettura con 5 layer e 4 layer.

DEVIAZIONE STANDARD

		DICE	PRECISION	RECALL
5 LAYER	TRAIN	0,1640	0,1518	0,2036
	TEST	0,2064	0,1707	0,2537
4 LAYER	TRAIN	0,1759	0,1341	0,2165
	TEST	0,2076	0,1334	0,2494

Tabella 5. Deviazione standard di Dice, Precision e Recall di train set e test set rispettivamente della architettura con 5 layer e 4 layer.

MEDIANA

		DICE	PRECISION	RECALL
5 LAYER	TRAIN	0,8280	0,8721	0,8374
	TEST	0,7367	0,9005	0,6737
4 LAYER	TRAIN	0,7708	0,9332	0,6861
	TEST	0,6667	0,9235	0,5452

Tabella 6. Mediana di Dice, Precision e Recall di train set e test set rispettivamente della architettura con 5 layer e 4 layer.

Per confrontare le performance delle due architetture si è tenuto conto della mediana, perchè più rappresentativa della distribuzione degli indicatori.

Da *Figura 19* e *Figura 20* risulta evidente che la rete con 5 layer presenta performance migliori, sia sul Train set che sul Test set, in percentuali di metastasi individuate, Dice e Recall. La Precision risulta leggermente migliore nell'architettura con 4 layer, ma a discapito di un'ampia diminuzione della Recall, che indica una tendenza a sotto-segmentare le lesioni. Motivo per cui la rete con architettura a 5 layer è quella che presenta il migliore compromesso tra Precision e Recall.

CONFRO LAYER SUL TEST SET

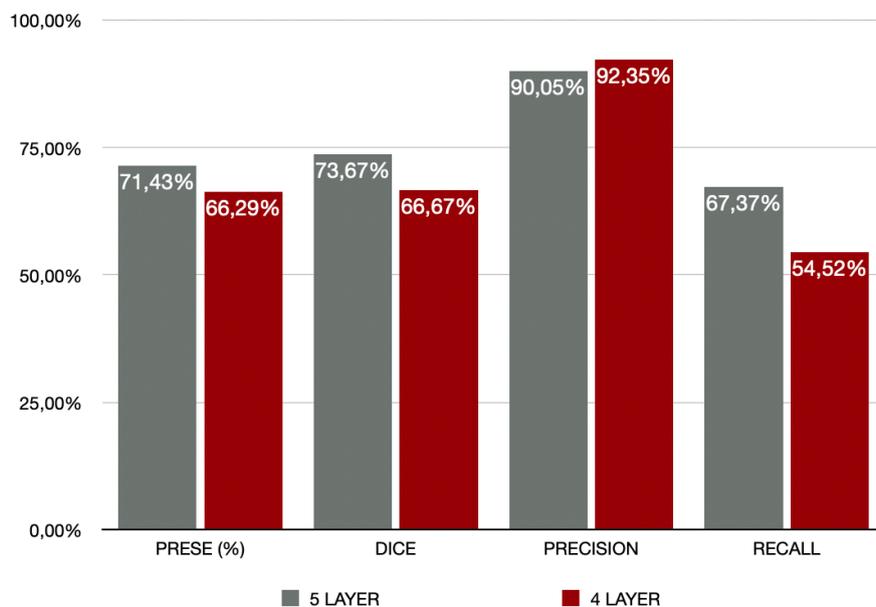


Figura 20. Percentuale metastasi prese, mediana Dice, Precision e Recall sul Test set rispettivamente dell'architettura con 5 e 4 layer

CONFRO LAYER SUL TRAIN SET

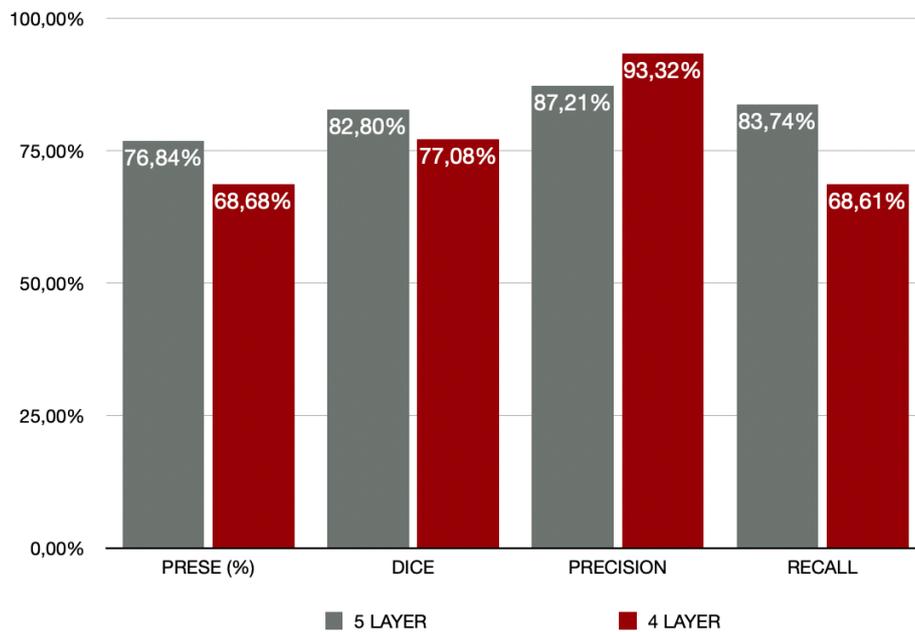


Figura 19. Percentuale metastasi prese, mediana Dice, Precision e Recall sul Train set rispettivamente dell'architettura con 5 e 4 layer

Successivamente è stata fatta un'analisi sui pazienti, per determinare il numero di lesioni individuate, sempre considerando un $Dice \geq 0,2$. Si sono presi in considerazione i risultati sia del Train set che del Test set rispettivamente delle due architetture. Seguono tabelle con i risultati.

Paziente	Metastasi Individuate	Paziente	Metastasi Individuate
29	0/1 (0%)	C1015	7/8 (87.5%)
C1002	0/1 (0%)	N1016	9/10 (90.0%)
C2002	0/1 (0%)	N2003	2/2 (100.0%)
N1022	0/3 (0%)	62	6/6 (100.0%)
N1038	1/7 (14.3%)	1	3/3 (100.0%)
27	1/4 (25.0%)	6	8/8 (100.0%)
33	1/3 (33.3%)	25	1/1 (100.0%)
N1049	1/3 (33.3%)	C1016	5/5 (100.0%)
55	1/3 (33.3%)	50	6/6 (100.0%)
23	2/5 (40.0%)	N1002	6/6 (100.0%)
N1040	4/10 (40.0%)	N1006	2/2 (100.0%)
C1010	7/13 (53.0%)	N1014	2/2 (100.0%)
N1015	1/2 (50.0%)	N1020	4/4 (100.0%)
N1057	4/7 (57.1%)	N1029	1/1 (100.0%)
18	2/3 (66.7%)	C1005	2/2 (100.0%)
C1001	7/10 (70.0%)	N1036	5/5 (100.0%)
N1017	3/4 (75.0%)	N1043	6/6 (100.0%)
N1055	3/4 (75.0%)	N2002	4/4 (100.0%)
37	7/9 (77.8%)	N2004	1/1 (100.0%)

Tabella 7. Percentuale di lesioni individuate di ogni paziente appartenente al Test set, derivanti dalla rete con 5 layer.

Paziente	Metastasi Individuate (%)	Paziente	Metastasi Individuate (%)
29	0/1 (0%)	C1015	6/8 (75.0%)
C1002	0/1 (0%)	N1016	3/4 (75.0%)
C2002	0/3 (0%)	N2003	3/4 (75.0%)
N1022	0/1 (0%)	62	7/9 (78.0%)
N1038	0/3 (0%)	1	2/2 (100.0%)
27	1/7 (14.0%)	6	6/6 (100.0%)
33	1/4 (25.0%)	25	3/3 (100.0%)
55	1/3 (33.0%)	50	1/1 (100.0%)
N1049	1/3 (33.0%)	C1005	6/6 (100.0%)
18	5/13 (38.0%)	C1016	6/6 (100.0%)
23	2/5 (40.0%)	N1002	4/4 (100.0%)
N1040	3/7 (43.0%)	N1006	2/2 (100.0%)
C1010	5/10 (50.0%)	N1014	2/2 (100.0%)
N1015	1/2 (50.0%)	N1020	1/1 (100.0%)
N1057	2/4 (50.0%)	N1029	2/2 (100.0%)
C1001	3/5 (60.0%)	N1036	5/5 (100.0%)
N1017	2/3 (67.0%)	N1043	6/6 (100.0%)
N1055	7/10 (70.0%)	N2002	10/10 (100.0%)
37	6/8 (75.0%)	N2004	1/1 (100.0%)

Tabella 8. Percentuale di lesioni individuate di ogni paziente appartenente al Test set, derivanti dalla rete con 4 layer.

Paziente	Metastasi Individuate (%)	Paziente	Metastasi Individuate (%)
12	0/2 (0%)	N1046	8/9 (89.0%)
20	0/2 (0%)	N1042	9/10 (90.0%)
46	0/1 (0%)	N1044	9/10 (90.0%)
C1003	0/2 (0%)	2	2/2 (100.0%)
C1024	0/1 (0%)	4	1/1 (100.0%)
C1027	0/1 (0%)	22	1/1 (100.0%)
N1034	0/1 (0%)	26	5/5 (100.0%)
C1017	1/6 (17.0%)	30	1/1 (100.0%)
C1008	1/4 (25.0%)	36	1/1 (100.0%)
C1011	1/4 (25.0%)	52	2/2 (100.0%)
17	1/3 (33.0%)	54	5/5 (100.0%)
C2004	1/3 (33.0%)	57	3/3 (100.0%)
N1009	1/3 (33.0%)	C1004	1/1 (100.0%)
N1041	1/3 (33.0%)	C1006	1/1 (100.0%)
5	2/4 (50.0%)	C1007	10/10 (100.0%)
N1024	1/2 (50.0%)	C1009	1/1 (100.0%)
N1056	2/4 (50.0%)	C1012	1/1 (100.0%)
N2001	1/2 (50.0%)	C1014	2/2 (100.0%)
28	4/7 (57.0%)	C1018	2/2 (100.0%)
38	4/7 (57.0%)	C1019	1/1 (100.0%)
41	4/7 (57.0%)	C1025	3/3 (100.0%)
C1021	4/7 (57.0%)	C1026	4/4 (100.0%)
45	3/5 (60.0%)	C2005	5/5 (100.0%)
49	2/3 (67.0%)	N1007	2/2 (100.0%)
C1020	2/3 (67.0%)	N1011	4/4 (100.0%)
N1008	2/3 (67.0%)	N1013	1/1 (100.0%)
N1012	2/3 (67.0%)	N1018	2/2 (100.0%)
N1048	4/6 (67.0%)	N1019	4/4 (100.0%)
N1053	4/6 (67.0%)	N1021	4/4 (100.0%)
31	5/7 (71.0%)	N1023	3/3 (100.0%)
63	6/8 (75.0%)	N1025	4/4 (100.0%)
66	9/12 (75.0%)	N1026	1/1 (100.0%)
C1023	3/4 (75.0%)	N1027	6/6 (100.0%)
32	7/9 (78.0%)	N1030	5/5 (100.0%)
10	4/5 (80.0%)	N1032	2/2 (100.0%)
C1013	4/5 (80.0%)	N1033	3/3 (100.0%)
C2003	8/10 (80.0%)	N1035	6/6 (100.0%)
N1004	8/10 (80.0%)	N1037	3/3 (100.0%)
N1010	4/5 (80.0%)	N1039	4/4 (100.0%)
N1005	5/6 (83.0%)	N1045	6/6 (100.0%)
19	6/7 (86.0%)	N1047	1/1 (100.0%)
C2001	7/8 (88.0%)	N1050	1/1 (100.0%)
N1003	7/8 (88.0%)	N1051	1/1 (100.0%)
N1031	7/8 (88.0%)	N1054	3/3 (100.0%)

Tabella 9. Percentuale di lesioni individuate di ogni paziente appartenente al Train set, derivanti dalla rete con 5 layer.

Paziente	Metastasi Individuate (%)	Paziente	Metastasi Individuate (%)
12	0/2 (0%)	C2003	7/10 (70.0%)
20	0/2 (0%)	N1004	7/10 (70.0%)
46	0/1 (0%)	19	5/7 (71.0%)
C1003	0/2 (0%)	63	6/8 (75.0%)
C1011	0/4 (0%)	N1011	3/4 (75.0%)
C1017	0/6 (0%)	N1021	3/4 (75.0%)
C1024	0/1 (0%)	C1013	4/5 (80.0%)
C1027	0/1 (0%)	N1010	4/5 (80.0%)
N1008	0/3 (0%)	N1044	8/10 (80.0%)
N1034	0/1 (0%)	N1005	5/6 (83.0%)
N1050	0/1 (0%)	N1027	5/6 (83.0%)
38	1/7 (14.0%)	C2001	7/8 (88.0%)
C1008	1/4 (25.0%)	N1031	7/8 (88.0%)
17	1/3 (33.0%)	N1042	9/10 (90.0%)
N1009	1/3 (33.0%)	4	1/1 (100.0%)
N1041	1/3 (33.0%)	22	1/1 (100.0%)
C1021	3/7 (43.0%)	26	5/5 (100.0%)
2	1/2 (50.0%)	30	1/1 (100.0%)
5	2/4 (50.0%)	31	7/7 (100.0%)
52	1/2 (50.0%)	36	1/1 (100.0%)
C1014	1/2 (50.0%)	54	5/5 (100.0%)
C1023	2/4 (50.0%)	C1004	1/1 (100.0%)
N1024	1/2 (50.0%)	C1006	1/1 (100.0%)
N1032	1/2 (50.0%)	C1007	10/10 (100.0%)
N1045	3/6 (50.0%)	C1009	1/1 (100.0%)
N1056	2/4 (50.0%)	C1012	1/1 (100.0%)
N2001	1/2 (50.0%)	C1018	2/2 (100.0%)
N1046	5/9 (56.0%)	C1019	1/1 (100.0%)
28	4/7 (57.0%)	C1026	4/4 (100.0%)
41	4/7 (57.0%)	C2005	5/5 (100.0%)
10	3/5 (60.0%)	N1007	2/2 (100.0%)
45	3/5 (60.0%)	N1013	1/1 (100.0%)
N1003	5/8 (63.0%)	N1018	2/2 (100.0%)
57	2/3 (67.0%)	N1019	4/4 (100.0%)
32	6/9 (67.0%)	N1023	3/3 (100.0%)
49	2/3 (67.0%)	N1025	4/4 (100.0%)
66	8/12 (67.0%)	N1026	1/1 (100.0%)
C1020	2/3 (67.0%)	N1030	5/5 (100.0%)
C1025	2/3 (67.0%)	N1033	3/3 (100.0%)
C2004	2/3 (67.0%)	N1035	6/6 (100.0%)
N1012	2/3 (67.0%)	N1037	3/3 (100.0%)
N1048	4/6 (67.0%)	N1039	4/4 (100.0%)
N1053	4/6 (67.0%)	N1047	1/1 (100.0%)
N1054	2/3 (67.0%)	N1051	1/1 (100.0%)

Tabella 10. Percentuale di lesioni individuate di ogni paziente appartenente al Train set, derivanti dalla rete con 4 layer.

Si sono poi creati dei range per vedere la distribuzione delle percentuali dei pazienti.

I range sono:

- Paziente con nessuna lesione individuata (0,0%)
- Paziente con alcune lesioni individuate (0,1% - 49,9%)
- Paziente con la maggior parte delle lesioni individuate (50,0% - 99,9%)
- Paziente con tutte le lesioni individuate (100,0%)

Con questa analisi non si riscontrano apprezzabili differenze tra le due architetture in termini di performance, anche se si ha una leggera diminuzione di pazienti con nessuna lesione individuata con l'architettura da 5 layer.

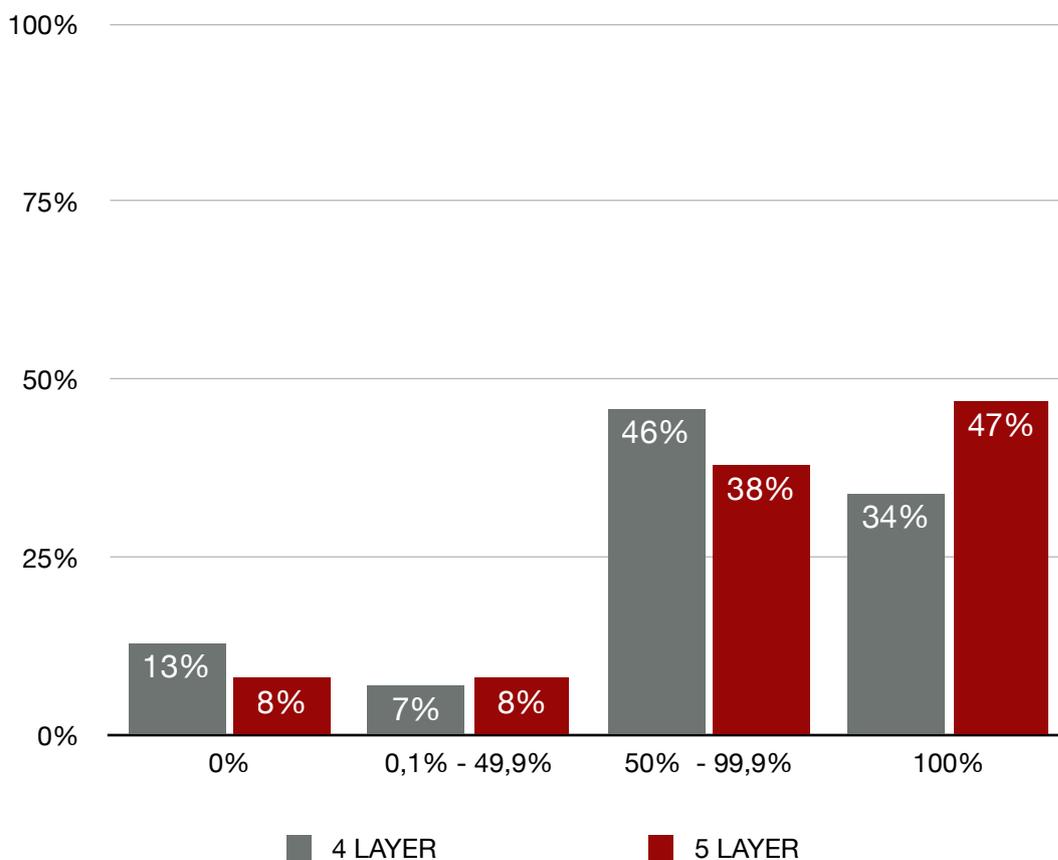


Figura 21. Percentuale di pazienti del Train test appartenete ai range indicanti le percentuali di lesioni individuate per paziente.

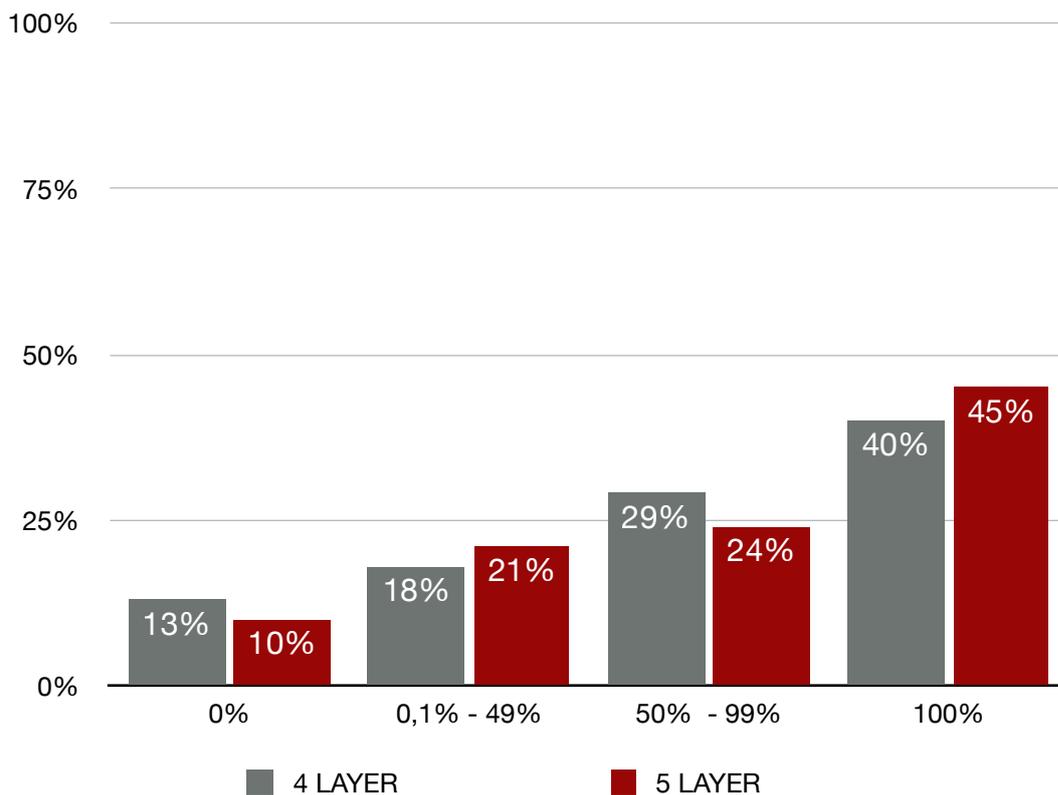


Figura 22. Percentuale di pazienti del Test test appartenete ai range indicanti le percentuali di lesioni individuate per paziente.

In ultima analisi, combinando le prove effettuate, la rete con 5 layer ha le performance migliori. Tutte le analisi che seguono sono state fatte con questo tipo di rete.

Analisi per pool di provenienza

Dopo aver individuato l'architettura migliore, si è andato a valutare se le performance della rete sono influenzate dalla provenienza dei pazienti.

Per fare ciò si sono divisi i pazienti tra pazienti di prima linea e pazienti di terza linea. I pazienti di prima linea sono stati ulteriormente divisi per centro di provenienza. Più nel dettaglio si ha:

- Pazienti di terza linea
- Pazienti di prima linea dell'Istituto Di Candiolo
- Pazienti di prima linea dell'Ospedale Niguarda

Come analogamente fatto nel confronto delle architetture della rete, si sono contate il numero di lesioni individuate e non individuate (*Tabella 11*), media (*Tabella 12*) e mediana (*Tabella 15*) degli indicatori di Dice, Precision e Recall.

Le analisi sono state fatte sia sul Train set che sul Test set.

METASTASI INDIVIDUATE E NON INDIVIDUATE

		NON INDIVIDUATE	INDIVIDUATE
TRAIN	TERZA LINEA	32/110	78/110
	PRIMA LINEA CANDIOLO	26/89	63/89
	PRIMA LINEA NIGUARDA	24/166	142/166
TEST	TERZA LINEA	23/62	39/62
	PRIMA LINEA CANDIOLO	10/39	29/39
	PRIMA LINEA NIGUARDA	17/74	57/74

Tabella 11. Numero di metastasi considerate individuate e non individuate, rispettivamente dei Train set e Test set appartenenti ai pazienti di terza linea, prima linea dell'Istituto di Candiolo e prima linea dell'Ospedale Niguarda.

MEDIA PER POOL DI PROVENIENZA

		DICE	PRECISION	RECALL
TRAIN	TERZA LINEA	0,7312	0,7949	0,7452
	PRIMA LINEA CANDIOLO	0,7903	0,8187	0,8195
	PRIMA LINEA NIGUARDA	0,7833	0,8603	0,7650
TEST	TERZA LINEA	0,6365	0,7800	0,6135
	PRIMA LINEA CANDIOLO	0,6846	0,8124	0,6378
	PRIMA LINEA NIGUARDA	0,6721	0,9138	0,5982

Tabella 12. Media di Dice, Precision e Recall delle lesioni di Train set e Test set appartenenti rispettivamente ai pazienti di terza linea, prima linea dell'Istituto di Candiolo e prima linea dell'Ospedale Niguarda

MEDIANA PER POOL DI PROVENIENZA

		DICE	PRECISION	RECALL
TRAIN	TERZA LINEA	79,96%	82,44%	82,95%
	PRIMA LINEA CANDIOLO	84,75%	87,92%	88,78%
	PRIMA LINEA NIGUARDA	84,26%	88,23%	82,36%
TEST	TERZA LINEA	70,58%	79,86%	65,09%
	PRIMA LINEA CANDIOLO	72,31%	88,09%	65,36%
	PRIMA LINEA NIGUARDA	75,97%	96,42%	67,47%

Tabella 13. Mediana di Dice, Precision e Recall delle lesioni di Train set e Test set appartenenti rispettivamente ai pazienti di terza linea, prima linea dell'Istituto di Candiolo e prima linea dell'Ospedale Niguarda

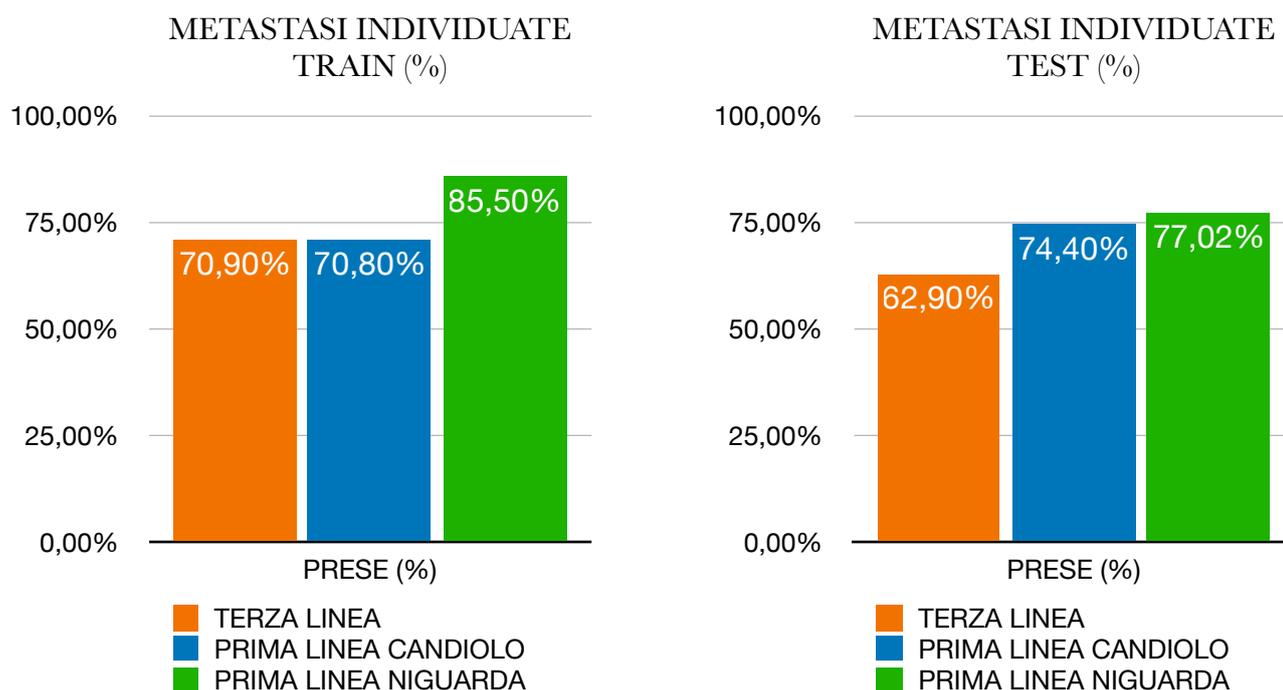


Figura 23. Percentuali di lesioni individuate di Train set e Test set appartenenti rispettivamente ai pazienti di terza linea, prima linea dell'Istituto di Candiolo e prima linea dell'Ospedale Niguarda

MEDIANA PER POOL DI APPARTENENZA TRAIN SET

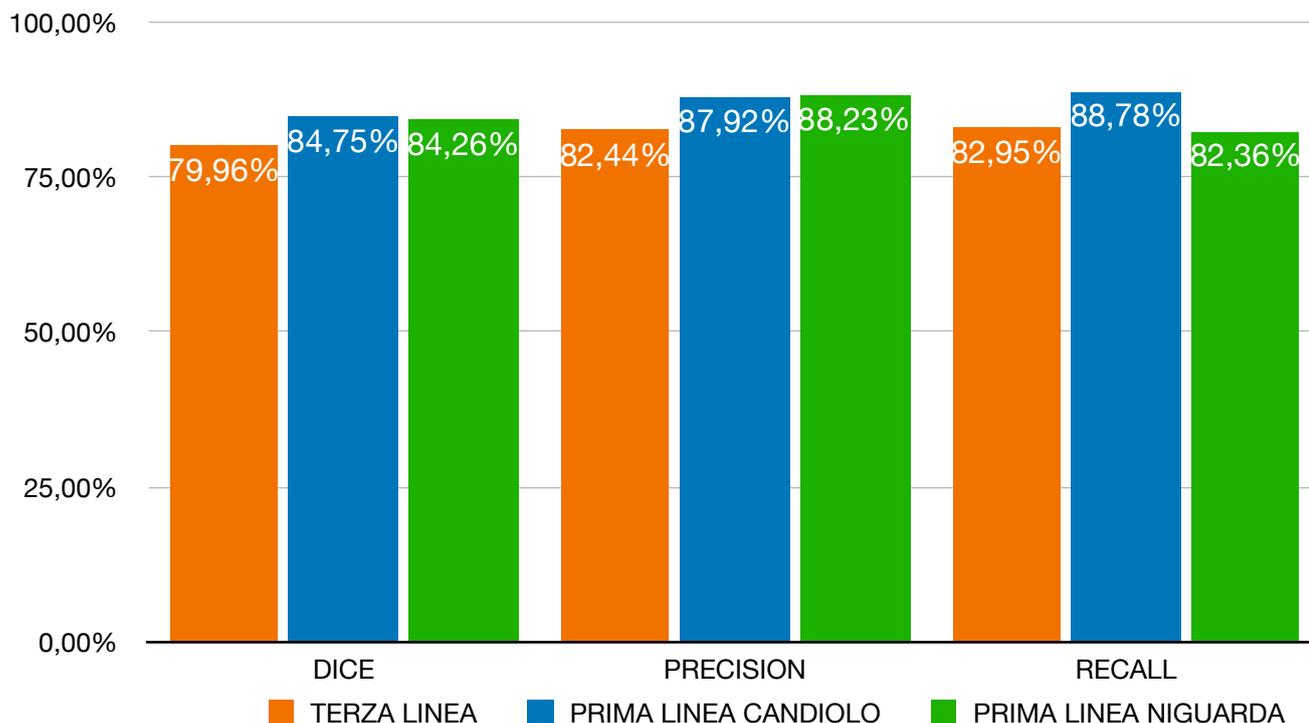


Figura 24. Mediana di Dice, Precision e Recall delle lesioni del Train set appartenenti rispettivamente ai pazienti di terza linea, prima linea dell'Istituto di Candiolo e prima linea dell'Ospedale Niguarda

MEDIANA PER POOL DI APPARTENENZA TEST SET

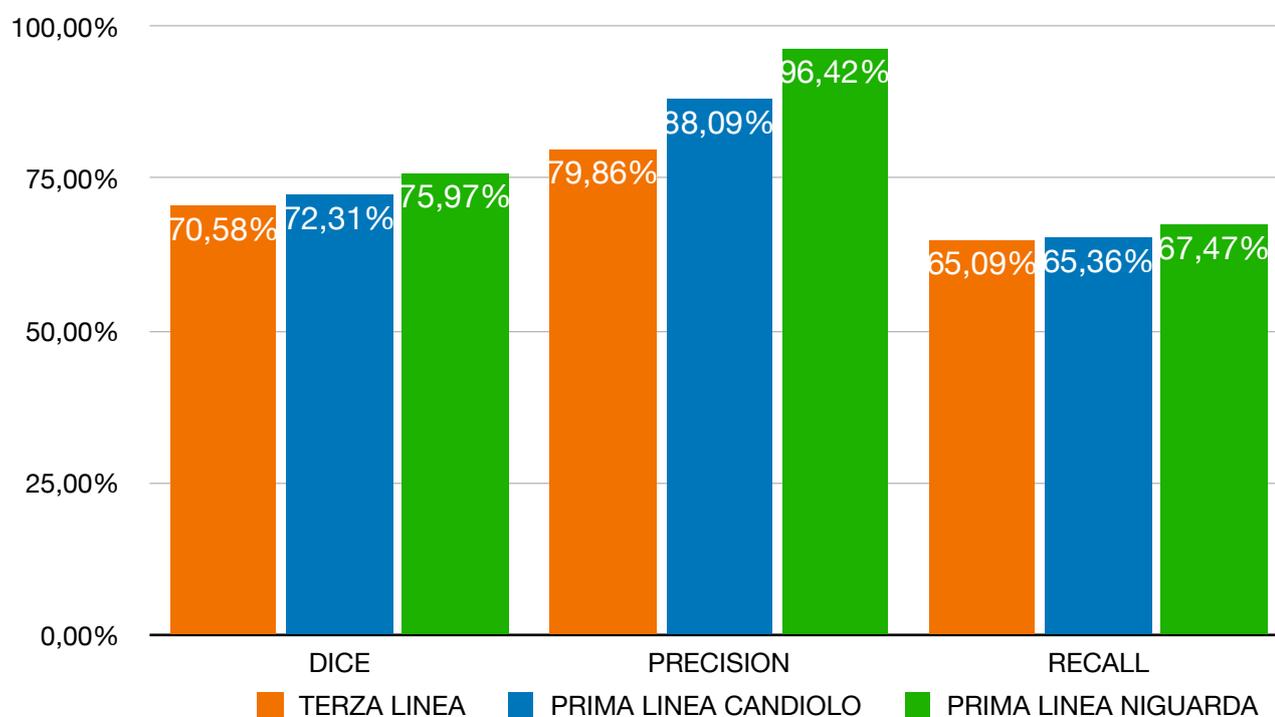


Figura 25. Mediana di Dice, Precision e Recall delle lesioni del Test set appartenenti rispettivamente ai pazienti di terza linea, prima linea dell'Istituto di Candiolo e prima linea dell'Ospedale Niguarda

Dai risultati ottenuti si evince che i pazienti di terza linea hanno performance generalmente inferiori rispetto a quelli di prima linea (*Figura 23*). Questo era un risultato atteso in quanto, essendo pazienti con alle spalle diversi cicli di terapia, presentano delle CT con un fegato meno uniforme e con più metastasi. Tutto questo aumenta notevolmente la complessità dell'immagine analizzata dalla rete, fattore che si traduce in un abbassamento delle performance di segmentazione.

Per quanto riguarda i pazienti di prima linea, questi hanno una performance paragonabile, leggermene a favore dei pazienti provenienti dall'Ospedale Niguarda. Fenomeno che viene spiegato dal maggiore numero di pazienti disponibili per l'allenamento della rete provenienti da quest'ultimo centro: 166 lesioni dai pazienti dell'Ospedale Niguarda contro le 89 lesione provenienti dall'Istituto di Candiolo.

Analisi intensità dei pixel

L'intensità di ogni pixel è un parametro fondamentale nella parte di apprendimento di una rete e le metastasi possono presentare valori di intensità molto diversi tra di loro. Si è perciò indagato se l'intensità media dei pixel della metastasi incidesse in qualche modo sulle performance della rete.

Per questa analisi si sono usate le immagini normalizzate (*Equazione 3*), per eliminare la dipendenza dell'intensità dalla qualità della CT.

Si è calcolata la media dell'intensità di tutti i pixel appartenenti alla metastasi, per ogni lesione. Successivamente si è calcolata media e deviazione standard su tutti i valori medi di metastasi, per Train set e Test set.

A questo punto si sono prese in considerazione i Dice di quelle metastasi con una intensità media non inclusa nell'intervallo definito dall'*Equazione 6*. Questo per vedere se esiste una dipendenza tra il rate di individuazione della metastasi e la sua intensità media.

$$If(mean - 2 \cdot STD \leq Intensità_{lesione} \leq mean + 2 \cdot STD)$$

Equazione 6. Intervallo di valori assunto dal 95% di una distribuzione Normale

Si evince dunque (*Tabella 14*) che i valori di intensità delle metastasi non sono un fattore che incide in maniera negativa sulle performance della rete.

INTENSITÀ MEDIA PIXEL

	MEDIA TUTTE METASTASI	DEV TUTTE METASTASI	PERCENTUALE METASASI INDIVIDUATE
TRAIN	0,4416	0,0933	89/119 (74,8%)
TEST	0,384	0,281	2/5 (40%)

Tabella 14. Intensità media e deviazione standard di tutte le metastasi unitamente al numero di lesioni fuori dal range, considerate prese, sul numero di lesioni totali rispettivamente del Train set e del Test set

INTENSITÀ MEDIA LESIONI TRAIN SET

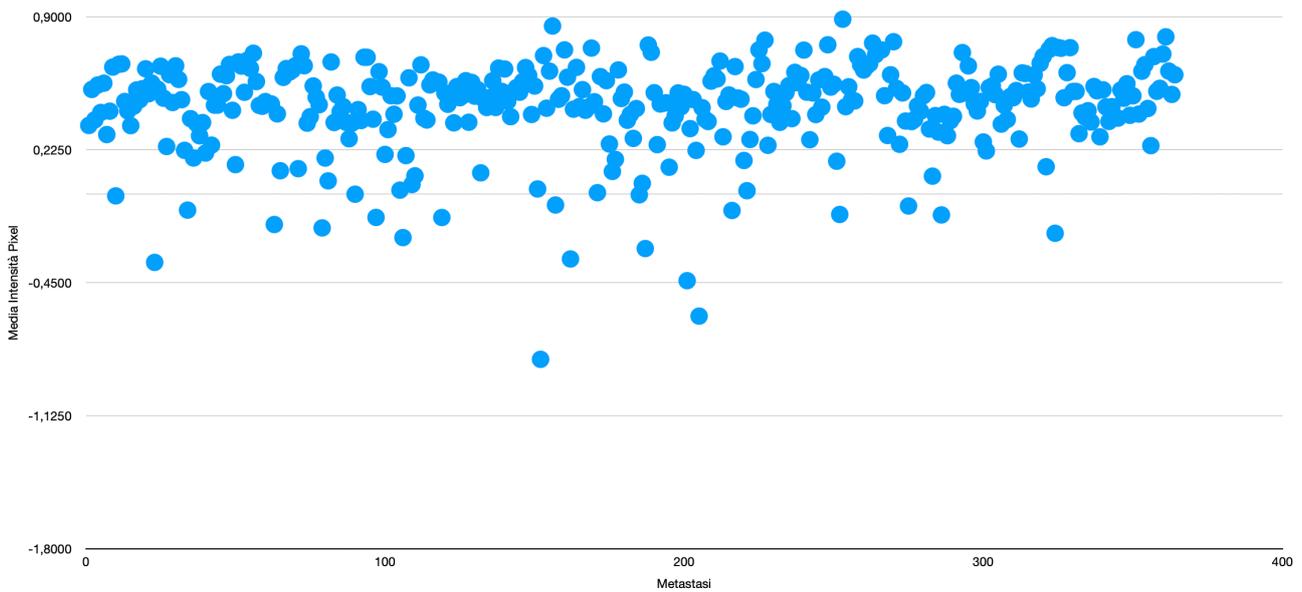


Figura 26. Rappresentazione della distribuzione dei valori di intensità dei pixel delle metastasi appartenenti al Train set

INTENSITÀ MEDIA LESIONI TEST SET

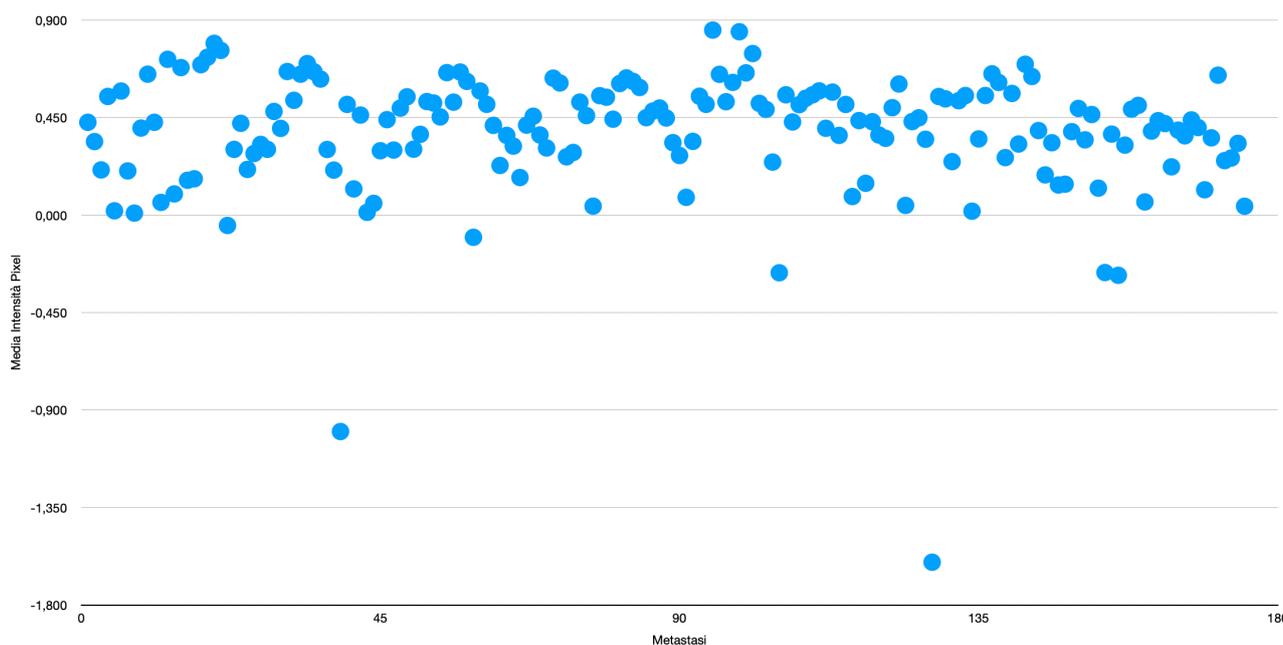


Figura 27. Rappresentazione della distribuzione dei valori di intensità dei pixel delle metastasi appartenenti al Test set

Analisi per grandezza della lesione

Altro parametro che si è analizzato è la grandezza della lesione. Per identificare la grandezza di una lesione si è preso in considerazione la baseline del diametro (in mm) e la profondità del volume 3D (in mm) della metastasi.

Il diametro è stato calcolato in fase di segmentazione manuale dall'operatore, mentre la profondità è stata calcolata conoscendo le dimensioni dei pixel in mm della CT, che sono state moltiplicate per il numero di slice in cui è presente la lesione.

Per quanto riguarda la baseline si sono distinti due range:

- lesioni piccole (baseline < 20 mm)
- lesioni grandi (baseline \geq 20mm)

Successivamente si sono andate a contare le lesioni all'interno dei due range con un Dice maggiore di 0,2 che le identifica come metastasi individuate.

Dai risultati ottenuti (Tabella 15, Figura 28) si evince che esiste una dipendenza evidente della performance rispetto alla baseline della metastasi, a favore delle metastasi grandi.

METASTASI INDIVIDUATE PER GRANDEZZA

	PICCOLE (<20mm)	GRANDI (\geq 20mm)
TRAIN	84/139 (60,42%)	190/211 (90,05%)
TEST	20/45 (44,44%)	105/130 (80,77%)

Tabella 15. Numero e relativa percentuale di metastasi piccole e grandi individuate, rispettivamente per Train set e Test set

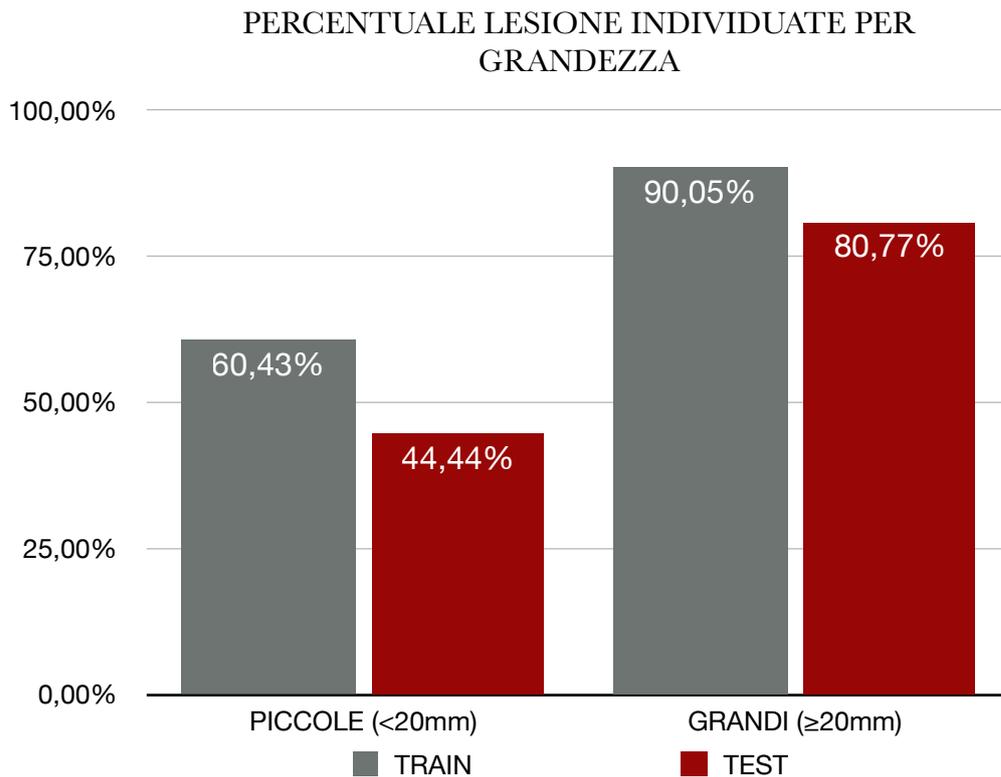


Figura 28. Percentuale di lesioni individuate per grandezza appartenenti rispettivamente del Train set e del Test set

Per quanto riguarda invece la profondità delle lesioni sono stati individuati 3 range differenti, più nel dettaglio:

- lesioni piccole (profondità \leq 10mm)
- lesioni medie (10mm<profondità<20mm)
- lesioni grandi (profondità \geq 20mm)

Anche in questo caso è chiara la dipendenza dalla dimensione della lesione (*Tabella 16, Figura 29*)

METASTASI INDIVIDUATE PER PROFONDITÀ

	PICCOLE ($\leq 10\text{mm}$)	MEDIE (10 - 20 mm)	GRANDI ($\geq 20\text{mm}$)
TRAIN	25/56 (44,64%)	97/138 (70,29%)	160/170 (94,12%)
TEST	10/26 (38,46%)	40/61 (65,57%)	75/88 (85,23%)

Tabella 16. Numero e relativa percentuale di metastasi individuate per ogni range, rispettivamente per Train set e Test set

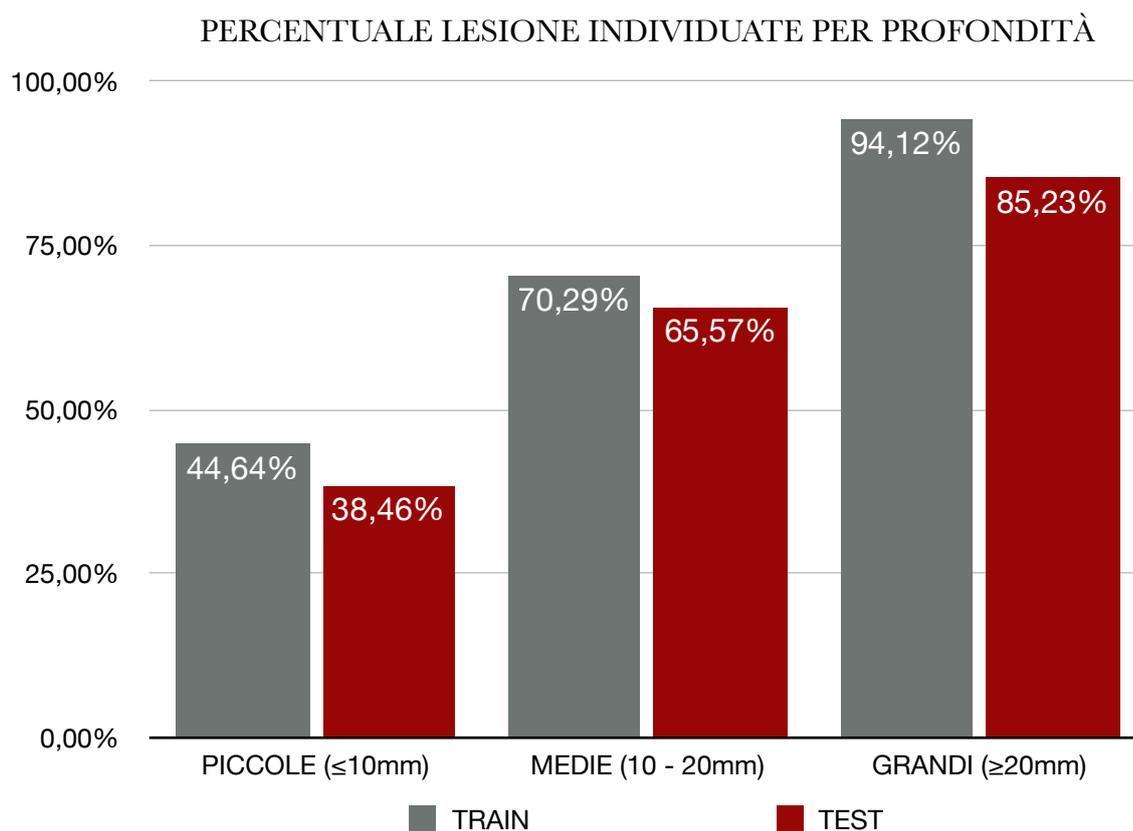


Figura 29. Percentuale di lesioni individuate per profondità appartenenti rispettivamente del Train set e del Test set

Analisi per tipologia di lesione

Dalle analisi precedenti risulta chiaramente che le metastasi con dimensione maggiore vengono individuate meglio dalla rete. Dunque si è preso questo gruppo di lesione, con baseline ≥ 20 e si sono divise le lesioni in ulteriori 4 gruppi, in base alla tipologia. In particolare si hanno:

- Metastasi cistiche (1)
- Metastasi a bersaglio (2)
- Metastasi disomogenee (3)
- Metastasi omogenee (4)

Si sono andate a contare il numero di metastasi considerate individuate per tutte e quattro le tipologie di lesioni.

I risultati sono riportati nella (Tabella 17, Figura 30).

Dopo aver effettuato il χ^2 test (9), risulta che non esiste una differenza statisticamente significativa tra le percentuali di lesioni identificate nelle diverse tipologie di metastasi.

METASTASI INDIVIDUATE PER TIPOLOGIA

	CISITCHE (1)	A BERSAGLIO (2)	DISOMOGENEE (3)	OMOGENEE (4)
TRAIN	27/29 (93,10%)	23/27 (85,19%)	37/40 (92,50%)	45/47 (95,74%)
TEST	14/16 (87,50%)	6/9 (66,67%)	35/43 (81,40%)	17/18 (94,44%)

Tabella 17. Numero e relativa percentuale di metastasi individuate per ogni tipo, rispettivamente per Train set e Test set

PERCENTUALE LESIONE INDIVIDUATE PER TIPOLOGIA

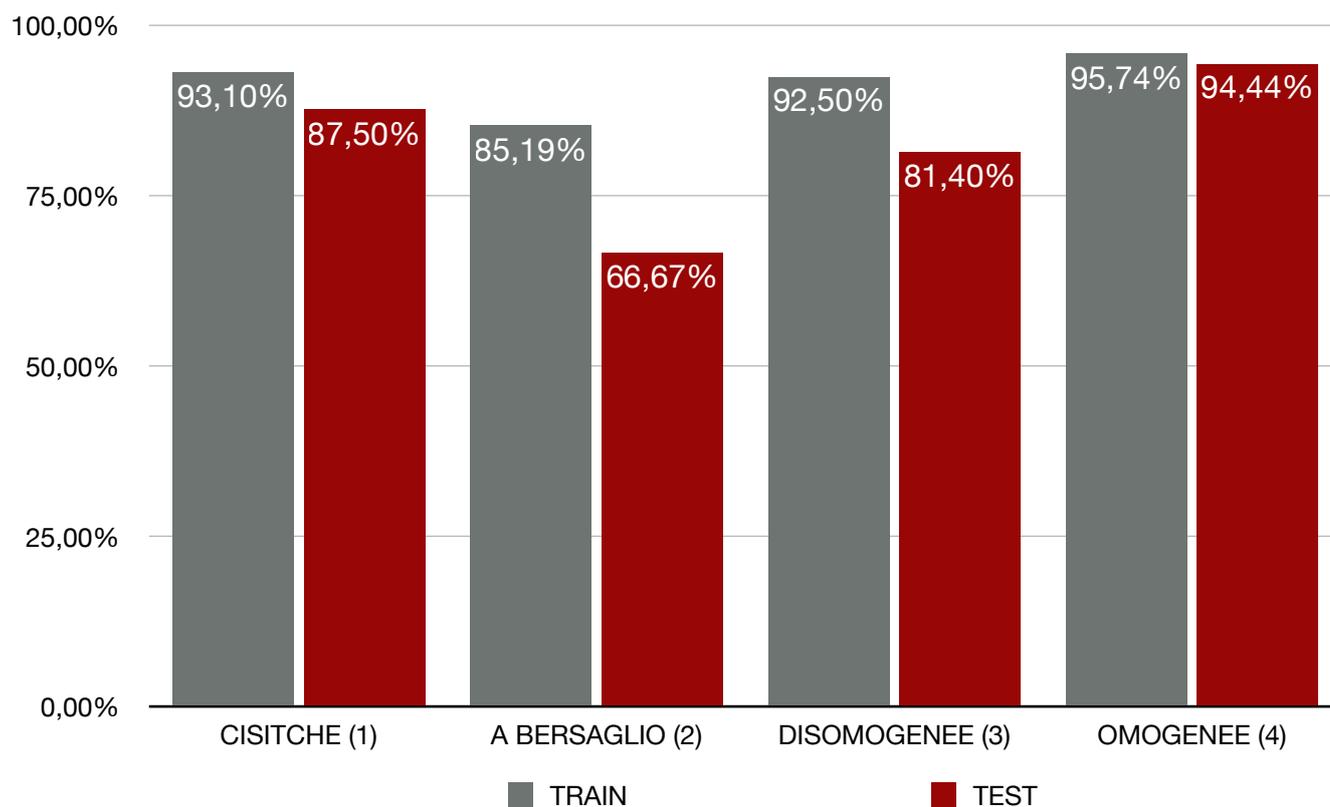


Figura 30. Percentuale di lesioni individuate per tipologia appartenenti rispettivamente del Train set e del Test set

Esempi di segmentazione

Di seguito sono riportati alcuni esempi di segmentazione automatica di metastasi compiuti dalla rete, messe a confronto con la segmentazione manuale. Si riporta un paziente per ogni gruppo di provenienza.

Nella *Figura 31* è rappresentato un paziente di terza linea con 2 lesioni:

- La lesione 1 (L1) con un Dice = 0,851;
- La lesione 2 (L2) con un Dice = 0,772;

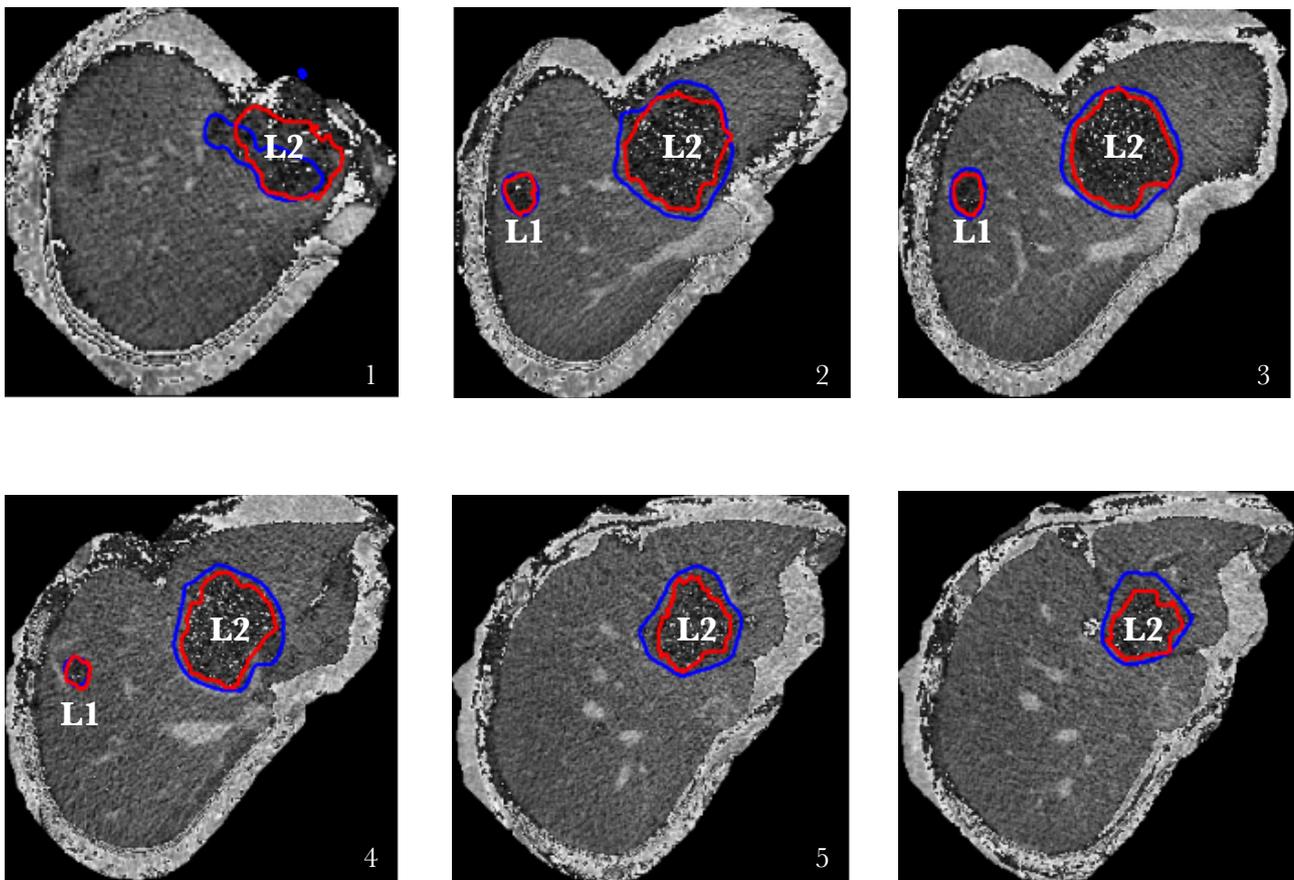


Figura 31. Paziente 1, in rosso la segmentazione manuale e in blu la segmentazione automatica.

Nella *Figura 32* è rappresentato un paziente di prima linea proveniente dall'Ospedale Niguarda. Questo paziente presenta una lesione, che viene segmentata con un $Dice = 0,748$. Il Dice è influenzato dal fatto che nella prima parte della CT la metastasi si trova in una parte non individuata dall'algoritmo di segmentazione del fegato, e quindi non viene considerata dalla rete. Questo spiega il Dice relativamente più basso delle aspettative rispetto all'analisi qualitativa della segmentazione.

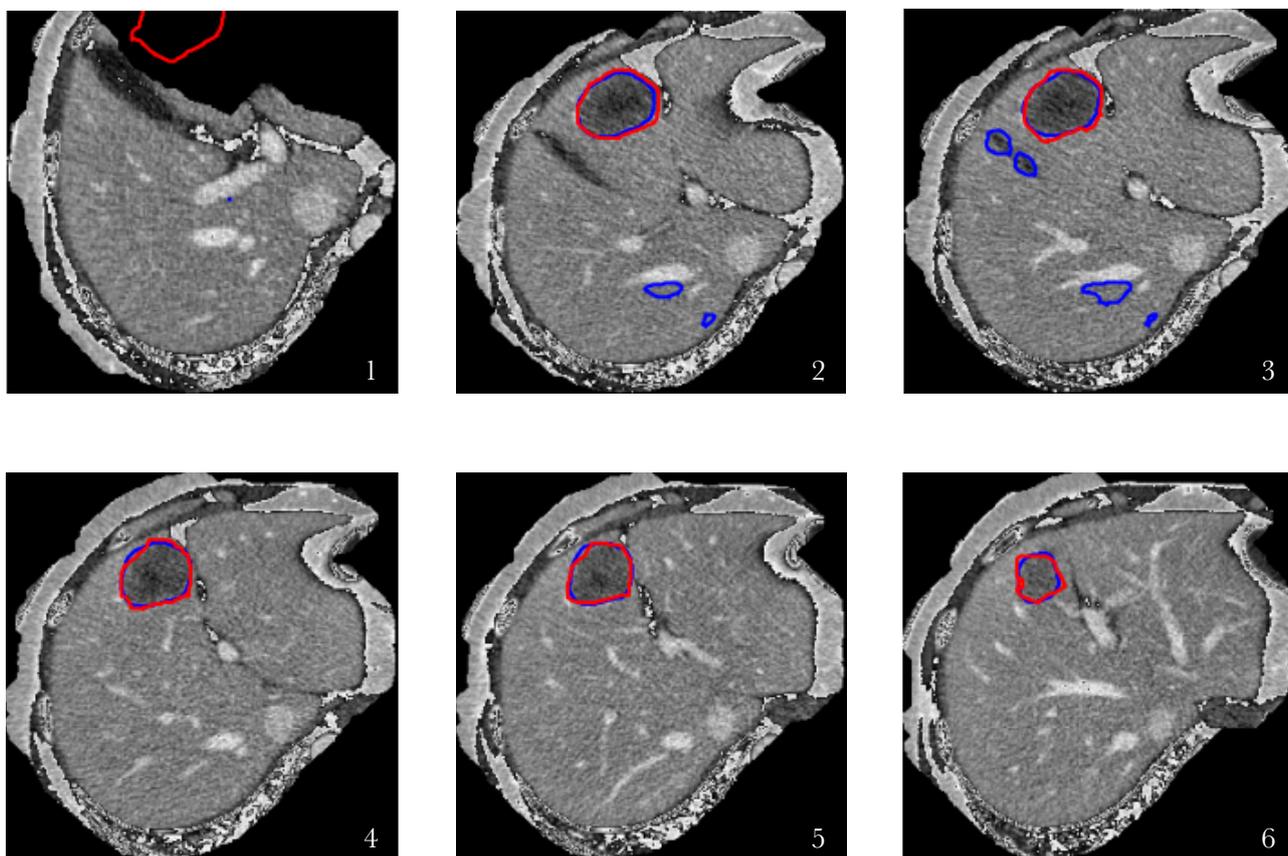


Figura 32. Paziente N2004 in rosso la segmentazione manuale e in blu la segmentazione automatica.

Nella *Figura 33* è rappresentato un paziente di prima linea dell'Istituto di Candiolo. Questo paziente presenta in totale 6 lesioni, nello specifico:

- Lesione 1 con un $Dice = 0,858$ (L1)
- Lesione 2 con un $Dice = 0,919$ (L2)
- Lesione 3 con un $Dice = 0,911$ (L3)
- Lesione 4 con un $Dice = 0,913$ (L4)
- Lesione 5 con un $Dice = 0,585$ (L5)
- Lesione 6 con un $Dice = 0,864$ (L6)

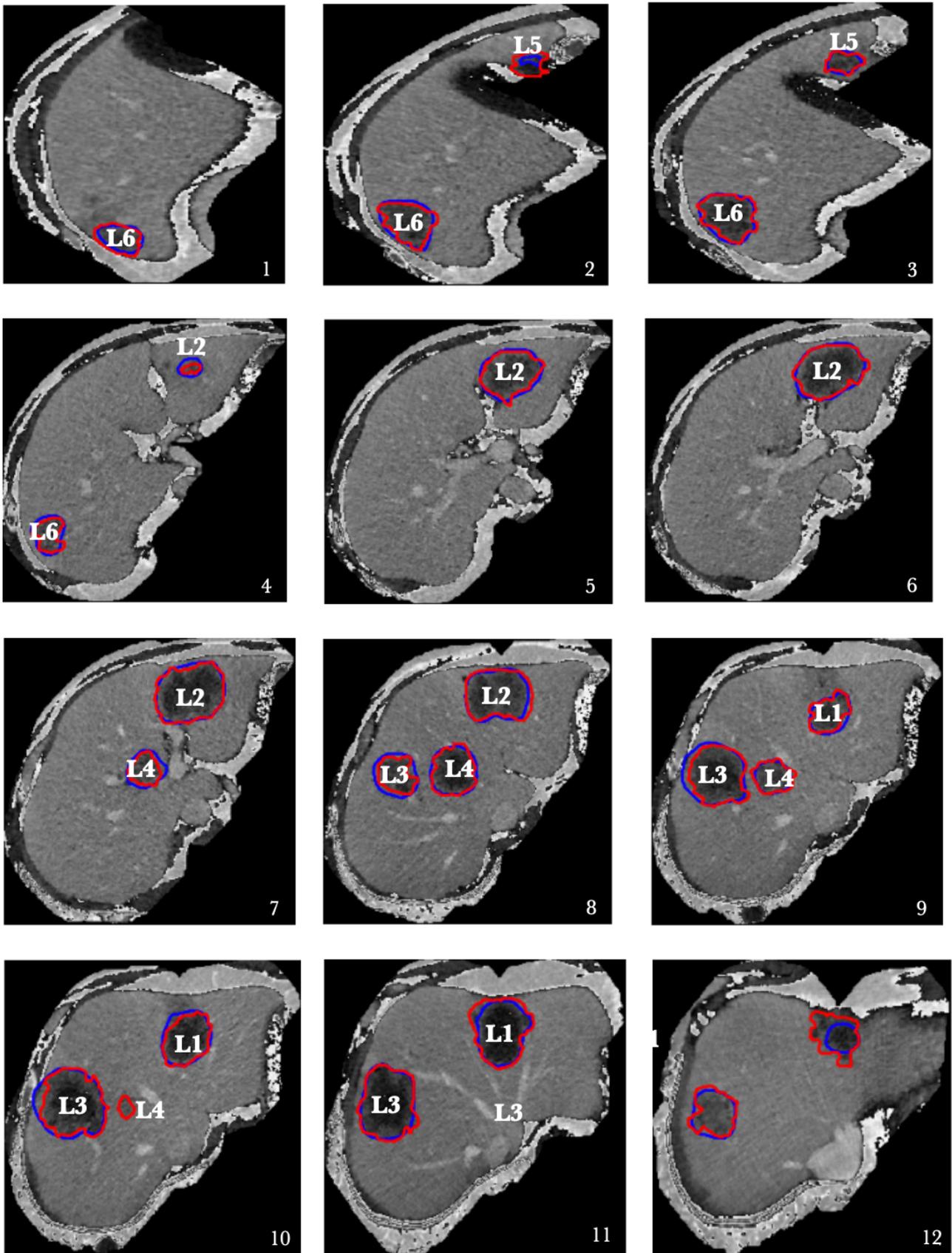


Figura 33. Paziente C1005, in rosso la segmentazione manuale e in blu la segmentazione automatica

Analisi lesioni critiche

Dalle analisi svolte sui pazienti che presentano solo lesioni non individuate, si sono ricondotte le cause della mancata individuazione ad alcune cause principali.

Le cause riscontrate sono:

- Lesioni poste al bordo del fegato (1)
- Intensità e distribuzione pixel simile al fegato (2)
- Segmentazione del fegato errata (3)
- Altre cause (4)

Si è andata a verificare questa ipotesi indagando sul gruppo di lesioni classificate meglio, cioè quelle con baseline $\geq 20\text{mm}$, per vedere se la mancata individuazione di quest'ultime può essere imputata ad una delle cause riportate.

Dai risultati ottenuti (*Figura 34*) si può concludere che la rete incontra delle difficoltà nell'individuare le metastasi che si trovano sul bordo del fegato.

Di seguito sono riportati degli esempi per ognuno dei gruppi illustrati precedentemente. Degno di nota è la *Figura 35* in cui si ha la segmentazione manuale di una delle due lesioni fuori dai limiti individuati dall'algoritmo di segmentazione del fegato discusso in precedenza; questo ha portato la rete a ignorare l'esistenza della lesione. Invece si può notare come l'altra lesione presenti una segmentazione accurata.

CAUSE DI CRITICITÀ

Bordo fegato (1)	Intensità pixel (2)	Segmentazione fegato (3)	Altro (4)
28/46 (60,87%)	5/46 (10,87%)	4/46 (8,70%)	9/46 (19,57%)

Tabella 18. Cause di criticità con relativo numero e percentuale di metastasi appartenenti ad ogni gruppo.

PERCENTUALI DI LESIONE PER CRITICITÀ

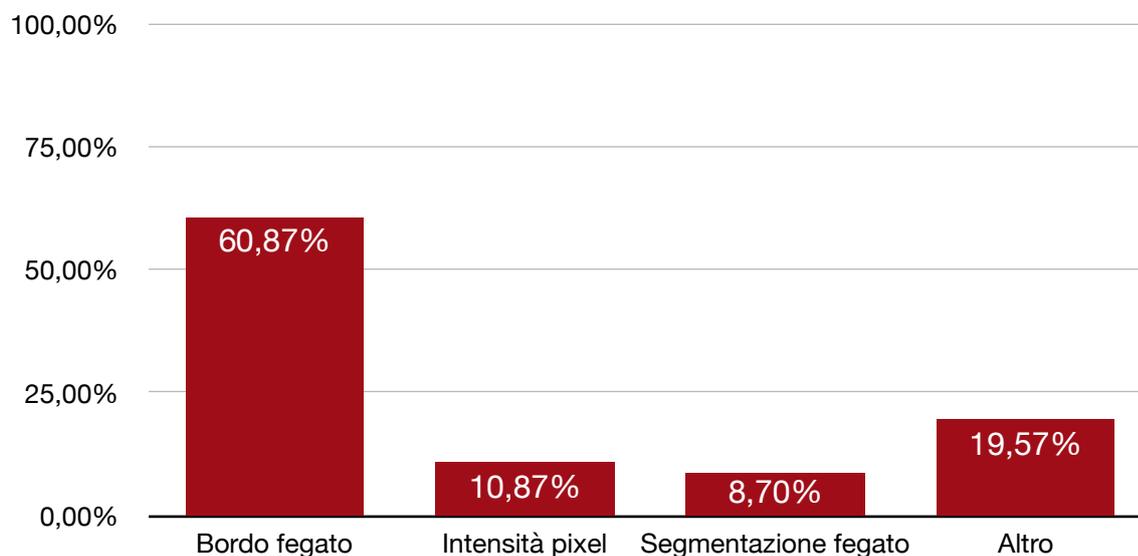


Figura 34. Percentuale di metastasi appartenenti alle diverse criticità

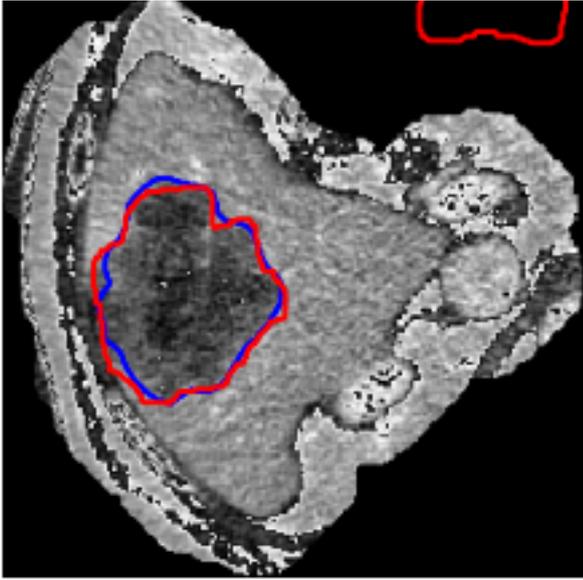


Figura 35. Paziente 49, esempio di lesione non individuata a causa della segmentazione del fegato non precisa, dove il rosso è la segmentazione manuale, il blu la segmentazione automatica

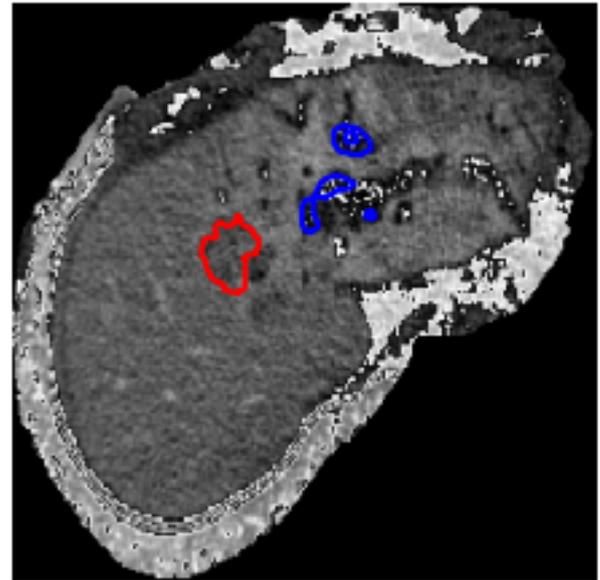


Figura 36. Paziente 12, esempio di lesione non segmentata dove il rosso è la segmentazione manuale, il blu la segmentazione automatica

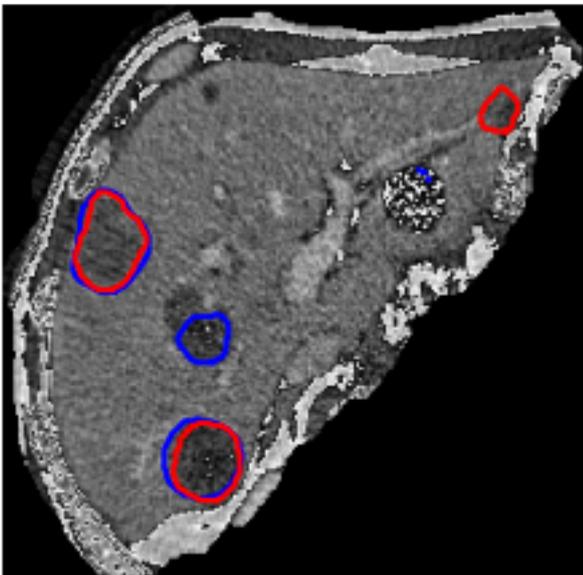


Figura 37. Paziente N1031, esempio di lesione non individuata sul bordo del fegato, dove il rosso è la segmentazione manuale, il blu la segmentazione automatica

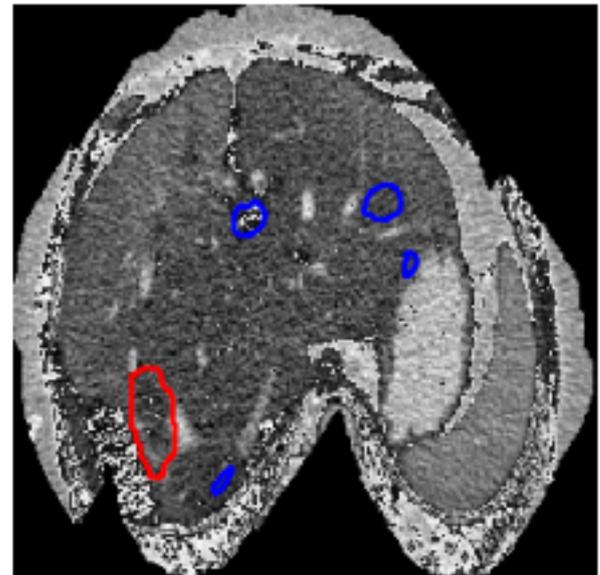


Figura 38. Paziente 46, esempio di lesione non individuata con intensità e distribuzione pixel simile al fegato dove il rosso è la segmentazione manuale, il blu la segmentazione automatica

DISCUSSIONE

In questa Tesi si è sviluppato un sistema automatico di segmentazione di metastasi epatiche tramite l'implementazione di una rete Unet, da affiancare ad un sistema di caratterizzazione delle lesioni epatiche per prevedere la risposta di quest'ultime alle terapie tumorali.

I valori ottenuti sono comparabili a quelli dello stato dell'arte. In particolare Vorontsov *et al* (6), su un totale di 75 lesioni (35 con dimensione $<20\text{mm}$ e 40 con dimensione $\geq 20\text{mm}$), ottiene rispettivamente per le lesioni piccole e per quelle grandi un Dice medio del 53% e del 68%; una Precision rispettivamente del 83% e del 94%; una Recall rispettivamente del 71% e del 85%.

La rete sviluppata in questa Tesi, su un insieme di 175 campioni esterni alla fase di allenamento (45 con dimensione $<20\text{mm}$ e 130 con dimensione $\geq 20\text{mm}$), individua il 71% di metastasi e le segmenta con dei valori medi di Dice, Precision e Recall rispettivamente dello 67%, 85% e 64%. Mentre presenta valori mediani di Dice, Precision e Recall rispettivamente di 74%, 90% e 67%.

Successivamente si sono anche indagate le cause di criticità del modello, ponendo le basi per miglioramenti futuri. In particolare si è stabilito che le performance della rete sono indipendenti dall'intensità media dei pixel delle metastasi, ma sono fortemente dipendenti dalle dimensioni fisiche di quest'ultime. In particolare le metastasi con baseline $< 20\text{ mm}$ sono individuate dalle rete con una percentuale del 44%, mentre quelle con baseline $\geq 20\text{ mm}$ hanno una percentuale del 81%.

Inoltre non si sono riscontrate differenze statisticamente significative nella qualità della segmentazione tra le varie tipologie di metastasi epatiche.

Un miglioramento attuabile sarebbe, disponendo delle segmentazione manuali, quello di passare ad un sistema di segmentazione per il fegato supervisionato. Questo migliorerebbe le prestazioni della segmentazione della regione epatica, rendendo l'input della rete ancora più specifico.

Inoltre, con la conoscenza raccolta nelle fasi di analisi si potrebbe costruire un Train set più significativo. Infatti si è riscontrato un calo nella percentuale di metastasi individuate nei pazienti di terza linea. Questo può essere evitato allargando il numero di pazienti e includendo, in fase di apprendimento, quei casi limite su cui l'algoritmo fallisce la segmentazione, così da rendere più robusta la rete.

BIBLIOGRAFIA

1. Bray, F., Ferlay, J., Soerjomataram, *et al*: Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA CANCER J CLIN, 2018.
2. Sheth KR e Clary BM: Management of hepatic metastases from colorectal cancer. Clin Colon Rectal Surg, 2005.
3. Yingyi Qi1 e Wei Xiong: Semi-automatic Segmentation of Liver Tumors from CT Scans Using Bayesian Rule-based 3D Region Growing. MICCAI, 2008.
4. Vorontsov, E., Tang, *et al*:. Metastatic liver tumour segmentation with a neural network-guided 3D deformable model. Medical & Biological Engineering & Computing, 2017.
5. Sun, C., Guo, S. *et al*: Automatic segmentation of liver tumors from multiphase contrast-enhanced CT images based on FCNs. Artificial Intelligence in Medicine, 2017.
6. Vorontsov, E., Cerny, M., *et val*: Deep Learning for Automated Segmentation of Liver Lesions at CT in Patients with Colorectal Cancer Liver Metastases. Radiology: Artificial Intelligence, 2019
7. Ronneberger O, Fischer P and Brox T: U-net: Convolutional networks for biomedical image segmentation. Lect Notes Comput Sci, 2015.
8. Diederik P. Kingma and Jimmy Lei Ba. Adam: a Method for Stochastic Optimization. International Conference on Learning Representations, 2015.
9. Richardson, J. T. E.: The analysis of 2×2 contingency tables-Yet again. Statistics in Medicine, 2011.