

UNIVERSITY OF MODENA AND REGGIO EMILIA

Department of Engineering

**Corso di laurea magistrale in
Ingegneria Meccatronica**

**Progettazione e validazione di un algoritmo di
denoising adattivo real time per segnali di
auscultazione polmonare**

Relatore:

Prof. Fabrizio Pancaldi

Candidato:

Simone Dolci

Anno Accademico 2024/2025

Sommario

L'auscultazione digitale rappresenta uno strumento fondamentale per la valutazione non invasiva delle condizioni respiratorie, ma risulta fortemente limitata dalla presenza di rumore ambientale e artefatti di acquisizione. In questo lavoro viene proposto un framework di denoising basato su filtro di Wiener, progettato per operare in modo adattivo su segnali respiratori non stazionari. Il metodo sfrutta l'identificazione automatica delle regioni di rumore all'interno del segnale, utilizzate per stimare dinamicamente lo spettro del disturbo e aggiornare il modello di rumore nel dominio tempo-frequenza. L'elaborazione è strutturata in modalità frame-wise, rendendo l'approccio compatibile con implementazioni in tempo reale su sistemi embedded. I risultati mostrano una significativa riduzione del rumore con preservazione delle caratteristiche spettrali del segnale respiratorio, evidenziando l'efficacia del metodo rispetto agli approcci tradizionali

Indice

Introduzione	1
0.1 Il ruolo dell'auscultazione nella diagnostica respiratoria	1
0.2 Evoluzione degli stetoscopi digitali	2
0.3 Problemi di rumore nell'auscultazione reale	2
0.4 Obiettivo della tesi	3
1 Analisi dei Suoni Respiratori e del Rumore	5
1.1 Caratteristiche acustiche dei suoni polmonari	5
1.2 Sorgenti di rumore nell'auscultazione digitale	6
1.3 Sovrapposizione spettrale tra segnale e rumore	7
1.4 Implicazioni per il denoising adattivo	7
2 Tecniche di denoising dei suoni respiratori	9
2.1 Limiti dei filtri lineari	9
2.2 Approcci tempo-frequenza	10
2.2.1 Trasformata wavelet	10
2.2.2 Empirical Mode Decomposition (EMD)	11
2.3 Metodi di stima del rumore	11
2.3.1 Spectral subtraction	12
2.3.2 Modelli statistici del rumore	12
2.4 Approcci basati sul deep learning	13
2.5 Motivazioni per l'utilizzo del filtro di Wiener	14
3 Sistema di acquisizione dei segnali respiratori	15
3.1 Stetoscopio digitale	15
3.2 Architettura del sistema di acquisizione	16
3.3 Catena di acquisizione del segnale	17
3.3.1 Sensore acustico e microfono	17
3.3.2 Front-end analogico	18
3.3.3 Conversione analogico-digitale	20
3.3.4 Salvataggio dei dati e parsing python	21

4	Analisi sperimentale del rumore	23
4.1	Misure in ambiente acusticamente controllato	23
4.2	Caratterizzazione spettrale del rumore	23
4.3	Verifica delle possibili sorgenti di rumore	25
4.3.1	Verifica della catena analogica	25
4.3.2	Verifica della pipeline di salvataggio e parsing	26
4.3.3	Analisi della digitalizzazione ADC	26
4.4	Riduzione del rumore mediante oversampling	27
4.5	Esempio di rumore ambientale reale	29
4.5.1	Esempio di auscultazione reale	30
4.5.2	Comportamento spettrale durante le pause respiratorie	31
4.5.3	Implicazioni per il denoising adattivo	31
5	Framework proposto per il denoising adattivo dei suoni respiratori	33
5.1	Panoramica del framework	33
5.2	Modello del segnale	34
5.3	Analisi spettrale frame-wise	35
5.4	Discriminante spettrale	35
5.5	Identificazione delle regioni di rumore	37
5.6	Stima adattiva della baseline di rumore	38
5.7	Filtraggio mediante Wiener	38
5.8	Esempio di comportamento del filtro	41
5.9	Ricostruzione del segnale	42
6	Creazione del dataset sperimentale e metodologia di valutazione	43
6.1	Creazione del dataset	43
6.2	Configurazione dell'analisi frame-based	45
6.3	Filosofia di valutazione	46
6.4	Metriche quantitative	46
6.4.1	Noise Suppression Index (NSI)	47
6.4.2	Respiratory Attenuation Index (RAI)	47
6.4.3	Respiratory–Background Contrast Improvement (ΔC)	48
6.5	Aggregazione e analisi statistica	49
7	Risultati sperimentali	51
7.1	Esempio qualitativo di denoising in auscultazione reale	51
7.2	Interpretazione del comportamento osservato	52

7.3	Setup sperimentale	52
7.4	Attenuazione del rumore di fondo (NSI)	52
7.4.1	Comportamento nella banda respiratoria	52
7.5	Coerenza tra baseline e componente rimossa	54
7.6	Comportamento del guadagno di Wiener	55
7.7	Discussione dei risultati	56
8	Conclusioni	59
8.1	Sviluppi futuri	60
	Bibliografia	62

Elenco delle figure

Figura 1	Prototipo stetoscopio digitale	15
Figura 2	Assieme smontato stetoscopio digitale	16
Figura 3	Microfono MEMS	18
Figura 4	Schema del front-end analogico dello stetoscopio digitale.	19
Figura 5	Uscita analogica jack per il monitoraggio del segnale tramite cuffie.	20
Figura 6	Spettrogramma del segnale acquisito in camera anecoica prima dell'introduzione dell'oversampling ADC.	24
Figura 7	Densità spettrale di potenza del segnale acquisito in camera anecoica prima dell'introduzione dell'oversampling ADC.	24
Figura 8	Andamento temporale dei conteggi ADC acquisiti in camera anecoica prima dell'introduzione dell'oversampling.	25
Figura 9	Spettrogramma del segnale acquisito in camera anecoica dopo l'introduzione dell'oversampling ADC.	27
Figura 10	Andamento dei conteggi ADC acquisiti in camera anecoica dopo l'introduzione dell'oversampling.	28
Figura 11	Densità spettrale di potenza del segnale acquisito in camera anecoica dopo l'introduzione dell'oversampling ADC.	28
Figura 12	Spettrogramma del segnale acquisito in una stanza silenziosa con rumori ambientali deboli provenienti da ambienti adiacenti.	29
Figura 13	Spettrogramma di una registrazione auscultatoria reale. Le variazioni dell'energia spettrale evidenziano le fasi di attività respiratoria alternate a pause respiratorie.	30
Figura 14	Framework proposto	34
Figura 15	Andamento del discriminante $r(t)$ nelle diverse fasi respiratorie.	36
Figura 16	Definizione automatica delle regioni dominate dal rumore	37
Figura 17	Spettrogramma pre e post applicazione del guadagno di wiener	40
Figura 18	Esempio di applicazione del filtro di Wiener su un frame spettrale.	41
Figura 19	Siti di Auscultazione.	44

Figura 20	Esempio di denoising in auscultazione reale con il framework proposto . .	51
Figura 21	Confronto tra RAI e ΔC calcolati sull'intero dataset.	53
Figura 22	Distribuzione spettrale mediana dei frame RESP e NOISE (con intervallo interquartile).	54
Figura 23	Guadagno di Wiener in funzione della frequenza per frame NOISE e RESP.	55

Elenco delle tabelle

Tabella 1	Parametri principali del microfono SPM0687LR5H-1	18
Tabella 2	Dataset	45
Tabella 3	Prestazioni di attenuazione del rumore di fondo (NSI) sull'intero dataset.	52
Tabella 4	Similarità tra componente rimossa e baseline di rumore stimata.	55

Introduzione

0.1 Il ruolo dell'auscultazione nella diagnostica respiratoria

L'auscultazione rappresenta da oltre due secoli uno degli strumenti fondamentali per la valutazione clinica dell'apparato respiratorio. Attraverso l'ascolto dei suoni generati dal flusso d'aria nelle vie respiratorie e dalle vibrazioni dei tessuti toracici, il medico può ottenere informazioni utili per l'identificazione di numerose condizioni patologiche. Nonostante i significativi progressi delle tecniche di imaging medico, come la tomografia computerizzata e l'ecografia polmonare, l'auscultazione continua a costituire una procedura diagnostica di primo livello grazie alla sua rapidità, non invasività e al basso costo operativo [1].

I suoni respiratori derivano principalmente dall'interazione tra il flusso d'aria e le strutture anatomiche delle vie aeree. In condizioni fisiologiche essi presentano una distribuzione spettrale relativamente ampia, con componenti energetiche tipicamente comprese tra circa 100 Hz e 800–1000 Hz [1]. In presenza di alterazioni patologiche possono emergere fenomeni acustici caratteristici, tra cui i *wheezes*, segnali quasi tonali associati a restringimenti bronchiali, e i *crackles*, eventi impulsivi brevi legati alla riapertura improvvisa delle vie aeree periferiche [1], [2].

L'interpretazione dei suoni respiratori rappresenta tuttavia un processo complesso e fortemente dipendente dall'esperienza dell'operatore. Diversi studi hanno evidenziato una significativa variabilità inter-operatore nella valutazione auscultatoria, in particolare quando i segnali sono deboli o contaminati da disturbi ambientali [3]. Questa limitazione ha motivato lo sviluppo di approcci quantitativi basati sull'analisi digitale dei segnali respiratori, con l'obiettivo di supportare l'interpretazione clinica e migliorare la riproducibilità delle valutazioni auscultatorie [3].

0.2 Evoluzione degli stetoscopi digitali

Lo sviluppo delle tecnologie elettroniche e dei sistemi di acquisizione digitale ha portato negli ultimi decenni alla diffusione degli stetoscopi digitali, dispositivi progettati per convertire le vibrazioni acustiche generate nel torace in segnali elettrici successivamente elaborabili mediante tecniche di elaborazione digitale del segnale. A differenza degli stetoscopi acustici tradizionali, nei quali il segnale viene trasmesso direttamente all'orecchio dell'operatore attraverso un sistema meccanico, gli stetoscopi digitali introducono una catena di acquisizione composta da un sensore acustico, uno stadio di amplificazione analogica, un convertitore analogico-digitale e un sistema di elaborazione del segnale.

Questa architettura consente di ottenere una rappresentazione digitale dei suoni respiratori, rendendo possibile l'applicazione di strumenti avanzati di analisi computazionale [3]. In particolare, è possibile applicare tecniche di analisi spettrale, rappresentazioni tempo-frequenza e algoritmi di classificazione automatica finalizzati all'identificazione di pattern acustici associati a specifiche condizioni patologiche [2], [3].

Inoltre, la possibilità di memorizzare e trasmettere le registrazioni auscultatorie ha favorito lo sviluppo di sistemi di telemedicina e la creazione di dataset utilizzabili per la ricerca e lo sviluppo di algoritmi di analisi automatica dei suoni respiratori [4]. Tuttavia, la disponibilità di segnali digitali introduce anche nuove sfide legate alla qualità delle registrazioni e alla gestione delle diverse sorgenti di disturbo presenti durante l'acquisizione.

0.3 Problemi di rumore nell'auscultazione reale

Nonostante i vantaggi offerti dalla digitalizzazione dei segnali auscultatori, l'acquisizione dei suoni respiratori in condizioni reali risulta frequentemente compromessa dalla presenza di contributi indesiderati. Il segnale registrato include infatti componenti provenienti dall'ambiente circostante, dall'elettronica interna e dal contatto meccanico tra lo stetoscopio e il torace del paziente [2], [3].

Tali componenti presentano spesso una distribuzione spettrale parzialmente sovrapponibile a quella dei suoni respiratori, rendendo inefficaci strategie di filtraggio basate esclusivamente sulla separazione in banda.

Un ulteriore elemento di complessità è rappresentato dalla natura non stazionaria delle condi-

zioni di acquisizione, che comporta variazioni temporali del rapporto segnale-rumore e rende poco realistica l'ipotesi di rumore stazionario su lunghi intervalli [5], [6].

Per queste ragioni, è necessario ricorrere a tecniche di miglioramento del segnale in grado di adattarsi dinamicamente alle condizioni di acquisizione.

In questo contesto si inserisce il lavoro presentato in questa tesi, che propone un framework di denoising progettato per stimare in modo adattivo il rumore nelle registrazioni auscultatorie e migliorare la qualità dei segnali respiratori acquisiti mediante stetoscopi digitali.

0.4 Obiettivo della tesi

Alla luce delle criticità descritte, l'obiettivo principale di questa tesi è lo sviluppo e la validazione di un framework di denoising adattivo per segnali di auscultazione polmonare acquisiti in condizioni reali.

Il metodo proposto sfrutta la struttura temporale del ciclo respiratorio per identificare automaticamente intervalli dominati dal rumore di fondo, utilizzati per stimarne le caratteristiche spettrali e aggiornare dinamicamente un modello di rumore.

Particolare attenzione è dedicata alla compatibilità con implementazioni embedded: l'algoritmo è progettato secondo una struttura causale e frame-wise, nella quale le operazioni di stima del rumore e filtraggio utilizzano esclusivamente l'informazione disponibile nei frame precedenti.

La tesi include inoltre lo sviluppo di una pipeline sperimentale per l'acquisizione e l'analisi dei segnali respiratori, basata su uno stetoscopio digitale prototipale e strumenti di elaborazione sviluppati in ambiente Python.

I principali contributi di questo lavoro possono essere riassunti nei seguenti punti:

- analisi delle caratteristiche di rumore presenti nelle registrazioni di suoni respiratori acquisite mediante stetoscopio digitale;
- progettazione di un metodo di denoising adattivo basato sull'identificazione automatica di regioni temporali dominate dal rumore;
- integrazione di un modello spettrale del rumore all'interno di un framework di filtraggio nel dominio tempo-frequenza;
- sviluppo di una pipeline sperimentale per l'acquisizione e l'elaborazione dei segnali respi-

ratori;

- valutazione sperimentale del metodo proposto su registrazioni auscultatorie acquisite in condizioni realistiche.

Attraverso queste attività, il lavoro contribuisce allo sviluppo di tecniche di elaborazione del segnale più robuste e interpretabili per l'analisi dei suoni respiratori, con particolare attenzione alle applicazioni su dispositivi embedded.

1. Analisi dei Suoni Respiratori e del Rumore

1.1 Caratteristiche acustiche dei suoni polmonari

I suoni respiratori sono generati dall'interazione tra il flusso d'aria all'interno delle vie respiratorie e le strutture anatomiche del sistema polmonare. Durante il processo di ventilazione, il flusso d'aria attraverso trachea, bronchi e bronchioli produce vibrazioni turbolente che si propagano attraverso i tessuti polmonari e la parete toracica fino alla superficie del torace, dove possono essere rilevate mediante stetoscopi [1]. Tali vibrazioni acustiche contengono informazioni rilevanti sullo stato fisiologico delle vie aeree e rappresentano una fonte importante di segnali diagnostici.

Dal punto di vista acustico, i suoni respiratori possono essere suddivisi in due categorie principali: suoni respiratori normali e suoni respiratori avventizi. I suoni normali sono generati principalmente dal flusso d'aria turbolento nelle grandi vie aeree e presentano una distribuzione spettrale relativamente ampia e continua, risultando generalmente più intensi durante la fase inspiratoria [1].

L'energia dei suoni respiratori fisiologici è concentrata prevalentemente nella banda compresa tra circa 100 Hz e 800 Hz, con una attenuazione delle componenti a frequenza più elevata dovuta alle proprietà di filtraggio dei tessuti polmonari e della parete toracica [7]. La distribuzione spettrale può tuttavia variare in funzione della posizione di auscultazione, delle caratteristiche anatomiche del paziente e delle condizioni respiratorie.

Accanto ai suoni normali, l'auscultazione può evidenziare la presenza di suoni avventizi, associati a condizioni patologiche o alterazioni del funzionamento delle vie aeree. Tra i fenomeni più rilevanti rientrano i wheezes e i crackles.

I wheezes sono segnali quasi periodici caratterizzati da componenti tonali relativamente stabili nel tempo, generalmente associati a restringimenti o ostruzioni delle vie aeree. Essi presentano tipicamente frequenze comprese tra circa 100 Hz e 1000 Hz e durate relativamente lunghe [3].

I crackles sono invece eventi impulsivi di breve durata, caratterizzati da transitori acustici con rapida crescita e decadimento dell'ampiezza. Essi sono associati alla riapertura di piccole vie

aeree o alla presenza di secrezioni e presentano una distribuzione spettrale ampia [1].

Un aspetto rilevante è la variabilità temporale del segnale respiratorio. Il segnale non è infatti stazionario, ma evolve nel tempo in funzione delle diverse fasi del ciclo respiratorio. Durante inspirazione ed espirazione il contributo acustico è generalmente dominante, mentre negli intervalli tra le due fasi il livello energetico si riduce significativamente.

Dal punto di vista dell'elaborazione del segnale, i suoni respiratori sono spesso analizzati mediante rappresentazioni tempo-frequenza, come la Short-Time Fourier Transform (STFT), che consentono di osservare l'evoluzione spettrale nel tempo e di evidenziare eventi transitori o componenti quasi periodiche [2].

La conoscenza delle caratteristiche spettrali e temporali dei suoni respiratori costituisce il riferimento necessario per analizzare l'effetto del rumore e valutare l'impatto delle tecniche di elaborazione del segnale sulle componenti di interesse.

1.2 Sorgenti di rumore nell'auscultazione digitale

L'acquisizione dei suoni respiratori mediante stetoscopi digitali avviene spesso in ambienti caratterizzati da condizioni acustiche non controllate. In tali contesti, il segnale registrato risulta contaminato da diverse sorgenti di disturbo che possono degradare la qualità dell'auscultazione [1], [2].

Le principali sorgenti di rumore possono essere suddivise in tre categorie.

La prima è costituita dal rumore ambientale, che include conversazioni, sistemi di ventilazione e altre sorgenti sonore presenti nell'ambiente. Tali contributi presentano spesso una distribuzione spettrale ampia, che può sovrapporsi alla banda dei suoni respiratori [3].

Una seconda categoria è rappresentata dal rumore elettronico del sistema di acquisizione. Il microfono, lo stadio di amplificazione analogica e il convertitore analogico-digitale possono introdurre componenti di disturbo legate a fenomeni quali rumore termico o errori di quantizzazione [2].

Infine, gli artefatti meccanici derivano dall'interazione tra il sensore e la superficie toracica. Movimenti del dispositivo o variazioni nella pressione di contatto possono generare vibrazioni che si traducono in transitori impulsivi o componenti broadband.

Nel complesso, il segnale auscultatorio può essere modellato come la somma del segnale respi-

ratorio di interesse e di un contributo di rumore composto da componenti eterogenee, la cui presenza rappresenta una delle principali difficoltà nell'elaborazione dei segnali auscultatori.

1.3 Sovrapposizione spettrale tra segnale e rumore

Un aspetto critico nell'elaborazione dei segnali auscultatori è rappresentato dalla sovrapposizione spettrale tra il segnale respiratorio e le componenti di rumore. A differenza di altri problemi di elaborazione del segnale, nei quali il disturbo può essere separato mediante filtri passa-banda, nel caso dei suoni respiratori questa condizione raramente si verifica [1], [7].

Molte sorgenti di rumore occupano infatti la stessa regione spettrale del segnale respiratorio, rendendo inefficaci strategie di filtraggio basate esclusivamente sulla separazione in frequenza.

Questa difficoltà è ulteriormente accentuata dalla variabilità temporale del segnale respiratorio, che comporta una evoluzione dinamica del rapporto segnale-rumore durante la registrazione.

In questo contesto, rappresentazioni tempo-frequenza come la STFT risultano strumenti fondamentali per analizzare il comportamento locale del segnale e del rumore [8].

1.4 Implicazioni per il denoising adattivo

Le caratteristiche descritte evidenziano che il denoising dei suoni respiratori richiede approcci in grado di adattarsi alle condizioni locali del segnale.

La sovrapposizione spettrale tra segnale e rumore e la variabilità temporale del ciclo respiratorio rendono infatti inefficaci strategie di filtraggio statiche basate unicamente sulla separazione in frequenza.

Queste considerazioni motivano l'impiego di tecniche di denoising nel dominio tempo-frequenza e basate su stima adattiva del rumore, che saranno analizzate nel capitolo successivo.

2. Tecniche di denoising dei suoni respiratori

L'elaborazione dei suoni respiratori acquisiti mediante stetoscopi digitali richiede tecniche di miglioramento del segnale in grado di operare in presenza di rumore e artefatti di acquisizione. Nel corso degli anni sono stati proposti numerosi approcci per affrontare questo problema, dai filtri lineari classici a metodi più avanzati basati su rappresentazioni tempo-frequenza, modelli statistici del rumore e tecniche di apprendimento automatico.

Nel caso dei segnali respiratori, la scelta del metodo di denoising dipende non solo dalle sue prestazioni teoriche, ma anche dalla robustezza rispetto a condizioni di acquisizione variabili e dai vincoli computazionali del sistema. In questo capitolo vengono analizzate le principali famiglie di metodi utilizzate in letteratura, con particolare attenzione ai loro punti di forza e ai limiti applicativi, al fine di motivare la scelta dell'approccio adottato in questa tesi.

2.1 Limiti dei filtri lineari

I metodi di filtraggio lineare rappresentano uno degli approcci più semplici e storicamente utilizzati per la riduzione del rumore nei segnali biomedicali. Tali tecniche si basano sull'applicazione di filtri progettati per attenuare specifiche bande di frequenza, assumendo implicitamente che il rumore occupi regioni spettrali distinte rispetto al segnale di interesse.

Nel contesto dei suoni respiratori, questi approcci vengono tipicamente impiegati sotto forma di filtri passa-banda, con l'obiettivo di isolare la banda caratteristica del segnale utile e attenuare componenti indesiderate a frequenze molto basse o molto elevate. Tali metodi risultano efficaci come fase preliminare di preprocessing, in particolare per la rimozione di disturbi semplici e ben separati spettralmente.

Tuttavia, nelle condizioni reali di auscultazione, il rumore presenta spesso una distribuzione spettrale sovrapposta a quella dei suoni respiratori. In presenza di tale sovrapposizione, il filtraggio basato esclusivamente sulla frequenza non consente di distinguere tra segnale utile e disturbo senza introdurre una perdita di informazione.

Un ulteriore limite è rappresentato dalla natura statica di questi approcci: la risposta del filtro

rimane costante nel tempo e non si adatta alle variazioni locali del rapporto segnale-rumore. Di conseguenza, un filtro a parametri fissi può risultare efficace solo in alcune porzioni del segnale, ma non in altre.

Per queste ragioni, i filtri lineari risultano insufficienti per affrontare in modo robusto il problema del denoising dei suoni respiratori in scenari realistici.

2.2 Approcci tempo-frequenza

Le limitazioni dei filtri lineari hanno portato allo sviluppo di tecniche di denoising più avanzate basate su rappresentazioni tempo-frequenza. A differenza dei metodi di filtraggio statici nel dominio della frequenza, questi approcci consentono di analizzare e modificare il contenuto spettrale del segnale in modo localizzato nel tempo.

Nel caso dei suoni respiratori, questa caratteristica è particolarmente rilevante, poiché la distribuzione energetica del segnale varia nel corso del ciclo respiratorio. Le rappresentazioni tempo-frequenza permettono quindi di descrivere il comportamento locale del segnale e di intervenire selettivamente sulle componenti dominate dal rumore.

Nel contesto dell'elaborazione dei segnali respiratori, diversi lavori hanno proposto l'utilizzo di trasformate multirisoluzione o decomposizioni adattive del segnale. Tra gli approcci più diffusi rientrano la trasformata wavelet e la *Empirical Mode Decomposition* (EMD), che consentono di analizzare il segnale su diverse scale temporali e frequenziali.

2.2.1 Trasformata wavelet

La trasformata wavelet rappresenta una tecnica di analisi multirisoluzione che consente di decomporre un segnale in componenti localizzate nel tempo e nella frequenza. A differenza della trasformata di Fourier, che utilizza funzioni sinusoidali di durata infinita, le wavelet impiegano funzioni base localizzate, risultando particolarmente adatte allo studio di segnali caratterizzati da eventi transitori o variazioni locali [8].

Nel contesto del denoising, i metodi basati su wavelet sfruttano la possibilità di separare il segnale in coefficienti associati a diverse scale di frequenza. Applicando tecniche di soglia (*thresholding*) ai coefficienti wavelet, è possibile attenuare le componenti di rumore preservando al contempo le strutture principali del segnale respiratorio.

Nonostante i risultati promettenti riportati in letteratura, l'efficacia di questi metodi dipende in modo significativo dalla scelta della famiglia di wavelet, dal numero di livelli di decomposizione e dai parametri di soglia adottati. Inoltre, la loro implementazione può risultare relativamente onerosa dal punto di vista computazionale, rendendone meno immediata l'integrazione in sistemi embedded a bassa potenza.

2.2.2 Empirical Mode Decomposition (EMD)

Un'altra tecnica frequentemente utilizzata nell'elaborazione di segnali non stazionari è la *Empirical Mode Decomposition* (EMD), proposta da Huang et al. [9]. Questo metodo consente di decomporre un segnale in una serie di componenti elementari denominate *Intrinsic Mode Functions* (IMF), ciascuna caratterizzata da una specifica banda di frequenza.

A differenza delle tecniche basate su trasformate predefinite, la decomposizione EMD è completamente adattiva e dipende direttamente dalle caratteristiche del segnale analizzato. Questo rende il metodo particolarmente interessante per segnali complessi e non stazionari.

Nel contesto del denoising, l'approccio tipico consiste nell'analizzare le componenti IMF ottenute dalla decomposizione e rimuovere o attenuare quelle associate principalmente al rumore. In molti casi, le componenti ad alta frequenza risultano maggiormente contaminate dal disturbo e possono essere modificate prima della ricostruzione del segnale.

Anche in questo caso, tuttavia, emergono criticità rilevanti. La decomposizione può essere sensibile al rumore e produrre fenomeni di *mode mixing*, in cui componenti di frequenza differenti vengono rappresentate all'interno della stessa IMF. Inoltre, il processo iterativo richiesto per ottenere la decomposizione può risultare computazionalmente costoso, limitandone l'uso in applicazioni in tempo reale o su piattaforme embedded.

2.3 Metodi di stima del rumore

Nei sistemi di miglioramento del segnale contaminato da rumore, una componente fondamentale dell'algoritmo è rappresentata dalla stima delle caratteristiche statistiche del disturbo presente nella registrazione. In particolare, molti metodi di denoising operano nel dominio tempo-frequenza e richiedono una stima dello spettro di potenza del rumore per poter attenuare selettivamente le componenti indesiderate del segnale [5], [6].

Nel caso dei segnali auscultatori, la stima del rumore risulta particolarmente complessa a causa della sovrapposizione spettrale tra il segnale respiratorio e molte delle sorgenti di disturbo presenti durante l’acquisizione, oltre che della possibile variabilità temporale del rumore stesso. Per questo motivo, numerosi metodi di denoising si basano su strategie di stima adattiva del rumore, che permettono di aggiornare dinamicamente il modello del disturbo durante l’analisi del segnale.

Tra gli approcci più diffusi si trovano i metodi basati sulla sottrazione spettrale (*spectral subtraction*) e quelli basati su modelli statistici del segnale e del rumore.

2.3.1 Spectral subtraction

La tecnica di *spectral subtraction* rappresenta uno dei primi metodi proposti per la riduzione del rumore nei segnali audio. Introdotto da Boll nel 1979 [5], questo approccio si basa sull’idea di stimare lo spettro di potenza del rumore e sottrarlo dallo spettro del segnale osservato nel dominio tempo-frequenza.

In questo metodo, il segnale viene inizialmente trasformato mediante STFT. Successivamente, lo spettro del rumore viene stimato utilizzando segmenti del segnale in cui il contributo del segnale utile è minimo. Lo spettro stimato viene quindi sottratto da quello del segnale osservato, producendo una attenuazione delle componenti associate al disturbo.

La *spectral subtraction* è concettualmente semplice e relativamente facile da implementare, ma presenta limiti significativi. In particolare, la qualità del risultato dipende fortemente dall’accuratezza della stima del rumore. Errori nella stima possono introdurre artefatti nel segnale ricostruito, spesso percepiti come rumore residuo di tipo “musicale”. Inoltre, in presenza di rumore non stazionario o di forte sovrapposizione spettrale tra segnale e disturbo, le prestazioni del metodo possono risultare limitate [10].

2.3.2 Modelli statistici del rumore

Un’altra famiglia di tecniche di denoising è basata su modelli statistici del segnale e del rumore. In questi approcci, il problema del miglioramento del segnale viene formulato come un problema di stima, in cui l’obiettivo è ottenere la miglior stima possibile del segnale pulito a partire dal segnale osservato.

Tra i metodi più noti in questa categoria si trovano gli algoritmi basati sulla minimizzazione dell'errore quadratico medio tra il segnale stimato e quello originale. In questo contesto, il filtraggio di Wiener rappresenta una delle soluzioni più diffuse e teoricamente fondate, poiché consente di ottenere una stima ottimale del segnale in presenza di rumore additivo sotto specifiche ipotesi statistiche [6], [11].

Questi approcci risultano particolarmente efficaci quando è disponibile una buona stima dello spettro di potenza del rumore. Il loro principale limite non risiede quindi nel formalismo del filtro, ma nella qualità della stima del disturbo su cui esso si basa.

2.4 Approcci basati sul deep learning

Negli ultimi anni, le tecniche di *deep learning* hanno ottenuto risultati significativi nel campo del miglioramento dei segnali audio e della soppressione del rumore. In particolare, modelli basati su reti neurali profonde sono stati ampiamente utilizzati nel contesto dello *speech enhancement*, dove l'obiettivo è stimare una versione pulita del segnale a partire da un'osservazione rumorosa [12], [13].

Molti dei modelli più efficaci proposti in letteratura utilizzano architetture profonde basate su reti convoluzionali, reti ricorrenti o modelli generativi. Queste soluzioni possono raggiungere prestazioni elevate nel miglioramento del segnale, ma richiedono generalmente risorse computazionali significative sia nella fase di addestramento sia durante l'inferenza.

Per applicazioni embedded sono stati quindi proposti modelli più leggeri, tra cui diverse varianti di *autoencoder*. In particolare, i *denoising autoencoder* (DAE) apprendono una rappresentazione del segnale pulito a partire da versioni contaminate dal rumore [14].

Nonostante la maggiore compattezza rispetto ad altre architetture, l'utilizzo di questi modelli presenta comunque alcune criticità nel contesto dei segnali respiratori. In particolare, l'addestramento efficace richiede la disponibilità di dataset contenenti registrazioni rumorose e corrispondenti segnali di riferimento puliti. Nel caso dei suoni respiratori acquisiti in ambiente reale, ottenere tali riferimenti può risultare difficile.

Per queste ragioni, nonostante il crescente interesse verso le tecniche di deep learning, in molte applicazioni embedded risulta ancora preferibile l'utilizzo di approcci basati su modelli statistici del segnale e del rumore, che consentono di ottenere un buon compromesso tra efficacia del denoising e complessità computazionale.

2.5 Motivazioni per l'utilizzo del filtro di Wiener

Come discusso nelle sezioni precedenti, ciascuna famiglia di metodi presenta limitazioni specifiche quando viene considerata nel contesto dei segnali respiratori acquisiti in condizioni reali e su dispositivi portatili.

I filtri lineari risultano efficaci soltanto quando il rumore occupa bande nettamente separate rispetto al segnale utile, condizione che raramente si verifica nelle registrazioni auscultatorie. Gli approcci basati su trasformate wavelet o decomposizioni adattive offrono una maggiore flessibilità, ma introducono dipendenze significative dalla scelta dei parametri e possono risultare onerosi dal punto di vista computazionale. I metodi basati su deep learning richiedono invece dataset di addestramento ampi e ben annotati, oltre a risorse di calcolo spesso non compatibili con piattaforme embedded a bassa potenza.

Alla luce di queste considerazioni, gli approcci basati su modelli statistici del segnale e del rumore rappresentano un compromesso particolarmente interessante tra efficacia del denoising, interpretabilità del metodo e complessità computazionale.

In questo contesto, il filtro di Wiener costituisce una soluzione rilevante, poiché consente di effettuare una stima ottimale del segnale in presenza di rumore additivo sulla base della minimizzazione dell'errore quadratico medio [11]. Inoltre, la sua formulazione nel dominio tempo-frequenza consente di adattare localmente il guadagno applicato al segnale in funzione della stima del rumore.

Per queste ragioni, il lavoro presentato in questa tesi adotta un approccio basato sul filtro di Wiener, integrato con una strategia di stima adattiva del rumore. Nel capitolo successivo viene descritto nel dettaglio il framework proposto per il miglioramento dei segnali auscultatori.

3. Sistema di acquisizione dei segnali respiratori

3.1 Stetoscopio digitale

Nel lavoro presentato in questa tesi, è stato utilizzato il prototipo di uno stetoscopio digitale come piattaforma di acquisizione per la registrazione dei segnali respiratori impiegati nelle successive fasi di sviluppo e valutazione dell'algoritmo di denoising. Le sezioni successive descrivono l'architettura del sistema di acquisizione utilizzato, la catena di acquisizione del segnale e le procedure adottate per la registrazione e l'organizzazione dei dati.



Figura 1: Prototipo stetoscopio digitale

3.2 Architettura del sistema di acquisizione

Il sistema di acquisizione adottato in questo lavoro segue una struttura tipica dei dispositivi di acquisizione audio biomedicali, composta da quattro blocchi principali: sensore acustico, front-end analogico, conversione analogico-digitale ed elaborazione digitale.

La rilevazione dei suoni respiratori avviene mediante un microfono MEMS analogico integrato nella cavità acustica dello stetoscopio digitale. Il segnale analogico generato dal sensore viene condizionato attraverso uno stadio di amplificazione differenziale e successivamente digitalizzato mediante il convertitore analogico-digitale (ADC) integrato nel microcontrollore STM32 utilizzato come piattaforma di acquisizione.

I campioni digitalizzati vengono memorizzati nella memoria flash del dispositivo in formato grezzo e successivamente esportati per l'elaborazione offline mediante strumenti software dedicati. Questa architettura consente di separare la fase di acquisizione del segnale dalla successiva fase di analisi e sviluppo degli algoritmi di elaborazione.

Nelle sezioni successive verrà descritta in modo più dettagliato la catena di acquisizione del segnale, analizzando il percorso seguito dal segnale acustico dalla fase di trasduzione fino alla registrazione dei dati digitali.



Figura 2: Assieme smontato stetoscopio digitale

3.3 Catena di acquisizione del segnale

La catena di acquisizione del segnale ha il compito di convertire le vibrazioni acustiche generate dai suoni respiratori in una rappresentazione digitale utilizzabile per l'elaborazione numerica del segnale. Questo processo avviene attraverso tre stadi principali: la rilevazione acustica mediante microfono, il condizionamento analogico del segnale e la conversione analogico-digitale. Le sezioni seguenti descrivono nel dettaglio i principali elementi della catena di acquisizione

3.3.1 Sensore acustico e microfono

La rilevazione dei suoni respiratori avviene mediante un microfono MEMS analogico differenziale SPM0687LR5H-1, basato su tecnologia SiSonic. Il sensore MEMS converte le variazioni di pressione acustica generate dalle vibrazioni toraciche in un segnale elettrico analogico proporzionale.

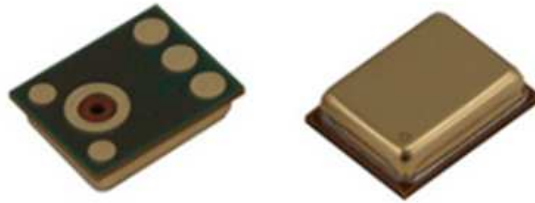
I microfoni MEMS integrano all'interno dello stesso package sia la struttura micromeccanica di rilevazione sia l'elettronica di condizionamento del segnale, consentendo dimensioni ridotte, elevata stabilità e buona immunità ai disturbi ambientali. Il dispositivo utilizzato presenta un elevato rapporto segnale-rumore e un'ampia dinamica, caratteristiche particolarmente rilevanti nelle applicazioni auscultatorie, dove i segnali respiratori possono avere ampiezza ridotta e contenuto spettrale distribuito.

Il microfono è utilizzato in configurazione differenziale (OUT+, OUT-), che consente una maggiore reiezione dei disturbi di modo comune e una migliore robustezza rispetto al rumore elettrico e alle interferenze ambientali.

Il sensore è montato all'interno della cavità acustica dello stetoscopio digitale, dove rileva le vibrazioni sonore generate dal flusso d'aria nelle vie respiratorie e trasmesse attraverso la parete toracica. Il segnale analogico differenziale prodotto dal sensore costituisce quindi il primo stadio della catena di acquisizione del segnale respiratorio.

Tabella 1: Parametri principali del microfono SPM0687LR5H-1

Parametro	Valore
Sensibilità	-34.7 dBV/Pa
SNR (A-weighted)	69.5 dB
AOP	130 dB SPL
Banda utile	13 Hz – 13 kHz
Configurazione uscita	Differenziale

**Figura 3:** Microfono MEMS

3.3.2 Front-end analogico

Il segnale analogico generato dal microfono presenta un'ampiezza relativamente ridotta e non è direttamente compatibile con il range dinamico del convertitore analogico-digitale. Per questo motivo viene sottoposto a una fase di condizionamento analogico mediante un opportuno front-end.

Nel sistema utilizzato, il segnale differenziale proveniente dal microfono MEMS viene amplificato tramite uno stadio di amplificazione differenziale. L'analisi del circuito, condotta sia in condizioni di polarizzazione continua sia mediante modello a piccoli segnali, consente di ricavare un guadagno complessivo pari a

$$G \approx 24$$

tra la tensione di ingresso V_{in} e la tensione di uscita V_{out} . Questo stadio ha lo scopo di adattare l'ampiezza del segnale al range operativo dell'ADC e, allo stesso tempo, migliorare la reiezione dei disturbi di modo comune eventualmente presenti sull'alimentazione o introdotti

dall'ambiente circostante.

Il segnale analogico amplificato viene quindi distribuito verso due percorsi distinti: da un lato viene inviato all'ingresso del convertitore analogico-digitale del microcontrollore, dall'altro viene reso disponibile su un'uscita analogica jack per cuffie. Questa uscita consente di monitorare direttamente il segnale analogico acquisito prima della conversione analogico-digitale, permettendo una verifica immediata della qualità del segnale proveniente dal microfono.

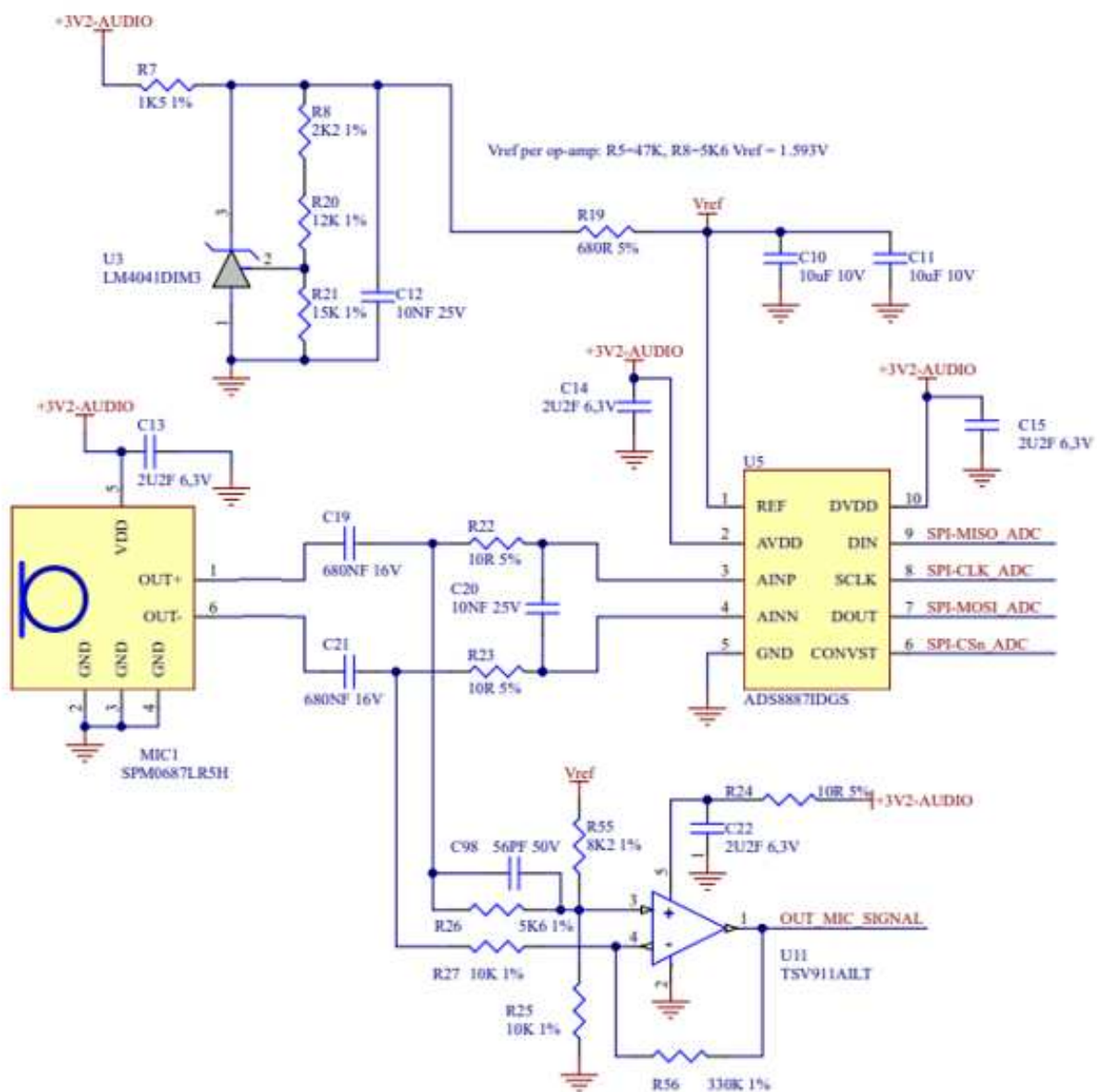


Figura 4: Schema del front-end analogico dello stetoscopio digitale.

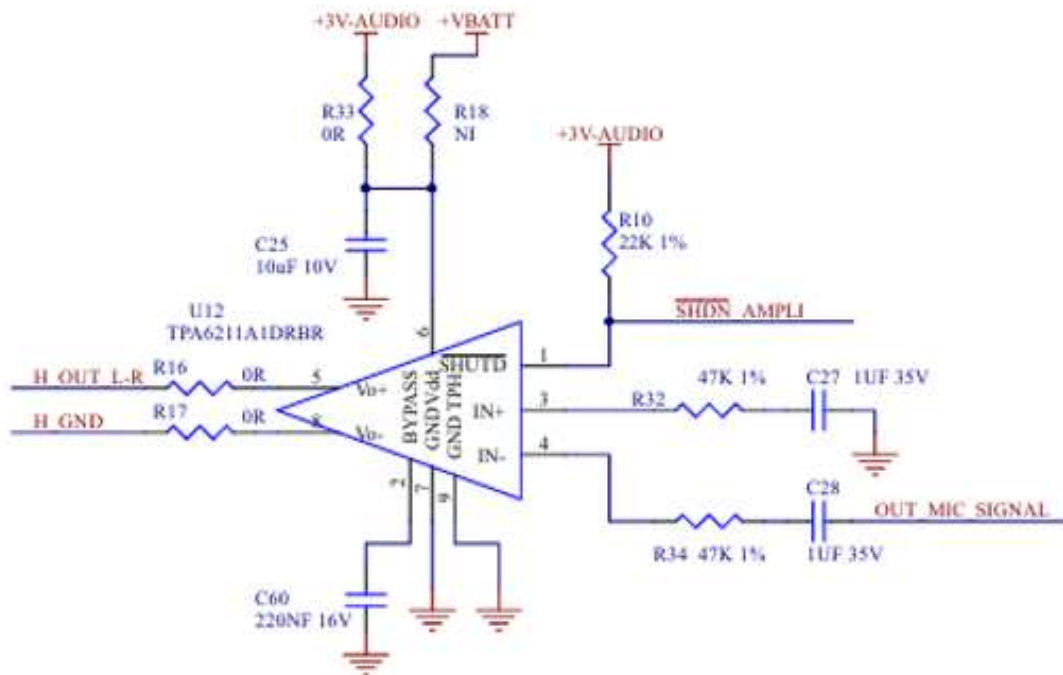


Figura 5: Uscita analogica jack per il monitoraggio del segnale tramite cuffie.

3.3.3 Conversione analogico-digitale

Il segnale analogico amplificato viene successivamente digitalizzato mediante il convertitore analogico-digitale integrato nel microcontrollore STM32L4R9ZIJ6, utilizzato come piattaforma di acquisizione del dispositivo.

Il convertitore opera con una risoluzione nominale di 12 bit, mentre il segnale viene campionato a una frequenza pari a $f_s=5\text{KHz}$.

Tale valore è stato scelto in quanto i suoni respiratori sono completamente rappresentati sotto i 2500Hz. I campioni acquisiti vengono successivamente memorizzati nella memoria del dispositivo e utilizzati per le successive fasi di elaborazione e analisi del segnale.

È importante osservare che, sebbene la risoluzione nominale del convertitore sia pari a 12 bit, il numero effettivo di bit utili (Effective Number Of Bits, ENOB) risulta inferiore a causa del rumore intrinseco del convertitore analogico-digitale. Questo aspetto può introdurre una componente di rumore di quantizzazione distribuita sull'intero spettro del segnale digitalizzato.

L'impatto di questo fenomeno sulla qualità delle registrazioni auscultatorie verrà analizzato in dettaglio nella sezione dedicata all'analisi sperimentale del rumore del sistema di acquisizione.

3.3.4 Salvataggio dei dati e parsing python

I campioni acquisiti dal canale audio dell'ADC non vengono salvati in Flash come messaggi del protocollo di comunicazione con header e CRC, ma come blocchi binari a struttura fissa. Il firmware accumula i campioni del microfono e, ogni volta che ne raccoglie 2000, crea un record contenente una firma iniziale (0xAAAA), un identificativo del blocco, la dimensione del record, un contatore progressivo e il vettore dei 2000 campioni memorizzati come valori uint16. Ogni blocco ha quindi dimensione costante pari a 4008 byte e la memoria Flash contiene una sequenza di record consecutivi dello stesso tipo. Il parsing in Python consiste nel leggere il file binario come successione di questi blocchi, verificarne la firma, estrarre i campi di intestazione e interpretare la parte finale come array di campioni audio a 16 bit, che vengono poi concatenati per ricostruire il segnale completo.

4. Analisi sperimentale del rumore

Durante le prime prove di acquisizione dei segnali respiratori è stato osservato che le registrazioni presentavano un livello di rumore di fondo significativamente elevato. Anche in ambienti apparentemente silenziosi il segnale risultava contaminato da una componente broadband percepibile all'ascolto come un rumore continuo di tipo “gracchiante”.

Per investigare l'origine di questa componente di disturbo è stata condotta un'analisi sperimentale del rumore generato dal sistema di acquisizione utilizzato in questo lavoro. L'obiettivo dell'analisi era caratterizzare le proprietà spettrali del segnale in condizioni acusticamente controllate e verificare quale elemento della catena di acquisizione potesse essere responsabile del rumore osservato. Le registrazioni sono state effettuate mediante il prototipo di stetoscopio digitale descritto nel capitolo precedente e successivamente estratto dalla memoria flash del dispositivo. I dati sono stati ricostruiti mediante una pipeline di parsing sviluppata in Python per il salvataggio dei file su pc in formato Raw.

4.1 Misure in ambiente acusticamente controllato

Per isolare il contributo del rumore generato dal sistema di acquisizione, alcune registrazioni sono state effettuate in condizioni acusticamente controllate all'interno di una piccola camera anecoica. In tali condizioni il contributo del rumore ambientale dovrebbe risultare significativamente ridotto e il segnale acquisito in tali condizioni è dominato solamente dal rumore elettronico intrinseco del sistema.

L'analisi delle registrazioni ha tuttavia evidenziato la presenza di un livello di rumore significativamente superiore alle aspettative.

4.2 Caratterizzazione spettrale del rumore

Per analizzare il comportamento del segnale acquisito è stata calcolata la rappresentazione tempo–frequenza mediante Short-Time Fourier Transform (STFT). Come mostrato in Fig. 6,

lo spettrogramma evidenzia la presenza di componenti broadband persistenti anche in assenza di sorgenti acustiche esterne.

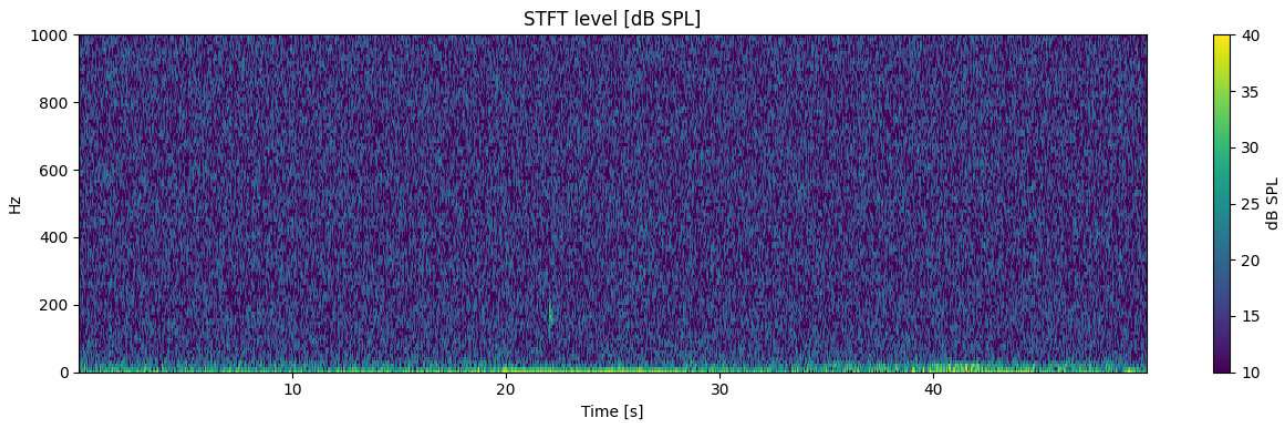


Figura 6: Spettrogramma del segnale acquisito in camera anecoica prima dell'introduzione dell'oversampling ADC.

L'energia del segnale risulta più pronunciata alle basse frequenze (sotto i 50 Hz) e successivamente distribuita in modo relativamente uniforme su gran parte della banda utile dei suoni respiratori.

Per ottenere una stima più stabile della distribuzione energetica del rumore è stata inoltre calcolata la densità spettrale di potenza (Power Spectral Density, PSD) del segnale registrato in condizioni di silenzio acustico.

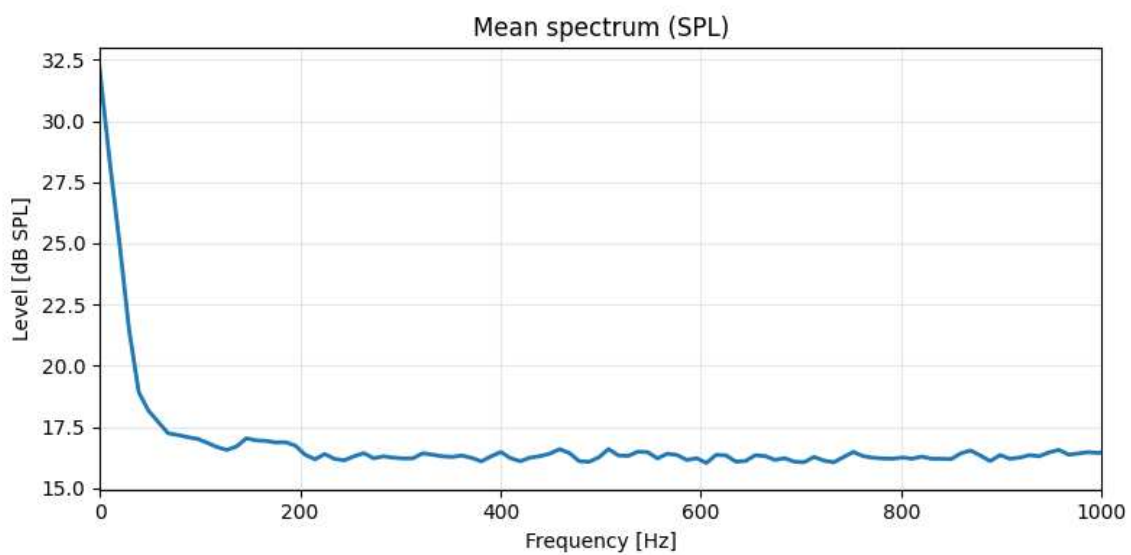


Figura 7: Densità spettrale di potenza del segnale acquisito in camera anecoica prima dell'introduzione dell'oversampling ADC.

Come mostrato in Fig. 7, la PSD evidenzia una componente di rumore distribuita su gran parte della banda analizzata, con una maggiore concentrazione energetica alle frequenze più basse. Questa distribuzione spettrale suggerisce la presenza di una sorgente di disturbo broadband non direttamente riconducibile a componenti acustiche ambientali.

Per comprendere meglio la natura di questo rumore è stato inoltre analizzato l'andamento temporale dei valori digitali prodotti dall'ADC. In Fig. 8 è mostrata la sequenza dei conteggi digitali acquisiti in camera anecoica.

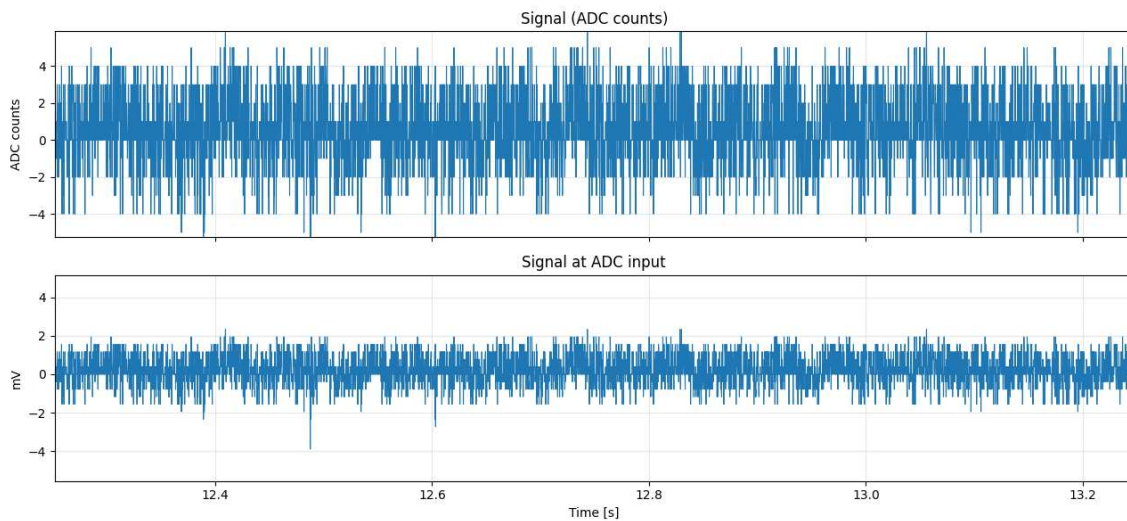


Figura 8: Andamento temporale dei conteggi ADC acquisiti in camera anecoica prima dell'introduzione dell'oversampling.

Si osserva come il segnale presenti variazioni rapide tra livelli digitali adiacenti anche in assenza di un segnale acustico rilevante. Il valore Δcount tra campioni consecutivi risulta tipicamente di circa ± 4 count rispetto alla media centrata in zero.

4.3 Verifica delle possibili sorgenti di rumore

Per individuare l'origine della componente broadband osservata è stata condotta una procedura di verifica progressiva dei principali elementi della catena di acquisizione del segnale.

4.3.1 Verifica della catena analogica

Il prototipo di stetoscopio dispone di un'uscita analogica che consente di monitorare direttamente il segnale proveniente dal microfono tramite un jack per cuffie come mostrato da schema

elettrico in Fig. 5. L'ascolto diretto del segnale analogico non ha evidenziato la presenza del rumore broadband osservato nelle registrazioni digitali.

Questa verifica ha quindi permesso di escludere come principali sorgenti di disturbo il microfono MEMS, lo stadio analogico di amplificazione e l'alimentazione del sistema.

4.3.2 Verifica della pipeline di salvataggio e parsing

Per escludere eventuali errori nella fase di salvataggio o ricostruzione dei dati è stata implementata a livello firmware una procedura di test che genera artificialmente una sinusoide digitale a frequenza nota.

Il segnale generato è stato salvato nella memoria del dispositivo e successivamente ricostruito mediante il parser Python utilizzato per l'estrazione dei dati. L'analisi spettrale del segnale ricostruito ha mostrato esclusivamente la componente sinusoidale alla frequenza impostata, senza l'introduzione di componenti spurie.

Questo test ha quindi permesso di verificare il corretto funzionamento della catena di memorizzazione dei dati e del processo di parsing.

4.3.3 Analisi della digitalizzazione ADC

Una volta escluse possibili cause nella catena analogica e nella pipeline di acquisizione, l'attenzione si è concentrata sulla fase di digitalizzazione del segnale.

Come spiegato in precedenza, il sistema utilizza un convertitore analogico-digitale a 12 bit integrato nel microcontrollore STM32L4R9ZIJ6. Tuttavia, il datasheet del dispositivo indica un numero effettivo di bit utili (Effective Number Of Bits, ENOB) pari a circa 10.5 bit.

La riduzione della risoluzione effettiva comporta la presenza di un errore di quantizzazione che potrebbe essere la causa più probabile del rumore broadband analizzato in camera anecoica.

Nel caso di segnali di ampiezza molto ridotta, come quelli tipicamente rilevati durante l'auscultazione respiratoria, il rumore di quantizzazione può diventare comparabile o addirittura dominante rispetto al segnale utile.

4.4 Riduzione del rumore mediante oversampling

Per verificare sperimentalmente questa ipotesi è stata modificata la configurazione di acquisizione dell'ADC introducendo una tecnica di oversampling hardware. In questa modalità ogni campione acquisito è ottenuto come media di 16 conversioni consecutive effettuate dal convertitore analogico-digitale.

Questa tecnica consente di ridurre l'impatto del rumore di quantizzazione aumentando la risoluzione effettiva del convertitore.

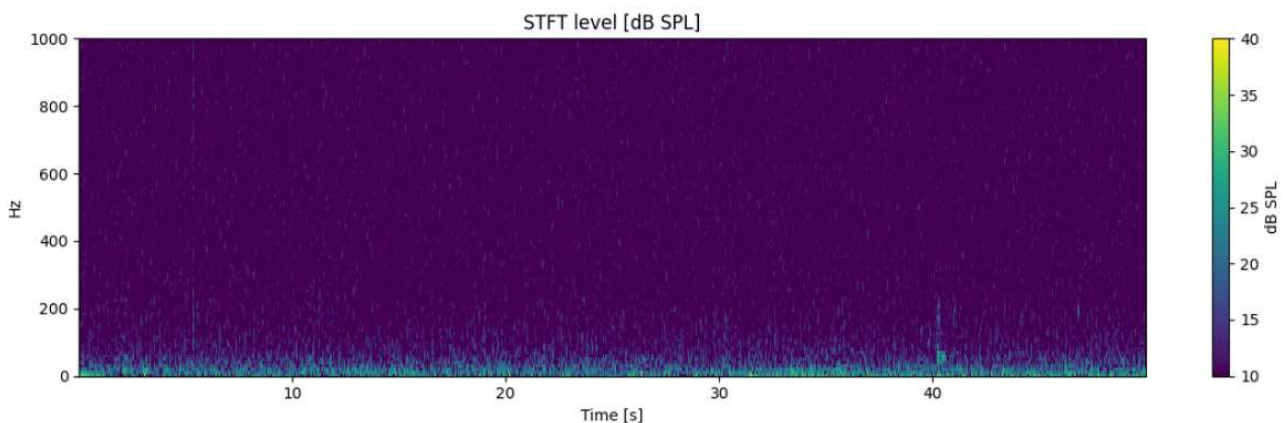


Figura 9: Spettrogramma del segnale acquisito in camera anecoica dopo l'introduzione dell'oversampling ADC.

L'analisi delle registrazioni effettuate con questa configurazione ha evidenziato una significativa riduzione della componente broadband precedentemente osservata.

Per analizzare il comportamento del convertitore nel dominio temporale è stato inoltre osservato l'andamento dei conteggi digitali prodotti dall'ADC dopo l'introduzione dell'oversampling.

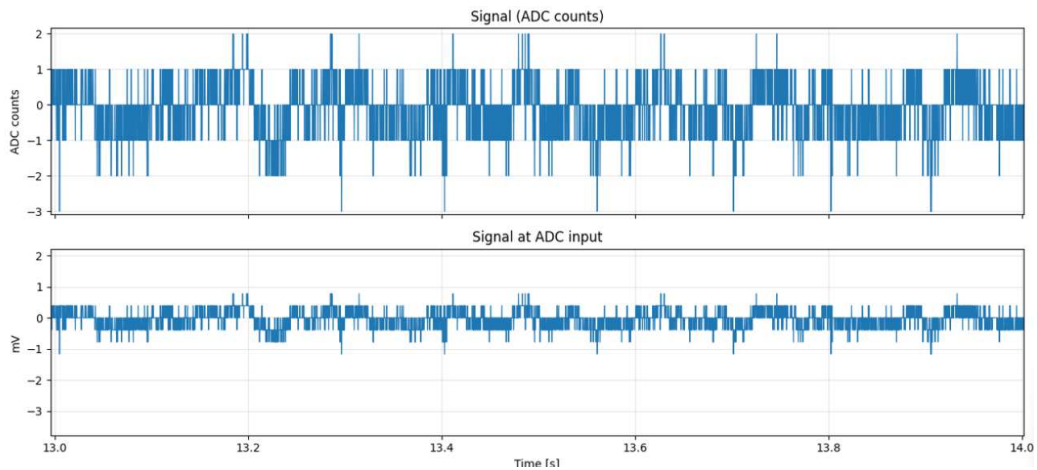


Figura 10: Andamento dei conteggi ADC acquisiti in camera anecoica dopo l'introduzione dell'oversampling.

Come mostrato in Fig. 10, l'ampiezza delle fluttuazioni tra campioni consecutivi risulta sensibilmente ridotta rispetto alle acquisizioni precedenti, passando da circa ± 4 a circa ± 2 count confermando la diminuzione del rumore introdotto dalla fase di digitalizzazione.

Per osservare in modo più stabile la distribuzione energetica residua del segnale è stata infine calcolata la densità spettrale di potenza (PSD) delle registrazioni ottenute con la nuova configurazione di acquisizione.

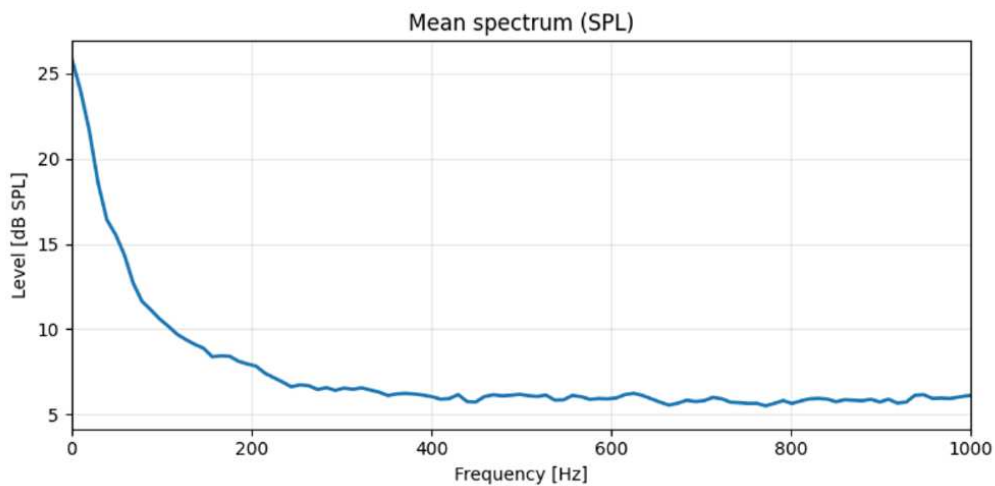


Figura 11: Densità spettrale di potenza del segnale acquisito in camera anecoica dopo l'introduzione dell'oversampling ADC.

Il confronto con le misure precedenti mostra come l'introduzione dell'oversampling riduca in modo significativo la componente broadband osservata nelle registrazioni in camera anecoica. Questo risultato conferma sperimentalmente che la principale sorgente del rumore "gracchiante"

osservato nelle acquisizioni iniziali era associata alla fase di digitalizzazione del segnale e, in particolare, alla limitata risoluzione effettiva del convertitore analogico-digitale ($\text{ENOB} \approx 10.5$ bit).

In condizioni di camera anecoica, la nuova configurazione consente di ottenere registrazioni in cui il livello di rumore residuo non risulta più percettibile all'ascolto e non evidenzia componenti broadband significative nelle rappresentazioni spettrali. Questo risultato fornisce quindi una validazione sperimentale della catena di acquisizione e consente di procedere con la fase successiva del lavoro, ovvero la costruzione del dataset di registrazioni auscultatorie utilizzato per lo sviluppo e la valutazione dell'algoritmo di denoising.

4.5 Esempio di rumore ambientale reale

Per valutare il comportamento del sistema in condizioni di acquisizione più realistiche, sono state effettuate alcune registrazioni ambientali in una stanza domestica silenziosa. Durante la misura erano presenti deboli sorgenti sonore provenienti dall'ambiente circostante, tra cui il rumore di ventilazione della stanza adiacente e occasionali passi o brevi segmenti di parlato, situazione comparabile con un ambulatorio medioco piuttosto silenzioso.

La rappresentazione tempo–frequenza del segnale acquisito è mostrata in Fig. 12. A differenza delle misure effettuate in camera anecoica, la STFT evidenzia la presenza di componenti spettrali variabili nel tempo associate alle sorgenti acustiche ambientali.

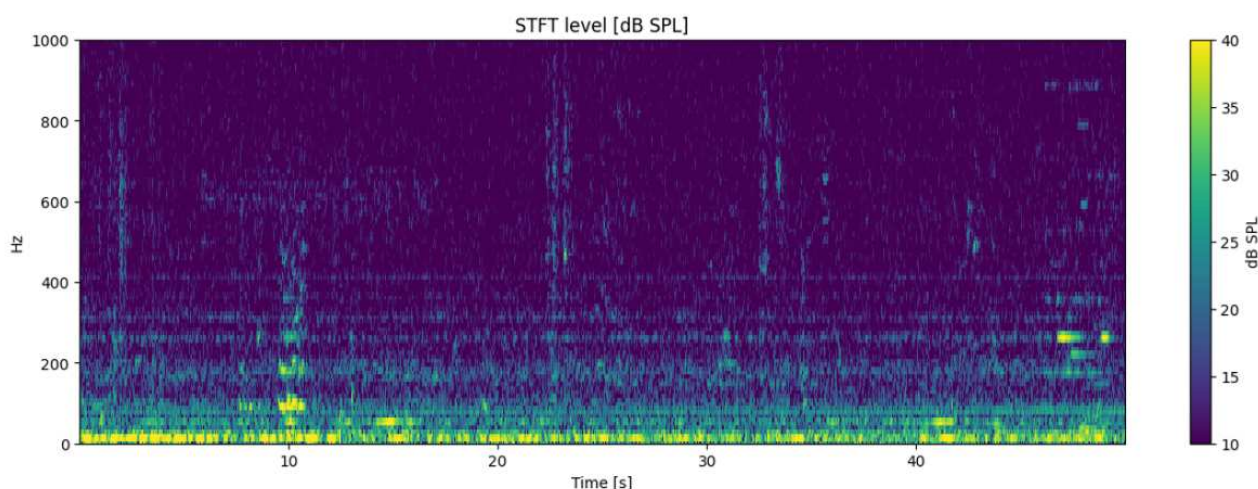


Figura 12: Spettrogramma del segnale acquisito in una stanza silenziosa con rumori ambientali deboli provenienti da ambienti adiacenti.

Si osserva come l'energia del segnale sia principalmente concentrata alle basse frequenze, con componenti diffuse fino a circa 400–500 Hz e occasionali eventi impulsivi associati a rumori transitori dell'ambiente. Questo comportamento è tipico del rumore acustico ambientale e conferma l'elevata sensibilità del sensore acustico utilizzato.

Questa osservazione evidenzia come, nelle registrazioni auscultatorie reali, il segnale possa essere facilmente contaminato da sorgenti di rumore ambientale anche in condizioni apparentemente silenziose.

4.5.1 Esempio di auscultazione reale

Per comprendere meglio il comportamento del segnale acquisito durante una registrazione auscultatoria reale, è stata analizzata la rappresentazione tempo–frequenza di una registrazione effettuata mediante lo stetoscopio digitale descritto nelle sezioni precedenti.

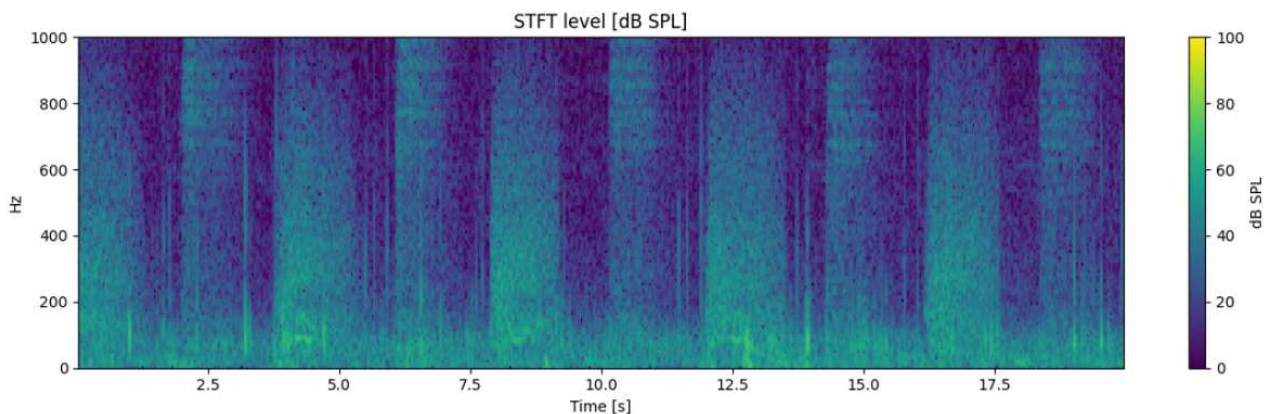


Figura 13: Spettrogramma di una registrazione auscultatoria reale. Le variazioni dell'energia spettrale evidenziano le fasi di attività respiratoria alternate a pause respiratorie.

La Fig. 13 mostra lo spettrogramma di una registrazione acquisita durante un ciclo respiratorio. Le fasi di inspirazione ed espirazione risultano caratterizzate da un aumento dell'energia spettrale nella banda tipicamente associata ai suoni polmonari, compresa approssimativamente tra 100 e 800 Hz.

Durante queste fasi il segnale presenta una struttura spettrale ben definita, associata alle vibrazioni generate dal flusso d'aria nelle vie respiratorie e trasmesse attraverso la parete toracica fino allo stetoscopio digitale.

4.5.2 Comportamento spettrale durante le pause respiratorie

Osservando la Fig. 13 è possibile notare che tra le fasi di inspirazione ed espirazione sono presenti intervalli temporali in cui l'attività respiratoria risulta assente o molto ridotta. Queste regioni corrispondono alle pause respiratorie naturali del ciclo respiratorio.

In tali intervalli il segnale acquisito non risulta tuttavia completamente silenzioso. Lo spettrogramma evidenzia infatti la presenza di una componente broadband relativamente uniforme sull'intera banda respiratoria.

Questa componente è principalmente associata a diversi contributi di rumore presenti durante l'acquisizione. Tra le possibili sorgenti si possono includere il rumore elettronico residuo della catena di acquisizione, seppur molto ridotto, eventuali code respiratorie dovute alla propagazione delle vibrazioni toraciche, rumori ambientali presenti nell'ambiente di misura e componenti acustiche generate da attività viscerali o da movimenti interni del soggetto.

Dal punto di vista spettrale, il segnale osservato durante le pause respiratorie presenta quindi caratteristiche significativamente diverse rispetto alle fasi di attività respiratoria, risultando dominato da componenti di rumore broadband piuttosto che da strutture spettrali tipiche dei suoni polmonari.

4.5.3 Implicazioni per il denoising adattivo

L'analisi delle registrazioni auscultatorie evidenzia quindi una proprietà importante del segnale respiratorio acquisito. Durante le pause respiratorie il contributo delle componenti respiratorie risulta fortemente ridotto, mentre il segnale è dominato prevalentemente da componenti di rumore indesiderate.

Queste regioni del segnale possono quindi essere utilizzate per ottenere una stima della distribuzione spettrale del rumore presente nella registrazione.

Questa osservazione rappresenta il principio alla base del framework di denoising proposto in questo lavoro. In particolare, il metodo sfrutta l'identificazione automatica delle regioni dominate dal rumore per aggiornare dinamicamente una baseline spettrale del rumore e applicare successivamente un filtraggio adattivo basato su filtro di Wiener.

Nel capitolo successivo verrà quindi descritto nel dettaglio il framework sviluppato per l'identificazione delle regioni dominate dal rumore e per l'attenuazione adattiva delle componenti

broadband nelle registrazioni auscultatorie.

5. Framework proposto per il denoising adattivo dei suoni respiratori

5.1 Panoramica del framework

Come discusso nei capitoli precedenti, i segnali respiratori acquisiti mediante stetoscopi digitali possono essere contaminati da diverse sorgenti di disturbo. In generale, tali componenti possono essere suddivise in due categorie principali: rumore di fondo quasi stazionario e artefatti impulsivi.

Il rumore di fondo (*background noise*) include principalmente il rumore ambientale presente nell'ambiente di acquisizione e il rumore elettronico introdotto dalla catena di misura, come il rumore del microfono, dello stadio di amplificazione analogica e del convertitore analogico-digitale. Queste componenti tendono ad avere una distribuzione spettrale relativamente ampia e una variazione temporale lenta.

Una seconda categoria è costituita dagli artefatti impulsivi, generalmente associati a vibrazioni meccaniche, movimenti dello stetoscopio o variazioni della pressione di contatto tra il sensore e la superficie toracica. Questi fenomeni si manifestano tipicamente come transitori di ampiezza elevata e durata limitata nel tempo.

Il framework di denoising proposto in questa tesi è stato progettato specificamente per ridurre la componente di *background noise*. Gli artefatti impulsivi di natura meccanica non rappresentano invece il target principale del metodo e non vengono esplicitamente trattati dal processo di filtraggio. Il sistema è stato inoltre concepito per un'implementazione in tempo reale su piattaforme embedded. Per questo motivo l'elaborazione avviene in modalità *frame-wise*: il segnale viene suddiviso in finestre temporali successive e ciascun frame viene processato sequenzialmente producendo direttamente un frame filtrato in uscita.

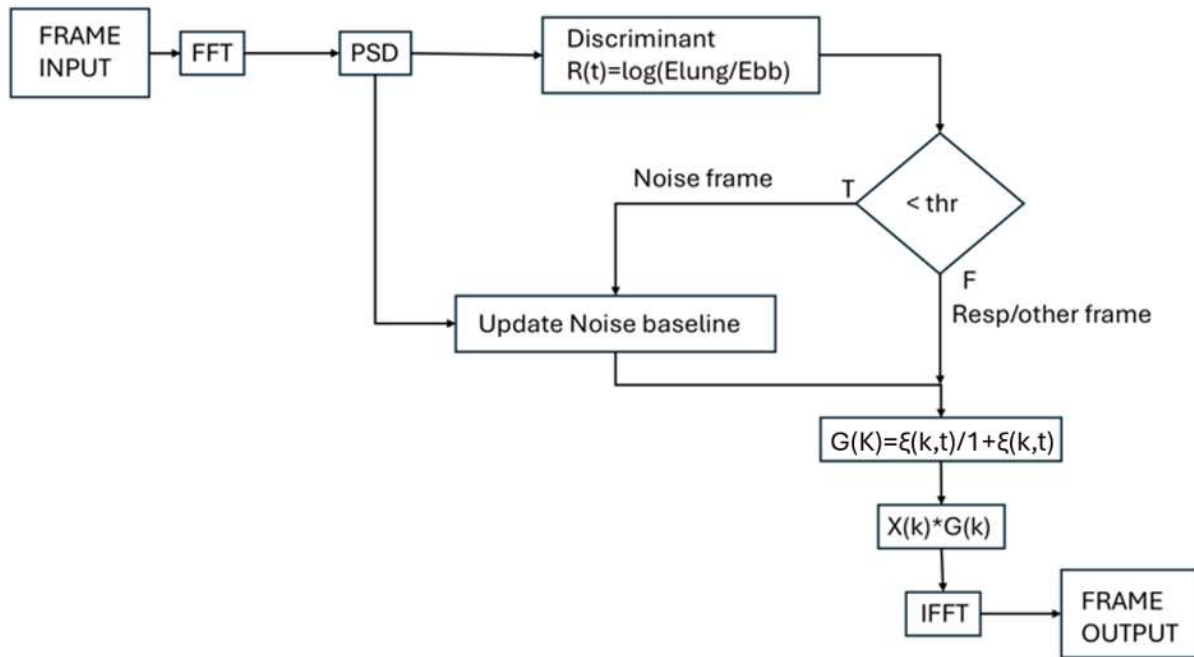


Figura 14: Framework proposto

5.2 Modello del segnale

Il segnale auscultatorio osservato può essere modellato come la somma tra il contributo respiratorio utile e una componente di rumore di fondo:

$$y(t) = x(t) + n(t)$$

dove $x(t)$ rappresenta il segnale respiratorio e $n(t)$ il rumore di background.

Nel framework proposto si assume che il rumore di fondo possa essere considerato localmente stazionario su intervalli temporali limitati, ovvero tra due aggiornamenti consecutivi della baseline di rumore. Questa ipotesi è ragionevole poiché le principali componenti del rumore ambientale ed elettronico tendono a variare lentamente nel tempo.

Durante le pause respiratorie il contributo del segnale utile diventa trascurabile e il segnale osservato può essere approssimato come

$$y(t) \approx n(t)$$

Queste regioni temporali possono quindi essere utilizzate per stimare direttamente lo spettro di potenza del rumore e aggiornare progressivamente la baseline spettrale utilizzata dal filtro di Wiener.

5.3 Analisi spettrale frame-wise

Il segnale acquisito viene suddiviso in frame temporali consecutivi e per ciascun frame viene calcolata la trasformata discreta di Fourier (FFT). A partire dai coefficienti spettrali viene quindi stimata la densità spettrale di potenza:

$$P(k, t) = |X(k, t)|^2$$

dove $X(k, t)$ rappresenta il coefficiente spettrale del bin k nel frame t .

La rappresentazione tempo-frequenza ottenuta consente di analizzare l'evoluzione spettrale del segnale respiratorio e costituisce la base per le successive fasi di discriminazione tra segnale utile e rumore.

5.4 Discriminante spettrale

Per distinguere automaticamente i frame dominati dal segnale respiratorio da quelli dominati dal rumore viene introdotto un discriminante spettrale $r(t)$.

Il discriminante è definito come il logaritmo del rapporto tra l'energia spettrale del frame nella banda polmonare e l'energia broadband:

$$r(t) = \log \left(\frac{E_{\text{lung}}}{E_{\text{BB}}} \right)$$

dove

$$E_{\text{BB}}(t) = \frac{1}{|\mathbf{B}_{\text{BB}}|} \sum_{k \in \mathbf{B}_{\text{BB}}} P(k, t)$$

$$E_{\text{lung}}(t) = \frac{1}{|\mathbf{B}_L|} \sum_{k \in \mathbf{B}_L} P(k, t)$$

Come si può osservare in Fig 15, il discriminante $r(t)$ assume valori più elevati durante le fasi di inspirazione ed espirazione, quando l'energia respiratoria domina nella banda polmonare. Al contrario, durante le pause respiratorie il valore di $r(t)$ tende a ridursi poiché il segnale osservato è costituito prevalentemente dal rumore di fondo.

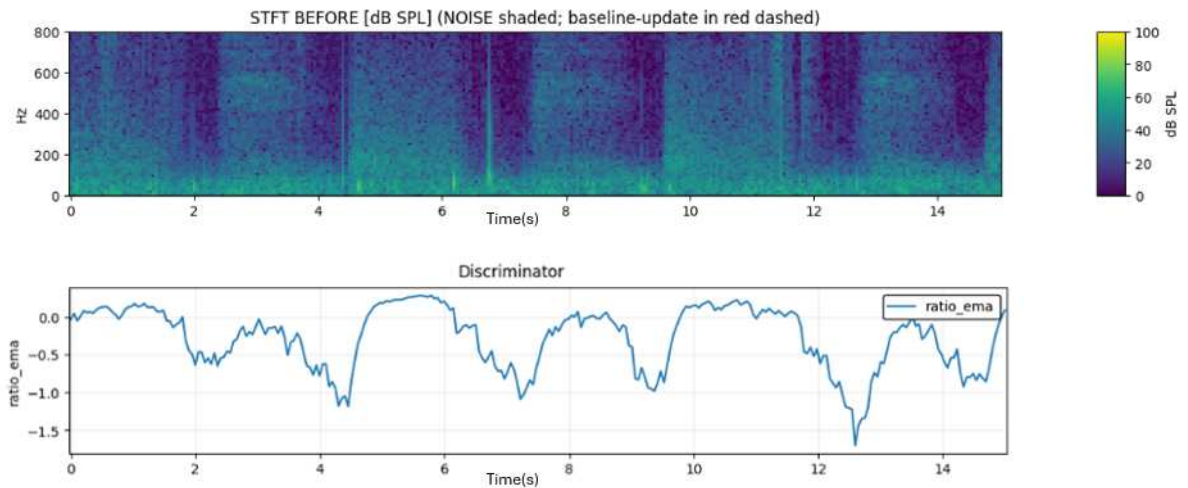


Figura 15: Andamento del discriminante $r(t)$ nelle diverse fasi respiratorie.

In questo modo il discriminante consente di identificare automaticamente le regioni temporali

dominate dal rumore.

5.5 Identificazione delle regioni di rumore

Per rendere robusta la classificazione dei frame viene utilizzata una soglia adattiva con isteresi. In particolare vengono definite due soglie: una soglia di ingresso e una soglia di uscita dalla regione *NOISE*. Le soglie sono ricavate statisticamente dalla distribuzione del discriminante $r(t)$.

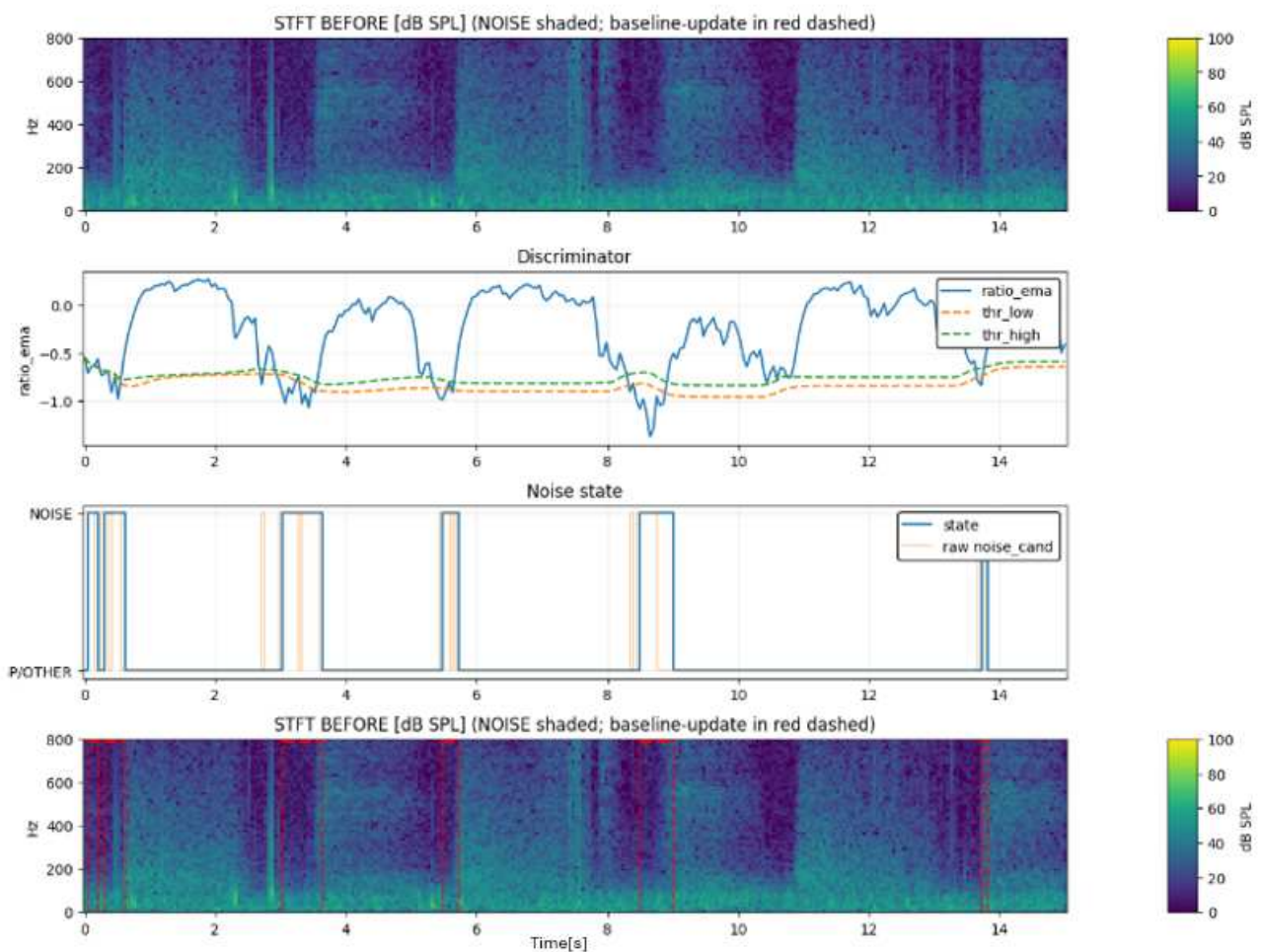


Figura 16: Definizione automatica delle regioni dominate dal rumore

Nel framework implementato in questo lavoro le soglie sono state fissate empiricamente ai quantili Q_{10} e Q_{15} della distribuzione di $r(t)$.

Quando il valore del discriminante scende al di sotto della soglia di ingresso il frame corrente viene classificato come appartenente alla regione *NOISE*. Finché il discriminante rimane al

di sotto della soglia di uscita il sistema mantiene lo stato di rumore, evitando commutazioni instabili tra stati adiacenti.

5.6 Stima adattiva della baseline di rumore

Una volta individuati i frame dominati dal rumore, il loro contenuto spettrale viene utilizzato per aggiornare progressivamente una baseline di rumore nel dominio della frequenza.

La baseline rappresenta la stima adattiva dello spettro di potenza del background noise e viene aggiornata esclusivamente durante i frame classificati come *NOISE*.

Per ciascun bin di frequenza k lo spettro di rumore stimato $N(k, t)$ viene aggiornato mediante una media esponenziale:

$$N(k, t) = \alpha N(k, t - 1) + (1 - \alpha)P(k, t)$$

dove $P(k, t)$ rappresenta la potenza spettrale osservata nel frame corrente e α è un coefficiente di smoothing che controlla la velocità di adattamento del modello di rumore.

5.7 Filtraggio mediante Wiener

Una volta stimata la baseline di rumore $N(k, t)$, il filtro di Wiener viene applicato nel dominio spettrale.

A partire dalla potenza osservata viene calcolato il rapporto segnale-rumore a posteriori:

$$\gamma(k, t) = \frac{P(k, t)}{N(k, t)}$$

Una stima della potenza del segnale utile può essere ottenuta come

$$\hat{S}(k, t) = \max(P(k, t) - N(k, t), 0)$$

da cui segue la stima del rapporto segnale-rumore a priori

$$\xi(k, t) = \frac{\hat{S}(k, t)}{N(k, t)}$$

Il guadagno del filtro di Wiener viene quindi espresso come

$$G(k, t) = \frac{\xi(k, t)}{1 + \xi(k, t)}$$

e applicato ai coefficienti spettrali del frame corrente:

$$\hat{X}(k, t) = G(k, t)X(k, t)$$

In questo modo il filtraggio viene effettuato in maniera adattiva: quando il rapporto segnale-rumore è elevato il guadagno tende all'unità e il coefficiente spettrale viene preservato; viceversa, quando la potenza osservata risulta comparabile con quella del rumore il guadagno diminuisce attenuando la componente spettrale corrispondente. Il risultato esposto in figura 17, presenta la STFT del segnale prima e dopo l'applicazione del guadagno di wiener.

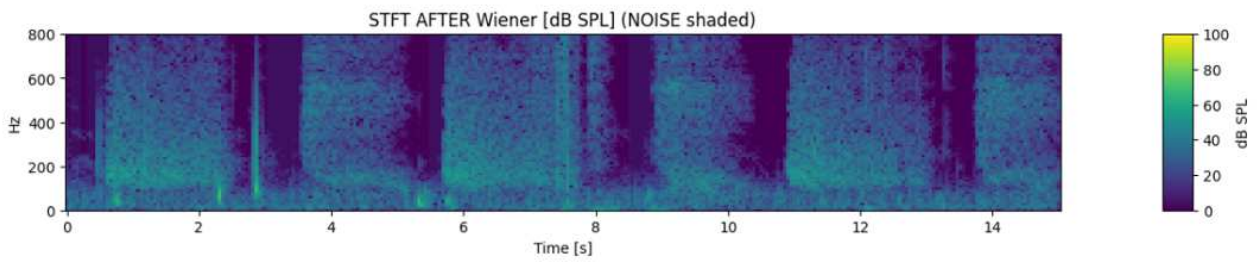


Figura 17: Spettrogramma pre e post applicazione del guadagno di wiener

5.8 Esempio di comportamento del filtro

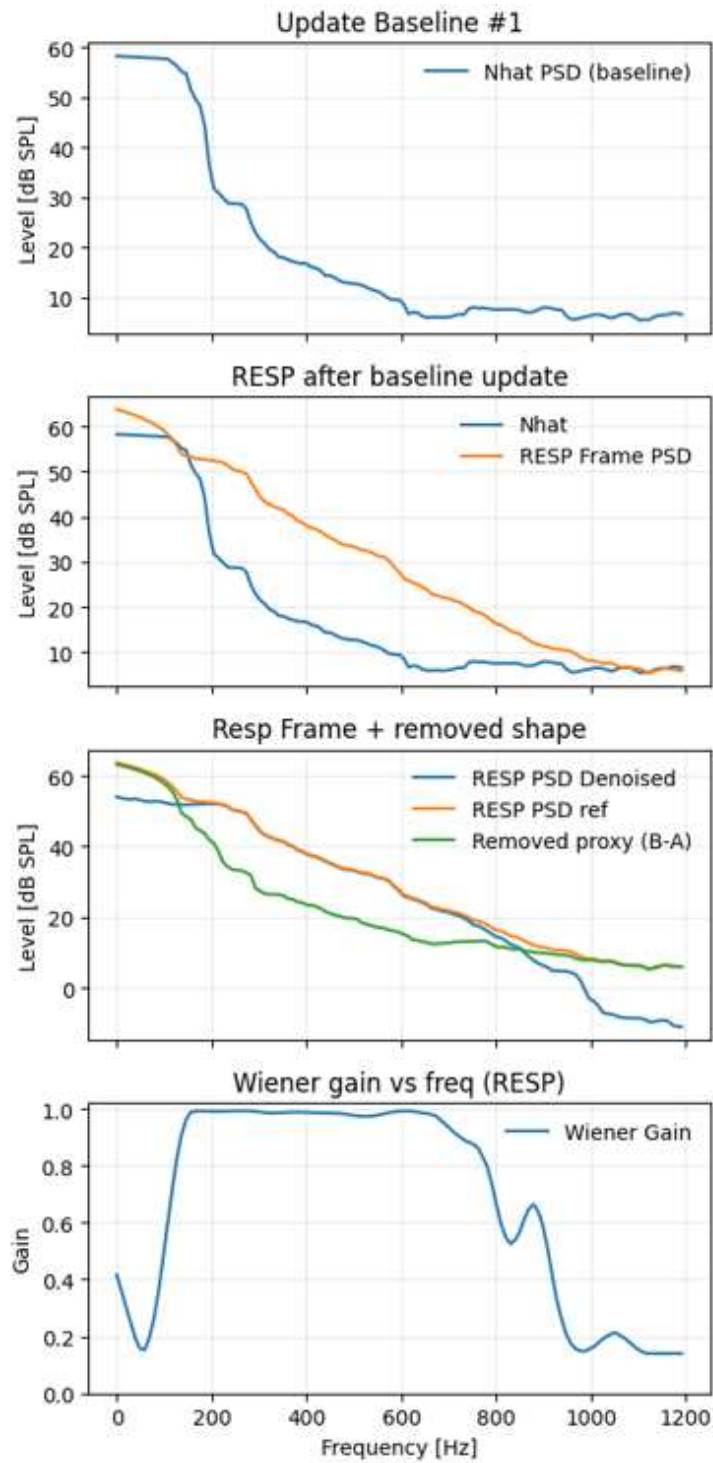


Figura 18: Esempio di applicazione del filtro di Wiener su un frame spettrale.

La Fig. 18 mostra un esempio del processo di filtraggio applicato a un frame spettrale. Nella figura è riportato l'ultimo aggiornamento della baseline di rumore una regione *NOISE* e il frame successivo classificato come *RESP*. Il frame classificato come *RESP* presenta un aumento dell'energia spettrale nella banda tipica dei suoni respiratori, compresa approssimativamente tra 100 Hz e 800 Hz.

Nelle regioni spettrali esterne alla banda respiratoria, dove la potenza osservata risulta comparabile con la baseline di rumore stimata, il filtro di Wiener riduce significativamente il guadagno attenuando tali componenti.

Al contrario, nella banda polmonare il rapporto segnale-rumore risulta più elevato e il guadagno del filtro rimane vicino all'unità, preservando le componenti spettrali associate al respiro.

5.9 Ricostruzione del segnale

I coefficienti spettrali filtrati vengono infine riportati nel dominio temporale mediante trasformata inversa (IFFT). I frame temporali ottenuti vengono quindi ricombinati tramite una procedura di sovrapposizione e somma (*overlap-add*), garantendo la continuità del segnale ricostruito.

Il flusso complessivo del framework può essere riassunto nei seguenti passi:

- acquisizione di un frame del segnale;
- calcolo della FFT e della PSD;
- valutazione del discriminante spettrale $r(t)$;
- classificazione del frame come *RESP* o *NOISE*;
- aggiornamento della baseline di rumore durante i frame *NOISE*;
- calcolo del guadagno del filtro di Wiener;
- applicazione del guadagno allo spettro del frame;
- ricostruzione del segnale nel dominio temporale.

Il risultato finale di questo processo è una versione denoised del segnale respiratorio, ottenuta mediante attenuazione selettiva delle componenti spettrali associate al rumore e preservazione delle caratteristiche acustiche del segnale di interesse.

6. Creazione del dataset sperimentale e metodologia di valutazione

6.1 Creazione del dataset

Il dataset utilizzato in questo lavoro è stato generato attraverso sessioni controllate di auscultazione polmonare effettuate mediante lo stetoscopio digitale descritto nel Capitolo 3. L'obiettivo della fase di acquisizione era ottenere registrazioni rappresentative delle condizioni realistiche in cui uno stetoscopio digitale viene utilizzato nella pratica clinica o in contesti di monitoraggio respiratorio.

Le registrazioni sono state effettuate in ambienti interni non acusticamente isolati, al fine di riprodurre condizioni di acquisizione realistiche piuttosto che scenari artificialmente controllati. Durante le sessioni di registrazione erano presenti diverse sorgenti di rumore ambientale, tra cui riverberazione della stanza, sistemi di ventilazione e attività di fondo a distanza. Questa scelta consente di ottenere segnali caratterizzati da condizioni acustiche non stazionarie, simili a quelle che possono verificarsi durante l'utilizzo reale del dispositivo.

Per ciascun soggetto sono state effettuate più sessioni di registrazione, coprendo diverse posizioni auscultatorie sul torace. Durante l'acquisizione, lo stetoscopio è stato mantenuto in contatto stabile con la parete toracica senza l'utilizzo di sistemi di isolamento meccanico o smorzamento delle vibrazioni oltre alle normali modalità di utilizzo clinico. Tutte le registrazioni sono state effettuate dallo stesso operatore, al fine di garantire condizioni di acquisizione il più possibile uniformi.

Le posizioni di auscultazione sono state definite secondo un insieme di otto siti anatomici standardizzati comunemente utilizzati negli studi sui suoni respiratori e coerenti con la configurazione riportata in letteratura [15].

Gli otto siti di riferimento includono:

- lobi superiori paravertebrali,
- lobi medi paravertebrali,
- lobi inferiori ascellari,
- lobi inferiori paravertebrali.

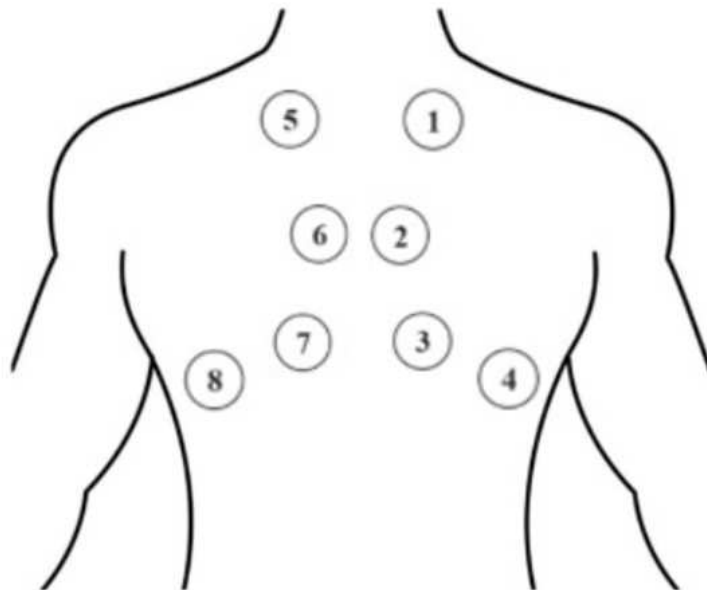


Figura 19: Siti di Auscultazione.

Per ciascun soggetto sono stati selezionati due siti tra queste posizioni standardizzate.

Durante le registrazioni, ai soggetti è stato richiesto di eseguire cicli di inspirazione ed espirazione a ritmo naturale inspirando profondamente con il naso ed espirando con la bocca. Ogni ciclo respiratorio include naturalmente brevi pause tra le fasi respiratorie, che risultano composte prevalentemente da rumore di fondo, rumore degli organi interni e rumore elettronico residuo del sistema di acquisizione. Queste pause naturali sono coerenti con le assunzioni alla base del framework proposto, che utilizza tali intervalli per stimare la baseline spettrale del rumore.

Tutti i soggetti coinvolti nello studio erano volontari adulti, con un'età compresa tra 22 e 87 anni, senza diagnosi di patologie respiratorie acute al momento delle registrazioni.

Tabella 2: Dataset

Numero di soggetti	15
Range di età	22–87
Auscultazioni totali	30
Durata media dell'auscultazione	≈ 1 min
Durata totale delle auscultazioni	30 minutes
frequenza di campionamento	5 kHz
siti di auscultazione	2 siti

6.2 Configurazione dell'analisi frame-based

Tutti i segnali del dataset sono stati elaborati utilizzando una configurazione di analisi spettrale basata su frame temporali, coerentemente con l'architettura del framework di denoising descritto nel Capitolo 3.

In particolare, il segnale è stato suddiviso in finestre temporali successive sulle quali è stata calcolata la trasformata di Fourier discreta (FFT). I parametri dell'analisi sono stati scelti in modo da ottenere un compromesso tra risoluzione in frequenza, stabilità delle stime spettrali e complessità computazionale, in modo da mantenere la compatibilità con possibili implementazioni embedded in tempo reale.

La configurazione adottata è la seguente:

- Lunghezza FFT: $N = 512$ campioni
- Tipo di finestra: finestra di Hann
- Hop size: 256 campioni (overlap del 50%)

Considerando una frequenza di campionamento pari a $f_s = 5$ kHz, questa configurazione fornisce una risoluzione in frequenza pari a

$$\Delta f = \frac{f_s}{N} \approx 9.77 \text{ Hz}$$

Tale risoluzione risulta adeguata per l'analisi della banda spettrale dei suoni respiratori, generalmente compresa tra circa 100 Hz e 800 Hz.

La scelta di una lunghezza di finestra relativamente contenuta consente inoltre di mantenere una latenza ridotta.

6.3 Filosofia di valutazione

Nel contesto delle registrazioni auscultatorie reali, la disponibilità di segnali di riferimento completamente privi di rumore è generalmente limitata. In letteratura esistono alcuni dataset di suoni respiratori considerati *clean*, ma sono stati acquisiti con dispositivi e catene di acquisizione differenti. Poiché le caratteristiche del segnale registrato dipendono anche dall'hardware di acquisizione e dall'ambiente, in questo lavoro si è scelto di operare direttamente su registrazioni ottenute con lo stesso stetoscopio digitale utilizzato nello sviluppo del framework. Di conseguenza non è disponibile una coppia di segnali rumore-pulito che permetta un confronto diretto con un riferimento ideale. Per questo motivo la valutazione del metodo si basa su un'analisi di consistenza interna del segnale, verificando se il comportamento del sistema di denoising risulta coerente con le proprietà attese del framework proposto. In particolare, la valutazione si concentra su tre aspetti fondamentali:

1. attenuazione efficace del rumore di fondo stazionario,
2. aumento del contrasto tra contenuto respiratorio e background noise,
3. coerenza tra le componenti spettrali rimosse e il modello di rumore stimato.

Tutte le metriche di valutazione sono calcolate automaticamente a partire dalle uscite dell'algoritmo, senza richiedere annotazioni manuali a livello di frame. Questo approccio consente di ottenere una valutazione oggettiva e riproducibile delle prestazioni del sistema.

6.4 Metriche quantitative

Le prestazioni del sistema di denoising vengono valutate mediante un insieme di metriche quantitative progettate per caratterizzare il comportamento di attenuazione del rumore e la redistribuzione dell'energia spettrale introdotta dal framework adattivo basato su filtro di Wiener.

Le metriche adottate consentono di analizzare tre aspetti complementari: la riduzione del rumore di fondo, l'impatto del filtraggio nella banda respiratoria e il miglioramento relativo della dominanza spettrale dei suoni respiratori.

6.4.1 Noise Suppression Index (NSI)

Il Noise Suppression Index misura l'attenuazione del segnale durante i frame classificati come dominati dal rumore:

$$NSI_{dB} = 10 \log_{10} \left(\frac{E_{B,after}}{E_{B,before}} \right)$$

dove E_B rappresenta l'energia spettrale calcolata nell'intervallo di frequenze compreso tra 80 Hz e 1000 Hz, corrispondente a un intorno esteso della banda respiratoria. Tale scelta consente di valutare in modo mirato la soppressione del rumore di fondo nelle componenti spettrali più rilevanti per i suoni respiratori. Con questa convenzione, valori negativi dell'indice corrispondono a una riduzione dell'energia dopo il denoising, mentre valori prossimi a zero indicano una variazione trascurabile.

6.4.2 Respiratory Attenuation Index (RAI)

Il Respiratory Attenuation Index quantifica l'attenuazione dell'energia spettrale nella banda dominata dai suoni respiratori durante i frame classificati come respiratori:

$$RAI_{dB} = 10 \log_{10} \left(\frac{E_{L,after}}{E_{L,before}} \right)$$

Valori negativi indicano attenuazione dell'energia respiratoria, mentre valori prossimi a zero corrispondono a una sostanziale preservazione del contenuto spettrale nella banda di interesse. Poiché le componenti di rumore e di segnale respiratorio sono spesso sovrapposte nello spettro, una certa attenuazione nella banda respiratoria è auspicabile anche quando il sistema opera correttamente.

6.4.3 Respiratory–Background Contrast Improvement (ΔC)

Per valutare se il denoising aumenta effettivamente la separazione tra contenuto respiratorio e background noise, viene definita una metrica di contrasto basata sulla caratterizzazione dei segmenti respiratori e dei corrispondenti segmenti di rumore.

In assenza di un segnale di riferimento privo di rumore, la stima diretta del rapporto segnale-rumore (SNR) non è possibile. Per questo motivo, viene introdotta una metrica proxy basata sul contrasto tra l'energia spettrale delle regioni respiratorie e quella della baseline di rumore stimata.

Tale approccio sfrutta la segmentazione del segnale in intervalli dominati dal segnale respiratorio e intervalli dominati dal rumore, consentendo una stima locale e adattiva del rapporto tra le due componenti. La variazione del contrasto tra queste regioni, prima e dopo il denoising, fornisce quindi una misura indiretta del miglioramento del SNR, coerente con le condizioni di acquisizione reali.

Per ciascun segmento classificato come RESP, si considera l'energia nella banda polmonare B_L :

$$E_{L,resp} = \sum_{k \in B_L} P_{resp}(k)$$

dove $P_{resp}(k)$ rappresenta la densità spettrale di potenza media del segmento respiratorio.

Analogamente, si considera l'energia del background locale nella stessa banda, stimata a partire dall'ultimo aggiornamento della baseline di rumore disponibile per quel segmento:

$$E_{L,noise} = \sum_{k \in B_L} P_{noise}(k)$$

dove $P_{noise}(k)$ rappresenta lo spettro della baseline di rumore, assunto costante tra due aggiornamenti successivi e derivato dai frame classificati come NOISE.

Il contrasto tra contenuto respiratorio e background noise è quindi definito come:

$$C = 10 \log_{10} \left(\frac{E_{L,resp}}{E_{L,noise}} \right)$$

Questa quantità può essere interpretata come una misura del rapporto tra energia respiratoria e rumore nella banda di interesse, risultando quindi analoga a una stima locale di SNR.

Tale quantità viene calcolata sia prima che dopo il denoising, considerando rispettivamente gli spettri originali e quelli filtrati. La variazione di contrasto introdotta dall'algoritmo è quindi definita come:

$$\Delta C = C_{after} - C_{before}$$

Valori positivi di ΔC indicano un aumento del rapporto tra contenuto respiratorio e rumore di fondo, ovvero un miglioramento del contrasto equivalente a un incremento di SNR nelle condizioni di acquisizione reale.

È importante sottolineare che questa metrica tiene conto simultaneamente sia della riduzione del rumore sia dell'eventuale attenuazione del contenuto respiratorio, fornendo quindi una misura diretta dell'efficacia del framework in termini di separazione tra le due componenti.

L'analisi della distribuzione di ΔC sui segmenti respiratori del dataset consente di valutare in modo robusto l'efficacia del metodo nell'amplificare il contrasto tra contenuti polmonari e rumore di fondo in condizioni realistiche.

6.5 Aggregazione e analisi statistica

Le metriche descritte nella sezione precedente vengono inizialmente calcolate a livello di singolo frame e successivamente aggregate su diverse scale temporali.

In particolare, i risultati vengono analizzati:

- a livello di segmento temporale, tra intervalli consecutivi dominati dal rumore,
- a livello di singola registrazione,

- sull'intero dataset.

Per ciascun livello di aggregazione vengono calcolate statistiche descrittive quali mediana, intervallo interquartile e distribuzioni empiriche delle metriche.

È importante sottolineare che i parametri dell'algoritmo rimangono invariati per tutte le registrazioni del dataset. Non viene effettuata alcuna ottimizzazione specifica per singolo file, in modo da garantire che le prestazioni osservate riflettano la robustezza intrinseca del metodo piuttosto che un adattamento specifico ai dati.

Questo approccio consente di valutare il comportamento del sistema su più scale temporali, analizzando sia la risposta istantanea dell'algoritmo sia la sua stabilità complessiva su registrazioni auscultatorie reali.

7. Risultati sperimentali

7.1 Esempio qualitativo di denoising in auscultazione reale

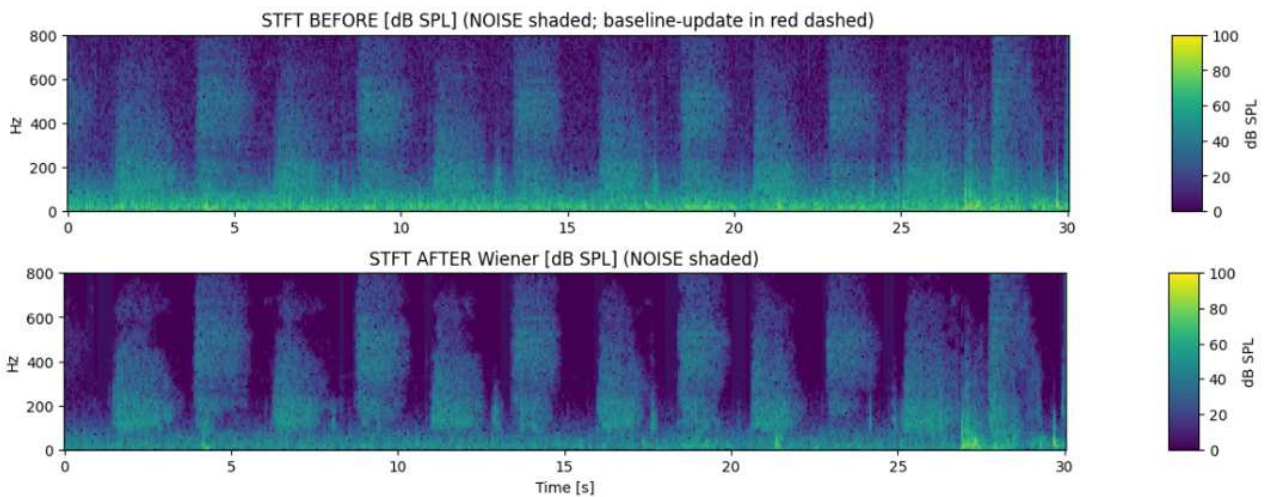


Figura 20: Esempio di denoising in auscultazione reale con il framework proposto

La Figura 20 mostra lo spettrogramma di una registrazione auscultatoria prima e dopo il denoising.

Nel segnale originale sono presenti componenti broadband, particolarmente evidenti durante le pause respiratorie, dove il contenuto è dominato dal rumore di fondo. Dopo l'elaborazione, tali componenti risultano significativamente attenuate, mentre la struttura spettrale dei suoni respiratori rimane riconoscibile.

In questo capitolo vengono presentati i risultati ottenuti applicando il framework al dataset descritto nel Capitolo 6. L'analisi combina osservazioni qualitative e metriche quantitative, con l'obiettivo di valutare l'efficacia della soppressione del rumore e il comportamento nella banda respiratoria.

7.2 Interpretazione del comportamento osservato

L'esempio qualitativo evidenzia la riduzione del contenuto broadband nelle pause respiratorie e la conservazione della struttura spettrale del segnale utile. Per quantificare tali effetti, vengono analizzate le metriche definite nel Capitolo 6, che permettono di valutare sia l'attenuazione del rumore sia la variazione del contrasto tra segnale respiratorio e background.

7.3 Setup sperimentale

Il framework è stato applicato all'intero dataset mantenendo invariati tutti i parametri dell'algoritmo. L'elaborazione è completamente automatica e non richiede annotazioni manuali dei segmenti.

7.4 Attenuazione del rumore di fondo (NSI)

L'attenuazione del rumore stazionario è stata valutata tramite il Noise Suppression Index (NSI), calcolato sui frame classificati come dominati dal rumore.

La Tab. 3 riporta le statistiche aggregate sull'intero dataset. Il valore mediano pari a circa -9 dB indica una riduzione consistente dell'energia broadband nelle regioni dominate dal rumore. La bassa dispersione dei valori evidenzia un comportamento stabile tra diverse registrazioni, nonostante l'assenza di ottimizzazione specifica per singola traccia.

Questi risultati indicano che il meccanismo di aggiornamento della baseline è in grado di adattarsi in modo consistente alle condizioni di rumore presenti nelle registrazioni.

Tabella 3: Prestazioni di attenuazione del rumore di fondo (NSI) sull'intero dataset.

Metrica	Mediana	Q1	Q3	Media	Dev. Std
NSI (dB)	-8.97	-9.18	-8.96	-9.08	0.38

7.4.1 Comportamento nella banda respiratoria

Il comportamento nella banda respiratoria è stato analizzato tramite il Respiratory Attenuation Index (RAI) e la metrica di contrasto ΔC .

Il valore mediano del RAI è circa -2 dB, indicando una moderata attenuazione anche nella banda polmonare. Questo effetto è coerente con la sovrapposizione spettrale tra segnale respiratorio e rumore, che rende inevitabile una riduzione parziale dell'energia utile durante la soppressione del disturbo.

La metrica ΔC assume valori prevalentemente positivi, con mediana pari a circa 6.9 dB, indicando un aumento del contrasto tra segmenti respiratori e segmenti dominati dal rumore.

La Fig. 21 riassume il comportamento delle due metriche, mostrando come a una moderata attenuazione energetica corrisponda un miglioramento della separazione tra segnale e rumore.

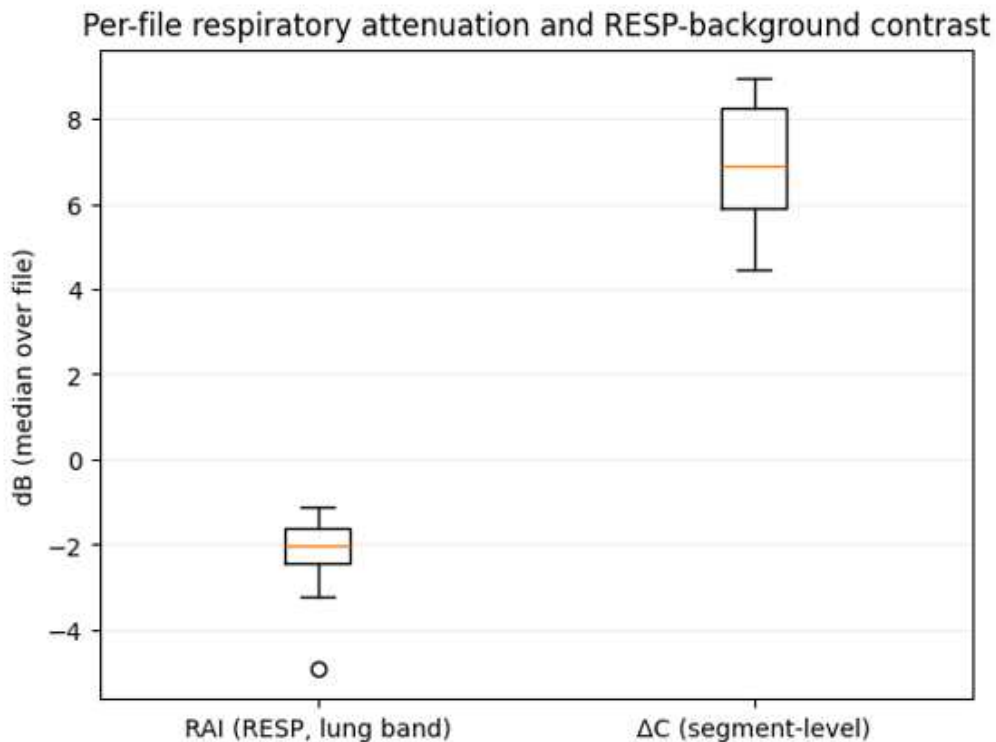


Figura 21: Confronto tra RAI e ΔC calcolati sull'intero dataset.

Per interpretare questo comportamento, la Fig. 22 mostra la distribuzione spettrale mediana dei frame RESP e NOISE.

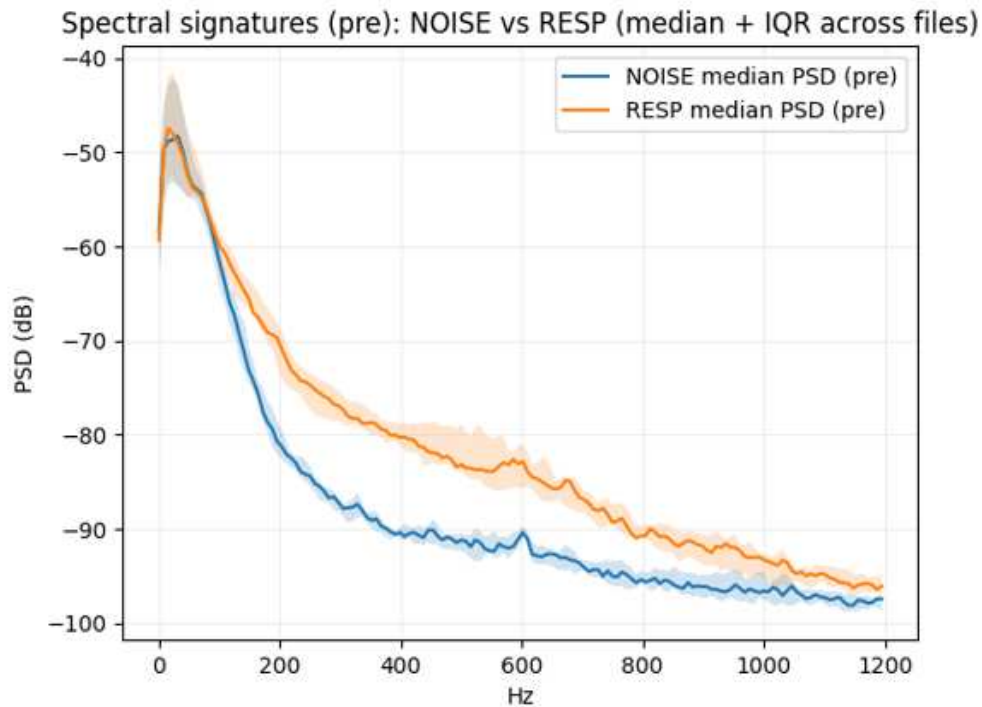


Figura 22: Distribuzione spettrale mediana dei frame RESP e NOISE (con intervallo interquartile).

I frame respiratori presentano un incremento energetico nella banda polmonare, mentre i frame di rumore mostrano una distribuzione più uniforme. Questa differenza statistica spiega il comportamento del sistema: nelle regioni in cui il segnale eccede la baseline, il guadagno rimane elevato, mentre nelle regioni dominate dal rumore viene ridotto.

7.5 Coerenza tra baseline e componente rimossa

La coerenza tra la componente rimossa e la baseline stimata è stata valutata tramite cosine similarity e correlazione di Pearson.

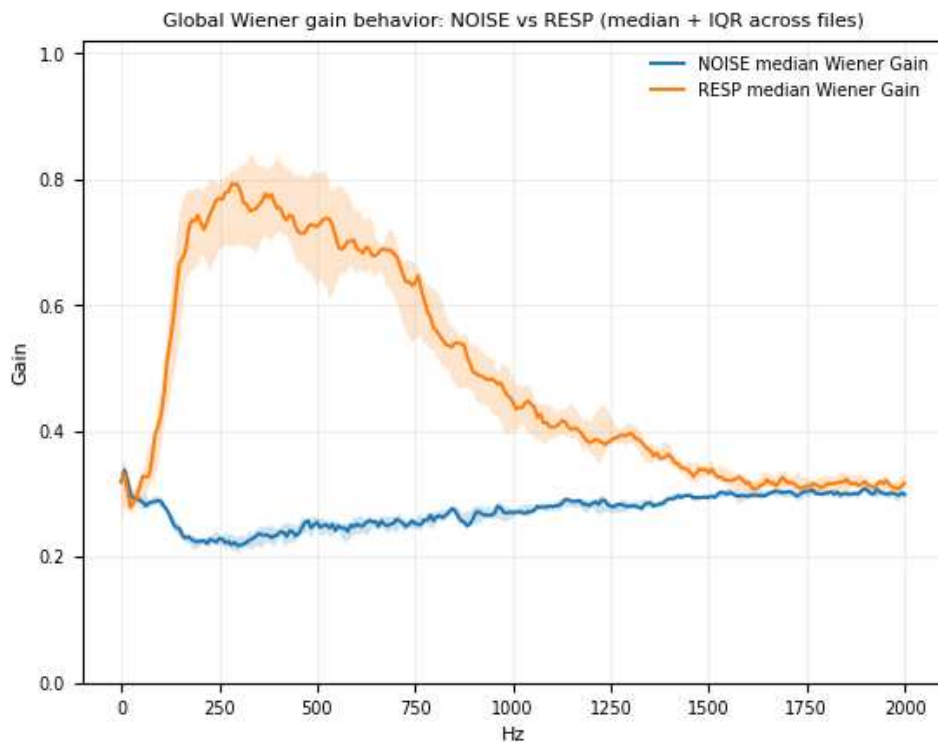
La Tab. 4 mostra valori mediani elevati (circa 0.93), indicando che l'energia attenuata è fortemente allineata con il modello di rumore stimato. Questo risultato conferma che la soppressione avviene principalmente sulle componenti coerenti con il disturbo.

Tabella 4: Similarità tra componente rimossa e baseline di rumore stimata.

Metrica	Mediana	Q1	Q3	Media	Dev. Std
Cosine Similarity	0.93	0.90	0.96	0.92	0.06
Pearson Correlation	0.94	0.90	0.96	0.92	0.06

7.6 Comportamento del guadagno di Wiener

La Fig. 23 mostra il guadagno del filtro di Wiener per frame NOISE e RESP.

**Figura 23:** Guadagno di Wiener in funzione della frequenza per frame NOISE e RESP.

Nei frame dominati dal rumore, il guadagno assume valori bassi su gran parte dello spettro. Nei frame respiratori, invece, risulta più elevato nella banda polmonare (100–800 Hz), dove l’energia del segnale supera la baseline.

Il profilo del guadagno varia in modo continuo in funzione del rapporto tra spettro osservato e rumore stimato, evidenziando il comportamento adattivo del framework.

7.7 Discussione dei risultati

I risultati sperimentali confermano che il framework proposto è in grado di operare efficacemente in condizioni reali di auscultazione, caratterizzate da rumore non stazionario e sovrapposizione spettrale tra segnale respiratorio e disturbo.

L'analisi tempo-frequenza evidenzia una riduzione significativa delle componenti broadband presenti nelle pause respiratorie, indicando che la stima della baseline di rumore risulta coerente con le caratteristiche statistiche del segnale nelle regioni identificate come NOISE. Questo comportamento è ulteriormente supportato dall'elevata similarità tra la componente spettrale rimossa e la baseline stimata, suggerendo che il processo di attenuazione è guidato da un modello di rumore adattivo e non da una soppressione indiscriminata delle componenti a bassa energia.

Dal punto di vista quantitativo, i valori dell'indice NSI mostrano una riduzione consistente del rumore di fondo sull'intero dataset, mentre l'indice RAI evidenzia una attenuazione contenuta della componente respiratoria. Questo compromesso indica che il filtro opera in una regione di funzionamento favorevole, nella quale la soppressione del disturbo non avviene a scapito della preservazione delle informazioni utili.

In questo contesto, l'incremento del contrasto tra frame respiratori e baseline di rumore, misurato tramite ΔC , rappresenta un risultato particolarmente rilevante. Più che massimizzare il rapporto segnale-rumore in senso assoluto, il framework migliora la separabilità tra le due classi di interesse (RESP e NOISE), facilitando sia l'interpretazione visiva dello spettrogramma sia eventuali successive fasi di analisi automatica.

L'analisi del guadagno di Wiener fornisce un'ulteriore chiave interpretativa del comportamento del sistema. Nei frame classificati come NOISE, il guadagno tende a valori prossimi allo zero su gran parte della banda, mentre nei frame RESP mantiene valori significativamente più elevati nelle regioni a maggiore energia spettrale. Questo conferma che il filtro non agisce come un semplice passa-banda, ma implementa una attenuazione selettiva dipendente dal contenuto spettrale locale e dalla stima adattiva del rumore.

Un aspetto centrale emerso dai risultati riguarda il ruolo dell'identificazione delle regioni di rumore. A differenza dei metodi basati su ipotesi di stazionarietà globale, il framework sfrutta informazioni temporali intrinseche al segnale respiratorio per aggiornare dinamicamente il modello di rumore. Questo approccio consente di seguire variazioni lente del contesto acustico e di mantenere prestazioni stabili anche in presenza di condizioni di acquisizione variabili.

Nel complesso, i risultati indicano che il metodo proposto si comporta come un meccanismo di miglioramento del contrasto tra segnale respiratorio e rumore di fondo, piuttosto che come un filtro progettato per la sola attenuazione energetica. Tale caratteristica risulta particolarmente rilevante in applicazioni auscultatorie, dove l'obiettivo principale non è la ricostruzione perfetta del segnale, ma la preservazione delle caratteristiche acustiche distintive utili alla diagnosi.

Infine, è importante sottolineare che l'intero framework è stato progettato secondo vincoli di causalità e complessità computazionale compatibili con implementazioni embedded. La natura frame-wise e l'assenza di operazioni che richiedano conoscenza futura del segnale rendono il metodo adatto a scenari real-time, rappresentando un ulteriore elemento di valore rispetto a soluzioni più complesse ma difficilmente trasferibili su dispositivi portatili.

8. Conclusioni

In questo lavoro è stato sviluppato e validato un framework di denoising adattivo per segnali respiratori acquisiti mediante stetoscopi digitali, basato su una formulazione Wiener guidata da regioni di rumore (*noise-region-driven*).

L'approccio proposto sfrutta le pause respiratorie per identificare automaticamente le regioni dominate dal rumore e aggiornare in modo adattivo una baseline spettrale del disturbo. Questa stima viene quindi utilizzata per applicare un filtraggio selettivo nel dominio tempo–frequenza, senza richiedere dati di riferimento puliti né procedure di addestramento.

I risultati sperimentali su registrazioni reali hanno evidenziato una riduzione consistente e stabile delle componenti broadband associate al rumore di fondo, con valori di attenuazione medi nell'ordine di circa -9 dB nelle regioni dominate dal rumore. Nonostante la presenza di una moderata attenuazione anche nella banda respiratoria, dovuta alla sovrapposizione spettrale tra segnale utile e rumore, è stato osservato un incremento del contrasto tra i frame respiratori e il background noise, quantificato tramite la metrica ΔC . Questo risultato indica che il framework non opera come un semplice filtro passa-banda, ma come un meccanismo adattivo in grado di aumentare la separazione tra contenuto respiratorio e disturbo di fondo, anche in presenza di riduzione energetica nella banda di interesse.

L'elevata coerenza tra la componente spettrale rimossa e la baseline di rumore stimata conferma inoltre che l'attenuazione è guidata da un modello statistico del rumore, derivato dagli aggiornamenti selettivi nei frame NOISE, e non da strategie euristiche o da filtri a risposta fissa.

Dal punto di vista implementativo, il framework è stato progettato per operare in modalità frame-wise con una complessità computazionale ridotta, risultando quindi compatibile con sistemi embedded e applicazioni in tempo reale su stetoscopi digitali.

Nel complesso, il lavoro dimostra come un approccio basato su stima adattiva del rumore e filtraggio di Wiener possa costituire una soluzione efficace, robusta e interpretabile per il miglioramento dei segnali auscultatori in condizioni realistiche.

8.1 Sviluppi futuri

Sebbene i risultati ottenuti siano promettenti, sono possibili diverse estensioni del lavoro.

In primo luogo, il framework attuale è stato progettato per la riduzione di rumore stazionario o lentamente variabile e non affronta esplicitamente artefatti impulsivi, come vibrazioni meccaniche o variazioni della pressione di contatto dello stetoscopio. Un'estensione naturale consiste nell'integrare moduli specifici per la rilevazione e la soppressione di tali artefatti.

Un ulteriore sviluppo riguarda la validazione su dataset clinici annotati, includendo segnali patologici, al fine di valutare l'impatto del denoising sulla rilevazione automatica di anomalie respiratorie.

Dal punto di vista metodologico, potrebbe essere interessante combinare l'approccio proposto con tecniche di apprendimento automatico, ad esempio utilizzando modelli di deep learning per stimare in modo più accurato la distribuzione del rumore o per adattare dinamicamente i parametri del filtro.

Infine, l'integrazione completa del framework all'interno di una piattaforma embedded per stetoscopi digitali rappresenta un passo fondamentale verso applicazioni reali in ambito clinico e di telemedicina, consentendo l'elaborazione e la trasmissione di segnali respiratori migliorati in tempo reale.

Bibliografia

- [1] H. Pasterkamp, S. S. Kraman e G. R. Wodicka, «Respiratory sounds: advances beyond the stethoscope», *American Journal of Respiratory and Critical Care Medicine*, 1997.
- [2] S. Reichert, R. Gass, C. Brandt e E. Andrès, «Analysis of respiratory sounds: state of the art», *Clinical Medicine: Circulatory, Respiratory and Pulmonary Medicine*, 2008.
- [3] R. X. A. Pramono et al., «Automatic adventitious respiratory sound analysis: a systematic review», *PLoS ONE*, 2017.
- [4] B. M. Rocha et al., «A Respiratory Sound Database for the Development of Automated Classification», in *Precision Medicine Powered by pHealth and Connected Health*, ser. IFMBE Proceedings, vol. 66, Springer, 2017, pp. 33–37. DOI: 10.1007/978-3-319-66275-0_55.
- [5] S. F. Boll, «Suppression of acoustic noise in speech using spectral subtraction», *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1979.
- [6] Y. Ephraim e D. Malah, «Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator», *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1984.
- [7] R. Murphy, «Computerized multichannel lung sound analysis», *Respiratory Care*, 2007.
- [8] S. Mallat, *A Wavelet Tour of Signal Processing*. gen. 1999, ISBN: 0-12-466606-X.
- [9] N. Huang et al., «The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis», *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 903–995, mar. 1998. DOI: 10.1098/rspa.1998.0193.
- [10] P. Loizou, *Speech Enhancement: Theory and Practice*. CRC Press, 2013.
- [11] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. MIT Press, 1949.
- [12] Y. Xu, J. Du, L.-R. Dai e C.-H. Lee, «A Regression Approach to Speech Enhancement Based on Deep Neural Networks», *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, pp. 7–19, gen. 2015. DOI: 10.1109/TASLP.2014.2364452.

- [13] S. Pascual, A. Bonafonte e J. Serrà, *SEGAN: Speech Enhancement Generative Adversarial Network*, 2017. arXiv: 1703.09452 [cs.LG]. indirizzo: <https://arxiv.org/abs/1703.09452>.
- [14] P. Vincent, H. Larochelle e P.-A. Manzagol, «Extracting and composing robust features with denoising autoencoders», gen. 2008, pp. 1096–1103. DOI: 10.1145/1390156.1390294.
- [15] F. Pancaldi et al., «VECTOR: An algorithm for the detection of COVID-19 pneumonia from velcro-like lung sounds», *Computers in Biology and Medicine*, vol. 142, p. 105220, 2022, ISSN: 0010-4825. DOI: <https://doi.org/10.1016/j.combiomed.2022.105220>. indirizzo: <https://www.sciencedirect.com/science/article/pii/S0010482522000129>.

Acknowledgments