



**Politecnico  
di Torino**

POLITECNICO DI TORINO

**Corso di Laurea Magistrale  
in Ingegneria Biomedica**

Tesi di Laurea Magistrale

**Valutazione dell'impatto di metodi di preprocessing  
sulla ricostruzione real-time  
di immagini optoacustiche usando reti neurali profonde**

**Relatori**

Prof.ssa Kristen Mariko Meiburger  
Dott.ssa Silvia Seoni

**Candidato**

Marco Pellerito

Marzo 2023

# Sommario

La tomografia optoacustica multispettrale (MSOT) è una tecnica innovativa che sfrutta l'effetto fotoacustico. La MSOT si basa infatti sull'assorbimento da parte dei tessuti di onde luminose a diversa lunghezza d'onda, i quali rispondono con variazioni meccaniche sufficienti a generare delle onde ultrasonore rilevabili. Lo scopo delle tecniche di ricostruzione di immagini optoacustiche è quello di riprodurre le immagini anatomiche a partire dai sinogrammi, ovvero i segnali generati dal rilevamento di queste onde da parte di sonde piezoelettriche, del tutto analoghe a quelle usate in ecografia.

In questo lavoro di tesi sono stati esaminati punti di forza e limitazioni di due delle tecniche di ricostruzione più efficaci: l'algoritmo di Back-Projection e il Model-Based iterativo. Il primo è un algoritmo real-time che ha come principale limitazione quella di assumere la velocità del suono costante all'interno dell'organismo: ciò diminuisce notevolmente le performance. Il secondo, invece, si avvicina di più al caso reale ma richiede tempi di elaborazione molto lunghi, incompatibili con il real-time.

Queste problematiche hanno spinto verso l'utilizzo di metodi basati su reti neurali, come il DeepMB. Questo modello di ricostruzione, oggetto della presente tesi, è stato proposto dal gruppo di ricerca i-Thera, collaboratore del Politecnico di Torino. Il modello riceve in ingresso i sinogrammi e restituisce le immagini ricostruite grazie all'applicazione di una rete neurale con struttura U-Net. Le immagini di riferimento utilizzate per l'allenamento sono quelle ottenute tramite il metodo Model-Based iterativo: questo permette alla rete, una volta allenata, di ricostruire le immagini con elevata accuratezza e in tempi brevi, tali da permetterne l'utilizzo real-time. Questo metodo risulta inoltre innovativo poiché prevede la possibilità di rendere la velocità del suono un parametro allenabile, aspetto che aumenta di molto le prestazioni.

In prima battuta è stata valutata la ripetibilità del modello proposto, per poi esaminare quali set-up potessero essere indagati e quali caratteristiche del modello potessero essere modificate, valutando se ciò porta ad un effettivo miglioramento delle prestazioni. Il punto su cui ci si è soffermati è l'effetto in termini di prestazioni di algoritmi di preprocessing applicati ai sinogrammi. Un altro fattore esaminato è stato il costo computazionale: il modello richiede tempi di allenamento molto lunghi, dunque sono state valutate condizioni di stop non troppo stringenti che potessero ridurre i tempi inficiando meno possibile sulle prestazioni.

I risultati ottenuti dimostrano le elevate prestazioni e l'effettiva ripetibilità del modello rispetto ai dati riportati in letteratura. Dalla prove effettuate l'utilizzo di algoritmi di preprocessing applicati ai sinogrammi non apporta differenze significative in termini di performance.

# Indice

<b>Elenco delle tabelle</b>	4
<b>Elenco delle figure</b>	5
<b>1 Introduzione</b>	8
1.1 Panoramica	8
1.2 Tomografia Optoacustica Multispettrale	8
1.2.1 Principio fisico	8
1.2.1.1 Illuminazione	8
1.2.1.2 Detezione	9
1.2.2 Prestazioni	9
1.2.3 Principali cromofori	10
1.2.4 Applicazioni	10
1.3 Tipi di sonde	11
1.3.1 Effetto piezoelettrico	11
1.3.2 Sector	12
1.3.3 Linear	13
1.3.4 Curved	13
1.4 Back-Projection Algorithm	14
1.5 Model-Based Reconstruction	15
1.6 Accenni su altri metodi basati su reti neurali	17
1.6.1 CNN - UNET	17
1.6.2 RES - UNET	18
1.7 Deep Model-Based Reconstruction	19
<b>2 Materiali e metodi</b>	20
2.1 Struttura del modello	20
2.2 Dataset	21
2.3 Data generation	22
2.3.1 Forward model	23
2.3.2 Model-based iterativo	23
2.3.3 Backprojection method	25
2.3.4 Parametri del modello	26
2.3.5 Parametri della sonda	27

2.4	Sonda . . . . .	28
2.4.1	MSOT Acuity Echo . . . . .	28
2.4.2	Verasonics L11-5v . . . . .	29
2.5	DeepMB . . . . .	32
2.5.1	Struttura generale . . . . .	33
2.5.2	Dominio del segnale . . . . .	33
2.5.3	Trasformazione del dominio . . . . .	33
2.5.4	Dominio dell'immagine . . . . .	34
2.5.4.1	Codifica e concatenazione velocità del suono . . . . .	34
2.5.4.2	Rete . . . . .	34
2.5.5	Allenamenti . . . . .	35
2.5.6	Parametri generali . . . . .	36
2.6	Metriche . . . . .	36
<b>3</b>	<b>Risultati</b> . . . . .	<b>37</b>
3.1	Identity . . . . .	37
3.1.1	Prestazioni: Training Set . . . . .	37
3.1.2	Prestazioni: Validation Set . . . . .	38
3.1.3	Prestazioni: Test Set . . . . .	40
3.1.4	Riepilogo . . . . .	42
3.2	Conv2D . . . . .	43
3.2.1	Prestazioni: Training Set . . . . .	43
3.2.2	Prestazioni: Validation Set . . . . .	45
3.2.3	Prestazioni: Test Set . . . . .	46
3.2.4	Riepilogo . . . . .	48
3.3	Condizioni di stop . . . . .	49
3.3.1	Prestazioni: Training Set . . . . .	49
3.3.2	Prestazioni: Validation Set . . . . .	51
3.3.3	Prestazioni: Test Set . . . . .	52
3.3.4	Riepilogo . . . . .	54
3.3.5	Approfondimento . . . . .	55
3.4	Confronto . . . . .	56
3.4.1	Metriche . . . . .	56
3.4.2	Training Set . . . . .	56
3.4.3	Validation Set . . . . .	58
3.4.4	Testing Set . . . . .	59
<b>4</b>	<b>Conclusioni</b> . . . . .	<b>60</b>

# Elenco delle tabelle

1.1	Principali cromofori e corrispettiva lunghezza d'onda d'eccitazione. [1] . . .	10
2.1	Parametri del modello impostati per la sonda Acuity Echo, iThera Medical. [13] . . . . .	28
2.2	Parametri della sonda Acuity Echo, iThera Medical. [7] . . . . .	28
2.3	Parametri del modello impostati per la sonda Verasonics L11-5v. . . . .	30
2.4	Parametri della sonda Verasonics L11-5v. [21][22] . . . . .	31
2.5	Parametri definiti per il pre-processing dei sinogrammi e dove vengono utilizzati. [13] . . . . .	33
2.6	Confronto tra DeepMB con velocità del suono allenabile e senza. [7] . . . .	34
2.7	Parametri definiti per la rete e dove vengono utilizzati. [13] . . . . .	35
2.8	Parametri generali. [13] . . . . .	36
3.1	Media e deviazione standard delle metriche calcolate sul training set, primo allenamento. . . . .	38
3.2	Media e deviazione standard delle metriche calcolate sul validation set, primo allenamento. . . . .	38
3.3	Media e deviazione standard delle metriche calcolate sul test set, primo allenamento. . . . .	40
3.4	Media e deviazione standard delle metriche calcolate sul training set, secondo allenamento. . . . .	43
3.5	Media e deviazione standard delle metriche calcolate sul validation set, secondo allenamento. . . . .	45
3.6	Media e deviazione standard delle metriche calcolate sul test set, secondo allenamento. . . . .	46
3.7	Media e deviazione standard delle metriche calcolate sul training set, terzo allenamento. . . . .	49
3.8	Media e deviazione standard delle metriche calcolate sul validation set, terzo allenamento. . . . .	51
3.9	Media e deviazione standard delle metriche calcolate sul test set, terzo allenamento. . . . .	52

# Elenco delle figure

1.1	Principio di funzionamento MSOT applicato alla proteina DsRed2. [1] . . .	9
1.2	A) Andamento del segnale acustico in relazione alla profondità. B) Andamento del segnale acustico in relazione alla grandezza del campione. [1] . . . . .	10
1.3	A) Immagine ecografica dell'avambraccio. B), C), D) Immagini MSOT della stessa regione. [2] . . . . .	11
1.4	Descrizione effetto piezoelettrico. [4] . . . . .	12
1.5	Sonda Sector. [5] . . . . .	12
1.6	Sonda Linear. [5] . . . . .	13
1.7	Sonda Curved. [5] . . . . .	14
1.8	A) Immagine di un fantoccio. B) Ricostruzione ottenuta tramite back-projection. [8] . . . . .	15
1.9	Decomposizione della risposta all'impulso totale. [9] . . . . .	15
1.10	A) Ricostruzione dell'immagine tramite altri metodi. B) Ricostruzione dell'immagine tramite model-based. [9] . . . . .	16
1.11	A) Ricostruzione dell'immagine tramite correzione EIR. B) Ricostruzione dell'immagine tramite correzione iEIR. [10] . . . . .	16
1.12	Architettura CNN-UNET. [11] . . . . .	17
1.13	A) Immagine di un dito. B) Ricostruzione tramite CNN-UNET. [11] . . . .	17
1.14	Architettura RES-UNET. [12] . . . . .	18
1.15	A) Immagine fantoccio. B) Immagine ricostruita tramite RES-UNET. [12]	18
1.16	DeepMB reconstruction pipeline. [7] . . . . .	19
2.1	Esempi immagine dataset. [14] . . . . .	21
2.2	Esempio applicazione toolbox (1). . . . .	22
2.3	Esempio applicazione toolbox (2). . . . .	22
2.4	A) Immagine di partenza. B) Sinogramma calcolato tramite forward model.	24
2.5	Ricostruzione tramite model-based iterativo. . . . .	24
2.6	A) Sinogramma acquisito in vivo di un bicipite sinistro. B) Ricostruzione tramite model-based iterativo. [13] . . . . .	25
2.7	Ricostruzione tramite backprojection method. . . . .	26
2.8	Sonda MSOT Acuity Echo, iThera Medical. [19] . . . . .	29
2.9	A) EIR originale della sonda Verasonics L11-5v. [20] B) EIR ricostruito. .	29
2.10	Sonda Verasonics L11-5v. [20] . . . . .	30

2.11	A) Immagine di partenza. B) Ricostruzione con i parametri della sonda Verasonics L11-5v . . . . .	31
2.12	Struttura della rete [7] . . . . .	32
3.1	A) Immagine target. B) Immagine ricostruita con DeepMB. Training, primo allenamento. MAE = 0,61; RMSE = 0,95. . . . .	38
3.2	A) Immagine target. B) Immagine ricostruita con DeepMB. Training, primo allenamento. MAE = 5,69; RMSE = 7,84. . . . .	39
3.3	A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, primo allenamento. MAE = 0,62; RMSE = 0,87. . . . .	39
3.4	A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, primo allenamento. MAE = 5,68; RMSE = 8,08. . . . .	40
3.5	A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, primo allenamento. MAE = 0,49; RMSE = 0,78. . . . .	41
3.6	A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, primo allenamento. MAE = 2,03; RMSE = 4,91. . . . .	41
3.7	Istogramma percentuale dei valori MAE, primo allenamento. . . . .	42
3.8	Istogramma percentuale dei valori RMSE, primo allenamento. . . . .	42
3.9	A) Immagine target. B) Immagine ricostruita con DeepMB. Training, secondo allenamento. MAE = 0,76; RMSE = 1,06. . . . .	44
3.10	A) Immagine target. B) Immagine ricostruita con DeepMB. Training, secondo allenamento. MAE = 6,15; RMSE = 8,65. . . . .	44
3.11	A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, secondo allenamento. MAE = 0,78; RMSE = 1,09. . . . .	45
3.12	A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, secondo allenamento. MAE = 6,14; RMSE = 9,20. . . . .	46
3.13	A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, secondo allenamento. MAE = 0,55; RMSE = 0,89. . . . .	47
3.14	A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, secondo allenamento. MAE = 2,15; RMSE = 5,08. . . . .	47
3.15	Istogramma percentuale dei valori MAE, secondo allenamento. . . . .	48
3.16	Istogramma percentuale dei valori RMSE, secondo allenamento. . . . .	48
3.17	A) Immagine target. B) Immagine ricostruita con DeepMB. Training, terzo allenamento. MAE = 0,69; RMSE = 1,02. . . . .	50
3.18	A) Immagine target. B) Immagine ricostruita con DeepMB. Training, terzo allenamento. MAE = 5,77; RMSE = 8,62. . . . .	50
3.19	A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, terzo allenamento. MAE = 0,70; RMSE = 0,95. . . . .	51
3.20	A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, terzo allenamento. MAE = 5,77; RMSE = 8,63. . . . .	52
3.21	A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, terzo allenamento. MAE = 0,58; RMSE = 0,89. . . . .	53
3.22	A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, terzo allenamento. MAE = 2,06; RMSE = 5,04. . . . .	53
3.23	Istogramma percentuale dei valori MAE, terzo allenamento. . . . .	54
3.24	Istogramma percentuale dei valori RMSE, terzo allenamento. . . . .	54

3.25	Andamento valori medi della metrica MAE in funzione delle epoche. . . . .	55
3.26	Andamento valori medi della metrica RMSE in funzione delle epoche. . . . .	55
3.27	Confronto tra i tre allenamenti. A) Training Set, B) Validation Set, C) Testing Set. . . . .	56
3.28	Istogramma percentuale dei valori MAE per le tre reti, training set. . . . .	57
3.29	Istogramma percentuale dei valori RMSE per le tre reti, training set. . . . .	57
3.30	Istogramma percentuale dei valori MAE per le tre reti, validation set. . . . .	58
3.31	Istogramma percentuale dei valori RMSE per le tre reti, validation set. . . . .	58
3.32	Istogramma percentuale dei valori MAE per le tre reti, testing set. . . . .	59
3.33	Istogramma percentuale dei valori RMSE per le tre reti, testing set. . . . .	59

# Capitolo 1

## Introduzione

### 1.1 Panoramica

Come suggerisce il nome, la tecnica MSOT nasce dall'unione delle tecniche di imaging ottico e di quelle tomografiche, L'obiettivo è quello di sfruttare i principali vantaggi delle tecniche ottiche, su tutti la capacità di distinguere biomarcatori cromoforici all'interno dei tessuti biologici, a livello macro, visto che le risoluzioni più spinte per queste applicazioni si ottengono nel campo della microscopia.

La bassa attenuazione dei tessuti nei campi dell'infrarosso permette di sfruttare marker fluorescenti che presentano spettri di eccitazione ed assorbimento proprio nell'intorno di questa banda, ottenendo risoluzioni sufficienti per portare avanti studi clinici su malattie neurodegenerative e cardiovascolari, oltre che sui tumori.

La tomografia ottica permette invece di osservare porzioni di tessuto più grandi, perdendo però nettamente di risoluzione in profondità.

L'approccio ibrido ha portato a risultati sorprendenti [1].

### 1.2 Tomografia Optoacustica Multispettrale

#### 1.2.1 Principio fisico

La tecnologia optoacustica si compone di due fasi: illuminazione e detezione.

##### 1.2.1.1 Illuminazione

La fase di illuminazione prevede il susseguirsi in rapida successione di impulsi laser a diversa lunghezza d'onda di durata variabile tra 1 ns e 100 ns (rappresentati in verde in Fig. 1.1c): come si nota in Fig. 1.1, infatti, la percentuale relativa di assorbimento della proteina DsRed2 varia fortemente con la lunghezza d'onda, dunque si usano impulsi diversi per campionare le differenti risposte. Questa particolare proprietà è utile per la successiva elaborazione dell'immagine: poiché il tessuto presenta uno spettro di assorbimento più piatto, sottraendo le immagini ottenute a due lunghezze d'onda consecutive si tende ad eliminare l'assorbimento di fondo (che provoca rumore). Ovviamente, più è

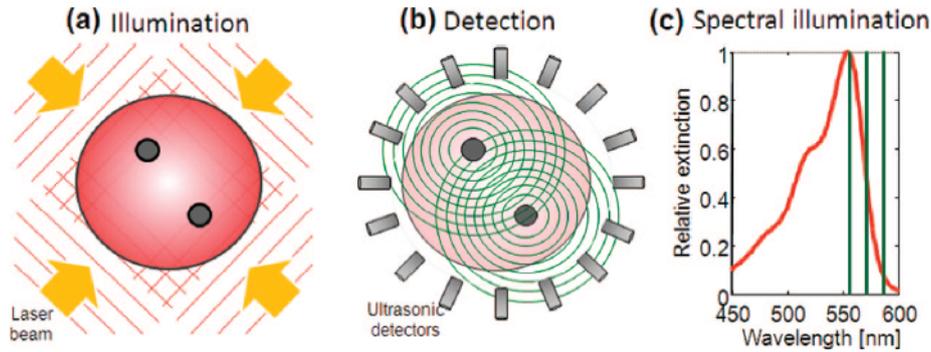


Figura 1.1: Principio di funzionamento MSOT applicato alla proteina DsRed2. [1]

stretta la banda spettrale di assorbimento della molecola fluorescente, maggiore è l'effetto dell'elaborazione. Altri fattori da tenere in considerazione sono gli effetti di disturbo in termini di propagazione della luce, che aumentano all'aumentare della profondità [1].

### 1.2.1.2 Detezione

Le molecole fotoassorbenti rispondono ai transitori di assorbimento generando onde acustiche per effetto termoacustico. Esse subiscono infatti una piccola espansione, non percepibile dall'uomo. Le onde generate vengono rilevate da trasduttori acustici sensibili [1].

## 1.2.2 Prestazioni

Uno dei più grandi vantaggi della tomografia optoacustica è il fatto che fornisce informazioni sia anatomiche che funzionali del tessuto analizzato. Le risoluzioni sono molto spinte, nell'ordine delle decine di micrometri, e si possono analizzare campioni di diversi centimetri.

E' una tecnica molto veloce, in quanto l'acquisizione avviene in maniera simile ad una acquisizione ecografica: come quest'ultima, inoltre, ha il vantaggio di essere una tecnica non invasiva, fattore che riduce sia il rischio clinico che l'impatto psicologico sul soggetto/paziente.

Come accennato, la principale limitazione consiste nel fatto che all'aumentare della profondità le proprietà ottiche dei tessuti diventano via via più sfavorevoli. In Fig. 1.2 è mostrato come si comporta il segnale optoacustico generato da una particolare molecola (Cy5.5) all'aumentare della profondità (A) e all'aumentare delle dimensioni del target a diverse profondità (B): come ci si può aspettare, entrambi i grafici suggeriscono che il segnale diventa poco rilevabile a profondità maggiori, seppur migliorando notevolmente in intensità quando si analizzano campioni più grandi [1].

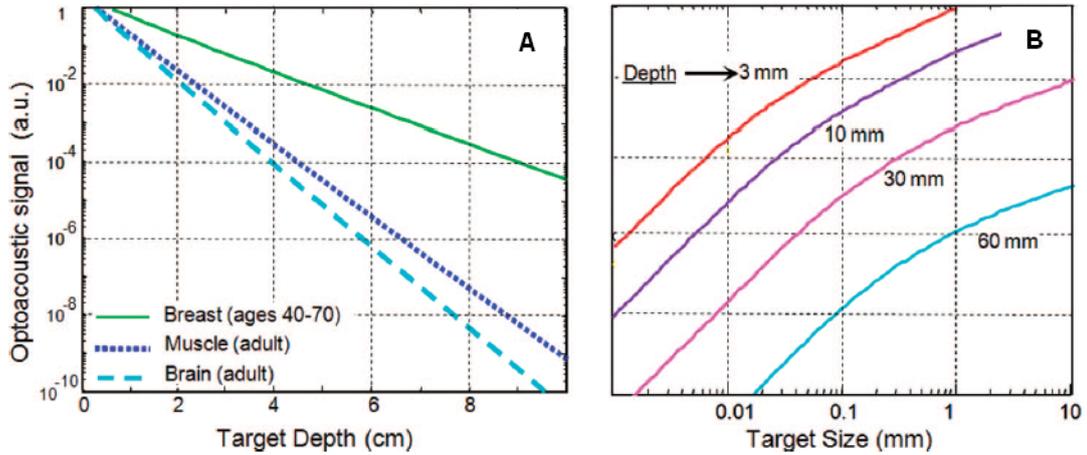


Figura 1.2: A) Andamento del segnale acustico in relazione alla profondità. B) Andamento del segnale acustico in relazione alla grandezza del campione. [1]

### 1.2.3 Principali cromofori

Di seguito è riportata una tabella riassuntiva dei principali cromofori e agenti di contrasto che vengono utilizzati nella tecnica MSOT e la corrispettiva lunghezza d'onda d'eccitazione laser [Tab. 1.1].

Tabella 1.1: Principali cromofori e corrispettiva lunghezza d'onda d'eccitazione. [1]

Cromoforo	Lunghezza d'onda di eccitazione (nm)
HbO <sub>2</sub>	750
Eumelanina	750
AlexaFluor750	750
mCherry	587
Nanotubi di carbonio a singola parete	750
Nanocilindri d'oro	750
X-gal	650
Quantum dots nel vicino infrarosso	840

### 1.2.4 Applicazioni

La tomografia optoacustica si inserisce nel panorama della diagnostica medica come tecnica complementare, e non sostitutiva, alle tradizionali tecniche in uso. Aiuta ad ottenere informazioni riguardo malattie multifattoriali come la distrofia muscolare di Duchenne o il tumore al seno, oltre che fornire supporto sull'analisi di classici fattori patofisiologici come la vascolarizzazione e l'ossigenazione [Fig. 1.3] [2].

La ricostruzione real-time delle immagini optoacustiche è un importante obiettivo per poter utilizzare la tecnica nell'ambito dell'imaging clinico, e questo ha spinto verso la ricerca di tecniche di ricostruzione che garantissero prestazioni ottimali in tempo reale.

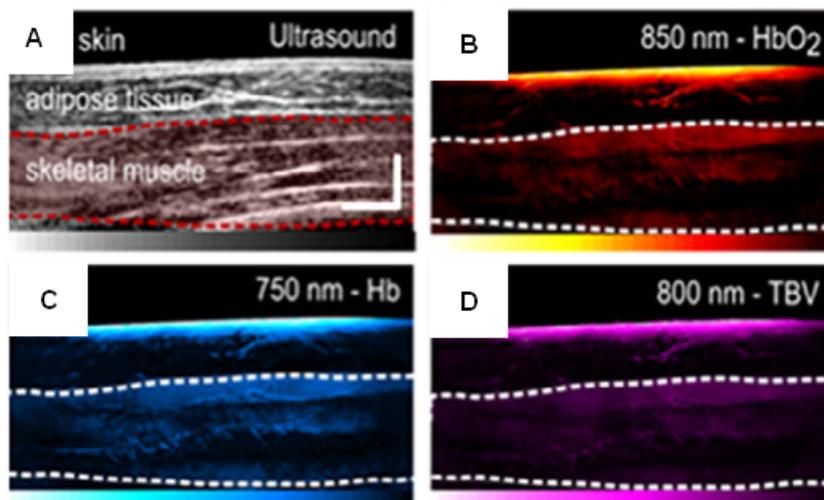


Figura 1.3: A) Immagine ecografica dell'avambraccio. B), C), D) Immagini MSOT della stessa regione. [2]

In Fig. 1.3 si può notare come in base alla differente lunghezza d'onda si metta in risalto il maggiore assorbimento della luce da parte dell'emoglobina ossigenata (B) o deossigenata (C), o l'uguale assorbimento (D, TBV sta per Total Blood Volume) [2].

## 1.3 Tipi di sonde

I detettori delle onde sonore generate dall'effetto fotoacustico sono normali cristalli piezoelettrici, dunque vengono usate semplici sonde ecografiche. L'utilizzo di una sonda rispetto ad un'altra cambia le caratteristiche geometriche dell'acquisizione (ad esempio il Field of View), oltre che ovviamente i parametri del sistema d'acquisizione (frequenza di campionamento, frequenze di taglio, numero di campioni acquisiti da ciascun cristallo). Di seguito, un piccolo approfondimento sull'effetto piezoelettrico e una panoramica delle principali sonde in commercio, con qualche accenno sull'evoluzione dell'imaging acustico nel tempo.

### 1.3.1 Effetto piezoelettrico

L'effetto piezoelettrico è un principio fisico noto da tempo tale per cui se un cristallo con proprietà piezoelettriche è interessato da uno stress meccanico, esso produce una differenza di potenziale ai suoi estremi. Il cristallo piezoelettrico più usato e noto ai più è il quarzo. In generale, la proprietà piezoelettrica è tipica di quei cristalli che non hanno un centro di simmetria [3]. Lo stress meccanico provoca all'interno del cristallo lo spostamento delle

cariche, che si raccolgono nelle due facce generando un campo elettrico [4]. In Fig. 1.4 è schematizzato quanto appena descritto.

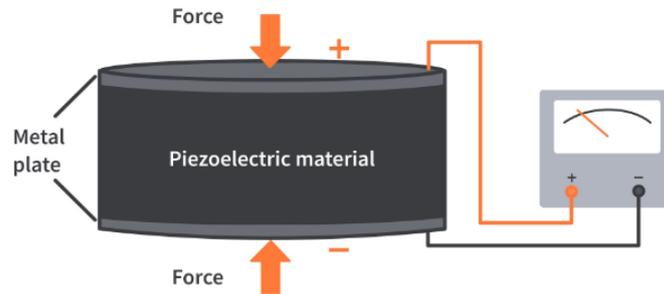


Figura 1.4: Descrizione effetto piezoelettrico. [4]

Nella normale ecografia, viene sfruttato anche l'effetto piezoelettrico inverso, ovvero la compressione e la dilatazione del cristallo attraverso l'utilizzo di un campo elettrico che insiste su di esso. Questi cambiamenti meccanici generano onde sonore, in analogia a quanto visto prima per i tessuti del corpo irradiati da luce laser.

### 1.3.2 Sector

Le prime sonde utilizzate nel campo dell'imaging acustico furono le sonde Sector [Fig. 1.5].

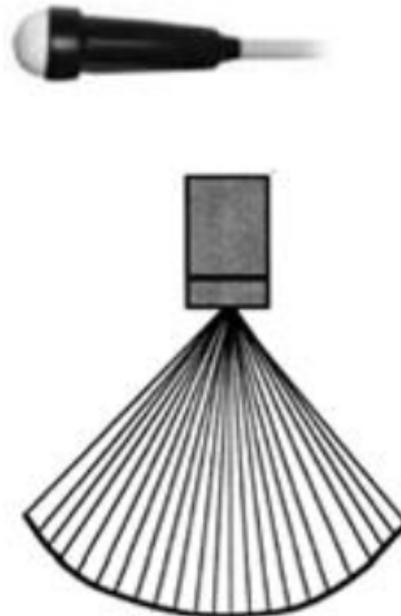


Figura 1.5: Sonda Sector. [5]

Inizialmente realizzate con un unico cristallo piezoelettrico, successivamente sono state aggiornate con un array circolare di detettori concentrici. Come si può notare dalla Fig. 1.5, il Field of View è particolarmente stretto all'inizio (near field) e aumenta notevolmente all'aumentare della profondità (far field) [5].

### 1.3.3 Linear

Per osservare l'immediato sottocute, si è allora pensato di allineare più elementi lungo un'unica direzione. Quella descritta è la sonda Linear [Fig. 1.6].

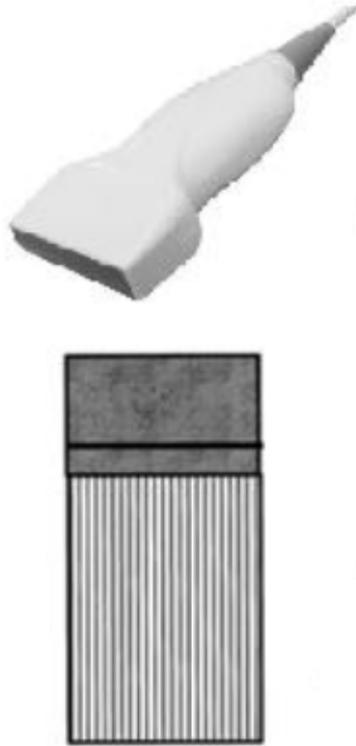


Figura 1.6: Sonda Linear. [5]

Così facendo, il risultato è un field of view rettangolare cui larghezza dipende dalla dimensione della sonda stessa, che permette di esaminare più attentamente il near field. Ovviamente, questo comporta un campo ridotto nel far field rispetto la sonda sector [5].

### 1.3.4 Curved

Una filosofia simile a quella appena descritta, ma che prova a superare la limitazione nel far field, ha portato all'avvento della sonda Curved [Fig. 1.7].



Figura 1.7: Sonda Curved. [5]

Questa evoluzione porta a mantenere una certa visione del near field, senza rinunciare alla visione in profondità. La principale limitazione di questo tipo di sonda, che condivide con la sonda Sector, è una risoluzione laterale peggiore di quella Linear [5].

## 1.4 Back-Projection Algorithm

La ricostruzione di immagini fotoacustiche consiste nel ricreare le immagini a partire dai sinogrammi acquisiti dalla sonda. Un semplice modello per la ricostruzione real-time delle immagini optoacustiche è l'algoritmo di retroproiezione [6].

Xu e Wang proposero una formula che fosse valida per le tre principali geometrie: piana, sferica e cilindrica. Questo algoritmo universale si basa però su una semplificazione che ne limita fortemente le prestazioni, ovvero quella di assumere che la velocità del suono nei tessuti sia costante e precisamente pari a 1500 m/s. Il principio su cui si basa questa tecnica, dunque, è solo quello di sfruttare i differenti comportamenti dei tessuti al passaggio del campo elettromagnetico, ovvero quando essi sono colpiti dalle radiazioni laser pulsate.

Questa limitazione comporta una serie di problemi a cascata che compromettono la qualità dell'immagine ricostruita, la quale soffre in particolar modo di una bassa risoluzione spaziale e di contrasto [Fig. 1.8] [7].

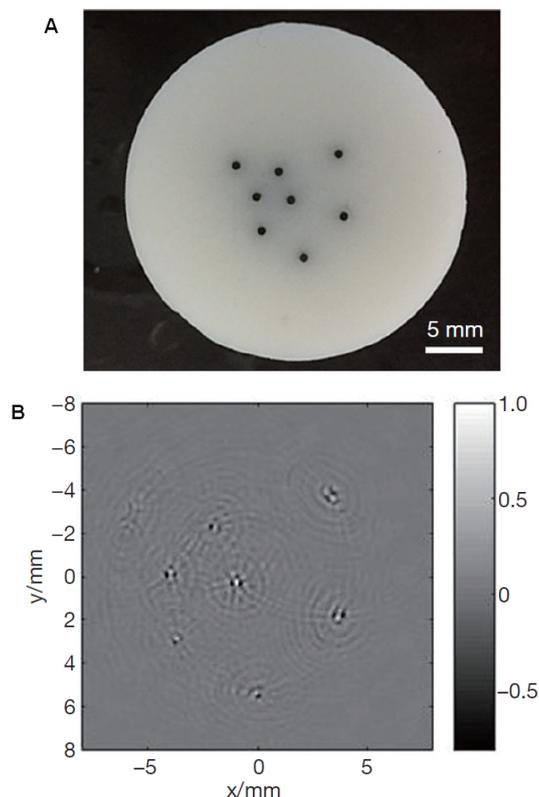


Figura 1.8: A) Immagine di un fantoccio. B) Ricostruzione ottenuta tramite back-projection. [8]

## 1.5 Model-Based Reconstruction

Un importante passo in avanti nella ricostruzione delle immagini fotoacustiche è stato fatto con l'introduzione del model-based reconstruction [9][10]. Questo modello si basa sulla cosiddetta risposta all'impulso totale (TIR), la quale tiene conto di una grande quantità di fattori che influenzano, o possono influenzare, l'imaging fotoacustico. La TIR nasce dalla convoluzione di due termini: la risposta all'impulso elettrica (EIR) e la risposta all'impulso spaziale (SIR) [Fig. 1.9].

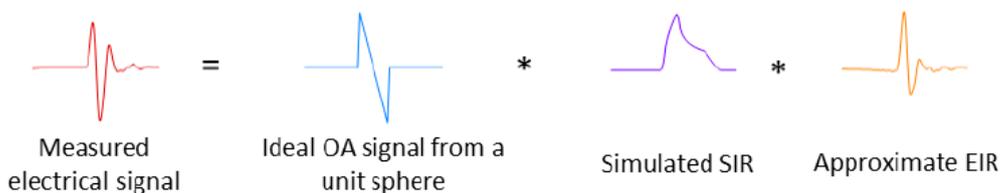


Figura 1.9: Decomposizione della risposta all'impulso totale. [9]

Si è dimostrato come l'effetto combinato della "correzione SIR" e della "correzione EIR" ha portato ad enormi miglioramenti nella ricostruzione dell'immagine. In particolare modo, questi fattori aiutano a riprodurre i disallineamenti acustici all'interno dei tessuti, superando l'assunzione fortemente limitativa del back-projection method di considerare la velocità del suono costante [Fig. 1.10].

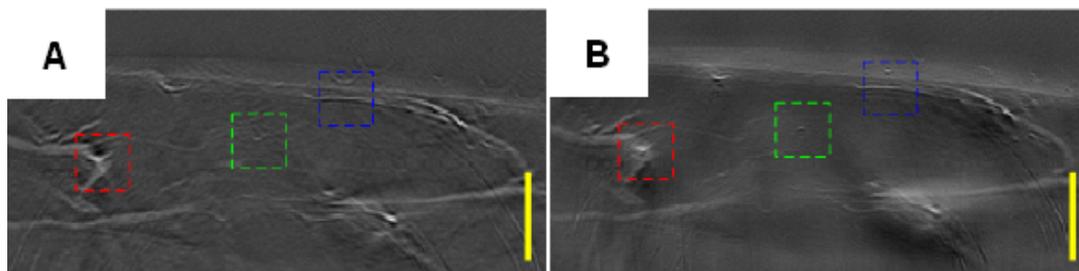


Figura 1.10: A) Ricostruzione dell'immagine tramite altri metodi. B) Ricostruzione dell'immagine tramite model-based. [9]

In [10] viene preso in esame un altro importante fattore che avvicina il modello alla condizione reale. La risposta elettrica all'impulso, infatti, è una rappresentazione mediata di come l'array di trasduttori risponde alla stimolazione elettrica: in base alla geometria della sonda, però, ciascun trasduttore risponde in maniera leggermente diversa, introducendo di fatto il concetto di risposta all'impulso elettrica individuale (iEIR) [Fig. 1.11].

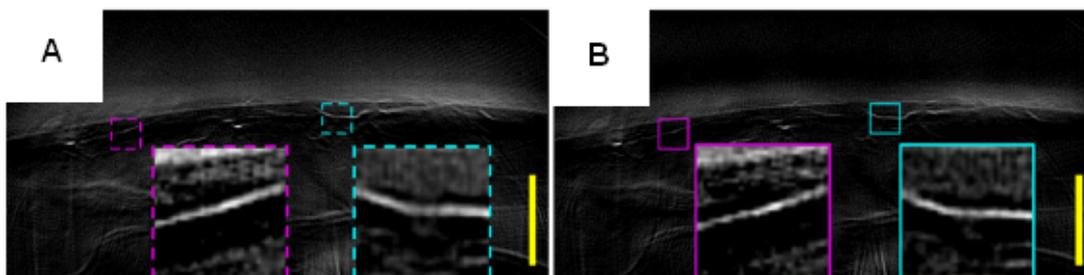


Figura 1.11: A) Ricostruzione dell'immagine tramite correzione EIR. B) Ricostruzione dell'immagine tramite correzione iEIR. [10]

La principale limitazione di questa tecnica risiede nel costo computazionale. Essendo un processo iterativo, richiede dispendio di risorse e di tempo, limitandone quindi l'applicazione real-time, ingrediente fondamentale come già discusso per potersi attestare come standard diagnostico nel panorama dell'imaging biomedico [7].

## 1.6 Accenni su altri metodi basati su reti neurali

Per rispondere all'esigenza crescente di tecniche real-time che offrirono prestazioni adeguate, sono stati ovviamente vagliati metodi che integrassero l'intelligenza artificiale, campo sempre più ingombrante nelle applicazioni riguardanti l'imaging medico. Di seguito, un accenno a due modelli proposti, prima di concentrarsi sulla rete soggetto della presente tesi.

### 1.6.1 CNN - UNET

In [11] è presentato un modello basato su una rete neurale convoluzionale (CNN) profonda con architettura UNET [Fig. 1.12].

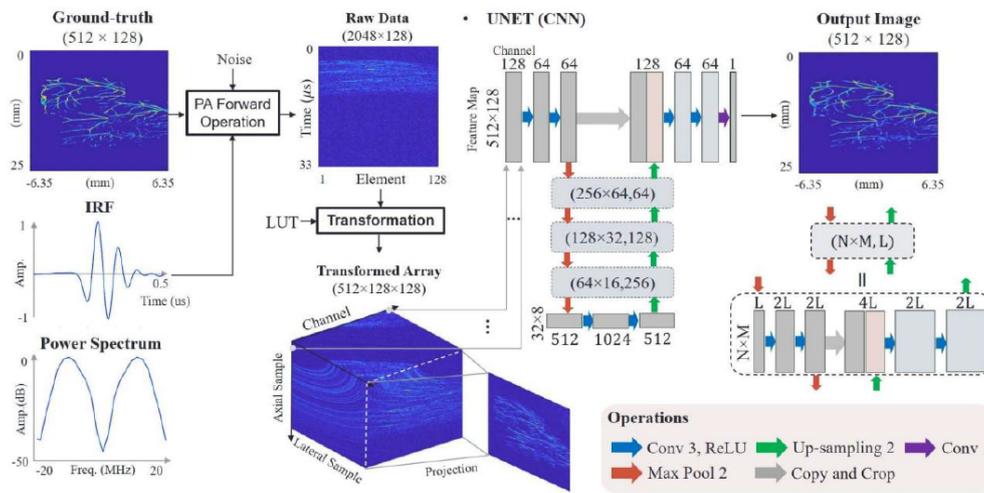


Figura 1.12: Architettura CNN-UNET. [11]

La scelta di utilizzare reti neurali profonde è stata effettivamente vincente in termini sia di prestazioni che di tempistiche, ma la rete mostrata sopra presenta poca capacità di generalizzare e, soprattutto, le prestazioni diminuiscono fortemente quando testata con dati reali [Fig. 1.13] [7][11].

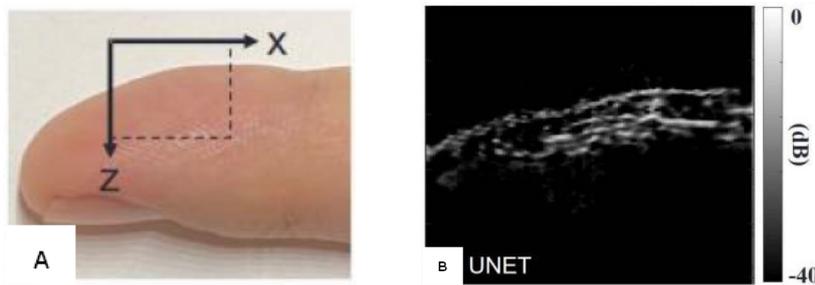


Figura 1.13: A) Immagine di un dito. B) Ricostruzione tramite CNN-UNET. [11]

### 1.6.2 RES - UNET

In [12] è proposto un modello basato sempre su CNN profonda ma con struttura RES-UNET [Fig. 1.14].

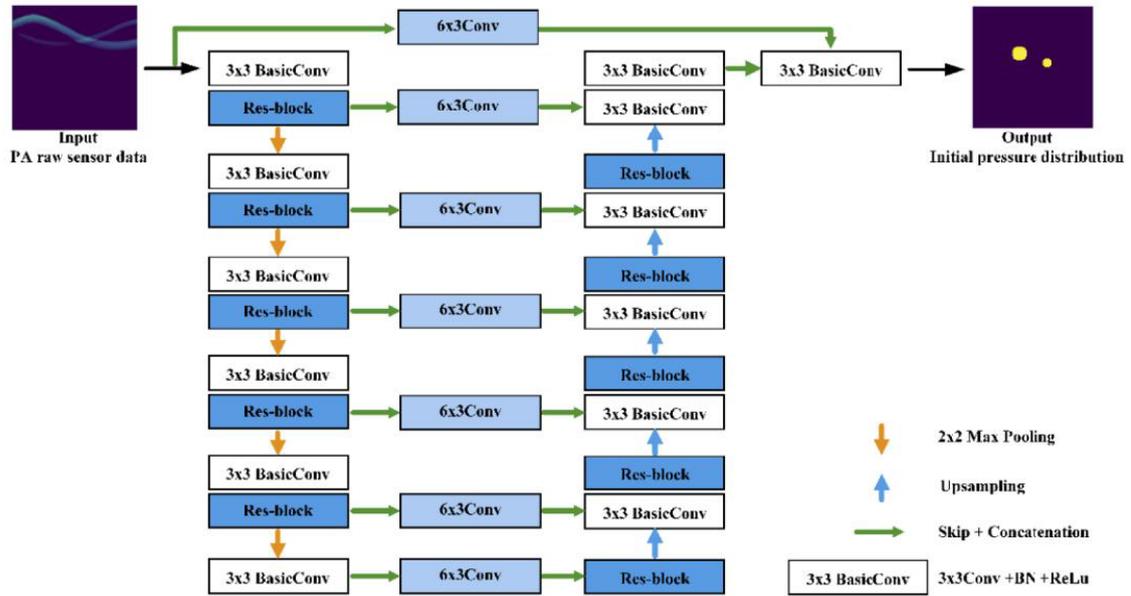


Figura 1.14: Architettura RES-UNET. [12]

Nonostante i promettenti test effettuati su immagini fantoccio [Fig. 1.15], l'approccio non è mai stato tentato su immagini in vivo [7].

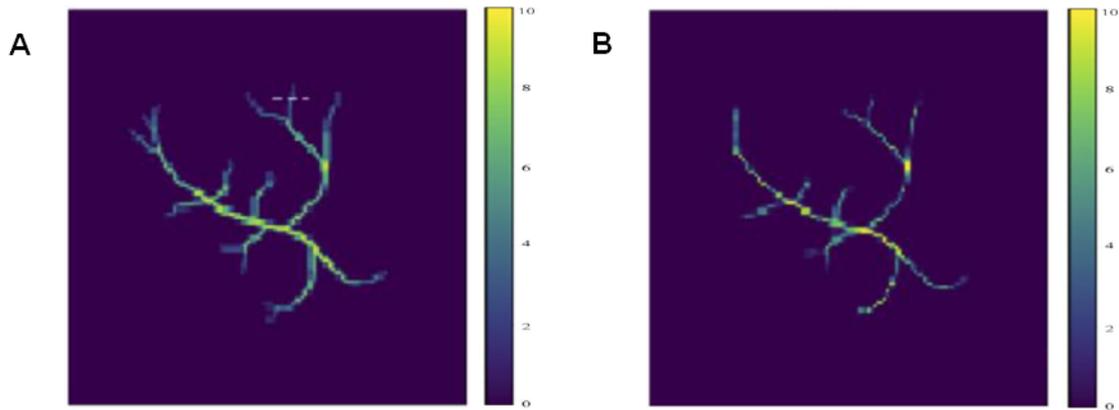


Figura 1.15: A) Immagine fantoccio. B) Immagine ricostruita tramite RES-UNET. [12]

## 1.7 Deep Model-Based Reconstruction

Nel 2022 è stato proposto un modello basato su rete neurale profonda che ha superato tutti i precedenti approcci simili, ottenendo una capacità di generalizzazione su dati reali sorprendente. Il modello permette di elaborare un'immagine in un tempo inferiore ai 10 ms, contro i 30-60 s del model-based iterativo, mantenendo prestazioni estremamente elevate [7][13].

Nei successivi paragrafi si discuterà in maniera approfondita di questo approccio e dei metodi utilizzati per la validazione dello stesso. Al contempo, verrà indagato come diversi allenamenti possono influenzare le prestazioni della rete. Infine, si mostrerà il confronto tra i set-up sulla base di metriche ben precise.

In Fig. 1.16 è mostrata la pipeline del modello, che è il punto di partenza per la disamina che seguirà.

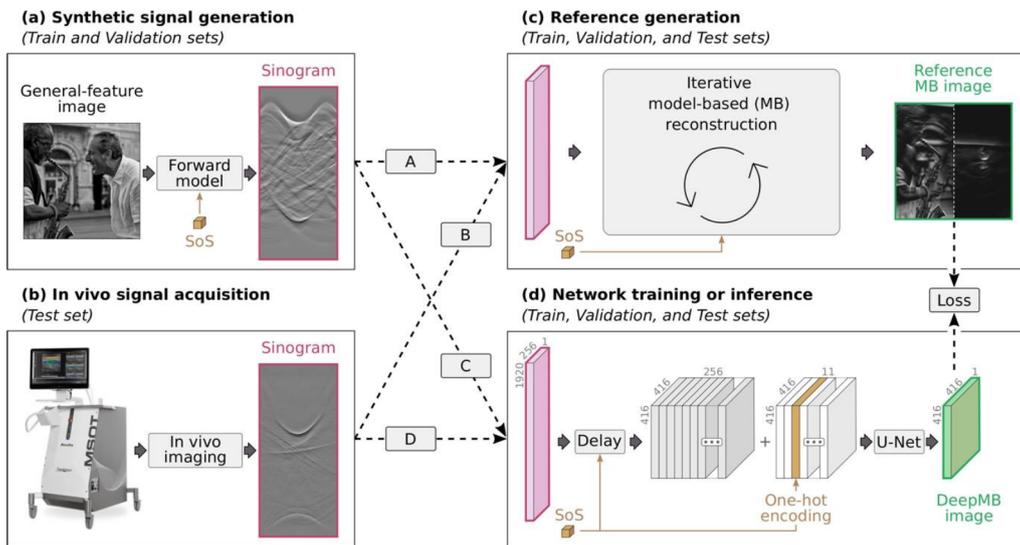


Figura 1.16: DeepMB reconstruction pipeline. [7]

# Capitolo 2

## Materiali e metodi

I metodi descritti in seguito sono stati sviluppati tramite i linguaggi di programmazione Matlab e Python, quest'ultimo in ambiente Spyder-Anaconda. Su Matlab sono stati alternati i sistemi operativi Windows e Linux, mentre su Python si è lavorato esclusivamente con il sistema operativo Linux. Il processo di data generation, che verrà descritto in dettaglio, non è stato implementato, bensì sono state utilizzate le ricostruzioni e i sinogrammi gentilmente forniti dai collaboratori.

### 2.1 Struttura del modello

Come si può notare dalla Fig. 1.16, la rete neurale accetta in ingresso i sinogrammi, i quali vengono generati in maniera differente in base al fatto che si tratti di dati sintetici o in vivo:

- Per i dati sintetici, si parte da un'immagine e si applica il forward model (che verrà approfondito in seguito). Tipicamente, anche per motivi di quantità di dati, i sinogrammi così generati vengono usati come training set e validation set per l'addestramento della rete;
- Per i dati in vivo, i sinogrammi derivano direttamente dall'acquisizione tramite sonda. Tipicamente questi sinogrammi vengono usati come test set per valutare la capacità della rete di generalizzare con dati provenienti dal mondo reale. Verrà accennato in seguito come variano le prestazioni della rete se si decide di allenarla con immagini in vivo.

L'output della rete, invece, è la ricostruzione dell'immagine optoacustica. Come riferimento sono state prese in considerazione le ricostruzioni ottenute tramite model-based iterativo, che è stato dunque applicato a tutti i sinogrammi di training set, validation set e test set. Nel toolbox, le performance sono calcolate tramite "data residual norm" e confrontate rispetto al target e, per completezza e come ulteriore confronto, rispetto alle immagini ricostruite tramite back-projection method; per necessità dovute a motivi di riservatezza, ci si è limitati all'utilizzo di metriche più convenzionali e al confronto solamente rispetto al target.

## 2.2 Dataset

Le immagini da cui partire per generare il dataset da usare per l'addestramento della rete sono state ricavate da un dataset pubblico contenente 16135 immagini di diversa e svariata natura [Fig. 2.1] [14].



Figura 2.1: Esempi immagine dataset. [14]

Del totale, 10000 sono state utilizzate per creare il dataset e sono state suddivise come segue:

- 8000 per il training set;
- 2000 per il validation set.

Ad ognuna di queste immagini è stato applicato il model-based reconstruction toolbox, approfondito nel capitolo successivo, ricavando così il sinogramma, la ricostruzione tramite model-based iterativo e quella tramite backprojection method.

Il toolbox mette a disposizione un dataset per l'allenamento alternativo, generato esclusivamente da immagini acquisite in vivo con una divisione 4/1/1 rispettivamente per training set, validation set e test set. Si è effettivamente dimostrato come la scelta di partire da immagini sintetiche aiuta la rete e aumenta le sue prestazioni, principalmente a causa degli artefatti delle immagini in vivo non presenti nei sinogrammi ricostruiti a partire dalle immagini del dataset pubblico [7][13].

## 2.3 Data generation

Per costruire le immagini target e i sinogrammi da mandare come input alla rete, è stato usato un toolbox predisposto per linguaggio Matlab [9][10][15][16][17][18]. Si precisa nuovamente che questo lavoro non è stato effettuato per la presente tesi ma si è partiti da un dataset fornito dai collaboratori. Il toolbox permette all'utente di seguire due strade:

- Partire da un'immagine, applicare il forward model per ricavarne il sinogramma e successivamente usare il model-based iterativo per la ricostruzione [Fig. 2.2];
- Partire direttamente dal sinogramma e applicare su di esso il model-based iterativo [Fig. 2.3].

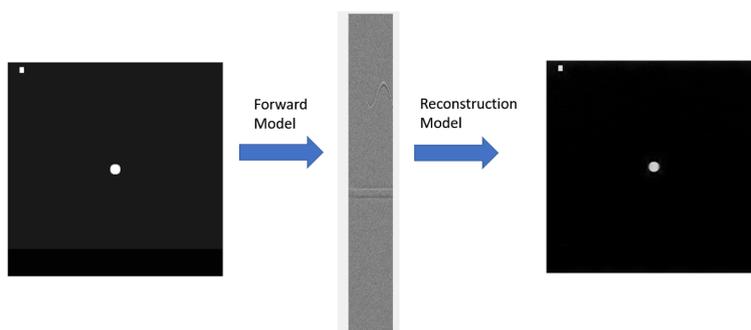


Figura 2.2: Esempio applicazione toolbox (1).

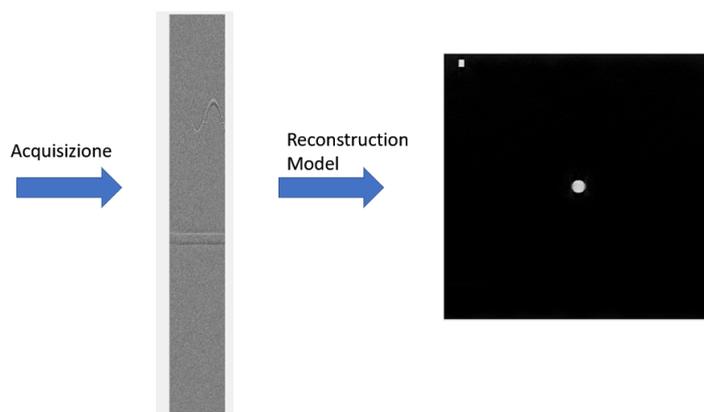


Figura 2.3: Esempio applicazione toolbox (2).

Il toolbox permette, in prima battuta, di simulare sia un'immagine che un sinogramma. Il calcolo del sinogramma (nel primo caso) o la simulazione dello stesso (nel secondo) viene fatto sulla base delle caratteristiche della sonda, che possono essere customizzate (il toolbox propone come default una sonda 'standard').

### 2.3.1 Forward model

Per la creazione dei sinogrammi viene sfruttato un accurato forward model [9], il quale viene applicato alle immagini del dataset pubblico opportunamente convertite in scala di grigio e ridimensionate secondo i parametri di input della rete [Tab. 2.1 e 2.3] [7].

Il segnale optoacustico si basa sul fatto che i tessuti vengono investiti da una determinata quantità di energia, generata dagli impulsi laser pulsati, la quale genera una pressione iniziale  $p_0$ . Questa pressione si propaga nell'organismo fino a quando non è intercettata dai cristalli piezoelettrici della sonda. Il segnale generato da ciascun trasduttore cambia, dipendentemente dalla sua posizione nella sonda. Si può quindi dedurre che il segnale  $s$  è funzione della pressione iniziale tramite un operatore che modella la risposta all'impulso dei cristalli della sonda:

$$s = Mp_0, [9] \quad (2.1)$$

dove  $M$  è la matrice del forward model.

Il forward model utilizzato è particolarmente avanzato per ciò che è stato visto nel capitolo precedente, ovvero perché include sia la risposta all'impulso elettrica sia la risposta all'impulso spaziale:

$$s = M^{TIR}p_0. \quad (2.2)$$

Per avere un grado di precisione ancora più elevato, il toolbox mette a disposizione la possibilità di inserire la risposta elettrica all'impulso individuale, cioè quella di ogni singolo trasduttore [10]:

$$s = M^{iTIR}p_0. \quad (2.3)$$

La velocità del suono è stata calcolata randomicamente nell'intervallo [1475 m/s - 1525 m/s] e considerata costante [7]. In Fig. 2.4 è mostrato un esempio di applicazione del modello.

### 2.3.2 Model-based iterativo

Il processo di ricostruzione dell'immagine consiste invece nel procedimento inverso a quello descritto precedentemente, ovvero trovare la pressione iniziale  $p_0$  a partire dal segnale  $s$ .

Il toolbox mette a disposizione più regolarizzazioni per risolvere il problema inverso, e la scelta è ricaduta su Shearlet L1 [7].

La formula per ricavare  $p_0$  può essere esplicitata come segue:

$$p_0 = \arg \min_{p \geq 0} \|M_{SOSP}p - s\|_2^2 + \lambda \|SH(p)\|_1, [7] \quad (2.4)$$

dove  $\|\cdot\|_n$  è la distanza n-norm,  $\lambda$  un parametro di regolarizzazione e SH la trasformata di Shearlet.

Il model-based iterativo è stato applicato a tutti i sinogrammi, sia quelli generati a partire dalle immagini del dataset pubblico [Fig. 2.5], sia quelli acquisiti in vivo [Fig. 2.6].



Figura 2.4: A) Immagine di partenza. B) Sinogramma calcolato tramite forward model.

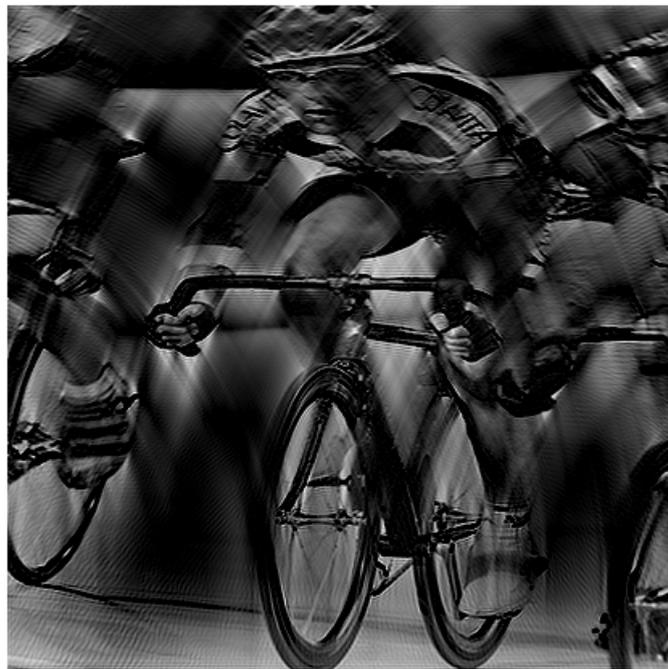


Figura 2.5: Ricostruzione tramite model-based iterativo.

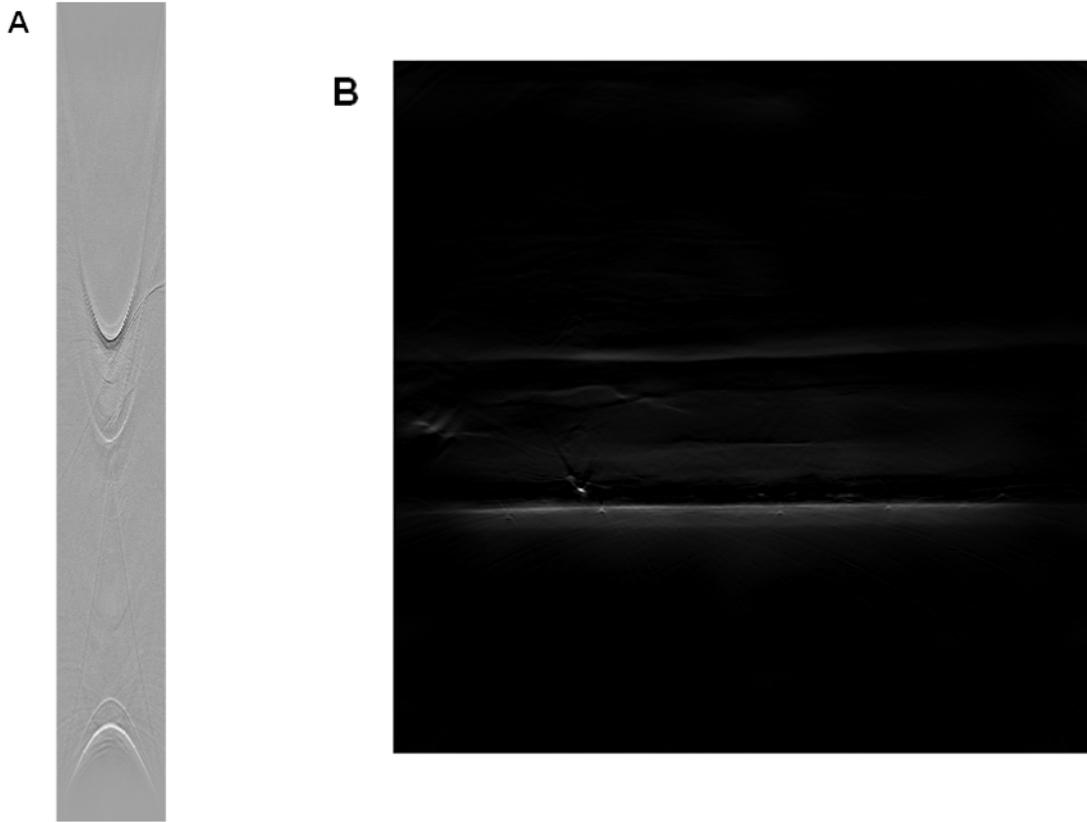


Figura 2.6: A) Sinogramma acquisito in vivo di un bicipite sinistro. B) Ricostruzione tramite model-based iterativo. [13]

### 2.3.3 Backprojection method

Il toolbox prevede anche la possibilità di ricostruire l'immagine tramite la formula del backprojection method:

$$I(x, y) \approx \sum_{i=1}^N S_i[(R - x \cos \theta_i - y \sin \theta_i)/v], [6] \quad (2.5)$$

dove  $i$  rappresenta i cristalli piezoelettrici,  $S_i$  il segnale optoacustico dell' $i$ -esimo cristallo,  $R$  il raggio del piano di scansione della sonda,  $v$  la velocità del suono (considerata costante) e  $\theta_i$  la coordinata angolare dell' $i$ -esimo trasduttore. In Fig. 2.7 è mostrata l'applicazione del metodo sull'immagine del dataset vista in precedenza.

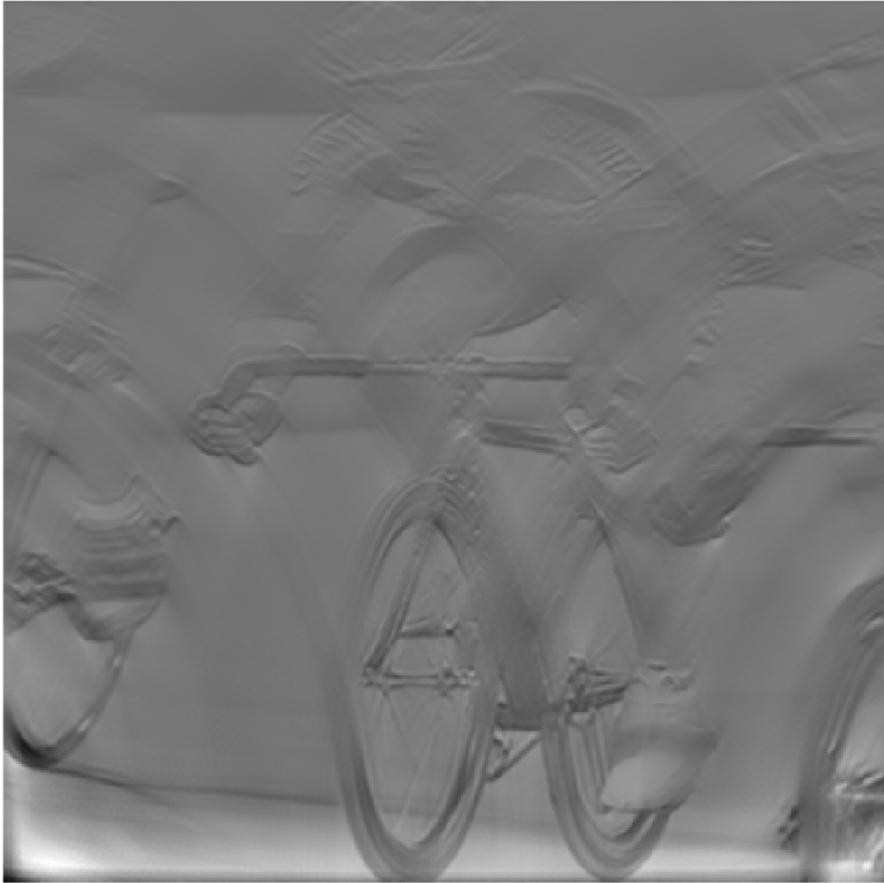


Figura 2.7: Ricostruzione tramite backprojection method.

### 2.3.4 Parametri del modello

Per adattare il modello al set-up sperimentale, il toolbox richiede di settare svariati parametri, di cui vengono elencati i principali [15]:

- ID della sonda, che deve essere definito in una classe apposita e di cui verrà discusso successivamente;
- Se usare o meno la 'correzione EIR';
- Se usare o meno la 'correzione EIR individuale';
- Se usare o meno la 'correzione SIR';
- Velocità del suono;
- Frequenze di taglio massima e minima del filtro passa-banda da applicare al segnale;

- Dimensioni del Field of View;
- Dimensioni dell'immagine;
- Valore del parametro di regolarizzazione  $\lambda$ ;
- Numero di iterazioni della ricostruzione model-based.

Nel capitolo successivo saranno esplicitati i parametri impostati in base alla sonda utilizzata: alcuni dipendono infatti indirettamente da essa (frequenze di taglio e dimensioni del Field of View). In generale, è importante che siano coerenti coi parametri di input della rete usata nel DeepMB toolbox.

### 2.3.5 Parametri della sonda

Il toolbox permette di customizzare i parametri della sonda, allo scopo di riprodurre i sinogrammi tramite forward model in maniera più fedele possibile alla sonda a disposizione. I parametri richiesti sono i seguenti [15]:

- Numero di trasduttori;
- Raggio del piano di scansione della sonda;
- Copertura angolare dei detettori in semicerchio;
- Coordinate x e z dei trasduttori;
- Raggio di elevazione della curvatura del singolo rilevatore;
- Altezza del singolo trasduttore;
- Larghezza del singolo trasduttore sommata alla distanza tra i trasduttori;
- Distanza tra i trasduttori;
- Numero totale di campioni registrati in una scansione da un trasduttore;
- Campioni di tempo di ritardo da parte del sistema di acquisizione;
- Frequenza di campionamento.

Nel prossimo capitolo saranno analizzate più da vicino due tipi di sonde e, in riferimento alla sonda Verasonics L11-5v, verranno presentati i tentativi di adattare il modello alla sonda stessa, obiettivo che verrà portato avanti in futuro.

## 2.4 Sonda

Il tipo di sonda utilizzata e le sue caratteristiche sono ovviamente fondamentali per determinare le prestazioni. Di seguito sono presentate la sonda utilizzata dai collaboratori, e dunque le caratteristiche usate per questa tesi, e la sonda a disposizione del Politecnico di Torino.

### 2.4.1 MSOT Acuity Echo

La sonda usata dai collaboratori iThera [Fig. 2.8] è di tipo curved array e prevede già al suo interno un sistema di illuminazione a impulsi laser pulsati. Dal punto di vista della rilevazione si affida ad un detettore customizzato della IMASONIC SAS. In Tab. 2.1 sono elencati i parametri del modello scelti con l'utilizzo di questa sonda, mentre in Tab. 2.2 sono elencate le sue caratteristiche [7]:

Tabella 2.1: Parametri del modello impostati per la sonda Acuity Echo, iThera Medical. [13]

Parametro	Unità di misura	Valore
device_id	NA	Acuity_ithera
use_eir	NA	true
use_indiv_eir	NA	false
use_sir	NA	true
filt_min	Hz	1e5
filt_max	Hz	12e6
field_of_view	m	[-0.02075 0.02075 -0.02075 0.02075]
speed_of_sound_range	m/s	1475:5:1525
N	px	[416 416]
lambda_shearlet	NA	1e-2
num_iterations_mb	NA	50

Tabella 2.2: Parametri della sonda Acuity Echo, iThera Medical. [7]

Parametro	Unità di misura	Valore
numOfTransducers	NA	256
radius	cm	4
angularCoverage	°	125
numRecordedSamplesPerTransducer	NA	1920
delayBeforeRecording	NA	8
frequency	Hz	4e7



Figura 2.8: Sonda MSOT Acuity Echo, iThera Medical. [19]

### 2.4.2 Verasonics L11-5v

La sonda a disposizione del Politecnico di Torino [Fig. 2.10] è il modello Verasonics L11-5v, un array di tipo lineare compatibile con le configurazioni standard e ad alta frequenza del Vantage Systems, sviluppato anch'esso da Verasonics [20][21]. Essendo di tipo lineare, la sfida principale è stata adattare il toolbox, pensato per una sonda avente posizioni angolari dei trasduttori differenti, a questo tipo di detettore.

Il primo passaggio è stato riprodurre la risposta elettrica all'impulso della sonda [20]:

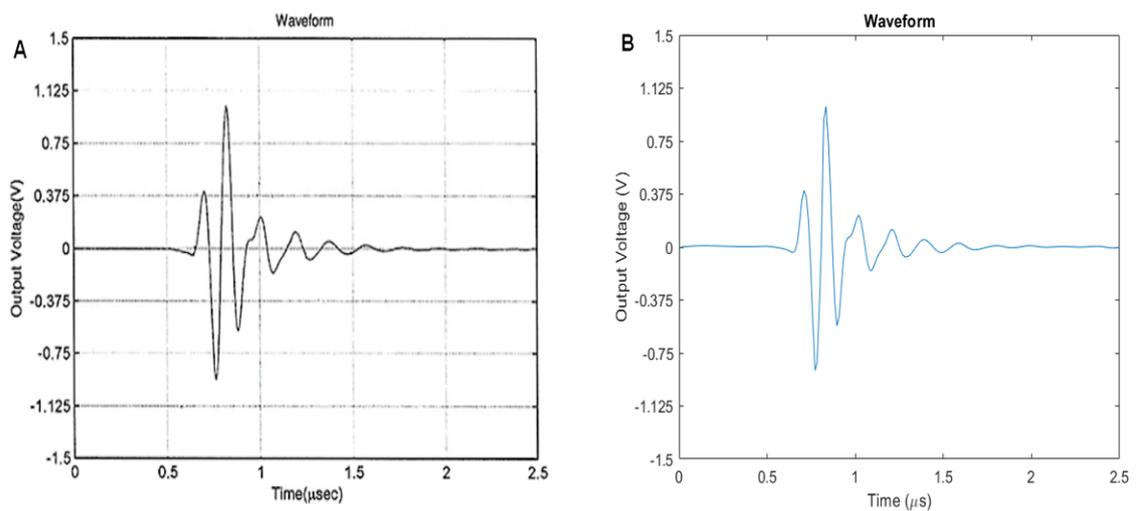


Figura 2.9: A) EIR originale della sonda Verasonics L11-5v. [20] B) EIR ricostruito.



Figura 2.10: Sonda Verasonics L11-5v. [20]

In Tab. 2.3 sono invece elencati i parametri del modello scelti con l'utilizzo di questa sonda, mentre in Tab. 2.4 sono elencate le sue caratteristiche:

Tabella 2.3: Parametri del modello impostati per la sonda Verasonics L11-5v.

Parametro	Unità di misura	Valore
device_id	NA	L11_5v
use_eir	NA	true
use_indiv_eir	NA	false
use_sir	NA	true
filt_min	Hz	8e5
filt_max	Hz	21e6
field_of_view	m	[-0.01905 0.01905 0.005 0.060]
N	px	[416 416]
lambda_shearlet	NA	1e-2
num_iterations_mb	NA	50

Tabella 2.4: Parametri della sonda Verasonics L11-5v. [21][22]

Parametro	Unità di misura	Valore
numOfTransducers	NA	128
radius	cm	0
angularCoverage	°	180
zPositionsOfTransducers	m	0x128
elevationRadiusOfSingleTransducer	m	18e-3
heightofSingleTransducer	m	5e-3
pitchOfTransducers	m	0.3e-3
seperationBetweenTransducers	m	0.27e-3
numRecordedSamplesPerTransducer	NA	1920
delayBeforeRecording	NA	0
frequency	Hz	5e7

Testando i parametri inseriti con l'immagine esempio del toolbox, purtroppo, è stato chiaro che l'adattamento alle caratteristiche della sonda non è andato a buon fine [Fig. 2.11]. L'ostacolo principale è stato la mancanza di fonti dal quale ricavare le caratteristiche della sonda stessa richieste dal toolbox. Dopo svariati tentativi, si è deciso di lasciare questo obiettivo come step successivo, concentrandosi sulla valutazione di diversi set-up del metodo DeepMB.

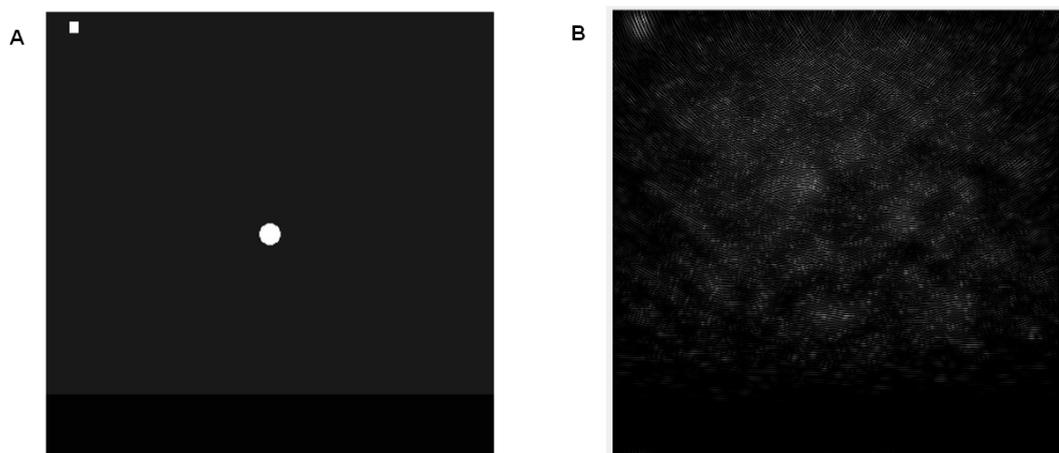


Figura 2.11: A) Immagine di partenza. B) Ricostruzione con i parametri della sonda Verasonics L11-5v .

## 2.5 DeepMB

Conclusa la fase di data generation, il materiale a disposizione è il seguente:

- 8000 sinogrammi generati tramite forward model da mandare alla rete come training set. Di ognuno di questi sinogrammi sono stati ricavati la ricostruzione tramite model-based iterativo e tramite back-projection;
- 2000 sinogrammi generati tramite forward model da mandare alla rete come validation set. Di ognuno di questi sinogrammi sono stati ricavati la ricostruzione tramite model-based iterativo e tramite back-projection;
- 44 sinogrammi generati tramite acquisizione in vivo [13] da usare come training set per valutare la capacità di generalizzare della rete. Di ognuno di questi sinogrammi sono stati ricavati la ricostruzione tramite model-based iterativo e tramite back-projection.

I codici per l'allenamento e l'utilizzo della rete sono stati sviluppati su Python. Il toolbox mette a disposizione infatti tre codici principali:

- Allenamento della rete;
- Test del modello su una singola immagine;
- Test del modello su un set di immagini.

Per lanciare l'allenamento della rete, è richiesto di impostare svariati parametri. Nei seguenti sottocapitoli si analizzeranno i principali.

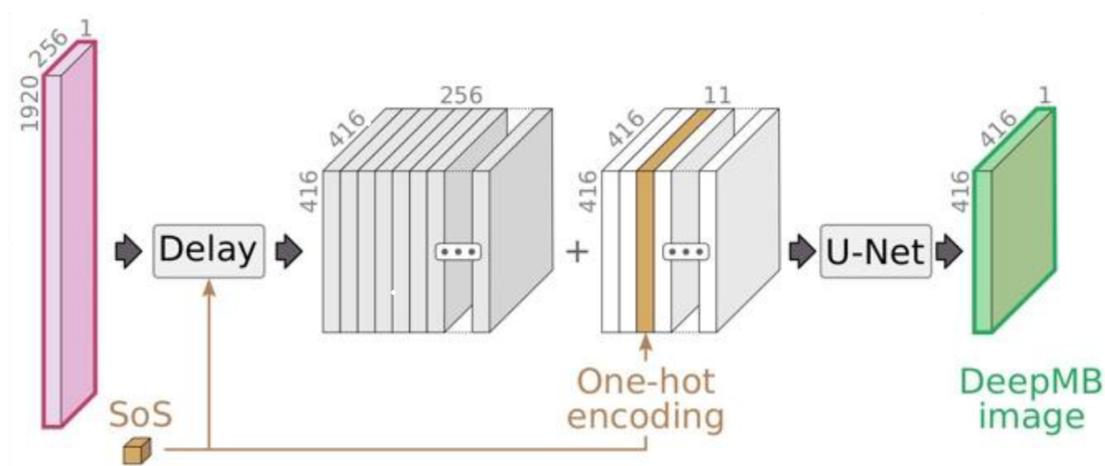


Figura 2.12: Struttura della rete [7]

### 2.5.1 Struttura generale

Come si nota in Fig. 2.12, l'allenamento si compone di più fasi. Il sinogramma, che ha dimensioni dipendenti dalla sonda e dal numero di campioni registrati da ciascun trasduttore, subisce un cambiamento di dominio, da quello del segnale a quello delle immagini. A loro volta, le velocità del suono vengono codificate e concatenate ai layer presenti. A questo punto una rete neurale, che si basa tipicamente sull'architettura U-NET, viene allenata avendo come immagini target quelle acquisite tramite model-based iterativo [7].

### 2.5.2 Dominio del segnale

Il toolbox propone una serie di reti neurali da applicare al sinogramma di partenza come preprocessing, prima quindi di passare al dominio delle immagini. Questo è il principale fattore su cui si è deciso di soffermarsi, provando a valutare come cambiano le performance in base all'applicazione di una di queste tecniche di preprocessing, in particolare la convoluzione 2D.

Le reti proposte sono [13]:

- IDENTITY (nessun pre-processing);
- UNET;
- CONV2D.

In Tab. 2.5 vengono elencati i principali parametri definiti per le reti precedenti e in quali casi vengono utilizzati.

Tabella 2.5: Parametri definiti per il pre-processing dei sinogrammi e dove vengono utilizzati. [13]

Parametro	Valore	Identity	U-Net	Conv2D
KERNEL_SIZE_UNET_SIG	(3, 3)	X	✓	✓
PAD_SIZE_UNET_SIG	(1, 1)	X	✓	✓
NB_LAYERS_UNET_SIG	3	X	✓	X
NGF_UNET_SIG	128	X	✓	X

### 2.5.3 Trasformazione del dominio

Lo step successivo è quindi il passaggio di dominio dal segnale all'immagine. Questo avviene tramite una funzione che tiene conto della frequenza di campionamento della sonda utilizzata e dalla distanza in pixel tra i trasduttori, ovvero dalla risoluzione [13].

## 2.5.4 Dominio dell'immagine

### 2.5.4.1 Codifica e concatenazione velocità del suono

Una delle principali innovazioni della tecnica presentata è la possibilità di rendere la velocità del suono un parametro allenabile, concatenando appositi layer codificati prima di allenare la rete. In [7] è stato infatti dimostrato come non fornire alcun valore oppure fornire un valore fisso di velocità del suono abbia effettivamente diminuito le prestazioni [Tab. 2.6].

Tabella 2.6: Confronto tra DeepMB con velocità del suono allenabile e senza. [7]

	DeepMB	DeepMB <sub>NO-SOS</sub>	DeepMB <sub>SCALAR-SOS</sub>
Data residual norm	0.156±0.088	0.164±0.083	0.169±0.081

Il data residual norm è una metrica che valuta quanto la ricostruzione si discosta dal corrispettivo sinogramma [7] e viene calcolata come:

$$R = \frac{\|M_{SOS} p_0 - s\|_2^2}{\|s\|_2^2}. \quad [7] \quad (2.6)$$

Come già accennato, per motivi di riservatezza legati ai dati della sonda iThera, le metriche utilizzate in questo lavoro sono differenti e valutano classicamente quanto la ricostruzione si discosta dal metodo di riferimento.

### 2.5.4.2 Rete

Nel dominio dell'immagine il toolbox propone più reti che si basano quasi tutte su una struttura di tipo U-Net.

Le reti proposte sono [13]:

- UNET;
- DAMPING+UNET (D+U-Net);
- SUM+UNET (S+U-Net);
- SUM (S);
- DAMPING (D).

In Tab. 2.7 vengono elencati i principali parametri definiti per le reti precedenti e in quali casi vengono utilizzati.

Tabella 2.7: Parametri definiti per la rete e dove vengono utilizzati. [13]

Parametro	Valore	U-Net	D+U-Net	S+U-Net	S	D
KERNEL_SIZE_UNET_IMA	(3, 3)	✓	✓	✓	X	X
PAD_SIZE_UNET_IMA	(1, 1)	✓	✓	✓	X	X
NB_LAYERS_UNET_IMA	5	✓	✓	✓	X	X
NGF_UNET_IMA	64	✓	✓	✓	X	X
CH_STEPS_DAMPING	[208, 160, 112]	X	✓	X	X	✓
KERNEL_SIZE_DAMPING	(3, 3)	X	✓	X	X	✓
PAD_SIZE_DAMPING	(1, 1)	X	✓	X	X	✓

Come verrà discusso più avanti, il costo computazione dell'allenamento non è per nulla trascurabile. Infatti, anche con calcolatori potenti, la durata è di più di 3 giorni.

### 2.5.5 Allenamenti

Anche se sono fornite più reti sia nel dominio delle immagini che nel dominio del segnale, la rete U-Net + Damping nel dominio delle immagini è stata valutata essere estremamente efficace [7]. Si è quindi deciso di concentrarsi sulla rete nel dominio del segnale, valutando due fattori:

- La ripetibilità del modello, non applicando alcun preprocessing ai sinogrammi e calcolando le prestazioni;
- Come cambiano le prestazioni applicando l'algoritmo CONV2D nel dominio del segnale, mantenendo costante la rete nel dominio delle immagini.

Questo algoritmo di preprocessing prevede in primo luogo di concatenare al sinogramma le velocità del suono e successivamente tramite convoluzione ridurre nuovamente il tutto ad un singolo layer, su cui viene applicata la trasformazione al dominio delle immagini.

Una volta ricavati i risultati, presentati nel prossimo capitolo, ci si è domandati quale impatto avesse l'introduzione di condizioni di stop, non previste originariamente. E' stato dunque proposto un terzo allenamento uguale al primo ma che si arrestasse dopo 10 epoche consecutive senza miglioramento in termini di funzione di perdita calcolata sul validation.

E' giusto precisare che i risultati che verranno mostrati fanno riferimento al modello salvato in corrispondenza dell'epoca che presenta il valore minimo in termini di funzione di perdita sul validation, e non al modello corrispondente all'ultima epoca.

## 2.5.6 Parametri generali

Di seguito vengono riportati i parametri definiti per entrambi i domini:

Tabella 2.8: Parametri generali. [13]

Parametro	Valore
BATCH_SIZE_TRN	4
BATCH_SIZE_VAL	4
BATCH_SIZE_TST	4
OPTIMIZER	SGD
NB_EPOCH	300
LEARNING_RATE	0.01
MOMENTUM	0.99
SCHEDULER_STEP_SIZE	1
SCHEDULER_GAMMA	0.99
LOSS_NICKNAME	MSE
OUTPUT_ACTIVATION	ABS

## 2.6 Metriche

Per calcolare le prestazioni dei test effettuati sono state prese in considerazione due metriche che danno informazioni su quanto la ricostruzione si discosta da un riferimento.

Il Mean Absolute Error restituisce la distanza media, in modulo, in termini di intensità della scala di grigi:

$$MAE = \frac{\sum_{i=1}^N |X_i - Y_i|}{N}, \quad (2.7)$$

dove  $N$  è il numero totale di pixel,  $X_i$  l'intensità dell' $i$ -esimo pixel del target e  $Y_i$  l'intensità dell' $i$ -esimo pixel della ricostruzione.

Il Root Mean Square Error fornisce un'informazione analoga a quella fornita dal Mean Absolute Error ma da un maggior peso agli errori più elevati, permettendo di distinguerli più facilmente. E' calcolata come la radice quadrata della media della differenza, al quadrato, dell'intensità:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (X_i - Y_i)^2}{N}}. \quad (2.8)$$

# Capitolo 3

## Risultati

In questo capitolo verranno presentati i risultati ottenuti da tre differenti allenamenti: nel primo non è stato applicato nessun preprocessing nel dominio del segnale, utilizzando dunque la rete IDENTITY; nel secondo è stata applicata una rete 2D convoluzionale (CONV2D) nel dominio del segnale; nel terzo è stato ripetuto il primo allenamento con l'aggiunta di condizioni di stop.

### 3.1 Identity

Per stimare in prima battuta la ripetibilità del modello e avere una base numerica su cui confrontare i successivi allenamenti, è stato effettuato un allenamento come proposto di default dal toolbox, ovvero:

- sig\_net\_archi = IDENTITY;
- ima\_net\_archi = DAMPING + UNET.

L'allenamento è durato poco più di tre giorni. Il modello allenato è stato applicato alle 8000 immagini di training set, alle 2000 immagini di validation set e alle 44 immagini di test set.

#### 3.1.1 Prestazioni: Training Set

Osservando le immagini target di partenza, si è notata la presenza di alcune immagini totalmente nere, o quasi. Il model-based iterativo non aveva dunque fornito un buon esito. Nel DeepMB queste immagini restituiscono di fatto le performance migliori, perché anche la rete ricrea immagini di bassissima intensità media: questo però avrebbe falsato in parte i risultati ottenuti. Si è deciso di comprendere nell'analisi dal decimo al novantesimo percentile del valore di MAE. Di seguito sono riportati media e deviazione standard delle metriche:

Tabella 3.1: Media e deviazione standard delle metriche calcolate sul training set, primo allenamento.

	Mean	Standard Deviation
MAE	2,62	1,39
RMSE	3,88	2,06

Il valore massimo osservato, in termini di intensità, tra le immagini target e le ricostruzioni per il training set è di 578,3: questo evidenzia come i risultati ottenuti siano notevoli, considerando anche l'enorme quantità di immagini. Di seguito sono riportate le immagini aventi le performance migliori [Fig. 3.1] e peggiori [Fig. 3.2].

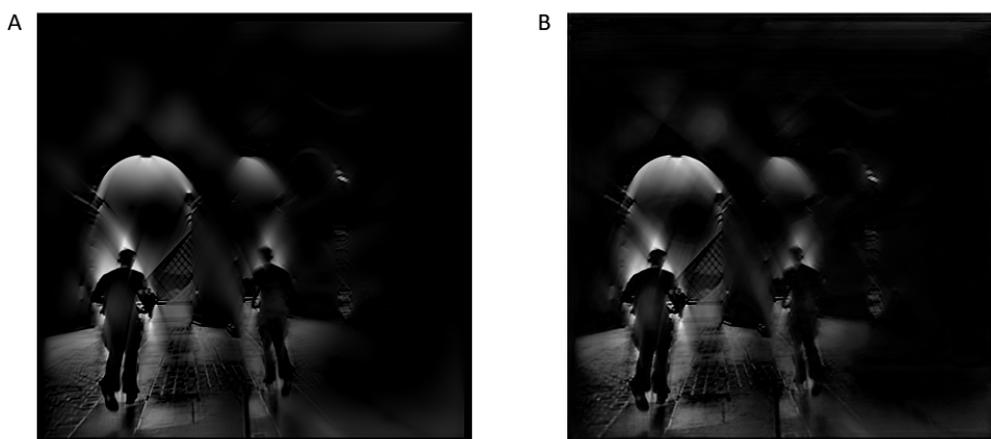


Figura 3.1: A) Immagine target. B) Immagine ricostruita con DeepMB. Training, primo allenamento. MAE = 0,61; RMSE = 0,95.

### 3.1.2 Prestazioni: Validation Set

Per il validation set valgono le considerazioni fatte precedentemente per il training set. Sono state quindi considerate le immagini comprese tra il decimo e il novantesimo percentile. Di seguito sono riportati la loro media e la deviazione standard:

Tabella 3.2: Media e deviazione standard delle metriche calcolate sul validation set, primo allenamento.

	Mean	Standard Deviation
MAE	2,60	1,36
RMSE	3,85	2,04

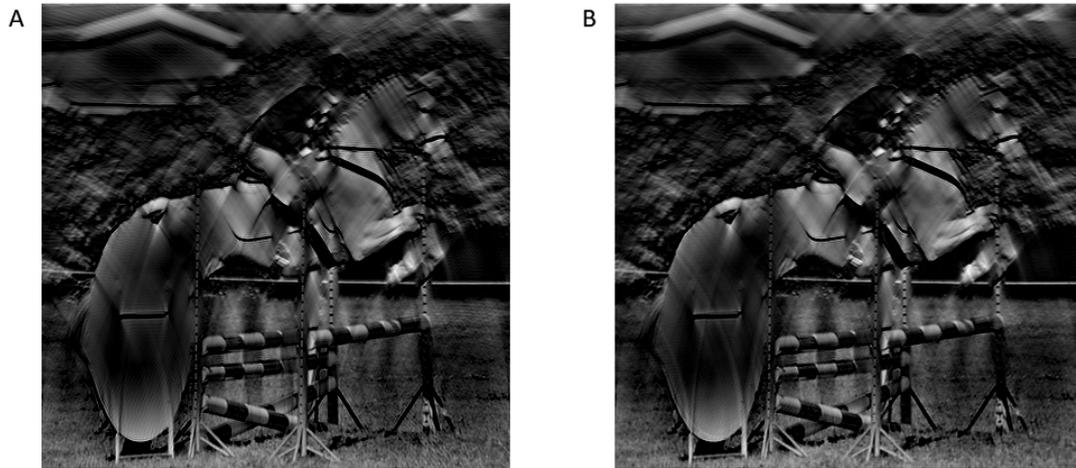


Figura 3.2: A) Immagine target. B) Immagine ricostruita con DeepMB. Training, primo allenamento. MAE = 5,69; RMSE = 7,84.

Il valore massimo osservato nel validation set è di 551,7: ancora una volta questo è sintomo della bontà della stima. I valori medi e le deviazioni standard ricalcano perfettamente quelli del training set. Anche in questo caso sono riportate le immagini aventi le performance migliori [Fig. 3.3] e peggiori [Fig. 3.4].

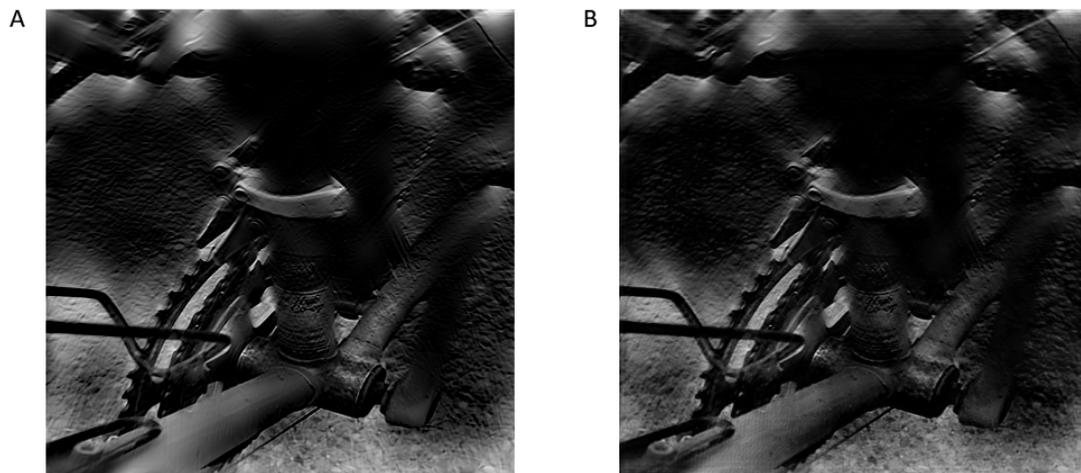


Figura 3.3: A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, primo allenamento. MAE = 0,62; RMSE = 0,87.

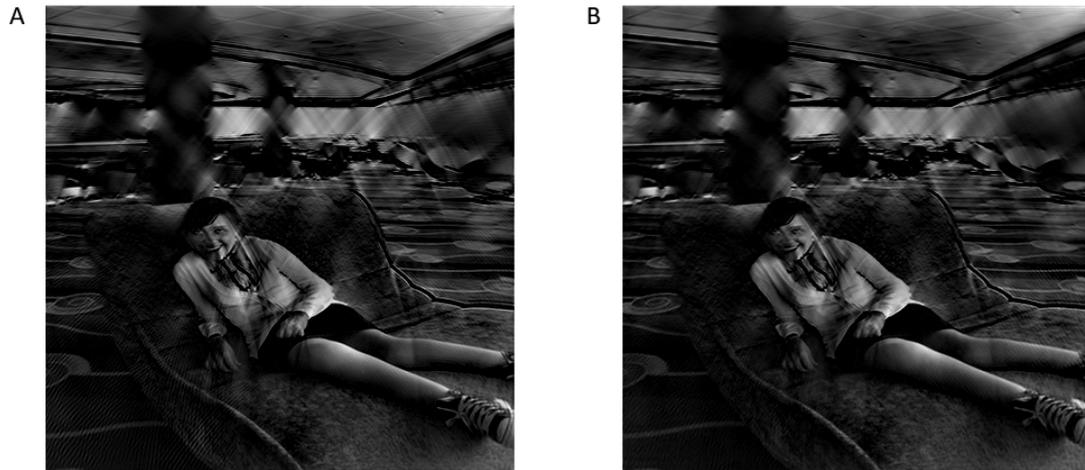


Figura 3.4: A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, primo allenamento. MAE = 5,68; RMSE = 8,08.

### 3.1.3 Prestazioni: Test Set

Visto lo scarso numero di dati, per il testing set si è deciso di considerare l'intera osservazione. Di seguito sono riportati media e deviazione standard delle metriche calcolate.

Tabella 3.3: Media e deviazione standard delle metriche calcolate sul test set, primo allenamento.

	Mean	Standard Deviation
MAE	0,86	0,37
RMSE	1,75	0,98

Il valore massimo osservato per il test set è 807,5. Medie e deviazioni standard sono inferiori a quelle calcolate per training e validation e questo conferma la capacità di generalizzare del DeepMB. Come di consueto, sono riportate a seguire le immagini aventi le performance migliori [Fig. 3.5] e peggiori [Fig. 3.6].

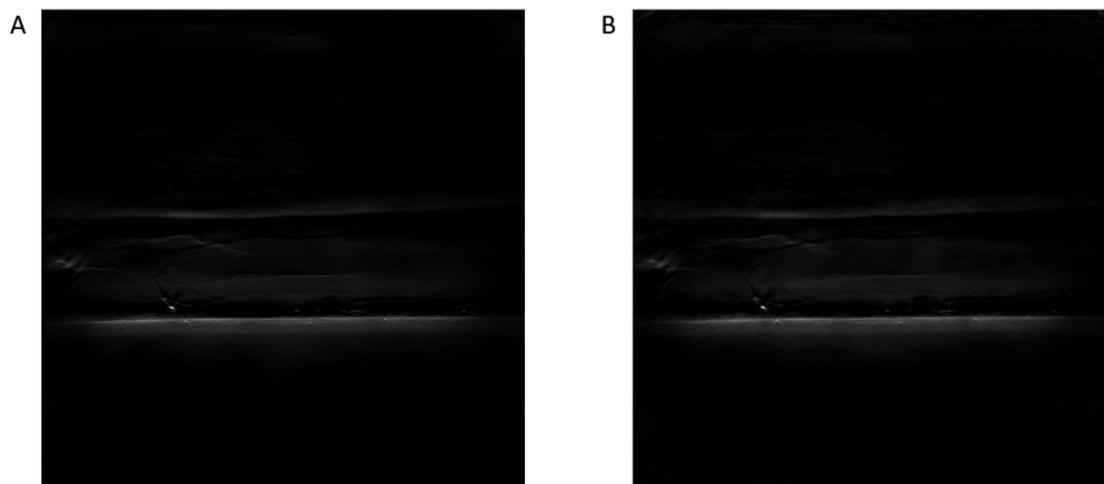


Figura 3.5: A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, primo allenamento. MAE = 0,49; RMSE = 0,78.

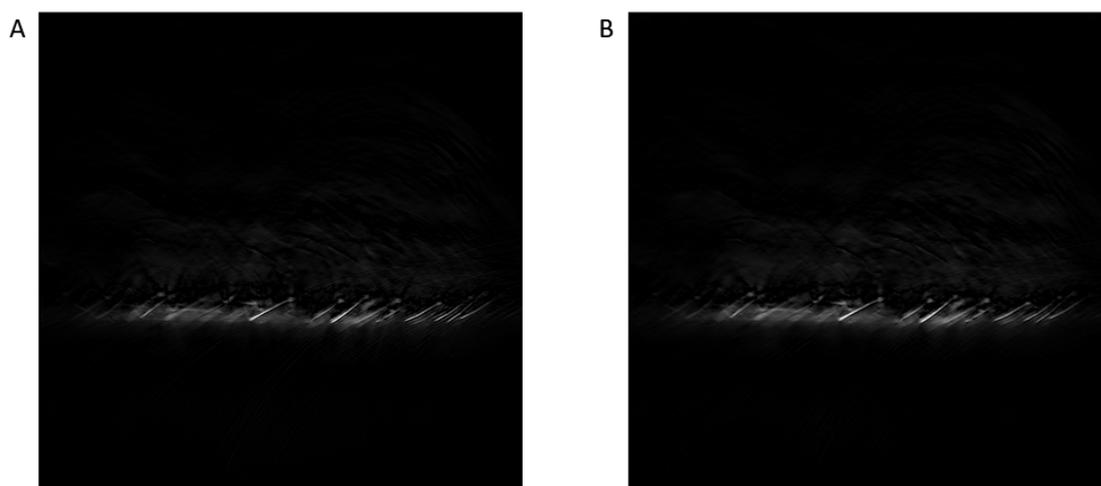


Figura 3.6: A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, primo allenamento. MAE = 2,03; RMSE = 4,91.

### 3.1.4 Riepilogo

Per rappresentare meglio i risultati ottenuti, si è pensato di rappresentare attraverso degli istogrammi la distribuzione dei valori medi delle metriche, in percentuale, tra training, validation e testing [Fig. 3.7 - 3.8].

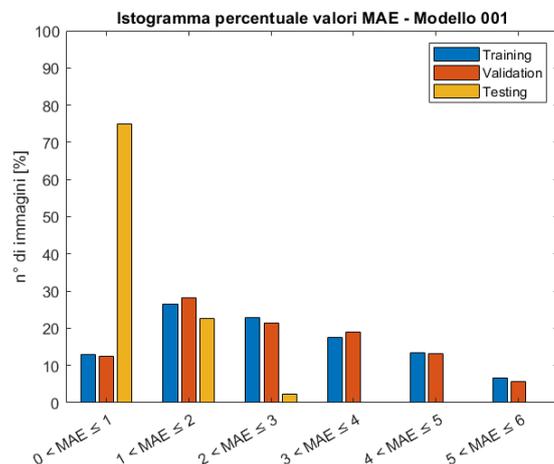


Figura 3.7: Istogramma percentuale dei valori MAE, primo allenamento.

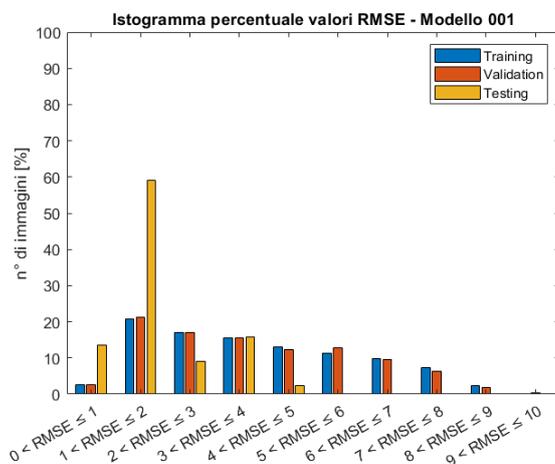


Figura 3.8: Istogramma percentuale dei valori RMSE, primo allenamento.

I grafici confermano quanto detto precedentemente, infatti il testing presenta un gran numero di immagini, in percentuale, negli intervalli più bassi. C'è da tenere in considerazione che il numero di immagini del testing è però molto inferiore, per cui anche considerando i valori percentuali sono dati da confermare con un maggior numero di osservazioni.

## 3.2 Conv2D

Si è pensato di applicare un preprocessing nel dominio del segnale, allo scopo di valutare le ricadute in termini di prestazioni. In particolare, si è applicata una convoluzione 2D dei sinogrammi, proposta dal toolbox. I parametri impostati sono:

- sig\_net\_archi = CONV2D;
- ima\_net\_archi = DAMPING + UNET.

L'allenamento è durato circa 4 giorni. Anche in questo caso il modello è stato applicato a tutte le immagini di training set, validation set e test set.

### 3.2.1 Prestazioni: Training Set

Visto che il dataset di partenza non è cambiato, sono stati effettuati ragionamenti analoghi a quanto visto per il primo allenamento. I parametri calcolati e mostrati successivamente [Tab. 3.4] sono stati quindi valutati tra il decimo e il novantesimo percentile del valore MAE.

Tabella 3.4: Media e deviazione standard delle metriche calcolate sul training set, secondo allenamento.

	Mean	Standard Deviation
MAE	2,87	1,47
RMSE	4,29	2,21

Anche in questo caso, il valore massimo osservato, in termini di intensità, tra le immagini target e le ricostruzioni per il training set è di 578,3. Anticipando il confronto che verrà approfondito successivamente, le prestazioni dimostrano la robustezza dell'algoritmo seppur peggiorando di poco in termini numerici. La differenza più evidente è nel fatto che il secondo allenamento ha impiegato circa un giorno in più: questo ha fornito lo spunto per un allenamento successivo, come discusso in seguito. Sono adesso riportate le immagini aventi le performance migliori [Fig. 3.9] e peggiori [Fig. 3.10].

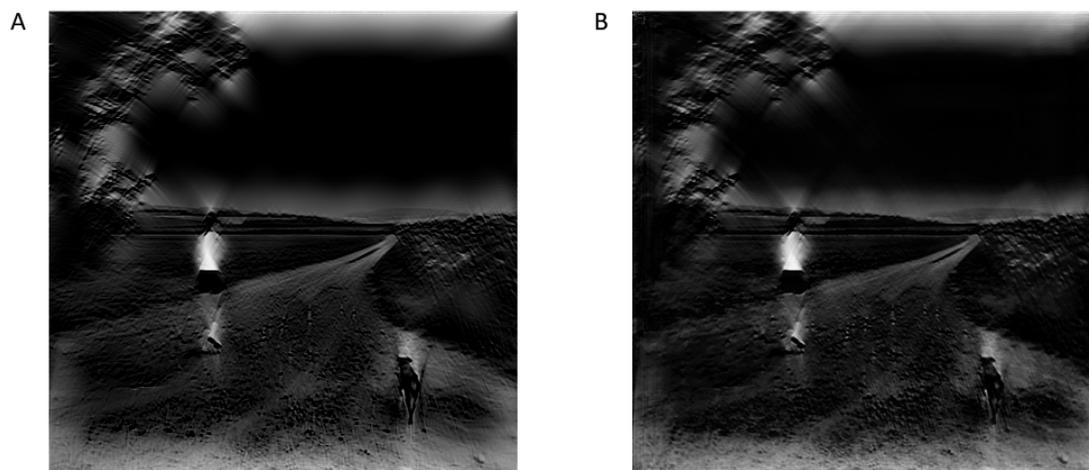


Figura 3.9: A) Immagine target. B) Immagine ricostruita con DeepMB. Training, secondo allenamento. MAE = 0,76; RMSE = 1,06.

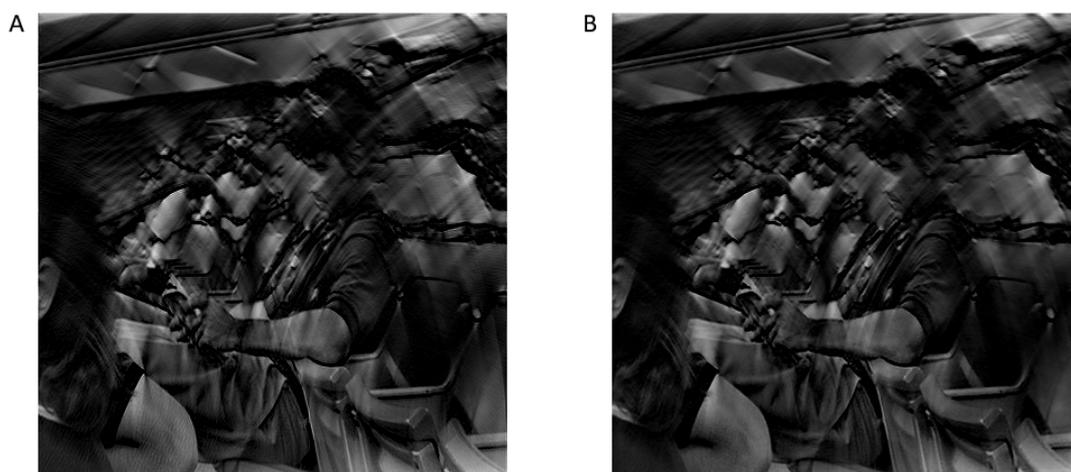


Figura 3.10: A) Immagine target. B) Immagine ricostruita con DeepMB. Training, secondo allenamento. MAE = 6,15; RMSE = 8,65.

### 3.2.2 Prestazioni: Validation Set

Per il validation set valgono le solite considerazioni. Di seguito sono riportate media e deviazione standard delle metriche.

Tabella 3.5: Media e deviazione standard delle metriche calcolate sul validation set, secondo allenamento.

	Mean	Standard Deviation
MAE	2,83	1,44
RMSE	4,25	2,19

Il valore massimo osservato tra ricostruzioni e immagini target nel validation set è di 568,6. Anche in questo caso i valori medi e le deviazioni standard ricalcano perfettamente quelli del training set. In seguito sono riportate le immagini aventi le performance migliori [Fig. 3.11] e peggiori [Fig. 3.12].

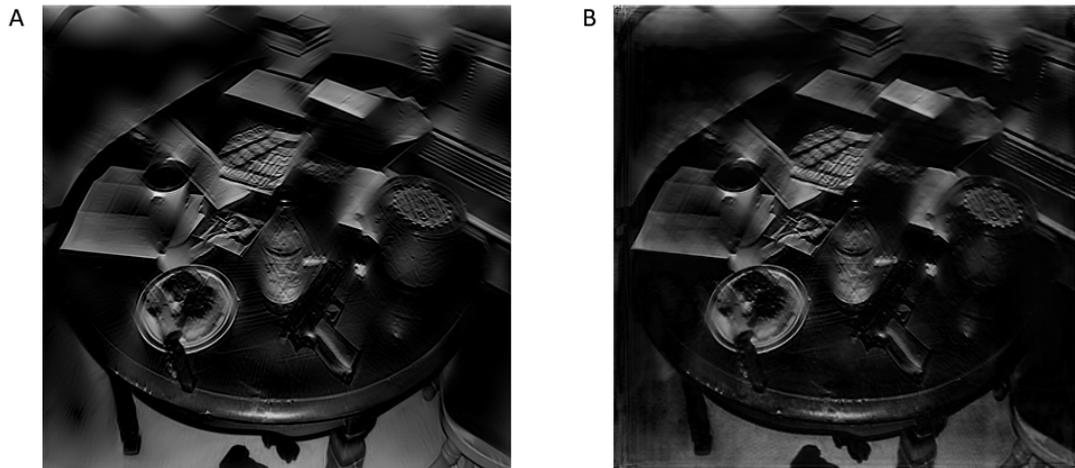


Figura 3.11: A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, secondo allenamento. MAE = 0,78; RMSE = 1,09.

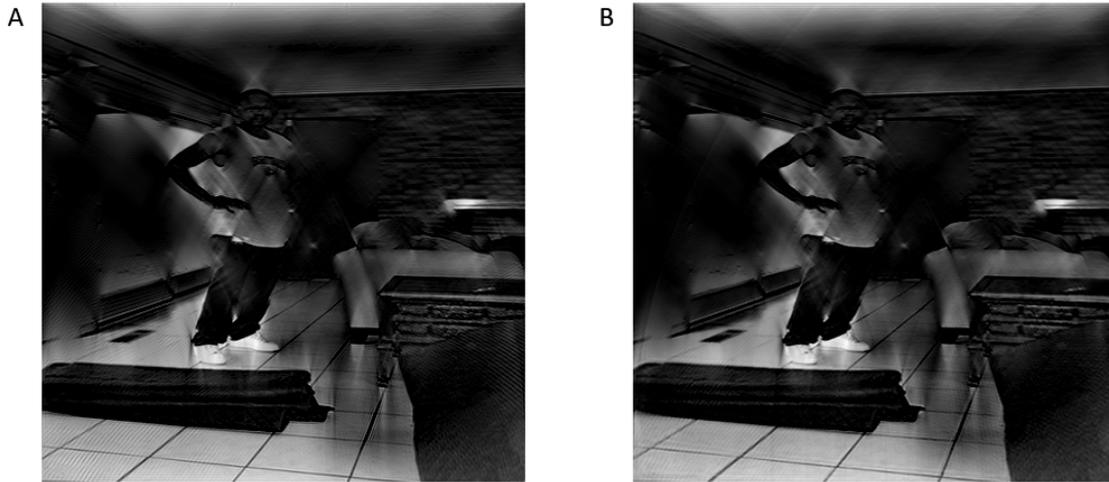


Figura 3.12: A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, secondo allenamento. MAE = 6,14; RMSE = 9,20.

### 3.2.3 Prestazioni: Test Set

Per il testing set si è nuovamente mantenuto l'intero numero di osservazioni. Di seguito sono riportati i risultati ottenuti:

Tabella 3.6: Media e deviazione standard delle metriche calcolate sul test set, secondo allenamento.

	Mean	Standard Deviation
MAE	0,93	0,40
RMSE	1,97	1,09

Il valore massimo osservato per il test set è ancora 807,5. Medie e deviazioni standard si confermano inferiori a quelle di training set e validation set, seppur di poco peggiori rispetto al primo allenamento. Come di consueto, sono riportate a seguire le immagini aventi le performance migliori [Fig. 3.13] e peggiori [Fig. 3.14].

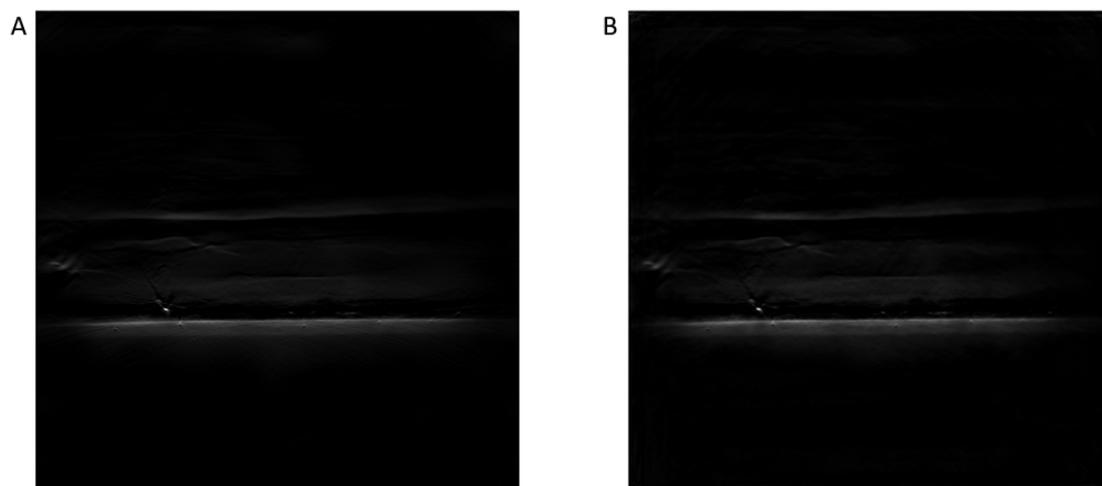


Figura 3.13: A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, secondo allenamento. MAE = 0,55; RMSE = 0,89.

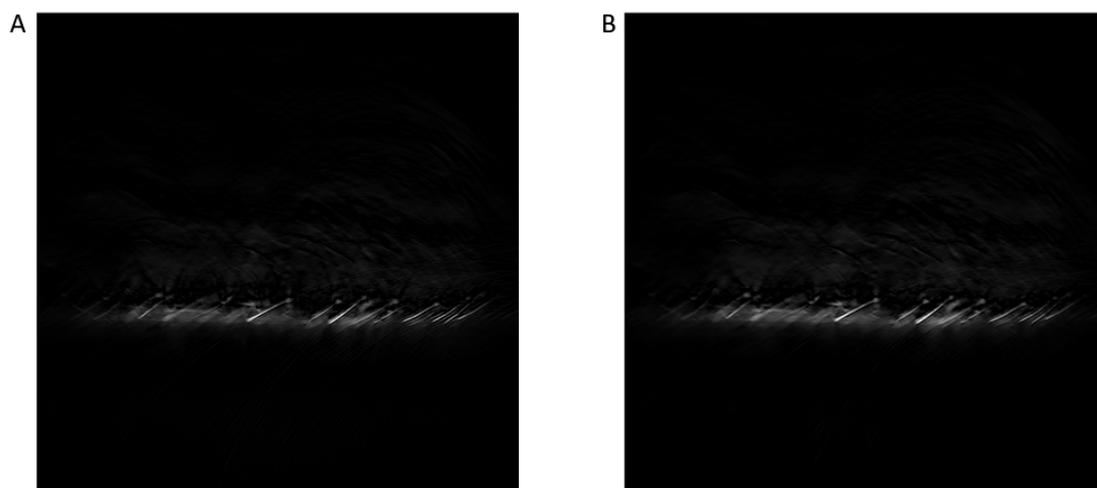


Figura 3.14: A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, secondo allenamento. MAE = 2,15; RMSE = 5,08.

### 3.2.4 Riepilogo

Come prima, sono rappresentate di seguito le distribuzioni percentuali dei valori tra training, validation e testing [Fig. 3.15 - 3.16].

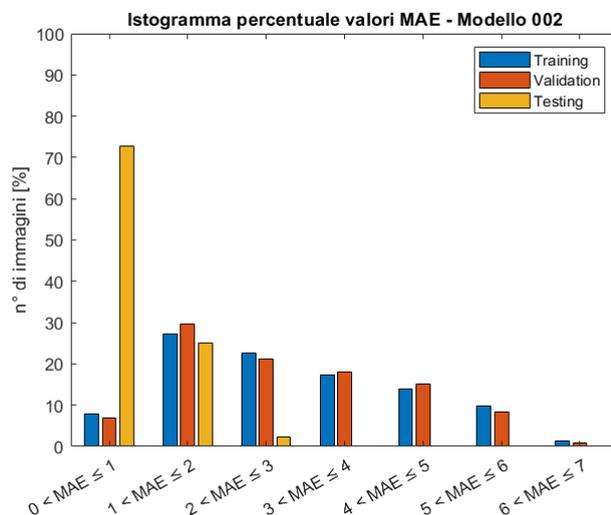


Figura 3.15: Istogramma percentuale dei valori MAE, secondo allenamento.

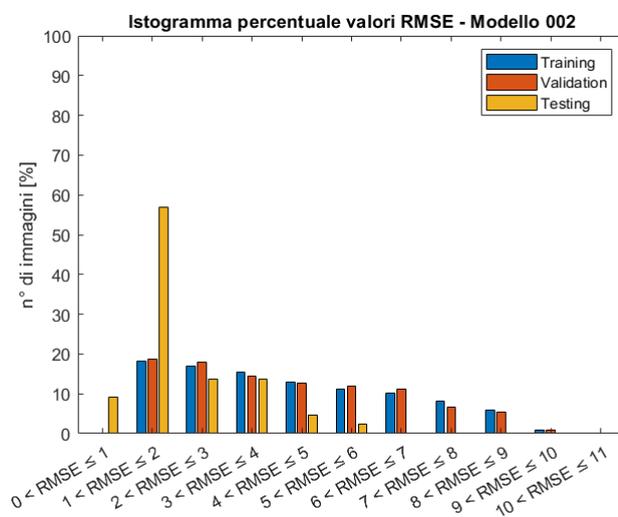


Figura 3.16: Istogramma percentuale dei valori RMSE, secondo allenamento.

Le considerazioni fatte precedentemente valgono anche in questo caso.

### 3.3 Condizioni di stop

Dai risultati riportati precedentemente, si è potuto constatare che l'applicazione dell'algoritmo nel dominio del segnale non ha cambiato significativamente le prestazioni. L'attenzione è stata allora spostata sulla durata dell'allenamento: nel primo il tempo richiesto è stato di tre giorni, nel secondo addirittura di quattro. Si è pensato dunque di applicare una condizione di stop, non troppo stringente, per valutare se riducendo il costo computazionale le prestazioni restassero elevate. Si è scelto di lavorare senza preprocessing visto che dai test precedenti era risultato il set-up con performance migliori. I parametri impostati sono dunque:

- sig\_net\_archi = IDENTITY;
- ima\_net\_archi = DAMPING + UNET;
- Condizione di stop: 10 epoche consecutive senza miglioramenti in termini di valore di loss sul validation.

I risultati che verranno analizzati adesso sono il frutto di circa due giorni di allenamento, visto che l'algoritmo si è arrestato all'epoca 187, risparmiando quindi circa un terzo delle epoche previste e di conseguenza anche di tempo. Il modello è stato applicato come di consueto a tutte le immagini di training set, validation set e test set.

#### 3.3.1 Prestazioni: Training Set

Anche in questo caso i parametri mostrati successivamente [Tab. 3.7] sono stati valutati tra il decimo e il novantesimo percentile del valore di MAE.

Tabella 3.7: Media e deviazione standard delle metriche calcolate sul training set, terzo allenamento.

	Mean	Standard Deviation
MAE	2,68	1,39
RMSE	3,96	2,08

Il valore massimo osservato, in termini di intensità, tra le immagini target e le ricostruzioni per il training set è di 578,3. Le prestazioni ancora una volta dimostrano la bontà e la ripetibilità del modello. Anticipando l'analisi di confronto, i dati ricalcano quelli ottenuti dal primo allenamento e, considerato che è durato quasi un giorno in meno, questo può essere visto come un miglioramento in termini di efficienza. Sono adesso riportate le immagini aventi le performance migliori [Fig. 3.17] e peggiori [Fig. 3.18].

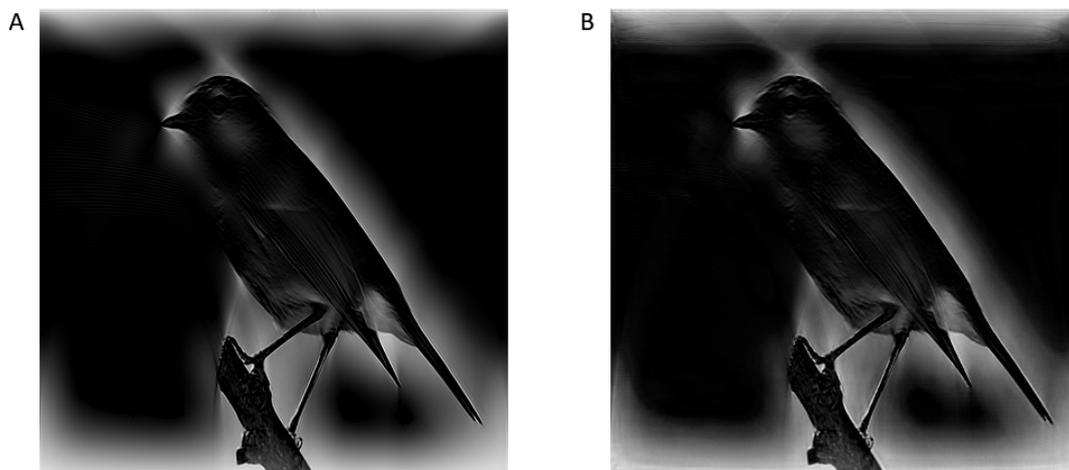


Figura 3.17: A) Immagine target. B) Immagine ricostruita con DeepMB. Training, terzo allenamento. MAE = 0,69; RMSE = 1,02.

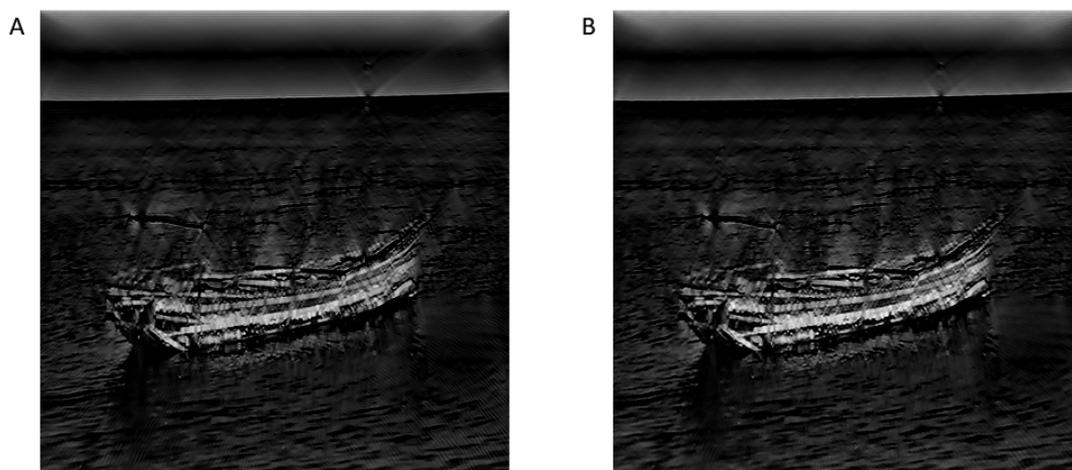


Figura 3.18: A) Immagine target. B) Immagine ricostruita con DeepMB. Training, terzo allenamento. MAE = 5,77; RMSE = 8,62.

### 3.3.2 Prestazioni: Validation Set

Per il validation set valgono le solite considerazioni. Di seguito sono riportate media e deviazione standard delle metriche calcolate:

Tabella 3.8: Media e deviazione standard delle metriche calcolate sul validation set, terzo allenamento.

	Mean	Standard Deviation
MAE	2,65	1,36
RMSE	3,92	2,07

Il valore massimo osservato tra ricostruzioni e immagini target nel validation set è di 551,7. Anche in questo caso i valori medi e le deviazioni standard ricalcano perfettamente quelli del training set. In seguito sono riportate le immagini aventi le performance migliori [Fig. 3.19] e peggiori [Fig. 3.20].

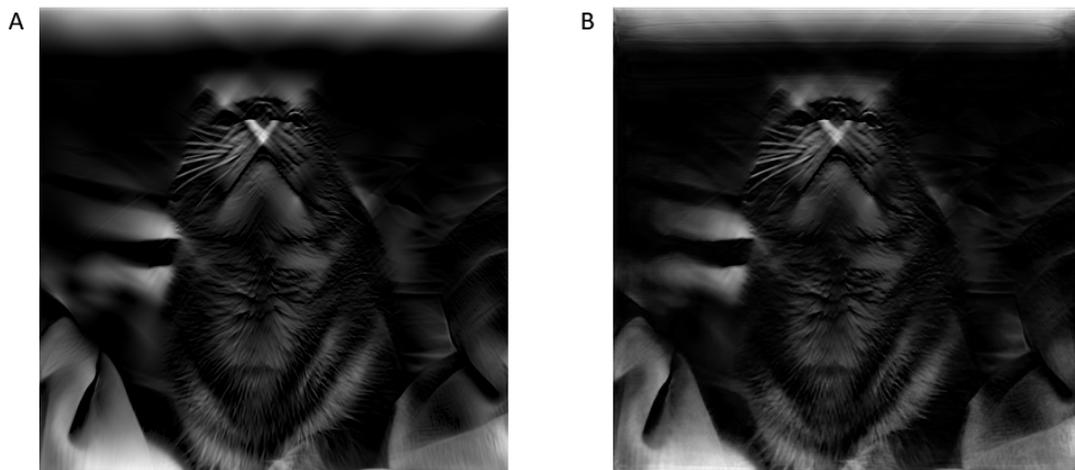


Figura 3.19: A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, terzo allenamento. MAE = 0,70; RMSE = 0,95.

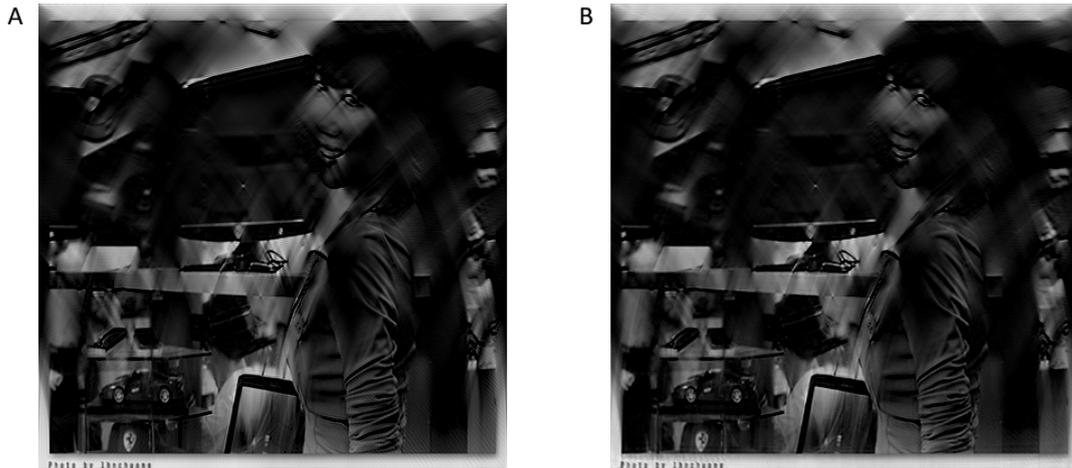


Figura 3.20: A) Immagine target. B) Immagine ricostruita con DeepMB. Validation, terzo allenamento. MAE = 5,77; RMSE = 8,63.

### 3.3.3 Prestazioni: Test Set

Di seguito sono riportati i risultati ottenuti applicando le metriche a tutte le immagini del testing set:

Tabella 3.9: Media e deviazione standard delle metriche calcolate sul test set, terzo allenamento.

	Mean	Standard Deviation
MAE	0,92	0,35
RMSE	1,81	0,98

Il valore massimo osservato per il test set è ancora 807,5. Ancora una volta medie e deviazioni standard si confermano inferiori a quelle di training set e validation set. I dati ricalcano quelli del primo allenamento, anche se ovviamente delle prestazioni peggiori, seppur di poco, sulle immagini in vivo hanno un peso maggiore. Come di consueto, sono riportate a seguire le immagini aventi le performance migliori [Fig. 3.21] e peggiori [Fig. 3.22].

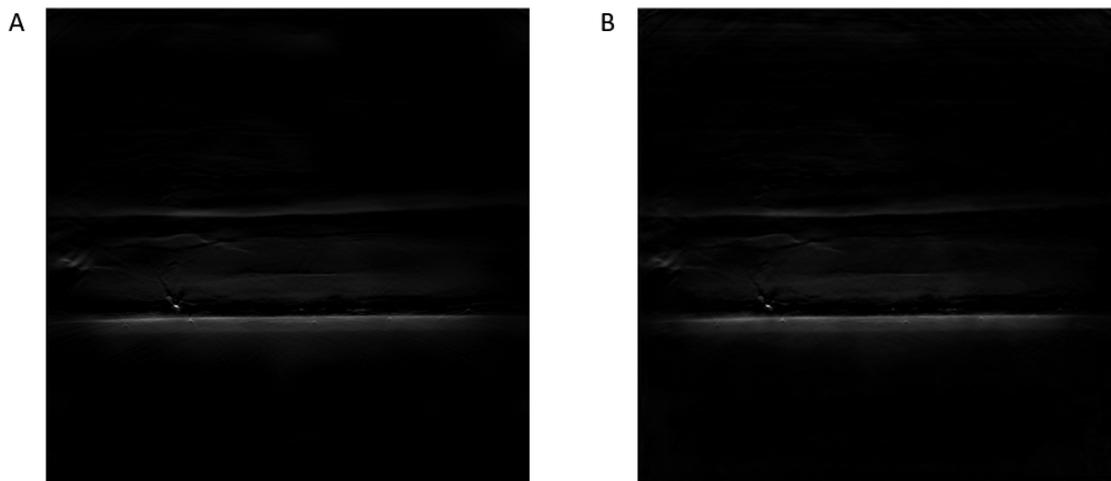


Figura 3.21: A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, terzo allenamento. MAE = 0,58; RMSE = 0,89.

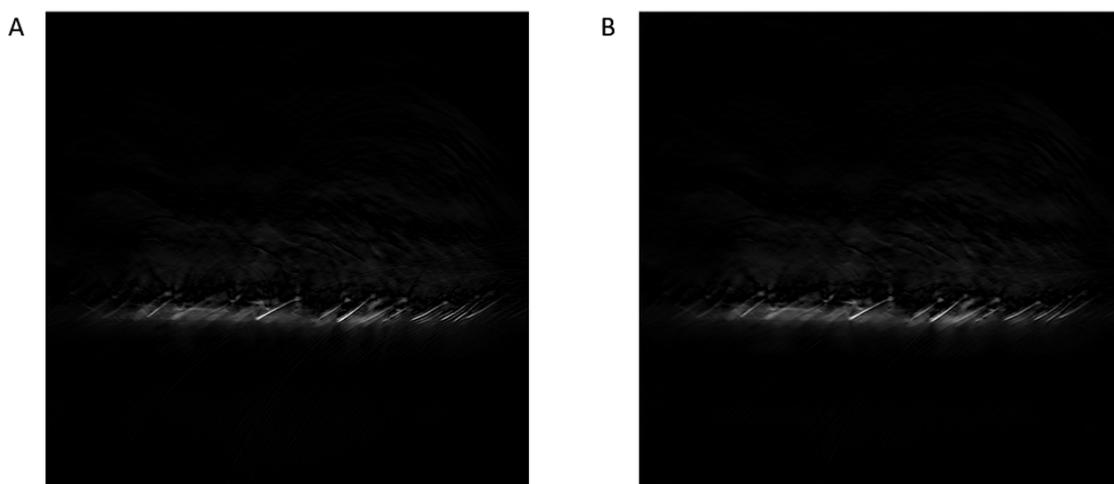


Figura 3.22: A) Immagine target. B) Immagine ricostruita con DeepMB. Testing, terzo allenamento. MAE = 2,06; RMSE = 5,04.

### 3.3.4 Riepilogo

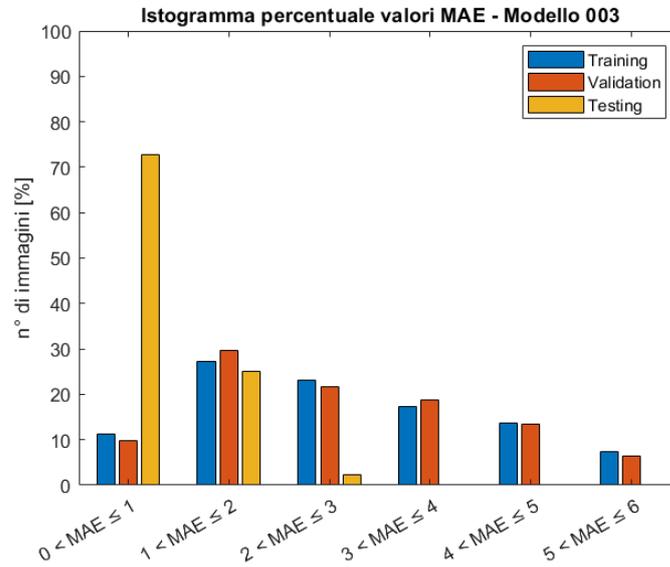


Figura 3.23: Istogramma percentuale dei valori MAE, terzo allenamento.

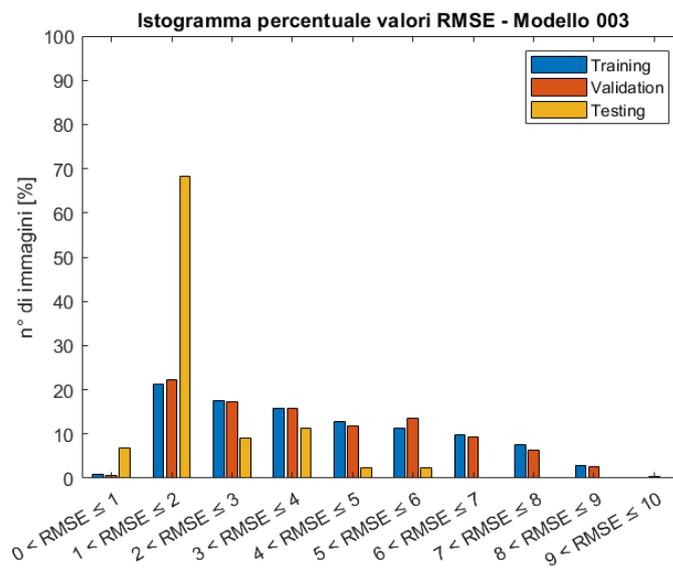


Figura 3.24: Istogramma percentuale dei valori RMSE, terzo allenamento.

I grafici confermano le considerazioni fatte precedentemente.

### 3.3.5 Approfondimento

Per valutare come cambiano le prestazioni andando avanti con l'allenamento e valutare eventuali problemi di overfitting, sono stati salvati per il terzo set-up i modelli ogni 20 epoche e riportate le medie delle metriche calcolate su Training Set, Validation Set e Testing Set [Fig. 3.25 - 3.26]. Per questa valutazione sono state considerate le prestazioni calcolate su tutte le immagini.

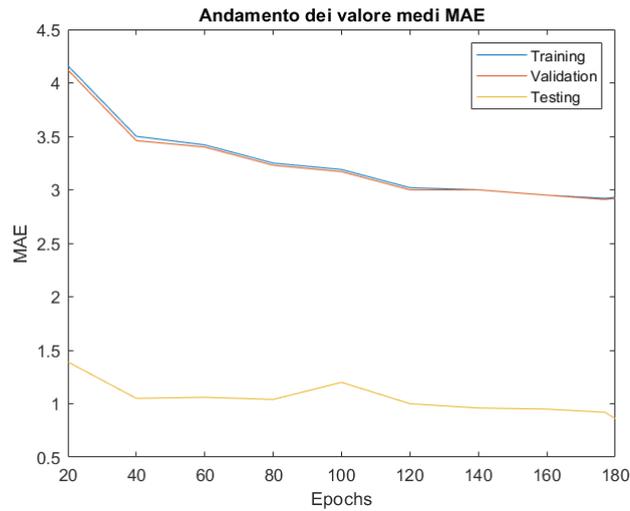


Figura 3.25: Andamento valori medi della metrica MAE in funzione delle epoche.

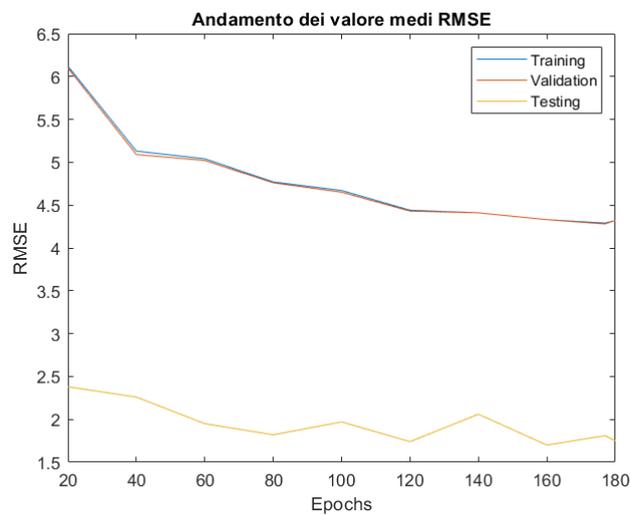


Figura 3.26: Andamento valori medi della metrica RMSE in funzione delle epoche.

Si può notare come gli andamenti di training set e validation set sono quasi perfettamente sovrapponibili. Diverso invece è il comportamento del testing set, che presenta un trend irregolare e che non segue training set e validation set. A dimostrazione di ciò, si può notare come dopo l'epoca 177, che è l'epoca su cui sono stati calcolati i risultati precedentemente riportati, il testing set ha riportato un miglioramento delle prestazioni all'epoca 180.

In definitiva, non si registra un fenomeno di overfitting, bensì un disallineamento sugli andamenti delle prestazioni calcolate sui dati reali rispetto alle immagini sintetiche. Un altro aspetto importante da notare è che si registrano fin da subito ottimi risultati sulle immagini in vivo.

## 3.4 Confronto

### 3.4.1 Metriche

In Fig. 3.27 sono riassunti i valori delle metriche medi estratti dai tre allenamenti:

	Modello 001			Modello 002			Modello 003		
	Mean	Standard Deviation		Mean	Standard Deviation		Mean	Standard Deviation	
A)	MAE	2,62	1,39	MAE	2,87	1,47	MAE	2,68	1,39
	RMSE	3,88	2,06	RMSE	4,29	2,21	RMSE	3,96	2,08
B)	MAE	2,60	1,36	MAE	2,83	1,44	MAE	2,65	1,36
	RMSE	3,85	2,04	RMSE	4,25	2,19	RMSE	3,92	2,07
C)	MAE	0,86	0,37	MAE	0,93	0,40	MAE	0,92	0,35
	RMSE	1,75	0,98	RMSE	1,97	1,09	RMSE	1,81	0,98

Figura 3.27: Confronto tra i tre allenamenti. A) Training Set, B) Validation Set, C) Testing Set.

Si nota come il primo allenamento riporta le performance migliori per tutti e tre i casi. Il secondo, invece, che applica il preprocessing risulta sempre il peggiore ma avvicina il terzo nel testing. Quest'ultimo riesce quasi a ricalcare il corrispettivo modello senza condizioni di stop durando circa un giorno in meno.

### 3.4.2 Training Set

In Fig. 3.28 - 3.29 sono riportati in un unico subplot gli istogrammi riportati precedentemente riferiti al training set per le tre reti. Si nota come la prima e la terza rete contengono più immagini nell'intervallo MAE compreso tra 0 e 1; al contrario, solo la seconda rete presenta un insieme di immagini, seppur piccolo, compreso nell'intervallo  $6 \div 7$ . Un discorso analogo vale per quanto riguarda il valore RMSE.

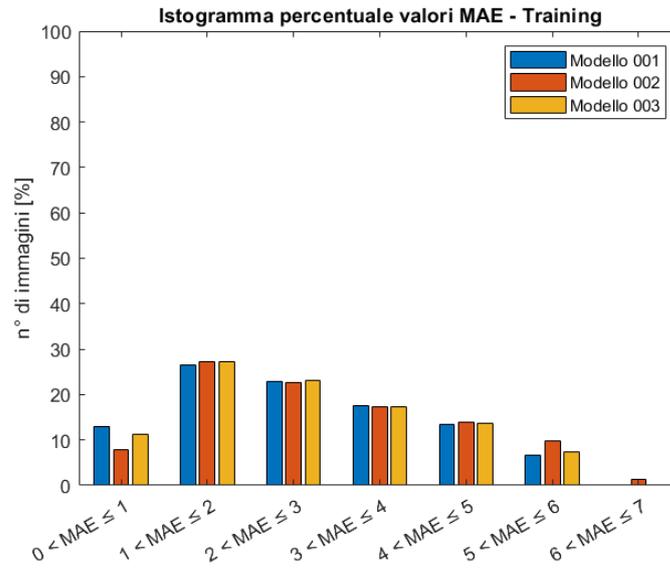


Figura 3.28: Istogramma percentuale dei valori MAE per le tre reti, training set.

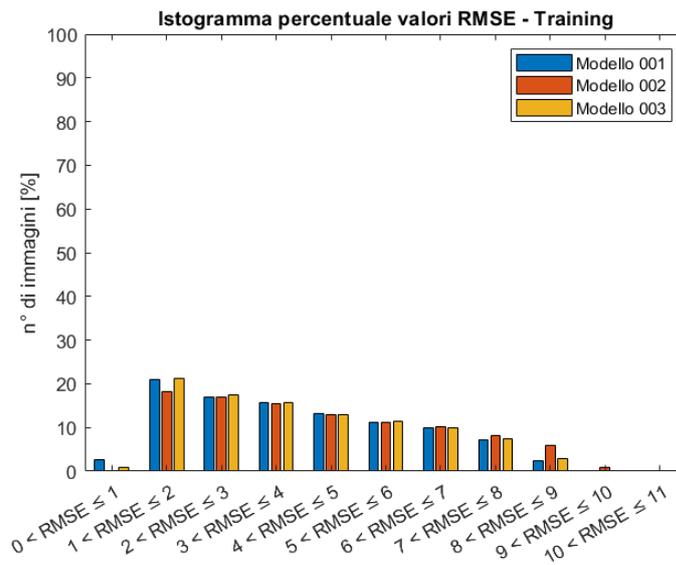


Figura 3.29: Istogramma percentuale dei valori RMSE per le tre reti, training set.

Primo e terzo allenamento si confermano molto simili tra loro.

### 3.4.3 Validation Set

Lo stesso trend si conferma per il validation set [Fig. 3.30 - 3.31].

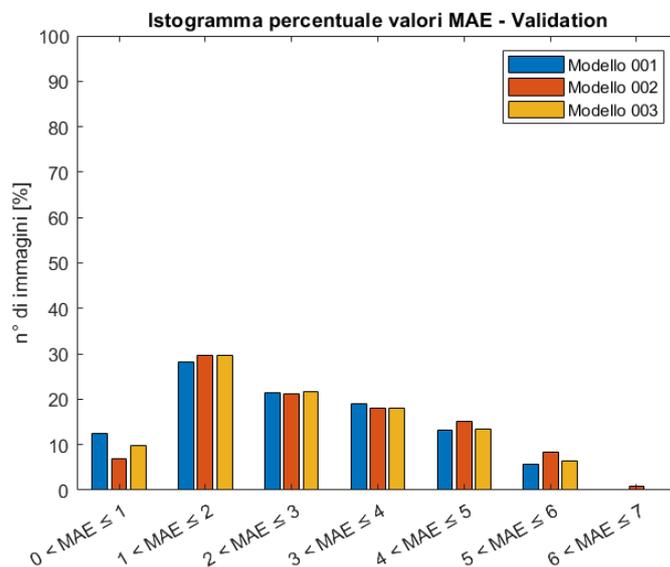


Figura 3.30: Istogramma percentuale dei valori MAE per le tre reti, validation set.

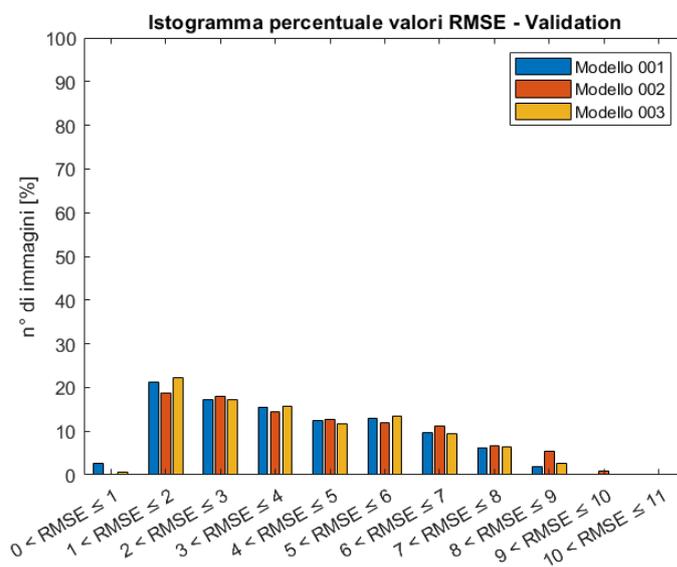


Figura 3.31: Istogramma percentuale dei valori RMSE per le tre reti, validation set.

### 3.4.4 Testing Set

Per il testing set [Fig. 3.32 - 3.33] la differenza è meno evidente, anche se soprattutto dal grafico riferito al valore RMSE si evidenzia il migliore rendimento della prima rete. Da notare come per la prima volta il numero di immagini del secondo allenamento compreso nell'intervallo RMSE  $0 < RMSE \leq 1$  è migliore rispetto al terzo.

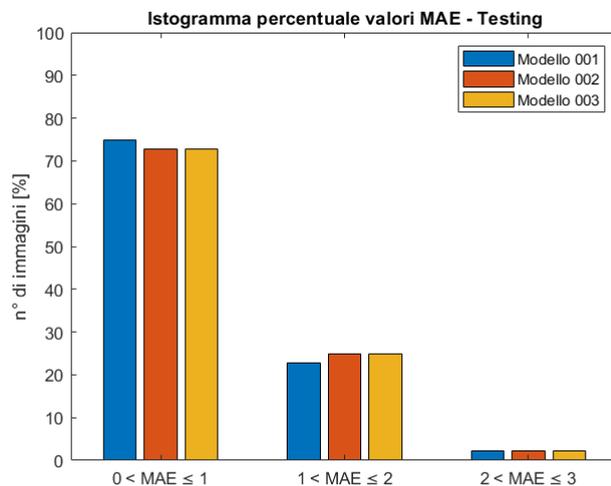


Figura 3.32: Istogramma percentuale dei valori MAE per le tre reti, testing set.

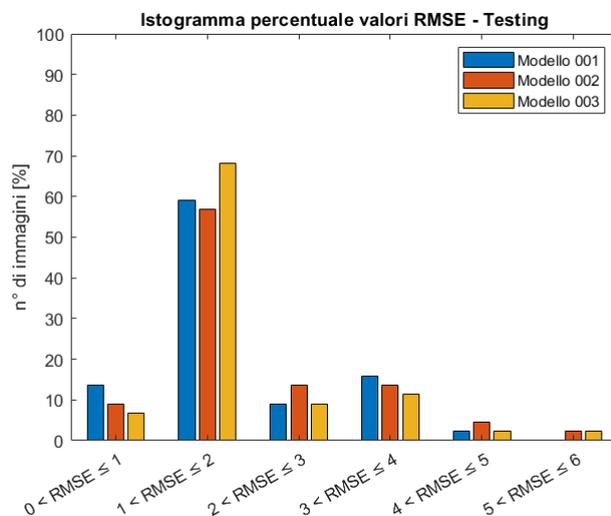


Figura 3.33: Istogramma percentuale dei valori RMSE per le tre reti, testing set.

## Capitolo 4

# Conclusioni

Dai test effettuati, il primo allenamento conferma la bontà dei risultati descritti in [7], sebbene sarebbe più opportuno per un'analisi più puntuale riferirsi alla stessa metrica. Il secondo e il terzo allenamento, a loro volta, dimostrano la robustezza del modello, visto che seppur modificando di poco il set-up i valori medi e le deviazioni standard delle metriche sono molto vicini tra loro.

Tutti gli allenamenti dimostrano un'elevata capacità di generalizzazione quando applicati ai dati in-vivo; un fattore da considerare è però il numero molto minore di immagini di testing rispetto a training e validation.

L'utilizzo di un algoritmo di preprocessing non ha portato a cambiamenti significativi; i risultati sono stati anzi leggermente peggiori. Ciò che è balzato all'occhio è stata la differenza in termini di costo computazionale, visto che il modello con preprocessing deve di fatto implementare due reti e questo si traduce in un allungamento delle tempistiche, già di per sè molto elevate.

Una possibile soluzione per ridurre il costo computazionale è l'utilizzo di reti pre-allenate; la seconda possibilità è l'implementazione di condizioni di stop non troppo stringenti. Il terzo allenamento sembra confermare questa versione, visto che un risparmio in termini di tempo di circa un terzo ha portato a risultati comparabili al corrispettivo allenamento senza condizioni di stop. Inoltre, non è stato osservato nessun fenomeno di overfitting.

Il maggiore punto di forza del modello è la sua capacità di elaborare le immagini in pochissimo tempo e con elevata accuratezza. L'inferenza su un elevato numero di immagini (più di 10000) è stato effettivamente molto rapido e questo lo rende un algoritmo real-time per l'elaborazione di immagini optoacustiche estremamente competitivo.

Come ulteriori sviluppi, l'obiettivo è quello di affinare il metodo, appena accennato in questo elaborato, per adattare il modello alla sonda Verasonics a disposizione, così da poter sfruttare le potenzialità del DeepMB. Inoltre, questo darebbe la possibilità di aumentare il numero di immagini acquisite in vivo.

# Bibliografia

- [1] Ntziachristos, V. & Razansky D. "*Molecular imaging by means of multispectral optoacoustic tomography (MSOT)*". Chem Rev 110, 2783-2794 (2010)
- [2] Karlas, Angelos, Michael Kallmayer, Nikolina-Alexia Fasoula, Evangelos Liapis, Michail Bariotakis, Markus Krönke, Maria Anastasopoulou, Josefine Reber, Hans-Henning Eckstein & Vasilis Ntziachristos. "*Optoacoustic Tomography of Muscle Perfusion and Oxygenation under Arterial and Venous Occlusion: A Human Pilot Study*". Journal of Biophotonics 13.6 (2020)
- [3] Binoy Bera & Madhumita Das Sarkar. "*Piezoelectric Effect, Piezotronics and Piezophotonics: A Review*". Imperial Journal of Interdisciplinary Research, Vol-2, Issue-11, 2016
- [4] <https://www.circuitbread.com/ee-faq/what-is-piezoelectric-effect>
- [5] De Dionig, S., & Bianchessi, M. *Studio e simulazione di sistemi di microbeamforming per ecografia 2D/3D con sonda CMUT*. Tesi di Laurea Magistrale, Politecnico di Milano. 2014
- [6] Xu, M. & Wang, L.V. "*Universal back-projection algorithm for photoacoustic computed tomography*". Phys Rev E Stat Nonlin Soft Matter Phys 71, 016706 (2005)
- [7] Christoph Dehner, Guillaume Zahnd, Dominik Jüstel & Vasilis Ntziachristos. "*DeepMB: Deep neural network for real-time optoacoustic image reconstruction with adjustable speed of sound*" (2022)
- [8] Bo Wang, Tianning Su, Weiran Pang, Ningning Wei, Jiaying Xiao & Kuan Peng. "*Back-projection algorithm in generalized form for circular-scanning-based photoacoustic tomography with improved tangential resolution*". Quantitative Imaging in Medicine and Surgery (2019)
- [9] Chowdhury, Kaushik Basak, Jaya Prakash, Angelos Karlas, Dominik Justel & Vasilis Ntziachristos. "*A Synthetic Total Impulse Response Characterization Method for Correction of Hand-Held Optoacoustic Images*". IEEE Transactions on Medical Imaging 39.10 (2020): 3218-230

- [10] Chowdhury, Kaushik Basak, Maximilian Bader, Christoph Dehner, Dominik Jüstel & Vasilis Ntziachristos. *"Individual Transducer Impulse Response Characterization Method to Improve Image Quality of Array-based Handheld Photoacoustic Tomography"*. Optics Letters 46.1 (2021): 1-4
- [11] M. Kim, G. -S. Jeng, I. Pelivanov & M. O'Donnell. *"Deep-Learning Image Reconstruction for Real-Time Photoacoustic System," in IEEE Transactions on Medical Imaging*. Vol. 39, no. 11, pp. 3379-3390, Nov. 2020
- [12] Feng, Jinchao, Jianguang Deng, Zhe Li, Zhonghua Sun, Huijing Dou & Kebin Jia. *"End-to-end Res-UNet Based Reconstruction Algorithm for Photoacoustic Imaging"*. Biomedical Optics Express 11.9 (2020): 5321-340
- [13] <https://github.com/juestellab/deepmb>
- [14] <http://host.robots.ox.ac.uk/pascal/VOC/index.html>
- [15] <https://github.com/juestellab/mb-rec-msot>
- [16] [http://field-ii.dk/?downloading\\_2021.html](http://field-ii.dk/?downloading_2021.html)
- [17] <http://shearlab.math.lmu.de/software>
- [18] <http://www.lx.it.pt/~mtf/SpaRSA>
- [19] <https://ithera-medical.com/products/clinical-research/msot-acuity-echo>
- [20] <https://verasonics.com/verasonics-transducers>
- [21] <https://verasonics.com/vantage-systems>
- [22] [https://www.biomecardio.com/MUST/functions/html/getparam\\_doc.html](https://www.biomecardio.com/MUST/functions/html/getparam_doc.html)