

Università di Pisa



Facoltà di Scienze Matematiche Fisiche e Naturali

Corso di Laurea Specialistica in Scienze Fisiche

Anno Accademico 2012/2013

Tesi di Laurea Magistrale:

*Un web service per l'utilizzo clinico di un
sistema CAD (Computer Aided Detection)
per l'analisi automatica di tomografie
polmonari*

*Relatore
Dott.ssa M. E. Fantacci*

*Candidato
Marco Saletta*

Indice

1	Contesto clinico e sistemi CAD	13
1.1	Introduzione	13
1.2	Cancro e Tumore al Polmone	13
1.3	Screening	17
1.4	Tomografia Computerizzata	22
1.5	Dose, Dose Equivalente e Dose effettiva .	24
1.6	Anatomia del polmone	27
1.7	CAD ed utilizzo in screening	30
1.8	Valutazione delle prestazioni di un CAD .	32
1.9	Datasets	35
1.9.1	ITALUNG-CT	36
1.9.2	ANODE09	37
1.9.3	LIDC-IDRI	38
2	Descrizione CAD VBNA	42
2.1	Introduzione	42
2.2	Segmentazione del polmone	42
2.3	Ricerca noduli interni (CAD_I)	46
2.4	Ricerca noduli pleurici (CAD_{JP})	48
2.5	Riduzione falsi positivi	51
2.6	Combinazione CAD_I e CAD_{JP}	54
3	Descrizione CAD CAM	55
3.1	Introduzione	55
3.2	Segmentazione polmone e strutture interne	55
3.3	Ricerca dei noduli	59
3.4	Features	59
3.5	Filtraggio	61
3.6	Classificatore neurale	63
4	Validazione CAD VBNA su database LIDC-IDRI	66
4.1	Introduzione	66
4.2	Operazioni preliminari	66
4.3	Utilizzo del CAD	67
4.4	Combinazione dei risultati	71
4.5	Realizzazione della curva FROC	72

4.6	Risultati su database LIDC-IDRI	73
5	Combinazione CAD CAM e CAD VBNA e validazione su database LIDC-IDRI	80
5.1	Introduzione	80
5.2	Utilizzo di più CAD combinati	80
5.3	Combinazione CAD CAM e CAD VBNA	81
5.4	Risultati combinazione CAD CAM e CAD VBNA (CAD M5L)	82
6	Test su database ANODE09	84
6.1	Introduzione	84
6.2	Validazione CAD M5L con radiologi	84
6.3	Risultati validazione con radiologi	86
6.4	Confronto con CAD commerciale	91
6.5	Analisi database ANODE09 con il CAD Lung Care SIEMENS	93
6.6	Risultati Lung Care SIEMENS (LCS) su ANODE09 e confronto con il CAD M5L	96
7	Servizio di CAD on-line	100
7.1	Introduzione	100
7.2	Cloud Computing	100
7.3	Implementazione CAD CAM e CAD VBNA in cloud	103
7.4	Motivazioni della realizzazione del servizio web	105
7.5	Utilizzo del servizio web	107
7.6	Test del servizio web	110
8	Conclusioni	117
A	Utilizzo plugin Osirix	121
A.1	Introduzione	121
A.2	Operazioni Iniziali	121
A.3	Refertazione senza CAD	122
A.4	Refertazione con CAD	125

Elenco delle figure

1	Principali tipi di cancro per insorgenza e causa di morte in U.S. per il 2013 [21] (*Le stime sono arrotondate alla decina più prossima ed escludono alcune tipologie di cancri alla pelle e carcinomi in situ tranne quello alla vescica)	15
2	Curve di mortalità maschile e femminile per vari tipi di cancro al variare dell'anno [20].	16
3	Differenze nel numero di decessi per tumori al polmone nel confronto tra pazienti sottoposti a screening con CT e non (sopra) . Tasso di sopravvivenza per paziente con tumori al polmone di stadio I nel caso sia stato effettuato o meno un intervento di rimozione (sotto). Dati tratti da [23].	19
4	Rappresentazione schematica del funzionamento di un sistema CT: acquisizione di una fetta (sinistra) e acquisizione a spirale con movimento costante del lettino (destra).	23
5	Spettro, prima e dopo la filtrazione, per un tubo a raggi x a diversi valori di kV e mAs	24
6	Valori del fattore di peso w_r della dose equivalente per diversi tipi di radiazione (sinistra) e grafico della variazione del fattore di peso in funzione dell'energia per neutroni (destra) [58].	26
7	Rappresentazione schematica delle varie componenti della struttura polmonare nel suo complesso.	28
8	Alcune strutture anatomiche visibili in un esame CT: 1 inizio divisione bronchi a fine trachea (1) e prime generazioni (2), colonna (3), costole (4), sterno (5), aria esterna (6), lettino (7), cuore e strutture connesse (8).	29
9	Immagine contenente un nodulo, all'interno del cerchio indicato dalla freccia, attaccato ad un vaso. Si può notare come la struttura nel complesso risulti molto simile alle altre presenti nel polmone, rendendo difficile l'individuazione.	29
10	Esempi di visualizzazioni dei risultati dell'analisi delle immagini tramite CAD, con indicazioni su zone sospette.	31
11	Esempi di curve ROC e FROC	35
12	Distribuzione delle percentuali dei 7371 noduli annotati da diversi numeri di radiologi (in grigio) e delle annotazioni nel complesso, comprendendo anche i "non noduli" (nero) [32].	40

13	Ricostruzione della superficie del polmone dopo il ri-campionamento e la prima operazione dei ricerca di componenti connesse. In trasparenza si possono vedere i vasi e la varie vie aeree.	43
14	Segmentazione del polmone prima della rimozione delle pareti delle vie aeree e vasi (a) e dopo la loro rimozione (b).	45
15	Fasi della segmentazione tridimensionale del polmone: immagine di input (a), dopo il processo di tresholding (b), dopo la ricerca delle componenti connesse (c), output dell'algoritmo (d).	45
16	Rappresentazione della maschera finale della segmentazione del polmone.	46
17	Rappresentazione di un nodulo interno: ricostruzione di una parte interna del polmone con nodulo evidenziato in rosso (a), sezione di un esame CT contenente nodulo interno (b).	47
18	Schema delle 15 combinazioni di poligoni uniche utilizzate nell'algoritmo "marching cube" [61].	49
19	Rappresentazione di un nodulo pleurico: rendering della superficie con nodulo evidenziato in verde (a), sezione estratta da un esame CT (b).	50
20	Rappresentazione della matrice score $S(x, y, z)$ con le intersezioni delle normali.	50
21	Esempio di costruzione del vettore dei livelli di grigio.	52
22	Diagramma a blocchi del CAD CAM.	56
23	Grafici raggio <i>vs</i> sfericità, per le CT del database LIDC-IDRI, nel caso di pixel da $1.25mm$ (a) e $2.5mm$ (b). I quadrati neri rappresentano candidati corrispondenti a veri noduli.	62
24	Schema degli strati di neuroni che costituiscono una rete neurale [54]. Lo strato nascosto può anche essere composto da più layers di neuroni.	64
25	Schermata del visualizzatore DICOM Osirix con la lista dei casi caricati nel programma.	70
26	Schermata di Osirix durante la visualizzazione di un esame. La finestra sulla destra è quella relativa al plugin per la visualizzazione dei findings del CAD. Il candidato nodulo selezionato nell'elenco è evidenziato dal cerchio rosso sulla fetta dell'esame. Come si può notare sotto la lista dei findings, è possibile variare la soglia sulla probabilità di quelli che vengono elencati.	71

27	Caso 176 del database LIDC-IDRI. Il paziente presenta sostanzialmente una sola struttura polmonare.	74
28	Caso 262 del database LIDC-IDRI. L'immagine è stata acquisita con il paziente girato al contrario.	75
29	Confronto FROC del CAD_I e del CAD_{JP} sul database LIDC-IDRI escludendo il training.	76
30	FROC della combinazione del CAD_I con il CAD_{JP} sul database LIDC-IDRI esclusi i casi di training.	77
31	FROC della combinazione del CAD_I con il CAD_{JP} sul database LIDC-IDRI esclusi i casi di training. Scala estesa fino a $80^{FPs}/scan$	77
32	Confronto tra le FROC del CAD VBNA ottenute escludendo solo i 69 casi di training o escludendo il training e i casi per cui la ricerca dei noduli è fallita.	79
33	Confronto delle FROC dei CAD CAM, VBNA e loro combinazione sul database LIDC-IDRI, escludendo il training. . . .	83
34	Confronto delle FROC dei CAD CAM, VBNA e loro combinazione sul database LIDC-IDRI, escludendo il training. Scala estesa fino a $80^{FPs}/scan$	83
35	Schermata di Osirix con la finestra del plugin utilizzata per le refertazioni. Le colonne in basso mostrano le annotazioni del radiologo (a sinistra) e i findings del CAD (a destra). Cliccando su una delle linee, come mostrato nell'immagine la ROI corrispondente viene visualizzata ed evidenziata sulla fetta corrispondente.	86
36	Distribuzione delle distanze tra i noduli annotati da ogni coppia di radiologi, utilizzando le refertazioni realizzate senza l'ausilio del CAD.	89
37	Distribuzione delle distanze tra i noduli annotati da ogni coppia di radiologi, utilizzando le refertazioni realizzate con l'ausilio del CAD.	89
38	Schermata output del CAD LCS.	95
39	Confronto delle distribuzioni delle distanze tra le annotazioni di ogni radiologo, ottenute senza il CAD M5L, rispetto al CAD LCS.	97
40	Confronto delle distribuzioni delle distanze tra le annotazioni di ogni radiologo, ottenute con l'ausilio del CAD M5L, rispetto al CAD LCS.	97

41	Schema della struttura di un servizio cloud, con i vari soggetti coinvolti [64].	101
42	Esempio di come sono riportati i findings all'interno del report <i>.pdf</i>	106
43	Esempio di visualizzazione dei risultati del CAD on-line tramite l'apposito plugin di Osirix. La freccia indica il cursore che permette di variare la soglia sulla probabilità per la visualizzazione dei findings.	106
44	Pagina di accesso del sito per l'utilizzo del servizio di CAD on-line.	107
45	Pagina per l'upload di esami da analizzare attraverso il servizio web.	109
46	Pagina di riepilogo dei casi inviati dall'utente con relativi risultati. Le frecce rosse indicano le liste di findings in formato <i>.pdf</i> e <i>.xml</i>	109
47	Esame, realizzato presso l'Ospedale di Cuneo, con nodulo di dimensioni macroscopiche non riconosciuto dal CAD.	114
48	Menù per la selezione dei Plugins	122
49	Selezione Tools e disegno ROI sull'immagine	123
50	Salvataggio della ROI indicante il finding trovato	124
51	Visualizzazione e valutazione findings CAD.	126

Elenco degli Acronimi

AIRC	Associazione Italiana per la Ricerca sul Cancro
CAD	Computer Aided Detection
CAD LCS	CAD Lung Care SIEMENS
CAD M5L	CAD CAM+VBNA
CAM	Channeler Ant Model
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine format
FN	False Negative
FP	False Positive
FROC	Free Receiver Operating Characteristic
HU	Hounsfield Unit
IDRI	Image Database Resource Initiative
LDCT	Low Dose Computed Tomography
LIDC	Lung Image Database Consortium
M5L	Magic 5 Lung
NNS	Number of Needed Scan
NSCLC	Non-Small Cell Lung Cancer
RGVP	Region Growing Volume Plateau
ROC	Receiving Operating Characteristic
ROI	Region Of Interest
SaaS	Software as a Service
SCLC	Small Cell Lung Cancer

TAC	Tomografia Assiale Computerizzata
TN	True Negative
TP	True Positive
USPSTF	U. S. Preventive Service Task Force
VBNA	Voxel Based Neural Approach

Introduzione

Negli ultimi anni grande interesse è stato dimostrato dalla comunità sanitaria internazionale verso la possibilità di sviluppo di programmi su calcolatore, chiamati CAD da “Computer Aided Detection”, che consentano la rivelazione automatica di anomalie all’interno di esami clinici. La complessità e il numero di informazioni da esaminare in molte metodologie diagnostiche è tale da suggerire l’utilizzo di sistemi computerizzati di questo tipo come supporto al lavoro del medico. Tali sistemi possono trovare ulteriore validità di utilizzo all’interno di programmi di screening, dove i medici sono chiamati ad analizzare un grande numero di esami, molti dei quali negativi.

Uno degli ambiti più promettenti per l’utilizzo di sistemi CAD è quello delle patologie tumorali polmonari, sia per le motivazioni sopra indicate e l’intrinseco tasso di mortalità di tali patologie, sia per i recenti sviluppi sulla possibilità di screening.

Sulla base di diversi studi la U.S. Preventive Services Task Force (USPSTF) ha infatti recentemente rilevato forti indicazioni dell’efficacia di uno screening tramite Tomografia Computerizzata a bassa dose (LDCT) per la diminuzione del numero di decessi legati a questa patologia [6].

Tale conclusione, negli anni a venire, fornirà un’importante spinta per lo sviluppo di programmi di screening sia in America che in Europa.

In questa tesi sono state verificate le prestazioni di due sistemi CAD per l’individuazione automatica di noduli all’interno di immagini TAC polmonari, sia singolarmente che combinati tra loro, e della loro implementazione all’interno di un servizio web di CAD on-line.

I CAD utilizzati sono il Voxel Based Neural Approach (VBNA) [49], sviluppato presso Dipartimento di Fisica e l’INFN di Pisa, e il Channeler Ant Model (CAM), realizzato presso l’Università di Torino [47].

Nel *Capitolo 1* di questo testo viene introdotto il contesto generale in cui si inserisce il lavoro di tesi, sia a livello clinico che per quanto concerne i sistemi CAD.

Insieme alle descrizioni delle patologie tumorali polmonari, dell’anatomia del polmone, dei programmi di screening e della tomografia computerizzata con le relative grandezze dosimetriche è quindi riportata un’introduzione a tali sistemi, al loro utilizzo e alle metodologie di stima delle prestazioni.

Nel *Capitolo 2* e nel *Capitolo 3* vengono descritte in dettaglio le strutture concettuali rispettivamente del CAD VBNA e del CAD

CAM. Entrambi i CAD si compongono, analogamente a tutti i sistemi di questo tipo, sostanzialmente di tre moduli principali (segmentazione, ricerca dei candidati noduli e riduzione dei falsi positivi), ma utilizzano approcci molto diversi.

Con il *Capitolo 3* si conclude la parte descrittiva di questa tesi; quanto riportato nei capitoli che seguono è stato interamente realizzato durante il suo svolgimento.

Il *Capitolo 4* riporta la validazione del CAD VBNA eseguita su i 1018 casi del database internazionale LIDC-IDRI, la più grande raccolta di esami CT polmonari refertati e pubblicamente disponibili. Di questi, 69 esami sono stati utilizzati come set di training e 949 come set di validazione.

Nel capitolo è riportato sia il lavoro di preparazione che è stato necessario compiere, sia quello vero e proprio per l'analisi dei casi tramite il CAD con la descrizione delle procedure di utilizzo dello stesso. I risultati ottenuti sono analizzati nel dettaglio e riportati sotto forma di curve FROC.

Sullo stesso database è stata eseguita inoltre la validazione delle combinazioni tra il CAD VBNA e il CAD CAM, come riportato nel *Capitolo 5*.

L'utilizzo di più sistemi combinati risulta molto interessante in quanto, come mostrato in [29], porta ad un miglioramento delle prestazioni offerte singolarmente da ogni programma. Poiché ogni CAD adotta strategie diverse per l'individuazione e la selezione dei candidati noduli, nella combinazione i vari sistemi possono risultare complementari tra loro.

Sempre nel *Capitolo 5*, oltre ai risultati ottenuti, sono riportati gli aspetti teorici e le operazioni effettuate per la realizzazione della combinazione.

Il sistema ottenuto dalla combinazione del CAD VBNA con il CAD CAM è stato denominato CAD M5L (Magic 5 Lung), nome della collaborazione nell'ambito di cui sono nati i due sistemi [27].

Nel *Capitolo 6* viene descritto l'utilizzo del CAD M5L sul database ANODE09, una raccolta di 50 esami appositamente realizzata per permettere il confronto tra le prestazioni di diversi sistemi CAD.

La validazione dei CAD CAM, VBNA e della loro combinazione su questo database era già stata effettuata ([29, 47]), i casi disponibili sono stati quindi utilizzati per testare l'efficacia del CAD M5L come seconda lettura per la refertazione degli esami. Tre radiologi hanno

esaminato i 50 esami separatamente ed in seguito hanno rivisito le proprie refertazioni utilizzando le indicazioni fornite dal CAD.

Nello stesso capitolo è riportato anche il test di confronto tra il CAD M5L e un CAD commerciale, il CAD Lung Care SIEMENS (CAD LCS), installato sulle consolle delle relative macchine TAC. Per entrambi i test è riportata la descrizione di quanto è stato compiuto per acquisire i dati, elaborarli e i risultati ottenuti.

Al momento tutte le liste di noduli dei radiologi e findings dei CAD sono state inviate a curatori del database perchè siano valutate rispetto al gold standard dello stesso, siamo in attesa di risposta.

Infine, nel *Capitolo 7*, è descritta la realizzazione di un servizio web di CAD on-line, basato sul CAD M5L, che permette l'analisi di esami CT polmonari semplicemente tramite un qualsiasi browser ed una connessione internet.

Questo approccio rappresenta un punto di originalità rispetto alle classiche modalità di impiego di sistemi CAD in ambito ospedaliero e permette di aumentare notevolmente semplicità e flessibilità di utilizzo. L'implementazione del CAD M5L all'interno del servizio è avvenuta tramite un approccio di Cloud Computing, utilizzando le risorse hardware messe a disposizione dall'INFN di Torino.

Nel capitolo è riportata una breve introduzione al Cloud Computing e alle motivazioni che hanno portato all'utilizzo di tale approccio per la realizzazione del servizio, la descrizione di quanto compiuto per realizzare l'implementazione e i risultati ottenuti su 14 esami provenienti dall'Ospedale S. Croce di Cuneo.

I test di utilizzo del servizio stanno attualmente proseguendo all'interno di diversi centri ospedalieri.

Le considerazioni conclusive di questo lavoro di tesi sono riportate nel *Capitolo 8*, mentre nell'appendice si può trovare una guida per l'utilizzo del plugin per il visualizzatore DICOM Osirix.

Parte I:
Introduzione Contesto

1 Contesto clinico e sistemi CAD

1.1 Introduzione

I paragrafi riportati in questo primo capitolo forniscono una descrizione dell'ambito clinico in cui si collocano i sistemi CAD e di come in esso possano essere utilizzati.

Per inquadrare gli aspetti anatomici e clinici sono riportate le descrizioni generali di patologie tumorali, dello screening come strumento di prevenzione, della tomografia computerizzata e dell'anatomia del polmone.

Per quanto invece concerne i sistemi CAD sono riportate una descrizione introduttiva (con riferimento al loro utilizzo nello screening) e le definizioni delle grandezze e delle metodologie con cui è possibile stimarne le prestazioni.

Sono infine descritte le caratteristiche generali dei tre diversi database di immagini CT che sono stati utilizzati nello sviluppo e nel test di sistemi per la ricerca dei noduli polmonari.

1.2 Cancro e Tumore al Polmone

Con cancro si intende la famiglia di patologie caratterizzata da un'incontrollata crescita e diffusione di cellule malate. In queste cellule vengono a mancare o ad essere mal funzionanti i meccanismi interni che controllano la riproduzione e la durata della vita, risultando molto più prolifiche di quelle sane.

Oltre ad espandersi nella zona di origine è possibile che avvenga una migrazione, definita metastatizzazione, con conseguente colonizzazione anche di tessuti situati in altre parti del corpo. La presenza di metastasi rappresenta lo stadio più avanzato della patologia tumorale e la causa diretta di molti decessi [19].

Il comportamento anomalo delle cellule malate è indotto da mutazioni geniche che sono principalmente dovute all'esposizione a fattori esterni quali fumo di sigaretta, amianto, combustione di idrocarburi, raggi UV, sostanze e solventi chimici, alcol, dieta, gas radon ecc ecc., ma possono anche avere origine genetica.

Per esempio si stima che un fumatore abbia probabilità di sviluppare un cancro al polmone 23 volte superiore rispetto ad un non fumatore.

In America il cancro è secondo alle patologie cardiache come numero di decessi dovuti a malattia. Nel 2013 è prevista la morte di 580,350 americani,

circa 1,600 al giorno a causa del cancro, ovvero approssimativamente una morte ogni quattro [20].

Il tasso di sopravvivenza a 5 anni tra il 2002 e il 2008, ottenuto confrontando pazienti malati di cancro e persone sane e considerando tutti i possibili tipi di cancro, è stato del 68%. Oltre il 30% degli individui a cui viene diagnosticata tale patologia quindi decede entro 5 anni.

Le *Figura 1* e la *Figura 2* mostrano rispettivamente la situazione per l'anno 2013 e l'evoluzione nel corso degli anni di insorgenza e mortalità relativa ai vari tipi di cancro. L'andamento mostrato nelle statistiche può essere considerato valido anche per il nostro paese.

Si può notare come il cancro al polmone sia e sia stato la principale causa di morte tra tutte le tipologie citate.

Se la curva della mortalità maschile nella *Figura 2* mostra una diminuzione di decessi negli ultimi 10 anni quella femminile mostra una minima discesa solo in anni ancora più recenti. Le diminuzioni sono comunque inferiori a quelle di altre tipologie.



Secondo le stime riportate nella *Figura 1* il tumore al polmone rappresenterà il 14% della totalità dei nuovi casi diagnosticati nel 2013.

I tassi di incidenza e mortalità si mantengono così elevati, nonostante la crescente attenzione verso i fattori di esposizione, per le carenze di diagnosi e terapia. Rispetto ad altre patologie anche tumorali infatti le terapie portano a risultati nettamente inferiori in quanto effettuate non precocemente.

Nello caso specifico del tumore polmonare la crescita incontrollata di determinati tipi di cellule provoca la comparsa di una massa solida che può ostruire le vie respiratorie o provocare emorragie polmonari o bronchiali.

Il 95% dei tumori polmonari è rappresentato dai cosiddetti broncògeni che insorgono sulla superficie dell'epitelio, ovvero del tessuto che ricopre alveoli, bronchi e bronchioli, mentre nei restanti casi vengono interessati altri tessuti quali quello nervoso e quello endocrino. Tra i broncògeni il 45 – 50% si manifesta su vie aeree di dimensioni maggiori ed è strettamente legato all'esposizione al fumo di sigaretta. La restante parte si sviluppa invece su bronchi con dimensioni minori. All'interno di questi gruppi i casi specifici si distinguono soprattutto per essere caratterizzati da cellule di dimensioni piccole (SCLC, da Small Cell Lung Cancer) o non piccole (NSCLC, da Non-Small Cell Lung Cancer), fattore che comporta differenze di crescita e diffusione con conseguenti differenze di trattamento.

Ricordiamo che, anche per soggetti sani, all'interno dei polmoni vi è comunque un certo numero di strutture nodulari che hanno caratteristiche tali

Estimated New Cases*							
			Males	Females			
Prostate	238,590	28%			Breast	232,340 29%	
Lung & bronchus	118,080	14%			Lung & bronchus	110,110 14%	
Colorectum	73,680	9%			Colorectum	69,140 9%	
Urinary bladder	54,610	6%			Uterine corpus	49,560 6%	
Melanoma of the skin	45,060	5%			Thyroid	45,310 8%	
Kidney & renal pelvis	40,430	5%			Non-Hodgkin lymphoma	32,140 4%	
Non-Hodgkin lymphoma	37,600	4%			Melanoma of the skin	31,630 4%	
Oral cavity & pharynx	29,620	3%			Kidney & renal pelvis	24,720 3%	
Leukemia	27,880	3%			Pancreas	22,480 3%	
Pancreas	22,740	3%			Ovary	22,240 3%	
All Sites	854,790	100%			All Sites	805,500	100%



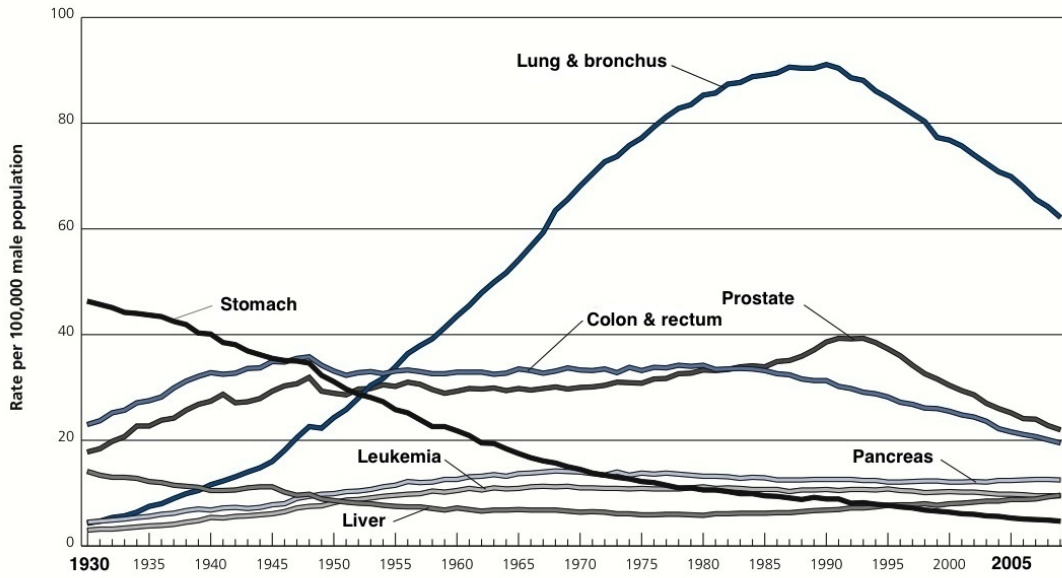
Estimated Deaths							
			Males	Females			
Lung & bronchus	87,260	28%			Lung & bronchus	72,220 26%	
Prostate	29,720	10%			Breast	39,620 14%	
Colorectum	26,300	9%			Colorectum	24,530 9%	
Pancreas	19,480	6%			Pancreas	18,980 7%	
Liver & intrahepatic bile duct	14,890	5%			Ovary	14,030 5%	
Leukemia	13,660	4%			Leukemia	10,060 4%	
Esophagus	12,220	4%			Non-Hodgkin lymphoma	8,430 3%	
Urinary bladder	10,820	4%			Uterine corpus	8,190 3%	
Non-Hodgkin lymphoma	10,590	3%			Liver & intrahepatic bile duct	6,780 2%	
Kidney & renal pelvis	8,780	3%			Brain & other nervous system	6,150 2%	
All Sites	306,920	100%			All Sites	273,430	100%

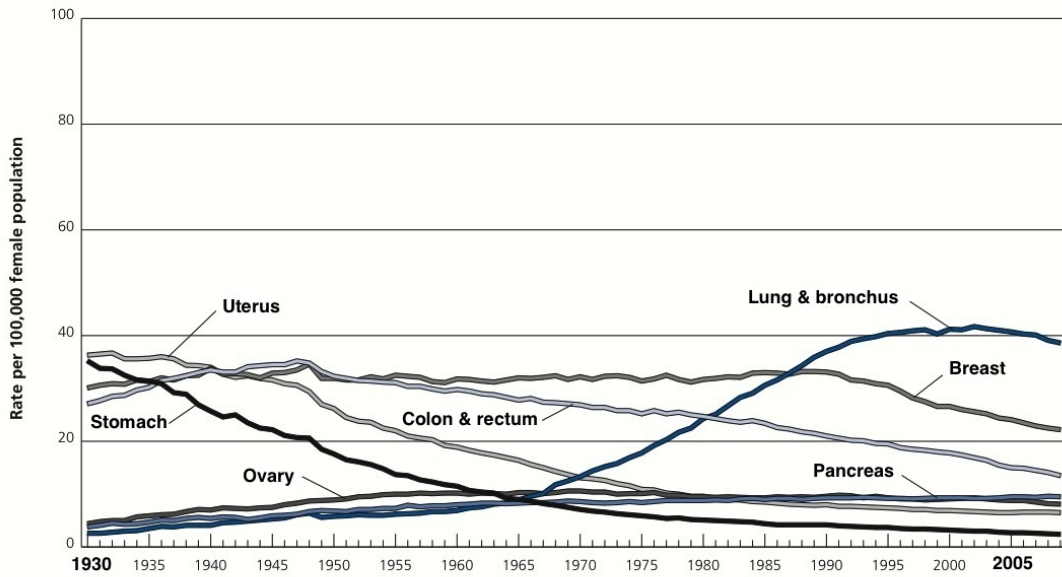
Figura 1: Principali tipi di cancro per insorgenza e causa di morte in U.S. per il 2013 [21] (*Le stime sono arrotondate alla decina più prossima ed escludono alcune tipologie di cancro alla pelle e carcinomi in situ tranne quello alla vescica)

Age-adjusted Cancer Death Rates*, Males by Site, US, 1930-2009



*Per 100,000, age adjusted to the 2000 US standard population.

Age-adjusted Cancer Death Rates*, Females by Site, US, 1930-2009



*Per 100,000, age adjusted to the 2000 US standard population. Uterus refers to uterine cervix and uterine corpus combined.

Figura 2: Curve di mortalità maschile e femminile per vari tipi di cancro al variare dell'anno [20].

da non risultare patologiche. Nelle popolazioni a rischio il numero di tali strutture può aumentare ed hanno inoltre una maggiore probabilità di sviluppare caratteristiche patologiche. Per questo motivo risultano comunque di interesse per l'analisi del radiologo.

Introduciamo infine il concetto di stadio del tumore ovvero la definizione, tramite un'opportuna classificazione, della dimensione e dell'estensione del cancro al momento della sua diagnosi. L'individuazione dello stadio corretto ha un'importanza fondamentale per la definizione delle terapie che dovranno essere adottate. Tipicamente il sistema di staging utilizzato è quello detto TNM: per un SCLC T definisce la dimensione del tumore, N il coinvolgimento dei linfonodi e M la presenza di metastasi distanti [22].

T varia tra 1 e 4 a seconda delle dimensioni del tumore primario.

N ha intervallo 0 – 3, dove:

- $N = 0$: cellule tumorali assenti dai linfonodi
- $N = 1$: metastasi presenti nel linfonodo più prossimo o in un piccolo numero di linfonodi locali
- $N = 2$: estensione del tumore intermedia tra $N = 1$ e $N = 3$
- $N = 3$: il tumore è esteso a linfonodi distanti o a molti locali.

Per quanto riguarda M invece:

- $M = 0$: non ci sono metastasi distanti
- $M = 1$: sono presenti metastasi in organi distanti

La *Tabella 1* mostra i principali quattro stadi del cancro a seconda della diffusione della patologia.

1.3 Screening

Con screening si intende la procedura diagnostica di controllo periodico, tramite opportuni esami diagnostici, di soggetti asintomatici appartenenti a sottogruppi dell'intera popolazione che hanno maggiore probabilità di sviluppare una data patologia. Questa maggiore predisposizione può provenire da variabili quali sesso, età, stile di vita, tipo di lavoro, provenienza geografica, anamnesi ecc ecc.

Stage	TNM subset	Stage	TNM subset
0	carcinoma in situ	IIIB	T4N0M0
IA	T1 N0 M0		T4 N1 M0
IB	T2 N0 M0		T4 N2 M0
IIA	T1 N1 M0		T1 N3 M0
IIIB	T2 N1 M0		T2 N3 M0
	T3 N0 M0		T3N3M0
IIIA	T3 N1 M0		T4 N3 M0
	T1 N2 M0	IV	Any T Any N M1
	T2 N2 M0		
	T3 N2 M0		

Tabella 1: Stadiazione del cancro secondo la classificazione TNM [22].

L'obiettivo di tale pratica è diagnosticare la malattia quando ancora si trova in uno stadio precoce ed è possibile intervenire in modo efficace. Se applicata alle patologie tumorali, tra le principali cause di morte in Europa e America, una diagnosi tempestiva può diminuire sensibilmente il numero di decessi o perlomeno aumentare l'aspettativa e la qualità di vita di un grande numero di persone.

La *Figura 3* mostra rispettivamente l'incremento delle probabilità di sopravvivenza per pazienti con tumore al polmone diagnosticato tramite screening e per cui sia stato possibile intervenire celermente (entro un mese) quando ancora la patologia si trovava alla stadio I .

La struttura generale di un programma di screening, a grandi linee, è la seguente:

- Definizione delle caratteristiche della popolazione interessata, durata del programma, tecnica diagnostica, intervalli di ripetizione ecc ecc.
- Comunicazione ai soggetti interessati che possono accettare o rifiutare
- Primo round diagnostico, detto "*baseline*". In questo contesto possono anche essere affiancati più metodi diagnostici, in modo da avere un quadro più completo.
- Controllo dei risultati ottenuti: in caso di risultato positivo il paziente viene chiamato ad eseguire ulteriori accertamenti. Se la diagnosi viene confermata si procede con la terapia. I pazienti negativi, o con anomalie di cui bisogna seguire lo sviluppo, vengono rimandati al round successivo.

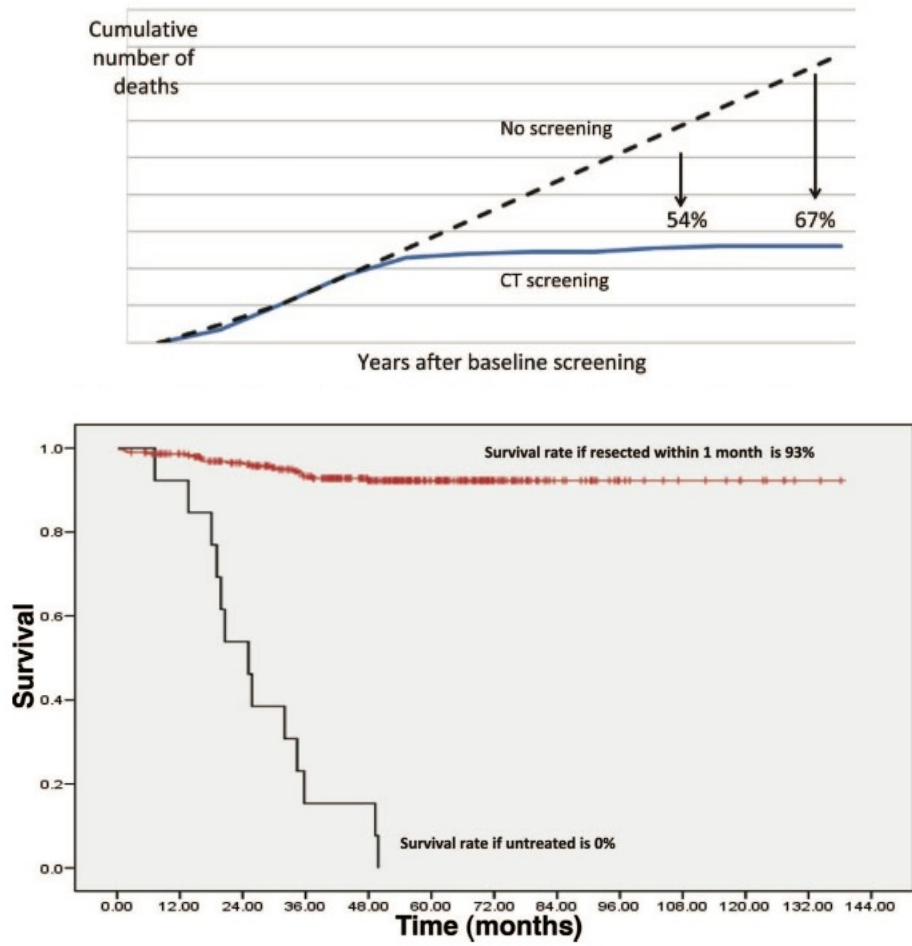


Figura 3: Differenze nel numero di decessi per tumori al polmone nel confronto tra pazienti sottoposti a screening con CT e non (sopra) . Tasso di sopravvivenza per paziente con tumori al polmone di stadio I nel caso sia stato effettuato o meno un intervento di rimozione (sotto). Dati tratti da [23].

- Si lascia trascorrere l'intervallo di tempo previsto.
- Secondo round di controllo utilizzando solo la metodologia diagnostica specifica del programma.
- Controllo dei risultati ottenuti con le stesse modalità del primo round.
- Si continua a procedere come descritto fino alla conclusione del programma, con prolungamento del “*follow-up*” nei casi in cui si renda necessario.

Esempio pratico è lo screening, correntemente attuato in Italia, per i tumori alla mammella, eseguito tramite mammografia bilaterale con frequenza biennale su donne di età compresa tra 50 e 69 anni.

Essendo una procedura molto impegnativa per il tempo richiesto per la refertazione degli esami, le competenze richieste, il numero di pazienti coinvolti e costo per il sistema sanitario, la possibilità di intraprendere una campagna di screening deve essere valutata attentamente.

Specialmente nel caso di patologie tumorali devono essere inoltre valutate possibili problematiche collaterali quali sovra-diagnosi, ovvero la rivelazione di noduli che non sarebbero mai diventati sintomatici, e la sovra-esposizione a radiazioni con il rischio di tumori radio indotti.

Diversi studi nel corso degli scorsi anni hanno vagliato l'efficacia di un programma di screening per tumori al polmone su individui con un'importante storia da fumatori ed un opportuno intervallo di età: NLST [7], DANTE [10], DLCST [9], MILD [8] e ITALUNG [11]. I valori di età e consumo di sigarette adottati da ogni studio per la selezione dei partecipanti sono elencanti nella *Tabella 2*.

Il numero *pacchetti-anno* è calcolato moltiplicando il numero di pacchetti fumati al giorno da una persona per il numero di anni in cui il soggetto ha fumato. A titolo esemplificativo il valore *20pacchetti – anno* può corrispondere ad un consumo di un pacchetto al giorno per 20 anni o a 2 pacchetti al giorno per 10 anni e via dicendo.

Tra i lavori sopra citati il più rilevante risulta essere quello effettuato dallo NLST [7]. I ricercatori hanno preso in esame persone di età compresa tra i 55 e i 74 anni (al momento dell'inizio dello screening) e con un consumo di almeno 30 pacchetti annui, esaminandoli tre volte a distanza di un anno. Sono stati considerati anche soggetti che avessero smesso di fumare da non

Studio	Pacchetti-anno	Intervallo età (anni)	Ex-fumatori da (anni)
NLST	≥30	55-74	<15
DANTE	≥20	60-74	<10
DLCST	≥20	50-70	<10 e oltre i 50 anni
MILD	≥20	≥49	<10
ITALUNG	≥20	55-69	<10

Tabella 2: Tabella di riepilogo delle caratteristiche dei partecipanti ad ogni studio: consumo di sigarette, intervallo di età e limite temporale entro cui sono accettati anche gli ex-fumatori. Dati tratti da [7, 10, 9, 8, 26].

più di 15 anni, in quanto diversi studi certificano la permanenza del rischio anche dopo aver smesso [12, 13, 14].

Sono stati invece escluse le persone con un cancro già diagnosticato, con sintomi sospetti o che erano già state sottoposte ad una CT toracica nei 18 mesi precedenti. Tra il 2002 e il 2004 sono stati iscritti nel programma 53,454 individui; lo screening è proceduto da agosto 2002 a settembre 2007, il follow-up per alcuni casi è stato protratto fino al 31 dicembre 2009. I pazienti sono stati divisi in modo casuale tra CT a bassa dose (26,722), eseguite con scanner a rivelatori multipli, e radiografia toracica (26,732), utilizzata come parametro di confronto.

Nel caso in cui un esame risultasse positivo sono stati eseguiti opportuni accertamenti, interventi o terapie.

I risultati ottenuti possono essere sintetizzati brevemente in questo modo: il numero di screening risultati positivi nei tre round è stato del 24% per la CT e del 7% per la radiografia. Di questi si sono rivelati essere falsi positivi il 96% e il 95% rispettivamente, con un incidenza risultante di cancro al polmone di 645 e 572 casi per 100000 persone-anno rispettivamente.

Il numero di decessi dovuti alla patologia (data una mortalità causata dagli interventi clinici trascurabile) è stato di 247 per 100000 persone-anno per il gruppo controllato tramite CT a bassa dose e di 309 per 100000 persone-anno nel gruppo della radiografia.

E' stato quindi riscontrato che lo screening tramite tomografia computerizzata ha determinato una diminuzione dei decessi per cancro al polmone del 20%, con una diminuzione di ~ 7% della mortalità dovuta ad altre cause. Il numero stimato di screen necessari per salvare una persona (NNS) è risultato essere di 320 (219 se si considerano tutte le possibili cause di morte).

Per confronto il valore di NNS per prevenire una morte dovuta a tumore al seno è di 1339, per donne di età compresa tra i 50 e i 59 anni dopo 11 o 20 anni di follow-up [16, 17].

Il 30 Luglio 2013 la U. S. Preventive Service Task Force (USPSTF) ha aggiornato la review del 2004 per lo screening al polmone con LDCT nell'articolo [6]. La nuova review, a fronte di 8000 articoli vagliati, è stata basata sui risultati degli studi NLST, DANTE, DLCST e MILD sopra citati, la cui rilevanza è stata valutata a seconda dei risultati ottenuti e delle scelte su età dei partecipanti, consumo di sigarette, lunghezza screening e follow-up effettuate dai vari studi.

Sulla base dei risultati esposti dagli studi considerati la USPSTF ha concluso (testualmente) che forti evidenze mostrano come lo screening tramite LDCT possa ridurre la mortalità dovuta al cancro al polmone e quella per altre cause, raccomandando sostanzialmente l'utilizzo.

Se questa review sarà approvata lo screening sarà inserito automaticamente tra le procedure del programma sanitario americano, andando ad interessare un grandissimo numero di cittadini [15].

1.4 Tomografia Computerizzata

I sistemi TAC sono tra gli strumenti più utilizzati nella normale pratica diagnostica ed hanno ampia diffusione all'interno di strutture di qualsiasi livello. Il concetto alla base di questa tecnica di imaging è la ricostruzione del volume di interesse partendo da un numero finito di proiezioni. Il corpo del paziente viene irraggiato tramite un tubo a raggi X e si raccolgono i dati relativi all'interazione della radiazione con il volume bersaglio [1].

L'irraggiamento avviene ruotando la sorgente intorno al corpo del paziente ad angoli predefiniti e traslandola longitudinalmente, come mostrato nella *Figura 4*.

I dati raccolti ad ogni angolo rappresentano le proiezioni da cui, con l'utilizzo di opportune tecniche di elaborazione tramite calcolatore, è possibile ricostruire le sezioni di interesse [4, 3, 2].

Le moderne apparecchiature utilizzano una modalità di acquisizione definita elicale o a spirale: per acquisire l'intero volume il lettino su cui si trova il paziente viene fatto avanzare in modo continuo all'interno dell'anello di rivelatori. Il moto della sorgente dal punto di vista del paziente risulta quindi a spirale, come mostrato nell'immagine di destra nella *Figura 4*.

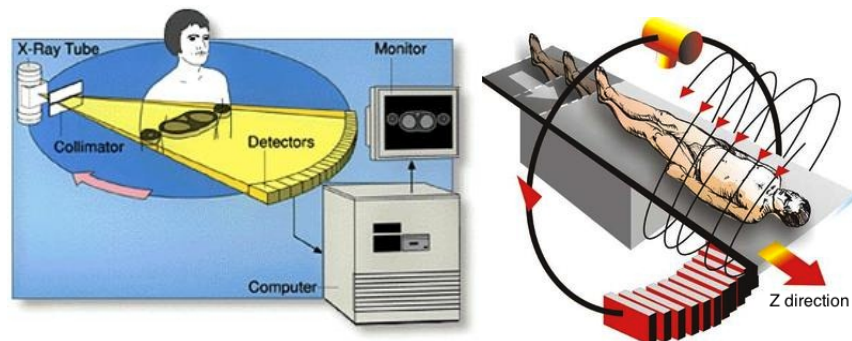


Figura 4: Rappresentazione schematica del funzionamento di un sistema CT: acquisizione di una fetta (sinistra) e acquisizione a spirale con movimento costante del lettino (destra).

Gli sviluppi più recenti hanno portato inoltre all'utilizzo di una struttura definita multi-slice o multi-detectors: un tubo per raggi x tramite un fascio a ventaglio illumina diversi anelli (4 o più) di rivelatori affiancati nella direzione longitudinale. In questo modo è possibile l'acquisizione di più fette contemporaneamente.

Nel tubo a raggi x un fascio di elettroni viene generato tramite un filamento riscaldato, viene accelerato tramite una differenza di potenziale e poi viene fatto impattare su un materiale ad alto z . In questo modo si ottiene la produzione di raggi X per bremsstrahlung ed emissione caratteristica, con energie che variano tra zero e il valore della differenza di potenziale utilizzata per accelerare gli elettroni. I valori più bassi di energia vengono però assorbiti all'interno del tubo o filtrati in quanto rappresentano una quantità di dose assunta dal paziente non associata ad un aumento dell'informazione raccolta. Nella *Figura 5* è riportato lo spettro per entrambi i casi per diversi valori di kV per l'accelerazione e mAs ovvero il prodotto tra il tempo di esposizione e la corrente all'interno del filamento.

La configurazione multislice-elica sopra descritta oltre ad un'ampia copertura volumetrica ed una grande velocità di acquisizione consente anche un'efficace gestione dello spessore delle sezioni che vengono acquisite, variando il numero di rivelatori utilizzati per la lettura di ogni fetta.

Si noti che data la natura dell'acquisizione elicale le fette non sono più definite tramite collimazione del fascio, ma sono ricostruite tramite interpolazione dei dati acquisiti. Il numero medio di sezioni così ottenute varia

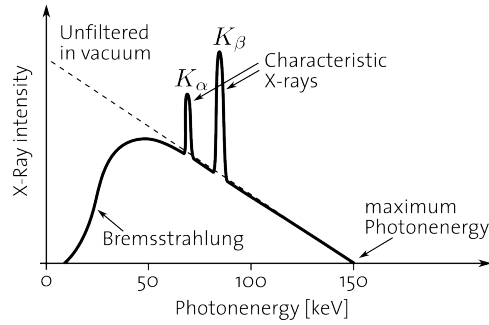


Figura 5: Spettro, prima e dopo la filtrazione, per un tubo a raggi x a diversi valori di kV e mAs

tra 300 e 400 per ogni esame, composte da voxel solitamente anisotropi e con ordini di grandezza per le dimensioni pari a $v_x = v_y \sim 0,5mm$ (intesi come pixel $2D$) e $v_z \sim 1mm$ (inteso come spessore della fetta). Negli esami acquisiti con macchine di ultima generazione invece i voxel risultano essere isotropi.

Trascurando lo spessore ogni sezione è una immagine due-dimensionale a livelli di grigio dove l'intensità di ogni pixel è il coefficiente di Hounsfield (HU), definito come

$$HU = \frac{\mu - \mu_{H_2O}}{\mu_{H_2O}} \times 1000 \quad (1)$$

dove μ è il coefficiente di attenuazione del materiale preso in esame e μ_{H_2O} è il coefficiente di attenuazione dell'acqua.

Tutte le immagini ottenute in ogni esame sono solitamente salvate nel formato DICOM, ovvero Digital Imaging and Communications in Medicine, che funge da standard per l'utilizzo e la gestione di questi tipi di file [5].

1.5 Dose, Dose Equivalente e Dose effettiva

Per descrivere il trasferimento di energia che avviene nell'interazione tra radiazione ionizzante e materia si utilizza la quantità chiamata dose. Nel Sistema Internazionale l'unità di misura della dose è il gray (Gy), dove un gray rappresenta l'assorbimento di un joule in un kg di materia.

La definizione della dose tramite il gray risulta però incompleta in quanto non tiene in considerazione che diverse tipologie di radiazione possono provocare effetti diversi a livello biologico. Per esempio a parità di dose un fascio

di protoni può provocare molte più rotture a livello dei filamenti di DNA rispetto ad un fascio di elettroni.

Si introduce quindi la grandezza definita dose equivalente:

$$H = w_r D \quad (2)$$

dove D è la dose e w_r è un fattore di peso che tiene conto del tipo di radiazione che rilascia la dose. La dose equivalente è misurata in sievert (Sv).

Nella *Figura 6* sono riportati valori del fattore di peso per diversi tipi di radiazione [58].

Bisogna tenere in considerazione anche che a parità di dose e radiazione l'effetto dell'esposizione ha effetti diversi a seconda dell'organo che viene interessato, avendo i vari tessuti del corpo radio-sensibilità diverse.

Viene quindi introdotta la dose efficace:

$$E = \sum_{t=0}^N w_t H_t \quad (3)$$

dove H_t è la dose equivalente e w_t è il fattore di peso che tiene conto della differente sensibilità dei vari tessuti, sommati su tutti quelli presenti nella zona irraggiata.

La *Tabella 3* riporta valori di w_t per diversi tipi di tessuto [58].

Tessuto	w_T	$\sum w_T$
Midollo osseo (rosso), Colon, Polmone, Stomaco Seno, Tessuti rimamenti *	0,12	0,72
Gonadi	0,08	0,08
Vescica, esofago, fegato, tiroide	0,04	0,16
Superficie dell'osso, cervello, ghiandole salivari, pelle	0,01	0,04
	Totale	1,00

Tabella 3: Diversi valori del fattore di peso della dose efficace per vari tessuti [58].

Tipo di radiazione	Fattore di ponderazione della radiazione, w_R
Fotoni	1
Elettroni ^a e muoni	1
Protoni e pioni carichi	2
Particelle alfa, frammenti di fissione, ioni pesanti	20
Neutroni	Funzione continua dell'energia del neutrone

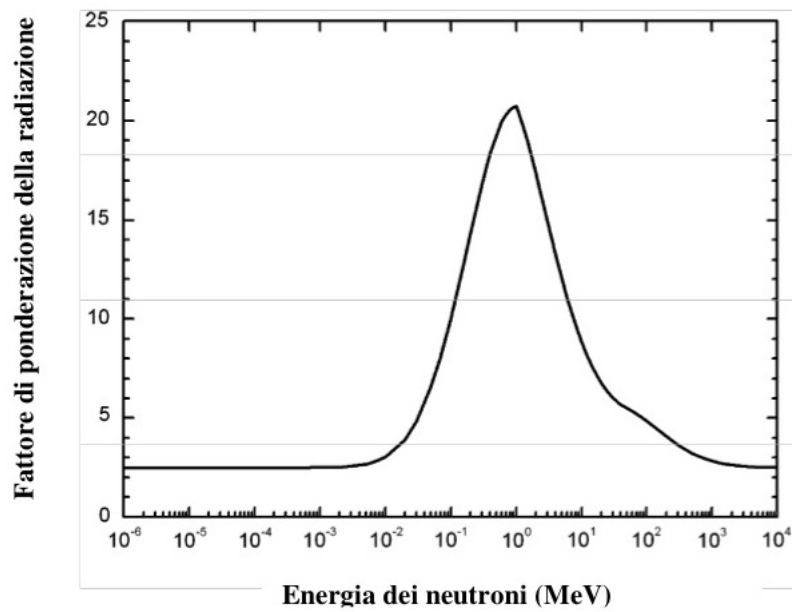


Figura 6: Valori del fattore di peso w_r della dose equivalente per diversi tipi di radiazione (sinistra) e grafico della variazione del fattore di peso in funzione dell'energia per neutroni (destra) [58].

1.6 Anatomia del polmone

I polmoni sono i principali organi deputati alla respirazione nell'essere umano e consentono di importare ossigeno nel sangue espellendo l'anidride carbonica.

Sono situati all'interno della gabbia toracica in modo simmetrico e sono composti da un materiale spugnoso rivestito da un uno strato di tessuto pleurico.

Il polmone destro è diviso in tre lobi (superiore, medio ed inferiore), mentre il sinistro in due (superiore ed inferiore). I solchi che dividono fisicamente i vari lobi sono detti scissure.

L'aria inalata a livello di bocca e naso viene condotta verso i polmoni per mezzo della trachea e portata al loro interno tramite i bronchi. La trachea è un tubo semi-rigido lungo circa 12cm formato da $15 - 20$ anelli cartilaginei sovrapposti e ricoperti da tessuto fibroso. Gli anelli sono incompleti posteriormente e sono chiusi da una membrana che consente l'espansione dell'esofago durante la deglutizione.

La trachea è chiaramente visibile negli esami CT come una struttura tubolare piena di aria dalla parte della testa del paziente.

In un punto detto carena la trachea si divide nei 2 bronchi che si dirigono verso il rispettivo polmone penetrandovi in una zona detta ilo polmonare. Il diametro del bronco di sinistra risulta essere decisamente inferiore a quello di destra a causa della presenza del cuore che, essendo per $\frac{2}{3}$ ruotato proprio verso sinistra, diminuisce il volume del polmone e il conseguente apporto di aria.

Una volta penetrati nei polmoni i bronchi iniziano una serie di ramificazioni: il primo ordine è formato da tre rami a destra e da due a sinistra, come il rispettivo numero di lobi, andando poi a diminuire nelle dimensioni all'aumentare dell'ordine fino a diventare microscopici. Tipicamente risultano esserci 23 ordini di bronchi con una relazione tra dimensione ed ordine esprimibile come:

$$d_n = d_0 2^{-\frac{n}{3}} \quad (4)$$

dove d_n è il diametro del n -esimo [34, 35].

Questa relazione vale fino al 16-esimo ordine, dopo di che le dimensioni restano circa costanti fino al 23-esimo.

Utilizzando le informazioni riportate in [35] il diametro della trachea in un adulto è di $18 \pm 2 \text{ mm}$, mentre per la primo ordine dei bronchi vale $14 \pm 1.6 \text{ mm}$. I valori possono variare a seconda di sesso, altezza ed età.

La *Figura 7* mostra in maniera schematica tutte le varie componenti della struttura polmonare descritte, mentre la *Figura 8* mostra diverse strutture anatomiche visibili all'interno di una immagine CT.

Infine, nella *Figura 9*, è mostrato un esempio di esame CT contenente un nodulo (nello specifico connesso ad un vaso) che risulta quasi indistinguibile rispetto alle strutture anatomiche presenti nell'immagine. Questa problematica è tipica degli esami CT, a causa della rappresentazione del volume per mezzo di fette bidimensionali. In sezione i noduli e le altre strutture risultano spesso molto simili, rendendo complessa la distinzione.

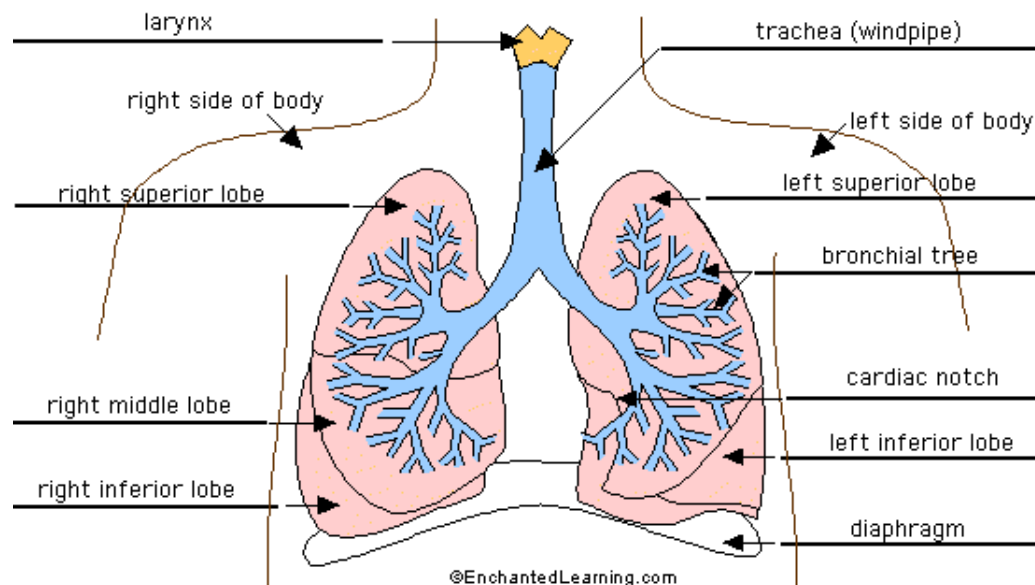


Figura 7: Rappresentazione schematica delle varie componenti della struttura polmonare nel suo complesso.

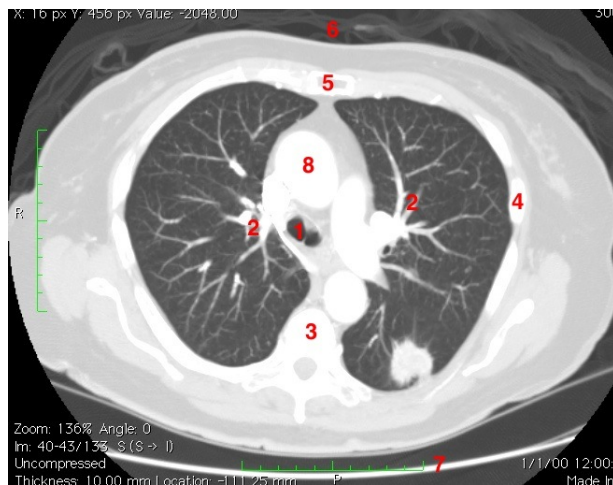


Figura 8: Alcune strutture anatomiche visibili in un esame CT: 1 inizio divisione bronchi a fine trachea (1) e prime generazioni (2), colonna (3), costole (4), sterno (5), aria esterna (6), lettino (7), cuore e strutture connesse (8).

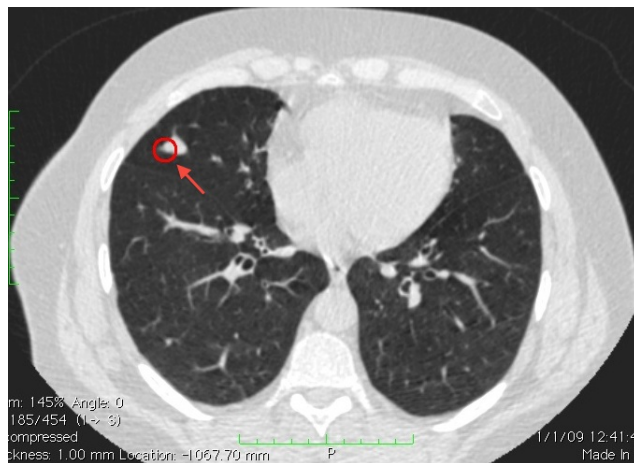


Figura 9: Immagine contenente un nodulo, all'interno del cerchio indicato dalla freccia, attaccato ad un vaso. Si può notare come la struttura nel complesso risulti molto simile alle altre presenti nel polmone, rendendo difficile l'individuazione.

1.7 CAD ed utilizzo in screening

Con CAD, o Computer Aided Detection, si fa riferimento ai programmi su calcolatore che possono affiancare il medico aiutandolo nella refertazione di tipologie di esami in cui il numero di immagini e informazioni da esaminare è molto alto, come CT, MRI ed ecografia.

In questi casi può accadere facilmente che il medico non si accorga della presenza di alcune lesioni o non riesca a valutarle correttamente rispetto alle altre strutture. Il problema è ancora più rilevante nel caso in cui si voglia andare discriminare la presenza di una patologia in uno stadio precoce del suo sviluppo.

Un programma CAD è pensato per analizzare le immagini andando a ricercare caratteristiche peculiari che consentano di identificare la presenza della patologia considerata. Al medico sono rese una serie di indicazioni sulle zone delle immagini ritenute sospette. Alcune di queste indicazioni possono essere errate, riportando la possibile presenza della patologia in punti dove invece è assente.

Al medico spetta quindi la revisione delle indicazioni ottenute, andando a valutare quali tra esse sono corrette e possono essere utilizzate per migliorare la sua refertazione originale.

Nella *Figura 10* sono mostrati due esempi di indicazioni visive rese al medico dall'analisi di un esame CT al polmone tramite differenti sistemi CAD. Nel caso specifico del tumore al polmone il programma viene istruito per la ricerca di zone con un'elevata densità e forma tendenzialmente sferica, in quanto sono caratteristiche che in prima analisi possono far sospettare la presenza di un nodulo.

Situazioni in cui l'utilizzo di un sistema CAD risulta molto interessante sono i programmi di screening e la modalità di refertazione a doppia lettura, utilizzata, nel caso del tumore al polmone, per ridurre il numero di noduli che non vengono rivelati da un solo medico.

In entrambi i casi il programma può essere utilizzato in due modalità:

1. Prima lettura: in questa modalità gli esami sono prima analizzati dal CAD ed in seguito mostrati al medico, che può solamente andare a modificare i findings segnalati dal programma. In sostanza la rivelazione della patologia è demandata al computer ed il medico assume un ruolo da revisore. Questa sarebbe la modalità più indicata per i programmi di screening, in quanto al medico vengono presentati solamente i casi

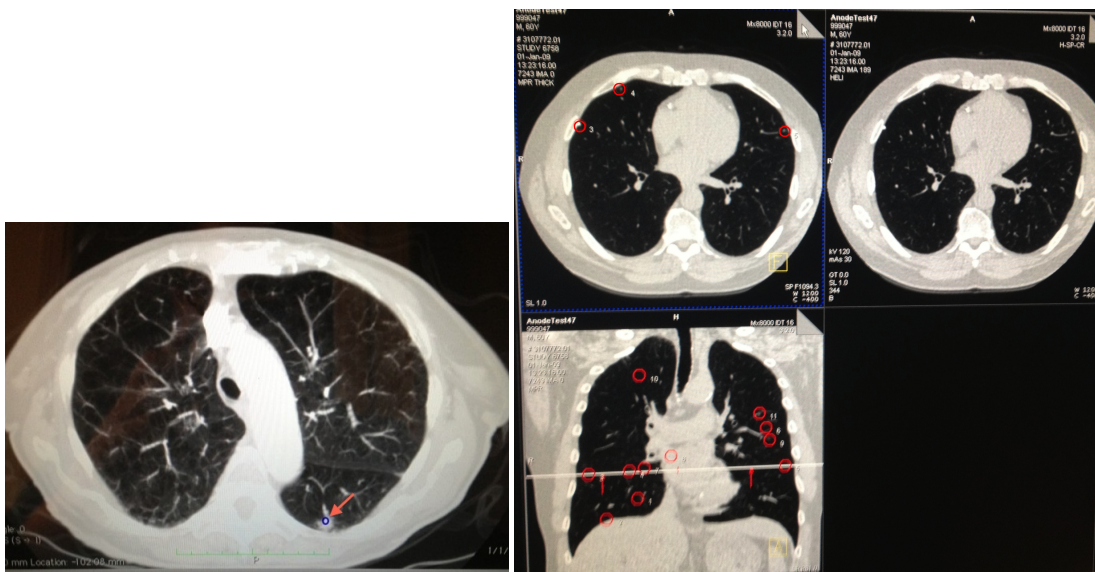


Figura 10: Esempi di visualizzazioni dei risultati dell'analisi delle immagini tramite CAD, con indicazioni su zone sospette.

sospetti e si evita che l'analisi di un gran numero di esami negativi provochi sviste e cali dell'attenzione. Di fatto però si limita la sensibilità del medico all'estremo superiore di quella del CAD. Di conseguenza questa modalità è generalmente rifiutata. Talvolta il CAD viene usato ma limitandolo al ruolo di selettore, per presentare al medico per primi gli esami che hanno maggiore probabilità di risultare positivi .

2. Seconda lettura: è la modalità più diffusa in quanto il CAD viene utilizzato solo dopo che il medico ha effettuato la refertazione in modo classico. Sulla base di quanto segnalato dal programma può poi andare eventualmente a rivalutare le proprie considerazioni. Questa modalità è più lenta della refertazione classica, ma può essere vantaggiosa se permette di aumentare la sensibilità dell'analisi. L'aumento di tempo richiesto è comunque proporzionale al numero di indicazioni errate rese dal CAD. Si noti che attualmente la doppia lettura indipendente è obbligatoria all'interno dei programmi di screening.

In conclusione un sistema CAD affiancando il medico, agevolandone il lavoro e migliorandone le prestazioni, potrebbe consentire l'impiego di un numero inferiore di personale qualificato o l'utilizzo di persone che non abbiano an-

cora raggiunto l'alto livello di esperienza richiesto per un corretto lavoro di refertazione.

1.8 Valutazione delle prestazioni di un CAD

Vi sono diversi metodi che consentono di valutare le prestazioni di un sistema CAD e permettono di confrontarle con altri sistemi o con i risultati ottenuti da un medico refertatore. E' però necessario far lavorare l'osservatore su un set di casi per cui sia disponibile un'analisi da utilizzare come riferimento per la qualità del lavoro. In medicina si definisce “*gold standard*” lo standard di riferimento rappresentato dall'esame diagnostico più accurato che permette di risolvere un dato dubbio diagnostico, e a cui ogni altra tipologia di analisi deve essere rapportata per valutarne la validità diagnostica. Molto spesso il gold standard è rappresentato dalla biopsia, in quanto permette una valutazione diretta della presenza della patologia.

Nel caso di un CAD per la ricerca dei noduli polmonari il gold standard è rappresentato da una lista dei noduli presenti negli esami annotati da radiologi esperti. Per questa patologia la biopsia non è attuabile (o comunque fortemente sconsigliata), a meno di un elevato rischio di malignità del nodulo già accertato, a causa dell'elevato rischio di collasso polmonare.

Nel proseguo del capitolo si farà riferimento come “osservatore” di un esame indipendentemente che esso sia un medico o un sistema CAD.

Prima di procedere è necessario introdurre alcune definizioni. Confrontando le annotazioni dell'osservatore e del gold standard si indica:

- Vero Positivo (TP, da true positive) un'annotazione positiva dell'osservatore corrispondente ad una positiva del gold standard
- Falso Positivo (FP, da false positive) un'annotazione positiva dell'osservatore corrispondente ad una negativa del gold standard
- Vero Negativo (TN, da true negative) un'annotazione negativa dell'osservatore corrispondente ad una negativa del gold standard
- Falso Negativo (FN, da false negative) un'annotazione negativa dell'osservatore corrispondente ad una positiva del gold standard

La *Tabella 4* riporta schematicamente quanto descritto.

annotazioni osservatore	annotazioni gold stand.	
	Positiva	Negativa
Positiva	TP	FP
Negativa	FN	TN

Tabella 4: Schematizzazione possibili risultati di un test

Introduciamo inoltre le seguenti definizioni:

$$sensibilità = \frac{TP}{TP + FN} \quad (5)$$

$$specificità = \frac{TN}{TN + FP} \quad (6)$$

Partendo dalle definizioni sopra riportate è possibile realizzare due curve che forniscono una stima delle prestazioni dell'osservatore: le curve ROC e FROC.

Per ottenere la curva ROC (Receiving Operating Characteristic) si parte da una classificazione binaria per gli esami presi in considerazione. Per esempio si chiede all'osservatore di distinguere quali esami hanno un'alta probabilità di essere positivi per cancro al polmone rispetto a quelli senza tracce di patologia. Viene inoltre richiesto di assegnare ad ogni scelta un livello di sospetto p , compreso tra zero e uno, che rappresenti un livello di confidenza sulla risposta data. Se si introduce una soglia $t \in [0; 1]$ è possibile osservare la variazione dei valori sopra descritti con $p \geq t$, per diversi valori di t .

La distinzione tra TP e FP viene effettuata confrontando la classificazione dell'osservatore con il gold standard.

La curva ROC è quindi realizzata plottando il valore della *sensibilità* rispetto a quello di $1 - specificità$ al variare di t [1].

L'area sotto la curva è una buona stima delle prestazioni dell'osservatore, poichè rappresenta la probabilità che un esame positivo scelto a caso abbia un livello di sospetto maggiore di uno negativo sempre scelto in modo casuale [24].

Data la natura binaria della sua definizione la curva ROC è adatta ad essere utilizzata in contesti in cui è possibile definire senza ambiguità la presenza di una data patologia in seguito all'individuazione di una lesione.

Per esempio se eseguendo un'ecografia al fegato viene riscontrata una lesione allora sicuramente l'organo presenta una qualche patologia ed è possibile attuare una classificazione binaria come "sano" o "malato" degli esami refertati.

Nel caso di problematiche come l'identificazione di un nodulo all'interno di un CT polmonare una classificazione di questo è insufficiente. All'interno di ogni polmone, anche di un soggetto sano, vi sono infatti noduli che non hanno caratteristiche patologiche e strutture che possono essere avere sembianze simili. La distinzione tra noduli rilevanti e non a livello patologico non è definibile in modo assoluto e varia molto da medico a medico.

Una classificazione binaria della presenza della patologia risulta quindi ambigua: due medici potrebbero diagnosticare un tumore polmonare ad un paziente pur avendo riscontrato lesioni completamente diverse. Inoltre potrebbero essere presenti più noduli rilevanti all'interno di ogni esame.

Per risolvere i problemi descritti e valutare correttamente le prestazioni dell'osservatore viene introdotta la curva FROC (Free Receiver Operating Characteristic) [1]. In questa metodologia l'osservatore non è più limitato ad indicare solamente se l'esame risulta positivo ma può indicare liberamente tutti i punti che ritiene sospetti, riportando la posizione della lesione individuata oltre al livello di sospetto.

Confrontando le annotazioni dell'osservatore con un gold standard costituito da refertazioni di radiologi esperti è possibile valutare nodulo per nodulo TP e FP. Se un'annotazione non risulta ragionevolmente vicina, definita una distanza limite, alla "reale" posizione della lesione riportata sul gold-standard, questa viene considerata come falso positivo.

La curva viene poi ottenuta in modo analogo a quanto descritto per la ROC, plottando la sensibilità rispetto al numero di FP.

In questa tesi le curve FROC saranno confrontate come proposto nell'articolo [29], rappresentando la sensibilità rispetto al rapporto FP/CT e mediando i valori corrispondenti a frazioni definite.

La *Figura 11* mostra due esempi di curve ROC e FROC.

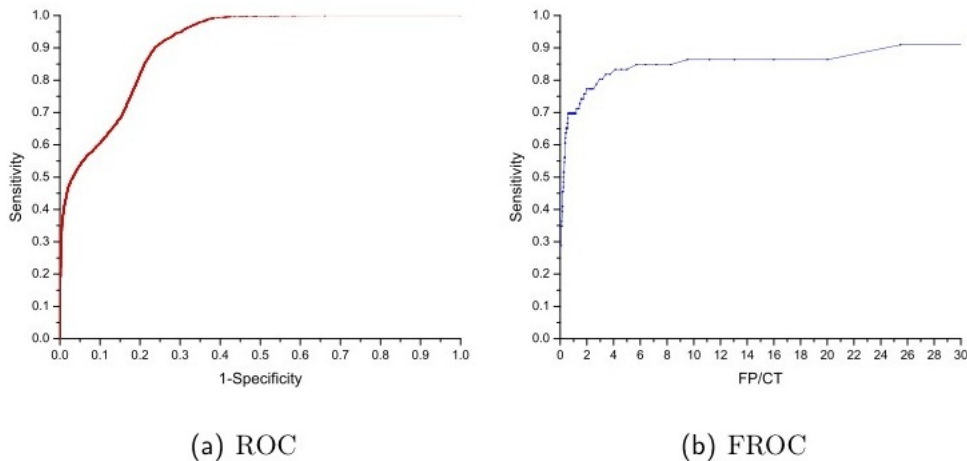


Figura 11: Esempi di curve ROC e FROC

1.9 Datasets

Nel paragrafo precedente è stato evidenziato come, per riuscire ad ottenere le curve ROC e FROC, sia necessaria una serie di esami di cui si conosce il gold standard.

Ricordiamo che nel caso del tumore al polmone questo è rappresentato da refertazioni di radiologi esperti contenenti le annotazioni delle posizioni di tutti i noduli che sono stati individuati all'interno di ogni esame. Sono solitamente utilizzate le refertazioni di più di un radiologo in quanto, come già riportato, la variabilità delle annotazioni tra i vari medici può essere molto alta.

Sviluppare una database di esami con annotazioni affidabili realizzate da radiologi esperti per il tumore al polmone risulta essere estremamente impegnativo. Non è possibile utilizzare quanto prodotto dalla normale pratica clinica principalmente in quanto gli esami non sono di screening, ma anche perchè le annotazioni relative riportano una valutazione incompleta delle lesioni presenti. Una volta che il radiologo individua un nodulo che ritiene rilevante il paziente viene comunque inviato ad ulteriori analisi. Non è quindi necessario che sia indicata la totalità delle lesioni o strutture rilevanti presenti e le annotazioni riportano posizione e tipologia dei noduli solo in maniera qualitativa.

Diversi studi hanno inoltre dimostrato come le prestazioni di un CAD

possano essere influenzate dal dataset su cui viene utilizzato [30, 31]. Ad esempio, se in quello su cui è stato ottimizzato il CAD vi è una prevalenza di tumori grandi, le sue prestazioni risulteranno ridotte se viene utilizzato su un set di esami con prevalenza di tumori di piccole dimensioni.

Per queste ragioni molto spesso per lo sviluppo di un CAD si fa affidamento sui dati database pubblici, in modo che i risultati ottenuti su questi database siano direttamente confrontabili con quelli ottenuti da altri sistemi presenti in letteratura.

Di seguito vengono riportate le caratteristiche principali di tre diversi datasets: ITALUNG-CT, ANODE09 e LIDC-IDRI.

1.9.1 ITALUNG-CT

L'ITALUNG-CT trial è stato il primo programma di screening trial randomizzato controllato per tumore al polmone in Italia [26].

Gli esami facenti parte di questo database sono stati realizzati con un protocollo di bassa dose e con fette sottili.

Nello specifico sono stati effettuati utilizzando uno scanner multislice a spirale con 4 anelli di rivelatori (Siemens Volume Zoom) con le seguenti impostazioni: $140kV$ di voltaggio, $20mAs$ di corrente di tubo, $0,6mSv$ di dose media e con $1.25mm$ di collimazione per fetta.

Tutte le immagini sono state salvate in formato DICOM.

La fette sono poi state ricostruite a $1mm$ di spessore, utilizzando un opportuno kernel (Siemens B50f). Il numero medio di fette per scansione è stato di circa 300, dove ogni fetta risulta essere una matrice $512 \times 512pixel$, con dimensioni dei pixel tra $0,53$ e $0,74mm$ e livello grigio a 12 bit in Hounsfield Units.

Ogni esame è stato refertato in maniera indipendente da 2 radiologi, che in caso di opinioni discordi hanno raggiunto un consenso. Per ogni esame viene quindi fornita una singola lista di annotazioni.

In totale sono stati coinvolti 18 radiologi, tre hanno effettuato la prima lettura e 15 la seconda. Tutti con esperienza superiore a 4 anni nell'ambito delle CT al torace.

Il CAD VBNA descritto nel seguito di questa tesi è stato in origine sviluppato utilizzando proprio questo database.

1.9.2 ANODE09

Il database ANODE09 è una raccolta di esami realizzata per fornire uno strumento di confronto diretto ed oggettivo tra le prestazioni di diversi CAD per la ricerca di noduli polmonari, con un unico protocollo di valutazione [29].

Gli esami disponibili sono 55, per la maggior parte forniti dal Nelson study (il più grande trial di screening per tumori al polmone con CT in Europa), realizzati secondo un protocollo a bassa dose e con fette sottili .

Nello specifico sono stati acquisiti tramite scanner a 4 o 16 sclices (Philips Medical System) utilizzando $30mAs$ e $120 - 140kVp$, a seconda del peso del paziente. Le singole fette, matrici 512×512 con dimensioni dei pixel tra 0.59 e $0.83mm$, sono state ricostruite con spessore di $1mm$ e $0,7mm$ di incremento tramite un opportuno kernel (Philips B). Il numero medio di slices per ogni esame è pari a 430.

Le annotazioni sono state realizzate in modo indipendente da 2 radiologi di diversa esperienza, un terzo medico è intervenuto per dirimere i casi di disaccordo. E' quindi riportata un'unica lista di annotazioni per ogni caso.

Sono state utilizzate come base le refertazioni del Nelson study a cui sono stati aggiunti gli eventuali nuovi findings trovati dai radiologi.

Durante l'analisi di ogni esame sono state annotate due categorie di oggetti: noduli rilevanti e noduli irrilevanti. Come irrilevanti sono state annotate tutte le strutture che hanno l'aspetto di un nodulo ma che il radiologo ritiene non essere tali, i noduli con caratteristiche benigne e i noduli troppo piccoli per essere considerati rilevanti. Per questo database la soglia sulla dimensione affinché un nodulo sia considerato rilevante è stata fissata a $4mm$.

Questa classificazione è stata introdotta in quanto per tali tipologie di strutture la valutazione può risultare molto diversa da radiologo a radiologo. Ne consegue che sarebbe scorretto considerare come FP findings del CAD situati in prossimità di tali punti. Alcuni radiologi anzi potrebbero apprezzare che siano segnalati dal programma.

I findings coincidenti con noduli annotati come non rilevanti non vengono quindi considerati né TP né FP, e sono perciò esclusi dalla realizzazione della curva FROC.

I noduli presenti nel database ANODE09 sono inoltre divisi in 6 categorie distinte a seconda delle loro dimensioni e della posizione, in modo che le prestazioni possano essere valutate in modo più preciso per ogni tipologia.

Le categorie sono: noduli più grandi di $5mm$, più piccoli di $5mm$, noduli isolati, pleurici, vascolari e perifissural.

L'iniziativa ANODE09 è stata presentata nel 2009 come una sorta di competizione a cui potevano prendere parte CAD sia realizzati da gruppi di ricerca che commerciali con l'obiettivo di confrontare le prestazioni fornite dai vari sistemi. Da quella data è stato sempre attivo il sito dedicato [28], da cui è possibile scaricare i 55 casi disponibili.

Di questi, 5 esami sono forniti completi delle annotazioni come esempio, in modo che possano essere usati come training per ottimizzare eventuali parametri di lavoro del CAD. I restanti 50 esami sono forniti senza annotazioni e costituiscono il set per la validazione del CAD.

Sono invece note le percentuali delle dimensioni dei noduli presenti: il 40% è sotto i $4mm$, il 40% ha diametro tra 4 e $6mm$ e il restante 20% è più grande di $6mm$.

Per quanto riguarda invece i 5 casi di training sono presenti 39 noduli rilevanti che si possono dividere, in modo visivo, in 36 interni e 3 pleurici.

I findings del CAD sui 50 casi di validazione, salvati su un file riportando posizione e grado di sospetto, possono essere caricati nuovamente sul sito sopra citato: i risultati ottenuti dal CAD sono analizzati dai curatori del programma ANODE09 e vengono notificati sotto forma di curva FROC.

1.9.3 LIDC-IDRI

Questo database è stato creato dal Lung Database Consortium e dal Image Database Resource Initiative con l'obiettivo di essere un supporto per lo sviluppo di sistemi CAD. E' stato realizzato con esami provenienti da diversi centri ed acquisiti utilizzando scanner CT di vari produttori. Attualmente risulta essere la più grande raccolta di esami annotati e pubblicamente disponibili, con 1018 esami relativi a 1010 pazienti (8 casi presentano un doppio studio)[33, 32].

A differenza dei database ITALUNG e ANODE09 gli esami raccolti non sono realizzati solamente a bassa dose e con fette sottili, ma presentano ampi intervalli di valori per tensione del tubo ($120 - 140kV$), corrente ($40 - 627mA$), spessore delle fette ($0.6 - 5.0mm$) e ricostruzione ($0.45 - 5.0mm$). Le dimensioni dei pixel spaziano da 0.461 fino a $0.977mm$. I kernel utilizzati per ricostruzione variano a seconda del produttore e della tipologia di console utilizzata.

Nel database sono stati inoltre inclusi esami in cui è possibile riscontrare altre patologie, alti livelli di rumore o artefatti, qualora queste fonti di “disturbo” non siano tali da compromettere l’individuazione dei noduli.

Le annotazioni sono realizzate complessivamente da 12 radiologi esperti appartenenti a 4 strutture diverse. Ogni struttura ha fornito una refertazione per ogni caso. La modalità di refertazione utilizzata è stata a doppia lettura con un prima fase in cui ogni radiologo ha analizzato l’esame in maniera indipendente, e una seconda in cui ha avuto accesso alle annotazioni anonimizzate dei 3 radiologi delle altre strutture.

Confrontando la propria analisi con le altre ogni refertatore ha potuto modificare o implementare le proprie annotazioni ottenendo una lista definitiva.

Si noti quindi che, a differenza di ITALUNG e ANODE09, non viene utilizzato il consenso forzato tra i radiologi: per ogni caso sono riportate 4 liste di annotazioni e non una singola. In questo modo si preserva la variabilità dell’interpretazione tra medico e medico.

Gli oggetti segnalati nelle annotazioni sono di tre tipologie:

- noduli con diametro $\geq 3mm$
- noduli con diametro $< 3mm$
- “non noduli” con diametro $> 3mm$

dove con nodulo si intende qualsiasi lesione considerata essere tale dal medico, infatti nessuna definizione a priori è stata imposta.

Sono invece indicate come “non-noduli” tutte le lesioni polmonari con dimensioni superiori ai $3mm$ che però non presentano le caratteristiche peculiari delle strutture nodulari.

Particolare attenzione è stata riservata durante la refertazione ai noduli con diametro $\geq 3mm$, in quanto sono la categoria con il più alto grado di pericolosità. Per questa categoria, dopo la seconda fase di lettura, a tutti e quattro i medici è stato chiesto di definire il contorno del nodulo e 8 caratteristiche soggettive con scala 1–5: sottigliezza, struttura interna, calcificazione, sfericità, margini, spigolosità, texture e malignità. Per le altre tipologie invece è stato chiesto di identificare approssimativamente solo il centro di massa della lesione.

Se ci si limita a propositi di ricerca dei noduli le informazioni sono ridondanti: sono infatti necessari solo il centroide e una stima della dimensione della lesione.

Considerando tutti i casi presenti nel database, il numero totale di noduli annotati (ovvero le lesioni indicate come nodulo da almeno un radiologo, indipendentemente dalla dimensione) è 7371. Di questi 2669 sono considerati avere diametro $\geq 3mm$.

Le lesioni che tutti i 4 medici hanno ritenuto essere noduli (sia $< 3mm$ che $\geq 3mm$) sono 1940, a dimostrazione di quanto possa variare la valutazione da soggetto a soggetto.

Considerando i 7371 noduli presenti nel database, nella *Figura 12* sono riportate le percentuali di quanti sono stati annotati da un radiologo, quanti da due e via dicendo. Sono inoltre riportate le analoghe percentuali per le annotazioni nel complesso, ovvero considerando sia quanto segnalato come nodulo (indipendentemente dalla dimensione) che come “non nodulo”. Le percentuale di noduli annotati da almeno due radiologi è ricavabile dalla somma delle percentuali relative a 2, 3 e 4 radiologi, quella dei noduli annotati dal almeno tre dalla somma delle percentuali di 3 e 4.

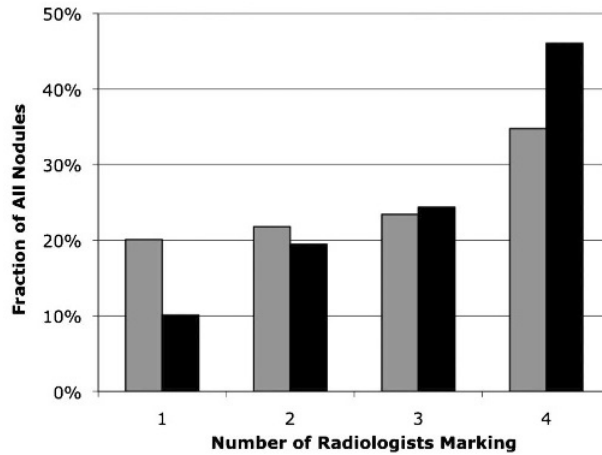


Figura 12: Distribuzione delle percentuali dei 7371 noduli annotati da diversi numeri di radiologi (in grigio) e delle annotazioni nel complesso, comprendendo anche i “non noduli” (nero) [32].

Parte II:
Descrizione CAD

2 Descrizione CAD VBNA

2.1 Introduzione

Nelle sezioni che seguono viene descritto il CAD per la ricerca di noduli polmonari VBNA.

Questo CAD, come molti altri programmi analoghi, si suddivide in tre moduli principali:

- Segmentazione
- Ricerca dei candidati noduli
- Classificazione

Nelle relative sezioni sono quindi descritti il ruolo, il funzionamento e le caratteristiche di ogni modulo di cui è composto.

2.2 Segmentazione del polmone

La segmentazione è una procedura utilizzata comunemente nell'ambito dell'analisi di immagini e consiste nella partizione dell'immagine in regioni significative. I pixel che la compongono vengono raggruppati in sottogruppi con caratteristiche comuni (colore, intensità o texture) che li distinguono gli uni dagli altri.

Considerando i pixel non singolarmente ma raggruppati si semplifica la rappresentazione dell'immagine a favore delle analisi successive. Tipicamente gli algoritmi di segmentazione vengono utilizzati per individuare e riunire tutti i pixel appartenenti ad una superficie o ad un oggetto presente nell'immagine.

Nell'ambito del CAD VBNA la segmentazione viene utilizzata per ricostruire le strutture del polmone, specialmente la pleura, che deve essere rappresentata in modo accurato per la ricerca dei noduli sulla sua superficie e per evitare che il sistema vada ad individuare punti all'esterno del polmone.

Prima di iniziare con la segmentazione vera e propria l'immagine TAC passa attraverso un processo finalizzato a rendere isotropi i voxel che la compongono. Questo processo si rende necessario in quanto le prestazioni degli algoritmi morfologici risultano migliori su immagini isotropiche, ovvero con voxel cubici, mentre quelle derivanti dalla tomografia computerizzata tipicamente non lo sono.

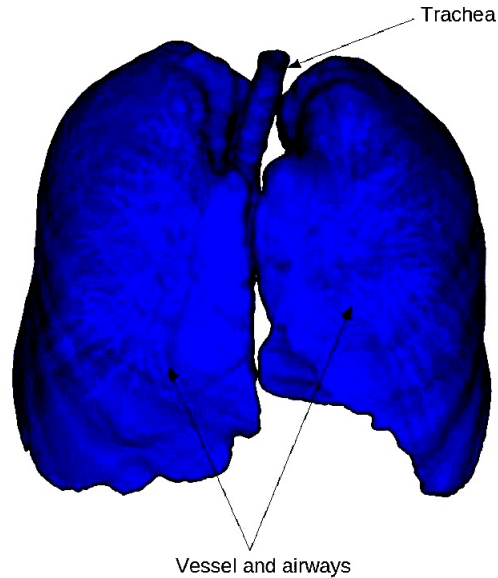


Figura 13: Ricostruzione della superficie del polmone dopo il ri-campionamento e la prima operazione di ricerca di componenti connesse. In trasparenza si possono vedere i vasi e le varie vie aeree.

Si è scelto quindi di ri-campionare l'immagine in modo da portare a 1mm le dimensioni dei voxel. Per eseguire correttamente il ri-campionamento è stato applicato un filtro gaussiano tridimensionale (ovvero uno "smoothing") con ampiezza proporzionale alle dimensioni originali dell'immagine [39, 38].

Si procede quindi cercando le regioni a bassa intensità all'interno dell'immagine tramite l'utilizzo di due soglie a -600HU e -1000HU , seguite da un'identificazione delle componenti connesse in tre dimensioni. La più grande componente che non giace sul bordo del volume rappresenta il polmone con la trachea, le vie aeree più grandi e l'aria in esse contenuta.

Partendo da questa componente viene creata una maschera dove i voxel a bassa intensità sono visualizzati in primo piano ed il resto è considerato come sfondo. Il risultato è mostrato nella *Figura 13*.

Poiché la trachea funge ancora da ponte tra i due polmoni va rimossa insieme alle altre vie aeree maggiori in modo da poterli analizzare separatamente.

Per segmentare la trachea e le altre vie in modo che sia possibile rimuoverle dalla maschera appena ricavata si utilizza il fatto che sono volumi

contenenti aria.

Utilizzando una decomposizione a due componenti ed una soglia opportuna è possibile quindi distinguerle dalle altre strutture [36].

La soglia adatta viene ricavata andando a ricercare il numero di voxel contenenti aria all'interno della maschera, ricavabile considerando le intensità e il numero di voxel dell'aria, delle restanti strutture e i valori medi su tutta l'immagine. Si ottiene quindi la soglia T_{Air} come

$$T_{Air} = \underset{T \in [T_{min}, T_{max}]}{max} \{n_T \leq n_{Air}\} \quad (7)$$

dove n_T è il numero dei voxel all'interno della maschera che hanno intensità $\leq T$ mentre T_{min} e T_{max} sono gli estremi dell'intervallo di intensità.

Identificata la T_{Air} si procede ad un'operazione di soglia tra T_{Air} e T_{min} seguita da una analisi delle componenti connesse in modo da scartare i voxel di aria isolati.

Per selezionare la componente connessa rappresentante la trachea si cerca una struttura circolare nei pressi della testa del paziente, la quale rappresenta la sezione della trachea stessa. La ricerca viene effettuata tramite un algoritmo che utilizza una trasformazione di Hough, capace di riconoscere diverse forme quali cerchi, ellissi ecc ecc [37].

Identificata correttamente la trachea si procede ad applicare un operatore di dilatazione con kernel di 2 mm per tenere anche conto delle pareti ad alta intensità delle vie aeree. La maschera della trachea così ottenuta è sottratta alla maschera complessiva ricavata sopra.

Se i polmoni risultano ancora non separati al termine di questa fase si procede cercando, per ogni fetta, il percorso che si snoda dalla parte anteriore a quella posteriore del paziente massimizzando i valori di HU dei punti che lo compongono. Questa metodologia si basa sull'ipotesi che le componenti che dividono i due polmoni abbiano intensità più alte rispetto ai polmoni stessi.

Ottenuto il percorso per ogni fetta lo si sottrae alla maschera senza la trachea e si procede quindi ad un'analisi delle componenti connesse con connettività a 6 - *pixel*.

Come si può osservare nella *Figura 14*, in questo stadio le pareti delle vie aeree sono ancora escluse dalla maschera ottenuta. Vanno quindi aggiunte facendo però attenzione a non modificare la morfologia della superficie della pleura, ovvero senza andare a modificare la forma di possibili noduli situati su di essa.



Figura 14: Segmentazione del polmone prima della rimozione delle pareti delle vie aeree e vasi (a) e dopo la loro rimozione (b).

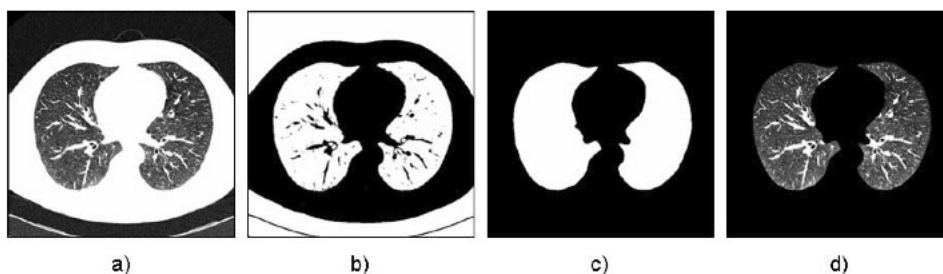


Figura 15: Fasi della segmentazione tridimensionale del polmone: immagine di input (a), dopo il processo di tresholding (b), dopo la ricerca delle componenti connesse (c), output dell'algoritmo (d).

Per questo scopo si utilizza una combinazione di operatori morfologici quali uno di erosione e uno di dilatazione seguiti da un operatore logico OR tra la maschera così ottenuta e quella iniziale. La maschera definitiva ottenuta comprende le pareti delle varie vie pur mantenendo la conformazione della superficie come richiesto. La *Figura 15* mostra un riepilogo delle varie fasi del processo.

La pleura è quindi definita considerando la superficie che divide il parenchima dal resto dell'immagine (*Figura 16*).

Al termine di tutta la procedura di segmentazione descritta è stato implementato un controllo automatico per verificare eventuali errori, che possono accadere a causa di differenza anatomiche dei vari pazienti.

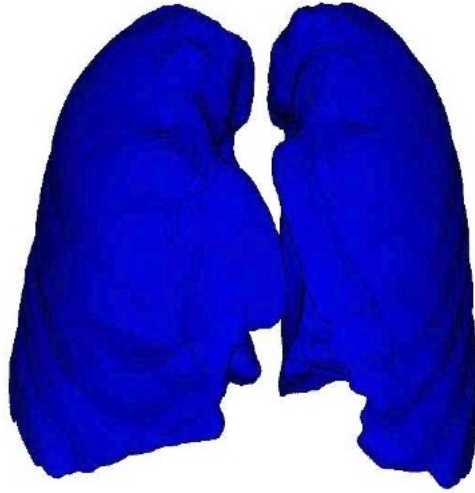


Figura 16: Rappresentazione della maschera finale della segmentazione del polmone.

2.3 Ricerca noduli interni (CAD_I)

Passiamo ora alla descrizione della parte di codice deputata della ricerca dei noduli presenti all'interno del polmone, distinta da quella a cui invece viene affidata la ricerca dei noduli presenti sulla superficie della pleura.

L'individuazione dei noduli risulta essere la parte cruciale di un sistema CAD in quanto, per non limitarne a priori le prestazioni, si deve cercare di avere una sensibilità più prossima possibile al 100%.

Il punto di partenza per questo algoritmo è considerare i noduli con simmetria sferica, aventi cioè un profilo gaussiano tridimensionale, come proposto in [40]. La *Figura 17* mostra un esempio di nodulo interno.

E' stato quindi implementato un filtro "*dot-enhancement*" basato sul calcolo degli autovalori della matrice Hessiana che consente di distinguere tra strutture sferiche, cilindriche e planari.

L'utilizzo diretto della matrice Hessiana senza uno smoothing comporterebbe un aumento del rumore, dato il necessario utilizzo di derivate.

Le caratteristiche locali di ogni voxel sono quindi determinate tramite la funzione

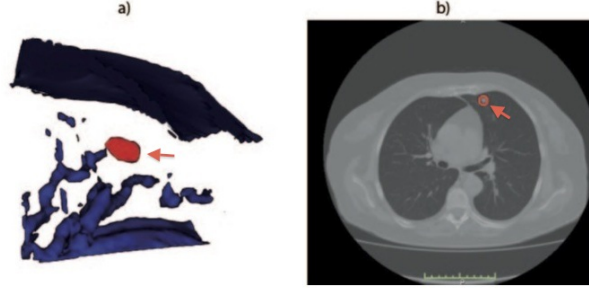


Figura 17: Rappresentazione di un nodulo interno: ricostruzione di una parte interna del polmone con nodulo evidenziato in rosso (a), sezione di un esame CT contenente nodulo interno (b).

$$z_{dot}(\lambda_1, \lambda_2, \lambda_3) = \begin{cases} |\lambda_3|^2/|\lambda_1|, & \text{se } \lambda_1 < 0, \lambda_2 < 0, \lambda_3 < 0 \\ 0, & \text{altrimenti} \end{cases} \quad (8)$$

dove $\lambda_1, \lambda_2, \lambda_3$ sono gli autovalori della matrice Hessiana ordinati in modo tale che $|\lambda_1| > |\lambda_2| > |\lambda_3|$. I segni sono tali da distinguere strutture ad alta intensità su di uno sfondo a bassa intensità.

E' possibile dimostrare che z_{dot} si annulla per strutture planari o cilindriche mentre per strutture sferiche con profilo gaussiano vale $z_{dot} \sim \frac{1}{\sigma^2}$.

Poiché i noduli presenti all'interno degli esami presentano dimensioni diverse è necessario implementare in questo tipo di analisi un approccio multiscala.

Si combina quindi la funzione z_{dot} con uno smoothing gaussiano a più scale [40, 41]. Partendo dalla conoscenza dell'intervallo di dimensioni dei noduli presenti all'interno del set di esami (qualora sia un'informazione accessibile) ed assumendo di approssimare il nodulo con una gaussiana 3D con parametro di scala σ , considerando 4σ è possibile comprendere il 95% del volume del nodulo.

Se il valore dei diametri dei noduli varia nell'intervallo $[d_{min}, d_{max}]$, la scala dello smoothing varia all'interno di $[\sigma_{min}, \sigma_{max}]$, dove $\sigma_{min, max} = s \cdot d_{min, max}/4$, con $s = \sqrt{\frac{3}{2}}$ fattore di normalizzazione [42].

All'interno degli intervalli descritti si considerano quindi N scale intermedie con $\sigma_i = r^{i-1} \sigma_{min}$ con $i = 1 \rightarrow N$ e $r = (d_{max}/d_{min})^{1/(N-1)}$.

Il filtro definitivo risulta essere

$$z_{max} = \max(\sigma_i^2 \cdot z_{dot}(\sigma_i)) \quad (9)$$

con $i = 1 \rightarrow N$.

Un algoritmo per la ricerca di picchi è utilizzato per trovare i massimi locali nella matrice output del filtro.

Al termine di tutta la procedura l'output complessivo è una lista di candidati noduli ordinata secondo il valore di z_{max} . La lista contiene anche un buon numero di falsi positivi, oltre ai noduli veri e propri.

2.4 Ricerca noduli pleurici (CAD_{JP})

La ricerca dei noduli situati sulla superficie della pleura si divide in 2 parti: prima la superficie stessa viene rappresentata attraverso un mesh poligonale ed in seguito i noduli vengono ricercati considerando le regioni in cui si accumulano le intersezioni tra le normali alla superficie.

Come scritto nel paragrafo precedente la pleura è definita come la superficie separante il parenchima dalle altre strutture anatomiche presenti nel resto dell'immagine. Prima di procedere risulta necessario rappresentarla in modo opportuno per la ricerca dei candidati noduli che si possono trovare sulla sua superficie.

Si utilizza quindi un algoritmo per trasformare la maschera binaria in un mesh polinomiale, procedura comunemente utilizzata nell'ambito della rappresentazione di oggetti 3D in computer grafica. Con mesh si intende la raccolta di spigoli, vertici e facce che definisce la forma di un oggetto poligonale in tre dimensioni.

Il mesh può essere fatto utilizzando triangoli, quadrati e altre forme semplici; in questo caso specifico è stato utilizzato un mesh triangolare.

Uno degli algoritmi più comunemente utilizzati è il cosiddetto "*marching cube algorithm*" [43]: data la superficie che si vuole rappresentare questo algoritmo ne seleziona una porzione in un cubo immaginario, costruito considerando le posizioni di 8 voxel vicini, e ricerca quali poligoni sono necessari per rappresentare la superficie passante nel cubo. I singoli poligoni vengono quindi uniti per creare la superficie considerata.

La ricerca dei poligoni viene fatta creando un indice ad una lista pre-costituita di 256 possibili configurazioni all'interno di un cubo: ad ognuno degli 8 voxel è assegnato un bit all'interno di un numero a *8bit* ($2^8 = 256$). Se il valore di HU del voxel è superiore a quello della superficie che attraversa

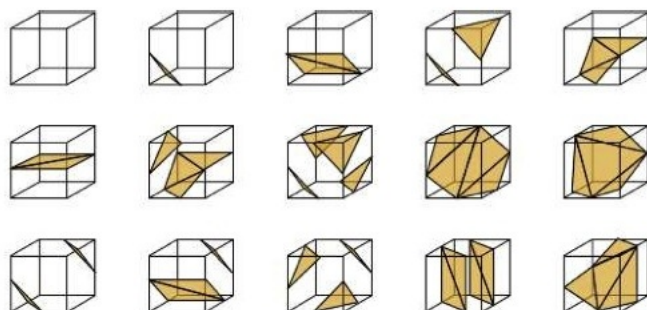


Figura 18: Schema delle 15 combinazioni di poligoni uniche utilizzate nell’algoritmo “marching cube” [61].

il cubo il bit è impostato a 1, altrimenti a 0. La sequenza specifica ottenuta da ogni cubo indica quale configurazione sia necessario utilizzare.

Se si considerano le possibili rotazioni e riflessioni il numero di possibili configurazioni può essere ridotto a 15 casi unici, come mostrato nella *Figura 18*.

Rappresentata la superficie della pleura è possibile passare alla ricerca dei noduli.

Essendo questi sostanzialmente strutture semisferiche attaccate alla pleura, come mostrato nella *Figura 19*, l’idea utilizzata è quella di tracciare le normali alla superficie e ricercare i punti in cui vi è un accumulo di intersezioni.

Le normali sono calcolate partendo dal mesh triangolare della superficie utilizzando una procedura chiamata “Gouraud shading” [44]. La normale di ogni triangolo viene calcolata facendo il prodotto vettoriale tra i suoi bordi, quindi la normale di ogni vertice è calcolata dalla media di quelle dei poligoni che vi convergono.

Poichè la ricerca delle intersezioni in tre dimensioni può risultare molto onerosa a livello computazionale, questa è stata implementata a livello dei voxel. In questo modo ogni voxel raccoglie un risultato pari al numero di normali che lo attraversano. I risultati sono raccolti all’interno della matrice di score $S(x, y, z)$ come mostrato nella *Figura 20*.

Per ovviare agli errori numerici che si possono generare utilizzando le normali nello spazio dei voxel è preferibile utilizzare cilindri Gaussiani invece che segmenti lineari. Questa operazione risulta equivalente ad operare uno

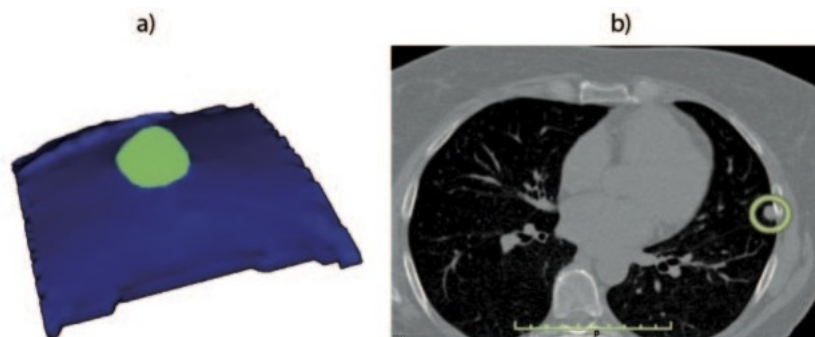


Figura 19: Rappresentazione di un nodulo pleurico: rendering della superficie con nodulo evidenziato in verde (a), sezione estratta da un esame CT (b).

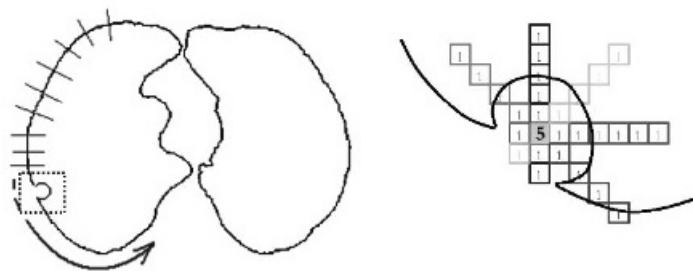


Figura 20: Rappresentazione della matrice score $S(x, y, z)$ con le intersezioni delle normali.

smoothing gaussiano con scala $\sigma_{cilindro}$ sulla matrice $S(x, y, z)$ [45]:

$$A(x, y, z) = \int \int \int S(x', y', z') \frac{e^{-\left[\frac{(x-x')^2 + (y-y')^2 + (z-z')^2}{2\sigma_{cilindro}^2}\right]}}{(2\pi \cdot \sigma_{cilindro})^3} dx' dy' dz' \quad (10)$$

Al termine di questa procedura è possibile ottenere una lista di candidati noduli cercando i massimi locali all'interno della matrice $A(x, y, z)$.

Come nel caso precedente la lista conterrà anche molti falsi positivi che andranno ridotti con la procedura descritta nel prossimo paragrafo.

2.5 Riduzione falsi positivi

Come messo in evidenza al termine delle descrizioni dei moduli per la ricerca dei noduli interni e pleurici, nelle liste di noduli output delle procedure vi è un grande numero di falsi positivi. Per ottimizzare la prestazione del CAD e limitare il tempo necessario ad un radiologo per visionare i risultati tale numero deve essere ridotto il più possibile, preservando però i findings corretti.

La selezione di quali candidati noduli siano da mantenere e quali da eliminare viene effettuata tramite l'utilizzo di un classificatore [55].

Il punto fondamentale per l'efficace utilizzo di un classificatore è la definizione di un numero finito di informazioni, chiamate features, che possano sintetizzare le informazioni principali contenute nell'immagine.

L'approccio utilizzato è quello del Voxel Based Neural Approach ([47, 49]), che consiste nell'assegnare ad ogni voxel di ogni candidato nodulo un vettore contenente features quali livelli di grigio, informazioni morfologiche dei voxel confinanti, informazioni tratte dai processi di ricerca dei noduli e informazioni sulla ROI.

L'idea è di fornire al classificatore informazioni che permettano l'apprendimento di correlazioni non evidenti tra i vari dati.

Partendo dalle liste dei massimi locali delle matrici $Z(x, y, z)$ e $A(x, y, z)$ si utilizza una soglia per ciascuna in modo da ottenere solo i findings con più alta probabilità di essere veri noduli. Inoltre una segmentazione di massima viene effettuata per eliminare i findings con dimensioni troppo piccole o troppo grandi.

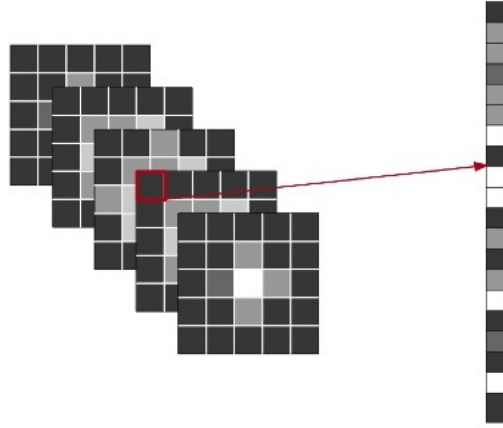


Figura 21: Esempio di costruzione del vettore dei livelli di grigio.

Per quanto riguarda le features legate ai livelli di grigio, per ogni voxel di una ROI si inseriscono in un vettore i livelli dei due pixel confinanti in ogni direzione in $3D$, per 125 valori complessivi, come mostrato nella *Figura 21*.

Le features legate alle caratteristiche morfologiche, per ogni voxel, sono ricavate utilizzando gli autovalori di matrice hessiana e gradiente. La necessità di utilizzare uno smoothing prima di procedere con i vari ordini di derivate viene soddisfatta utilizzando una scala pari al valore di z_{max} per i noduli interni e pari a $\sigma_{cilindro}$ per quelli pleurici. I sei valori ricavati per ogni voxel esprimono quindi le caratteristiche del voxel e di quelli confinanti e vengono aggiunti alle features già trovate.

Vengono inoltre aggiunti i valori, relativi ad ogni voxel, estratti dalle matrici $Z(x, y, z)$ e $A(x, y, z)$. Le features sono completate dall'aggiunta del volume in mm^3 del candidato.

Si noti che tutte le features fanno riferimento a informazioni ricavate dai voxel appartenenti ad ogni candidato. La segmentazione degli stessi è quindi un passaggio necessario per escludere i voxel non appartenenti alla strutture di interesse.

Il vantaggio del metodo VBNA deriva dal fatto che nessuna feature legata alla forma della ROI viene utilizzata per il classificatore. In questo modo non è necessario avere una segmentazione molto accurata: finché la differenza tra la segmentazione “ottimale” del nodulo e quella ottenuta non sono troppo differenti i risultati sono praticamente inalterati. Al contrario, se

fossero utilizzate features legate alla forma, piccole differenze nella segmentazione potrebbero comportare valori completamente diversi delle informazioni raccolte e conseguenti differenze di classificazione.

Definite le features, i vettori vengono analizzati tramite una SVM (support vector machines) [55, 56], utilizzata per eseguire la classificazione. Le SVM sono modelli per l'apprendimento supervisionato utilizzati per la classificazione.

A livello di classificazione, ricevendo in ingresso una serie di dati relativi ad un oggetto, sono in grado di eseguire una classificazione binaria e separare gli oggetti esaminati in due classi distinte. L'addestramento, o training, di una SVM viene effettuato fornendo alla macchina una serie di esempi di cui si conosce la classificazione: un algoritmo di apprendimento apposito permette di ricavare un modello che assegni ogni nuovo caso ad una delle due categorie.

Per una SVM un modello è una rappresentazione degli esempi come punti nello spazio, dove la posizione di ogni punto e tale da consentire la migliore separazione tra i punti delle due classi. Ogni nuovo caso viene quindi classificato in base a dove si colloca il punto che lo rappresenta nello spazio rispetto ai gruppi di punti degli esempi.

Oltre ad eseguire la classificazione una SVM può assegnare ad ogni caso analizzato un punteggio in base a come si colloca nello spazio. Il punteggio è calcolato utilizzando una funzione "sigmoide", del tipo

$$f(x) = \frac{1}{1 + e^{-ax}} \quad (11)$$

dove a è un parametro e x è la distanza con segno dall'iperpiano che rappresenta la migliore separazione.

Nel caso specifico del CAD VBNA la SVM riceve in ingresso il vettore contenente le features di ogni voxel appartenente ad un candidato nodulo. Analizzando questi dati per ogni candidato nodulo, la SVM riesce a separare una parte dei falsi positivi dal resto findings.

Ad ogni finding associa inoltre un "livello di sospetto", ottenuto come media del punteggio assegnato a tutti i voxel che lo compongono, che rappresenta la probabilità di rilevanza del candidato nodulo.

Il training è stato eseguito separatamente per il CAD_I e il CAD_{JP} utilizzando 69 casi del database LIDC, contenenti 96 noduli rilevanti interni e 42 pleurici.

2.6 Combinazione CAD_I e CAD_{JP}

Il CAD_I e il CAD_{JP} sono complementari uno all'altro e molto spesso, più che alle liste separate per le due categorie di noduli, si ha interesse ad avere una lista complessiva di findings. E' necessario quindi combinare i risultati, e l'operazione risulta non banale in quanto la semplice fusione (o "merging") dei findings porterebbe a risultati non controllabili. Avendo algoritmi diversi il comportamento dei due CAD può differire da caso a caso e con esso, per esempio, può variare il livello di probabilità assegnato ai findings. Facendo un semplice merging la lista di findings complessiva sarebbe dominata sempre dai candidati noduli con probabilità più alta.

L'approccio utilizzato, descritto in [57, 29, 47], si compone di due parti: normalizzazione e clustering. Può anche essere usato per combinare i risultati di due CAD in generale.

La normalizzazione serve a rendere direttamente confrontabili i risultati di CAD diversi. Il grado di sospetto di ogni finding viene sostituito con un nuovo valore calcolato come

$$p \rightarrow f(p) = \frac{TP(p)}{TP(p) + FP(p) + 1} \quad (12)$$

dove $TP(p)$ e $FP(p)$ sono i veri positivi e i falsi positivi con livello di sospetto $\geq p$. Si noti che senza il fattore +1 la 12 sarebbe il valore predittivo positivo con soglia p .

Per come è definita, la funzione di normalizzazione deve prima essere calcolata su di un database per cui si ha un gold standard, successivamente può essere utilizzata anche su datasets diversi.

Normalizzati i findings dei CAD è possibile combinarli facendo un clustering di quelli che si trovano più vicini di una distanza predefinita e sommando i relativi valori di $f(p)$.

3 Descrizione CAD CAM

3.1 Introduzione

Nelle sezioni che seguono viene descritto il CAD CAM dedicato, come il VBNA, alla ricerca di noduli polmonari. Nel seguito sarà poi discussa la combinazione di questi due sistemi e il loro utilizzo nell'ambito di un servizio web.

Anche il CAD CAM si suddivide in tre moduli principali:

- Segmentazione
- Ricerca dei noduli
- Classificazione

Nelle relative sezioni sono quindi descritti il ruolo, il funzionamento e le caratteristiche di ogni modulo di cui è composto.

3.2 Segmentazione polmone e strutture interne

Il CAD Channeler Ant Model (CAM) descritto in questo capitolo è stato sviluppato, come il VBNA, nell'ambito del progetto MAGIC5 dall'INFN, ma si distingue da quanto descritto nel capitolo precedente per le modalità con cui viene segmentato il polmone e per la classificazione dei noduli. La segmentazione delle strutture interne al polmone, e quindi anche degli stessi noduli, avviene generando in ogni voxel che abbia determinate caratteristiche una colonia di formiche. La vita e l'evoluzione della colonia avviene secondo regole precise e al termine della procedura si ottiene una mappa di ferormone che consente di ricavare le ROI.

Un descrizione più dettagliata di quanto esposto in questo capitolo è riportata in [29, 47] e la *Figura 22* mostra schematicamente il diagramma a blocchi dei vari stadi di elaborazione.

Prima di iniziare la segmentazione tramite le formiche le immagini passano attraverso due stadi paralleli, che hanno come punto di partenza le immagini in formato DICOM. Nel primo ogni fetta delle immagini CT viene processata attraverso un filtro di Savitzky-Golay che consente di diminuire il rumore senza una perdita in risoluzione [50, 51]. Questo filtro è l'unica componente del CAD CAM ad essere bidimensionale, tutto il resto è intrinsecamente 3D.

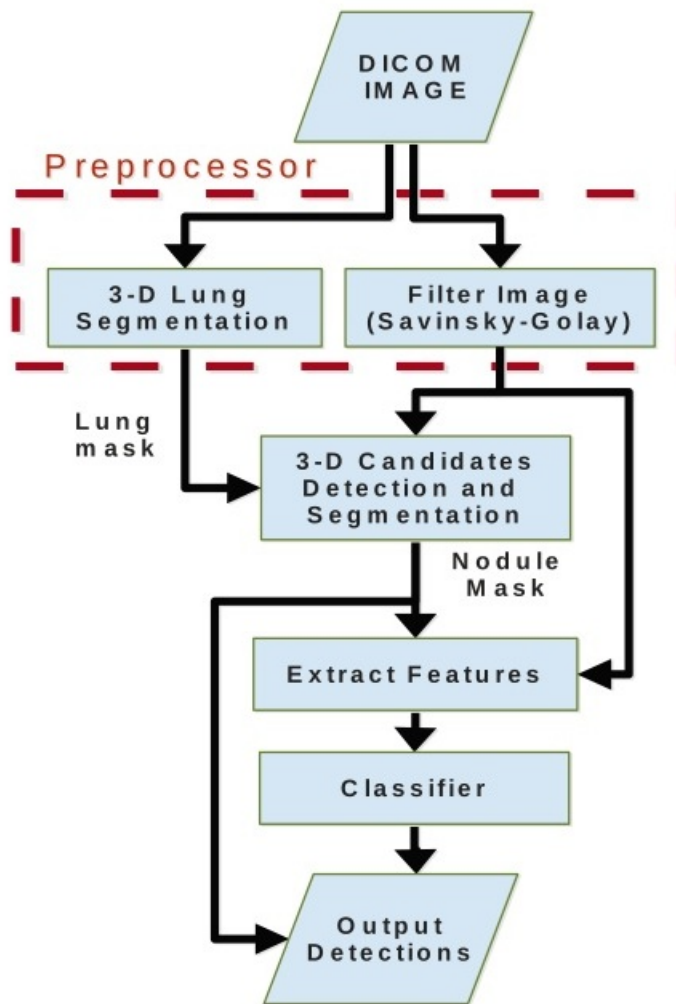


Figura 22: Diagramma a blocchi del CAD CAM.

Nel secondo il polmone viene segmentato con una soglia opportuna ricercata tramite l'analisi della distribuzione dei valori HU ed utilizzando un procedura di “*region-growing*” [53]. Successivamente un algoritmo di “*wave-front*” si occupa di ricercare la superficie della pleura dall'interno e rimuovere trachea e bronchi. Infine un operatore morfologico è usato sull'esterno del polmone per ricercare i noduli pleurici e chiudere i “buchi” lasciati dalla segmentazione dei vasi e delle vie aeree.

Come è stato anticipato, la segmentazione dei noduli avviene tramite l'utilizzo di colonie di formiche virtuali. Questo metodo, pensato per la segmentazione di strutture con forme ed intensità diverse in ambienti 3D, è stato scelto per la capacità delle formiche virtuali di modificare l'ambiente in cui si trovano e comunicare tra loro tramite la deposizione di ferormone.

La vita delle formiche è costituita da una serie di time step atomici in cui le formiche si muovono, rilasciano ferormone, si nutrono, variano la loro energia e muoiono secondo regole predefinite [52]. Nello stesso articolo è riportata la validazione di questo metodo su una serie di oggetti 3D artificiali.

La probabilità che un voxel dell'immagine diventi destinazione delle formiche è definita come

$$P_{ij}^1(v_i \rightarrow v_j) = 1 - P_{ij}(v_i \rightarrow v_j) + P_{ij}^{Int}(v_i \rightarrow v_j) \quad (13)$$

con

$$P_{ij}(v_i \rightarrow v_j) = \frac{W(\sigma_j)}{\sum_{n=1,26} W(\sigma_n)} \quad (14)$$

dove $W(\sigma_j)$ dipende dalla quantità di ferormone nel voxel v_j e $P_{ij}^{Int}(v_i \rightarrow v_j)$ è lo stesso $P_{ij}(v_i \rightarrow v_j)$ con l'intensità del voxel I_j al posto di $W(\sigma_j)$.

L'espressione (14) era la regola originalmente utilizzata per la probabilità che un voxel diventi destinazione delle formiche. Il passaggio a (13) è stato effettuato per velocizzare l'esplorazione del polmone favorendo come destinazione voxel con livelli di ferormone bassi ma HU alti, ovvero che hanno ricevuto poche visite. In questo modo l'espansione nell'ambiente 3D avviene più velocemente.

Il massimo numero di visite che un voxel può ricevere è stato definito come

$$N_\nu(I_j) = N_{min} \left(1 + \frac{I_{max} - I_j}{I_{max}} \right) \quad (15)$$

dove I_{max} è il valore massimo dell'intensità nel polmone e N_{min} è un parametro libero fissato a 5.

Un punto cruciale del CAD è che le formiche abbiano la capacità di rilasciare in noduli piccoli e di bassa intensità abbastanza ferormone perché vengano rivelati applicando una soglia. Poiché il ferormone rilasciato come funzione dell'intensità del voxel è una funzione fissata, bisogna lavorare sulla variazione di energia, che definisce quanti step possono fare le formiche prima che l'energia scenda sotto il livello di morte.

Inoltre per noduli molto piccoli il movimento iniziale può portare le formiche al di fuori della struttura, provocando un'estinzione prematura della colonia.

La variazione dell'energia è stata quindi definita come

$$\epsilon_{t+1}^k - \epsilon_t^k = -\alpha \left(1 - \frac{\Delta_{ph}^k}{\langle \Delta_{ph} \rangle} \right) \quad (16)$$

dove $\epsilon_{t,t+1}^k$ indica l'energia della formica k ai time-step t e $t + 1$, Δ_{ph}^k è il valore di ferormone rilasciato dalla formica nello step attuale e $\langle \Delta_{ph} \rangle$ il valore medio di ferormone rilasciato dalle formiche per ogni step dall'inizio dell'evoluzione della colonia. Il parametro α e il valore dell'energia iniziale ϵ_0 vengono utilizzati per bilanciare il rapporto tra morti e nascite e vengono fissati al momento della generazione della colonia. Se la colonia supera le 100 unità allora i parametri vengono impostati a valori di default per evitare la crescita esponenziale. Si ottiene in questo modo una migliore segmentazione delle strutture più piccole senza intaccare quella delle più grandi.

Se il numero di formiche sale sopra 1000 allora il valore $\langle \Delta_{ph} \rangle$ in (16) viene sostituito con il valore più probabile di rilascio di ferormone, valutato per la singola colonia. Questo risulta più stabile della media in quanto non affetto da bassi valori dovuti all'esplorazione dei voxel a bassa intensità nelle vicinanze.

Le formiche sono applicate in maniera iterativa ad entrambi i polmoni. La prima colonia esplora le vie aeree partendo da un formicaio posto in prossimità dell'inizio della trachea: trovato un voxel che soddisfa le condizioni per divenire formicaio, le formiche agiscono secondo le varie regole fino a che non sono tutte morte. I voxel toccati dalla colonia estinta sono tolti dall'immagine e il ciclo viene ripetuto in maniera finché all'interno di entrambi i polmoni non rimane nessun voxel in grado di soddisfare le condizioni

per essere formicaio. Si procede quindi a analizzare la mappa di ferormone ottenuta.

3.3 Ricerca dei noduli

Per analizzare la mappa di ferormone si ricerca in maniera iterativa una soglia opportuna per distinguere le ROIs relative a candidati noduli. Ogni voxel con un contenuto di ferormone superiore ad un valore minimo viene utilizzato come seme per un algoritmo di region-growing. Il valore della soglia è abbassato ad ogni passo dell'espansione e viene scelta come soglia migliore (T_{ph}) quella che garantisce la crescita minore. Si scelgono quindi le regioni connesse con valori superiori a T_{ph} , ottenendo la migliore separazione tra segnale e background.

Poichè il 20% circa di noduli polmonari rilevanti sono segmentati con porzioni di strutture vascolari a cui sono connesse, e non passerebbero i successivi stadi di classificazione, è stata implementata una procedura apposita per queste strutture. Ogni zona della mappa di ferormone con raggio superiore a $10mm$ (valore trovato empiricamente, in corrispondenza del quale si iniziano ad osservare variazioni delle features) viene analizzata alla ricerca di strutture semisferiche. Ogni voxel di queste strutture viene mediato con i voxel confinanti all'interno di una sfera di raggio R e nuovamente sottoposto alla soglia T_{ph} . Le strutture con dimensioni minori di R di conseguenza svaniscono, mentre i noduli risultano rimpiccioliti.

Per recuperare una parte della dimensione originale si ripete sostanzialmente la procedura considerando sfere di raggio $R/2$ e soglia $T_{ph} + T_{ph}/3$: i punti sopra a questa soglia sono reintegrati all'interno dell'immagine. Si itera la procedura tre volte con sfere di raggio crescente per aumentare mano mano la dimensione e l'output dei tre passaggi è combinato per ottenere la maschera definitiva del nodulo.

3.4 Features

Uno dei problemi più comuni che affligge i sistemi CAD è “*l'overtraining*”, ovvero la scelta di utilizzare un grande numero di features su un set di casi di training relativamente piccolo. Se l'incremento del numero di informazioni teoricamente permette una migliore separazione tra le strutture che sono noduli e quelle che non lo sono, per contro rischia di far perdere generalità al sistema di classificazione. In un dataset di training piccolo le features

estratte saranno molto specifiche e il classificatore tenderà a riconoscere bene solo quella data configurazione di caratteristiche, che difficilmente è costante su un insieme di casi più esteso. Di conseguenza le prestazioni saranno molto buone su quel dataset ma diminuiranno appena il numero di casi su cui si usa il sistema aumenta.

Per evitare questo problema il CAD CAM utilizza 13 features per la ricerca dei noduli, appartenenti a queste categorie:

- features 3D che, essendo invarianti per rotazioni e traslazioni, consentono di distinguere strutture sferiche da ROI invece appartenenti a vasi o parenti del polmone
- features basate sui livelli HU dei voxel (valori che non sono stati utilizzati per la ricerca dei noduli)
- frazione dei voxel della ROI connessi con il polmone: permette di distinguere tra i noduli interni e pleurici, che hanno forme diverse, utilizzando la stessa rete neurale per entrambi

Poichè tutto il CAD CAM opera in 3D non vengono utilizzate features bidimensionali.

Per questo CAD il raggio, per cui non vi è una definizione standard data la forma irregolare dei noduli, è considerato come $4/3$ della distanza media dei voxel della ROI dal centro di gravità. Risulta quindi uguale al raggio della sfera per una ROI con precisamente questa forma.

Il centro di gravità è calcolato utilizzando i valori HU dei voxel come fattori di peso, più un ulteriore fattore 0.1 per i voxel sulla superficie della struttura:

$$X_i = \frac{\sum_{K=1}^N I_k r_{k,i} \epsilon_k}{\sum_{K=1}^N I_k \epsilon_k} \quad (17)$$

dove

$$\epsilon_k = \begin{cases} 1 & k \in \text{voxel interni} \\ 0.1 & k \in \text{voxel sulla superficie} \end{cases} \quad (18)$$

con k che scorre su tutti i voxel facenti parte della ROI, I_k è il livello HU del voxel k e $i = 1, 2, 3$ a seconda che si consideri x,y o z.

Il fattore aggiuntivo ϵ_k consente di calcolare in maniera migliore il centro di gravità nel caso di ROI che comprendono pezzi di pleura o vasi.

Features Geometriche	Features Intensità
centro di gravità $X_{i=x,y,z}$ (★)	Media
Raggio (mm)	Media fuori dalla maschera
Sfericità	Deviazione Std.
Skewness distanza da X_i	Dev. Std. fuori dalla maschera
Curtoi distanza da X_i	Massimo
Volume (mm^3) (★)	Entropia
Frazione voxel conn. alla pleura	Entropia fuori dalla maschera

Tabella 5: Lista delle features scelte per il CAD CAM. Quelle contrassegnate con (★) non sono utilizzate nel processo di classificazione.

La *Tabella 5* mostra un elenco delle features utilizzate nel CAD CAM, di seguito la descrizione di alcune di esse.

La *Sfericità* è definita come il rapporto tra il volume della ROI e quello di una sfera di uguale raggio.

La *Skewness* e *Curtoi* sono i momenti normali standardizzati terzo e quarto. L'entropia (di Shannon) delle intensità in HU è definita come

$$H(I) = \sum_{j=1}^N P(I_j) \log_2 \frac{1}{P(I_j)} \quad (19)$$

dove I_j è l'intensità del voxel e $P(I_j)$ è la probabilità di avere quel valore considerando la distribuzione dei livelli HU presenti nella ROI.

La *Frazione connessa alla pleura* è calcolata dividendo il numero di voxel (della superficie della ROI) connessi alla pleura rispetto al totale di quelli presenti sulla superficie della ROI.

Le features della *Tabella 5* indicate come “fuori dalla maschera” sono ottenute allargando la maschera della segmentazione tramite un elemento strutturante sferoidale (ovvero un operatore morfologico) di raggio $1,5mm$.

3.5 Filtraggio

Definite le features viene applicata una fase di filtraggio atta a diminuire il numero di candidati da inviare al classificatore. Per ogni esame infatti il numero di candidati noduli alla fine della fase di ricerca varia da centinaia a migliaia a seconda del numero delle fette, numero troppo grande per l'utilizzo diretto del classificatore.

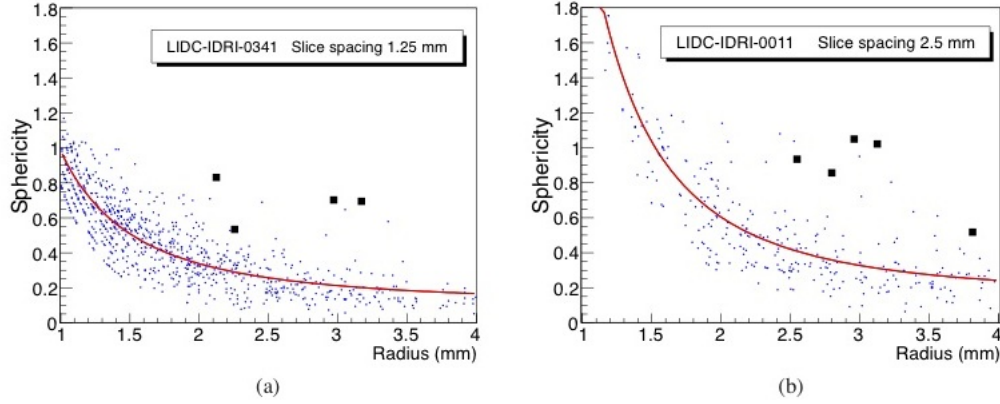


Figura 23: Grafici raggio *vs* sfericità, per le CT del database LIDC-IDRI, nel caso di pixel da 1.25mm (a) e 2.5mm (b). I quadrati neri rappresentano candidati corrispondenti a veri noduli.

Molti di questi candidati sono facilmente distinguibili osservando features quali raggio, sfericità e percentuale di superficie connesse alla pleura.

Nella *Figura 23* è possibile osservare come che la relazione tra raggio e sfericità consenta di distinguere molti dei candidati noduli che sono falsi positivi. Per poter usare questa relazione come filtro è però necessario equalizzare l'effetto della dimensione del voxel, come si nota nel confronto tra i due grafici della *Figura 23*. Per questo motivo per ogni CT si esegue un fit della correlazione tra raggio e sfericità con la funzione

$$S = \frac{a}{R^b} + c \quad (20)$$

L'equalizzazione si ottiene considerando la distanza tra la sfericità originale e il valore letto dal fit.

Oltre al filtraggio collegato alla sfericità si richiedono inoltre condizioni sul valore del raggio e della frazione di voxel connessa alla pleura.

Sebbene questo stadio di selezione comporti la perdita di alcuni noduli veri nel caso in cui mal segmentati, esso permette di ridurre il numero complessivo che viene inviato allo stadio di classificazione ad alcune decine per ogni CT.

3.6 Classificatore neurale

Per la classificazione dei noduli il CAD CAM utilizza una rete neurale “*feed forward*” [55, 54]. E’ possibile trovare la descrizione della rete in [47].

Una rete neurale feed forward è un caso particolare di rete neurale. In generale, le reti neurali sono la trasposizione in ambito informatico della struttura del sistema nervoso umano. Tra le varie attività per cui possono essere utilizzate vi sono la percezione di immagini e il riconoscimento di forme.

Come il modello biologico a cui si ispira, una rete neurale artificiale è composta da una serie di neuroni interconnessi da una rete di collegamenti per lo scambio di informazioni. Ogni neurone riceve in ingresso le informazioni provenienti da altri neuroni, le elabora pesandole in maniera diversa a seconda della provenienza ed invia a sua volta l’output ai neuroni a cui è connesso. Il peso assegnato ad ogni percorso dai singoli neuroni viene determinato durante l’addestramento della rete neurale, ed è la base della sua capacità di elaborazione delle informazioni. L’obiettivo del training è ottenere una rete neurale che sia in grado di cogliere relazioni non ovvie tra le informazioni che le sono fornite.

Normalmente una rete neurale è composta da tre strati di più neuroni, come mostrato nella *Figura 24*. Lo strato di input adegua le informazioni in ingresso alla rete alle esigenze dei neuroni dello strato nascosto, che le analizzano. I risultati dall’analisi vengono inviati allo strato di output, che produce l’output vero e proprio della rete.

Una rete neurale feed forward è una rete in cui il flusso di informazioni tra i neuroni avviene solo in direzione diretta, dall’input verso l’output. Il sistema quindi non ha nessun tipo di “*feedback*” interno.

Nello specifico del CAD CAM la rete neurale utilizzata ha la seguente struttura: 13 neuroni di input, 1 strato da 25 neuroni nascosti e 1 neurone di output che rende la probabilità di rilevanza del findings.

Il training della rete è stato effettuato utilizzando 20 esami del database ITALUNG, 5 di ANODE09 e 69 di LIDC-IDRI.

Allo stato attuale la rete non è ottimizzata ma l’ottimizzazione potrebbe essere fatta, qualora impiegata nella pratica clinica, utilizzando campioni di dimensione crescente per il training e la validazione.

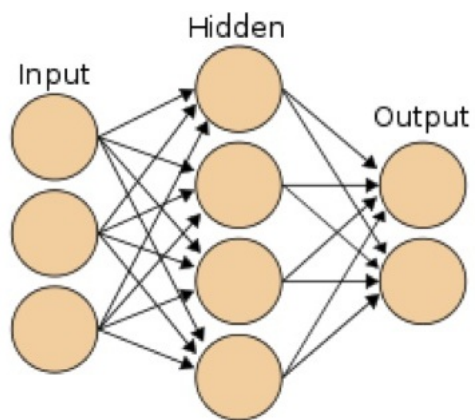


Figura 24: Schema degli strati di neuroni che costituiscono una rete neurale [54]. Lo strato nascosto può anche essere composto da più layers di neuroni.

Parte III:
Misure e risultati

4 Validazione CAD VBNA su database LIDC-IDRI

4.1 Introduzione

Nelle sezioni che seguono viene descritto l'utilizzo pratico del CAD VBNA sui casi del database LIDC-IDRI e i risultati ottenuti. Il lavoro di preparazione e utilizzo del CAD sui 1018 casi di questo database è stato interamente compiuto nell'ambito delle realizzazioni di questa tesi.

L'analisi del funzionamento e di quanto ottenuto risulta particolarmente interessante considerando la dimensione del database e la composizione molto eterogenea del database LIDC-IDRI.

I risultati precedentemente ottenuti dal CAD VBNA su un numero ristretto di casi di LIDC-IDRI e su altri database sono esposti in [47, 49].

4.2 Operazioni preliminari

Il database LIDC-IDRI introdotto nel *Paragrafo* 1.9.3 è stato utilizzato nella versione descritta in [32], contenente 1018 casi. Il numero è dovuto alla presenza nominale di 1012 esami di cui 8 contengono un doppio studio. Due esami mancano completamente.

Gli esami che presentano un doppio studio sono i numeri 132, 151, 315, 332, 355, 365, 442, 484.

I casi mancanti invece sono il numero 238 e il numero 585.

All'interno della cartella complessiva del database i singoli casi hanno path *LIDC - IDRI/LIDC - IDRI - num.caso/1.3.6.1...../000000/ROIs/*. La terza cartella è nominata come il codice identificativo dell'esame, il *studyinstanceUID* presente nei DICOM metadata. Questo codice consente di distinguere gli esami nei casi con doppio studio.

La cartella 000000 è quella che contiene le fette dell'esame, mentre nella cartella *ROIs* è presente il file contenente le annotazioni dei radiologi e vi viene anche salvato il file dei findings del CAD.

Per semplicità di lavoro in questa prima fase di utilizzo del CAD VBNA sul database si è scelto di non utilizzare tutta la struttura delle cartelle, in modo da non dover riportare oltre al numero del caso anche il codice identificativo.

Si è quindi provveduto a copiare tutti i file di ogni caso all'interno della cartella *LIDC – IDRI – #####*. I casi che presentano doppio studio sono stati quindi gestiti a mano.

La struttura originaria con il path completo dei singoli esami sarà reintrodotta in seguito.

Gli script che saranno citati nel corso della descrizione sono stati realizzati in bash o in python.

4.3 Utilizzo del CAD

Il funzionamento del CAD è stato descritto dal punto di vista concettuale nella *Capitolo 2*. Nel seguito sono invece riportate le operazioni necessarie per l'utilizzo pratico del programma.

Quanto viene descritto è un resoconto del lavoro svolto, nell'ambito di questa tesi, durante l'utilizzo del CAD VBNA sui 1018 casi del database LIDC-IDRI.

Oltre alle procedure strettamente necessarie all'utilizzo del CAD, viene riportata anche la descrizione di quanto è stato realizzato per la gestione del lavoro su un numero di casi così elevato.

Per comodità e per limitare il tempo di esecuzione si è scelto di lavorare su gruppi di 200 esami alla volta utilizzando i moduli *CAD_I* e *CAD_{JP}* separatamente: prima è stata effettuata la ricerca dei noduli interni su tutti gli esami, quindi si è passati alla ricerca di quelli pleurici.

Il primo passo per utilizzare il CAD è la creazione di un file testuale contenente l'identificazione del database, coerente con il nome della cartella contenente tutti gli esami, e la lista di casi che sono presi in considerazione. Uno script opportuno è stato realizzato per generare con un ciclo questo file per ogni gruppo di esami che si vuole considerare.

Avendo trascurato il codice identificativo dei singoli esami è sufficiente usare il numero del caso.

Il file viene passato ad un programma insieme al path della cartella in cui si vorrà salvare quanto prodotto dal lavoro del CAD sui casi presi in esame e il path del cartella contenente il CAD stesso. Questo script genera la cartella di output e le sotto directory utilizzate nel proseguimento dell'analisi, predisponendo quanto necessario per gli step successivi.

Le sotto directory sono:

- *Classification*, in cui saranno inserite le liste di findings dopo lo stadio di classificazione
- *Config*, in cui sono inseriti i file di configurazione del CAD
- *Data*, in cui sono copiate le cartelle contenenti i singoli esami considerati
- *FeaturesExtraction*, utilizzata solo se si sta eseguendo il training del CAD
- *RoiHunter*, in cui sono salvate le liste dei noduli output del Roi Hunter, i file stdout e immagini intermedie dell'elaborazione
- *Scripts*, in cui sono salvati gli script utilizzati per le varie operazioni eseguite dal CAD
- *Segmentations*, in cui sono salvate le segmentazioni dei singoli esami e file stdout

La descrizione dettagliata di quanto riportato è fornita nel seguito.

Il passo successivo è inserire all'interno della cartella Config i file di configurazione, contenenti alcuni dei parametri e dei path necessari al funzionamento del CAD. I parametri sono stati lasciati invariati rispetto a [47], mentre alcuni path sono stati sistemati in modo opportuno.

A questo punto, aperto un terminale, è possibile spostarsi nella cartella Scripts e lanciare i programmi relativi alle fasi di elaborazione descritte nella *Capitolo 2*.

Gli script in questione sono quelli relativi a :

- Segmentation, che esegue la segmentazione del polmone
- Roi Hunter, che compie l'identificazione delle ROI dei candidati noduli
- Classification, che compie la classificazione dei noduli con la riduzione dei falsi positivi
- Decision Rule, che converte la lista dei noduli trovati per ogni esame in formato Anode (*.anode*), in modo che sia comprensibile

Per le liste di noduli viene utilizzata la dicitura *.anode* in quanto storicamente è stato il primo database su cui il CAD è stato utilizzato.

Gli script che eseguono le date operazioni per il CAD_I e per il CAD_{JP} sono differenti, come spiegato in precedenza. Per la segmentazione non è necessaria questa distinzione in quanto l'operazione è identica per entrambi i CAD e viene eseguita una sola volta. Utilizzato il CAD_I su tutti i casi, il CAD_{JP} viene eseguito in maniera analoga ripetendo tutto quanto descritto fino ad ora. La fase di segmentazione è saltata copiando direttamente quanto ottenuto per i noduli interni.

Vengono quindi realizzate due liste di noduli per ogni esame, che vanno combinate per ottenere i risultati complessivi del CAD VBNA.

E' possibile convertire le liste di findings in modo che sia possibile osservarle insieme all'esame corrispondente attraverso un apposito visualizzatore DICOM.

In questo modo la posizione dei noduli viene indicata visivamente da un cerchio disegnato sulla fetta corrispondente.

Il visualizzatore DICOM utilizzato in questa tesi è Osirix, un programma molto comune per questo tipo di applicazioni, disponibile per computer Mac [59].

Convertendo le liste di findings dal formato *.anode* al formato *.xml* è possibile visualizzarle su Osirix tramite un apposito plugin.

La *Figura 25* e la *Figura 26* mostrano, a titolo di esempio, due schermate del programma.

Per realizzare la conversione bisogna utilizzare uno script apposito passando in ingresso il file *.anode* che si vuole convertire, il path dove salvare il file *.xml* prodotto, il label che precede i findings e altri parametri tra cui la dimensione del cerchio che verrà visualizzato sullo schermo.

Il file *.xml* creato va poi inserito nella cartella ROIs dell'esame affinché sia rilevato dal plugin.

L'operazione va ripetuta in modo iterativo per ogni esame, e può essere effettuata tramite uno script appositamente creato che esegue un ciclo su tutti i casi presi in considerazione.

Poichè si è scelto di trascurare per questa prima parte il numero di identificazione dei singoli studi, i casi contenenti un doppio studio sono stati trattati a mano. Per questi casi il CAD, fatto operare su tutto il database in modo indistinto, ha selezionato uno solo dei due studi. Si è quindi provveduto a verificare a quale studi facessero riferimento i risultati ottenuti utilizzando il file stdout della segmentazione. Isolati gli studi "mancanti" è stata applicata anche su di essi tutta la procedura sopra descritta. I risultati ottenuti sono quindi stati aggiunti a quelli complessivi.

Da questo punto si è scelto di ritornare ad utilizzare anche l'identificazione del singolo studio. Tramite uno script realizzato appositamente sono state create delle liste contenenti i path dalla cartella del singolo caso caso fino a quella contenente le fette per la totalità dei casi, i casi usati come training e tutti meno il training. In questo modo si tengono conto dei doppi studi ed è possibile, passando queste liste in ingresso, rendere molto più efficienti i vari script.

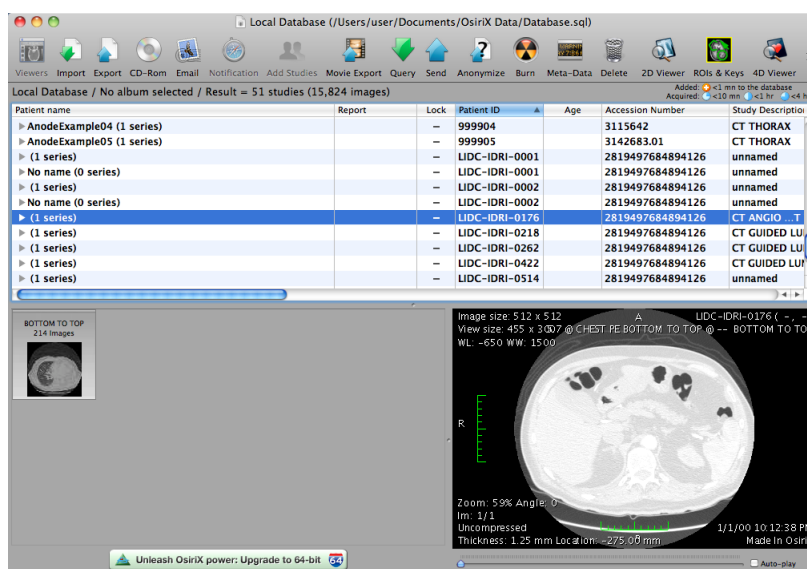


Figura 25: Schermata del visualizzatore DICOM Osirix con la lista dei casi caricati nel programma.

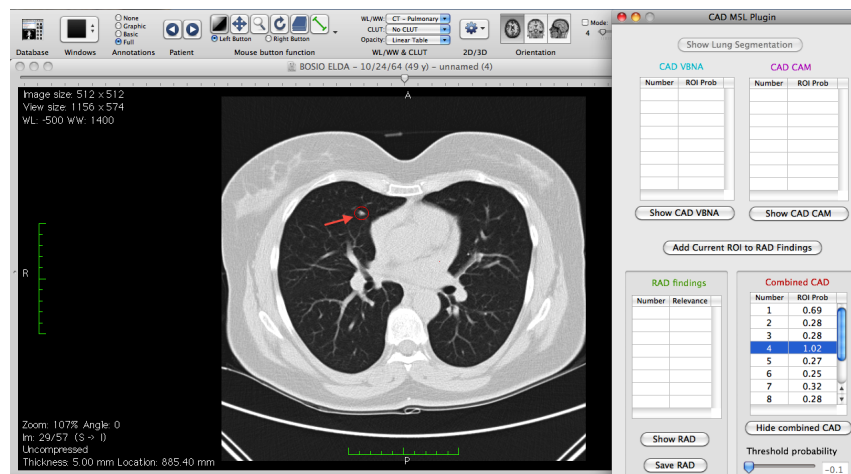


Figura 26: Schermata di Osirix durante la visualizzazione di un esame. La finestra sulla destra è quella relativa al plugin per la visualizzazione dei findings del CAD. Il candidato nodulo selezionato nell’elenco è evidenziato dal cerchio rosso sulla fetta dell’esame. Come si può notare sotto la lista dei findings, è possibile variare la soglia sulla probabilità di quelli che vengono elencati.

4.4 Combinazione dei risultati

Come spiegato nella *Paragrafo 2.6*, per combinare i risultati dei CAD_I e CAD_{JP} sono necessarie operazioni quali normalizzazione e merging.

La normalizzazione viene effettuata tramite uno script che richiede in ingresso una lista di findings, il dizionario di normalizzazione, lo spacing file e un’identificazione.

La liste di findings sono ottenute mettendo insieme tutti i file *.anode* separatamente per il CAD_I e il CAD_{JP} , mentre il dizionario di normalizzazione è ottenuto da una curva Froc realizzata su un set di dati che si ritiene opportuno. Tipicamente il set scelto è quello utilizzato per il training del CAD, ma può essere anche utilizzato un dataset più esteso in modo da aumentare la statistica di casi.

Lo spacing file è la lista delle dimensioni di voxel e fetta degli esami i cui candidati noduli sono inseriti nella lista di findings, con denominazioni consistenti tra le due. La lista delle dimensioni è ottenuta partendo dai files di report del processo di segmentazione (ovvero i file di “*standard output*” dello script), che contengono questi dati in una stringa specifica. Per recupe-

rarli è stato realizzato un programma che, in maniera iterativa per ogni caso considerato, ricerca la stringa specifica all'interno del file di report e la salva nello spacing file modificandone cambiandone l'intestazione e la struttura in modo da rendere evidenti i valori numerici e il caso a cui corrispondono.

I risultati dello script di normalizzazione di entrambi i CAD possono ora essere combinati utilizzando uno script che ne fa la combinazione. Questo programma richiede in ingresso le liste normalizzate e la distanza di cluster citata al termine del *Paragrafo 2.6*.

In questo modo si ottiene la lista complessiva dei findings del CAD VBNA per il gruppo di esami considerato.

4.5 Realizzazione della curva FROC

Come descritto nella *Paragrafo 1.8* la curva FROC è un metodo di valutazione delle prestazioni di un sistema CAD. Tutte le curve FROC presentate in questa tesi sono state ottenute per mezzo di uno script apposito secondo la procedura descritta di seguito.

Per poter utilizzare questo script è necessario avere per i casi presi in considerazione, oltre alla lista dei findings e a quello dello spacing, anche l'elenco delle annotazioni dei radiologi. Per questa lista è stato realizzato uno script ausiliario che per ogni caso va a ricercare la refertazione nella cartella ROIs la copia in una lista unica sistemando il layout in maniera opportuna.

Le tre liste sono passate in input allo script per realizzare la FROC insieme a diversi altri parametri che definiscono quale sia il formato delle annotazioni, come considerare le stesse, come trattare eventuali findings coincidenti, se escludere noduli irrilevanti e il nome che si vuole dare ai file di output.

Il formato delle annotazioni si distingue tra lo standard utilizzato nel database ANODE09 e quello usato in LIDC-IDRI. Entrambe le modalità di annotazione riportano il nome del caso, la coordinate (x, y, z) del nodulo e il suo raggio, ma si differenziano per i dati successivi. Nel caso di ANODE09 viene solo aggiunto l'indice 1 o 2 per indicare un nodulo come rilevante o irrilevante (come descritto nel *Paragrafo 1.9.2*), mentre nel caso di LIDC-IDRI viene riportato il numero di radiologi che hanno annotato il nodulo e un valore da 1 a 3 per la tipologia del nodulo (come descritto nel *Paragrafo 1.9.3*).

E' quindi necessario indicare al programma quale formato viene usato affinché legga correttamente le informazioni e, nel caso LIDC-IDRI, anche

il numero minimo di radiologi che si richiede abbiano annotato un nodulo. Data l'ampia variabilità delle annotazioni tra i diversi radiologi tipicamente si richiede l'accordo di almeno due di essi.

La lista delle annotazioni utilizzata è stata ottenuta come ibrido tra le due modalità di annotazione: partendo dalle liste proprie del database LIDC-IDRI si è passati ad un formato ANODE09 considerando come noduli non rilevanti secondo questo database tutti quelli annotati come “irrilevanti” o “non noduli” dai 4 radiologi di LIDC-IDRI. Inoltre sono stati considerati come non rilevanti anche tutti i noduli annotati da un solo radiologo. In questo modo, pur richiedendo l'accordo di almeno due refertatori, si evita di considerare come falsi positivi noduli che hanno indotto almeno un radiologo a segnalarli.

Inoltre le annotazioni risultano più simili agli standards di ANODE09 e ITALUNG, entrambi basati su un sistema a doppia lettura.

Per convertire le annotazioni in questo modo è stato realizzato uno script apposito che legge ogni riga della lista di annotazioni, individua tutti i campi contenuti e la stampa all'interno di una nuova lista compiendo le sostituzioni spiegate sopra.

Dal programma per la FROC si ottengono 3 file i output: un file testuale con l'elenco dei noduli utilizzati per la curva e le soglie sulla sensibilità, un file *matlab* [60] e un nuovo dizionario che potrebbe essere utilizzato per la normalizzazione.

Utilizzando il file *matlab* è possibile realizzare i grafici della curva ottenuta.

4.6 Risultati su database LIDC-IDRI

In questo paragrafo sono descritti i risultati ottenuti dall'utilizzo CAD VBNA sul database LIDC-IDRI. Sono stati analizzati 949 casi, ovvero 1018 totali meno i 69 casi utilizzati per il training. I casi di training sono esclusi per avere un set di validazione che non comporti bias. Su questi 949 casi il CAD ha completato con successo l'analisi a meno di 20 casi, pari ad un tasso di mancato funzionamento del 2%. Per la maggior parte di questi casi il problema è stato il mancato funzionamento della segmentazione. Essendo il primo passo nell'analisi dell'immagine, ed essendo condivisa da entrambi i moduli, per questi esami non si hanno findings ne per il CAD_I ne per il CAD_{JP} . I casi in cui si sono verificati i mancati funzionamenti sono stati analizzati singolarmente, ottenendo le seguenti cause:

- caso 176: il paziente risulta avere un solo polmone, come visibile dalla *Figura 27*. Il parere di un radiologo interpellato a riguardo è che sia stato asportato chirurgicamente e successivamente si sia riempito di liquido. Osservando l'immagine con un finestra di livelli di grigio opportuna è possibile distinguere i punti di sutura a livello dei bronchi.
- casi 332, 355, 365, 442 e 514 (uno dei due studi se doppi): lo spacing a livello di spessore fetta risulta essere anomalo, probabilmente per un problema a livello di attributi del DICOM. A livello esemplificativo per l'esame 514 a livello di segmentazione risulta essere di $38mm$, valore molto più elevato della media del database. Per alcuni di questi casi la segmentazione non è fallita, ma l'analisi non ha comunque portato all'individuazione di alcun nodulo sempre, a causa dello spacing errato.
- casi 218, 262, 291, 311, 313, 415, 422, 423, 671, 697, 785, 796, 850 e 938: le immagini risultano girate al contrario, ovvero acquisite con il paziente steso sulla pancia. Questo provoca problemi di identificazione delle strutture a livello della segmentazione. Quale che sia la motivazione per cui queste immagini sono state acquisite in questo modo comunque risultano un'eccezione rispetto alla normale pratica diagnostica. Un esempio è mostrato nella *Figura 28*



Figura 27: Caso 176 del database LIDC-IDRI. Il paziente presenta sostanzialmente una sola struttura polmonare.

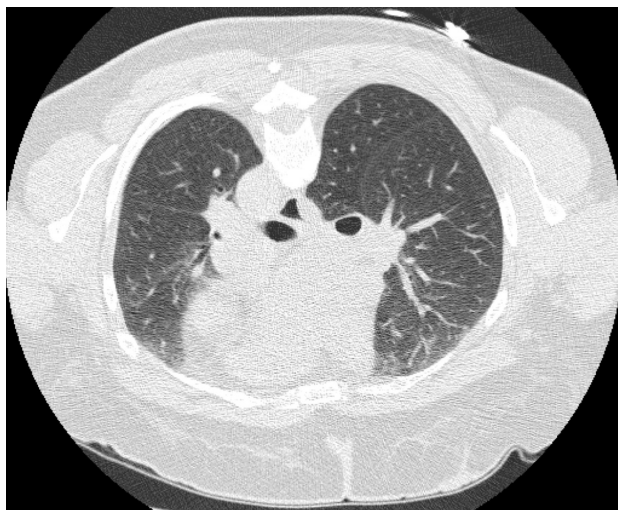


Figura 28: Caso 262 del database LIDC-IDRI. L'immagine è stata acquisita con il paziente girato al contrario.

Considerando le varie tipologie delle problematiche riscontrate, si può notare come siano tutte da considerare come anomalie dell'anatomia, dei file DICOM o dell'acquisizione. Sui restanti casi, che rappresentano la normalità per esami di questo tipo, il CAD è sempre giunto al termine dell'analisi producendo i candidati noduli sia per il CAD_I che per il CAD_{JP} .

E' possibile supporre che alcuni dei casi sopra citati facciano parte del gruppo di esami con "fonti di disturbo" (alti livelli di rumore, artefatti, altre patologie ecc ecc.) che sono stati lasciati di proposito all'interno del database, come spiegato nel *Paragrafo 1.9.3*.

Le tipologie di anomalie riscontrate nei diversi casi, seppure diverse, sono tutte tali da compromettere in partenza il funzionamento del CAD. In tutti i casi infatti viene a mancare una delle ipotesi di base con cui viene affrontata la segmentazione delle strutture. Tali ipotesi si basano sullo standard di questa tipologia di esami, e non sono quindi soddisfatte da casi con caratteristiche peculiari molto diverse dalla normalità.

Non è possibile pertanto ottimizzare il CAD in maniera che possa risultare funzionante anche in condizioni analoghe, che risultano comunque essere molto distanti dall'ambito di utilizzo per cui è stato realizzato.

Considerando tutti i risultati ottenuti si possono realizzare le curve FROC separatamente per il CAD_I e il CAD_{JP} . Sono inclusi anche casi in cui

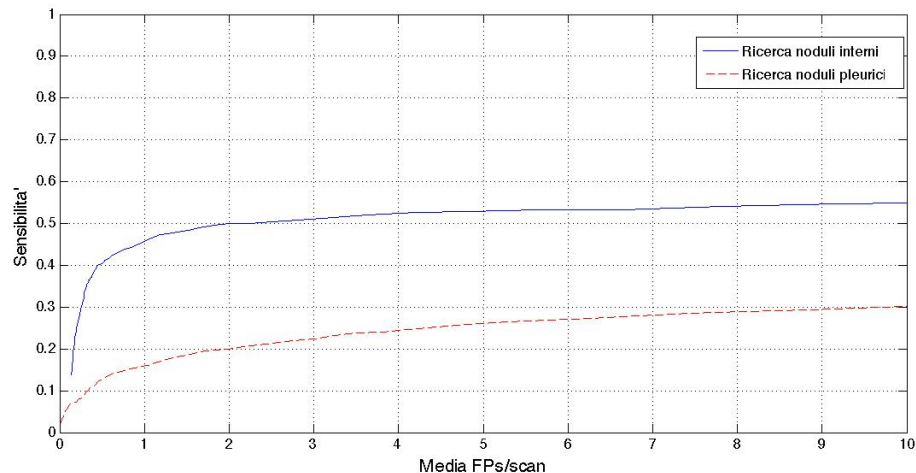


Figura 29: Confronto FROC del CAD_I e del CAD_{JP} sul database LIDC-IDRI escludendo il training.

l'analisi è fallita l'analisi del CAD, che quindi non hanno nessun findings elencato nella lista relativa.

Le curve ottenute sono mostrate nella *Figura 29*.

La curva del CAD_I non parte da zero in quanto se esiste un FP con punteggio maggiore del TP con probabilità più alta allora il primo punto della FROC non sarà nell'origine. Si può notare che la FROC del CAD_{JP} è decisamente più bassa rispetto a quella del CAD_I , ma presenta un incremento maggiore spostandosi verso alti valori di $FPs/scan$. La ragione di questa minore prestazione può essere ricondotta allo stadio di classificazione dei noduli pleurici, che non risulta totalmente efficiente nel discriminare i noduli reali dai falsi positivi peggiorando il risultato del CAD_{JP} . Un futuro sviluppo del sistema potrebbe quindi passare attraverso un'ottimizzazione della classificazione dei noduli pleurici.

Le *Figure 30* e *31* mostrano invece le FROC del CAD VBNA nel suo complesso ottenute, sul database LIDC-IDRI escludendo i casi di training, con la combinazione del CAD_I e del CAD_{JP} .

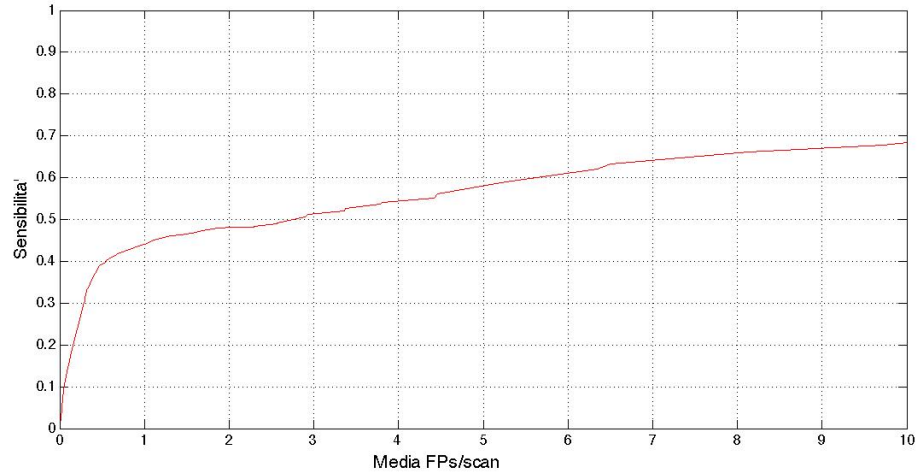


Figura 30: FROC della combinazione del CAD_I con il CAD_{JP} sul database LIDC-IDRI esclusi i casi di training.

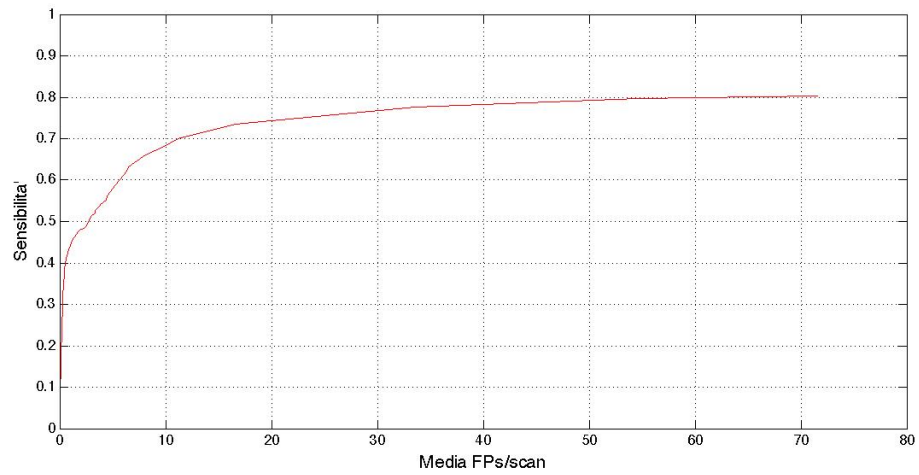


Figura 31: FROC della combinazione del CAD_I con il CAD_{JP} sul database LIDC-IDRI esclusi i casi di training. Scala estesa fino a $80^{FPs/scan}$.

Nella *Figura 30* si può notare come la crescita della prima parte della curva sia resa più graduale per effetto del CAD_{JP} , che come spiegato risulta meno efficiente a bassi livelli di falsi positivi. Nonostante questo la *Figura 31* mostra come la sensibilità complessiva del CAD VBNA sia molto prossima al 80% intorno a $50^{FPs/scan}$ e sia superiore al 70% poco oltre a $10^{FPs/scan}$.

Questi risultati sono confrontabili e migliori rispetto a quelli pubblicati su [47, 49] ottenuti su un numero limitato di esami del database, rispettivamente 69 e 85, considerando sempre l'accordo tra almeno due radiologi. Quelli riportati in [49] differiscono da quanto riportato in questa tesi anche per il set di training utilizzato per l'ottimizzazione della classificazione, in quanto è stato utilizzato il database ITALUNG. Per i risultati ottenuti in [47] è stato utilizzato lo stesso set di 69 esami del database LIDC-IDRI citato nel *Paragrafo 2.5*.

Il fatto che le prestazioni del sistema si mantengano paragonabili variando set di training e di validazione è prova della stabilità di funzionamento del CAD VBNA. Data l'ampia variabilità garantita dal grande numero di esami, i risultati ottenuti sui 949 casi del database LIDC-IDRI permettono di ipotizzare concretamente l'utilizzo di questo CAD anche su esami provenienti dalla normale pratica diagnostica.

La *Figura 32* mostra il confronto tra la FROC riportata nella *Figura 30* e quella ottenuta escludendo, oltre ai casi utilizzati per il training, anche tutti i casi descritti in precedenza per cui il CAD ha fallito la ricerca dei candidati noduli.

La curva ottenuta escludendo i casi problematici risulta più alta in quanto nella FROC della *Figura 30* per la mancanza di findings tutti i noduli annotati come rilevanti vengano conteggiati come falsi negativi nella definizione della sensibilità, abbassandola.

La differenza è comunque quantificabile in un paio di centesimi ad alti valori di $FPs/scan$, mentre a valori più bassi le due curve sono più vicine.

Si riporta questo confronto per completezza in quanto nel proseguimento del lavoro sono stati sempre considerati anche i casi falliti. Questo perchè, per quanto inusuali, problemi analoghi possono essere riscontrati qualora il CAD venga utilizzato nella normale pratica clinica.

Le problematiche sono state comunque parzialmente risolte tramite l'utilizzo del CAD VBNA in combinazione con il CAD CAM, come descritto nella *Capitolo 5*.

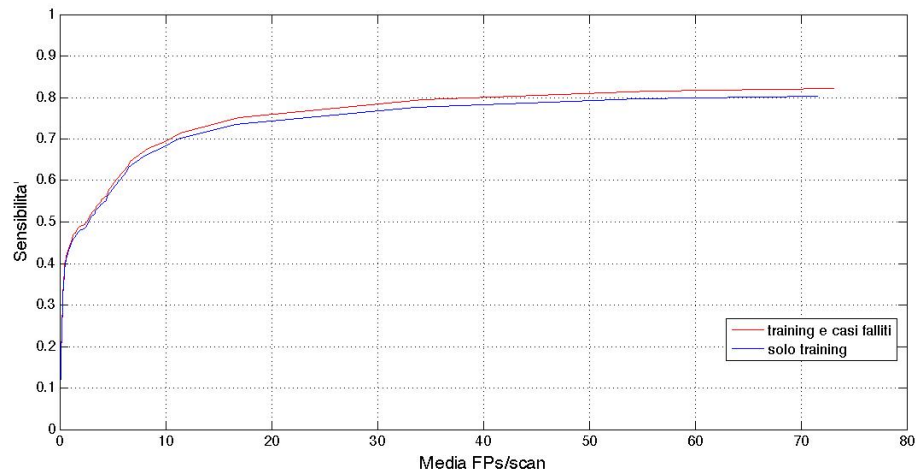


Figura 32: Confronto tra le FROC del CAD VBNA ottenute escludendo solo i 69 casi di training o escludendo il training e i casi per cui la ricerca dei noduli è fallita.

5 Combinazione CAD CAM e CAD VBNA e validazione su database LIDC-IDRI

5.1 Introduzione

Nei capitoli che seguono viene descritta la combinazione del CAD VBNA con il CAD CAM. Come viene spiegato nel paragrafo successivo, l'utilizzo di CAD diversi combinati tra loro consente di migliorare le prestazioni dei singoli sistemi. I diversi approcci adottati fanno sì che i vari sistemi risultino parzialmente complementari tra loro.

Nell'ambito dello sviluppo di questa tesi è stata quindi eseguita la combinazione dei risultati ottenuti singolarmente dai CAD VBNA e CAM sui 1018 casi dal database LIDC-IDRI, ottenendo le FROC relative all'utilizzo in parallelo dei due sistemi.

Nel proseguimento della tesi sarà descritto l'utilizzo combinato di CAM e VBNA nell'ambito di un innovativo servizio web di CAD on-line.

5.2 Utilizzo di più CAD combinati

Come risulta chiaro dalle descrizioni nei *Capitoli 2 e 3*, le strutture degli algoritmi CAD VBNA e CAD CAM sono molto diverse. Questo è valido anche più in generale: essendo l'individuazione dei noduli all'interno delle immagini TAC una tematica molto complessa, quasi tutti i sistemi CAD affrontano il problema con approcci diversi e presentano caratteristiche originali. Si può quindi immaginare che l'utilizzo combinato di sistemi diversi possa portare vantaggi, qualora si rivelino complementari nell'andare a sopperire alle singole mancanze o inefficienze.

Questa ipotesi viene verificata all'interno dell'articolo [29], dove diversi sistemi vengono utilizzati singolarmente e in combinazione sui casi del database ANODE09. La conclusione riportata è che la combinazione di più sistemi risulta sempre migliore del sistema più performante preso singolarmente. Altre evidenze sono presentate in [57].

La motivazione per cui le sensibilità della combinazione dei sistemi migliora rispetto ai risultati dei singoli è riconducibile al fatto che, a fronte di un aumento dei TP, il numero di FP resta all'incirca costante.

I diversi approcci consentono quindi ai vari CAD di essere complementari nella rivelazione dei noduli, senza incrementare troppo il numero di falsi positivi. Alla diversità di funzionamento può aggiungersi quella dei set di

esami utilizzati per il training, ma l'incremento delle prestazioni permane anche se addestrati sullo stesso dataset [47].

La riduzione dei falsi positivi permette inoltre di diminuire il tempo necessario ad un radiologo per rivedere i findings del sistema ed operare la selezione manualmente.

5.3 Combinazione CAD CAM e CAD VBNA

Come scritto in precedenza, i risultati presentati in [47] sono stati ottenuti su di una versione del database LIDC-IDRI contenente 138 casi, di cui 69 sono stati utilizzati per la validazione. Negli ultimi mesi sia per il CAD VBNA che per il CAD CAM è stata nuovamente eseguita la validazione utilizzando 949 sei 1018 casi presenti attualmente all'interno del database, non facenti parte del training. Avendo a disposizione questi dati si è scelto di eseguire nuovamente la combinazione, per confermare i risultati ottenuti su un gruppo di casi molto più ampio ed eterogeneo.

Per poter gestire in maniera semplice i risultati dei singoli CAD e quelli della combinazione si è deciso creare un archivio con esattamente la stessa struttura di cartelle e percorsi presente all'interno del database LIDC-IDRI. In questo modo per entrambi i sistemi è possibile utilizzare tutti i programmi e le procedure che si userebbero sul database stesso.

Per ogni caso i files dei findings del CAD VBNA e dello spacing sono stati ottenuti partendo dalle liste costruite per ottenere la FROC, descritte nei *Paragrafi* 4.4 e 4.5. Tramite uno script apposito è stato realizzato un dizionario dove l'identificazione dei casi funge da chiave per ricercare all'interno delle liste le stringhe contenenti le informazioni ad essa associate. Le stringhe raccolte per ogni chiave sono state stampate su files di testo situati nelle posizioni opportune all'interno dell'archivio. Le annotazioni dai radiologi sono state invece semplicemente copiate per ogni caso dalla cartella ROIs del database a quella dell'archivio.

I risultati del CAD CAM sono stati salvati in maniere analoga.

L'utilizzo dell'archivio consente inoltre maggiore flessibilità nel creare le nuove liste necessarie alla combinazione dei due CAD e alla realizzazione della FROC. Partendo da un elenco contenente numero e codice identificativo degli esami che si vogliono utilizzare è possibile, tramite uno script opportuno, generare le liste di findings, spacing e annotazioni con i file salvati all'interno dell'archivio.

La combinazione dei risultati dei due CAD è stata ottenuta con la stessa procedura descritta nel *Paragrafo* 4.4 nell'ambito della combinazione del CAD_I e del CAD_{JP} .

5.4 Risultati combinazione CAD CAM e CAD VBNA (CAD M5L)

La curva FROC ottenuta sui 949 casi non facenti parte del training è mostrata nella *Figura* 33, mentre la *Figura* 34 mostra la curva con scala estesa fino a $80^{FPs/scan}$.

Osservando entrambe le figure è possibile notare come la combinazione dei due CAD migliora sensibilmente la prestazione rispetto ai risultati ottenuti dai sistemi separatamente. Rispetto a quanto ottenuto in [47] inoltre l'incremento non si limita alla fascia compresa tra 0 e $10^{FPs/scan}$ ma, come risulta dalla *Figura* 34, si mantiene rilevante anche a valori più elevati di falsi positivi. Ne risulta una sensibilità di $\sim 90\%$ a $50^{FPs/scan}$.

Nell'intervallo più interessante per i radiologi, tra 0 e $10^{FPs/scan}$, la sensibilità è maggiore al 70% già oltre a $2^{FPs/scan}$ ed è $\sim 80\%$ a $7^{FPs/scan}$.

Questi risultati sono quindi molto interessanti sia per i valori di sensibilità ottenuti a numeri bassi di falsi positivi che per il miglioramento che rappresentano rispetto ai singoli CAD (quantificabile in poco meno del 10% nell'intervallo $0 - 10^{FPs/scan}$ per quanto riguarda il CAD CAM).

L'incremento è ancora maggiore considerando il CAD VBNA che però, come è già stato scritto, paga una prestazione inferiore a bassi numeri di falsi positivi.

Il CAD CAM utilizzato per ottenere i risultati riportati nelle *Figure* 33 e 34 è in una versione successiva e migliorata rispetto a questo esposto in [47].

Nei capitoli che seguano il CAD formato dalla combinazione tra CAD CAM e CAD VBNA sarà denominato come CAD M5L (Magic 5 Lung), dal nome della collaborazione nel cui ambito sono stati sviluppati questi CAD [27].

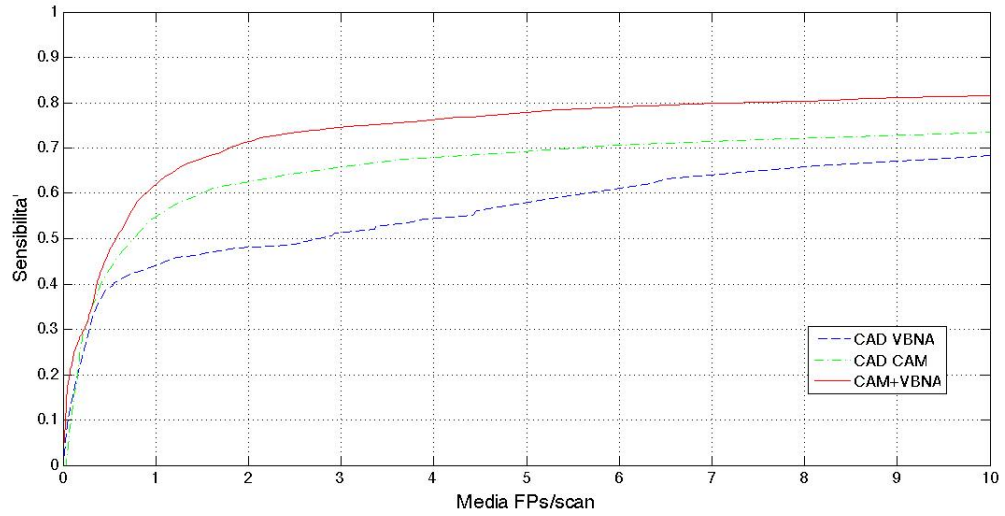


Figura 33: Confronto delle FROC dei CAD CAM, VBNA e loro combinazione sul database LIDC-IDRI, escludendo il training.

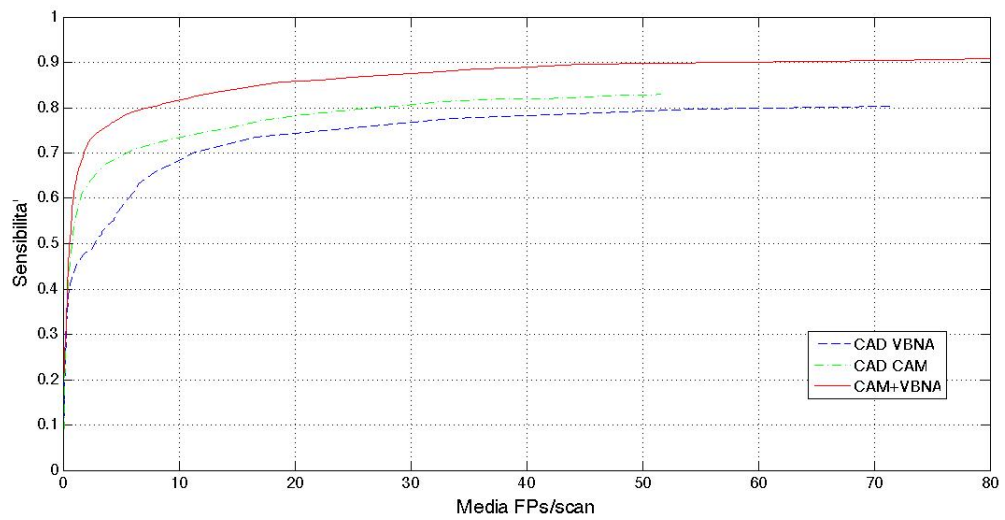


Figura 34: Confronto delle FROC dei CAD CAM, VBNA e loro combinazione sul database LIDC-IDRI, escludendo il training. Scala estesa fino a $80^{FPs/scan}$

6 Test su database ANODE09

6.1 Introduzione

Nei capitoli che seguono viene descritto l'utilizzo della combinazione dei CAD CAM e VBNA sul database ANODE09. Poichè la validazione dei due sistemi sui casi del database è già stata effettuata all'interno di diversi articoli ([29, 47]), in questa tesi viene testato l'utilizzo dei CAD a livello pratico, ovvero come supporto alla lavoro del radiologo.

A questo proposito tre radiologi hanno eseguito singolarmente la refertazione dei 50 casi del database in due fasi distinte. Nella prima fase i radiologi hanno esaminato normalmente gli esami, nella seconda hanno rivisto le proprie refertazioni utilizzando i findings segnalati dai CAD. Considerando la variazione nei risultati ottenuti con e senza CAD è possibile ottenere una stima dell'efficacia del sistema come seconda lettura.

E' stato inoltre effettuato il confronto, sui casi del database, tra le prestazioni fornite dalla combinazione di CAM e VBNA e quelle di un sistema commerciale che viene fornito in associazione ad uno scanner CT tra i più utilizzati nei reparti di radiologia.

I dati e i risultati riportati nei seguenti capitoli sono stati sono stati interamente raccolti nell'ambito della realizzazione di questa tesi.

Ricordiamo che il CAD M5L è la denominazione utilizzata per indicare la combinazione tra CAD CAM e CAD VBNA.

6.2 Validazione CAD M5L con radiologi

La validazione della combinazione del CAD VBNA e del CAD CAM sul database ANODE09 è stata effettuata in varie modalità ed è riportata in diversi articoli [29, 47]. In questo paragrafo viene invece descritto il test dell'utilizzo, da parte di diversi radiologi, dei CAD combinati come seconda lettura per la refertazione dei casi del dataset.

Ricordiamo che in questa modalità di refertazione, dopo aver analizzato gli esami normalmente, il radiologo controlla nuovamente l'esame tramite i findings del CAD, rivedendo le proprie annotazioni qualora lo ritenga necessario

Il lavoro si è svolto nel seguente modo: tre radiologi hanno refertato i 50 esami del database ANODE09 separatamente, distinguendo tra noduli rilevanti e irrilevanti. Ricordiamo che sono considerati irrilevanti, nell'accezione

del database, tutti i noduli con caratteristiche benigne, troppo piccoli per essere rilevanti o le strutture che mimano un nodulo ma il medico ritiene non essere tali.

Completata questa prima fase, e lasciato passare qualche tempo per evitare bias, i radiologi hanno ricontrollato gli esami osservando però solamente i findings del CAD. Gli esami non sono stati rivisti completamente, al medico è stato richiesto soltanto di valutare i candidati noduli del programma, aggiungendoli eventualmente alla sua lista.

Si noti che il fatto di valutare solo i findings del CAD è una condizione più restrittiva rispetto ad una seconda lettura, ovvero ad eseguire completamente una seconda refertazione. In questo modo è possibile osservare solo variazioni della sensibilità come conseguenza all'aggiunta di findings del CAD nelle annotazioni, quanto segnalato dal medico nella prima refertazione non viene modificato. Di conseguenza il numero di FP può solamente restare uguale o aumentare.

Rivedendo completamente l'esame invece al medico sarebbe permesso modificare le proprie annotazioni originarie, per esempio eliminando annotazioni dalla propria lista. In tale caso il numero di falsi positivi potrebbe anche diminuire.

Per limitare il numero di findings mostrati ai radiologi è stato deciso di adottare una soglia sulla probabilità, fissata al 5%. In questo modo il numero di candidati noduli segnalati dal CAD M5L è 588.

Poiché le refertazioni dei tre radiologi sono avvenute in momenti diversi, le modalità in con cui sono state ottenute le liste di annotazioni differiscono leggermente.

Per il lavoro effettuato dai primi due radiologi non è stato possibile osservare contemporaneamente i punti annotati dal medico con i findings del CAD. I candidati noduli segnalati dal programma sono stati visionati separatamente e sono stati indicati come rilevanti o non rilevanti. Solo in un secondo momento, tramite uno script apposito, è stata realizzata la fusione tra la refertazione di ogni radiologo e la lista delle sue valutazioni dei findings del CAD. Alle liste stilate dai radiologi sono stati aggiunti sia i findings rilevanti del CAD che non coincidevano (all'interno di un intervallo di tolleranza) con nessuna annotazione del medico, sia i findings che, pur coincidendo, mostravano che l'opinione del medico era passata da non rilevante a rilevante.

Nel caso che più findings del CAD fossero abbastanza vicini da risultare coincidenti con un nodulo della refertazione si è scelto di utilizzare quello con

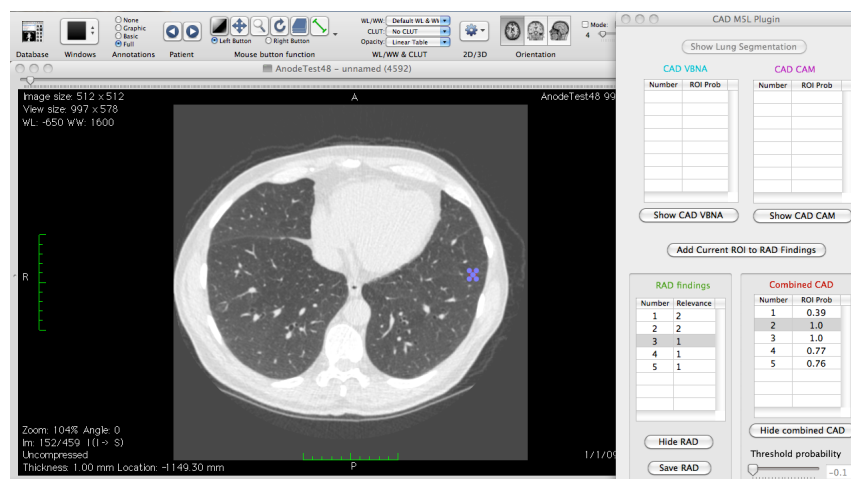


Figura 35: Schermata di Osirix con la finestra del plugin utilizzata per le refertazioni. Le colonne in basso mostrano le annotazioni del radiologo (a sinistra) e i findings del CAD (a destra). Cliccando su una delle linee, come mostrato nell'immagine la ROI corrispondente viene visualizzata ed evidenziata sulla fetta corrispondente.

la distanza minore dal nodulo annotato dal radiologo.

Il lavoro del terzo radiologo è stato invece più diretto in quanto al momento della refertazione era disponibile un plugin di Osirix appositamente realizzato, mostrato nella *Figura 35*.

Questa applicazione permette di creare una lista di annotazioni direttamente dal programma, indicando con una ROI le strutture sospette. Inoltre è possibile visualizzare i findings del CAD contemporaneamente a quanto già annotato ed eventualmente aggiungerli alla lista di annotazioni.

Le istruzioni per l'utilizzo del plugin sono riportate nell'*Appendice A*.

6.3 Risultati validazione con radiologi

Raccolte le refertazioni (senza e con l'ausilio del CAD) dei tre radiologi sui singoli esami, per ogni medico sono state realizzate due liste complessive di annotazioni utilizzando uno script apposito.

Prima di inviare i risultati ai curatori del database ANODE09 per l'analisi (come spiegato nel *Paragrafo 1.9.2*) si è provveduto a realizzare una serie di confronti in modo da verificare la consistenza tra le varie liste di annotazioni.

Radiologo	Ann. senza CAD	Ann . con CAD
1	371	442
2	442	484
3	308	427

Tabella 6: Numero di noduli annotati dai tre radiologi, sui 50 casi del database ANODE09, senza e con l’ausilio del CAD M5L.

I risultati sono raccolti nelle *Tablelle* 6 e 7: nella prima sono riportati il numero di noduli complessivamente annotati dai tre radiologi nelle due fasi della refertazione, nella seconda il numero di annotazioni coincidenti tra le refertazioni di ciascuna possibile coppia.

I numeri riportati nella Tabella 6 possono sembrare elevati, in considerazione del fatto che sono stati ottenuti su 50 esami, come anche i 588 findings (con probabilità $\geq 5\%$) riportati dal CAD M5L. Bisogna però ricordare che, come spiegato nel *Paragrafo* 1.2, ogni polmone, anche se appartenente ad un soggetto “sano”, contiene qualche struttura nodulare. Principalmente tali strutture hanno caratteristiche benigne. Poichè i casi provengono da una popolazione a rischio, il numero di tali strutture è più elevato e vi è una maggiore probabilità che alcune presentino caratteristiche patologiche. Per questo motivo sono comunque rilevate e segnalate dai radiologi.

I confronti riportati nella *Tabella* 7 sono stati ottenuti tramite uno script apposito, analogo a quello utilizzato per la fusione dei findings del CAD con le refertazioni dei primi due radiologi. Sono considerati coincidenti i noduli con una distanza quadratica

$$dist = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \leq 12pixel \quad (21)$$

Confr. radiologi	Coinc. solo rad.	Coinc. rad+CAD
1 vs 2	123	130
1 vs 3	124	152
2 vs 3	135	173
1 vs 2 vs 3	62	78

Tabella 7: Confronti tra le annotazioni, eseguite sia senza che con l’ausilio del CAD M5L, dei radiologi presi a coppie. Nell’ultima riga sono riportati i noduli coincidenti tra tutti e tre i radiologi.

Questa soglia sulla distanza è stata scelta considerando le distribuzioni delle distanze tra tutti i noduli annotati per ogni caso da ogni coppia di radiologi, riportate nella *Figura 36* e nella *Figura 37*. Le distribuzioni sono state ricavate calcolando la distanza tra ogni nodulo annotato dal primo radiologo e ogni nodulo annotato dal secondo, per ognuno dei 50 casi. In questo modo si ottengono delle distribuzioni che hanno un massimo vicino a zero, dato dalle distanze tra le coppie di noduli effettivamente coincidenti, e una coda a distanze elevate, costituita dai valori ottenuti su coppie di annotazioni che riguardano noduli diversi. La distanza a cui si trova il minimo, come si può osservare all'incirca a $12pixel$, può quindi essere scelta come taglio minimizzando il rischio di escludere noduli effettivamente coincidenti o includerne coincidenze fasulle.

Osservando la *Tabella 6* è possibile notare come i numeri di noduli annotati siano abbastanza dissimili nelle refertazioni iniziali, ma diventino molto simili al termine della revisione con l'ausilio del CAD.

Le differenze tra i radiologi 2 e 3, quelli che in entrambi i casi presentano numeri più diversi, sono di 134 noduli nella prima refertazione e di 57 a seguito della revisione tramite il CAD. Considerando i valori percentuali delle differenze rispetto ai valori medi si ottiene una differenza del 36% nel primo caso e del 13% nel secondo.

Nonostante le differenze, nella seconda colonna della *Tabella 7* si può notare come il numero di noduli coincidenti tra le varie coppie di radiologi nelle refertazioni senza il CAD sia sostanzialmente lo stesso.

Questi risultati sono spiegabili considerando che non esiste una definizione univoca del concetto di nodulo e che la valutazione può variare da radiologo a radiologo a seconda dell'esperienza, dalla formazione e da molti altri fattori.

I numeri paragonabili nei confronti a coppie dimostrano che vi è sicuramente un gruppo di anomalie che tutti e tre i radiologi hanno rilevato e su cui concordano, probabilmente formata dai noduli più evidenti e con caratteristiche più tipiche, ma sulla restante parte le annotazioni variano a seconda del medico.

Dall'ultima riga della *Tabella 7* è possibile vedere come i noduli su cui concordano tutti i radiologi, senza l'utilizzo del CAD, sono soltanto 62.

La *Tabella 8* mostra invece il livello di accordo tra le refertazioni, eseguite senza l'ausilio del CAD, dei radiologi e i findings del programma: si può notare come i numeri siano paragonabili a quelli relativi all'accordo tra le coppie di radiologi. In questo senso è quindi ipotizzabile che il CAD possa

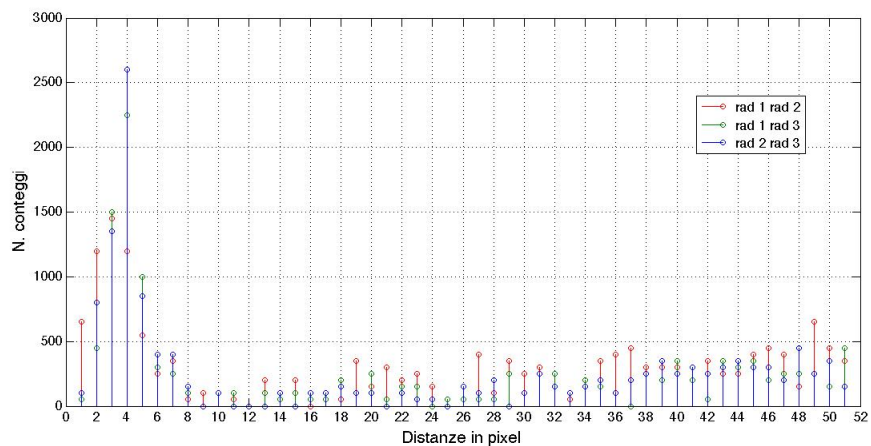


Figura 36: Distribuzione delle distanze tra i noduli annotati da ogni coppia di radiologi, utilizzando le refertazioni realizzate senza l'ausilio del CAD.

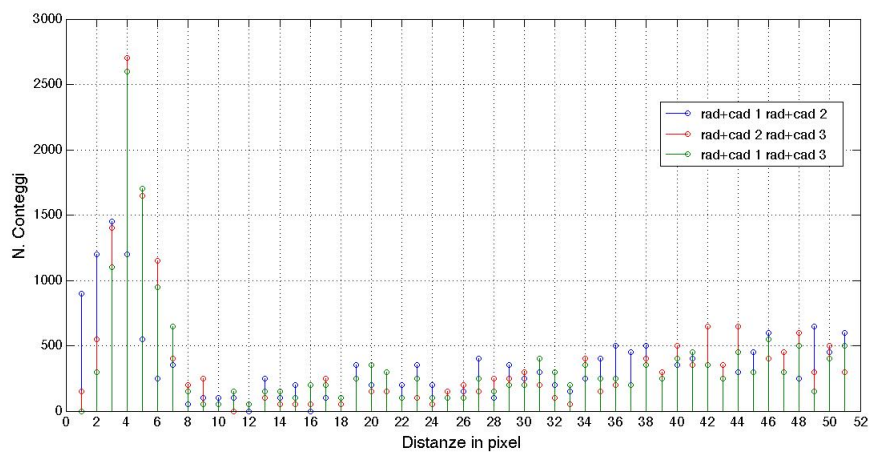


Figura 37: Distribuzione delle distanze tra i noduli annotati da ogni coppia di radiologi, utilizzando le refertazioni realizzate con l'ausilio del CAD.

CAD M5L vs rad.	Ann. Coinc.
CAD vs rad 1	131
CAD vs rad 2	142
CAD vs rad 3	183

Tabella 8: Numero di annotazioni coincidenti tra le refertazioni di ogni radiologo, eseguite senza l'ausilio del CAD M5L, e i findings dello stesso CAD.

essere considerato come un ulteriore radiologo, giustificando in parte il suo utilizzo come secondo lettore.

Il fatto che le refertazioni, a seguito dell'utilizzo del CAD, abbiano numeri di noduli annotati più paragonabili tra loro è da attribuirsi alla base comune fornita dai findings del programma.

Tale conseguenza può essere ovvia, ma l'aumento di noduli coincidenti tra le varie coppie a seguito dell'utilizzo del CAD, riportato nella *Tabella 7*, risulta meno banale.

Questo incremento indica che ogni radiologo, durante la prima analisi, non ha considerato alcune anomalie che sono comunque abbastanza significative da ricevere una valutazione analoga da almeno un altro radiologo.

Dall'ultima riga della *Tabella 7* si può notare come anche i noduli su cui tutti i radiologi concordano passino da 62 a 78, con un incremento di oltre il 25%.

L'aumento medio di accordo tra le varie coppie, a seguito dell'utilizzo del CAD, risulta essere di circa il 19%.

Come raffronto, ricordiamo i dati riportati in [32] riguardo all'accordo tra i radiologi del database LIDC-IDRI sui casi dello stesso. Su 7371 lesioni annotate come nodulo (indipendentemente dalla dimensione) da almeno un radiologo, solo 1940 risultavano avere l'accordo di tutti i quattro radiologi.

L'esperienza diretta durante le refertazioni su questi esami (soprattutto per quanto riguarda il radiologo di Cuneo, di cui è stato possibile seguire sostanzialmente tutto il lavoro) ha mostrato come, nonostante la particolare cura con cui sono stati esaminati i casi, alcuni noduli rilevanti non siano stati individuati durante la prima lettura. Solo in seguito, con l'ausilio del CAD, il radiologo ha notato la presenza del nodulo.

Questo margine di errore è comunque sempre presente indipendentemente dall'esperienza o preparazione del radiologo, ed è la motivazione che spinge

ad adottare una metodologia di doppia lettura come strumento per migliorare la sensibilità delle refertazioni [48].

6.4 Confronto con CAD commerciale

Come spiegato nel *Paragrafo* 1.9.2, il database ANODE09 è stato appositamente realizzato per permettere il confronto tra le prestazioni di CAD diversi. Nell'articolo [29] tra i vari sistemi paragonati viene preso in considerazione anche un CAD commerciale della Philips, il Lung Nodule CAD. Sistemi di questo tipo sono realizzati da diverse case produttrici di macchine CT e sono installati sulle consolle ad esse associate.

Il confronto tra le prestazioni di sistemi di questo tipo e CAD realizzati da gruppi di ricerca può rivelarsi molto interessante dato che i programmi commerciali possono essere utilizzati nella pratica clinica in quanto direttamente installati sulla postazione che viene utilizzata per la visualizzazione e la refertazione di tutti gli esami effettuati con la relativa CT.

Sebbene molti CAD commerciali possano risultare utili durante la pratica diagnostica, risultano limitati dalla mancanza delle informazioni di posizione e probabilità per ogni candidato nodulo segnalato. Questa mancanza di informazioni, come evidenziato nell'articolo, rende il confronto di non facile attuazione.

Per quanto riguarda la posizione, molti sistemi commerciali restituiscono l'output della ricerca dei noduli solo come visualizzazione a schermo. I findings sono elencati con una denominazione d'ordine generica e segnalati sulle immagini delle fette, ma non vengono riportate le coordinate della posizione.

Inoltre non viene associata ad essi nessuna probabilità. In definitiva le informazioni fornite sono di carattere qualitativo e il loro utilizzo è limitato alla consolle sulla quale vengono visualizzati tutti gli esami, non solo le CT polmonari, effettuati con quella macchina.

Oltre alla mancanza di informazioni sulla probabilità, la maggior parte dei sistemi commerciali non consente di variare la soglia di sensibilità del CAD, fissata al punto di lavoro ritenuto migliore dalla casa produttrice. Questo ne rende scomodo l'uso, in quanto non è possibile fare nessun tipo di selezione sui findings che vengono visualizzati.

Inoltre ottenere direttamente la curva FROC come descritto in precedenza risulta impossibile.

Per ovviare a questi problemi nell'articolo [29] vengono adottate due soluzioni che sono riportate di seguito. Dato che costituiscono un precedente

presente in letteratura, tali soluzioni sono state adottate anche nell'ambito di questa tesi.

Per quanto riguarda le coordinate è possibile ottenerle approssimativamente andando a ricercare le posizioni dei findings della consolle osservando gli stessi esami tramite un visualizzatore opportuno su un altro pc. La procedura deve essere fatta manualmente e risulta quindi molto laboriosa.

Per quanto riguarda invece la probabilità viene invece introdotta l'ipotesi che l'ordine con cui vengono elencati i candidati noduli dal CAD sia corrispondente alla probabilità. Questa ipotesi, per quanto non possa essere verificata, risulta compatibile anche con quanto osservato sulla consolle SIEMENS durante il lavoro descritto nel prossimo paragrafo, in quanto l'ordine assegnato ai candidati noduli non è riconducibile a nessun'altra caratteristica evidente (posizione, grandezza, ecc.) e visivamente i primi candidati sembrano essere quelli maggiormente sospetti.

Ne segue che il primo finding elencato, secondo la denominazione d'ordine resa dal CAD, verrà considerato avere livello di sospetto maggiore e probabilità pari a 1, il secondo avrà probabilità $1/2$ e così via.

Si noti che in questo modo il primo finding di ogni esame ha sempre la stessa probabilità.

In questo modo è possibile a ricavare una curva FROC: il cui primo punto sarà dato dal solo primo finding, il secondo punto dai primi due findings ecc ecc. Per come è ricavata la curva ottenuta non coincide con la curva reale del sistema con differenze imprevedibili tra le prestazioni reali e quelle stimate.

In conclusione, ricapitolando, i principali problemi riscontrabili nell'utilizzo di un CAD commerciale risultano essere:

- mancanza di informazioni su posizione e probabilità
- scarsa libertà di variare le impostazioni
- vincolo dell'utilizzo ad una data consolle

Considerando questi punti potrebbe risultare molto promettente la realizzazione di soluzioni alternative basate su sistemi sperimentali che siano utilizzabili nella quotidiana pratica diagnostica.

A questo proposito nel *Capitolo 7* viene descritto il servizio on-line di CAD per noduli polmonari, basato sulla combinazione del CAD CAM del CAD VBNA, oggetto di questa tesi. In questo modo è possibile avere a disposizione le prestazioni e i vantaggi che la combinazione di due sistemi di

questo tipo garantiscono, rendendo il programma accessibile da qualsiasi pc e in qualsiasi momento.

Il medico viene quindi svincolato dall'utilizzo di una specifica consolle e l'ospedale dall'acquisto di una postazione, che può divenire obsoleta con lo scorrere del tempo.

6.5 Analisi database ANODE09 con il CAD Lung Care SIEMENS

Al fine di confrontare i risultati della combinazione dei CAD VBNA e CAM con un prodotto commerciale, il sistema SIEMENS denominato *Lung Care* è stato testato sul database ANODE09. Questo CAD si può trovare installato sulle consolle *Leonardo syngo* della SIEMENS, associate alla TAC della stessa casa.

Due di queste consolle si trovano presso l'Ospedale di Cisanello a Pisa e presso l'Ospedale di Careggi a Firenze.

Questo sistema è stato scelto sia per la possibilità di accesso, sia perché rappresenta un buon esempio dei sistemi che normalmente usati nei reparti di radiologia.

La consolle Leonardo viene principalmente utilizzata per la visualizzazione degli esami eseguiti con la macchina CT associata. Oltre al visualizzatore, su di essa sono installati programmi di supporto specifici per diversi ambiti diagnostici.

Il CAD Lung Care è il programma dedicato alla ricerca dei noduli polmonari, e può essere utilizzato sia su esami eseguiti direttamente dalla TAC associata sia su esami caricati dall'esterno. Per praticità nei paragrafi che seguono questo sistema sarà indicato come CAD LCS, acronimo per Lung Care Siemens.

Nel nostro caso gli esami del database ANODE09 erano salvati su una memoria USB e da questa sono stati importati all'interno della memoria della consolle. Una volta importati è stato possibile visualizzarli ed utilizzare su di essi il CAD LCS.

Durante il caricamento degli esami è emersa un'elevata sensibilità alle caratteristiche del DICOM da parte della consolle: la mancanza di alcuni attributi può infatti causare problemi nel caricamento (e conseguente visualizzazione) delle fette nell'ordine corretto o nel riconoscimento stesso dei file delle immagini. Questa caratteristica è probabilmente da attribuirsi al fat-

to che la consolle è pensata soprattutto per essere utilizzata con gli esami realizzati dalla CT collegata.

Utilizzando gli esami originali del database, soltanto resi anonimi, non vi sono stati problemi sulla consolle dell'Ospedale di Careggi. Quella presente presso l'Ospedale di Cisanello invece ha continuato a mostrare un comportamento anomalo: in tutti i diversi tentativi effettuati venivano importate sistematicamente solo la metà delle fette di ogni esame, saltandone sistematicamente una ogni due.

Avendo effettuato dei confronti si presume che questo comportamento in sia dovuto ad una configurazione specifica per come vengono realizzate e salvate le serie di esami della CT del reparto, ma non è stato possibile neppure con il supporto dell'assistenza SIEMENS venire a capo della problematica.

Anche tenendo in considerazione l'eventualità di un malfunzionamento, quanto accaduto è sintomatico dei limiti di sistemi di questo tipo, pensati come "accessorio" della consolle più che come supporto efficace e flessibile al lavoro del radiologo.

Tutti i 50 casi del database ANODE09 sono stati quindi caricati sulla consolle dell'Ospedale di Careggi e sono stati visualizzati per assicurarsi che non vi fossero stati problemi di caricamento. Per ogni caso è stato quindi lanciato il CAD LCS, che ha mostrato un tempo di elaborazione di qualche decina di minuti per ognuno degli esami.

I risultati dell'elaborazione del programma sono mostrati nella schermata mostrata nella *Figura 38*.

Il riquadro in basso nell'immagine riporta la posizione di tutti i findings e consente, facendo scorrere la linea orizzontale indicata dalle 2 frecce rosse, di selezionare una sezione del polmone. La fetta selezionata ed i noduli eventualmente riscontrati dal CAD su di essa, indicati dal cerchio rosso, vengono visualizzati nel riquadro in alto a sinistra. Alzando o abbassando il cursore si scorrono le fette lungo tutto il polmone. L'intensità del colore del cerchietto indicante la posizione del nodulo diminuisce gradualmente mentre ci si allontana dalla fetta su cui è stato situato, sparendo del tutto dopo tre o quattro. Il raggio è uguale per tutti i findings, e non fornisce indicazioni sulla dimensione del nodulo.

Le scritte che si possono vedere a schermo sono legate all'identificazione del caso, l'unica informazione mostrata e la posizione della fetta visualizzata in quel momento. Non sono indicate le posizioni dei noduli e la loro probabilità, come già anticipato nel *Paragrafo 6.4*.

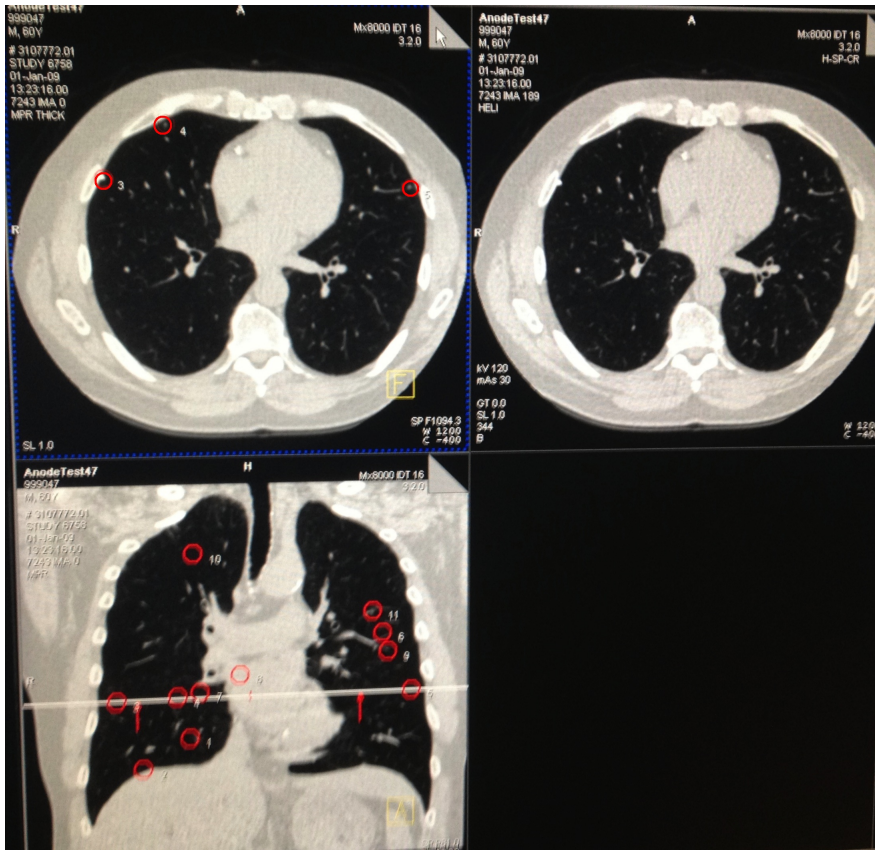


Figura 38: Schermata output del CAD LCS.

Utilizzando la posizione della fetta su un altro pc, tramite il software Osirix, sono stati visualizzati gli stessi esami e si è provveduto a cercare i punti corrispondenti a candidati noduli mostrati dal CAD LCS. Le coordinate mostrate da Osirix, indicanti approssimativamente la posizione del centro del nodulo, sono state riportate su un file di testo.

I candidati noduli sono stati trascritti seguendo l'ordine dato dai numeri visibili accanto ai cerchi nel riquadro in basso a sinistra della schermata. Seguendo quanto proposto in [29], la probabilità è stata assegnata ad ogni annotazione facendo l'inverso della sua posizione nella lista di findings di ogni caso.

Uno script apposito è stato realizzato per calcolare in questo modo le probabilità e aggiungerle sui file di testo accanto alle relative coordinate.

Nel complesso, sui 50 casi del database ANODE09, i findings trovati dal CAD LCS sono 288.

6.6 Risultati Lung Care SIEMENS (LCS) su ANODE09 e confronto con il CAD M5L

Partendo dalle liste di ogni caso, ottenute come descritto nel paragrafo precedente, tramite uno script apposito è stata realizzata la lista complessiva dei findings del CAD LCS sugli esami ANODE09. Analogamente a quanto fatto nel *Paragrafo* 6.3, prima di procedere con l'invio della lista ai curatori del database sono stati realizzati alcuni confronti per verificare la consistenza dai risultati. I confronti e le analisi descritti in seguito sono state effettuate "off-line", ovvero dopo aver acquisito tutti i dati sia per i radiologi che per il CAD LCS.

Nello specifico i findings del CAD LCS sono stati confrontati con le annotazioni dei radiologi, realizzate sia senza che con l'ausilio del CAD M5L, e con i findings ottenuti dal CAD M5L stesso.

Per definire la soglia da utilizzare per la ricerca dei noduli coincidenti, sono state nuovamente utilizzate le distribuzioni delle distanze tra le annotazioni di diverse liste, sempre come spiegato nel *Paragrafo* 6.3. In questo caso sono state confrontate le annotazioni, con e senza CAD M5L, dei radiologi rispetto ai findings del CAD LCS. I risultati sono mostrati nelle *Figure* 39 e 40.

Come si può notare confrontandole con le *Figure* 36 e 37, il minimo delle distribuzioni risulta meno definito, differendo leggermente tra la coppia dei

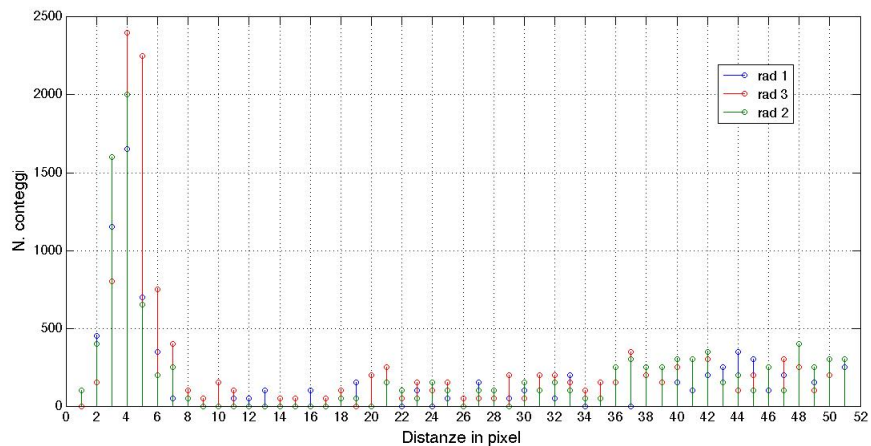


Figura 39: Confronto delle distribuzioni delle distanze tra le annotazioni di ogni radiologo, ottenute senza il CAD M5L, rispetto al CAD LCS.

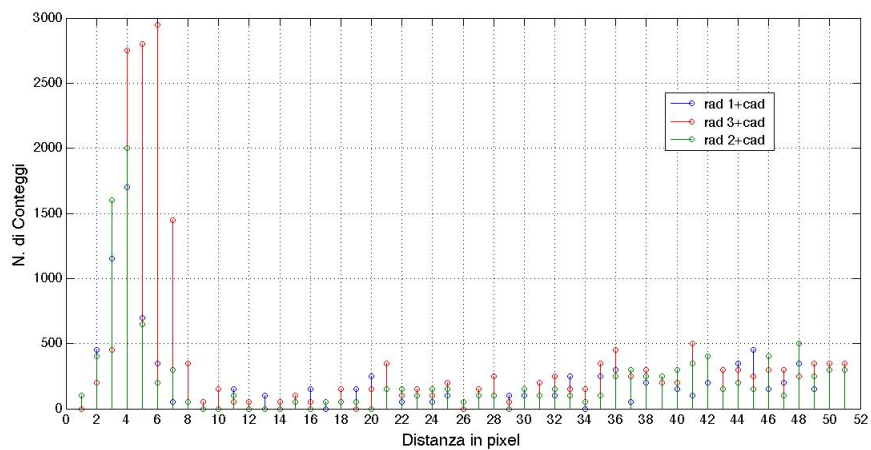


Figura 40: Confronto delle distribuzioni delle distanze tra le annotazioni di ogni radiologo, ottenute con l'ausilio del CAD M5L, rispetto al CAD LCS.

Rad. vs LCS	Coinc. (solo rad)	Coinc. (rad+M5L)
1 vs LCS	92	94
2 vs LCS	105	108
3 vs LCS	143	224

Confr. tra CAD	Find. Coinc.
CAD M5L vs LCS	238

Tabella 9: Numero di annotazioni coincidenti tra le annotazioni, realizzate senza e con l’ausilio del CAD M5L, dei radiologi e i findings del CAD LCS. Nella seconda tabella è riportato il numero di findings coincidenti tra il CAD M5L e il CAD LCS.

radiologi 1 e 2 rispetto al radiologo numero 3. Come compromesso si è scelto di utilizzare una soglia fissata a 10pixel . Abbiamo comunque verificato, come risulta anche ipotizzabile osservando le differenze di conteggi tra massimo e minimo nei vari grafici, che la scelta della distanza di soglia (restando nei pressi del minimo) modifica in maniera minima i risultati ottenuti. Una differenza di 2pixel su tale valore comporta una variazione, in media, di 2 noduli nel computo totale dei coincidenti.

I numeri di noduli coincidenti tra le annotazioni dei radiologi e i findings del CAD LCS sono raccolti nella *Tabella 9*.

Nell’ultima riga della tabella è riportato il numero di findings coincidenti tra quelli trovati dal CAD LCS e dal CAD M5L: come si può notare risultano coincidenti 238 dei 288 candidati noduli trovati dal CAD LCS, pari ad una percentuale del $\simeq 83\%$.

Ricordiamo che i findings del CAD M5L sui 50 casi ANODE09 sono 588, numero ottenuto applicando una soglia sulla probabilità al 5%.

Un tale livello di sovrapposizione tra i due CAD è significativo, in quanto utilizzando il CAD M5L si tiene comunque conto della grande maggioranza dei findings del CAD LCS. In tale considerazione bisogna comunque ricordare la grande differenza nel numero complessivo di findings tra i due CAD.

Una possibile ipotesi è che il CAD LCS si limiti ad evidenziare solamente i findings che presentano, in qualche misura, un livello di probabilità maggiore.

Data la limitatezza delle informazioni disponibili sul dettaglio del suo funzionamento e la scarsa libertà di impostazioni, tale ipotesi non è direttamente verificabile.

Per contro, nei confronti con le annotazioni dei radiologi, riportati nella

Tabella 9, si può notare come il livello di accordo sia inferiore rispetto ai valori analoghi mostrati nella *Tabella 8* (*Paragrafo 6.3*).

Questi risultati sembrano indicare una minore sensibilità verso una parte dei noduli riconosciuti da almeno un radiologo.

7 Servizio di CAD on-line

7.1 Introduzione

Nei paragrafi che seguono vengono descritti la realizzazione ed il funzionamento del servizio di CAD on-line basato sul CAD M5L, costituito dalla combinazione dei CAD CAM e VBNA, accessibile all'indirizzo <https://mag03xl.to.infn.it/m5l/user>.

Il servizio è stato realizzato utilizzando un approccio di “*cloud computing*”, che consente ad ogni utente di accedervi utilizzando solo una connessione ad Internet, da qualsiasi computer e senza dover installare nessun software.

Questo approccio rappresenta un punto di originalità rispetto ai normali metodi di impiego di CAD in ambito clinico, favorendo semplicità e flessibilità di utilizzo.

Una descrizione più dettagliata di cosa si intende con “*cloud computing*” è fornita nel prossimo paragrafo.

Riportiamo inoltre la descrizione del lavoro fatto per realizzare l'implementazione dei CAD come servizio automatico e le motivazioni che hanno portato alla realizzazione di questo progetto.

Sono infine descritti i test di utilizzo effettuati da diverse strutture ospedaliere utilizzando esami CT acquisiti nei rispettivi reparti di radiologia. I risultati ottenuti da questi test forniscono un'indicazione del funzionamento e dell'efficacia del servizio come strumento di supporto al lavoro dei radiologi.

7.2 Cloud Computing

Il *cloud computing* è un'evoluzione relativamente recente dell'informatica nell'ambito della condivisione delle risorse tramite l'utilizzo della rete [63, 62].

In prima approssimazione è possibile considerare il *cloud computing* come un'estensione di quanto è possibile realizzare con diversi computers collegati all'interno di una rete locale. E' noto che in tale configurazione un dispositivo può accedere ai dati presenti sugli altri ed anche utilizzare la loro capacità di calcolo.

In una cosiddetta *cloud* un gestore fornisce al cliente la possibilità di accedere a risorse di archiviazione ed elaborazione semplicemente tramite una connessione ad internet, senza necessità di possedere risorse fisiche o software.

Scendendo nello specifico, solitamente tra il soggetto che fornisce le risorse hardware (ovvero i server) e l'utente finale si inserisce un terzo soggetto,

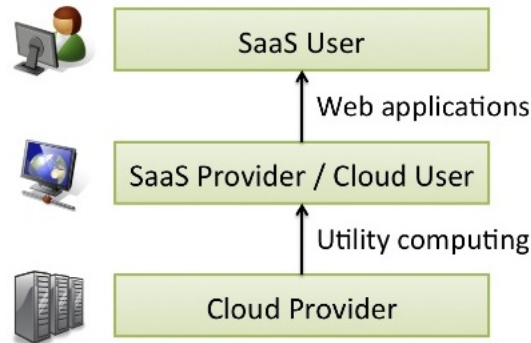


Figura 41: Schema della struttura di un servizio cloud, con i vari soggetti coinvolti [64].

che utilizza le risorse hardware per fornire un servizio che viene utilizzato dall'utente finale.

Solitamente infatti chi possiede le risorse hardware non gestisce direttamente nessun servizio, ma si limita a fornire i server al soggetto intermedio.

A livello di terminologia si possono introdurre le seguenti definizioni [64]:

- “*Cloud*” è definito l’insieme di hardware e software forniti dal soggetto proprietario dei server (o “*cloud provider*”). L’insieme di queste risorse è anche chiamato “*cloud farm*”
- “*Utility Computing*” è definito il servizio venduto, inteso come possibilità di accesso ai server e ai software situati su di essi, da chi possiede i server al soggetto intermedio
- SaaS (“*Software as a Service*”) è definito il servizio fornito dal soggetto intermedio all’utente finale utilizzando le risorse della cloud. Sostanzialmente con questo termine si intende l’utilizzo in remoto, da parte dell’utente finale, di software installati sulla cloud, realizzati e gestiti dal soggetto intermedio.

Nella *Figura 41* è schematizzato il rapporto gerarchico tra i vari utenti.

Utilizzare un approccio di cloud computing per fornire un servizio tramite internet permette di ottenere diversi vantaggi. Il primo, già evidenziato, è che l’utente finale è svincolato dalla necessità di possedere le risorse hardware e software per provvedere direttamente alle proprie esigenze.

Analogamente, anche per il soggetto intermedio non è più necessario possedere direttamente i server che utilizza.

Per quanto riguarda il cloud provider, il vantaggio risiede nella possibilità di utilizzare un'economia di scala, offrendo le risorse hardware e software in suo possesso a più soggetti intermedi che le utilizzino in momenti diversi.

La chiave per poter gestire in questo modo i server è la possibilità di riallocare dinamicamente le risorse a seconda di come vengono richieste. In questo modo sia il soggetto intermedio che l'utente finale possono disporre potenzialmente di tutti i server in un dato momento, con i relativi vantaggi in termini di potenza e rapidità di calcolo, e successivamente le stesse risorse possono essere utilizzate da altri.

Per poter riallocare dinamicamente le risorse si ricorre alla virtualizzazione [65], ovvero all'utilizzo di “*virtual machines*” analoghe a quelle utilizzate comunemente per simulare, nel caso sia necessario utilizzare applicazioni che non siano supportate, un sistema operativo differente da quello del pc su cui si sta lavorando. A seconda delle necessità degli utenti vengono create macchine virtuali con le caratteristiche richieste che vengono poi eliminate una volta terminato l'utilizzo, liberando le risorse hardware per altre applicazioni.

Per quanto riguarda il servizio di CAD on-line descritto nei capitoli successivi, i soggetti coinvolti sono i seguenti:

- Cloud Provider è il Centro di Calcolo della sezione di Torino dell'INFN, che ha scelto di utilizzare un approccio di Cloud Computing per la gestione della propria infrastruttura di “Calcolo Scientifico”. Tale infrastruttura fornisce diverse applicazioni, tra cui servizi per esperimenti in corso al CERN ed altri [25].
- Soggetto intermedio attualmente può essere considerato l'insieme di persone che ha curato l'implementazione dei due CAD, la realizzazione del sito web e che sta seguendo i test del funzionamento del servizio.
- Utenti finali del servizio sono gli ospedali, nello specifico i medici radiologi che si occupano di diagnosi per patologie tumorali al polmone tramite esami CT.

Utilizzando un'interfaccia web (disponibile all'indirizzo <https://mag03xl.to.infn.it/m5l/user>) gli utenti del servizio possono caricare sulla cloud gli esami che vogliono analizzare ai CAD, ricevendo sulla stessa interfaccia i risultati di tale analisi.

7.3 Implementazione CAD CAM e CAD VBNA in cloud

Utilizzando i server messi a disposizione dall'INFN di Torino è stato possibile implementare in cloud il CAD CAM e il CAD VBNA. Su un server apposito sono stati quindi copiati tutti gli script necessari al funzionamento dei due programmi singolarmente e quelli necessari per ottenere la combinazione dei risultati.

La natura di questi script ed il loro funzionamento è stato descritto in precedenza nei relativi capitoli. Nessuna variazione di rilievo è stata effettuata su di essi; è stato sufficiente modificare i vari percorsi utilizzati per renderli consistenti con le nuove collocazioni di cartelle e file.

Poichè questi programmi erano stati realizzati per essere utilizzati manualmente, la parte principale del lavoro di implementazione è stata la realizzazione di una serie di script che rendano automatiche tutte le procedure.

In questo modo per ogni esame inviato in ingresso, in automatico viene eseguita l'analisi da parte di entrambi i CAD e la combinazione dei findings. I risultati complessivi sono quindi resi sia sotto forma di *.pdf* che di *.xml* per la visualizzazione su Osirix.

Come citato nel paragrafo precedente, gli esami possono essere sottoposti all'analisi dei CAD utilizzando un'interfaccia web su un apposito sito. L'interfaccia web, e la gestione dei file tramite essa, è stata realizzata nell'ambito della collaborazione M5L dell'INFN e sarà descritta più nel dettaglio nel *Paragrafo 7.5*.

Una volta che l'esame è stato caricato sul sito viene salvato all'interno di una apposita cartella sul server della cloud e vengono lanciati gli script per l'esecuzione automatica sia del CAD CAM che del CAD VBNA.

Per quanto riguarda il CAD VBNA, gli script citati nel *Paragrafo 4.3* sono richiamati da uno script dedicato prima per il modulo relativo ai noduli interni e successivamente per quello relativo ai pleurici. Al termine delle procedure un altro script controlla che entrambi i moduli abbiano prodotto una lista di findings e richiama quanto necessario ad effettuare la normalizzazione e combinazione, come descritto nel *Paragrafo 4.4*.

Avendo a disposizione i risultati del CAD VBNA sul database LIDC-IDRI, i dizionari di normalizzazione sono stati realizzati utilizzando le curve FROC ottenute separatamente dai moduli CAD_I e CAD_{JP} su questi casi. L'ampiezza e l'eterogeneità di questo database consentono di aumentare la statistica utilizzata per il calcolo delle probabilità normalizzate, secondo l'espressione (12), introdotta nel *Paragrafo 2.6*.

Realizzata la combinazione dei due moduli, i findings complessivi del CAD VBNA sono scritti in un file di testo.

Analogamente a quanto avviene per il CAD VBNA uno script apposito gestisce l'esecuzione del CAD CAM.

Al termine dell'elaborazione dei due CAD i risultati sono combinati da uno script analogo a quello utilizzato per i due moduli del CAD VBNA, ottenendo la lista complessiva che viene resa come output. Viene anche effettuato nuovamente un controllo per verificare che entrambi i programmi abbiano prodotto una lista di findings. In caso contrario la lista mancante viene sostituita da un file vuoto e un messaggio di errore viene salvato in un opportuno file di testo. In questo modo la combinazione può comunque procedere ma è possibile controllare quale sia la natura della problematica.

Per le stesse motivazioni descritte nel caso di CAD_I e CAD_{JP} , il dizionario di normalizzazione per la combinazione di CAM e VBNA è stato realizzato utilizzando la curva FROC ottenuta dal CAD VBNA su tutti i 1018 casi del database LIDC-IDRI.

La lista complessiva dei findings viene sia salvata all'interno di un file di testo sia convertita in formato *.xml*, in modo che sia possibile visualizzarla utilizzando Osirix.

Come ultima operazione, partendo dalla lista dei candidati noduli, viene creato un file *.pdf* che riporta i findings con probabilità maggiore o uguale al 25%.

Questi sono indicati sia con le coordinate sia con una piccola immagine della loro posizione all'interno della fetta corrispondente. Un esempio è mostrato nella *Figura 42*.

La soglia del 25% di probabilità corrisponde, per la curva FROC riportata nella *Figura 33 (Paragrafo 5.4)*, a considerare $6^{FP/CT}$.

Sul file realizzato per Osirix invece sono riportati tutti i noduli; utilizzando l'apposito plugin è possibile variare successivamente la soglia che si vuole utilizzare sulla probabilità, come è possibile vedere nella *Figura 43*.

Il file *.pdf* ed il file *.xml* sono compressi in formato *.zip* ed inviati al server del sito web, su cui vengono decompressi. La realizzazione della cartella compressa e il suo trasferimento viene effettuata attraverso uno script dedicato.

L'utente finale, avvisato tramite e-mail della fine dell'analisi, può quindi visualizzare o scaricare i risultati utilizzando l'interfaccia web.

Per i soli amministratori del servizio, sul sito sono anche accessibili alcuni report per la verifica degli errori in caso di eventuali problemi.

Insieme ai vari scripts per la gestione dei file e dell'analisi è stato realizzato un file che, in automatico, permette di installare il necessario per il funzionamento dei CAD ogni volta che viene creata una macchina virtuale all'interno della cloud o è necessario intervenire sul server stesso.

7.4 Motivazioni della realizzazione del servizio web

In questo paragrafo sono descritte le motivazioni che hanno portato alla realizzazione di un servizio on-line di CAD in cloud, metto in evidenza quelli che riteniamo essere i principali punti di interesse.

Come più volte sottolineato nel *Capitolo 6* i CAD commerciali presentano limitazioni a livello di informazioni rese, libertà di impostazioni e vincolo di utilizzo su una data consolle.

L'idea alla base della realizzazione del servizio di CAD on-line è di andare a sopperire ad alcune delle mancanze e limitazioni mostrate dai sistemi normalmente utilizzati nei reparti di radiologia, aggiungendo inoltre alcuni aspetti innovativi.

Un primo punto di interesse è dato dal fatto che vengono utilizzati due sistemi combinati tra loro, a differenza dei normali CAD commerciali dove viene utilizzato un singolo programma. I vantaggi forniti da questa configurazione sono stati discussi nel *Capitolo 5*, e possono essere sintetizzati come un aumento delle prestazioni (rispetto ai singoli sistemi) dovuto alla complementarità dei rispettivi approcci alla ricerca dei noduli.

Bisogna inoltre considerare che la validazione effettuata sul database LIDC-IDRI permette di utilizzare i risultati ottenuti, nelle modalità descritte nel paragrafo precedente, anche durante l'analisi di esami provenienti dalla pratica clinica come base per il calcolo delle probabilità dei candidati noduli. In questo modo l'ampiezza del database LIDC-IDRI fornisce stabilità all'analisi, che può risultare corretta per un ampio insieme di anatomie ed espressioni della patologia.

La struttura del servizio tramite web rende anche possibile in futuro l'eventuale aggiunta di altri algoritmi, con i relativi miglioramenti delle prestazioni complessive, e l'addestramento del software su un numero più consistente di casi.

Un secondo punto di interesse è dato dal fatto che l'utilizzo del cloud computing, come già spiegato nel *Paragrafo 7.2*, svincola l'utente del servizio dall'acquisto di postazioni specifiche o computers dedicati. Non risulta neanche necessario l'acquisto di nessun software addizionale, è sufficiente

1. Tracking Identifier = "Nodule # 1"

Properties:

- Center: (323,261) Radius=15
- Slice: 18, Z-Position: -267.2
- Certainty of Feature: 95%



Figura 42: Esempio di come sono riportati i findings all'interno del report .pdf.

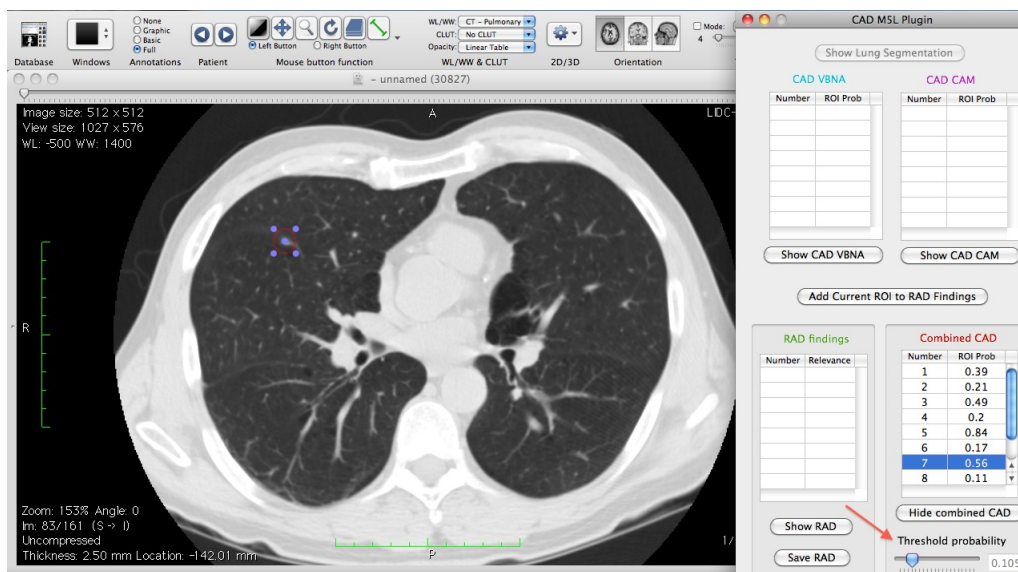


Figura 43: Esempio di visualizzazione dei risultati del CAD on-line tramite l'apposito plugin di Osirix. La freccia indica il cursore che permette di variare la soglia sulla probabilità per la visualizzazione dei findings.

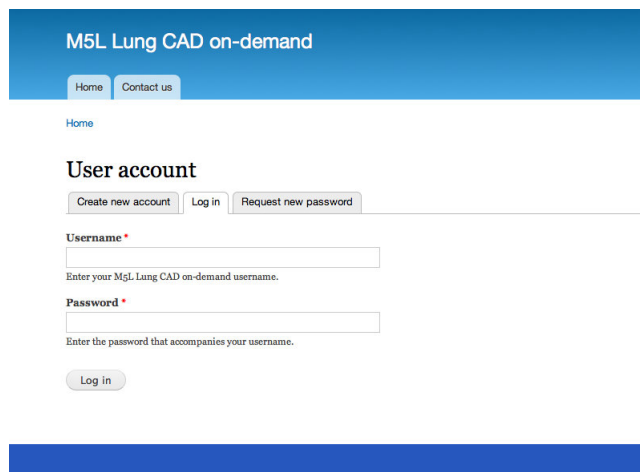
possedere un visualizzatore DICOM (software comunque di comune utilizzo e spesso, come nel caso di Osirix, free) se si vogliono visualizzare direttamente i findings sui relativi esami. Al momento questa possibilità, dato che il servizio è in fase sperimentale, è fornita per il solo programma Osirix, ma in futuro sarà possibile espandere l'opzione anche ad altri programmi.

Un terzo punto di interesse, infine, è la possibilità di accedere al servizio attraverso un qualsiasi pc dotato di connessione ad internet. Questo permette al radiologo, per esempio, di inviare esami e ricevere i risultati su computer diversi, senza dover per forza utilizzare una specifica postazione in reparto. I risultati restano comunque accessibili sul sito in qualsiasi momento e possono quindi anche essere condivisi tra più persone, fornendo una grande flessibilità al lavoro del medico.

7.5 Utilizzo del servizio web

Il servizio web è accessibile all'indirizzo <https://mag03xl.to.infn.it/m5l/user>, da cui è possibile fare il login tramite la finestra mostrata nella *Figura 44*.

Nel caso in cui non si sia ancora registrati al servizio è possibile utilizzare la stessa pagina per registrarsi e richiedere l'autorizzazione all'utilizzo. Diverse e-mail saranno inviate all'utente per segnalare l'invio della richiesta delle credenziali di accesso e l'avvenuta autorizzazione.



The image shows a web interface for 'M5L Lung CAD on-demand'. At the top, there is a blue header with the text 'M5L Lung CAD on-demand' and two buttons: 'Home' and 'Contact us'. Below the header, there is a 'Home' link. The main section is titled 'User account' and contains three buttons: 'Create new account', 'Log in', and 'Request new password'. Below these buttons are two input fields: 'Username *' and 'Password *'. The 'Username *' field has a placeholder text 'Enter your M5L Lung CAD on-demand username.' and the 'Password *' field has a placeholder text 'Enter the password that accompanies your username.' At the bottom of the form is a 'Log in' button.

Figura 44: Pagina di accesso del sito per l'utilizzo del servizio di CAD on-line.

Una volta ottenute le credenziali ed effettuato l'accesso, all'interno del sito sono disponibili diverse sezioni.

Utilizzando la pagina denominata “*Submit*” è possibile compiere l'upload di un esame CT che si vuole far analizzare. La *Figura 45* mostra la pagina evidenziando i campi da compilare.

Nel campo indicato come 1 è necessario inserire un'identificazione per il caso che si intende caricare, utilizzando solo caratteri alfanumerici senza spazi. Tramite il pulsante 2 è invece possibile selezionare, attraverso un menù, l'esame direttamente da dove è salvato sul pc che si sta utilizzando. Tutte le fette che compongono un singolo esame devono essere contenute in una cartella compressa in formato *.zip*, non è possibile utilizzare cartelle che contengono più di una serie di immagini.

Utilizzando i testi 3 e 4 in questo ordine è quindi possibile prima eseguire l'upload e quindi, terminato il caricamento, sottomettere il caso al CAD. Un messaggio a schermo indica se quest'ultima operazione è avvenuta correttamente.

E' importante specificare che durante la procedura di upload gli esami vengono anonimizzati, in rispetto della privacy.

Spostandosi nella sezione indicata come “*My CASES*” è possibile trovare l'elenco di tutti gli esami che sono stati inviati al CAD con i relativi risultati, come mostrato dalla *Figura 46*.

Fintanto che un esame è in elaborazione nella casella indicata come “*status*” viene visualizzato il messaggio “*in progress*”. Appena il programma termina l'elaborazione l'utente viene informato tramite mail. Utilizzando il link contenuto nella mail o tornato alla pagina del sito è quindi possibile accedere ai risultati dell'analisi. Sempre nella casella “*status*” viene indicato se l'analisi è terminata in modo corretto o si sono verificati degli errori.

Nella *Figura 46* con le due frecce rosse sono evidenziati i file *.xml* e *.pdf* contenenti i risultati, che sono quelli di maggiore interesse per l'utente. Entrambi i file possono essere scaricati, il file *.pdf* inoltre può anche essere visualizzato direttamente.

La *Figura 42* mostra un esempio di come sono visualizzati i findings all'interno di questo report: per ogni candidato nodulo vengono mostrate le coordinate del centroide, l'immagine della fetta su cui è stato identificato e uno zoom intorno alla porzione che lo contiene.

Ricordiamo che nel report, per motivi di spazio, vengono mostrati solo i findings con probabilità $\geq 25\%$.

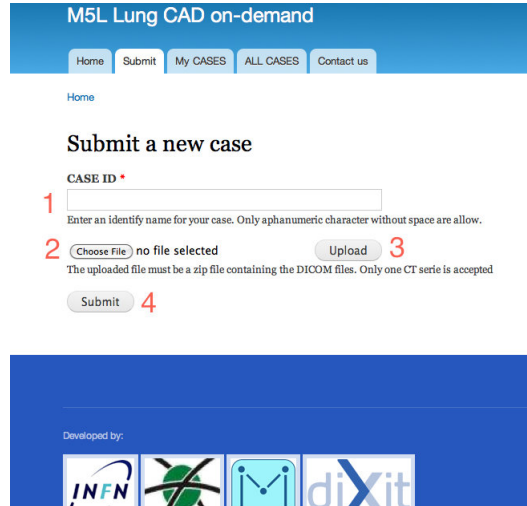


Figura 45: Pagina per l'upload di esami da analizzare attraverso il servizio web.

CASE ID	Submitter	Date	Upload file	Status	Results
torace_20_pulito	marco	12/30/2013 - 18:01	DICOM zip Open with Osirix	Success	<ul style="list-style-type: none"> DICOM Chest CAD SR CAD Report (HTML) CAD Report (PDF) Osirix plugin (XML)
torace_19_pulito	marco	12/30/2013 - 13:06	DICOM zip Open with Osirix	Success	<ul style="list-style-type: none"> DICOM Chest CAD SR CAD Report (HTML) CAD Report (PDF) Osirix plugin (XML)
torace_18_pulito	marco	12/27/2013 - 11:50	DICOM zip Open with Osirix	Success	<ul style="list-style-type: none"> DICOM Chest CAD SR CAD Report (HTML) CAD Report (PDF) Osirix plugin (XML)

Figura 46: Pagina di riepilogo dei casi inviati dall'utente con relativi risultati. Le frecce rosse indicano le liste di findings in formato *.pdf* e *.xml*.

Per visualizzare i findings del CAD direttamente sull'esame tramite Osirix è sufficiente creare una cartella denominata "ROIs", allo stesso livello di quella contenente le fette dell'esame, e scaricarvi dentro il file *.xml*. Successivamente importando entrambe le cartelle su Osirix è possibile osservare contemporaneamente esame e findings utilizzando l'apposito plugin, come mostrato nella *Figura 43*.

Le altre voci tra i risultati dell'esame, visibili nella *Figura 46* non indicati dalle frecce, sono l'analogo del report *.pdf* in formato *.html* e un file contenente informazioni riguardanti gli attributi del DICOM.

Consideriamo infine le voci contenute nella sezione indicata come "Upload file", che fanno riferimento ai file dell'esame che è stato inviato al CAD. Utilizzando la funzione "Dicom zip" è possibile scaricare dal sito la cartella compressa contenente le fette che compongono l'esame, mentre utilizzando il pulsante "Open with Osirix" è possibile aprire l'esame direttamente all'interno di questo visualizzatore DICOM.

In questo modo è possibile visualizzare l'esame anche se non lo si ha in origine sul computer che si sta utilizzando, favorendo flessibilità di lavoro e condivisione.

Per visualizzare i findings del CAD è comunque necessario utilizzare la procedura esposta poco sopra.

Dato che al momento il servizio è in fase sperimentale, con relativamente pochi utenti e impostazioni della cloud che potranno subire variazioni, è difficile fornire valutazioni corrette su quelli che potranno essere i tempi necessari ad eseguire l'analisi di un esame in futuro.

Al momento inviando un singolo esame sono necessari pochi minuti per l'upload, a seconda della connessione che si utilizza, e una decina per l'analisi. In media è possibile stimare 15 minuti tra l'inizio della procedura e l'arrivo dell'email di notifica.

Poichè tale intervallo di tempo è comunque paragonabile con quello necessario ad un radiologo per eseguire la refertazione, può essere ipotizzabile un utilizzo in parallelo, rendendo disponibili i findings del CAD al termine dell'analisi del medico.

7.6 Test del servizio web

Al fine di verificare la possibilità di utilizzo e l'efficacia del servizio web come effettivo strumento di supporto al lavoro dei radiologi, sono stati realizzati alcuni test presso diversi centri ospedalieri.

L'idea è stata di utilizzare esami provenienti direttamente dalla pratica clinica, realizzati quindi all'interno delle strutture stesse, da sottoporre al CAD M5L tramite la procedura spiegata nel paragrafo precedente.

Il primo test è stato effettuato presso l'Ospedale Sante Croce di Cuneo, grazie alla collaborazione della Fisica Sanitaria e di un medico radiologo. Utilizzando il servizio web sono stati inviati all'analisi del CAD 14 CT di pazienti dell'ospedale.

Gli esami provengono da un progetto di screening per tumore al polmone realizzato all'interno della struttura. I pazienti partecipanti sono stati sottoposti ad una tomosintesi [46] come esame di "baseline" e successivamente, in caso di esito positivo, sono stati richiamati per due CT di verifica.

Gli esami utilizzati provengono dal primo round di verifica e sono quindi forniti sia della refertazione realizzata per la tomosintesi, sia di quella relativa all'esame CT stesso.

In origine il numero di esami presi in considerazione era maggiore di 14, ma alcuni sono stati esclusi successivamente in quanto acquisiti in condizioni particolari o poichè presentavano problemi tecnici. Sui 14 casi presi in considerazione, l'analisi del CAD è sempre stata portata a termine con successo.

In media, ogni esame è costituito da circa 70 sezioni, con uno spessore di fetta di $5mm$ e pixel di $0,71mm$.

Dopo l'invio e l'elaborazione di tutti gli esami tramite il servizio web, i risultati sono stati controllati con l'ausilio del radiologo. I findings del CAD sono stati sia confrontati con le refertazioni di CT e tomosintesi, sia controllati singolarmente.

I findings del CAD sono stati visualizzati attraverso l'apposito plugin di Osirix, mostrato nella *Figura 43*. Come è visibile nella figura, tramite il plugin è possibile variare la soglia della probabilità per i findings che vengono visualizzati, mentre nel report *.pdf* dei risultati la soglia è fissa al 25%.

Per coerenza tra le due modalità di visualizzazione, si è scelto di utilizzare la stessa soglia per la selezione dei findings attraverso il plugin.

Questa scelta trova comunque, a posteriori, anche un riscontro pratico: fissando la soglia al 25% il numero di findings mostrato, rispetto a qualche centinaio trovati in media per ogni esame, risulta essere dell'ordine di 15 – 20. Nell'ottica dell'utilizzo del servizio web nella normale pratica diagnostica, un tale numero di findings da esaminare può risultare un buon compromesso tra la possibilità di miglioramento della refertazione e il tempo richiesto per eseguirla.

Un numero maggiore di findings verosimilmente, a fronte possibili miglioramenti della refertazione, potrebbe essere poco gradito dal medico chiamato a verificarli e indurlo a terminare la verifica prima di aver esaminato tutti i findings.

Per quanto riguarda i 14 esami considerati, poichè sono stati chiamati ad eseguire la CT soltanto i pazienti trovati positivi tramite la tomosintesi di baseline, quasi tutti i soggetti presentano uno o più noduli chiaramente identificabili. E' stato quindi generalmente possibile accoppiare senza ambiguità i noduli annotati nelle refertazioni con i rispettivi findings del CAD. Nei pochi casi dubbi, un paio sul totale dei noduli, si è fatto riferimento all'opinione del radiologo presente.

Dopo aver verificato se i noduli annotati nelle refertazioni fossero stati individuati dal CAD, i restanti findings sono stati esaminati uno ad uno dal radiologo.

La *Tabella 10* e la *Tabella 11* riportano i risultati ottenuti.

Nella *Tabella 10* sono riportati i findings, con probabilità maggiore del 25%, trovati dal CAD per ogni caso. Ogni candidato nodulo è stato conteggiato come vero positivo (TP) se corrispondeva con un nodulo annotato nelle refertazioni o è stato giudicato come tale dal radiologo, altrimenti è stato considerato come falso positivo (FP).

In media, ogni esame CT contiene $2TP$ e $11FP$.

Tali valori, in linea con quanto discusso sopra, in prima analisi possono risultare ragionevoli per utilizzare il servizio come supporto alle refertazioni senza richiedere un eccessivo aumento del tempo di refertazione.

La *Tabella 11* riporta invece, per ogni esame, rispettivamente il numero di noduli annotati nelle refertazioni, quello di findings che sono risultati coincidenti con uno di essi, il numero di findings relativi a noduli non presenti nelle refertazioni ma giudicati essere tali dal radiologo e, infine, il numero di noduli presenti nelle refertazioni ma non trovati dal CAD.

Come si può osservare sono stati trovati 21 dei 27 noduli segnalati all'interno delle refertazioni, pari ad una percentuale di $\simeq 78\%$.

In aggiunta il CAD ha segnalato 8 noduli, confermati come tali dal radiologo, non presenti all'interno delle refertazioni.

In sei casi invece un nodulo annotato dai radiologi non è stato trovato dal CAD. Su questi noduli è possibile però fare alcune considerazioni.

Due di essi sono stati effettivamente trovati dal CAD, ma con una probabilità molto bassa ($< 10\%$). Questo ha reso di fatto impossibile il loro riscontro, se non controllando un numero molto alto di findings. In un al-

CASO	N. Find (>25%)	TP	FP
1	5	1	4
2	12	1	11
3	9	1	8
4	11	2	9
5	7	1	6
6	13	1	12
7	17	2	15
8	10	0	10
9	32	2	30
10	23	8	15
11	8	1	7
12	7	4	3
13	14	3	11
14	13	1	12
Somme	181	28	153

Tabella 10: Tabella di sintesi dei risultati ottenuti tramite il servizio web di CAD sui 14 esami forniti dall’Ospedale di Cuneo. La seconda colonna riporta il numero di findings con probabilità > 25% trovati dal CAD, la terza e la quarta quanti si sono rivelati essere veri positivi e quanti falsi positivi.

CASO	Nod. Ann.	Nod. Conf.	Nod. Agg.	Nod. Persi
1	1	1	0	0
2	1	1	0	0
3	1	1	0	0
4	1	1	1	0
5	2	1	1	1
6	1	1	0	0
7	2	2	0	0
8	1	0	0	1
9	2	1	1	1
10	6	6	2	0
11	1	1	0	0
12	4	3	1	1
13	3	1	2	2
14	1	1	0	0
Somme	27	21	8	6

Tabella 11: Tabella di sintesi dei risultati ottenuti tramite il servizio web di CAD sui 14 esami forniti dall’Ospedale di Cuneo. Per ogni caso sono riportati il numero di findings che sono stati confermati tramite le refertazioni, il numero di noduli trovati in aggiunta alle refertazioni e infine quello dei noduli refertati che però non sono stati trovati dal CAD.

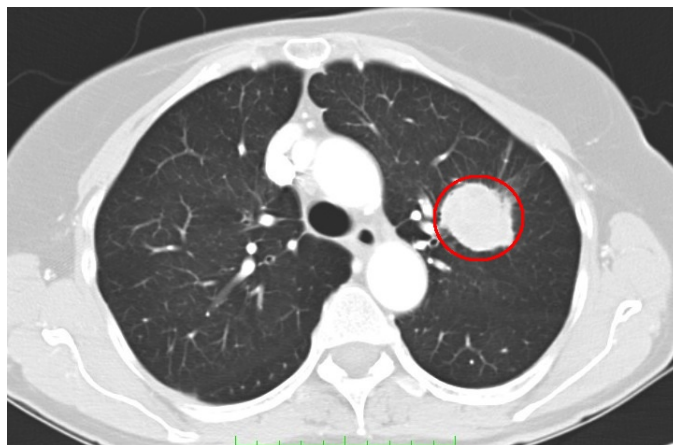


Figura 47: Esame, realizzato presso l’Ospedale di Cuneo, con nodulo di dimensioni macroscopiche non riconosciuto dal CAD.

tro esame è stato riscontrato un nodulo, più precisamente una opacità, di dimensioni macroscopicamente maggiori rispetto alla media, come possibile osservare nella *Figura 47*.

Le motivazioni per cui tale nodulo non è stato rilevato sono diverse per i due sistemi. Nel caso del CAD VBNA il nodulo è stato escluso nella segmentazione effettuata per la riduzione di falsi positivi dove, come spiegato nel *Paragrafo 2.5*, sono eliminati i candidati noduli troppo grandi o troppo piccoli.

Per quanto riguarda invece il CAD CAM il nodulo, date le dimensioni, è stato escluso durante la fase di segmentazione iniziale insieme alle strutture anatomiche.

Il mancato rilevamento non pregiudica comunque la diagnosi in quanto è decisamente inverosimile che un’anomalia di tali dimensioni non venga riconosciuta da un radiologo.

Infine un ulteriore nodulo “mancato” dal CAD è stato valutato come non rilevante dal radiologo.

Complessivamente, ricapitolando, risulta che il totale di noduli presenti all’interno degli esami è 35: 27 annotati all’interno delle refertazioni e 8 aggiunti in seguito all’analisi con il CAD.

Avendo riconosciuto 29 noduli su 35, per il CAD è possibile stimare una sensibilità di $\simeq 83\%$ a $11^{FPs}/scan$. Quella dei radiologi, calcolata dalle refertazioni, risulta invece del 77% riferita a $0^{FPs}/scan$ per convenzione.

Tali risultati sono molto interessanti, a maggior ragione considerando la doppia refertazione effettuata su ogni esame tra baseline e CT. Dal confronto dei valori numerici di conteggi e sensibilità risulta come il CAD possa essere sostanzialmente considerato a livello di un radiologo. Questa considerazione è già stata formulata all'interno del *Paragrafo 6.3* e trova qui maggiore sostegno.

Inoltre, la complementarità tra CAD e radiologi riscontrata su un porzione significativa di noduli (circa il 20% del totale) avvalorava l'utilizzo del programma come seconda lettura per il miglioramento delle refertazioni dei radiologi.

Si noti che il valore $\simeq 83\%$ a $11^{FPs}/scan$ ottenuto per la sensibilità è sostanzialmente in linea a quanto ottenuto sul database LIDC-IDRI e riportato nel *Paragrafo 5.4*.

Nel complesso i risultati attualmente ottenuti sono stati valutati positivamente da radiologi provenienti da diverse strutture, che hanno fornito disponibilità per proseguire ed ampliare il lavoro di test del sistema.

I centri che hanno fino ad ora aderito a questa fase di sperimentazione sono l'Ospedale di Cisanello a Pisa, il Niguarda di Milano e il Fatebenefratelli dell'Isola Tiberina a Roma.

L'Ospedale di Cuneo sta intanto procedendo con l'upload di esami di esami sempre provenienti dal programma di screening tramite tomosintesi, per un totale di 132 casi disponibili.

Parte IV:
Conclusioni

8 Conclusioni

Il lavoro svolto in questa tesi riguarda il test di due sistemi CAD per l'individuazione automatica di noduli polmonari all'interno di esami CT, il CAD VBNA e il CAD CAM, e l'implementazione della loro combinazione (CAD M5L) all'interno di un servizio web di CAD on-line.

Dalla validazione del CAD VBNA e del CAD M5L sui 1018 casi del database internazionale LIDC-IDRI sono emersi risultati interessanti, sia a livello di sensibilità sia a livello della stabilità di funzionamento. Nello specifico i valori di sensibilità ottenuti dal CAD M5L sono molto buoni, confermando come la combinazione di sistemi con modalità di funzionamento diverse permetta un incremento delle prestazioni rispetto agli stessi sistemi presi singolarmente.

Il test clinico di utilizzo del CAD M5L come secondo lettore sui 50 casi del database ANODE09 ha invece evidenziato come, a seguito dell'utilizzo del CAD, siano stati rilevati noduli significativi non riconosciuti dai radiologi durante la prima refertazione.

Dal confronto, sempre sui 50 casi del database ANODE09, tra il CAD M5L e il CAD commerciale Lung Care della SIEMENS (CAD LCS), è risultata evidente la carenza di prestazioni e praticità di utilizzo di tali sistemi commerciali.

Tutte le liste di annotazioni e findings ottenute sui casi di ANODE09 sono state inviate ai curatori del database perchè siano confrontate con il gold standard dello stesso. Al momento siamo in attesa di risposta.

Infine, la possibilità di utilizzo del CAD M5L come supporto al lavoro dei radiologi è stata concretizzata con la sua implementazione all'interno di un servizio web di CAD on-line. Tale servizio consente di far analizzare al CAD esami CT polmonari, grazie ad un approccio di *Cloud Computing*, utilizzando semplicemente un browser ed una connessione ad Internet.

Questo approccio rappresenta un punto di originalità rispetto all'utilizzo di CAD installati su consolle fisse nei reparti di radiologia, che devono essere condivise tra tutti gli utilizzatori e per tutte le applicazioni possibili della macchina CT a cui sono collegate. L'obiettivo è fornire un sistema con elevate prestazioni e che consenta di migliorare semplicità e flessibilità di utilizzo.

Attualmente il servizio è in fase di test presso diversi centri ospedalieri.

Sono già disponibili i risultati ottenuti su 14 esami, refertati in precedenza, provenienti dall'ospedale di Cuneo. Su tali esami la sensibilità del

CAD risulta in linea con quanto ottenuto sul database LIDC-IDRI e migliore rispetto a quella stimabile per i radiologi.

Complessivamente i risultati ottenuti fino ad ora dal CAD M5L possono essere considerati indicativi della stabilità di funzionamento e dell'efficacia di utilizzo come supporto al lavoro di refertazione.

Dalle esperienze dirette effettuate con il radiologo di Cuneo è risultato evidente l'apprezzamento del medico per le indicazioni fornite dal CAD, che in diverse occasioni hanno permesso l'individuazione di noduli sfuggiti durante la refertazione.

In base a tali risultati e all'interesse mostrato nei confronti del web service, è prevedibile che tale servizio possa diventare a tutti gli effetti un efficace strumento al servizio del lavoro dei radiologi.

Ringraziamenti

Vi sono veramente molte persone che devo ringraziare per la collaborazione ed il supporto durante la realizzazione di questa tesi, mi scuso quindi per eventuali dimenticanze.

Innanzitutto vorrei esprimere la mia gratitudine alla Dott.ssa Maria Evelina Fantacci, per avermi seguito e consigliato durante tutto il lavoro e la stesura di questo testo.

Un grande ringraziamento al Dott. Piergiorgio Cerello (INFN di Torino) e al Dott. Ernesto Lopez Torres (CAEDEN, Cuba), per la collaborazione e l'aiuto su tutto quanto concerne il CAD CAM, la realizzazione del web service ed il suo utilizzo. Ugualmente voglio ringraziare anche il Dott. Niccolò Camarlinghi, per il supporto sull'utilizzo del CAD VBNA e la programmazione.

Vorrei quindi ringraziare i radiologi che hanno effettuato le refertazioni sui 50 casi del database ANODE09: Dott. G. Gallarato (Ospedale Santa Croce di Cuneo), Dott.ssa I. Zecca (Ospedale Fazzi di Lecce) e Dott.ssa D. Cataldo (Ospedale di Sanremo).

Devo un ulteriore ringraziamento al Dott. Gallarato per le preziose spiegazioni sulle modalità di refertazione e per la collaborazione durante i test di utilizzo del web-service. Per lo stesso motivo voglio ringraziare il Dott. S. Chauvie e il Dott. E. Roberto (Fisica Sanitaria dell'Ospedale Santa Croce di Cuneo).

Un ringraziamento inoltre a tutte le persone ed ai radiologi che hanno collaborato per le varie fasi di test: il Dott. F. Falaschi (Ospedale di Cisanello a Pisa), il Dott. M. Torsello (Ospedale Fazzi di Lecce), la Dott.ssa R. Cataldo (Università del Salento e INFN di Lecce) e il Dott. M. Mascalchi ed il Dott. L. Sali (Ospedale di Careggi a Firenze).

Grazie anche alla Dott.ssa C. Carpentieri (ISS, Roma) per l'ospitalità e la disponibilità durante la mia breve visita.

Infine vorrei ringraziare tutte le persone che mi sono state accanto in questo percorso, specialmente Elisa e la mia famiglia.

Appendice

A Utilizzo plugin Osirix

A.1 Introduzione

Nei paragrafi che seguono sono descritte le istruzioni di utilizzo del plugin M5L per il visualizzatore DICOM Osirix, appositamente realizzato, nell'ambito della collaborazione M5L dell'INFN, per la visualizzazione dei findings del CAD M5L. Questo plugin permette, oltre alla visualizzazione dei findings del sistema, di annotare noduli ed eseguire delle refertazioni direttamente sul visualizzatore Osirix.

Le istruzioni sono state scritte durante il test del CAD M5L come secondo lettore per la refertazione dei 50 casi del database ANODE09, come descritto nel *Paragrafo* 6.2 per quanto riguarda il terzo radiologo. Nello specifico quindi forniscono una guida per la realizzazione della refertazione iniziale e della successiva revisione tramite i findings del CAD, ma sono comunque valide per tutti i possibili ambiti di utilizzo del plugin.

A.2 Operazioni Iniziali

1. Aprire tramite Osirix l'esame che si vuole refertare trascinando la cartella che lo contiene nella finestra del programma
2. Scegliere di copiare solamente il link e non le immagini
3. Visualizzare le immagini facendo doppio click sulla linea corrispondente all'esame
4. Nella schermata di visualizzazione immagini il plugin M5L si apre seguendo il percorso **Plugins**→**ROI Tools**→**CAD M5L**, partendo dal menù a tendina in alto a destra come mostrato nella *Figura* 48
5. Verrà aperta la finestra **CAD M5L Plugin** con cui è possibili refertare e visualizzare i findings del cad

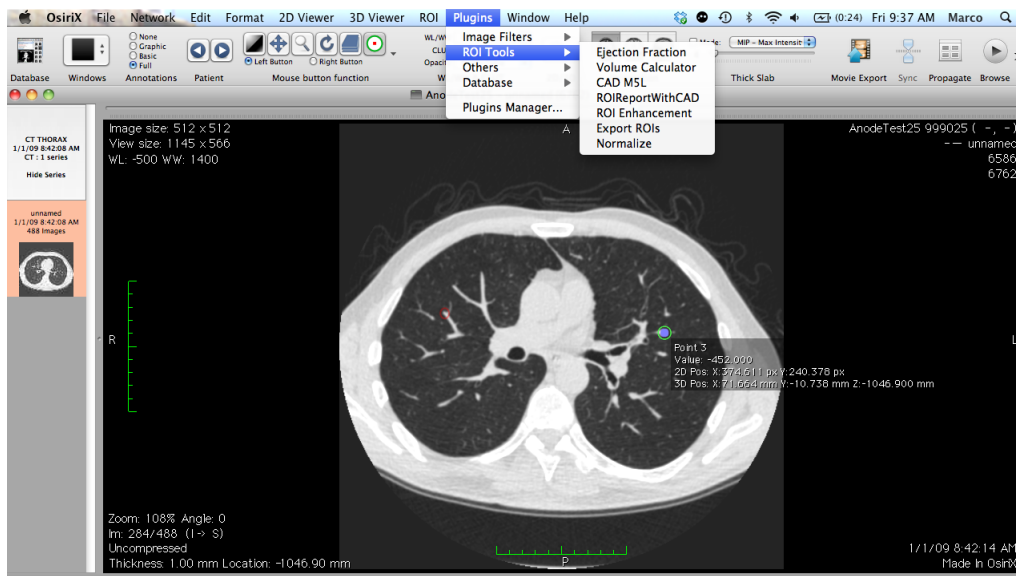


Figura 48: Menù per la selezione dei Plugins

A.3 Refertazione senza CAD

1. Scorrere le immagini per cercare eventuali noduli
2. Individuato un candidato nodulo selezionare tramite il pulsante **ROI Tolls**, posizionato sulla barra grigia grande sopra l'immagine, l'ellisse o il punto e disegnarlo intorno alla lesione individuata (*Figura 49*)
3. Premere sul pulsante **Add Current ROI to RAD Findings** presente nella finestra del Plugging: nella casella **RAD Findings** in basso a sinistra apparirà un numero relativo al findings selezionato (*Figura 50*)
4. Cliccare sulla casella bianca accanto al numero per inserire il numero 1 o 2 a seconda che si ritenga il nodulo rilevante o meno, secondo la definizione di ANODE09 (vedi NOTA) (*Figura 50*)
5. Premere sul tasto **Save RAD** per salvare il findings. La lista di findings creata in questo modo sarà salvata nel file RAD.xml dentro alla cartella ROIs dell'esame refertato (*Figura 50*)

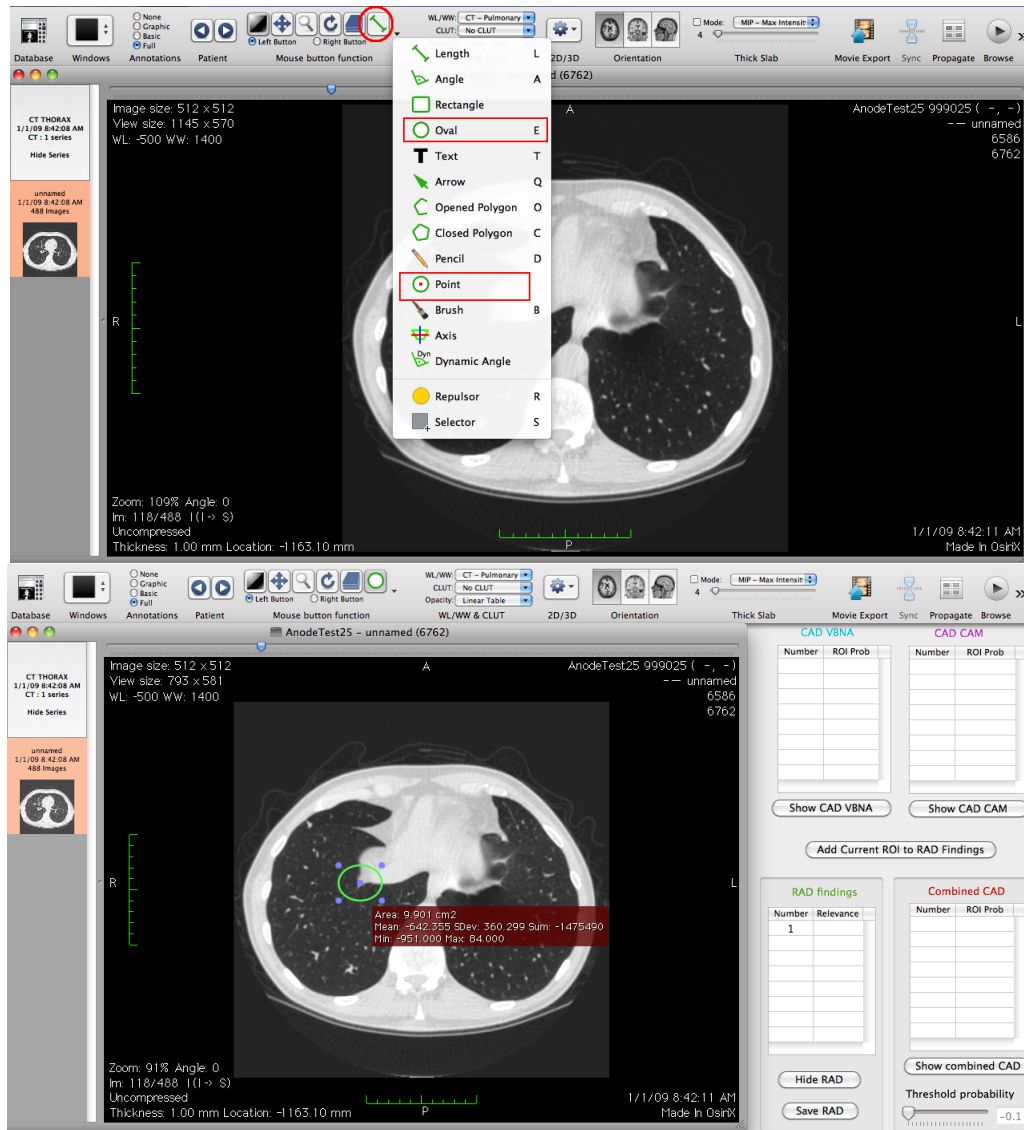


Figura 49: Selezione Tools e disegno ROI sull'immagine

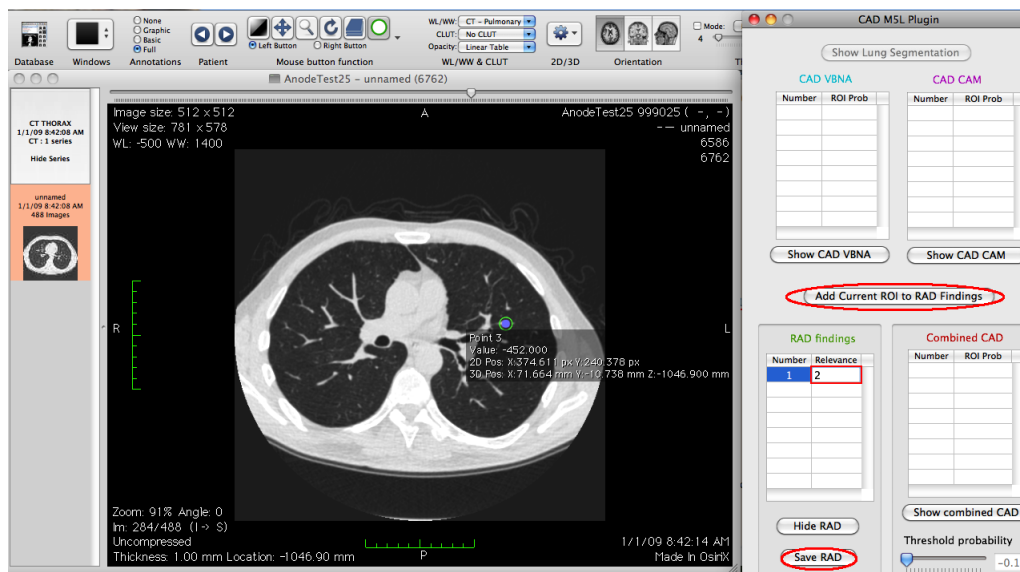


Figura 50: Salvataggio della ROI indicante il finding trovato

NOTA: secondo il sistema adottato per il database ANODE09 i findings sono contrassegnati nel seguente modo:

1. se il candidato viene considerato essere un nodulo effettivamente rilevante
2. se il candidato è posto su una struttura che potrebbe essere scambiata per un nodulo, ma il radiologo ritiene essere una qualche altra struttura anatomica. Sono inseriti in questa categoria anche i noduli troppo piccoli per essere considerati rilevanti e noduli con caratteristiche benigne.

Questa secondo categoria di findings del CAD è irrilevante a livello terapeutico ma va comunque annotata in quanto, per la difficoltà di distinzione rispetto ad un nodulo rilevante, non devono essere considerati come falsi positivi nell'elaborazione delle prestazioni del sistema.

A.4 Refertazione con CAD

IMPORTANTE: prima di procedere con la refertazione con l'ausilio del CAD bisogna inviare indietro i file RAD.xml creati come descritto sopra. Nel caso sia possibile utilizzarlo verrà fornito uno script (backup_RAD.xml.py) per generare automaticamente tutti i file da inviare, rinominati in modo che sia possibile distinguere l'esame.

In caso contrario è necessario compiere l'operazione manualmente, rinominando i file per esempio come RAD_1.xml, RAD_2.xml ecc ecc.

1. Premendo sul pulsante Show Combined Findings del Plugin nella casella Combined CAD compare la lista di findings del CAD (*Figura 51*)
2. Cliccando su una voce dell'elenco sarà mostrata la posizione del finding sulla fetta relativa (*Figura 51*)
3. Se si ritiene che il finding sia effettivamente un candidato nodulo si procede come descritto nella sezione precedente premendo **Add Current ROI to RAD Findings** e proseguendo nelle operazioni
4. Facendo lo stesso per tutti i findings si ottiene la lista dei candidati annotati in partenza con l'aggiunta di quelli del CAD ritenuti corretti.

NOTA: poichè la localizzazione del candidato nodulo a livello di fetta può essere variabile è opportuno scorrere alcune fette prima o dopo quella su cui viene visualizzato il findings del CAD per controllare la coincidenza con findings annotato in precedenza.

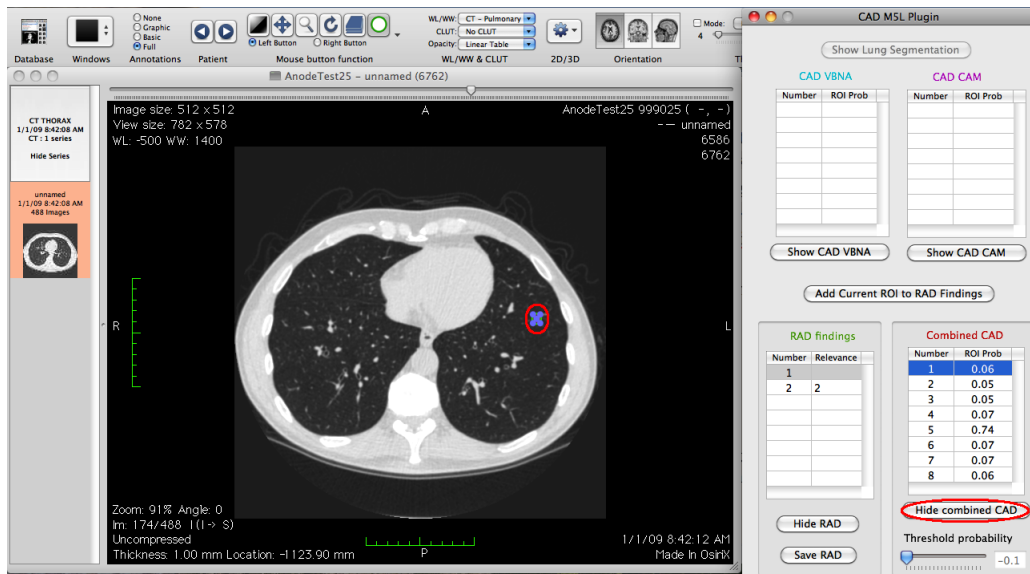


Figura 51: Visualizzazione e valutazione findings CAD.

Riferimenti bibliografici

- [1] J. Beutel. *Handbook of medical imaging: Physics and psychophysics*. Number V. 1. SPIE Press, 2000.
- [2] Brooks RA, Di Chiro G. *Theory of image reconstruction in Computed Tomography*. Radiology 117:516-572, 1975.
- [3] Rossmann K. *Point Spread Function, Line Spread Function, and Modulation Transfer Function: tools for the study of imaging system*. Radiology 93:257-272, 1969.
- [4] Gordon R, Herman GT, Johnson Sa. *Image reconstruction from projections*. Scientific American, October:56-58, 1975.
- [5] DICOM. Informazioni e materiale visibile su <http://medical.nema.org/>
- [6] Linda L. Humphrey, Mark Deffebach, Miranda Pappas, et al. *Screening for lung cancer with low dose computer tomography: a systematic review to update the U.S. Preventive Services Task Force Recommendation*. Annals of Internal Medicine (www.annals.org), 30 July 2013.
- [7] Aberle DR., Adams AM., Berg CD., et al; National Lung Screening Trial Research Team. *Reduced lung-cancer mortality with low-dose computed tomographic screening*. New England Journal of Medicine, 365:395-409, 2011. [PMID : 21712237]
- [8] U. Pastorino, M. Rossi, V. Rosato, A. Marchianò, N. Sverzellati, C. Morosi, et al. *Annual or biennial CT screening versus observation in heavy smokers: 5 years result of the MILD trial*. Eur J Cancer Prev, 21:308-15, 2012. [PMID : 22465911]
- [9] Z. Sagir, A. Dirksen, H. Asharf, KS. Bach, J. Brodersen, PF. Clemenssen, et al. *CT screening for lung cancer brings forward early disease. The randomised Danish Lung Cancer Screening Trial: status after five annual screening rounds with low dose CT*. Thorax, 67:296-301, 2012. [PMID : 22286927]
- [10] M. Infante, S. Cavuto, FR. Lutman, G. Brambilla, G. Chiesa, G. Ceresoli, et al. DANTE Study Group. *A randomized study of lung*

- cancer screening with spiral computed tomography : three-year results from DANTE trial.* Am. J. Respir. Crit. Med., 180:445-53, 2009. [PMID : 19520905]
- [11] AL. Pegna, G. Picozzi, F. Falaschi, L. Carrozzi, M. Falchini, FM. Carozzi, et al. ITALUNG Study Research Group. *Four-Year Results of Low-Dose CT Screening and Nodule Management in the ITALUNG Trial.* J Thorac Oncol, 8(7):866-875, 2013 Jul.
- [12] DM. Burns. *Primary prevention, smoking and smoking cessation: implications for future trends in lung cancer prevention.* Cancer, 89:2506-9, 2000. [PMID : 11147637]
- [13] MT. Halpern, BW. Gillespie, KE. Warner. *Patterns of absolute risk for lung cancer mortality in former smokers.* J. Natl. Cancer Inst. 85:457-64, 1993. [PMID : 8445673]
- [14] L. Tong, MR. Spitz, JJ. Fueger, CA Amos. *Lung carcinoma in former smokers.* Cancer, 78:1004-10, 1996. [PMID : 8780538]
- [15] E. Barnes. AuntMinnie.com staff. *USPSTF recommends CT lung cancer screening.* AuntMinnie.com, 29 luglio 2013
- [16] LL. Humphrey, M. Halfand, BK. Chan, SH. Woolf. *Brest Cancer screening: a summary of the evidence for the U. S. Preventive Service Task Force.* Ann. Int. Med., 137:347-60, 2002. [PMID : 12204020]
- [17] H. Nelson, K. Tyne, A. Naik, C. Bougatos, B. Chan, P. Nygren, et al. *Screening for breast cancer: systematic evidence review update for the U. S. Preventive Service Task Force.* Rockville, MD: Agency for Healthcare Research and Quality; Report 10-05142-EF-1, 2009.
- [18] RE. Schoen, PF. Pinsky, JL. Weissfeld, LA. Yokochi, T. Church, AO. Laiyemo, et al. PLCO project team. *Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy.* N. Engl. J. Med., 366:2345-57, 2012. [PMID : 22612596]
- [19] Informazioni visibili sul sito dell' Associazione Italiana Ricerca Cancro <http://www.airc.it>

- [20] Cancer Facts and Figures 2013 informazioni e materiale disponibile sul sito <http://www.cancer.org/>
- [21] R. Siegel, D. Naishadham, A. Jemal: American Cancer Society. *Cancer Statistics*. CA Cancer. J. Clin., 63:11-30, 2013.
- [22] C. F. Mountain. *Staging classification of lung cancer. A critical evaluation*. Clinics in Chest Medicine, 23(1):103-121, 2002.
- [23] C. I. Henschke, D. F. Yankelevitz. *CT screening for lung cancer: update 2007*. The Oncologist, 13:65-78, 2008.
- [24] D. P. Chakraborty, L. H. Winter. *Free-response methodology: alternate analysis and a new observer-performance experiment*. Radiology, 174(3 Pt 1):873-881, March 1990. [PMID : 2305073]
- [25] Per informazioni contattare stefano.bagnasco@to.infn.it
- [26] A. L. Pegan, G. Picozzi, M. Mascalchi, F. Mira, F. M. Carozzi, L. Carozzi, et al. *Design, recruitment and baseline results of the ITALUNG trial for lung cancer screening with low dose CT*. Lung Cancer (Amsterdam, Netherlands), 64(1):34-40, April 2009. [PMID : 18723240]
- [27] R. Bellotti, P. Cerello, S. Tangaro, V. Bevilacqua, M. Castellano, G. Mastronardi, et al. *Distributed medical images analysis on a grid infrastructure*. Future Gener. Comput. Syst., 23(3):475-484, 2007.
- [28] Informazioni disponibili su <http://anode09.isi.uu.nl/>.
- [29] B. van Ginneken, S. G. Armato III, B. de Hoop, S. van Amelsvoort-van de Vorst, T. Diundam, et al. *Comparing and combining algorithms for computer-aided detections of pulmonary nodules in computed tomography scans: The ANODE09 study*. Medical Image Analysis 14, 707-722, 2010.
- [30] R. M. Nishikawa, L. M. Yarusso. *Variations in measured performance of CAD schemes due to database composition and scoring protocol*. Proc. SPIE 3338, 840-844, 1998.
- [31] R. M. Nishikawa et al. *Effect of case selection on the performance of computer-aided detection schemes*. Med. Phys. 21, 265-269 (1999).

- [32] S. G. Armato III, et al. *The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDR): a completed reference database of lung nodules on CT scans*. Med. Phys., 38 (2), February 2011.
- [33] Informazioni disponibili su <https://wiki.nci.nih.gov/display/CIP/LIDC>
- [34] E. R. Weibel, D. M. Gomez. *Architecture of the human lung. Use of quantitative methods establishes fundamental relations between size and number of lung structures*. Science (New York, N.Y.), 137:577–585, August 1962. [PMID : 14005590]
- [35] E. Breatnach, G. C. Abbott, R. G. Fraser. *Dimensions of the normal human trachea*. AJR. American Journal of Roentgenology, 142(5):903–906, May 1984. [PMID : 6609569]
- [36] C. Vinhais, A. Campilho. *Lung parenchyma segmentation from ct images based on material decomposition*. In Aurelio Campilho and Mohamed Kamel, editors, Image Analysis and Recognition, volume 4142 of Lecture Notes in Computer Science, pages 624–635. Springer Berlin Heidelberg, 2006. 10.1007/11867661_56.
- [37] R. C. Gonzalez, R. E. Woods. *Digital image processing*. Prentice Hall, 2008.
- [39] Informazioni disponibili sulla *guida di itk* presente sul sito "www.itk.org/ItkSoftwareGuide.pdf" alle pagine 257–252.
- [38] Informazioni disponibili sul sito "<http://www.itk.org/>".
- [40] Q. Li , F. Li, K. Doi. *Computerized Detection of Lung Nodules in Thin-Section CT Images by use of Selective Enhancement Filters and an Automated Rule-Based Classifier*. Acad.Radiol.,15 (2008) 165. [PMID : 18206615, PMCID : 2266079]
- [41] T. Lindeberg. *On scale selection for different operators*. Proceedings of the Eighth Scandinavian Conference on Image Analysis, Tromso, Norway, pp. 857–866, 1993 .

- [42] Y. Sato, S. Nakajima, N. Shiraga, H. Atsumi, S. Yoshida, T. Koller, G. Gerig, R. Kikinis. *Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images*. Med. Image Anal., 2 (1998) 143.
- [43] W. Lorensen, H. Cline. *Marching cubes: A high resolution 3D surface construction algorithm*. SIGGRAPH '87: Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, Vol. 21, pp. 169, 163. ACM Press, 1987.
- [44] Informazioni disponibili su http://it.wikipedia.org/wiki/Gouraud_shading.
- [45] D. S. Paik, C. F. Beaulieu, G. D. Rubin, B. Acar, R. B. Jeffrey, J. Yee, J. Dey, S. Napel. *Surface normal overlap: a computer-aided detection algorithm with application to colonic polyps and lung nodules in helical CT*. IEEE Trans. Med. Imaging, 23 (2004) 661. [PMID : 15191141]
- [46] Informazioni disponibili su <http://en.wikipedia.org/wiki/Tomosynthesis>
- [47] N. Camarlinghi, I. Gori, A. Retico, R. Bellotti, P. Bosco, P. Cerello, G. Gargano, E. L. Torres, R. Megna, M. Peccarisi, M. E. Fantacci. *Combination of computer-aided detection algorithms for auto-matic lung nodule identification*. International Journal of Computer Assisted Radiology and Surgery, July 2011. [PMID : 21739112]
- [48] D. Wormanns, K. Ludwig, F. Beyer, et al. *Detection of pulmonary nodules at multirow-detector CT: effectiveness of double reading to improve sensitivity at standard-dose and low-dose chest CT*. Eur Radiol, 15:14–22, 2005
- [49] N. Camarlinghi, M. E. Fantacci, I. Gori, A. Retico. *A computer-aided detection system for lung nodules in CT images*. Il Nuovo Cimento, 01/03/2011.
- [50] A. Savitzky, M. J. E. Golay. *Smoothing and differentiation of data by simplified least squares procedures*. Analytical Chemistry 36, 1627–1639, 1964.
- [51] S. Rajagopalan, R. A. Robb. *Image smoothing with Savitzky-Golay filters*. Proc. SPIE 5029, Medical Imaging 2003: Visualization, image-guided procedures , 773–781.

- [52] P. Cerello, S. C. Cheran, S. Bagnasco, R. Bellotti, et al. *3-D object segmentation using ant colonies*. Pattern Recognition 43, 1476–1490, 2010.
- [53] Informazioni disponibili su http://it.wikipedia.org/wiki/Region_growing.
- [54] Informazioni disponibili su http://en.wikipedia.org/wiki/Feedforward_neural_network
- [55] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 2 edition, July 1998.
- [56] V. Kecman. *Learning and Soft Computing, Support Vector Machines, Neural Networks and Fuzzy Logic Models*. MIT Press, 2001.
- [57] M. Niemeijer, M. Loog, M.D. Abramoff, M.A. Viergever, M. Prokop, B. van Ginneken. *On combining computer-aided detection systems*. Medical Imaging, IEEE Transactions on, 30(2):215 –223, feb. 2011
- [58] *Pubblicazione ICRP 103*. Testo disponibile su http://www.icrp.org/docs/P103_Italian.pdf
- [59] Informazioni disponibili su <http://www.osirix-viewer.com/>
- [60] Informazioni disponibili su <http://www.mathworks.it/products/matlab/>
- [61] Informazioni disponibili su http://en.wikipedia.org/wiki/Marching_cubes
- [62] Informazioni disponibili su http://it.wikipedia.org/wiki/Cloud_computing
- [63] Informazioni disponibili su http://en.wikipedia.org/wiki/Cloud_computing
- [64] M. Armbrust, A. Fox, R. Griffith, et al. *Above the Clouds: A Berkeley View of Cloud Computing*. Technical Report No. UCB/EECS-2009-28, February 10, 2009.
- [65] Informazioni disponibili su <http://en.wikipedia.org/wiki/Virtualization>