



**UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA**



**DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE**

**CORSO DI LAUREA IN INGEGNERIA INFORMATICA**

**INTELLIGENZA ARTIFICIALE NEL MONITORAGGIO ACUSTICO  
PASSIVO**

**Relatore: Prof. Loris Nanni**

**Laureanda: Ioana Munteanu**

**ANNO ACCADEMICO 2021 – 2022**

**Data di laurea 14/03/2022**



# Abstract

In questo documento si analizza un metodo per la classificazione delle specie tramite il suono. L'idea prevede l'utilizzo di un software che elabora lo spettrogramma del suono al fine di ricavare più informazioni possibili.

Il monitoraggio acustico passivo (Passive Monitoring Acoustic) permette di osservare l'ecosistema marino e la localizzazione degli animali mammiferi. Grazie all'individuazione della posizione geografica si possono evitare le collisioni fra balene e navi. Lo studio del habitat e delle abitudini consente di controllare la salute degli animali.

Lo scopo di questo elaborato è descrivere come può essere utilizzata la tecnica PAM nell'ambito marino attraverso dei filtri. I filtri PAM sono stati creati come protocolli sperimentali con lo scopo di sfruttare al meglio tutti i campioni disponibili e studiandone al meglio le prestazioni del sistema.





# Indice

Abstract .....	3
Indice.....	6
Introduzione .....	7
Capitolo 1.....	8
1.1 I mammiferi marini .....	8
1.2 Raccolta dati .....	9
1.3 Il ruolo di Machine Learning .....	10
Capitolo 2.....	13
2.1 Filtro.....	13
2.2 Protocolli Sperimentali Watkins .....	13
2.2.1 Protocollo sperimentale Watkins 1 .....	14
2.2.2 Protocollo sperimentale Watkins 2 .....	15
2.2.3 Protocollo sperimentale Watkins 3 .....	17
2.2.4 Caratteristiche .....	17
Capitolo 3.....	21
3.1 Esperimento numero 1 .....	21
3.2 Esperimento numero 2 .....	24
3.3 Conclusioni .....	25
Bibliografia .....	27

# Introduzione

Le località in cui sono state registrati i suoni marini sono in quantità minore rispetto al numero di campioni disponibili per cui diventa rilevante la suddivisione dei suoni in sottoinsiemi.

Uno fra i ruoli fondamentali è la rappresentazione in immagine di un segnale audio, ovvero lo spettrogramma ed estrarre più informazioni possibili.

Nella prima parte di questo elaborato vengono descritti i mammiferi marini, come avviene la raccolta di informazioni ed il ruolo dell'intelligenza artificiale.

Successivamente vengono esposti tre protocolli pensati per la suddivisione dei dati a tal fine di non avere lo stesso suono dell'individuo sia nell'addestramento e sia nel test

L'estrapolazione delle informazioni sono state ricavate tramite il software Matlab, tutti i dati sono stati salvati in un file con estensione "mat" per poi essere caricato nella fase di addestramento e test.

L'addestramento è stato eseguito tramite reti neurali convoluzionali, in particolare Alexnet.

In fine viene discusso le performance ottenute e le modifiche eseguite in base ai risultati.

# Capitolo 1

## 1.1 I mammiferi marini

Gli esseri viventi sono molto importanti per il corretto funzionamento dell'ecosistema terrestre, ogni anno ci sono sempre più animali in via di estinzione, come ad esempio le balene. Una fra le maggiori cause è il riscaldamento globale, i cambiamenti climatici provocano la mutazione del habitat naturale e degli esseri viventi, la temperatura di anno in anno aumenta e ci sono sempre meno ghiacciai. Questo è uno dei motivi per cui le specie migrano in altre zone geografiche "insolite"; tuttavia, non c'è solo un cambiamento geografico ma anche fisico, infatti, il corpo degli esseri viventi si adatta a sopravvivere ad alte temperature.

Fra le altre cause c'è l'inquinamento marino, causato dal petrolio e anche dalla plastica, la prima è molto tossica per l'ambiente, la seconda può essere responsabile dell'incagliamento tra le reti degli animali, o ancor peggio, può essere causa di malformazioni fisiche. Un'altra minaccia è la pesca industriale, nel passato e per molto tempo, le balene erano animali sconosciuti e molto affascinanti, purtroppo l'uomo ha eseguito delle competizioni per la loro cattura, contribuendone all'estinzione di alcune specie.

Secondo il comitato di tassonomia di The Society for Marine Mammalogy [1], i mammiferi marini sono classificati in tre ordini e diverse famiglie: generi, specie e sottospecie; alcuni di loro potrebbero essere già estinti. La ricerca effettuata su questi animali è ardua, proprio perché vivono nell'acqua la maggior parte del tempo, questo rende difficile la raccolta dei dati, come ad esempio una vasta gamma di immagini per ciascun individuo e le registrazioni audio dei "gemiti".

Dalle analisi di Sayigh [2] è emerso che i delfini con il naso a forma di bottiglia utilizzano i fischi per identificarsi univocamente rispetto agli altri individui; lo stesso Sayigh ha ammesso che questo non accade con altre registrazioni [2]. Sebbene consideriamo la stessa specie [3] il ruolo del suono è complesso da definire nell'ambito sociale e delle relazioni fra gli animali, un'altra specie, i delfini gobbi, è stata osservata per studiarne la classificazione delle registrazioni audio per determinare il comportamento e la distribuzione in base a quest'ultimo [4].



Oltre al mondo dei mammiferi, un altro mondo interessante è lo studio dei pesci, per esempio nelle lagune o in acqua dolce. Fra i pesci non vi sono una serie di suoni standard, ma un'elevata varietà di suoni ed in quantità minore, tra determinate specie. Spesso e volentieri sono diversi tant'è che [5] afferma che vi è presente un "dialetto" fra i pesci.

## 1.2 Raccolta dati

Il monitoraggio acustico passivo (Passive Acoustic Monitoring) è una tecnica che permette l'osservazione, in particolare acustica, delle specie in modo passivo per cui non invade l'ambiente; ci permette di raccogliere dati in "totale" sicurezza ambientale. L'osservazione acustica è preferibile in alcune situazioni, perché i mammiferi marini spendono la maggior parte del tempo sott'acqua per cui è difficile ottenere delle immagini.

Le registrazioni audio dei mammiferi sono memorizzate e disponibili nel database "Watkins Marine Mammal Sound Database", indicativamente i dati sono stati registrati dal 1940 in poi, fino agli anni 2000. Nella base di dati ci sono 1600 interi nastri, ognuno è composto da diversi minuti di audio in cui possono sentirsi differenti specie, identificate dai biologi marini. Questi nastri sono stati suddivisi in tagli, il quale contiene il suono di un solo individuo.

La qualità del suono non è sempre ottimale, oltre all'audio dell'animale può esserci il rumore presente nell'ambiente oppure può verificarsi un non corretto funzionamento del dispositivo, dunque, la qualità si abbassa.

Ci sono principalmente due tipologie di tagli: tagli di qualità migliore ed i rimanenti che sono di un livello minore ma che contribuiscono a formare l'intera gamma a disposizione. I primi sono detti di qualità maggiore poiché hanno minori disturbi, i secondi presentano interferenze dovute all'ambiente circostante, come ad esempio le condizioni meteorologiche o come anticipato precedente, guasti sui sensori dei dispositivi di rilevazione acustica.

Ci sono 1694 file di qualità alta tra cui sono state identificate 32 specie di mammiferi marini. Alcuni di questi campioni contenevano più di una specie, e quindi sono stati rimossi. Collegandosi al sito web, dove sono reperibili le registrazioni, oltre a ciascun taglio audio, sono anche disponibili i relativi metadati; per esempio, la posizione geografica, la data di registrazione e la specie marina.

Dai dati presenti si può sottolineare che la maggior parte delle specie è stata registrata quasi sempre nella stessa località geografica.

Lo scopo è addestrare un sistema intelligente sulla base dei dati a disposizione, in navigazione il segnale acustico viene registrato tramite un dispositivo, installato a bordo dell'imbarcazione, ed è in grado di trasmettere i dati tramite una connessione attraverso la rete internet. Questo segnale viene trasformato dal sistema intelligente in uno spettrogramma, dopo l'elaborazione viene restituita come risposta alla trasmissione, la specie corrispondente al suono. I spettrogrammi sono rappresentati come la frequenza del segnale in funzione del tempo.

Il motivo per cui vengono utilizzati gli spettrogrammi è dovuto al fatto che la registrazione dura pochissimi secondi, in alcuni casi può durare anche meno di un secondo. Da un audio digitale, è possibile ottenere uno spettrogramma utilizzando la trasformata discreta di Fourier (DFT).

### **1.3 Il ruolo di Machine Learning**

Machine learning gioca un ruolo fondamentale nel riconoscimento della specie marina, uno dei punti cruciali legati all'audio è il rumore, quest'ultimo può portare ad ovviare la classificazione e allo stesso tempo ribaltarne il risultato. Infatti, prima di addestrare il sistema si implementano dei metodi che tolgono il rumore. Un altro punto cruciale è l'addestramento del sistema, per far sì che il sistema non impari "a memoria" a riconoscere l'individuo, ma bensì la specie, bisogna evitare di avere più volte lo stesso animale mammifero sia nella fase di addestramento che nella fase di test. Una delle minacce per le balene è la collisione con le navi; infatti, nel 2013 è stata lanciata una sfida internazionale [6] per identificare le balene franche del nordatlantico. Essa consisteva nell'identificare automaticamente se un audio conteneva o meno le registrazioni di tale specie. Questo tipo di riconoscimento permetterebbe alle navi di cambiare rotta per evitare una possibile collisione. I dati per questa sfida sono stati raccolti da boe galleggianti, tramite un particolare

dispositivo al loro interno. Le balene franche sono fra le specie in pericolo, tant'è che nel 2015 è stata proposta un'altra competizione [7], in quest' ultimo caso il database si componeva di immagini scattate tramite degli elicotteri per giungere allo stesso obbiettivo di aiutare e tenere traccia della salute e dello stato generale degli individui.



# Capitolo 2

## 2.1 Filtro

Definiamo due insiemi principali, uno è l'insieme contenente i campioni per allenamento e l'altro l'insieme di test. Nella fase di test viene valutata l'accuratezza, ovvero il numero di pattern che il nostro sistema riesce a riconoscere sul totale del numero di pattern dell'insieme test.

Un "pattern" rappresenta un campione di tale registrazione audio, il termine "classe" rappresenta la specie dell'individuo.

Dagli esperimenti fatti è emerso che il sistema ha raggiunto dei risultati alti in corrispondenza della suddivisione casuale in due insiemi, rispettivamente l'addestramento ed il test.

Dai metadati del WMMSD possiamo notare che ci sono molti campioni degli stessi individui, inoltre lo stesso modello di rumore era presente in diversi campioni della stessa specie.

Nella classificazione del genere musicale, Flexer [9] ha definito il concetto di "artista – filtro", che significa avere tutti i campioni dello stesso artista, sia nell'addestramento che nel test. Le prestazioni ottenute erano molto alte e con deviazioni standard inferiori; questo significa che il sistema impara a "memoria" a riconoscere l'artista invece del genere musicale.

Il concetto di "filtro – artista" ha portato alla nascita dei filtri PAM [9], ovvero lo stesso individuo compare nell'insieme di addestramento o nell'insieme di test, ma mai in entrambi.

## 2.2 Protocolli Sperimentali Watkins

Per studiare i filtri PAM al meglio ed il loro impatto, sono stati creati tre diversi protocolli dello stesso database, con e senza i filtri. Il protocollo WEP#1 (Watkins Experimental Protocols) non dispone di filtri PAM applicati, mentre il protocollo WEP#2 e WEP#3 si. Va sottolineato che è stato cercato di utilizzare il maggior numero di campioni e classi disponibili.

## 2.2.1 Protocollo sperimentale Watkins 1

Il primo protocollo consiste nella divisione del database in modo casuale in 10 parti per la convalida incrociata, tutte le classi con meno di 10 campioni sono state rimosse.

Nella figura 1 sono rappresentate le specie del database: la prima colonna è rappresentativa delle specie, la seconda colonna del numero delle posizioni geografiche in cui vi sono state le registrazioni, la terza colonna rappresenta il numero di campioni per ogni posizione geografica, mentre la quarta colonna rappresenta il numero totale di campioni per quella determinata specie.

Species	Number of Locations	Samples by Location	Number of Samples in WEP#1
<i>Balaena mysticetus</i>	2	1; 49	50
<i>Balaenoptera acutorostrata</i>	1	17	17
<i>Balaenoptera physalus</i>	2	5; 45	50
<i>Delphinapterus leucas</i>	4	1; 6; 16; 27	50
<i>Delphinus delphis</i>	3	2; 15; 35	52
<i>Erignathus barbatus</i>	3	3; 9; 15	27
<i>Eubalaena australis</i>	2	7; 18	25
<i>Eubalaena glacialis</i>	4	3; 12; 19; 20	54
<i>Globicephala macrorhynchus</i>	4	5; 16; 18; 26	65
<i>Globicephala melas</i>	4	11; 12; 14; 28	65
<i>Grampus griseus</i>	3	1; 21; 45	67
<i>Hydrurga leptonyx</i>	1	10	10
<i>Lagenodelphis hosei</i>	1	87	87
<i>Lagenorhynchus acutus</i>	3	12; 12; 31	55
<i>Lagenorhynchus albirostris</i>	2	20; 37	57
<i>Megaptera novaeangliae</i>	3	1; 17; 46	64
<i>Monodon monoceros</i>	3	4; 10; 36	50
<i>Odobenus rosmarus</i>	3	1; 16; 21	38
<i>Ommatophoca rossi</i>	3	11; 19; 20	50
<i>Orcinus orca</i>	5	1; 2; 5; 8; 19	35
<i>Pagophilus groenlandicus</i>	1	47	47
<i>Peponocephala electra</i>	1	56	56
<i>Physeter macrocephalus</i>	6	2; 2; 2; 9; 12; 33	60
<i>Pseudorca crassidens</i>	2	11; 48	59
<i>Stenella attenuata</i>	2	11; 54	65
<i>Stenella clymene</i>	2	14; 49	63
<i>Stenella coeruleoalba</i>	4	8; 12; 27; 34	81
<i>Stenella frontalis</i>	1	58	58
<i>Stenella longirostris</i>	2	1; 113	114
<i>Steno bredanensis</i>	1	50	50
<i>Tursiops truncatus</i>	3	1; 10; 13	24
Number of samples			1645

Figura 1 le registrazioni audio per ciascuna specie presente nel database Watkins Marine Mammal Sound Database

### 2.2.2 Protocollo sperimentale Watkins 2

Alcune specie, per esempio la “Balaenoptera acutorostrata” ha tutti i suoni registrati nella stessa posizione geografica. Se l’audio contiene del rumore generato dall’ambiente circostante o dal dispositivo stesso, significa che la maggior parte dei dati è affetta da rumore. Tali specie sono state rimosse in questo protocollo per la problematica appena citata.

Per il secondo protocollo sperimentale utilizzando i filtri PAM, sono stati scansionati i metadati dei file per suddividere il database in due insiemi corrispondenti alle posizioni geografiche. Per ciascuna specie tutte le registrazioni audio in una determinata località geografica sono assegnate esclusivamente al primo insieme oppure al secondo.

Il primo insieme corrisponde al gruppo di addestramento mentre il secondo a quello di test, quest’ultimo serve a convalidare le prestazioni del sistema.

Nel caso in cui le registrazioni sono avvenute in sole due località, come ad esempio la specie “Stenella attenuata”, quella col maggior numero di registrazioni i cui campioni vengono assegnati all’insieme di addestramento, invece i restanti all’insieme di test.

Nel caso in cui vi siano più di due località, le registrazioni audio sono state distribuite nei due insiemi, tale che un 70% delle registrazioni viene assegnato all’ambiente di addestramento ed il restante 30% a quello di test.

Nella figura 2 viene illustrato il numero di campioni fra tutte le specie utilizzati per questo protocollo ed il successivo.

Species	Number of Samples			
	Train	WEP#2	Test	WEP#3
<i>Balaena mysticetus</i>	49		1	-
<i>Balaenoptera acutorostrata</i>	-		-	-
<i>Balaenoptera physalus</i>	45		5	-
<i>Delphinapterus leucas</i>	27		23	27
<i>Delphinus delphis</i>	35		17	35
<i>Erignathus barbatus</i>	15		12	15
<i>Eubalaena australis</i>	18		7	-
<i>Eubalaena glacialis</i>	31		23	31
<i>Globicephala macrorhynchus</i>	34		31	34
<i>Globicephala melas</i>	37		28	37
<i>Grampus griseus</i>	45		22	45
<i>Hydrurga leptonyx</i>	-		-	-
<i>Lagenodelphis hosei</i>	-		-	-
<i>Lagenorhynchus acutus</i>	31		24	31
<i>Lagenorhynchus albirostris</i>	37		20	37
<i>Megaptera novaeangliae</i>	46		18	46
<i>Monodon monoceros</i>	36		14	36
<i>Odobenus rosmarus</i>	21		17	21
<i>Ommatophoca rossi</i>	30		20	30
<i>Orcinus orca</i>	19		16	19
<i>Pagophilus groenlandicus</i>	-		-	-
<i>Peponocephala electra</i>	-		-	-
<i>Physeter macrocephalus</i>	33		27	33
<i>Pseudorca crassidens</i>	48		11	48
<i>Stenella attenuata</i>	54		11	54
<i>Stenella clymene</i>	49		14	49
<i>Stenella coeruleoalba</i>	42		39	42
<i>Stenella frontalis</i>	-		-	-
<i>Stenella longirostris</i>	113		1	-
<i>Steno bredanensis</i>	-		-	-
<i>Tursiops truncatus</i>	13		11	13
Sums	908		412	542

Figura 2 è rappresentata la suddivisione in insiemi del protocollo WEP#2 e WBP#3

Le specie escluse da questo protocollo sono *Balaenoptera acutorostrata*, *Hydrurga leptonyx*, *Lagenodelphis hosei*, *Pagophilus groenlandicus*, *Peponocephala electra*, *Physeter macrocephalus*, *Stenella coeruleoalba* ed *Steno bredanensis*.

Per un totale di 24 specie, ovvero 24 classi con 908 campioni per l'insieme di addestramento e 412 per quello di test.



### **2.2.3 Protocollo sperimentale Watkins 3**

Il terzo e ultimo protocollo prevede una doppia convalida incrociata. È stato creato per usare tutti i campioni per l'addestramento o per il test, migliorandone l'affidabilità dei risultati.

Il terzo protocollo sperimentale si differisce dal secondo perché sono state rimosse tutte le classi per cui per ciascuna località vi sono meno di 10 campioni.

Ad esempio la specie *Balaenoptera physalus* in una località ci sono 5 campioni ed in un'altra 45 campioni per cui è stata rimossa da questo protocollo.

Nel protocollo WEP#3 ci sono in totale 20 classi di cui 542 per l'addestramento e 539 per il test.

### **2.2.4 Caratteristiche**

L'estrapolazione delle caratteristiche viene fatta attraverso lo spettrogramma, esso ci permette di visualizzare il segnale in base al tempo, alla frequenza e all'intensità del segnale. La texture visiva è stata usata in diversi ambiti di classificazione come l'identificazione del genere musicale [10], identificare delle specie di uccelli [11] e delle balene [12].

Nella sottostante Figura 4 si può osservare gli spettrogrammi di due differenti specie di balene.

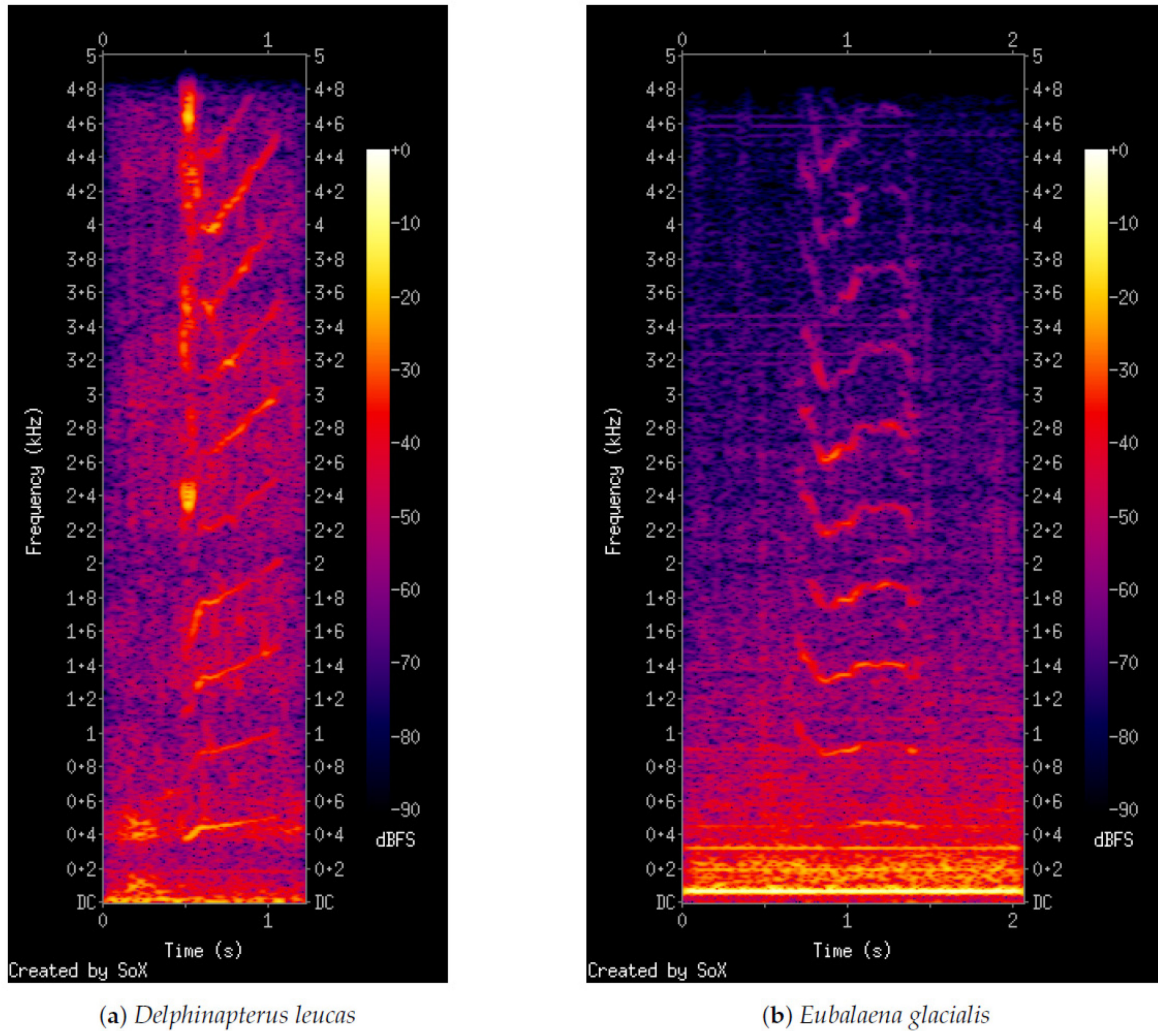


Figura 4 Rappresentano gli spettrogrammi di due balene.

In [13] gli autori hanno proposto di utilizzare il Local Binary Pattern (LBP). La trama di un'immagine viene descritta tramite un istogramma; ogni cella rappresenta il numero di occorrenze di un determinato valore.

Il Local Binary Pattern viene calcolato nel seguente modo:

$$LBP_{P,R} = \sum_{p=0}^{p-1} s(v_p - v_c) 2^p$$

Il parametro P rappresenta il numero di pixel totali, R rappresenta il numero di pixel come raggio, di solito P vale 8 e R vale 2,  $v_p$  rappresenta il valore del pixel p ed  $v_c$  rappresenta il valore del pixel centrale.

La funzione  $s$  vale 1 oppure 0 ed è definita nel seguente modo:

$$s(x) = \begin{cases} 1, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases}$$

Per chiarire le formulazioni precedenti verrà svolto un esercizio:

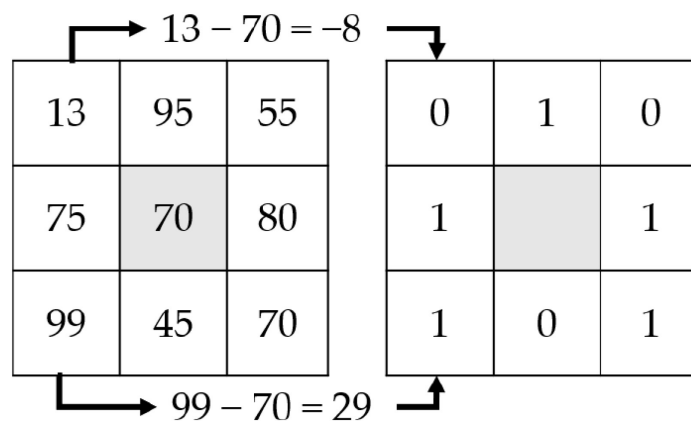


Figura 5 Il riquadro di destra sono 8 pixel ed ognuno col suo valore

Nella figura 5 ci sono 8 pixel, il pixel centrale vale 70, per ciascun pixel diverso da quello centrale si applica la funzione  $s$  (definita sopra) della differenza del valore del pixel più in alto a sinistra ed il valore del pixel centrale, e così via in senso orario; se la differenza è maggiore di 0 allora la funzione  $s$  vale 1 altrimenti 0.

Il Local Binary Pattern del pixel centrale è la somma di ciascuna coppia moltiplicata per 2 elevato al numero del pixel  $i$ -esimo. Il valore del LBP in questo esercizio viene calcolato:

$$\begin{aligned} \sum_{p=0}^{p-1} s(v_p - v_c)2^p &= \\ &= s(13 - 70)2^0 + s(95 - 70)2^1 + s(55 - 70)2^2 + s(80 - 70)2^3 + s(70 - 70)2^4 \\ &\quad + s(45 - 70)2^5 + s(99 - 70)2^6 + s(75 - 70)2^7 \\ &= 0 \times 2^0 + 1 \times 2^1 + 0 \times 2^2 + 1 \times 2^3 + 0 \times 2^4 + 0 \times 2^5 + 1 \times 2^6 + 1 \times 2^7 = 202 \end{aligned}$$

Il valore calcolato è 202, al variare di P cambiamo le possibili soluzioni, in questo caso con 8 pixel vi sono  $2^8$ , ovvero 256 possibili soluzioni. Dopo aver calcolato LBP per tutti i pixel otteniamo un istogramma con il numero di occorrenze per ciascun valore.

## Capitolo 3

### 3.1 Esperimento numero 1

Ho scelto di sperimentare il protocollo WEP#3, come è stato detto precedentemente sono stati divisi i file delle registrazioni in due insiemi. Per questo protocollo ho seguito un file di testo preimpostato in questo modo:

```
CLASS;SAMPLE;FOLD
Megaptera_novaeangliae;55113001;0
Megaptera_novaeangliae;55113002;0
Megaptera_novaeangliae;55113003;0
Eubalaena_glacialis;5601601O;0
Eubalaena_glacialis;56017007;0
Eubalaena_glacialis;56017008;0
..
Erignathus_barbatus;78018005;1
Erignathus_barbatus;7801800M;1
Lagenorhynchus_acutus;79004007;1
Lagenorhynchus_acutus;79006018;1
Lagenorhynchus_acutus;79006019;1
Megaptera_novaeangliae;8002001M;1
Globicephala_macrorhynchus;81037001;1
Globicephala_macrorhynchus;81037002;1
Globicephala_macrorhynchus;81037004;1
Globicephala_macrorhynchus;81037005;1
..
```

In questo protocollo ci sono 20 specie utilizzate, di conseguenza l'addestramento ed il test viene effettuato su 20 classi, per un totale di 1081 campioni.

Le seguenti specie sono state utilizzate in questa sperimentazione: *Megaptera novaeangliae*, *Eubalaena glacialis*, *Globicephala macrorhynchus*, *Stenella coeruleoalba*, *Grampus griseus*, *Lagenorhynchus acutus*, *Delphinapterus leucas*, *Stenella attenuata*, *Ommatophoca rossi*, *Erignathus barbatus*, *Lagenorhynchus albirostris*, *Odobenus rosmarus*, *Delphinus delphis*, *Monodon monoceros*, *Globicephala melas*, *Stenella clymene*, *Physeter macrocephalus*, *Tursiops truncatus*, *Pseudorca crassidens*, *Orcinus orca*.

Questo file è composto dalla prima riga di intestazione e dalla seconda riga in poi vi sono i dati. Come primo dato viene scritta la specie, per secondo dato viene scritto il nome del file senza l'estensione, il terzo dato vale 0 oppure 1, e rappresenta in quale insieme va inserito lo spettrogramma.

Ho scritto uno script in Matlab (R2021b) il quale legge questo file, per ciascuna riga (esclusa l'intestazione) salva la classe di quel campione ed il nome del file, in seguito avviene la lettura in memoria del file e successivamente tramite le funzioni di Matlab viene ricavato lo spettrogramma corrispondente. In base al terzo valore della riga viene inserito nel primo o nel secondo insieme. Il risultato più alto ottenuto in questa fase è 21,59 %, il che è molto basso.

Fra molti errori che il sistema fa, uno di questi confonde la prima classe, ovvero i campioni della specie *Megaptera novaeangliae* con quelli della classe 9 oppure 10, rispettivamente *Delphinapterus leucas* e *Stenella attenuata*. All'orecchio umano le registrazioni di queste ultime specie sono affette da rumore, per un totale di 18 specie. Il risultato più alto ottenuto è stato 23,23 %.

Le prestazioni sono basse, questo significa che il sistema "confonde" campioni di una determinata specie piuttosto che un'altra, studiando la matrice di confusione, in particolare, le classi che il sistema sbaglia quasi sempre sono le seguenti: *Lagenorhynchus albirostris*, *Delphinus delphis*, *Monodon monoceros*, *Eubalaena glacialis*.

In seguito ho rimosso queste classi, dunque il sistema è stato addestrato e testato per un totale di 16 specie, le performance del sistema sono incrementate del 13,10 %, l'accuratezza massima ottenuta vale 36,34 %.

Di seguito viene riportato lo pseudocodice utilizzato

---

```
apro file in modalità lettura
skip_chars ← conto quanti caratteri devo saltare per la prima riga (dovuta all'intestazione)
from = skip_chars + 1 ← il puntatore del file parte dal carattere successivo
indexs_dot_coma ← sono tutti gli indici in cui nel file compare ";" escluso la prima riga
classes_label ← inizializzo l'array vuoto delle classi
index_label = 1 ← inizializzo l'indice delle classi che leggerò dal file
index_spettrogramma = 1 ← inizializzo l'indice relativo al numero di spettrogramma
i = 1 ← indice dell'array indexs_dot_coma
n ← numero di ";" presenti nel file
while (i <= n)
    mi prendo l'indice del prossimo ";"
    leggo la classe del file da from a next_dot_coma -1
    controllo se questa classe l'ho già inserita nell'array classes_label; se sì mi prendo l'indice
    corrispondente;
    se no la aggiungo all'array ed incremento index_label

    salvo pattern_class

    from = next_dot_coma + 1 ← from parte dopo il ;
    next_dot_coma leggo il prossimo ;
    file_name ← leggo il nome del file da from a next_dot_coma -1

    leggo il file tramite pattern_class e file_name
    eseguo lo spettrogramma
    fold ← leggo il valore del file da from a from

    a seconda se vale 0 o 1 li inserisco o meno in uno dei fold

end while

salvo i dati in un file database.mat
```

---

### 3.2 Esperimento numero 2

In questo esperimento è stato elaborato uno script che legge come definito precedentemente i file del protocollo WEP#3, in particolare lo spettrogramma viene costruito da due parti. La prima parte è la registrazione audio dell'animale la cui parte finale viene tagliata in base ad una determinata percentuale (nel primo esperimento è 15%). La seconda parte è sempre lo stesso suono in cui la prima parte viene tagliata in base sempre alla stessa percentuale. Infine lo spettrogramma è l'unione della prima parte e della seconda parte, il suono risulta "due gemiti" invece che uno ed all'orecchio umano appare continuativo.

Al variare della percentuale del taglio è stato eseguito il test con tutte classi, ovvero 20 e successivamente con 16 classi (senza le specie *Lagenorhynchus albirostris*, *Delphinus delphis*, *Monodon monoceros*, *Eubalaena glacialis*).

Qui sotto sono riportati i risultati ottenuti in base alla percentuale di taglio ed il numero di classi.

Numero esecuzione	Percentuale di taglio	Accuratezza con 20 classi	Accuratezza con 16 classi
1	15 %	20,48 %	31,44 %
2	18 %	21,03 %	32,22 %
3	21 %	21,96 %	30,15 %
4	24 %	21,03 %	31,19 %
5	27 %	21,96 %	34,02 %
6	30 %	23,62 %	32,47 %
7	33 %	21,22 %	32,73 %



### 3.3 Conclusioni

Dagli esperimenti eseguiti, nel caso in cui la percentuale di taglio sia il 15%, accuratezza con 20 classi vale 20,48%, mentre nel caso di 16 classi vale 31,44%.

Fino al 21% e dal 27% al 33% l'accuratezza aumenta in entrambi i casi.

Anche solo il fatto di rimuovere 4 classi dal sistema ha incrementato del 10% le prestazioni, essendo queste, le classi che il sistema non riusciva a riconoscere nella fase di test.

I punteggi sono bassi perché gli spettrogrammi non sono abbastanza discriminanti per ogni specie, questo significa che il sistema può confondere campioni di una specie rispetto ad un'altra.

Un'altra problematica legata alla discriminazione dei dati è il rumore; può influenzare lo spettrogramma ottenuto, con conseguenze di errori nella valutazione.

Per ottenere delle alte performance si può lavorare su tecniche che rimuovono il rumore ed altre che ne estraggono più informazioni possibili dal segnale audio.

Si potrebbe pensare di combinare lo spettrogramma con filtri che aggiungono informazioni univoche per quella determinata classe, un'altra ipotesi sarebbe la possibilità di eseguire tecniche di manipolazione del segnale acustico come ad esempio: il time-stretching dell'audio e la regolazione del volume.

L'immagine del campione può essere costruita tramite una tecnica o più, 0.3273 combinate insieme con lo scopo di discriminare al meglio le classi.



## Bibliografia

- [1] T. S. for Marine Mammalogy., «List of Marine Mammal Species and Subspecies,» 2021. [Online]. Available: <https://marinemammalscience.org/science-and-publications/list-marine-mammal-species-subspecies/>.
- [2] L. Sayigh, R. Wells e V. Janik, «What's in a voice? Dolphins do not use voice cues for individual recognition,» 2017.
- [3] S. King, W. Friedman, S. Allen, L. Gerber, F. Jensen, S. Wittwer, R. Connor e Krützen, «Dolphins Retain Individual Vocal Labels in Multi-level Alliances,» 2018.
- [4] F. Caruso, L. Dong, M. Lin, M. Liu, Z. Gong, W. Xu, G. Alonge e S. Li, «Monitoring of a Nearshore Small Dolphin Species Using,» 2020.
- [5] «Geographical variation in sound production in the anemonefish Amphiprion,» *Proc. Royal Soc. B Biological Sci.*, p. 272, 2005.
- [6] «Whale Challenge: Develop Recognition Solutions to Detect and Classify Right Whales for BIG Data,» 2013. [Online]. Available: <https://www.kaggle.com/c/the-icml-2013-whale-challenge-right-whale-redux>.
- [7] «Right Whale Recognition: Identify Endangered Right Whales in Aerial Photographs.,» [Online]. Available: <https://www.kaggle.com/c/noaa-right-whale-recognition>.
- [8] G. M. ., L. N. ., Y. C. a. C. S. J. Rafael Aguiar, «On the Importance of Passive Acoustic Monitoring Filters,» [Online]. Available: <https://www.mdpi.com/2077-1312/9/7/685>.
- [9] A. Flexer, «Closer Look on Artist Filters for Musical Genre Classification. In Proceedings of the 8th International Conference on Music Information Retrieval,» p. 341–344, 2007.
- [10] Y. Costa, L. Oliveira, A. Koerich e F. Gouyon, «Music Genre Recognition Using Spectrograms,» Sarajevo, Bosnia and Herzegovina., 2011.
- [11] A. Montalvo, Y. Costa e J. Calvo, Language identification using spectrogram texture, Uruguay, : 543–550, 2015.
- [12] L. Nanni, R. Aguiar, Y. Costa, S. Brahnham, C. J. Silla, R. Brattin e Z. Zhao, «Bird and whale species identification,» *IET Comput. Vis.*, p. 178–184, 2018.
- [13] T. Ojala, M. Pietikainen e D. Harwood, «Performance evaluation of texture measures with classification based on Kullback,» p. 582–585, 1994.
- [14] M. Bittle e A. Duncan, «A review of current marine mammal detection and classification algorithms for use in automated passive acoustic monitoring.,» 17 November 2013.