

POLITECNICO DI TORINO

Laurea Magistrale in Ingegneria Biomedica



Tesi di Laurea Magistrale

**STUDIO DI ALGORITMI PER LA
SEGMENTAZIONE AUTOMATICA DI
LESIONI POLMONARI SU IMMAGINI
CT**

Relatori

Prof. Massimo SALVI

Dr.ssa Matilde COSTA

Candidata

Cristina Pia PROCHILO

Anno Accademico 2023/2024

Abstract

Il cancro al polmone non a piccole cellule (NSCLC) rappresenta l'85% delle neoplasie polmonari che colpiscono la popolazione mondiale. La diagnosi prevede l'esecuzione di una tomografia computerizzata (CT) con mezzo di contrasto seguita dall'esame di un campione biotico che confermerà o meno la presenza del tumore. La segmentazione accurata della lesione polmonare rappresenta non solo un passo rilevante per una corretta caratterizzazione del tumore, ma consente di eseguire analisi avanzate come la Radiomica, che consiste nell'estrazione di un elevato numero di caratteristiche quantitative dalle immagini mediche, come la densità, la forma, la texture. Queste possono essere correlate con informazioni cliniche, permettendo di ottenere una migliore comprensione dell'avanzamento del tumore e delle risposte terapeutiche, migliorando la pianificazione e la personalizzazione delle cure. A tal proposito, l'obiettivo principale di questo studio è sviluppare una soluzione che consenta un'identificazione automatica delle lesioni polmonari, riducendo così i tempi di intervento manuale da parte del clinico e limitando la variabilità inter-operatore. A tale scopo, viene proposto un metodo basato sul Deep Learning. Lo studio è stato eseguito su CT toraciche di 88 pazienti, con le corrispondenti segmentazioni manuali delle zone di interesse utilizzate come riferimento. Durante l'analisi sono state individuate diverse criticità e si sono sviluppate soluzioni mirate a garantire un corretto funzionamento dell'algoritmo. Sono stati presi in considerazione tre approcci che differiscono per tipologia di taglio effettuato sulla CT, quali ROI attorno al polmone, ROI attorno al centro di massa ed estrazione del parenchima polmonare. Quest'ultimo metodo si è rilevato essere la strategia più efficace. Lavorando sulle fette bidimensionali di ogni volume, il primo step di quest'analisi è stato il pre-processing, necessario per ottenere immagini con la stessa risoluzione, dimensione e intensità simile. Le immagini sono state ricampionate a una risoluzione comune, è stata valutata una finestra del range dinamico ed è stata eseguita una normalizzazione basata sul min-max scaling. Segue una fase di bilanciamento dei dataset di training in modo da garantire un equilibrio tra le immagini contenenti le lesioni e quelle senza lesione. La rete neurale utilizzata è di tipo U-Net che prende in input immagini 2D e restituisce in output la maschera automatica delle lesioni presenti. L'affidabilità

del modello è stata valutata utilizzando un test set composto dal 10% dei pazienti. Le predizioni della rete sono state sottoposte a un processo di post-elaborazione per ridurre le misclassificazioni, che spesso derivano dalla presenza di strutture interne al polmone erroneamente classificate. Si basa su una strategia di analisi della connettività dei volumi identificati, attraverso la quale vengono isolate e valutate le regioni contigue. Questo approccio consente di preservare solo i volumi significativi ottenendo un Dice Similarity Coefficient medio di 0,68, una Recall media di 0,91 e una Precision media di 0,58. Questi risultati suggeriscono una buona capacità di individuare le lesioni, mettono però in evidenza difficoltà nell'accuratezza e nella precisione delle predizioni, che potrebbero essere dovute al numero limitato di dati a disposizione. Sono tuttavia necessari ulteriori miglioramenti per adattare l'algoritmo alla variabilità clinica e renderlo utilizzabile nella pratica medica.

Indice

Elenco delle figure	VI
Acronimi	IX
1 INTRODUZIONE	1
1.1 I polmoni	1
1.2 Cancro al polmone non a piccole cellule	2
1.3 Tomografia computerizzata	5
1.4 Segmentazione di immagini mediche	7
1.5 Deep Learning nella segmentazione delle immagini mediche	10
1.6 Radiomica	14
1.7 Obiettivo dello studio	15
2 MATERIALI E METODI	16
2.1 Descrizione del data-set	16
2.2 Associazione serie e RTSTRUCT	20
2.3 Pre-processing	23
2.3.1 Metodo 1 di pre-processing: ROI attorno al polmone	25
2.3.2 Metodo 2 di pre-processing: ROI attorno al centro di massa	40
2.3.3 Metodo 3 di pre-processing: estrazione del parenchima polmonare	41
2.4 Suddivisione e bilanciamento del data set	42
2.5 Data Augumentation	45
2.6 Rete Neurale: U-Net	45
2.7 Post-processing	48
2.8 Metriche di valutazione	50
3 RISULTATI	52
3.1 Pre-processing	52
3.1.1 Risultati del metodo 1	52
3.1.2 Risultati del metodo 2	57

3.1.3	Risulati del metodo 3	58
3.2	Rete Neurale	59
3.2.1	Risulati del metodo 1	59
3.2.2	Risulati del metodo 2	63
3.2.3	Risulati del metodo 3	63
4	DISCUSSIONE	71
5	SVILUPPI FUTURI	78
	Bibliografia	80

Elenco delle figure

1.1	Organizzazione dell'apparato respiratorio.	1
1.2	Nuovi casi di diagnosi di tumore al polmone.	2
1.3	Pie Chart delle percentuali di tutti i tipi di tumore al polmone [5].	3
1.4	Stadiazione del carcinoma polmonare[7].	4
1.5	Macchinario della Tomografia Computerizzata.	5
1.6	Funzionamento della Tomografia Computerizzata.	6
1.7	Volumi considerati in RT[14].	7
1.8	Esempio di OAR e GTV.	8
1.9	Correlazione tra AI, ML e DL.	10
1.10	Neurone biologico e neurone artificiale[19].	11
1.11	Funzionamento dei neuroni negli strati nascosti.	12
1.12	Architettura di una basica CNN[21].	13
1.13	Workflow Radiomico [23].	14
2.1	Pie Chart rappresentativo della posizione del tumore.	17
2.2	Grafico a dispersione rappresentativo del volume della lesione.	18
2.3	Grafico rappresentativo della frequenza dei volumi in intervalli differenti.	18
2.4	Grafico rappresentativo delle posizioni delle lesioni.	19
2.5	Flowchart relativo al complessivo flusso di lavoro.	20
2.6	Flowchart del pre-processing comune a tutti i metodi successivamente sviluppati.	23
2.7	Flowchart del pre-processing delle CT del primo approccio.	25
2.8	Slice d'esempio di una scansione CT toracica.	26
2.9	Volume intero della scansione CT.	27
2.10	Esempio di slice di output dell'operazione di sogliatura.	28
2.11	Scala Hounsfield.	28
2.12	Istogramma rappresentativo di un volume di CT in scala logaritmica.	29
2.13	Esempi di slices di output dell'operazione di labelling.	30
2.14	Esempi di slices di output dell'operazione di estrazione della maschera polmonare.	31

2.15	Rappresentazione dei contorni della maschera polmonare su diverse slices.	34
2.16	Istogramma rappresentativo dell'andamento delle risoluzioni spaziali delle CT.	35
2.17	Window Level e Window Width.	37
2.18	Confronto tra istogrammi delle lesioni di diversi pazienti.	38
2.19	Flowchart del pre-processing delle CT del secondo approccio.	40
2.20	Flowchart del pre-processing delle CT del terzo approccio.	41
2.21	Architettura della Rete U-Net.	46
2.22	Architettura della Rete.	46
2.23	Esempio in un caso bidimensionale del parametro di connectivity.	49
3.1	Risultati dell' algoritmo automatico di segmentazione polmonare.	53
3.2	Rappresentazione del crop effettuato su diverse slices dello stesso paziente.	55
3.3	Rappresentazione del windowing con conseguente padding.	56
3.4	Casi rappresentativi di diversi pazienti con taglio attorno al centro di massa della lesione.	57
3.5	Casi rappresentativi dell'estrazione del parenchima polmonare.	58
3.6	Validation-Snapshot della prova 1.	61
3.7	Validation-Snapshot della prova 2.	61
3.8	Validation-Snapshot della prova 3.	62
3.9	Validation-Snapshot della prova 4.	63
3.10	Validation-Snapshot della prova 5.	64
3.11	Validation-Snapshot della prova 6.	64
3.12	Box Plot di DSC, RECALL e PRECISION dei dati di testing.	66
3.13	Esempi di pazienti del dataset di testing raffigurante l'area.	67
3.14	Esempi di pazienti del dataset di testing raffigurante solo i contorni.	68
3.15	Esempi di pazienti del dataset di testing: ROI in rosso rappresenta la maschera automatica.	69
3.16	Grafico a barre di confronto tra le metriche calcolate con l'output della rete e quelle calcolate in output dal post-processing.	70

Acronimi

AI

Artificial Intelligence

CAD

Computer Aided Diagnosis and Detection

CNN

Convolutional Neural Network

CT

Tomografia Computerizzata

CTV

Clinical Tumor Volum

DICOM

Digital Imaging COmmunications in Medicine

DL

Deep Learning

DSC

Dice Similarity Coefficient

FN

False Negative

FP

False Positive

GTV

Gross Tumor Volume

HU

Unità Hounsfield

LDCT

Low Dose CT

ML

Machine Learning

MONAI

Medical Open Network for AI

MRI

Magnetic Resonance Imaging

NN

Neural Network

NSCLC

Non Small Cell Lung Cancer

OAR

Organi A Rischio

PET

Tomografia a Emissione di Positroni

PTV

Planning Tumor Volume

ReLu

Rectifier Linear Unit

ROI

Region Of Interest

RT

RadioTerapia

SCLC

Small Cell Lung Cancer

STP

Temperatura e Pressione in condizioni Standard

TP

True Positive

TPS

Treatment Planning System

VOI

Volume of Interest

WL

Window Level

WW

Window Width

Capitolo 1

INTRODUZIONE

1.1 I polmoni

L'apparato respiratorio, visualizzato in Figura 1.1, si divide in due vie principali:

- le vie aeree superiori, ovvero gli organi collocati esternamente al torace, quali naso, faringe e laringe;
- le vie aeree inferiori, che comprendono gli organi situati all'interno del torace, quali trachea, bronchi e polmoni.

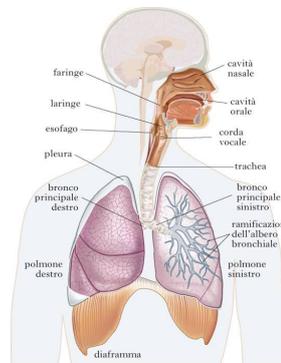


Figura 1.1: Organizzazione dell'apparato respiratorio.

In particolare, i polmoni sono due organi che hanno la funzione di fornire O_2 all'organismo ed eliminare CO_2 dal sangue. Essi hanno una forma a cono e sono costituiti da tessuto spugnoso. Al loro interno sono presenti circa 2400 km di vie aeree e sono separati tra loro dal cuore e dal mediastino. Ognuno dei due polmoni ha, all'estremità superiore, un apice che si estende verso l'alto fino alla trachea e, l'estremità inferiore che poggia sul muscolo diaframmatico. La superficie interna

della parete toracica e quella esterna dei polmoni sono rivestite rispettivamente da due pleure polmonari: parietale (più robusta), e viscerale (più sottile). Le pleure sono costituite da uno strato di cellule epiteliali e da tessuto connettivo elastico. Tra le due è presente la cavità pleurica in cui risiede un fluido pleurico lubrificante che permette agli strati di scivolare l'uno sull'altro nel corso della respirazione [1].

1.2 Cancro al polmone non a piccole cellule

Fino ai primi anni del '900 il cancro al polmone era una malattia poco comune, ma oggi a causa di più fattori di rischio rappresenta la principale causa di morte in Europa per neoplasie. Nel mondo, nel 2012, si sono registrati 1.8 milioni di casi di cui 1.6 milioni di morti. In Europa la sopravvivenza media a 5 anni è del 11.2% negli uomini e 13,9% nelle donne [2]. Il carcinoma polmonare, seconda neoplasia più frequente dopo il cancro alla prostata, tende a manifestare i suoi sintomi dopo un lungo periodo di tempo dall'insorgenza della malattia. La causa maggiore di questo cancro, responsabile del 85% dei casi, è il fumo di sigaretta. Tuttavia, il 10%-15% dei casi coinvolge persone che non hanno mai fumato [3]. In questi casi i fattori di rischio possono essere principalmente genetici, di esposizione a sostanze cancerogene (come amianto, radon, arsenico ecc.), fumo passivo o agenti atmosferici inquinanti [4].

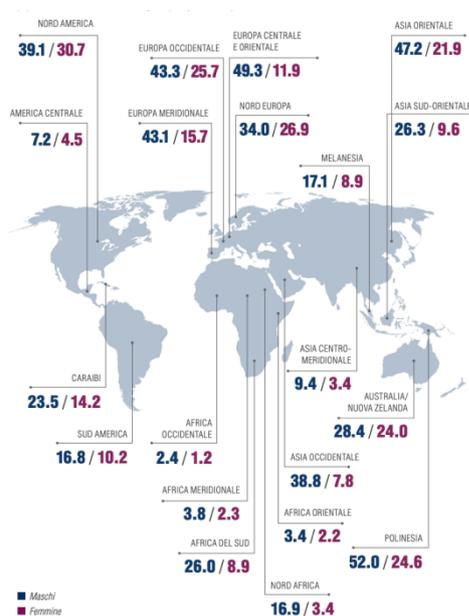


Figura 1.2: Nuovi casi di diagnosi di tumore al polmone effettuati nel 2018 ogni 100.000 persone sulla base delle statistiche effettuate da Ferlay et al.[3]

Il cancro al polmone origina tipicamente nelle cellule epiteliali che rivestono i bronchi, i bronchioli o gli alveoli. La crescita incontrollata di queste cellule può condurre alla formazione di metastasi nei tessuti circostanti. Viene comunemente distinto in due macro-categorie:

- Cancro al polmone a piccole cellule (SCLC);
- Cancro al polmone non a piccole cellule (NSCLC).

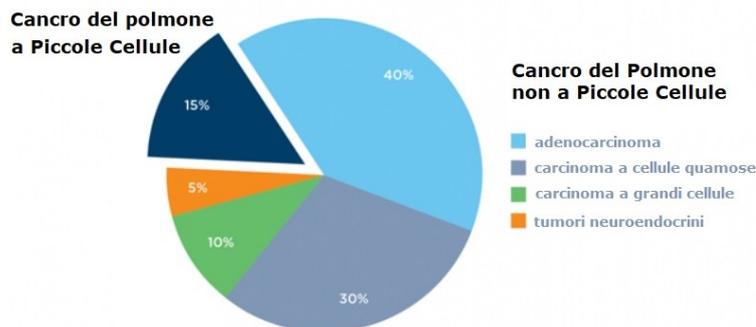


Figura 1.3: Pie Chart delle percentuali di tutti i tipi di tumore al polmone [5].

Il primo rappresenta circa il 15% di tutte le tipologie tumorali. Si distingue per la minor dimensione delle cellule di cui è formato e per il luogo di origine. Lo sviluppo avviene principalmente nella sottomucosa in sede ilare o centrale con coinvolgimento dei bronchi e con comparsa di piccole cellule rotonde o affusolate con margini ben delineati [5]. SCLC è un tumore molto aggressivo, ha una crescita rapida e si manifesta spesso nei fumatori.

Il secondo rappresenta l'80-90% di tutte le neoplasie polmonari maligne ed è più variabile in base al tipo istologico. Il 40% dei pazienti ha sviluppato metastasi al momento della diagnosi. I tre principali sottotipi istologici di NSCLC sono [6]:

- Adenocarcinoma polmonare: rappresenta il 40% di tutti i tumori polmonari, origina dalle cellule secernenti muco delle vie aeree. Nella maggior parte dei casi progredisce nella porzione periferica del polmone, a causa di infezioni croniche o in presenza di aree necrotiche;
- Il carcinoma a cellule squamose: rappresenta il 25%-30% di tutti i tumori, prolifera dalle cellule epiteliali danneggiate dal fumo, solitamente nella porzione centrale del polmone. Esso è classificato come tale in base all'aspetto a squama di pesce delle cellule al microscopio;
- Il carcinoma a grandi cellule (indifferenziato): rappresenta il 10%-15% di tutti i tumori e solitamente compare nella zona periferica del polmone con

coinvolgimento subpleurico. La sua diagnosi avviene per esclusione essendo scarsamente differenziato.

Il sistema di stadiazione per i tumori polmonari non a piccole cellule è il TNM [6] dove:

- T: indica le dimensioni e il grado di estensione del tumore principale (da T1 a T4);
- N: rappresenta il grado di coinvolgimento dei linfonodi vicini (da N0 a N3);
- M: indica la presenza di metastasi (M0 a M1b).

Per la valutazione della terapia si considerano diversi stadi progressivi:

- STADIO 0: viene anche chiamato malattia in situ poichè il cancro non ha ancora invaso i tessuti circostanti;
- STADIO 1: le cellule tumorali non hanno coinvolto i linfonodi vicini e sono limitate ad un solo polmone.
- STADIO 2: le cellule tumorali sono ancora presenti in unico polmone, ma sono di dimensioni maggiori e potrebbero aver coinvolto altre aree circostanti o i linfonodi vicini. Non sono presenti metastasi.
- STADIO 3: si tratta di una malattia localmente avanzata, in cui il cancro è ancora presente in un solo polmone, ma ha dimensioni superiori. Potrebbe invadere organi circostanti o coinvolgere i linfonodi vicini, senza metastasi.
- STADIO 4: malattia avanzata o metastatica in quanto i tumori sono di qualsiasi dimensione e hanno coinvolto l'altro polmone con metastasi.

	T	N	M
Carcinoma occulto	TX	N0	M0
Stadio 0	Tis (carcinoma in situ)	N0	M0
Stadio IA	T1a,b	N0	M0
Stadio IB	T2a	N0	M0
Stadio IIA	T2b	N0	M0
	T1a,b	N1	M0
Stadio IIB	T2a	N1	M0
	T2b	N1	M0
Stadio IIIA	T3	N0	M0
	T1a,b, T2a,b	N2	M0
	T3	N1, N2	M0
Stadio IIIB	T4	N0, N1	M0
	T4	N2	M0
Stadio IIIB	Qualunque T	N3	M0
Stadio IV	Qualunque T	Qualunque N	M1

■ Malattia iniziale
 ■ Malattia localmente avanzata*
 ■ Malattia metastatica

* Alcune classificazioni inquadrano lo stadio II come malattia localmente avanzata

Figura 1.4: Stadiazione del carcinoma polmonare[7].

La diagnosi del carcinoma polmonare può essere effettuata attraverso diverse modalità per pianificare in modo appropriato il trattamento. I passaggi da seguire per arrivare alla diagnosi comprendono: raccolta della storia clinica, esami di imaging, analisi di laboratorio, biopsia e test per identificare la presenza di biomarcatori [8]. Gli esami di imaging sono fondamentali per determinare la posizione e l'estensione della massa tumorale.

1.3 Tomografia computerizzata

Gli esami di imaging vengono effettuati per individuare eventuali aree sospette, diffusione del cancro, per monitorare il trattamento e il post trattamento[8]. Le principali tecniche di imaging includono la radiografia, la tomografia computerizzata (TC), la risonanza magnetica (MRI), la tomografia ad emissioni di positroni (PET) e la scansione ossea.

A differenza di una radiografia tradizionale, che scatta una singola foto, uno scanner CT cattura più immagini. La tomografia computerizzata è un esame radiologico che prevede la raccolta di dati relativi al passaggio di vari fasci di raggi X nell'area interessata e la loro rielaborazione da parte di un computer, in modo da ricostruire un'immagine tridimensionale dei diversi tipi di tessuto analizzati [9]. Oggi esistono macchinari multistrato più moderni e la CT non è più solo assiale, ma le immagini sono acquisite con una tecnica spirale. Il tubo che emette i raggi X ruota attorno al paziente durante l'esame, inviando i fasci di raggi attraverso il corpo da diverse angolazioni. I raggi X sono una forma di radiazione ionizzante che attraversa i tessuti del corpo, la cui quantità assorbita varia a seconda della densità dei tessuti: ossa, muscoli, grasso e liquidi assorbono i raggi X in maniera differente. I tessuti con densità più alta, come le ossa, assorbono più raggi X rispetto ai tessuti molli, come i muscoli o i grassi, che lasciano passare più radiazioni.



Figura 1.5: Macchinario della Tomografia Computerizzata.

I raggi X, che passano attraverso il corpo, vengono raccolti da una serie di rilevatori posti sull'anello (detto gantry) opposto al tubo. Questi misurano l'intensità dei raggi che li raggiungono, convertendo i dati in segnali elettronici. Il paziente viene disteso su un lettino che si muove progressivamente all'interno del gantry, permettendo alla macchina di eseguire diverse scansioni attraverso il corpo [10].

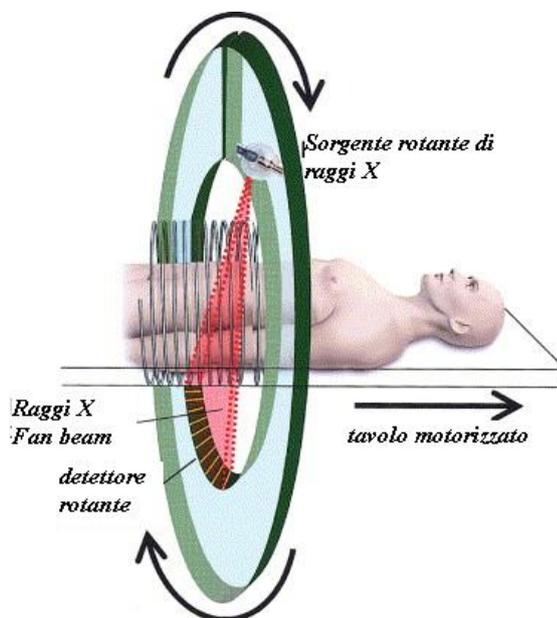


Figura 1.6: Funzionamento della Tomografia Computerizzata.

Una CT consente di ricostruire la dimensione, la forma e la posizione di strutture anatomiche, come organi interni e tessuti. Inoltre visualizza anche eventuali masse o noduli tumorali presenti in diverse parti del corpo. Perciò, è in grado di mostrare anche le lesioni polmonari, fornendo informazioni sulla loro localizzazione e grandezza. Oggi viene spesso utilizzata una LDCT low dose-CT scan (dose pari ad un quinto dello scanner normale) per lo screening polmonare e i follow up per monitorare l'efficacia del trattamento [11]. Al paziente può essere somministrato un mezzo di contrasto, solitamente a base di iodio, per migliorare la visibilità di alcuni tessuti. Questo materiale ha la capacità di assorbire i raggi X in misura maggiore rispetto ai tessuti circostanti, grazie alla sua elevata densità. Questo maggiore assorbimento provoca una riduzione della quantità di radiazioni che raggiunge i rilevatori, creando un'immagine più chiara e contrastata nella zona in cui si trova il mezzo di contrasto. Si rendono così più evidenti dettagli che altrimenti potrebbero risultare difficili da distinguere, come tumori o malformazioni vascolari.

La Scala Hounsfield viene utilizzata per rappresentare attraverso l'unità Hounsfield (HU) il coefficiente di attenuazione del fascio di radiazioni nell'esame di CT. Questa

è una trasformazione lineare della misurazione del coefficiente di attenuazione lineare del materiale rispetto alla radiodensità dell'acqua distillata a pressione e temperatura in condizioni standard (STP) che è definita come 0 HU, mentre la radiodensità dell'aria in condizioni standard è definita come -1000 HU. In un voxel, con coefficiente di assorbimento μ , il corrispondente valore HU è quindi dato da:

$$HU = 1000 \times \frac{\mu - \mu_{\text{acqua}}}{\mu_{\text{acqua}} - \mu_{\text{aria}}}$$

Quindi, una variazione di una unità Hounsfield rappresenta una variazione dello 0,1% del coefficiente di attenuazione dell'acqua poiché il coefficiente di attenuazione dell'aria è quasi zero. Per definizione, la calibrazione degli scanner per tomografia computerizzata viene fatta con riferimento all'acqua.

1.4 Segmentazione di immagini mediche

Nel campo della radioterapia (RT), per garantire un trattamento efficace è fondamentale un targeting preciso del tumore, preservando al contempo gli organi a rischio a causa delle complicanze associate alle radiazioni. I tre volumi principali da considerare nella RT, come mostrato nella Figura 1.7: il Gross Tumor Volume (GTV), che indica la posizione e l'estensione del tumore primario; il Clinical Tumor Volume (CTV), che comprende la posizione e l'estensione del tumore primario, includendo anche l'estensione della diffusione microscopica del tumore non visibile tramite diagnostica per immagini; il Planning Tumor Volume (PTV) viene aggiunto ai due precedenti per tener conto delle incertezze nella pianificazione e nella somministrazione del trattamento[12] [13].

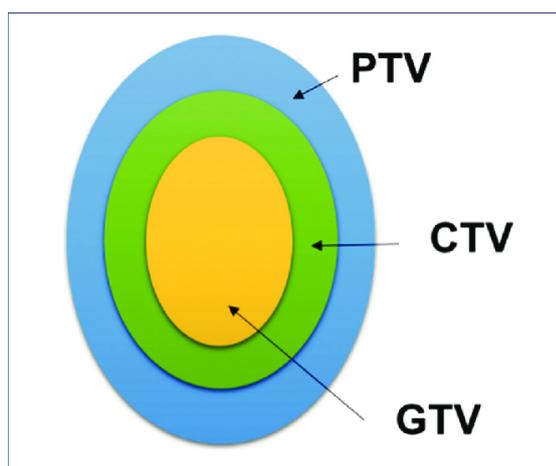


Figura 1.7: Volumi considerati in RT[14].

Nella Figura 1.8 è illustrato un esempio di segmentazione di organi a rischio (OAR) e GTV.

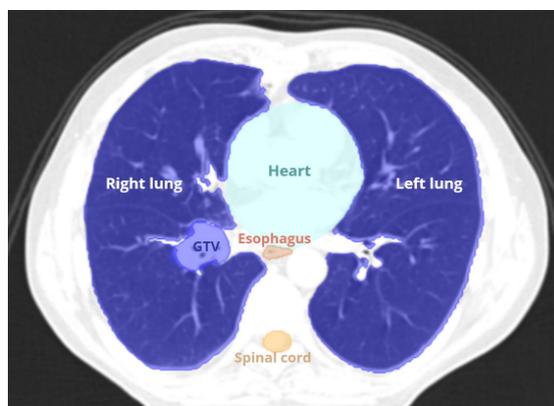


Figura 1.8: Esempio di OAR e GTV.

Tuttavia, la segmentazione manuale del GTV e dell'OAR è un compito noioso e richiede tempo tedioso per il clinico. Questo può causare ritardi nel trattamento e influire negativamente sui tassi di sopravvivenza. Inoltre, la qualità della segmentazione manuale è soggetta all'esperienza dell'operatore[15].

La fase di segmentazione delle immagini è una fase cruciale nell'elaborazione delle immagini mediche che prevede la classificazione di un'immagine digitale in più classi, ciascuna delle quali comprende insiemi di pixel. Permette di identificare la regione di interesse (Region Of Interest, ROI) per immagini in 2D oppure il volume di interesse (Volume Of Interest, VOI) per immagini in 3D, identificando la parte che verrà sottoposta successivamente al processo di estrazione delle feature. Esistono tre tipologie di segmentazione delle immagini:

- segmentazione manuale, che viene effettuata da un clinico che annota manualmente i pixel di interesse;
- segmentazione semi-automatica, che prevede l'uso di algoritmi che aiutano nel processo di segmentazione definendone il contorno;
- segmentazione automatica, che non richiede l'input dell'utente e può essere classificata in metodi basati sull'apprendimento e non[16].

Tuttavia, il clinico che effettua manualmente la segmentazione è anche il responsabile dell'accettazione della delineazione del volume del tumore, anche se è stata utilizzata la segmentazione automatica. Ciò è dovuto al fatto che non sono ancora stati sviluppati modelli affidabili di autosegmentazione del GTV.

I metodi di segmentazione automatica convenzionali si basano sulle caratteristiche di superficie dell'immagine, tra cui la scala di grigi, la texture e il gradiente, per segmentare l'oggetto desiderato. Il metodo della soglia, il metodo dell'atlante e il metodo del regione growing sono alcuni degli approcci più comuni utilizzati nelle tecniche di segmentazione automatica tradizionali[15]. Il metodo della soglia seleziona una soglia appropriata in base al target e allo sfondo, in modo che tutti i pixel dell'immagine, in base al loro valore, vengono classificati nella classe del target o in quella dello sfondo. Il metodo dell'atlante, invece, allinea la nuova immagine di input con un'immagine di riferimento nota come modello di atlante. Le etichette dell'atlante vengono quindi applicate alla nuova immagine di input per realizzare il compito di segmentazione. Il metodo Region Growing prevede la delineazione manuale di sottoregioni, seguita dall'unione di pixel vicini con attributi simili nella regione predeterminata e dalla segmentazione dell'area target dallo sfondo.

Con il progredire della metodologia del Deep Learning (DL), i modelli basati su DL hanno dimostrato un notevole potenziale nella segmentazione automatica delle immagini mediche. I modelli DL apprendono in modo indipendente la rappresentazione delle caratteristiche e utilizzano queste features per segmentare senza bisogno di interazione manuale[15]. L'uso del DL nella pratica clinica ha il potenziale di ridurre i tempi inutili e di alleggerire il carico di lavoro del personale addetto.

1.5 Deep Learning nella segmentazione delle immagini mediche

La relazione gerarchica Intelligenza Artificiale (AI), Machine Learning (ML) e DL è riassunta nel diagramma illustrato nella Figura 1.9[17].

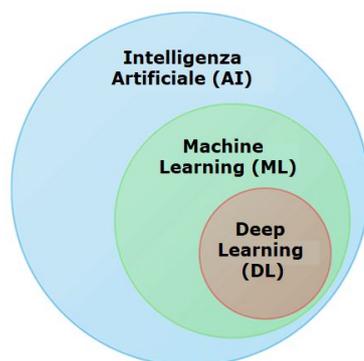


Figura 1.9: Correlazione tra AI, ML e DL.

L'intelligenza artificiale (IA) si riferisce alla capacità dei dispositivi di simulare comportamenti umani e di svolgere compiti complessi con un intervento umano ridotto o minimo. Il machine learning (ML) è un ambito specifico che mira a migliorare l'abilità di un software di eseguire determinate operazioni. Questo avviene attraverso algoritmi che apprendono da un insieme di dati di addestramento pertinenti al problema, consentendo ai computer di riconoscere schemi e tendenze nascoste senza necessità di programmazione diretta. Il Deep learning (DL), invece, è una branca del ML che utilizza reti neurali artificiali per elaborare le informazioni. L'architettura del deep learning (DL) si rifà ai principi di elaborazione delle informazioni tipici delle reti neurali biologiche presenti nel cervello umano. Il neurone, che rappresenta l'unità funzionale del sistema nervoso, è costituito da tre elementi principali, come illustrato nella Figura 1.10:

- il soma, ovvero il corpo della cellula, che contiene il nucleo e altre strutture di supporto cellulare;
- i dendriti, che hanno il compito di ricevere i segnali elettrochimici dai neuroni circostanti;
- l'assone, che trasmette i segnali verso il neurone successivo[18].

I dendriti di neuroni contigui sono connessi tramite una sinapsi. Quando il segnale ricevuto è sufficientemente intenso da superare una certa soglia, il neurone si attiva e il segnale viene trasferito al neurone seguente.

Un neurone artificiale ha l'obiettivo di riprodurre la struttura e il funzionamento di quelli biologici. Infatti, le reti neurali si avvicinano al funzionamento del cervello per due aspetti principali:

- Acquisiscono la conoscenza attraverso un processo di apprendimento;
- La conoscenza viene memorizzata utilizzando i pesi delle sinapsi.

Un algoritmo di apprendimento modifica i pesi di connessione tra i neuroni per raggiungere un obiettivo specifico. Esistono due tipi di apprendimento:

- Apprendimento supervisionato: i pesi sono regolati per ridurre al minimo l'errore tra l'uscita e l'obiettivo fornito;
- Apprendimento non supervisionato: l'obiettivo è raggruppare gli elementi del dataset in gruppi omogenei.

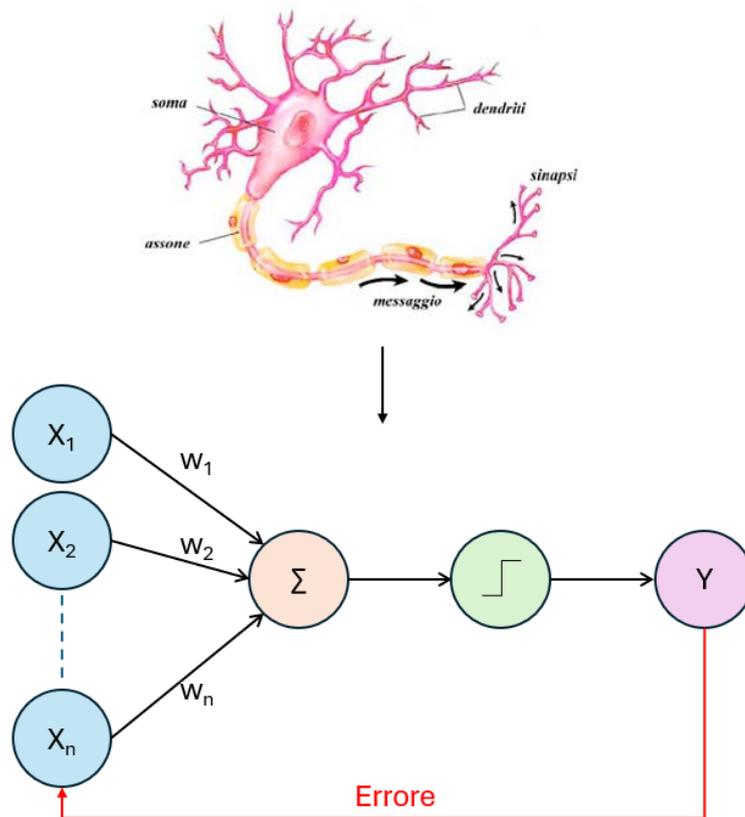


Figura 1.10: Neurone biologico e neurone artificiale[19].

La rete neurale artificiale più semplice è il perceptron mostrato nella Figura 1.10. Questo classifica gli input in una delle due classi utilizzando una funzione di attivazione che riceve la somma ponderata degli input. Tuttavia, poiché il perceptron è una rete neurale (NN) a singolo strato, non può eseguire classificazioni non lineari, il che può essere superato aggiungendo strati nascosti per creare una rete neurale profonda. L'architettura di una rete neurale profonda è strutturata in modo altamente stratificato, poiché sono richiesti più livelli per processare i dati e produrre l'output finale. Ogni livello della rete è formato da un certo numero di neuroni e segue l'organizzazione mostrata nella 1.11:

- Strato di input: contiene neuroni responsabili di ricevere i dati in ingresso dall'esterno;
- Strato di output: formato dai neuroni che generano il risultato finale della rete;
- Strati intermedi (o nascosti): composti da neuroni che non interagiscono direttamente con l'ambiente esterno e che si collegano, in modo completo o parziale, con gli strati adiacenti.

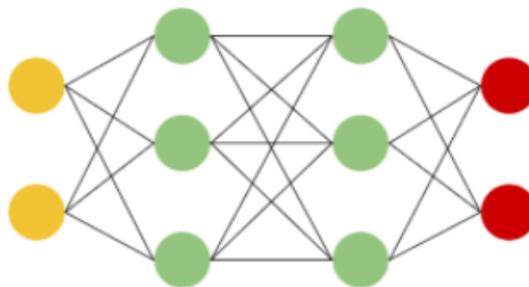


Figura 1.11: Funzionamento dei neuroni negli strati nascosti (in verde) che elaborano tra loro gli input ricevuti dagli strati di input (in giallo) prima di trasmetterli allo strato di output (in rosso).

L'apprendimento in una NN può essere suddiviso in due fasi:

- I modelli di ingresso vengono presentati allo strato di ingresso e poi propagati da uno strato all'altro fino a generare un modello di uscita;
- Se il modello di uscita generato differisce dall'uscita target, viene calcolato un errore e propagato a ritroso dal livello di uscita al livello di ingresso, modificando di conseguenza i pesi di input.

L'utilizzo delle tradizionali reti neurali artificiali completamente connesse è difficile a causa dell'elevato numero di parametri di rete richiesti, con conseguente complessità computazionale e perdita di informazioni. Per affrontare queste sfide, vengono utilizzate le reti neurali convoluzionali (CNN). Queste utilizzano l'operatore di convoluzione, che sostituisce la moltiplicazione delle matrici con un insieme di kernel di convoluzione. Questo passaggio riduce la complessità computazionale[20].

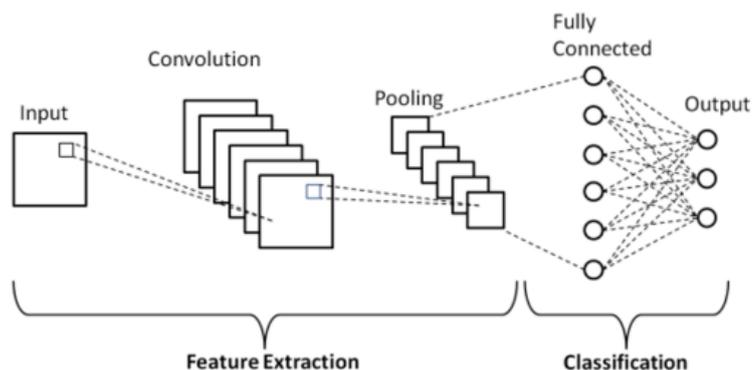


Figura 1.12: Architettura di una base CNN[21].

I principali elementi costitutivi di una CNN sono gli strati convoluzionali, gli strati di pooling e gli strati completamente connessi [20], come illustrato nella Figura 1.12. Gli strati convoluzionali applicano una serie di filtri a un'immagine di ingresso, i quali estraggono varie caratteristiche dall'immagine. Lo stride e il padding sono due parametri della convoluzione, che determinano, rispettivamente, il modo in cui il filtro viene spostato lungo la matrice di ingresso, e l'aggiunta di uno zeropadding alla mappa delle caratteristiche di input per assicurare che la finestra possa essere sempre centrata sul valore della matrice di ingresso. Le mappe di caratteristiche risultanti vengono poi inserite nei layer di pooling, che le ricampionano per ridurre le dimensioni e rendere la rete più efficiente. Dopo diversi strati convoluzionali e di pooling l'output viene appiattito e alimentato in strati completamente connessi, che eseguono la classificazione finale dell'input. Questi strati completamente connessi imparano a combinare le caratteristiche estratte dagli strati precedenti per fare una previsione finale.

La convoluzione è un'operazione lineare, quindi l'applicazione di una non linearità è necessaria per aumentare la potenza espressiva del modello. Pertanto, gli strati convoluzionali incorporano anche funzioni di attivazione non lineari, come ad esempio Rectifier Linear Unit (ReLU). Nell'architettura di una CNN si utilizzano anche strati di normalizzazione per normalizzare l'output degli strati precedenti sottraendo la media e dividendo per la deviazione standard delle attivazioni. Questo migliora la stabilità e la velocità dell'addestramento[22].

1.6 Radiomica

La radiomica è un campo di ricerca nell'ambito delle analisi di immagine mediche che permette l'estrazione ad alto rendimento di feature quantitative che convertono le immagini in dati estraibili al fine di analizzarli per migliorare diagnosi, prognosi e l'accuratezza predittiva. Tale processo può giocare un ruolo centrale nel lavoro clinico giornaliero consentendo la costruzione di database e di report contenenti dati e analisi utili per il supporto di decisioni future. I sistemi CAD (Computer-Aided Diagnosis and Detection) vengono utilizzati per la rilevazione o la diagnosi di malattie, ma risultano essere sistemi standalone, in quanto hanno come risultato un'unica risposta, ad esempio la presenza o meno di una lesione. La radiomica al contrario è un processo che comporta l'estrazione di una grande quantità di feature da immagini digitali. Essa nasce al fine di generare biomarcatori di imaging attraverso l'elaborazione e l'analisi delle immagini mediche, utilizzabili come strumenti a supporto delle decisioni riuscendo a combinare fra loro sia dati estratti dalle immagini sia altre caratteristiche dei pazienti.

In oncologia, la radiomica permette l'estrazione e l'analisi di feature quantitative legate alla biologia del tumore, come ad esempio intensità, la dimensione, il volume, la forma e la texture, aiutando nel rilevamento della malattia, nella diagnosi, nella valutazione della prognosi, nella previsione della risposta ai trattamenti nonché al monitoraggio dello stato della malattia. La radiomica è inoltre uno strumento fondamentale per la nuova frontiera della medicina di precisione, dove con medicina di precisione si intende il tentativo di personalizzare il più possibile prevenzione, diagnosi e cure in base al singolo paziente, abbandonando concetti vecchi basati su protocolli standardizzati utilizzati in passato.

Il processo di estrazione ed analisi di feature da immagini mediche segue diversi passaggi rappresentati in Figura 1.13

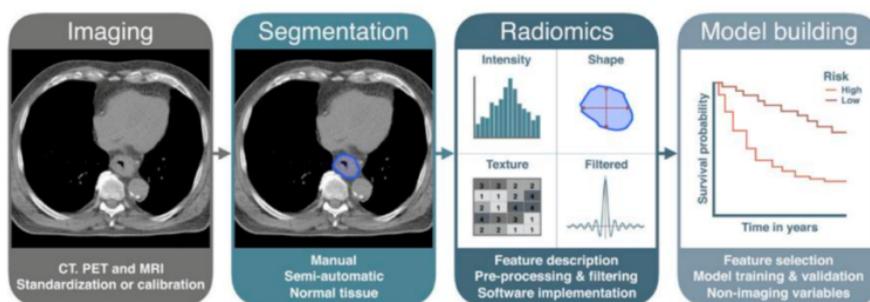


Figura 1.13: Workflow Radiomico [23].

1.7 Obiettivo dello studio

Nel ramo dell'imaging medicale, come è stato visto, sta rapidamente emergendo la Radiomica che permette di sviluppare strumenti a supporto delle decisioni degli esperti. Per manipolare e analizzare un tale numero di dati si ricorre spesso a tecniche avanzate, con lo scopo finale di estrarre direttamente dalle immagini indicazioni sull'aggressività della malattia, sulle terapie più indicate e sulla risposta alle cure. Pertanto, l'obiettivo di questo studio è quello di trovare un algoritmo automatico che identifichi le lesioni tumorali presenti nei polmoni da immagini CT, affinché sia uno strumento utile per gli studi di Radiomica. Questo aiuterebbe a ridurre significativamente i tempi di segmentazione manuale, un processo tradizionalmente lungo e soggetto a variabilità tra operatori. L'idea non è solo quella di rendere il processo più efficiente, ma anche di garantire maggiore coerenza e riproducibilità nella segmentazione. Un modello efficace deve ridurre al minimo i falsi negativi, assicurando che tutte le aree tumorali vengano identificate correttamente (alta sensibilità), e i falsi positivi, evitando di includere tessuti sani all'interno della segmentazione (alta specificità). L'obiettivo finale è ottenere una segmentazione automatica che velocizzi l'intero processo clinico, mantenendo una qualità elevata, così che i medici possano concentrarsi sulla valutazione dei risultati piuttosto che sulla segmentazione manuale delle immagini. Questo approccio non è finalizzato alla diagnosi, ma a ottimizzare il flusso di lavoro e migliorare l'efficienza delle operazioni cliniche. L'automatizzazione dell'identificazione e della segmentazione del tumore non sostituirà completamente l'intervento del medico, poiché il clinico avrà comunque il compito di verificare e, se necessario, rifinire la segmentazione automatica. Tuttavia, il tempo richiesto per questa revisione sarà significativamente inferiore rispetto alla segmentazione manuale completa. L'obiettivo principale è fornire una base solida e affidabile che il clinico possa rapidamente confermare, riducendo il carico di lavoro e velocizzando l'intero processo.

Capitolo 2

MATERIALI E METODI

2.1 Descrizione del data-set

Per questo studio sono stati forniti, dall'azienda Tecnologie Avanzate, 88 pazienti con diagnosi di NSCLC al terzo stadio, con dati relativi al periodo compreso tra il 2012 e il 2022. I pazienti inclusi nello studio hanno un'età compresa tra i 45 e gli 85 anni e includono soggetti di entrambi i sessi. Le scansioni CT toraciche sono state eseguite con l'utilizzo di mezzo di contrasto. Il protocollo di acquisizione delle scansioni CT varia sia in base al tipo di macchinario utilizzato, realizzato da differenti produttori, sia per le specifiche tecniche del protocollo stesso, che possono differire in termini di energia irradiata e altre impostazioni operative.

Le caratteristiche cliniche dei pazienti variano in modo significativo per dimensione e posizione delle lesioni, nonché per il tipo istologico del tumore: 65 soggetti presentano adenocarcinoma, 19 soggetti presentano carcinoma squamoso, 3 soggetti con carcinoma indifferenziato e 1 soggetto con carcinoma neuroendocrino a grandi cellule.

La dimensione e la posizione del tumore hanno avuto un impatto rilevante sull'andamento dello studio, influenzando in modo sostanziale le conclusioni derivanti dall'analisi dei dati. Infatti, come illustrato in Figura 2.4 e nel Grafico 2.1, il data set è composto da tumori localizzati per il 34% internamente al polmone, 31% a contatto con il mediastino e 35% aderenti alla pleura.

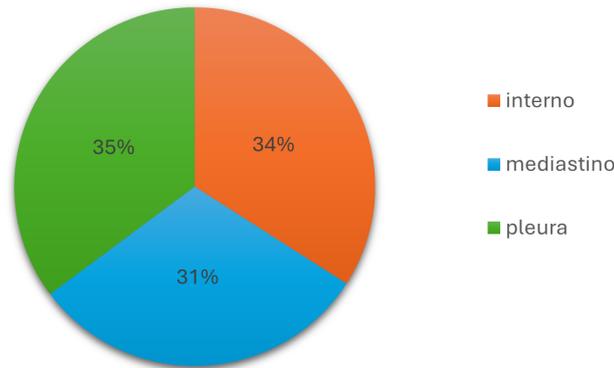


Figura 2.1: Pie Chart rappresentativo della posizione del tumore.

Anche le dimensioni della lesione differiscono in modo notevole tra i pazienti. Come si può notare dal grafico in Figura 2.2 esse variano da casi estremamente piccoli (498 mm^3) a casi molto più considerevoli (361.132 mm^3). Questa ampia gamma di valori evidenzia che le dimensioni non si concentrano all'interno di un singolo intervallo ben definito, ma piuttosto si distribuiscono su una scala molto ampia, quindi si rileva una marcata eterogeneità tra i pazienti. Tale variabilità suggerisce che la distribuzione delle dimensioni è caratterizzata da un'ampia dispersione, senza un'unica fascia predominante.

In relazione a quanto detto, si nota dal grafico in Figura 2.3 che le dimensioni delle lesioni possono essere distinte in intervalli diversi:

- $0 - 4000 \text{ mm}^3$: 15 pazienti rientrano in questo intervallo, che indica lesioni di volume molto piccolo;
- $4000 \text{ mm}^3 - 11000 \text{ mm}^3$: questo è l'intervallo più frequente, con 25 pazienti;
- $11000 \text{ mm}^3 - 30000 \text{ mm}^3$: si osserva una frequenza di 21 pazienti, indicando lesioni di dimensioni medio-grandi;
- $30000 \text{ mm}^3 - 80000 \text{ mm}^3$: altri 16 pazienti rientrano in questo intervallo, suggerendo un ulteriore gruppo di lesioni di dimensioni più grandi;
- $80000 \text{ mm}^3 - 400000 \text{ mm}^3$: 11 pazienti presentano lesioni molto grandi.

Il grafico mostra che la maggior parte dei pazienti ha lesioni di dimensioni comprese tra 4000 mm^3 e 80000 mm^3 mentre un numero minore di pazienti presenta lesioni estremamente piccole o molto grandi.

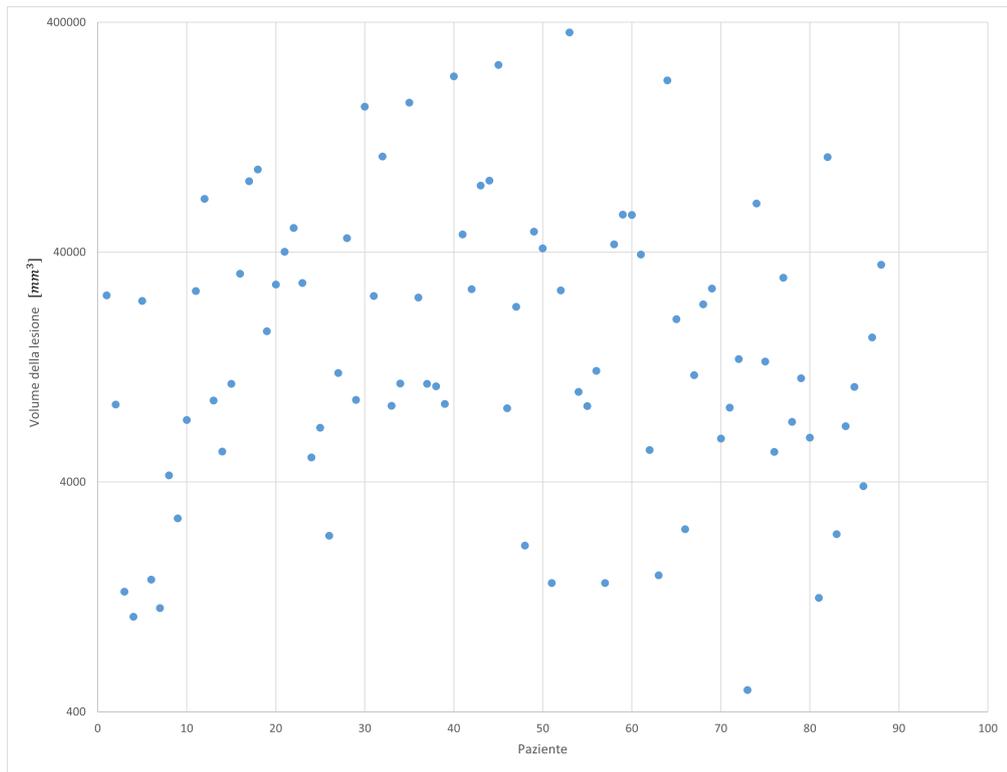


Figura 2.2: Grafico a dispersione rappresentativo del volume della lesione di ogni singolo paziente. Sull'asse x si hanno i pazienti, sull'asse y il volume della lesione.

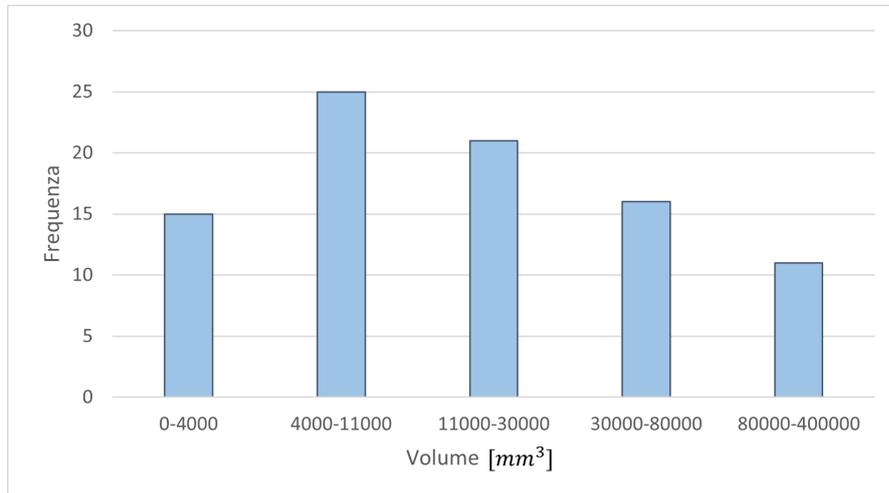


Figura 2.3: Grafico rappresentativo della frequenza dei volumi in intervalli differenti.

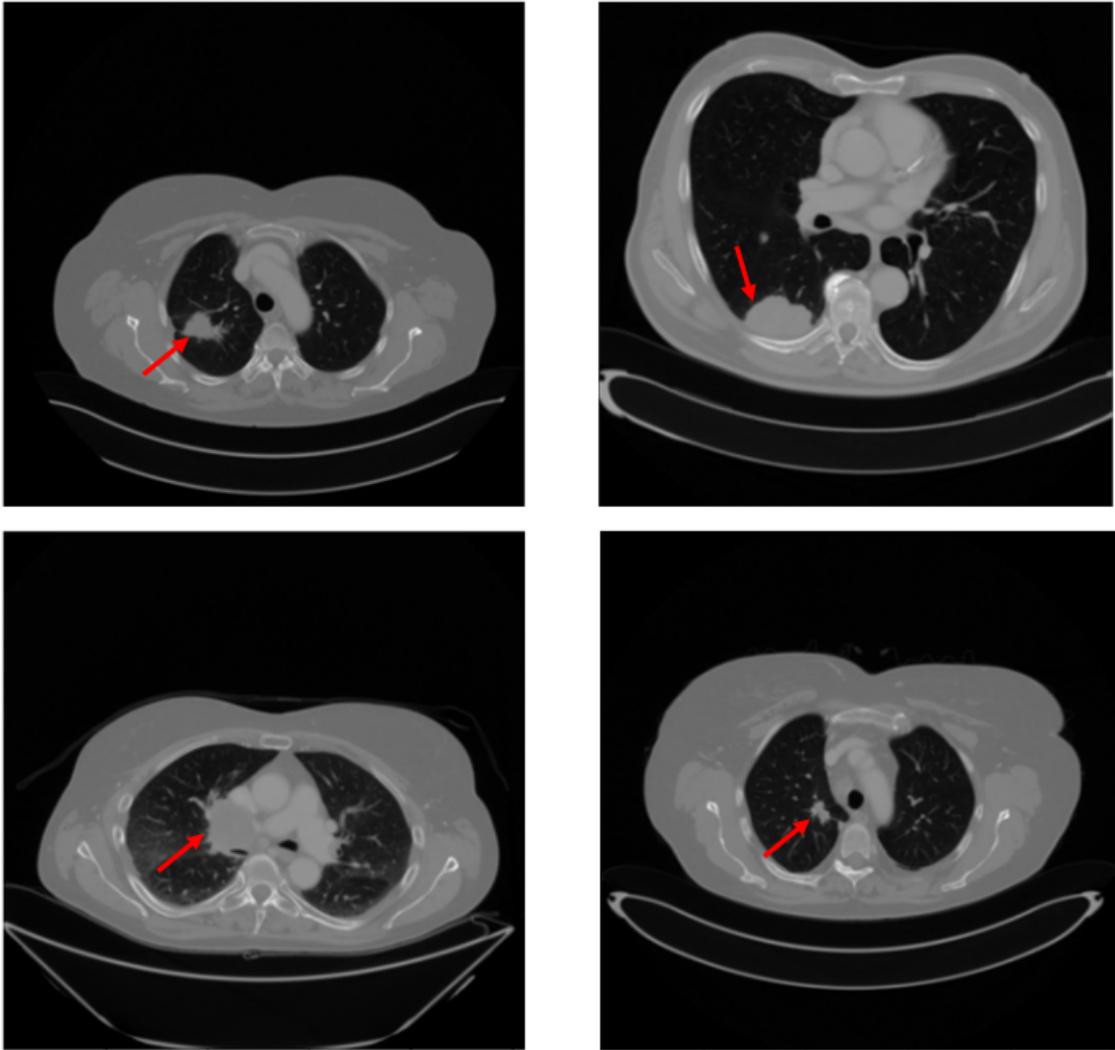


Figura 2.4: Grafico rappresentativo delle posizioni delle lesioni: (a) slice di paziente con lesione in posizione interna al polmone, (b) slice di paziente con lesione alla pleura, (c) slice di paziente con lesione al mediastino, (d) slice di paziente con lesione interna di piccole dimensioni.

Per la gestione e l'elaborazione delle immagini, sono stati utilizzati diversi software esterni. In particolare, *MicroDicom*, *LIFEx* e *3DSlicer*, software open-source, sono stati impiegati durante la fase iniziale di osservazione, consentendo un'analisi visiva delle immagini e delle maschere associate. D'altra parte, *RayStation* è un software commerciale messo a disposizione da Tecnologie Avanzate ed è un Treatment Planning System (TPS). Offre funzionalità avanzate di segmentazione finalizzata alla definizione dei margini dei tumori e degli organi sani. Esso ha fornito supporto

nella fase di pre-processing, facilitando la preparazione e l'organizzazione dei dati per ulteriori analisi. L'integrazione di questi strumenti ha permesso di ottimizzare il flusso di lavoro e migliorare la qualità complessiva delle operazioni svolte.

Il flowchart, in Figura 2.5, rappresentativo del progetto prevede più fasi: associazione della serie DICOM con l'RTSTRUCT, pre-processing, bilanciamento del dataset, data augmentation, rete neurale, testing, post-processing e calcolo delle metriche.

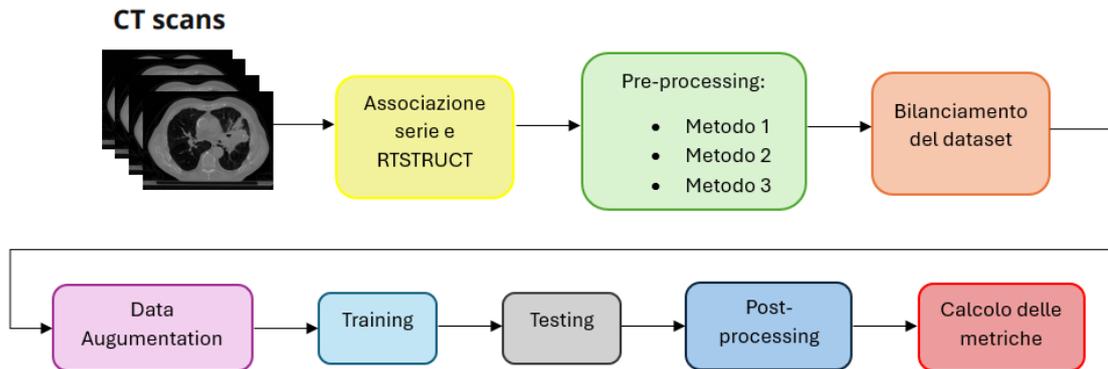


Figura 2.5: Flowchart relativo al complessivo flusso di lavoro.

2.2 Associazione serie e RTSTRUCT

Per ogni paziente sono stati messi a disposizione:

- le serie di CT toraciche in formato DICOM ¹;
- l'RTSTRUCT, contenente i contorni manuali effettuati dal radiologo del volume 3D del tumore macroscopico primario e di lesioni secondarie;
- un file '.xlsx' che indica i nomi delle regioni di interesse (ROI) corrispondenti alle lesioni target primarie per ciascun paziente, da tenere in considerazione in questa prima fase del lavoro.

I dati sopra descritti sono stati organizzati tramite diverse fasi di pre-elaborazione prima che venissero utilizzati per l'addestramento e l'inferenza.

In una CT, una serie si riferisce a un insieme di scansioni acquisite durante l'intero esame, che rappresentano diverse fasi o angolazioni dell'area da esaminare. Ogni paziente può avere più serie, a seconda della complessità dell'esame e delle tecniche

¹Digital Imaging COmmunications in Medicine

utilizzate, come l'acquisizione di immagini senza mezzo di contrasto seguite da immagini con mezzo di contrasto.

In questo studio, per ogni paziente sono state fornite più serie, ma non tutte state oggetto di studio. Infatti, per ogni paziente è stato eseguito un processo di estrazione e caricamento, per combinare i dati della lesione manuale e della serie corrispondente da file diversi in un'unica directory che consente una gestione più semplice e strutturata.

Per identificare la serie su cui è stata tracciata la segmentazione manuale è stata creata una procedura automatica che esplorasse la cartella principale, in cui sono presenti i dati di ogni paziente in modo distinto, e tutte le sue sottocartelle per ricercare i file ".dcm". Per ogni file DICOM individuato, vengono estratte informazioni significative dalle relative tag DICOM, in particolare è stato sfruttato l'UID della serie con *SeriesInstanceUID*. In seguito, si sono esaminati i file RTSTRUCT per individuare la serie a cui essi sono associati: ogni RTSTRUCT contiene informazioni su una o più serie di immagini tramite la tag *ReferencedFrameOfReferenceSequence*. A questo punto, per ogni serie DICOM identificata, l'algoritmo confronta l'UID della serie con quelli elencati nel file RTSTRUCT. Quando viene trovata una corrispondenza, vengono estratti e registrati i dettagli della serie.

Successivamente si caricano sia le immagini DICOM della serie che il relativo file RTSTRUCT in cui sono presenti le ROI. Utilizzando il file '.xlsx' fornito, si seleziona il nome della ROI corrispondente alla lesione di interesse, che viene poi utilizzato per estrarre la maschera 3D della lesione. Quest'ultima viene elaborata per ricavare le sue slices 2D lungo l'asse z. Tutte le maschere, così come i file DICOM associati, hanno dimensioni di 512x512 pixel.

I dettagli della serie registrati sono:

- manufacturer, informazione utilizzata esclusivamente per identificare il produttore del dispositivo di acquisizione, non impiegata direttamente nell'analisi, ma utile per comprendere la presenza di più macchinari diversi che possono influenzare la qualità e le caratteristiche delle immagini prodotte;
- rescale intercept e rescale slope per convertire i valori dei pixel in valori in HU;
- slice thickness per ottenere la spaziatura dei voxel, quindi la risoluzione spaziale lungo la direzione z;
- pixel spacing per ottenere la risoluzione spaziale nella direzione x-y;
- rows e columns per determinare la dimensione delle immagini in termini di numero di pixel;

- window level e window width per definire la gamma di valori di intensità visualizzati, permettendo una migliore osservazione delle strutture di interesse nelle immagini.

Anche queste informazioni, in aggiunta a quelle descritte nel paragrafo 2.1, mostrano come ogni dato sia differente dall'altro, evidenziando la varietà intrinseca del dataset. Questa diversità sottolinea la necessità di procedere con un'accurata procedura per uniformare i dati, al fine di ridurre le discrepanze e garantire una maggiore coerenza nell'analisi e nei risultati. Solo attraverso un processo di normalizzazione appropriato sarà possibile ottenere un dataset più omogeneo e adatto a un'analisi comparativa affidabile. Questo processo comprende una conversione in HU, seguita da un crop uniforme basato su un'area di scansione identica per tutti i pazienti, da un ricampionamento delle immagini per garantire coerenza nella risoluzione e da un adattamento finale per rendere tutte le immagini della stessa dimensione.

2.3 Pre-processing



Figura 2.6: Flowchart del pre-processing comune a tutti i metodi successivamente sviluppati.

Per ogni paziente, è stata caricata l'intera serie di immagini DICOM rappresentante una scansione volumetrica della CT utilizzando la libreria Pydicom.

Le immagini caricate sono state poi ordinate lungo l'asse z , ovvero la direzione di acquisizione della scansione, utilizzando la tag DICOM *ImagePositionPatient*, che indica la posizione spaziale di ciascuna slice rispetto alle altre. Questo passaggio è cruciale per ottenere la corretta ricostruzione tridimensionale del volume. Per garantire un esatto ordinamento delle immagini lungo l'asse z , è stata utilizzata un'altra tag DICOM, *InstanceNumber*, come secondo controllo, assicurando che ogni slice fosse inclusa correttamente nella ricostruzione del volume.

Successivamente, le immagini sono state convertite in HU, la scala di misura che standardizza i valori di attenuazione dei raggi X per rappresentare in modo consistente le proprietà dei tessuti corporei in termini di densità. Questa conversione è stata effettuata applicando i valori di *RescaleIntercept* e *RescaleSlope* forniti nei metadati DICOM di ciascuna serie, correggendo così i valori di intensità per riflettere le HU corrette. La formula per convertire i valori di pixel grezzi in HU è la seguente:

$$HU = (ValorePixel \times RescaleSlope) + RescaleIntercept$$

- Valore del pixel indica il valore di intensità grezzo per ciascun pixel della slice;
- *RescaleSlope* è un fattore di scala che modula il valore del pixel;
- *RescaleIntercept* è un valore che viene aggiunto dopo la scalatura, per convertire i valori di pixel nel range HU corretto.

Come parte della stessa pipeline, le serie DICOM sono state salvate come immagini in scala di grigi 2D nel formato '.png'. Parallelamente, le maschere manuali di segmentazione del GTV sono state archiviate come immagini binarie, mantenendo le stesse coordinate delle scansioni originali ordinate. Questo processo garantisce che tutte le maschere delle lesioni e le immagini della serie siano organizzate e

pronte per ulteriori analisi, rendendo così facilmente identificabili l'immagine e la sua maschera corrispondente, poiché condividono la stessa posizione lungo l'asse z.

Successivamente alle elaborazioni comuni sopra elencate, sono stati implementati tre approcci di pre-processing distinti. Ogni scelta metodologica è stata effettuata sulla base delle conclusioni e delle criticità ottenute dai risultati del metodo precedente.

2.3.1 Metodo 1 di pre-processing: ROI attorno al polmone

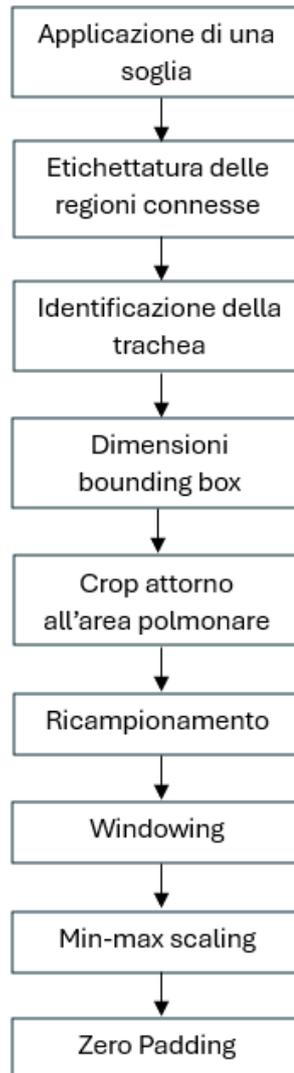


Figura 2.7: Flowchart del pre-processing delle CT del primo approccio.

In questo metodo, finalizzato a effettuare un ritaglio sulle immagini in una zona centrata sull'area dei polmoni, è stato necessario implementare un processo di estrazione della maschera polmonare per isolare le sole strutture polmonari all'interno del volume di imaging. Questa procedura prevede più passaggi fondamentali schematizzati in Figura 2.7.

Focalizzandoci sulla Figura 2.8, possiamo osservare un'immagine di CT toracica

che mette in evidenza principalmente i polmoni. Queste due strutture, di forma simile e posizionate al centro dell'immagine, sono ben definite e occupano una porzione significativa del volume toracico. I polmoni sono circondati dalla cavità toracica, che costituisce il contorno esterno dell'immagine, e all'interno di essa si trovano altri organi vitali, come il cuore e il midollo spinale, che sono localizzati nella parte mediana. Nella parte inferiore dell'immagine, delle linee orizzontali rappresentano il lettino su cui è posizionato il paziente durante l'esame.



Figura 2.8: Slice d'esempio di una scansione CT toracica.

In Figura 2.9 si ha l'esempio di un volume visualizzato da 3DSlicer, che potrebbe simulare ciò che la CT rileva in una scansione toracica. Da questa si comprende l'importanza di isolare esclusivamente i polmoni. Infatti, da quest'immagine si nota la presenza di ulteriori strutture corporee, che potrebbero confondere gli algoritmi e complicare l'analisi. In 2D questo significa che ci sono tante slices in cui i polmoni non sono nemmeno presenti, poiché rappresentano sezioni del corpo in cui si trovano altri organi situati sopra o sotto i polmoni. Queste slices, che possono essere in grande quantità, non sono utili per l'analisi specifica dei polmoni. Di tutte le slice della scansione, solo circa il 30-40% rappresenta effettivamente i polmoni, un numero nettamente inferiore rispetto al totale delle sezioni fornite dal dataset originale. Questo rende essenziale un attento processo di selezione per isolare le sole slice pertinenti.

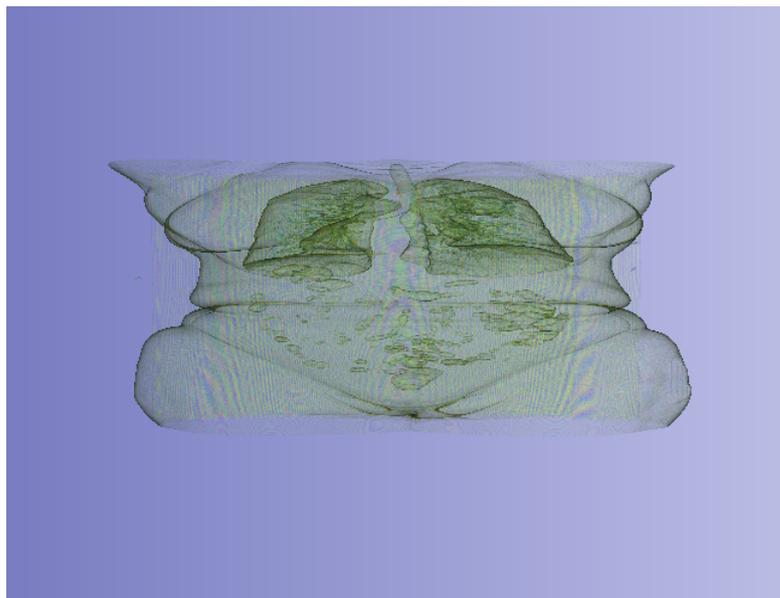


Figura 2.9: Volume intero della scansione CT.

Pertanto, è stata estratta la maschera polmonare dal volume 3D. Per iniziare, partendo da immagini CT di dimensione 512x512, viene generata una maschera binaria sulla base dei valori HU del volume. I pixel che hanno un valore in HU superiore a -400 sono stati considerati come appartenenti ai polmoni, mentre i valori inferiori vengono classificati come sfondo:

$$\begin{cases} \text{valore pixel} > -400 & \text{polmoni} \\ \text{valore pixel} < -400 & \text{sfondo} \end{cases}$$

In Figura 2.10 è rappresentato l'esempio di una slice di output dell'operazione di sogliatura.

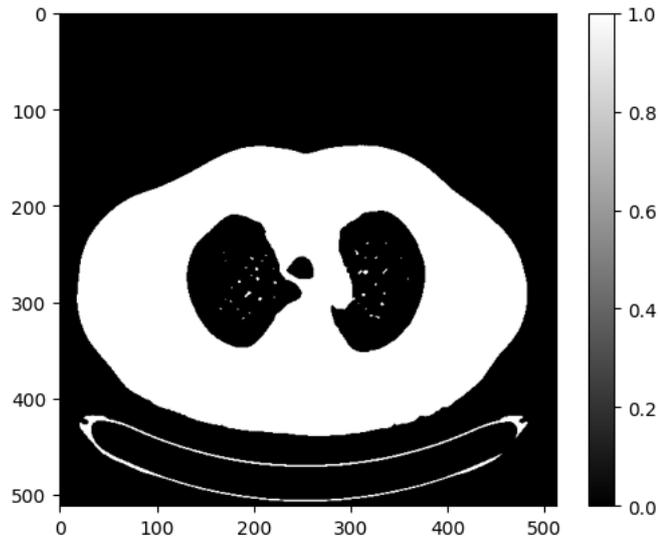


Figura 2.10: Esempio di slice di output dell'operazione di sogliatura.

La scelta di -400 HU come soglia è motivata dal fatto che il valore associato alla radiodensità dei tessuti polmonari nelle scansioni CT si aggira tra $\{-500, 900\}$ HU. Come si vede dalla Figura 2.11, i valori inferiori a -400 HU tendono a rappresentare aree non polmonari, come il tessuto adiposo o le strutture ossee circostanti.

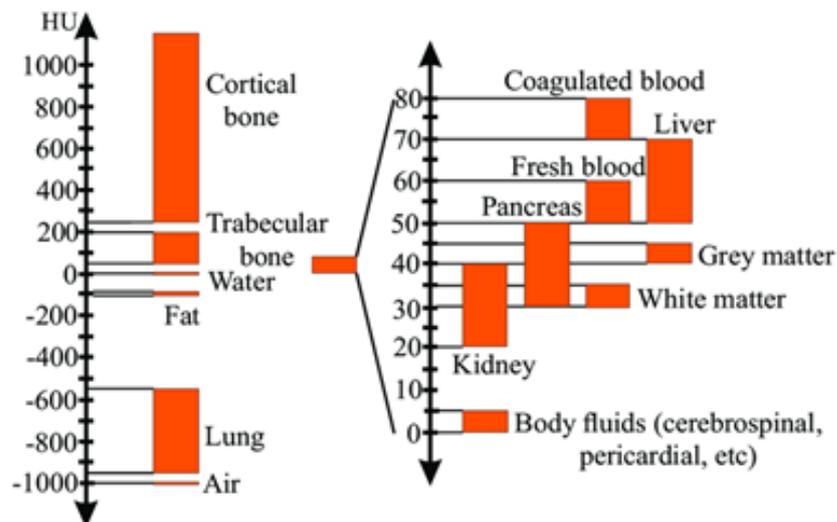


Figura 2.11: Scala Hounsfield.

Questa scelta è confermata anche dall'istogramma del volume di una CT, rappresentato in Figura 2.12. Questo riproduce la distribuzione dei valori HU all'interno

del volume. Sull'asse x viene indicato il valore HU e sull'asse y la frequenza, in scala logaritmica, con cui questi valori si verificano nel volume in esame. La distribuzione dell'istogramma mostra 2 picchi principali: un picco attorno a valori negativi che rappresenta la forte presenza di aria e suggerisce che nel volume è presente un'ampia area polmonare e di vuoto, e un picco vicino allo 0 indicativo dei tessuti molli.

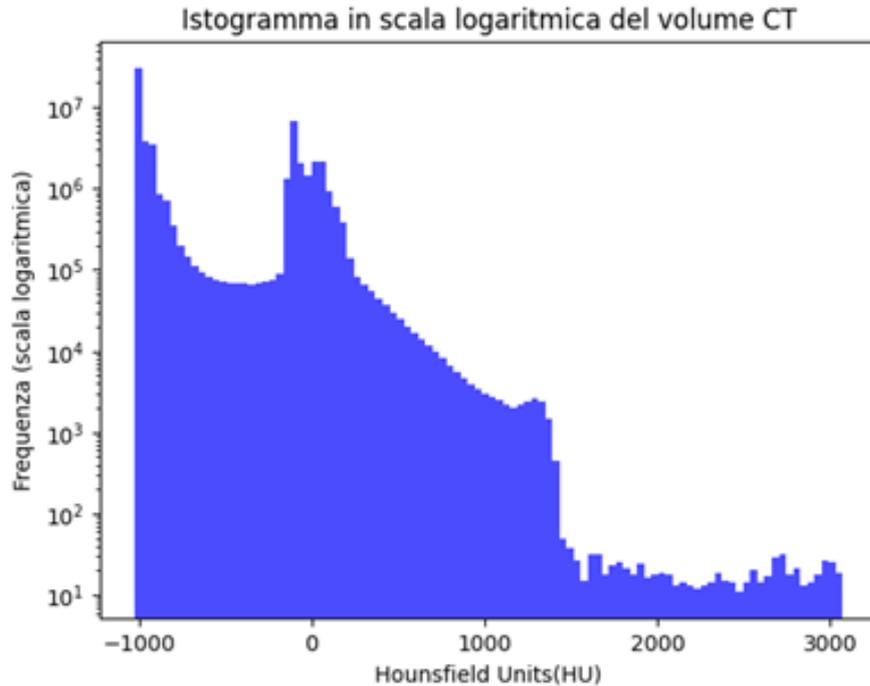


Figura 2.12: Istogramma rappresentativo di un volume di CT in scala logaritmica.

A questo punto, è essenziale distinguere correttamente lo sfondo dal resto del volume per evitare che esso venga erroneamente considerato come parte delle strutture anatomiche. A tale scopo, viene impiegata la tecnica di etichettatura delle regioni connesse. Quest'ultimo è un metodo utilizzato nelle elaborazioni delle immagini per differenziare le diverse aree connesse di un'immagine binaria, nel nostro caso di un volume binario. Le regioni connesse rappresentano gruppi di pixel bianchi o neri che sono contigui, ovvero collegati direttamente l'uno all'altro. Questo processo restituirà più regioni connesse, identificate da un'etichetta univoca, associando a ogni pixel l'etichetta della regione di cui fa parte. In questo modo si separano, si differenziano e si assegna un identificatore univoco a diverse aree connesse di un'immagine binaria. È una tecnica che permette di trattare regioni composte da pixel adiacenti come entità separate.

In Figura 2.13 è mostrato il risultato prodotto da questa metodologia. Sebbene, visivamente, la distinzione non sia percepita in modo ottimale, è comunque evidente

che ad ogni regione connessa viene assegnata un'etichetta, rappresentata da un colore diverso per ciascuna parte.

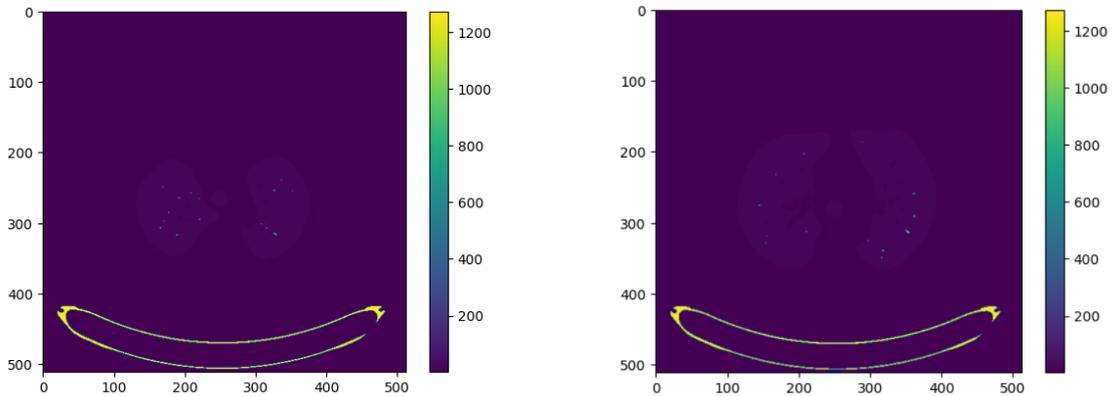


Figura 2.13: Esempi di slices di output dell'operazione di labelling in cui si distingue la regione dello sfondo, del lettino, dei polmoni e delle sue strutture interne rappresentate da piccoli puntini impercettibili.

Una volta create le etichette per ogni regione connessa all'interno del volume, per identificare lo sfondo è stato sfruttato il fatto che l'angolo superiore sinistro del volume, con coordinate $[0,0,0]$ tipicamente si trova al di fuori del corpo del paziente, in un'area che potrebbe rappresentare l'aria circostante o il vuoto. Si assume quindi che l'intera regione connessa a questo punto appartenga allo sfondo, e tramite l'etichetta associata a esso vengono identificate tutti gli altri punti che appartengono allo sfondo e classificati come tali.

Per migliorare ulteriormente la precisione della segmentazione, si procede al riempimento delle eventuali cavità presenti all'interno dei polmoni. Queste, se non trattate, potrebbero causare errori nella segmentazione, poiché potrebbero essere erroneamente considerate come aree esterne ai polmoni. Utilizzando nuovamente la tecnica di etichettatura delle regioni connesse, vengono identificate e riempite tutte le aree che non appartengono alla regione polmonare principale, garantendo una segmentazione accurata che include l'intero volume polmonare senza interruzioni.

Successivamente, è necessario escludere altre parti indesiderate nella scansione CT, come il lettino, il torace e altre strutture anatomiche. Dal momento che i polmoni rappresentano il volume più grande in una scansione 3D, essi possono essere individuati come la regione connessa di maggiore dimensione, facilitando la loro identificazione attraverso l'etichetta corrispondente.

Dopo aver completato questi passaggi per estrarre la maschera 3D del volume polmonare, il risultato ottenuto su una singola slice è mostrato in Figura 2.14.

Dalle immagini del labelling, precedentemente mostrate, è stato quindi eliminato lo sfondo, riempite le strutture interne ai polmoni, è stato eliminato il lettino prendendo la struttura connessa di maggior dimensione. Si può osservare come venga isolata esclusivamente la maschera dei polmoni rispetto alla rappresentazione iniziale.

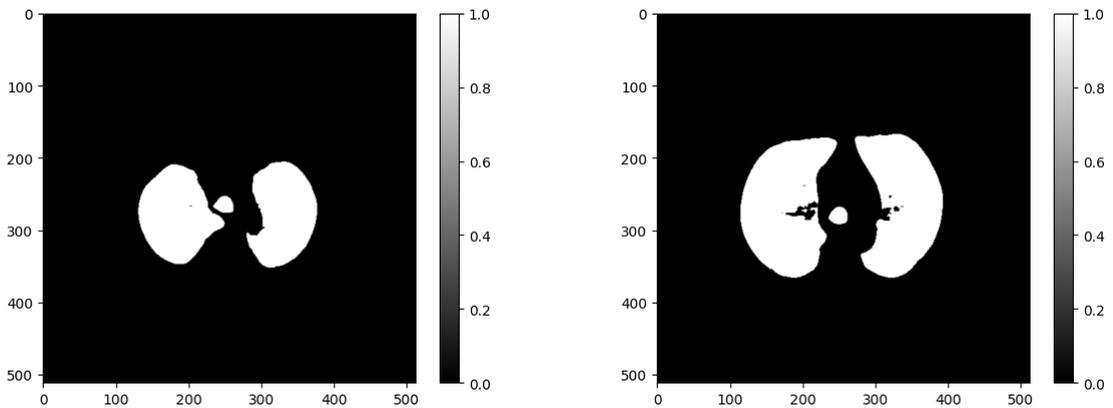


Figura 2.14: Esempi di slices di output dell'operazione di estrazione della maschera polmonare. Rappresentano sezioni del polmone differenti.

La piccola regione mascherata, quasi in posizione centrale alle slices, rappresentata da un cerchio, corrisponde alla trachea. Poiché nelle slices della parte apicale della maschera era presente solo quest'organo, queste sono state eliminate, in quanto non contenevano ancora i polmoni.

Infatti, nel processo di segmentazione polmonare è comune che strutture come la trachea vengano erroneamente incluse nella maschera segmentata dei polmoni, in quanto anche nella trachea è presente aria e perciò nella CT i suoi valori HU si avvicinano a quelli dei polmoni. Inoltre la trachea è un volume connesso ai polmoni. Questo significa che la maschera polmonare finora realizzata presenta porzioni del volume identificate come polmoni ma che in realtà appartengono alla trachea. Per cui, il passaggio seguente comporta la rimozione di quest'organo al fine di eliminare successivamente le porzioni del volume lungo l'asse z dove i polmoni non sono presenti. Il metodo implementato per rimuovere la trachea si basa su una semplice osservazione morfologica anatomica. In particolare, considerando le slices 2D, la trachea occupa un'area significativamente inferiore rispetto ai polmoni, pertanto viene adottato un criterio basato sulla somma dei pixel segmentati, quelli con valore 1, in ciascuna fetta assiale. Se il numero di pixel segmentati in una fetta è inferiore a una soglia di 4500 pixel in quella fetta è presente solo la trachea e viene quindi

rimossa assegnando un valore di 0 a tutti i pixel della fetta:

$$\begin{cases} \text{somma dei pixel bianchi} > 4500 & \text{polmoni} \\ \text{somma dei pixel bianchi} < 4500 & \text{trachea} \end{cases}$$

La soglia di 4500 pixel è stata selezionata sulla base dell'osservazione delle maschere nella parte superiore del volume. In un'immagine di dimensione 512x512, un'area di 4500 pixel è significativamente inferiore rispetto alla dimensione effettiva delle strutture polmonari. Pertanto, un numero così ridotto di pixel non può rappresentare correttamente i polmoni, ma è più probabile che corrisponda a strutture non polmonari, come la trachea.

L'estrazione del volume polmonare, lavorando in 2D, si traduce in un'eliminazione di tutte le slices che includono parti del corpo non rilevanti per la nostra analisi, ovvero in un crop lungo l'asse z secondo il salvataggio della prima e dell'ultima slice in cui sono presenti i polmoni. Per concentrare il lavoro successivo sulle porzioni di volume che contengono solo i polmoni, le slice assiali superiori e inferiori ai polmoni stessi vengono eliminate. Per delimitarne l'area lungo l'asse z ed effettuare il successivo ritaglio, viene esaminata ogni slice della maschera, finora creata, verificando la presenza di almeno un pixel con valore pari a 1. Se questa condizione fosse soddisfatta, indicherebbe che i polmoni erano presenti in quella fetta. Questo approccio ha consentito di salvare la posizione della prima e dell'ultima slice in cui erano presenti i polmoni, permettendo successivamente di effettuare il crop del volume eliminando le porzioni al di fuori di queste coordinate. Questo riduce la dimensione del volume lungo l'asse z, comportando una diminuzione del numero di slices bidimensionali, focalizzando l'analisi sulle strutture polmonari di reale interesse clinico per lo studio.

Per lo stesso obiettivo, anche sul piano x-y è stata delimitata l'area che racchiudesse i polmoni. In particolare, per ogni fetta della maschera nel range definito al passo precedente, viene calcolato il gradiente utilizzando il filtro di Sobel². Questo passaggio consente di rilevare i bordi delle strutture presenti all'interno delle fette, cioè i bordi dei polmoni. Da questi contorni si identificano le coordinate dei pixel che corrispondono al valore massimo e minimo del contorno stesso. Per ogni fetta si avranno 4 coordinate, 2 lungo x e 2 lungo y, che rappresentano i punti nel piano x-y in cui sono compresi i polmoni.

²Operatore differenziale che calcola un valore approssimato del gradiente di una funzione che rappresenta la luminosità dell'immagine.

I risultati di questo passaggio sono mostrati in Figura 2.15. Nella prima immagine si vedono chiaramente i contorni delle vie aeree superiori e di una piccola porzione dei polmoni. I bordi indicano che la slice è situata nelle parti superiori del volume polmonare, dove i polmoni sono meno sviluppati e sono visibili porzioni di trachea o bronchi principali. La seconda immagine mostra una porzione più centrale dei polmoni, infatti il loro contorno è ben delineato. La terza slice rappresenta la parte bassa dei polmoni, dove si nota la caratteristica forma a "U" degli organi.

Per effettuare un crop uniforme su tutto il volume, sono state scelte come dimensioni della bounding box le coordinate massime e minime ottenute tra tutte le fette del volume dello stesso paziente. Si è stato eseguito un crop sul volume.

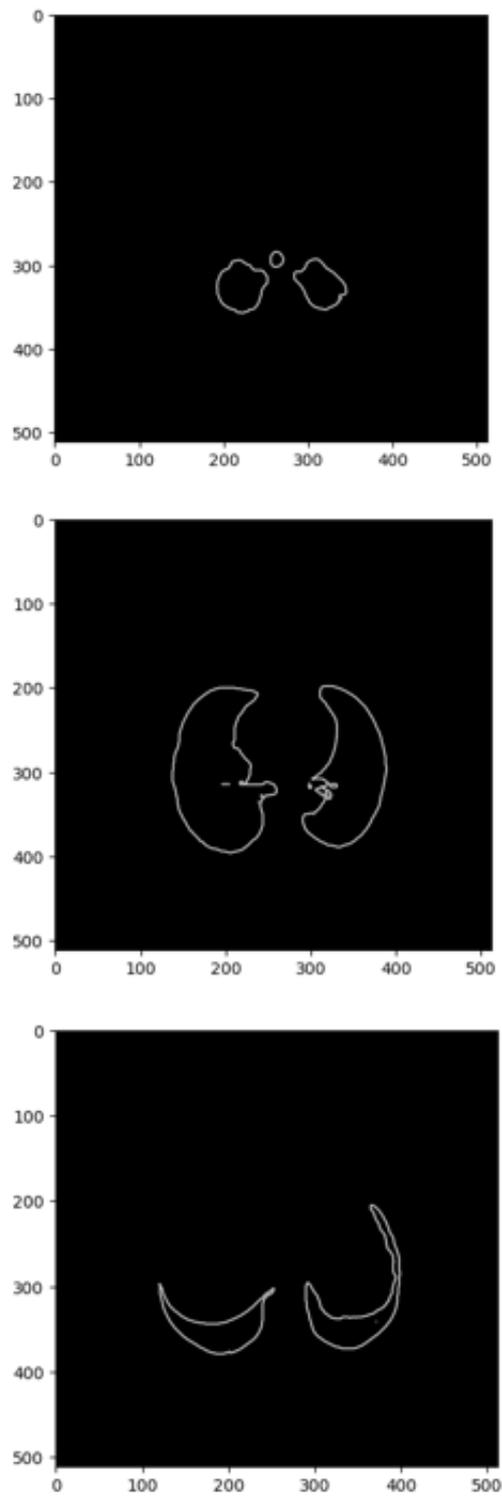


Figura 2.15: Rappresentazione dei contorni della maschera polmonare su diverse slices.

Per garantire l'uniformità delle immagini tomografiche e facilitare la successiva elaborazione con la rete neurale, è stato eseguito un processo di ricampionamento. Come descritto nel paragrafo 2.2, sono stati estratti i valori della spaziatura dei pixel, ovvero della risoluzione spaziale, fondamentali per effettuare il ricampionamento. Per standardizzare la risoluzione spaziale delle immagini e rendere comparabili le diverse scansioni, è stata scelta come nuova risoluzione spaziale la media tra tutti i valori della vecchia risoluzione di ogni paziente. La scelta è ricaduta su un valore medio, pari a 0,787 mm, in quanto la risoluzione spaziale variava molto tra un paziente e l'altro, come è figura nell'istogramma 2.16, con un valore minimo di 0,607 mm e un massimo di 0,976 mm.

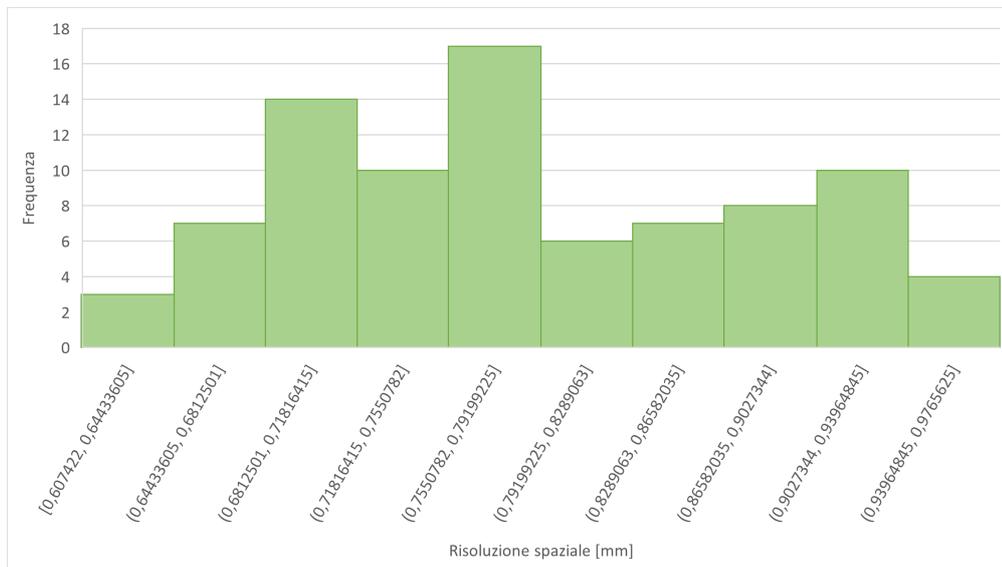


Figura 2.16: Istogramma rappresentativo dell'andamento delle risoluzioni spaziali delle CT.

La scelta errata della nuova risoluzione spaziale avrebbe potuto causare 2 effetti indesiderati:

- Sotto-ricampionamento, causando la perdita di dettagli;
- Sovra-ricampionamento, generando falsi dettagli e aumento del rumore.

Questo ricampionamento è stato essenziale per ottenere immagini con una risoluzione uniforme su tutto il volume. Il ricampionamento è stato quindi effettuato applicando un fattore di ridimensionamento alle immagini in base alla spaziatura media dei pixel calcolata. Questo rapporto indica quanto le dimensioni dell'immagine

devono essere scalate per ottenere la risoluzione desiderata:

$$\text{Resize Factor} = \frac{\text{Old spacing}}{\text{New spacing}}$$

- Old Spacing indica la risoluzione spaziale iniziale
- New Spacing indica la risoluzione spaziale che si vuole ottenere

È stato utilizzato un algoritmo di interpolazione per ridimensionare le immagini lungo gli assi x e y e per riempire correttamente i pixel nella nuova griglia di risoluzione, mantenendo la qualità delle immagini e garantendo che i dettagli anatomici fossero preservati. Nel nostro caso si è utilizzata un'interpolazione spline di terzo ordine. Essa utilizza una funzione polinomiale cubica per determinare i valori dei pixel nelle nuove coordinate spaziali. Questo metodo è scelto per bilanciare qualità e prestazioni:

- alta qualità: riducendo al minimo la sfocatura e preservando i dettagli fini, rispetto a metodi di ordine inferiore come l'interpolazione lineare (order=1);
- riduzione degli artefatti visivi: come aliasing che può verificarsi con interpolazioni di ordine inferiore.

Dopo aver effettuato il ricampionamento delle immagini, si è valutata l'applicazione di una tecnica di windowing per migliorare la visualizzazione delle strutture anatomiche di interesse. Essa viene utilizzata per ottimizzare la rappresentazione dei valori di intensità di strutture di interesse in modo che siano facilmente interpretabili dalla rete, facilitando così la sua capacità di analizzare e classificare le immagini in modo più accurato. Nelle immagini CT, i valori delle unità Hounsfield possono variare ampiamente, da valori negativi (aria, -1000 HU) a valori positivi (osso, da +400 a +3000). Tuttavia, per visualizzare alcune strutture anatomiche specifiche, è sufficiente considerare solo una parte dei valori delle unità Hounsfield. In questo contesto, si seleziona una porzione specifica della gamma totale di HU, mappando i valori al di fuori di questo intervallo come completamente neri per quelli superiori al valore massimo, e bianchi per quelli inferiori al valore minimo, e rappresentando la gamma interna con una scala di grigi:

$$\begin{cases} \text{valori pixel} > \text{max} = \text{max} \\ \text{valori pixel} < \text{min} = \text{min} \end{cases}$$

I parametri chiave del windowing, visualizzati in Figura 2.17, sono:

- Window Level (WL) rappresenta il valore centrale della gamma di intensità che si desidera visualizzare. È il valore HU intorno al quale viene centrata la finestra.
- Window Width (WW) definisce l'intervallo di valori HU che sarà mappato sulla scala dei grigi. Una window width stretta permette di visualizzare con maggiore contrasto piccoli dettagli in una ristretta gamma di HU, mentre una window width più ampia è utile per visualizzare un'ampia gamma di tessuti con un contrasto ridotto.

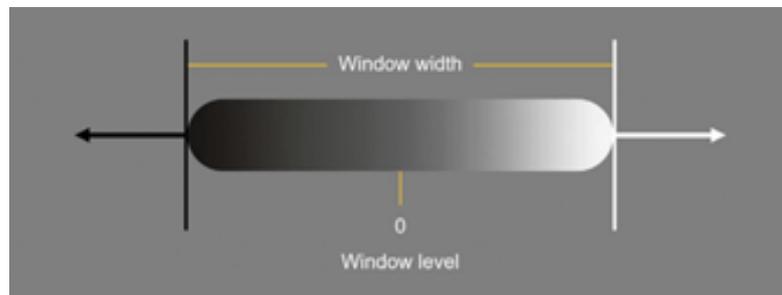


Figura 2.17: Window Level e Window Width.

Nel presente studio è stato valutato un windowing sul range di valori Hounsfield { -300, 400 } per evidenziare le lesioni. Per cui si è impostato:

- WW = 700
- WL = 50

Nella Figura 2.18 è rappresentata la sovrapposizione degli istogrammi delle lesioni di diversi pazienti in cui si verifica che sono concentrate tutte intorno allo 0, tra -200 e 200. L'area della curva di ogni istogramma è differente in quanto dipende dalla grandezza della lesione.

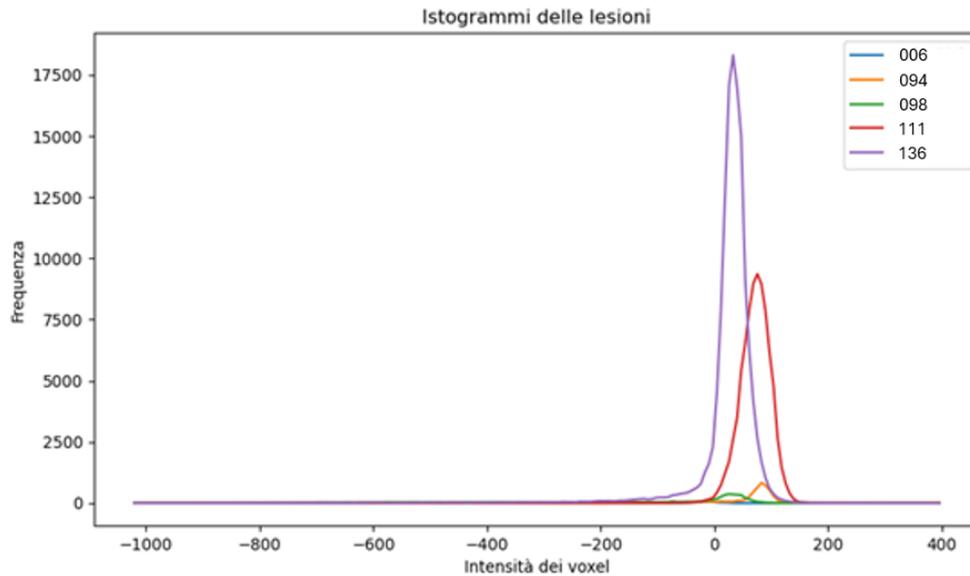


Figura 2.18: Confronto tra istogrammi delle lesioni di diversi pazienti.

A questo punto è stata applicata la normalizzazione min-max scaling con l'obiettivo di ridimensionare i valori di intensità delle immagini in un intervallo specifico compreso tra 0 e 1:

$$X_{normalizzato} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

- X è il valore originale del pixel
- X_{min} è il valore minimo del volume
- X_{max} è il valore massimo del volume

Il min-max scaling è particolarmente utile in contesti di apprendimento automatico poiché consente di ridurre la sensibilità del modello a variazioni nei valori di intensità, migliorando così la stabilità e la velocità della convergenza durante il processo di ottimizzazione. Inoltre, il min-max scaling aiuta a prevenire che alcune caratteristiche dominino il processo di apprendimento, consentendo alla rete di apprendere in modo più efficace le informazioni rilevanti presenti nelle immagini.

Come conclusione alle operazioni di pre-processing mirate a uniformare il dataset è stato necessario applicare lo zero padding. Questa scelta è dovuta al fatto che, in seguito al processo di ricampionamento è emerso che le dimensioni delle immagini differivano tra i vari pazienti, a causa della risoluzione spaziale originale delle scansioni e di quella attuale. In particolare, per ogni paziente, si è scelto di uniformare le dimensioni di larghezza e altezza nel piano x-y a quelle massime riscontrate tra tutti i pazienti. Questa operazione ha permesso di ricondurre le slices di tutti i pazienti a una stessa dimensione pari a 449x315.

Pre-processing delle maschere manuali

Per le maschere, sono state effettuate tre operazioni di pre-processing principali: il crop attorno all'area polmonare considerando le stesse coordinate del volume CT, il ricampionamento e lo zero padding.

A differenza delle immagini CT, nel ricampionamento si è utilizzata l'interpolazione nearest-neighbor, fondamentale per preservare la precisione delle aree segmentate. Dopo il ricampionamento, è stato eseguito un riempimento dei buchi per riparare eventuali danni derivanti dall'ultimo processo, garantendo così che le maschere risultassero complete e coerenti. Questo passaggio è stato essenziale per mantenere l'integrità delle informazioni.

2.3.2 Metodo 2 di pre-processing: ROI attorno al centro di massa

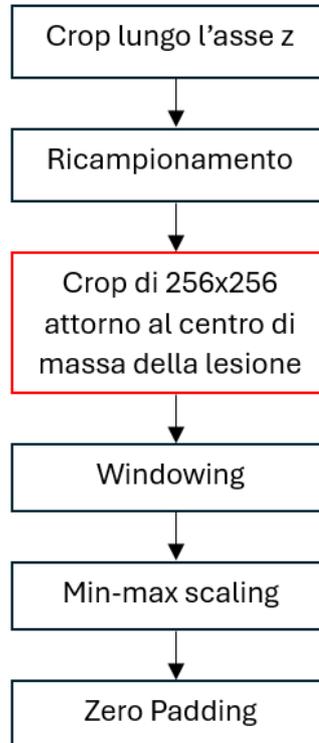


Figura 2.19: Flowchart del pre-processing delle CT del secondo approccio.

Al primo approccio ne è seguito un secondo con l'obiettivo di rendere le immagini non solo più uniformi e standardizzate tra loro, ma anche più informative a livello visivo. A tal fine, è stato effettuato un crop attorno al centro di massa delle lesioni. Quest'operazione ha permesso di focalizzare l'attenzione sulle aree in cui il tumore era più evidente, migliorando così la visibilità di eventuali anomalie. Gli step seguiti sono schematizzati in Figura 2.19.

Uguualmente all'approccio precedentemente descritto, inizialmente è stato effettuato un crop sull'asse z per delimitare l'area polmonare. Per fare ciò, sono state utilizzate le posizioni della prima e dell'ultima slices in cui sono presenti i polmoni, salvate dall'algoritmo spiegato nel paragrafo 2.3.1. È seguito il ricampionamento svolto con le stesse modalità. A questo punto, le immagini CT sono state ritagliate utilizzando come punto centrale la coordinata del centro di massa del tumore, ottenuta dalla maschera manuale anch'essa ricampionata. Per il ritaglio, è stata utilizzata una bounding box di dimensioni 256x256 pixel. Questa scelta dimensionale ha consentito

di mantenere una visione adeguata delle strutture circostanti al tumore. Se si fosse optato per una dimensione più piccola, si sarebbe corso il rischio di escludere parti della lesione.

Alle stesse condizioni e per gli stessi obiettivi del caso precedente sono stati eseguiti windowing, min-max-scaling e zero padding.

2.3.3 Metodo 3 di pre-processing: estrazione del parenchima polmonare

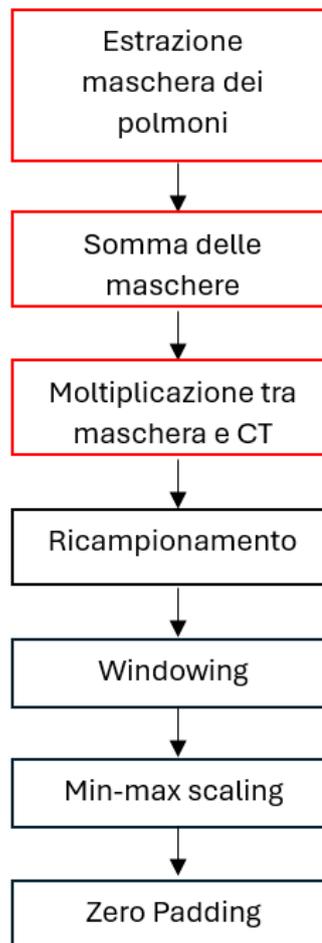


Figura 2.20: Flowchart del pre-processing delle CT del terzo approccio.

Questo terzo approccio prevede l'utilizzo di un software esterno, RayStation. Inserendo la serie del paziente, il programma ha elaborato i dati per generare una nuova RTSTRUCT. Inoltre in input all'applicativo è stato dato anche l'RTSTRUCT

contenente la segmentazione manuale del clinico, per assicurarsi, attraverso un'osservazione visiva, che la maschera dei polmoni generata dal software comprendesse effettivamente le lesioni. L'RTSTRUCT finale contiene i contorni di 2 ROI essenziali: una delinea i contorni del polmone sinistro e l'altra i contorni del polmone destro, rispettivamente denominate 'Lung_L' e 'Lung_R'. Tramite i nomi delle ROI, sono state estratte le maschere binarie del polmone destro e del polmone sinistro. Per facilitare l'analisi le due maschere sono state sommate ottenendone solo una che rappresenti la maschera di segmentazione dei polmoni. Per ottenere la segmentazione vera e propria, la maschera 2D è stata moltiplicata con la CT corrispondente, entrambe di dimensione 512x512:

$$maschera \times CT = \begin{cases} \text{if maschera} == 0 \Rightarrow CT = \text{valore minimo CT} \\ \text{if maschera} == 1 \Rightarrow CT = \text{valore CT} \end{cases}$$

Nel caso in cui la maschera fosse uguale a 0, il valore del pixel dell'immagine di output è stato impostato come il valore minimo della CT affinché i pixel al di fuori dell'area polmonare siano neri. Se, al contrario, fosse stato impostato 0, ciò non era garantito.

In seguito, sono stati eseguiti ricampionamento, windowing, normalizzazione e zero padding alle stesse condizioni degli altri due metodi.

2.4 Suddivisione e bilanciamento del data set

Prima di procedere con il bilanciamento, i pazienti sono stati suddivisi in tre gruppi: training, validation e test. La percentuale di divisione, avvenuta in modo randomico, è la seguente:

- 10% dei pazienti nel test set;
- 20% dei pazienti nel validation set;
- 70% dei pazienti nel training set.

In questo modo, si garantisce che ogni set abbia pazienti distinti, prevenendo la contaminazione tra i dati di allenamento e quelli di validazione.

Una volta completata la suddivisione, tutte le slice dei pazienti appartenenti a ciascun set sono state raggruppate per formare i rispettivi dataset.

Si è ritenuto importante bilanciare il dataset poiché, nonostante fossero già state eliminate tutte le slices in cui i polmoni non erano presenti, le lesioni erano comunque distribuite in un numero relativamente ridotto di fette rispetto al totale. Il bilanciamento delle slices è un passo fondamentale per garantire che il numero di immagini

con lesioni sia equamente distribuito rispetto a quelle senza lesioni. Questo aiuta a prevenire che il modello apprenda con un bias verso una classe specifica, soprattutto se ci sono molte più immagini di un tipo rispetto all'altro. Il bilanciamento avviene secondo una regola ben definita e strutturata rappresentata in Tabella 2.1. Per ogni paziente si analizzano tutte le maschere manuali per determinare la presenza della lesione, controllando se contengono pixel bianchi, che indicano l'esistenza del tumore. In tal modo, le maschere vengono differenziate in due gruppi, quelle contenenti la lesione e quelle completamente nere che corrispondono a slice CT in cui il tumore non è presente.

A questo punto, i due sottoinsiemi vengono trattati in maniera differente:

- Il primo gruppo, contenente le maschere con lesioni, viene mantenuto nella sua totalità, in quanto rappresenta le immagini clinicamente rilevanti per l'addestramento del modello;
- Il secondo gruppo, formato dalle maschere nere, quindi senza lesioni, viene bilanciato selezionando un numero di maschere pari al numero di maschere con lesioni presenti per lo stesso paziente.

Le maschere nere che compongono il secondo gruppo sono state selezionate secondo la seguente logica:

- metà sono state scelte dall'insieme che rappresenta le sezioni più apicali del polmone;
- metà sono state scelte dall'insieme che rappresenta le sezioni più basali del polmone.

In questo modo, viene garantito che la distribuzione delle immagini senza tumore sia bilanciata sia tra la parte superiore che inferiore del polmone. Questo approccio assicura che il dataset rimanga rappresentativo dell'intera anatomia, evitando che il modello sia esposto solo a una porzione limitata delle immagini senza lesione.

Il bilanciamento completo del dataset è stato effettuato in modo da ridurre la predominanza di maschere completamente nere. Questo processo ha garantito un dataset più bilanciato, prevenendo il rischio che il modello impari a classificare tutte le immagini come appartenenti alla classe senza lesioni (sebbene sia più comune). Questo passaggio è cruciale per garantire un training robusto.

Maschere e CT con lesioni	Nella loro totalità
Maschere e CT senza lesioni	50% delle sezioni più apicali del polmone
	50% delle sezioni più basali del polmone

Tabella 2.1: Distribuzione delle maschere e delle CT in relazione alle lesioni polmonari.

La divisione del dataset, e il successivo bilanciamento, ha avuto i seguenti risultati:

- **Training:** 63 pazienti \rightarrow 1705 slices
- **Validation:** 16 pazienti \rightarrow 501 slices
- **Testing:** 9 pazienti \rightarrow 216 slices

2.5 Data Augmentation

Il Data Augmentation è una tecnica utilizzata nel campo del machine learning e della visione artificiale per aumentare artificialmente la quantità e la varietà dei dati di addestramento. Essa consiste nell'applicare trasformazioni casuali o controllate alle immagini di input, come rotazioni, traslazioni, scaling, modifiche di luminosità e contrasto, aggiunta di rumore e flipping, per creare nuove versioni leggermente modificate dei dati originali. Questo processo aiuta il modello a imparare in maniera più efficace, poiché lo espone a una gamma più ampia di variazioni rispetto alle immagini originali, rendendolo meno sensibile a dettagli specifici o variazioni minime che non influenzano la performance reale.

Per migliorare la robustezza e la generalizzazione del modello, sono state implementate una serie di tecniche di data augmentation sui dati di input, quindi sulle immagini e sulle corrispondenti segmentazioni manuali. Le tecniche di data augmentation applicate includono:

- **RandFlipd**: esegue un ribaltamento casuale delle immagini e delle segmentazioni lungo l'asse specificato, in questo caso l'asse 0, con una probabilità del 50%. Questo aiuta a introdurre variazioni orizzontali e verticali, aumentando la diversità del dataset;
- **RandRotate90d**: rotazioni casuali di 90 gradi sulle immagini e le segmentazioni, con una probabilità del 50% e un massimo di 3 rotazioni. Questa operazione è utile per rendere il modello meno sensibile all'orientamento delle immagini;
- **RandAffined**: una trasformazione alle immagini e alle segmentazioni che con una probabilità del 50% permette di effettuare traslazioni, fino a 10 pixel in entrambe le direzioni, e rotazioni fino a 0.1 radiante.

2.6 Rete Neurale: U-Net

La U-Net è un'architettura di rete neurale convoluzionale particolarmente utilizzata per compiti di segmentazione in immagini biomediche. È stata introdotta da Olaf Ronneberger et al. nel 2015 e ha guadagnato rapidamente popolarità per la sua capacità di segmentare in maniera efficace preservando dettagli importanti a diverse scale spaziali. Il nome "U-Net" deriva dalla forma a "U" che caratterizza la struttura della rete, avendo una prima fase di downsampling a cui segue la fase di upsampling. Per quest'implementazione si è utilizzato MONAI (Medical Open Network for AI), una libreria sviluppata specificamente per applicazioni di intelligenza artificiale nel campo della medicina, che offre strumenti e funzionalità avanzate per la segmentazione e l'analisi delle immagini biomediche.

L'architettura della rete U-Net generica è rappresentata in Figura 2.21. In Figura 2.22 è presente il modello della rete utilizzato per questo studio.

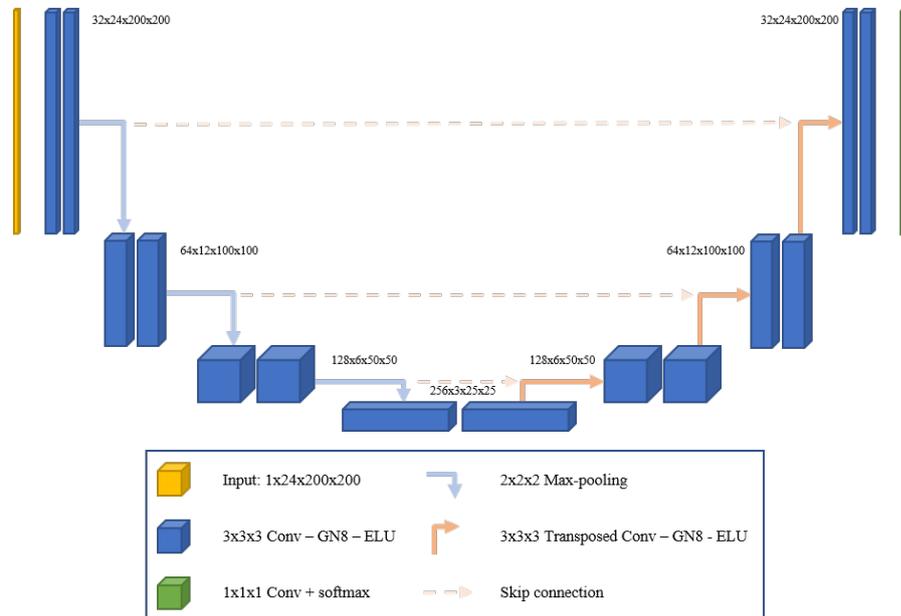


Figura 2.21: Architettura della Rete U-Net.

```

model = UNet (
    spatial_dims = 2,
    in_channels = 1,
    out_channels = 1,
    channels = (16, 32, 64, 128),
    strides = (2, 2, 2),
    num_res_units = 2,
    norm = Norm.BATCH
).to(device)

```

Figura 2.22: Architettura della Rete.

In questo studio, i dati di input sono 2D (immagini bidimensionali), e il modello è monocanale sia in input, siccome si hanno in ingresso immagini di CT in scala di grigi, che in output in quanto si ha in uscita una segmentazione binaria dell'immagine, dove ogni suo pixel indica la probabilità che esso appartenga a una determinata

classe o a una regione specifica. Nella fase di estrazione delle caratteristiche la funzione di convoluzione applica una serie di filtri, chiamati anche *kernel*, che eseguono una scansione ogni immagine, producendo mappe di caratteristiche. La rete utilizzata viene avviata con 16 filtri (canali) nella prima convoluzione, raddoppiando il numero di filtri a ogni passaggio nel percorso di contrazione, seguendo la sequenza (16, 32, 64, 128). Questi numeri indicano quanti filtri vengono applicati ad ogni livello della rete, cioè la dimensione della mappa di caratteristiche in uscita da ciascuno strato convoluzionale.

Lo *stride* definisce di quanto il filtro si sposta quando scorre sull'immagine durante l'operazione di convoluzione. Si imposta uno stride di (2, 2, 2), il che significa che il kernel si sposta di 2 pixel alla volta lungo ogni asse dell'immagine 2D e ogni operazione di pooling ridurrà le dimensioni spaziali della mappa di input della metà.

Dopo ogni convoluzione è applicata la *Batch Normalization*. Questa tecnica normalizza l'output della convoluzione basato sulle medie e sulle varianze calcolate in un mini-batch di dati. Aiuta a stabilizzare e accelerare l'allenamento della rete, riducendo il rischio di scomparsa del gradiente.

Nella fase di espansione (decoding) della U-Net, la rete utilizza la convoluzione trasposta per aumentare la risoluzione spaziale delle mappe di caratteristiche. Questa operazione è essenzialmente l'inverso della convoluzione standard e viene utilizzata per l'upsampling delle feature map, ripristinando le dimensioni originali.

Il modello utilizza una DiceLoss come funzione di perdita, ottimizzato tramite l'algoritmo Adam con un tasso di apprendimento di 0.001. Sono state testate anche la Binary Cross Entropy e le due funzioni combinate, ma alla fine è stata scelta la Dice Loss per la sua maggiore efficacia nella valutazione delle segmentazioni.

Nella Tabella 2.2 sono descritti alcuni esempi delle prove svolte nella fase di addestramento, con i tre metodi e i relativi parametri.

Esempi delle prove di addestramento svolte				
Numero Prova	Input della rete	Loss	Numero di Epoche	Tipologia di Finestratura
1	Immagini 2D con ROI attorno all'area polmonare	DiceLoss	100	[-300,400]
2	Immagini 2D con ROI attorno all'area polmonare	DiceLoss	100	[-600,400]
3	Immagini 2D con ROI attorno all'area polmonare	DiceLoss	100	NO
4	Immagini 2D con ROI attorno al CM	DiceLoss	20	[-300,400]
5	Immagini 2D con estrazione parenchima polmonare	DiceLoss	100	[-300,400]
6	Immagini 2D con estrazione parenchima polmonare	DiceLoss	100	NO

Tabella 2.2: Esempi delle prove di addestramento svolte con diverse configurazioni di input, funzioni di perdita e tipologie di finestratura.

2.7 Post-processing

Per migliorare la qualità delle segmentazioni automatiche generate dalla rete neurale, è stata implementata una procedura di post-processing sulla ricostruzione 3D delle maschere automatiche. Questo processo è fondamentale per affinare i risultati ed è mirato a ridurre i falsi positivi. L'obiettivo principale di questo post-processing è rimuovere le piccole regioni che non soddisfano un volume minimo specificato. Questa operazione è fondamentale per garantire che le regioni significative siano preservate, per ridurre le misclassificazioni e per migliorare l'accuratezza dell'analisi.

La tecnica di post-processing si basa sul concetto di connettività. In un primo momento, sono state identificate le componenti connesse del volume binario. Successivamente per ciascun volume connesso viene valutata l'estensione e fatta una

selezione.

Considerando quanto segue:

- Connettività = 1 (6-connesso): i voxel sono considerati connessi solo se condividono un lato. Questo significa che i voxel possono connettersi solo attraverso le facce, escludendo i bordi e gli angoli. Questa configurazione è più restrittiva e tende a produrre componenti connesse più isolate.
- Connettività = 2 (26-connesso): i voxel sono considerati connessi se condividono un lato, un bordo o un angolo. In questo contesto, ogni voxel può connettersi a voxel su ciascun piano (XY, XZ, YZ) che condividono una faccia. Questa impostazione consente una connessione più ampia e comprende un totale di 26 voxel potenzialmente connessi, il che facilita l'identificazione di componenti più grandi e coese all'interno del volume.

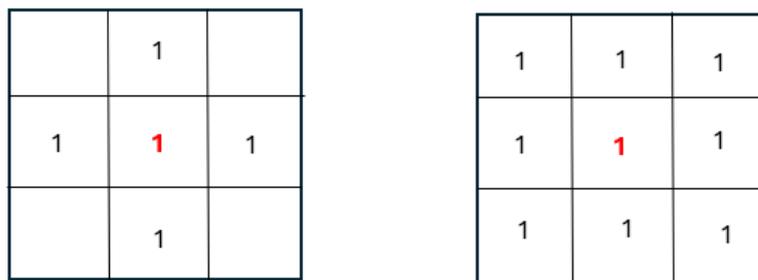


Figura 2.23: Esempio in un caso bidimensionale del parametro di connectivity. Il valore centrale, in rosso, è quello di cui si deve valutare la connessione con i vicini. A sinistra è rappresentata la connessione = 1, in 2D il pixel centrale ha 4 vicini. A destra è rappresentata la connessione = 2, in 2D il pixel centrale ha 8 vicini.

Si è imposto un parametro di connectivity=2.

Durante questo processo, sono stati estratti e stampati vari dettagli relativi a ciascuna regione connessa ed è stato scelto come criterio di eliminazione il volume delle ROI trovate. Analizzando i volumi di tutte le lesioni di riferimento correttamente segmentate dal clinico, quella minore è risultata essere di 566 voxel; per tale motivo e per essere più conservativi, evitando di perdere le lesioni sottosegmentate, è stata imposta una soglia di 300 voxel. Sono state eliminate tutte le ROI che avessero una dimensione inferiore a questo valore, la maggior parte delle quali rappresentavano spot spuri.

2.8 Metriche di valutazione

Per valutare i risultati ottenuti dall'algoritmo, sono stati messi a confronto i volumi delle maschere manuali, il grou truth, con quelli delle maschere automatiche. Questo confronto è basato sul calcolo delle seguenti metriche:

- **PRECISION**: misura la qualità delle previsioni positive fatte dal modello, ossia quanti dei campioni che il modello ha classificato come positivi sono effettivamente corretti. E' rapporto tra il numero di veri positivi (TP) e la somma di tutti i casi segnalati dal modello come positivi (TP + FP):

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

La precision è importante per minimizzare i falsi positivi (FP), cioè quei casi in cui il modello segnala una lesione che in realtà non esiste. Si immagini di usare un modello per identificare tumori in una CT e che il modello indichi una lesione in una zona in cui essa non è presente; in un contesto clinico, se la precision è bassa, il medico potrebbe ordinare esami inutili su regioni sane, aumentando i costi e lo stress per il paziente.

- **RECALL**: valuta la capacità del modello di identificare correttamente tutti i campioni positivi. Essa è una misura della copertura dei risultati positivi effettivi. E' IL rapporto tra il numero di veri positivi (TP) e la somma di tutti i casi realmente positivi (TP + FN):

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

La Recall diventa essenziale per ridurre i falsi negativi (FN), ovvero quei casi in cui il modello non riesce a rilevare una lesione che invece è presente. Un'alta recall significa che il modello è in grado di trovare la maggior parte delle lesioni reali presenti nelle scansioni. In un contesto di screening, per un medico, questo è cruciale: trascurare una lesione potrebbe avere gravi conseguenze, come ritardi nella diagnosi di un tumore maligno.

- **DICE COEFFICIENT**: è una misura di similarità utilizzata spesso per valutare il confronto tra set predetti e set reali. Il coefficiente varia da 0 a 1, con 1 che indica una perfetta somiglianza:

$$Dice\ score = \frac{2|A \cap B|}{|A| + |B|} = \frac{2TP}{(TP + FP) + (TP + FN)}$$

Il Dice score combina precision e recall in quanto considera sia quante lesioni il modello ha trovato (recall) sia quanto queste siano precise (precision). Un

alto Dice score indica che il modello non solo trova un numero significativo di lesioni, ma lo fa anche in modo accurato, con una buona corrispondenza tra la lesione predetta e quella reale.

Nella rilevazione di lesioni su immagini CT, la Precision è utile per evitare falsi allarmi, la Recall garantisce che non vengano trascurate lesioni importanti, e il Dice score dà una visione d'insieme della qualità del modello, bilanciando le due metriche.

Capitolo 3

RISULTATI

3.1 Pre-processing

3.1.1 Risultati del metodo 1

Inizialmente, la maschera del volume polmonare doveva essere utilizzata per isolare ed estrarre dalle slices il solo parenchima polmonare, escludendo tutte le altre strutture non rilevanti, come i tessuti circostanti o gli organi adiacenti. Tuttavia, come evidenziato in Figura 3.1 non è stato possibile ottenere questo risultato a causa dell'inadeguatezza dell'algoritmo automatico di segmentazione polmonare per i pazienti con lesioni localizzate in prossimità della pleura o del mediastino. Il problema principale risiede nel fatto che l'algoritmo utilizzato si basa su una tecnica di sogliatura, che classifica i tessuti in base ai valori di intensità dei pixel. Già nelle prime fasi di segmentazione, questo metodo ha portato a errori significativi nel riconoscimento di tali lesioni. In particolare, le lesioni situate in aree adiacenti alla pleura o al mediastino sono state erroneamente identificate come parte integrante delle strutture anatomiche a cui erano adese, come il torace o le strutture mediastiniche, e sono quindi state escluse dall'area di interesse. Questo ha compromesso l'efficacia dell'estrazione del parenchima polmonare e la corretta identificazione delle lesioni in quelle regioni, rendendo il risultato finale inadatto per un'analisi accurata delle lesioni polmonari vicine a tali strutture. Al contrario, l'algoritmo ha avuto più successo con le lesioni situate all'interno del polmone, in quanto queste erano correttamente incluse nella maschera generata.

Il fatto che le lesioni adiacenti alla pleura e al mediastino vengano escluse rappresenta un significativo limite dell'algoritmo, impedendo una segmentazione completa e precisa del polmone, specialmente in quei pazienti con lesioni in posizioni critiche. Questo limite compromette anche la possibilità di ottenere una visualizzazione ottimale e accurata di tutte le lesioni presenti nel volume polmonare.

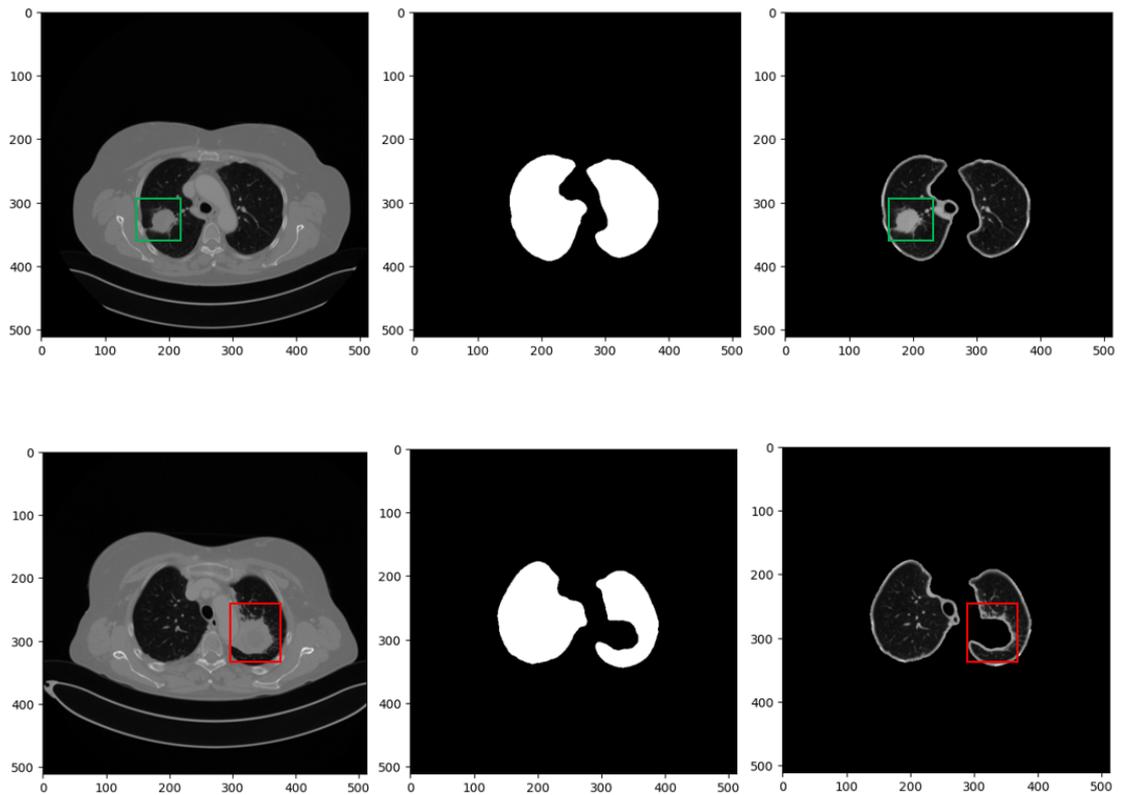


Figura 3.1: Risultati dell'algorithm automatico di segmentazione polmonare: nella parte a sinistra è rappresentata l'iniziale scansione CT di un paziente con lesione all'interno del polmone (in alto) e un paziente con lesione al mediastino (in basso); nella parte centrale si trova la corrispondente maschera polmonare dei due casi; nella parte a destra è rappresentata il risultato della sovrapposizione tra slice e maschera. Nel caso in alto la ROI in verde mostra come la lesione viene compresa nella segmentazione; nel caso in basso la ROI rossa mostra il contrario.

Per questo motivo, si è deciso di utilizzare comunque la maschera polmonare volumetrica per estrarre i contorni dei polmoni da ogni slice, effettuando successivamente un crop sul piano x-y, al fine di ridurre l'area di analisi anche su questo piano. Per ogni paziente, le coordinate delle ROI sono state individuate utilizzando i contorni ottenuti dall'applicazione del Filtro di Sobel.

L'immagine 3.2 mostra una serie di slice CT dello stesso paziente, tutte ritagliate utilizzando una dimensione uniforme. Tuttavia, poiché il crop è stato determinato in base alle coordinate massime e minime che includono i polmoni in tutte le slice, si osserva come nelle diverse immagini vengano visualizzate porzioni differenti del corpo. Nella prime slice, il campo visivo comprende principalmente strutture al di fuori dell'area polmonare, come tessuti molli; mentre nelle slices centrali è più evidente la presenza dei polmoni. Nell'ultima slice, si nota chiaramente la fine dei polmoni e l'inizio di altre strutture anatomiche, come il fegato. Questo effetto è dovuto al fatto che le dimensioni del crop rimangono costanti per lo stesso paziente, anche quando la struttura polmonare non occupa l'intera area visualizzata.

Le immagini in Figura 3.3 mostrano l'applicazione del windowing, che ottimizza il contrasto e migliora la visibilità di specifiche strutture anatomiche, rispetto a quelle in cui non viene applicato. Sono state prese le stesse slice dell'immagine precedente.

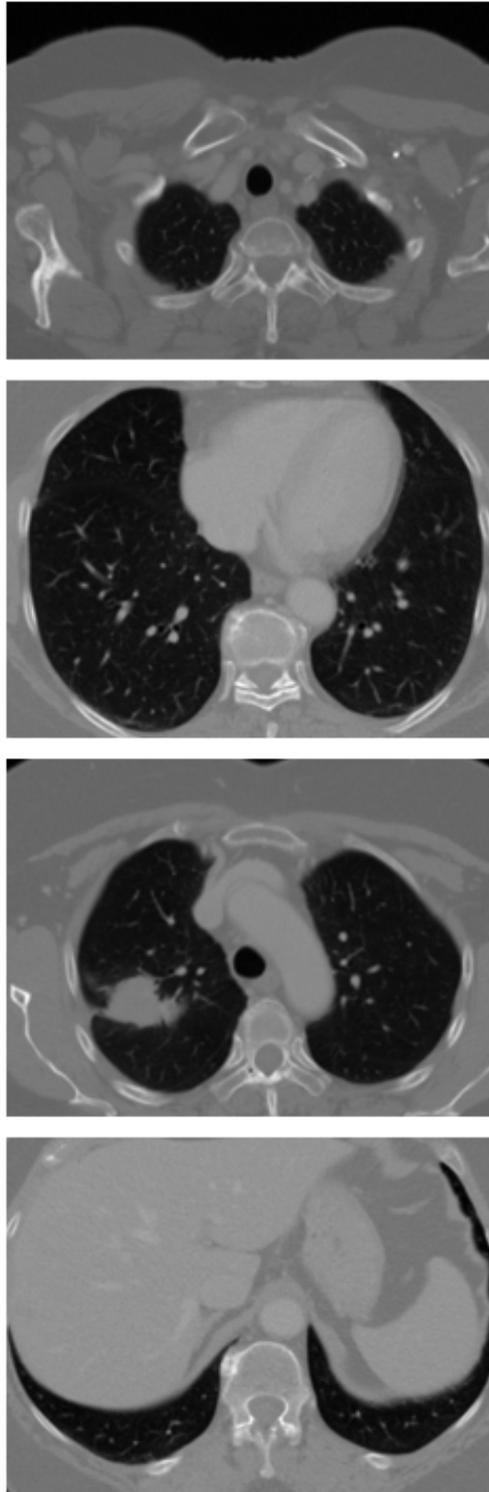


Figura 3.2: Rappresentazione del crop effettuato su diverse slices dello stesso paziente.



Figura 3.3: Rappresentazione del windowing con conseguente padding.

3.1.2 Risultati del metodo 2

La differenza con il Metodo 1 risiede nelle coordinate e nelle dimensioni della bounding box utilizzata per ritagliare attorno al centro di massa della lesione. La Figura 3.4 mostra l'output finale di questo metodo: immagini di dimensione 256x256, con il tumore in posizione centrale. Chiaramente sono mostrate le slices in cui è presente il tumore, ma questo processo è stato applicato a tutte le slice contenenti i polmoni.

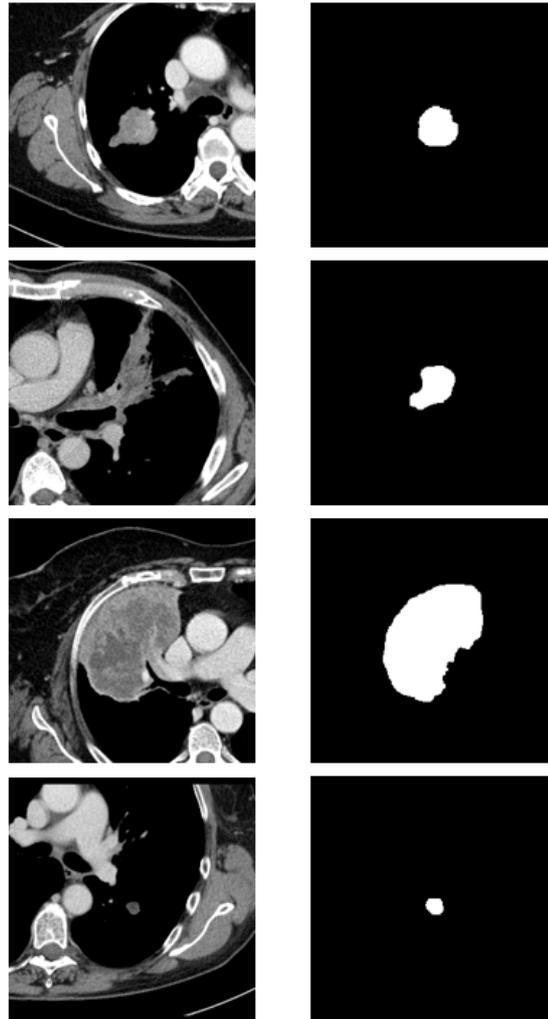


Figura 3.4: Casi rappresentativi di diversi pazienti. A ogni slices è affiancata la rispettiva maschera manuale; in ordine, dall'alto verso il basso, si ha una lesione interna di medie dimensioni, una lesione adiacente al mediastino di medie dimensioni, una lesione adesa alla pleura di grandi dimensioni, una lesione interna di piccole dimensioni.

3.1.3 Risultati del metodo 3

I risultati ottenuti dal Metodo 3 sono rappresentati in Figura 3.5. Si può notare come sia stato isolato esclusivamente il parenchima polmonare, diversamente dagli altri due metodi. Le maschere dei polmoni sono state generate direttamente dal Software Raystation, mentre le immagini CT sono il risultato dell'applicazione delle maschere sulle rispettive scansioni CT.

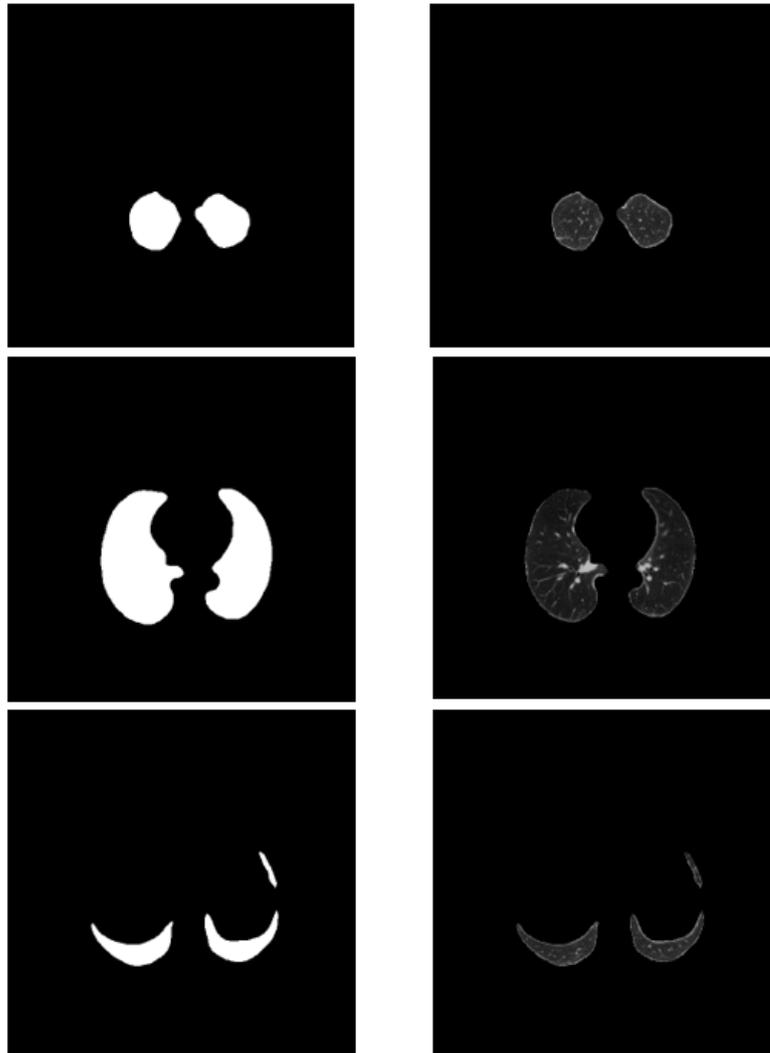


Figura 3.5: A sinistra è rappresentata la maschera polmonare, a destra la corrispondente slice CT dello stesso caso. Le immagini mostrano gli output relativi a diverse parti del polmone, includendo una sezione centrale, una superiore e una inferiore, per evidenziare come la segmentazione si comporti nelle varie regioni polmonari.

Quest'immagine porta alla luce come una CT che include l'intero volume toracico presenta una complessità visiva significativa, poiché mostra una varietà di tessuti e organi con densità e caratteristiche diverse. Questo può ostacolare l'analisi specifica del tessuto polmonare, rendendo più difficile la visualizzazione e l'esaltazione delle lesioni. D'altra parte, un'immagine che isola esclusivamente il parenchima polmonare evidenzia solo il polmone, eliminando distrazioni e semplificando notevolmente l'identificazione di anomalie. Focalizzarsi solo sul parenchima consente una maggiore facilità nell'identificazione della lesione.

3.2 Rete Neurale

Nei paragrafi seguenti, del Metodo 1 e del Metodo 2 vengono riportati i risultati dei Validation-Snapshot di alcune delle prove di addestramento effettuate. I Validation-Snapshot sono campioni di valutazione utilizzati durante il processo di addestramento per monitorare le prestazioni del modello su dati non visti, ossia sui dati provenienti dal set di validazione. Questi giocano un ruolo cruciale nell'assicurare che il modello stia apprendendo in modo efficace, poiché forniscono un feedback immediato sulle sue capacità predittive in condizioni che non ha incontrato in fase di addestramento. Questo approccio aiuta a garantire che il modello possa considerarsi generalizzabile, aumentando così la sua affidabilità quando sarà applicato al set di test, dove le performance verranno valutate in modo definitivo.

Del Metodo 3 è stata svolta la fase di inferenza e di post-processing, e i risultati proposti si riferiscono a queste.

3.2.1 Risultati del metodo 1

Si prende in considerazione il metodo 1, ovvero la scelta di una ROI attorno al polmone, e gli esempi delle sue prove 'PROVA 1', 'PROVA 2', 'PROVA 3', annotate in 2.2, rispettivamente con finestra di $[-300,400]$, $[-600,400]$ e senza alcuna finestra. Nelle Figure 3.6, 3.7 e 3.8 sono stati riportati degli esempi di riferimento di risultati ottenuti dalle 3 prove sopra menzionate. Tutte le 3 figure mostrano un confronto tra 2 pazienti, selezionando per ognuno di essi la stessa slice: il primo con lesione di grandi dimensioni in posizione adiacente alla pleura e il secondo con lesione interna al polmone e di dimensioni estremamente piccole.

Si nota come il windowing, Figura 3.6 e Figura 3.7, possa aiutare a riconoscere e segmentare i tumori dei pazienti che hanno una lesione medio-grande. Questa efficacia può essere attribuita al fatto che il tumore, indipendentemente dalla posizione (pleura, mediastino, interno al polmone) possiede dei valori di pixel

distintivi rispetto il resto del torace e di ciò che è presente nell'immagine, come ad esempio cuore e midollo. In particolare, i tumori tendono a presentare una densità e un'intensità di pixel che si discostano nettamente dai tessuti circostanti. Queste differenze possono manifestarsi in variazioni di contrasto, rendendo i tumori più facilmente identificabili nelle scansioni.

Al contrario, l'applicazione del windowing nei pazienti con lesioni di dimensioni minori si rivela insufficiente affrontare le problematiche legate alla piccola dimensione delle lesioni. Si evidenzia innanzitutto che non solo la lesione piccola non viene identificata in modo definito, ma si registrano numerosi falsi positivi, ovvero regioni che vengono identificate dalla rete ma che non corrispondono a tumori. Questo potrebbe derivare dal fatto che la rete è stata principalmente addestrata con immagini di pazienti affetti da lesioni medio-grandi, il che potrebbe compromettere l'apprendimento delle caratteristiche associate a lesioni più piccole. Un'altra possibile spiegazione è che il windowing non fornisce un supporto adeguato per i tumori di dimensioni ridotte, in quanto tende a eliminare troppe informazioni utili. In particolare, restringendo i valori di pixel a un intervallo troppo limitato, si possono perdere dettagli fondamentali necessari per valutare correttamente l'espansione reale del tumore. Di conseguenza, l'uso del windowing in questi casi potrebbe ridurre la visibilità della lesione e dei suoi contorni, allontanandoli dalla maschera manuale elaborata dal radiologo. Questo introduce una maggiore complessità nell'addestramento della rete, ostacolando ulteriormente la sua capacità di riconoscere accuratamente le lesioni di dimensioni minori.

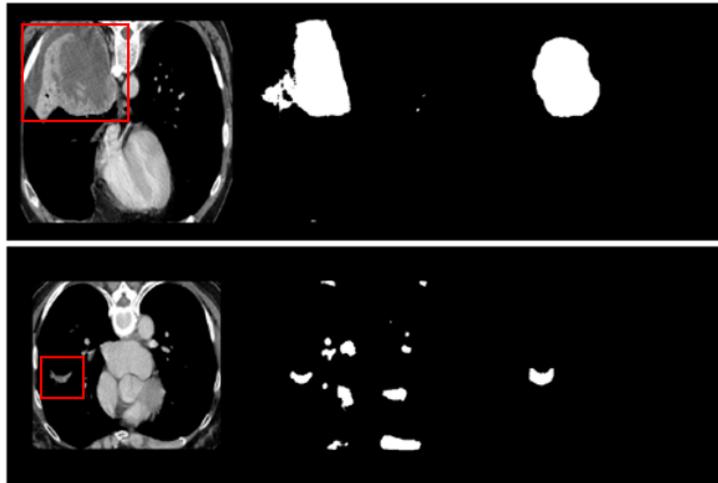


Figura 3.6: Validation-Snapshot della prova 1. In alto è presente un caso di paziente con lesione di grandi dimensioni in posizione adiacente alla pleura sinistra (evidenziato dal quadrato rosso); in basso è presente un paziente con lesione di piccole dimensioni con lesioni interna al polmone (evidenziato dal quadrato rosso). L'immagine CT a sinistra è quella di input alla rete, l'immagine binaria a destra è la maschera manuale, l'immagine bianca centrale è l'output della rete.

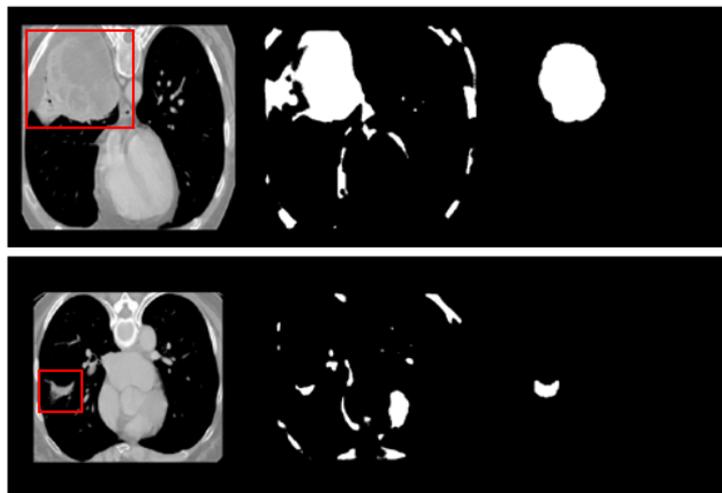


Figura 3.7: Validation-Snapshot della prova 2. In alto è presente un caso di paziente con lesione di grandi dimensioni in posizione adiacente alla pleura sinistra (evidenziato dal quadrato rosso); in basso è presente un paziente con lesione di piccole dimensioni con lesioni interna al polmone (evidenziato dal quadrato rosso). L'immagine CT a sinistra è quella di input alla rete, l'immagine binaria a destra è la maschera manuale, l'immagine bianca centrale è l'output della rete.

A conferma di tale tesi, nella Figura 3.8, in cui non è presente il windowing, si osserva come i tumori piccoli e interni al polmone vengano maggiormente riconosciuti. Per contro, le lesioni medio-grandi e soprattutto adiacenti a pleura e mediastino, senza windowing non vengono identificate. Una spiegazione valida potrebbe essere che la rete riconosca tali lesioni come facenti parte della struttura anatomica con cui sono a contatto in quanto i pixel hanno simile intensità.

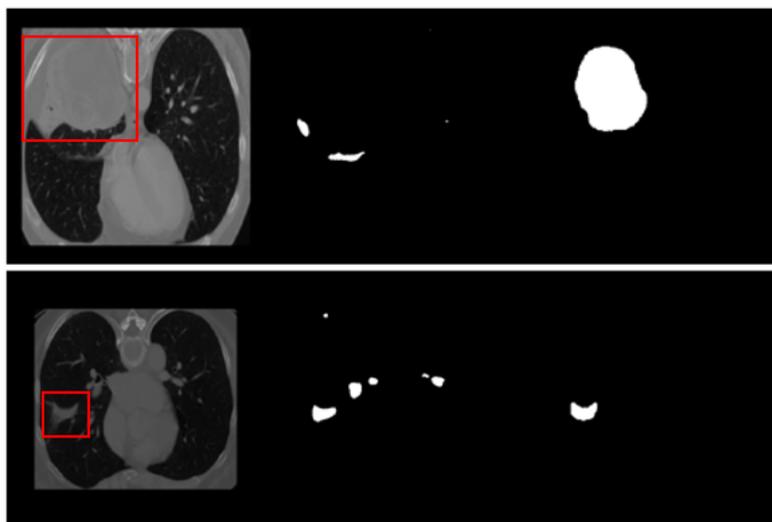


Figura 3.8: Validation-Snapshot della prova 3. In alto è presente un caso di paziente con lesione di grandi dimensioni in posizione adiacente alla pleura sinistra (evidenziato dal quadrato rosso); in basso è presente un paziente con lesione di piccole dimensioni con lesione interna al polmone (evidenziato dal quadrato rosso). L'immagine CT a sinistra è quella di input alla rete, l'immagine binaria a destra è la maschera manuale, l'immagine bianca centrale è l'output della rete.

3.2.2 Risultati del metodo 2

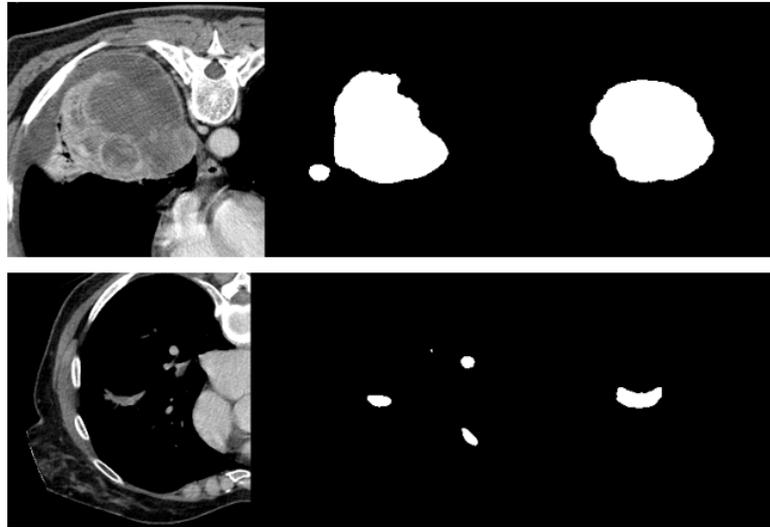


Figura 3.9: Validation-Snapshot della prova 4. In alto è presente un caso di paziente con lesione di grandi dimensioni in posizione adiacente alla pleura sinistra; in basso è presente un paziente con lesione di piccole dimensioni con lesioni interna al polmone. L'immagine CT a sinistra è quella di input alla rete, l'immagine binaria a destra è la maschera manuale, l'immagine bianca centrale è l'output della rete.

3.2.3 Risultati del metodo 3

Prendendo in considerazione il metodo 3, ovvero immagini in cui viene visualizzato solo il parenchima polmonare, e le sue prove 'PROVA 5' e 'PROVA 6', rispettivamente con windowing $[-300,400]$ in Figura 3.10 e senza windowing in Figura 3.11, non si riscontra una differenza significativa tra gli output delle due configurazioni. In entrambe le prove, le lesioni di dimensioni medio-grandi, così come quelle di dimensioni più ridotte, vengono identificate correttamente. Questo risultato potrebbe essere attribuito al fatto che le immagini fornite in input alla rete sono già sufficientemente pulite e prive di disturbi, semplificando così il processo di apprendimento del modello. Nonostante ciò, è possibile osservare la presenza di outliers in entrambi i casi, anche se la loro quantità risulta significativamente inferiore rispetto a quanto osservato con il metodo 1. Questo indica che, sebbene la pulizia delle immagini abbia migliorato il riconoscimento complessivo, alcuni elementi indesiderati continuano a essere rilevati, ma in misura decisamente più contenuta.

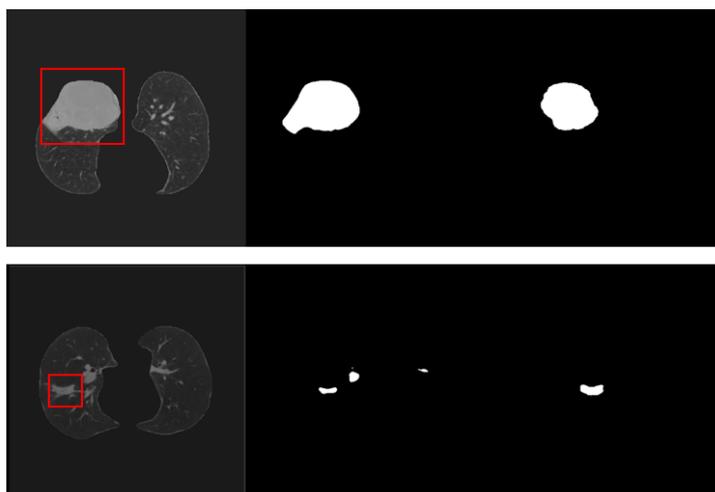


Figura 3.10: Validation-Snapshot della prova 5. In alto è presente un caso di paziente con lesione di grandi dimensioni in posizione adiacente alla pleura sinistra (evidenziato dal quadrato rosso); in basso è presente un paziente con lesione di piccole dimensioni con lesioni interna al polmone (evidenziato dal quadrato rosso). L'immagine CT a sinistra è quella di input alla rete, l'immagine binaria a destra è la maschera manuale, l'immagine bianca centrale è l'output della rete.



Figura 3.11: Validation-Snapshot della prova 6. In alto è presente un caso di paziente con lesione di grandi dimensioni in posizione adiacente alla pleura sinistra (evidenziato dal quadrato rosso); in basso è presente un paziente con lesione di piccole dimensioni con lesioni interna al polmone (evidenziato dal quadrato rosso). L'immagine CT a sinistra è quella di input alla rete, l'immagine binaria a destra è la maschera manuale, l'immagine bianca centrale è l'output della rete.

Inferenza

Nella Tabella 3.1 sono rappresentate le metriche calcolate sul data set di training, validation e testing, con il modello della PROVA 6. Si evidenzia come i risultati di ogni metrica, ovvero DSC, Recall e Precision, siano simili tra i tre dataset, suggerendo una buona generalizzazione del modello. Questo indica l'assenza di overfitting, una condizione in cui il modello performa molto bene sul set di training, ma non riesce a generalizzare su dati non visti, come quelli di validation e testing. Il fatto che le prestazioni siano omogenee tra i tre set suggerisce che il modello non si sia adattato eccessivamente ai dati di training, ma piuttosto ha appreso delle caratteristiche significative che possono essere applicate anche su nuovi dati. Perciò, il modello sembra avere un buon equilibrio tra bias e varianza, ed è in grado di catturare le regole sottostanti senza sovradimensionare le specifiche peculiarità del training set. Questo è un aspetto molto positivo in fase di sviluppo, in quanto un modello privo di overfitting è più robusto e affidabile nel momento in cui viene applicato su dati reali. L'assenza di overfitting, dunque, garantisce una migliore capacità predittiva e una maggiore possibilità di trasferire il modello a contesti applicativi diversi da quelli usati per l'addestramento.

	DSC	RECALL	PRECISION
TRAINING	64%	90%	54%
VALIDATION	60%	88%	49%
TESTING	63%	91%	53%

Tabella 3.1: Metriche su training, validation e testing utilizzando il modello della PROVA 6.

In Figura 3.12 sono rappresentati, tramite box plot, gli andamenti dei valori delle metriche calcolate sul dataset di testing. Questi mostrano un primo confronto tra la maschera manuale e l'output della rete.

I valori del DSC si concentrano principalmente tra 0,60 e 0,75 che rappresentano l'intervallo interquartile. Questo evidenzia come la maggior parte delle previsioni del modello hanno una buona corrispondenza con il ground truth. La mediana, di 0,65, suggerisce che il 50% delle previsioni del modello ha un DSC superiore o uguale a questo valore. Questo indica una performance ragionevole ma non eccellente. I baffi del box plot mostrano che vi sono risultati peggiori (0,45) e alcuni estremamente buoni (0,85). Media e mediana sono simili, per cui i dati non sono asimmetrici. Questo è positivo in quanto la distribuzione non è sbilanciata in modo estremo.

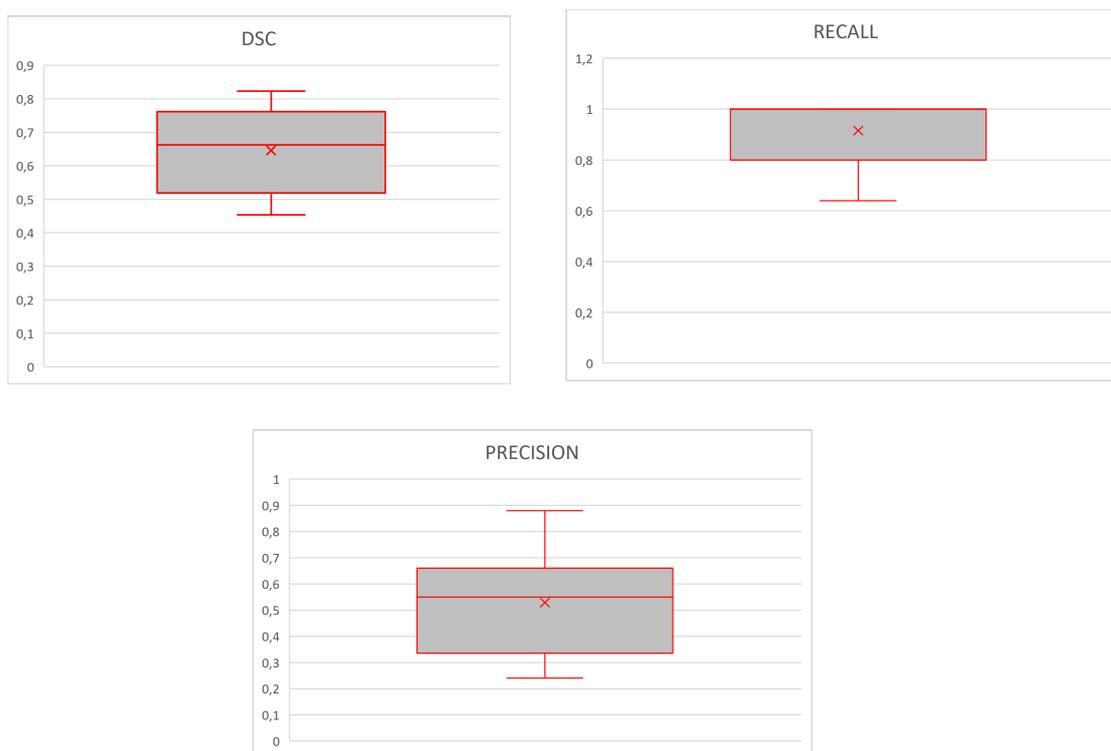


Figura 3.12: Box Plot di DSC, RECALL e PRECISION dei dati di testing.

Nel risultato della RECALL, si può notare che il box plot si estende fino a 1. Ciò indica che una parte corposa delle predizioni raggiunge un'alta recall, riuscendo a identificare correttamente tutti i positivi. Questo è sostenuto anche dal valore di mediana, pari a 0,95. Si arriva anche al valore di 0,65, ma considerando l'andamento complessivo si può affermare che non sia estremamente basso.

Il valore di PRECISION si estende fino a quasi 0,9, indicando che in alcuni casi il modello ottiene un'ottima precisione, individuando più veri positivi rispetto i falsi positivi. Media e mediana suggeriscono una leggera asimmetria verso valori più alti. La mediana vicina a 0,55 indica che il modello non è sempre molto preciso e in più della metà dei casi sono presenti in percentuale maggiore falsi positivi rispetto ai veri positivi.

A livello visivo, in Figura 3.13 sono rappresentate le sovrapposizioni tra immagine di input alla rete, maschera manuale e maschera in output della rete. In Figura 3.14, invece, c'è la sovrapposizione tra slice CT originale, maschera ottenuta dal clinico e maschera predetta dalla rete. Diversamente dall'immagine precedente, in cui sono mostrate le aree delle segmentazioni della lesione, qui se ne evidenziano i contorni. Si evince come, indipendentemente dalla posizione della lesione tumorale

(interna al polmone, adiacente a pleura o mediastino), il riconoscimento della massa avviene correttamente. Questo è a supporto dei valori di Recall alti. Ciò si traduce nel fatto che la maggior parte dei veri positivi sono stati identificati dall'algoritmo in modo efficiente.

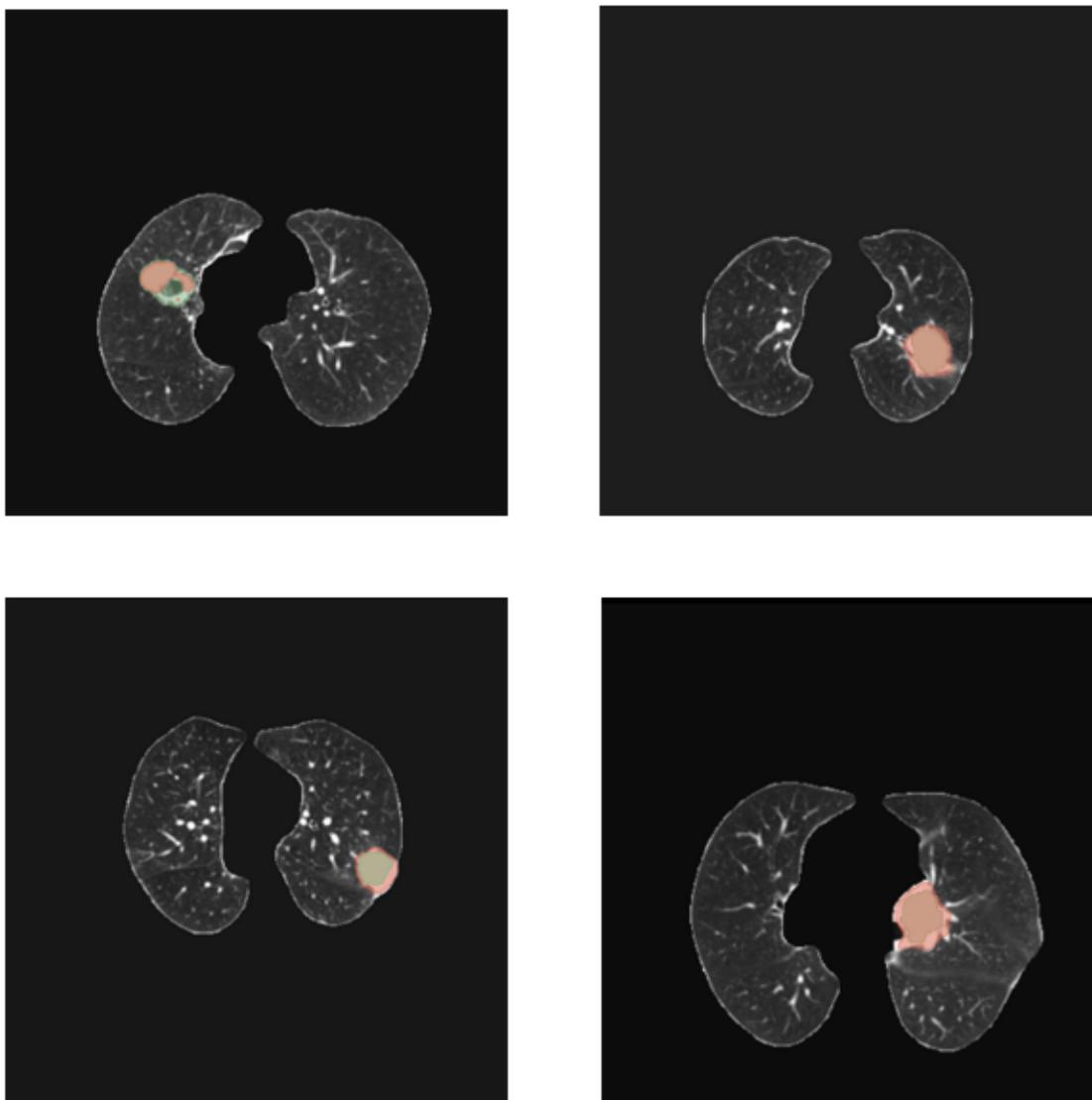


Figura 3.13: Esempi di pazienti del dataset di testing. Sovrapposizione tra slice di input alla rete, maschere create dal clinico e maschera predetta dalla rete. ROI in verde rappresenta il ground truth, ROI in rosso rappresenta la maschera automatica in output dalla rete. In alto sono rappresentati due pazienti con lesione tumorale interna al polmone, in basso a sinistra un paziente con lesione adiacente alla pleura, in basso a destra un paziente con lesione adiacente al mediastino.

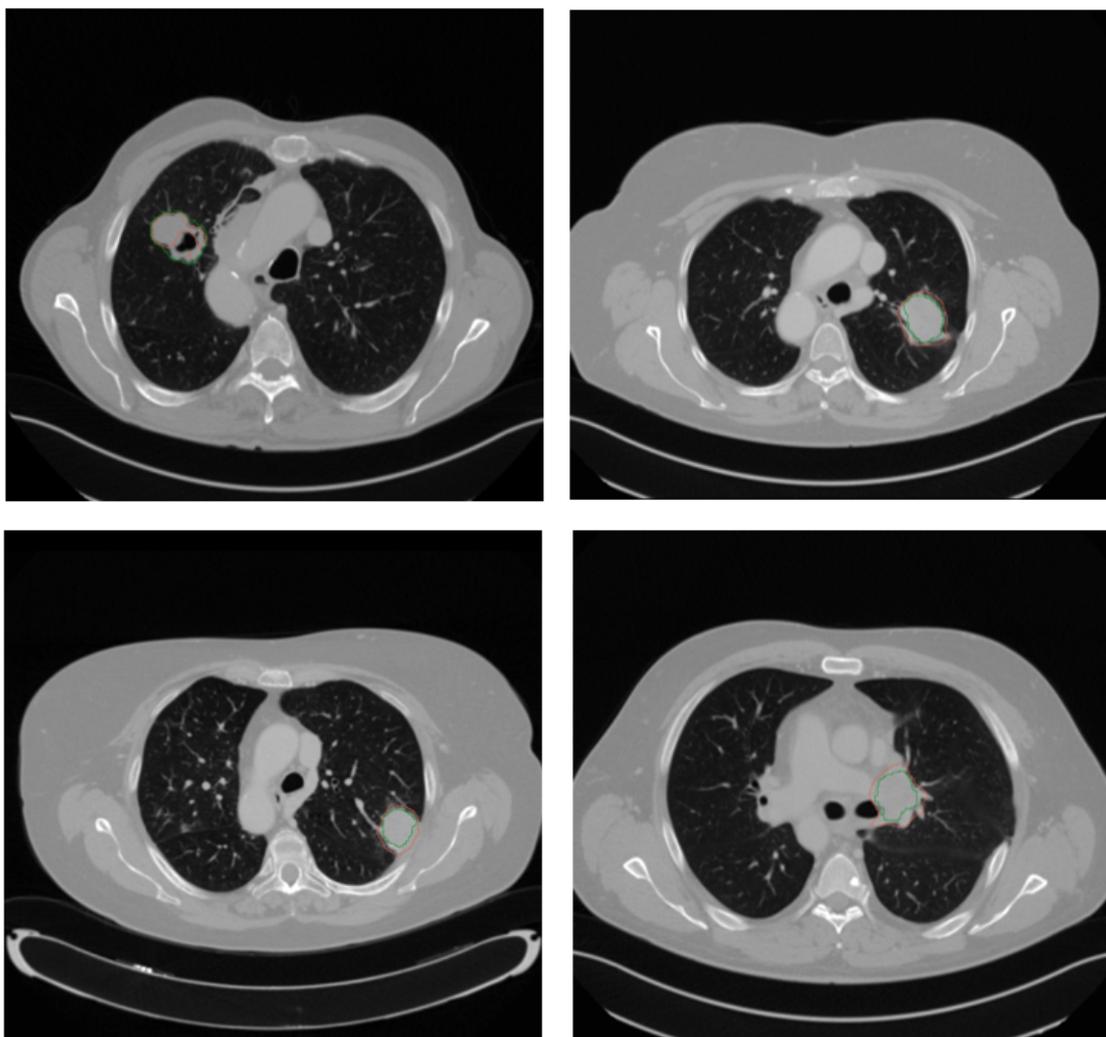


Figura 3.14: Esempi di pazienti del dataset di testing. Sovrapposizione tra CT originale, maschera creata dal clinico e maschera predetta dalla rete. Il contorno in verde rappresenta il ground truth, il contorno in rosso rappresenta la maschera automatica in output dalla rete. In alto sono rappresentati due pazienti con lesione tumorale interna al polmone, in basso a sinistra un paziente con lesione adiacente alla pleura, in basso a destra un paziente con lesione adiacente al mediastino.

In Figura 3.15, invece, si nota come l'output della rete neurale riconosca come tumorali anche regioni del polmone, quali bronchi e bronchioli. Questo è in accordo con una bassa percentuale di Precision ottenuta. In seguito a queste valutazioni, è stato necessario effettuare un'analisi di post-processing che permetta di migliorare le prestazioni.

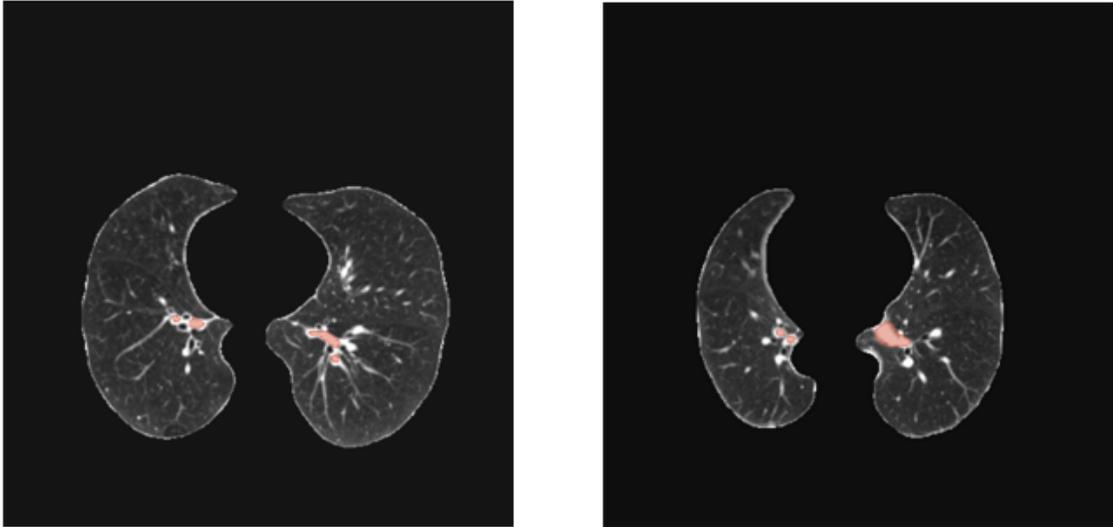


Figura 3.15: Esempi di pazienti del dataset di testing; ROI in rosso rappresenta la maschera automatica.

Post-processing

Il post-processing ha permesso di affinare le prestazioni del modello, riducendo, in particolare, il numero di falsi positivi, ovvero le lesioni erroneamente detectate, senza compromettere la capacità del modello di identificare correttamente i veri positivi, ovvero le lesioni annotate. Questi risultati, rappresentati in Figura 3.16 e nella Tabella 3.2, suggeriscono che l'applicazione di tecniche di post-processing è stata efficace per ottimizzare le prestazioni complessive del modello.

	DSC	RECALL	PRECISION
TESTING	68%	91%	58%

Tabella 3.2: Metriche su testing sviluppando il post-processing.

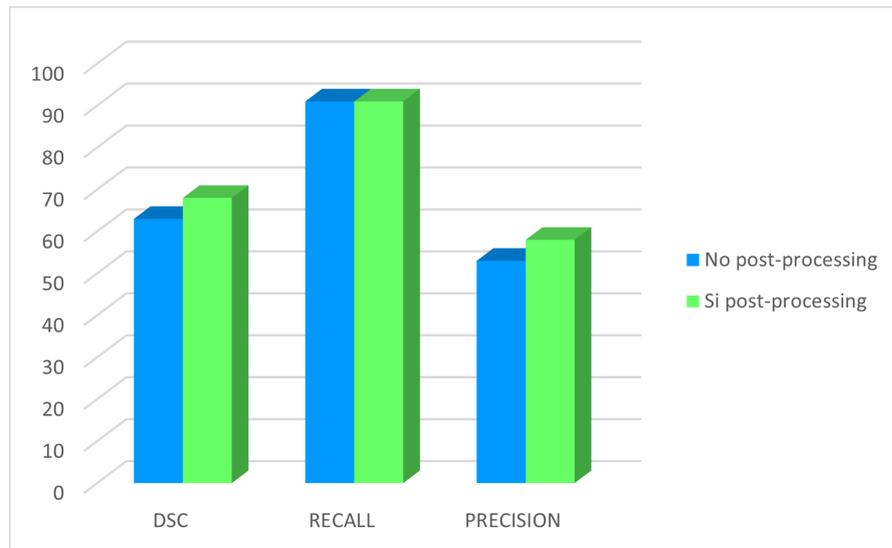


Figura 3.16: Grafico a barre di confronto tra le metriche calcolate con l'output della rete e quelle calcolate in output dal post-processing.

Il miglioramento più evidente si registra nella Precision, che è passata dal 53% al 58.57% per il test set. Si nota come l'eliminazione di tutte le ROI di grandezza inferiore alla soglia impostata, non impattano nel valore di Recall. Questo significa che per questo set di dati, il post processing ha eliminato solo le aree non rilevanti senza compromettere l'identificazione corretta dei veri positivi, preservando quindi l'identificazione della lesione annotata. Il post-processing è stato efficace nel rimuovere le false identificazioni senza influire negativamente sulle vere lesioni. Il miglioramento della Precision senza compromessi sulla Recall indica un bilanciamento efficace tra la riduzione dei falsi positivi e il mantenimento di un'alta capacità di identificazione delle lesioni.

Capitolo 4

DISCUSSIONE

Il tumore al polmone è una delle principali cause di morte per cancro a livello globale, con una prevalenza particolarmente elevata nei paesi industrializzati. Il carcinoma polmonare non a piccole cellule rappresenta circa l'80-90% di tutte le neoplasie polmonari maligne, configurandosi come la forma più comune di cancro al polmone e la principale causa di morte oncologica sia negli uomini che nelle donne. Il tumore al polmone non a piccole cellule si distingue per il suo decorso clinico più lento, ma spesso viene diagnosticato in stadio avanzato o metastatico a causa della mancanza di sintomi specifici nelle fasi iniziali.

Proprio perché la diagnosi è tardiva, è fondamentale puntare sul trattamento e sulle cure per aumentare le percentuali di sopravvivenza. A questo proposito, la segmentazione del tumore è un argomento di grande interesse nell'analisi delle immagini mediche, in quanto fornisce ai medici informazioni quantitative significative e affidabili, e una migliore comprensione della fisiopatologia del cancro, per il monitoraggio della malattia, per la pianificazione di interventi terapeutici mirati e per la previsione dell'esito del paziente. Attraverso la segmentazione delle immagini ottenute con esami come la tomografia computerizzata, è possibile isolare le aree tumorali dai tessuti circostanti, permettendo una valutazione delle dimensioni, della forma e della localizzazione del tumore.

La segmentazione manuale del tumore, però, pur essendo una pratica comune, presenta diversi svantaggi: è un processo lungo, laborioso e fortemente soggetto a variabilità tra operatori, il che può compromettere l'accuratezza e la riproducibilità dei risultati. Questi fattori rendono il procedimento meno affidabile e poco efficiente in un contesto clinico, dove la rapidità e la precisione sono essenziali. In questo scenario, l'adozione di tecniche di segmentazione automatica o semi-automatica rappresenta un notevole progresso. Questi metodi riducono significativamente il tempo necessario per l'analisi, eliminando la componente ripetitiva e soggettiva della

segmentazione manuale. Di conseguenza, si accelera l'intero flusso di lavoro clinico, consentendo di ottenere informazioni sul tumore in tempi ridotti. Tuttavia, la segmentazione automatica dei tumori polmonari è impegnativa a causa dell'elevata varietà nell'aspetto, nella forma, nella posizione e nella dimensione dei tumori. Questo rende difficile l'elaborazione di regole di segmentazione efficaci e univoche.

Qui entra in gioco il Deep Learning, che rappresenta un avanzamento cruciale per la segmentazione automatica delle immagini mediche. A differenza dei metodi tradizionali, il Deep Learning si avvale di reti neurali profonde capaci di apprendere modelli complessi direttamente dai dati, senza la necessità di definire manualmente criteri rigidi. Può analizzare non solo l'intensità dei pixel, ma anche caratteristiche come la forma, la texture e il contesto anatomico, migliorando l'accuratezza della segmentazione. Grazie alla sua capacità di apprendere dai grandi volumi di dati disponibili, riesce a generalizzare meglio rispetto ai metodi tradizionali. Tuttavia, l'adozione del Deep Learning presenta anche alcune sfide. Richiede una grande quantità di dati per l'addestramento e risorse computazionali significative. Inoltre, può essere percepito come una "scatola nera", in quanto non sempre è chiaro come vengano prese le decisioni all'interno delle reti neurali. Nonostante questi svantaggi, i benefici che offre in termini di affidabilità lo rendono una soluzione sempre più indispensabile per la segmentazione delle lesioni tumorali.

Alla luce di queste premesse, l'obiettivo di questo progetto è identificare e segmentare automaticamente le lesioni tumorali polmonari per ridurre significativamente i tempi di segmentazione manuale, un processo lungo e soggetto a variabilità tra operatori. L'obiettivo finale è ottenere una segmentazione automatica che velocizzi l'intero processo clinico, mantenendo una qualità elevata.

Nel corso dello studio sono state utilizzate scansioni toraciche CT per la segmentazione delle lesioni tumorali polmonari di stadio III. Questo stadio è caratterizzato da un avanzato grado di diffusione della malattia, in cui il tumore può aver invaso strutture adiacenti. I tumori di stadio III presentano una variabilità morfologica significativa, con contorni irregolari e dimensioni variabili, il che può complicare ulteriormente l'accuratezza della segmentazione. La presenza di aree di necrosi o infiltrazione nei tessuti circostanti rende queste lesioni ancora più difficili da identificare e delineare. In questo contesto, una segmentazione precisa è cruciale, poiché informazioni dettagliate sulle dimensioni e sulla forma delle lesioni possono influenzare l'analisi quantitativa dei dati, migliorando la valutazione delle caratteristiche morfologiche e strutturali.

Per ottimizzare la segmentazione, sono stati sviluppati tre metodi distinti, ciascuno caratterizzato da un diverso campo visivo (FOV) per valutare l'impatto di diverse aree di interesse sull'accuratezza della rete neurale. I tre approcci includevano: una ROI attorno all'intera area polmonare, una ROI centrata sul centro di massa

della lesione e infine una ROI limitata esclusivamente al parenchima polmonare. L'analisi delle diverse configurazioni di FOV ha rivelato un impatto significativo sulle prestazioni della rete neurale U-Net 2D utilizzata per la segmentazione. In particolare, è emerso che un pre-processing dell'immagine che consenta una delimitazione dei polmoni e, quindi, dell'area di addestramento della rete, abbia aiutato in modo rilevante a migliorare l'identificazione della lesione di interesse.

Analizzando i risultati, è emerso chiaramente che il primo approccio, che utilizzava una ROI attorno all'intera area polmonare, presentava prestazioni nettamente inferiori rispetto agli altri due. Nello specifico, il modello identifica un elevato numero di falsi positivi al di fuori delle aree polmonari, perciò identificava erroneamente strutture anatomiche non pertinenti come possibili lesioni, confondendo ad esempio tessuti mediastinici o muscolari per anomalie polmonari. Questo errore comprometteva significativamente l'accuratezza della segmentazione, rendendo difficile isolare le vere lesioni. Inoltre, il modello ha mostrato anche una tendenza preoccupante a produrre falsi negativi, ossia a non rilevare correttamente le lesioni reali, in alcuni casi arrivando persino a ignorare completamente l'area di interesse. In questi casi estremi, la lesione di interesse veniva completamente cancellata dal processo di segmentazione, vanificando l'intero obiettivo e riducendo drasticamente l'efficacia del metodo. Questo scenario ha evidenziato i limiti significativi dell'approccio con una ROI troppo estesa, che, anziché migliorare la capacità di individuazione delle lesioni, portava a un aumento degli errori sia in termini di identificazioni errate che di mancata rilevazione delle anomalie reali. La complessità e la variabilità intrinseche di queste immagini ampie, con inclusione di tessuti sani e aree non tumorali, rendevano difficile per la rete neurale identificare e segmentare le lesioni tumorali, concludendo che questo metodo non fosse idoneo per l'obiettivo.

Il secondo approccio, che si concentrava sulla ROI centrata attorno al centro di massa della lesione, ha mostrato prestazioni accettabili, ma la sua efficacia era limitata dalla sua natura semplificata. Sebbene la rete fosse in grado di identificare la lesione in modo più diretto, non riusciva a cogliere appieno le complessità morfologiche del tumore e il contesto circostante. Questa mancanza di complessità nel FOV non permetteva al modello di apprendere e adattarsi alle variazioni più sottili, risultando in una segmentazione che, sebbene corretta in alcuni casi, era meno robusta rispetto alle aspettative.

Al contrario, il terzo approccio, che limitava l'analisi esclusivamente al parenchima polmonare, ha fornito risultati notevolmente migliori. Questo FOV ridotto ha effettivamente minimizzato il rumore di fondo, consentendo alla rete di focalizzarsi sulle strutture polmonari rilevanti. Tuttavia, pur riducendo la complessità generale delle immagini, l'approccio non era focalizzato unicamente sulla lesione, ma su tutto il tessuto polmonare. Pertanto, quella che può sembrare una semplificazione,

riferita alla restrizione dell'analisi al solo parenchima polmonare, eliminando altre strutture anatomiche potenzialmente distraenti, in realtà non è così significativa. Infatti, anche se il focus è sulla porzione polmonare, l'intero polmone è comunque incluso nell'analisi, il che implica che la rete deve gestire in ogni caso una variabilità all'interno di quest'area. Sebbene più efficace nell'identificare e segmentare le lesioni con maggiore precisione rispetto agli approcci precedenti, restava ancora un'analisi che coinvolgeva l'intero polmone, mantenendo un livello di complessità sufficiente a garantire robustezza nei risultati e una riduzione dei falsi positivi e negativi.

Questa ricerca ha dimostrato l'importanza della scelta del campo visivo nel migliorare le prestazioni delle reti neurali nella segmentazione delle immagini mediche. L'ottimizzazione del FOV non influisce solo sulla qualità della segmentazione, ma ha anche implicazioni pratiche per l'efficienza del flusso di lavoro clinico. Configurazioni più semplici, come quelle adottate nel terzo approccio, possono ridurre significativamente il tempo necessario per il training della rete e il successivo controllo da parte dei clinici.

Nel corso del progetto, si è utilizzato un approccio che ha previsto l'analisi di sezioni CT sia contenenti tumori che prive di anomalie, evidenziando così una differenza cruciale rispetto a uno studio che si concentra esclusivamente sulla segmentazione. L'inclusione di slice con e senza tumore consente di affrontare non solo il compito di segmentazione, ma anche quello di detection, rendendo lo studio più complesso e completo. Questo approccio consente di addestrare la rete neurale a riconoscere e distinguere tra le aree tumorali e quelle sane, affinando la sua capacità di rilevare anomalie in contesti variabili. Utilizzando immagini che rappresentano sia tumori che tessuti polmonari normali, il modello sviluppato non solo impara a segmentare correttamente le lesioni, ma acquisisce anche competenze nella rilevazione di regioni di interesse.

Tuttavia, per garantire l'efficacia del modello, è stata cruciale la fase di bilanciamento del dataset. È importante evitare situazioni in cui il dataset fosse composto dall'80% di immagini prive di lesioni e solo dal 20% di immagini contenenti anomalie, poiché un tale squilibrio potrebbe compromettere significativamente il processo di addestramento. Infatti, una distribuzione così asimmetrica porterebbe il modello a imparare in modo predominante le caratteristiche delle immagini normali, riducendo la sua capacità di generalizzare e riconoscere le lesioni tumorali. Questo fenomeno, noto come bias del dataset, si tradurrebbe in una scarsa sensibilità nella rilevazione delle anomalie e in un aumento dei falsi negativi, poiché il modello potrebbe ignorare o sottovalutare la presenza di tumori a causa della sua predominante esposizione a tessuti sani.

Nel corso dello studio, si è trovati a valutare se adottare o meno la tecnica del windowing, dato il suo potenziale impatto sulla qualità dell'analisi. Il windowing è

una tecnica frequentemente utilizzata nell'analisi delle immagini CT per migliorare la visibilità delle strutture anatomiche e delle lesioni. Questa metodologia consente di ottimizzare il contrasto e la luminosità delle immagini, adattando i valori di densità per mettere in evidenza particolari aree di interesse. Tuttavia, come evidenziato in questo studio, il windowing presenta sia vantaggi che svantaggi, specialmente in contesti complessi come la segmentazione delle lesioni tumorali.

Da un lato, uno dei principali vantaggi del windowing è la sua capacità di migliorare la visualizzazione di specifici tessuti o anomalie. In genere, questo approccio permette ai clinici e ai radiologi di esaminare dettagli che potrebbero rimanere nascosti in un'immagine non elaborata, rendendo più facile identificare lesioni e altre strutture critiche. In situazioni in cui le lesioni presentano contrasto elevato rispetto ai tessuti circostanti, il windowing può essere uno strumento utile per facilitare la diagnosi.

Tuttavia, nel contesto della segmentazione delle lesioni tumorali polmonari in questo studio, il windowing ha dimostrato di avere effetti negativi significativi. In particolare, la riduzione della gamma di densità delle immagini ha portato a una visibilità limitata del tumore stesso, rendendo difficile per il modello riconoscere e segmentare correttamente le lesioni. La scelta di impostazioni di windowing non ottimali ha causato una compressione delle informazioni visive, il che ha compromesso la capacità della rete neurale di apprendere le caratteristiche distintive dei tumori. Questo problema si è tradotto nell'utilizzo di maschere manuali che non erano identificative delle lesioni, poiché il windowing aveva ridotto il contrasto tra il tumore e i tessuti circostanti. In altre parole, i tumori, che avrebbero dovuto apparire come aree ben definite, risultavano quasi indistinguibili all'interno dell'immagine elaborata. Tale mancanza di definizione ha portato a una variabilità nelle maschere generate manualmente, ostacolando ulteriormente l'efficacia della segmentazione automatica. Questo problema risultava particolarmente critico per i tumori di piccole dimensioni, che sono già di per sé difficili da rilevare. In queste circostanze, i tumori più piccoli tendevano a diventare quasi invisibili all'interno dell'immagine elaborata, riducendo drasticamente la possibilità di una segmentazione accurata.

In sintesi, mentre il windowing offre indubbi vantaggi nella visualizzazione delle immagini, in questo caso ha presentato criticità tali da compromettere la qualità delle maschere di segmentazione, evidenziando l'importanza di una scelta ponderata delle tecniche di elaborazione delle immagini in relazione agli obiettivi specifici dello studio.

Le predizioni della rete neurale sono state sottoposte a un processo di post-elaborazione per ridurre le misclassificazioni. La strategia adottata si basa sull'analisi della connettività dei volumi identificati, un approccio che esamina le regioni contigue per separare i diversi volumi identificati e poterli valutare singolarmente,

al fine di poter differenziare i volumi tumorali dalle altre strutture identificate. Nella pratica, in un primo momento, vengono separate le regioni che presentano una connessione spaziale coerente e plausibile; successivamente, si applica un filtro sulle dimensioni, escludendo le aree che risultano disconnesse o di dimensioni non significative. Tale analisi risulta particolarmente efficace per ridurre falsi positivi, preservando solo i volumi ritenuti clinicamente rilevanti. Questo step permette di eliminare le regioni erroneamente classificate e focalizzare l'attenzione sui volumi che rappresentano potenziali lesioni. Il contributo del post-processing ha comunque permesso un miglioramento rispetto ai risultati grezzi generati dalla rete: si genera un incremento del 5% sul DSC e sulla PRECISION. Senza questa fase, le prestazioni del modello, soprattutto in contesti complessi come la segmentazione di lesioni polmonari, risulterebbero sensibilmente inferiori, compromettendo la robustezza complessiva del sistema.

Le prestazioni del modello nella segmentazione delle lesioni polmonari variano in base alle dimensioni delle lesioni stesse. Le lesioni di dimensioni estese mostrano una buona capacità di identificazione da parte della rete, grazie alla loro maggiore visibilità nei dati. Questa maggiore visibilità è attribuibile alla dimensione delle lesioni, che le rende più facilmente distinguibili. Tuttavia, la segmentazione di queste lesioni richiede un accorgimento, poiché tende a sovrasegmentarsi rispetto alla segmentazione manuale. La sovrasegmentazione si verifica quando il modello identifica aree più ampie di quelle effettivamente occupate dalla lesione, il che può portare a una rappresentazione imprecisa della sua estensione reale. Questa difficoltà è accentuata quando il tumore è a contatto con tessuti circostanti, cioè si trova attaccato al contorno del parenchima; in questi casi i margini della lesione non sono ben definiti.

Al contrario, la segmentazione delle lesioni di dimensioni più piccole presenta sfide significative. Anche se possono essere identificate correttamente, queste lesioni sono più suscettibili a confusione con il rumore di fondo o con piccole anomalie nei tessuti polmonari, portando a un aumento dei falsi negativi. La loro ridotta superficie e la minore estensione rendono difficile per la rete apprendere le caratteristiche distintive necessarie per una corretta segmentazione. Inoltre, il fatto che le lesioni più piccole siano spesso associate a margini poco definiti complica ulteriormente la loro segmentazione. Questi fattori possono causare perdite nel rilevamento delle lesioni.

Parallelamente all'approccio di segmentazione delle lesioni tumorali, è stata intrapresa anche l'estrazione automatica della maschera dei polmoni, un passo cruciale per una corretta analisi delle immagini. Tuttavia, la posizione delle lesioni ha avuto un impatto significativo su questo processo, in particolare per quelle adiacenti alla pleura e al mediastino. Queste lesioni, infatti, presentavano un'intensità di pixel simile a quella del tessuto circostante, portando a un problema notevole

nell'applicazione di metodi di sogliatura e di labeling. Su un totale di 88 pazienti analizzati, ben il 66% delle lesioni presentavano questa problematica, evidenziando l'entità della sfida. Quando l'algoritmo tentava di segmentare automaticamente i polmoni, ma le lesioni che si trovavano in queste posizioni critiche venivano scartate, poiché il loro profilo di intensità dei pixel si sovrapponeva a quello che veniva considerato sfondo. Di conseguenza, le maschere generate non riuscivano a riconoscere correttamente le lesioni, portando a falsi negativi significativi. Questa sfida evidenzia come le tecniche basate su sogliatura possano avere limitazioni significative quando applicate a contesti complessi come la segmentazione dei polmoni. L'incapacità di distinguere tra le lesioni e il tessuto sano circostante, a causa di questa similitudine di intensità, rappresentava un ostacolo importante non solo per l'estrazione della maschera dei polmoni, ma anche per l'accuratezza generale della successiva segmentazione delle lesioni tumorali. Di conseguenza, l'estrazione automatica della maschera polmonare, sebbene promettente, ha rivelato la necessità di sviluppare approcci più sofisticati in grado di gestire le complessità delle variazioni di intensità.

In conclusione, la segmentazione automatica delle lesioni polmonari rappresenta un significativo passo avanti nell'ambito dell'analisi medica, poiché offre un'opportunità per migliorare l'efficienza del flusso di lavoro clinico. I risultati ottenuti dal progetto evidenziano chiaramente l'importanza cruciale di diversi fattori, come la scelta del campo visivo, la gestione del dataset e le tecniche di post-processing, tutte variabili essenziali per garantire segmentazioni affidabili e accurate. In particolare, la corretta selezione del FOV consente di concentrare l'analisi sulle aree più rilevanti, riducendo il rumore e migliorando la qualità dei dati. La gestione adeguata del dataset, che include la raccolta di immagini eterogenee e rappresentative, è fondamentale per addestrare modelli robusti e generalizzabili. Inoltre, le tecniche di post-processing, possono notevolmente aumentare la capacità del modello di affrontare le variazioni nelle immagini cliniche.

L'integrazione della segmentazione automatica delle lesioni polmonari nelle pratiche cliniche quotidiane promette di trasformare il panorama della diagnostica oncologica, portando a un miglioramento della qualità delle cure e, di conseguenza, a un aumento della qualità della vita dei pazienti. La continua ricerca e sviluppo in questo campo sono fondamentali per superare le sfide esistenti e massimizzare il potenziale di questa tecnologia innovativa. Con un approccio sistematico e integrato, si può sperare di assistere a un significativo miglioramento negli esiti clinici per i pazienti affetti da tumore polmonare nei prossimi anni.

Capitolo 5

SVILUPPI FUTURI

Come obiettivi futuri, una possibile strategia di miglioramento potrebbe essere l'addestramento di un classificatore dedicato sulle ROI rimaste dopo il processo di post-elaborazione. Questo classificatore potrebbe essere allenato per estrarre caratteristiche distintive, come l'intensità, la texture, la forma e la densità delle strutture presenti, al fine di distinguere in modo più preciso tra vere lesioni tumorali e altre strutture polmonari non rilevanti. L'analisi di queste feature potrebbe fornire informazioni più dettagliate sulle caratteristiche morfologiche e radiomiche delle lesioni, migliorando la capacità del sistema di rilevare anche sottili variazioni che potrebbero sfuggire a una semplice segmentazione basata su regole geometriche. L'introduzione di questo ulteriore livello di analisi permetterebbe di affinare le predizioni, riducendo ulteriormente i falsi positivi e aumentando la precisione globale del sistema. Un approccio di questo tipo, combinando diverse tipologie di feature, consentirebbe al modello di adattarsi meglio alle variazioni fisiologiche e patologiche delle strutture polmonari, garantendo un'identificazione più robusta e affidabile anche in contesti clinici complessi.

Un ulteriore sviluppo futuro dovrebbe focalizzarsi sull'incremento del dataset utilizzato per l'addestramento della rete. È fondamentale assicurarsi che il dataset contenga un numero sufficiente di immagini che rappresentino equamente le diverse dimensioni del tumore all'interno dei polmoni. Attualmente, esiste un significativo squilibrio, soprattutto per quanto riguarda le lesioni di dimensioni più piccole, che sono altamente sottorappresentate. Questo squilibrio può influenzare negativamente le prestazioni del modello, poiché il sistema potrebbe non essere in grado di generalizzare adeguatamente per identificare e segmentare lesioni in posizioni meno comuni o più difficili da rilevare. Aumentare la varietà e la quantità di dati, includendo immagini di lesioni di diverse dimensioni, permetterebbe di addestrare un modello più robusto e performante. Un dataset più equilibrato non solo migliorerebbe le metriche di performance, ma renderebbe anche il modello più utile

e applicabile in contesti clinici, aumentando la sua affidabilità.

Bibliografia

- [1] *Polmoni*. Humanitas. 2024. URL: <https://www.humanitas.it/enciclopedia/anatomia/apparato-respiratorio/polmoni/> (cit. a p. 2).
- [2] *Statistiche di incidenza*. International Respiratory Coalition. 2024. URL: <https://international-respiratory-coalition.org/lung-facts/> (cit. a p. 2).
- [3] Jacques Ferlay, Murielle Colombet, Isabelle Soerjomataram, Colin Mathers, Donald M Parkin, Marlon Piñeros, Ariana Znaor e Freddie Bray. «Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods». In: *International journal of cancer* 144.8 (2019), pp. 1941–1953 (cit. a p. 2).
- [4] *Fattori di rischio*. AIRC. 2024. URL: <https://www.airc.it/cancro/informazioni-tumori/guida-ai-tumori/tumore-al-polmone> (cit. a p. 2).
- [5] *Types of Lung Cancer*. LUNGEVITY, Transforming Lung Cancer. 2024. URL: <https://www.lungevity.org/lung-cancer-basics/types-of-lung-cancer> (cit. a p. 3).
- [6] *Cancro del polmone non a piccole cellule*. ESMO, Good Science Better Medicine Best Practice. 2024. URL: <https://www.esmo.org/content/download/87433/1608958/1/IT-Cancro-del-Polmone-non-a-Piccole-Cellule-NSCLC-Guida-per-il-Paziente.pdf> (cit. alle pp. 3, 4).
- [7] *Cosa fare in caso di sospetto di tumore polmonare?* 2021. URL: https://www.gastroepato.it/tumore_del_polmone3.htm (cit. a p. 4).
- [8] *Diagnosis Lung Cancer*. LUNGEVITY, Transforming Lung Cancer. 2024. URL: <https://www.lungevity.org/lung-cancer-basics/diagnosing-lung-cancer> (cit. a p. 5).
- [9] *TC (Tomografia Computerizzata)*. AIRC. 2024. URL: <https://www.airc.it/cancro/affronta-la-malattia/guida-agli-esami/tc-tomografia-computerizzata#:~:text=%C3%88%20un%20esame%20radiologico%20che,diversi%20tipi%20di%20tessuto%20analizzati>. (cit. a p. 5).

-
- [10] *Tomografia Computerizzata Assiale*. European Hospital. 2024. URL: <https://www.europeanhospital.it/index.php/en/tomografia-assiale-computerizzata> (cit. a p. 6).
- [11] *Programma di Screening Del Tumore Polmonare con Tc Low-Dose*. IEO, Istituto Europeo di Oncologia. 2024. URL: <https://www.iew.it/it/PREVENZIONE/In-primo-piano/TAC-a-basso-dosaggio/SCREENING-TUMORE-POLMONARE-CON-TC-LOW-DOSE/> (cit. a p. 6).
- [12] Elizabeth Escobar-Peralta, Rebeca Gil-García, Abril Quezada e Otoniel López-Riverol. «Helical tomotherapy: advanced radiotherapy technology». In: *Revista Modica del Hospital General de Mexico* 85 (dic. 2022). DOI: 10.24875/HGMX.22000021 (cit. a p. 7).
- [13] Neil G Burnet, Simon J Thomas, Kate E Burton e Sarah J Jefferies. «Defining the tumour and target volumes for radiotherapy». In: *Cancer Imaging* 4.2 (2004), p. 153 (cit. a p. 7).
- [14] Elizabeth Escobar-Peralta, Rebeca Gil-García, Abril A Quezada-Bautista e Otoniel López-Riverol. «Helical tomotherapy: advanced radiotherapy technology». In: *Revista médica del Hospital General de México* 85.4 (2022), pp. 179–184 (cit. a p. 7).
- [15] Xi Liu, Kai-Wen Li, Ruijie Yang e Li-Sheng Geng. «Review of deep learning based automatic segmentation for lung cancer radiotherapy». In: *Frontiers in oncology* 11 (2021), p. 717039 (cit. alle pp. 8, 9).
- [16] Martijn PA Starmans, Sebastian R van der Voort, Jose M Castillo Tovar, Jifke F Veenland, Stefan Klein e Wiro J Niessen. «Radiomics: data mining using quantitative medical image features». In: *Handbook of medical image computing and computer assisted intervention*. Elsevier, 2020, pp. 429–456 (cit. a p. 8).
- [17] Christian Janiesch, Patrick Zschech e Kai Heinrich. «Machine learning and deep learning». In: *Electronic Markets* 31.3 (2021), pp. 685–695 (cit. a p. 10).
- [18] Harsh Kukreja, N Bharath, CS Siddesh e S Kuldeep. «An introduction to artificial neural network». In: *Int J Adv Res Innov Ideas Educ* 1.5 (2016), pp. 27–30 (cit. a p. 10).
- [19] *What is neuron and artificial neuron in deep learning?* 2017. URL: <http://www.mplsvpn.info/2017/11/what-is-neuron-and-artificial-neuron-in.html> (cit. a p. 11).
- [20] K O’Shea. «An introduction to convolutional neural networks». In: *arXiv preprint arXiv:1511.08458* (2015) (cit. a p. 13).

- [21] Van Hiep Phung e Eun Joo Rhee. «A deep learning approach for classification of cloud image patches on small datasets». In: *Journal of information and communication convergence engineering* 16.3 (2018), pp. 173–178 (cit. a p. 13).
- [22] *Batch Normalization in Convolutional Neural Networks*. 2024. URL: <https://www.baeldung.com/cs/batch-normalization-cnn> (cit. a p. 13).
- [23] Ruben THM Larue, Gilles Defraene, Dirk De Ruyscher, Philippe Lambin e Wouter Van Elmpt. «Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures». In: *The British journal of radiology* 90.1070 (2017), p. 20160665 (cit. a p. 14).