



UNIVERSITÀ
DEGLI STUDI DI BARI
ALDO MORO



SERLAB
Software Engineering Research
UNIVERSITÀ DEGLI STUDI DI BARI



Sistemi e Tecnologie Industriali Intelligenti
per il Manifatturiero Avanzato
Consiglio Nazionale delle Ricerche

AN AUTOMATIC SYSTEM FOR ODONTOCETE WHISTLES DETECTION AND ENUMERATION

Relatore:

Prof. Giovanni Dimauro

Correlatori:

Dott.ssa Rosalia Maglietta

Dott.ssa Elena Papale

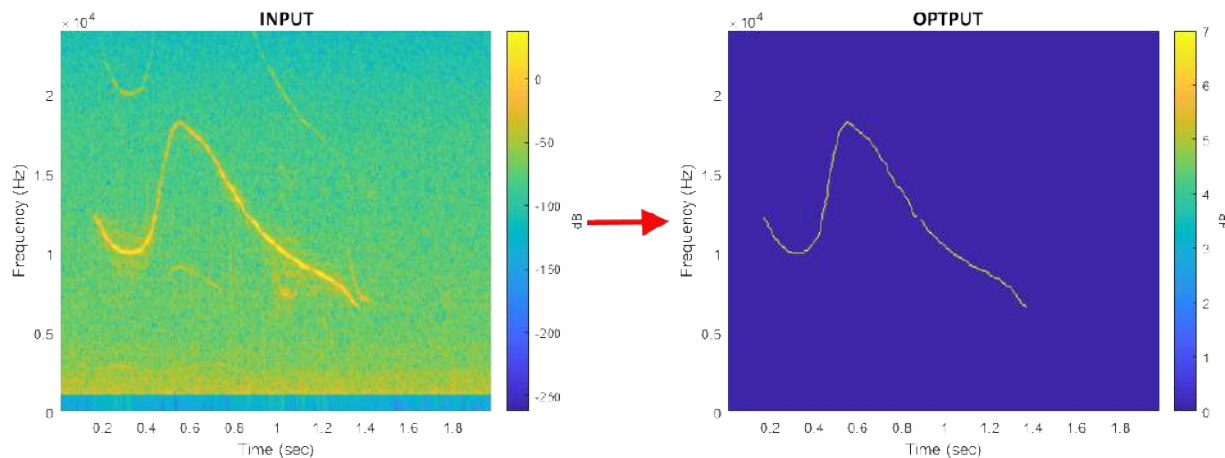
Laureando:

Ferrulli Vito

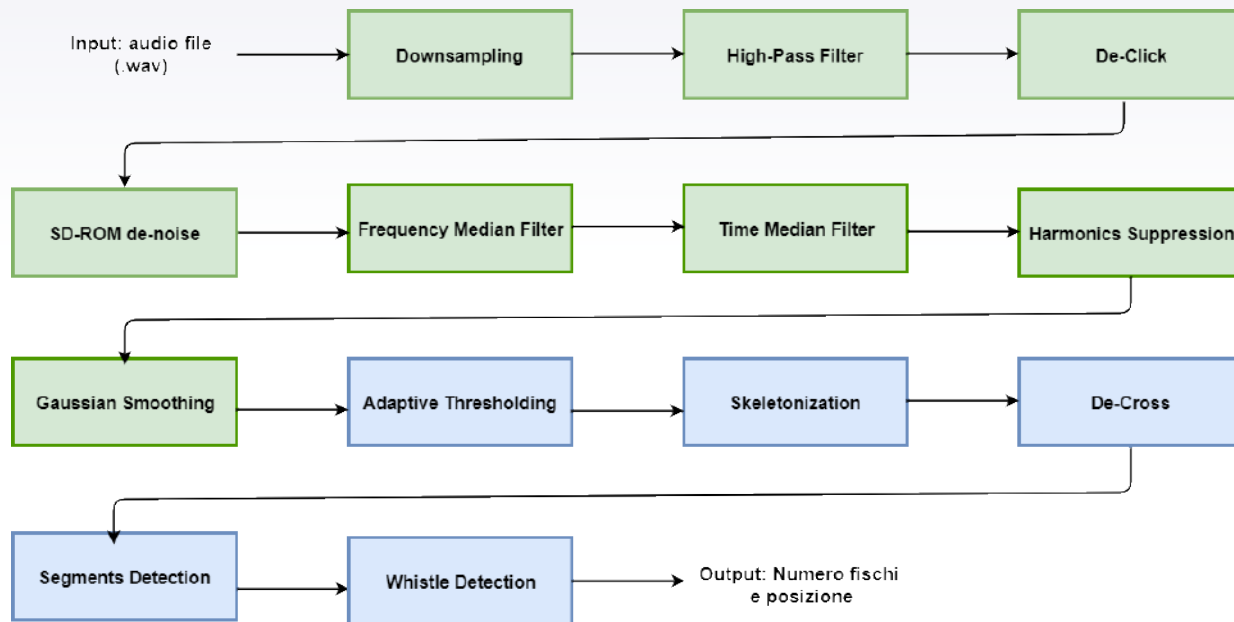
Obiettivo



Sviluppare un sistema software capace di rilevare ed estrarre fischi di odontoceti presenti in file audio attraverso l'analisi dello spettrogramma per **automatizzare l'estrazione dei fischi** da file audio di grandi dimensioni.



Software Pipeline

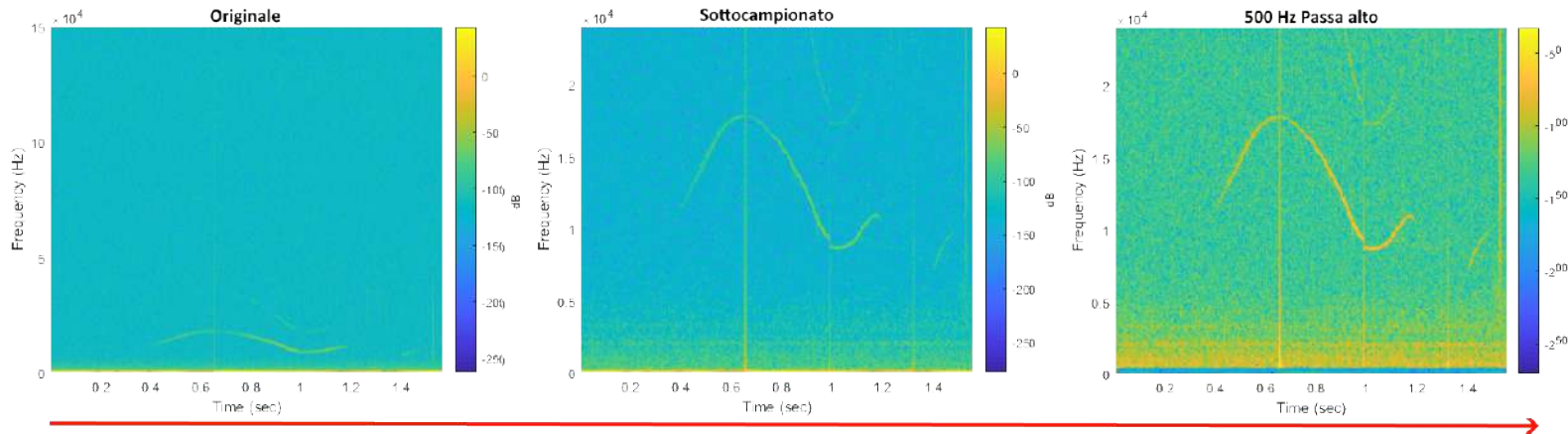


□ Preprocessing

□ Algoritmo SB-WDC

Preprocessing: Downsampling and High-pass

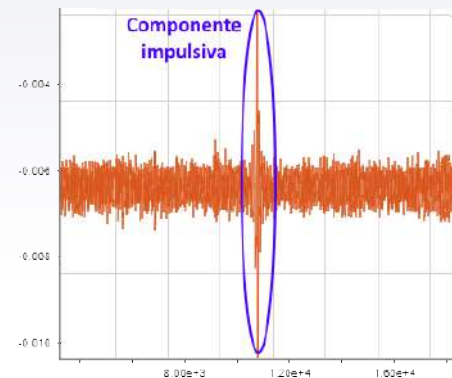
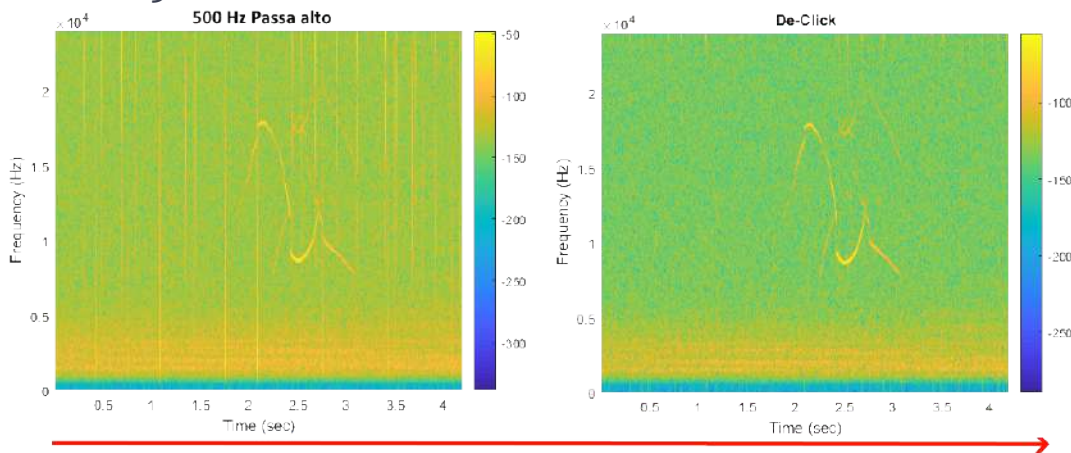
- ▶ **Downsampling** : sottocampionato del segnale a 24kHz in quanto l'armonica fondamentale dei fischi è ben contenuta nella banda di frequenza 1-24kHz
- ▶ **High-pass**: applicazione di un filtro passa alto a 500Hz per rimuovere l'intenso rumore presente alle basse frequenze e rendere i fischi più visibili



Preprocessing: De-click & SD-ROM de noise

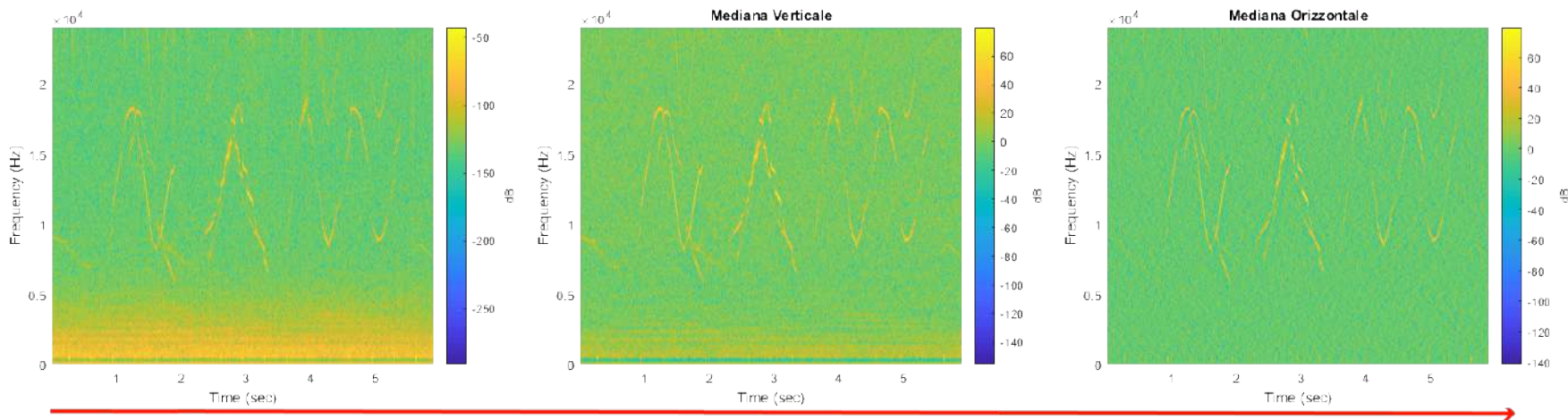
Algoritmi utilizzati per eliminare dal segnale **rumore impulsivo** che può presentarsi sotto forma di click di odontoceti o disturbi elettronici.

Questi algoritmi operano analizzando la distribuzione temporale dei campioni del segnale.



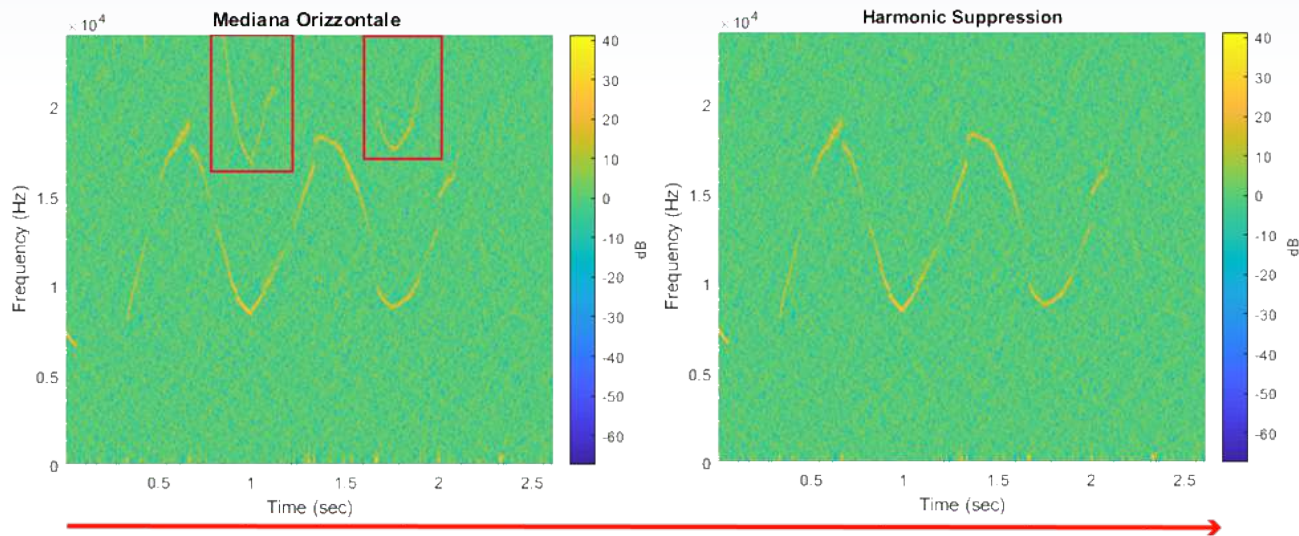
Preprocessing: Filtri mediana

Applicazione dei filtri mediana verticale (in **frequenza**) ed orizzontale (nel **tempo**) per omogenizzare il background e rimuovere componenti rumorose costanti verticalmente ed orizzontalmente.



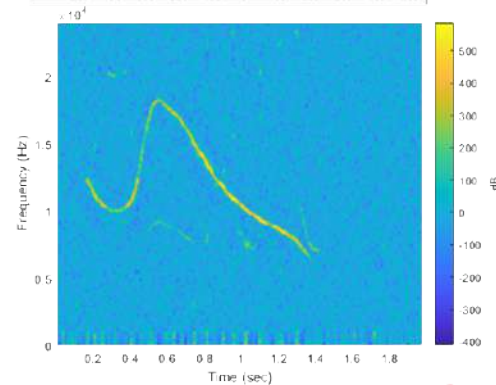
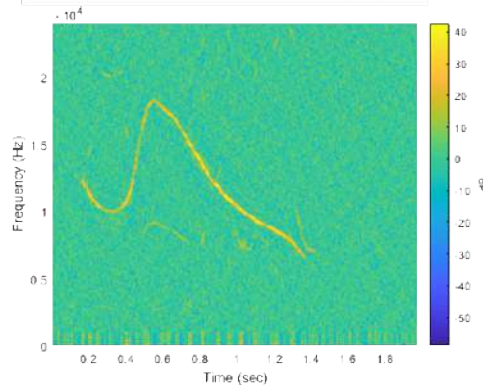
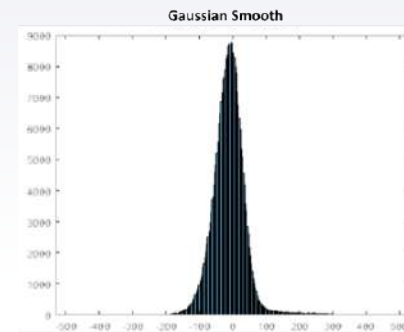
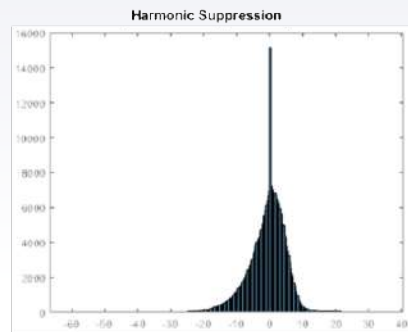
Preprocessing: Harmonic Suppression

Algoritmo utilizzato per **rimuovere le armoniche** della frequenza fondamentale del fischio.



Preprocessing : Gaussian Smooth

Smooth effettuato con un Kernel 3x3 e $\sigma=0.8$, utilizzato per separare ulteriormente i fischi dal background e normalizzare la **distribuzione delle intensità** per la fase di *Thresholding*.



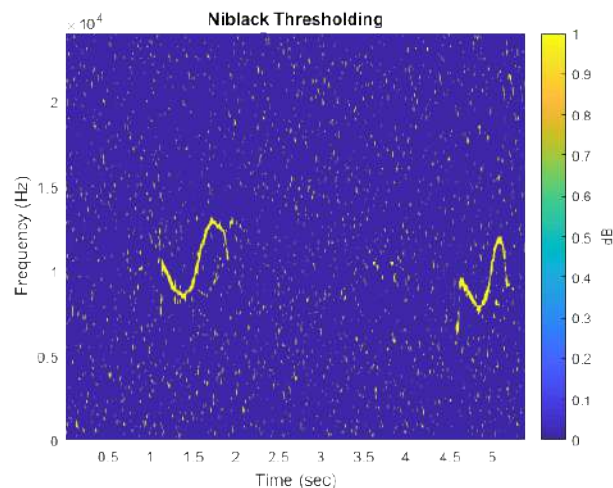
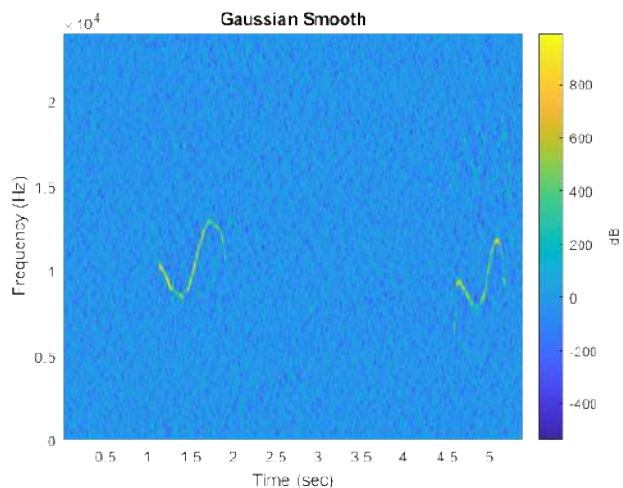
SB-WDC: Adaptive Niblack Thresholding

Utilizzato per separare i fischi dal rumore di fondo. Considera la distribuzione delle intensità come una normale, assumendo che **oltre il 90%** delle intensità appartengano al **background**.

$$T = \mu + k\sigma$$

$$\int_{-\infty}^{\mu+k\sigma} N_{\mu,\sigma}(x)dx = p$$

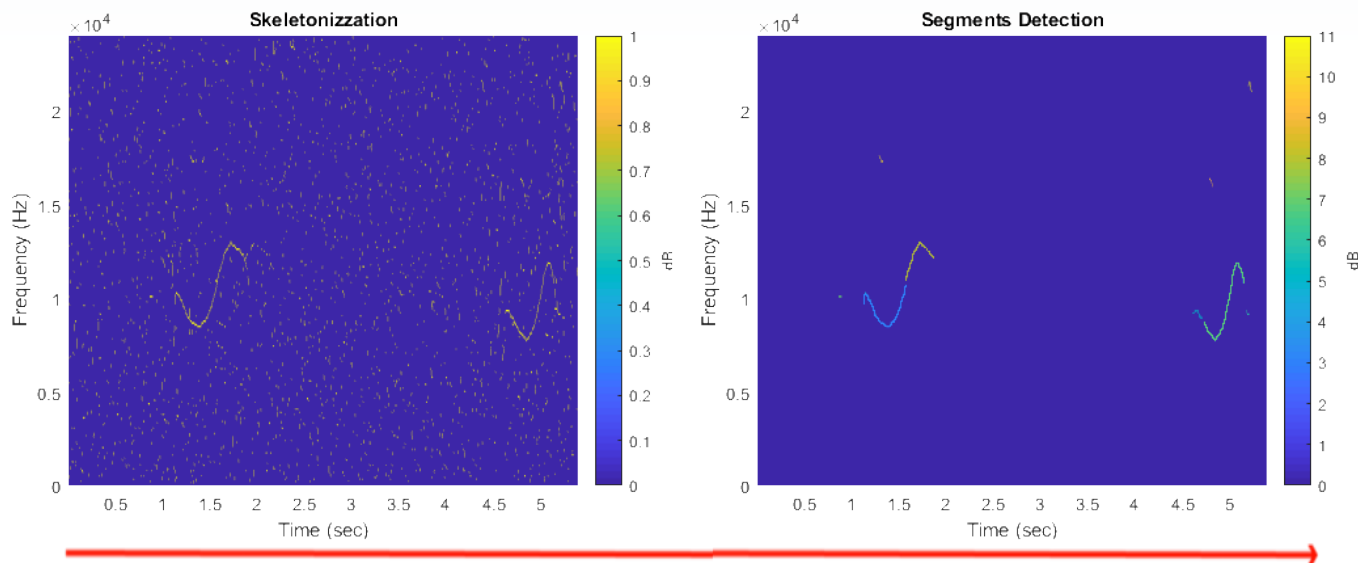
$$T = \text{norminv}(93, \mu, \sigma)$$



SB-WDC: Skeletonization & Segments Detection

I segmenti dei fischi vengono successivamente *scheletrizzati* e vengono identificate e salvate le **coordinate** di inizio e fine di ogni segmento.

Inoltre, si eliminano i segmenti dalla durata inferiore a 30 ms.

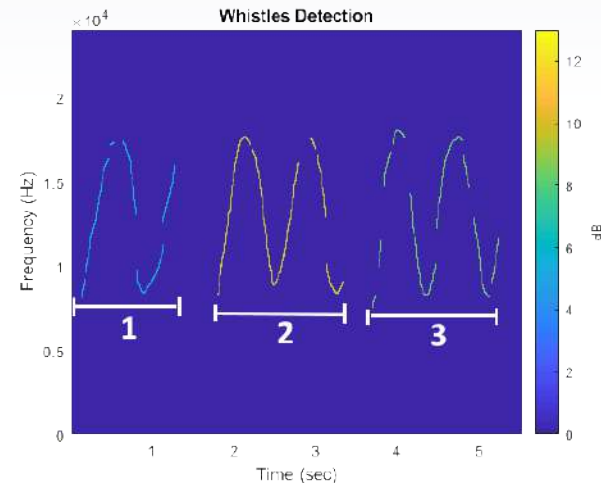
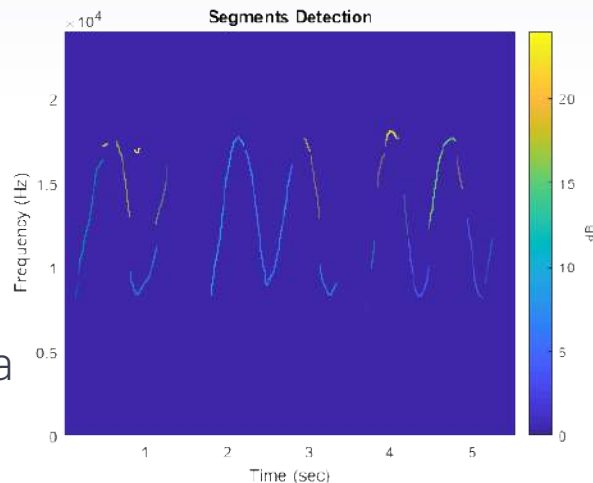


SB-WDC: Whistles Detection

I segmenti trovati vengono assegnati ai rispettivi fischi sulla base della loro **distribuzione temporale**.

Due segmenti appartengono allo stesso fischio se:

- ▶ La distanza fra i due è al più di 20 ms
- ▶ La differenza in frequenza è massimo 7KHz



Dataset

Il Dataset utilizzato è composto da **159 registrazioni** appositamente selezionate contenenti un totale di **307 fischi**.

Per ogni registrazione è stato annotato:

- ▶ Numero di fischi presenti
- ▶ Se i fischi sono sovrapposti (Booleano 0/1)
- ▶ Se ha un buon rapporto segnale rumore **SNR** (Booleano 0/1)

Risultati - 1° Test

Il **primo test** è stato effettuato sull'intero dataset differenziando le registrazioni contenenti fischi isolati e quelle contenenti fischi sovrapposti.

Fischi Sovrapposti	TP	FP	FN
77	60	11	17

PRECISION \cong 80%

RECALL \cong 75%

Fischi Isolati	TP	FP	FN
230	203	41	27

PRECISION \cong 79%

RECALL \cong 82%

Risultati - 2° Test

Il **secondo test** è stato effettuato sulle registrazioni **aventi un buon SNR** differenziando quelle contenenti fischi isolati e quelle contenenti fischi sovrapposti.

Fischi Sovrapposti	TP	FP	FN
44	35	5	9

PRECISION \cong 90%

RECALL \cong 79%

Fischi Isolati	TP	FP	FN
134	124	20	10

PRECISION \cong 93%

RECALL \cong 92%

Si denota aumento delle prestazioni del **~10%** con valori di **precision** simili in entrambi i casi ma con maggior differenza in **recall**.

Risultati - 3° Test

Il **terzo test** è stato effettuato sulle registrazioni **non aventi un buon SNR** differenziando quelle contenenti fischi isolati e quelle contenenti fischi sovrapposti.

Fischi Sovrapposti	TP	FP	FN
33	25	6	8

PRECISION \cong 80%

RECALL \cong 75%

Fischi Isolati	TP	FP	FN
96	79	21	17

PRECISION \cong 79%

RECALL \cong 82%

Come aspettato, si evince un calo delle prestazioni del **~10%** con valori di **precision** simili in entrambi i casi ma con maggior differenza in **recall**.

Sviluppi futuri

- ▶ Migliorare le performances generali del sistema
- ▶ Migliorare le performances nel caso di fischi sovrapposti
- ▶ Quantizzare la precisione con cui i fischi vengono tracciati
- ▶ Riconoscere i falsi positivi
- ▶ Ricostruire i fischi danneggiati dal rumore



GRAZIE PER L'ATTENZIONE