

UNIVERSITÀ DEGLI STUDI DI PISA  
FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI

Corso di Laurea in Informatica

**Analisi sperimentali di modelli  
neurali per la classificazione in  
immagini di risonanza  
magnetica del seno**

Candidato  
Carmela Luongo

Relatori  
Franco Alberto Cardillo  
Alessio Micheli

Controrelatore  
Umberto Barcaro

Anno Accademico 2010/2011



*Alla mia famiglia*



# Ringraziamenti

Desidero esprimere i miei più calorosi ringraziamenti al Prof. Franco Alberto Cardillo ed al Prof. Alessio Micheli. La buona riuscita del mio lavoro è dovuta soprattutto alla fiducia che mi hanno dimostrato in ogni momento, al loro sostegno costante, ai loro preziosi suggerimenti che mi hanno aiutato nei momenti di difficoltà e mi hanno indicato la giusta via da seguire durante tutto lo svolgimento della tesi.



# Indice

<b>1</b>	<b>Introduzione</b>	<b>1</b>
<b>2</b>	<b>Tumore al seno</b>	<b>9</b>
2.1	Oncologia del seno . . . . .	9
2.2	Statistiche . . . . .	12
2.3	Esami del seno . . . . .	13
2.3.1	Mammografia . . . . .	14
2.3.2	Ecografia . . . . .	16
2.3.3	Tomografia Computerizzata . . . . .	18
2.3.4	Risonanza Magnetica . . . . .	19
2.3.5	Confronti . . . . .	19
<b>3</b>	<b>Contrast-Enhanced MRI</b>	<b>21</b>
3.1	Introduzione storica della CE-MRI . . . . .	21
3.2	Descrizione . . . . .	22
3.2.1	Principi fisici . . . . .	22
3.2.2	Tecnica e requisiti . . . . .	23
3.2.3	Agente di contrasto . . . . .	23
3.2.4	Pregi . . . . .	24
3.2.5	Limiti . . . . .	25
3.2.6	Osservazioni . . . . .	25
3.3	Protocolli diagnostici . . . . .	27
3.3.1	Criteri morfologici . . . . .	27
3.3.2	Criteri dinamici . . . . .	29
3.3.3	Combinazione di criteri . . . . .	32
3.4	Parametri dell'esame . . . . .	33
<b>4</b>	<b>Analisi automatica</b>	<b>39</b>
4.1	Applicazione . . . . .	40
4.1.1	Linguaggio di programmazione e libreria grafica . . . . .	40
4.1.2	Interfaccia grafica . . . . .	41

## INDICE

---

4.1.3	Region Growing . . . . .	44
4.2	Classificatori . . . . .	47
4.2.1	Classificatore a soglia . . . . .	49
4.2.2	Multilayer Perceptron . . . . .	50
<b>5</b>	<b>Sperimentazione</b>	<b>57</b>
5.1	Criteri di analisi . . . . .	57
5.1.1	Analisi dei criteri . . . . .	59
5.2	Costruzione dataset . . . . .	60
5.3	Scelta e stima del modello . . . . .	66
5.4	Ambiente di calcolo . . . . .	68
5.5	Valutazione dei risultati . . . . .	70
5.6	Sperimentazioni . . . . .	71
5.7	Risultati . . . . .	74
5.7.1	Initial Slope . . . . .	75
5.7.2	Washout Ratio . . . . .	75
5.7.3	Relative Enhancement . . . . .	76
5.7.4	Relative Enhancement-Early-increase . . . . .	76
5.7.5	Relative Enhancement-Initial Slope . . . . .	77
5.7.6	Relative Enhancement-Washout Ratio . . . . .	78
5.7.7	WashoutCurve . . . . .	79
5.7.8	Relative Enhancement-WashoutPeak4 . . . . .	79
5.7.9	Relative Enhancement-WashoutCurve . . . . .	80
5.7.10	E-W-S . . . . .	81
5.7.11	E-W-EI-S . . . . .	81
5.7.12	E-C . . . . .	82
5.7.13	E-S-C . . . . .	83
5.7.14	E-W-C . . . . .	83
5.7.15	S-EI-W-C . . . . .	84
5.8	Discussione . . . . .	84
<b>6</b>	<b>Conclusioni</b>	<b>105</b>
<b>A</b>	<b>Lo standard DICOM</b>	<b>111</b>
	<b>Bibliografia</b>	<b>114</b>



# Elenco delle figure

3.1	Classificazione schematica delle curve tempo-enhancement. Curve di tipo I si trovano principalmente nelle lesioni benigne. Curve di tipo III si trovano soprattutto nelle lesioni maligne. Curve di tipo II si trovano in entrambi i casi, per questo sono incerte e possono, però, essere utilizzate come supporto al sospetto di malignità. . . . .	37
3.2	Piani di acquisizione. . . . .	38
4.1	Interfaccia grafica MRM-Works. Immagini visualizzate: Slice 21; Serie post-contrasto 3. Parametri immagine pre-contrasto: Zoom 150; Window Center 127; Window Width 255. Parametri immagine sottratta: Zoom 150; Window Center 30; Window Width 50. . . . .	42
4.2	Pannello dell'immagine sottratta con visualizzata la curva di relative enhancement di una ROI di dimensione $3 \times 3$ voxel all'interno del carcinoma. . . . .	43
4.3	Curve di relative enhancement classificate utilizzando MRM-Works . . . . .	44
4.4	Architettura di un Multilayer Perceptron . . . . .	52
5.1	Curve estratte dal dataset . . . . .	60
5.2	Area di enhancement selezionata manualmente. . . . .	62
5.3	Area di enhancement estratta con region growing. . . . .	63
5.4	Curve eliminate . . . . .	64
5.5	Immagine appartenente ad un esame di risonanza magnetica diagnosticato istologicamente come displasia. . . . .	65
5.6	Curve di learning. . . . .	67
5.7	Istogrammi dei criteri dinamici initial slope, washout ratio, early-increase e curvature at the peak of enhancement. . . . .	85

## ELENCO DELLE FIGURE

---

5.8	Esempio di curve di relative enhancement normalizzate con la regola 5.5. Le due curve di Figura 5.8a, dopo la normalizzazione, producono la stessa curva normalizzata (Figura 5.8b). . . . .	87
5.9	Pattern classificati erroneamente dalla rete neurale. . . . .	88
5.10	Istogramma delle accuratèzze ottenute dal multilayer perceptron sulle diverse combinazioni di input con le rispettive deviazioni standard. I numeri indicati sull'asse dei criteri dinamici, per migliorare la leggibilità del grafico, hanno le seguenti corrispondenze: 1 = E, 2 = E-EI, 3 = E-S, 4 = E-W, 5 = WC, 6 = E-W4, 7 = E-WC, 8 = E-W-S, 9 = E-W-EI-S, 10 = E-C, 11 = E-S-C, 12 = E-W-C, 13 = S-EI-W-C . . . . .	90
5.11	Istogrammi di accuratezza, sensitività e specificità ottenute con il classificatore a soglia (CS) applicato al criterio dinamico washout ratio, con il multilayer perceptron (MLP) applicato alla combinazione relative enhancement-washout e dai risultati presenti in letteratura (L) con le rispettive deviazioni standard.	92

# Elenco delle tabelle

3.1	Schema di Fischer. . . . .	32
5.1	Iperparametri e relativi valori testati nel multilayer perceptron. . . . .	72
5.2	Configurazione base dei multilayer perceptron. . . . .	73
5.3	Accuratezza (%), sensitività (%) e specificità (%), comprensive di deviazione standard, del classificatore a soglia sui criteri dinamici initial slope e washout ratio nelle dieci iterazioni della 10-fold cross-validation. . . . .	93
5.4	Accuratezza (%), comprensiva di deviazione standard, ottenuta con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement, relative enhancement-early-increase, relative enhancement-initial slope e relative enhancement-washout ratio. . . . .	94
5.5	Accuratezza (%), comprensiva di deviazione standard, ottenuta con il multilayer perceptron applicato alle combinazioni di criteri washoutcurve, relative enhancement- $W_{peak-4}$ , relative enhancement-washoutcurve, relative enhancement-washout ratio-initial slope. . . . .	95
5.6	Accuratezza (%), comprensiva di deviazione standard, ottenuta con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-washout ratio-early-increase-initial slope, relative enhancement-initial slope-curvature at the peak of enhancement, relative enhancement-washout ratio-curvature at the peak of enhancement. . . . .	96
5.7	Accuratezza (%), comprensiva di deviazione standard, della combinazione di criteri initial slope-early-increase-washout ratio-curvature . . . . .	97
5.8	Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle curve di relative enhancement. . . . .	97

ELENCO DELLE TABELLE

---

5.9	Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-early-increase, relative enhancement-initial slope. . . . .	98
5.10	Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-washout ratio, washoutcurve. . . . .	99
5.11	Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement- $W_{peak-4}$ , relative enhancement-washoutcurve. . . . .	100
5.12	Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-washout ratio-initial slope, relative enhancement-washout ratio-early-increase-initial slope. . . . .	101
5.13	Sensitività(%) e specificità(%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-curvature at the peak of enhancement, relative enhancement-initial slope-curvature at the peak of enhancement. . . . .	102
5.14	Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-washout ratio-curvature at the peak of enhancement, initial slope-early-increase-washout ratio-curvature at the peak of enhancement. . . . .	103
A.1	Moduli di <i>MR Image IOD</i> . . . . .	113

# Capitolo 1

## Introduzione

Il tema trattato in questa tesi nasce dall'incontro di due fattori basilari. Da una parte c'è l'interesse per lo studio e la sperimentazione di modelli neurali. Dall'altra c'è la volontà di provare a dare un contributo, seppur minimo, agli studi che riguardano il problema del cancro al seno, ancora oggi argomento di ricerca del tutto aperto, continuando ad essere una delle maggiori cause di mortalità femminile nei paesi industrializzati. La gravità del tumore al seno dipende dallo stadio della malattia. Tipicamente esso non produce sintomi quando è agli stadi iniziali, cioè quando è piccolo e maggiormente curabile. Inoltre in base allo stadio in cui viene diagnosticato ed alle sue caratteristiche biologiche cambia anche il tipo di terapia a cui deve essere sottoposta la paziente. Per questi motivi è molto importante riuscire ad effettuare una diagnosi precoce, cioè prima che il cancro inizi a causare dei sintomi e che, quindi, sia ad uno stadio più avanzato e più difficilmente curabile. La ricerca su questo tipo di malattia è un settore molto ampio e complesso sotto diversi aspetti, alcuni dei quali possono essere affrontati e migliorati attraverso un approccio automatizzato. L'introduzione della risonanza magnetica come tecnica di acquisizione delle immagini e gli sviluppi tecnologici hanno generato dataset contenenti centinaia di immagini e tale numero è in continua crescita. Questo rende la diagnosi medica e il relativo piano terapeutico fortemente dipendenti dallo studio delle immagini. La stessa analisi delle immagini costituisce un settore molto complesso e ricco di sfaccettature. L'aspetto che ci si propone di trattare durante questa tesi è la classificazione delle immagini ottenute da un esame di risonanza magnetica per fornire ai radiologi una "seconda opinione" che possa essere di aiuto nell'elaborazione della diagnosi del tumore al seno.

La risonanza magnetica è una modalità di acquisizione delle immagini non invasiva che viene utilizzata nei casi in cui ci sono ambiguità derivanti dalla mammografia e dall'ecografia oppure nei casi in cui la paziente presenta

cicatrici dovute alla radio-terapia o ad un intervento chirurgico conservativo al seno [21]. Tale tecnica utilizza potenti campi magnetici ed onde radio per creare immagini che riproducono sezioni molto dettagliate del corpo. L'esame di risonanza magnetica si basa sul fatto che i tumori sono caratterizzati da una diversa angiogenesi (formazione di nuovi vasi sanguigni) e permeabilità vascolare (presentano un'organizzazione vascolare anarchica) rispetto ai tessuti normali [4]. Per questo motivo si utilizza un agente di contrasto paramagnetico che viene iniettato in una piccola vena del braccio prima o durante l'esame e che permette una migliore visualizzazione delle lesioni. Nelle zone tumorali, infatti, a causa della maggiore vascolarizzazione, si verifica una più alta concentrazione dell'agente di contrasto, che viene definita *enhancement*. Quando viene usato in combinazione con l'agente di contrasto, l'esame di risonanza magnetica viene chiamato *Contrast-Enhanced MRI* (CE-MRI) [15, 38], e le immagini che saranno analizzate nello studio di questa tesi appartengono a tale tipologia di esame. In Contrast-Enhanced MRI l'intero volume del seno è acquisito una volta prima della somministrazione del mezzo di contrasto (serie pre-contrasto) e diverse volte dopo (serie post-contrasto). Tutte le immagini collezionate devono essere analizzate ed interpretate dagli esperti per caratterizzare le eventuali lesioni in esse individuate. Per determinare se una lesione sospetta è benigna o maligna è possibile utilizzare diversi insiemi di criteri [50]:

- criteri morfologici;
- criteri dinamici;
- combinazione di criteri.

I criteri morfologici studiano la forma e la struttura interna della lesione. I criteri dinamici studiano la diffusione e la concentrazione dell'agente di contrasto nei tessuti. La differenziazione tra lesioni benigne e lesioni maligne rimane tuttora una sfida aperta in quanto entrambe presentano, a differenza dei tessuti normali, una maggiore concentrazione dell'agente di contrasto. Per l'importanza del problema in questione e per la sua molteplice natura, sono stati effettuati diversi studi con lo scopo di testare varie strategie e metodi di analisi delle diverse categorie di criteri in modo da raggiungere, nella discriminazione tra lesioni maligne e lesioni benigne, risultati confrontabili con quelli ottenuti dai radiologi esperti. In questa tesi verranno presi in considerazione i criteri dinamici. Il primo passo nell'interpretazione di un esame di CE-MRI consiste nella ricerca di regioni di *enhancement* in ogni immagine delle serie post-contrasto, chiamate *Region Of Interest* (ROI). Dall'evoluzione nel tempo dell'intensità del segnale di risonanza, dovuta alla diffusione

dell'agente di contrasto in tutte le serie post-contrasto, si ottengono le curve "tempo-relative enhancement", corrispondenti a tali ROI, che poi possono essere analizzate sia quantitativamente che qualitativamente. L'analisi qualitativa consiste nella classificazione delle lesioni in tre categorie differenti in base alla forma della curva di relative enhancement ad esse associata [28]. L'analisi quantitativa consiste nel calcolo di parametri che misurano il comportamento dell'intensità del segnale di risonanza magnetica durante le diverse fasi dell'esame [26, 64, 11].

Poiché ogni immagine deve essere ispezionata, il tempo necessario per completare l'analisi di tutte le immagini acquisite durante un esame di risonanza magnetica può essere molto lungo. L'intero processo di ispezione, inoltre, può essere soggetto ad errore a causa della stanchezza e della distrazione del radiologo. Per ridurre la probabilità di errore, di solito si ricorre ad una doppia lettura dello stesso esame da parte di due radiologi che analizzano l'esame in questione in modo indipendente e poi confrontano i risultati. Tale approccio soffre di variabilità inter-osservatore a causa della scarsa standardizzazione di interpretazione del protocollo. Il lavoro di questa tesi si propone di realizzare uno strumento per *Computer-Aided Diagnosis* (CAD) che sia di supporto ai medici per l'analisi delle immagini di risonanza magnetica del seno e per la diagnosi del cancro al seno. Tali sistemi CAD possono essere impiegati per diversi scopi: supportare i radiologi durante la fase di ispezione delle immagini, fornire una propria diagnosi, confrontare diversi insiemi di criteri diagnostici. Tramite il loro supporto il numero di errori causati dalla stanchezza può essere ridotto e le prestazioni umane nella lettura ed interpretazione delle immagini possono essere significativamente migliorate. La tesi si sviluppa in due stadi:

- Identificazione e selezione delle ROI su cui vengono calcolati i vari criteri dinamici.
- Classificazione dei criteri precedentemente calcolati per formulare una diagnosi.

Per la raccolta dei dati si realizza un'interfaccia grafica attraverso la quale si ha la possibilità di visualizzare le immagini di risonanza magnetica di un intero esame e di selezionare all'interno di tali immagini le Region Of Interest che si intende analizzare. Per la visualizzazione e l'elaborazione delle immagini si utilizza un'estensione della piattaforma Java, la *Java Advanced Imaging* (JAI) *Application Program Interface* (API). Tale estensione fornisce un insieme di classi ed operatori per mezzo dei quali vengono realizzate alcune funzionalità di manipolazione delle immagini (zoom, modifica delle caratteristiche di luminosità/contrasto dell'immagine) con lo scopo di migliorarne

la visibilità e, di conseguenza, l'analisi. L'implementazione di tali strumenti migliora la qualità di analisi di ogni singola immagine e velocizza l'ispezione di tutte le immagini appartenenti ad un esame. L'interfaccia, quindi, oltre a contribuire alla raccolta dei dati di questa tesi, può anche essere di supporto ai radiologi nell'identificazione e nella selezione di nuovi dati su cui effettuare ulteriori valutazioni.

La classificazione dei criteri dinamici avviene mediante l'utilizzo di reti neurali artificiali. Una rete neurale, nella sua forma più generale, è un sistema di calcolo progettato per modellare il modo in cui il cervello esegue una particolare funzione di interesse, interagendo con l'ambiente esterno attraverso un processo di apprendimento [20]. Il paradigma di apprendimento utilizzato in questa tesi è l'apprendimento supervisionato. In questo paradigma il dataset è composto da un insieme di esempi input-output, in cui ogni esempio è costituito da un vettore di input e il corrispondente vettore di output desiderato. Tale risposta desiderata rappresenta l'azione ottima che dovrebbe essere eseguita dalla rete neurale. Lo scopo della procedura di apprendimento è trovare una configurazione dei parametri liberi della rete neurale tale che, per ogni vettore di input, il vettore di output prodotto dalla rete neurale sia uguale o sufficientemente vicino alla risposta desiderata.

Un aspetto importante dello studio di questa tesi riguarda la costruzione del dataset che sarà poi utilizzato in fase di sperimentazione per ottenere la classificazione dei criteri dinamici. Il comportamento del segnale di risonanza magnetica nel tempo, misurato per mezzo dei criteri dinamici, contiene informazione significativa per la distinzione tra lesioni benigne e lesioni maligne [34, 9, 60]. In generale le masse maligne sono caratterizzate da un più forte e veloce aumento della concentrazione dell'agente di contrasto nella prima fase post-contrasto rispetto alle lesioni benigne, seguito da un rapido *washout*, cioè una veloce fuoriuscita dello stesso mezzo di contrasto dal sito del tumore, nelle successive fasi post-contrasto. Tale distinzione, nella realtà, non è così ben definita in quanto la diffusione dell'agente di contrasto all'interno delle lesioni è molto spesso eterogenea, soprattutto nelle masse maligne. Esistono, infatti, molte lesioni maligne che producono solo un lento e minimo enhancement e, contemporaneamente, esistono molte lesioni benigne che presentano un aumento del segnale di risonanza magnetica nella prima fase post-contrasto molto simile a quello delle lesioni maligne [34, 68, 28]. Questo comportamento irregolare dell'enhancement causa una larga sovrapposizione tra i parametri estratti da esami maligni e quelli estratti da esami benigni. Il dataset di questa tesi è stato costruito con l'obiettivo di selezionare un insieme di dati quanto più possibile realistico. Tale insieme deve, quindi, rispecchiare la situazione reale in cui patologie benigne e tumori maligni presentano delle caratteristiche molto simili tra loro, rendendo molto difficile la



distinzione tra le diverse lesioni. Inoltre la definizione della Region Of Interest avviene manualmente, o in maniera semiautomatica, senza il supporto di un radiologo esperto. Questo ha reso il dataset soggetto all'introduzione di dati erronei e fuorvianti che sicuramente rendono più complicato il processo di classificazione e quindi di discriminazione tra patologie benigne e patologie maligne. Le sperimentazioni eseguite su un insieme di dati di questo tipo potrebbero produrre risultati non ottimali, ma la sfida della tesi è proprio quella di provare a sperimentare la performance diagnostica dei classificatori utilizzando un dataset più "ambizioso" rispetto a quelli notevolmente filtrati esistenti in letteratura [1, 34, 66, 68].

Una volta effettuata la raccolta delle regioni di maggiore enhancement e il calcolo delle relative curve e parametri dinamici, il passo successivo consiste nella classificazione di tali criteri per ottenere una diagnosi delle lesioni sospette. I parametri da tenere in considerazione nella formulazione di una diagnosi sono numerosi ed uno degli obiettivi che si propone questa tesi consiste nel realizzare un'ampia sperimentazione con le reti neurali di diverse combinazioni dei criteri dinamici per poter ottenere un buon risultato dal punto di vista diagnostico. Per eseguire una buona diagnosi bisogna prendere in considerazione e combinare molte caratteristiche presenti nelle lesioni. Un radiologo esperto si concentrerà solo su alcuni dei criteri a disposizione, in base alla propria esperienza personale. Con le reti neurali, invece, è possibile sperimentare diverse combinazioni dei parametri dinamici per testare la loro abilità discriminativa e selezionare le variabili di input più rilevanti, cioè quelle che permettono di ottenere le migliori prestazioni da un punto di vista diagnostico. In letteratura sono stati proposti diversi modelli per classificare i tumori al seno che analizzano i criteri dinamici. Durante la tesi si esegue una ricca sperimentazione in cui vengono testate anche nuove combinazioni di tali parametri e si effettua il collaudo di diverse reti neurali in base alle differenti combinazioni dei criteri dinamici. In questo modo è possibile effettuare una buona selezione delle caratteristiche dinamiche in modo da costruire un sistema efficace che possa essere di supporto all'analisi radiologica. Due particolari criteri dinamici (washout ratio ed initial slope, descritti nella sezione 3.3 della tesi), importanti nell'interpretazione radiologica delle immagini di risonanza magnetica, vengono utilizzati, a scopo di riferimento, anche in due classificatori a soglia che mostrano il loro potere predittivo.

Data la complessità e l'eterogeneità del dataset in questione, la ripartizione dei casi durante le sperimentazioni può essere molto significativa ed influenzarne il risultato. Inoltre una suddivisione fissata dei dati comporterebbe l'impiego, nella rete neurale, di insiemi di allenamento, validazione e test fissati e quindi più ristretti. Per questi motivi, durante la tesi si è deciso di utilizzare un processo di validazione più lungo ed articolato, la tecnica

*double k-fold cross-validation* [40, 12], che viene dettagliatamente descritta nella sezione 5.3. Tale procedura, per mezzo di due cross-validation annidate, permette un maggior e miglior utilizzo dei dati. Con la cross-validation esterna, infatti, il processo di test non viene eseguito su una singola porzione dei dati ma, utilizzando di volta in volta un differente sottoinsieme del dataset, viene alla fine eseguito su tutto il dataset a disposizione. Inoltre, con la cross-validation interna si realizza una procedura di model selection ampia e corretta, in quanto anch'essa eseguita su tutti i dati a disposizione, e staccata dal test set. L'utilizzo della tecnica double k-fold cross-validation, quindi, permette di effettuare un'indagine accurata ed affidabile, dalla quale non si ottiene un unico modello finale ma si ricava una stima rigorosa ed estesa dell'errore di predizione perché è calcolata su tutto il dataset ed è indipendente dalla particolare ripartizione dei dati. Al termine di questa approfondita valutazione, si sceglie la combinazione di criteri dinamici che produce i migliori risultati, e con questa si costruisce il modello finale fissato, di cui è già stata fornita la stima dell'errore di predizione durante la fase di double k-fold cross-validation. Il modello prescelto, infine, viene inserito all'interno dell'interfaccia grafica in modo che il radiologo possa chiedere la classificazione di una data curva di relative enhancement ed ottenere quindi una seconda "opinione automatica" come supporto nell'elaborazione della diagnosi.

La tesi si articola nel seguente modo. Il capitolo 2 introduce il problema del cancro al seno dal punto di vista medico e sociale. In questo capitolo si descrive lo sviluppo del cancro, le statistiche realizzate nei paesi di maggiore incidenza della malattia, i paesi occidentali, e i diversi esami utilizzati in ambito medico per identificare la presenza del cancro al seno.

Nel capitolo 3 viene effettuata una descrizione più approfondita dell'esame di risonanza magnetica, le cui immagini sono oggetto di studio di questa tesi. All'interno di questo capitolo vengono presentati anche i protocolli diagnostici più frequentemente utilizzati dai medici per la caratterizzazione di una lesione sospetta.

Il capitolo 4 illustra i sistemi CAD realizzati ed utilizzati nello studio di questa tesi per l'analisi delle immagini di risonanza magnetica e per la diagnosi del cancro al seno. Nel capitolo, quindi, viene descritta l'interfaccia grafica utilizzata per la visualizzazione e la manipolazione delle immagini e vengono presentati i classificatori che saranno poi testati in fase di sperimentazione per classificare i vari criteri dinamici.

Il capitolo 5 mostra tutto il processo di sperimentazione effettuato durante questa tesi. Nel capitolo viene descritta la fase di costruzione del dataset, il processo di validazione delle reti neurali e i risultati di tutte le sperimentazioni eseguite sulle diverse combinazioni di criteri dinamici.

Nel capitolo 6, infine, vengono spiegate le conclusioni raggiunte su tutto il lavoro effettuato e vengono poste le premesse per ulteriori miglioramenti e sviluppi futuri.



# Capitolo 2

## Tumore al seno

Il cancro al seno è una delle più antiche forme tumorali, se ne trova una sua descrizione già nel *papiro egiziano di Edwin Smith* risalente al XVII secolo a.C. e probabilmente copia di un lavoro ancora più antico. Il papiro descrive 8 casi di tumori o ulcere della mammella che venivano trattati mediante cauterizzazione.

A partire dagli anni '70, probabilmente a causa dei cambiamenti nello stile di vita nel mondo occidentale, il numero di casi di carcinoma mammario è aumentato in maniera significativa, al punto da costringere i paesi industrializzati a studiare programmi di monitoraggio che permettano di individuare il cancro nelle sue fasi iniziali. Tali programmi, insieme ai progressi tecnologici nelle modalità diagnostiche, consentono oggi una diminuzione del tasso di mortalità nonostante il numero di donne colpite sia in continua crescita.

In questo capitolo viene introdotto il problema del cancro al seno dal punto di vista medico e sociale. La prima sezione descrive l'anatomia del seno e discute lo sviluppo del cancro al seno evidenziando l'importanza di una diagnosi precoce. La sezione 1.2 riassume le statistiche calcolate dai paesi occidentali, dove questa malattia ha la maggiore incidenza. La sezione 1.3, infine, descrive le diverse modalità diagnostiche attualmente utilizzate discutendone i rispettivi limiti e vantaggi. In questa sezione viene soltanto presentata la tecnica di risonanza magnetica che, essendo l'oggetto di studio di questo progetto, verrà descritta con maggiore precisione nel capitolo ad essa dedicato (capitolo 3).

### 2.1 Oncologia del seno

Per comprendere gli esami mammografici e sviluppare algoritmi per la loro analisi automatica è utile conoscere l'anatomia del seno femminile e capire

come un tumore vi si sviluppa.

Il seno è costituito da un insieme di strutture ghiandolari, chiamate lobuli, unite tra loro a formare un lobo. In un seno vi sono da 15 a 20 lobi. Il latte prodotto dalle ghiandole giunge al capezzolo dai lobuli attraverso piccoli tubi chiamati dotti galattofori (o lattiferi). Tessuti fibrosi e adiposi riempiono gli spazi tra i lobuli e i dotti. Il seno contiene anche vasi linfatici connessi a piccoli aggregati di tessuto linfatico detti linfonodi.

Con il termine tumore si fa riferimento ad un'intera classe di malattie caratterizzate da un'incontrollata riproduzione di alcune cellule dell'organismo che smettono di rispondere ai meccanismi fisiologici di controllo cellulare a seguito di danni al loro patrimonio genetico. Normalmente le cellule del nostro corpo si riproducono attraverso un processo ordinato di crescita: nuove cellule sane rimpiazzano quelle morte. Nel tempo, mutazioni o cambiamenti anormali nei geni all'interno del nucleo di una cellula possono causare danni al suo sistema di controllo della riproduzione, facendola diventare tumorale. Tali cellule tumorali hanno l'abilità di moltiplicarsi senza ordine o controllo, per cui un maggior numero di cellule tumorali si moltiplica ed un minor numero di esse muore, mentre quelle che sopravvivono continuano a moltiplicarsi.

Esistono molti tipi diversi di tumore al seno, i cui nomi derivano dalla parte del seno in cui ha inizio la crescita delle cellule tumorali. I più frequenti sono quelli che nascono dalle cellule ghiandolari (carcinoma lobulare) e quelli che nascono dalle cellule che formano la parete dei dotti (carcinoma duttale). Con il progredire della malattia, le cellule tumorali possono invadere i tessuti del seno ad esse vicini fino a giungere ai linfonodi, dai quali poi possono diffondersi nel resto del corpo attraverso il sistema linfatico. La gravità del tumore dipende dallo stadio della malattia. Ci sono diversi sistemi di staging per il tumore, di cui il principale è "The American Joint Committee on Cancer (AJCC) classification of tumors" che utilizza informazioni riguardanti la dimensione del tumore, quanto il tumore si è diffuso all'interno del seno e degli organi adiacenti, il coinvolgimento dei linfonodi e l'eventuale presenza di metastasi. Riassumendo, la progressione del tumore può essere descritta nei seguenti stadi:

**Stadio 0** E' usato per descrivere cellule anormali che non sono invasive: carcinoma in situ.

**Stadio 1** Il tumore è in fase iniziale, con meno di 2 cm di diametro e confinato all'interno della regione in cui ha avuto inizio, senza coinvolgimento dei linfonodi.

**Stadio 2** Il tumore non ha iniziato a diffondersi nei tessuti circostanti ma ha un diametro maggiore di 2 cm; oppure spesso questo stadio indica che il tumore ha ancora una dimensione di al massimo 2 cm, ma le cellule tumorali hanno raggiunto i linfonodi più vicini.

**Stadio 3** Il tumore è più grande ed ha iniziato ad invadere i tessuti circostanti e i linfonodi della regione.

**Stadio 4** Quest'ultimo stadio rappresenta la metastasi, cioè uno stadio in cui le cellule tumorali si sono diffuse in altri organi del corpo attraverso il sistema linfatico o il flusso sanguigno.

Un tumore viene definito *benigno* quando le sue cellule non crescono in maniera incontrollata, non invadono i tessuti circostanti e non si diffondono in altre parti del corpo.

Il tumore al seno, tipicamente, non produce sintomi quando è agli stadi iniziali della malattia, cioè quando è piccolo e maggiormente curabile, ma solo quando si trova ad uno stadio più avanzato e quindi più rischioso. Per questo motivo sono importanti gli esami mammografici che permettono di diagnosticare precocemente il cancro, anche se ancora non ha mostrato sintomi evidenti. Secondo gli studi effettuati dall'American Cancer Society, se la diagnosi avviene quando il cancro è nelle prime fasi (1 e 2), la probabilità di sopravvivenza della donna per almeno cinque anni dopo l'operazione è del 98%. Tale probabilità diventa l'84% per diagnosi nel terzo stadio. Per un cancro diagnosticato nell'ultimo stadio la probabilità scende al 23% [62].

In base allo stadio in cui viene diagnosticato il cancro ed alle sue caratteristiche biologiche cambia anche il tipo di terapia a cui deve essere sottoposta la paziente. Quasi tutte le donne con un tumore al seno, indipendentemente dallo stadio, subiscono un intervento chirurgico per rimuovere i tessuti malati. Se, però, il cancro viene diagnosticato nella fase iniziale del suo sviluppo, è possibile ricorrere alla chirurgia conservativa, cioè si asporta tutta la parte in cui si trova la lesione ma si salva il seno. L'intervento deve poi essere seguito dalla radioterapia per distruggere eventuali cellule cancerose rimaste nel seno dopo l'operazione. Se la diagnosi avviene quando il cancro è già nelle fasi più avanzate, diventa necessario asportare una parte (mastectomia parziale o segmentale) o l'intero seno (mastectomia radicale modificata). Quest'ultimo caso prevede l'asportazione della ghiandola, dei linfonodi sotto l'ascella, di parte o di tutto il muscolo pettorale e spesso anche della pelle sovrastante. In entrambi i casi, la donna deve essere sottoposta a cicli di radioterapia in seguito all'intervento chirurgico. Malgrado l'asportazione chirurgica del tumore, c'è sempre il rischio di una ricaduta dovuta a cellule che si sono staccate dalla massa d'origine e si sono sparse per il corpo. Per questa ragione alla

maggior parte delle pazienti viene prescritta anche una terapia con farmaci anticancro in grado di eliminarle (chemioterapia). Negli ultimi anni si è diffuso anche l'utilizzo della chemioterapia neoadiuvante, ovvero somministrata prima dell'intervento chirurgico per ridurre la dimensione e l'aggressività del tumore. In alcuni casi questo permette l'utilizzo della chirurgia conservativa per delle donne che altrimenti avrebbero dovuto subire la mastectomia.

## 2.2 Statistiche

L'importanza di una diagnosi precoce è evidenziata anche dalle statistiche effettuate in diversi paesi occidentali. I principali fattori che vengono analizzati sono l'incidenza e la mortalità. L'incidenza è il numero di nuovi casi di cancro che si verificano in una specifica popolazione durante un periodo di tempo. La mortalità è il numero di morti provocati dal cancro nello stesso periodo [46].

Escludendo il cancro alla pelle, il cancro al seno è il più comune tra le donne (rappresenta il 23% del numero totale di casi di cancro), e la seconda causa di morte dopo il cancro ai polmoni (14% del numero di morti per cancro tra le donne). Più della metà di tutti i casi si riscontrano nei paesi industrializzati: il 27.3% in Europa e il 31.3% in Nord America [46]. L'incidenza è più modesta nei paesi dell'Europa dell'Est, del SudAmerica, dell'Africa meridionale e dell'Asia occidentale, ma anche in queste regioni il cancro al seno continua a rappresentare il cancro più comune tra le donne. In parte, la più alta incidenza nelle aree del mondo più sviluppate è probabilmente dovuta alla presenza di programmi di screening che permettono di individuare precocemente tumori invasivi, alcuni dei quali sarebbero altrimenti stati diagnosticati più tardi o per niente.

Secondo l'American Cancer Society la percentuale di incidenza del cancro al seno invasivo ha subito diverse fasi dal 1975, quando cioè è iniziato lo studio del cancro in base alla popolazione [62]:

- 1975-1980: l'incidenza era quasi costante.
- 1980-1987: l'incidenza aumentava del 4.0% all'anno.
- 1987-1994: l'incidenza era quasi costante.
- 1994-1999: l'incidenza aumentava del 1.6% all'anno.
- 1999-2006: l'incidenza diminuiva del 2.0% all'anno.



Il rapido aumento tra il 1980 e il 1987 è dovuto principalmente al maggiore utilizzo di screening mammografico che ha permesso di individuare un maggior numero di tumori molto piccoli; infatti in quegli anni l'incidenza dei tumori più piccoli è raddoppiata, mentre quella dei tumori più grandi ( $> 3.0$  cm) è diminuita del 27%.

Lo stabilizzarsi ed il successivo declino della mortalità causata dal cancro al seno si sono verificati a partire dagli anni '90:

- 1975-1990: la percentuale di morti aumentava dello 0.4% annualmente.
- 1990-1995: la mortalità diminuiva dell'1.8% annualmente.
- 1995-1998: la mortalità diminuiva del 3.3% all'anno.
- 1998-2006: la mortalità diminuiva dell'1.9% annualmente.

La diminuzione si registra maggiormente tra le donne di età inferiore a 50 anni. La riduzione della mortalità viene attribuita al progresso delle terapie e all'aumento di diagnosi precoci con conseguente possibilità di intervenire sulla malattia quando questa è ancora in uno stadio meno rischioso e più facile da trattare.

La maggior parte dei fattori di rischio attualmente conosciuti non è facilmente modificabile: alcuni di questi sono l'età, la predisposizione genetica (mutazioni o alterazioni dei geni BRCA1 e BRCA2), la densità del seno, la menopausa in età avanzata. Ci sono, poi, altri fattori associati all'aumento di rischio di cancro al seno su cui si può intervenire, come l'obesità, l'assunzione di alcol e l'inattività fisica. Comunque non c'è, in questo momento, un modo sicuro per prevenire il cancro al seno. Tali considerazioni, unite ai dati statistici esistenti, confermano la grandezza del problema del cancro al seno e la conseguente importanza di esami che consentano una diagnosi precoce.

## 2.3 Esami del seno

Metodi tradizionali di screening per identificare il cancro al seno in fase iniziale includono la mammografia e l'ecografia. Attualmente stanno ricevendo grande attenzione nuove tecniche: la risonanza magnetica e la tomografia computerizzata. Per confrontare la performance dei diversi metodi diagnostici bisogna valutare la loro accuratezza, sensibilità e specificità. Tali misure sono calcolate nel seguente modo [35]:

- *Sensibilità*: Pazienti con sospetto cancro / Pazienti con diagnosi istologica di cancro;

- *Specificità*: Pazienti con sospetto tumore benigno / Pazienti con diagnosi istologica di tumore benigno;
- *Accuratezza*: Pazienti con diagnosi corretta (sia positiva che negativa) / Pazienti con diagnosi istologica di cancro;
- *Falso negativo*: Diagnosi errata in cui un tumore maligno (diagnosticato istologicamente) viene, invece, classificato come benigno;
- *Falso positivo*: Diagnosi errata in cui un tumore benigno (diagnosticato istologicamente) viene, invece, classificato come maligno.

La sensibilità esprime la probabilità che un test risulti positivo quando la malattia è effettivamente presente; la specificità esprime la probabilità che un test risulti negativo quando la malattia non è presente. Affinchè una procedura diagnostica sia efficace è importante che entrambe siano alte: anche se un falso negativo è più grave di un falso positivo, quest'ultimo comporta che la donna sia sottoposta ad ulteriori accertamenti, non necessari e spesso maggiormente invasivi (ad esempio, la biopsia).

### 2.3.1 Mammografia

La mammografia è una radiografia che permette la visualizzazione della struttura interna del seno. Si tratta di un test non invasivo in cui le mammelle vengono esposte ad una bassa dose di raggi X per produrre immagini dettagliate del seno. Durante la procedura il seno è compresso tra due lastre collegate ad una fotocamera altamente specializzata che raccoglie immagini del seno da due direzioni (frontale e laterale). Fino a pochi anni fa le immagini acquisite con la mammografia venivano registrate su pellicole che poi venivano sviluppate come normali fotografie. Recenti sviluppi hanno portato alla mammografia digitale.

La *mammografia digitale*, anche conosciuta come FFDM (Full Field Digital Mammography) è un sistema in cui le pellicole sono sostituite da sensori digitali che convertono i raggi X in segnali elettrici. Quest'ultimi sono usati per produrre immagini che possono essere viste al computer oppure stampate come nelle mammografie convenzionali.

Esistono due tipi di esami mammografici:

- *Mammografia di screening*: è una radiografia al seno eseguita quando la donna non ha alcun sintomo. Ha lo scopo di individuare il cancro in una fase in cui è ancora troppo piccolo da poter essere rilevato dalla donna stessa o dal suo medico.

- *Mammografia diagnostica*: è una radiografia al seno eseguita quando nella donna è stata già individuata qualche anormalità. Viene utilizzata per determinare l'esatta dimensione e posizione del tumore e per fotografare i tessuti circostanti e i linfonodi.

Differenti parti del corpo assorbono i raggi X in maniera diversa. Le ossa assorbono la maggior parte delle radiazioni, mentre i tessuti soffici, come il grasso e gli organi, si lasciano attraversare dalle radiazioni. Il risultato è che le ossa appaiono nella radiografia come zone più chiare, i tessuti soffici, invece, in diverse gradazioni di grigio ed infine l'aria appare come una zona nera. I tessuti tumorali, sia benigni che maligni, hanno la caratteristica di essere radioopachi. Se il seno è composto principalmente da tessuti adiposi, quindi relativamente trasparenti ai raggi X, il tumore appare nella radiografia come una zona più chiara.

## Pregi

**Risoluzione.** Con le moderne mammografie è possibile acquisire immagini di alta qualità utilizzando una dose molto bassa di raggi X. Tali immagini consentono l'individuazione di eventuali microcalcificazioni sospette, piccoli depositi di calcio nel tessuto soffice del seno che possono indicare la presenza di un cancro nella sua fase iniziale. Le microcalcificazioni sono presenti in circa il 20% di tutti i carcinomi e nel 40%-50% dei carcinomi scoperti in fase preclinica (cioè non palpabili) [23].

**Sensitività.** La sua sensitività è di circa l'80%-90% nelle donne in menopausa e nelle donne il cui seno è composto prevalentemente da tessuti adiposi. Con i moderni sistemi di mammografia digitale, la sensitività aumenta anche nel caso di donne non in menopausa e di donne con tessuti del seno più densi [49].

**Costo.** La mammografia tradizionale ha un costo molto basso e questo la rende adatta anche ai programmi di monitoraggio di massa. I sistemi digitali, invece, risultano essere da 1.5 a 4 volte più costosi [49].

## Limiti

**Applicabilità limitata.** La sua accuratezza è molto bassa sui seni densi e sui tessuti ghiandolari perché, essendo questi radioopachi, possono nascondere eventuali tumori oppure causare falsi positivi. La mammografia non può essere applicata sui seni che hanno già subito interventi chirurgici

per la presenza di tessuti cicatrizzati opachi e nemmeno sui seni che hanno impianti di silicone, soprattutto dopo la mastectomia, perché questi non sono trasparenti ai raggi X e possono impedire una chiara visione dei tessuti circostanti [23, 24].

**Radiazioni.** Le radiazioni sono un potenziale fattore cancerogeno soprattutto per le donne più giovani e per quelle con predisposizione genetica al cancro (mutazioni ai geni BRCA1 e BRCA2). Questo limita l'età e la frequenza con cui la mammografia può essere utilizzata [71].

**Falsi positivi.** I falsi positivi costituiscono un possibile rischio della mammografia a causa della sua bassa specificità (circa il 70%), a differenza di altre tecniche di screening. I falsi positivi comportano che la donna si sottoponga ad esami aggiuntivi (ulteriori radiografie o biopsie) pur non avendo il cancro. Secondo degli studi dell'American Cancer Society, circa il 10% delle donne che ha subito una mammografia dovrà sottoporsi ad ulteriori test. Solo l'8%-10% di queste donne avrà bisogno di una biopsia e circa l'80% di queste biopsie saranno negative (daranno come risultato l'assenza di cancro) [35, 14].

**Overdiagnosi.** Si riferisce alle donne che ricevono una diagnosi di cancro invasivo o non invasivo a causa di anomalie trovate durante la mammografia che probabilmente non sarebbero mai diventate clinicamente evidenti in assenza di screening. Stime sull'overdiagnosi variano tra l'1% e il 10% [14].

**Dolore.** Durante la mammografia viene effettuata una compressione del seno per creare densità uniforme, per ridurre lo spessore del seno ed appiattare i tessuti in modo da avere immagini più chiare e ridurre la dose di radiazioni. Questa compressione può causare dolore e disagio in diverse donne [14].

### 2.3.2 Ecografia

L'ecografia è una tecnica di imaging che permette di ottenere immagini dell'interno del corpo umano attraverso l'utilizzo di ultrasuoni. Onde sonore ad alta frequenza sono trasmesse attraverso il seno da un trasduttore posto sulla pelle. I segnali di ritorno (echi) sono raccolti dallo stesso trasduttore e convertiti da un computer in immagini in bianco e nero visualizzate sullo schermo.

## Pregi

**Sensitività sui tessuti densi.** L'alta sensitività dell'ecografia nel trovare il cancro nelle donne con seni densi è dovuta al contrasto: nell'ecografia, infatti, il tessuto denso appare bianco mentre il cancro è nero. Nella mammografia, invece, sia i tessuti densi che il cancro appaiono come zone più chiare, per cui risulta più difficile distinguerli.

**Evitare inutili biopsie.** Se con la mammografia viene rilevata un'anormalità, l'ecografia è il modo migliore per capire se tale anormalità è una massa solida (ad esempio fibroadenoma benigno oppure cancro) che potrebbe richiedere una biopsia per determinarne la malignità, oppure è piena di liquido (ciste benigna). In quest'ultimo caso, essendo certi della non malignità, non è necessaria la biopsia [24].

**Immagini “real-time”.** Con l'ecografia è possibile ottenere immagini del seno in quasi tutte le direzioni e le immagini vengono visualizzate “real-time”. Il medico, nel frattempo, può memorizzare quelle che ritiene rilevanti. Grazie a tale proprietà, l'ecografia è utile nel guidare una biopsia, sia nel determinare dove posizionare l'ago, che nel determinare se un'anormalità è cancerogena.

**Nessuna controindicazione.** L'ecografia non utilizza radiazioni e non ha effetti collaterali.

**Costo.** L'esame ha un costo molto basso perché l'attrezzatura utilizzata è ampiamente disponibile e flessibile.

## Limiti

**Non approvata come strumento di screening.** L'ecografia non è una tecnica che può essere utilizzata da sola, ma serve per completare altri test di screening. Essa, infatti, non può determinare se un nodulo solido è cancerogeno e non è in grado di individuare microcalcificazioni.

**Risoluzione.** Pur avendo un'eccellente risoluzione per quanto riguarda il contrasto, l'ecografia, a differenza della mammografia, manca di buona risoluzione spaziale.

**Operatore-dipendente.** I risultati dell'esame riflettono l'abilità del radiologo (o del tecnico) di maneggiare appropriatamente l'attrezzatura.

### 2.3.3 Tomografia Computerizzata

La tomografia computerizzata (CT) è una tecnica diagnostica che, attraverso i raggi X, produce immagini dettagliate degli organi interni del corpo. Durante l'esame, la paziente passa attraverso uno scanner a raggi X che raccoglie molte fotografie del corpo da diverse angolazioni. Un computer, poi, combina tali immagini per riprodurre sezioni corporee della paziente. Il test può essere migliorato con l'infusione di un agente di contrasto endovenoso organo-iodato, che consente una migliore differenziazione di strutture con densità simile, o della stessa struttura in tempi diversi.

#### Pregi

**Risoluzione.** La tomografia computerizzata ha un'alta risoluzione spaziale e temporale che, a differenza di altre tecniche di imaging, permette di visualizzare una combinazione di tessuto soffice, ossa e vasi sanguigni. Nelle donne con il cancro al seno, quindi, questo test viene utilizzato per esaminare il torace e l'addome in modo da verificare se il cancro ha raggiunto altri organi.

**Imaging multiplanare.** I dati di un singolo esame possono essere visti come immagini nel piano assiale, coronale o sagittale, in base alla diagnosi che bisogna effettuare. In questo modo, la tomografia non solo rivela la presenza del tumore, ma anche la sua dimensione, la sua collocazione spaziale e la sua estensione.

#### Limiti

**Radiazioni.** Le radiazioni a cui è sottoposta una donna durante un esame di tomografia computerizzata sono stimate essere 10 volte superiori a quelle di una mammografia standard, 20 volte per una CT pre-contrasto e post-contrasto, circa 70 volte per una CT dinamica [58]. Per questo motivo dovrebbe essere assolutamente evitata in soggetti ad alto rischio genetico [61]. Comunque gli sviluppi tecnologici stanno diminuendo queste proporzioni, con lo scopo di arrivare ad un esame di tomografia computerizzata che utilizzi un dose di radiazioni che sia nello stesso range di quella impiegata nella mammografia convenzionale.

**Agente di contrasto.** Siccome l'agente di contrasto viene iniettato per via endovenosa in modo da ottenere una migliore qualità dell'immagine, c'è il rischio di provocare nella paziente reazioni allergiche anche piuttosto serie [58].

### 2.3.4 Risonanza Magnetica

La risonanza magnetica (MRI) è una modalità di imaging non invasiva che sta assumendo grande importanza negli ultimi anni. Attualmente viene molto spesso utilizzata nei casi in cui ci sono ambiguità derivanti dalla mammografia e dall'ecografia oppure nei casi in cui la paziente presenta cicatrici dovute alla radio-terapia o ad un intervento chirurgico conservativo al seno [21]. La risonanza magnetica utilizza potenti campi magnetici ed onde radio per creare immagini che riproducono sezioni molto dettagliate del corpo. Il sistema è in grado di cambiare i campi magnetici e le onde radio in modo da ottenere immagini da ogni piano ed ogni orientazione. Essa è, inoltre, una tecnica di imaging multiparametrica, in quanto i parametri di riferimento utilizzabili sono sia la densità protonica che i tempi di rilassamento  $T_1$  e  $T_2$ . L'esame di risonanza magnetica al seno si basa sul fatto che i tumori aumentano la vascolarizzazione per supportare la loro crescita. Per questo motivo si utilizza un agente di contrasto paramagnetico (di solito gadolinium DTPA) che viene iniettato in una piccola vena del braccio prima o durante l'esame e che permette una migliore visualizzazione delle lesioni. Nelle zone tumorali, infatti, a causa della maggiore vascolarizzazione, si verifica una più alta concentrazione dell'agente di contrasto, che viene definita *enhancement*. Quando è usato in combinazione con l'agente di contrasto, l'esame di risonanza magnetica viene chiamato *Contrast-Enhanced MRI* (CE-MRI) [15, 38]. Nel capitolo seguente la tecnica della risonanza magnetica, che è oggetto dello studio di questa tesi, verrà descritta in maniera più dettagliata.

### 2.3.5 Confronti

Lo screening mammografico può ridurre la mortalità a causa del cancro al seno del 30%-50%, ma l'ecografia si dimostra superiore nel caso di donne giovani o con seni densi e nella differenziazione tra tumori solidi (maligni) e cisti (benigni). Secondo uno studio dell'American Cancer Society, la combinazione di mammografia ed ecografia è molto efficace nell'individuazione del cancro al seno tra le donne con seni densi. In tale studio risulta che l'80% dei tumori rilevati attraverso la combinazione di queste due tecniche è nello stadio 0 oppure nello stadio 1, quindi più facilmente trattabili.

Diverse ricerche hanno confermato che un'eccellente sensibilità ed una moderata specificità possono essere raggiunte combinando l'informazione ricavata dalla risonanza magnetica con quella ottenuta attraverso le tecniche di imaging convenzionali. Anche se la risonanza magnetica fornisce informazione molto utile e spesso fondamentale, a causa dei suoi limiti essa può essere considerata solo un test aggiuntivo ma non può completamente sostituirsi.

tuire mammografia ed ecografia (per un'accurata descrizione della risonanza magnetica e della sua applicabilità si veda la sezione 3.2 del capitolo 3).

Nella sezione è stata presentata anche la tomografia computerizzata. Nonostante sia una tecnica che sta ricevendo notevole attenzione, a causa dei suoi limiti non può sostituire gli altri test diagnostici, ma deve essere utilizzata solo nel caso in cui mammografia ed ecografia non danno risultati utili e la risonanza magnetica ha delle forti controindicazioni.



# Capitolo 3

## Contrast-Enhanced MRI

Come già accennato nella sezione 2.3.4 del capitolo 2, la risonanza magnetica è una tecnica di generazione di immagini, usata prevalentemente a scopi diagnostici in campo medico, che produce immagini di alta qualità della struttura interna del corpo umano. Essa viene utilizzata per caratterizzare le lesioni sospette rilevate dalla mammografia e dall'ecografia. Un importante valore aggiuntivo dell'esame di risonanza magnetica si trova nella sua capacità di individuare tumori multicentrici e multifocali, di trovare componenti invasive nel carcinoma duttale in situ e di mostrare il tumore in un'immagine tridimensionale [48]. La sua alta sensibilità nell'individuare il carcinoma, in particolar modo nei tessuti densi, ha portato al ruolo emergente della risonanza magnetica nello screening su donne considerate ad alto rischio di cancro al seno.

Il capitolo è organizzato come segue: la prima parte contiene una breve introduzione storica della risonanza magnetica, la seconda parte presenta una sua descrizione sia dal punto di vista della tecnica che dal punto di vista dell'applicabilità, la terza parte descrive i diversi criteri diagnostici e la quarta ed ultima parte spiega i diversi parametri che caratterizzano un particolare protocollo di risonanza magnetica ed indica il loro valore nei dati utilizzati in questa tesi.

### 3.1 Introduzione storica della CE-MRI

CE-MRI è stata introdotta alla fine degli anni '80, quando è stato dimostrato che il cancro al seno presenta un enhancement significativo in seguito alla somministrazione dell'agente di contrasto. In seguito, ulteriori analisi hanno evidenziato che non solo le lesioni maligne ma anche quelle benigne possono mostrare un grado simile di enhancement, causando quindi una sovrapposi-

zione. I primi esami di CE-MRI al seno furono eseguiti con una sequenza “spin-echo  $T_1$  pesata” prima e dopo la somministrazione dell’agente di contrasto, con un tempo di imaging di almeno 5 minuti e spessore della slice di 5 mm (per la definizione del termine slice si veda la sezione 3.3). Con lo sviluppo delle sequenze “gradient-echo pulse  $T_1$  pesate” più veloci, divenne possibile l’imaging dell’intero seno in sezioni contigue più sottili. Con questa nuova tecnica si poteva ripetere la stessa immagine a piccoli intervalli di tempo e quindi caratterizzare l’enhancement di una lesione su un più breve intervallo di tempo. Utilizzando questi dati, veniva generata una curva time-intensity di cui si poteva studiare la velocità e la percentuale di enhancement. Ulteriori studi con tecniche di imaging più veloci ( $T_2$  pesata,  $T_2^*$  pesata) hanno dimostrato che la fase iniziale di rapido assorbimento del contrasto, durante i primi due o tre minuti, contiene importante informazione per distinguere tra tessuti benigni e maligni [39].

## 3.2 Descrizione

### 3.2.1 Principi fisici

La risonanza magnetica si basa sul principio fisico della risonanza magnetica nucleare, una tecnica spettroscopica utilizzata per ottenere informazione microscopica fisica e chimica riguardo le molecole. Per effettuare un’indagine di risonanza magnetica il paziente viene immerso in un potente campo magnetico statico e vengono indirizzati al paziente impulsi radio a determinate frequenze dette “di risonanza”. Il fenomeno di risonanza magnetica ha l’effetto di fornire energia che viene assorbita dal corpo del paziente. Quando l’impulso di radiofrequenza viene interrotto, il sistema restituisce l’energia assorbita, sotto forma di onde elettromagnetiche della stessa frequenza dell’impulso. Tali impulsi (di ritorno) vengono misurati tramite una bobina che funge da antenna ricevente e l’informazione ottenuta viene memorizzata per ricostruire l’immagine dell’area del corpo scansionata. Questa fase di ritorno viene definita rilassamento ed avviene con due costanti di tempo distinte,  $T_1$  e  $T_2$ . La generazione di immagini avviene attraverso la ripetuta acquisizione dei segnali provenienti dal corpo. Una sequenza di acquisizione riguarda l’invio degli impulsi a radiofrequenza, che hanno la funzione di generare il segnale di risonanza, e l’applicazione di gradienti, che influenzano la frequenza del segnale ricevuto e sono necessari alla formazione dell’immagine.

### 3.2.2 Tecnica e requisiti

Non esiste ancora una tecnica standard e generalmente accettata per la risonanza magnetica al seno, ma vengono utilizzati diversi protocolli in base all'attrezzatura hardware e software disponibile, alle indicazioni cliniche ed ai risultati desiderati. Comunque per raggiungere un'alta sensibilità e accuratezza, CE-MRI deve essere eseguita con alta qualità.

La componente principale della maggior parte dei sistemi di risonanza magnetica è un magnete cilindrico. Durante l'esame (definizione in sezione 3.3) la paziente è posizionata su un lettino all'interno dello scanner, dove un campo magnetico è creato dal magnete. Un trasmettitore di segnale radio viene acceso e spento in modo da produrre un campo elettromagnetico e l'energia riflessa dal corpo viene continuamente misurata dallo scanner. Un computer digitale rielabora tali echi per riprodurre immagini del seno. Altri campi magnetici, generati da diverse bobine, sono applicati durante l'esame per modificare la forza del campo magnetico.

Il requisito fondamentale per raggiungere una buona risoluzione spaziale e temporale è avere una bobina dedicata al seno. Questo permette una sufficiente omogeneità di ricezione così che l'enhancement non sia sottostimato (nelle zone di cattiva ricezione) o sovrastimato (nelle zone di buona ricezione). Per ottenere una maggiore sensibilità è preferibile avere due bobine dedicate al seno, in modo da individuare, o eventualmente escludere, fuochi aggiuntivi del cancro nell'altro seno, che si suppone sano [24].

L'accuratezza della CE-MRI è fortemente influenzata anche dallo spessore della slice (sezione 3.3). Se una lesione occupa solo una parte della slice, solo questa percentuale dell'enhancement sarà visibile o misurabile. Altrettanto importante è la riduzione del movimento. Quest'ultimo complica la valutazione della MRI in maniera significativa [24].

Un altro fattore molto importante è la scelta di una adeguata sequenza di imaging (ad esempio, gradiente 2D oppure acquisizione 3D) [15].

### 3.2.3 Agente di contrasto

Diversi studi sulla risonanza magnetica al seno [8, 10] hanno evidenziato che l'esame di risonanza magnetica senza la somministrazione dell'agente di contrasto (plain MRI) apporta solo limitate informazioni in più rispetto ai metodi convenzionali di imaging (mammografia, ecografia). Ad esempio, in [8] viene osservato che un grande svantaggio della risonanza magnetica senza agente di contrasto è la sua incapacità di evidenziare le calcificazioni, sia benigne che maligne. Soltanto dopo l'introduzione degli agenti di contrasto, la risonanza magnetica ha assunto un ruolo molto importante nella diagnosi

del cancro al seno. Gli agenti di contrasto maggiormente utilizzati in questo esame sono le sostanze paramagnetiche, come ad esempio il gadolinio. Tali sostanze riducono i tempi di rilassamento T1 e T2, aumentando, così, l'intensità del segnale nelle sequenze T1-pesate [15]. In questo modo viene sensibilmente aumentato il contrasto tra tessuti normali e tessuti patologici, anche a bassissima concentrazione. Infatti, già il primo studio in cui fu utilizzato il gadolinio come mezzo di contrasto nella risonanza magnetica del seno riportò un forte enhancement nei tessuti patologici, mentre i tessuti normali mostravano solo un debole o nullo enhancement [22]. In conclusione, senza la somministrazione dell'agente di contrasto, un esame di risonanza magnetica al seno può evidenziare la densità dei tessuti, cisti, ematomi, impianti lesionati, dotti o linfonodi ingrossati. Con l'utilizzo del mezzo di contrasto, confrontando le immagini prima e dopo la sua somministrazione, l'esame di risonanza magnetica può determinare la presenza di anomalie al seno, se tali anomalie sono benigne o maligne, la dimensione e la posizione delle anomalie che appaiono come maligne.

#### 3.2.4 Pregi

**Sensitività.** Diversi studi hanno provato che la CE-MRI ha una eccellente sensitività (88%-100%): è in grado di individuare piccole lesioni che non sono state rilevate con altri esami, come la mammografia e l'ecografia [21].

**Applicabilità.** La risonanza magnetica ha un'applicabilità molto vasta: è in grado di individuare lesioni maligne anche in donne con tessuti del seno molto densi, con predisposizione genetica al cancro, con impianti di silicone e con cicatrici dovute a precedenti interventi chirurgici. Tutti questi casi costituiscono un limite per le tradizionali tecniche di imaging [24].

**Valutazione preoperativa.** Importante caratteristica della CE-MRI è la sua capacità di individuare tumori multicentrici e multifocali nello stesso seno o nel seno opposto. Essa può, inoltre, determinare l'estensione e la grandezza del tumore (staging) ed identificare la presenza e la grandezza di residui di tessuti patologici. Tutte queste informazioni si rivelano molto utili alla diagnosi, ad una valutazione preoperativa, che potrebbe comprendere il trattamento del paziente con chemioterapia preoperativa, e per l'elaborazione di un adeguato piano chirurgico [24, 38, 39].

### 3.2.5 Limiti

**Specificità.** Il maggior limite della risonanza magnetica è la sua bassa specificità, che varia tra il 37% e il 97%. Spesso la sua distinzione tra tumori maligni e tumori benigni non è attendibile a causa del fatto che alcune lesioni benigne hanno un grado di enhancement simile a quello delle lesioni maligne. Tale incertezza può condurre a risultati falsi positivi e, di conseguenza, ad inutili biopsie. Per questo motivo non è raccomandata come unico test di screening per le donne ad alto rischio di cancro al seno.

**Costo.** A causa dell'attrezzatura e del tempo necessari per produrre immagini di alta qualità, la risonanza magnetica è un esame molto costoso. Per questo motivo è indicato solo nei casi in cui le altre modalità di imaging non sono applicabili oppure presentano delle ambiguità nei risultati [15].

**Assenza di standardizzazione.** Una delle questioni maggiormente controverse per quanto riguarda l'utilizzo della risonanza magnetica al seno nella routine clinica è la mancanza di standardizzazione del protocollo di imaging sia in termini di tecnica che in termini di linee guida di interpretazione. Esiste una grande varietà di tecniche di MRI che portano a diversi criteri di interpretazione delle lesioni benigne e maligne. Questo rende difficile il consolidamento dei dati risultanti dagli esami. Un ulteriore problema è dovuto alla mancanza di un linguaggio standard nel descrivere dati morfologici e cinetici delle lesioni. Tale mancanza di consenso danneggia il confronto tra i dati di diversi studi, diminuisce l'applicabilità di ciascuna tecnica ed impedisce la valutazione scientifica dell'efficacia della MRI per il cancro al seno [25, 29].

### 3.2.6 Osservazioni

Come finora descritto, la risonanza magnetica, utilizzata in combinazione con l'agente di contrasto, è una modalità di imaging non invasiva che permette di evidenziare la fisiologia della microcircolazione all'interno della massa tumorale con un'ottima risoluzione spaziale. La microcircolazione si divide in tre compartimenti: la microvascolatura, lo spazio interstiziale o extravascolare extracellulare (EES), e lo spazio intracellulare. Gli agenti che raggiungono i tessuti o il tumore attraverso la circolazione devono passare attraverso questi compartimenti. La struttura dei tumori solidi è fondamentalmente caotica, a differenza della struttura anatomica ordinata degli organi e dei tessuti normali. La microcircolazione all'interno del tumore differisce profondamente da quella degli organi normali in tre modi: nelle caratteristiche del flusso

e, a volte, del volume di sangue della microvascolatura; nella permeabilità microvascolare; e, per molti tumori maligni, nell'aumento del volume dello spazio extravascolare extracellulare. Il tumore causa la formazione di nuovi piccoli vasi sanguigni e ne influenza la maturazione. Questa nuova rete di vasi sanguigni produce un flusso di sangue che è spazialmente e temporalmente più eterogeneo rispetto all'irroramento efficiente ed uniforme dei normali organi e tessuti. Inoltre il tumore differisce marcatamente rispetto ai normali organi circostanti nella permeabilità dei suoi capillari. Tale alterata permeabilità è importante perché cambia le regole che governano il trasferimento dei composti tra il sangue e il tessuto tumorale. Infine un'ulteriore anormalità della microcircolazione all'interno del tumore è data dalla marcata alterazione nel volume dei maggiori compartimenti del tessuto, con l'espansione del volume dello spazio extravascolare extracellulare [67]. L'angiogenesi, quindi, è un passo essenziale nello sviluppo del cancro invasivo, essa è necessaria per supportare la crescita del tumore. Perché un tumore cresca più di 1-2 mm, le cellule tumorali non solo devono proliferare, ma devono anche essere in grado di indurre la formazione e la crescita di nuovi vasi sanguigni [63]. L'angiogenesi può avere un ruolo essenziale anche nella diffusione del mezzo di contrasto nei vasi sanguigni. Diversi studi hanno dimostrato che la rapida distribuzione dell'agente di contrasto all'interno dei tumori maligni è dovuta all'aumento della densità microvascolare, con il conseguente aumento dell'irroramento dell'area, in combinazione con l'iperpermeabilità di questi nuovi vasi sanguigni. La maggiore permeabilità dei vasi sanguigni dei tessuti tumorali permette una più alta diffusione dell'agente di contrasto attraverso il muro capillare, producendo un più forte aumento dell'intensità del segnale di risonanza magnetica. La risonanza magnetica, quindi, con l'utilizzo del mezzo di contrasto, può evidenziare l'attività angiogenica del tumore fino alla rete dei capillari, dimostrandosi di grande utilità nella caratterizzazione delle lesioni sospette. Inoltre diversi studi hanno dimostrato una correlazione tra le caratteristiche dell'enhancement nella Contrast-Enhanced MRI e fattori istologici prognostici per il cancro al seno. Sulla base del fatto che l'angiogenesi supporta la crescita del cancro e le metastasi, può essere ipotizzato che un tumore con una più alta angiogenesi possa facilitare la diffusione delle cellule cancerogene e così rendere più probabile la metastasi ai linfonodi ed una successiva e precoce ricomparsa del tumore. La risonanza magnetica, quindi, può essere utile al fine di predire una prognosi, di pianificare trattamenti e di monitorare gli effetti di una terapia antiangiogenica [6, 2, 41, 3, 65].

### 3.3 Protocolli diagnostici

In Contrast-Enhanced MRI l'intero volume del seno è acquisito una volta prima della somministrazione del mezzo di contrasto e diverse volte dopo. Ogni acquisizione produce lo stesso numero di immagini. Di seguito viene introdotta l'organizzazione di tali acquisizioni, che viene poi utilizzata nei protocolli diagnostici e nel resto della tesi.

- Slice: sezione dell'intero volume del seno acquisito.
- Serie: sequenza ordinata delle immagini appartenenti ad ogni acquisizione.
- Serie pre-contrasto: la prima serie, cioè quella acquisita prima dell'iniezione del mezzo di contrasto. Questa serie viene anche definita serie morfologica e corrisponde ad un esame di risonanza magnetica *plain*, cioè senza l'utilizzo dell'agente di contrasto.
- Serie post-contrasto: serie acquisita dopo la somministrazione dell'agente di contrasto. Il numero di serie post-contrasto e l'intervallo di tempo che intercorre tra due serie consecutive dipendono dal particolare protocollo utilizzato nel centro clinico e dalla specifica procedura di acquisizione.
- Studio o esame: insieme ordinato di tutte le serie acquisite (su un paziente). Le immagini appartenenti ad uno stesso studio hanno le stesse dimensioni, cioè lo stesso numero di voxel, e la stessa profondità di colore, cioè lo stesso numero di livelli di grigio.

Per caratterizzare una lesione sospetta, quindi per determinare se tale lesione è benigna o maligna, è possibile utilizzare diversi insiemi di criteri [50]:

- criteri morfologici;
- criteri dinamici;
- combinazione di criteri.

#### 3.3.1 Criteri morfologici

I criteri morfologici studiano la forma e la struttura interna della lesione. Il primo modello predittivo costruito sulla base di feature architettoniche estratte dalle immagini ottenute dalla risonanza magnetica fu introdotto in [43] ed in seguito aggiornato in [42]. In uno studio condotto da un gruppo di

### 3. Contrast-Enhanced MRI

---

esperti di risonanza magnetica al seno (Lesion Diagnosis Working Group) è stato sviluppato e standardizzato un lessico comune per la descrizione delle lesioni [25]. Le lesioni possono essere suddivise in tre categorie principali:

- *Foci*: aree di enhancement che misurano meno di 5 mm di diametro e che non mostrano masse nelle immagini precontrasto;
- *Mass*: massa tridimensionale che occupa uno spazio all'interno del seno;
- *Non-mass-like enhancement*: area di enhancement che non presenta una massa tridimensionale rilevabile.

I descrittori della massa includono:

- *Forma*: circolare, ovale, lobulata (con contorni ondeggianti), irregolare, aghiforme (coperta di punti sulla superficie);
- *Margini*: regolari (circolari oppure ovali), lobulati, irregolari, aghiforme.
- *Enhancement*: omogeneo, eterogeneo, rim enhancement (principalmente concentrato ai margini della massa), setti interni, setti interni scuri (setti di non-enhancement in una massa di enhancement).

Per descrivere non-mass-like enhancement bisogna valutare la sua distribuzione, il suo enhancement interno e la sua simmetria. Essa viene suddivisa in:

- lineare,
- lineare-duttale,
- settoriale,
- regionale,
- diffusa,
- diffusa-disomogenea
- asimmetrica.

L'enhancement interno può essere definito: omogeneo, eterogeneo, punteggiato, raggruppato. I risultati pubblicati dal working group indicavano che, sebbene la forma delle lesioni fosse relativamente predittiva, gli esperti avevano difficoltà nel distinguere la forma dai margini; quindi alla fine la forma della lesione fu tolta dal lessico.



Da uno studio effettuato in [59] risulta che le caratteristiche maggiormente predittive di malignità sono i margini della massa e l'omogeneità dell'enhancement interno alla massa. Rim enhancement, anche se è una feature con bassa prevalenza, ha un alto valore predittivo positivo. Infine, la mancanza di enhancement è fortemente predittiva di benignità, anche se non esclude il cancro invasivo.

### 3.3.2 Criteri dinamici

I criteri dinamici studiano la diffusione e la concentrazione dell'agente di contrasto nei tessuti. Essi si basano sul fatto che i tumori sono caratterizzati da una diversa angiogenesi (formazione di nuovi vasi sanguigni) e permeabilità vascolare (sono caratterizzati da un'organizzazione vascolare anarchica) rispetto ai tessuti normali [4]. Questi ultimi non presentano enhancement, quindi il loro livello di grigio non cambia tra le immagini pre-contrasto e quelle post-contrasto. Nelle lesioni, sia maligne che benigne, invece, l'aumentata densità vascolare causa una maggiore concentrazione dell'agente di contrasto dovuta ad un maggior irroramento dell'area. Inoltre la maggiore permeabilità vascolare è responsabile di una più veloce fuoriuscita dell'agente di contrasto dai vasi sanguigni (washout) [30, 27]. Per questo motivo il primo passo nell'interpretazione di un esame di CE-MRI consiste nella ricerca di regioni di enhancement in ogni immagine delle serie post-contrasto, chiamate *Region Of Interest* (ROI). Tale ricerca viene eseguita sulle immagini sottratte, cioè immagini ottenute sottraendo da un'immagine post-contrasto la corrispondente immagine nella serie pre-contrasto. Dall'evoluzione nel tempo dell'intensità del segnale di risonanza, dovuta alla diffusione dell'agente di contrasto in tutte le serie post-contrasto, si ottengono le curve tempo-enhancement che poi possono essere analizzate sia quantitativamente che qualitativamente.

#### Analisi Quantitativa

L'analisi quantitativa consiste nel calcolo di parametri che descrivono il comportamento dell'intensità del segnale di risonanza magnetica nelle diverse fasi dell'esame (prima e seconda fase dopo la somministrazione dell'agente di contrasto). Di seguito sono spiegati i criteri dinamici quantitativi più diffusi.

**Relative enhancement** Relative enhancement, introdotto per la prima volta in [26], è uno dei criteri dinamici maggiormente utilizzato. Esso rappresenta l'incremento dell'intensità del segnale rispetto alla fase pre-contrasto:

$$E_k^{\%} = \frac{I_k(R) - I_0(R)}{I_0(R)} * 100 \quad k \in [1, n) \quad (3.1)$$

### 3. Contrast-Enhanced MRI

---

dove  $I_k$  è l'intensità del segnale di risonanza nell'immagine esaminata,  $k$  è l'indice della serie,  $I_0$  è l'intensità del segnale nella corrispondente immagine pre-contrasto,  $R$  è la Region Of Interest. Interpolando i valori ottenuti si ricava la curva di tempo-enhancement.

**Early-increase** Early-increase analizza il comportamento dell'intensità del segnale nella fase iniziale subito dopo la somministrazione dell'agente di contrasto, enhancement dopo  $s$  secondi dalla somministrazione, dove  $s$  di solito viene posto a 60 secondi oppure 80 secondi.

**Initial Slope** Questo criterio inizialmente fu introdotto in [5] con il nome di steepest slope e fu poi modificato in [64] in initial slope. Esso definisce la pendenza della curva di enhancement nel picco:

$$Slope_i = \frac{E_{peak}}{T_{peak}} \quad (3.2)$$

dove  $E_{peak}$  è il massimo relative enhancement e  $T_{peak}$  è il tempo (espresso in indice di serie) che trascorre dall'iniezione dell'agente di contrasto fino al picco nell'enhancement.

**Washout Ratio** Il parametro quantitativo washout ratio, descritto in [64], definisce la discesa della curva tempo-relative enhancement:

$$W_{peak-k} = \frac{I_{peak}(R) - I_k(R)}{I_{peak}(R)} * 100 \quad (3.3)$$

dove  $R$  è la Region Of Interest,  $k$  è l'indice di una serie post-contrasto e  $peak$  è l'indice dell'immagine con la massima intensità del segnale. Questo parametro descrive il decremento dell'intensità del segnale dal punto di picco fino al punto scelto ( $k$ ).

**Curvature at the peak of enhancement** Questo criterio è stato introdotto in [11] per distinguere tra le curve che hanno il picco dell'enhancement nello stesso momento ma hanno comportamenti di washout completamente differenti. Le curve, al picco dell'enhancement, possono piegarsi in maniera più o meno brusca, e questo può essere descritto matematicamente con la curvatura. In questo studio la concentrazione dell'agente di contrasto viene espressa come una funzione continua del tempo calcolata su una piccola Region Of Interest contenente i pixel che presentano il più rapido e forte enhancement.

$$k(t) = \frac{\left| \frac{d^2 C}{dt^2} \right|}{\left( 1 + \left( \frac{dC}{dt} \right)^2 \right)^{3/2}} = \frac{|C(D^2 + \frac{dD}{dt})|}{(1 + (CD)^2)^{3/2}} \quad (3.4)$$

dove:

- $t = T_{peak}$ ,
- $C$  è la curva della concentrazione dell'agente di contrasto:

$$C(t) = A(1 - e^{-\alpha t})^q e^{-\beta t} \frac{1 + e^{-\gamma t}}{2} \quad (3.5)$$

dove  $A$  è il limite superiore della concentrazione del mezzo di contrasto,  $\alpha$  è il tasso di assorbimento dell'agente di contrasto,  $\beta$  è la percentuale complessiva di washout dell'agente di contrasto,  $\gamma$  è il tasso iniziale di washout e  $q$  è correlato allo slope dell'assorbimento iniziale e alla curvatura della transizione dall'assorbimento al washout.

- $\frac{dC(t)}{dt} = C(t)D(t)$
- $D(t) = \frac{q\alpha e^{-\alpha t}}{1 - e^{-\alpha t}} - \beta - \frac{\gamma e^{-\gamma t}}{1 + e^{-\gamma t}}$

### Analisi Qualitativa

L'analisi qualitativa studia la morfologia della curva di relative enhancement. Uno dei criteri più utilizzati è stato descritto in [28]. Esso classifica le lesioni in tre categorie differenti in base alla forma della curva di relative enhancement ad esse associata. I tre tipi di curve differiscono nell'intensità nelle fasi post-contrasto intermedia e finale (Figura 3.1):

- Tipo I. Curve steady. L'intensità del segnale continua a crescere nel tempo, oppure si appiattisce nell'ultimo periodo post-contrasto per effetti di saturazione.
- Tipo II. Curve plateau. Dopo una crescita iniziale, l'enhancement si interrompe bruscamente e nelle fasi post-contrasto intermedia e finale l'intensità del segnale rimane costante nel tempo.
- Tipo III. Curve washout. Dopo una crescita iniziale, l'enhancement si blocca e l'intensità del segnale decresce nelle fasi post-contrasto intermedia e finale.

Il comportamento dinamico dell'enhancement, analizzato sulla base delle curve tempo-enhancement, differisce significativamente tra lesioni benigne e maligne, quindi fornisce informazione molto utile dal punto di vista diagnostico. Le curve di tipo I sono considerate indicative di una lesione benigna, le curve di tipo III sono considerate dei forti indicatori di lesione maligna, mentre le curve di tipo II si riscontrano sia in lesioni benigne che maligne, per cui vengono considerate incerte o sospette di malignità.

### 3.3.3 Combinazione di criteri

Esistono schemi di classificazione di una regione di enhancement che combinano l'informazione ricavata da entrambi i criteri, morfologici e dinamici. Uno degli schemi più diffusi è lo schema di Fischer, descritto in [13]. In questo studio, una lesione viene analizzata in base a tutte le caratteristiche dei diversi criteri. Ad ogni caratteristica viene assegnato un punteggio (0-2 punti) in base a quanto essa sia indicativa di malignità. Nella Tabella 3.1 viene presentato il sistema a punteggio di Fischer. Early-increase rappresenta l'incremento dell'intensità del segnale nei primi 3 minuti dopo la somministrazione dell'agente di contrasto rispetto alla fase pre-contrasto. La morfologia della curva è considerata continua se l'incremento dell'intensità del segnale dopo i primi 3 minuti dalla somministrazione dell'agente di contrasto è maggiore del 10% relativamente alla fase di early-increase; è considerata plateau se l'intensità del segnale ha una deviazione del  $\pm 10\%$  rispetto alla fase di early-increase; è considerata washout se l'intensità del segnale decresce di più del 10% relativamente alla fase di early-increase.

Criterio/Punti	0	1	2
Early-increase	< 50%	50% – 100%	> 100%
Morfologia curva	Continua(tipo I)	Plateau(tipo II)	Washout(tipo III)
Margini	Ben definiti	Mal definiti	-
Forma	Circolare,ovale, lobulare	Lineare,dendritica, stellata	-
Enhancement	Omogeneo	Non omogeneo	Ad anello

Tabella 3.1: Schema di Fischer.

Infine il numero totale di punti per tutti i criteri viene sommato per ottenere la classificazione finale. Indicando con  $s$  la somma totale, la classificazione della lesione si ricava in questo modo:

$$0 \leq s \leq 2 \rightarrow \textit{benigna}$$

$s = 3 \rightarrow \textit{incerta}$

$s \geq 4 \rightarrow \textit{maligna}$

## 3.4 Parametri dell'esame

Come introdotto nella sezione 3.2 di questo capitolo, per la risonanza magnetica al seno non esiste ancora una tecnica standard, ma vengono utilizzati diversi protocolli e procedure di acquisizione, in base all'attrezzatura hardware e software disponibile, alle indicazioni cliniche ed ai risultati desiderati. Nel resto della sezione vengono presentati i diversi parametri che caratterizzano un particolare protocollo e viene indicato il loro valore nei dati utilizzati nello studio di questa tesi, che sono stati forniti dal Dipartimento di Radiologia dell'Università di Pisa.

### Sequenza di impulsi

Una successione di impulsi a radiofrequenza e di applicazione di gradienti costituisce le sequenze di impulsi. Gli impulsi a radiofrequenza hanno la funzione di eccitare i protoni e di generare il segnale di risonanza. Gli impulsi di gradiente influenzano la frequenza e la fase del segnale ricevuto e sono necessari per ottenere l'immagine.

Il dataset a disposizione è stato acquisito con uno scanner General Electric 1.5T Signa Contour che utilizza sequenze di impulsi “*3D fast spoiled gradient echo sequences (FSPGR)*” senza soppressione del segnale proveniente dai tessuti adiposi.

### Livelli di grigio

Il livello di grigio di ogni voxel delle immagini a disposizione è memorizzato su 9 bit, quindi il numero di livelli di grigio è  $2^9 = 512$ .

### Risoluzione delle immagini

La risoluzione, cioè il numero di voxel (elemento di volume) dell'immagine, è la stessa per tutte le immagini a disposizione ed è pari a 256x256 voxel. Gli attuali dispositivi possono raggiungere risoluzioni di 512x512 o 1024x1024 voxel.

#### **Numero di serie**

Ogni esame è composto da sei serie di immagini: una serie pre-contrasto e cinque serie post-contrasto, acquisite a 0, 2, 4, 6 e 8 minuti dopo la somministrazione dell'agente di contrasto.

**Numero di immagini per serie** Il numero di immagini per serie non è lo stesso per tutti gli esami, ma dipende dalla dimensione del volume che deve essere acquisito.

Le serie degli esami utilizzati in questo studio comprendono dalle 30 alle 60 immagini per serie. Tutte le serie di uno stesso studio hanno lo stesso numero di immagini.

#### **Risoluzione temporale**

La risoluzione temporale è l'intervallo di tempo che intercorre tra le acquisizioni di due serie differenti. Essa dovrebbe essere sufficiente a permettere l'acquisizione sia della fase di washin che della fase di washout dell'agente di contrasto.

Nei dati forniti dal Dipartimento di Radiologia dell'Università di Pisa l'intervallo di tempo tra due diverse scansioni è fissato a 2 minuti.

#### **Risoluzione spaziale**

La risoluzione spaziale è definita dalla grandezza di ogni pixel che compone l'immagine e dal volume acquisito. In [23] vengono elencati dei requisiti che un'immagine di risonanza magnetica dovrebbe avere affinché tale tecnica risulti efficace in campo clinico:

- **Slice thickness:** ha un grande impatto sull'accuratezza della Contrast-Enhanced MRI. Se una lesione occupa solo una parte dello spessore della slice, soltanto questa percentuale di enhancement sarà visibile o misurabile. I tumori al seno più larghi di 5 mm dovrebbero essere diagnosticati con un'alta affidabilità ed inoltre diversi tumori nello stadio iniziale crescono lungo i dotti con un diametro di 1-2 mm. Viene, quindi, raccomandato che le slice abbiano uno spessore di circa 2-3 mm.
- **Gap:** è importante che non ci siano spazi vuoti tra le slice, in modo che l'intero volume del seno sia acquisito per assicurare che le eventuali lesioni siano completamente visibili.

- In-plane spatial resolution: la risoluzione spaziale sul piano di acquisizione dovrebbe essere inferiore ad 1 mm in modo da permettere la valutazione della struttura interna di una lesione e dei suoi bordi.

I dati utilizzati in questa tesi hanno le seguenti caratteristiche:

- Slice thickness: 3 mm;
- nessun gap tra immagini successive e adiacenti;
- grandezza di ogni pixel: 1.5mm x 1.5mm.

Dunque la dimensione di ogni voxel dell'immagine è di 1.5mm x 1.5mm x 3mm.

### Piano di acquisizione

I possibili piani di acquisizione sono il piano coronale, il piano trasversale ed il piano sagittale (Figura 3.2).

- *Piano coronale o piano frontale.* Il piano coronale è un piano verticale che corre parallelo alla fronte e divide il corpo in parte anteriore e parte posteriore. L'acquisizione sul piano coronale non visualizza bene patologie duttali in quanto potrebbe troncare i dotti lattiferi. Inoltre tale piano di acquisizione può sacrificare nella visualizzazione anomalie lineari e i tessuti della parte estrema posteriore del seno, in particolare i tessuti della coda ascellare. Un ulteriore problema è il fatto che la maggior parte dei vasi sanguigni viaggiano nel piano coronale, così può diventare difficile distinguere una lesione focale dalla sezione di una vena. Il vantaggio dell'orientazione coronale è la possibilità di acquisire entrambi i seni all'interno di una stessa immagine e quindi di ridurre il tempo di acquisizione [39].
- *Piano trasversale o piano assiale.* Il piano trasversale è un piano orizzontale che divide il corpo in due metà, una superiore ed una inferiore. L'acquisizione nel piano trasversale visualizza bene i dotti lattiferi ed ha il vantaggio di corrispondere alle proiezioni ottenute dalla mammografia per cui è possibile una diretta correlazione. Inoltre i movimenti causati dal respiro occorrono soprattutto nel piano coronale, quindi con l'acquisizione nel piano trasversale si troveranno all'interno del piano di una data sezione, senza influenzare particolarmente il risultato dell'acquisizione [39].

- *Piano sagittale.* Il piano sagittale è un piano verticale che decorre in senso antero-posteriore e divide il corpo in due parti, parte destra e parte sinistra. L'acquisizione nel piano sagittale visualizza bene i dotti lattiferi e, come per il piano trasversale, ha il vantaggio di corrispondere alle proiezioni ottenute dalla mammografia per cui è possibile una diretta correlazione. Tale piano, però, permette l'acquisizione di immagini di un solo seno alla volta [39].

Gli esami a disposizione sono stati acquisiti lungo il piano coronale.

#### **Formato delle immagini**

Le immagini sono memorizzate in formato digitale secondo lo standard DICOM, che fornisce i criteri per il trattamento, la memorizzazione, stampa e comunicazione di informazioni di tipo biomedico.



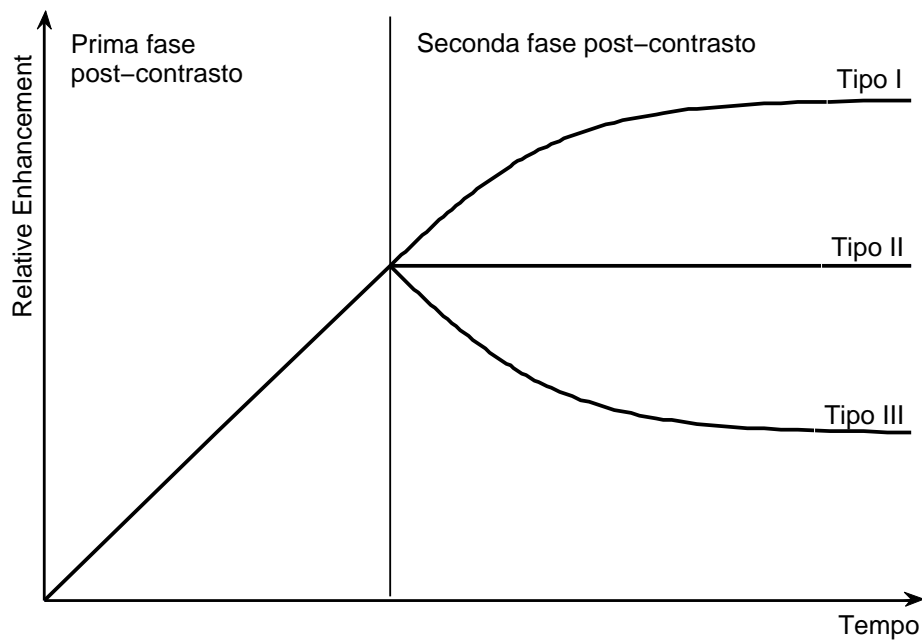


Figura 3.1: Classificazione schematica delle curve tempo-enhancement. Curve di tipo I si trovano principalmente nelle lesioni benigne. Curve di tipo III si trovano soprattutto nelle lesioni maligne. Curve di tipo II si trovano in entrambi i casi, per questo sono incerte e possono, però, essere utilizzate come supporto al sospetto di malignità.

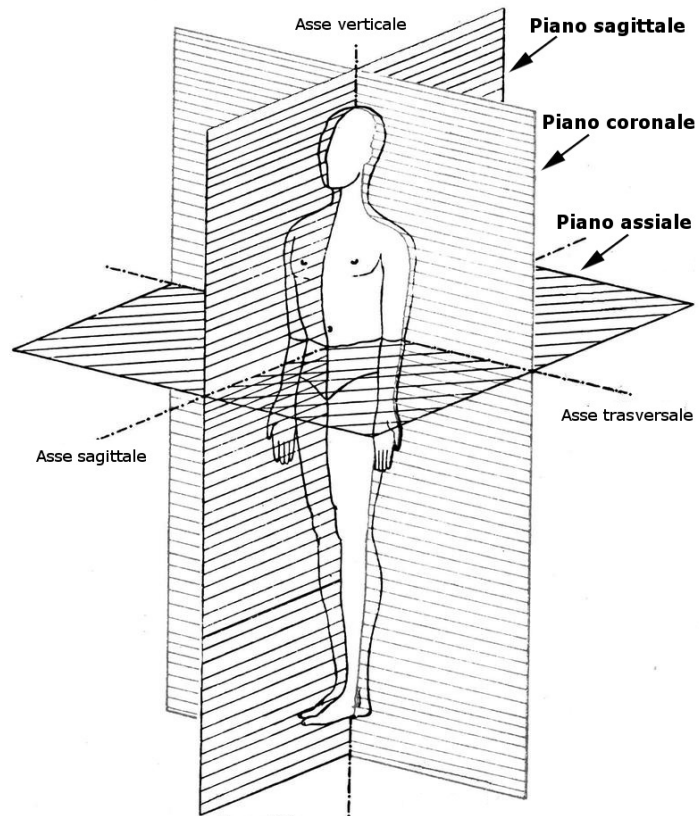


Figura 3.2: Piani di acquisizione.

# Capitolo 4

## Analisi automatica

La tecnica Contrast-Enhanced MRI comporta l'acquisizione di un largo numero di immagini ad alta risoluzione per fornire informazioni sulle caratteristiche dinamiche e morfologiche delle lesioni sospette. Le attuali tecniche di acquisizione delle immagini producono dataset contenenti centinaia di immagini e tale numero è continuamente in crescita con gli sviluppi tecnologici. L'analisi di un così elevato numero di immagini, necessaria per eseguire una diagnosi, richiede molto tempo ed inoltre è soggetta ad errori a causa della stanchezza e della distrazione del radiologo. Per ridurre la probabilità di errore di solito si sottopone un esame ad una doppia lettura da parte di due radiologi differenti. I radiologi analizzano l'esame in modo indipendente l'uno dall'altro e poi confrontano i risultati. Tale approccio comporta un'alta variabilità inter-osservatore dovuta alla scarsa standardizzazione di linee guida di interpretazione del protocollo: a causa della complessa struttura radiografica del seno e della difficile caratterizzazione del cancro al seno nelle sue fasi iniziali, ogni radiologo, seppur esperto, si concentrerà solo su alcune delle caratteristiche disponibili in base alla propria esperienza. L'implementazione di strumenti per la valutazione automatica dei dati, sistemi *Computer-Aided Diagnosis* (CAD), può avere un importante valore pratico nella minimizzazione di tutti questi problemi. Programmi CAD possono essere di grande aiuto ai medici nell'ispezione dei dati, quindi localizzazione e visualizzazione delle lesioni, in quanto permettono un'osservazione più rapida ed efficace delle immagini che costituiscono ogni esame. Inoltre essi, analizzando tutte le immagini in un tempo molto ridotto rispetto a quello necessario ad un radiologo, possono fornire una propria diagnosi che può rappresentare per i medici una seconda opinione da tenere in considerazione nella classificazione delle lesioni sospette. L'utilizzo di sistemi CAD nella pratica clinica ha, quindi, diversi vantaggi. L'analisi delle immagini mediche risulta essere maggiormente obiettiva in quanto è minimizzata l'influenza di fattori sog-

gettivi come l'esperienza e la stanchezza del radiologo. Inoltre è evitata la variabilità dell'interpretazione radiologica causata dalla doppia lettura delle immagini da parte di medici differenti. L'efficienza dell'analisi aumenta in quanto viene notevolmente ridotto il tempo necessario per l'elaborazione di un intero dataset, velocizzando così anche la formulazione di un'eventuale ulteriore diagnosi con cui il medico può confrontarsi per migliorare i risultati nella distinzione tra lesioni maligne e lesioni benigne.

In questo capitolo vengono presentati i sistemi CAD realizzati ed utilizzati nello studio di questa tesi per l'analisi di immagini di risonanza magnetica del seno e per la diagnosi del cancro al seno. La sezione 4.1 descrive l'applicazione realizzata per visualizzare le immagini di risonanza magnetica, per selezionare le zone di enhancement da processare e per visualizzare la diagnosi ottenuta per una data lesione sospetta. La sezione 4.2 descrive i classificatori che saranno poi testati nelle sperimentazioni per classificare le curve di relative enhancement e gli altri criteri dinamici precedentemente descritti.

## 4.1 Applicazione

Le immagini di risonanza magnetica sono memorizzate in formato digitale secondo lo standard DICOM, che fornisce i criteri per incapsulare i dati e per definire come questi debbano essere codificati o interpretati. L'applicazione MRM-Works è costituita da un'interfaccia grafica attraverso la quale il medico ha la possibilità di visualizzare le immagini di risonanza magnetica di un intero esame e di selezionare all'interno di tali immagini delle zone che ritiene sospette. Durante lo studio di questa tesi l'applicazione è stata utilizzata per costruire il dataset poi analizzato in fase di sperimentazione (la raccolta e la selezione dei dati sarà in seguito descritta nella sezione 5.2). Infine, l'applicazione MRM-Works permette al medico di richiedere la classificazione di una data lesione e di visualizzarne i risultati.

### 4.1.1 Linguaggio di programmazione e libreria grafica

Il linguaggio di programmazione utilizzato per l'implementazione dell'applicazione MRM-Works è il linguaggio Java. I sistemi informativi ospedalieri sono molto eterogenei, per cui si presenta la necessità per i medici di poter eseguire l'applicazione sulle diverse macchine che hanno a disposizione. Il linguaggio Java è adatto a questo scopo grazie alla sua caratteristica di essere indipendente dalla piattaforma, quindi grazie alla sua portabilità su piattaforme hardware e sistemi operativi diversi.

Per la visualizzazione e l'elaborazione delle immagini, in questa tesi viene utilizzata un'estensione della piattaforma Java, la *Java Advanced Imaging (JAI) Application Program Interface (API)*. Tale estensione fornisce un insieme di interfacce orientate agli oggetti che consente la lettura, la visualizzazione e la manipolazione di diversi formati di immagine. JAI è un insieme di classi ed operatori che permette di inserire all'interno di programmi Java elaborazione sofisticata di immagini ottenendo alte prestazioni. Inoltre, l'estensione JAI, come il linguaggio Java, è indipendente dalla piattaforma, quindi rispetta la necessità dei medici di poter eseguire l'applicazione su qualsiasi computer. All'interno di questa tesi, JAI è stata utilizzata per visualizzare le immagini di risonanza magnetica e per realizzare alcune funzionalità di modifica delle caratteristiche delle immagini stesse con lo scopo di migliorarne la visibilità e, di conseguenza, l'analisi. In particolar modo, per mezzo delle classi e degli operatori di JAI, sono state implementate le procedure di *zoom* e *windowing* (modifica delle caratteristiche di luminosità/contrasto dell'immagine). Quest'ultima funzionalità, che con la combinazione dei due parametri *Window Width* e *Window Level* oppure *Window Center* determina l'intervallo di rappresentazione dei livelli di grigio, è molto importante nell'ambito delle immagini di risonanza magnetica. Essa, infatti, non solo permette di evidenziare le zone di maggiore enhancement, ma serve anche per la visualizzazione stessa delle immagini su schermi con 8 bit per colore, in quanto le immagini di risonanza magnetica sono ottenute con una codifica del segnale di risonanza su un numero di bit superiore. L'implementazione di tali strumenti migliora la qualità di analisi di ogni singola immagine e velocizza l'ispezione di tutte le immagini appartenenti ad ogni esame.

### 4.1.2 Interfaccia grafica

L'interfaccia grafica permette l'interazione con l'utente per la visualizzazione e l'analisi delle immagini. In Figura 4.1 viene mostrata l'interfaccia durante l'ispezione di un esame.

Come si vede dalla figura, nell'interfaccia è visualizzata un'immagine pre-contrasto e l'immagine sottratta, cioè l'immagine ottenuta sottraendo da un'immagine post-contrasto la corrispondente immagine della serie pre-contrasto. Attraverso l'interfaccia è possibile analizzare un intero esame, scorrendo tutte le serie e tutte le immagini per ogni serie. La figura mostra l'immagine 21 della serie pre-contrasto e la corrispondente immagine sottratta della terza serie post-contrasto. Per migliorare la visibilità delle immagini, sono presenti le funzionalità di zoom e windowing. Nella figura sono visualizzate le immagini appartenenti ad un esame diagnosticato istologicamente come carcinoma. Modificando i parametri *Window Width* e *Window Cen-*

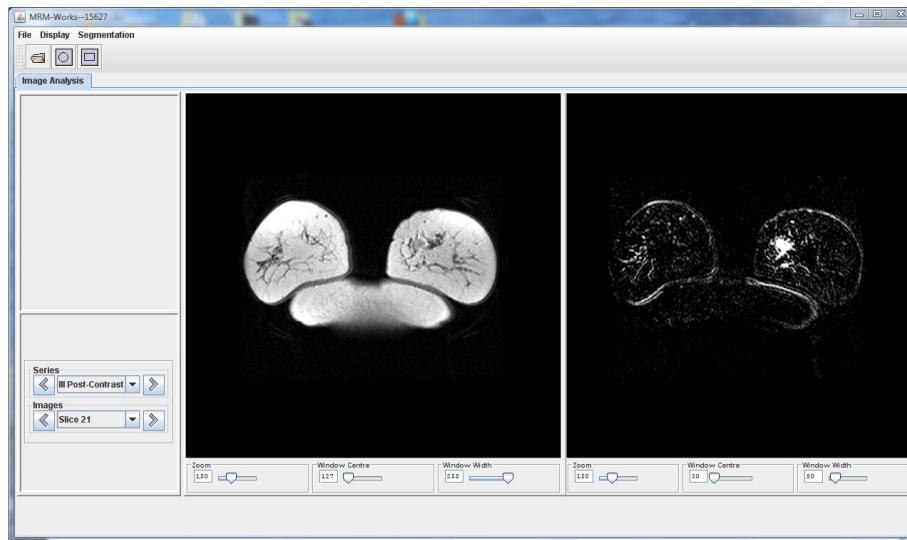


Figura 4.1: Interfaccia grafica MRM-Works. Immagini visualizzate: Slice 21; Serie post-contrasto 3. Parametri immagine pre-contrasto: Zoom 150; Window Center 127; Window Width 255. Parametri immagine sottratta: Zoom 150; Window Center 30; Window Width 50.

ter sull'immagine sottratta, viene evidenziato il carcinoma (zona di più forte enhancement, che appare bianca) nel quadrante sinistro. Come sarà meglio descritto nella sezione 5.2, tale interfaccia è stata utilizzata per la preparazione del dataset in quanto fornisce funzionalità di selezione ed estrazione, all'interno delle immagini, delle zone di maggiore enhancement. La Figura 5.2 mostra una regione di enhancement selezionata manualmente per mezzo dello strumento "Select Zone" dell'interfaccia. La Figura 5.3, invece, mostra la stessa area di enhancement estratta con la tecnica del region growing, che verrà descritta nella sezione 4.1.3, mediante lo strumento "Extract Region" dell'interfaccia. Sui dati ricavati da queste regioni di enhancement sono stati calcolati i vari criteri dinamici poi utilizzati nelle sperimentazioni. Le funzionalità Select Zone ed Extract Region oltre a contribuire alla costruzione del dataset di questa tesi, possono agevolare il compito del radiologo nel caso in cui questi volesse aggiungere nuovi dati ad un dataset su cui effettuare ulteriori valutazioni. Il radiologo, inoltre, ha la possibilità, sempre per mezzo dell'interfaccia, di associare ai dati selezionati la loro diagnosi istologica per eventuali confronti successivi. L'interfaccia offre anche la possibilità di calcolare e visualizzare la curva di relative enhancement su una Region Of Interest (ROI) di dimensione  $3 \times 3$  voxel per mezzo dello strumento "Square ROI". La Figura 4.2 mostra solo una sezione dell'interfaccia grafica, in cui

è visualizzata, sull'immagine sottratta a cui appartiene, la curva di relative enhancement di una ROI selezionata nella regione del carcinoma. In rosso sono indicate le coordinate, all'interno dell'immagine, del voxel centrale della ROI di dimensione  $3 \times 3$ .

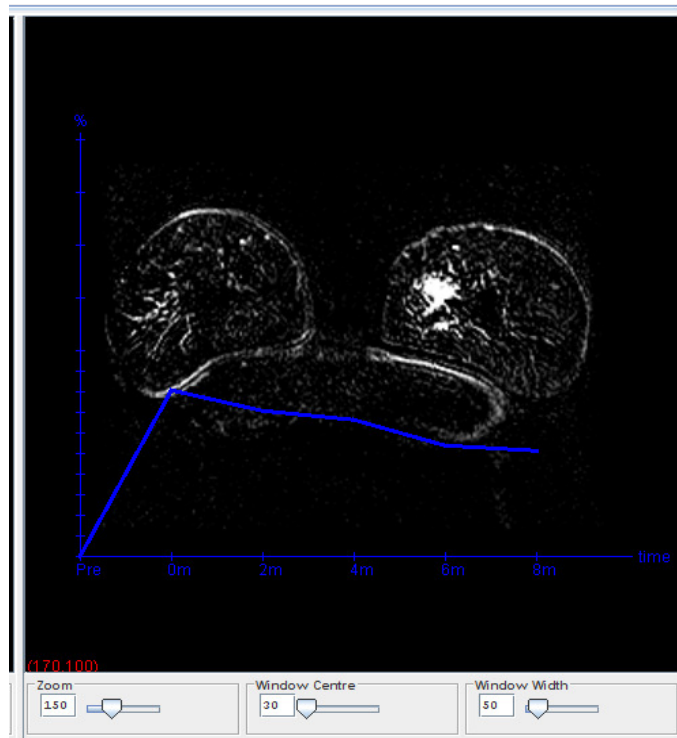


Figura 4.2: Pannello dell'immagine sottratta con visualizzata la curva di relative enhancement di una ROI di dimensione  $3 \times 3$  voxel all'interno del carcinoma.

Il radiologo può chiedere la classificazione della curva di relative enhancement visualizzata. Tale curva diventerà rossa se la classificazione è positiva, verde se è negativa. In Figura 4.3 sono mostrati due esempi di curve di relative enhancement classificate per mezzo dell'applicazione. La Figura 4.3a rappresenta la stessa curva della Figura 4.2, corrispondente ad una ROI selezionata nella regione del carcinoma (tumore maligno), colorata di rosso ad indicare che è stata classificata come positiva. La classificazione risultante è dunque corretta. La Figura 4.3b rappresenta, invece, una curva di relative enhancement di una ROI selezionata su un'immagine di un esame diagnostico istologicamente come displasia (tumore benigno). Anche in questo caso, con la curva colorata di verde, la classificazione ottenuta per mezzo dell'applicazione risulta essere corretta.

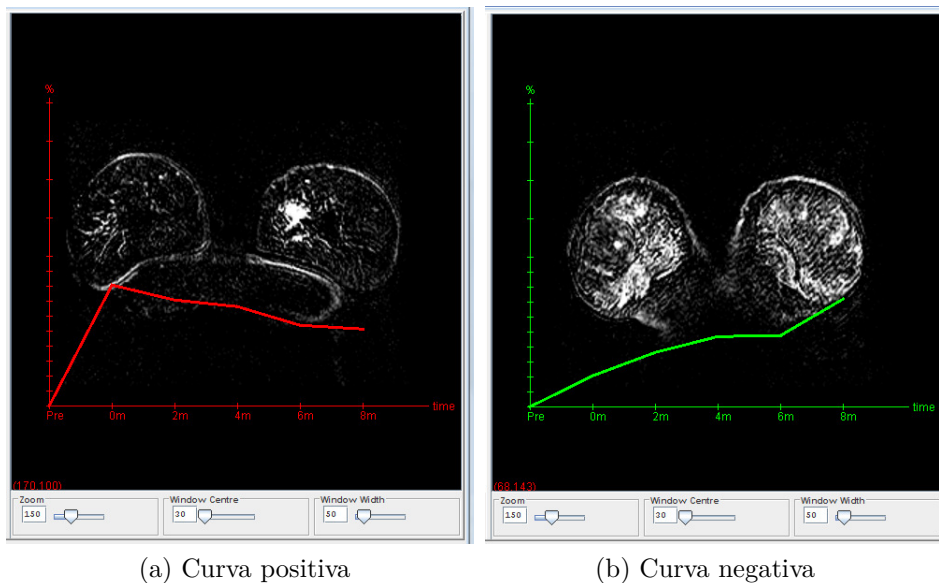


Figura 4.3: Curve di relative enhancement classificate utilizzando MRM-Works

### 4.1.3 Region Growing

La segmentazione è un aspetto importante nell'elaborazione delle immagini mediche. Essa costituisce la base dell'estrazione di regioni patologicamente interessanti dalle immagini originali, con lo scopo di ottenere maggiore informazione diagnostica, utile anche per l'analisi e il trattamento della malattia. Il processo di segmentazione consiste nel suddividere un'immagine in regioni separate, in modo che ogni regione sia omogenea rispetto ad una particolare proprietà (intensità, colore, etc.) [17]. Le tecniche di segmentazione delle immagini possono essere suddivise nelle seguenti classi principali.

*Metodi basati su soglia.* L'immagine è composta da un numero di oggetti di intensità costante, ben separati dallo sfondo. Tali metodi utilizzano una soglia determinata dall'istogramma dei livelli di intensità di grigio dell'immagine. Questi metodi funzionano bene solo quando i livelli di grigio degli oggetti sono chiaramente distinti dai livelli di grigio dello sfondo [19].

*Metodi basati sull'estrazione dei bordi.* Tali metodi si basano sulla localizzazione di variazioni significative nei livelli di grigio dell'immagine (bordi) [75]. L'immagine ottenuta dall'estrazione dei bordi deve poi subire ulteriori passi di elaborazione, in cui i bordi selezionati vengono concatenati per formare i contorni chiusi degli oggetti.

*Metodi basati su regioni.* L'obiettivo di tali metodi è partizionare l'imma-



gine in regioni che soddisfano un criterio di omogeneità prestabilito [19]. A differenza dell'approccio precedente, che cerca i bordi tra le regioni, questa tecnica costruisce le regioni direttamente.

*Metodi basati su modelli deformabili.* I *modelli deformabili* sono curve o superfici definite sull'immagine che possono muoversi sotto l'influenza di forze interne, definite all'interno della curva o superficie stessa, e di forze esterne, calcolate dall'immagine. L'influenza combinata di tali forze rende i modelli deformabili robusti rispetto ad eventuale rumore nell'immagine ed interruzioni nei bordi [74].

*Metodi basati su modelli markoviani.* Si assume che l'immagine sia una realizzazione di un *Markov random field* con una distribuzione che cattura il contesto spaziale dell'immagine [32]. Data la distribuzione a priori dell'immagine, il problema di segmentazione è formulato come un problema di ottimizzazione.

*Metodi ibridi.* La maggior parte di questi metodi riguarda l'integrazione di tecniche basate sull'estrazione dei bordi e tecniche basate su regioni [19, 47].

La tecnica del *region growing*, utilizzata in questa tesi per estrarre le zone di più forte enhancement, è un metodo di segmentazione delle immagini appartenente alla classe dei metodi basati su regioni. Tale procedura si basa sull'accrescimento delle regioni mediante progressiva fusione di pixel o regioni adiacenti. L'approccio base è costituito da due fasi:

- Inizializzazione: si sceglie un insieme di punti o regioni base, detti “*seed*”, da cui iniziare la costruzione delle regioni.
- Accrescimento iterativo: ad ogni seed vengono aggiunti i pixel vicini che hanno proprietà simili al seed (ad esempio rientrano in uno specifico intervallo di livelli di grigio o colore). Questo processo è ripetuto finché nessun pixel può più essere aggiunto alla regione. Infine, regioni adiacenti simili possono essere fuse se la regione risultante soddisfa il criterio di omogeneità.

La scelta dei seed è un passo chiave in questa tipologia di segmentazione, in quanto il risultato dell'algoritmo dipende dalla selezione dei punti iniziali da cui parte il processo di crescita della regione.

L'algoritmo di *region growing* implementato in questa tesi prende spunto dall'algoritmo proposto in [73]. Viene scelta manualmente la zona di maggiore enhancement che si vuole estrarre. Il seed iniziale è una ROI di dimensione  $3 \times 3$  centrata nel punto selezionato con il mouse. Come misura di similarità si considera la media dei valori di grigio della regione costruita fino a quel punto. La regione viene accresciuta iterativamente aggiungendo

ad essa i pixel vicini che soddisfano il criterio di similarità, e ad ogni iterazione viene aggiornata la media dei valori di grigio in base ai nuovi pixel aggregati. Al termine del processo, quando nessun pixel adiacente può più essere aggiunto, vengono applicati alla regione costruita operatori di morfologia matematica per riempire eventuali spazi vuoti, probabilmente causati da rumore, all'interno della regione estratta. Prima di descrivere l'algoritmo in maniera dettagliata, viene introdotto il concetto di morfologia matematica e i relativi operatori utilizzati [18, 45].

La morfologia matematica si basa sull'algebra di operatori non lineari che agiscono sulla forma degli oggetti. La sua assunzione iniziale è che le immagini reali possono essere modellate con insiemi di punti ed essa opera sulle proprietà di tali insiemi, sulla loro connettività e sulla loro forma. Una trasformazione morfologica  $\Psi$  è data dalla relazione dell'immagine (insieme di punti  $X$ ) con un altro piccolo insieme di punti  $B$ , chiamato elemento strutturale.  $B$  è espresso rispetto ad un punto di origine locale,  $O$ , chiamato punto rappresentativo. Applicare la trasformazione morfologica  $\Psi(X)$  all'immagine  $X$  vuol dire muovere l'elemento strutturale  $B$  sistematicamente sull'intera immagine. La morfologia matematica è nata per operare su immagini binarie, ed in seguito i suoi operatori sono stati estesi per trattare anche i livelli di grigio. Si consideri lo spazio 2D di tutti gli interi,  $E = Z^2$ , di cui le immagini binarie sono sottoinsiemi; matematicamente un pixel è un punto in tale spazio. Le operazioni morfologiche primarie sono:

- *Dilatazione.* La trasformazione morfologica dilatazione  $\oplus$  consiste nell'aggiunta di un dato pattern ad una piccola area. Essa combina due insiemi usando la somma di vettori. La dilatazione  $X \oplus B$  è l'insieme di punti di tutte le possibili somme di vettori di coppie di elementi, ognuno dei quali appartenente rispettivamente agli insiemi  $X$  e  $B$ .

$$X \oplus B = \{p \in E : p = x + b, x \in X \wedge b \in B\} \quad (4.1)$$

La forma di  $B$  definisce la natura della dilatazione. Questo tipo di trasformazione morfologica è utilizzata per riempire piccoli buchi e spazi vuoti negli oggetti. Essa aumenta la dimensione dell'oggetto.

- *Erosione.* La trasformazione morfologica erosione  $\ominus$  consiste nell'eliminazione di un insieme di pixel aventi un dato pattern. Essa combina due insiemi usando la sottrazione di vettori tra gli elementi dei due insiemi.

$$X \ominus B = \{p \in E : p + b \in X, \forall b \in B\} \quad (4.2)$$

L'erosione è utilizzata per semplificare la struttura di un oggetto: oggetti o loro parti con larghezza unitaria spariscono. Essa è, inoltre, usata per decomporre oggetti complicati in parti più semplici.

Dalla combinazione di queste operazioni base scaturiscono le seguenti operazioni morfologiche più complesse.

- *Apertura.* Tale trasformazione morfologica consiste nell'applicazione di erosione seguita da dilatazione utilizzando lo stesso elemento strutturale. L'apertura di un'immagine  $X$  per mezzo dell'elemento strutturale  $B$  è denotata da  $X \circ B$  ed è definita come

$$X \circ B = (X \ominus B) \oplus B \quad (4.3)$$

Quest'operazione tende ad aprire piccoli spazi tra gli oggetti adiacenti in un'immagine. Dopo l'applicazione dell'apertura gli oggetti sono isolati meglio e possono essere contati o classificati.

- *Chiusura.* Tale trasformazione morfologica consiste nell'applicazione di dilatazione seguita da erosione usando lo stesso elemento strutturale. La chiusura di un'immagine  $X$  per mezzo dell'elemento strutturale  $B$  è denotata da  $X \bullet B$  ed è definita come

$$X \bullet B = (X \oplus B) \ominus B \quad (4.4)$$

Quest'operazione connette oggetti che sono vicini tra di loro, riempie piccoli buchi e regolarizza i contorni degli oggetti.

Viene ora presentato l'algoritmo di region growing implementato in questa tesi (Algoritmo 1).

## 4.2 Classificatori

Sono stati proposti diversi modelli per classificare i tumori al seno che analizzano i vari criteri introdotti nel capitolo 3. In [7] gli autori presentano una sperimentazione di due classificatori a soglia per classificare i criteri dinamici initial slope e washout ratio, e una rete neurale multistrato, allenata con l'algoritmo della back-propagation, per classificare le curve di relative enhancement. In [1], [34], [9], [69] e [66] gli autori allenano dei multilayer perceptron per classificare caratteristiche dinamiche (le curve tempo-intensità del segnale e diverse caratteristiche estratte da tali curve) e/o caratteristiche statiche (morfologiche) estratte dalle ROI. In [60] gli autori propongono un metodo basato su alberi di decisione costruiti utilizzando caratteristiche dinamiche, sia qualitative che quantitative, e caratteristiche morfologiche estratte dalle ROI. In [68] gli autori confrontano diversi metodi di classificazione dei tumori sulla base delle curve tempo-intensità del segnale e di altri

---

**Algorithm 1** Region growing

---

```
Input: seed, coeff.  
// seed è il seed iniziale da cui parte il processo di crescita della regione  
// coeff è il coefficiente che pesa il grado di similarità  
Output: la regione estratta, ExtractedRegion  
  
// Inizializzazione  
ExtractedRegion  $\leftarrow$  seed  
// neighboringList contiene tutti i pixel già esaminati e che appartengono  
alla regione estratta  
neighboringList  $\leftarrow$  seed  
// Calcola la media dei valori di grigio della regione iniziale, avg  
avg  $\leftarrow$  calculateAverageGray(ExtractedRegion)  
// Accrescimento iterativo  
while  $\neg$ neighboringList.isEmpty do  
    Estrai il primo elemento di neighboringList e seleziona i pixel adiacenti  
    che non sono ancora stati esaminati, currentPixels  
    for all pixel  $\in$  currentPixels do  
        if  $avg - coeff \times avg < pixel < avg + coeff \times avg$  then  
            ExtractedRegion  $\leftarrow$  pixel  
            neighboringList  $\leftarrow$  pixel  
            avg  $\leftarrow$  calculateAverageGray(ExtractedRegion)  
        end if  
    end for  
end while  
// Applica la chiusura alla regione risultante per eliminare buchi  
calculateClosing(ExtractedRegion)  
return ExtractedRegion
```

---

criteri dinamici. I metodi testati sono: classificatore a soglia sul parametro di minimo enhancement, funzione discriminante lineare di Fisher, rete neurale artificiale probabilistica, multilayer perceptron allenato con l'algoritmo della backpropagation ed infine un criterio basato sul Coefficiente di Correlazione. Il lavoro presentato in [16] utilizza l'analisi strutturale per differenziare tra lesioni maligne e lesioni benigne estratte da immagini di risonanza magnetica. I parametri strutturali vengono combinati in un modello di regressione logistica per massimizzare l'accuratezza diagnostica. In [70] gli autori sperimentano un tipo speciale di rete neurale artificiale supervisionata, chiamata "Temporal Associative Subjective Memory with Bimodal Activation", per studiare caratteristiche dinamiche e morfologiche estratte dalle lesioni. In [37] viene presentata un'equazione logistica modificata a cinque parametri che descrive il comportamento dinamico dell'agente di contrasto. In [11] gli autori propongono un Modello Matematico Empirico (EMM) della diffusione dell'agente di contrasto per distinguere tra lesioni benigne e lesioni maligne.

In questa tesi la classificazione delle lesioni sospette avviene mediante *apprendimento supervisionato*. In questo paradigma di apprendimento, anche conosciuto come *learning with a teacher*, il dataset è composto da un insieme di esempi input-output. Ogni esempio, quindi, è una coppia costituita da un vettore di input ed il corrispondente vettore di output desiderato. Nello studio viene utilizzato un classificatore a soglia per classificare le feature initial slope (equazione 3.2) e washout ratio (equazione 3.3) e vengono testati diversi multilayer perceptron, allenati con l'algoritmo della backpropagation, con varie combinazioni dei criteri dinamici precedentemente descritti.

### 4.2.1 Classificatore a soglia

Gli esperimenti per il classificatore a soglia sono stati eseguiti utilizzando la tecnica k-fold cross-validation [36, 20]. Questo metodo consiste nel suddividere gli  $n$  dati a disposizione in  $k$  sottoinsiemi disgiunti, ognuno di dimensione  $n/k$ . Il classificatore viene allenato per  $k$  volte, in ognuna delle quali si utilizza un diverso sottoinsieme come test set e i restanti  $k - 1$  sottoinsiemi come training set. In questa sperimentazione i dati sono stati suddivisi in  $k = 10$  folder, ognuna contenente  $1200/10 = 120$  pattern, 60 estratti da due esami negativi e 60 estratti da due esami positivi. Ad ognuna delle  $k$  iterazioni una folder è stata utilizzata come test set e le restanti  $k - 1 = 9$  folder come training set. Tutti i dati appartenenti ad uno stesso esame sono stati inseriti all'interno di un'unica folder e tale suddivisione è stata conservata durante tutta la k-fold cross-validation, in modo che ogni esame si trovasse, di volta in volta, nel training set oppure nel test set. Ad ogni iterazione della cross-validation viene calcolata l'accuratezza del classificatore sul training set

e si sceglie la soglia che permette al classificatore di raggiungere la migliore accuratezza sul training set. Tale soglia viene, poi, applicata al test set per calcolarne l'accuratezza. Alla fine di tutta la k-fold cross-validation si sono, quindi, ricavate  $k = 10$  soglie, e l'accuratezza del classificatore è stimata come la media delle accuratezze sui diversi test set. Per ottenere il modello finale, il classificatore è stato nuovamente allenato su tutto il dataset come training e su questo è stata scelta la soglia migliore. L'accuratezza rimane quella stimata in precedenza per mezzo della cross-validation. Si noti che in tale procedura, in realtà, non viene utilizzato il validation set, in quanto il classificatore a soglia è un modello di bassa complessità in cui non ci sono iperparametri da fissare. Di seguito è presentato l'algoritmo del classificatore a soglia (Algoritmo 2).

### 4.2.2 Multilayer Perceptron

Il perceptron semplice (single-layer) fu introdotto da Rosenblatt come nuovo modello di apprendimento supervisionato [54]. Rosenblatt progettò il perceptron cercando di imitare l'organizzazione e le regole del sistema visivo biologico, costruendo un dispositivo che fosse in grado di imparare a riconoscere classi di pattern visivi [51, 53]. Il perceptron è la più semplice forma di rete neurale (consiste di un singolo neurone) capace di classificare pattern linearmente separabili. Rosenblatt sviluppò la regola di correzione dell'errore per aggiustare i parametri liberi di tale rete neurale, conosciuta come procedura di convergenza del perceptron. Inoltre dimostrò, nel teorema di convergenza del perceptron, che se il dataset è linearmente separabile allora l'algoritmo del perceptron converge dopo un numero finito di iterazioni [52]. Il *multilayer perceptron* (MLP) è una generalizzazione del perceptron semplice. Esso è composto da:

- un insieme di nodi sensori (nodi sorgente) che costituiscono lo strato di input;
- uno o più strati intermedi di nodi di elaborazione; queste unità non fanno parte né dell'input né dell'output della rete, per questo vengono definite unità nascoste;
- uno strato di output di nodi di elaborazione.

Non ci sono connessioni all'interno di uno stesso strato oppure da uno strato successivo ad uno precedente, ma una connessione può saltare qualche strato intermedio. Il segnale di input si propaga "in avanti" attraverso la rete neurale di strato in strato, dallo strato di input fino allo strato di output

---

**Algorithm 2** Classificatore a soglia

---

**Input:** il dataset ( $D$ ) composto da  $n$  coppie  $\langle \text{valore}, \text{target} \rangle$ **Output:** la migliore soglia,  $\theta_{BEST}$ Dividi  $D$  in  $k$  sottoinsiemi ( $D_{j \in [1..k]}$ ) di dimensione  $n/k$ .**for**  $j = 1..k$  **do**    //  $j$ -esima iterazione della  $k$ -fold cross-validation    //  $D_j$  è il test set     $TR \leftarrow \cup_{i \neq j} D_i$  // training set    // Calcola la migliore soglia su  $TR$  utilizzando i valori come possibili

soglie

**for all** valori  $\theta_m \in TR$  **do**         $acc_{TR}(\theta_m) \leftarrow$  accuratezza del classificatore su  $TR$  rispetto a  $\theta_m$     **end for**     $\theta_{best_j} \leftarrow \arg \max_{\theta_m} acc_{TR}(\theta_m)$     // Applica il classificatore con la soglia  $\theta_{best_j}$  al test set    //  $c$  è la classificazione ottenuta    **for all** valori  $\in testset$  **do**        **if**  $valore \geq \theta_{best_j}$  **then**             $c \leftarrow +1$         **else**             $c \leftarrow -1$         **end if**    **end for**     $acc_{D_j} \leftarrow$  accuratezza del classificatore sul test set  $D_j$  rispetto a  $\theta_{best_j}$ **end for**

// Calcola accuratezza del classificatore

 $ACC \leftarrow \frac{1}{k} \sum_{j=1}^k acc_{D_j}$ 

// Costruisci modello finale

**for all** valori  $\theta_m \in D$  **do**     $acc_D(\theta_m) \leftarrow$  accuratezza del classificatore su  $D$  rispetto a  $\theta_m$ **end for** $\theta_{BEST} \leftarrow \arg \max_{\theta_m} acc_D(\theta_m)$ **return**  $\theta_{BEST}$ 

---

attraversando gli strati intermedi di nodi nascosti. Le unità in ogni strato, quindi, hanno i loro stati determinati dall'input che ricevono dalle unità negli strati precedenti. Per questo motivo tali reti neurali vengono definite *multilayer feedforward networks*. Per la presenza di uno o più strati di unità nascoste l'informazione che arriva alle unità di input è ricodificata in una rappresentazione interna nelle unità nascoste e gli output sono generati da tale rappresentazione interna piuttosto che dai pattern di input originali. In Figura 4.4 è mostrata l'architettura di un multilayer perceptron.

L'algoritmo più popolare utilizzato per allenare un multilayer perceptron è conosciuto come *error back-propagation algorithm* o, più comunemente, *back-propagation*.

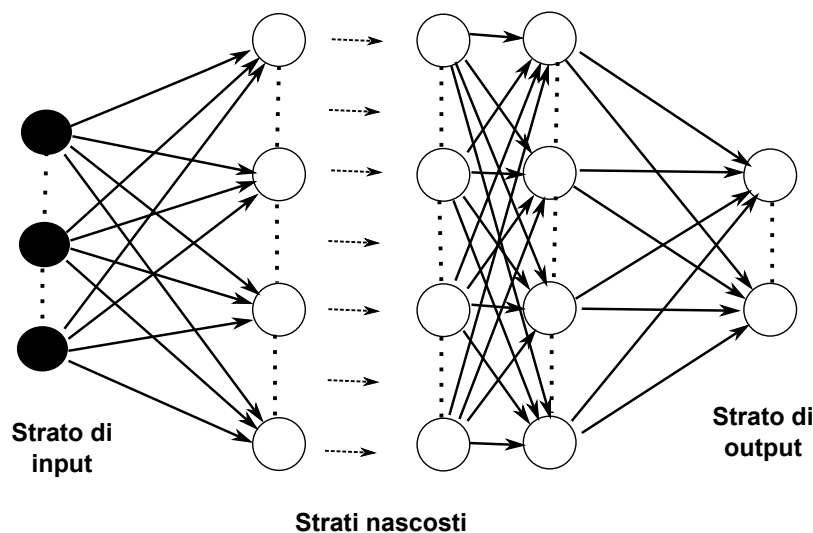


Figura 4.4: Architettura di un Multilayer Perceptron

### Error back-propagation algorithm

La back-propagation è una procedura di apprendimento per reti neurali che fu sviluppata nel 1986 da Rumelhart, Hinton e Williams [55, 56, 57]. Negli stessi anni varianti dell'algoritmo per allenare reti multistrato erano state scoperte indipendentemente in posti diversi. Parker nel 1985 derivò una regola di apprendimento che chiamò "learning-logic" [44]. Le Cun, nello stesso anno, studiò uno schema di apprendimento simile alla back-propagation [31]. Dopo la scoperta dell'algoritmo venne fuori che, in realtà, la back-propagation era stata già precedentemente descritta da Werbos nella sua tesi di dottorato nel 1974 [72]. La tesi di dottorato di Werbos fu la prima descrizione documentata del processo di training delle reti neurali artificiali attraverso



la propagazione all'indietro dell'errore. La procedura della back-propagation modifica iterativamente i pesi sulle connessioni della rete in modo da minimizzare una misura della differenza tra il vettore di output attuale della rete e il vettore di output desiderato. Con l'aggiustamento dei pesi le regolarità del task in questione sono catturate dalle interazioni tra le unità e si sviluppa, così, una rappresentazione interna che è appropriata per quel particolare dominio. Lo scopo dell'algoritmo è trovare un insieme di pesi che assicuri che, per ogni vettore di input, il vettore di output prodotto dalla rete neurale sia uguale o sufficientemente vicino al vettore di output desiderato. La risposta desiderata rappresenta l'azione ottima che dovrebbe essere eseguita dalla rete neurale. L'algoritmo back-propagation consiste di due fasi: una fase forward e una fase backward.

- Fase forward: un vettore di input è applicato ai nodi sensori della rete neurale e il suo effetto si propaga in avanti attraverso tutta la rete, di strato in strato. Alla fine un vettore di output viene prodotto come risposta attuale della rete. Tale risposta è sottratta alla risposta desiderata per produrre il segnale di errore per ogni neurone di output. Durante questa fase i pesi sinaptici delle connessioni sono tutti fissati.
- Fase backward: il segnale di errore viene propagato "all'indietro" attraverso tutta la rete neurale, di strato in strato, nella direzione opposta a quella delle connessioni sinaptiche. Da questa fase l'algoritmo prende il nome di error back-propagation. I pesi sinaptici vengono tutti aggiustati in base alla regola di correzione dell'errore dell'algoritmo, in modo che la risposta attuale della rete neurale si avvicini sempre di più alla risposta desiderata.

Di seguito è presentato l'algoritmo back-propagation che applica il metodo della discesa del gradiente per calcolare la correzione dell'errore dei pesi sinaptici (Algoritmo 3).

---

**Algorithm 3** Algoritmo Error Back-propagation

---

**Input:**  $TR, \eta, n_{in}, n_{hidden}, n_{out}, \varphi_j$ .  
 //  $TR$  è l'insieme degli esempi di training. Ogni esempio di training è una coppia della forma  $\langle \vec{x}, \vec{t} \rangle$ , dove  $\vec{x}$  è il vettore dei valori di input della rete e  $\vec{t}$  è il relativo vettore dei valori di output target.  
 //  $\eta > 0$  è il learning rate: determina la dimensione del passo nella discesa del gradiente.  
 //  $n_{in}$  è il numero di input della rete.  
 //  $n_{hidden}$  è il numero di unità nello strato nascosto.  
 //  $n_{out}$  è il numero di unità di output.  
 //  $\varphi_j$  è la funzione di attivazione associata al neurone  $j$ .

Crea una rete feedforward con  $n_{in}$  input,  $n_{hidden}$  unità nascoste e  $n_{out}$  output.

Inizializza tutti i pesi della rete a piccoli valori random.

**while** condizione di terminazione **do**

**for all**  $\langle \vec{x}, \vec{t} \rangle$  in  $TR$  **do**

    // Propaga l'istanza di input  $\vec{x}$  in avanti attraverso tutta la rete e calcola il segnale di output per ogni unità della rete.

$$v_j \leftarrow \sum_{i=0}^m w_{ji} y_i \quad (4.5)$$

dove  $v_j$  è il campo locale indotto del neurone  $j$ ,  $m$  è il numero totale di input applicati al neurone  $j$ ,  $w_{ji}$  è il peso sinaptico della connessione dal neurone  $i$  al neurone  $j$  e  $y_i$  è il segnale di output del neurone  $i$  dello strato precedente in input al neurone  $j$ .

Il segnale di output del neurone  $j$  è

$$y_j \leftarrow \varphi_j(v_j) \quad (4.6)$$

Se il neurone  $j$  appartiene al primo strato nascosto della rete, allora

$$y_i \leftarrow x_i$$

dove  $x_i$  è l' $i$ -esimo elemento del vettore di input  $\vec{x}$ .

Se il neurone  $j$  appartiene allo strato di output della rete, allora

$$y_j \leftarrow o_j$$

dove  $o_j$  è il  $j$ -esimo elemento del vettore di output.

// Calcola il segnale di errore  $e_j$

$$e_j \leftarrow t_j - o_j \quad (4.7)$$


---

---

// Propaga gli errori all'indietro attraverso la rete e calcola il gradiente locale  $\delta_j$  per ogni neurone  $j$ .

$$\delta_j \leftarrow \begin{cases} e_j \varphi'_j(v_j) & \text{se } j \text{ è un neurone di output} \\ \varphi'_j(v_j) \sum_k \delta_k w_{kj} & \text{se } j \text{ è in uno strato nascosto} \end{cases} \quad (4.8)$$

dove il neurone  $k$  appartiene allo strato successivo a quello in cui si trova  $j$ .

// Aggiorna ogni peso della rete  $w_{ji}$

$$w_{ji} \leftarrow w_{ji} + \Delta w_{ji} \quad (4.9)$$

dove

$$\Delta w_{ji} = \eta \delta_j y_i \quad (4.10)$$

**end for**  
**end while**

---



# Capitolo 5

## Sperimentazione

### 5.1 Criteri di analisi

In questa tesi la caratterizzazione delle lesioni sospette avviene mediante lo studio della diffusione e della concentrazione dell'agente di contrasto nei tessuti, quindi attraverso criteri dinamici. Le regioni di enhancement vengono analizzate sia dal punto di vista quantitativo che dal punto di vista qualitativo. Esistono diverse misure per classificare le zone di enhancement, qui vengono utilizzate quelle già descritte nella sezione 3.3 del capitolo 3. In questa sezione tali misure vengono solo elencate in quanto per alcune di esse è necessario fornire qualche osservazione o chiarimento.

#### Parametri quantitativi

Ricordiamo che Early-increase ed Initial Slope descrivono il comportamento dell'intensità del segnale nella prima fase dopo la somministrazione dell'agente di contrasto. Washout Ratio descrive il comportamento dell'intensità del segnale nella seconda fase post-contrasto (periodo post-contrasto intermedio e finale).

- Relative enhancement (RE)
- Early-increase (EI)
- Initial Slope (IS)
- Washout Ratio (WR)
- Curvature at the peak of enhancement (CP): questo criterio verrà di seguito dettagliatamente descritto in quanto, nel caso della sperimen-

tazione effettuata nella tesi, il parametro è stato calcolato in maniera diversa da quanto presentato nella formula 3.4.

### Early-increase

Nelle sperimentazioni di questa tesi viene utilizzato l'enhancement dopo 80 secondi dalla somministrazione dell'agente di contrasto.

### Washout Ratio

Nelle sperimentazioni di questa tesi, in cui ogni esame è composto da sei serie di immagini,  $k \in [0, 6)$ , il parametro washout ratio ( $W_{peak-k}$ ) è stato calcolato per diversi indici  $k$ .  $W_{peak-5}$ , cioè il decremento tra la massima intensità del segnale e l'intensità del segnale nell'ultima immagine post-contrasto, è stato classificato con il classificatore a soglia e combinato con altri criteri dinamici come input ai multilayer perceptron.  $W_{peak-k}$  per altri indici  $k$  sono stati calcolati ed utilizzati nelle reti neurali in combinazione tra loro o con altri parametri. Il criterio maggiormente utilizzato in queste sperimentazioni è  $W_{peak-5}$  perché è quello che si è rivelato più significativo. Per semplicità di notazione vengono omessi gli indici per  $W_{peak-5}$ , quindi dove non specificato, ma indicato solo come washout ratio, si dovrà intendere  $W_{peak-5}$ , altrimenti sarà esplicitamente indicato l'indice di serie a cui si fa riferimento.

### Curvature at the peak of enhancement

Come descritto nella sezione 3.3 del capitolo 3, il parametro curvature at the peak of enhancement ( $k(t)$ ) viene calcolato mediante la formula 3.4. In questo caso, la concentrazione dell'agente di contrasto  $C(t)$  viene calcolata come una funzione continua del tempo. Le curve di relative enhancement utilizzate nel dataset di questa tesi, invece, sono funzioni discrete del tempo, espresso in termini di indice di serie. Per questo motivo, la curvatura viene qui approssimata per mezzo della derivata seconda discreta dell'enhancement nel suo picco.

$$k(T_{peak}) = \frac{RI_{peak+T} - RI_{peak}}{T} \quad (5.1)$$

dove

- $T$  è l'incremento del tempo rispetto al picco, espresso come differenza tra gli indici di serie,
- $RI_{peak} = \frac{E_{peak+T} - E_{peak}}{T}$  è il rapporto incrementale dell'enhancement nel suo picco relativamente all'incremento di tempo  $T$ .

Nelle sperimentazioni la curvatura viene calcolata anche in punti (indici di serie) diversi dal picco di enhancement.

### Parametri qualitativi

La caratteristica più utilizzata della curva di relative enhancement è la sua morfologia. Essa può essere suddivisa in tre categorie, come raffigurato nella Figura 3.1. Questo schema di classificazione è solo ideale e non corrisponde precisamente alle curve estratte dalle immagini reali. Nella realtà, infatti, la forma delle curve di enhancement non è così ben definita; ad esempio è difficile trovare curve plateau in cui l'intensità del segnale nelle fasi post-contrasto intermedia e finale rimane costante nel tempo. Inoltre le curve presentano enhancement iniziali molto diversificati tra loro.

#### 5.1.1 Analisi dei criteri

Per comprendere a fondo il risultato delle sperimentazioni effettuate, è importante fare alcune osservazioni generali sulla natura e sulla complessità del problema che si sta per affrontare.

Come riportato in letteratura [34, 9, 60], i criteri dinamici, quindi il comportamento del segnale di risonanza magnetica nel tempo, contengono importante informazione per la distinzione tra lesioni benigne e lesioni maligne. Le lesioni maligne rilasciano fattori angiogenici che causano la crescita dei capillari già esistenti ed inducono alla formazione di nuovi vasi sanguigni. Tale attività angiogenica patologica comporta un duplice effetto: si verifica un accrescimento della vascolarizzazione che porta ad un maggiore afflusso dell'agente di contrasto nella zona della lesione e, contemporaneamente, si verifica un aumento della permeabilità vascolare che causa un'accelerata fuoriuscita dello stesso mezzo di contrasto dal sito del tumore (washout) [29]. Il rischio di cancro per una lesione che ha una curva washout rispetto al rischio di cancro per una lesione che ha una curva persistente è di circa 5 a 1 [59]. Il problema della classificazione delle lesioni rimane comunque molto complesso in quanto il comportamento dell'enhancement del materiale di contrasto all'interno delle lesioni stesse non è ben definito. L'aumento locale della vascolarizzazione e della permeabilità vascolare, infatti, non è affatto una caratteristica specifica dei tessuti maligni. Ad esempio, quasi tutte le lesioni benigne neoplastiche presentano una significativa ipervascolarizzazione. In generale le masse maligne sono caratterizzate da un più veloce e più forte aumento dell'intensità del segnale nella prima fase post-contrasto rispetto alle lesioni benigne. Tale fenomeno è subito seguito da un rapido washout nelle successive fasi post-contrasto. Esistono, però, molte lesioni

maligne che producono solo un lento e minimo enhancement. Inoltre esiste una grande varietà di lesioni benigne che presentano un rapido e marcato aumento dell'intensità del segnale nella prima fase post-contrasto, quindi un initial slope molto simile a quello delle lesioni maligne, che è una causa della bassa specificità della risonanza magnetica [34, 68, 28]. In Figura 5.1 sono riportate, come esempio, alcune curve reali tratte dal dataset oggetto di studio di questa tesi. Nella Figura 5.1a sono riprodotte alcune curve estratte da un esame di risonanza magnetica con diagnosi istologica di displasia. La displasia è un'anomalia dello sviluppo degli elementi cellulari costituenti un tessuto o un organo presente nella maggior parte dei tumori benigni. Dalla figura, però, è possibile notare la presenza di curve con comportamento dinamico tipicamente maligno. Nella Figura 5.1b sono riportate alcune curve estratte da un esame diagnosticato istologicamente come carcinoma. Il carcinoma è un tumore maligno molto frequente, eppure nella figura si possono riscontrare diverse curve dalla morfologia tipicamente benigna.

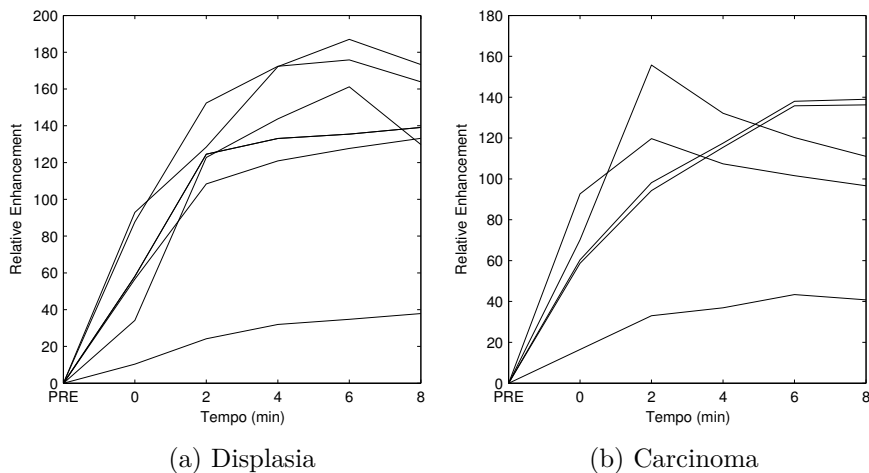


Figura 5.1: Curve estratte dal dataset

## 5.2 Costruzione dataset

Il dataset analizzato in questo studio è stato costruito utilizzando l'applicazione descritta nel capitolo 4 per la raccolta dei dati. Dal Dipartimento di Radiologia dell'Università di Pisa sono stati forniti 85 esami di risonanza magnetica contenenti, per ogni esame, la diagnosi fornita dai medici e la diagnosi istologica, che, fornendo la valutazione dell'effettiva natura del-



la lesione, rappresenta la diagnosi definitiva. Per l'applicazione di una rete neurale è importante preparare un dataset bilanciato, che, nel processo di training, comprenda tutti i casi che dovranno essere classificati. La valutazione delle reti neurali, infatti, potrebbe essere inappropriata se il dataset non è bilanciato, cioè se un tipo di pattern è sottorappresentato o sovrarappresentato, in quanto ne sarebbe influenzata. Per questo motivo, dopo una prima fase di "data-cleaning" in cui sono stati scartati gli esami corrotti o quelli la cui interpretazione risultava troppo rischiosa senza l'aiuto di un esperto, sono stati scelti 40 esami, di cui 20 esami contengono tumori maligni e gli altri 20 contengono tumori benigni. Le immagini scelte presentano tutte delle zone di enhancement visibili. Per il dataset sono stati scelti gli esami contenenti le immagini con zone di forte enhancement e, all'interno di tali immagini, sono state selezionate le zone di maggiore enhancement. Tra i 40 esami sono presenti due falsi negativi, cioè 2 esami diagnosticati come negativi dai medici ma che poi si sono rivelati positivi all'analisi istologica. In entrambi i casi si tratta di carcinoma che, come è stato spiegato nella sezione precedente, può presentare delle caratteristiche tipicamente benigne, pur essendo un tumore maligno. Dopo aver rimosso il background, vengono estratte, all'interno delle immagini, le zone di più forte enhancement attraverso la tecnica del region growing oppure selezionando a mano l'intera area di enhancement. In Figura 5.2 si vede un esempio di regione di enhancement selezionata manualmente. La figura mostra un'immagine di risonanza magnetica appartenente al dataset di questa tesi. Nel quadrante sinistro è presente un carcinoma, come risulta dalla diagnosi istologica. Il rettangolo rosso rappresenta la zona selezionata a mano per mezzo dell'applicazione. Come si vede dalla figura, si cerca di selezionare la più piccola regione che includa l'intera zona con più forte enhancement.

Alcune regioni di enhancement vengono, invece, estratte in maniera semi-automatica con la tecnica del region growing, descritta nel capitolo 4. La Figura 5.3 mostra la stessa immagine di risonanza magnetica della figura precedente (Figura 5.2). In questo caso, la stessa regione di enhancement che nell'immagine precedente era stata selezionata manualmente, viene estratta con il region growing. In tal modo viene selezionata più precisamente solo la zona di enhancement e non i tessuti circostanti, come invece accadeva nel caso della selezione completamente manuale. In alto a sinistra nella figura è possibile vedere l'intera regione estratta dal region growing.

Come è evidente dalle figure descritte, in entrambi i casi la regione estratta può essere anche piuttosto grande, in base all'estensione della zona di maggiore enhancement. Per questo motivo ogni regione estratta viene, a sua volta, suddivisa in Region Of Interest (ROI) di dimensione  $3 \times 3$  voxel. Su ognuna di tali ROI viene calcolata la curva di relative enhancement in modo

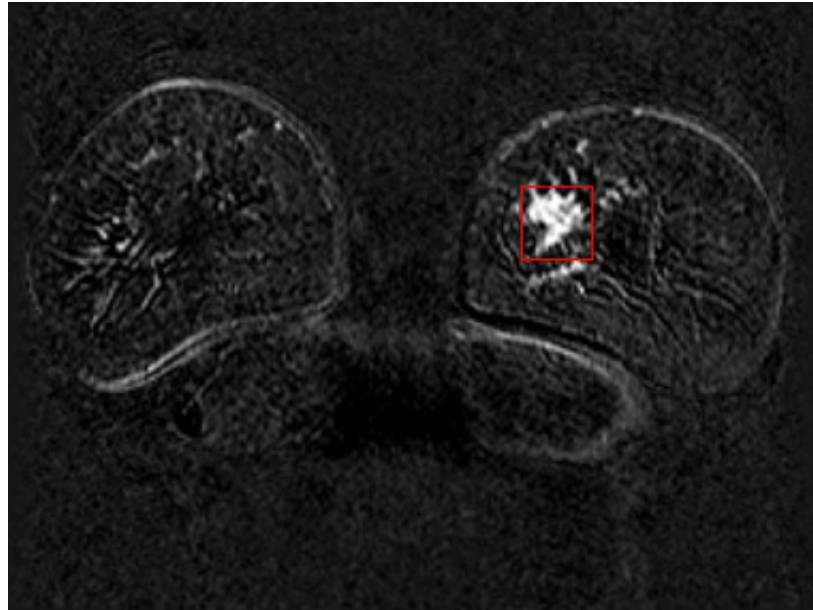


Figura 5.2: Area di enhancement selezionata manualmente.

da ottenere le curve in ogni punto della regione estratta. Non tutte le curve vengono inserite nello studio, sono eliminate quelle che indicano un enhancement molto basso o addirittura nullo e quelle che presentano un andamento molto irregolare in quanto potrebbero derivare da una selezione del tutto sbagliata (ad esempio potrebbero essere rappresentative della sezione di una vena invece che di una lesione, come sarà meglio descritto in seguito). In particolare:

- vengono eliminate le curve di relative enhancement in cui tutti i punti sono uguali a 0 (Figura 5.4a);
- vengono eliminate le curve corrispondenti a ROI in cui, dopo due minuti dalla somministrazione dell'agente di contrasto, si verifica un enhancement relativo inferiore a  $\theta = 30\%$  (Figura 5.4b);
- vengono eliminate le curve di relative enhancement corrispondenti a ROI che presentano fasi di crescita dell'intensità del segnale alternate a fasi di intensità del segnale nulla, quindi un andamento a zig-zag in cui alcuni punti sono uguali a 0 (Figura 5.4c).

Ad ogni curva scelta è associata la diagnosi istologica relativa all'esame da cui è stata estratta. I criteri di selezione appena descritti sono gli unici applicati per filtrare i dati a disposizione, in quanto, nella costruzione del dataset per questa tesi, ci si è posti l'obiettivo di creare un dataset quanto più

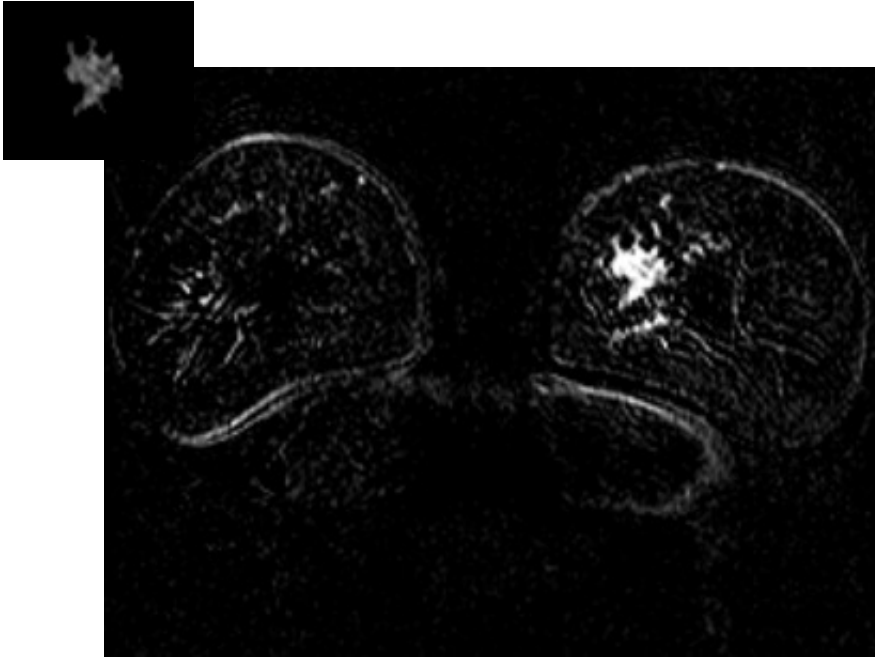


Figura 5.3: Area di enhancement estratta con region growing.

possibile realistico. Tale scelta dipende dal fatto che esistono numerose sperimentazioni sui criteri diagnostici precedentemente descritti, utilizzando però dataset notevolmente filtrati. In questa tesi si vuole provare a sperimentare la performance diagnostica dei classificatori utilizzati analizzando tali criteri su un insieme di dati realistico e dunque più complesso. Il dataset risultante, quindi, contiene curve scelte in modo da rispecchiare la situazione reale, in cui patologie benigne (ad es. fibroadenoma, che è caratterizzato da un rapido e forte enhancement) mostrano intensità del segnale molto simili a quelle dei tumori maligni causando un largo overlap tra curve di relative enhancement maligne e curve di relative enhancement benigne [50]. In Figura 5.1, dove sono riportate alcune curve tratte dal dataset oggetto di studio, si è già visto come patologie benigne e tumori maligni presentino nella realtà delle caratteristiche, nella morfologia delle curve, molto simili tra loro, comportando una vera e propria sovrapposizione e rendendo molto difficile la distinzione tra masse benigne e masse maligne.

Un aspetto interessante del dataset di questa tesi consiste, appunto, nella definizione manuale o semi-automatica della regione di enhancement e, di conseguenza, delle relative Region Of Interest. La definizione della dimensione, della forma e della posizione della ROI influenza criticamente il profilo tempo-intensità del segnale, in quanto gli effetti dei tessuti con scarso enhan-

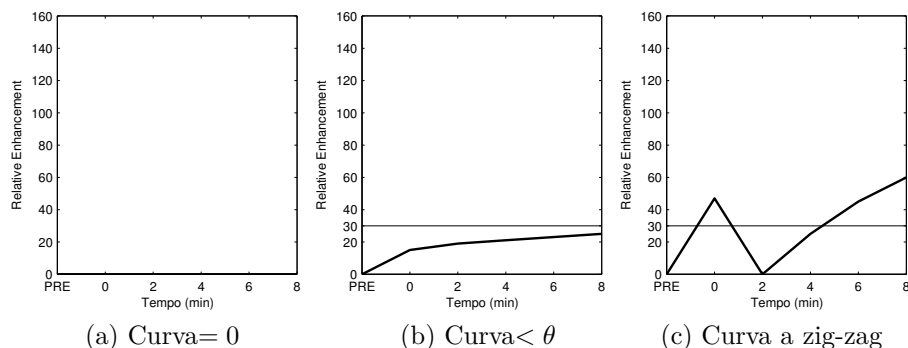


Figura 5.4: Curve eliminate

cement o del tutto privi di enhancement potrebbero essere inavvertitamente inclusi nella ROI a causa della sovrapposizione o vicinanza tra tessuti tumorali e tessuti sani. Tale overlap comporta l'appiattimento dell'intensità del segnale influenzando, così, la forma della curva tempo-enhancement e tutti i parametri quantitativi che derivano dall'intensità del segnale. La scelta della posizione della ROI, in particolare, può condurre alla selezione di dati molto fuorvianti. Come è stato descritto nella sezione 3.4 del capitolo 3, gli esami a disposizione sono stati acquisiti lungo il piano coronale. Uno dei problemi di questa tipologia di acquisizione è il fatto che la maggior parte dei vasi sanguigni viaggia sul piano coronale, quindi nelle immagini di risonanza magnetica una sezione di una vena può apparire molto simile ad una lesione focale rendendone difficile la distinzione. Questa eventuale confusione può indurre al posizionamento erroneo di una ROI sulla sezione di una vena piuttosto che sulla lesione, introducendo nel dataset dati sbagliati. Anche la definizione semi-automatica, attraverso la tecnica del region growing, della regione di enhancement è soggetta al rischio di errore. La scelta della zona di enhancement che verrà estratta automaticamente, infatti, avviene in maniera manuale e questo comporta la possibilità di selezionare, come regione da estrarre, la sezione di una vena piuttosto che la zona di una lesione. Inoltre, come già spiegato nella sezione 4.1.3, il risultato stesso di tale metodo è sensibile alla scelta del punto iniziale da cui costruire la regione da estrarre. La tecnica del region growing, infine, non può essere applicata in tutti gli esami, in quanto, in alcune patologie, l'enhancement è estremamente diffuso e non risulta possibile selezionare delle singole regioni. La Figura 5.5 mostra un'immagine di risonanza magnetica tratta dall'esame di una paziente affetta da una grave forma di displasia. Dalla figura è evidente che la patologia è diffusa su entrambe le mammelle e, di conseguenza, anche l'enhancement, rendendo

impossibile l'estrazione di una singola regione per mezzo della tecnica del region growing.

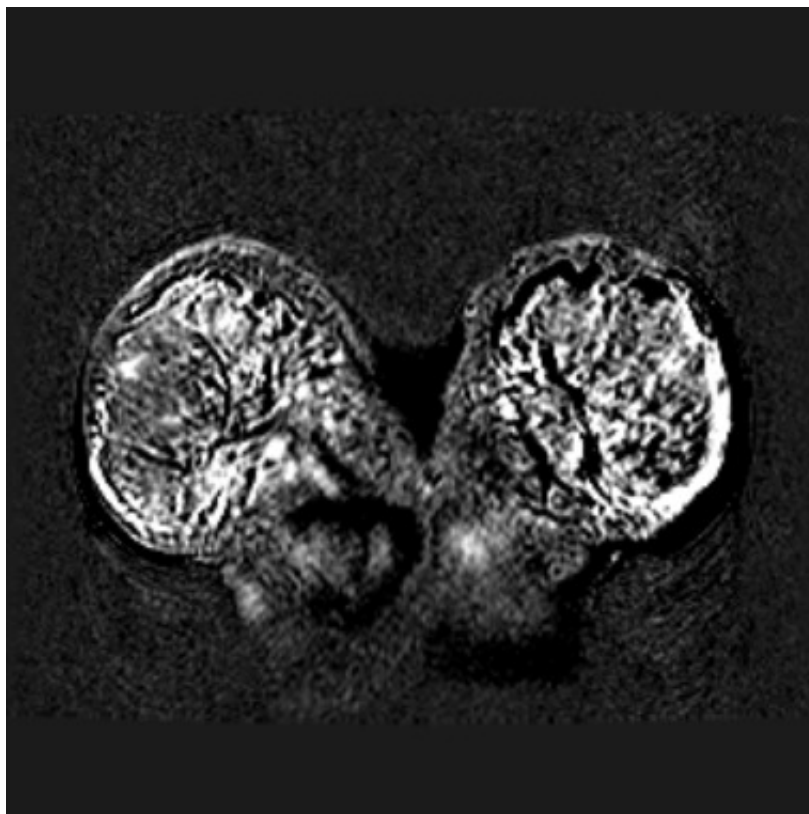


Figura 5.5: Immagine appartenente ad un esame di risonanza magnetica diagnosticato istologicamente come displasia.

In conclusione, per lo studio effettuato in questa tesi è stato costruito un dataset più complesso rispetto a quelli esistenti in letteratura sotto diversi aspetti. Il dataset è fortemente soggetto ad errori a causa della selezione manuale della ROI, sia per quanto riguarda la possibile sovrapposizione tra tessuti sani e tessuti tumorali, sia per quanto riguarda la scelta del posizionamento della ROI stessa. Sono state estratte curve di relative enhancement dalle regioni di enhancement indicate dal referto istologico ma, per i motivi precedentemente descritti, tale selezione potrebbe introdurre nel dataset falsi positivi o falsi negativi. I dati scelti, quindi, sono stati solo minimamente filtrati, cioè non sono state eliminate curve che non presentano l'apparenza tipica, in modo da rispecchiare le reali sovrapposizioni tra patologie benigne e patologie maligne e le conseguenti problematiche di interpretazione affrontate anche dagli stessi radiologi. Tutti gli altri dataset esistenti sono stati

notevolmente filtrati con il supporto di un radiologo esperto. Questo rende la correttezza di tali dataset comunque dipendente dalla scelta manuale del posizionamento della ROI e dalla correttezza dell'interpretazione radiologica, ma sicuramente meno soggetta ad errori rispetto al dataset costruito per le sperimentazioni di questa tesi in cui è più probabile l'introduzione di falsi positivi e falsi negativi. La presenza all'interno del dataset di numerosi casi "difficili", dovuti alle motivazioni precedentemente descritte, rende le sperimentazioni su questo insieme di dati una sfida interessante che potrebbe anche portare al conseguimento di risultati non ottimali.

Riassumendo la costruzione del dataset, le immagini scelte appartengono a 40 esami di risonanza magnetica contenenti la diagnosi istologica. Venti esami contengono tumori maligni e gli altri 20 contengono tumori benigni. Da ogni esame sono state estratte 30 ROI all'interno delle regioni di enhancement su ognuna delle quali è stata calcolata la curva di relative enhancement e gli altri parametri quantitativi precedentemente descritti. Il dataset è, quindi, composto da  $30 \times 20 = 600$  pattern positivi (maligni) e  $30 \times 20 = 600$  pattern negativi (benigni), per un totale di 1200 pattern. Ad ogni pattern è associata la diagnosi istologica relativa all'esame da cui è stato estratto.

### 5.3 Scelta e stima del modello

Un aspetto delicato nel creare una rete neurale è stabilire il training set ed il test set. La qualità della rete neurale risultante e del processo di validazione può dipendere dalla qualità della distribuzione dei casi all'interno del training set. Per questo motivo i diversi multilayer perceptron sono stati testati utilizzando una variante della k-fold cross-validation, la tecnica *double k-fold cross-validation* [40, 12]. Questo metodo consiste di due cross-validation annidate: quella più esterna serve per stimare l'errore di predizione del modello su un insieme indipendente di dati non visti durante la fase di addestramento, il test set. La cross-validation interna, invece, è utilizzata per selezionare la miglior configurazione di parametri della rete neurale e, in questo caso, per poter applicare la tecnica di *early stopping* per ridurre il problema dell'overfitting, come sarà meglio spiegato nel resto della sezione. La double k-fold cross-validation, quindi, permette un maggior utilizzo dei dati, in quanto, grazie alle due cross-validation annidate, tutti i dati a disposizione vengono sfruttati per tutte le fasi del processo di allenamento, validazione e test della rete neurale. In questo modo si realizza un'indagine ampia e corretta, dunque più affidabile. Più precisamente, la strategia double k-fold cross-validation consiste nel suddividere gli  $n$  dati a disposizione in  $k$  sottoinsiemi disgiunti, ognuno di dimensione  $n/k$ . Ad ognuna delle  $k$  iterazioni si utilizza un diverso

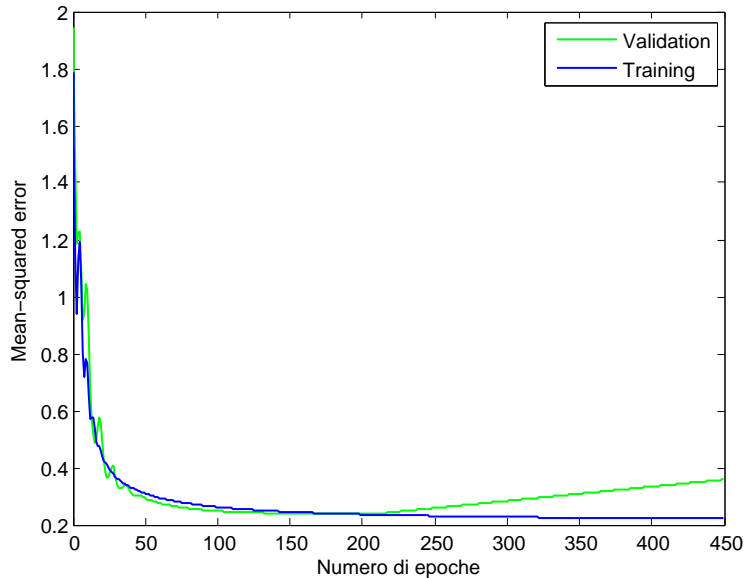


Figura 5.6: Curve di learning.

sottoinsieme come test set e i restanti  $k - 1$  sottoinsiemi come calibration set. Sul calibration set viene eseguita un'ulteriore  $j$ -fold cross-validation. Il calibration set, quindi, viene a sua volta suddiviso in  $j$  sottoinsiemi disgiunti. Il classificatore viene allenato per  $j$  volte, in ognuna delle quali si utilizza un diverso sottoinsieme come validation set e i restanti  $j - 1$  sottoinsiemi come training set. Infine l'accuratezza del modello viene calcolata sul test set. Con questa tecnica non si ottiene un modello, ma si ricava solo una stima dell'errore di predizione, calcolata su  $k \times j$  diversi modelli. Tale stima è, però, accurata ed affidabile perché determinata per mezzo di un'estesa valutazione effettuata con la double k-fold cross-validation. Nello studio di questa tesi i dati a disposizione sono stati suddivisi in  $k = 10$  folder, ognuna contenente  $1200/10 = 120$  pattern, 60 estratti da due esami negativi e 60 estratti da due esami positivi. Ad ognuna delle  $k$  iterazioni una folder è stata utilizzata come test set e le restanti  $k - 1 = 9$  folder come calibration set. Il calibration set è stato, a sua volta, suddiviso con una ulteriore  $j$ -fold cross-validation, con  $j = 9$ . Ad ognuna delle  $j$  iterazioni una folder è stata utilizzata come validation set e le restanti  $j - 1 = 8$  folder come training set. Dunque la rete neurale è stata allenata su  $j - 1 = 8$  folder (training set) e l'errore di validazione è stato calcolato sulla folder rimanente (validation set). Il multilayer perceptron allenato con l'algoritmo della back-propagation è soggetto al pro-

blema dell'overfitting se il training non viene fermato al momento giusto: con il procedere della sessione di training, l'errore sul training set diminuisce, ma da un certo punto in poi la rete neurale inizia ad imparare caratteristiche di alcuni esempi di training che non sono rappresentative della distribuzione generale degli esempi e quindi non è più in grado di generalizzare bene. Con la tecnica di early stopping l'algoritmo utilizza il training set per guidare la discesa del gradiente e il validation set per monitorare l'errore. L'algoritmo dovrebbe fermarsi dopo un numero di iterazioni che produce il minimo errore sul validation set, in quanto questo è il miglior indicatore del comportamento della rete sugli esempi non visti. In Figura 5.6 è mostrato un esempio di curve di apprendimento. Nella figura si vede che l'errore sul training set diminuisce monotonamente. L'errore sul validation set, invece, diminuisce fino ad un certo punto e poi inizia ad aumentare a causa dell'overfitting sugli esempi di training. La strategia double k-fold cross-validation ha prodotto ad ogni sperimentazione  $k \times j = 90$  modelli (Algoritmo 4). Per costruire il classificatore finale, che è stato poi inserito all'interno dell'interfaccia grafica descritta nel capitolo 4, sono stati scelti gli iperparametri con la cross-validation interna e tali iperparametri sono stati utilizzati per allenare una rete neurale su tutto il dataset. La rete neurale è stata allenata con la tecnica 10-fold cross-validation da cui si sono ottenuti dieci modelli, che sono stati organizzati in un comitato per realizzare un unico modello finale. Si è ottenuto, così, un modello fissato di cui si conosce la stima dell'errore di predizione fatta in precedenza. Tutti i dati appartenenti ad uno stesso esame sono stati inseriti all'interno di un'unica folder e tale suddivisione è stata mantenuta per tutta la double k-fold cross-validation, in modo che ogni esame si trovasse, di volta in volta, nel training set oppure nel validation set oppure nel test set.

### 5.4 Ambiente di calcolo

In questa tesi, per effettuare le sperimentazioni, è stato utilizzato l'ambiente di calcolo Matlab. Il motivo di tale scelta è dovuto al fatto che Matlab contiene un toolbox specifico per le reti neurali, Neural Network Toolbox. Tale toolbox possiede gli strumenti per la progettazione, l'implementazione, la visualizzazione e la simulazione di reti neurali. Neural Network Toolbox fornisce un supporto completo per molti paradigmi di rete neurale noti ed inoltre semplifica la progettazione e la gestione di reti neurali personalizzate. Con le reti neurali è possibile costruire modelli predittivi basati su un numero ottimamente minimizzato di caratteristiche di input [33]. La selezione delle variabili di input è, dunque, un problema rilevante perché i risultati del test sono fortemente influenzati dalla combinazione dei valori di input du-



---

**Algorithm 4** Double k-fold cross-validation

---

**Input:** il dataset ( $D$ ) composto da  $n$  pattern

Dividi  $D$  in  $k$  sottoinsiemi ( $D_{i \in [1..k]}$ ) di dimensione  $n/k$ .

**for**  $i = 1..k$  **do**

    //  $i$ -esima iterazione della double k-fold cross-validation

    //  $D_i$  è il test set

$CAL \leftarrow \cup_{v \neq i} D_v$  // calibration set

    Dividi  $CAL$  in  $j$  sottoinsiemi ( $C_{l \in [1..j]}$ ) di dimensione  $|CAL|/j$ .

    // Esegui una  $j$ -fold cross-validation per ogni iperparametro

**for**  $l = 1..j$  **do**

        //  $l$ -esima iterazione della  $j$ -fold cross-validation interna

        //  $C_l$  è il validation set

$TR \leftarrow \cup_{v \neq l} C_v$  // training set

        Allena la rete neurale sul training set ( $TR$ ).

        Calcola l'errore di validazione sul validation set ( $C_l$ ).

**end for**

    Calcola l'errore totale di validazione su tutti i validation set.

    Scegli il modello con il minor errore di validazione.

    Applica le  $j$  reti al test set ( $D_i$ ) e calcola la media delle accurattezze.

**end for**

---

rante la fase di training. Mediante il supporto del Neural Network Toolbox è stato, quindi, possibile effettuare numerose sperimentazioni testando diverse combinazioni dei vari criteri dinamici al fine di trovare la migliore selezione delle variabili di input per ottenere la migliore performance diagnostica. Inoltre i diversi multilayer perceptron sono stati collaudati testando diverse combinazioni degli iperparametri al fine di trovare un setting adeguato di iperparametri per il problema in questione.

### 5.5 Valutazione dei risultati

I risultati dei vari classificatori sono stati analizzati e confrontati tra di loro in base alle misure statistiche della performance di un test di classificazione introdotte nel paragrafo 2.3. Tali misure sono calcolate nel seguente modo:

- *Vero positivo*: diagnosi corretta in cui un tumore maligno (diagnosticato istologicamente) viene correttamente classificato come positivo (TP);
- *Falso positivo*: diagnosi errata in cui una lesione benigna (diagnosticata istologicamente) viene erroneamente classificata come positiva (FP);
- *Vero negativo*: diagnosi corretta in cui una lesione benigna (diagnosticata istologicamente) viene correttamente classificata come negativa (TN);
- *Falso negativo*: diagnosi errata in cui un tumore maligno (diagnosticato istologicamente) viene erroneamente classificato come negativo (FN);
- *Sensitività*: riguarda l'abilità del test nell'identificare i risultati positivi; essa è la probabilità che il test risulti positivo nel caso in cui la malattia sia effettivamente presente:

$$Sens = \frac{TP}{TP + FN} \quad (5.2)$$

- *Specificità*: riguarda l'abilità del test nell'identificare i risultati negativi; essa è la probabilità che il test risulti negativo nel caso in cui la malattia non sia presente:

$$Spec = \frac{TN}{TN + FP} \quad (5.3)$$

- *Accuratezza*: riguarda l'abilità del test nel restituire una classificazione corretta; essa è la proporzione di risultati corretti (sia positivi che negativi) rispetto a tutti i dati classificati:

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \quad (5.4)$$

## 5.6 Sperimentazioni

Nelle sperimentazioni sono stati studiati i vari criteri dinamici descritti all'inizio del capitolo ed è stato analizzato il loro potere predittivo. Le caratteristiche initial slope e washout ratio sono state classificate mediante un classificatore a soglia. Gli altri parametri sono stati combinati in diversi modi, e con initial slope e washout ratio, e su tali combinazioni sono stati testati diversi multilayer perceptron allenati con l'algoritmo della back-propagation. Gli esperimenti con il classificatore a soglia sono stati eseguiti utilizzando la tecnica k-fold cross-validation, descritta nel paragrafo 4.2.1.

I multilayer perceptron implementati in questo studio sono composti da tre strati. Il primo strato è costituito da un numero variabile di nodi di input, in base alle diverse combinazioni di caratteristiche, ognuno corrispondente ad uno dei diversi criteri dinamici utilizzati per la classificazione. Il secondo strato, lo strato nascosto, è costituito da un numero variabile di nodi di elaborazione, in base al modello scelto durante la fase di model selection della cross-validation interna per quella data combinazione di caratteristiche in input. Infine, l'ultimo strato, lo strato di output, è costituito da un singolo neurone che produce una classificazione binaria (maligno/benigno). Per determinare la migliore struttura di ogni rete neurale, è stato effettuato un gran numero di simulazioni variando i diversi iperparametri della rete (unità nascoste, iterazioni, learning rate, momento, funzione di attivazione, inizializzazione). Durante la fase di model selection gli iperparametri sono stati modificati gli uni rispetto agli altri, in maniera incrociata, in modo da individuare un sottoinsieme per cui è possibile fissare un valore ragionevole per tutte le configurazioni. Infine il modello è stato ottimizzato, rispetto ai rimanenti iperparametri considerati più sensibili, al variare delle configurazioni di input. Come già spiegato nella sezione 5.3, la selezione degli iperparametri della rete neurale e la stima del modello ottenuto sono state effettuate mediante la tecnica double 10-fold cross-validation. Per mezzo della cross-validation interna, è stato selezionato, di volta in volta, il modello che ha prodotto il minor errore di validazione. Tale errore è stato calcolato, sul validation set, come l'errore quadratico medio (mse) tra l'output della rete neurale e l'output desiderato. Per costruire il modello finale, è stata allenata

## 5. Sperimentazione

---

una rete neurale, con gli iperparametri selezionati, su tutto il dataset per mezzo di una 10-fold cross-validation, in modo da poter applicare la tecnica di early-stopping. Si sono, così, ottenuti dieci modelli con i quali è stato realizzato un comitato che rappresenta il modello finale. La stima dell'accuratezza di tale modello è quella fatta in precedenza per mezzo della double 10-fold cross-validation. Di seguito sono riportati i diversi valori testati con lo scopo di trovare un setting adeguato di iperparametri per il problema in questione (tabella 5.1).

Iperparametri	Valori
Funzione attivazione	Funzione logistica, funzione tangente iperbolica
Neuroni nascosti	5, 7, 10, 15, 20, 25, 30, 50, 100
Learning rate	0.01, 0.05, 0.1, 0.5
Momento	0.5, 0.7, 0.8, 0.9
Numero epoche	10000, 15000, 20000, 25000, 30000, 50000, 100000
Early stopping	100, 200, 300, 400, 500

Tabella 5.1: Iperparametri e relativi valori testati nel multilayer perceptron.

Nella tabella degli iperparametri, numero epoche indica il numero massimo di epoche di training testate ed early stopping indica il numero massimo di epoche di training durante le quali l'errore di validazione continua ad aumentare. Inoltre sono state provate diverse normalizzazioni dei dati di input, di seguito descritte.

1. Ogni vettore di input è normalizzato in  $[-1, 1]$  secondo la regola

$$y = (ymax - ymin) * (x - xmin) / (xmax - xmin) + ymin \quad (5.5)$$

dove

- $ymax$  e  $ymin$  rappresentano il massimo e il minimo valore dell'intervallo di normalizzazione (qui  $ymax = 1$  e  $ymin = -1$ ),
  - $xmax$  e  $xmin$  rappresentano il massimo e il minimo valore della sequenza di input considerata,
  - $x$  è il valore di input che deve essere normalizzato.
2. Tutti i dati sono normalizzati rispetto al massimo e minimo valore sull'intero dataset prima di darli alla rete neurale.
  3. Ogni vettore di input è normalizzato in modo che abbia media 0 e deviazione standard 1.

Durante la fase di model selection della cross-validation interna sono state sperimentate numerose combinazioni di tali iperparametri, anche con diverse inizializzazioni, fino ad individuare una configurazione base, composta solo da alcuni di essi, che ha prodotto risultati ragionevoli per tutte le diverse combinazioni di input e che, quindi, è stato possibile fissare. In tabella 5.2 è riportata la configurazione degli iperparametri selezionata per essere mantenuta costante nel resto della sperimentazione.

Funzione attivazione	Learning rate	Momento	Numero epoche	Early stopping
Funzione logistica	0.1	0.8	30000	300

Tabella 5.2: Configurazione base dei multilayer perceptron.

La normalizzazione dei dati di input dipende dalla particolare combinazione dei criteri dinamici che si intende classificare, così varia in base alla configurazione di input utilizzata. Tale configurazione di input, ed in particolare il numero di neuroni di input, potrebbe influire anche sul numero di neuroni nascosti necessari, per cui il setting di questo iperparametro è stato fatto di volta in volta al variare delle combinazioni di input. Inoltre, avendo verificato con molte sperimentazioni che il multilayer perceptron allenato con l'algoritmo della back-propagation è soggetto ad overfitting se il training non viene fermato al momento giusto, è stata utilizzata la tecnica di early stopping per tutte le sessioni di addestramento delle reti neurali. Il numero di iterazioni oltre il quale l'algoritmo di training dovrebbe fermarsi, cioè il numero di iterazioni che produce il minor errore sul validation set, dipende dal particolare modello utilizzato. Per decidere in che punto fermare la sessione di training, l'errore sul validation set è monitorato durante il processo di addestramento. Quando l'errore di validazione continua ad aumentare per un numero specificato di epoche, l'algoritmo di training viene fermato e vengono restituiti i parametri liberi della rete neurale relativi al punto in cui l'errore di validazione era minimo. Una volta terminata la fase di model selection delle reti neurali, quindi una volta stabilito il setting adeguato per gli iperparametri presi in esame, per ogni diversa configurazione di input sono state effettuate 10 diverse inizializzazioni dei multilayer perceptron. Alla fine la stima dell'accuratezza del modello è stata calcolata come la media delle accuratèzze sui test set ottenute dalle singole inizializzazioni.

Nella sezione 5.7 sono riportati i risultati delle sperimentazioni per tutte le combinazioni dei criteri dinamici che sono state testate con i multilayer perceptron ed i risultati ottenuti con i classificatori a soglia.

## 5.7 Risultati

In questa sezione sono mostrate le accuratezze, sensitività e specificità stimate, in base alle diverse configurazioni di input, e le stesse misure ottenute con i classificatori a soglia per i criteri initial slope e washout ratio. Come precedentemente detto, le sperimentazioni con il classificatore a soglia sono state effettuate utilizzando la tecnica 10-fold cross-validation, e quelle con i multilayer perceptron con la tecnica double 10-fold cross-validation. Qui sono, quindi, riportati i risultati dei classificatori per le 10 iterazioni. Nella sezione 5.3 è stato spiegato che la cross-validation esterna serve per stimare il rischio del modello, selezionato per mezzo della cross-validation interna, su un insieme di dati mai visti finora, il test set. In questo caso, ad ogni iterazione della cross-validation esterna si ottengono 9 reti neurali che vengono applicate al test set calcolando la media delle accuratezza, delle sensitività e delle specificità. Nelle sperimentazioni viene realizzato anche un comitato delle 9 reti neurali: viene calcolato l'output medio delle 9 reti neurali applicate al test set e su tale output medio vengono calcolate accuratezza, sensitività e specificità. Per quanto riguarda i parametri classificati con i multilayer perceptron, le misure statistiche di seguito riportate costituiscono la media calcolata sulle dieci diverse inizializzazioni. Per tutte le misure, oltre alla media, è riportata anche la deviazione standard, per mostrare la dispersione dei risultati rispetto alle diverse inizializzazioni e rispetto alle diverse iterazioni della double k-fold cross-validation. Di seguito sono elencate le combinazioni testate nelle sperimentazioni con le relative abbreviazioni in alcuni casi utilizzate per semplificare la notazione.

- Initial Slope (S)
- Washout Ratio (W)
- Relative Enhancement (E)
- Relative Enhancement-Early-increase (E-EI)
- Relative Enhancement-Initial Slope (E-S)
- Relative Enhancement-Washout Ratio (E-W)
- WashoutCurve (WC)
- Relative Enhancement- $W_{peak-4}$  (E-W4)
- Relative Enhancement-WashoutCurve (E-WC)

- Relative Enhancement-Washout Ratio-Initial Slope (E-W-S)
- Relative Enhancement-Washout Ratio-Early-increase-Initial Slope (E-W-EI-S)
- Relative Enhancement-Curvature at the peak of enhancement (E-C)
- Relative Enhancement-Initial Slope-Curvature at the peak of enhancement (E-S-C)
- Relative Enhancement-Washout Ratio-Curvature at the peak of enhancement (E-W-C)
- Initial Slope-Early-increase-Washout Ratio-Curvature at the peak of enhancement (S-EI-W-C)

### 5.7.1 Initial Slope

Il criterio initial slope definisce la pendenza della curva di relative enhancement nel picco (sezione 3.3.2). Come già detto, tale parametro, oltre ad essere combinato con gli altri criteri dinamici in input alla rete neurale, è stato classificato con un classificatore a soglia. In tabella 5.3a sono riportate le misure statistiche ottenute dal classificatore a soglia sul parametro initial slope in ciascuna delle dieci iterazioni della 10-fold cross-validation. In fondo sono calcolate la media e la deviazione standard sulle dieci iterazioni. Dai risultati mostrati nella tabella è evidente che tale parametro non contiene in sé informazione discriminativa rilevante. Questo è dovuto al fatto che, come si vede nell'istogramma di Figura 5.7a, i valori di tale parametro per i pattern benigni e i pattern maligni si trovano tutti concentrati nelle stesse regioni. Inoltre la bassa specificità ottenuta conferma l'esistenza di una grande varietà di lesioni benigne (ad esempio, fibroadenoma) che presentano un rapido e marcato aumento dell'intensità del segnale nella prima fase post-contrasto, quindi un initial slope molto simile a quello delle lesioni maligne.

### 5.7.2 Washout Ratio

Il parametro washout ratio definisce la discesa della curva tempo-relative enhancement (sezione 3.3.2). Come per il criterio initial slope, anche questo parametro, oltre ad essere combinato con gli altri nella rete neurale, è stato classificato con un classificatore a soglia. Nella tabella 5.3b sono riportate le misure statistiche del classificatore a soglia relative a tale criterio. A differenza del criterio initial slope, il washout ratio con il classificatore a soglia ha prodotto risultati decisamente migliori, dimostrando che tale criterio

ha di per sé una significativa abilità discriminativa (accuratezza = 88.83%, sensitività = 88.16%, specificità = 90.5%). Questo è evidente anche dall'istogramma di Figura 5.7b, in cui, nonostante una zona di sovrapposizione, i pattern negativi sono abbastanza separati da quelli positivi.

### 5.7.3 Relative Enhancement

Come già spiegato nella sezione 3.3.2, questo criterio rappresenta l'incremento dell'intensità del segnale rispetto alla fase pre-contrasto. Gli esami utilizzati nello studio di questa tesi sono composti da cinque serie post-contrasto, acquisite a 0, 2, 4, 6 e 8 minuti dopo la somministrazione dell'agente di contrasto (sezione 3.4). Interpolando i valori ottenuti si ricava la curva tempo-relative enhancement. Il dataset costruito per classificare tali curve è, quindi, composto da sequenze di cinque valori, corrispondenti ai valori dell'enhancement nelle cinque serie post-contrasto. Il modello che ha prodotto i migliori risultati sul validation set, durante la fase di model selection, è costituito da 30 nodi nello strato nascosto. In conclusione, la struttura che emerge dalle diverse prove della cross-validation interna è in media un multilayer perceptron così formato:

- 5 nodi nello strato di input
- 30 nodi nello strato nascosto
- 1 nodo nello strato di output

Ogni sequenza dei dati di input è stata normalizzata secondo la regola 5.5.

Nella tabella 5.4a è riportata la stima dell'accuratezza dei multilayer perceptron per il criterio relative enhancement. Nella tabella 5.8 sono riportate le stime della sensitività e della specificità. L'accuratezza del comitato delle nove reti neurali applicate a tale criterio dinamico, pari al 91.15%, dimostra che la morfologia della curva di relative enhancement ha un significativo potere predittivo nella differenziazione delle lesioni. La sensitività (88.17%) ottenuta risulta essere più bassa a causa delle curve "difficili" introdotte nel dataset. Per quanto riguarda la specificità, invece, sono stati ottenuti risultati promettenti (93.67%).

### 5.7.4 Relative Enhancement-Early-increase

Questa configurazione rappresenta la combinazione dei criteri dinamici relative enhancement ed early-increase. Come già spiegato nella sezione 3.3.2, il parametro early-increase è un valore che esprime il comportamento dell'intensità del segnale di risonanza magnetica nella fase immediatamente successiva



alla somministrazione dell'agente di contrasto. Il dataset costruito per tale combinazione è, quindi, composto da pattern di sei valori. Il modello che ha comportato il minor errore di validazione durante la fase di model selection della rete neurale, è costituito da 20 nodi nello strato nascosto. La struttura del multilayer perceptron che viene fuori dalle diverse prove della cross-validation interna è, in media, formata da:

- 6 nodi nello strato di input
- 20 nodi nello strato nascosto
- 1 nodo nello strato di output

Ogni sequenza di input è stata così normalizzata: la curva di relative enhancement è stata normalizzata in  $[-1, 1]$  secondo la regola 5.5, mentre ogni valore di early-increase è stato diviso per il massimo valore che tale parametro assume su tutto il dataset in modo da riportarlo nello stesso intervallo di valori della curva di relative enhancement.

Nelle tabelle 5.4b e 5.9a sono riportate l'accuratezza, la sensibilità e la specificità relative a tale configurazione di criteri. Come si vede nella tabella 5.4b, l'accuratezza (91.13%) ottenuta dalla rete neurale per tale combinazione di criteri è simile a quella ottenuta per le curve di relative enhancement da sole. Questo dipende dal fatto che il parametro early-increase non apporta informazione aggiuntiva alla curva, come si può notare nell'istogramma di Figura 5.7c, in cui è evidente la sovrapposizione tra pattern benigni e pattern maligni.

### 5.7.5 Relative Enhancement-Initial Slope

Tale configurazione rappresenta la combinazione dei criteri relative enhancement e initial slope. In fase di model selection è emerso che una struttura con 20 nodi nello strato nascosto produce il minor errore di validazione. Il multilayer perceptron costruito in seguito alle diverse prove della cross-validation interna è, in media, così formato:

- 6 nodi nello strato di input
- 20 nodi nello strato nascosto
- 1 nodo nello strato di output

Ogni vettore di input è stato normalizzato nel seguente modo: la curva di relative enhancement è stata normalizzata in  $[-1, 1]$  secondo la regola 5.5,

mentre il parametro initial slope è stato diviso per il massimo valore che tale parametro prende sull'intero dataset, in modo da riportarlo nell'intervallo  $[-1, 1]$  come i valori della curva.

Nelle tabelle 5.4c e 5.9b sono elencate le stime dell'accuratezza, sensibilità e specificità relative a tale combinazione di parametri. Come per la combinazione dei criteri relative enhancement ed early-increase, anche in questo caso l'accuratezza ottenuta (91.28%) non è molto aumentata rispetto a quella stimata per le curve di relative enhancement da sole. Il parametro initial slope, come è stato già spiegato, non è particolarmente discriminativo, essendo inoltre legato al parametro early-increase; entrambi, infatti, costituiscono una misura correlata alla fase di washin dell'agente di contrasto. La sensibilità è confrontabile con quella ottenuta per le curve di relative enhancement, mentre la specificità è leggermente superiore.

### 5.7.6 Relative Enhancement-Washout Ratio

Tale configurazione rappresenta la combinazione dei criteri relative enhancement e washout ratio. Il modello che, durante la fase di model selection, ha prodotto i migliori risultati sul validation set è costituito da 20 nodi nello strato nascosto. La struttura che emerge dalla cross-validation interna dopo le diverse prove di model selection, è un multilayer perceptron formato da:

- 6 nodi nello strato di input
- 20 nodi nello strato nascosto
- 1 nodo nello strato di output

Come per la configurazione relative enhancement-initial slope, ogni vettore di input è stato così normalizzato: la curva di relative enhancement è stata normalizzata in  $[-1, 1]$  secondo la regola 5.5, mentre il parametro washout ratio è stato diviso per il massimo valore che tale parametro prende sull'intero dataset.

Nelle tabelle 5.4d e 5.10a sono riportate le misure statistiche relative a tale configurazione di criteri dinamici in input alla rete neurale. Le stime espresse nelle tabelle mostrano che tale combinazione ha prodotto risultati interessanti. Come già spiegato in precedenza, sia il washout ratio che la curva di relative enhancement sono criteri che hanno evidenziato significativa abilità predittiva. In particolare accuratezza (92.55%) e specificità (95.6%) sono molto promettenti. Si noti nella tabella 5.10a che su quattro sottoinsiemi di test (Test1, Test5, Test7, Test10) tale combinazione di criteri raggiunge una specificità pari al 100%. La sensibilità è comunque superiore a quella

ottenuta con la rete neurale applicata alle curve di relative enhancement e a quella ottenuta con il classificatore a soglia con il washout. Inoltre è importante notare che il multilayer perceptron con tale configurazione di input ha classificato correttamente i due falsi negativi dei medici presenti all'interno del dataset di questa tesi.

### 5.7.7 WashoutCurve

WashoutCurve rappresenta il criterio washout ratio calcolato in tutte le serie post-contrasto ( $W_{peak-k}$ , con  $k \in [1, 5]$ ). Gli esami utilizzati in queste sperimentazioni sono composti da cinque serie post-contrasto, quindi il dataset per tale criterio è costituito da sequenze di cinque valori, corrispondenti al valore del washout ratio in ciascuna delle cinque serie post-contrasto. Il modello che ha prodotto i migliori risultati sul validation set, durante la fase di model selection, è costituito da 20 nodi nello strato nascosto. In conclusione, il multilayer perceptron in seguito alle diverse prove della cross-validation interna è, in media, così formato:

- 5 nodi nello strato di input
- 20 nodi nello strato nascosto
- 1 nodo nello strato di output

Ogni sequenza dei dati di input è stata normalizzata secondo la regola 5.5.

Nelle tabelle 5.5a e 5.10b sono riportati i risultati ottenuti per il criterio dinamico washoutCurve. L'accuratezza ottenuta dal comitato delle 9 reti neurali, pari al 91.08%, è confrontabile a quella ottenuta con le curve di relative enhancement ma non mostra miglioramenti. Anche le stime di sensibilità e specificità sono molto simili a quelle calcolate per la morfologia delle curve, quindi inferiori a quelle ottenute con la combinazione dei criteri relative enhancement e washout ratio.

### 5.7.8 Relative Enhancement-WashoutPeak4

Tale configurazione rappresenta la combinazione dei criteri dinamici relative enhancement e  $W_{peak-4}$ .  $W_{peak-4}$  indica il decremento tra la massima intensità del segnale di risonanza magnetica e l'intensità del segnale nella quarta serie post-contrasto. Il dataset costruito per questa configurazione di input è, quindi, composto di sequenze di sei valori. La struttura del multilayer perceptron che ha prodotto il minor errore di validazione è costituita da 20

nodi nello strato nascosto. Dunque la rete neurale che viene fuori dalle diverse prove della cross-validation interna per tale combinazione di criteri è, in media, formata da:

- 6 nodi nello strato di input
- 20 nodi nello strato nascosto
- 1 nodo nello strato di output

La curva di relative enhancement di ogni pattern di input è stata normalizzata secondo la regola 5.5, mentre il parametro  $W_{peak-4}$  di ogni sequenza è stato diviso per il massimo valore che tale parametro assume su tutto il dataset.

Nelle tabelle 5.5b e 5.11a sono mostrate le stime relative alla performance diagnostica ottenuta dal multilayer perceptron applicato a tale configurazione di input. Tali stime (accuratezza = 91.31%, sensibilità = 88.62%, specificità = 94%) mostrano che la morfologia delle curve combinata con il washout produce dei buoni risultati. Comunque la maggiore informazione diagnostica è contenuta nel washout ratio calcolato rispetto all'ultima serie post-contrasto, come si vede dal confronto tra i risultati.

### 5.7.9 Relative Enhancement-WashoutCurve

Come è stato già spiegato, entrambi i criteri, relative enhancement e washoutcurve, sono sequenze di cinque valori, corrispondenti rispettivamente ai valori dell'enhancement e del washout ratio nelle cinque serie post-contrasto. Il dataset utilizzato per tale configurazione è, quindi, composto da pattern di dieci valori. Il modello che ha prodotto i migliori risultati in fase di validazione è formato da 30 neuroni nello strato nascosto. La rete neurale che emerge dalle diverse prove della cross-validation interna ha, in media, la seguente struttura finale:

- 10 nodi nello strato di input
- 30 nodi nello strato nascosto
- 1 nodo nello strato di output

La normalizzazione effettuata su ogni sequenza di input è la seguente: entrambi i criteri sono stati normalizzati in  $[-1, 1]$  secondo la regola 5.5, ma utilizzando per ognuno dei criteri i corrispettivi valori massimo e minimo.

Nelle tabelle 5.5c e 5.11b sono riportati i risultati ottenuti con la rete neurale avente tale configurazione di input.

### 5.7.10 E-W-S

Tale configurazione rappresenta la combinazione dei criteri dinamici relative enhancement, washout ratio e initial slope. Il dataset in questione è composto da sequenze di sette valori. Il modello che ha prodotto il minor errore di validazione è quello con 20 nodi nello strato nascosto. La rete neurale applicata a tale combinazione di input, in seguito alle diverse prove della cross-validation interna, è, quindi, così strutturata:

- 7 nodi nello strato di input
- 20 nodi nello strato nascosto
- 1 nodo nello strato di output

Ogni sequenza di input è stata normalizzata nel seguente modo: alla curva di relative enhancement è stata applicata la regola 5.5, i parametri washout ratio e initial slope sono stati divisi rispettivamente per il massimo valore che ognuno di essi assume sull'intero dataset.

Nelle tabelle 5.5d e 5.12a sono indicate le stime ottenute dal multilayer perceptron applicato a tale configurazione di input. I risultati riportati nelle tabelle (accuratezza = 91.72%, sensibilità = 89.11%, specificità = 94.33%) mostrano che tale combinazione di criteri dinamici ha un buon potere predittivo, anche se l'incertezza introdotta dal criterio initial slope peggiora leggermente la situazione rispetto alla configurazione composta solo dalle curve di relative enhancement con il washout ratio. Si noti che per alcune suddivisioni di dati (Test5, Test7, Test10) questa combinazione di criteri raggiunge una specificità del 100%.

### 5.7.11 E-W-EI-S

Tale configurazione rappresenta la combinazione dei criteri relative enhancement, washout ratio, early-increase e initial slope. Il dataset costruito per tale combinazione è, dunque, composto da sequenze di otto valori. La struttura della rete neurale che ha prodotto i migliori risultati in fase di validazione è costituita da 20 nodi nello strato nascosto. La struttura che emerge dalle diverse prove della cross-validation interna è, in media, un multilayer perceptron così fatto:

- 8 nodi nello strato di input
- 20 nodi nello strato nascosto
- 1 nodo nello strato di output

La normalizzazione applicata ad ogni pattern di input è la seguente: la curva di relative enhancement è stata normalizzata in  $[-1, 1]$  secondo la regola 5.5, gli altri parametri quantitativi sono stati divisi rispettivamente per il massimo valore che ognuno di essi assume sull'intero dataset. In questo modo tutti i valori sono stati riportati nell'intervallo  $[-1, 1]$ .

Le tabelle 5.6a e 5.12b mostrano i risultati ottenuti con la rete neurale avente come input tale combinazione di criteri. L'accuratezza (90.33%) e la sensibilità (86.58%) ottenute evidenziano che la presenza dei criteri early-increase ed initial slope peggiora un po' la situazione rispetto ad altre configurazioni di input. Ciò è dovuto all'incertezza introdotta da questi due criteri che, come è stato visto negli istogrammi di Figura 5.7, comportano un'ampia sovrapposizione tra i pattern benigni e i pattern maligni.

### 5.7.12 E-C

Questa configurazione rappresenta la combinazione dei criteri dinamici relative enhancement e curvature at the peak of enhancement. Come già spiegato nella sezione 3.3, il parametro curvature at the peak of enhancement è un valore che esprime la curvatura della curva di relative enhancement nel suo picco. Il dataset costruito per tale combinazione di criteri è, quindi, composto da sequenze di sei valori. Il modello che ha prodotto i migliori risultati sul validation set è costituito da 20 nodi nello strato nascosto. In conclusione, il multilayer perceptron che emerge dalle diverse prove della cross-validation interna è, in media, così formato:

- 6 nodi nello strato di input
- 20 nodi nello strato nascosto
- 1 nodo nello strato di output

Per ogni sequenza di input, la curva di relative enhancement è stata normalizzata in  $[-1, 1]$  secondo la regola 5.5, mentre il valore della curvatura è stato diviso per il massimo valore che tale parametro prende sull'intero dataset.

Nelle tabelle 5.6b e 5.13a sono riportate le stime delle misure statistiche calcolate per tale configurazione di input. Tali stime (accuratezza = 90.64%, sensibilità = 87.17%, specificità = 94.11%) mostrano che la curvatura in combinazione con la morfologia delle curve, pur producendo dei risultati confrontabili con quelli ottenuti per le curve di relative enhancement da sole, introduce una certa confusione nella rete neurale dovuta alla considerevole sovrapposizione tra pattern benigni e pattern maligni, come è evidente dall'istogramma di Figura 5.7d.

### 5.7.13 E-S-C

Questa configurazione è data dalla combinazione dei criteri dinamici relative enhancement, initial slope e curvature at the peak of enhancement. Il dataset costruito è, quindi, composto da sequenze di sette valori. In fase di model selection è emerso che una struttura con 20 nodi nello strato nascosto produce il minor errore di validazione. Il modello su cui si sono calcolate le stime è così composto:

- 7 nodi nello strato di input
- 20 nodi nello strato nascosto
- 1 nodo nello strato di output

Ogni vettore di input è stato normalizzato nel seguente modo: la curva di relative enhancement è stata normalizzata in  $[-1, 1]$  secondo la regola 5.5, mentre i parametri initial slope e curvature at the peak of enhancement sono stati divisi per il massimo valore che ognuno di loro prende sull'intero dataset, in modo da riportarli entrambi nell'intervallo  $[-1, 1]$ , come la curva.

Nelle tabelle 5.6c e 5.13b sono mostrate accuratezza, sensibilità e specificità relative a tale configurazione di input.

### 5.7.14 E-W-C

Tale configurazione rappresenta la combinazione dei criteri dinamici relative enhancement, washout ratio e curvature at the peak of enhancement. Il dataset costruito in questo caso è, dunque, composto da sequenze di sette valori. Il modello che, in fase di model selection, ha prodotto il minor errore di validazione è costituito da 20 nodi nello strato nascosto. La struttura che emerge dalle diverse prove della cross-validation interna è, in media, un multilayer perceptron formato da:

- 7 nodi nello strato di input
- 20 nodi nello strato nascosto
- 1 nodo nello strato di output

Ogni vettore di input è stato normalizzato nel seguente modo: la curva di relative enhancement è stata normalizzata in  $[-1, 1]$  secondo la regola 5.5, mentre i parametri washout ratio e curvature at the peak of enhancement sono stati divisi per il massimo valore che rispettivamente prendono sull'intero dataset.

Nelle tabelle 5.6d e 5.14a sono riportate le stime calcolate sui multilayer perceptron allenati con questa combinazione di input. Tali stime (accuratezza = 91.67%, sensitività = 89.06%, specificità = 94.28%) mostrano che questa configurazione di criteri dinamici produce dei buoni risultati, anche se l'incertezza introdotta dal parametro curvature at the peak of enhancement diminuisce leggermente la performance diagnostica rispetto alla combinazione relative enhancement con washout ratio. Comunque su tre sottoinsiemi di test (Test5, Test7, Test10) si riesce ad ottenere una specificità del 100%.

### 5.7.15 S-EI-W-C

Tale configurazione rappresenta la combinazione dei criteri dinamici initial slope, early-increase, washout ratio e curvature at the peak of enhancement. In questo caso non si utilizza la morfologia della curva di relative enhancement ma solo parametri quantitativi, estratti dalla curva stessa, per testare la loro abilità discriminativa. Durante la fase di model selection, il modello che ha prodotto il minor errore di validazione è composto di 15 nodi nello strato nascosto. La struttura che viene fuori dalle diverse prove della cross-validation interna è, dunque, così formata:

- 4 nodi nello strato di input
- 15 nodi nello strato nascosto
- 1 nodo nello strato di output

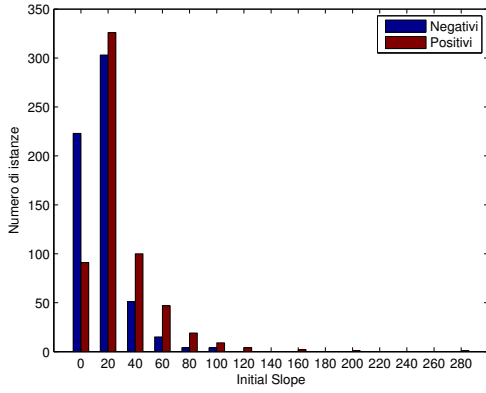
Ogni sequenza di input è stata normalizzata dividendo ognuno dei parametri della configurazione per il suo massimo valore sull'intero dataset, in modo da riportarli tutti nell'intervallo  $[-1, 1]$ .

Nelle tabelle 5.6a e 5.12b sono mostrati i risultati ottenuti con le sperimentazioni su tale combinazione di criteri dinamici in input. Le stime ottenute (accuratezza = 89.17%, sensitività = 86.08%, specificità = 92.25) sono inferiori a quelle ottenute con le curve di relative enhancement. Questo è dovuto al fatto che i parametri early-increase ed initial slope non hanno mostrato particolare potere predittivo mentre la morfologia delle curve, non presente in questa combinazione di criteri, ha in sé significativa informazione diagnostica.

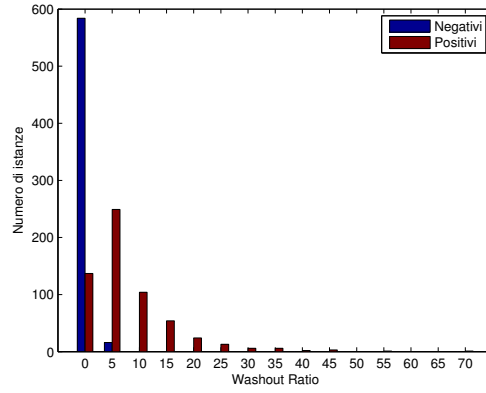
## 5.8 Discussione

Come precedentemente detto, due classificatori a soglia sono stati applicati ai criteri dinamici initial slope e washout ratio, a scopo di riferimento.

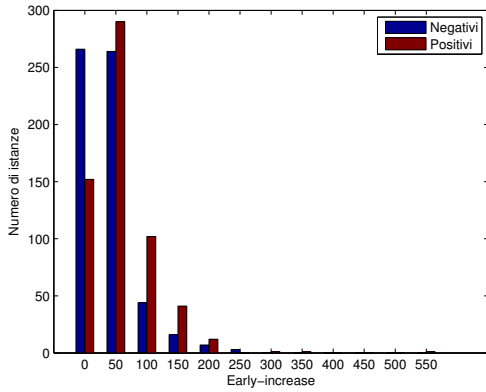




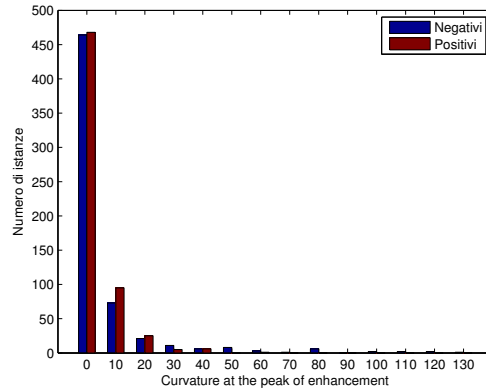
(a) Istogramma initial slope



(b) Istogramma washout ratio



(c) Istogramma Early-increase



(d) Istogramma Curvature at the peak of enhancement

Figura 5.7: Istogrammi dei criteri dinamici initial slope, washout ratio, early-increase e curvatura at the peak of enhancement.

L'istogramma di Figura 5.7a della distribuzione delle istanze di initial slope mostra che i pattern benigni e i pattern maligni sono concentrati nelle stesse regioni, quindi tale criterio non fornisce in sé una buona discriminazione, come risulta evidente anche dalle stime ottenute con il classificatore a soglia riportate nella tabella 5.3a. La soglia finale scelta dal classificatore è 14.99. Nell'istogramma si può vedere che tale soglia causa molti falsi negativi e falsi positivi sul dataset di studio in questa tesi. Nell'istogramma costruito per il criterio washout ratio, riportato in Figura 5.7b, si vede, invece, che i pattern benigni e maligni sono maggiormente separati rispetto a quanto accade per initial slope, pur mantenendo una zona di sovrapposizione. Il criterio washout ratio, quindi, contiene in sé importante informazione

discriminativa per la differenziazione delle lesioni, come dimostrano anche i risultati ottenuti con il classificatore a soglia riportati in tabella 5.3b. La soglia finale scelta dal classificatore per il criterio washout ratio è 1.05. Come si vede già dall'istogramma, questa soglia produce una specificità (90.50%) molto più alta di quella ottenuta con il criterio initial slope. I risultati prodotti dal classificatore a soglia applicato al criterio washout ratio sono stati confrontati con quelli ottenuti con le reti neurali per valutare se un modello più complesso del classificatore a soglia possa essere in grado di produrre una migliore performance diagnostica.

La morfologia della curva di relative enhancement, in input al multilayer perceptron, ha mostrato importante potere predittivo (tabelle 5.4a e 5.8), anche in accordo con i risultati della letteratura [60, 28]. La regola 5.5, che viene qui utilizzata per normalizzare i pattern di relative enhancement, pur essendo quella che ha prodotto i migliori risultati tra tutte le diverse regole di normalizzazione testate, ha, però, il problema che tende ad appiattire il dataset. Ciò vuol dire che curve di relative enhancement che differiscono tra di loro solo per una costante moltiplicativa, dopo la normalizzazione risultano essere lo stesso esperimento (in Figura 5.8 è mostrato un esempio di questo tipo di normalizzazione). A causa di tale appiattimento, viene preservata la forma della curva ma se ne perde l'altezza, che è un'importante informazione per quanto riguarda questo caso di studio in quanto rappresenta la velocità e l'intensità della diffusione dell'agente di contrasto nei tessuti. In Figura 5.9a e in Figura 5.9b sono mostrati i grafici dei pattern di relative enhancement classificati erroneamente dalla rete neurale. Come si vede dalla figura, le curve false positive e quelle false negative si trovano quasi tutte nella stessa regione. Inoltre, a causa della normalizzazione, la rete neurale non può tener conto della fase di washin dell'agente di contrasto. Infine la Figura 5.9b mostra che i casi classificati erroneamente come negativi dalla rete neurale appartengono, in effetti, alla categoria delle curve steady oppure a quella delle curve plateau (si veda la sezione 3.3.2). Per risolvere tali problemi vengono introdotti dei parametri quantitativi, in combinazione con la curva di relative enhancement in input al multilayer perceptron, che possano apportare maggiore informazione discriminativa.

Il parametro quantitativo early-increase indica la fase di washin dell'agente di contrasto, quindi potrebbe ridurre il problema dell'appiattimento del dataset dovuto alla normalizzazione. Nelle tabelle 5.4b e 5.9a, però, si vede che l'introduzione di tale parametro non apporta miglioramenti alla performance diagnostica. Questo è causato dal fatto che tale criterio non fornisce una buona discriminazione tra i pattern benigni e maligni, come è mostrato nell'istogramma di Figura 5.7c, in cui è evidente che le istanze sono distribuite all'interno delle stesse regioni.

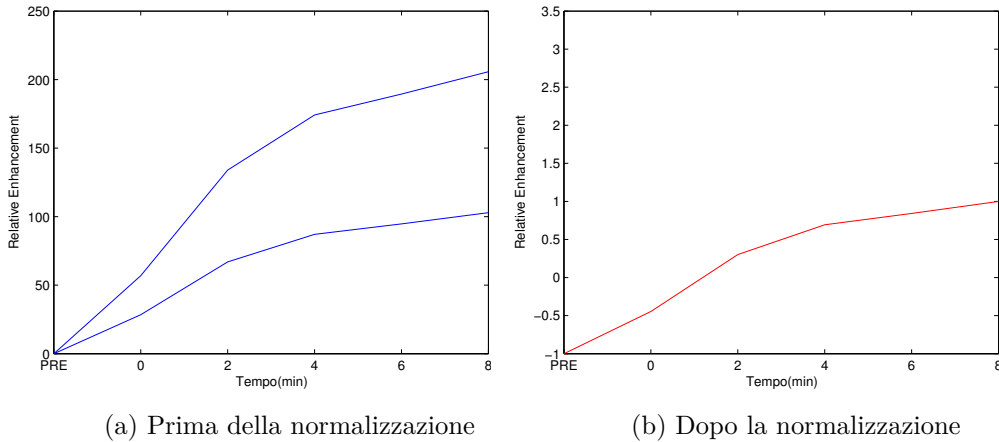


Figura 5.8: Esempio di curve di relative enhancement normalizzate con la regola 5.5. Le due curve di Figura 5.8a, dopo la normalizzazione, producono la stessa curva normalizzata (Figura 5.8b).

Il parametro initial slope, in combinazione con relative enhancement, non produce ottimi risultati in quanto, come è stato già visto nell'istogramma di Figura 5.7a, c'è un'ampia sovrapposizione tra i pattern benigni e i pattern maligni. Tale parametro è, infatti, correlato al parametro early-increase in quanto entrambi descrivono il comportamento del segnale nella prima fase post-contrasto.

L'introduzione del parametro washout ratio, invece, apporta dei miglioramenti alla performance diagnostica stimata con la morfologia delle curve. I risultati ottenuti con il comitato delle nove reti neurali sono i seguenti (tabella 5.4d e tabella 5.10a): accuratezza = 92.55%, sensibilità = 89.9%, specificità = 95.6%. La specificità, in particolare, è molto alta, infatti nella tabella sono riportati quattro sottoinsiemi di test in cui il multilayer perceptron raggiunge una specificità del 100%. La sensibilità, comunque, risulta maggiore del 98% su tre sottoinsiemi di test. La rete neurale, applicata alla combinazione del criterio washout ratio con le curve di relative enhancement, riesce a risolvere alcuni falsi positivi e alcuni falsi negativi. Come si vede dai dati indicati nella Figura 5.9, tale combinazione di criteri risolve 13 falsi negativi e 12 falsi positivi rispetto alle curve di relative enhancement da sole. Anche se non si tratta di numeri elevatissimi, se contestualizzati all'ambito dello studio di questa tesi, costituiscono comunque dei risultati importanti. La diminuzione dei falsi positivi, infatti, significa la diminuzione del numero di persone che verranno mandate ad eseguire una inutile biopsia, e, la diminuzione dei falsi negativi rappresenta l'aumento delle probabilità di sopravvivenza per

## 5. Sperimentazione

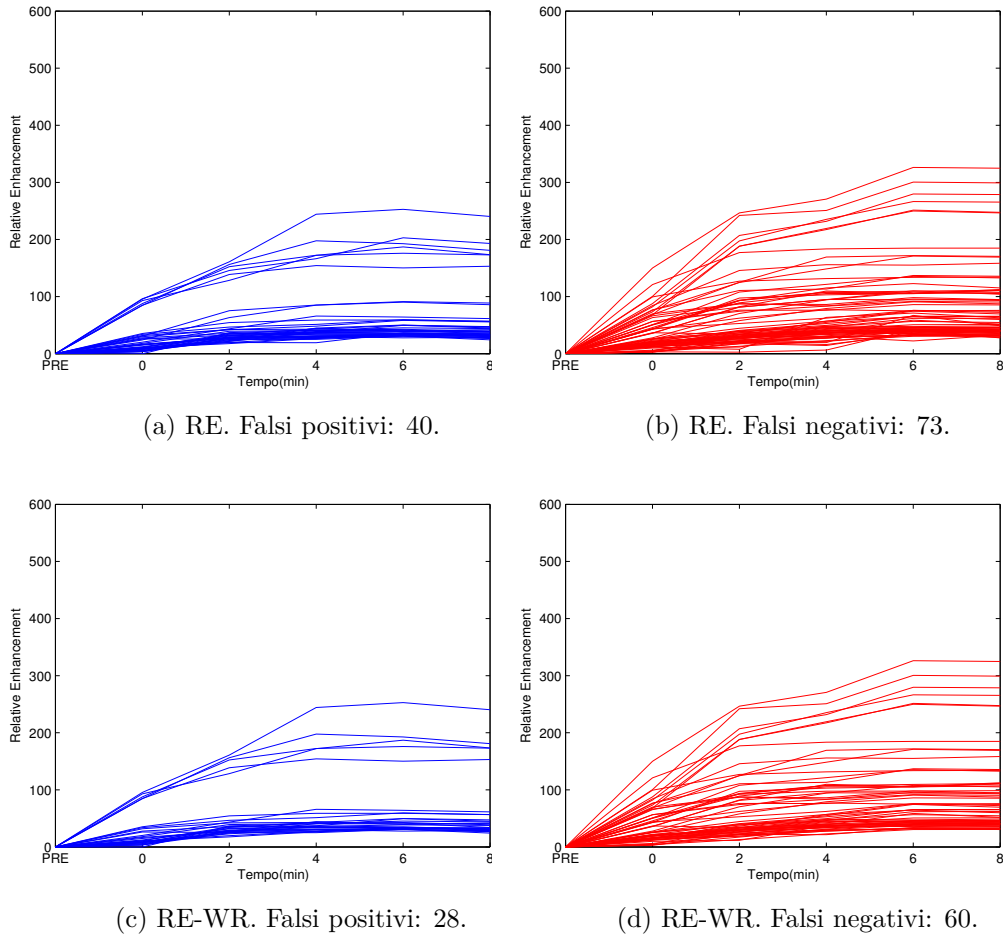


Figura 5.9: Pattern classificati erroneamente dalla rete neurale.

quelle persone a cui il cancro è stato diagnosticato correttamente. Inoltre è importante notare che tra i falsi negativi risolti dalla rete neurale con tale configurazione di criteri in input, vi sono anche i due falsi negativi dei medici presenti nel dataset di questa tesi. Dal grafico di Figura 5.9d risulta evidente che i pattern classificati erroneamente come negativi dalla rete neurale hanno il washout ratio uguale a zero, che è una caratteristica indicativa di benignità. Per risolvere questo problema sono state introdotte combinazioni di criteri basati sul washout ratio calcolato in diversi istanti di tempo. I risultati delle tabelle 5.5a, 5.10b, 5.5b, 5.11a, 5.5c e 5.11b, in cui sono riportate le stime delle combinazioni delle curve di relative enhancement con il washout ratio calcolato in diverse serie post-contrasto, mostrano che il washout ratio calcolato rispetto all'intensità del segnale dell'ultima immagine post-contrasto è il washout maggiormente predittivo. Tali combinazioni,

comunque, raggiungono la massima specificità (100%) e sensibilità maggiore del 98% su tre differenti test set.

La configurazione di input relative enhancement-washout ratio è stata anche estesa con l'aggiunta dei criteri initial slope e early-increase che contengono informazione riguardante la fase di washin dell'agente di contrasto. Tali combinazioni, come si vede dalle stime riportate nelle tabelle 5.5d, 5.12a, 5.6a e 5.12b, pur non migliorando i risultati ottenuti con la sola morfologia delle curve, sempre a causa dell'incertezza introdotta dai criteri early increase e initial slope, hanno riportato ugualmente risultati con essa confrontabili.

Le curve nel picco dell'enhancement possono piegarsi in maniera più o meno brusca, e questo può essere descritto matematicamente con la curvatura. Sono, quindi, state introdotte, in input alla rete neurale, le combinazioni che prevedono anche l'utilizzo della curvatura nel picco dell'enhancement. Tali combinazioni hanno prodotto risultati confrontabili con quelli ottenuti con le curve di relative enhancement da sole ma non hanno provocato sostanziali miglioramenti. Questo è dovuto al fatto che nel dataset utilizzato in questa tesi, relativamente al criterio curvature at the peak of enhancement, c'è un'ampia sovrapposizione tra pattern benigni e pattern maligni, come dimostra l'istogramma in Figura 5.7d. Le stime ottenute dalla combinazione dei criteri relative enhancement, washout ratio e curvature at the peak of enhancement (accuratezza = 91.67%, sensibilità = 89.06%, specificità = 94.28%) mostrano che tale combinazione produce dei buoni risultati nonostante l'incertezza introdotta dal parametro curvature at the peak of enhancement che diminuisce solo leggermente la performance diagnostica rispetto alla combinazione costituita solo dalla morfologia delle curve e il washout ratio.

Durante le sperimentazioni è stata testata anche una configurazione di input composta solo da parametri quantitativi. Tale combinazione non si è rivelata molto interessante, come dimostrano i risultati riportati in tabella 5.7 e in tabella 5.14b. Questo potrebbe essere dovuto all'incertezza causata dal dataset "ambizioso" utilizzato in questa tesi, ricco di casi difficili, e all'importanza della morfologia della curva di relative enhancement che è stata esclusa dalla configurazione.

In Figura 5.10 è mostrato l'istogramma delle accuratezze ottenute con il multilayer perceptron applicato alle diverse configurazioni di input. I risultati sono tutti maggiori del 90%, tranne la combinazione che utilizza solo i parametri quantitativi. Le configurazioni con initial slope e early increase hanno performance diagnostica più bassa dovuta all'incertezza che tali criteri introducono a causa della larga sovrapposizione tra i pattern positivi e i pattern negativi. La combinazione relative enhancement-washout ratio produce i risultati migliori, anche tenendo conto della deviazione standard. Dalle ac-

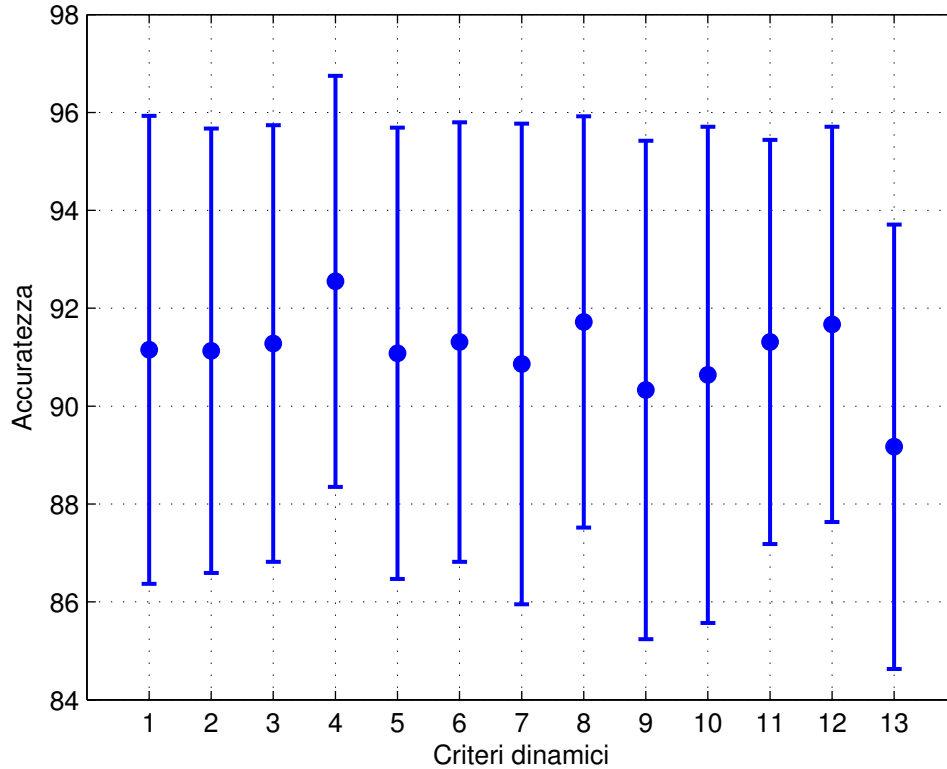
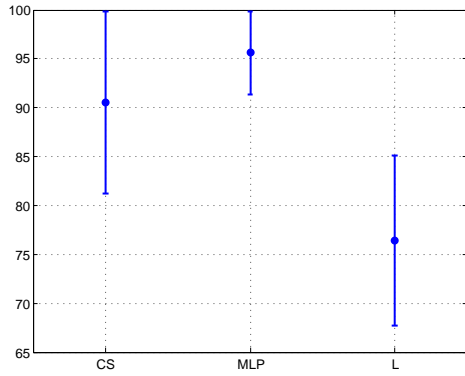


Figura 5.10: Istogramma delle accuratèzze ottenute dal multilayer perceptron sulle diverse combinazioni di input con le rispettive deviazioni standard. I numeri indicati sull'asse dei criteri dinamici, per migliorare la leggibilità del grafico, hanno le seguenti corrispondenze: 1 = E, 2 = E-EI, 3 = E-S, 4 = E-W, 5 = WC, 6 = E-W4, 7 = E-WC, 8 = E-W-S, 9 = E-W-EI-S, 10 = E-C, 11 = E-S-C, 12 = E-W-C, 13 = S-EI-W-C

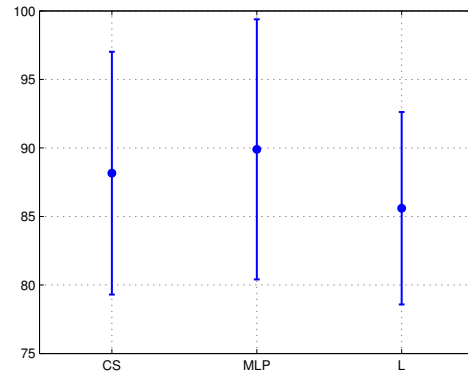
curatezze riportate nelle tabelle delle diverse combinazioni risulta che esiste un particolare sottoinsieme di test (Test5) piú “sfortunato” rispetto agli altri, che influisce negativamente sulla media e sulla deviazione standard. La combinazione costituita dalla morfologia delle curve con il washout ottiene i migliori risultati rispetto a tutte le altre combinazioni anche su questo particolare test set “difficile”. Questo comporta che anche la deviazione standard di tale combinazione è tra le piú basse e quindi il minimo valore della barra della deviazione standard, come si vede dall'istogramma, risulta comunque al di sopra dei range inferiori di tutte le altre combinazioni.

In Figura 5.11 vengono confrontate le stime di accuratezza, sensitività e specificità ottenute con il classificatore a soglia applicato al washout ratio, con

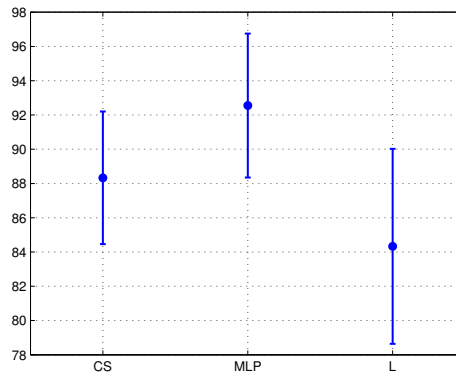
il multilayer perceptron applicato alla combinazione relative enhancement-washout ratio e dai risultati esistenti in letteratura [1, 34, 66, 68]. Bisogna precisare che il confronto con i risultati della letteratura è solo indicativo in quanto le sperimentazioni sono state eseguite su dataset differenti e, in alcuni casi, su criteri dinamici differenti. Dall'istogramma è evidente che con un modello di calcolo più complesso del classificatore a soglia, il multilayer perceptron, con un processo di validazione ampio ed accurato, la double k-fold cross-validation, è possibile realizzare un buon miglioramento nella performance diagnostica complessiva. La figura mostra che tale risultato è molto competitivo anche rispetto alle stime esistenti in letteratura per sperimentazioni simili. La sensibilità, ottenuta con il multilayer perceptron realizzato in questa tesi, non apporta miglioramenti significativi rispetto a quelle esistenti in letteratura, soprattutto a causa dell'alta deviazione standard, ma risulta comunque confrontabile con esse. Bisogna, infatti, ricordare che il dataset utilizzato in queste sperimentazioni è ricco di casi "difficili", molti dei quali sono stati, però, risolti dalla rete neurale, come precedentemente spiegato. La specificità ottenuta con il multilayer perceptron è confrontabile con quella ricavata con il classificatore a soglia, soprattutto considerando l'alta deviazione standard di quest'ultimo. Tale specificità, compresa di deviazione standard, risulta superiore alla stima media ottenuta con il classificatore a soglia. Essa è, invece, nettamente superiore alla stima della specificità ricavata dai dati della letteratura, anche tenendo presente la deviazione standard di entrambi. Questo è un importante risultato in quanto uno dei limiti delle immagini acquisite con la tecnica della risonanza magnetica è proprio la bassa specificità. Infine, anche l'accuratezza del multilayer perceptron mostra miglioramenti significativi sia rispetto al classificatore a soglia che rispetto alle stime della letteratura. La deviazione standard dei risultati ottenuti con il multilayer perceptron risulta, infatti, superiore alle stime medie degli altri classificatori.



(a) Specificità



(b) Sensitività



(c) Accuratezza

Figura 5.11: Istogrammi di accuratezza, sensitività e specificità ottenute con il classificatore a soglia (CS) applicato al criterio dinamico washout ratio, con il multilayer perceptron (MLP) applicato alla combinazione relative enhancement-washout e dai risultati presenti in letteratura (L) con le rispettive deviazioni standard.



<b>S</b>	Accuratezza	Sensitività	Specificità
Test1	70	91.67	48.34
Test2	53.34	35	71.67
Test3	48.34	68.34	28.34
Test4	75.84	85	66.67
Test5	70.84	58.34	83.34
Test6	70	71.67	68.34
Test7	57.5	61.67	53.34
Test8	75	91.67	58.34
Test9	47.5	46.67	48.34
Test10	58.34	50	66.67
$M \pm std$	$62.67 \pm 10.33$	$66 \pm 19.39$	$59.34 \pm 15.48$

(a) Initial Slope

<b>W</b>	Accuratezza	Sensitività	Specificità
Test1	95	95	95
Test2	94.17	95	93.33
Test3	85	80	90
Test4	86.67	81.67	91.67
Test5	85.83	71.67	100
Test6	89.17	83.34	95
Test7	93.33	88.33	98.33
Test8	86.67	98.33	75
Test9	83.34	98.33	68.33
Test10	89.16	80	98.33
$M \pm std$	$88.33 \pm 3.87$	$88.16 \pm 8.86$	$90.50 \pm 10.51$

(b) Washout Ratio

Tabella 5.3: Accuratezza (%), sensitività (%) e specificità (%), comprensive di deviazione standard, del classificatore a soglia sui criteri dinamici initial slope e washout ratio nelle dieci iterazioni della 10-fold cross-validation.

## 5. Sperimentazione

<b>E</b>	Accuratezza		<b>E-EI</b>	Accuratezza	
	Test	Comitato		Test	Comitato
Test1	98.43 ± 0.34	99.25 ± 0.61	Test1	98.38 ± 0.2	98.75 ± 0.59
Test2	88.7 ± 2.27	90.83 ± 1.11	Test2	90.65 ± 0.39	91.25 ± 0.59
Test3	92.98 ± 0.61	93.92 ± 0.79	Test3	91.16 ± 1.9	92.08 ± 0.59
Test4	89.39 ± 0.79	90.92 ± 0.73	Test4	91.62 ± 0.2	92.5 ± 0
Test5	81.88 ± 1.12	82.75 ± 1.11	Test5	83.24 ± 0.13	83.75 ± 0.59
Test6	86.07 ± 1.28	87.33 ± 1.02	Test6	86.62 ± 2.03	88.33 ± 1.18
Test7	96.52 ± 1.27	97.5 ± 0.39	Test7	97.04 ± 0.65	97.92 ± 0.59
Test8	89.56 ± 0.74	90.17 ± 0.66	Test8	89.49 ± 0.72	89.58 ± 0.59
Test9	88.18 ± 1.52	89.33 ± 2.03	Test9	87.64 ± 0.46	87.92 ± 0.59
Test10	88.28 ± 1.07	89.5 ± 0.81	Test10	88.29 ± 0.07	89.17 ± 0
<i>M ± std</i>	90 ± 4.91	91.15 ± 4.78	<i>M ± std</i>	90.41 ± 4.57	91.13 ± 4.54

(a)

(b)

<b>E-S</b>	Accuratezza		<b>E-W</b>	Accuratezza	
	Test	Comitato		Test	Comitato
Test1	98.12 ± 0.63	98.92 ± 0.56	Test1	98.84 ± 0.61	99.79 ± 0.42
Test2	89.65 ± 2.13	91.42 ± 1.47	Test2	92.69 ± 1.48	93.8 ± 1.44
Test3	92.49 ± 0.42	93 ± 0.81	Test3	92.55 ± 0.9	93.8 ± 1.44
Test4	90.67 ± 0.98	91.83 ± 0.86	Test4	90.63 ± 0.37	91.46 ± 0.48
Test5	82.54 ± 0.58	83.83 ± 0.58	Test5	84.7 ± 0.14	84.7 ± 0.14
Test6	86.98 ± 0.71	88.17 ± 1.17	Test6	90.46 ± 0.11	90.42 ± 0.48
Test7	97.19 ± 0.44	97.58 ± 0.26	Test7	96.57 ± 1.71	97.92 ± 0.48
Test8	89.63 ± 0.62	90.33 ± 0.43	Test8	91.97 ± 0.14	90.63 ± 0.42
Test9	88.81 ± 0.78	88.92 ± 1.18	Test9	92.62 ± 1.44	92.5 ± 0.68
Test10	88.09 ± 0.88	88.83 ± 0.7	Test10	91.3 ± 0.37	90.63 ± 0.42
<i>M ± std</i>	90.42 ± 4.63	91.28 ± 4.46	<i>M ± std</i>	92.23 ± 3.75	92.55 ± 4.2

(c)

(d)

Tabella 5.4: Accuratezza (%), comprensiva di deviazione standard, ottenuta con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement, relative enhancement-early-increase, relative enhancement-initial slope e relative enhancement-washout ratio.

<b>WC</b>	Accuratezza		<b>E-W4</b>	Accuratezza	
	Test	Comitato		Test	Comitato
Test1	98.35 ± 0.48	99.29 ± 0.31	Test1	98.15 ± 0.48	98.96 ± 0.42
Test2	90.33 ± 0.68	91.19 ± 0.81	Test2	90.25 ± 0.83	91.25 ± 0.83
Test3	92.99 ± 1.06	93.45 ± 0.31	Test3	93.29 ± 0.56	93.96 ± 0.42
Test4	88.06 ± 1.28	89.76 ± 1.15	Test4	90.42 ± 0.7	90.83 ± 1.18
Test5	82.96 ± 0.81	83.93 ± 0.41	Test5	82.8 ± 1.02	83.96 ± 0.42
Test6	86.46 ± 0.84	87.5 ± 1.36	Test6	87.15 ± 0.86	88.54 ± 1.25
Test7	96.43 ± 1.19	97.5 ± 0.48	Test7	96.32 ± 1.2	97.5 ± 0
Test8	89.72 ± 0.72	90.24 ± 0.41	Test8	89.98 ± 0.09	90.63 ± 0.42
Test9	87.62 ± 1.88	88.1 ± 1.25	Test9	88.1 ± 0.96	87.71 ± 0.42
Test10	89.01 ± 0.83	89.88 ± 0.31	Test10	89.19 ± 0.44	89.79 ± 0.42
<i>M ± std</i>	90.19 ± 4.62	91.08 ± 4.61	<i>M ± std</i>	90.57 ± 4.45	91.31 ± 4.49

(a)

(b)

<b>E-WC</b>	Accuratezza		<b>E-W-S</b>	Accuratezza	
	Test	Comitato		Test	Comitato
Test1	98.52 ± 0.16	99.44 ± 0.96	Test1	98.86 ± 0.3	99.44 ± 0.48
Test2	89.66 ± 0.27	90.83 ± 0	Test2	88.58 ± 1.93	91.11 ± 0.48
Test3	93.18 ± 0.05	93.33 ± 0	Test3	92.16 ± 0.86	92.5 ± 0.83
Test4	90.37 ± 0.48	90.56 ± 0.48	Test4	91.33 ± 0.8	91.67 ± 0.83
Test5	81.3 ± 1.44	81.94 ± 1.92	Test5	81.36 ± 2.78	81.94 ± 1.92
Test6	86.42 ± 0.43	87.22 ± 0.48	Test6	87.84 ± 0.68	88.61 ± 1.27
Test7	95.9 ± 1.39	96.94 ± 0.48	Test7	95.8 ± 1.09	97.22 ± 0.48
Test8	89.54 ± 0.32	90 ± 0	Test8	90.52 ± 0.19	90.28 ± 0.48
Test9	89.17 ± 0.8	88.61 ± 1.92	Test9	88.67 ± 1.37	88.06 ± 1.73
Test10	88.49 ± 0.05	89.72 ± 0.48	Test10	95.46 ± 6.2	96.39 ± 6.25
<i>M ± std</i>	90.26 ± 4.82	90.86 ± 4.91	<i>M ± std</i>	91.06 ± 4.95	91.72 ± 5.09

(c)

(d)

Tabella 5.5: Accuratezza (%), comprensiva di deviazione standard, ottenuta con il multilayer perceptron applicato alle combinazioni di criteri washout-curve, relative enhancement- $W_{peak-4}$ , relative enhancement-washoutcurve, relative enhancement-washout ratio-initial slope.

## 5. Sperimentazione

---

<b>E-W-EI-S</b>	Accuratezza		<b>E-C</b>	Accuratezza	
	Test	Comitato		Test	Comitato
Test1	$98.7 \pm 0.13$	$100 \pm 0$	Test1	$97.56 \pm 1.08$	$98.06 \pm 1.27$
Test2	$90.93 \pm 1.44$	$93.75 \pm 0.59$	Test2	$90.28 \pm 3.89$	$92.5 \pm 1.67$
Test3	$88.24 \pm 1.31$	$89.17 \pm 1.18$	Test3	$88.52 \pm 4.77$	$91.39 \pm 2.41$
Test4	$91.67 \pm 0.65$	$92.92 \pm 0.59$	Test4	$90.15 \pm 0.54$	$91.39 \pm 0.48$
Test5	$77.64 \pm 0.33$	$77.5 \pm 1.18$	Test5	$79.91 \pm 2.89$	$81.11 \pm 2.93$
Test6	$85.42 \pm 0.2$	$85.83 \pm 0$	Test6	$84.44 \pm 2.47$	$85.56 \pm 1.73$
Test7	$96.99 \pm 0.46$	$97.5 \pm 0$	Test7	$96.7 \pm 1.55$	$97.78 \pm 0.48$
Test8	$90.05 \pm 0.46$	$90 \pm 0$	Test8	$89.07 \pm 1.13$	$89.72 \pm 0.96$
Test9	$87.27 \pm 0.72$	$87.08 \pm 0.59$	Test9	$89.63 \pm 0.94$	$90 \pm 1.67$
Test10	$87.04 \pm 1.18$	$89.58 \pm 0.59$	Test10	$87.99 \pm 1.26$	$88.89 \pm 0.48$
$M \pm std$	$89.4 \pm 5.94$	$90.33 \pm 6.33$	$M \pm std$	$89.43 \pm 5.15$	$90.64 \pm 5.07$

(a)

(b)

<b>E-S-C</b>	Accuratezza		<b>E-W-C</b>	Accuratezza	
	Test	Comitato		Test	Comitato
Test1	$96.67 \pm 1.98$	$97.5 \pm 0.83$	Test1	$97.19 \pm 0.19$	$97.78 \pm 0.96$
Test2	$90.43 \pm 0.84$	$91.67 \pm 0.83$	Test2	$89.66 \pm 1.02$	$91.67 \pm 0$
Test3	$92.07 \pm 0.79$	$92.78 \pm 1.27$	Test3	$92.72 \pm 0.42$	$93.89 \pm 0.48$
Test4	$90.71 \pm 1.18$	$91.39 \pm 1.27$	Test4	$90.9 \pm 0.48$	$91.11 \pm 0.48$
Test5	$82.78 \pm 0$	$83.89 \pm 0.48$	Test5	$84.51 \pm 1.66$	$84.17 \pm 0$
Test6	$87.96 \pm 0.98$	$88.89 \pm 0.48$	Test6	$88.95 \pm 2.14$	$90 \pm 1.44$
Test7	$96.54 \pm 1.21$	$97.78 \pm 0.48$	Test7	$97.5 \pm 0.49$	$97.78 \pm 0.48$
Test8	$90.37 \pm 0.16$	$90.83 \pm 0$	Test8	$88.67 \pm 2.3$	$90.56 \pm 0.48$
Test9	$88.8 \pm 0.52$	$89.72 \pm 0.48$	Test9	$88.92 \pm 0.8$	$89.72 \pm 0.48$
Test10	$88.12 \pm 1.02$	$88.61 \pm 0.96$	Test10	$89.66 \pm 1.26$	$90 \pm 0.83$
$M \pm std$	$90.45 \pm 4.1$	$91.31 \pm 4.13$	$M \pm std$	$90.87 \pm 3.99$	$91.67 \pm 4.04$

(c)

(d)

Tabella 5.6: Accuratezza (%), comprensiva di deviazione standard, ottenuta con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-washout ratio-early-increase-initial slope, relative enhancement-initial slope-curvature at the peak of enhancement, relative enhancement-washout ratio-curvature at the peak of enhancement.

S-EI-W-C	Accuratezza	
	Test	Comitato
Test1	96.32 ± 0.16	96.04 ± 0.42
Test2	92.43 ± 0.63	93.33 ± 0.96
Test3	81.46 ± 0.6	81.04 ± 0.42
Test4	88.13 ± 0.49	89.17 ± 0
Test5	85.3 ± 0.09	85.83 ± 0
Test6	89.35 ± 0.13	90 ± 0
Test7	93.22 ± 0.25	94.38 ± 0.42
Test8	89.4 ± 0.49	89.17 ± 0
Test9	86.18 ± 0.4	86.04 ± 0.8
Test10	84.03 ± 3.28	86.67 ± 1.18
<i>M ± std</i>	88.58 ± 4.54	89.17 ± 4.54

Tabella 5.7: Accuratezza (%), comprensiva di deviazione standard, della combinazione di criteri initial slope-early-increase-washout ratio-curvature

E	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	98.52 ± 0.59	99.67 ± 0.7	98.24 ± 0.64	98.83 ± 0.81
Test2	81.54 ± 3.97	83.67 ± 1.89	93.59 ± 7.77	96.5 ± 4.68
Test3	91.24 ± 1.23	93 ± 1.31	94.57 ± 1.17	94.83 ± 0.95
Test4	89.89 ± 2.08	90.67 ± 0.86	89.3 ± 2.05	91.33 ± 2.05
Test5	63.89 ± 2.29	65.5 ± 2.23	96.28 ± 11.32	96.67 ± 10.54
Test6	75.22 ± 2.53	77.67 ± 1.96	94.93 ± 5.96	95.33 ± 5.43
Test7	93.39 ± 2.42	95 ± 0.79	99.15 ± 1.61	99.5 ± 1.58
Test8	97.57 ± 0.75	98.67 ± 1.31	82.98 ± 5.32	83.5 ± 5.8
Test9	96.69 ± 2.07	98.33 ± 0	81.63 ± 6.11	82.5 ± 6.77
Test10	78.5 ± 2.26	79.5 ± 1.37	96.5 ± 5.57	97.67 ± 5.68
<i>M ± std</i>	86.65 ± 11.45	88.17 ± 11.26	92.72 ± 6.12	93.67 ± 6.07

Tabella 5.8: Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle curve di relative enhancement.

## 5. Sperimentazione

---

<b>E-EI</b>	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	$98.7 \pm 0.26$	$99.17 \pm 1.18$	$98.06 \pm 0.13$	$98.33 \pm 0$
Test2	$84.35 \pm 0.65$	$85 \pm 0$	$96.94 \pm 0.13$	$97.5 \pm 1.18$
Test3	$89.63 \pm 3.93$	$91.67 \pm 0$	$92.69 \pm 0.13$	$92.5 \pm 1.18$
Test4	$90 \pm 0$	$90 \pm 0$	$93.24 \pm 0.39$	$95 \pm 0$
Test5	$66.48 \pm 0.26$	$67.5 \pm 1.18$	$100 \pm 0$	$100 \pm 0$
Test6	$76.2 \pm 1.7$	$78.33 \pm 2.36$	$97.04 \pm 2.36$	$98.33 \pm 0$
Test7	$94.81 \pm 1.31$	$95.83 \pm 1.18$	$99.26 \pm 0$	$100 \pm 0$
Test8	$97.31 \pm 1.44$	$97.5 \pm 1.18$	$81.67 \pm 0$	$81.67 \pm 0$
Test9	$98.24 \pm 0.13$	$98.33 \pm 0$	$77.04 \pm 1.05$	$77.5 \pm 1.18$
Test10	$77.22 \pm 0.52$	$89.26 \pm 10.66$	$92.88 \pm 8.13$	$100 \pm 0$
<i>M ± std</i>	$87.29 \pm 10.98$	$89.26 \pm 10.05$	$92.88 \pm 7.66$	$94.08 \pm 8.06$

(a)

<b>E-S</b>	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	$98.39 \pm 0.63$	$99.5 \pm 0.81$	$97.85 \pm 0.85$	$98.33 \pm 0.79$
Test2	$84.11 \pm 3.11$	$85.5 \pm 2.84$	$95.19 \pm 1.87$	$97.33 \pm 0.86$
Test3	$91.06 \pm 0.92$	$92.17 \pm 1.12$	$93.93 \pm 0.49$	$93.83 \pm 0.81$
Test4	$89.57 \pm 2.33$	$90.17 \pm 0.53$	$91.76 \pm 1.66$	$93.5 \pm 2.14$
Test5	$65.07 \pm 1.16$	$67.67 \pm 1.17$	$100 \pm 0$	$100 \pm 0$
Test6	$76.87 \pm 1.73$	$78.17 \pm 2.54$	$97.09 \pm 0.92$	$98.17 \pm 0.53$
Test7	$94.67 \pm 0.86$	$95.17 \pm 0.53$	$99.7 \pm 0.2$	$100 \pm 0$
Test8	$97.78 \pm 0.91$	$99 \pm 0.86$	$81.48 \pm 0.66$	$81.67 \pm 0$
Test9	$97.83 \pm 0.77$	$98.33 \pm 0$	$79.8 \pm 1.79$	$79.5 \pm 2.36$
Test10	$76.83 \pm 1.76$	$77.67 \pm 1.41$	$99.35 \pm 0.5$	$100 \pm 0$
<i>M ± std</i>	$87.22 \pm 11.25$	$88.34 \pm 10.83$	$93.62 \pm 7.34$	$94.23 \pm 7.58$

(b)

Tabella 5.9: Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-early-increase, relative enhancement-initial slope.

E-W	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	$98.61 \pm 1.22$	$99.58 \pm 0.83$	$99.07 \pm 0$	$100 \pm 0$
Test2	$66.73 \pm 40.46$	$88.96 \pm 2.67$	$97.96 \pm 0$	$98.54 \pm 0.42$
Test3	$90.32 \pm 1.87$	$92.5 \pm 1.67$	$94.77 \pm 0.35$	$95 \pm 1.36$
Test4	$90.23 \pm 0.78$	$90.23 \pm 0.78$	$91.02 \pm 0.56$	$92.98 \pm 0.83$
Test5	$69.4 \pm 0.28$	$69.4 \pm 0.28$	$100 \pm 0$	$100 \pm 0$
Test6	$83.06 \pm 0.19$	$83.33 \pm 0$	$97.78 \pm 0.37$	$97.5 \pm 0.96$
Test7	$93.24 \pm 3.53$	$95.83 \pm 0.96$	$99.95 \pm 0.093$	$100 \pm 0$
Test8	$98.98 \pm 0.19$	$98.52 \pm 0.37$	$84.95 \pm 0.093$	$84.95 \pm 0.09$
Test9	$98.29 \pm 0.28$	$98.54 \pm 0.42$	$86.25 \pm 4.54$	$86.79 \pm 0.25$
Test10	$82.59 \pm 0.74$	$82.08 \pm 2.1$	$100 \pm 0$	$100 \pm 0$
$M \pm std$	$87.15 \pm 11.65$	$89.9 \pm 9.49$	$95.18 \pm 5.8$	$95.6 \pm 5.66$

(a)

WC	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	$98.17 \pm 1.16$	$99.76 \pm 0.63$	$97.88 \pm 1.2$	$98.81 \pm 0.81$
Test2	$83.97 \pm 1.43$	$84.05 \pm 1.63$	$96.19 \pm 2.59$	$98.1 \pm 0.63$
Test3	$90.74 \pm 2.54$	$92.62 \pm 1.31$	$95.13 \pm 0.88$	$94.29 \pm 1.31$
Test4	$88.23 \pm 1.78$	$88.57 \pm 0.63$	$88.78 \pm 2.05$	$90.48 \pm 2.49$
Test5	$65.95 \pm 1.61$	$68.1 \pm 0.63$	$100 \pm 0$	$100 \pm 0$
Test6	$75.93 \pm 1.84$	$77.62 \pm 2.86$	$96.72 \pm 1.29$	$97.14 \pm 0.81$
Test7	$93.2 \pm 2.44$	$94.76 \pm 1.15$	$99.55 \pm 0.24$	$100 \pm 0$
Test8	$97.88 \pm 0.73$	$98.81 \pm 0.81$	$81.85 \pm 0.76$	$81.67 \pm 0$
Test9	$96.32 \pm 2.66$	$98.33 \pm 0$	$78.97 \pm 2.31$	$79.05 \pm 3.58$
Test10	$79.07 \pm 1.1$	$79.52 \pm 0.81$	$98.47 \pm 1.5$	$100 \pm 0$
$M \pm std$	$86.95 \pm 10.65$	$88.21 \pm 10.65$	$93.35 \pm 7.54$	$93.95 \pm 7.79$

(b)

Tabella 5.10: Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-washout ratio, washoutcurve.

## 5. Sperimentazione

---

<b>E-W4</b>	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	$97.69 \pm 1.07$	$98.75 \pm 0.83$	$98.61 \pm 0.39$	$99.17 \pm 0.96$
Test2	$83.7 \pm 0.76$	$84.58 \pm 1.6$	$96.81 \pm 1.77$	$97.92 \pm 0.83$
Test3	$92.27 \pm 1.77$	$93.33 \pm 2.36$	$94.31 \pm 1.14$	$94.58 \pm 1.6$
Test4	$90.56 \pm 0.59$	$90 \pm 0$	$90.28 \pm 1.85$	$91.67 \pm 2.36$
Test5	$65.6 \pm 2.04$	$67.92 \pm 0.83$	$100 \pm 0$	$100 \pm 0$
Test6	$77.18 \pm 1.95$	$79.17 \pm 2.89$	$97.13 \pm 1.05$	$97.92 \pm 0.83$
Test7	$92.87 \pm 2.54$	$95 \pm 0$	$99.77 \pm 0.23$	$100 \pm 0$
Test8	$98.43 \pm 0.44$	$99.58 \pm 0.83$	$81.53 \pm 0.38$	$81.67 \pm 0$
Test9	$97.82 \pm 0.63$	$98.33 \pm 0$	$78.38 \pm 1.41$	$77.08 \pm 0.83$
Test10	$78.94 \pm 0.98$	$79.58 \pm 0.83$	$99.44 \pm 0.34$	$100 \pm 0$
<i>M ± std</i>	$87.51 \pm 10.88$	$88.62 \pm 10.53$	$93.63 \pm 7.82$	$94 \pm 8.23$

(a)

<b>E-WC</b>	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	$98.27 \pm 0.11$	$99.44 \pm 0.96$	$98.77 \pm 0.21$	$99.44 \pm 0.96$
Test2	$82.04 \pm 1.6$	$83.33 \pm 0$	$97.28 \pm 2.14$	$98.33 \pm 0$
Test3	$92.41 \pm 0.96$	$93.33 \pm 0$	$93.95 \pm 1.07$	$93.33 \pm 0$
Test4	$90.8 \pm 0.11$	$90 \pm 0$	$89.94 \pm 1.07$	$91.11 \pm 0.96$
Test5	$62.59 \pm 2.89$	$63.89 \pm 3.85$	$100 \pm 0$	$100 \pm 0$
Test6	$75.8 \pm 1.18$	$77.78 \pm 0.96$	$97.04 \pm 0.32$	$96.67 \pm 0$
Test7	$92.16 \pm 2.46$	$93.89 \pm 0.96$	$99.63 \pm 0.32$	$100 \pm 0$
Test8	$97.28 \pm 0.75$	$98.33 \pm 0$	$81.79 \pm 0.11$	$82.22 \pm 0.96$
Test9	$98.09 \pm 0.11$	$98.33 \pm 0$	$80.25 \pm 1.71$	$78.33 \pm 2.89$
Test10	$77.35 \pm 0.43$	$79.44 \pm 0.96$	$99.63 \pm 0.32$	$100 \pm 0$
<i>M ± std</i>	$86.68 \pm 11.85$	$87.78 \pm 11.51$	$93.83 \pm 7.43$	$93.94 \pm 7.86$

(b)

Tabella 5.11: Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement- $W_{peak-4}$ , relative enhancement-washoutcurve.



<b>E-W-S</b>	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	98.7 ± 0.32	99.44 ± 0.96	99.01 ± 0.47	99.44 ± 0.96
Test2	80.37 ± 2.73	83.89 ± 0.96	96.79 ± 1.62	98.33 ± 0
Test3	91.54 ± 0.47	91.11 ± 0.96	92.78 ± 2.09	93.89 ± 2.55
Test4	89.81 ± 0.98	90 ± 0	92.84 ± 2.5	93.33 ± 1.67
Test5	62.72 ± 5.56	63.89 ± 3.85	100 ± 0	100 ± 0
Test6	78.09 ± 0.75	79.44 ± 2.55	97.59 ± 0.64	97.78 ± 0.96
Test7	92.35 ± 2	94.44 ± 0.96	99.26 ± 0.19	100 ± 0
Test8	97.96 ± 1.03	97.78 ± 0.96	83.09 ± 1.23	82.78 ± 0.96
Test9	97.9 ± 0.43	98.33 ± 0	79.44 ± 3.03	77.78 ± 3.47
Test10	91.85 ± 12.51	92.78 ± 12.51	99.07 ± 0.67	100 ± 0
<i>M ± std</i>	88.13 ± 11.34	89.11 ± 10.91	93.99 ± 7.22	94.33 ± 7.88

(a)

<b>E-W-EI-S</b>	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	98.61 ± 0.13	100 ± 0	98.8 ± 0.39	100 ± 0
Test2	86.11 ± 0.79	89.17 ± 1.18	95.74 ± 2.1	98.33 ± 0
Test3	82.78 ± 1.57	84.17 ± 1.18	93.7 ± 1.05	94.17 ± 1.18
Test4	90.28 ± 0.13	90 ± 0	93.06 ± 1.44	95.83 ± 1.18
Test5	55.28 ± 0.65	55 ± 2.36	100 ± 0	100 ± 0
Test6	73.15 ± 0	73.33 ± 0	97.69 ± 0.39	98.33 ± 0
Test7	95.09 ± 0.65	96.67 ± 0	98.89 ± 0.26	98.33 ± 0
Test8	97.78 ± 1.05	98.33 ± 0	82.31 ± 0.13	81.67 ± 0
Test9	97.78 ± 0.26	98.33 ± 0	76.76 ± 1.7	75.83 ± 1.18
Test10	79.17 ± 1.44	80.83 ± 1.18	94.91 ± 3.8	98.33 ± 0
<i>M ± std</i>	85.6 ± 13.75	86.58 ± 14.1	93.19 ± 7.66	94.08 ± 8.38

(b)

Tabella 5.12: Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-washout ratio-initial slope, relative enhancement-washout ratio-early-increase-initial slope.

<b>E-C</b>	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	97.84 ± 0.39	98.33 ± 0	97.28 ± 2.17	97.78 ± 2.55
Test2	84.26 ± 4.94	86.67 ± 3.33	96.3 ± 2.89	98.33 ± 0
Test3	85.8 ± 4.76	87.78 ± 5.09	91.23 ± 6.06	95 ± 1.67
Test4	90.19 ± 1.21	90.56 ± 0.96	90.12 ± 1.6	92.22 ± 0.96
Test5	60.86 ± 5.06	62.22 ± 5.85	98.95 ± 1.82	100 ± 0
Test6	72.35 ± 4.32	73.89 ± 4.19	96.54 ± 0.7	97.22 ± 0.96
Test7	94.38 ± 2.06	96.11 ± 0.96	99.01 ± 1.13	99.44 ± 0.96
Test8	97.59 ± 0.81	98.89 ± 0.96	80.56 ± 1.45	80.56 ± 0.96
Test9	97.65 ± 0.53	98.33 ± 0	81.6 ± 2.32	81.67 ± 3.33
Test10	77.9 ± 0.91	78.89 ± 0.96	98.09 ± 1.78	98.89 ± 0.96
<i>M ± std</i>	85.88 ± 12.4	87.17 ± 12.23	92.97 ± 6.95	94.11 ± 7.22

(a)

<b>E-S-C</b>	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	96.6 ± 2.21	97.22 ± 1.92	96.73 ± 1.78	97.78 ± 0.96
Test2	84.38 ± 0.6	85.56 ± 0.96	96.48 ± 1.11	97.78 ± 0.96
Test3	91.11 ± 2.31	92.22 ± 2.55	93.02 ± 1.66	93.33 ± 0
Test4	90.37 ± 0.67	90 ± 0	91.05 ± 2.99	92.78 ± 2.55
Test5	65.62 ± 0.11	67.78 ± 0.96	99.94 ± 0.11	100 ± 0
Test6	78.58 ± 1.4	79.44 ± 0.96	97.35 ± 0.75	98.33 ± 0
Test7	93.33 ± 2.43	95.56 ± 0.96	99.75 ± 0.11	100 ± 0
Test8	98.7 ± 0.49	100 ± 0	82.04 ± 0.67	81.67 ± 0
Test9	97.78 ± 0.64	98.33 ± 0	79.81 ± 0.49	81.11 ± 0.96
Test10	76.3 ± 1.95	77.22 ± 1.92	99.94 ± 0.11	100 ± 0
<i>M ± std</i>	87.28 ± 10.86	88.33 ± 10.63	93.61 ± 7.31	94.28 ± 7.26

(b)

Tabella 5.13: Sensitività(%) e specificità(%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-curvature at the peak of enhancement, relative enhancement-initial slope-curvature at the peak of enhancement.

<b>E-W-C</b>	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	96.85 ± 0.49	97.78 ± 0.96	97.53 ± 0.21	97.78 ± 0.96
Test2	83.46 ± 0.47	85.56 ± 0.96	95.86 ± 2.03	97.78 ± 0.96
Test3	91.6 ± 0.53	93.33 ± 0	93.83 ± 0.43	94.44 ± 0.96
Test4	90 ± 0.49	90 ± 0	91.79 ± 0.57	92.22 ± 0.96
Test5	69.14 ± 3.42	68.33 ± 0	99.88 ± 0.11	100 ± 0
Test6	80.49 ± 2.99	82.22 ± 1.92	97.41 ± 1.28	97.78 ± 0.96
Test7	95.25 ± 0.75	95.56 ± 0.96	99.75 ± 0.28	100 ± 0
Test8	97.04 ± 2.57	99.44 ± 0.96	80.31 ± 2.04	81.67 ± 0
Test9	98.27 ± 0.11	98.33 ± 0	79.57 ± 1.5	81.11 ± 0.96
Test10	79.51 ± 2.51	80 ± 1.67	99.81 ± 0	100 ± 0
<i>M ± std</i>	88.16 ± 9.66	89.06 ± 10.01	93.57 ± 7.66	94.28 ± 7.24

(a)

<b>S-EI-W-C</b>	Sensitività		Specificità	
	Test	Comitato	Test	Comitato
Test1	96.06 ± 0.18	95.42 ± 0.83	96.57 ± 0.19	97.08 ± 0.83
Test2	88.52 ± 0.34	89.58 ± 0.83	96.34 ± 1.28	96.67 ± 2.36
Test3	82.04 ± 0.66	82.08 ± 0.83	80.88 ± 0.92	80 ± 0
Test4	80.14 ± 1.09	81.67 ± 0	96.11 ± 0.15	97.5 ± 1.67
Test5	70.6 ± 0.18	71.67 ± 0	100 ± 0	99.17 ± 1.67
Test6	79.12 ± 0.46	80 ± 0	99.58 ± 0.23	99.17 ± 1.67
Test7	91.25 ± 0.38	92.08 ± 0.83	95.19 ± 0.55	97.5 ± 1.67
Test8	96.02 ± 0.78	96.67 ± 0	82.78 ± 0.34	81.67 ± 0
Test9	97.87 ± 0.32	98.33 ± 0	74.49 ± 0.87	73.75 ± 1.6
Test10	71.48 ± 2.09	73.33 ± 2.36	96.57 ± 4.98	100 ± 0
<i>M ± std</i>	85.31 ± 10.09	86.08 ± 9.66	91.85 ± 8.97	92.25 ± 9.76

(b)

Tabella 5.14: Sensitività (%) e specificità (%) ottenute con il multilayer perceptron applicato alle combinazioni di criteri relative enhancement-washout ratio-curvature at the peak of enhancement, initial slope-early-increase-washout ratio-curvature at the peak of enhancement.



# Capitolo 6

## Conclusioni

In questa tesi è stato realizzato un sistema per la classificazione delle immagini ottenute da un esame di risonanza magnetica per fornire ai medici un supporto automatico nell'elaborazione della diagnosi del cancro al seno.

Il cancro al seno è tra le maggiori cause di mortalità femminile nei paesi industrializzati. La diagnosi precoce del tumore aiuta a ridurre la necessità di trattamenti terapeutici radicali ed aumenta le probabilità di sopravvivenza. La risonanza magnetica in combinazione con l'agente di contrasto (Contrast-Enhanced MRI) costituisce una valida modalità complementare ai tradizionali metodi diagnostici (mammografia ed ecografia) per l'individuazione e la diagnosi del cancro al seno. La risonanza magnetica mostra i tumori maligni evidenziando, per mezzo della diffusione dell'agente di contrasto, la loro vascolarizzazione patologica. Tale tecnica comporta l'acquisizione di un elevato numero di immagini per fornire informazioni sulle caratteristiche dinamiche e morfologiche delle lesioni sospette. La formulazione di una diagnosi richiede l'analisi e l'interpretazione di tutte le immagini appartenenti ad un esame. Questo processo necessita di un notevole impiego di tempo ed è soggetto ad errori a causa della complessa struttura radiografica del seno, della difficile individuazione delle caratteristiche del cancro nelle sue fasi iniziali, e della stanchezza e della distrazione del radiologo. I sistemi Computer-Aided Diagnosis (CAD) costituiscono un valido approccio per supportare i medici nell'ispezione di un così elevato numero di immagini e per fornire una propria diagnosi.

Nel lavoro svolto durante questa tesi è stato realizzato uno strumento CAD basato su un'ampia sperimentazione con le reti neurali di diverse combinazioni dei criteri dinamici per poter ottenere buoni risultati dal punto di vista diagnostico. La sfida affrontata in questa tesi è stata quella di eseguire un'analisi sperimentale su un insieme di dati realistico, in cui non esiste una ben definita separazione tra il comportamento dinamico dei tumori maligni

e quello dei tessuti benigni. Un dataset di questo tipo risulta, quindi, ricco di casi “difficili” che potrebbero compromettere la bontà dei risultati. Con l’obiettivo di realizzare un’indagine ampia ed accurata, e quindi più affidabile, l’intero processo di validazione e di stima dell’errore di predizione delle reti neurali è stato realizzato per mezzo della tecnica double k-fold cross-validation. Inoltre è stata implementata un’interfaccia grafica che permette di visualizzare le immagini di risonanza magnetica, selezionare le zone di enhancement da processare ed ottenere una diagnosi per una data lesione sospetta. Tale interfaccia è stata utilizzata per la raccolta dei dati in seguito impiegati nelle sperimentazioni.

Per la costruzione del dataset sono stati utilizzati gli esami di risonanza magnetica forniti dal Dipartimento di Radiologia dell’Università di Pisa, contenenti, per ogni esame, la diagnosi attribuita dai medici e la diagnosi istologica. Sono stati selezionati 40 esami, di cui 20 contengono tumori benigni e 20 contengono tumori maligni. All’interno delle immagini sono state estratte le zone di più forte enhancement (Region Of Interest) e su tali zone è stata calcolata la curva di relative enhancement e i parametri quantitativi descritti nella sezione 3.3. La definizione delle Region Of Interest è stata fatta manualmente o in maniera semiautomatica, per mezzo della tecnica del region growing (descritta nella sezione 4.1.3), senza il supporto da parte di un radiologo. I dati, inoltre, sono stati solo minimamente filtrati, cioè sono state eliminate solo le curve che rappresentano un enhancement molto basso o addirittura nullo e quelle derivanti probabilmente da una selezione errata. In questo modo è stato costruito un dataset “ambizioso”, di difficile classificazione, diverso da quelli esistenti in letteratura la cui composizione avviene sotto la guida di radiologi esperti e i cui dati sono notevolmente filtrati. La sfida interessante è stata proprio quella di sperimentare diversi modelli neurali su un insieme di dati di questo tipo per riuscire ad ottenere dei buoni risultati dal punto di vista diagnostico, magari confrontabili con quelli esistenti in letteratura.

Per ogni combinazione di input, corrispondente ad una combinazione di criteri dinamici, è stato effettuato il collaudo della rete neurale, per mezzo della tecnica double k-fold cross-validation, in modo da trovare una configurazione adeguata di iperparametri per il problema in questione. Tale tecnica ha permesso di sfruttare tutti i dati a disposizione per ognuna delle fasi di allenamento, validazione e test della rete neurale. Senza una suddivisione fissata dei dati, infatti, è stato possibile allenare la rete neurale utilizzando tutto il dataset. Inoltre con la cross-validation interna è stato eseguito un processo di model selection ampio e corretto, perché eseguito su un gran numero di dati e perché indipendente dalla particolare ripartizione dei dati stessi, e separato dalla fase di test. Infine, la cross-validation esterna ha per-

messo di calcolare una stima dell'errore di predizione accurata ed affidabile in quanto il test non è stato eseguito su una singola porzione del dataset ma su tutti i dati a disposizione, di volta in volta inseriti all'interno del test set. Al termine della fase di model selection per tutti gli iperparametri presi in esame, per ogni diversa configurazione di input sono state effettuate 10 differenti inizializzazioni delle reti neurali. Infine si è calcolata la stima dell'accuratezza del modello come la media delle accuratèzze sui test set ottenute dalle singole inizializzazioni. In questo modo è stata effettuata una stima accurata, rigorosa ed affidabile dell'accuratezza del modello. Tale stima, infatti, oltre a non dipendere dalla particolare ripartizione dei dati, è anche indipendente dalla particolare inizializzazione della rete neurale, essendo stata calcolata come la media delle dieci diverse inizializzazioni.

Nelle sperimentazioni di questo studio sono stati testati diversi multilayer perceptron, allenati con l'algoritmo della back-propagation, per classificare le caratteristiche dinamiche estratte dalle immagini di risonanza magnetica. Nell'ambito di un'ampia sperimentazione, sono state provate alcune combinazioni di input già esistenti in letteratura ma sul dataset più complesso costruito in questa tesi per verificare la possibilità di ottenere risultati almeno confrontabili. Sono state, inoltre, introdotte nuove combinazioni di criteri dinamici utilizzando anche alcuni criteri poco sperimentati e comunque non in combinazione con altri (si consideri, ad esempio, il criterio dinamico curvature at the peak of enhancement descritto con l'equazione 5.1). I risultati dei vari classificatori, con le differenti configurazioni dei criteri dinamici in input, sono stati analizzati e confrontati tra di loro in base alle misure statistiche della performance di un test di classificazione descritte nella sezione 5.5. In seguito a tale analisi sono state ricavate le caratteristiche dinamiche che hanno mostrato una maggiore abilità discriminativa. Con questa combinazione è stato costruito il modello finale che poi è stato inserito all'interno dell'interfaccia grafica in modo da permettere ai medici di richiedere la classificazione di una data curva di relative enhancement e di visualizzarne la diagnosi corrispondente ottenuta dalla rete neurale. Dal confronto tra i risultati è emerso che la combinazione di criteri dinamici più rilevante è quella costituita dalla curva di relative enhancement combinata con il parametro washout ratio. Il washout si è rivelato una caratteristica significativa nella definizione della natura di una formazione patologica anche in diversi altri studi presenti in letteratura [50, 60, 9, 11]. Nell'ambito stesso della tesi il parametro washout ratio è stato utilizzato, a scopo di riferimento, anche in un classificatore a soglia che ha mostrato come tale parametro contenga in sé importante informazione discriminativa tra pattern benigni e pattern maligni. Con la classificazione ottenuta dal classificatore a soglia, infatti, sono state calcolate le seguenti misure statistiche: accuratezza = 88.83%, sensitività =

88.16%, specificità = 90.5%. La morfologia della curva di relative enhancement ha mostrato un significativo potere predittivo nella differenziazione delle lesioni anche in altre sperimentazioni presenti in letteratura [60, 28]. La performance diagnostica ottenuta in questa tesi dal multilayer perceptron con la combinazione di input formata dalle caratteristiche relative enhancement e washout ratio è data dalle seguenti misure statistiche: accuratezza  $\approx 92.5$ , sensitività  $\approx 89.9$ , specificità  $\approx 95.6$ . Tra i risultati presenti in letteratura per sperimentazioni simili, cioè multilayer perceptron allenato con l'algoritmo back-propagation con criteri dinamici in input, si trova: accuratezza compresa tra 78% e 89%, sensitività compresa tra 78% e 97%, specificità compresa tra 62% e 85% [1, 34, 66, 68]. Come si può vedere da tali misure statistiche, la performance diagnostica raggiunta dalla rete neurale realizzata in questa tesi è competitiva con quelle esistenti, soprattutto tenendo presente che tali risultati sono stati ottenuti su un dataset più "difficile" rispetto a quelli analizzati in letteratura. Il miglioramento ottenuto, seppur piccolo, è significativo se contestualizzato all'ambito in cui si svolgono le sperimentazioni di questo studio, in cui ogni singolo passo avanti costituisce un importante contributo all'aumento delle probabilità di sopravvivenza delle donne colpite dal cancro al seno. L'incremento della specificità, che in questo caso è rilevante, implica la riduzione dei casi benigni che vengono sottoposti ad una biopsia non necessaria. La sensitività ottenuta in queste sperimentazioni non ha apportato miglioramenti radicali rispetto a quelle esistenti in letteratura, ma risulta comunque confrontabile con esse. Questo risultato è dovuto all'utilizzo di un dataset realistico e soggetto ad errore, come è stato spiegato nella sezione 5.2. Dai grafici dei pattern classificati erroneamente come negativi dalla rete neurale, mostrati nella sezione 5.8, risulta infatti evidente che tali pattern hanno, in effetti, un comportamento chiaramente negativo oppure un comportamento dubbio. Bisogna, però, evidenziare che nel dataset analizzato in queste sperimentazioni sono presenti due falsi negativi dei medici, cioè 2 esami diagnosticati come negativi dai medici ma che poi si sono rivelati positivi all'analisi istologica. Tali esami sono, invece, stati classificati correttamente dalla rete neurale. Questo costituisce un importante risultato, anche se valutato solo su due esempi a disposizione, in quanto dimostra come uno strumento automatico possa essere di supporto ai medici nella diagnosi del tumore, altrimenti condizionata dall'esperienza del radiologo, dal tempo impiegato nell'interpretazione di un esame, dai criteri utilizzati per estrarre i dati.

Nonostante, in generale, i tumori maligni presentino un più forte e veloce afflusso dell'agente di contrasto e, contemporaneamente, un'accelerata fuoriuscita dello stesso mezzo di contrasto dal sito del tumore (washout) rispetto ai tessuti benigni, in realtà ci sono numerose eccezioni che causano



una considerevole sovrapposizione nel comportamento dinamico delle patologie benigne e dei tumori maligni. La combinazione delle curve di relative enhancement con i parametri quantitativi analizzati dalla rete neurale, come è stato sperimentato in questa tesi, può ridurre la sovrapposizione tra pattern benigni e maligni. Questo è dovuto al potere delle reti neurali nel fare associazioni tra diversi parametri indipendenti. Le caratteristiche dinamiche da sole, però, non forniscono un criterio discriminativo definitivo proprio a causa di tale sovrapposizione che non viene completamente risolta. Per ottenere le migliori prestazioni dal punto di vista diagnostico è necessaria l'analisi combinata di criteri dinamici e morfologici. Ad esempio, bordi irregolari e spigolosi sono tipici delle lesioni maligne; al contrario bordi regolari sono indicativi di masse benigne. Per combinare in modo ottimo sia i criteri dinamici che quelli morfologici è importante avere contemporaneamente alta risoluzione temporale ed alta risoluzione spaziale. L'infusione attraverso i tessuti dell'agente di contrasto dura solo pochi secondi, quindi un'alta risoluzione temporale è necessaria per ottenere immagini che catturino sia la fase di washin che quella di washout dell'agente di contrasto. L'alta risoluzione temporale, però, limita la dimensione della regione che può essere acquisita, causando la perdita di alcuni piccoli ma importanti dettagli della morfologia della lesione. Per questo motivo è necessario avere contemporaneamente anche un'alta risoluzione spaziale.

In conclusione, la rete neurale realizzata in questa tesi, con un ampio ed accurato processo di validazione, ha prodotto risultati competitivi anche su un dataset "ambizioso". Dunque l'analisi automatica delle immagini di risonanza magnetica con le reti neurali ha dimostrato di essere un approccio promettente per il supporto ai medici nella diagnosi del cancro al seno.

I possibili sviluppi futuri sono di seguito elencati:

- Integrazione di criteri dinamici e morfologici, possibile anche grazie agli sviluppi tecnologici nelle modalità di acquisizione delle immagini, per ottenere una diagnosi più accurata.
- Estrazione automatica dei dati per minimizzare la dipendenza dell'output della rete neurale dalla correttezza dell'interpretazione radiologica.
- Sperimentazione utilizzando altri modelli di machine learning.

## *6. Conclusioni*

---

# Appendice A

## Lo standard DICOM

Lo standard DICOM (Digital Imaging and Communications in Medicine) fornisce i criteri per il trattamento, la memorizzazione, la stampa e la comunicazione di informazioni di tipo biomedico. Tale standard permette lo scambio di informazioni digitali tra i dispositivi di acquisizione delle immagini mediche ed altri sistemi specificando:

- un insieme di protocolli che devono essere seguiti per la comunicazione;
- una sintassi ed una semantica di comandi associati a questi protocolli;
- un insieme di servizi di memorizzazione dei dati conformi allo standard, come ad esempio il formato del file per facilitare l'accesso alle immagini ed alle relative informazioni.

Lo standard DICOM si basa su una definizione astratta delle entità del mondo reale applicabile soprattutto alle immagini di ambito medico. Ogni entità del mondo reale, ad esempio un'immagine o un paziente, viene modellata con un *Information Object Definition* (IOD), un oggetto informativo che consiste di una descrizione degli attributi che definiscono quella specifica entità. Uno IOD può essere di due tipi:

- IOD normalizzato, include solo gli attributi inerenti l'entità del mondo reale rappresentata. Ad esempio l'oggetto informativo "Paziente" è normalizzato in quanto contiene solo gli attributi riguardanti il paziente (nome, data di nascita, etc.).
- IOD composto, può includere attributi aggiuntivi che sono correlati con l'entità del mondo reale ma non strettamente inerenti ad essa. Ad esempio, l'oggetto informativo "Risonanza Magnetica" è composto in

quanto contiene non solo attributi inerenti all'immagine (data dell'immagine, etc.) ma anche attributi che sono solo correlati con l'immagine stessa (nome del paziente, etc.).

Gli attributi di ogni IOD sono partizionati in gruppi di attributi simili che sono specificati come moduli indipendenti (*information object modules* (IOM)) e possono essere riutilizzati da altri IOD composti. Gli attributi sono le entità di informazione base. Per ogni attributo lo standard DICOM definisce le seguenti caratteristiche:

- *Attribute Name*, unico
- *Attribute Tag*, unico
- *Attribute Description*, semantica
- *Value Representation*, sintassi
- *Value Multiplicity*, quanti valori può assumere l'attributo

Attribute Tag è un identificatore unico composto da una coppia di numeri (gruppo, elemento). Dopo che l'informazione è stata acquisita, viene codificata nei formati definiti dallo standard DICOM usando Attribute Tag e Value Representation per creare un DICOM *data set*, in cui ogni attributo è codificato in un *data element*. La caratteristica Value Representation descrive come un attributo è codificato in un data element.

Le immagini di risonanza magnetica sono modellate con uno IOD composto (*MR Image IOD*). La tabella A.1 mostra lo IOD che specifica un'immagine creata con un dispositivo di acquisizione basato su risonanza magnetica:

- Information Entity (IE) rappresenta il sottoinsieme di attributi definiti da uno IOD composto correlato ad una specifica entità del mondo reale.
- Module rappresenta un insieme di attributi all'interno di una Information Entity che sono logicamente correlati gli uni con gli altri.
- Usage:
  1. Mandatory (M), campo obbligatorio; tutte le istanze dello IOD avranno tale campo specificato.
  2. Conditional (C), campo obbligatorio sotto alcune condizioni; se le condizioni specificate non sono verificate il modulo in questione potrebbe non essere supportato e quindi l'informazione riguardante quel modulo potrebbe non essere specificata.
  3. User Option (U), campo opzionale.

<b>IE</b>	<b>Module</b>	<b>Usage</b>
Patient	Patient	M
	Clinical Trial Subject	U
Study	General Study	M
	Patient Study	U
	Clinical Trial Study	U
Series	General Series	M
	Clinical Trial Series	U
Frame of Reference	Frame of Reference	M
Equipment	General Equipment	M
Image	General Image	M
	Image Plane	M
	Image Pixel	M
	Contrast/bolus	C
	Device	U
	MR Image	M
	Overlay Plane	U
	VOI LUT	U
	SOP Common	M

Tabella A.1: Moduli di *MR Image IOD*



# Bibliografia

- [1] P. Abdolmaleki, L. D. Buadu, and H. Naderimansh. Feature extraction and classification of breast cancer on dynamic magnetic resonance imaging using artificial neural network. *Cancer Letters*, 171(2):183 – 191, 2001.
- [2] Y. Ando, H. Fukatsu, T. Ishiguchi, T. Ishigaki, T. Endo, and M. Miyazaki. Diagnostic utility of tumor vascularity on magnetic resonance imaging of the breast. *Magnetic Resonance Imaging*, 18(7):807 – 813, 2000.
- [3] S. Bahri, J.-H. Chen, H. J. Yu, A. Kuzucan, O. Nalcioglu, and M.-Y. Su. Can dynamic contrast-enhanced mri (dce-mri) predict tumor recurrence and lymph node status in patients with breast cancer? *Annals of Oncology*, 2008.
- [4] G. Brix, F. Kiessling, R. Lucht, S. Darai, K. Wasser, S. Delorme, and J. Griebel. Microcirculation and microvasculature in breast tumors: Pharmacokinetic analysis of dynamic mr image series. *Magnetic Resonance in Medicine*, 52(2):420–429, 2004.
- [5] L. D. Buadu, J. Murakami, S. Murayama, N. Hashiguchi, S. Sakai, K. Masuda, S. Toyoshima, S. Kuroki, and S. Ohno. Breast lesions: correlation of contrast medium enhancement patterns on mr images with histopathologic findings and tumor angiogenesis. *Radiology*, 200(3):639–649, 1996.
- [6] D. L. Buckley, P. J. Drew, S. Mussurakis, J. R. T. Monson, and A. Horsman. Microvessel density in invasive breast cancer assessed by dynamic gd-dtpa enhanced mri. *Journal of Magnetic Resonance Imaging*, 7(3):461–464, 1997.
- [7] F. A. Cardillo, A. Starita, D. Caramella, A. Cilotti, and F. Odoguardi. The segmentation and visualization of human organs based on adaptive

## BIBLIOGRAFIA

---

- region growing method. *European Conference on Emergent Aspects in Clinical Data Analysis*, 1:103–110, 2005.
- [8] N. Dash, A. R. Lupetin, R. H. Daffner, Z. L. Deeb, R. J. Sefczek, and R. L. Schapiro. Magnetic resonance imaging in the diagnosis of breast disease. *Am. J. Roentgenol.*, 146(1):119–125, 1986.
- [9] A. Degenhard, C. Tanner, C. Hayes, D. J. Hawkes, M. O. Leach, and the UK MRI Breast Screening Study. Comparison between radiological and artificial neural network diagnosis in clinical screening. *Physiological Measurement*, 23(4):727–739, 2002.
- [10] S. J. El Yousef, R. H. Duchesneau, R. J. Alfydi, J. R. Haaga, P. J. Bryan, and J. P. LiPuma. Magnetic resonance imaging of the breast. work in progress. *Radiology*, 150(3):761–766, 1984.
- [11] X. Fan, M. Medved, G. S. Karczmar, C. Yang, S. Foxley, S. Arkani, W. Recant, M. A. Zamora, H. Abe, and G. M. Newstead. Diagnosis of suspicious breast lesions using an empirical mathematical model for dynamic contrast-enhanced mri. *Magnetic Resonance Imaging*, 25(5):593 – 603, 2007.
- [12] P. Filzmoser, B. Liebmann, and K. Varmuza. Repeated double cross validation. *Journal of Chemometrics*, 23(4):160–171, 2009.
- [13] U. Fischer, L. Kopka, and E. Grabbe. Breast carcinoma: Effect of pre-operative contrast-enhanced mr imaging on the therapeutic approach. *Radiology*, 213(3):881–888, 1999.
- [14] U.S. Preventive Services Task Force. Screening for breast cancer: U.s. preventive services task force recommendation statement. *Annals of Internal Medicine*, 151(10):716–726, 2009.
- [15] M. Friedrich. Mri of the breast: state of the art. *European Radiology*, 8(5):707–725, 1998.
- [16] P. Gibbs and L. W. Turnbull. Textural analysis of contrast-enhanced mr images of the breast. *Magnetic Resonance in Medicine*, 50(1):92–98, 2003.
- [17] R. C. Gonzales and R. E. Woods. *Digital Image Processing*. Prentice Hall, 2002.



- 
- [18] R. M. Haralick and L. G. Shapiro. *Computer and Robot Vision*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1992.
- [19] K. Haris, S. N. Efstratiadis, N. Maglaveras, and A. K. Katsaggelos. Hybrid image segmentation using watersheds and fast region merging. *IEEE transactions on Image Processing*, 7:1684–1699, 1998.
- [20] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 1999.
- [21] T. H. Helbich. Contrast-enhanced magnetic resonance imaging of the breast. *European Journal of Radiology*, 34(3):208 – 219, 2000.
- [22] S. H. Heywang, D. Hahn, H. Schmidt, I. Krischke, W. Eiermann, R. Basermann, and J. Lissner. Mr imaging of the breast using gadolinium-dtpa. *Journal of Computer Assisted Tomography*, 10(2):199–204, 1986.
- [23] S. H. Heywang-Köbrunner. *Contrast-Enhanced Mri of the Breast*. Springer-Verlag, 1995.
- [24] S. H. Heywang-Köbrunner, P. Viehweg, A. Heinig, and Ch. Kuchler. Contrast-enhanced mri of the breast: accuracy, value, controversies, solutions. *European Journal of Radiology*, 24(2):94 – 108, 1997.
- [25] D. M. Ikeda, N. M. Hylton, K. Kinkel, M. G. Hochman, C. K. Kuhl, W. A. Kaiser, J. C. Weinreb, S. F. Smazal, H. Degani, P. Viehweg, J. Barclay, and M. D. Schnall. Development, standardization, and testing of a lexicon for reporting contrast-enhanced breast magnetic resonance imaging studies. *Journal of Magnetic Resonance Imaging*, 13:889–895, 2001.
- [26] W. A. Kaiser and E. Zeitler. Mr imaging of the breast: fast imaging sequences with and without gd-dtpa. *Radiology*, 170(3):681–686, 1989.
- [27] M.V. Knopp, E. Weiss, H.P. Sinn, J. Mattern, H. Junkermann, J. Radeleff, A. Magener, G. Brix, S. Delorme, I. Zuna, and G. van Kaick. Pathophysiologic basis of contrast enhancement in breast tumors. *Journal of Magnetic Resonance Imaging*, 10(3):260–266, 1999.
- [28] C. K. Kuhl, P. Mielcareck, S. Klaschik, C. Leutner, E. Wardelmann, J. Gieseke, and H. H. Schild. Dynamic breast mr imaging: Are signal intensity time course data useful for differential diagnosis of enhancing lesions? *Radiology*, 211(1):101–110, 1999.

## BIBLIOGRAFIA

---

- [29] C. K. Kuhl and H. H. Schild. Dynamic image interpretation of mri of the breast. *Journal of Magnetic Resonance Imaging*, 12(6):965–974, 2000.
- [30] M. Leach. Breast imaging technology: Application of magnetic resonance imaging to angiogenesis in breast cancer. *Breast Cancer Res*, 3(1):22–27, 2001.
- [31] Y. LeCun. Une procédure d'apprentissage pour réseau a seuil asymetrique (a learning scheme for asymmetric threshold networks). In *Proceedings of Cognitiva 85*, pages 599–604, Paris, France, 1985.
- [32] S. Z. Li. Markov random field models in computer vision. In *Proceedings of the third European conference on Computer Vision (Vol. II)*, pages 361–370, Secaucus, NJ, USA, 1994. Springer-Verlag New York, Inc.
- [33] J. Y. Lo, J. A. Baker, P. J. Kornguth, C. E. Floyd, and Jr. Computer-aided diagnosis of breast cancer: Artificial neural network approach for optimized merging of mammographic features. *Academic Radiology*, 2(10):841 – 850, 1995.
- [34] R. E. A. Lucht, M. V. Knopp, and G. Brix. Classification of signal-time curves from dynamic mr mammography by neural networks. *Magnetic Resonance Imaging*, 19(1):51 – 57, 2001.
- [35] S. Malur, S. Wurdinger, A. Moritz, W. Michels, and A. Schneider. Comparison of written reports of mammography, sonography and magnetic resonance mammography for preoperative evaluation of breast lesions, with special emphasis on magnetic resonance mammography. *Breast Cancer Res*, 3:55–60, 2001.
- [36] T. M. Mitchell. *Machine Learning*. McGraw-Hill, 1997.
- [37] P. J. Moate, L. Dougherty, M. D. Schnall, R. J. Landis, and R. C. Boston. A modified logistic model to describe gadolinium kinetics in breast tumors. *Magnetic Resonance Imaging*, 22(4):467 – 473, 2004.
- [38] E. A. Morris. Breast cancer imaging with MRI. *Radiologic Clinics of North America*, 40(3):443–466, 2002.
- [39] E. A. Morris and L. Liberman. *Breast MRI: Diagnosis and Intervention*. Springer Science-Business Media, 2005.
- [40] C. I. Mosier. I. problems and designs of cross-validation 1. *Educational and Psychological Measurement*, 11(1):5–11, 1951.

- 
- [41] H. Narisada, T. Aoki, T. Sasaguri, H. Hashimoto, T. Konishi, M. Morita, and Y. Korogi. Correlation between numeric gadolinium-enhanced dynamic mri ratios and prognostic factors and histologic type of breast carcinoma. *Am. J. Roentgenol.*, 187(2):297–306, 2006.
- [42] L. W. Nunes, M. D. Schnall, and S. G. Orel. Update of breast mr imaging architectural interpretation model. *Radiology*, 219(2):884–894, 2001.
- [43] L. W. Nunes, M. D. Schnall, S. G. Orel, M. G. Hochman, C. P. Langlotz, C. A. Reynolds, and M. H. Torosian. Breast mr imaging: interpretation model. *Radiology*, 202(3):833–841, 1997.
- [44] D. B. Parker. *Learning-logic: Casting the Cortex of the Human Brain in Silicon*. Technical Report TR-47, Center for Computational Research in Economics and Management Science, MIT, Cambridge, MA, 1985.
- [45] J. R. Parker. *Algorithms for Image Processing and Computer Vision*. John Wiley & Sons, Inc., New York, NY, USA, 1st edition, 1996.
- [46] D. M. Parkin and L. M. G. Fernández. Use of statistics to assess the global burden of breast cancer. *The Breast Journal*, 12, 2006.
- [47] T. Pavlidis and Y.T. Liow. Integrating region growing and edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(3):225–233, mar 1990.
- [48] N. Peters, I. Borel Rinkes, W. Mali, M. van den Bosch, R. Storm, P. Plaisier, E. de Boer, A. van Overbeeke, and P. Peeters. Breast mri in nonpalpable breast lesions: a randomized trial with diagnostic and therapeutic outcome - monet - study. *Trials*, 8(1):40, 2007.
- [49] E. D. Pisano, R. E. Hendrick, M. J. Yaffe, J. K. Baum, S. Acharyya, J. B. Cormack, L. A. Hanna, E. F. Conant, L. L. Fajardo, L. W. Bassett, C. J. D’Orsi, R. A. Jong, M. Rebner, A. N. A. Tosteson, and C. A. Gatsonis. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in dmist. *Radiology*, 246(2):376–383, 2008.
- [50] G. Potente, D. Messineo, C. Maggi, and S. Savelli. Practical application of contrast-enhanced magnetic resonance mammography [ce-mrm] by an algorithm combining morphological and enhancement patterns. *Computerized Medical Imaging and Graphics*, 33(2):83–90, 2009.

## BIBLIOGRAFIA

---

- [51] F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408, 1958.
- [52] F. Rosenblatt. On the convergence of reinforcement procedures in simple perceptrons. *Cornel Aeronautical Laboratory Report VG-1916-G-4.*, 1960.
- [53] F. Rosenblatt. Perceptron simulation experiments. *Proceedings of the IRE*, 48(3):301–309, march 1960.
- [54] F. Rosenblatt. *Principles of neurodynamics; perceptrons and the theory of brain mechanisms*. Spartan Books Washington,, 1962.
- [55] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. *Learning internal representations by error propagation*, pages 318–362. MIT Press, Cambridge, MA, USA, 1986.
- [56] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, 1986.
- [57] D.E. Rumelhart and J.L. McClelland. *Parallel distributed processing: explorations in the microstructure of cognition. Volume 1. Foundations*. MIT Press, Cambridge, MA, 1986 Jan 01.
- [58] F. Sardanelli and A. Fausto. Computed tomography in breast cancer imaging? *Surgery*, 134(5):844, 2003.
- [59] M. D. Schnall, J. Blume, D. A. Bluemke, G. A. DeAngelis, N. DeBruhl, S. Harms, S. H. Heywang-Köbrunner, N. Hylton, C. K. Kuhl, E. D. Pisano, P. Causer, S. J. Schnitt, D. Thickman, C. B. Stelling, P. T. Weatherall, C. Lehman, and C. A. Gatsonis. Diagnostic architectural and dynamic features at breast mr imaging: Multicenter study. *Radiology*, 238(1):42–53, 2006.
- [60] M. D. Schnall, S. Rosten, S. Englander, S. G. Orel, and L. W. Nunes. A combined architectural and kinetic interpretation model for breast mr images. *Academic Radiology*, 8(7):591 – 597, 2001.
- [61] S. K. Sharan, M. Morimatsu, U. Albrecht, DS. Lim, E. Regel, C. Dinh, A. Sands, G. Eichele, P. Hasty, and A. Bradley. Embryonic lethality and radiation hypersensitivity mediated by rad51 in mice lacking brca2. *Nature*, 386:804–810, 1997.

- 
- [62] American Cancer Society. Breast cancer facts & figures 2009-2010. 2009.
- [63] P. C. Stomper, J. S. Winston, S. Herman, D. L. Klippenstein, M. A. Arredondo, and L. E. Blumenson. Angiogenesis and dynamic mr imaging gadolinium enhancement of malignant and benign breast lesions. *Breast Cancer Research and Treatment*, 45:39–46, 1997. 10.1023/A:1005897227030.
- [64] B. K. Szabó, P. Aspelin, M. Kristoffersen Wiberg, and B. Boné. Dynamic mr imaging of the breast. *Acta Radiologica*, 44(4):379–386, 2003.
- [65] B. K. Szabó, P. Aspelin, M. Kristoffersen Wiberg, T. Tot, and B. Boné. Invasive breast cancer: correlation of dynamic mr features with prognostic factors. *European Radiology*, 13:2425–2435, 2003. 10.1007/s00330-003-2000-y.
- [66] B. K. Szabó, M. K. Wiberg, B. Boné, and P. Aspelin. Application of artificial neural networks to the analysis of dynamic mr imaging features of the breast. *European Radiology*, 14:1217–1225, 2004.
- [67] J. S. Taylor, P. S. Tofts, R. Port, J. L. Evelhoch, M. Knopp, W. E. Reddick, V. M. Runge, and N. Mayr. Mr imaging of tumor microcirculation: Promise for the new millenium. *Journal of Magnetic Resonance Imaging*, 10(6):903–907, 1999.
- [68] G. Torheim, F. Godtlielsen, D. Axelson, K. A. Kvistad, O. Haraldseth, and P. A. Rinck. Feature extraction and classification of dynamic contrast-enhanced t2-weighted breast image data. *IEEE Trans. Med. Imaging*, pages 1293–1301, 2001.
- [69] A. A. Tzacheva, K. Najarian, and J. P. Brockway. Breast cancer detection in gadolinium-enhanced mr images by static region descriptors and neural networks. *Journal of Magnetic Resonance Imaging*, 17(3):337–342, 2003.
- [70] T. W. Vomweg, M. Buscema, H. U. Kauczor, A. Teifke, M. Intraligi, S. Terzi, C. P. Heussel, T. Achenbach, O. Rieker, D. Mayer, and M. Thelen. Improved artificial neural networks in prediction of malignancy of lesions in contrast-enhanced mr-mammography. *Medical Physics*, 30(9):2350–2359, 2003.
- [71] R. Warren. Screening women at high risk of breast cancer on the basis of evidence. *European Journal of Radiology*, 39(1):50 – 59, 2001.

## BIBLIOGRAFIA

---

- [72] P. Werbos. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. PhD thesis, Harvard University, Cambridge, MA, 1974.
- [73] J. Wu, F. Ye, J.L. Ma, X.P. Sun, J. Xu, and Z.M. Cui. The segmentation and visualization of human organs based on adaptive region growing method. *2008 IEEE 8th International Conference on Computer and Information Technology Workshops*, pages 439–443, 2008.
- [74] C. Xu, D. L. Pham, and J. L. Prince. Image segmentation using deformable models. In *Handbook of Medical Imaging. Vol.2 Medical Image Processing and Analysis*, pages 175–272.
- [75] D. Ziou and S. Tabbone. Edge detection techniques - an overview. *International Journal of Pattern Recognition and Image Analysis*, 8:537–559, 1998.