

POLITECNICO DI MILANO - Polo di Como  
Scuola di Ingegneria Industriale e dell'Informazione  
Corso di Laurea Magistrale in Ingegneria Informatica  
Dipartimento di Elettronica, Informazione e Bioingegneria



## Riconoscimento automatico del distretto anatomico in immagini TAC

Relatore: Prof. Giuseppe Pozzi  
Correlatore: Dr. Angelo Ostinelli  
Dr. Antonio Vassena

Tesi di laurea di:  
Francesco Colombo Matr. 797540

Anno Accademico 2013–2014

*«If you do something and it turns out pretty good, then you should go do something else wonderful, not dwell on it for too long. Just figure out what's next.»*

Steve Jobs

*«Se fai qualcosa e risulta essere molto buona, poi dovresti fare qualcos'altro di meraviglioso, non soffermarti per molto tempo su ciò che hai fatto. Bisogna solo capire cosa ci sarà dopo.»*

Steve Jobs

# Ringraziamenti

## GRAZIE

Al Professor Giuseppe Pozzi, per i suoi utili consigli durante lo sviluppo dell'intero progetto.

Al Dr. Angelo Ostinelli, per essersi sempre dimostrato disponibile e benevolo ed avermi accolto a braccia aperte durante le mie ripetute visite all'Ospedale S. Anna di Como.

Al Dr. Antonio Vassena, per avermi assistito più volte, nei molti problemi che si sono presentati, durante l'intero periodo del lavoro di tesi.

Ai miei genitori, per avermi dato la possibilità di proseguire negli studi, per essermi stati vicini nei momenti più difficili, per aver sempre creduto in me e per avermi sopportato: non sarei dove sono oggi senza di loro.

A tutti i miei amici ed ai compagni di studio conosciuti in questi anni di Politecnico, per avermi aiutato durante questi lunghi anni attraverso il corso di studi e per aver portato nella mia vita tante risate e distrazioni.

# Sommario

All'Azienda Ospedaliera S.Anna di Como, ogni anno vengono effettuati circa 26500 esami TAC. Un esame TAC è costituito da un numero variabile di immagini digitali in un formato specifico chiamato DICOM che contiene un'intestazione con informazioni su paziente ed esame svolto per l'ottenimento dell'immagine. Nel caso della pianificazione di una terapia con radiazioni ionizzanti oppure per lo svolgimento di ulteriori esami, lo studio della distribuzione della dose di radiazioni nel paziente è molto importante, perché fornisce una stima dell'irradiazione dei tessuti tumorali e sani, permettendo una maggiore capacità di previsione sulla risposta e sulla tossicità della dose di radiazione subita dal corpo.

Per il calcolo della dose è in uso all'Ospedale un Software chiamato ARGO, già frutto della collaborazione con il Politecnico di Milano, che permette l'ottenimento di un fattore piuttosto preciso chiamato SSDE ( Size-Specific Dose Estimates) che tiene conto di età, dimensioni e distretto anatomico analizzato. Tale programma non è però in grado di rilevare in automatico il distretto anatomico preso in esame: l'etichettatura dei file DICOM, svolta dai tecnici, è spesso imprecisa o sbagliata; nasce così l'idea del seguente lavoro di tesi.

L'obbiettivo è trovare un metodo affidabile per il riconoscimento automatico dei distretti anatomici principali (encefalo, torace ed addome) analizzando direttamente il contenuto visivo delle immagini radiologiche TAC. Il problema, che si è dimostrato essere non banale, è stato risolto modificando l'uso di un metodo noto di confronto tra immagini attraverso l'estrazione di features: Histograms of Oriented Gradients (HOG). Viene creato un dataset calcolando HOG per porzioni significative da alcune immagini e l'analisi è poi svolta confrontando l'HOG di porzioni delle immagini da etichettare con il dataset mediante il calcolo della loro simmetria, espressa dal coefficiente di una retta nel piano ottenuta calcolando un istogramma in funzione dell'altro. La posizione all'interno dell'immagine in cui il coefficiente della retta è più vicina al valore 1 indica la zona dove si è trovato il valore più simile ad uno degli elementi del dataset. Tale risultato permette non solo di capire la tipologia di distretto anatomico presente nell'immagine, ma anche la sua posizione permettendo calcoli di dose ancora più precisi.

Questo metodo è stato sviluppato mediante una lunga fase di testing ed è basato

su un approccio completamente empirico al problema successivo al fallimento dei metodi tradizionali. L'accuratezza del sistema si attesta al 91% circa mentre il tempo di elaborazione medio, utilizzando i PC in dotazione all'Ospedale, per una singola immagine è pari ad 1 minuto e 40 secondi. Tali risultati sono stati ritenuti accettabili dall'Azienda Ospedaliera comasca.

# Abstract

Every year St. Anna Hospital in Como performs some 26500 CT (Computerized Tomography) exams. A CT exam entry is composed by a variable number of DICOM digital images. DICOM is a file format for medical use; it containing the visual images of the body region analyzed and additional information about the patient and the exam. The estimation of the toxic level of a radiation dose administered to a patient is a mandatory parameter to consider when a patient undergoes anti-tumor radiation therapy, besides being a precautionary measure for the patient's safety. The dose can be estimated by using the ARGO system, a special software developed by Politecnico di Milano in cooperation with the St. Anna Hospital. ARGO provides an accurate toxicity prediction value called SSDE (Size Specific Dose Estimate). SSDE is calculated based on the additional information provided by DICOM headers such as age, body region size and type. These DICOM header values are manually inserted by technicians during the exam; these values, in certain cases may be erroneous due to typing errors or may lack the required precision. The aim of this thesis is to correct technicians' errors by providing a method for automatic body region recognition (head, thorax and abdomen) by analyzing CT image contents. This problem is not as trivial as one may think, and the results are obtained by using a modified version of Histogram of Oriented Gradients (HOG) technique. A pre-labeled dataset is created by obtaining HOG features vectors from some relevant image portions and the labeling procedure is done by comparing the dataset with HOG feature vectors taken from different regions inside the not labeled sample images. The level of similarity between HOG feature vectors is computed using the symmetry level expressed by the coefficient of a straight line (generated by plotting a features vector in terms of another one) in a 2D space: the close to 1 this coefficient is, the more similar the vectors are. The position inside the image which provides the best "near to 1" coefficient is the result of the computation. This result not only provides the right body region type of an image, but also the exact position inside a larger portion of the body. This extra feature can also improve the precision of dose estimates calculation. This method was conceived after a long testing phase and it's completely empirical. System accuracy is about 91% and the execution average time for each image on Hospital PCs is about 1m40s. These results

have been considered acceptable by St. Anna Hospital Staff.

# Indice

<b>1</b>	<b>Introduzione</b>	<b>1</b>
1.1	Importanza della diagnostica per immagini . . . . .	1
1.2	Breve storia della diagnostica per immagini . . . . .	2
1.3	Tomografia computerizzata (CT) . . . . .	2
1.4	Ecografia (Eco) . . . . .	4
1.5	Risonanza magnetica (RM) . . . . .	6
1.6	Pericolo per la salute delle immagini radiologiche . . . . .	7
1.7	Punto di partenza dello studio . . . . .	8
<b>2</b>	<b>Cenni di radiologia</b>	<b>10</b>
2.1	Fondamenti di fisica nucleare . . . . .	10
2.2	Le unità di misura fondamentali della radiologia: Gray e Sievert a confronto . . . . .	11
2.3	Dose Assorbita . . . . .	12
2.4	Dose Equivalente . . . . .	12
2.5	Dose Efficace . . . . .	13
2.6	Indici di Tomografia computerizzata . . . . .	14
2.6.1	Region of Interest (ROI) . . . . .	16
2.6.2	CTDI . . . . .	17
2.6.3	CTDI <sub>100</sub> . . . . .	17
2.6.4	CTDI <sub>W</sub> . . . . .	17
2.6.5	CTDI <sub>Vol</sub> . . . . .	18
2.6.6	Dose-Length Product (DLP) . . . . .	18
2.6.7	Dose Efficace in CT . . . . .	19
2.6.8	Size Specific Dose Estimates . . . . .	19
<b>3</b>	<b>Stato dell'arte</b>	<b>21</b>
3.1	Descrittori alfanumerici nelle immagini DICOM . . . . .	22
3.2	Analisi del contenuto delle immagini radiologiche . . . . .	23
3.3	L'utilizzo delle immagini Scout: un approccio ibrido . . . . .	24
3.4	Riconoscimento di immagini . . . . .	24



---

3.4.1	Indici di somiglianza . . . . .	25
3.4.1.1	Distanza euclidea ed IMED . . . . .	25
3.4.1.2	Color Histogram . . . . .	26
3.4.1.3	Structural Similarity (SSIM) . . . . .	27
3.4.2	Features detection . . . . .	28
3.4.2.1	Speeded Up Robust Features (SURF) . . . . .	28
3.4.2.2	Histogram Of Oriented Gradients (HOG) . . . . .	29
3.4.3	Reti Neurali . . . . .	32
3.4.3.1	Self-Organizing Map (SOM) . . . . .	33
3.4.3.2	Rete di Hopfield . . . . .	34
3.5	Valutazione delle prestazioni di un sistema di classificazione: la matrice di confusione . . . . .	36
<b>4</b>	<b>Obiettivo della tesi</b> . . . . .	<b>40</b>
4.1	Requisiti . . . . .	40
4.1.1	PACS . . . . .	41
4.2	Vincoli di progetto . . . . .	42
4.3	Strumenti . . . . .	43
4.3.1	Linguaggi di programmazione . . . . .	43
4.3.1.1	Java ed ImageJ . . . . .	43
4.3.1.2	MATLAB . . . . .	44
4.3.2	Ambienti di sviluppo . . . . .	45
4.3.2.1	Ambiente Java . . . . .	45
4.3.2.2	Ambiente MATLAB . . . . .	45
4.3.3	Archiviazione e condivisione del codice sorgente . . . . .	45
<b>5</b>	<b>Approccio utilizzato</b> . . . . .	<b>47</b>
5.1	Insieme di immagini test . . . . .	48
5.2	Valutazione preliminare delle prestazioni . . . . .	48
5.3	Metodi della letteratura . . . . .	49
5.3.1	Risultati del metodo “Distanza Euclidea” . . . . .	49
5.3.2	Risultati del metodo “IMED” . . . . .	50
5.3.3	Risultati del metodo “Color Histogram” . . . . .	50
5.3.4	Risultati del metodo “SSIM” . . . . .	51
5.3.5	Risultati del metodo “SURF” . . . . .	51
5.3.6	Risultati del metodo “HOG” . . . . .	52
5.3.7	Risultati del metodo “SOM” . . . . .	52
5.3.8	Risultati del metodo “rete di Hopfield” . . . . .	53
5.4	Valutazione complessiva dei metodi della letteratura . . . . .	53
5.5	Definizione di nuovi metodi . . . . .	54
5.5.1	Dimensioni medie del corpo umano . . . . .	54

5.5.1.1	Risultati dimensioni medie del corpo umano . . . . .	55
5.5.1.2	Valutazione dimensioni medie del corpo umano . . . . .	55
5.5.2	Distribuzione spaziale dei livelli di grigio nell'immagine . . . . .	55
5.5.2.1	Risultati distribuzione spaziale dei livelli di grigio nell'immagine . . . . .	56
5.5.2.2	Valutazione distribuzione spaziale dei grigi nell'immagine . . . . .	56
5.5.3	Valore della simmetria in regioni d'interesse con HOG . . . . .	57
5.5.3.1	Creazione del Dataset . . . . .	57
5.5.3.2	Processo di etichettatura . . . . .	58
5.5.3.3	Risultati e valutazione . . . . .	60
5.5.3.4	Struttura ed utilizzo del programma MATLAB . . . . .	60
<b>6</b>	<b>Risultati</b>	<b>64</b>
6.1	Matrice di confusione per il metodo «Valore della Simmetria in regioni d'interessi con HOG» . . . . .	64
6.2	Studio degli errori più frequenti . . . . .	68
6.2.1	Errori encefalo . . . . .	68
6.2.2	Errori toraci antero-posteriori . . . . .	68
6.2.3	Errori addomi antero-posteriori . . . . .	69
6.3	Tempi di elaborazione . . . . .	69
<b>7</b>	<b>Conclusioni e sviluppi futuri</b>	<b>72</b>
7.1	Conclusioni . . . . .	72
7.2	Sviluppi futuri . . . . .	73
7.2.1	Sistema aperto vs. sistema chiuso . . . . .	73
7.2.2	Pazienti pediatrici . . . . .	74
7.2.3	Estrazione ed analisi automatica delle immagini direttamente sul PACS . . . . .	74
7.2.4	Supporto su immagini di più produttori . . . . .	75
7.2.5	Aumentare le prestazioni temporali . . . . .	75
	<b>Bibliografia</b>	<b>78</b>
	<b>A Glossario</b>	<b>79</b>

# Elenco delle figure

1.1	Apparecchiatura per TAC . . . . .	3
1.2	Apparecchiatura per Ecografia . . . . .	4
1.3	Apparecchiatura per RMI . . . . .	6
2.1	Fondamenti di radiologia . . . . .	11
2.2	Fantoccio acrilico per CT . . . . .	15
2.3	Esempio ROI in ARGO . . . . .	16
3.1	Header DICOM . . . . .	22
3.2	Esemplificazione visiva di SURF . . . . .	29
3.3	Esemplificazione visiva di HOG . . . . .	30
3.4	Esempio visivo di SVM . . . . .	31
3.5	Reti neurali naturali ed artificiali . . . . .	32
3.6	Struttura SOM . . . . .	33
3.7	Struttura Rete di Hopfield . . . . .	35
4.1	PACS . . . . .	41
5.1	Tabella dimensioni corpo . . . . .	55
5.2	Funzioni distribuzione spaziale dei livelli di grigio nell'immagine . . . . .	56
5.3	Tipiche ROI encefalo . . . . .	58
5.4	Tipiche ROI torace . . . . .	59
5.5	Tipiche ROI addome . . . . .	60
5.6	Interfaccia MATLAB . . . . .	61
5.7	Cartella del Dataset . . . . .	62
5.8	Grafici risultati elaborazione . . . . .	63
5.9	File Results.txt . . . . .	63
6.1	Risultato arti inferiori . . . . .	65
6.2	Errore etichettatura encefalo . . . . .	68
6.3	Errori etichettatura toraci antero-posteriori . . . . .	69
6.4	Errore etichettatura addomi antero-posteriori . . . . .	70

7.1 Encefalo e torace pediatrico . . . . . 74

# Elenco delle tabelle

2.1	Fattori sensibilità dei diversi tessuti corporei ICRP103 2007 . . . . .	14
2.2	Fattori di conversione per il calcolo della Dose Efficace a partire da DLP	19
3.1	Prototipo matrice di confusione . . . . .	37
3.2	Prototipo tabella di confusione . . . . .	38
5.1	Tabella riassuntiva metodi letteratura . . . . .	53
6.1	Matrice di confusione . . . . .	65
6.2	Tabella di confusione . . . . .	66
6.3	Analisi del tempo di esecuzione . . . . .	71
6.4	Tempi analisi annui . . . . .	71

# Capitolo 1

## Introduzione

Il lavoro oggetto della presente tesi è stato svolto in collaborazione con l'Azienda Ospedaliera Sant'Anna di Como. L'obiettivo della collaborazione tra Politecnico di Milano e l'Ospedale, è di rispondere a specifiche esigenze cliniche che mirano a individuare e riconoscere in modo automatico, senza il diretto intervento di un operatore, i distretti anatomici delle immagini radiologiche in possesso della Struttura Medica di Como. L'attività si è incentrata sulla ricerca di un metodo specifico di riconoscimento, ottenuto implementando tecniche di confronto tra immagini così da individuare un metodo affidabile per poter eseguire il riconoscimento automatico.

### **1.1 Importanza della diagnostica per immagini**

La Radiologia è quella branca della Medicina Clinica che fornisce informazioni diagnostiche fondate su immagini e, a questo scopo, impiega radiazioni ionizzanti (RX) per la diagnostica convenzionale e la Tomografia Computerizzata (TC), ma anche onde ultrasoniche (ecografia: Eco) e la Risonanza Magnetica (RM). Dunque, oggi è più corretto parlare di Diagnostica per immagini, piuttosto che semplicemente di Radiologia. La Radiologia è contenuta nella moderna Diagnostica per immagini. Una immagine diagnostica è simile ad una mappa, in quanto custodisce e trasmette informazioni in codice e deve essere decodificata. La diagnostica per immagini, oggi rappresenta la principale possibilità di raccordo culturale, di sintesi diagnostica o di dialogo tra diversi specialisti. La possibilità di ottenere immagini dell'interno del corpo permette anche di raggiungere l'importante obiettivo di creare un database della normale anatomia e fisiologia umana per poter poi facilitare l'individuazione di anomalie all'interno degli organi.

## 1.2 Breve storia della diagnostica per immagini

Le origini della radiologia sono riconducibili al 22 dicembre 1895, quando W.C. Röntgen eseguì la prima radiografia guidato dall'intuizione che, le radiazioni da lui chiamate 'X' perché allora erano di natura ed origine sconosciute, avevano la proprietà di poter attraversare i corpi solidi e quindi permettere una visione dell'interno del corpo umano senza bisogno di sezionarlo impressionando inoltre l'emulsione di una pellicola fotografica. L'evidenza radiologica dei diversi distretti anatomici umani e di eventuali patologie si basa sulla differenza di contrasto tra tessuti diversi, cioè sul fatto che i vari tessuti hanno caratteristiche di trasparenze diverse al passaggio dei raggi X. Inizialmente queste tecniche furono impiegate per lo studio delle ossa, strutture che, data la differenza di radiopacità rispetto ai tessuti molli, risultano contrastate e facilmente evidenziabili in radiografia. Quasi nello stesso periodo è iniziato il lungo cammino della radiologia toracica e polmonare allo scopo d'individuare la tubercolosi: la bassissima opacità radiologica del parenchima polmonare, contenente aria, fa risaltare per contrasto la maggiore opacità del cuore e dei vasi mediastinici e quella ancora maggiore delle costole e delle vertebre, che sono ricche di calcio. La scoperta dei raggi X fu solo la prima delle molte scoperte ed innovazioni tecnologiche applicate alla diagnostica per immagini. Ad essa sono seguite, ricordando solo le più importanti, la tomografia computerizzata (CT), resa possibile dall'applicazione dell'informatica alle tecniche di elaborazione delle immagini, l'ecografia e la risonanza magnetica. Le varie tecniche di imaging hanno particolarità e prestazioni diverse fra loro: ciascuna di esse presenta dei limiti di 'potenzialità informativa', non esiste quindi una metodica che sia in grado di rendere le altre obsolete. Questa situazione ha di fatto determinato il passaggio dall'approccio 'monofasico' della radiologia tradizionale all'attuale approccio 'polifasico', che prevede la combinazione delle varie metodologie ed ha portato alla creazione dei dipartimenti di diagnostica per immagini, dotati di sistemi digitali per la gestione della mole sempre crescente di informazioni disponibili [1].

## 1.3 Tomografia computerizzata (CT)

La tomografia computerizzata (TC o CT, dall'inglese Computed tomography, o TAC, Tomografia assiale computerizzata) è una tecnica radiologica con la quale si ottengono immagini di sezioni del corpo umano. La tecnica si fonda sulla ricostruzione della densità dei diversi tessuti del corpo umano attraversati da un fascio di radiazioni X. Tale fascio, di energia nota, viene attenuato in misura direttamente proporzionale alla densità dei differenti tessuti corporei; questo valore d'attenuazione viene misurato da appositi rilevatori, situati dalla parte opposta della sorgente di raggi X rispetto al paziente. Numerose componenti tra loro collegate costituiscono



Figura 1.1: Un comune macchinario per esami TAC

no l'apparecchiatura CT: è presente un tubo radiogeno che emette raggi X ad alta intensità, una serie di rivelatori di raggi X (detettori), un adeguato calcolatore con un sistema di visualizzazione delle immagini. Il tubo a raggi X si muove lungo una traiettoria circolare, puntato verso il centro di rotazione, con il paziente disposto lungo l'asse di rotazione. I rivelatori di raggi X ruotano sullo stesso piano di rotazione, simultaneamente al tubo radiogeno, ma opposti a esso. Il tubo radiogeno è simile a quelli utilizzati nelle comuni indagini radiografiche, ma è caratterizzato da un fuoco di piccole dimensioni (macchia focale di 0,5-1 mm), con emissione di raggi pulsati, elevata capacità di dissipazione del calore per sopportare alti carichi di lavoro. I detettori oggi impiegati sono costituiti da un cristallo scintillatore allo stato solido (ioduro di sodio), oppure possono basare il loro funzionamento sulla ionizzazione di gas (gas xenon). Gli scintillatori allo stato solido emettono un impulso luminoso nello spettro visibile ogni qualvolta sono colpiti da un fotone X e questo impulso luminoso è convertito da un fotomoltiplicatore in segnale elettrico, che successivamente viene amplificato. Si misura l'energia del fascio di raggi X che fuoriesce dal corpo del paziente e, quindi, viene calcolata l'attenuazione subita dal fascio lungo il suo percorso attraverso i tessuti corporei. La trasformazione della misurazione dell'attenuazione del fascio di raggi X nella rappresentazione visiva del corpo umano in sezione, si attua in varie fasi: si ottiene il valore di attenuazione di ciascun punto all'interno del corpo con misurazioni effettuate da punti di vista diversi; un computer risolve rapidamente il sistema di equazioni risultante (circa 760 misurazioni negli apparecchi più moderni) ed infine assegna, a ciascuno dei blocchi (voxel) in cui è stato diviso il corpo, il suo valore di attenuazione. L'immagine TC è di tipo digitale, ma, dato che la sezione corporea in esame ha un suo spessore, ai pixel corrispondono volumi unitari del corpo, che prendono il nome di voxel. Il corpo è quindi suddiviso dal calcolatore in tante piccole unità tridimensionali, i voxel, le quali successivamente sono rappresentate sull'immagine in tante piccole unità bidimensionali, cioè in pixel. Ad ogni voxel viene assegnato un valore di attenuazione che viene normalizzato in maniera tale da raggruppare l'intera gamma dei valori tra gli



estremi di -1000 (per l'aria) e +1000 (per l'osso): le unità di questa tecnica sono state definite unità Hounsfield (UH). Ad ogni singolo pixel, che corrisponde ad un voxel nell'immagine in 3D, si assegna una tonalità di grigio proporzionale all'attenuazione del voxel corrispondente, ottenendo in tal modo una visualizzazione della sezione del corpo umano su uno schermo televisivo. Tutta la scala di unità Hounsfield potrebbe essere così rappresentata con 2001 diverse tonalità di grigio, ma l'occhio umano è in grado di differenziare solo una ventina di diverse gradazioni, quindi si cerca di sfruttare l'elevata risoluzione di contrasto della CT e di rilevare pertanto anche piccole differenze nella densità dei diversi pixel. Inoltre, possono essere effettuate, in modo automatico, misure quantitative del valore di attenuazione dei diversi organi o tessuti, selezionando una regione di interesse mediante la console di comando. È anche possibile misurare le dimensioni di strutture normali o patologiche visibili sulle immagini. L'immagine, nei moderni apparecchi, viene acquisita in tempi di 2-7 secondi, ma, in maniera preliminare all'esame, viene effettuata una foto del distretto anatomico che verrà analizzata ed utilizzata poi come riferimento: tale immagine è chiamata "Scout" (o Localizer in inglese). Durante l'acquisizione, al paziente viene chiesto di rimanere immobile e, per esami del torace e dell'addome, di non respirare. Questa collaborazione è necessaria, altrimenti le immagini risultano di cattiva qualità: i movimenti del paziente, infatti, determinano sfasamenti nel calcolo dell'attenuazione dei raggi X, che sull'immagine ricostruita appaiono come strisce o zone ombrate. Vengono inoltre iniettate sostanze a base di iodio poiché i diversi organi e tessuti, pur avendo valori densitometrici differenti, presentano anche ampie aree di sovrapposizione [1].

## 1.4 Ecografia (Eco)



Figura 1.2: Un comune macchinario per esami ecografici

L'ecografia è un sistema d'indagine diagnostica che utilizza gli ultrasuoni per evidenziare e visualizzare gli organi interni del corpo umano: si basa sul principio

dell'emissione e della trasmissione delle onde ultrasonore a frequenze tra i 2 e 20MHz. La formazione dell'immagine è infatti basata sulla differente capacità di riflessione dell'onda ultrasonora da parte della superficie di separazione tra le varie strutture (interfaccia). L'innocuità dell'ecografia, determinata dall'utilizzo di una forma di energia diversa dalle radiazioni ionizzanti, ha reso estremamente diffuso il suo impiego diagnostico. Un apparecchio ecografico è essenzialmente costituito da: una sonda che trasmette e riceve il segnale; un sistema elettronico che pilota il trasduttore, genera l'impulso di trasmissione, riceve l'eco di ritorno della sonda e tratta il segnale ricevuto, temporizzando opportunamente l'azione di questi elementi; un sistema di visualizzazione dell'immagine ecografica. A seconda dell'esame da eseguire, per migliorare i risultati diagnostici, si possono usare vari tipi di sonde o trasduttori, ma in generale hanno caratteristiche simili. Le onde ultrasonore utilizzate in ecografia sono generate per effetto piezoelettrico da ceramiche o cristalli, quali il quarzo, dotati della proprietà di deformarsi quando venga loro applicata una tensione elettrica e generare una tensione elettrica proporzionale alla compressione cui vengono sottoposti. Usando questi materiali è possibile ottenere trasduttori che, eccitati da una tensione alternata, generano un'onda ultrasonora di intensità direttamente proporzionale alla tensione cui viene sottoposto il trasduttore stesso e inversamente proporzionale allo spessore della ceramica. Gli ultrasuoni si propagano con velocità e assorbimento legati, secondo leggi note, alle peculiari caratteristiche del mezzo. La velocità di propagazione è proporzionale alla radice quadrata del rapporto tra la specifica costante elastica di interesse e la sua densità. L'assorbimento degli ultrasuoni è correlato al tipo di tessuto e aumenta con l'aumentare della frequenza ultrasonora. Il fascio ultrasonoro subisce quindi, durante il suo percorso, un'attenuazione, che aumenta all'aumentare del percorso e della frequenza. La riflessione, cui si deve l'eco di ritorno al trasduttore, è tanto maggiore quanto più marcata è la differenza di impedenza acustica tra i due mezzi e dipende dall'angolo di incidenza della radiazione sull'interfaccia. La capacità di risoluzione di un apparecchio a ultrasuoni ne definisce la possibilità di rappresentare fedelmente un determinato oggetto. La risoluzione di contrasto dipende dalla capacità di riconoscere due zone con differente impedenza acustica, anche quando tale differenza è minima. Infatti, l'interfaccia che separa due aree aventi densità molto simile, induce una minima riflessione del fascio. La risoluzione di contrasto è tanto maggiore quanto più il sistema è capace di rilevare anche echi riflessi di bassissima intensità. La risoluzione di contrasto dipende, tra l'altro, dal numero di livelli di grigio che il sistema è in grado di presentare e, quindi, in ultima analisi, dall'accuratezza del sistema di rilevazione. La risoluzione spaziale, cioè la capacità di visualizzare un bersaglio in maniera il più possibile rispondente alla realtà, migliora anche aumentando il numero di linee che compongono l'immagine ecografica. Per fare ciò è possibile aumentare il numero di cristalli che compongono il trasduttore stesso (trasduttore ad alta densità) [1].

## 1.5 Risonanza magnetica (RM)



Figura 1.3: Un comune macchinario per la risonanza magnetica

La risonanza magnetica (RM) è una tecnica diagnostica che consente l'acquisizione di immagini di sezioni del corpo umano attraverso l'analisi del segnale elettromagnetico prodotto da nuclei atomici indotti ad assumere situazioni energetiche di equilibrio in un campo magnetico statico e successivamente eccitati da onde di radiofrequenza. Il principio su cui si basa la risonanza magnetica è, che se un fascio di radiazioni elettromagnetiche di grande lunghezza d'onda attraversa un materiale in presenza di un campo magnetico, le modalità di assorbimento delle radiazioni danno informazioni sulla struttura della materia attraversata. In particolare il fenomeno della risonanza magnetica si fonda sulle proprietà magnetiche dei nuclei dell'idrogeno (cioè i protoni), l'elemento più rappresentativo del corpo umano, che è formato in percentuale rilevante da acqua. I protoni sono particelle cariche positivamente, che si comportano come se fossero in costante movimento rotatorio attorno al proprio asse; il momento angolare intrinseco che è necessario attribuire alla particella, indipendentemente dal suo moto nello spazio, viene definito spin. Quando si pone un paziente in un apparecchio RM i protoni, assimilabili a dipoli magnetici, si allineano nella direzione del campo magnetico esterno, orientandosi secondo un verso parallelo o antiparallelo. I protoni, inoltre, non sono statici, ma hanno un movimento (detto di precessione) attorno alle linee di campo magnetico, che è paragonabile a quello di una trottola. La velocità di precessione misurata come frequenza di precessione è direttamente proporzionale all'intensità del campo magnetico. Essendo però impossibile la rilevazione di tale segnale, è necessario perturbare questo equilibrio, al fine di rendere misurabile dall'esterno l'entità del fenomeno. A questo scopo viene inviato un impulso di radiofrequenza (RF), con frequenza uguale alla velocità di precessione dei protoni. L'effetto di questo impulso sui protoni è duplice: in primo luogo essi assumono energia, passando dal verso parallelo a quello antiparallelo, con conseguente diminuzione della magnetizzazione longitudinale e, in secondo luogo,

sincronizzano il loro movimento ('precedono in fase'). Quando l'impulso è interrotto, i protoni cominciano a cedere energia e si rilassano. Questo comporta il ripristino della magnetizzazione longitudinale che avviene in un determinato lasso di tempo. Si definisce quindi come tempo di rilassamento T1 il tempo necessario per riacquistare una frazione (il 63%) della magnetizzazione longitudinale, contemporaneamente l'intervallo di tempo necessario affinché la magnetizzazione decresca del 37% rispetto al valore originale viene definito come tempo di rilassamento T2. Questi parametri dipendono, in diversa misura, dal microambiente chimico che circonda ogni singolo nucleo e dalle interazioni tra i protoni. Più semplicemente, si può dire che l'impulso di RF ribalta la magnetizzazione di 90° su di un piano trasversale, rendendola quindi misurabile come segnale dall'esterno. Il tempo che intercorre tra i due impulsi è definito tempo di ripetizione (TR) e può essere variato dall'operatore. Le fasi di esecuzione di un esame RM possono essere così riassunte: il paziente viene posizionato in un magnete, viene inviato un impulso di radiofrequenza che viene poi interrotto, il paziente emette un segnale che viene rilevato e usato per la ricostruzione delle immagini. La RM fornisce immagini dei diversi organi e strutture che si basano sulla distribuzione nel loro contesto dei nuclei mobili contenuti nell'acqua e nei lipidi. Si possono ottenere immagini differenti variando alcuni parametri durante l'esecuzione dell'esame: la scelta dei parametri programmabili si ripercuote in maniera decisiva sui tempi di scansione, sulla qualità dell'immagine e sulla risoluzione spaziale. La RM fornisce rilievi morfologici spesso più risolutivi rispetto a quelli della TC. Le immagini del cervello, per esempio, superano di gran lunga quelle ottenute con la TC per ricchezza di informazione anatomica. Il vantaggio maggiore della RM è però correlato all'utilizzo di radiazioni non ionizzanti, che la rendono una tecnica sicura, al punto da poter essere utilizzata anche in gravidanza. La natura e l'intensità degli eventi fisici in gioco è tuttavia tale da prospettare alcuni eventi potenzialmente pericolosi: effetti magnetomeccanici (torsione delle clip chirurgiche), effetti magnetoelettrici (induzione di correnti elettriche endotessutali), riscaldamento tessutale. Per tali motivi l'esame è assolutamente controindicato in pazienti portatori di elementi metallici non compatibili (clips, spirali, filtri cavali) a livello vascolare [1].

## 1.6 Pericolo per la salute delle immagini radiologiche

La scoperta dei raggi X circa 120 anni fa e lo sviluppo delle tecniche radiologiche ha comportato una rivoluzione prima in ambito diagnostico e, una volta scoperto l'effetto biologico negativo delle radiazioni, anche in ambito terapeutico. L'effetto tossico delle radiazioni, sfruttato anche a fini terapeutici come antitumorale, è prevalentemente indiretto e comporta la idrolisi dell'acqua con formazione di radicali liberi. I pionieri della radiologia hanno subito pesanti danni da radiazioni: radiodermatiti nella migliore delle ipotesi, ma spesso anche malattie sistemiche come leucemie. Una

riduzione drastica fino quasi ad annullare gli effetti delle radiazioni sugli operatori sanitari risale al massimo a 40 anni fa quando, grazie alla ricerca sulle bombe atomiche ed in concomitanza ai primi viaggi spaziali, lo studio sul nucleare è stato potenziato. Oggi le apparecchiature utilizzate nella radiologia tradizionale sono molto sicure, ma, per quanto riguarda i pazienti, a tutt'oggi il 50% delle indagini diagnostiche sono ancora composte da tecniche radiologiche tradizionali ed esse sono le responsabili del 70/80 % della dose di radiazioni cui è sottoposta la popolazione. L'UE, con la Direttiva Europea dell'Euratom 97/43 del '97 che ha come scopo la radioprotezione del paziente [2], rende obbligatoria la limitazione della dose di radiazioni che dev'essere assorbita da un paziente. Tale principio viene ulteriormente espresso, in ambito italiano, dal Decreto Legislativo n.187 del 26 maggio 2000 che enuncia: "Tutte le dosi dovute a esposizioni mediche per scopi radiologici devono essere mantenute al livello più basso ragionevolmente ottenibile e compatibile con il raggiungimento dell'informazione diagnostica" [3]. Il calcolo della dose diventa quindi un passo fondamentale a cui attenersi nello svolgere gli esami radiologici.

## 1.7 Punto di partenza dello studio

Attualmente all'Ospedale Sant'Anna è in uso il software ARGO, sviluppato in passato, anch'esso da studenti del Politecnico di Milano. Questo software permette di studiare un livello di dose di radiazioni specifico e il conseguente rischio di danno, ma non permette di riconoscere i diversi distretti anatomici in modo automatico, in quanto ogni regione corporea ha diversi livelli di tolleranza sulla quantità di radiazioni ricevute durante l'esame medico o le terapie fatte al paziente, e pertanto con ARGO è obbligatoria e indispensabile l'inserimento manuale della zona da analizzare. Ci si aspetta di promuovere un impiego razionale della Diagnostica per immagini per ridurre il rischio legato alle radiazioni ionizzanti ed ai mezzi di contrasto e per contenerne i costi. Le leggi europee ed italiane (legge 187) conferiscono al Radiologo piena responsabilità ed autonomia di scelta. Il Radiologo non solo può, ma deve rifiutarsi di eseguire indagini che ritiene non adeguate o controindicate e ha la facoltà di sostituire l'indagine non appropriata con un'altra più adatta, dopo adeguata informazione e dopo avere ottenuto il consenso del paziente [3]. Pertanto il progetto redatto va in questa direzione: automatizzare il riconoscimento delle diverse regioni anatomiche.

Nel capitolo 2 della seguente tesi si elencano e spiegano alcuni concetti fondamentali riguardanti le scienze radiologiche.

Il capitolo 3 presenta lo stato dell'arte per quello che concerne le tecniche di riconoscimento delle immagini digitali e ne illustra alcune tipologie.

Obbiettivi, vincoli, requisiti e strumenti utilizzati nel progetto di tesi vengono approfonditi nel capitolo 4.

Il capitolo 5 contiene un breve riassunto del lavoro di tesi svolto e spiega il metodo di riconoscimento che è stato sviluppato per risolvere questo particolare problema, insieme alla struttura del programma.

Nel capitolo 6 viene proposta un'analisi dei risultati ottenuti a livello di percentuali d'errore e prestazioni temporali.

Infine, nel capitolo 7 si riassume il lavoro svolto e si discute di quali possano essere gli sviluppi futuri di questo metodo.

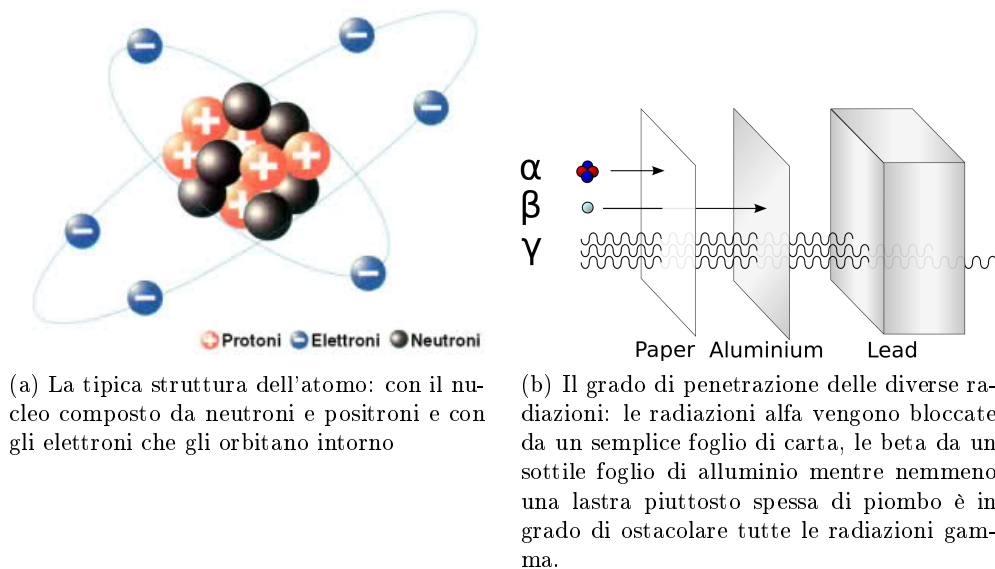
## Capitolo 2

# Cenni di radiologia

Le caratteristiche delle immagini radiologiche non hanno subito variazioni significative dalla loro comparsa nel 1895: anche le recenti immagini digitali sono sostanzialmente identiche a quelle tradizionali. Al contrario, ci sono invece state grandi innovazioni nell'ambito dell'elaborazione e dello studio delle immagini ed è un campo ancora in continua evoluzione. La radiologia convenzionale è rimasta per decenni l'unica metodica d'indagine disponibile ed è ancora oggi la più utilizzata nel settore della diagnostica per immagini. Con la T.A.C ( Tomografia Assiale Computerizzata) invece, inizia l'era della radiologia digitale che prosegue con l'introduzione di nuovi sistemi di acquisizione e trattamento dell'immagine, una pratica che si sta affermando sempre più in quasi tutti i campi della medicina. La problematica d'interesse più rilevante nell'ambito della diagnostica per immagini è quella relativa alla dose di radiazione impiegata durante l'acquisizione delle immagini. Poiché le radiazioni possono essere nocive per il corpo umano, di seguito vengono spiegati i concetti fondamentali legati alle radiazioni nucleari ed i metodi di ottenimento degli indici dosimetrici più comuni che vengono utilizzati dagli operatori sanitari per stimare ed abbassare il rischio di esposizione alle radiazioni. Dopo questi argomenti preliminari si considereranno in maniera particolare gli indici relativi alla TAC, in quanto l'oggetto principale della presente tesi è il confronto di immagini digitali tomografiche.

### 2.1 Fondamenti di fisica nucleare

Le particelle (protoni e neutroni) che costituiscono i nuclei degli atomi vengono tenute legate insieme da particolari forze, chiamate forze nucleari: ad una forza nucleare corrisponde un'energia di legame che fa stare unite tali particelle (Figura 2.1a). Un nucleo è più stabile quando è costituito da un numero uguale sia di elettroni che di neutroni, quando questo equilibrio viene alterato aggiungendo altri elettroni oppure positroni, il nucleo diventa instabile. Un nucleo instabile tende verso una naturale



(a) La tipica struttura dell'atomo: con il nucleo composto da neutroni e positroni e con gli elettroni che gli orbitano intorno

(b) Il grado di penetrazione delle diverse radiazioni: le radiazioni alfa vengono bloccate da un semplice foglio di carta, le beta da un sottile foglio di alluminio mentre nemmeno una lastra piuttosto spessa di piombo è in grado di ostacolare tutte le radiazioni gamma.

Figure 2.1: Fondamenti di radiologia

condizione di stabilità e per farlo libera energia, cioè radiazioni. Le radiazioni sono di 3 tipi: alfa (grande massa e carica positiva), beta (piccola massa e carica negativa) e gamma (pura energia senza massa e carica) (Figura 2.1b). Gli effetti indotti sulla materia dalle radiazioni dipendono dalla facoltà di ionizzare, cioè di rimuovere elettroni dagli atomi della materia che viene esposta ad esse. La ionizzazione, tipica delle radiazioni gamma, esprime il carattere nocivo delle radiazioni: nel corpo umano può portare a danneggiamenti del DNA e dell'RNA; vengono cioè spezzati alcuni legami chimici che tengono unite le catene molecolari. Rompendo le catene molecolari vengono prodotti radicali liberi che danneggiano le cellule portando alla loro morte repentina o ritardata. La quantità di ioni prodotti da una data radiazione diventa, quindi, un indice importante della pericolosità di una data radiazione [4].

## 2.2 Le unità di misura fondamentali della radiologia: Gray e Sievert a confronto

Nel Sistema Internazionale il Gray (Gy) è l'unità di misura della dose di radiazione ionizzante assorbita dalla materia. È una quantità fisica e non tiene in considerazione il contesto biologico; viene espressa dal rapporto Joule/Kilogrammo: 1 Gy rappresenta il deposito di 1 Joule di radiazioni in 1 Kg di materia [5]. Un'esposizione di tutto il corpo a 5Gy di una radiazione ad alta intensità può portare alla morte in due settimane. Tuttavia, il valore in Gy di radiazione assorbita non è un buon indicatore dei possibili effetti biologici: 1 Gy di radiazioni alfa è molto più dannoso di 1 Gy di radiazioni di fotoni. Per questo si usa il Sievert (Sv): l'unità di



misura derivata dalla dose di radiazione ionizzante che esprime gli effetti sulla salute umana in presenza di bassi livelli di radiazioni. Esprime un effetto biologico (e non una quantità fisica come Gy) ed è sempre espressa dal rapporto Joule/Kilogrammo. 1 Sv è l'effetto biologico equivalente di un deposito di 1 Joule di energia in 1 Kg di tessuto umano. Praticamente è data dal prodotto della dose in Gy moltiplicata per un fattore di qualità [6].

## 2.3 Dose Assorbita

La Dose Assorbita rappresenta la quantità media di energia assorbita dalla materia per unità di massa. Nel Sistema Internazionale l'unità di misura è il Gray espresso dal rapporto Joule/Kilogrammo. La quantità di dose assorbita è di fondamentale importanza per il calcolo della dose di radiazione anche se, usata da sola e senza modifiche, non è un indicatore adeguato del rischio per la salute di un paziente. Si è scoperto che nel calcolo della probabilità di contrarre un cancro o di avere danni genetici seguenti l'irraggiamento, bisogna tener conto anche del tipo di radiazione e della sensibilità del tipo di tessuto che viene irradiato: ciò richiede l'uso di fattori di modifica [6]. Convenzionalmente, quindi, la dose assorbita è esclusivamente utilizzata per il calcolo e la comparazione degli effetti deterministici e chiaramente visibili come nell'avvelenamento da radiazioni. Per rappresentare il rischio stocastico, cioè la probabilità di contrarre una malattia, sono invece utilizzate la Dose Equivalente e la Dose Efficace che vengono calcolate a partire dalla Dose Assorbita grazie all'uso di fattori e coefficienti appropriati. Se la dose assorbita non è uniforme o è applicata solo ad una parte di un oggetto, come nel caso del corpo umano, vale la seguente formula:

$$D_{ass,T} = \frac{\int_T D(x,y,z)\rho(x,y,z)dV}{\int_T \rho(x,y,z)dV}$$

dove T è l'oggetto interessato dalla radiazione,  $D(x,y,z)$  è una funzione della dose assorbita in funzione della posizione,  $\rho$  la funzione della densità dell'oggetto in funzione della posizione e V il volume dell'oggetto [7].

## 2.4 Dose Equivalente

La Dose Equivalente è una quantità di dose usata nell'ambito della protezione radiologica per rappresentare il rischio stocastico (ovvero la probabilità di avere danni genetici in seguito all'esposizione alle radiazioni) per radiazioni ionizzanti a basso livello. L'International Commission on Radiological Protection (ICRP) ha pubblicato delle linee guida per il calcolo della dose equivalente a partire dalla dose assorbita [6]: ha il significato di quantità massima per specificare i limiti di esposizione per

assicurarsi che i rischi per la salute siano a livelli accettabili e le reazioni all'interno dei tessuti siano nulle. Siccome questo è un valore calcolato e non può essere misurato, il suo scopo è quello di confrontare questa quantità di radiazioni assorbita con gli effetti osservabili sulla salute del paziente. La dose equivalente è calcolata moltiplicando la dose media assorbita da un tessuto o da un organo per un fattore di pericolosità  $W_R$ , che dipende dal tipo e dall'energia della radiazione. Il fattore di pericolosità rappresenta l'impatto biologico relativo e modifica la dose assorbita per tenere in considerazione anche gli effetti sul corpo delle varie radiazioni. L'ICRP ha pubblicato una tabella dove ha assegnato ogni fattore di pericolosità al tipo di radiazione in base al loro rischio relativo [7]. L'unità di misura della dose equivalente nel Sistema Internazionale è il Sievert che nel rapporto Joule/Kilogrammo esprime gli effetti biologici di 1 Joule di radiazioni in 1 Kg di tessuto umano. La formula per calcolarla è:

$$D_{eq,T} = \sum_R W_R \cdot D_{ass,T,R}$$

dove R è una tipologia di radiazione,  $D_{ass,T,R}$  è la dose assorbita di radiazione R dal tessuto T e  $W_R$  è fattore di pericolosità definito dalle regolamentazioni dell'ICRP [6, 7].

## 2.5 Dose Efficace

La quantità di radiazioni conosciuta come dose efficace è una somma pesata di diverse dosi equivalenti di tutti i tessuti e gli organi del corpo che vengono irradiati durante un esame e rappresenta il rischio stocastico per la salute, ovvero la probabilità di contrarre un cancro a seguito di una radiazione ionizzante ricevuta in quelle parti del corpo. È il valore fondamentale studiato dall'International Commission of Radiological Protection (ICRP) per limitare la dose di radiazioni assunta da un paziente in ambito di protezione radiologica [6]. La dose efficace è principalmente utilizzata per pianificare, valutare ed ottimizzare la protezione radiologica, inoltre è anche usata a scopi regolativi. La dose efficace ha portato un enorme contributo alla protezione radiologica, dal momento che ha permesso a diverse dosi, applicate a diverse parti del corpo, di essere sommate anche se provenienti da radiazioni esterne di vario tipo. L'unità di misura della dose efficace nel Sistema Internazionale è il Sievert che nel rapporto Joule/Kilogrammo esprime gli effetti biologici di 1 Joule di radiazioni in 1 Kg di tessuto umano. Il calcolo viene effettuato utilizzando dei fattori correttivi chiamati fattori di sensibilità: ogni organo o tessuto ha un diverso fattore [7]. I fattori di sensibilità dei vari tessuti corporei vengono elencati in una tabella redatta dall'ICRP, insieme alle direttive per effettuare i calcoli: ci sono alcuni tessuti molto più sensibili di altri alle radiazioni. La somma dei diversi fattori di sensibi-

lità risulta essere uguale ad 1.0 quindi, nel caso di irraggiamento uniforme su tutto il corpo, la dose efficace risulta essere uguale alla dose equivalente. E' importante considerare, inoltre, che l'irraggiamento può essere anche interno, causato cioè da isotopi radioattivi usati come mezzo di contrasto per permettere ad alcune strutture corporee di risaltare maggiormente. La formula per calcolare la dose efficace è:

$$D_{eff} = \sum_T W_T \cdot D_{eq,T}$$

dove T è il tessuto o l'organo d'interesse,  $W_T$  è il fattore di sensibilità legato a quel particolare tessuto (meglio illustrato in tabella 2.1) e  $D_{eq,T}$  è la dose equivalente assorbita dal tessuto T.

<b>Organo</b>	<b>Fattore di sensibilità</b>
Encefalo	0.01
Ghiandole salivari	0.01
Pelle	0.01
Ossa	0.01
Esofago	0.04
Tiroide	0.04
Fegato	0.04
Vescica	0.04
Gonadi	0.08
Mammelle	0.12
Polmoni	0.12
Stomaco	0.12
Colon	0.12
Midollo Osseo	0.12
Resto del Corpo	0.12
<b>Totale</b>	<b>1.0</b>

Tabella 2.1: In tabella vengono elencati i tessuti corporei ed i loro relativi fattori di sensibilità secondo le più recenti norme dell'ICRP del 2007 [6]. Si nota che la somma di tali fattori è uguale ad 1.0, quindi, se il corpo intero viene irradiato da una radiazione esterna ed omogenea la dose efficace è uguale alla dose equivalente per l'intero corpo [8].

## 2.6 Indici di Tomografia computerizzata

La Tomografia Computerizzata è attualmente la principale causa di esposizione a radiazioni nell'ambito della diagnostica per immagini ed il suo uso è sempre più in aumento. Quando una scansione TC viene eseguita, i parametri della dose vengono regolarmente segnalati, ma non sono facilmente reperibili al momento dell'analisi successiva. Malgrado le molteplici diversità tra i diversi produttori di macchinari

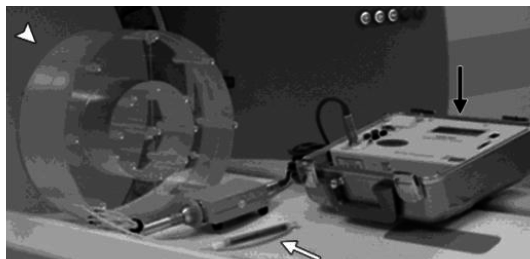


Figura 2.2: Esempio di strumentazione utilizzata per la caratterizzazione degli scanner CT: a sinistra un fantoccio acrilico composto di due sezioni che simulano tronco e testa umani e, a destra, uno strumento di rilevazione delle radiazioni.

TAC, ognuno dei quali utilizza un proprio formato per presentare i dati dosimetrici (Figura 2.2), le informazioni sono comunque sufficienti per consentire una valutazione della dose di radiazioni ed ottenere informazioni per migliorarla. In un esame TC il kilovolt di picco, cioè la massima tensione applicata attraverso il tubo radiogeno, è in proporzionalità diretta e determina il livello di energia del fascio di raggi X: modificare il valore del picco penalizza la qualità dell'immagine. Gli esami sugli adulti utilizzano 120 Kilovolt, ma un picco ha solitamente un voltaggio di 80-100 Kilovolt: studi eseguiti con l'imaging a basso voltaggio dimostrano riduzioni di dose del 45-50%. La carica anodica, cioè il prodotto di corrente-tempo, è un altro valore che può venire utilizzato nel calcolo della dose in quanto il dosaggio varia linearmente con la corrente che passa nel tubo. Meno carica anodica significa meno dose di radiazioni assunta dal paziente, ma maggior rumore presente nell'immagine. Nei sistemi moderni la corrente viene modulata in automatico per ridurre la dose al paziente dal 20 al 40%. Per poter costruire ed utilizzare gli indici di dose sui pazienti reali è necessario focalizzare lo studio nella regione anatomica che si vuole analizzare, per questo viene introdotto il concetto di ROI, regione d'interesse usata per isolare una parte dell'immagine sulla quale verrà approfondito lo studio. La TC ha degli indici di dose personalizzati: i CTDI (CT Dose Index) espressi in milligray ( $1 \text{ Gy} = 1 \text{ J/Kg}$ ) ed indicano la produzione di energia dello scanner TC. Il più semplice da calcolare tra questi è il  $\text{CTDI}_{100}$  ed indica l'energia ricevuta al centro di un fantoccio di forma cilindrica in acrilico (Figura 2.2) proveniente da una camera di ionizzazione da 100-mm.  $\text{CTDI}_W$  è una media ponderata dell'energia su tutto il fantoccio: viene cioè tenuto conto delle differenze di energia tra una zona periferica ed il centro. Infine il  $\text{CTDI}_{Vol}$ , è un calcolo affinato di  $\text{CTDI}_W$  che tiene conto dei kilovolt e della carica anodica (milliampere-secondi): rappresenta il rendimento energetico dello scanner e specifica l'intensità della radiazione usata. Per stimare la dose al paziente viene utilizzato anche il prodotto dose-lunghezza (DLP): rappresenta il prodotto tra la lunghezza della scansione in centimetri e il  $\text{CTDI}_{Vol}$ . Il DLP può essere convertito in una dose stimata efficace moltiplicandolo per un fattore di conversione di nome "k-factor". Ogni tipo di tessuto corporeo ha un diverso K-factor e tale valore è redatto

da ICRP. Il cambiamento di dose efficace dal gray al sievert riflette l'applicazione dei fattori di ponderazione tissutale [9]. Con DLP è possibile determinare una dose efficace stimata per l'esame CT di un paziente: abbassando la lunghezza di scansione si abbassa la dose al paziente, ma si rischia di non inquadrare la patologia.

Vi sono inoltre altri parametri che possono influenzare la dose e cioè età e dimensione del paziente: in pazienti più giovani è maggiore il rischio di essere soggetti agli effetti negativi delle radiazioni a causa del potenziale più alto di mutazioni. I protocolli TC dovrebbero sempre essere adattati alle dimensioni del paziente per ottimizzare la qualità delle immagini diagnostiche e per massimizzare il rapporto di beneficio-rischio per esposizioni alle radiazioni ionizzanti. Di seguito si spiega in maniera più dettagliata il significato dei termini utilizzati in campo TAC.

### 2.6.1 Region of Interest (ROI)



Figura 2.3: Una region of interest creata grazie al Software Argo: il calcolo di SSDE verrà svolto solo sulla zona dell'immagine all'interno del quadrato giallo, tenendo conto delle dimensioni e del tipo di tessuto al suo interno.

La ROI (Region of Interest) è l'area d'interesse dell'immagine sulla quale si vogliono effettuare rilevamenti, misure, elaborazioni [10]. Viene delimitata da un indicatore visivo che può essere poligonale (Figura 2.3), circolare o ellittico. Il tracciamento della ROI in un'immagine può essere effettuato sia da un utente utilizzando il mouse, oppure in maniera automatica fornendo al software le coordinate dell'area a cui si è interessati. L'utilità della ROI nasce dal fatto che, delimitando una sola parte dell'immagine e cioè quella più rilevante ad un'indagine, tutti gli sforzi di tipo osservativo e computazionale vengono concentrati solo su quell'area, tralasciando la restante parte dell'immagine che è poco interessante. Tutto ciò ha un grossissimo

impatto sul carico computazionale a cui sottoporre il calcolatore, rendendo più veloce l'attività da svolgere.

### 2.6.2 CTDI

Rappresenta la dose assorbita media lungo l'asse  $z$  di una serie di irradiazioni contigue (viene misurato per una singola rotazione della sorgente a raggi X). Viene calcolata dividendo la dose assorbita integrata dalla collimazione del fascio nominale totale (Capitolo 1 Sezione 1.3) ed offre un metodo facile di stimare questo valore risparmiando una notevole quantità di tempo [11]. La sua formula:

$$CTDI = \frac{1}{NT} \int_{-\infty}^{+\infty} D(z) dz$$

dove  $N$  è il numero di sezioni della tomografia (che al massimo è uguale al numero di canali usati in una sola scansione),  $T$  è lo spessore della singola sezione e  $D(z)$  è il profilo della dose di radiazioni lungo l'asse  $z$  del paziente [12].

### 2.6.3 CTDI<sub>100</sub>

E' un parametro misurato di esposizione alle radiazioni. E' più conveniente da usare del CTDI ed è sempre calcolato durante la preparazione clinica. E' una dose ottenuta mediante una camera di ionizzazione che integra l'esposizione alle radiazioni di una singola scansione assiale sulla lunghezza di 100mm [13]. Gli eventi di ionizzazione producono, nella camera, corrente in maniera proporzionale al loro numero. L'unità di misura nel Sistema Internazionale è coulomb/Kilogrammo. Rappresenta la distribuzione spaziale dell'esposizione alla radiazione. La formula:

$$CTDI_{100} = \frac{1}{NT} \int_{-50mm}^{+50mm} D(z) dz$$

dove  $N$  è il numero di sezioni della tomografia (che al massimo è uguale al numero di canali usati in una sola scansione),  $T$  è lo spessore della singola sezione e  $D(z)$  è il profilo della dose di radiazioni lungo l'asse  $z$  del paziente [14].

### 2.6.4 CTDI<sub>W</sub>

E' una media pesata di alcune misure di CTDI<sub>100</sub> al centro ed in posizioni periferiche del campo di vista. Riflette la dose media assorbita sulle dimensioni  $x$  ed  $y$  del fantoccio. E' usata per convertire CTDI che è in C/Kg in una dose assorbita che è in Gray (Gy). La formula:

$$CTDI_W = \frac{1}{3}CTDI_{100,center} + \frac{2}{3}CTDI_{100,edge}$$

dove  $CTDI_{100,center}$  è la misura di  $CTDI_{100}$  al centro del campo di vista e  $CTDI_{100,edge}$  è la misura di  $CTDI_{100}$  nelle zone periferiche del campo di vista [15].

### 2.6.5 $CTDI_{Vol}$

È un recente parametro per il calcolo della dose di radiazioni accettato dalla Commissione Elettrotecnica Internazionale. Viene calcolato direttamente da  $CTDI_W$ , da dati quindi facilmente calcolabili poiché derivati da  $CTDI_{100}$ . Questo Dose Index fa una media tra le 3 direzioni x, y e z ed è la migliore espressione della dose di radiazione in ambito di Dosimetria con TC. La maggior parte degli scanner TC mostrano il valore di  $CTDI_{Vol}$  sulla console dell'operatore per aiutare il medico a comparare le dosi di radiazioni che il paziente riceve da diversi protocolli di immagini. Anche in questo caso l'unità di misura è il Gray.  $CTDI_{Vol}$  ha tuttavia un grande svantaggio: il valore numerico dipende solo dalla distribuzione spaziale delle singole scan e non dal numero totale di successive scansioni durante un esame TC (cioè, la lunghezza totale della scansione). La formula è:

$$CTDI_{vol} = \frac{CTDI_W}{pitch} = \frac{NT}{I}CTDI_W$$

dove  $pitch$  è definito come il rapporto dello scostamento per rotazione (ovvero  $I$  che sarebbe lo spostamento assiale del lettino usato per la scansione in mm) e la larghezza del fascio nominale (data dal prodotto  $NT$  dove  $N$  è il numero di scansioni e  $T$  è lo spessore della singola scansione) [14]. Quindi  $pitch = \frac{I}{NT}$ .

### 2.6.6 Dose-Length Product (DLP)

DLP è un indicatore della dose di radiazione totale di un intero esame TC [12]. Il DLP incorpora in sé il numero di scansioni e la loro larghezza, quindi DLP aumenta all'aumentare della lunghezza totale della scansione, oppure, con le variabili che sono necessarie, al calcolo di  $CTDI_{Vol}$ . Dal momento che la lunghezza di scansione è espressa in centimetri, l'unità di misura per DLP nel Sistema Internazionale è il milligray per centimetro ( $mGy \cdot cm$ ). Un display con il valore di DLP è un requisito legale necessario in molti paesi Europei, tra cui l'Italia. Anche se il DLP riflette in maniera più efficace la dose di radiazione di un esame TC, il suo valore numerico è soggetto alle variazioni dell'anatomia del paziente (più alto per persone alte). Quindi,  $CTDI_{Vol}$  è più utile nella fase di imaging e per confrontare i risultati di differenti protocolli. La formula di DLP è:

$$DLP = CTDI_{vol} \cdot l$$

dove  $l$  è la lunghezza della scansione che viene effettuata [11].

### 2.6.7 Dose Efficace in CT

La dose efficace riflette l'assorbimento non uniforme di radiazione successiva a parziali esposizioni di zone del corpo permettendo le comparazioni di rischio tra diversi protocolli di esami in TC. L'unità di misura nel Sistema Internazionale è il sievert (Sv). Una tecnica chiamata simulazione Monte Carlo è usata per determinare le dosi specifiche per ogni organo, simulando l'assorbimento e la diffusione ottica di fotoni di raggi X nei vari tessuti [16]. Una ragionevole approssimazione usata in ambito CT per il calcolo della dose efficace è ottenibile moltiplicando DLP per un fattore  $k$  ( $\text{mSv} \cdot \text{mGy}^{-1} \cdot \text{cm}^{-1}$ ) che dipende dalla zona del corpo. Questi fattori  $k$  sono diversi da quelli usati per la Dose Efficace tradizionale. La semplice formula è:

$$D_{eff} = k \cdot DLP$$

dove  $k$  è un fattore numerico noto e che varia in base all'organo che viene analizzato seguendo i valori in tabella 2.2.

Età	Encefalo	Collo	Torace	Addome/Pelvi
0	0.011	0.017	0.039	0.049
1 anno	0.0067	0.012	0.026	0.03
5 anni	0.004	0.011	0.018	0.02
10 anni	0.0032	0.0079	0.013	0.015
adulto	0.0021	0.0059	0.014	0.015

Tabella 2.2: Il metodo più semplice per il calcolo della dose efficace a partire da DLP è usare i fattori di conversione noti e pubblicati che sono elencati in tabella. I valori numerici di tali fattori variano in base all'età (pediatrica o adulta) ed in base alla macroregione che viene analizzata durante un esame TC [17]. Tali fattori hanno come unità di misura  $\text{mSv} \cdot \text{mGy}^{-1} \cdot \text{cm}^{-1}$ .

### 2.6.8 Size Specific Dose Estimates

La dose ricevuta durante un esame TC dipende sia dalle dimensioni del paziente che dalla radiazione in uscita dallo scanner: gli indici CTDI forniscono solo informazioni riguardo i livelli di radiazione senza tener conto della dimensione del paziente. In caso di dimensioni non standard del paziente, usando CTDI e DLP, si può incorrere in sottostime addirittura di un fattore 2-3 volte inferiore a quanto invece è la realtà dei fatti. SSDE è un fattore che applicato alla dose CTDI, consente una stima di dose migliorata, poiché prende in considerazione anche le dimensioni del paziente. Vengono utilizzati dei calibri elettronici prima dell'esame per calcolare la dimensione del lato del corpo (LAT) e lo spessore laterale del corpo (AP) ed infine calcolare il diametro effettivo che ha la stessa area del prodotto di LAT per AP ed è a forma di ellisse. Il fattore di conversione SSDE è quindi calcolato per interpolazione da 2



set di dati : usando manichini di forma umana e mediante il metodo di simulazione Monte Carlo<sup>1</sup> [18].

---

<sup>1</sup>Una simulazione Monte Carlo consiste nello studio del comportamento di un sistema, in cui sono coinvolte anche molte variabili aleatorie, mediante la sua riproduzione in un contesto controllabile. Nella simulazione si costruisce un modello matematico costituito da equazioni che descrivono le relazioni tra le componenti del sistema oggetto di studio e il loro legame con il suo funzionamento/comportamento, con l'obiettivo di effettuare esperimenti sul modello matematico ed assumendo che i risultati di tali esperimenti costituiscano una riproduzione sufficientemente accurata del comportamento che avrebbe il sistema per carpirne il funzionamento. Si utilizza questo approccio quando la soluzione per via analitica risulta troppo complessa o impossibile.

## Capitolo 3

# Stato dell'arte

L'avvento dell'informatica ha influito in maniera crescente sull'evoluzione della Sanità. Tra le discipline mediche che hanno beneficiato di tali progressi figura la diagnostica per immagini (imaging).

La possibilità di avvalersi delle capacità di elaborazione dei computer ha portato ad una nuova generazione di apparecchiature radiologiche le cui prestazioni forniscono al medico un prezioso aiuto nel percorso diagnostico e clinico. Corollario di questa innovazione è il cambiamento delle modalità con cui le immagini radiologiche stesse vengono realizzate. Si è sostanzialmente abbandonato il supporto fisico su pellicola radiografica approdando ad una soluzione più pratica e di facile gestione: l'immagine digitale.

Se in ambito radiologico la "rivoluzione informatica" è stata universalmente riconosciuta, non altrettanto si può dire nel resto del mondo sanitario. L'avvento del Compact Disc (CD) e del Personal Computer (PC), dove spesso vengono immagazzinate le immagini, è per molti ancora problematico. Il divario tra l'imaging e la clinica è accresciuto dalla scarsa abilità in ambito informatico di buona parte della classe medica. Alla difficoltà di disporre ed utilizzare computer adeguati consegue una rinuncia alla valutazione diretta dell'immagine radiologica, *modus operandi* considerato opportuno se non fondamentale che deve esser proprio di qualunque specialista. Parallelamente ad una timida campagna di alfabetizzazione informatica dei medici, messa in atto da varie istituzioni, nel corso degli ultimi anni sono stati sviluppati programmi per computer di facile comprensione ed immediato utilizzo.

Un passo successivo per facilitare la lettura dei diversi supporti radiologici è stata l'introduzione dello standard per immagini DICOM, che oltre all'informazione visiva dell'immagine contiene diverse informazioni sul paziente.

Un ulteriore progresso della procedura è la memorizzazione tramite PACS (Picture Archiving and Communication System) che permette di velocizzare l'archiviazione e la consultazione dell'indagine radiologica senza l'utilizzo di un supporto "fisico" che può essere anche smarrito o rovinato, ma attraverso un sistema hardware e soft-

ware dedicato all'archiviazione, trasmissione, visualizzazione e stampa delle immagini diagnostiche digitali, visibili anche in sedi ospedaliere distanti tra loro.

Il riconoscimento delle immagini radiologiche avviene attualmente attraverso tre macro categorie che si differenziano per il tipo di approccio che viene utilizzato. La più semplice avviene attraverso la lettura dei descrittori alfanumerici ovvero degli header nei file DICOM; in alternativa si può utilizzare il confronto diretto del contenuto delle immagini, oppure, utilizzare un approccio ibrido tra le due prima menzionate.

### 3.1 Descrittori alfanumerici nelle immagini DICOM

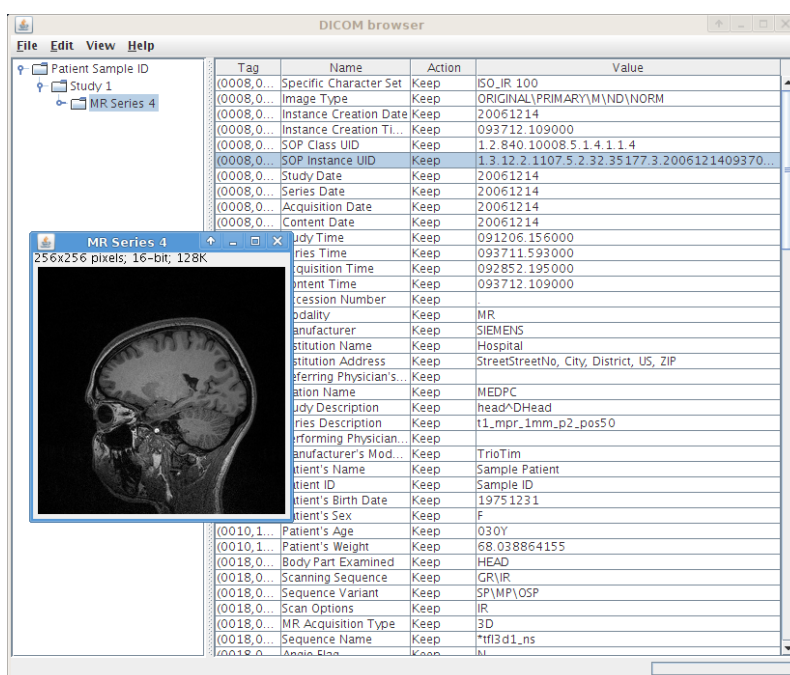


Figura 3.1: Un'immagine DICOM contiene un numero di tag che traghettano le informazioni su paziente ed esame: nell'illustrazione c'è un esempio di alcuni valori che compongono un'immagine DICOM

La necessità di inviare e scambiare immagini mediche con varie informazioni fondamentali allegate all'analisi, utilizzando reti telematiche, ha fatto sorgere l'esigenza di uno standard di comunicazione globale che permetta al radiologo la consultazione e l'accesso a tutti gli esami effettuati da uno stesso paziente, attraverso un collegamento alla rete locale. I nuovi sistemi radiografici digitali che operano seguendo gli standard DICOM (Digital Imaging and Communications in Medicine) rappresentano un modello di tale standardizzazione, la cui nascita è stata inizialmente ostacolata dalle ditte produttrici che avevano creato situazioni di completa incompatibilità tra apparecchiature di marchi differenti. Lo standard DICOM definisce alcuni criteri per

la comunicazione, la visualizzazione, l'archiviazione e la stampa di informazioni di tipo biomedico, quali ad esempio le immagini radiologiche, realizzando un chiaro e dettagliato modello per descrivere in maniera consona una serie di informazioni che, nella loro totalità, formano il dato radiologico ed esprimono il modo in cui diversi dati sono collegati tra loro. Un'entità del mondo reale (paziente, immagine, esame, etc.) viene definita come un oggetto che possiede molti attributi (Fiugra 3.1). Definiti gli oggetti e tutti i loro attributi, lo standard DICOM definisce quali operazioni possono essere eseguite e su quali oggetti: è molto simile al paradigma informatico di "Object Oriented" [19]. Il contenuto di questi file DICOM (DCM) include tutti i dati anagrafici del paziente insieme a molti altri dati medici pertinenti. Ad esempio un file di immagine a raggi x di un torace, in realtà contiene l'ID del paziente all'interno del file, in modo che l'immagine non possa mai essere separata da questa informazione, il protocollo della tipologia di esame (in questo caso RX\_Torace o un codice equivalente) ed un informazione che permette di ottenere le dimensioni dell'immagine reali in mm partendo dal numero di pixel. Lo standard DICOM è noto, ma ,solo alcune informazioni, vengono memorizzate come dati pubblici e sono accessibili utilizzando qualsiasi lettore DICOM, mentre per altre, definite come campi privati, è necessaria una cooperazione da parte dei costruttori delle apparecchiature radiologiche di marche diverse. Addirittura il formato stesso del file e i dati riguardanti il contenuto dell'immagine non solo non sono sempre accessibili, ma in base al produttore possono addirittura variare di posizione, di sintassi e di formattazione. La capillare diffusione di DICOM si rivela essere estremamente vantaggiosa proprio perché permette di avere una solida base di interscambio di informazioni tra apparecchiature di diversi produttori, server e PC, specifica per l'ambito biomedico. Occorre notare che DICOM non è solo uno standard industriale (ovvero usato da tutti i produttori di apparecchiature mediche per l'imaging), ma anche uno standard de jure (in riferimento alla commissione tecnica Health Informatics ISO 12052:2006) [20].

Il metodo che normalmente viene utilizzato nelle strutture ospedaliere è quello di accedere, con appositi programmi (ARGO, Osirix , Agnosco DICOM viewer, ecc.), all'header in cui vengono salvate le informazioni riguardanti la zona del corpo del paziente sottoposta ad esame radiologico e più precisamente all'identificativo ID di valore 0018,1030 "**ProtocolName**" che, univocamente, rileva l'esatta parte corporea analizzata. Questo metodo però non è attendibile, in quanto la creazione dell'immagine DICOM viene svolta manualmente dall'operatore sanitario, pertanto soggetta a sviste, dimenticanze ed errori umani.

## 3.2 Analisi del contenuto delle immagini radiologiche

Nel panorama informatico-medico esistono moltissimi algoritmi dedicati al ritrovamento automatico di macropatologie in vari tessuti corporei, oppure al processo

di registrazione, ovvero l'operazione che viene effettuata per ridurre la deformazione corporea nell'immagine dovuta alla respirazione del paziente, ai suoi movimenti, ecc. Tra quelli esistenti non è stato rilevato alcun algoritmo in grado di riconoscere automaticamente distretti anatomici in immagini radiologiche; nella normalità si dà credito all'operatore che ha effettuato l'esame e ha inserito l'informazione come sopra descritto. Occorre, quindi, trasferire il discorso a tecniche di confronto e riconoscimento immagini non strettamente legate all'ambito medico, ma applicabili in qualunque contesto.

### 3.3 L'utilizzo delle immagini Scout: un approccio ibrido

Il risultato di un esame TC è una serie di immagini (slice) allineate perpendicolarmente all'asse definito dalla spina dorsale del paziente. Ogni slice rappresenta una fetta del corpo del paziente di un determinato spessore. Per lo studio del riconoscimento dei vari distretti anatomici, invece di usare direttamente la serie di slices vengono utilizzate le immagini Scout (o Localizer in inglese): una vista esterna preliminare all'esame che viene sempre svolta per avere una visione d'insieme della parte da analizzare. Anche così facendo le problematiche e le eccezioni tuttavia sono le più svariate, in quanto si è rilevato, ad esempio, che la scout può essere presa di profilo o frontalmente, le braccia possono essere alzate, lungo i fianchi, incrociate, etc., parte della testa può essere tagliata nella foto, oppure storta e alcune volte sono presenti parti dell'attrezzatura o del lettino. Individuare le immagini Scout tuttavia facilita molto il lavoro di localizzazione del distretto anatomico e si parla di approccio ibrido in quanto, per individuare il Localizer da un'insieme di immagini radiologiche, è necessario leggere nell'header chiamato **ScanOptions** di valore 0018,0022 la dicitura "SCOUT MODE" ed una volta individuata si potranno usare varie tecniche di riconoscimento e confronto tra immagini usando solo le Scout .

### 3.4 Riconoscimento di immagini

Quanto viene qui di seguito spiegato è valido sia per riconoscere il soggetto in immagini radiologiche generiche, che con l'ausilio delle sole immagini scout. Alcuni passi che possono contribuire a un corretto confronto tra immagini, si possono riassumere nelle seguenti azioni di analisi:

1. pre-processare le immagini, aumentando il contrasto, filtrando il rumore
2. segmentare l'immagine: processo in cui l'immagine viene convertita in regioni, in cui i pixel della stessa regione sono simili a quelli della stessa regione di un'altra immagine e diversi da quelli delle altre regioni. Questo può essere fatto utilizzando degli algoritmi di classificazione oppure il clustering.

3. creare dei descrittori per le regioni: possono includere perimetro, area, colore, ecc.
4. risegmentazione delle immagini: processo in cui le diverse regioni sono riunite e possono essere considerate come appartenenti allo stesso oggetto
5. considerare solo le regioni rilevanti per quanto necessario, valutare eliminando le regioni più piccole o insignificanti dal punto di vista del confronto
6. salvare i descrittori delle immagini per ulteriori processi
7. utilizzare i descrittori per confrontare il contenuto di altre immagini utilizzando algoritmi di classificazione clustering, intelligenza artificiale, ecc.

Si deve tenere in considerazione che i passi appena descritti sono generici, applicabili in qualunque situazione o ambito, ma non danno garanzie di successo. In questo contesto di confronto di immagini, si deve tenere in considerazione che con il termine “similare” si intende il paragone tra immagini che sono alla vista non identiche, ma che hanno la medesima natura, appartenenti alla medesima classe di oggetti. In questo caso non cambiano solo le condizioni esterne o il punto di vista, ma spesso anche le caratteristiche intrinseche dell’oggetto. Verranno di seguito descritte in maniera generale e sintetica tutte le più importanti tecniche utilizzate per il riconoscimento di immagini.

### 3.4.1 Indici di somiglianza

Può essere utile per confrontare tra loro delle immagini, avere un valore numerico che dica in termini di quantità matematica quanto alcune immagini sono diverse tra loro. Qui di seguito gli indici di somiglianza più impiegati e noti.

#### 3.4.1.1 Distanza euclidea ed IMED

Il calcolo della distanza euclidea è il più immediato indice che permette di capire se due immagini sono simili a livello visivo, ma senza avere garanzie che quanto ci sia rappresentato nell’immagine contenga lo stesso oggetto. Rappresenta una misura quantitativa, per le immagini digitali, in quanto esse vengono espresse come matrici bidimensionali. Si definisce distanza euclidea  $D$  per immagini di dimensioni uguali, la classica misura geometrica ottenuta dalla nota relazione:

$$D = \sqrt{\sum_{k=1}^{MN} (x^k - y^k)^2}$$

dove  $x$  ed  $y$  sono due immagini  $M$  per  $N$  e  $x^k$  ed  $y^k$  rappresentano i livelli di grigio alla locazione  $k$  (con  $k$  pixel dell’immagine  $MN$ ) [21].

La distanza Euclidea è una misura semplice dal punto di vista intuitivo, ma dal punto di vista computazionale risulta onerosa a causa dell'operatore di radice e dei valori non interi di  $D$ . La distanza euclidea è sensibile alla deformazione, cioè confrontando un'immagine con una sua versione deformata, i risultati che si ottengono non sono affidabili come invece ci si potrebbe aspettare. In caso di immagini leggermente deformate non tiene nemmeno conto della distanza spaziale tra pixel. Se si vuole rendere la metrica insensibile a piccole deformazioni, è necessario tenere conto anche dell'informazione sulle relazioni spaziali tra pixel. Per questo è stata introdotta una versione modificata della distanza Euclidea ovvero la Image Euclidean Distance (IMED). IMED ha le seguenti proprietà: piccole deformazioni significano piccola distanza tra immagini e la distanza tra due immagini rimane invariata se viene effettuata traslazione, rotazione e riflessione. Inoltre la metrica si può applicare senza particolare perdita di bontà del risultato anche su immagini di ogni dimensione e risoluzione. Qui di seguito la formula:

$$D_{IME} = \sqrt{\frac{1}{2\pi} \sum_{i,j=1}^{MN} \exp\left\{-\frac{(P_i - P_j)^2}{2}\right\} (x^i - y^i)(x^j - y^j)}$$

dove  $x$  ed  $y$  sono due immagini  $M$  per  $N$ ,  $x^i, y^i, x^j$  e  $y^j$  rappresentano i livelli di grigio alla locazione  $k$  (con  $k$  pixel dell'immagine  $MN$ ) ed infine  $(P_i - P_j)$  è la distanza tra il pixel in posizione  $i$  ed il pixel in posizione  $j$  [21].

### 3.4.1.2 Color Histogram

Un altro metodo piuttosto immediato e poco conosciuto è l'istogramma dei colori: una rappresentazione della distribuzione dei colori dell'immagine. Nelle immagini digitali il color histogram esprime il numero di pixel che ha un valore specifico all'interno di una scala di colori; in caso di immagini monocromatiche (come sono quelle radiologiche) si parla di istogramma dell'intensità. L'istogramma fornisce un riassunto compatto della effettiva distribuzione dei dati all'interno dell'immagine ed è invariante alla rotazione e alla traslazione. Fare la comparazione di istogrammi tra due immagini è particolarmente efficace per il riconoscimento di oggetti in posizioni sconosciute e che possono essere anche rotati all'interno della scena [22]. E' da tenere in considerazione però che l'istogramma si concentra sulla proporzione del numero di differenti tipi di colore senza considerare la loro risoluzione spaziale. In generale, se la distribuzione dei colori dello sfondo e del primo piano è diversa, è molto probabile che si sia in presenza di una distribuzione bimodale nell'istogramma (due picchi con una valle al centro). Il principale difetto dell'istogramma consiste nel fatto che la rappresentazione dell'oggetto che viene studiato sta nel solo colore tralasciando forma e textures. Due istogrammi di colore possono essere potenzialmente identici per due immagini molto diverse, ma che per puro caso possono contenere la stes-

sa informazione sul colore. Come indice di similarità tra istogrammi vengono usati solitamente la Distanza Euclidea, l'intersezione tra istogrammi o le distanze coseno o quadratiche, ma la metrica che è risultata essere la più efficace è la distanza di Bhattacharyya utilizzabile quando la deviazione standard è la stessa per entrambi gli istogrammi che si stanno comparando. La formula di Bhattacharyya è:

$$d(H_1, H_2) = \sqrt{1 - \frac{1}{\sqrt{\bar{H}_1 \bar{H}_2 N^2}} \sum_I \sqrt{H_1(I) \cdot H_2(I)}}$$

dove  $H_1$  ed  $H_2$  sono due istogrammi con  $N$  canali,  $H_1(I)$  ed  $H_2(I)$  è il valore dell'istogramma nel canale  $I$ , e i valori a denominatore sono calcolati così  $\bar{H}_k = \frac{1}{N} \sum_J H_k(J)$  [23].

### 3.4.1.3 Structural Similarity (SSIM)

La structural similarity (SSIM) è un indice utilizzato per misurare la similarità tra immagini. La qualità dell'immagine si misura a partire da un'immagine non compressa e senza alcuna distorsione come riferimento. SSIM è stato progettato per migliorare i metodi tradizionali come segnale di picco di rumore e errore quadratico medio, che hanno dimostrato di essere in contrasto con la percezione dell'occhio umano. La differenza rispetto ad altre tecniche come MSE<sup>1</sup> o PSNR<sup>2</sup> è che questi approcci stimano errori percepiti; d'altra parte, in SSIM come degradazione dell'immagine s'intende il cambiamento percepito nell'informazione strutturale. L'informazione strutturale è l'idea che i pixel hanno forti dipendenze specialmente quando sono spazialmente vicini. Queste dipendenze trasportano importanti informazioni sulla struttura degli oggetti nella scena visiva. La metrica SSIM è calcolata in varie finestre di un'immagine. La misura tra le due finestre  $x$  ed  $y$  di dimensioni comuni  $N \times N$  è:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

dove  $\mu_x$  e  $\mu_y$  sono le medie di  $x$  ed  $y$ ,  $\sigma_x^2$  e  $\sigma_y^2$  sono la varianza di  $x$  ed  $y$ ,  $\sigma_{xy}$  è la covarianza di  $x$  ed  $y$  e  $c_1$  e  $c_2$  sono due variabili usate per stabilizzare la divisione [24].

Per valutare la qualità dell'immagine questa formula viene applicata solo sulla luminanza. L'indice SSIM risultante è un valore decimale compreso tra -1 e 1, e il valore 1 è raggiungibile solo nel caso di due insiemi di dati identici. In genere si è calcolato sulle regioni di dimensione  $8 \times 8$ . Alcune ricerche hanno dimostrato che SSIM non è molto preciso come ci si aspetterebbe ed in realtà fornisce alcuni

<sup>1</sup>MSE (mean square error): in statistica indica la discrepanza quadratica media fra i valori dei dati osservati e i valori dei dati stimati.

<sup>2</sup>PSNR (peak signal to noise ratio): termine ingegneristico per esprimere il rapporto tra la massima potenza di un segnale e la potenza di un rumore che può invalidare la fedeltà della sua rappresentazione compressa.



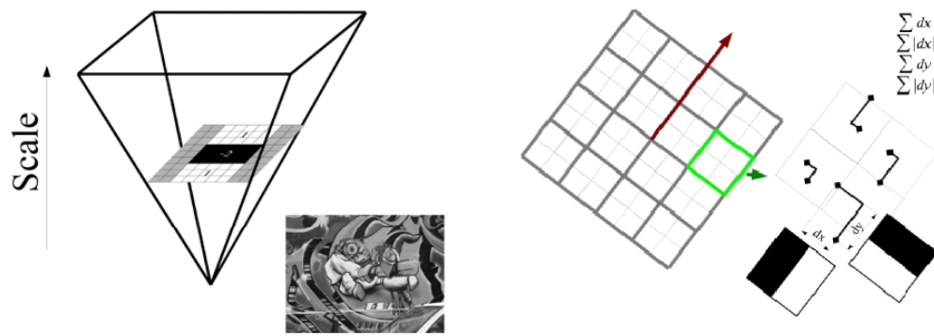
risultati non migliori rispetto ad MSE. Anche se SSIM dovrebbe riprodurre la percezione all'occhio umano, la sua formula non contiene nessun modello elaborato della percezione visiva ed invece utilizza calcoli puramente matematici.

### 3.4.2 Features detection

La feature detection (rilevamento di caratteristiche) permette attraverso l'utilizzo di diversi metodi, l'estrapolazione di informazioni da una immagine al fine di poter individuare una caratteristica particolare in una determinata regione. Le caratteristiche individuate sono usate spesso come punto di partenza di molti algoritmi. Pertanto la qualità stessa dell'algoritmo è spesso influenzata dalla bontà del suo feature detection. Una proprietà desiderabile per un feature detection è la sua ripetibilità, ovvero se la stessa feature sarà localizzata in due o più differenti immagini della stessa scena. La feature detection è un'operazione di image processing a basso livello ed esamina ogni pixel. È la prima operazione che si fa su un'immagine. Se invece fa parte di un algoritmo, allora, di solito, esamina solamente la regione individuata. Una volta che le caratteristiche importanti sono state individuate in una regione dell'immagine, tale caratteristica può essere estratta ed il risultato, chiamato feature descriptor, può essere salvato ed utilizzato per ulteriori analisi. Le più comuni tipologie di feature descriptor sono SURF (speeded up robust transform) ed HOG (Histogram of Oriented Gradients).

#### 3.4.2.1 Speeded Up Robust Features (SURF)

Questo algoritmo è nato dall'esigenza di migliorare le performance e la robustezza alle trasformazioni delle immagini di un altro feature descriptor (SIFT) particolarmente usato ed utile solo in ambito di riconoscimento facciale e per oggetti di una stessa scena visti da una prospettiva diversa. SURF lavora esclusivamente sulla luminanza dell'immagine, ovvero sui toni di grigio e mediante diverse fasi arriva al calcolo di un descrittore. Il primo passo è quello di calcolare una rappresentazione intermedia dell'immagine da processare, chiamata immagine integrale, data dalla somma dei valori dell'intensità dei pixel compresi tra un punto e l'origine. L'individuazione dei keypoints, cioè dei punti chiave per l'estrazione delle features, è basata su una matrice Hessiana, che viene approssimata mediante l'uso delle immagini integrali, e su un filtro Gaussiano per individuare punti d'interesse con fattori di scala differenti. Per realizzare il ridimensionamento di scala si passa alla rappresentazione dell'immagine nello Scale-space, ovvero una piramide rovesciata di immagini, che è ottenuta applicando diverse funzioni di tipo gaussiano con filtri sempre più grandi (Figura 3.2a). Per ogni scala si calcolano i massimi per interpolazione coi punti vicini che saranno i keypoints. Per rendere il descrittore invariante alla rotazione viene identificato un orientamento riproducibile, per ogni keypoints estratto nella prece-



(a) Estrazione dei keypoints in SURF: un'immagine (in basso a destra in questa figura) viene convertita in una matrice Hessiana approssimata (la figura all'interno della piramide rovesciata) e nella rappresentazione all'interno dello Scale-Space (la piramide rovesciata in figura) si calcolano i keypoints.

(b) Descrittore SURF: il centro di questa matrice in figura è uno dei keypoints estratti e la freccia rossa la sua direzione dominante ottenuta calcolando la risposta alle wavelet di Haar per ogni matrice nell'intorno di tale punto. Un esempio della risposta alle wavelet di Haar per la piccola porzione della matrice in verde, è descritta nella porzione d'immagine a destra.



(c) Alcuni esempi di vettori SURF diversi e rappresentati in base a diverse regioni considerate

Figure 3.2: Esemplicazione visiva di SURF

dente fase. Viene quindi stimata la direzione dominante calcolando la somma di tutte le wavelet di Haar nell'intorno di un punto massimo tra i keypoints, ovvero una sequenza di funzioni di "forma quadrata" simili alla trasformata di Fourier e viene assunta come orientamento locale (Figura 3.2b). Il descrittore viene infine determinato in modo da preservare il maggior numero di informazioni spaziali, costruendo una griglia quadrata centrata sul punto d'interesse ed orientata lungo l'orientamento principale assegnato al passo precedente. Da questa struttura viene calcolata la risposta della Wavelet di Haar centrata nel punto d'interesse ed ottenuta una serie di vettori (Figura 3.2c) risultato della descrizione [25].

### 3.4.2.2 Histogram Of Oriented Gradients (HOG)

L'algoritmo, che è definibile come un descrittore locale, ovvero estrae caratteristiche da regioni interne all'immagine, è costituito da due fasi distinte: l'addestramento dove viene creato un classificatore mediante un set di esempi positivi (ciò che vogliamo cercare nell'immagine) e negativi (ciò che non vogliamo trovare), e la fase di

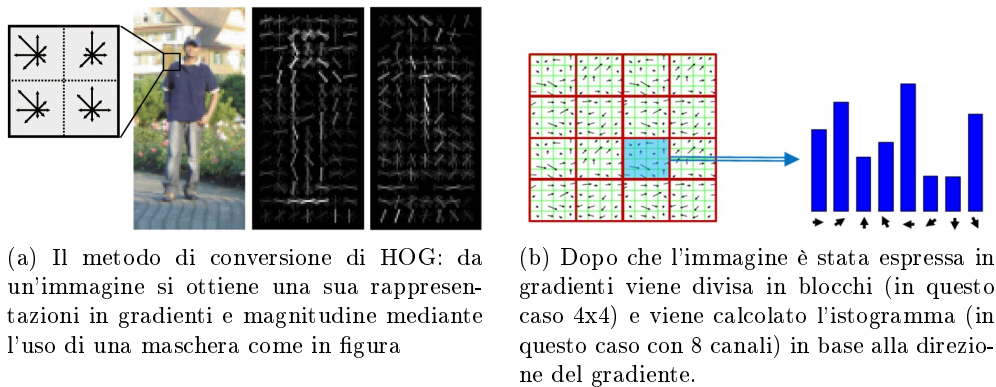
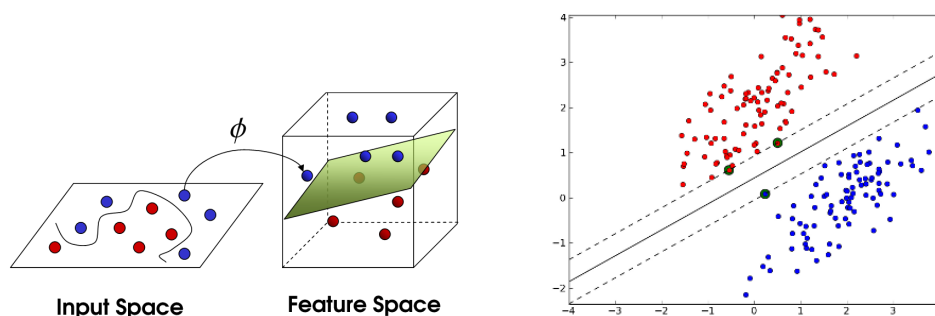


Figure 3.3: Esemplicazione visiva di HOG

riconoscimento in cui si cerca di capire se, il campione che deve essere analizzato, contenga o meno ciò che vogliamo cercare. La scelta di un insieme di dati da utilizzare durante la fase di addestramento per creare un training set, che verrà utilizzato successivamente per il riconoscimento, è un'operazione alquanto complessa perché se si hanno pochi dati o i dati stessi sono poco rilevanti, tutto il meccanismo fallisce il proprio compito. Un'immagine viene normalizzata utilizzando una correzione di gamma e convertita in scala di grigi (se già non lo è) per velocizzare tutto il processo. Vengono poi calcolati i valori dei gradienti applicando una maschera 1-D ad una o entrambe le direzioni (orizzontale e verticale) ad ogni pixel dell'immagine (Figura 3.3a). Gli istogrammi vengono costruiti dividendo l'immagine in celle rettangolari di  $N \times N$  pixels, tipicamente con  $N=8$  o  $16$ : ogni pixel all'interno della cella ha un voto pesato basato sul valore del gradiente calcolato precedentemente. Il peso del voto è dato dal valore della magnitudine del gradiente (modulo). L'istogramma viene costruito utilizzando 9 canali diversi (valore migliore per il riconoscimento umano ed anatomico) [26] dove si vanno ad accumulare le direzioni dei gradienti in base all'angolo compreso:  $[0^\circ-20^\circ]$  nel canale 1 fino a  $[160^\circ-180^\circ]$  (Figura 3.3b). E' quindi importante associare l'orientazione del gradiente in un dato intervallo ad un "voto" che equivale all'importanza del gradiente in un punto, ciò indica se ci sono o meno dei margini: il gradiente nell'intorno di un bordo è chiaramente molto più significativo di quello calcolato in una zona uniforme dell'immagine ed è un buon modo per estrarre informazioni. Gli istogrammi ottenuti vengono normalizzati mediante fattori di normalizzazione variabili che vengono calcolati per ogni blocco di celle  $2 \times 2$ : questo processo è necessario per renderlo invariante a diverse condizioni di luminosità, contrasto ed ombre. Può essere utile sovrapporre i blocchi tra loro ottenendo quindi dei dati ridondanti che aiutano a rendere l'algoritmo più preciso. Lungo l'immagine originale viene fatta scorrere una detection window che, per ogni porzione dell'immagine, fa quanto spiegato prima, riscaldando successivamente l'immagine finché non raggiunge le dimensioni della finestra stessa. Ogni finestra è ottenuta concatenando

gli istogrammi che la compongono creando un feature vector, che viene classificato tipicamente mediante un classificatore, ovvero Support Vector Machine.

### Fase di Riconoscimento: Support Vector Machine (SVM)



(a) Il tipico funzionamento di SVM: da un insieme di punti su un piano bidimensionale si passa ad un iperpiano ad una più alta dimensionalità (in questo caso tridimensionale) per migliorare e facilitare la separazione tra i punti

(b) In SVM l'etichettatura dei dati viene effettuata grazie alla retta rappresentata in figura: se un dato si trova al di sotto di essa sarà etichettato come blu altrimenti rosso

Figure 3.4: Esempio visivo di SVM

SVM è un modello di apprendimento supervisionato, ovvero una funzione che permette di dedurre un risultato a partire da un insieme di dati precedentemente etichettati ed usati per l'apprendimento, ed è utilizzato principalmente per la classificazione. Funziona generando un iperpiano di separazione in uno spazio ad alta dimensionalità in modo da aumentare e ottimizzare la distanza tra punti di due differenti classi di dati: ogni punto del feature vector ottenuto con HOG è visto e rappresentato come un punto in uno spazio di  $n$ -dimensioni dove  $n$  è la dimensione del vettore (Figura 3.4a). Nel caso in cui i diversi training set usati per indicare le classi non siano separabili linearmente mediante un iperpiano, allora i campioni vengono mappati in uno spazio a dimensione maggiore di quanto viene tipicamente usato. Per motivi di prestazioni, in SVM, i prodotti scalari vengono calcolati mediante l'uso di funzioni kernel, studiate appositamente e che si adattano al tipo di problema da risolvere [27]. SVM multi classe è un tipo di SVM dove si cerca di etichettare un insieme di oggetti, partendo da un insieme finito composto da molti elementi: solitamente il problema diventa di sola classificazione binaria ed è usato per distinguere se un oggetto appartiene o meno ad una categoria. Per questo tipo di approccio si adopera una strategia di voto: il massimo vince ed ogni classificatore associa l'oggetto ad una delle "due" classi, la classe più votata determina la classificazione (Figura 3.4b).

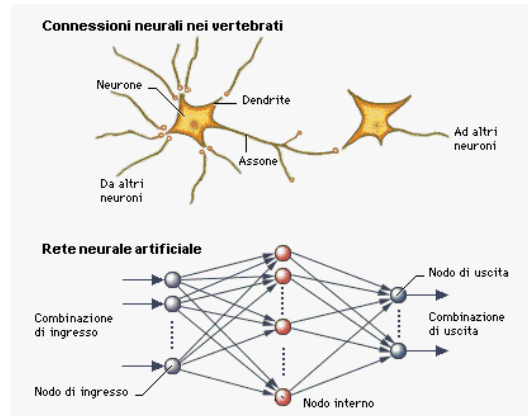


Figura 3.5: Un confronto visivo tra le reti naturali e quelle artificiali. Il loro comportamento è il medesimo: un numero di dati in ingresso si traduce in un numero di dati in uscita.

### 3.4.3 Reti Neurali

Le reti neurali sono dei sistemi di elaborazione dell'informazione. L'etimologia del nome deriva dall'unità elementare che caratterizza le strutture cerebrali, cioè il neurone: un'unità cellulare che è in grado di ricevere, integrare e trasmettere impulsi nervosi. Le reti neurali rappresentano quindi una tecnica di intelligenza artificiale, ovvero un metodo per simulare il comportamento dell'intelligenza umana. L'unità fondamentale della rete neurale è il neurone artificiale, ovvero un modello matematico che calcola una funzione che, come dati in ingresso, ha i risultati dell'elaborazione provenienti da altri neuroni e, come risultato, ha un serie di valori che verranno trasmessi, a loro volta, agli altri neuroni (esattamente come per i neuroni biologici) (Figura 3.5). Questa funzione, detta di attivazione, trasforma quindi l'insieme degli ingressi in corrispondenti valori nell'insieme delle uscite. Le reti neurali si ottengono dalla connessione di molti neuroni artificiali tra loro ed in questo modo si cerca di mimare i comportamenti del cervello: memorizzando dati, richiamando tali oggetti con una descrizione solo parziale (esattamente nello stesso modo in cui funziona la memoria umana), generalizzare o raggruppare oggetti partendo da alcuni oggetti d'esempio e trovare le somiglianze tra loro. Collegando i neuroni secondo schemi ordinati e non seguendo uno schema casuale, può aiutare a studiare le proprietà dei modelli computazionali che si vuole analizzare. Le reti neurali hanno un ampio numero di proprietà che le rende più attrattive rispetto all'uso di un approccio algoritmico. Usare degli algoritmi è migliore solo nel caso in cui ci sia abbastanza informazione sui dati e sulla teoria con la quale essi vengono generati, cioè se si può calcolare direttamente la soluzione partendo dallo spazio del problema. Invece, al contrario, quando non vi è abbastanza informazione per creare un modello affidabile, può essere utile creare un modello di rete neurale che imita il comportamento della deduzione umana, collezionando informazioni e testando le soluzioni in maniera più

diretta. Le reti neurali quindi, risultano utili in situazioni in cui c'è un'abbondanza di dati, ma poca teoria sulla quale affidarsi nell'ottenere risultati, come nel campo della comparazione tra immagini. I dati possono essere non lineari, non stazionari o caotici, quindi non facili da modellare. Le reti neurali non acquisiscono assunzioni a priori dallo spazio del problema, nemmeno a livello statistico. Il pattern recognition, cioè il riconoscimento di schemi ricorrenti nei dati (in questo caso nelle immagini) è l'area di più grande utilizzo delle reti neurali e consiste nella classificazione di oggetti, creando diverse classi di natura diversa in base alla loro somiglianza. Si parla di apprendimento supervisionato nel caso in cui i dati in ingresso vengono etichettati a priori, in una fase di training, per renderle parte di una stessa classe di oggetti poi successivamente confrontati con nuovi campioni che verranno etichettati in automatico. Per quanto riguarda il campo delle immagini, le tecniche di reti neurali più affidabili ed usate nel campo del confronto tra pixel o del confronto tra features, sono le Self-organasing feature maps e Hopfield [28].

### 3.4.3.1 Self-Organizing Map (SOM)

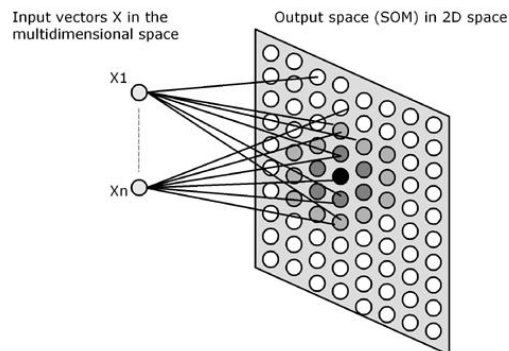


Figura 3.6: Struttura tipica di SOM con i due strati di input ed output: lo strato d'input è completamente connesso con quello di output mentre i neuroni nello strato di output sono connessi tra loro soltanto in zone adiacenti. Per esempio, il neurone nero è connesso con tutti i grigi che lo circondano.

La Self-Organizing Map (SOM) è una tipologia di reti neurali che viene allenata usando l'apprendimento non supervisionato (ovvero cercare di trovare una struttura ricorrente nascosta di dati che non vengono etichettati preventivamente) per produrre un sistema di riconoscimento affidabile. SOM è una tra le più importanti architetture di reti neurali: è stata sviluppata principalmente da Teuvo Kohonen tra il 1979 e il 1982, in parte sulla base di studi svolti precedentemente (Von der Malsburg 1973, Willshaw e Von der Malsburg 1976) [29]. Nelle SOM viene definito, come elemento caratteristico, uno strato, detto strato di Kohonen, costituito da neuroni artificiali disposti spazialmente in modo ordinato. Tale strato di processing elements (PE o neuroni artificiali) si evolve durante l'apprendimento, specializzando le posizioni dei

singoli PE come indicatori delle caratteristiche statistiche importanti degli stimoli di ingresso. Questo processo di organizzazione spaziale delle caratteristiche di dati di ingresso è chiamato anche “Feature Mapping”. In questo tipo di rete l’apprendimento non è supervisionato, quindi non bisogna conoscere a priori le classi in uscita (sarà la rete, dopo essere stata addestrata, a dirci a quale cluster appartiene l’input). Topologicamente questa rete si caratterizza per uno schema a due strati, uno di input ed uno di output (quest’ultimo comunemente chiamato strato di Kohonen). I neuroni dei due strati sono completamente connessi fra loro, mentre i neuroni dello strato di output sono connessi ciascuno con un vicinato di neuroni (guardare Output Space in figura 3.6), secondo un sistema di inibizione laterale che ha la forma di una Gaussiana (cappello messicano). Le interazioni laterali intorno al neurone vincitore sono funzioni della distanza: si eccitano i neuroni più vicini, si inibiscono quelli più lontani. Anche se esistono versioni più complesse, nella versione più semplice c’è un solo vincitore e l’apprendimento è di tipo competitivo. Il vincitore è il nodo che risulta avere il suo vettore dei pesi più simile al vettore in ingresso; questa modalità di scegliere il nodo vincitore è nota con il nome di winner takes all. Nella fase di addestramento, una volta scelto il vincitore, si regolano i pesi per avvicinarli ancora di più al vettore in ingresso. Le SOM a differenza delle altre reti adottano una funzione di contorno per preservare le proprietà topologiche dello spazio dei dati input. La SOM definisce un mapping dallo spazio dei dati di input  $x_1 \dots x_n$  ad un array monodimensionale o bidimensionale di nodi. Il mapping è eseguito in modo che la relazione topologica nello spazio n-dimensionale di ingresso si mantiene costante quando associata alla SOM. Inoltre, la densità locale dei dati in ingresso, si riflette anche sulla mappa: le zone dello spazio dei dati in input, che sono rappresentati da più informazioni, vengono mappati su una maggiore superficie della SOM. Ogni nodo sulla mappa è definito da un vettore di pesi  $w_{ij}$  che è modificato durante la fase di training, che è molto semplice e si riduce ai seguenti passi:

1. Si seleziona un oggetto dal training-set
2. Si trova il nodo che è più vicino ai dati, ad esempio usando la distanza tra il  $w_{ij}$  e il vettore in ingresso
3. Si regolano i vettori dei pesi del nodo vincitore e dei nodi che rientrano nel suo vicinato, in modo che il suo  $w_{ij}$  si muova verso il dato in ingresso
4. Si ripete dal passo 1 per un numero prefissato di volte. Ogni volta però si riduce sia il tasso di apprendimento sia la dimensione del vicinato.

### 3.4.3.2 Rete di Hopfield

La rete di Hopfield è un tipo di rete neurale ricorrente, cioè dove le connessioni tra i neuroni formano un ciclo direzionato (Figura 3.7). Viene garantita la convergenza

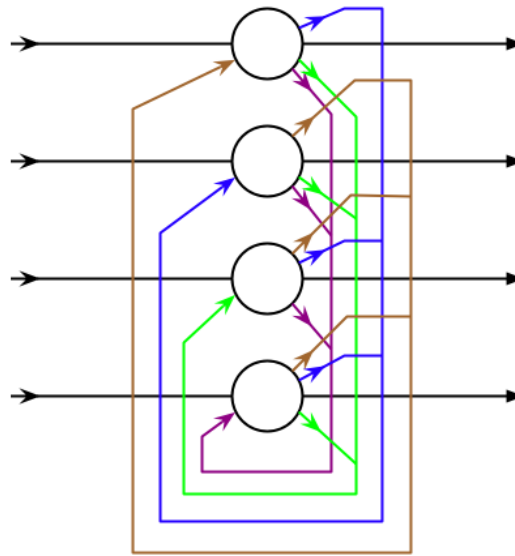


Figura 3.7: La tipica struttura di una rete di Hopfield con le frecce nere di input ed output e le frecce colorate che formano dei cicli direzionati all'interno della rete.

ad un minimo locale, ma non c'è garanzia della precisione del risultato. Le unità nelle reti di Hopfield hanno valori binari determinati dal fatto che l'input che ricevono ecceda o meno una determinata soglia, solitamente le unità hanno valore 0 o 1 o -1. La connessione tra due unità ha un costo, nel senso che una rete di Hopfield può essere descritta come un grafo pesato non direzionato. La funzione di aggiornamento può essere fatta in due diversi modi:

1. la modalità asincrona dove una unità viene aggiornata una alla volta: questa unità può essere scelta a caso, oppure in un'unità predefinita all'inizio
2. la modalità sincrona dove tutte le unità sono aggiornate in contemporanea: questo necessita di un orologio centrale per mantenere la sincronizzazione ed è il meno realistico.

Le reti di Hopfield hanno un valore scalare associato ad ogni stato della rete che è chiamato energia della rete. Questo valore è chiamato energia poiché per definizione, quando alcune unità sono scelte a caso per l'aggiornamento, l'energia rimarrà la stessa o diminuirà. Tipicamente i costi delle connessioni sono simmetrici, poiché garantiscono che la funzione dell'energia decresca in maniera monotona e poiché la rete potrebbe mostrare comportamenti periodici o caotici se non vengono usati costi non simmetrici. Tuttavia questo comportamento caotico è confinato a parti relativamente piccole dello spazio delle fasi e non compromette l'abilità della rete di funzionare come un sistema di memoria. L'inizializzazione di una rete di Hopfield è fatta inserendo i valori delle unità nello schema iniziale. Aggiornamenti continui sono fatti fintanto che la rete non converga in uno stato stabile, cioè uno schema che non può cambiare nessun valore al suo interno mediante l'aggiornamento. L'allenamento di una rete di



Hopfield necessita di abbassare l'energia degli stati che la rete dovrebbe "ricordare", questo permette alla rete di avere significato di un sistema in cui si può memorizzare del materiale, così la rete convergerà in uno stato "ricordato" [30]. Ci sono varie e differenti regole di apprendimento che possono essere usate per salvare informazioni in una rete di Hopfield, in generale devono avere le due seguenti proprietà:

1. Essere locale: una regola d'apprendimento è locale se ogni peso è aggiornato usando neuroni da entrambi i lati della connessione che è associata con quel particolare peso
2. Essere incrementale: nuovi schemi possono essere appresi senza usare l'informazione da vecchi schemi che sono stati precedentemente utilizzati per l'allenamento.

La capacità della rete di Hopfield è determinata dal numero di neuroni e dal numero di connessioni all'interno di una rete. E' quindi evidente che molti errori potrebbero verificarsi se si prova a salvare un gran numero di vettori.

### 3.5 Valutazione delle prestazioni di un sistema di classificazione: la matrice di confusione

Per riuscire a valutare in maniera precisa l'affidabilità di un sistema di etichettatura se ne osservano i risultati. Tali risultati possono essere corretti o errati ma, esprimendoli in maniera tabulare utilizzando la matrice di confusione, si può ottenere un'interpretazione più completa. Una matrice di confusione è una tabella che permette, appunto, di visualizzare le performance di un algoritmo in ambito di machine learning [31]. Ogni colonna della matrice rappresenta una classe di predicati che vengono predetti dal sistema mentre ogni riga rappresenta le istanze nella classe reale. In un'analisi predittiva, come questa, è necessario utilizzare anche la tabella di confusione: ovvero una tabella che mostra il numero di veri positivi, veri negativi, falsi positivi e falsi negativi (valori spiegati di seguito). Ciò permette un'analisi molto più dettagliata che tiene anche conto della distribuzione dei risultati non corretti. La matrice di confusione permette di estrarre alcuni parametri interessanti che verranno illustrati successivamente.

La costruzione della matrice di confusione (Tabella 3.1) permette di ottenere alcuni valori indispensabili per le analisi successive:

- veri positivi (true positives **TP**): è il numero di risultati correttamente classificati alla fine dell'elaborazione. Sono i valori che si trovano sulla diagonale della matrice. Per esempio per la classe A: TP<sub>A</sub>.
- falsi positivi (false positives **FP**): è il numero di risultati non appartenenti ad una data classe, ma che sono stati considerati tali. In questo problema

	PREDETTI						
		A	B	C	D	E	TP+FN
REALI	A	TPa	Eab	Eac	Ead	Eae	
	B	Eba	TPb	Ebc	Ebd	Ebe	
	C	Eca	Ecb	TPc	Ecd	Ece	
	D	Eda	Edb	Edc	TPd	Ede	
	E	Eea	Eeb	Eec	Eed	TPe	
	TP+FP						

Tabella 3.1: Matrice di confusione: leggendo una riga è possibile capire come si è distribuita la classificazione di una classe di oggetti. Una riga invece rappresenta il numero di oggetti di una classe in base a come li ha etichettati il sistema. Ogni casella della matrice rappresenta il numero di elementi reali classificati in un certo modo. Per esempio l'elemento Eab (Errore ab) ha il significato di « Numero di elementi reali della classe A che sono stati predetti come appartenenti alla classe B». Sulla diagonale ci sono il numero di true positives per una data classe: per esempio l'elemento 1,1 è in numero true positives di A (TPa). Con TP+FP s'intende la somma di true positives e false positives e con TP+FN la somma di true positives e false negatives (ciò è meglio spiegato qui di seguito).

multiclasse, il numero di FP per una data classe è dato dalla somma di una colonna della matrice a cui viene sottratto l'elemento TPx. Per la classe A:  $FPa = Eba + Eca + Eda + Eea$

- falsi negativi (false negatives **FN**): è il numero di elementi che appartengono ad una determinata classe, ma non sono stati considerati tali. In questo problema multiclasse, il numero di FN per una data classe è dato dalla somma di una riga della matrice a cui viene sottratto l'elemento TPx. Per la classe A:  $FNa = Eab + Eac + Ead + Eae$
- veri negativi (true negatives **TN**): è il numero di risultati correttamente predetti dal programma come non appartenenti ad una determinata classe. In questo problema multiclasse equivale alla somma di tutte le colonne e le righe della matrice che non appartengono alla classe che si sta considerando. Per la classe A:

$$TNa = TPb + Ebc + Ebd + Ebe + Ecb + TPc + Ecd + Ece + Edb + Edc + TPd + Ede + Eeb + Eec + Eed + TPe$$

Dopo aver effettuato il calcolo di questi valori li si può esprimere in maniera tabellare con la tabella di confusione (Tabella 3.2).

Dopo la costruzione di queste matrici, si calcolano i parametri più comuni per una situazione multi classe. Si nota che tutti questi parametri vanno calcolati sepa-

Distretto	TP	FP	FN	TN
A	TPa	FPa	FNa	TNa
B	TPb	FPb	FNb	TNb
C	TPc	FPc	FNc	TNc
D	TPd	FPd	FNd	TNd
E	TPe	FPe	FNe	TNe
<b>Totale</b>	<b>TPtot</b>	<b>FPtot</b>	<b>FNtot</b>	<b>TNtot</b>

Tabella 3.2: Nella tabella di confusione vengono espressi i valori di true positives (TP), false positives (FP), false negatives (FN) e true negatives (TN) per ogni classe del problema e per il sistema totale (TPtot, FPtot, FNtot e TNtot). Con TPa s'intende il numero di true positives per la classe A e, allo stesso modo, vanno interpretati il resto delle classi e dei valori in tabella. I valori TPtot, FPtot, FNtot e TNtot sono ottenuti sommando tutti i valori di una riga della matrice.

ratamente per ogni classe (a parte l'accuracy) e sono migliori tanto più si avvicinano al 100%.

**Accuracy:** è il grado di vicinanza che il risultato di un sistema ha con una quantità reale e misurabile, la differenza tra valor medio campionario e valor medio di riferimento. L'accuracy è quindi un valore medio e si calcola per valutare tutto il sistema. La sua formula è data da:

$$Accuracy = \frac{TPtot}{N}$$

dove N è il numero totale di elementi analizzati.

**Sensitivity:** è un parametro che serve per misurare la proporzione di oggetti realmente etichettati come tali. Per un problema multi classe va calcolata singolarmente per ogni classe che compone il sistema, la sua formula è:

$$Sensitivity = \frac{TPx}{TPx + FNx}$$

dove x è una classe generica di oggetti.

**Precision:** rappresenta il grado in cui i risultati ottenuti (in termini di numero di predizioni corrette) possono ripetersi in analisi successive su altri oggetti diversi da quelli presi in esame. Sottolinea quindi la ripetibilità e la riproducibilità di alcuni risultati simili. Anche se accuratezza e precisione sono spesso usati colloquialmente come sinonimi, non è questo il caso. Per un problema multi classe, ogni classe di elementi ha un proprio valore di precisione: calcolare il valore medio del sistema non rispecchia le sue reali caratteristiche. La formula:

$$Precision = \frac{TPx}{TPx + FPx}$$

dove  $x$  è una classe generica di oggetti.

**Specificity:** è un parametro che serve a misurare l'entità reale di risultati negativi identificati correttamente come tali; è praticamente il complementare della sensitivity. Si usa solitamente per problemi con 2 sole classi, ma può essere utilizzato anche in problemi multi classe [32]. La formula:

$$Specificity = \frac{TNx}{TNx + FPx}$$

dove  $x$  è una classe generica di oggetti.

## Capitolo 4

# Obiettivo della tesi

Il presente progetto di tesi è nato dalla collaborazione tra il Politecnico di Milano e l'Azienda Ospedaliera S. Anna di Como. L'obiettivo della tesi si incentra sulla ricerca di un metodo specifico di riconoscimento dei diversi distretti anatomici rappresentati nelle indagini di tomografia assiale computerizzata (TAC).

Si è lavorato per la creazione di un software flessibile, che possa essere utilizzato ed eseguito su vari sistemi operativi. Si è altresì valutata la possibilità di collocare il software direttamente sulla piattaforma PACS, in modo che le analisi venissero gestite in totale autonomia da eventuali sistemi gestionali presenti nell'Azienda Ospedaliera.

Tuttavia, al momento non si è riusciti a realizzare ciò, poiché il programma utilizzato per accedere al PACS non è di tipo open source. E' dunque necessario un supporto da parte del produttore, che al momento non ha dato disponibilità in tempi consoni al completamento del presente lavoro di tesi.

Il lavoro si è svolto implementando varie tecniche di confronto tra immagini con il fine ultimo di individuare un metodo affidabile per eseguire il riconoscimento automatico. Il presente capitolo analizza i requisiti che hanno portato alle scelte fatte e le caratteristiche del sistema software sviluppato.

### 4.1 Requisiti

In una prima fase della presente tesi, si era pensato di aggiungere nuove funzionalità al programma ARGO a disposizione dell'Azienda Ospedaliera, che viene utilizzato per studiare il livello di radiazione specifico, oggetto di una precedente tesi svolta da altri studenti del Politecnico di Milano. Pertanto si è avuto a disposizione il codice sorgente di detto software e circa 600 esami medici completi, estratti casualmente dal sistema PACS. Gli esami, 200 per ogni distretto anatomico principale (encefalo, torace ed addome) sono stati utilizzati per gli studi. In un successivo momento, a causa del riscontrato limite delle prestazioni di ARGO, si è passati a creare un programma ex novo, multi-piattaforma, intuitivo e leggero, specializzato nel ri-

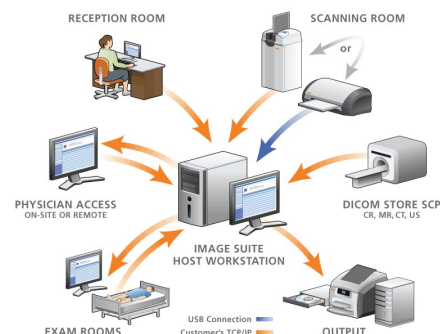
conoscimento automatico dei vari distretti anatomici. Il programma deve essere in grado di supportare il formato delle immagini che vengono fornite, cioè il DICOM, e deve essere capace di distinguere e processare immagini derivanti dalle seguenti macchine per la Tomografia Computerizzata dislocate nelle varie sedi dell'Azienda Ospedaliera comasca:

- PHILIPS Brilliance CT BR64 System - 64 slice (San Fermo)
- G.E. LightSpeed VCT - 64 slice (San Fermo)
- G.E. LightSpeed Ultra - 8 slice (Cantù)
- G.E. LightSpeed Ultra - 16 slice (Menaggio)

#### 4.1.1 PACS



(a) Un PACS molto simile a quello utilizzato presso l'Ospedale S. Anna di Como



(b) Un breve riassunto del workflow riguardante il PACS: è la macchina centrale che riceve e salva direttamente le informazioni su pazienti ed esami. Solo i dottori possono modificare le informazioni confidenziali in esso contenute. Il protocollo usato per la connessione è il TCP/IP anche via internet.

Figure 4.1: PACS

Il Picture Archiving and Communication System (Sistema di archiviazione e trasmissione di immagini) è un sistema hardware e software dedicato all'archiviazione, trasmissione, visualizzazione e stampa delle immagini diagnostiche digitali. Le immagini elettroniche e i rapporti-dose associate ad esse, sono trasmessi in digitale via PACS: questo elimina la necessità di recuperare i file manualmente usando supporti fisici. Il formato universale per la memorizzazione e il trasferimento delle immagini è DICOM. Un PACS consiste di quattro elementi principali:

1. Gestione modalità di imaging: radiografia (RX), tomografia computerizzata (CT) e risonanza magnetica (MRI)

2. Una rete protetta per la trasmissione di informazioni del paziente.
3. Stazioni di lavoro per l'interpretazione e la revisione immagini.
4. Archivi per la memorizzazione e il recupero di immagini e referti.

Il PACS ha quindi la capacità di fornire accesso tempestivo ed efficiente alle immagini e ai relativi dati. PACS rompe le barriere fisiche e di tempo associate alla pellicola tradizionale per quanto riguarda il recupero dell'immagine, la distribuzione e la visualizzazione. Dagli anni '80 ai PACS moderni ci sono delle infrastrutture molto solide e in grado di eseguire l'intera gamma di funzioni. L'imaging digitale e la comunicazione nell'ambito della medicina (standard DICOM 3.0) sono stati istituiti con la collaborazione della American College of Radiology e la National Equipment Manufacturing Association [33]. I requisiti per un buon PACS sono protocolli efficienti per la visualizzazione della scala di grigi, interfacce veloci e scalabilità.

## 4.2 Vincoli di progetto

Il numero di immagini rappresenta il maggiore vincolo di progetto per due motivi principali: in termini di tempistiche, in quanto il programma per l'estrazione degli esami dall'archivio PACS è molto lento, ed in termini di privacy dei pazienti (la cui identità deve rimanere anonima). Per l'estrazione delle immagini occorre collegarsi al PACS attraverso l'uso di credenziali in possesso del personale ospedaliero, filtrare il database per l'individuazione di uno specifico distretto anatomico, selezionare manualmente con un mouse tutte le immagini di un singolo esame radiologico, utilizzare un comando specifico e scaricarle in locale sul PC d'uso e infine utilizzare un ulteriore programma per la decompressione dei file scaricati. La procedura richiede tempistiche che vanno dai cinque ai dieci minuti per ciascun esame, quindi ha richiesto molte ore per l'estrazione dei suddetti 600 campioni di immagine. Altra complicanza riscontrata riguarda l'obbligatorietà del totale rispetto della privacy del paziente. Come noto le immagini DICOM riportano tra l'altro anche i dati anagrafici del paziente, dati che non possono essere in alcun modo diffusi e perciò oltre ad estrarli nel modo prima descritto, si è dovuto procedere attraverso l'applicazione di uno specifico programma a renderli anonimi, cancellando nome, cognome e data di nascita del paziente. Un'altra problematica deriva dal fatto che le immagini radiologiche non hanno dimensioni standard, ma molti formati diversi. Perciò è stato necessario effettuare un resize sulle immagini che vengono, quindi, deformate da tale operazione con conseguente aggiunta di rumore e casistiche da gestire.

## 4.3 Strumenti

Inizialmente si è pensato di operare attraverso il programma ARGO: il software in uso presso il Sant'Anna di Como. ARGO utilizza il linguaggio Java in combinazione con librerie per la gestione specifica delle immagini, la principale delle quali è ImageJ. Per velocizzare i test, si è dovuto cambiare piattaforma di sviluppo e linguaggio di programmazione, passando al più adatto Matlab. Questa piattaforma non necessita di ulteriori librerie per la gestione delle immagini, ma supporta direttamente il formato DICOM, oltre a racchiudere già in sé molte funzioni adatte al confronto di immagini.

### 4.3.1 Linguaggi di programmazione

#### 4.3.1.1 Java ed ImageJ

Java è uno dei linguaggi più diffusi in tutti gli ambiti dell'informatica: dallo sviluppo mobile (Android) alle applicazioni Enterprise, dal desktop al Web, applicando l'approccio di sviluppo e progettazione orientato agli oggetti che ha segnato gli standard industriali degli ultimi venti anni. Nell'approccio OO in Java non solo si definiscono gli oggetti su cui si intende lavorare, ma li si organizza in categorie: una classe è esattamente la definizione delle proprietà e dei metodi che avrà ogni elemento della categoria. La distinzione tra oggetto e classe è enfatizzata dalle seguenti definizioni: classe è una collezione di uno o più oggetti contenenti un insieme uniforme di attributi e servizi, insieme ad una descrizione su come creare nuovi elementi della classe stessa; un oggetto è dotato di stato, comportamento ed identità; la struttura ed il comportamento di oggetti simili sono definiti nelle loro classi comuni; i termini istanza ed oggetto sono intercambiabili. Java è un marchio registrato di Oracle.

Tra le librerie di Java, ImageJ è quella che è stata scelta per la gestione immagini in ARGO. ImageJ può visualizzare, editare, analizzare, processare, salvare e stampare immagini a livelli di grigio (8 bit, 16 bit e 32 bit) ed a colori (8 bit e 24 bit). Può leggere molti formati di immagine inclusi TIFF, PNG, GIF, JPEG, BMP, DICOM, FITS, così come anche alcuni formati "raw". ImageJ supporta gli stack di immagini, serie di immagini sovrapposte (sia spazialmente come "sezioni" di un corpo tridimensionale, sia temporalmente come una sequenza di immagini bi o tri-dimensionali, comunque denominate "slices") che condividono una singola finestra ed è abilitato al multithreading, in modo che alcune operazioni che richiedono molto tempo di calcolo possano essere eseguite in parallelo su hardware multi-CPU. Il software ImageJ è un programma di calcolo di analisi delle immagini mediche originariamente sviluppato dal NIH (National Institute of Health, ente statunitense) ed è in grado di leggere, elaborare e analizzare le immagini in formato DICOM. Questo software è di dominio pubblico e può essere scaricato attraverso il sito Internet all'in-



dirizzo “<http://rsbweb.nih.gov/ij/>”. ImageJ è stato progettato con un’architettura aperta che fornisce una flessibile estensibilità tramite plugin Java. L’acquisizione, l’analisi e l’elaborazione di plugin personalizzati possono essere sviluppati addirittura utilizzando l’editor incorporato di ImageJ e un compilatore Java. Un plugin sviluppato personalmente permette di risolvere molti problemi di elaborazione delle immagini e di analisi, dell’imaging tridimensionale e l’elaborazione di immagini radiologiche. ImageJ può calcolare area e statistiche di valore in pixel di selezioni definite dall’utente, può misurare distanze e angoli e può creare istogrammi di densità e profilo di linea. Supporta le funzioni di elaborazione delle immagini standard come operazioni logiche e aritmetiche tra le immagini, manipolazione, convoluzione, analisi di Fourier, affilatura, levigatura, rilevamento dei bordi e la mediana di filtraggio, trasformazioni geometriche come il ridimensionamento e la rotazione. Il programma può supportare qualsiasi numero di immagini simultaneamente, limitato solamente dalla memoria disponibile. In ARGO ImageJ è stato utilizzato per supportare lo sviluppo di un software per l’analisi di immagini CT, ovvero ha permesso lo sviluppo di un plug-in eseguibile dal programma stesso e, al contrario, di sviluppare un’applicazione che usa le funzioni di ImageJ attraverso le API.

#### 4.3.1.2 MATLAB

MATLAB (MATrix LABoratory) è un linguaggio di programmazione ad alto livello che consente di costruire e gestire facilmente matrici e, come casi particolari, vettori e scalari. La struttura dati di base è la matrice: ciò significa che durante l’elaborazione ogni quantità viene trattata dall’ambiente di calcolo come una matrice di dimensione  $N \times M$  (un vettore è una matrice  $1 \times N$ , mentre uno scalare è gestito come una matrice  $1 \times 1$ ). MATLAB fornisce un ambiente di calcolo, visualizzazione e programmazione scientifica, in cui è possibile:

1. calcolare direttamente espressioni matematiche
2. utilizzare il semplice ambiente di programmazione per costruire i propri algoritmi
3. sfruttare algoritmi di base già implementati (built-in functions)

Altre funzionalità di MATLAB sono:

- Matematica e calcolo
- Sviluppo di procedure e applicazioni
- Modellistica, simulazione e prototipizzazione
- Analisi di dati, esplorazione e visualizzazione

- Disegno industriale e scientifico
- Costruzione di interfacce utente
- Toolbox vari

E' un linguaggio interpretato e non compilato: questo significa che le istruzioni vengono tradotte in linguaggio macchina (il linguaggio "capito" dal processore) e subito eseguite una per volta. Utilizzando C, Fortran, C++, la traduzione da linguaggio ad alto livello a linguaggio macchina avviene invece nel processo di compilazione, in cui tutto il programma viene tradotto in linguaggio macchina e poi eseguito.

### 4.3.2 Ambienti di sviluppo

#### 4.3.2.1 Ambiente Java

NetBeans è un ambiente di sviluppo, multi linguaggio scritto interamente in Java. E' un progetto open source molto utilizzato a livello mondiale. Sun Microsystems ha fondato il NetBeans open source project nel Giugno del 2000 e continua ad essere lo sponsor principale del progetto. Quello utilizzato per lo svolgimento della presente tesi è la versione NetBeans IDE. NetBeans IDE è uno strumento destinato ai programmatori per scrivere, compilare ed eseguire il debug ed il deploy di programmi. E' scritto in Java, ma può supportare qualsiasi linguaggio di programmazione. Esiste anche un gran numero di moduli utili per estendere le funzionalità di NetBeans IDE. NetBeans IDE è un prodotto gratuito senza alcuna restrizione riguardante il suo uso. Possiede numerosi plug-in che lo rendono appetibile al pubblico e richiede 512 Megabyte di Ram a causa dell'uso delle librerie grafiche standard di Java (Swing). Per un progetto di questa portata, che comprende gestione grafica e gestione delle normali attività della programmazione ad oggetti (gestione errori ed eccezioni), programmare senza un IDE è impensabile.

#### 4.3.2.2 Ambiente MATLAB

L'ambiente di sviluppo MATLAB è scritto in C. Oltre al linguaggio di programmazione omonimo di cui abbiamo precedentemente parlato, supporta anche la possibilità di interfacciarsi a programmi scritti in C, C++, Java, Fortran e Python. Anche se MATLAB è principalmente pensato per il calcolo matematico, molte interessanti funzionalità sulla gestione delle immagini sono già integrate e non c'è bisogno di installare delle librerie aggiuntive.

### 4.3.3 Archiviazione e condivisione del codice sorgente

Si è scelto di utilizzare GitHub che è un servizio di hosting online basato su Git: un sistema per la revisione, lo sviluppo ed il controllo del software che pone parti-

colare enfasi sull'integrità dei dati e sulla velocità di esecuzione del codice. GitHub fornisce un'interfaccia grafica basata sul web e molte funzionalità come accesso a documentazione, gestione dei privilegi e individuazione dei bug. Si può avere accesso ai progetti su GitHub per manipolarli usando git da linea di comando oppure con i comandi standard. Permette inoltre ad utenti sia registrati che non, di visitare le cartelle pubbliche sul sito. Il sito <http://github.com/> fornisce funzionalità in stile social network con aggiornamenti, elenco dei follower di un progetto ed un grafico che spiega come gli sviluppatori lavorano sulle versioni del codice. Un utente deve necessariamente creare un account per poter aggiornare i contenuti sul sito, ma le cartelle pubbliche possono essere viste e scaricate da tutti. Oltre al codice sorgente supporta i seguenti formati e funzionalità:

- Creazione della documentazione con README
- Individuazione delle problematiche
- Creazione di una Wiki dedicata ad un progetto
- Hosting di piccoli siti internet
- Render di file 3D
- Preview di immagini Photoshop o di altri formati direttamente online

## Capitolo 5

# Approccio utilizzato

Il presente capitolo illustra i metodi utilizzati per l'individuazione del distretto anatomico coinvolto in esami TAC. Questo problema è poco considerato e documentato: in ambito ospedaliero si considera solitamente veritiero quanto viene riportato dai tecnici negli header DICOM al momento dell'esame. L' Ospedale S. Anna ha però confermato che gli errori di etichettatura sono abbastanza frequenti. Un altro motivo per cui è necessario utilizzare un metodo che automatizzi il riconoscimento del distretto anatomico è per il risparmio di tempo di cui i medici gioverebbero nell'utilizzo di software come ARGO. Tali programmi sono indispensabili per il calcolo della dose di radiazioni, ma richiedono una selezione manuale del distretto anatomico. In letteratura non è stato possibile rilevare pubblicazioni significative sull'argomento: gli algoritmi in ambito biomedico sono esclusivamente utilizzati per risolvere problemi di registratura (ovvero le problematiche legate al movimento del paziente durante un esame) oppure per aiutare nell'individuazione di macropatologie come infezioni in atto o tumori <sup>1</sup>. I metodi che si sono considerati, illustrati precedentemente nel Capitolo 3, sono tecniche generali e non proprie dell'ambito biomedico e sono stati scelti accuratamente perché i più promettenti in questa classe di problema. A tali metodologie, se ne aggiungono altre di nuova concezione e frutto dell'attività della presente tesi. La validazione dei risultati per la totalità di questi metodi è stata effettuata utilizzando un insieme di immagini campione prelevate dagli archivi PACS dell'Ospedale S. Anna.

---

<sup>1</sup>Tuttavia, l' Ing. Bai dell'Ordine degli Ingegneri di Milano lavora con un gruppo di ricerca per trovare una soluzione a tale problema, estendendolo però a tutte le tipologie di esami radiologici, da ormai 5 anni. Si è provato ad entrare in contatto con tale team di sviluppo per chiedere maggiori delucidazioni sui risultati da loro conseguiti, ma non si sono riuscite ad ottenere informazioni al riguardo.

## 5.1 Insieme di immagini test

A causa delle lunghe tempistiche per l'ottenimento delle immagini dovute ai limiti tecnici del PACS Siemens, ma anche alla necessità di cancellare alcuni header DICOM per renderle anonime, si è iniziato con un insieme significativo ma piuttosto ristretto di esami sui quali lavorare. E' stata intuita fin da subito la potenzialità di lavorare esclusivamente sulle immagini scout che mostrano il distretto anatomico nella sua interezza. Ogni esame possiede 2 immagini scout: in una è rappresentata la vista laterale del corpo mentre nell'altra la vista frontale (scout antero-posteriore). Dal momento che le 2 tipologie di vista presentano problematiche diverse, è necessario analizzare e distinguere le scout laterali e frontali separatamente. I distretti anatomici che l'Ospedale vuole che il sistema sia in grado di distinguere sono 3: encefalo, torace ed addome. Gli arti vengono esclusi da tale analisi perché, per gli esami TAC, non esistono direttive ufficiali sulla quantità consigliate di dose in tali parti del corpo. Questo aspetto può essere desunto dalla Tabella 2.2. Un'altra ragione è il numero annuo trascurabile di esami di questa tipologia effettuati dal S. Anna. Il problema viene quindi illustrato come multiclasse con 6 classi distinte: encefalo antero-posteriore, encefalo laterale, torace antero-posteriore, torace laterale, addome antero-posteriore ed addome laterale. Per iniziare la valutazione di un metodo si è deciso di utilizzare 25 immagini scout per ogni classe del problema per un totale di 180 immagini: tale numero, che corrisponde alla quantità di esami svolti dall'azienda ospedaliera comasca in 3 giorni, è stato considerato un numero sufficientemente rappresentativo e rilevante per effettuare valutazioni preliminari significative. Queste immagini sono state preventivamente scelte dallo staff dell'ospedale per le loro caratteristiche: sono stati scelti campioni di una stessa classe tutti visivamente simili per facilitare, almeno in linea di principio, la cernita dei metodi.

## 5.2 Valutazione preliminare delle prestazioni

Per poter testare i metodi del Capitolo 3 è necessario preparare preventivamente un dataset, ovvero un insieme di dati strutturati noti (immagini) che possano venire utilizzati per altre elaborazioni. In questo caso si può parlare di training-set: 5 immagini per ognuna delle 6 classi designate (per un totale di 30) vengono preventivamente etichettate correttamente da un operatore umano per poter effettuare confronti con le rimanenti che verranno analizzate in modo automatico. Il numero di immagini che si proverà a classificare utilizzando gli algoritmi è quindi 120 (ovvero 20 immagini per ogni classe). La valutazione avviene calcolando il valore di sensitivity per ogni classe, ovvero il rapporto tra il numero di veri positivi per una data classe e numero totale di immagini per quella classe (20). Tale indice è abbastanza indicativo della bontà di un metodo e permette di capire se è necessario insistere o

meno nel suo utilizzo nel caso ci si avvicini o si superi il valore di 0.8 Si vedono ora nel dettaglio i metodi utilizzati.

### 5.3 Metodi della letteratura

I metodi presenti in letteratura e già descritti nel Capitolo 3 utilizzati per queste indagini sono stati, in ordine cronologico:

1. Distanza Euclidea
2. IMage Euclidean Distance
3. Color Histogram
4. Structural Similarity
5. Speeded Up Robust Features
6. Histogram of Oriented Gradients
7. Self Organizing Map
8. Rete di Hopfield

Di seguito si riportano i risultati relativi al parametro di sensitivity (Paragrafo 3.5) ad ogni metodo.

#### 5.3.1 Risultati del metodo “Distanza Euclidea”

Presupposto per poter utilizzare il metodo della distanza euclidea e' che le immagini siano della stessa dimensione. Poiche' invece le immagini scout hanno dimensioni diverse , è necessario effettuare un ridimensionamento/sottocampionamento ( resizing/downsampling) delle stesse immagini. Questo metodo calcola la somma dei valori assoluti delle differenze di colore tra pixel omologhi tra due immagini diverse:il metodo individuerà il distretto anatomico appartenente alla immagine dataset per il quale e' stato ottenuto il valore minimo di tale somma. Questi sono i risultati ottenuti sulle 120 immagini:

- Su 20 encefali antero-posteriori 15 sono stati individuati: sensitivity al 75%
- Su 20 encefali laterali 16 sono stati individuati: sensitivity all' 80%
- Su 20 toraci antero-posteriori 10 sono stati individuati: sensitivity al 50%
- Su 20 toraci laterali 7 sono stati individuati: sensitivity al 35%
- Su 20 addomi antero-posteriori 9 sono stati individuati: sensitivity al 45%
- Su 20 addomi laterali 7 sono stati individuati: sensitivity al 35%

### 5.3.2 Risultati del metodo “IMED”

Presupposto per poter utilizzare il metodo IMED e' che le immagini siano della stessa dimensione. Poiche' invece le immagini scout hanno dimensioni diverse, è necessario effettuare un ridimensionamento/sottocampionamento ( *resizing/downsampling* ) delle stesse immagini. Il funzionamento considera oltre alla differenza di colore tra un pixel e il suo omologo nella immagine di confronto, anche la differenza di colore con tutti gli altri, pesandole in maniera inversamente proporzionale alla loro distanza dal pixel di partenza. L' algoritmo viene ripetuto per tutti i pixel e i risultati vengono sommati. Il metodo individua' il distretto anatomico appartenente alla immagine dataset per il quale e' stato ottenuto il valore minimo di tale somma. Questi sono i risultati ottenuti sulle 120 immagini:

- Su 20 encefali antero-posteriori 16 sono stati individuati: sensitivity al 80%
- Su 20 encefali laterali 17 sono stati individuati: sensitivity all' 85%
- Su 20 toraci antero-posteriori 10 sono stati individuati: sensitivity al 50%
- Su 20 toraci laterali 6 sono stati individuati: sensitivity al 30%
- Su 20 addomi antero-posteriori 9 sono stati individuati: sensitivity al 45%
- Su 20 addomi laterali 8 sono stati individuati: sensitivity al 40%

I risultati non sono molto diversi da quelli ottenuti con la distanza euclidea tradizionale che però, a livello di prestazioni è decisamente più performante di IMED che impiega un tempo di elaborazione considerevolmente superiore per arrivare al risultato.

### 5.3.3 Risultati del metodo “Color Histogram”

Consiste nell'analizzare per ciascuna gradazione di colore presente in una immagine il numero di pixel di tale colore ed esprimerne il valore in un istogramma, indipendentemente dalla posizione del pixel nella immagine. Il metodo calcola la distanza di Bhattacharyya tra gli istogrammi della immagine da analizzare e del dataset. Il metodo individua' il distretto anatomico appartenente alla immagine dataset per il quale e' stato ottenuto il valore minimo di tale distanza. Questi sono i risultati ottenuti sulle 120 immagini:

- Su 20 encefali antero-posteriori 11 sono stati individuati: sensitivity al 55%
- Su 20 encefali laterali 11 sono stati individuati: sensitivity all' 55%
- Su 20 toraci antero-posteriori 9 sono stati individuati: sensitivity al 45%
- Su 20 toraci laterali 9 sono stati individuati: sensitivity al 45%

- Su 20 addomi antero-posteriori 9 sono stati individuati: sensitivity al 45%
- Su 20 addomi laterali 8 sono stati individuati: sensitivity al 40%

#### 5.3.4 Risultati del metodo “SSIM”

Presupposto per poter utilizzare il metodo SSIM e' che le immagini siano della stessa dimensione. Poiche' invece le immagini scout hanno dimensioni diverse , è necessario effettuare un ridimensionamento/sottocampionamento ( resizing/downsampling) delle stesse immagini. SSIM è una misura che tiene conto dell'interdipendenza spaziale tra pixel facendo un calcolo sulla luminanza delle immagini. L'immagine nel dataset che avrà il valore di SSIM più vicino ad 1 sarà quella che darà l'indicazione sul distretto anatomico. Questi sono i risultati ottenuti sulle 120 immagini:

- Su 20 encefali antero-posteriori 15 sono stati individuati: sensitivity al 75%
- Su 20 encefali laterali 17 sono stati individuati: sensitivity all' 85%
- Su 20 toraci antero-posteriori 11 sono stati individuati: sensitivity al 55%
- Su 20 toraci laterali 7 sono stati individuati: sensitivity al 35%
- Su 20 addomi antero-posteriori 12 sono stati individuati: sensitivity al 60%
- Su 20 addomi laterali 7 sono stati individuati: sensitivity al 35%

#### 5.3.5 Risultati del metodo “SURF”

Con metodi di estrazione di caratteristiche si confrontano i features vectors e non direttamente le immagini : per questo il dataset che viene utilizzato è formato dai features vectors estratti dalle immagini dataset utilizzando il metodo SURF. Il confronto tra features vectors del dataset e il feature vector che si vuole testare avviene mediante SVM, che ,tramite un processo di etichettatura binario decide se un dato appartiene o meno ad un certo insieme. Questi sono i risultati ottenuti sulle 120 immagini:

- Su 20 encefali antero-posteriori 11 sono stati individuati: sensitivity al 55%
- Su 20 encefali laterali 12 sono stati individuati: sensitivity all' 60%
- Su 20 toraci antero-posteriori 12 sono stati individuati: sensitivity al 60%
- Su 20 toraci laterali 13 sono stati individuati: sensitivity al 65%
- Su 20 addomi antero-posteriori 12 sono stati individuati: sensitivity al 60%
- Su 20 addomi laterali 13 sono stati individuati: sensitivity al 65%



### 5.3.6 Risultati del metodo “HOG”

Presupposto per poter utilizzare il metodo HOG e' che le immagini siano della stessa dimensione. Poiche' invece le immagini scout hanno dimensioni diverse , è necessario effettuare un ridimensionamento/sottocampionamento ( *resizing/downsampling*) delle stesse immagini. Con metodi di estrazione di caratteristiche si confrontano i features vectors e non direttamente le immagini : per questo il dataset che viene utilizzato è formato dai features vectors estratti dalle immagini dataset con il metodo HOG. HOG estrae features analizzando piccole regioni dell'immagine presa in esame individuando il loro orientamento e la magnitudo. Il confronto tra features vectors del dataset e il feature vector che si vuole testare avviene mediante SVM, che, tramite un processo di etichettatura binario decide se un dato appartiene o meno ad un certo insieme. Questi sono i risultati ottenuti sulle 120 immagini:

- Su 20 encefali antero-posteriori 19 sono stati individuati: sensitivity al 95%
- Su 20 encefali laterali 18 sono stati individuati: sensitivity all' 90%
- Su 20 toraci antero-posteriori 13 sono stati individuati: sensitivity al 65%
- Su 20 toraci laterali 11 sono stati individuati: sensitivity al 55%
- Su 20 addomi antero-posteriori 15 sono stati individuati: sensitivity al 75%
- Su 20 addomi laterali 11 sono stati individuati: sensitivity al 55%

### 5.3.7 Risultati del metodo “SOM”

SOM, come tutte le tecniche di reti neurali, ha bisogno di una fase di apprendimento in cui viene costruita una mappa partendo dalle immagini in ingresso . Se i dati in ingresso sono pochi vengono somministrati più volte. Dopo la fase di apprendimento c'è il mapping in cui un nuova immagine viene classificata e categorizzata. Il vincitore di tale processo è dato dal calcolo della distanza euclidea tra i dati in ingresso ed il vettore dei pesi nella rete. Questi sono i risultati ottenuti sulle 120 immagini:

- Su 20 encefali antero-posteriori 17 sono stati individuati: sensitivity al 85%
- Su 20 encefali laterali 17 sono stati individuati: sensitivity all' 85%
- Su 20 toraci antero-posteriori 8 sono stati individuati: sensitivity al 40%
- Su 20 toraci laterali 7 sono stati individuati: sensitivity al 35%
- Su 20 addomi antero-posteriori 8 sono stati individuati: sensitivity al 40%
- Su 20 addomi laterali 8 sono stati individuati: sensitivity al 40%

### 5.3.8 Risultati del metodo “rete di Hopfield”

La rete di Hopfield, come tutte le tecniche di reti neurali ha bisogno di una fase di apprendimento in cui viene costruita una mappa partendo dalle immagini in ingresso. Se i dati in ingresso sono pochi vengono somministrati più volte. Dopo la fase di apprendimento c'è il mapping in cui una nuova immagine viene classificata e categorizzata. Le reti di Hopfield sono reti neurali interamente connesse: per riconoscere un insieme di input bisogna associarlo ad un punto fisso stabile nella funzione di energia che viene modellata grazie al training set. Questi sono i risultati ottenuti sulle 120 immagini:

- Su 20 encefali antero-posteriori 16 sono stati individuati: sensitivity al 80%
- Su 20 encefali laterali 17 sono stati individuati: sensitivity all' 85%
- Su 20 toraci antero-posteriori 9 sono stati individuati: sensitivity al 45%
- Su 20 toraci laterali 10 sono stati individuati: sensitivity al 50%
- Su 20 addomi antero-posteriori 12 sono stati individuati: sensitivity al 60%
- Su 20 addomi laterali 9 sono stati individuati: sensitivity al 45%

## 5.4 Valutazione complessiva dei metodi della letteratura

Metodo	Encefali AP	Encefali L	Toraci AP	Toraci L	Addomi AP	Addomi L
<b>Dist. Eucl.</b>	75%	80%	50%	35%	45%	35%
<b>IMED</b>	80%	85%	50%	30%	45%	40%
<b>C.Hist.</b>	55%	55%	45%	45%	45%	40%
<b>SSIM</b>	75%	85%	55%	35%	60%	35%
<b>SURF</b>	55%	60%	60%	65%	60%	65%
<b>HOG</b>	95%	90%	65%	55%	75%	55%
<b>SOM</b>	85%	85%	40%	35%	40%	40%
<b>Hopf.</b>	80%	85%	45%	50%	60%	45%

Tabella 5.1: I risultati di sensitivity per tutte le classi del problema e per tutti i metodi espressi in maniera tabellare: con AP s'intende antero-posteriore e con L laterale.

La valutazione di tali risultati è basata sul presupposto che i valori, per essere considerati buoni ed accettabili, devono superare la soglia dell'80% di sensitivity per tutte le classi del problema. Risultati inferiori o prossimi al 50% sono da considerarsi sostanzialmente casuali. Da tale analisi incrociata dei metodi si nota come solo l'encefalo, sia antero-posteriore che laterale, venga riconosciuto in maniera efficace: quasi tutti i metodi (tranne Color Histogram e SURF) permettono ottime percentuali di riconoscimento di questo distretto anatomico e con HOG si raggiunge addirittura

il 95%. Quello in cui tutti i sistemi hanno difficoltà è nel riuscire ad etichettare in maniera corretta toraci ed addomi: anche in questo caso i risultati migliori si hanno utilizzando HOG (orientamento antero-posteriore) e SURF (orientamento laterale). Per queste parti del corpo si può osservare che la casualità dei rilevamenti è sostanzialmente indipendente dall'orientamento dell'immagine analizzata. Tra questi metodi non è stato quindi possibile identificare un candidato alla risoluzione del problema. Per ovviare a questo inconveniente, su indicazione del personale ospedaliero, l'esperimento è proseguito cercando d'individuare e studiare nuovi metodi inediti.

## 5.5 Definizione di nuovi metodi

In questa fase del lavoro si è riusciti ad entrare in possesso di una quantità più sostanziosa di immagini scout, condizione indispensabile secondo una prassi corrente per poter validare una metodologia non precedentemente documentata. Si è quindi raggiunto un totale di 1200 immagini (200 per ogni classe del problema). La ricerca è cominciata mediante una fase di brainstorming durante la quale si sono osservate lungamente le caratteristiche delle immagini per riuscire a capire come procedere nel lavoro. Si sono sviluppati 3 diversi metodi :

1. Dimensioni medie del corpo umano
2. Distribuzione spaziale dei livelli di grigio nell'immagine
3. Valore della simmetria in regioni d'interesse con HOG

### 5.5.1 Dimensioni medie del corpo umano

Tra i valori degli header DICOM è presente un fattore di conversione che permette di trasformare le dimensioni di un'immagine in pixel nelle reali dimensioni in millimetri. Le immagini scout raffigurano soggetti grigi su uno sfondo nero: dall'immagine è quindi facile ottenere le dimensioni in pixel (e quindi in millimetri) della sola figura ritratta. Per comodità e per evitare qualunque possibile falso rilevamento, si è preferito ritagliare lo sfondo. Per questo metodo si è costruita una tabella, utilizzando informazioni ricavate da un libro d'arte [34], contenente informazioni sulla lunghezza media dei distretti anatomici al variare di sesso ed età per poter confrontare questi valori con la lunghezza di quanto raffigurato nelle indagini radiologiche. Le immagini DICOM infatti contengono anche informazioni su età e sesso del paziente. Il metodo si basa quindi sul confronto tra i valori in tabella ( Figura 5.1) con le dimensioni dell'immagine ritagliata: il risultato è dato dal distretto anatomico che più si avvicina al valore dell'immagine presa in esame, sempre tenendo in considerazione età e sesso.

età	head+neck				Thorax				Trunk			
	M min	M max	F min	F max	M min	M max	F min	F max	M min	M max	F min	F max
1	17,54	21,65	16,87	21,50	10,52	12,99	10,12	12,90	26,31	32,48	25,30	32,25
2	18,90	24,33	18,01	24,43	11,46	14,75	10,92	14,81	28,65	36,88	27,30	37,03
3	19,70	25,87	19,42	25,58	12,08	15,86	11,91	15,69	30,20	39,65	29,76	39,21
4	20,29	26,69	19,86	26,71	12,58	16,55	12,32	16,56	31,46	41,37	30,79	41,40
5	21,06	27,22	20,02	27,95	13,21	17,08	12,56	17,53	33,02	42,69	31,40	43,83
6	21,15	28,40	20,54	28,63	13,43	18,02	13,03	18,17	33,56	45,06	32,58	45,43
7	22,15	28,35	21,76	28,37	14,22	18,20	13,97	18,22	35,56	45,51	34,93	45,54
8	22,81	28,44	22,21	28,64	14,82	18,48	14,43	18,61	37,05	46,19	36,07	46,52
9	23,09	28,97	22,91	28,80	15,17	19,04	15,05	18,92	37,93	47,59	37,64	47,30
10	23,61	29,38	23,48	29,24	15,69	19,52	15,60	19,43	39,22	48,79	39,00	48,57
11	24,33	29,62	24,35	29,64	16,33	19,89	16,35	19,90	40,84	49,71	40,87	49,75
12	23,96	31,05	24,06	31,52	16,25	21,05	16,31	21,37	40,62	52,63	40,77	53,42
13	25,23	31,52	24,96	32,35	17,26	21,56	17,07	22,13	43,16	53,91	42,68	55,33
14	26,00	33,00	25,40	33,12	17,93	22,76	17,52	22,85	44,84	56,90	43,79	57,12
15	27,19	34,19	25,81	33,54	18,89	23,75	17,93	23,30	47,23	59,37	44,83	58,25
16	28,18	35,20	25,92	34,05	19,69	24,60	18,12	23,80	49,23	61,50	45,30	59,50
17	29,14	35,85	27,02	33,73	20,47	25,18	18,98	23,69	51,18	62,96	47,45	59,23
18	30,48	35,76	27,23	34,40	21,49	25,21	19,20	24,26	53,72	63,04	48,00	60,64

Figura 5.1: In tabella si può notare le dimensioni minime e massime dei distretti anatomici in cm al variare di sesso ed età. Con i 18 anni si considera fermato lo sviluppo dell'essere umano: qualunque persona al di sopra dei 18 anni è da confrontarsi con quei valori. La lunghezza dell'addome è calcolata sottraendo ad una casella di «Trunk» la corrispettiva di «Thorax».

#### 5.5.1.1 Risultati dimensioni medie del corpo umano

Questi sono i risultati ottenuti sulle 1200 immagini:

- Su 200 encefali antero-posteriori 152 sono stati individuati: sensitivity al 76%
- Su 200 encefali laterali 142 sono stati individuati: sensitivity all' 71%
- Su 200 toraci antero-posteriori 136 sono stati individuati: sensitivity al 68%
- Su 200 toraci laterali 130 sono stati individuati: sensitivity al 65%
- Su 200 addomi antero-posteriori 148 sono stati individuati: sensitivity al 74%
- Su 200 addomi laterali 138 sono stati individuati: sensitivity al 69%

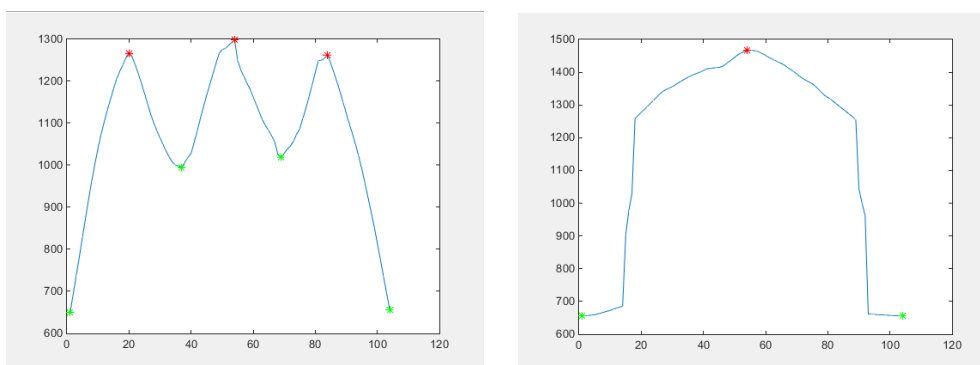
#### 5.5.1.2 Valutazione dimensioni medie del corpo umano

Neppure questo metodo ha dato gli esiti positivi che ci si aspettava: in molte immagini il confronto dimensionale dava risultato positivo non univoco, in quanto si applicava a più distretti anatomici assieme. La sola informazione sull'altezza, non è in grado di individuare in modo univoco il contenuto dell'immagine. Altro elemento fuorviante è anche la frequente presenza nelle radiografie di elementi estranei quali le attrezzature mediche o parti del lettino ove viene fatto sdraiare il paziente, che vanno a modificare la dimensione stimata del distretto anatomico. Questo metodo non è facilmente migliorabile ed il 70% circa di risultati corretti non lo rende accettabile.

#### 5.5.2 Distribuzione spaziale dei livelli di grigio nell'immagine

Dal momento che con HOG si sono avuti ottimi risultati nell'individuazione dell'encefalo, lo si usa per distinguere la testa dalle altre parti del corpo. Per le restanti

immagini si costruisce lungo tutta l'altezza dell'immagine, per ogni riga, una funzione matematica tenendo conto della distribuzione spaziale dei livelli di grigio in quella riga. La funzione matematica viene calcolata per interpolazione polinomiale dopo aver applicato più volte all'immagine un filtro per la rimozione del rumore (cioè per rendere i contorni della figura più netti). Questa operazione, che ha motivo di applicabilità solo in caso d'immagini antero-posteriori, viene svolta poichè si è notato che in corrispondenza dei polmoni e dell'addome l'andamento di tali funzioni è molto diverso (Figura 5.2). Questo andamento di funzione si verifica solo per immagini antero-posteriori in quanto, in quelle laterali, non è possibile vedere entrambi i polmoni e quindi non si verifica tale differenziazione. Malgrado la rimozione del rumore, per diversificare queste due funzioni, non è possibile utilizzare un modello matematico, ma solamente uno empirico che si basa sulla differenza di altezza tra i minimi ed i massimi delle funzioni.



(a) La tipica funzione della distribuzione dei livelli di grigio di un'immagine in corrispondenza dei polmoni (e quindi del torace)

(b) La tipica funzione della distribuzione dei livelli di grigio di un'immagine in corrispondenza dell'addome

Figure 5.2: Funzioni distribuzione spaziale dei livelli di grigio nell'immagine

### 5.5.2.1 Risultati distribuzione spaziale dei livelli di grigio nell'immagine

Come già specificato, questo metodo è applicabile solo a toraci ed addomi antero-posteriori. Questi sono i risultati ottenuti sulle 400 immagini:

- Su 200 toraci antero-posteriori 136 sono stati individuati: sensitivity al 68%
- Su 200 addomi antero-posteriori 148 sono stati individuati: 74%

### 5.5.2.2 Valutazione distribuzione spaziale dei grigi nell'immagine

I risultati di questo metodo non sono comunque accettabili. La ragione principale è la difficoltà ed impossibilità di applicare il modello matematico. L'applicazione di un modello empirico basato sulla altezza dei minimi e dei massimi si è dimostrata

non generalizzabile per immagini molto diverse tra loro. Presenta inoltre il limite di non poter essere utilizzata sulle immagini laterali (come spiegato prima) e ha un grado di complessità di valutazione superiore ad altre tecniche, senza fornire risultati migliori.

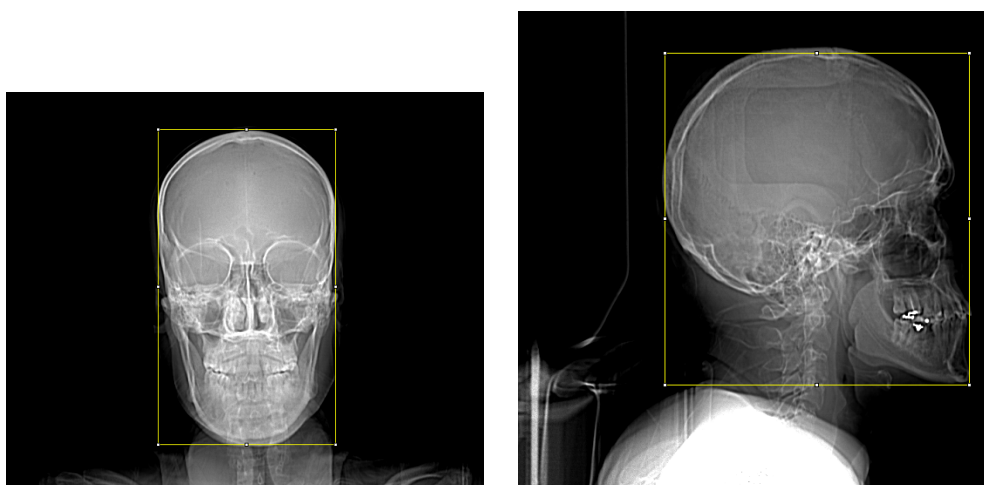
### 5.5.3 Valore della simmetria in regioni d'interesse con HOG

Considerando che le prestazioni di HOG appaiono le più attinenti alla risoluzione del problema, si è pensato di continuare a testare questa tecnica cercando di evitare il più possibile di deformare le immagini nell'acquisizione dei campioni.

#### 5.5.3.1 Creazione del Dataset

L'attività dell'utente umano risulta essere fondamentale: deve individuare, per il distretto anatomico che vuole etichettare nel dataset, un minimo di due ed un massimo di quattro immagini scout che, a livello visivo e dimensionale, siano simili tra loro, ma appartenenti a pazienti diversi. All'interno di una di queste immagini bisogna selezionare una "porzione" particolarmente significativa (Region Of Interest o ROI) che sia in grado di descrivere al meglio il distretto anatomico preso in analisi (Figure 5.3, 5.4, 5.5). Nelle altre immagini l'utente sarà chiamato a selezionare altre ROI (Sottosezione 2.6.1), scelte con lo stesso criterio, ma che dovranno avere una dimensione in pixel paritaria rispetto alla prima selezionata. Per motivi prestazionali, le ROI scelte vengono scalate ad un 64esimo della loro reale dimensione e viene calcolato il features vector mediante HOG (Sottosezione 3.4.2.2). Un elemento del dataset è quindi costituito da: l'informazione relativa al distretto anatomico analizzato con relativo orientamento ( laterale o antero-posteriore ), i features vectors di tutte le ROI, le dimensioni delle ROI che li hanno generati ( tutte uguali tra loro ) e il fattore di scala utilizzato. Più elementi del dataset possono essere utilizzati per rappresentare un medesimo distretto anatomico.

Al momento, per raggiungere l'obiettivo di questa tesi, sono stati utilizzati circa 5-6 elementi, così costruiti, per ogni categoria di distretto anatomico ( un totale di 10-15 immagini per ognuna delle 6 classi del problema ). E' stato inoltre aggiunto un elemento per gestire una casistica particolare: gli arti inferiori in zona pubica. L'aggiunta è stata necessaria poiché il programma inizialmente, in presenza di arti inferiori in immagini di addome antero-posteriore, riconosceva erroneamente il torace antero-posteriore: con questo accorgimento tale criticità è stata risolta. Il dataset contiene, quindi, un totale di 28 elementi in grado di riconoscere le 6 tipologie di distretti anatomici.



(a) La tipica ROI scelta nella creazione di un dataset per encefalo antero-posteriore

(b) La tipica ROI scelta nella creazione di un dataset per encefalo laterale

Figure 5.3: Tipiche ROI encefalo

### 5.5.3.2 Processo di etichettatura

La procedura di etichettatura automatica di un'immagine scout (che per semplicità chiameremo "campione" o  $C$ ) è effettuata confrontando porzioni di tale immagine con tutti gli elementi del dataset precedentemente costruito. Si verifica se  $C$  può contenere una ROI di dimensioni uguali a quelle di un elemento ( $E$ ) del dataset: se la risposta è affermativa si continua l'analisi, in caso contrario si testa l'elemento successivo.

Per tutta la lunghezza dell'asse verticale centrale del campione, ogni 8 pixel, vengono selezionate delle porzioni d'immagine ( $P$ ) con le stesse dimensioni della ROI ( $R$ ) dell'elemento  $E$  in analisi. Queste porzioni vengono scalate anch'esse ad un 64esimo della loro dimensione: la scelta del numero 64 è stata fatta perché, durante la fase di testing, si è notato che fosse il migliore compromesso per numero di risultati esatti e prestazioni del programma.

Dopo la scalatura, viene calcolato il valore dei features vectors di tutte le porzioni  $P$  utilizzando HOG. Tutto il processo viene ripetuto anche lungo altri assi verticali dell'immagine, situati nell'intorno di quello centrale, finché è possibile ottenere delle porzioni d'immagine con dimensioni paritarie ad  $R$ . Più ci si sposta dall'asse centrale di  $C$ , più si alza il rischio di ottenere risultati poco affidabili: tale fenomeno è dovuto alla presenza dello sfondo nero nelle zone laterali delle scout che influisce negativamente sul confronto.

L'algoritmo prosegue confrontando tutti i features vectors delle  $P$  ottenute con un feature vector  $M$ .  $M$  è calcolato come media tra tutti i vettori dell'elemento  $E$  preso in analisi. Il confronto avviene calcolando ognuno dei features vectors delle  $P$  in funzione di  $M$ : si ottengono delle rette che, nel caso in cui i vettori presi in

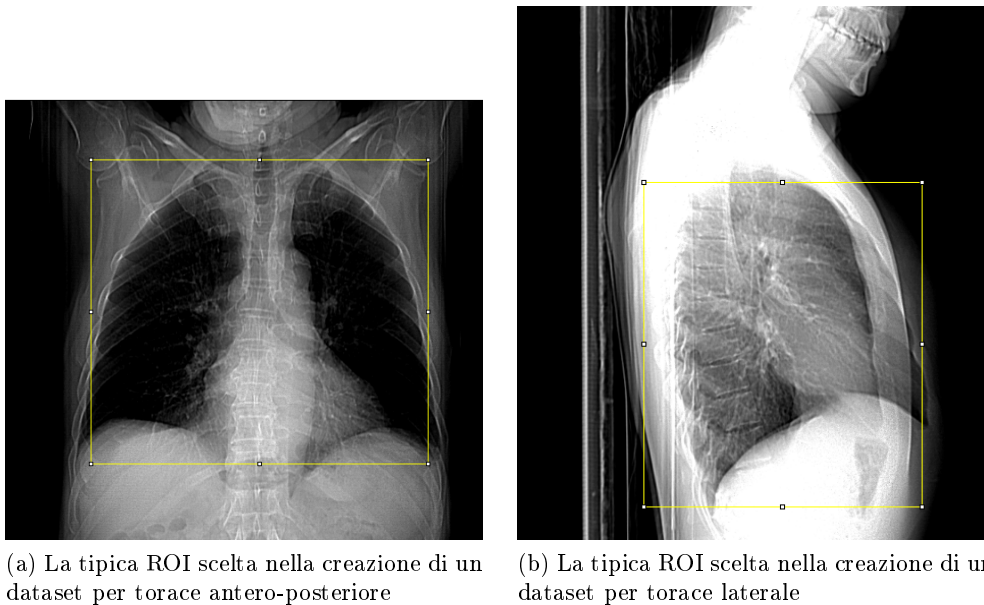


Figure 5.4: Tipiche ROI torace

esame fossero uguali, avrebbero un coefficiente angolare di valore 1. Si calcolano quindi tali rette per tutti i features vectors delle  $P$  e, tra tutti i coefficienti angolari di queste rette, si sceglie quello che ha valore più vicino ad 1. Questo lungo processo è ripetuto per tutti gli elementi del dataset: per ogni elemento  $E$  si otterrà quindi un solo coefficiente angolare  $e$ , tra tutti i coefficienti angolari ottenuti, si andrà ancora una volta a selezionare quello più vicino al valore 1.

Questo risultato permette non solo di etichettare correttamente il distretto anatomico rappresentato in immagine, ma anche d'individuare precisamente la sua posizione all'interno di un'immagine più ampia semplicemente memorizzando le coordinate della porzione dell'immagine che ha portato al miglior risultato.

L'operazione di ottenimento del coefficiente angolare, che equivale al calcolo della simmetria tra due insiemi di punti, è stata individuata dopo che le tecniche più accreditate sono fallite. Se si utilizza HOG, la tecnica più utilizzata per fare clustering, ovvero per etichettare dei campioni, è Support Vector Machine, ma, in questo caso, si è dimostrata completamente inefficace. Tale fallimento è dovuto al fatto che i features vectors degli elementi del dataset non hanno tutti le stesse dimensioni: la grandezza di un feature vector ottenuto con HOG dipende dal numero di pixel della ROI utilizzata per la sua creazione. La Distanza di Bhattacharyya, che è solitamente usata per il calcolo della somiglianza tra istogrammi, ha dato risultati equiparabili alla pura casualità. Infine anche la Distanza Euclidea, lavorando direttamente su porzioni d'immagini e ROI, ha fallito dando i medesimi risultati di Bhattacharyya.



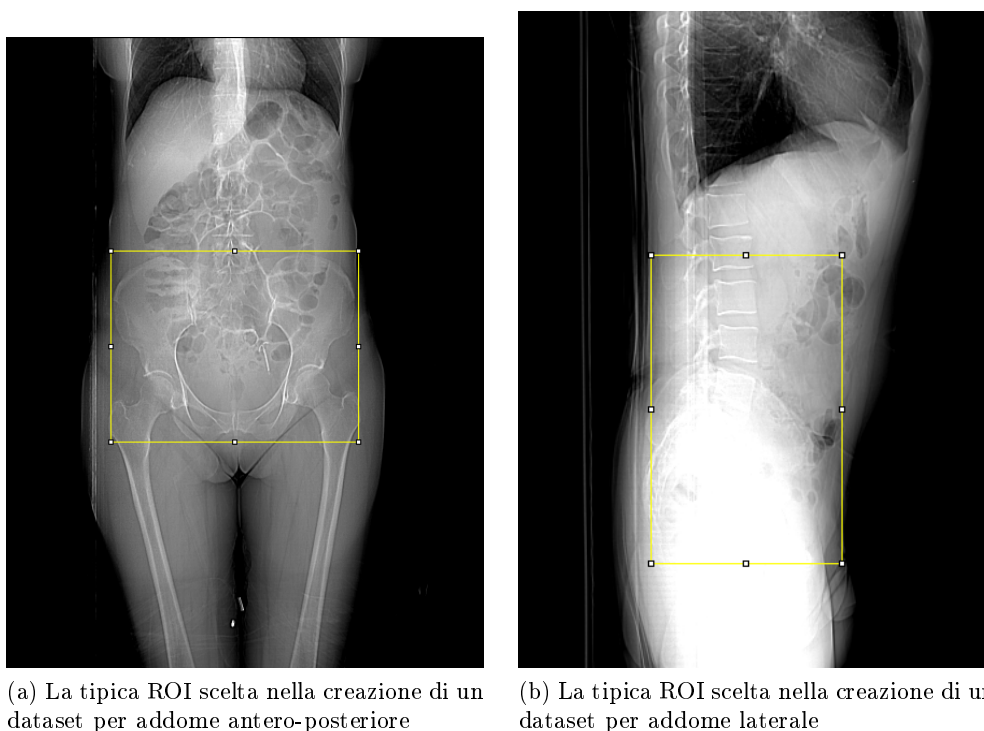


Figure 5.5: Tipiche ROI addome

### 5.5.3.3 Risultati e valutazione

Si rimanda al Capitolo 6.

### 5.5.3.4 Struttura ed utilizzo del programma MATLAB

Il programma scritto in MATLAB, che calcola la simmetria dei punti di interesse con HOG, è composto da 3 file<sup>2</sup> che rispecchiano le funzioni dell’algoritmo descritto nei precedenti paragrafi.

Il file **START.m** che permette il lancio del programma e contiene la logica relativa all’interfaccia grafica di base per la selezione della modalità d’utilizzo del programma (Figura 5.6).

Nel file **TRAINING.m** vi è la logica relativa al processo di creazione ed aggiunta di un elemento al dataset. L’utente deve precedentemente aver preparato una cartella con le immagini DICOM da analizzare. Il programma chiederà di selezionare tale cartella e tutte le immagini DICOM in essa presenti saranno utilizzate dal programma per creare un nuovo elemento di dataset. Dopo aver scelto il fattore di scala

<sup>2</sup> In realtà sono 4: 3 dei quali creati appositamente per questa tesi ed uno scaricato da Internet. Tale file permette il calcolo dei feature vectors con HOG. Le ultime versioni di MATLAB sono già aggiornate con tale funzione, ma purtroppo, in dotazione all’Ospedale S. Anna, è presente una versione obsoleta (MATLAB 2010b) e quindi è stato necessario integrare i software per poter testare le prestazioni del programma sui PC ospedalieri.

(64 come valore di default), un'interfaccia permetterà di selezionare direttamente su tali immagini le regioni d'interesse, una per ogni immagine. Per la prima immagine della serie bisognerà scegliere posizione e dimensioni della ROI, mentre per tutte le altre solo la loro posizione (le dimensioni saranno paritarie a quelle della prima ROI selezionata). Vengono quindi generati 2 files: NOMEELEMENTO.mat che conterrà i feature vectors generati con HOG a partire dalle ROI selezionate e un file NOMEELEMENTO.mat.txt con informazioni relative alle dimensioni in pixel della prima ROI selezionata ed il valore del fattore di scala. Al posto di "NOMEELEMENTO" l'utente può inserire il nome che preferisce. Questi due files sono usati per rappresentare un solo elemento del dataset e vanno inseriti, insieme a tutti gli altri, in una cartella dedicata (Figura 5.7).

Nel file `TEST_SAMPLES.m` è racchiusa la logica relativa all'identificazione del distretto anatomico. Bisogna selezionare la cartella coi file del dataset precedentemente creati ed accorpatisi e una cartella che contiene le immagini DICOM che si vogliono testare. Durante l'elaborazione verranno mostrate una serie d'immagini utili ad individuare, a posteriori, problematiche ricorrenti al fine di correggerle (Figura 5.8). Questa funzionalità, che appesantisce il programma allungando il tempo d'esecuzione, è disattivabile. I risultati finali dell'elaborazione, calcolati come spiegato sopra, vengono salvati all'interno di un file chiamato Results.txt. Questo file contiene informazioni sull'esecuzione del programma, ovvero tempo di esecuzione e nome degli elementi del dataset, ed i risultati della computazione mostrati in 3 colonne. Nella prima si trova l'elenco dei file testati, nella seconda il risultato dell'elaborazione (ovvero nome dell'elemento che è risultato essere più simile a quel file) e nella terza il valore numerico di tale risultato (coefficiente della retta dei vettori campione ed elemento sotto test) (Figura 5.9).

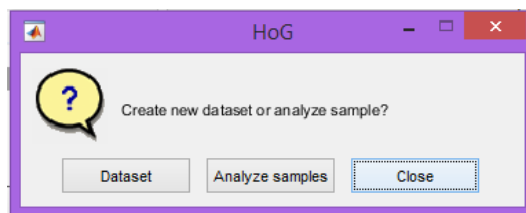


Figura 5.6: La tipica interfaccia del programma MATLAB

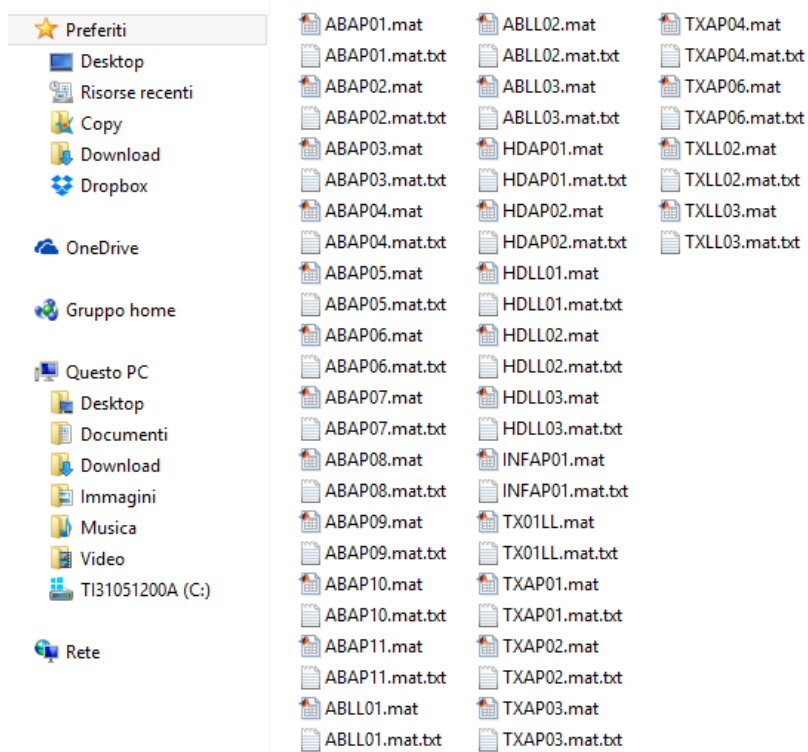


Figura 5.7: Aspetto di una cartella contenente i file del Dataset: ogni elemento è rappresentato da un file .mat ed uno .txt



## Capitolo 6

# Risultati

Una volta che la scansione di tutta una cartella di immagini è terminata, si può avere accesso al file Results.txt per verificarne la veridicità. Tale file, oltre al distretto anatomico col relativo grado di similarità, presenta informazioni riguardo il tempo di elaborazione e un elenco completo degli insiemi utilizzati per il confronto. Una fase comunque indispensabile e' quella di dover verificare direttamente se la classificazione proposta e' veritiera e valutare quantitativamente lo scostamento tra il contenuto reale delle immagini e quanto si e' ottenuto. In ultima analisi bisogna poi accertarsi quale sia l'entità dell'errore trovato.

Si è verificato empiricamente che un'analisi è affidabile se il risultato numerico dell'elaborazione è maggiore di 0,7 (coefficiente della retta calcolata): al di sotto di tale soglia la misurazione perde credibilità. Tipicamente ci sono due parametri di cui tenere conto in indagini effettuate su un insieme di immagini di questo tipo: l'accuratezza dei risultati, cioè quanto il sistema è affidabile nel suo funzionamento, ed i tempi di esecuzione per l'ottenimento del risultato. Mentre il primo parametro è di vitale importanza il secondo non è di particolare interesse, in quanto non è necessario che l'analisi venga effettuata in tempo reale sugli esami in archivio. Una tecnica efficace per valutare la bontà di un sistema a partire dai risultati è la matrice di confusione, che aiuta nell'analisi dei risultati fornendo parametri utili. Nei successivi paragrafi verranno analizzati e commentati i risultati ottenuti.

### **6.1 Matrice di confusione per il metodo «Valore della Simmetria in regioni d'interessi con HOG»**

Una spiegazione del metodo di costruzione ed estrapolazione dei dati con la matrice di conclusione è già stato affrontato nel Paragrafo 3.5.

In questo caso di studio, si è deciso di unificare le classi encefalo antero-posterio ed encefalo laterale in quanto quello che in realtà interessa è capire il distretto anatomico rappresentato nell'immagine e non il suo orientamento (laterale o frontale).

Considerati gli ottimi risultati con la tecnica HOG tradizionale e dal momento che gli errori riscontrati nell'etichettare tali distretti anatomici consistono esclusivamente nel loro orientamento (ed in un caso che verrà approfondito più avanti) non è opportuno dividere queste due classi.

Un'altra considerazione va fatta sugli arti inferiori in zona pubica usati per il training set: nel caso il sistema rilevi la presenza di arti inferiori, l'assunzione automatica è che si è in presenza dell'addome al di sopra di tale parte dell'immagine. Di conseguenza l'immagine viene etichettata come addome dal momento che i falsi positivi sulle gambe risultano uguali a zero (Figura 6.1).

Di seguito, in Tabella 6.1, i risultati ottenuti per questo sistema.

Sample file:Addome AP (130).dcm  
Training file:INFAP01.mat



Figura 6.1: Un esempio tipico dell'individuazione delle gambe: l'insieme delle gambe è INFAP01.mat e la zona che viene rilevata è quella verde in figura.

Distretto	Encefali	Toraci AP	Toraci L	Addomi AP	Addomi L	TP+FN
Encefali	399	0	1	0	0	400
Toraci AP	7	183	0	10	0	200
Toraci L	21	10	161	8	0	200
Addomi AP	4	4	2	190	0	200
Addomi L	9	10	4	15	162	200
TP+FP	440	207	168	223	162	1200

Tabella 6.1: Matrice di confusione (Tabella 3.1)costruita tenendo conto dei risultati ottenuti dal metodo: leggendo una riga è possibile capire come si è distribuita la classificazione di immagini appartenenti ad un determinato distretto anatomico. L'elemento (1,3) per esempio rappresenta il «Numero di Encefali che sono stati classificati come Toraci Laterali». Con AP s'intende immagini antero-posteriori, con L s'intende immagini laterali, con TP+FP s'intende la somma di true positives e false positives e con TP+FN la somma di true positives e false negatives.

Dopo la matrice di confusione viene effettuata la costruzione della tabella di confusione calcolando il numero di true positives (TP), false positives (FP), false negatives (FN) e true negatives (TN) (Tabella 6.2).

Distretto	TP	FP	FN	TN
Encefali	399	41	1	759
Toraci AP	183	24	17	976
Toraci L	161	7	39	993
Addomi AP	190	33	10	967
Addomi L	162	0	38	1000

Tabella 6.2: Tabella di confusione (Tabella 3.2): in tabella i valori di true positive (TP), false positives (FP), false negatives (FN) e true negatives (TN) per ogni classe del sistema. Con AP s'intende antero-posteriori e con L s'intende laterale.

Dopo la costruzione di queste matrici, si calcolano i parametri più comuni per una situazione multi classe come questa (Paragrafo 3.5). Si nota che tutti questi parametri sono migliori tanto più si avvicinano al 100%.

**Accuracy:** un valore medio che da una valutazione generica del sistema. La sua formula è data da:

$$Accuracy = \frac{TP_{tot}}{N}$$

dove  $TP_{tot}$  è la somma dei TP di ogni classe del problema ed  $N$  è il numero totale di elementi analizzati, in questo caso 1200.

L'accuratezza del sistema è quindi data dalla somma della prima colonna della tabella di confusione. Il suo calcolo è:  $Accuracy = \frac{399+190+162+183+161}{1200} = \frac{1095}{1200}$ . L'accuracy del sistema si attesta quindi al 91,25 %. Un ottimo risultato per un sistema di questo tipo.

**Sensitivity:** esprime la proporzione di oggetti che sono stati correttamente classificati sul totale. Per un problema multi classe va calcolata singolarmente per ogni classe che compone il sistema, la sua formula è:

$$Sensitivity = \frac{TP_x}{TP_x + FN_x}$$

dove  $x$  è una generica classe del problema.

In questo sistema:

$$Sensitivity_{encefali} = \frac{399}{400} = 99,75\%$$

$$Sensitivity_{toraciAP} = \frac{183}{200} = 91,5\%$$

$$Sensitivity_{toraciL} = \frac{161}{200} = 80,5\%$$

$$Sensitivity_{addomiAP} = \frac{190}{200} = 95\%$$

$$Sensitivity_{addomiL} = \frac{162}{200} = 81\%$$

L'individuazione degli encefali rasenta la perfezione: i buoni risultati di HOG tradizionale sono ulteriormente migliorati. Addomi e toraci antero posteriori superano il 90% ed è un ulteriore passo avanti rispetto al sistema che era prima quasi eguagliabile alla casualità, mentre sui distretti laterali c'è ancora qualche problema nel riconoscimento visto l'80% di match corretti. Risultati di gran lunga migliori se confrontati a quelli dei sistemi tradizionali (Tabella 5.1).

**Precision:** rappresenta il grado di ripetibilità e riproducibilità dei risultati ottenuti su altri oggetti di una stessa classe, ma non presenti nella prima indagine. Per un problema multi classe, ogni classe di elementi ha un proprio valore di precision: calcolare il valore medio del sistema non rispecchia le sue reali caratteristiche. La formula:

$$Precision = \frac{TPx}{TPx + FPx}$$

dove x è una generica classe del problema.

Risultati con calcoli in questo sistema:

$$Precision_{encefali} = \frac{399}{440} = 90,68\%$$

$$Precision_{toraciAP} = \frac{183}{207} = 88,41\%$$

$$Precision_{toraciL} = \frac{161}{168} = 95,83\%$$

$$Precision_{addomiAP} = \frac{190}{223} = 90,52\%$$

$$Precision_{addomiL} = \frac{162}{162} = 100\%$$

Si nota subito come sorprendentemente le misure più precise siano in realtà quelle col valore di sensitivity minore, ovvero quelle che sono più difficili da classificare correttamente. Senza alcuna sorpresa, si conferma che i valori più bassi si riscontrano tra addomi e toraci frontali che sono stati il motivo principale per continuare la ricerca con altre tecniche.

**Specificity:** è un parametro che serve a misurare l'entità reale di risultati negativi identificati correttamente come tali. Per un problema multi classe, ogni classe di elementi ha un proprio valore di specificity: calcolare il valore medio del sistema non rispecchia le sue reali caratteristiche. La formula:

$$Specificity = \frac{TNx}{TNx + FPx}$$

dove x è una generica classe del problema.

Risultati con calcoli:

$$Specificity_{encefali} = \frac{759}{800} = 94,88\%$$

$$Specificity_{toraciAP} = \frac{976}{1000} = 97,6\%$$

$$Specificity_{toraciL} = \frac{1000}{1000} = 100\%$$

$$Specificity_{addomiAP} = \frac{967}{1000} = 96,7\%$$

$$Specificity_{toraciAP} = \frac{993}{1000} = 99,3\%$$



Tutti i risultati sono molto alti. Ciò vuol dire che il sistema funziona correttamente ed è in grado di riconoscere gli elementi che non appartengono ad una determinata classe.

## 6.2 Studio degli errori più frequenti

Osservando dove il sistema commetteva errori in alcuni casi si è riuscito ad osservare schemi ricorrenti ed a individuare il perché di tale errore: si nota che non è stato possibile individuare schemi ricorrenti per le scout laterali. Di seguito l'elenco degli schemi riconosciuti.

### 6.2.1 Errori encefalo



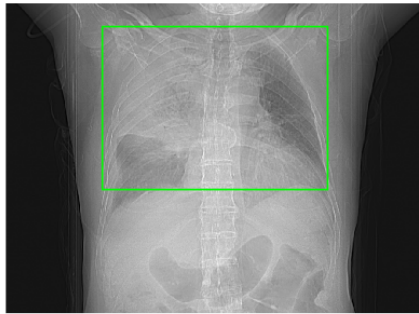
Figura 6.2: Come si può facilmente rilevare, la faccia della persona è inclinata verso destra (la sinistra nell'immagine): questa problematica non permette al sistema d'indovinare correttamente il distretto anatomico perchè non si notano i dettagli dello scheletro facciale.

Una sola immagine dell'encefalo è sbagliata (Figura 6.2) ed andando ad analizzarla nel dettaglio si nota che si tratta di un encefalo antero-posteriore con una particolarità: la faccia, a differenza di tutte le altre immagini, è girata di lato e non dritta come nel tipico esame TAC. Non essendoci un insieme in grado di gestire questa casistica fa sbagliare il sistema. Tuttavia il valore numero relativo a tale match è piuttosto basso ovvero 0,5710 per essere universalmente accettato come corretto.

### 6.2.2 Errori toraci antero-posteriori

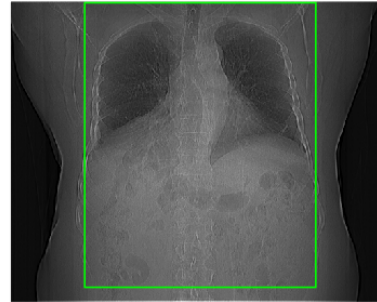
Quando si è in presenza di macropatologie sia polmonari (Figura 6.3a) che cardiache (Figura 6.3b) l'organo interessato si vede bianco nella radiologia (il cuore lo si vede normalmente bianco ma, in questo caso, è ingrossato ed occupa buona parte della cassa toracica): questo fenomeno non permette ai polmoni, solitamente molto scuri, di essere rilevati facilmente e l'etichettatura della zona ne risente.

Sample file: Torace AP (137).dcm  
Training file: ABAP01.mat



(a) Un esempio di polmoni malati. I polmoni sono tipicamente molto scuri mentre in questo caso non sono rilevabili dal programma dal momento che c'è un'infezione in atto. Il programma sbagliando rileva un addome antero-posteriore (ABAP01.mat) in corrispondenza del torace.

Sample file: Torace AP (102).dcm  
Training file: HDAP02.mat



(b) Un esempio di patologia cardiaca. Il cuore piuttosto gonfio a causa di una malattia occupa grande parte della cassa toracica e il sistema, sbagliando, a causa della sua forma sferica, lo etichetta come encefalo (HDAP02.mat).

Figure 6.3: Errori etichettatura toraci antero-posteriori

### 6.2.3 Errori addomi antero-posteriori

Alcune volte, quando si esegue un esame scout sull'addome, nella foto è presente anche parte del torace o addirittura tutto tale distretto anatomico ( Figura 6.4). Si è in presenza, quindi, di un doppio distretto anatomico e il sistema può essere che rilevi il torace con un livello di precisione maggiore rispetto all'addome: non è sbagliata la posizione del torace e la sua individuazione, ma il fatto che l'intera foto debba essere etichettata come torace.

## 6.3 Tempi di elaborazione

Il programma è stato testato su due PC molto diversi tra loro sia a livello di hardware (processore) che di sistema operativo: uno con un processore Intel Core i7-3630QM a 3,4 GHz x64 (con MATLAB 2014b a 64 bit) con Windows 64 ed un altro, in dotazione dell'ospedale, con un Intel® Pentium 4 da 3,0 GHz (usando MATLAB 2010b a 32 bit) con una versione di Linux a 32 bit.

Il tempo impiegato per l'esecuzione resta molto lungo in entrambi i casi e dipende da 4 fattori principali: numero delle immagini da testare, numero degli insiemi da cui è composto il dataset, numero dei controlli effettuati su un'immagine e presenza o meno dell'interfaccia grafica durante l'esecuzione.

Mentre i primi due fattori sono abbastanza chiari ed auto esplicativi, cioè all'aumentare del numero di tali elementi aumenta anche il tempo di computazione, il terzo ed il quarto richiedono una spiegazione più approfondita. Il numero dei con-

Sample file: Addome AP (78).dcm  
 Training file: TXAP06.mat

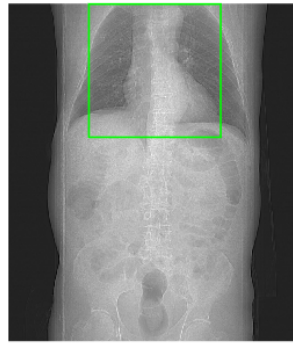


Figura 6.4: Quest'immagine deve venire etichettata come addome (oppure come torace+addome) ma il sistema rileva, anche se a ragione, solo il torace (TXAP06.mat). In immagine sono infatti presenti entrambi i due distretti anatomici ed il programma sceglie di etichettare solo quello che ha il match maggiore.

trolli effettuati sull'immagine dipende da due parametri ovvero dalla dimensione in pixel dell'immagine campione e dalla dimensione in pixel della ROI di base utilizzata per la costruzione del dataset. Le immagini possono avere un'altezza che varia dai 400 (ma c'è anche qualche caso di immagini di 200) ai 1500 pixel e, anche facendo un test solo per ogni riga (e non è mai così) ogni 8, si arriva facilmente a 200 test per ogni immagine.

Una ROI ha tipicamente una dimensione che è circa uguale ad un quarto o meno di un'immagine radiologica, quindi un'altezza che varia dai 50 pixel ai 400: nel caso di immagini con ROI particolarmente piccole il numero di test diventa piuttosto alto e difficilmente calcolabile se si tiene conto anche dei test effettuati sulla larghezza d'immagine.

La presenza dell'interfaccia grafica, che è necessaria per salvare alcune schermate dell'elaborazione, invece è un parametro molto più facilmente valutabile e calcolabile empiricamente: si guadagnano circa 10 secondi per ogni immagine che viene elaborata. L'interfaccia grafica è necessaria per vedere la posizione in cui si verifica un determinato match e quindi può essere utile in fase di testing o di comprensione degli errori (vedi sopra).

Nella Tabella 6.3 sono mostrati i tempi di elaborazione delle 1200 immagini per entrambi i computer prima descritti e si mettono i risultati per due casi di studio: il primo in presenza di un dataset di 28 insiemi (ovvero quello utilizzato per questa tesi) ed un altro con 51 insiemi (alcuni dei quali usati per tentare di correggere gli errori sopra citati ed altri introdotti esclusivamente per effettuare questo test).

Se le analisi vengono fatte senza interfaccia grafica bisogna togliere ai vari totali 3h20 (metodo puramente matematico e non testato).

Per stimare le prestazioni reali del sistema e dove verrà usato, si studiano i due

Processore + numero insiemi	Encefali AP	Encefali L	Toraci AP	Toraci L	Addomi AP	Addomi L	Tempo totale
Core i7 - 28 insiemi	2h44m	4h31m	5h04m	4h13m	6h27m	5h32m	<b>28h31 m</b>
Pentium 4 - 28 insiemi	3h52m	5h44m	6h25m	5h49m	7h45m	7h11m	<b>36h46 m</b>
Core i7 - 51 insiemi	10h31m	11h12m	18h47m	17h34m	19h02m	18h55m	<b>96h01 m</b>
Pentium 4 - 51 insiemi	13h11m	13h47m	21h09m	20h56m	22h36m	22h04m	<b>113h43 m</b>

Tabella 6.3: In tabella vengono descritti i tempi di computazione per le varie tipologie di immagini TAC sui diversi sistemi dove sono stati testati. Con AP s'intende antero-posteriore e con L s'intende laterale.

casi: quello pessimo nel caso in cui debbano essere introdotti ulteriori insiemi per migliorare il programma (caso con P4 e 51 insiemi) e quello ottimo in cui si fa affidamento sui 28 insiemi creati per questa tesi (P4 e 28 insiemi). In un anno all'Ospedale S.Anna vengono effettuati circa 26500 esami TC; utilizzando le misure in tabella si può calcolare il tempo medio di elaborazione di una immagine: i7-28=1m26, P4-28=1m50, i7-51=4m48, P4-51=5m41 per misure senza interfaccia grafica bisogna togliere 10secondi. Quindi, nel caso ottimo, ci vorranno circa 736 ore, ovvero poco meno di 31 giorni (un mese) per elaborare tutti gli esami di un intero anno, mentre, nel caso pessimo, il risultato da 2429 ore, ovvero circa 3 mesi e mezzo. Entrambe queste misure di tempo vengono accettate dai tecnici dell'Ospedale che, come già riferito prima, non hanno particolari vincoli temporali sull'elaborazione del loro storico. Per quanto riguarda l'analisi giornaliera degli esami, ovvero 73 al giorno (ed è una sovrastima) il sistema impiega poco più di 6h30 nel caso pessimo per finire la computazione ed è un tempo più che accettabile. In tabella 6.4, un riassunto di quanto detto prima.

Processore + numero insiemi	$T\_medio = \frac{Tempototale}{1200}$	$T\_medio_{sg} = T\_medio - 10s$	$T\_giorno = T\_medio_{sg} * 73$	$T\_anno = T\_medio_{sg} * 26500$
Pentium 4 - 28 insiemi	1m50s	1m40s	2h1m	<b>736h=31 giorni</b>
Pentium 4 - 51 insiemi	5m41s	5m31s	6h30m	<b>2429h=3,5 mesi</b>

Tabella 6.4: In tabella i tempi medi di esecuzione calcolati sulla base dei risultati empirici:  $T\_medio$  è il tempo medio calcolato dividendo il tempo totale per il numero d'immagini usate per la tesi,  $T\_medio_{sg}$  è una stima del tempo medio senza interfaccia grafica,  $T\_giorno$  stima il tempo medio per analizzare tutti gli esami svolti in una giornata (anche quello un numero stimato) e  $T\_anno$  il tempo per analizzare tutti gli esami di un intero anno.

## Capitolo 7

# Conclusioni e sviluppi futuri

### 7.1 Conclusioni

La presente tesi ha come obiettivo l'individuazione del distretto anatomico oggetto di indagini radiologiche di tipo TAC.

Tale passo e' fondamentale al fine di poter valutare in modo automatico la quantita' di dose effettivamente assorbita da un paziente durante l'esame radiologico. Una valutazione automatica di tale dose e' importante sia al fine di valutare la opportunita' di eseguire ulteriori analisi radiologiche sia al fine di stimare la quantita' di radiazioni da somministrare al paziente in caso di terapie radianti. L'esecuzione automatica delle misure, poi, permettera' di eliminare la variabilita' delle misure stesse introdotta dall'operatore.

Questa tesi s'inserisce in questo percorso di valutazione automatica della dose efficace coprendo l'area dell' identificazione del distretto anatomico che attualmente viene attribuito all'immagine in modo manuale.

La tesi ha considerato alcuni metodi gia' presenti in letteratura: tuttavia tali metodi non hanno fornito i risultati aspettati. La tesi ha quindi portato alla definizione di un nuovo metodo denominato "Valore della simmetria in regioni d'interesse con HOG", i cui risultati sembrano promettenti.

Per la definizione di tale metodo, sono stati individuati i punti di forza delle tecniche usate in precedenza ed in modo particolare il calcolo dei feature vectors con Histogram of Oriented Gradients. Il punto debole di tale metodo consisteva nella necessita' di modificare le dimensioni delle immagini in quanto utilizzava, come metodo di valutazione il Support Vector Machine (SVM) che ha senso solo con immagini di uguale dimensione. La nuova tecnica evita la necessita' di tale ridimensionamento. Essa infatti analizza piccole porzioni d'immagine ricavandone per ciascuna il feature vector e lo va a confrontare con la media di quelli del dataset, tramite il calcolo della loro simmetria, cercando quello che più si avvicina al coefficiente angolare 1 della retta bisettrice del primo e terzo quadrante. L'analisi viene reiterata spostando la

piccola area su cui si calcola il feature vector fino a coprire l'intera immagine da analizzare, senza variarne quindi le dimensioni originali.

L'applicazione di tale metodo ha consentito di ottenere prestazioni incoraggianti. Estendendo l'indagine ad un campione di 1200 immagini, si sono ottenuti i seguenti risultati:

- Accuracy globale del metodo: 91%
- Sensitivity nel caso pessimo: 80,5% per toraci laterali

Questi risultati sono stati considerati accettabili dall'Ente Ospedaliero per prendere in considerazione il metodo almeno come controllo incrociato rispetto alle assegnazioni effettuate con metodo manuale. Da questo buon risultato di partenza è pensabile che con un ulteriore miglioramento nella fase di scelta del posizionamento e dimensionamento delle ROI si possa giungere ad un risultato tale da poter considerare questo metodo come sostitutivo di quello di classificazione manuale.

## 7.2 **Sviluppi futuri**

Oltre chiaramente alla volontà di migliorare la precisione dell'indagine, in particolare sulle SCOUT che presentano viste laterali del corpo, vi sono ulteriori migliorie che sono apportabili a suddetto sistema.

### 7.2.1 **Sistema aperto vs. sistema chiuso**

Quando il programma restituisce un valore come risultato minore di 0,7 allora è molto probabile che il risultato dell'analisi non sia attendibile: in presenza di un caso simile bisogna chiedere una conferma all'operatore umano sull'entità dell'immagine e poter inserire una ROI presa dall'immagine relativa a tale risultato nel dataset, rendendolo aperto all'aggiunta di nuovi insiemi rispetto a quelli già presenti nel core del programma. Tuttavia, durante una fase di testing, dove si sono aggiunti altri insiemi oltre ai 28 canonici nati dall'esigenza di provare a migliorare ulteriormente il programma, si è notato che l'aumentare in modo sconsiderato le dimensioni del dataset è sia controproducente a livello di prestazioni temporali che a livello di risultati esatti. L'introduzione o meno di questa possibilità è ancora tuttora in discussione in quanto la capacità di un operatore d'intuire quando è necessario aggiungere le caratteristiche di un'immagine nel dataset è molto soggettiva ed introdurre immagini non necessarie può portare al peggioramento di tutte le prestazioni del programma.

Al momento si è pensato solo d'introdurre pochi altri insiemi con dati riguardanti le macropatologie polmonari e cardiache ed i primi test fatti sono promettenti.



Figura 7.1: La struttura corporea di un bambino di 1 anno: è abbastanza diversa sia per dimensioni che a livello osseo rispetto ad un adulto

### 7.2.2 Pazienti pediatrici

Nelle immagini esaminate sono presenti in piccolo numero anche immagini di bambini dagli 1 a 14 anni (esempio in figura 7.1): tali immagini presentano caratteristiche, sia a livello dimensionale che a livello morfologico, diverse da quelle degli adulti ed andrebbero, probabilmente, trattate come casi a parte. Il sistema al momento non ha particolari problemi nell'etichettatura corretta di tali immagini, ma, proprio a causa delle caratteristiche di tali immagini in futuro potrebbero emergere criticità date dalla differenza della struttura corporea dei bambini. Per questo sarebbe necessario un ulteriore studio per verificarne la reale entità e pensare a misure preventive, per esempio leggendo l'informazione riguardo all'età del paziente ed effettuando i test usando altri insiemi appositamente creati per pazienti pediatrici.

### 7.2.3 Estrazione ed analisi automatica delle immagini direttamente sul PACS

Ancora oggi per i tecnici è difficile sfruttare appieno qualsiasi software realizzato da sviluppatori di terze parti, poiché i sistemi PACS risultano essere particolarmente chiusi e mancano ancora di strumenti rapidi per l'esportazione organizzata delle immagini mediche: di fatto, il PACS è collegato con il server locale e tramite questo i dati in arrivo dagli scanner. L'operazione di download delle foto per l'analisi è molto lenta e macchinosa: sono necessari dai 5 ai 10 minuti per ottenere un solo esame radiologico completo. Una delle caratteristiche più desiderabili da un'applicazione

come questa è quella di rendere possibile l'analisi dei file direttamente dalla base di dati generale di tutto il reparto, preferibilmente poi memorizzando i risultati delle misure su una opportuna base di dati. L'Ospedale ha provveduto a contattare Siemens, il produttore nel PACS in loro possesso, al fine di realizzare un eseguibile in grado di funzionare in diretta unione con tale macchina, ma attualmente il supporto non è ancora disponibile.

#### **7.2.4 Supporto su immagini di più produttori**

Le immagini che sono state oggetto di questo studio provengono da macchine PHILIPS e GENERAL ELECTRIC, macchine in dotazione all'Ospedale S.Anna. Altri grandi produttori come SIEMENS o TOSHIBA, oltre ad avere una struttura diversa a livello degli header DICOM delle immagini, potrebbero avere anche caratteristiche visuali diverse rispetto alle sopracitate, rendendo difficilmente utilizzabile il sistema oggetto della presente tesi.

#### **7.2.5 Aumentare le prestazioni temporali**

Le prestazioni del programma a livello di tempo rimangono molto limitate: sarà necessario l'utilizzo di altri linguaggi di programmazione e la rifinitura del codice sorgente per migliorarle. Bisognerebbe inoltre ottenere un eseguibile per poter testare il programma fuori dal contesto S.Anna ed in PC che non hanno installato il developer kit di MATLAB.



# Bibliografia

- [1] R.Passariello et al. *Universo del Corpo*. Istituto dell'Enciclopedia Italiana, 1998.
- [2] Consiglio Euratom. Direttiva 97/43. *Gazzetta Ufficiale*, 180:22–27, luglio 1997.
- [3] Il Presidente della Repubblica. Decreto legislativo 241. *Gazzetta Ufficiale*, 203, agosto 2000.
- [4] D. A. Greenwood W. Noel Cottingham. *An introduction to the Standard Model of Particle Physics*. Cambridge University Press, 1999.
- [5] Bureau International des Poids et Mesures. The international system of units (si). 2010.
- [6] ICRP. Publication 103. 2007.
- [7] E. B. Podgorsak et al. *The 2007 Recommendations of the International Commission on Radiological Protection*. International Atomic Energy Agency, 2007.
- [8] United Nations Scientific Committee on the Effects of Atomic Radiation. *Sources and Effects of Ionizing Radiation*. UNSCEAR, 2008.
- [9] EJ Hall DJ Brenner. Computed tomography—an increasing source of radiation exposure. *N. Engl. J.Med.*, 357(22):2277–84, novembre 2007.
- [10] Ron Brinkmann. *The Art and Science of Digital Compositing*. Morgan Kaufmann, 1990.
- [11] American Association of Physicists in Medicine. The measurement, reporting and management of radiation dose in ct. *Report n.96*, 2008.
- [12] Jessen KA et al. European guidelines on quality criteria for computed tomography. *Office for Official Publications of the European Communities*, 2000.
- [13] American Association of Physicists in Medicine. Stanstandard methods for measuring diagnostic x-ray exposures. *AAPM*, 1990.

- 
- [14] Boone JM. The trouble with ctdi 100. *Medical Physics*, pages 1364–1371, 2007.
- [15] International Electrotechnical Commission. 60601-2-44.
- [16] Schueler BA McCollough CH. Calculation of effective dose. *Medical P*, pages 828–837, 2000.
- [17] Shrimpton P.C. et al. National survey of ddose from ct in the uk. *Br. J. Radiol.*, pages 968–980, 2003.
- [18] R. L. Morin J.A. Brink. Size-specific dose estimation for ct: How should it be used and what does it mean? *Radiology*, 2012.
- [19] National Electrical Manufacturers Association. Digital imaging and communications in medicine (dicom) part 1: Introduction and overview. 2006.
- [20] Konig H. *Access to persisten health information objects: Exchange of image and document data by the use of DICOM and HL7 standard*. International Congress Series, 2005.
- [21] Yan Zhang Liwei Wang and Jufu Feng. On the euclidean distance of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1335–1339, agosto 2005.
- [22] Noboyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems Man and Cybernetics*, 1979.
- [23] Bhattacharyya A. On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society*, 1943.
- [24] Loza et al. Structural similarity-based object tracking in video sequences. *Proc. of the 9th International Conf. on Information Fusion*, 2006.
- [25] Tinne Tuytelaars Herbert Bay and Luc Van Gool. Speeded up robust features. *Katholieke Universiteit Leuven*, 2008.
- [26] B. Triggs N. Dalal. Histograms of oriented gradients for human detection. *Computer Vision and Pattern Recognition*, 2005.
- [27] M.A. Aizerman et al. Theoretical foundations of the potential function method in pattern recognition learning". *Automation and Remote Control*, 1964.
- [28] H. Handels M. Egmont-Petersen, D. de Ridder. Image processing with neural networks - a review. *Pattern Recognition*, 2002.

- [29] T.Kohonen. *Self-organization and associative memory: 3rd edition*. 1989.
- [30] Donald O. Hebb. *The organization of behavior: A neuropsychological theory*. 2002.
- [31] Stephen V. Stehman. Selecting and interpreting measure of thematic classification accuracy. *Remote Sensing of the Environment*, 62(1):77–89, 1997.
- [32] Guy Lapalme Marina Sokolova. A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, 2009.
- [33] Spilker C. The acr-nema digital imaging and communications standard: a non technical description. *J Digit Imaging*, 1989.
- [34] Andrew Looms. *Figure drawing for all it's worth*. 1943.

# Appendice A

## Glossario

**API** Application Programming Interface

**ARGO** Un programma sviluppato dal Politecnico di Milano per il calcolo della dose di radiazioni in ambito CT

**CD** Compact Disk

**CTDI** Computed Tomography Dose Index

**CTDI<sub>100</sub>** Computed Tomography Dose Index misurato con una camera ionizzante da 100mm

**CTDI<sub>vol</sub>** Computed Tomography Dose Index Volumetrico

**CTDI<sub>W</sub>** Computed Tomography Dose Index calcolato media ponderata della radiazione su tutto il corpo

**DICOM/DCM** Digital Imaging and Communications in Medicine

**DLP** Dose-Length Product (Prodotto Dose-Lunghezza)

**ECO** Ecografia ED Effective Dose (Dose Efficace)

**FN** False Negatives

**FP** False Positives

**Gy** Gray: unità di misura della dose di radiazione ionizzante assorbita dalla materia

**HOG** Histogram of Oriented Gradients

**ICRP** International Commission on Radiological Protection

**ID** Identity Document

**IDE** Integrated Development Environment

---

**IMED** IMage Euclidean Distance

**MATLAB** MATrix LABoratory

**MSE** Mean Square Error

**NIH** National Institute of Health

**OO** Object Oriented

**PACS** Picture Archiving and Communication System

**PC** Personal Computer

**PSNR** Peak Signal to Noise Ratio

**RF** RadioFrequenza

**RM** Risonanza magnetica

**ROI** Region Of Interest

**RX** Radiazioni X

**SOM** Self-Organizing Map

**SSDE** Size-Specific Dose Estimates

**SSIM** Structural SIMilarity

**SURF** Speeded Up Robust Features

**Sv** Sievert: unità di misura derivata dalla dose di radiazione ionizzante che esprime gli effetti di tale radiazione sulla salute umana

**SVM** Support Vector Machine

**TAC/TC/CT** Tomografia Assiale Computerizzata/Tomografia Computerizzata/  
Computed Tomography

**TN** True Negatives

**TP** True Positives

**UH** Unità Hounsfield