

*Questa è la DEDICA:
ognuno può scrivere quello che vuole,
anche nulla . . .*

Introduzione

Questa è l'introduzione.

Indice

Introduzione	i
1 Rumore e Denoising	1
1.1 Imaging	1
1.2 Rumore	2
1.3 Denoising	5
1.3.1 Approccio Classico	5
1.3.2 Deep Learning	7
Conclusioni	15
A Prima Appendice	17
B Seconda Appendice	19
Bibliografia	21

Elenco delle figure

1.1	esempio di Feedforward Neural Network	10
1.2	Rappresentazione dell' operazione di convoluzione	12
1.3	Residual Learning	13

Elenco delle tabelle

Capitolo 1

Rumore e Denoising

Questo è il primo capitolo.

1.1 Imaging

Con il termine **medical imaging** ci si riferisce alla generica tecnica o processo di imaging attraverso il quale è possibile osservare un' area di un organismo non visibile dall'esterno, per analisi cliniche o una rappresentazione visiva delle funzioni di organi e tessuti.

Una **radiografia** è una tecnica di imaging che produce un radiogramma. Tale tecnica si basa sull' utilizzo di un fascio di fotoni, come raggi X o raggi Gamma, generati da una sorgente, passanti per della materia e catturati da un recettore. Durante il passaggio per il corpo interposto i raggi sono assorbiti o attenuati in maniera differente l'uno dall'altro creando un insieme di raggi di diversa intensità. Il recettore trasforma questo insieme di raggi di diversa concentrazione in un immagine "in negativo" del corpo, essendo catturati dal recettore i fotoni che non vengono assorbiti.

L'assorbimento dovuto alle diverse composizioni elementali del corpo gioca un ruolo fondamentale nell' imaging medico. Il processo di assorbimento è proporzionale al numero atomico della materia attraversata e questo permette di distinguere la materia grigia dalla materia bianca nell'immagine

generata. Per esempio tra il sangue e il muscolo non è osservabile un contrasto elevato, essendo la densità degli elementi che li compongono molto simile. Per generare un contrasto tra i due è necessario iniettare un liquido di contrasto composto da elementi con un alto numero atomico durante l'esposizione ai raggi.

I recettori localizzati in maniera opposta alla sorgente dei raggi ricevono i fasci di fotoni e generano un segnale elettrico. Questo segnale elettrico deve essere convertito in un segnale digitale da un convertitore analogico/digitale.

1.2 Rumore

Ci sono molti elementi che possono influenzare la qualità dell'immagine costruita, i fattori principali sono :

- Blurring: l' immagine potrebbe risultare sfocata per diverse ragioni, i motivi principali sono:
 1. Uso improprio della strumentazione.
 2. Movimento del paziente dovuto al respiro o al battito del cuore.
 3. Fluttuazioni dei valori dei pixel dell' immagine in una regione di materiale uniforme.
 4. Gli algoritmi usati per filtrare i raggi sfocano l'immagine.
- Field of View (FOV): La dimensione della regione anatomica inclusa nella scansione viene detta field of view. Se la dimensione della regione è troppo grande o troppo piccola, la qualità dell' immagine potrebbe risultare degradata.
- Artefatti: Gli artefatti sono distorsioni o errori nell' immagine digitale che non dipendono dall' oggetto scansionato. Gli artefatti più comuni sono:
 1. Beam Hardening: Beam Hardening avviene quando l'energia media di un raggio passante per il paziente aumenta. Per evitare che

ciò avvenga alcune soluzioni possono essere messe in pratica, aumentare il kvp(peak kilovoltage), decrementare la grandezza delle slice(in caso di tomografia computerizzata), pre-filtrare i raggi, evitare regioni ad alto assorbimento, soluzioni algorimiche.

2. Artefatti metallici: Materiali metallici possono causare soprattutto nelle CT i cosiddetti streaking artifacts (fasci di luce), dovuti dal blocco delle informazioni di proiezione. Per ridurre questo fenomeno è necessario rimuovere il materiale metallico se possibile.
 3. Movimento del paziente: Il movimenti volontari o involontari del paziente possono produrre streaking artifacts. Si può ridurre il tempo di scansione e migliorare il posizionamento del paziente.
 4. Software e Hardware: Artefatti derivanti da implementazioni o configurazioni scorrette di algoritmi di costruzione dell'immagine digitale o degradazione dei componenti elettronici e meccanici.
- Rumore Visivo: Il rumore visivo è una variazione casuale della luminosità dei singoli pixel dell' immagine. Questo tipo di rumore deriva di solito da errori nell' acquisizione, trasmissione del segnale ed errori computazionali matematici. Questo tipo di rumore ha una maggiore influenza su soggetti a basso contrasto.

La dose di radiazioni a cui è sottoposto il paziente è uno dei fattori più importanti nella generazione di rumore. Per la sicurezza del paziente è necessario minimizzare il tempo di esposizione alle radiazioni, ma la conseguenza principale nel ridurre la dose di radiazioni è un aumento del rumore dell'immagine facendo diminuire la qualità della diagnosi.

Ci sono due approcci principali per cercare di ridurre il rumore visivo senza sacrificare la qualità dell'immagine:

1. Facendo uno studio ed una ottimizzazione della dose di radiazioni le immagini a bassa dose di radiazioni possono essere migliorate.
2. La qualità dell' immagine può essere migliorata sviluppando algoritmi per la rimozione del rumore nelle immagini. Questi algoritmi possono

essere successivamente utilizzati per ridurre la dose di radiazioni. Il processo di rimozione di rumore da immagini è conosciuto con il termine di **image denoising**.

Le radiografie sono caratterizzate da una sensitibilità ai contrasti, che permettono di distinguere tra i tessuti molli del corpo umano. Questa caratteristica è particolarmente influenzata dal rumore che nuoce alla visualizzazione delle strutture a basso contrasto.

Prima di considerare tecniche di image denoising è necessario comprendere quali sono le sorgenti principali di rumore visivo:

Rumore randomico e statistico: Può essere generato dalla ricezione di un numero finito di quanti (pacchetti di energia) di raggi da parte del recettore. Durante il processo non c'è nessuna forza che distribuisce i fotoni in maniera uniforme sulla superficie, un area del recettore potrebbe percepire più fotoni di un'altra. La causa principale di questo rumore è la cosiddetta radiazione secondaria, prodotta dall'interazione dei fotoni emessi dal generatore con il corpo in scansione, disperdendo ulteriormente in maniera casuale le particelle. Il risultato del rumore viene percepito come una fluttuazione nella densità dell'immagine. Come risultato si ha che il cambiamento della densità dell'immagine è impredicibile con comportamento randomico. Questo tipo di rumore è anche chiamato **rumore quantico**.

Rumore Elettrico: Il processo di un circuito elettrico che riceve segnali analogici può generare un disturbo del segnale. La sorgente più comune di rumore è il rumore termico intrinseco di ogni elemento dissipativo (es. resistori) che si trovi ad una temperatura diversa dallo zero assoluto.

Errori di arrotondamento: I segnali analogici sono convertiti in segnali digitali usando algoritmi di quantizzazione del segnale. Per ricostruire l'immagine ora discreta viene elaborata da un calcolatore. La rappresentazione in memoria del segnale deve essere necessariamente finita, essendo limitata dal numero di bit usati dal sistema numerico implementato dal computer per il calcolo. La computazione matematica non è dunque possibile senza errori di arrotondamento.

Generalmente quindi il rumore nella ricostruzione di immagini mediche viene introdotto per due motivi. Il primo è un rumore variabile dovuto al rumore elettrico e dagli errori di arrotondamento e può essere modellato con un semplice rumore additivo. Il secondo è il rumore dovuto dalle fluttuazioni di fotoni percepiti nelle diverse zone del recettore, più complesso da modellizzare. Il rumore può essere calcolato statisticamente utilizzando la deviazione standard dell'intensità dei pixel in una regione fisica uniforme dell'immagine. L'analisi statistica dei valori dei pixel produce una distribuzione con forma a campana. La variazione standard, cioè la misura della variazione dei pixel nella zona di interesse, rappresenta il livello di rumore nell'immagine. Un valore alto della deviazione standard implica un alto livello di rumore. Questa distribuzione statistica mostra che il valore medio del rumore è uguale alla varianza del rumore e può essere analizzato che il rapporto segnale-rumore è proporzionale alla radice quadrata della dose di raggi emessi. Dunque la quantità relativa di rumore quantico percepita dai recettori diminuisce all'aumentare della dose di raggi emessi. Il problema di identificare la distribuzione esatta del rumore deriva dal fatto che tutti i passi intermedi nella generazione dell'immagine introducono una correlazione statistica con i valori contenenti rumore catturati dal recettore. A causa di queste dipendenze la distribuzione esatta del rumore nell'immagine finale risulta sconosciuta.

Sperimentalmente si è potuto osservare come il rumore quantico sia modellizzabile con una distribuzione di Poisson mentre il rumore termico con una distribuzione di Gauss.

1.3 Denoising

1.3.1 Approccio Classico

Con image denoising si intende il problema di rimuovere rumori e distorsioni da una immagine. Generalmente si considera l'immagine rumorosa come una somma tra l'immagine originale e una componente di rumore, que-

sto comporta che l'obbiettivo delle tecniche di denoising sia quello di ridurre la componente rumorosa cercando di mantenere nell'immagine pulita le caratteristiche dell' immagine originale.

Quando si tratta di denoising applicato a imaging medico l'obbiettivo del denoising è quello di ridurre il rumore mantendendo i dettagli clinici dell' immagine in modo tale da mantenere l'immagine utile per una corretta diagnosi. Nel tempo sono stati sviluppati filtri di lisciamento (smoothing) e di sharpening. I primi in caso di alti livelli di rumore tendono a sfocare l'immagine e a non preservare i bordi dell'oggetto, mentre i filtri di sharpening non eccellono nel mantenere piccoli dettagli dell' immagine. In caso di denoising per analisi di immagini mediche questo presenta un grande problema in quanto mantenere più dettagli possibili dell'immagine originale è di fondamentale importanza.

I problemi principali nel denoising di immagini mediche sono:

- Le regioni piatte (zone aventi una variazione minima dei valori dei pixel) dell'immagine devono rimanere zone piatte.
- I contorni degli oggetti devono rimanere preservati.
- I dettagli di struttura dell'oggetto non devono essere persi.
- Il contrasto dell'immagine deve essere preservato.
- Non devono formarsi nuovi artefatti.

Generalmente ci sono due approcci per il denoising di immagini : spatial domain filtering e transform domain filtering [A survey of edge preserving image denoising].

Filtri Spaziali

I filtri spaziali agiscono applicando un filtro direttamente all' immagine con rumore, e possono essere definiti in filtri lineari e non lineari.

Filtri Lineari

I filtri lineari come il filtro medio (Mean filter) o il filtro di Wiener tendono a sfocare i bordi dell' oggetto, distruggere linee e dettagli nell'immagine, e generalmente non performano bene nel caso di rumore dipendente dal segnale come il rumore poissoniano. L'idea del filtro medio è molto semplice, consiste nel sostituire ogni valore di ogni pixel con il valore medio dei pixel vicini, incluso se stesso. Il filtro di Wiener invece richiede informazioni sullo spettro del rumore e dell'immagine originale.

Filtri Non Lineari

I filtri non lineari invece rimuovono il rumore senza cercare di identificarlo. Per cercare di risolvere i problemi dei filtri lineari una serie di filtri non lineari sono stati sviluppati: filtro mediano, filtro mediano pesato e filtro mediano rilassato.

Transform Domain Filtering

A differenza dei metodi spaziali, i metodi di trasformazione prima ottengono una trasformazione dell' immagine rumorosa e successivamente applicano la procedura di denoising sulla trasformazione. Si dividono in fitri Data adaptive e Non-Data adaptive.

1.3.2 Deep Learning

I metodi di **Machine Learning** si dividono in supervisionati, semi-supervisionati e non supervisionati. I metodi supervisionati utilizzano delle informazioni associate all' input (labels) per imparare dei parametri e trainare il modello di denoising. Per esempio dato il seguente modello

$$y = x + \mu$$

dove x , y e μ rappresentano rispettivamente l' immagine originale, l' immagine rumorosa e un rumore additivo Gaussiano (AWGN) con deviazione

standard σ . Il modello di denoising utilizza le coppie $\{x_k, y_k\}_{k=1}^N$ per imparare i parametri della funzione f che cerca di rimuovere il rumore, dove x_k e y_k sono la k-esima immagine originale e la k-esima immagine con rumore; N è il numero di immagini totali. Il processo può essere descritto come

$$x_k = f(y_k, \theta, m)$$

dove θ sono i parametri e m la quantità di rumore.

I metodi non supervisionati invece utilizzano gli esempi di training per cercare patterns nel data, invece di fare label matching come nel caso supervisionato. Nel caso dei metodi semi-supervisionati invece si cerca di imparare una distribuzione del data in modo da poter assegnare labels ai dati che non ne hanno.

Le **Reti Neurali** sono alla base dei metodi di machine learning, che a loro volta sono alla base dei metodi di deep learning. La maggior parte delle reti neurali sono composte da neuroni, input X , funzioni di attivazione ϕ , pesi $W = [W^0, W^1, \dots, W^{n-1}]$ e bias $b = [b^0, b^1, b^n - 1]$. A differenza dei modelli lineari come la regressione lineare e la regressione logistica, che possono modellizzare solo funzioni lineari senza comprendere l'interazioni tra due qualsiasi variabili di input, le reti neurali non applicano un modello lineare direttamente ad \mathbf{x} . Viene utilizzata invece una trasformazione $\phi(\mathbf{x})$, dove ϕ è una trasformazione non lineare.

$$y = f(\mathbf{X}; \mathbf{b}, \mathbf{W}) = \phi(\mathbf{W}^T \mathbf{X} + \mathbf{b}). \quad (1.1)$$

Se la rete neurale è composta da più "livelli" di profondità, dati da funzioni di attivazione e pesi diversi, la rete viene definita con il termine **multi-layer perceptrons** (MLPs) oppure **Deep feedforward networks** o **feed-forward neural networks**.

L' obiettivo delle reti feedforward è quello di approssimare una funzione f^* . Per esempio nel caso di un classificatore , $y = f^*(\mathbf{x})$ mappa un input \mathbf{x} ad una categoria y . Una rete feedforward definisce un mapping $\mathbf{y} = f(\mathbf{x}; \boldsymbol{\theta})$ ed impara il valore dei parametri $\boldsymbol{\theta}$ che permettono a f di meglio approssimare

f^* .

Queste reti sono chiamate **feedforward** perchè le informazioni fluiscono dall' input \mathbf{x} , attraverso le funzioni intermedie che compongono f , per arrivare all' output \mathbf{y} . Non ci sono connessione **feedback** nelle quali l'output del modello viene restituito al modello stesso.

Le feedforward neural networks vengono chiamate **networks** perchè tipicamente sono rappresentate dalla composizione funzioni diverse. Per esempio potremmo avere tre funzioni f^1, f^2, f^3 composte tra di loro per formare $f(\mathbf{x}) = f^3(f^2(f^1(\mathbf{x})))$. In questo caso, f^1 è chiamato il primo layer, f^2 il secondo layer e così via. La lunghezza della catena definisce la **depth** del modello, da questa terminologia deriva il termine "deep learning". L'ultimo layer della rete neurale viene detto **output layer**. Durante il training della rete neurale , l' obiettivo è quello di far assomigliare $f(\mathbf{x})$ a $f^*(\mathbf{x})$. Il data di training è composto da dei punti \mathbf{x} , ognuno associato ad una label $y \approx f^*(x)$. Il data di training specifica quale deve essere il comportamento dell' output layer per ogni punto \mathbf{x} , cioè produrre valori vicini a y . L' algoritmo di training deve decidere come variare i parametri degli altri layer in modo tale da ottenere la migliore approssimazione di f^* , per fare questo viene utilizzata una **loss function** che descrive quanto il valore in output della rete su input x sia "distanse" secondo una qualche metrica dalla label associata. Siccome l'output dei layer intermedi non mostrano l'output desiderato, questo layer vengono chiamati **hidden layers**.

In fine questo tipo di reti vengono definite *neurali* perchè sono ispirate alle neuroscienze. Ogni hidden layer della rete prende tipicamente in input un vettore di valori. La grandezza di questo vettore determina la **width** della rete. Ogni elemento di questo vettore può essere interpretato come un segnale per un neurone. Invece di pensare ad un hidden layer come una funzione che mappa un vettore in un altro, possiamo considerare ogni layer come composto da un insieme di **units** che operano in parallelo. Ogni unit ricorda un neurone, nel senso che riceve degli input da molte altre unit e computa la propria funzione di attivazione. Nonostante questa intuizione non biso-

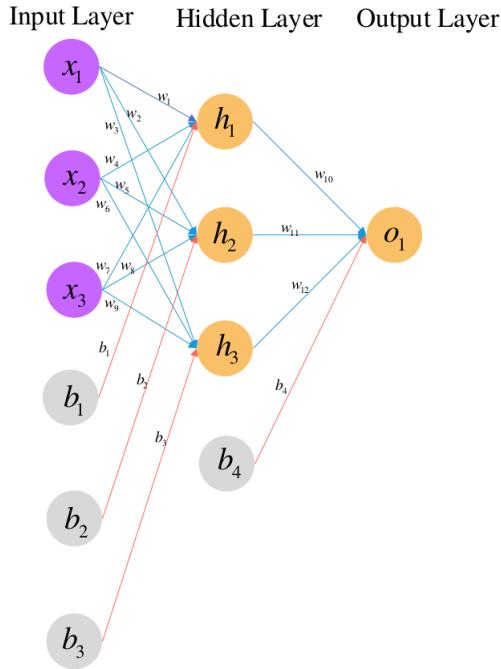


Figura 1.1: Rete neurale a due layer

gna pensare alle reti neurali come modellizzazioni del cervello, ma piuttosto come macchine che approssimano funzioni in modo tale da raggiungere una generalizzazione statistica del problema dato un insieme di esempi.

Una rete neurale può essere espressa nel seguente modo :

$$f(X; W; b) = \phi^{n-1}(W^{n-1}\phi^{n-2}(W^{n-2}\dots\phi^0(W^0X + b^0)\dots b^{n-2}) + b^{n-1}) \quad (1.2)$$

La rete neurale in figura 1.1 completamente connessa presenta due layer : un hidden layer e un layer di output (Il layer di input solitamente non viene considerato come layer). I parametri \$x_1, x_2, x_3\$ e \$o_1\$ rappresentano l'input e l'output della rete. \$w_1, w_2, \dots, w_{12}\$ sono i pesi e \$b_1, b_2, b_3, b_4\$ è il bias. Per esempio, l'output del neurone \$h_1\$ si ottiene come segue :

$$f(z_{h1}) = f(w_1x_1 + w_4x_2 + w_7x_3 + b_1) \quad (1.3)$$

$$o(h1) = f(z_{h1}) \quad (1.4)$$

Una volta calcolato l'output o_1 la rete usa un algoritmo di back propagation e una funzione di loss per imparare i parametri [back propagation citation]. Nell' ambito dell' elaborazioni delle immagini una tipologia di rete neurale molto utilizzata è la **Convolutional Neural Network** (CNN).

Le Convolutional Network sono delle reti neurali specializzate per processare del data con una topologia a griglia. Esempi includono serie temporali, che possono essere rappresentate come una griglia 1D prendendo osservazioni ad intervalli regolari, e immagini, rappresentabili come una griglia 2D di pixel. Il nome "convolutional network" indica che la rete utilizza una particolare operazione matematica lineare denominata **convoluzione**. Utilizziamo un esempio per spiegare l'operazione di convoluzione : supponiamo di voler tracciare la posizione di un astronave con un sensore laser. Il sensore restituisce un singolo output $x(t)$, la posizione dell' astronave al tempo t . sia x che t sono valori reali. Supponiamo che il sensore restituisca output rumoroso. Per ottenere una stima meno rumorosa della posizione della posizione dell' astronave vogliamo utilizzare una media pesata tra le misurazioni dando più importanza alle misurazioni recenti. Possiamo utilizzare una funzione di pesatura $w(\alpha)$, dove α è il tempo della misurazione. Applicando questa media pesata ad ogni istante, si ottiene una nuova funzione s che indica una stima regolarizzata della posizione.

$$s(t) = \int x(a)w(t-a)da \quad (1.5)$$

Questa operazione viene detta **convoluzione**. L' operazione viene tipicamente denotata con un asterisco.

$$s(t) = (x * w)(t)$$

Nel nostro esempio, w deve essere una funzione di densità valida, altrimenti l'output non sarebbe una media pesata. Ma questo non è necessario per altri esempi, in generale la convoluzione è definita per ogni coppia di funzioni dove l'integrale 1.5 è definito, e l'operazione può essere usata per altri scopi oltre alla media pesata. Nella terminologia delle convoluzioni il primo termine,

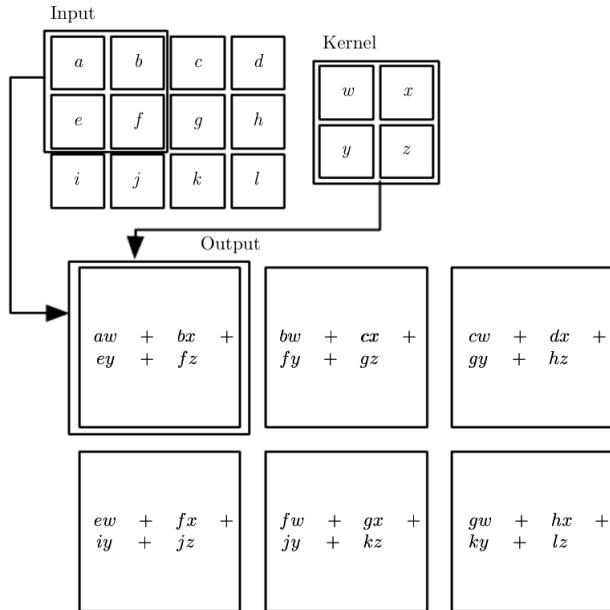


Figura 1.2: Un esempio di convoluzione 2D

in questo caso la funzione x prende il nome di **input**, il secondo argomento viene detto **kernel**. L'output viene indicato con il termine **feature map**. Nelle applicazioni di machine learning, l'input è solitamente discretizzato e rappresentato da un array multidimensionale, e il kernel è rappresentato da un array multidimensionale di parametri adattati dall'algoritmo di apprendimento. Nel caso di immagini vogliamo definire la convoluzione su più assi. Per esempio con un'immagine due-dimensionale I come input, vogliamo utilizzare un kernel due-dimensionale K :

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(m, n)K(i - m, j - n)$$

Le CNNs per come sono definite sono in grado di sfruttare le correlazioni spaziali dell'immagine [inserire citazione a deeplearning su proprietà cnns], preservando bordi e dettagli.

Un avanzamento importante alle CNNs è stato dato dalla funzione di attivazione ReLU [Rectified Linear Unit citation vedi deeplearning book], mi-

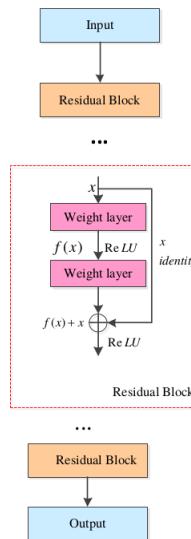


Figura 1.3: Un esempio di blocco residuo

giorando la velocità della discesa del gradiente stocastica (SGD) [citanzione a deeplearning], un metodo molto utilizzato per il training dei parametri del kernel. Sperimentalmente è stato osservato che se la rete è molto profonda, durante il training si possono presentare fenomeni di annullamento o esplosione del gradiente; e se la rete è molto ampia, si può presentare il fenomeno dell' overfitting [citanzione a deep learning su overfitting]. Per risolvere questi problemi nel 2016 è stato proposto un particolare tipo di rete, ResNet [citanzione a resnet]. ResNet introduce il concetto di *skip connection*: l'input di un layer è aggiunto all'output di un layer successivo. L' obiettivo di una rete neurale è quello di modellizzare una funzione target $h(\mathbf{x})$. Aggiungendo l' input \mathbf{x} all' output della rete, allora la rete è costretta ad imparare una funzione $f(\mathbf{x}) = h(\mathbf{x}) - \mathbf{x}$. Questa tecnica prende il nome di *residual learning*.

Figura 1.3.

Quando si inizializza una rete neurale i suoi pesi sono vicini a zero e la rete restituisce in output valori vicini a zero. Aggiungendo una skip connection, la rete restituirà in output una copia del suo input; modellando inizialmente la funzione identità. Se la funzione target è abbastanza vicina alla funzione

di identità (particolarmente vero per il problema del denoising), il residual learning velocizza l' apprendimento dei parametri in maniera considerevole.

Conclusioni

Queste sono le conclusioni.

In queste conclusioni voglio fare un riferimento alla bibliografia: questo è il mio riferimento [3, 4].

Appendice A

Prima Appendice

In questa Appendice non si è utilizzato il comando:

`\clearpage{\pagestyle{empty}\cleardoublepage}`, ed infatti l'ultima pagina 8 ha l'intestazione con il numero di pagina in alto.

Appendice B

Seconda Appendice

Bibliografia

- [1] Primo oggetto bibliografia.
- [2] Secondo oggetto bibliografia.
- [3] Terzo oggetto bibliografia.
- [4] Quarto oggetto bibliografia.

Ringraziamenti

Qui possiamo ringraziare il mondo intero!!!!!!!!!
Ovviamente solo se uno vuole, non è obbligatorio.