

Universidade Federal de Juiz de Fora
Instituto de Ciências Exatas / Faculdade de Engenharia
Programa de Pós-graduação em Modelagem Computacional

Edson Bruno Novais

**ALGORITMOS COMPUTACIONAIS PARA DETECÇÃO EFICIENTE DE
ODONTOCETOS EM DISPOSITIVOS FIXOS AUTÔNOMOS**

Juiz de Fora

2016

Edson Bruno Novais

**ALGORITMOS COMPUTACIONAIS PARA DETECÇÃO EFICIENTE DE
ODONTOCETOS EM DISPOSITIVOS FIXOS AUTÔNOMOS**

Tese apresentada ao Programa de Pós-graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora, na área de concentração em Modelagem Computacional, como requisito parcial para obtenção do título de Doutor em Ciências em Modelagem Computacional.

Orientador: Carlos Cristiano Hasenclever Borges

Coorientador: Artur Andriolo

Juiz de Fora

2016

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Novais, Edson Bruno.

Algoritmos computacionais para detecção eficiente de Odontocetos em dispositivos fixos autônomos / Edson Bruno Novais. -- 2016.

112 f. : il.

Orientador: Carlos Cristiano Hasenclever Borges

Coorientador: Artur Andriolo

Tese (doutorado) - Universidade Federal de Juiz de Fora, ICE/Engenharia. Programa de Pós-Graduação em Modelagem Computacional, 2016.

1. Aprendizado de máquina. 2. Bioacústica. 3. Processamento de sinais. I. Borges, Carlos Cristiano Hasenclever, orient. II. Andriolo, Artur, coorient. III. Título.

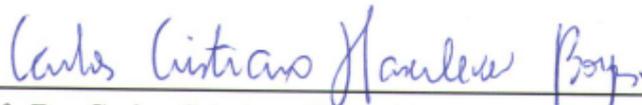
Edson Bruno Novais

ALGORITMOS COMPUTACIONAIS PARA DETECÇÃO EFICIENTE DE
ODONTOCETOS EM DISPOSITIVOS FIXOS AUTÔNOMOS

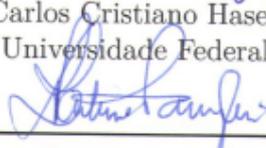
Tese apresentada ao Programa de Pós-graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora, na área de concentração em Modelagem Computacional, como requisito parcial para obtenção do título de Doutor em Ciências em Modelagem Computacional.

Aprovada em: 9 de março de 2016

BANCA EXAMINADORA



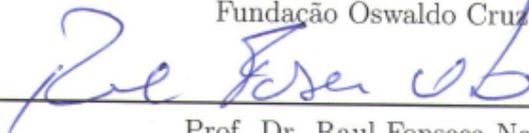
Prof. Dr. Carlos Cristiano Hasenclever Borges - Orientador
Universidade Federal de Juiz de Fora



Prof. Dr. Paulo Augusto Zaitune Pamplin
Universidade Federal de Alfenas



Prof. Dr. Douglas Adriano Augusto
Fundação Oswaldo Cruz



Prof. Dr. Raul Fonseca Neto
Universidade Federal de Juiz de Fora



Prof. Dr. Ciro de Barros Barbosa
Universidade Federal de Juiz de Fora

Dedico este trabalho a família e amigos, em especial a minha noiva Érica.

AGRADECIMENTOS

A minha noiva Érica por todo apoio, incentivo e paciência durante o desenvolvimento deste trabalho;

Aos meus orientadores Cristiano e Artur pela orientação durante a elaboração deste trabalho;

Aos membros da banca Paulo Augusto, Douglas Adriano, Raul Fonseca e Ciro Barbosa pela disponibilidade de participação e contribuições acerca da tese;

A banca de qualificação Renata Sousa-Lima, Flávio Barbosa e Henrique Hippert pela contribuição no exame de qualificação propondo melhorias ao trabalho;

Aos colegas de doutorado Fabrizzio, Aldemon, Bruno, Jonathan, Pedro, Rafael e Denis pela amizade e bom humor que amenizaram momentos difíceis;

Aos colegas da Ecologia e Comportamento Animal da UFJF, Tiago, Gustavo, Franciele e Frederico pela ajuda e disponibilidade no desenvolvimento do trabalho;

Aos funcionários do Programa de Pós-graduação em Modelagem Computacional, Reginaldo, Natália e Adriana pela disponibilidade e atenção nos serviços prestados;

Ao Departamento de Ciência da Computação da UFJF pelo auxílio em pesquisa durante os 2 anos como professor substituto;

Ao Programa de Pós-graduação em Modelagem Computacional;

À Universidade Federal de Juiz de Fora;

À FAPEMIG pelo apoio financeiro concedido.

“The fewer the facts, the stronger the opinion.”

Arnold H. Glasow

RESUMO

Considerando o sucesso dos dispositivos móveis e a evolução de suas tecnologias, o estudo de Odontocetos em tempo real é uma possibilidade emergente. Apesar desta evolução, a análise de grandes volumes de dados por algoritmos complexos requer considerável esforço computacional. A coleta de dados de Odontocetos é executada em ambiente marinho com recursos limitados, o que reduz o potencial de pesquisa. Sendo assim, a escolha dos algoritmos para a criação de um Fluxo de Trabalho deve manter um balanço entre a eficiência computacional e a eficácia de detecção. Esta tese tem como objetivo propor um modelo de Fluxo de Trabalho eficiente para o Monitoramento Acústico Passivo. Para tal, um Fluxo de Trabalho de referência comumente utilizado em campo por pesquisadores foi utilizado como base, sendo inserido uma nova etapa de pré-processamento das informações capturadas. A etapa de detecção, foco deste trabalho devido sua aplicabilidade e notável impacto nas próximas etapas, é responsável por analisar os sinais acústicos recebidos, filtrando boa parte dos dados. A próxima etapa trata da condensação dos dados de forma a facilitar a transferência destes para localidades remotas. Em sequência tem-se a etapa de identificação das informações recebidas a partir da etapa anterior. Por fim, a última etapa baseia-se em componentes de software para o estudo das informações relevantes adquiridas. A aplicação da etapa de detecção no Fluxo de Trabalho de referência apresentou um desempenho satisfatório acarretando em uma redução de 96,52% do volume total de dados a serem armazenados e processados, facilitando que informações relevantes da captura sejam identificadas e distribuídas *online* para estações de pesquisa remotas.

Palavras-chave: Aprendizado de Máquina. Bioacústica. Processamento de Sinais.

ABSTRACT

Considering the success of mobile devices and the evolution of its technologies, the study of Odontoceti in real time is an emerging possibility. Despite the evolution, analysis of big data chunks by complex algorithms requires considerable computing effort. Data collection of Odontoceti is executed in marine environment with limited resources, thus reducing research potential. Therefore, choosing the right algorithm to create the Workflow should maintain a balance between computational efficiency and detection accuracy. The goal of this thesis is to propose an efficient Workflow for Passive Acoustic Monitoring. For such, a common Workflow used in the field by researchers was used as base, adding a new step for preprocessing of captured data. The detection step, focus of this thesis due to its applicability and notable impact on the next steps, is responsible to analyse received acoustic signals, filtering a good amount of data. The next step condensates data in a way that facilitates transfer of captured information to remote locations. In sequence there is the step responsible for the identification of received information from the previous step. The last step is based on software components to study relevant information. The application of the detection step have shown a satisfactory performance providing a reduction of 96.52 of total data to be processed, making it easy for relevant information to be identified and distributed online to remote research stations.

Key-words: Machine Learning. Bioacoustics. Signal Processing.

LISTA DE ILUSTRAÇÕES

Figura 1 – Baleia Jubarte, extraída de [26].	22
Figura 2 – Golfinho Rotador, extraída de [27].	22
Figura 3 – Ecolocalização de um Odontoceti, extraída de [30].	24
Figura 4 – Exemplo de um sistema AR, extraída de [32].	25
Figura 5 – Fluxo de Trabalho geral utilizado como referência.	34
Figura 6 – Fluxo de Trabalho geral proposto para dispositivos fixos autônomos.	47
Figura 7 – Fluxo de Trabalho geral proposto com possibilidades emergentes.	49
Figura 8 – Sinal de um segundo contendo três cliques de Odontocetos.	54
Figura 9 – Fluxo de Trabalho geral para etapa de detecção.	55
Figura 10 – Divisão do sinal em segmentos menores.	57
Figura 11 – Conversão do sinal para Espaço de Amplitude.	58
Figura 12 – Espaço de Amplitude com e sem clique de cetáceos.	59
Figura 13 – Espaço de Amplitude gerado a partir do sinal normalizado.	60
Figura 14 – Vetor de ocorrências gerados como exemplo em um espectrograma.	61
Figura 15 – Variância dos Espaços de Amplitude gerados no tempo.	65
Figura 16 – Exemplo de distribuição normal.	66
Figura 17 – Regras utilizadas pela lógica difusa.	68
Figura 18 – Espaço de Amplitude sem ordenação.	69
Figura 19 – Espaço de Amplitude com ordenação.	70
Figura 20 – Máxima margem, extraída de [177].	77
Figura 21 – Unidades relevantes do Espaço de Amplitude.	80
Figura 22 – Locais de implantação do HARP, extraída de [181].	82
Figura 23 – Exemplo de um 10- <i>fold</i>	86

LISTA DE TABELAS

Tabela 1 – Espécies de Odontocetos capturadas.	81
Tabela 2 – Detalhes dos locais de implantação do HARP.	82
Tabela 3 – Dados do projeto SOCAL sítio R utilizados como treinamento do IMA.	84
Tabela 4 – Dados de análise por especialistas.	84
Tabela 5 – Dados de análise por especialistas comparado com o modelo proposto.	85
Tabela 6 – Matriz de confusão criada a partir dos dados de treinamento.	86
Tabela 7 – Avaliação dos modelos nos 10-fold do teste.	87
Tabela 8 – Valores percentuais da acurácia para os 10 grupos de teste.	87
Tabela 9 – Redução de dados por projeto e sítio.	88
Tabela 10 – Resultados detalhados por dia da captura no projeto CINMS sítio B.	106
Tabela 11 – Resultados detalhados por dia da captura no projeto CINMS sítio C.	107
Tabela 12 – Resultados detalhados por dia da captura no projeto DCPD sítio A.	108
Tabela 13 – Resultados detalhados por dia da captura no projeto DCPD sítio B.	109
Tabela 14 – Resultados detalhados por dia da captura no projeto DCPD sítio C.	110
Tabela 15 – Resultados detalhados por dia da captura no projeto SOCAL sítio E.	111
Tabela 16 – Resultados detalhados por dia da captura no projeto SOCAL sítio R.	112

LISTA DE ABREVIATURAS E SIGLAS

ACPI	Advanced Configuration and Power Interface
ANN	Artificial Neural Network
ANOVA	Análise de Variância
AR	Autonomous Recorders
BPMN	Business Process Model and Notation
CS-SVM	Class-Specific SVM
DCLDE	Detection, Classification, Localization, and Density Estimation
DFA	Discriminant Function Analysis
DFT	Discrete Fourier Transform
D_J	Divergência de Jeffrey
D_{JS}	Divergência de Jensen-Shannon
D_{KL}	Divergência de Kullback-Leibler
EM	Expectation Maximization
ECDF	Empirical Distribution Function
ERMA	Energy Ratio Mapping Algorithm
FCH	Fixed-cabled Hydrophone
FFT	Fast Fourier Transform
FMCD	Frequency Modulated Echolocation Click Detector
FMP	Fixed Margin Perceptron
HARP	High-frequency Acoustic Recording Package
HI	Histogram Intersection
HM	Histogram Match
HTK	Hidden Markov Model Toolkit
ICA	Independent Component Analysis
IMA	Incremental Margin Algorithm

JIP	Joint Industry Programme
GMM	Gaussian Mixture Model
GMMD	Gaussian Mixture Model Detector
GPLP	Generalized Perceptual Linear Prediction
KS	Kolmogorov-Smirnov
LSH	Locality Sensitive Hashing
MLE	Maximum Likelihood Estimator
OMG	Object Management Group
PAM	Passive Acoustic Monitoring
PCA	Principal Component Analysis
PLP	Perceptual Linear Prediction
RLH	Radio-linked Hydrophone
SNR	Signal-to-Noise Ratio
SMC	Sequential Monte Carlo
STFrFT	Short-time Fractional Fourier Transform
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
SWfMS	Scientific Workflow Management System
SWFSC	Southwest Fisheries Science Center
TDOA	Time Difference of Arrival
TFR	Time Frequency Representation
WAVE	Waveform Audio File Format (WAVE)
WSD	Warping-Based Signal Decomposition

SUMÁRIO

1	INTRODUÇÃO	15
1.1	OBJETIVOS	18
1.2	ESTRUTURA DO TRABALHO	18
2	MONITORAMENTO ACÚSTICO PASSIVO	19
2.1	CETÁCEOS	21
2.1.1	Cliques	23
2.2	DISPOSITIVOS FIXOS AUTÔNOMOS	24
2.3	VOLUME DE DADOS	28
3	UM MODELO DE FLUXO DE TRABALHO PARA MONI- TORAMENTO ACÚSTICO PASSIVO	31
3.1	FLUXO DE TRABALHO DE REFERÊNCIA	33
3.1.1	Detecção	33
3.1.1.1	Detecção de Limiar de Energia	34
3.1.1.2	Detector de Entropia	35
3.1.1.3	Teste de Page	35
3.1.1.4	Análise de Componente Principal e Análise de Componente Independente	36
3.1.1.5	Algoritmo de Schur	36
3.1.1.6	Transformada <i>Wavelet</i>	36
3.1.2	Vetor de Características	37
3.1.2.1	Traçado do Pico de Frequência	37
3.1.2.2	Ajuste da Curva	38
3.1.2.3	Variância Mínima de Capon	38
3.1.2.4	Decomposição do Sinal Baseado em Deformação	38
3.1.2.5	Filtro de Partículas	39
3.1.2.6	Características Cepstrais	39
3.1.2.7	Modelo de Predição Linear Perceptual Generalizada	40
3.1.3	Classificação	40
3.1.3.1	Redes Neurais Artificiais	41
3.1.3.2	Modelo de Mistura Gaussiana	42
3.1.3.3	Máquinas de Vetores de Suporte	43
3.1.3.4	Análise de Função Discriminante Multivariada	44
3.1.3.5	Picos Espectrais	44
3.1.4	Pacotes de Software	45
3.2	FLUXO DE TRABALHO PROPOSTO	47

3.2.1	Possibilidades Emergentes	48
4	DETECÇÃO EFICIENTE	52
4.1	FLUXO DE DETECÇÃO	55
4.2	ESPAÇO DE AMPLITUDE	57
4.2.1	Distâncias e Divergências	60
4.2.1.1	<i>Histogram Intersection</i>	61
4.2.1.2	<i>Histogram Match</i>	61
4.2.1.3	Distância Euclidiana	62
4.2.1.4	Distância Chi-quadrada	62
4.2.1.5	Divergência de Kullback-Leibler	63
4.2.1.6	Divergência de Jensen-Shannon	63
4.2.1.7	Divergência de Jeffrey	64
4.2.2	Variância	64
4.2.2.1	Testes Paramétricos	65
4.2.2.2	Testes não Paramétricos	66
4.2.2.3	Lógica Difusa	67
4.3	ORDENAÇÃO	68
4.3.1	Ordenação Simples	70
4.3.2	Ordenação Eficiente	71
4.3.3	Ordenação de Inteiros	73
4.4	APRENDIZADO DE MÁQUINA	76
5	RESULTADOS EXPERIMENTAIS	81
6	CONCLUSÃO	89
	REFERÊNCIAS	92
	APÊNDICE A – Resultados Completos de Detecção	105

1 INTRODUÇÃO

A inspeção de espécies marinhas em seu habitat natural é de enorme interesse para promover conhecimento de padrões comportamentais, estimativa populacional e até mesmo influência de efeitos antropogênicos na preservação das espécies. Devido às dificuldades naturais inerentes ao ambiente marinho, intensificado pela maior parte da vida destes animais ser debaixo da superfície do mar, estudos sobre os efeitos destas atividades se tornam viáveis somente com o uso de tecnologia [1]. A identificação de mamíferos marinhos normalmente é realizada através de longas horas de pesquisa visual e acústica. Embora a pesquisa visual desempenhe um papel vital no processo de identificação, mamíferos marinhos são difíceis de serem detectados, especialmente em condições meteorológicas adversas [2, 3]. Além disso, a pesquisa visual é capaz de detectar somente uma fração dos mamíferos marinhos presentes, dado que o tempo de permanência na superfície normalmente é curto para uma análise adequada [4].

Desta forma, resultados de pesquisas visuais podem ser altamente variáveis, dado tanto a aglomeração em grandes grupos, como a sua escala espacial e temporal relativamente limitada [4]. Odontocetos da família *Ziphiidae* por exemplo, são notoriamente evasivos e difíceis de serem estudados, especialmente com a utilização de pesquisas visuais [5]. Isso acontece por causa da sua capacidade de submersão, permanecendo normalmente por uma hora em profundidades de até mil metros [6], tendo registros de submersões de mais de duas horas a dois mil metros de profundidade [7].

Um dos objetivos de avaliações bioacústicas em ambientes marinhos é a avaliação de ruídos antropogênicos, que possuem efeitos extremamente prejudiciais aos mamíferos marinhos [8], levando até mesmo ao óbito em certos casos [9]. Preocupações sobre estes efeitos estão ganhando grande visibilidade, particularmente nos Odontocetos, que são conhecidos pela sua sensibilidade e capacidade acústica [10], sendo dependentes de sons para navegação e comunicação [11].

Devido aos desenvolvimentos computacionais nos últimos anos, métodos de Monitoramento Acústico Passivo (*Passive Acoustic Monitoring*, PAM), tornaram-se cada vez mais difundidos para observação de mamíferos marinhos [4, 5]. No ano 2000, o PAM foi integrado aos métodos de pesquisa do *Southwest Fisheries Science Center* (SWFSC) [12]. Isso resultou em um aumento tanto na quantidade e distância de detecção, como no entendimento do comportamento dos mamíferos marinhos [3]. Em procedimentos conjuntos utilizando PAM e pesquisa visual, a modalidade acústica foi capaz de detectar de uma até dez vezes mais grupos que a modalidade visual [4]. Além disso, observações acústicas possuem a vantagem de poderem continuar a noite, com mau tempo, e outras condições as quais observações visuais seriam ineficazes [4, 11].

Em um passado recente as limitações e alto custo em tecnologias, bem como as

restrições de métodos disponíveis e conhecimento necessário, dificultaram o desenvolvimento e aplicação do PAM para estudo de mamíferos marinhos [11]. Com os recentes avanços em usabilidade juntamente com a diminuição dos preços de equipamentos, o uso de métodos acústicos baseados em componentes eletrônicos e computacionais para detecção, localização e acompanhamento de mamíferos marinhos vêm ganhando cada vez mais espaço [13]. Com a difusão dessas técnicas, a quantidade de dados adquiridos aumentou de maneira expressiva, levando a uma necessidade emergente no desenvolvimento de sistemas para detecção e identificação automática¹ [4], particularmente sobre a vocalização dos Odontocetos [14]. A disponibilização de dados contínuos é um dos benefícios mais aparentes à pesquisa científica proporcionado por esse tipo de sistema [4], entretanto, existe uma diferença muito grande entre a capacidade de sistemas executados nos laboratórios em terra, com aqueles executados em dispositivos fixos autônomos no mar, que podem apresentar restrições em relação a capacidade de armazenamento, capacidade energética e acessibilidade de dados.

O uso de dispositivos fixos autônomos reduziu consideravelmente o custo e a quantidade de mão de obra especializada necessária para monitorar mamíferos marinhos por longos períodos de tempo. Para prover gravações acústicas de Odontocetos por longos períodos utilizando dispositivos fixos autônomos, existem alguns itens principais que devem ser levados em consideração. O balanceamento destes itens é de extrema importância na hora de escolher o melhor equipamento para uma aplicação em particular [11]. Algumas questões que devem ser avaliadas:

- Assim como em qualquer dispositivo suprido por baterias, componentes de baixo custo energético são essenciais para implantações de longa duração;
- Digitalização de alta velocidade para gravações de mamíferos marinhos que possuem uma largura de banda extensa, como é o caso de Odontocetos [15]; e
- Dispositivos de armazenamento de alta capacidade. Quando aumentamos a quantidade de dispositivos de armazenamento, temos um impacto direto no gasto energético [16].

Exemplificando a relevância destas questões, sabe-se que dispositivos fixos autônomos são extremamente úteis na detecção de mamíferos marinhos em áreas onde a ocorrência destes animais é baixa [17]. Apesar disso, algoritmos complexos estão sendo executados por processadores de baixa capacidade [5, 13, 18], dificultando assim a captura e disponibilização de informações relevantes.

¹ Entende-se como detecção, o reconhecimento de vocalização pertencente a um *Odontoceti* de interesse, e identificação a caracterização de qual espécie este animal pertence.

Ainda que exista a possibilidade técnica de implementação de detecção de qualidade, a identificação e distribuição automática dos dados capturados requer extensa atividade computacional. Sendo assim, o custo envolvido na implementação destas tecnologias em ambientes marinhos com recursos limitados pode inviabilizar seu uso efetivo. Entre as principais limitações, destacam-se os recursos energéticos. O maior desafio desses sistemas em relação ao gasto energético é a manipulação do grande volume de dados acústicos adquirido, que requer análises robustas para gerar detecções de espécies confiáveis [4]. Portanto, um sistema para detecção, identificação e distribuição *online* de dados bioacústicos em ambiente marinho deve fazer um balanço entre eficiência computacional e eficácia de identificação, bem como o momento ideal de distribuição das informações geradas através de redes que estejam disponíveis.

Existem diversos pacotes de software disponíveis para detecção e identificação de mamíferos marinhos, entretanto, poucos realizam essas tarefas efetivamente, e nenhum é capaz de classificar concorrentemente a vocalização de um número elevado de espécies com boa acurácia [11, 19, 20]. Métodos automáticos para detecção de Mysticetos são relativamente bem desenvolvidos [21], porém, detecções automáticas em altas frequências – como é o caso dos Odontocetos –, necessitam de mais estudo e desenvolvimento [11]. Para dispositivos fixos autônomos, juntamente com a acurácia de detecções em altas frequências, é importante ressaltar a implicação da frequência de Nyquist², que acaba aumentando o gasto computacional em função da elevação do volume de dados necessários à captura de interesse. De uma forma geral, os principais sistemas para PAM possuem vantagens e desvantagens. Diversos aspectos importantes devem ser considerados na escolha de qual dispositivo e método de análise deve ser utilizado. Por exemplo, a frequência dos sinais emitidos pela espécie de interesse influencia diretamente em qual método utilizar, bem como na capacidade de armazenamento, processamento e tempo de monitoramento [11].

Como pode ser visto em Sousa-Lima et al. [22], muitos dos dispositivos fixos autônomos utilizam modelos desenvolvidos com propósitos distintos, normalmente executados em equipamentos mais robustos. Acontece que neste caso, um modelo robusto está sendo executado por um dispositivo de baixa capacidade, sendo que dispositivos de maior capacidades foram utilizados no desenvolvimento do modelo. São raros os casos de dispositivos que consideram os fatores limitantes desde seu projeto até implementação, como é o caso do *High-frequency Acoustic Recording Package* (HARP), que tem como objetivo primário prover gravações de cliques de Odontocetos em altas frequências durante longos períodos de tempo [16].

² A quantidade de amostras por unidade de tempo de um sinal deve ser maior que o dobro da maior frequência contida no sinal a ser amostrado [11].

1.1 OBJETIVOS

Considerando as informações apresentadas, esta tese tem como objetivo propor um modelo de Fluxo de Trabalho eficiente para PAM. Para tal, o Fluxo de Trabalho comumente utilizado em campo por pesquisadores foi utilizado como base, sendo inserido uma nova etapa de pré-processamento das informações capturadas.

A etapa de pré-processamento facilita a utilização de modelos mais robustos em equipamentos mais simples, reduzindo a quantidade de dados processados pelas próximas etapas do Fluxo de Trabalho. Ao reduzir a quantidade de dados totais a serem manipulados sem perda de informações relevantes, a necessidade de processamento e armazenamento também é reduzida. Com isso, diminui-se a demanda computacional necessária dos equipamentos utilizados, facilitando a distribuição e identificação *online* de informações relevantes encontradas nas detecções.

1.2 ESTRUTURA DO TRABALHO

A presente tese está dividida em 6 capítulos, incluindo esta introdução que apresenta uma visão geral sobre o tema que será tratado durante os próximos capítulos. No Capítulo 2 é apresentado todo o embasamento teórico necessário para o entendimento do PAM, focando em especial nos dispositivos fixos autônomos e na quantidade de dados gerados pela captura. Posteriormente no Capítulo 3 é desenvolvido um Fluxo de Trabalho para detecção de dados capturados por dispositivos fixos autônomos. O Fluxo de Trabalho tem como intuito padronizar parte do processo computacional, reduzindo gastos e facilitando a distribuição, identificação e análise das capturas. No Capítulo 4 é delineado o algoritmo proposto para detecção, levando em consideração as diversas limitações impostas pelo uso de dispositivos fixos autônomos em alto mar. O Capítulo 5 apresenta os resultados obtidos a partir da aplicação da detecção em um conjunto de dados anotados por especialistas, onde são detectados diversos Odontocetos em áreas e momentos distintos. Finalmente no último capítulo é apresentado as conclusões obtidas com o desenvolvimento desta tese, bem como propostas para trabalhos futuros.

2 MONITORAMENTO ACÚSTICO PASSIVO

Em 1880, Pierre Curie e seu irmão mais velho Jacques Curie [23] descobriram que quando uma pressão mecânica é exercida em um cristal de Quartz, um potencial elétrico é produzido. Essa descoberta possibilitou que pesquisadores escutem, armazenem e analisem sons produzidos por mamíferos marinhos [11].

A introdução de sons antropogênicos no meio ambiente marinho começou a ser levado a sério com a revolução industrial. Os efeitos potenciais de tal som em mamíferos e ecossistemas marinhos não foram reconhecidos como uma preocupação até os anos de 1970, quando estudos na região do Ártico indicavam que algumas espécies de mamíferos marinhos estavam respondendo a atividades associadas à exploração e extração de petróleo e gás [24].

Expandindo um pouco os conceitos, o monitoramento acústico pode ser dividido em duas categorias, o ativo e o passivo. No Monitoramento Acústico Ativo – não será tratado neste trabalho –, os sons são transmitidos pelos equipamentos, onde o retorno é analisado. Este tipo de monitoramento é amplamente utilizado para observação de Zooplâncton e Micronecton (principalmente peixes, lulas e camarões), bem como na pesquisa de pesca [4]. No Monitoramento Acústico Passivo (*Passive Acoustic Monitoring*, PAM) - utilizado neste trabalho -, o equipamento não produz nenhum tipo de som [4].

O PAM é uma metodologia de trabalho que se embasa em equipamentos capazes de obterem informações acústicas passivamente, realizando a captura de sons ao redor do ambiente implantado. Ele é uma ferramenta valiosa para monitorar os mamíferos marinhos, visto que muitas espécies produzem sons distintivos [25]. Devido à recente redução dos preços de equipamentos, além da evolução computacional e usabilidade dos sistemas, o PAM se tornou largamente utilizado para observação de Cetáceos [4]. Neste tipo de monitoramento, destaca-se a eficiência alcançada na detecção de Baleias Bicudas em relação às técnicas baseadas em pesquisas visuais [18].

O PAM apresenta algumas variações em seus procedimentos, destacando a possível utilização de equipamentos móveis ou fixos [4]. Para equipamentos móveis por exemplo, existem hidrofones sendo rebocados por navios, afixados em planadores oceânicos, ou qualquer outro tipo de plataforma móvel capaz de amostrar uma grande área [4]. Para equipamentos fixos temos tipicamente hidrofones implantados em estações permanentes ou semi-permanentes, sendo mantido por períodos de até dois anos com um custo relativamente baixo [4]. Ambas as abordagens compartilham de um processo de captura similar, onde o projeto de mobilidade ou fixação deve ser realizado, indicando se haverá a recuperação dos equipamentos utilizados e como será realizada a extração e análise dos dados para uma interpretação robusta [4].

Em adição às abordagens citadas, existem também arquiteturas de hidrofones

fixos em boias capazes de se comunicarem através de satélites, possibilitando a troca de informações entre estações geograficamente dispersas [4]. Além disso, uma variante mais recente dada pela miniaturização de certos componentes eletrônicos, consiste na utilização de Cetáceos como plataforma acústica. Estes componentes são armazenados em etiquetas anexáveis, sendo capazes de capturar informações acústicas do ambiente onde o animal está exposto [4].

Sistemas de Hidrofones Fixos Cabeados (*Fixed-cabled Hydrophone*, FCH) são tipicamente permanentes ou semi-permanentes localizado no fundo do mar. Eles têm a capacidade de serem continuamente alimentados por uma fonte externa e podem enviar continuamente dados para uma estação de recepção, normalmente localizado em terra. Alguns exemplos são: SOSUS de baixa frequência utilizado pela marinha dos EUA, gama de testes subaquáticos com agregados de hidrofones (por exemplo, AUTECH nas Bahamas), observatórios de grande escala para o mar, redes de hidrofones de menor escala e matrizes destinadas à investigação mamíferos marinhos [11].

Sistemas de Hidrofones através de Rádio (*Radio-linked Hydrophone*, RLH) consistem em um ou mais hidrofones que estão ancorados ou fixados ao fundo do mar. Estes sistemas podem transmitir sinais acústicos através de ondas de rádio para uma estação de recepção em terra, o que permite que dados acústicos possam ser remotamente monitorados e processados em tempo real, ou seja, *online*. Alguns exemplos de RLHs incluem sistemas que foram desenvolvidos e implantados para monitorar grandes baleias nas rotas de navegação ao longo do porto de Boston, Massachusetts [11].

Hidrofones Fixos Cabeados, Sistemas de Hidrofones através de Rádio e Gravadores Autônomos (*Autonomous Recorder*, AR) – que será detalhado na Seção 2.2 –, possuem vantagens e desvantagens distintas. Em geral, os custos de instalação e de infra-estrutura são mais elevados para FCH e RLH sendo menor para AR. No entanto, a largura de banda de dados acústicos, recursos de coleta, longevidade de monitoração e capacidades em tempo real são maiores para os modelos FCH e RLH. Devido ao seu caráter autônomo e portátil, os ARs são mais flexíveis em suas possibilidades de configuração espacial e possíveis locais de implantação. No entanto, os ARs requerem a recuperação do aparelho para o acesso aos dados. Desta forma, devido as suas características, o monitoramento em tempo real de dados acústicos capturados por ARs é inviável. Os RLHs têm capacidade de aquisição de dados em tempo real, mas geralmente possuem taxas de largura de banda de frequência e de transferência de dados mais limitados do que os sistemas FCHs, porque precisam usar redes de transmissão via rádio. Sistemas RLHs tipicamente têm menor custo de instalação do que os FCHs, mas seus custos de desenvolvimento e manutenção tendem a ser maiores. RLHs estão normalmente localizados relativamente perto da costa e requerem uma estação de recepção em terra, mas os dados podem ser processados em tempo real ou pré-processados na boia de captura de dados. Sistemas híbridos combinam

elementos de ARs, RLHs e FCHs podendo fornecer um bom compromisso de custo e capacidade, permitindo à aquisição de dados em tempo real ou quase em tempo real, com a flexibilidade nas possibilidades de implantação de ARs [11].

2.1 CETÁCEOS

Os Cetáceos – do latim *cetus* (grande animal marinho) –, são animais com formas e tamanhos variados, desde os pequenos golfinhos que mal ultrapassam um metro de comprimento, até a Baleia Azul que mede cerca de 25 metros de comprimento, sendo o maior animal vivo da Terra.

Os Cetáceos formam um grupo dentro dos mamíferos marinhos e apresentam características próprias, tais como: presença de glândulas mamárias, respiração aérea feita através de pulmões, produção de leite para alimentar seus filhotes, orifício respiratório no topo da cabeça, corpo fusiforme com espessa camada de gordura, período de gestação longo e o fato de passarem toda a sua vida no ambiente marinho. Eles se distribuem por todos os oceanos e pela maioria dos grandes rios de todo o mundo, desde as águas quentes do equador até as águas frias dos polos. Os Cetáceos estão divididos em duas sub-ordens:

- Mistícetos (as baleias sem dentes): são caracterizadas por cerdas bucais, que são estruturas parecidas com peneiras localizadas na parte superior da boca feitas de queratina. As baleias utilizam tais cerdas para filtrar plâncton da água. Os Mistícetos são considerados a maior espécie de animais do planeta Terra. A Fig. 1 mostra a Baleia Jubarte, uma das representantes dos Mistícetos.
- Odontocetos (as baleias com dentes): caracterizam-se por apresentarem dentes, alimentando-se de peixes e lulas. Possuem crânio assimétrico e orifício respiratório único dirigido para cima. Normalmente os machos são maiores do que as fêmeas. Os Odontocetos possuem um sistema acústico conhecido como ecolocalização, que possibilita a localização de presas e objetos. Este sistema permite determinar a forma, tamanho, textura e até mesmo a distância do objeto avaliado. O sistema de ecolocalização não foi ainda comprovado nos Mistícetos. A Fig. 2 mostra um exemplo de Odontoceti, representado pelo Golfinho Rotador.

Para viver no ambiente aquático onde as condições de visibilidade são bastante limitadas, os Cetáceos tiveram que desenvolver alguns sentidos mais apurados do que a maioria dos mamíferos terrestres. O mundo das baleias, botos e golfinhos é basicamente um mundo de sons. Apesar de não possuírem orelhas, os Cetáceos possuem o sentido de audição muito eficiente. O ouvido é capaz de indicar claramente a direção de onde está vindo o som na água, mesmo sem a presença de orelhas. Os Cetáceos são capazes de ouvir e emitir uma gama de sons muito maior do que a faixa que o ouvido humano

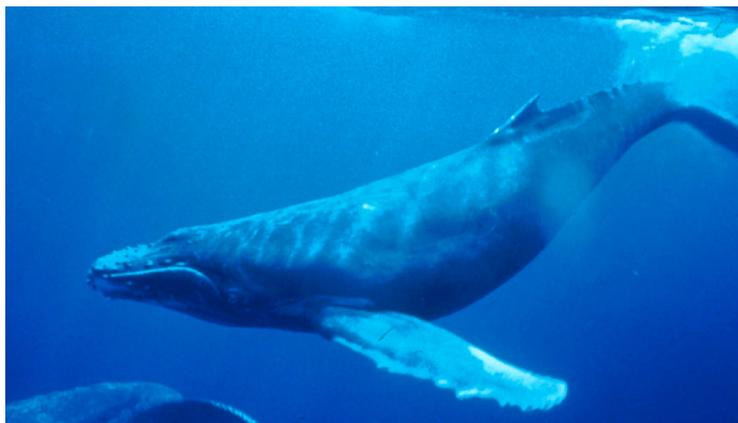


Figura 1 – Baleia Jubarte, extraída de [26].



Figura 2 – Golfinho Rotador, extraída de [27].

pode captar. A emissão dos sons serve para reunir os sexos para a reprodução, localizar e reunir indivíduos desgarrados, advertir do perigo ou intimidar inimigos e localizar objetos ou presas [28]. Para produzir esses sons, os Cetáceos não contam com cordas vocais e sim canais respiratórios.

Os Cetáceos podem se comunicar a quilômetros de distância, visto que na água o som se propaga mais rápido do que no ar. Os sons podem ser simples estalidos e ruídos ou até composição contínuas que se assemelham a canções. Esses sons podem ser escutados ou não, sendo que as canções podem durar de alguns minutos até horas. Para captar tais sons e gravá-los, utilizam-se de hidrofones - microfones adaptados à água.

Com o desenvolvimento de avançadas técnicas de estudo dos Cetáceos, descobriu-se que determinadas espécies apresentam vocalizações características. A maioria dos Odontocetos emitem basicamente três categorias de sons: os assobios, os cliques de ecolocalização e os sons pulsantes. Resumidamente, os assobios são utilizados para a comunicação

intraespecífica (ocorre entre indivíduos da mesma espécie) e os cliques de ecolocalização para a localização de presas. Os sons pulsantes parecem estar associados as emoções e são produzidos em situações onde se tem comportamentos agressivos, comportamentos pré-copulatórios, dentre outros. Os sons de ecolocalização e pulsantes se assemelham, sendo que um tipo de som pode ser originado de outro durante a emissão.

2.1.1 Cliques

Os mamíferos marinhos vivem a maior parte de suas vidas sob a superfície do oceano, fora do alcance visual dos seres humanos. Produzem sons que são altamente variáveis com características que se estendem por várias ordens de magnitude ao longo das dimensões de tempo, frequência e amplitude. O repertório de sinais acústicos de mamíferos marinhos inclui curta duração (por exemplo, 2 *ms*), maior duração (de frações de segundo a alguns segundos), entre outros. Algumas espécies produzem sons estereotipados (por exemplo, as Baleias Azuis e Cachalotes), enquanto outros produzem sinais com grande variabilidade sonora (por exemplo, algumas espécies de golfinhos e Baleias Jubarte).

Os sons utilizados na ecolocalização consistem em curtos estouros de cliques agudos que são repetidos em diferentes frequências. Cliques em baixa frequência têm um maior poder de penetração e podem viajar longas distâncias. Eles são refletidos por grandes feições e utilizados na obtenção de informações sobre a topologia nos arredores. Ondas sonoras de baixa frequência também podem ser utilizadas por algumas baleias com dentes para surpreender suas presas. Além dos cliques de ecolocalização, alguns animais emitem certas variações, como é o caso das Cachalotes, capazes de emitir quatro variações de cliques por razões sociais [29].

Esses cliques enviados ao ambiente pelos Cetáceos permitem que eles sejam capazes de encontrar objetos distantes. Os ecos que retornam servem para determinar a distância de um objeto baseado no tempo de reflexão do som. A Fig. 3 ilustra a ecolocalização de um Odontoceti. A ecolocalização é uma sofisticada capacidade biológica de detectar a posição e/ou a distância de objetos (obstáculos no meio ambiente) ou animais através de emissão de ondas ultrassônicas, no ar ou na água. Para diversos mamíferos, como morcegos, golfinhos e baleias, essa capacidade é de importância crucial em condições onde a visão é insuficiente, como no caso dos morcegos ou em águas escuras ou turvas para os golfinhos, seja para locomoção ou para captura de presas.

Assim que o animal ecolocalizador chega mais perto da presa, a frequência com que produz os cliques fica cada vez mais rápida. A medida que o intervalo desses cliques fica menor, ele começa a soar como um chiado. O som que retorna tem um som diferente do que o originalmente enviado. Essa diferença de som, fornece ao animal informação sobre o tamanho, forma, orientação, direção, velocidade e até mesmo sobre a composição do objeto.

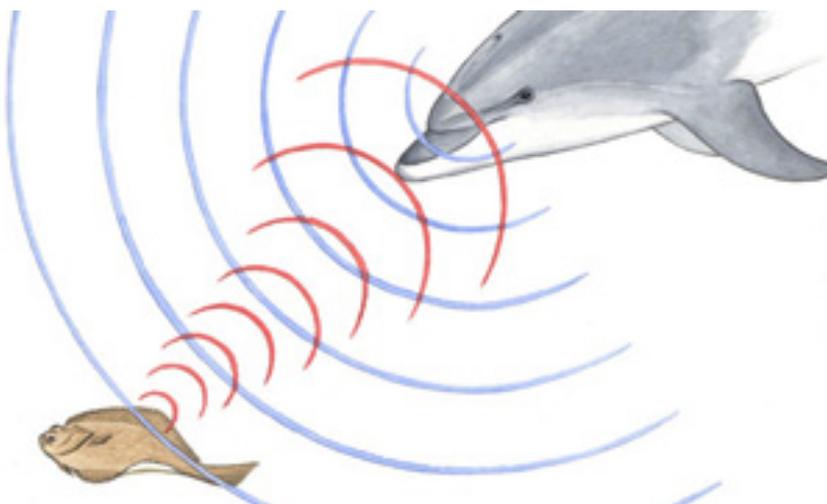


Figura 3 – Ecolocalização de um Odontoceti, extraída de [30].

2.2 DISPOSITIVOS FIXOS AUTÔNOMOS

Dispositivos de gravação autônomos fixos – Gravadores Autônomos (*Autonomous Recorders*, AR) – são definidos como qualquer dispositivo eletrônico de registo ou sistema que adquire e armazena dados acústicos internamente por conta própria, sem a necessidade de uma pessoa para executá-lo. É implantado semi-permanentemente debaixo d'água sendo caracterizado como de arquivo, ou seja, deve ser recuperado após o período de implantação para acesso dos dados [31]. A Fig. 4 ilustra um sistema AR e seus componentes, que incluem:

- Hidrofonos únicos ou múltiplos para aquisição de sinais;
- Armazenamento, sendo mais conhecidos os discos rígidos, cartões de memória *flash* e discos de estado sólido;
- Componentes eletrônicos internos para controle e condicionamento acústico como, por exemplo, amplificadores de sinal, *anti-aliasing*, filtros passa-banda, e conversores analógico-digitais; e
- Eletrônicos adicionais como, por exemplo, mecanismos de liberação acústica para permitir a recuperação do dispositivo.

O AR é um método eficaz de controle de mamíferos marinhos, especialmente para identificação de quais espécies estão presentes em uma determinada área e em um determinado período [33, 34, 35], para a localização e rastreamento de indivíduos [36], para a identificação de sons associados a diferentes regiões [37] e para a determinação de padrões de distribuição e abundância relativa [38]. O desenvolvimento deste dispositivo reduziu muito os custos e dados necessários para o monitoramento de mamíferos marinhos

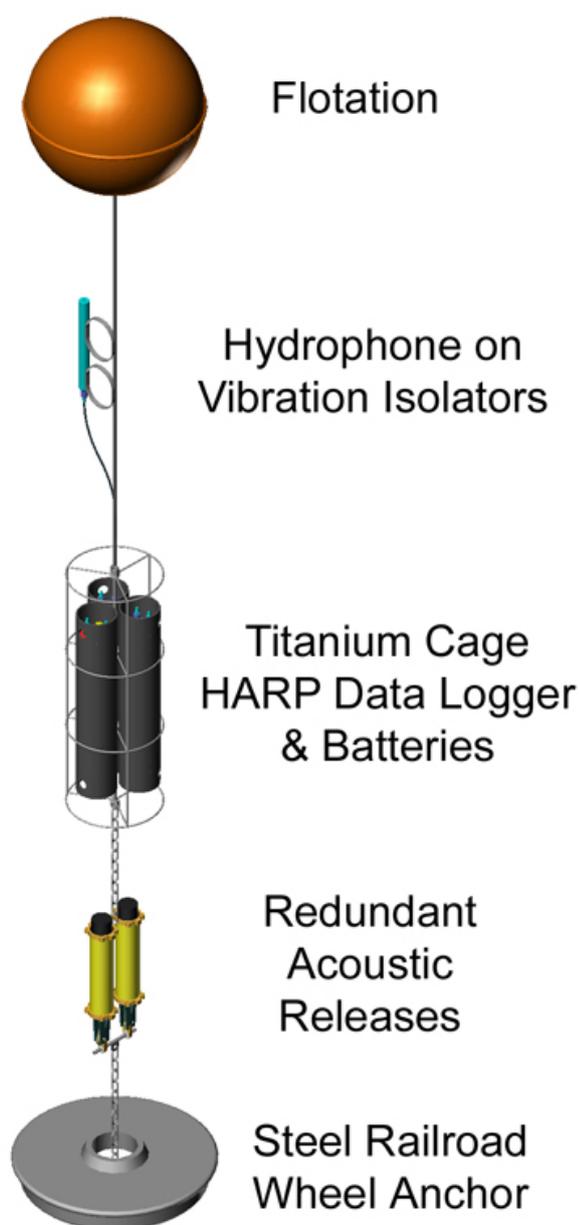


Figura 4 – Exemplo de um sistema AR, extraída de [32].

por períodos prolongados de tempo [22]. Os ARs também são importantes no estudo de peixes e outros organismos aquáticos e marinhos [39].

Um AR pode ser facilmente implantado no fundo do oceano para registrar dados acústicos por dias, semanas ou mesmo meses. Contudo, os arquivos devem ser recuperados para transferir dados para pós-processamento e análise por técnicos ou biólogos [31]. Os ARs são mais rentáveis quando utilizado em locais extremos ou remotos, onde o acesso é limitado ou difícil, bem como onde as condições ambientais existentes são muito duras para conduzir pesquisas a bordo de navios de investigação como, por exemplo, nas regiões polares [17, 35, 40].

Os ARs fornecem uma maneira custo-efetiva para determinar a presença, números relativos e distribuição da vocalização de mamíferos marinhos no espaço e no tempo. Os recursos dos ARs que são necessários para monitorar mamíferos marinhos variam de acordo com os objetivos e questões biológicas, o comportamento da produção de sons de espécies de interesse e as características do ruído do ambiente. Por exemplo, o monitoramento da ocorrência sazonal de Mysticetos geralmente requer monitoramento de vários meses a um ano. No entanto, como Mysticetos produzem sons de baixa frequência com boas características de propagação, as exigências para a cobertura espacial e taxas de amostragem é relativamente baixa (normalmente inferior a 1 kHz [41]) em comparação com as que seriam necessárias para controlar a maioria das espécies de Odontocetos em uma área semelhante.

Algumas regiões do mundo que são importantes para a exploração de petróleo e gás, também são áreas de alta densidade de mamíferos marinhos. Atividades relacionadas com a construção e instalação de plataformas e outras unidades de produção geram ruído subaquático [42]. O tráfego de navios tem demonstrado que causam distúrbios no comportamento de várias espécies de mamíferos marinhos, sendo o barulho do transporte o que mais contribuiu para o aumento dos níveis de ruído no fundo do oceano ao longo de uma década [43]. PAM fixo usando ARs em áreas de tais atividades humanas pode ser uma maneira eficaz para monitorar como o ruído está afetando potencialmente mamíferos marinhos, medindo quanto de seu habitat acústico está sendo perdido [44].

Baixas taxas de amostragem necessárias para gravar sons de baixa frequência também facilitam a economia de recursos de energia e capacidades de armazenamento para os ARs. Em geral, os Odontocetos produzem sons de médio (assobios em torno de 20 kHz) a alta frequência (cliques pulsados maior que 20 kHz) que não se propagam bem como os sons produzidos por Mysticetos (geralmente abaixo de 1 kHz). Isso ocorre porque frequências mais altas são atenuadas rapidamente e os sinais pulsados produzidos por Odontocetos muitas vezes têm modelos de feixe muito estreitos. Devido a estas características acústicas, sinais de Odontocetos poderão ser dispensados se os sensores não estiverem localizados perto do eixo do feixe. Como resultado, as escalas espaciais em que os Odontocetos podem ser monitorados são menores em comparação com os Mysticetos. Portanto, a cobertura de uma determinada área com sensores ARs para monitorar cliques de ecolocalização de Odontocetos (sons pulsados) deve ser relativamente densa, ou seja, mais sensores por unidade de área.

Os ARs possuem fonte de alimentação e armazenamento de dados eletrônicos independentes, e que limitam a capacidade destes sistemas, grande parte devido a necessidade de compensações entre fonte de alimentação, capacidade de armazenamento de dados, frequência de amostragem e tamanho do instrumento e classificação de profundidade. O aumento das necessidades de energia tem um efeito direto sobre o número de baterias

incluídas no pacote, por conseguinte, aumentando tanto o tamanho do instrumento quanto o requisito de flutuação. O tamanho do pacote do instrumento irá determinar os custos de instalação e remoção [11]. *Tradeoffs* entre estes componentes afetam o custo e a duração do período de captação [22]. Alguns elementos comumente levados em consideração na hora de escolher qual AR utilizar são:

- A capacidade do sistema exigida por uma aplicação específica. As principais limitações na duração de captação são a frequência de amostragem e a capacidade da bateria. Alguns ARs detectam ultra-som de alta frequência sem o uso de taxas de amostragem, permitindo longos períodos de execução com pequenas capacidades de bateria (por exemplo, A-TAG e C-POD). Outros ARs integram sensores adicionais para recolher dados não-acústico oceanográficos, tais como: a temperatura da água, a velocidade do som, a condutividade, a pressão e outros [45, 46];
- Condições ambientais e do tipo de substrato na área de implantação como, por exemplo, tipo de fundo do oceano, a presença de fortes correntes e ventos de superfície, batimetria, tráfego de navios e outros;
- Logística de implantação, que incluem o tipo de embarcação e equipamento de elevação disponível para implantação e recuperação como, por exemplo, guinchos, guindastes e mergulhadores. Todas estas características têm um impacto sobre o tipo e a configuração do instrumento (ou vice-versa) [22].

Entre as principais dificuldades para análise e interpretação dos dados acústicos recolhidos usando os ARs encontra-se a dificuldade de associar o número de sons gravados com o número de animais presentes, o volume de detecção associado à localização dos sons, assim como as variações sazonais, comportamentais e demográficas no padrão de emissão de diferentes espécies. Além disso, o fluxo e o ruído causado pelo movimento da água ao longo do hidrofone também pode ser um problema para os ARs em alguns ambientes [31].

Cada desenvolvedor de dispositivos ARs encontra uma solução diferente para gerenciar o tamanho, custo e a capacidade de compensação do sistema. As vantagens e desvantagens do sistema são aspectos críticos na escolha de um modelo de AR para aplicação durante a exploração de petróleo e gás, e atividades de produção ou qualquer estudo PAM. Tem-se por exemplo a disponibilidade de sistemas de amostragem (isto é, as gravações são feitas em intervalos predefinidos), onde um AR pode ser programado para gravar num ciclo de trabalho de 30 minutos com 30 minutos desativado (por exemplo, se o tempo de análise é uma restrição). Isto irá economizar energia, armazenamento e requisitos de pós-processamento.

Uma das principais limitações funcionais de um AR é o auto-ruído eletrônico [41]. Instrumentos que usam discos rígidos de giro ou outras partes mecânicas móveis podem

gerar sinais não desejados nas gravações que podem mascarar sons de interesse [22]. Além disso, o ruído de transporte (ou seja, o ruído do tráfego de navios) é o elemento mais importante para o aumento dos níveis de ruído no fundo do oceano nas últimas décadas [43].

Levando estas considerações em conta, como um usuário deve escolher um sistema AR? Em primeiro lugar, as questões e objetivos a serem abordados para um determinado estudo devem ser claramente definidas e consideradas. Isto irá por sua vez, determinar os requisitos do sistema de AR. Com base nos custos, capacidades e especificações de um sistema de AR, bem como problemas de implantação e recuperação relacionadas com a área monitorada, o usuário pode então considerar as opções disponíveis.

2.3 VOLUME DE DADOS

Independente da abordagem utilizada para captura, sempre gera-se um volume de dados considerável. Este volume de dados além de ser um dos maiores benefícios da utilização do PAM, dada a possibilidade de disponibilização de dados em praticamente tempo real, é também um dos maiores problemas, dada as limitações impostas pelo ambiente e as tecnologias disponíveis para processar toda essa quantidade de dados [4].

Dentre as limitações impostas, podemos citar como uma das mais importantes o gasto energético, que tem um maior impacto em abordagens móveis. Isso acontece devido a estações permanentes possuírem recursos de abastecimento energético não disponíveis nas estações móveis. Normalmente o abastecimento em dispositivos móveis é dado por baterias, com tempo de vida limitado e previamente estabelecido. Dessa forma, abordagens móveis normalmente acontecem de acordo com um plano de amostragem. Por outro lado, abordagens fixas normalmente realizam captura contínua [4].

Ainda considerando as limitações, mas dessa vez a capacidade de transferência de dados, temos que na maioria dos casos os dados são armazenados em disco, não sendo transferidos para estações remotas com especialistas capacitados. Isso acontece devido ao alto custo envolvido na implementação de transferência de dados no ambiente em questão. Dessa forma, os dados devem ser recuperados manualmente antes de iniciar uma análise de interesse [4]. Considerando a relação entre volume de dados processados e gasto energético como quase linear [47], a economia no processamento dos dados apresentada gera uma redução do gasto energético expressiva e de bastante interesse para dispositivos móveis.

Devido ao sucesso da computação móvel, um aumento expressivo em técnicas de gestão de energia emergiu nos últimos anos [47]. Convencionalmente, essas técnicas são focadas em diminuir a frequência do processador, acionando estados de dormência em consequência de inatividades. Esse tipo de abordagem tem sido tipicamente bem sucedido em reduzir o consumo global de energia [47]. Dependendo do equipamento e sistema

operacional utilizado, é possível manter o processador em estados de inatividades mais profundos. Para tal, a especificação aberta *Advanced Configuration and Power Interface* [48] (ACPI) padroniza o gerenciamento de energia em sistemas computacionais.

Mais recentemente os avanços em eletrônica de baixo consumo, armazenamento de dados de alta capacidade, tecnologia de processamento de computador, e as unidades de fornecimento de energia têm possibilitado o desenvolvimento e uso de ARs capazes de monitorar o ambiente acústico e comportamento de muitas espécies de mamíferos marinhos. Melhorias nas tecnologias de armazenamento de dados e baterias eletrônicas têm permitido o recolhimento de dados por longos períodos de tempo e com taxas de amostragem mais elevadas do que era possível anteriormente [22].

Para fornecer registros contínuos de sinais acústicos de Odontocetos por longos períodos identificáveis usando um instrumento autônomo, existem três requisitos principais para o sistema eletrônico de aquisição de dados [22]:

- Baixo consumo de energia;
- Digitalização de alta velocidade;
- Armazenamento de dados de alta capacidade.

Como acontece com qualquer instrumento autônomo alimentado por bateria, os componentes de baixa potência são essenciais para implantações de longa duração. Digitalização de alta velocidade é necessária para gravar os sinais acústicos de Odontocetos de banda larga e para fornecer largura de banda suficiente para a diferenciação de chamadas. A digitalização de alta velocidade juntamente com gravações de longa duração requerem armazenamento de dados de alta capacidade a ser utilizado. Na maioria dos casos, o armazenamento de dados de alta capacidade é conseguido usando vários discos rígidos ou cartões de memória, que exigem um microcontrolador e do *firmware* dedicado a controlar o processo de gravação de dados, como por exemplo a abordagem utilizada pelo *High-frequency Acoustic Recording Package* (HARP). A taxa de aquisição de dados é indiretamente relacionada com o consumo de energia, pois determina a frequência com que o disco rígido terá de ser acessado e gravado. Esses recursos são importantes para entender quando se escolhe o melhor AR disponível para uma determinada aplicação [22].

O disco rígido padrão *pop-up* armazena na ordem de centenas de GB de dados, por conseguinte, as altas taxas de amostragem, a capacidade de armazenamento de dados, em oposição ao fornecimento de energia, podem limitar a duração do monitoramento AR. A mudança de discos rígidos para cartões de memória *flash* de alta capacidade de armazenamento irá cuidar dessa limitação. O HARP tem uma capacidade de armazenamento muito maior (na ordem de alguns TB) do que os *pop-ups*. A configuração padrão

de energia (estimado em 330 Amp-horas usando pilhas alcalinas de tamanho D 192), grava continuamente a uma taxa de amostragem máxima preenchendo os discos rígidos antes da capacidade da bateria ser alcançada [16]. Nos últimos anos, a memória *flash* em estado sólido caiu significativamente de preço e aumento de capacidade, oferecendo uma alternativa aos discos rígidos que são maiores e mais pesados [22].

Ao determinar que tipo de sistema de amostragem ou ciclo de trabalho que se deve usar, é importante ter uma ideia sobre o comportamento acústico das espécies-alvo a ser monitorada. Notavelmente, estar ciente da faixa de frequência da espécie alvo e também do tempo da atividade de vocalização, irá apoiar decisões sobre os ARs a serem adotados [22].

Inicialmente a prospecção acústica de médio a longo prazo deve ser concluída para determinar quais espécies estão presentes, que tipos de sons elas produzem e quantas vezes eles são produzidos. No entanto, a quantidade de dados gerada por gravações contínuas é excessivamente grande (1 ano pode resultar em mais de 2 TB de dados em taxas de amostragem com menos de 2 kHz) de tal forma que esquemas de detecção ou de amostragem automáticas devem ser aplicadas durante o pós-processamento e análise [22].

3 UM MODELO DE FLUXO DE TRABALHO PARA MONITORAMENTO ACÚSTICO PASSIVO

O processo completo de Monitoramento Acústico Passivo é longo e complexo, podendo ser representado computacionalmente como um Fluxo de Trabalho (*Workflow*). O termo Fluxo de Trabalho originou-se pelo processo de automação de escritórios, tendo como principal objetivo sistematizar o processo de criação, armazenamento, compartilhamento e revisão de documentos. Por conseguinte, o tempo gasto pela troca e disponibilização de conhecimento entre os envolvidos no processo se tornou cada vez menor, otimizando desta forma a execução de tarefas e dissipando gastos desnecessários embutidos no processo [49].

A difusão computacional dos últimos anos resultou na criação de diversas ferramentas e notações para representação gráfica de um Fluxo de Trabalho. Apesar disso, nem todas estas ferramentas e notações são capazes de representar padrões comumente utilizados na modelagem de um Fluxo de Trabalho genérico [50]. Testes comparativos entre duas notações amplamente utilizadas foram realizados [51], onde considerou-se 20 padrões distintos [50], modelados utilizando Diagramas de Atividade [52] e Notação e Modelo de Processos de Negócio (*Business Process Model and Notation*, BPMN) [53], ambos mantidos pelo *Object Management Group* (OMG) [54]. O Diagrama de Atividades conseguiu modelar 19 dos 20 padrões propostos, apresentando dificuldades somente com Rotas Paralelas Intercaladas. Em contrapartida, o BPMN foi capaz de modelar todos os padrões propostos. Os Fluxos de Trabalho representados graficamente nesta tese se baseiam na premissa da simplicidade proposta pelo BPMN, bem como na vasta quantidade de ferramentas que implementam seu conjunto de notações.

Atualmente o conceito de Fluxo de Trabalho se estende a diversos ambientes, sendo utilizado para descrever as atividades necessárias para automatizar um determinado processo a partir de um conjunto de regras bem definidas. Neste âmbito o conceito de Fluxo de Trabalho se expande para o ambiente científico, focando principalmente em experimentos científicos de alta escala [55].

No ambiente científico temos que dentre as diversas tarefas desenvolvidas por pesquisadores, algumas dizem respeito a composição de algoritmos computacionais. Um algoritmo pode ser representando internamente em um Fluxo de Trabalho como uma atividade, gerando saídas que são encadeadas como entradas de outras atividades. Ressalta-se que o processo de composição não é uma tarefa trivial, e em muitos casos torna-se uma barreira para análises mais sofisticadas. Em contrapartida, a execução de um Fluxo de Trabalho previamente definido pode até mesmo ser utilizado por indivíduos sem conhecimento prévio do experimento científico [56].

Atualmente os experimentos científicos são divididos em quatro categorias, sendo duas mais tradicionais e utilizadas há décadas. A primeira delas é conhecida como *in*

vivo, onde os experimentos são realizados no ambiente natural do objeto em estudo. Já na segunda categoria, *in vitro*, os estudos são conduzidos em ambientes mais controlados, como laboratórios de pesquisa e universidades. As outras duas categorias são mais recentes, conhecidas como *in virtuo* e *in silico*. Os experimentos *in virtuo* são caracterizados pela participação interativa dos objetos de estudo em um ambiente simulado. A última delas conhecida como *in silico*, tanto os objetos de estudo como o ambiente são simulados computacionalmente [57].

Os experimentos *in silico* são comumente mais utilizados que *in virtuo*, impulsionados pelos avanços computacionais nas últimas décadas. Diversas áreas de pesquisa usufruem dos benefícios oriundos à execução de experimentos *in silico*, tais como meteorologia, bioinformática e exploração de petróleo. Este tipo de experimento executado via computador costuma demandar uma capacidade computacional extremamente elevada, podendo demorar meses para serem executados, grande parte devido à quantidade de dados e complexidade dos modelos a serem explorados [56, 58]. Apesar disso, um Fluxo de Trabalho quando bem planejado por especialistas computacionais juntamente com pesquisadores da área de interesse, provê a abstração necessária na criação de um ambiente robusto para execução de experimentos *in silico*. Promove-se assim soluções efetivas com recursos computacionais otimizados, sendo neste caso mais conhecido como Fluxo de Trabalho Científico [56].

Como o próprio nome sugere, um Fluxo de Trabalho Científico é composto por um conjunto de atividades científicas, e o respectivo fluxo de dados entre as mesmas. As atividades normalmente são vistas como tarefas abstratas, podendo ser um algoritmo computacional capaz de processar um conjunto de dados de entrada, bem como tarefas manuais executadas por cientistas. O fluxo entre as atividades é caracterizado pelas suas dependências, sendo a saída da atividade atual a entrada da próxima [59]. Além disso, existe a possibilidade de uma atividade de um Fluxo de Trabalho ser inteiramente composta por outro Fluxo de Trabalho, dessa forma, os experimentos simulados são compostos de diversos modelos computacionais embutidos [58].

Para gerenciar a complexidade e interação entre Fluxos de Trabalho, Sistemas de Gerência de Fluxo de Trabalho Científico (*Scientific Workflow Management System*, SWfMS) [60, 61, 62] foram criados. Os SWfMS são utilizados para orquestrar o fluxo de execução das atividades bem como os recursos utilizados, podendo ser locais ou remotos [58]. Algumas implementações de SWfMS tiveram seu potencial avaliado com sucessos em diversas áreas da ciência [63, 64, 65, 66].

Uma das maiores vantagens disponíveis em um SGWfC é a possibilidade de realizar a proveniência dos resultados a cada etapa de execução do experimento. No contexto da ciência em larga escala, a proveniência ajuda a determinar o histórico de um determinado experimento a partir de suas origens [67]. A proveniência pode ser caracterizada como

prospectiva ou retrospectiva. A prospectiva está relacionada com as atividades e suas características, tais como a relação entre as atividades. A retrospectiva por sua vez, considera mais os recursos computacionais como, por exemplo, o tempo de execução de cada atividade do experimento [59].

Apesar das grandes vantagens de utilização de um SWfMS, quando considera-se sua aplicação em um dispositivo fixo autônomo, sua implementação se torna muito difícil, principalmente devido ao gasto computacional e interação visual com o usuário na maioria dos casos. Tendo a eficiência computacional como foco deste trabalho, algoritmos computacionalmente custosos¹ para gerência de execução em SWfMS acabariam invalidando parte dos ganhos alcançados pelas abordagens propostas. Além disso, a dificuldade de interação visual necessária na maioria dos SWfMS, inviabilizaria sua implementação em ambientes não conectáveis.

3.1 FLUXO DE TRABALHO DE REFERÊNCIA

Basicamente o processo de referência envolve três etapas principais [11], como mostrado na Fig. 5. A primeira etapa conhecida como detecção, tem como objetivo avaliar se em determinada parte da captura existe algum sinal considerado de interesse. Antes da existência dos métodos de detecção automática, esta etapa era realizada por especialistas que revisavam as capturas a partir de análises auditivas em conjunto com anotações visuais [2, 3]. Contudo, devido ao grande volume de dados, a detecção manual se tornou impraticável, sendo cada vez mais executada por abordagens automáticas. Apesar disso, abordagens automáticas ainda requerem a validação de especialistas indicando que sinais de interesse não são perdidos [22]. Destacam-se neste caso as técnicas que dependem de limiares definidos previamente pelo usuário, onde é realizada uma validação visual de exemplares de modo a ajustar o limiar com intuito de obter boas detecções.

3.1.1 Detecção

O modelo de detecção utilizado é de interesse tanto pelo custo, quanto pela estratégia de gerência [11]. Neste caso, a etapa de detecção trabalha como um filtro no processo, sendo analisado somente uma fração reduzida dos dados obtidos. Com a redução da quantidade de dados totais que devem ser analisados, uma diminuição do tempo gasto por especialistas é esperada.

Uma das formas mais comuns de reduzir o volume de dados é a definição prévia de intervalos temporais para captura. Nesta abordagem, o dispositivo de captura é ativado somente em determinados momentos, conseqüentemente armazenando um volume de dados

¹ Por computacionalmente custoso entende-se elevado conjunto de instruções para executar funcionalidades oriundas a outros problemas, incorporadas de maneira genérica e normalmente executadas por máquinas virtuais como, por exemplo, o tradutor dinâmico do Java.

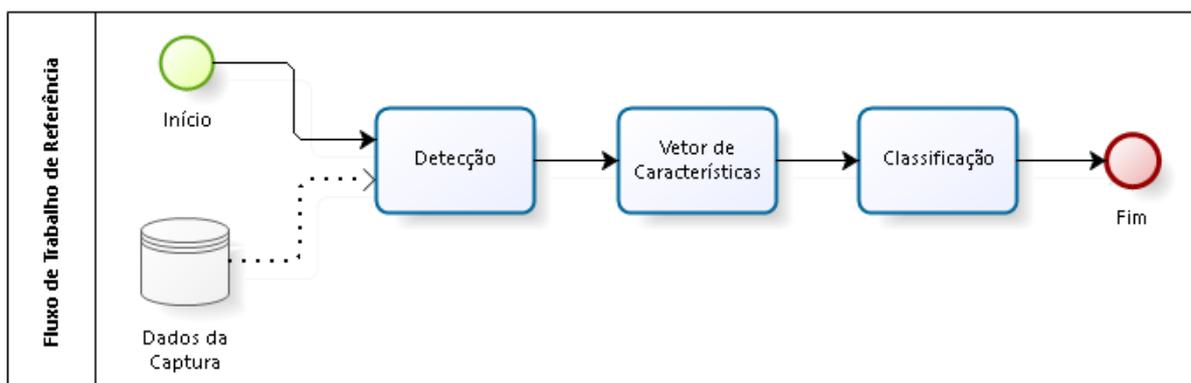


Figura 5 – Fluxo de Trabalho geral utilizado como referência.

menor. Abordagens baseadas em intervalos são propícias a uma perda de dados maior, podendo existir a possibilidade das gravações serem realizadas somente em momentos onde nenhum mamífero marinho está presente.

Com o auxílio da computação, este processo pode ser gradativamente substituído por algoritmos de ativação por detecção. Algoritmos de ativação por detecção são técnicas computacionais capazes de processar partes da captura somente caso exista algum animal presente. É importante ressaltar que a análise realizada a partir de algoritmos de ativação por detecção é feita nos laboratórios em terra, navios ou plataformas, diferenciando-se da proposta deste trabalho que é viabilizar sua implantação em dispositivos fixos autônomos. Para avaliar se um animal está presente ou não, diversos métodos podem ser utilizados [31].

Para a avaliação de presença, algoritmos de Detecção de Limiar de Energia (*Energy Threshold Detectors*) [68, 69, 70, 71, 72] são uma das formas mais comuns de detecção de mamíferos marinhos. Existem diversos outros métodos como, por exemplo, Detector de Entropia (*Entropy Detector*) [73, 74, 75], Teste de Page (*Page's Test*) [76, 77], Análise de Componente Principal (*Principal Component Analysis*, PCA) [78], Algoritmo de Schur (*Schur Algorithm*) [79, 80] e *Wavelets* [81]. Apesar de reduzirem a quantidade de dados a serem analisados, os métodos apresentados possuem uma complexidade computacional limitante quando executados em dispositivos fixos autônomos, em especial quando aplicados em altas frequências com intuito de detectar Odontocetos. Além disso, a maioria deles foram projetados para trabalhar com chamados de Mysticetos, apresentando resultados não satisfatórios quando aplicados em Odontocetos, em especial para cliques. Mais detalhes de cada um dos algoritmos apresentados podem ser encontrados em [11].

3.1.1.1 Detecção de Limiar de Energia

Um dos métodos mais comuns para a detecção de mamíferos marinhos é medir a energia e amplitude em uma frequência de banda específica do sinal de entrada e

determinar se ele excede um valor de limiar definido pelo usuário. Essa comparação é normalmente realizada consecutivamente em cada espaço de tempo, e o valor de limiar é tipicamente definido em relação ao ruído do ambiente na banda de frequência de interesse [68, 69, 70, 71, 72]. A primeira vez que um sinal excede o limite é considerado como sendo o início de um chamado. Informações relacionadas com a frequência de pico nos próximos espaços de tempo são armazenados e analisados. O fim de um chamado ocorre quando nenhum pico adicional excede o limiar definido pelo usuário. Embora seja um método simples e eficaz, o detector de limiar de energia sofre com sinais com uma taxa de SNR (*Signal-to-Noise Ratio*) baixa, bem como na presença de sinais sobrepostos na mesma banda de frequência a partir de múltiplas fontes. Normalmente o limiar de energia é definido pelo usuário após uma rápida análise visual e auditiva dos dados capturados.

3.1.1.2 Detector de Entropia

A entropia de Shannon é uma medida da quantidade de informação contida em um sinal. Ela tem sido usada para caracterizar e identificar os sinais de mamíferos marinhos, já que a maioria dos seus sons se diferem claramente do ruído ambiente quando calculado sua energia espectral [73, 74, 75].

Este método caracteriza a probabilidade de distribuições em termos do seu grau de energia espectral, onde uma forte distribuição é descrita como tendo baixa entropia. Dado que densidades espectrais acústicas são distribuições aditivas positivas, elas podem ser tratadas como distribuições probabilísticas enquanto seus somatórios são normalizados para 1 [82].

O detector de entropia superou as abordagens básicas de detecção de energia, inclusive para casos de SNR baixo. Outra vantagem dessa abordagem é sua execução consideravelmente rápida, podendo ser executada em tempo real. Entretanto, sua maior vantagem é sua capacidade de detectar chamados de diversas espécies, não sendo obrigatório uma configuração prévia da espécie de interesse. Um dos maiores problemas dessa abordagem é a quantidade elevada de alarmes falsos que são armazenados como detecções. É importante notar que esse tipo de detector é bom para chamados e sinais tonais, sendo ruim para detecções de sinais pequenos e com largura de banda alta, como é o caso dos cliques emitidos por Odontocetos.

3.1.1.3 Teste de Page

Uma técnica alternativa para determinar os tempos de início e fim de um sinal acústico é o Teste de Page [77, 76]. Page desenvolveu este procedimento para determinar o momento em que ocorre uma alteração nos dados obtidos sequencialmente. Ele é capaz de minimizar o tempo de atraso médio antes da detecção, enquanto se restringe o tempo médio entre os alarmes falsos. O teste de Page é apropriado para detecções de espécies

específicas, como as Baleias Bicudas de Blainville (*Mesoplodon densirostris*), chamados de Golfinhos de Risso (*Grampus griseus*) e Baleia Piloto de Aleta Curta (*Globicephala macrorhynchus*).

3.1.1.4 Análise de Componente Principal e Análise de Componente Independente

O PCA (*Principal Component Analysis*) em conjunto com o ICA (*Independent Component Analysis*) foram utilizados para detecção de chamados de Mysticetos [78]. O PCA é uma transformação linear simples que é usada para separar os dados. O ICA é similar ao PCA, e separa os sinais de múltiplos receptores em várias características não gaussianas (neste caso, os chamados dos Mysticetos) e ruídos gaussianos. O PCA foi utilizado como uma etapa de pré-processamento para o ICA, bem como testado sua efetividade sozinho. Sinais a partir de vários receptores foram pela primeira vez alinhados através de correlação cruzada, a fim de assegurar que uma determinada característica acústica aparecesse na mesma janela de tempo para todos os receptores. O algoritmo PCA foi utilizado para transformar os dados em janelas, seguidas da aplicação do ICA em cada posição de tempo, e os resultados estatísticos foram comparados com um valor limiar para determinar se características de interesse tinham sido detectadas.

3.1.1.5 Algoritmo de Schur

O algoritmo de Schur adaptativo baseia-se na normalidade em mínimos quadrados e no filtro de rede variante no tempo. Ele está bem adaptado para a análise de dados de séries temporais não estacionárias [79, 80]. O algoritmo é aplicado a cada espaço de tempo, e o sinal é ortogonalmente projetado sobre valores passados do sinal para produzir um conjunto de coeficientes de Schur variáveis no tempo.

O ruído ambiente resulta em pequenas variações nestes coeficientes, enquanto sinais como cliques de Cachalote produzem uma drástica mudança nestes coeficientes. Em ambientes simulados, o Algoritmo de Schur foi capaz de detectar 100% dos dados, com somente 10% de detecções incorretas (falso positivos). Entretanto, não há evidências de que quando aplicado em sinais com gravações reais - sons naturais e antropogênicos - os resultados apresentem a mesma taxa de detecção.

3.1.1.6 Transformada *Wavelet*

Embora os espectrogramas sejam um dos métodos mais comuns para representar sinais de bioacústica, eles estão sujeitos ao balanceamento de tempo-frequência devido ao princípio da incerteza. Este princípio afirma que é impossível saber a frequência exata e o tempo exato de ocorrência da frequência em um sinal. Em outras palavras, um sinal não pode ser representado como um ponto exato em um espaço tempo-frequência, e aumentar a resolução de um diminui a resolução do outro. A transformada *Wavelet* é

uma técnica desenvolvida para superar essa limitação. Na análise *Wavelet* o espectro do sinal é calculado com uma janela modulada (como um banco de filtros constantes), ao contrário de uma janela fixa como utilizado em transformadas de Fourier. A geração de espectro é repetida muitas vezes com janelas de tamanhos diferentes, resultando em um conjunto de representação do sinal com diferentes resoluções tempo-frequência. Em geral, as frequências mais altas são melhor resolvidas no tempo e frequências mais baixas são mais bem resolvidas na frequência [83].

3.1.2 Vetor de Características

A segunda etapa do Fluxo de Trabalho é definida pela criação de um vetor com características relevantes do sinal de interesse. O vetor de característica armazena os atributos de entrada para a próxima etapa, onde o sinal detectado é classificado em determinada espécie [4, 5]. Dentre as diversas características que podem ser utilizadas, o conjunto selecionado e a precisão da criação possui uma influência significativa no sucesso da classificação. As características escolhidas devem prover o máximo de informações, levando em consideração que a quantidade de informações deve ser suficiente para uma correta classificação, mas sem impactar no gasto computacional [11].

As características dos espectrogramas são provavelmente os atributos mais utilizados para classificação de mamíferos marinhos. Este tipo de característica normalmente é mais fácil de ser coletada, principalmente por técnicas bastante evoluídas de processamento de imagens. São algumas delas, frequência inicial, final e central em conjunção com outras variáveis temporais, tais como duração do sinal e sub-sinais e fase [11]. Características do espectrograma mais complexas também já foram utilizadas, como é o caso do Formato da Vocalização [84, 85, 86], Traçado do Pico de Frequência [87, 88, 89] e Ajuste da Curva [90].

Embora as características do espectrograma sejam as mais comuns, existem outras características que podem ser utilizadas para classificação de mamíferos marinhos, tais como, Variância Mínima de Capon (*Capon's Minimal-Variance*) [91], Decomposição do Sinal Baseado em Deformação (*Warping-Based Signal Decomposition*, WSD) [81], Filtro de Partículas (*Particle Filter*) [92], Características Cepstrais [93] e Predição Linear Perceptiva Generalizada (*Generalized Perceptual Linear Prediction Model*) [93]. Uma descrição mais abrangente de cada característica apresentada pode ser encontrada em [11].

3.1.2.1 Traçado do Pico de Frequência

O traçado do pico de frequência utiliza o contorno da frequência dos sinais detectados como de interesse na hora da criação do vetor de características. O método mais comum para realizar essa tarefa é a combinação em um contorno único dos espaços temporais que estão acima de um limiar de energia definido pelo usuário. Outra definição do usuário neste caso é a banda de frequência de interesse [87, 88, 89, 94].

Problemas aparecem quando múltiplos espaços temporais estão sobrepostos no mesmo espaço utilizado para criação do contorno. Critérios adicionais podem ser utilizados para melhorar a extração do contorno, obtendo de 80-90% no sucesso do processo [88, 94].

3.1.2.2 Ajuste da Curva

Quando separados em segmentos, os sinais de entrada podem ser utilizados para calcular modelos lineares utilizando Estimadores de Verossimilhança Máxima (Maximum Likelihood Estimator, MLE). Modelos contínuos que não se diferenciaram na frequência por mais que um limite definido previamente foram re combinados em um contorno completo, e os parâmetros do modelo linear foram utilizados como vetor de características. Este modelo é capaz de detectar somente um sinal em cada segmento, impossibilitando que seja utilizado para sinais com sobreposição de segmentos. Como essa técnica é aplicada sequencialmente para segmentos pequenos, ela não se aplica bem a sinais com SNR baixo [90].

3.1.2.3 Variância Mínima de Capon

Leprettre e Martin usaram um estimador espectral conhecido como Variância Mínima de Capon para criar uma representação de frequência de tempo (*Time Frequency Representation*, TFR) do sinal de entrada semelhante a um espectrograma, mas sem as mesmas limitações. Desenvolvida originalmente para a estimativa do número da onda no processamento de sinal em arrasto [95], a técnica mede a potência de um conjunto de filtros de banda estreita uniformes para se adaptar ao sinal e reduzir a resposta das componentes espectrais fora da banda de interesse. Após o estimador de Capon, é utilizado o conhecimento a priori de chamados do mamífero marinho de interesse para projetar um critério capaz de distinguir entre estes chamados e ruídos ambientes.

3.1.2.4 Decomposição do Sinal Baseado em Deformação

O processo de decomposição do sinal para análise da representação tempo-frequência de sinais com estruturas não lineares foi desenvolvido por Ioana *et al.* [81]. Chamados de Cetáceos são geralmente não lineares, e isto coloca um problema para análises típicas baseadas em tempo-frequência. A técnica WSD primeiro lineariza as estruturas de tempo-frequência através de uma bateria de operadores de deformação, e em seguida coleta parâmetros relevantes para as estruturas de tempo-frequência mais representativas que descrevem os chamados de interesse. A representação criada não possui as limitações de balanceamento devido a incerteza temporal, consequentemente resultando em parâmetros para o vetor de característica mais precisos.

3.1.2.5 Filtro de Partículas

A Transformada de *Fourier* Fracionada de Tempo Curto (*Short-time Fractional Fourier Transform*, STFrFT) pode ser conceituada como a representação da energia de um sinal em um determinado tempo-frequência e associado com uma taxa de alcance particular. O espectrograma clássico é um caso especial do STFrFT, em que a taxa de alcance é ignorada. Ao incluir a taxa de alcance no algoritmo, o STFrFT é capaz de aumentar o SNR de um sinal, desde que a taxa de alcance escolhida seja próxima da taxa de alcance real inerente aos chamados da espécie de interesse. Para criação de um vetor de características de assobio de golfinhos de água salgada (família Delphinidae), um filtro de partículas chamado de Re-amostragem de Importância Sequencial (*Sequential Importance Resampling*) foi utilizado. O filtro de partículas infere os parâmetros do sinal recursivamente, submetendo amostras de sinais de entrada ruidosos para um algoritmo de Monte Carlo Sequencial (*Sequential Monte Carlo*, SMC) e realiza o ajuste dos modificadores de semelhança de acordo com suas iterações. Apesar de ter mostrado na literatura bastante sucesso na extração de assobios de golfinhos, é possível observar que o método é computacionalmente caro e requerer modificações para otimizá-lo à execução em tempo real.

3.1.2.6 Características Cepstrais

Os coeficientes Cepstrais são características espectrais dos sinais bioacústicos comumente utilizadas no processamento de fala humana. Elas são bem adequadas para os modelos de reconhecimento de padrões estatísticos, já que os coeficientes Cepstrais tendem a ser não correlacionados uns com os outros [93]. Dessa forma, o número de parâmetros que devem ser estimados são reduzidos consideravelmente. As características Cepstrais são calculadas utilizando uma Transformada Rápida de Fourier (*Fast Fourier Transform*, FFT) em cada segmento do sinal, resultando em uma Transformada de Fourier de Tempo Curto (*Short-Time Fourier Transform*, STFT). O eixo de frequência é então deformado através da multiplicação do espectro com uma série de n filtros em frequências adequadamente espaçadas.

Ao utilizar características Cepstrais o timbre dos sons são capturados, sendo está uma característica perdida quando se utiliza espectrogramas. Roch desenvolveu um sistema de detecção automática baseado em características Cepstrais para detecção de assobios, sons pulsados e cliques produzidos por golfinhos [96]. Devido a sua generalidade, não dependendo de um tipo específico de sons, o sistema foi considerado bom ao obter uma taxa de classificação correta de 65-75%, dependendo do particionamento e dados de treinamento.

3.1.2.7 Modelo de Predição Linear Perceptual Generalizada

O modelo de Predição Linear Perceptual Generalizada (*Generalized Perceptual Linear Prediction*, GPLP) para criação do vetor de características, incorpora informações sobre as habilidades de percepção e as características do aparelho vocal (ou sacos de ar em mamíferos marinhos) de espécies em estudo para o cálculo de características relevantes. Ele baseia-se no modelo de Predição Linear Perceptual (*Perceptual Linear Prediction*, PLP) desenvolvido por Hermansky para a fala humana [97]. O primeiro passo no processo de criação do vetor de características é filtrar a vocalização utilizando um filtro de "pré-ênfase", que reduz o intervalo dinâmico do espectro de modo que seja mais facilmente aproximado através do passo de modelagem auto-regressivo da análise. A vocalização é então dividida em segmentos e o espectro de energia é estimado. Este espectro de energia é transformado para explicar vários fenômenos psicoacústicos.

Primeiro, um banco de filtros triangular é aplicado ao espectro de potência para contabilizar tanto o mascaramento da frequência de banda crítica, quanto o mapeamento não linear entre a posição e a sensibilidade de frequência. Uma forma de filtro triangular simples é usada, pois poucos dados estão disponíveis nas formas de filtro auditivas de animais, com exceção dos seres humanos. Uma vez que as energias de banco de filtro terem sido calculadas, uma curva de igual intensidade (aproximada do audiograma de uma espécie) é utilizada para normalização [93]. Finalmente, a modelagem auto-regressiva é realizada. Nesta etapa as energias dos bancos de filtros normalizados são aproximados por um modelo que utiliza o método de autocorrelação e as equações de Yule-Walker [98]. Os coeficientes gerados durante este passo são transformados no domínio Cepstral usando uma fórmula de recorrência. Estes coeficientes finais são em grande parte não correlacionados e representam a forma do filtro durante a produção de vocalização. Isso resulta em vetores de características que levam em conta as habilidades de percepção da espécie [93].

3.1.3 Classificação

A última etapa considera o vetor de característica criado anteriormente para classificação do sinal em uma espécie de mamífero marinho, ou grupo de espécies [22]. A etapa de classificação de sinais acústicos é um processo bem estudado, graças a sua ampla utilização em diversas áreas de conhecimento. Para sinais bioacústicos o processo possui um nível de complexidade distinto, devido a algumas características ambientais que afetam o sinal como, por exemplo, variação do caminho de propagação da onda, orientação do animal e similaridade do contorno de sinais produzidos por algumas espécies. Além disso, alguns sinais naturais ou antropogênicos dificultam a classificação, podendo até serem detectados como mamíferos marinhos [11].

Diversos métodos para classificação de mamíferos marinhos estão sendo desenvolvidos, podendo-se destacar as Redes Neurais Artificiais (*Artificial Neural Network*, ANN)

[99, 100], Máquinas de Vetores de Suporte (*Support Vector Machine*, SVM) [70, 101, 102], Análise de Função Discriminante Multivariada (*Discriminant Function Analysis*, DFA) [103, 104], Modelos de Mistura de Gauss (*Gaussian Mixture Model*, GMM) [96, 105] e Picos Espectrais (*Spectral Peak*) [106]. Uma apresentação mais detalhada é apresentada em [11].

3.1.3.1 Redes Neurais Artificiais

As Redes Neurais Artificiais (*Artificial Neural Networks*, ANNs) foram desenvolvidas por meio da modelagem de sistemas biológicos de processamento de informação e tornou-se muito popular na década de 1980 nas áreas de reconhecimento acústico de palavras [107] e de reconhecimento de caracteres ou imagem [108, 109]. Desde então, as ANNs têm sido bem sucedidas na classificação de um número variado de sinais complexos, incluindo a fala humana [110], cliques de ecolocalização [71, 111, 112], Cão da Pradaria [113], e Porcos domésticos [114].

Todas as ANNs são compostas de unidades chamadas de neurônios e as conexões entre elas. Elas consistem tipicamente de três ou mais camadas de neurônios: uma camada de entrada, uma camada de saída e uma ou mais camadas ocultas. A camada de entrada é constituída por n neurônios que codificam para n características no vetor de características. A camada de saída é constituída por k neurônios que representam as k classes que se deseja classificar. O número de camadas escondidas entre a camada de entrada e a camada de saída, bem como o número de neurônios por camada, são empiricamente escolhidos pelo investigador. Cada conexão entre os neurônios da rede está associada a um valor do peso que é modificado por interações sucessivas durante a formação da rede.

As ANNs são promissoras para a classificação automática de sinais por várias razões. Em primeiro lugar, a entrada para uma ANN pode variar desde vetores de características das medições realizadas a partir de espectrogramas quanto do formato da onda. É importante notar, contudo, que o balanceamento entre resolução temporal e frequência inerente a criação do espectrograma pode afetar drasticamente a representação de um sinal e como ele é classificado. Por exemplo, muitos espectrogramas diferentes podem ser produzidos a partir do mesmo sinal fazendo somente pequenas alterações no tamanho do segmento utilizado para aplicação da FFT [115]. Cuidados devem ser tomados, portanto, ao usar espectrogramas como entrada para as ANNs. Em segundo lugar, as ANNs servem como classificadores adaptativos que aprendem através de exemplos. Como resultado, para desenvolver um bom modelo matemático não são necessárias características do sinais antes que análises subjacentes comecem [116].

Além disso, as ANNs são estimadores não lineares. Isso significa que elas estão bem adaptadas para os problemas que envolvem distribuições arbitrárias e de entrada ruidosa [116, 117]. Uma desvantagem comum das ANNs é que elas são consideradas caixas preta

e, portanto, é difícil determinar exatamente como estão tomando decisões de classificação. No entanto, Potter *et al.* foram capazes de investigar os papéis dos neurônios escondidos por linearização das funções de transferência [117]. Usando essa técnica, os autores foram capazes de “interrogar” as ANNs para revelar seus paradigmas de operação e confirmar se foram classificadas com base em características relevantes.

Os cliques de ecolocalização de golfinhos foram classificados em sete categorias definidas pela frequência de pico e informações de largura de banda, utilizando uma rede de contra-propagação por Houser *et al.* [118]. Suas pontuações de classificação corretas variou de 45,5-82% para os cliques produzidos por diferentes golfinhos individuais.

Finalmente, Mellinger também usou uma ANN *feed-forward* treinada usando *Back-propagation* para classificar cliques de Baleias Bicudas de Blainville [71]. Espectrogramas condicionados (equalizado, retificado e normalizado) foram utilizados como entrada para a ANN. Dos 99% da taxa de classificação correta, apenas 0,6% dos cliques de Baleias Bicudas foram perdidos. Mellinger postulou que uma das razões para o bom desempenho deste classificador é que os dados de treinamento e de teste foram extraídos das mesmas gravações, sendo provavelmente das mesmas baleias.

3.1.3.2 Modelo de Mistura Gaussiana

O Modelo de Mistura Gaussiana (*Gaussian Mixture Models*, GMM) é comumente usado para modelar distribuições arbitrárias, como as combinações lineares de distribuições paramétricas. Ele é apropriado para a identificação de espécies quando não há expectativa quanto a sequência de chamados e quando vários chamados podem ocorrer simultaneamente [96]. Para criar um GMM, um conjunto de n distribuições normais com meios separados e matrizes de covariância diagonal são escalonados por fatores de peso c_i ($1 < i < n$). A soma dos c_i s deve ser 1 para assegurar que o GMM represente uma distribuição de probabilidade [96, 105, 106]. O número de misturas é escolhido empiricamente e os seus parâmetros são estimados utilizando um algoritmo iterativo, tal como o algoritmo de Maximização de Expectativa (*Expectation Maximization*, EM). Uma vez que o GMM foi treinado, a probabilidade é calculada para cada um dos vetores de características e um teste de Razão de Verossimilhança é usado para decidir quais as espécies a que pertence [106].

O GMM pode ser implementado usando o *Hidden Markov Model Toolkit* (HTK), que é uma conjunto de códigos abertos de programas projetados para o reconhecimento da fala humana [96]. Roch *et al.* usaram o GMM e vetores de características Cepstrais para classificar assobios, cliques e pulsos produzidos por golfinhos (*Delphinus delphis*, *Delphinus capensis* e *Lagenorhynchus obliquidens*) com 67-75% de classificação correta, dependendo do particionamento dos dados de treinamento e de teste e do número de misturas utilizadas.

Em um estudo posterior, Roche *et al.* classificaram cliques produzidos pelas Baleias Bicudas de Blainville, Baleias-Piloto (*Globicephala*) e Golfinhos de Risso usando GMM. Os pontos de classificação corretos para essas três espécies foram 96,7%, 83,2% e 99,9%, respectivamente [106]. O potencial do GMM também foi avaliado por Brown e Smaragdis para classificação de chamados de Orcas (*Orcinus orca*), resultando em concordância de até 92% com as categorias perceptivelmente criadas a partir de tipos de chamados.

O GMM também tem sido utilizado com sucesso para classificar os chamados do tipo A e B produzidos por Baleias Azuis no nordeste do Pacífico [119] e quatro espécies registradas no mar Chukchi (Baleia Branca, Baleia da Groenlândia, Orca e Baleia Jubarte).

3.1.3.3 Máquinas de Vetores de Suporte

Máquinas de Vetores de Suporte (*Support Vector Machines*, SVMs) são uma rica família de Algoritmos de Aprendizagem com base no trabalho de Vapnik e Chervonenkis em teoria de aprendizagem estatística. De acordo com essa teoria, o melhor classificador é o que minimiza tanto o erro de treinamento quanto a complexidade do classificador [101]. De acordo com esta ideia, as SVMs encontram o hiperplano ideal entre duas classes que maximiza a separação entre essas classes e tem o menor risco de erro.

Jarvis *et al.* desenvolveu um novo tipo de SVM *multi-class* chamado *Class-Specific SVM* (CS-SVM) [70]. Neste método, k SVMs binários são criados, onde cada SVM discrimina entre uma das k classes de interesse e uma classe de referência comum. A classe cuja função de decisão é maximizada em relação à classe referência é selecionada. Se todas as funções de decisão são negativas, a classe referência é selecionada. O CS-SVM foi utilizado também para discriminar cliques produzidos por Baleias Bicudas de Blainville a partir de ruídos do ambiente e obteve uma pontuação de classificação correta de 98%. O método também foi aplicado na classificação de Golfinhos Pintados com 95% de acurácia.

É importante notar que estes dados de treinamento foram incluídos nos dados de teste, o que provavelmente resultou em pontos inflacionados de classificação correta. Mazhar *et al.* [102] usou um CS-SVM e vetores de características com coeficientes Cepstrais para o reconhecimento dos chamados produzidos por sete Baleias Jubarte individuais. O SVM resultou em 99% de classificação correta, o que foi uma melhoria significativa sobre a classificação correta de 88% obtidos utilizando um GMM no mesmo conjunto de dados [120].

Os SVMs têm várias vantagens sobre o GMM. Em primeiro lugar, a melhoria observada por Mazhar *et al.* foi obtida com um tamanho muito reduzido de formação do conjunto de dados. Além disso, os GMMs dependem fortemente do cálculo preciso de agrupamentos para a geração de modelos adequados e a seleção dos parâmetros iniciais tem uma influência significativa no resultado do agrupamento. Os SVMs não estão vinculados por estas restrições. Com base nos dois estudos que testaram SVMs, eles parecem ter

potencial para identificação de espécies, mas precisam ser testados em um número maior de espécies e em diferentes situações de SNR. Mazhar *et al.* usou gravações com uma alta taxa de SNR, o que provavelmente contribuiu para os altos pontos de classificação corretas obtidos. Além disso, devido aos SVMs serem essencialmente classificadores binários, é questionável se eles serão úteis para problemas de classificação envolvendo um grande número de espécies.

3.1.3.4 Análise de Função Discriminante Multivariada

A Análise de Função Discriminante Multivariada (*Multivariate Discriminant Function Analysis*, DFA) é uma técnica estatística paramétrica que remonta desde 1935. Esta técnica determina combinações lineares de variáveis que melhor caracterizam as diferenças entre os grupos. Tais combinações são conhecidos como funções discriminantes canônicas. A primeira função discriminante canônica é a combinação linear de variáveis que maximiza as diferenças entre as médias dos grupos em uma dimensão. A segunda variável canônica representa o máximo de separação dos meios numa direção que é ortogonal à primeira. A terceira variável canônica representa a dispersão em uma dimensão independente dos dois primeiros, e assim por diante.

Na análise DFA, as variáveis podem ser utilizadas na análise de todos ao mesmo tempo, ou de forma gradual (quer para a frente ou para trás). Quando funções discriminantes canônicas foram calculadas, as variáveis mensuradas de indivíduos no teste de dados em conjunto são substituídas em cada função e os indivíduos são classificados de acordo com a função que produziu o maior resultado. Como o DFA é uma técnica paramétrica, é assumido que os dados utilizados têm uma distribuição normal multivariada com a mesma matriz de covariância [103]. A principal fraqueza do DFA para tarefas de classificação de mamíferos marinhos é pressupor que as classes são linearmente separáveis. Dado que acontece uma combinação linear de variáveis nesta análise, o espaço característico só pode ser separado em certos aspectos restritos que pode não ser apropriado para todas as vocalizações de mamíferos marinhos.

3.1.3.5 Picos Espectrais

O potencial da utilização de picos espectrais para classificar cliques de ecolocalização produzidos por golfinhos selvagens no sul do golfo da Califórnia foi examinado por Soldevilla *et al.* [106]. Eles examinaram cliques produzidos por cinco espécies da família *Delphinidae* comumente observados na área e quantificaram picos espectrais usando uma regressão de primeira ordem baseada em picos e algoritmos de seleção nos cliques espectrais normalizados, que foram criados a partir de séries de cliques concatenados.

Cliques de ecolocalização são direcionais e as suas características espectrais e de forma da onda têm sido utilizadas para alterar consideravelmente a orientação em relação

a um hidrofone. Algumas dessas características espectrais foram encontradas de maneira consistente em algumas das espécies registradas. No entanto, em alguns casos de outras espécies e tipos de sons essas características espectrais não foram encontradas. Isto sugere que as decisões de classificação podem ser baseadas nas características espectrais de cliques para apenas algumas espécies. Soldevilla *et al.* estão trabalhando em classificadores automatizados que incluem ambos os cliques e assobios.

3.1.4 Pacotes de Software

Apesar de ser mais utilizado na tarefa de classificação de Cetáceos, sistemas de identificação automática podem ser utilizados para outros estudos como, por exemplo, a localização espacial de Cetáceos. Enquanto o animal se move, sua posição exata pode ser calculada através da diferença no tempo de chegada do sinal acústico em cada hidrofone [4]. Para tal, existem pacotes de softwares que executam os algoritmos computacionais apresentados de forma a criar um Fluxo de Trabalho de interesse para pesquisas em campo.

Existem diversos detectores de múltiplos propósitos capazes de realizar tarefas associadas à manipulação de dados oriundos do PAM. ISHMAEL é um software livre para análise acústica de vários propósitos utilizado por alguns pesquisadores para detecção de Cetáceos. XBAT é um software extensível que recebeu algumas funcionalidades para análise de sons produzidos por diversos tipos de animais, incluindo a possibilidade para detecção de Cetáceos. O software PAMGUARD é uma iniciativa atualmente financiada pelo projeto OGP E&P Sound and Marine Life [121], com intuito de criar um sistema computacional padronizado capaz de atender tanto as necessidades de desenvolvedores, quanto de usuários das diversas aplicações do PAM. Uma comparação entre os pacotes mais utilizados pode ser vista em [5].

Além dos pacotes de software apresentados, existem diversas implementações capazes de identificar Cetáceos automaticamente. A maioria delas se baseia em séries temporais e espectrogramas [4]. Desenvolvido por Marie Roch, o *Gaussian Mixture Model Detector* (GMMD), como o nome implica, utiliza GMM para representar cada uma das espécies classificadas.

Existe também o *Frequency Modulated Echolocation Click Detector* (FMCD), desenvolvido por Steve Martin. O FMCD é capaz de realizar detecções de Baleias Bicudas em três estágios. O estágio inicial determina o desvio padrão de uma janela com um segundo. Antes de realizar o cálculo do desvio padrão, o sinal é filtrado por um filtro passa alta em 20 kHz . Em seguida são criadas janelas a cada 5 ms , sendo detectados sinais que excedam aproximadamente sete vezes o desvio padrão calculado. O segundo estágio procura por modulações de frequência com características dos cliques das Baleias Bicudas. O terceiro e último estágio valida a detecção através de comparações do intervalo entre os cliques das Baleias Bicudas, evitando ruídos ou cliques de outros Cetáceos [5].

O próximo detector é conhecido como *Energy Ratio Mapping Algorithm* (ERMA) [14]. O ERMA foi desenvolvido com intuito de aumentar o desempenho de algoritmos baseados em somatório de energia, em especial os utilizados para detecção de cliques de ecolocalização. Esses algoritmos são comumente aplicados em ambientes com capacidade computacional e energética reduzida, tal como planadores acústicos [14, 18]. Com procedimentos computacionais mais econômicos que diminuem o uso de energia, o ERMA apresenta um grande diferencial para execução em equipamentos autônomos por longos períodos de tempo [13, 18], incluindo dispositivos especializados na detecção de cliques como, por exemplo, o C-POD [122], os planadores acústicos e hidrofones fixos em boias com detectores em tempo real [4, 13, 18]. Contudo, devido ao número de detecções falso positivas, o detector ERMA pode trabalhar melhor quando utilizado em sistemas de identificação em duas etapas, onde o segundo passo é capaz de realizar tarefas mais sofisticadas para identificação automática. Essa metodologia é especialmente útil em aplicações com algoritmos mais sofisticadas para identificação que não podem operar continuamente [14]. Assim, cria-se um balanço entre eficiência computacional e eficácia de detecção [5, 14].

Em [5] é apresentada uma comparação entre os sistemas descritos. A comparação realizada é qualitativa, onde nove fatores são considerados, dentre eles: disponibilidade pública, plataforma computacional, processamento em tempo real, métodos de localização, custo, qualidade de detecção, detecção de cliques, habilidade de identificação e capacidade de exibição dos dados coletados. A comparação entre sistemas de identificação automática é bastante complexa, principalmente quando se avalia questões relacionadas ao objetivo de sua aplicação [5]. Sendo assim, definir a verdadeira qualidade de um detector é tarefa completamente dependente das características que estão sendo avaliadas [5]. A escolha do detector depende diretamente do seu uso, por exemplo, necessidade de detecção em tempo real, processamento posterior a recuperação de dados, determinação da presença de Cetáceos, estimação populacional e contagem total de cliques [5]. PAMGUARD e ISHMAEL são detectores livres, sendo que para todos os outros apresentados é necessária uma licença do software Matlab [123]. Outra desvantagem é que todos os detectores apresentados foram desenvolvidos com a necessidade de interação humano-computador através de uma interface gráfica, inviabilizando sua implementação em dispositivos autônomos. Além da interação necessária, softwares como o Matlab gastam uma quantia demasiada de recursos computacionais, tais como memória e processador. Considerando o objetivo deste trabalho como detecção eficiente em dispositivos autônomos, os métodos existentes apresentados seriam ineficientes, ou até mesmo inviáveis.

Independente do dispositivo utilizado, dois pontos são fundamentais para a definição da estratégia de detecção [4]. O primeiro é dado pelo tipo de vocalização a ser detectada, bem como a quantidade de variações nestas vocalizações. Algumas espécies possuem vocalizações altamente estereotipadas, sendo assim mais suscetíveis a detecções utilizando métodos de espectrogramas. Outras espécies, tais como os Golfinhos, produzem sons

com alta variabilidade tonal. Para este tipo de detecção, tipicamente é necessário utilizar somatório de energia, seguido de técnicas estatísticas para classificação. Contudo, as espécies que produzem sons com variações intermediárias são melhores detectadas utilizando Redes Neurais Artificiais. O segundo ponto em questão é a acurácia de detecção desejada. Em um cenário perfeito, um método de detecção seria capaz de detectar todas as instâncias de um determinado tipo de chamado, e desprezando as demais. Este ideal dificilmente é alcançado, em parte porque inevitavelmente existem chamados difíceis de classificar, até mesmo para especialistas humanos. O problema então se torna em configurar a sensibilidade do detector para obter um certo balanço entre chamados perdidos e detecções incorretas.

3.2 FLUXO DE TRABALHO PROPOSTO

Quando aplicado em dispositivos fixos autônomos, os algoritmos apresentados e utilizados pelo processo descrito na Fig. 5 acabam inviabilizando sua implementação, grande parte devido ao gasto de processamento computacional dos algoritmos, que possuem uma relação direta com o gasto energético do dispositivo. Considerando aspectos da Engenharia de Software, a Fig. 6 propõe um Fluxo de Trabalho capaz de facilitar o processo de PAM em dispositivos fixos autônomos para captura e análise de dados de Odontocetos, reduzindo a quantidade total de dados a serem armazenados e processados.

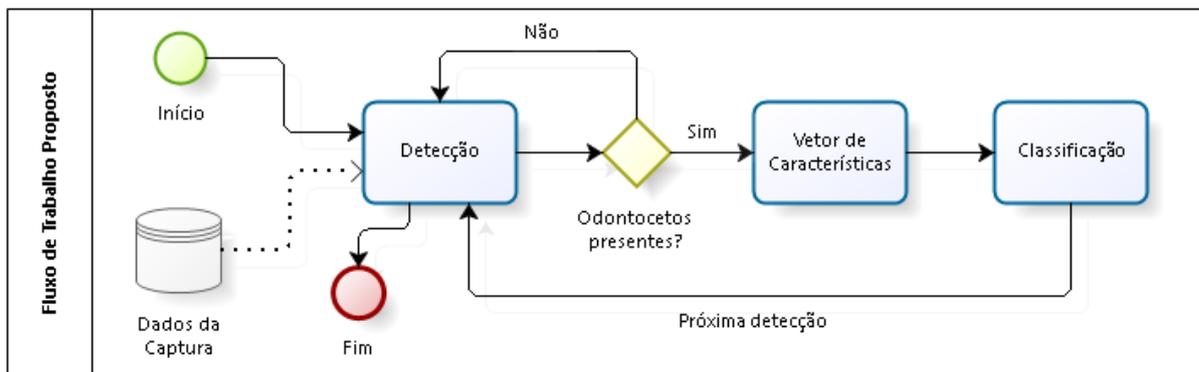


Figura 6 – Fluxo de Trabalho geral proposto para dispositivos fixos autônomos.

Assim como apresentado no Capítulo 2, quando a detecção necessita de trabalhar em faixas de frequências mais altas, como é o caso dos cliques emitidos por Odontocetos, o volume de dados aumenta consideravelmente. O grande volume de dados capturado inevitavelmente gera um gargalo computacional pela sua necessidade de análises robustas para gerar detecções de espécies confiáveis [4].

O Fluxo de Trabalho proposto na Fig. 6 utiliza uma visão diferenciada da abordagem utilizada que foi apresentada na Fig. 5. Conforme mostrado na Fig. 6, por se tratar de um processo de detecção em tempo real, o fluxo espera que sinais sejam recebidos a

partir de um dispositivo de captura. Ao serem recebidos estes sinais passam pela etapa de detecção, onde são analisados a procura de cliques.

Os cliques são considerados pela sua extrema importância no comportamento de um Odontoceti. Em especial, os cliques de ecolocalização, que são utilizados para localização em ambientes com visão reduzida, como é o caso de grandes profundidades no mar. Dessa forma, consideramos que antes da emissão de qualquer outra vocalização, os cliques provavelmente estão sendo emitidos [124, 125, 126, 127, 128, 129]. Baseando-se nessa premissa, algoritmos mais simples podem ser utilizados para detecção. Neste caso, a etapa de detecção funciona como um filtro, com objetivo primário de reduzir a quantidade de dados necessários de entrada aos algoritmos de maior custo computacional das próximas etapas.

Para concretizar a redução dos dados a serem pós-processados, duas abordagens de armazenamento podem ser consideradas. A primeira delas, diz respeito a realizar o descarte total dos dados que não forem detectados como de interesse, armazenando somente a fração considerada de interesse. Desenvolvimentos futuros para o sistema HARP planejam implementar esse tipo de detecção em seus dispositivos de armazenamento, resultando em quantidades de dados capturados muito menores [16]. A segunda abordagem trata da possibilidade de trabalho similar a algoritmos de ativação, onde ao se receber um clique, os dados podem ser gravados em disco enquanto não é alcançado um tempo pré-definido de ociosidade.

A maior preocupação em ambos os casos é garantir que nenhuma informação relevante seja perdida. Considerando que sinais capturados por dispositivos fixos autônomos são na maioria das vezes processos estacionários [130], onde cerca de 70-90% [16, 130] dos dados são ruídos naturais do ambiente, o espaço para redução sem perda é bastante amplo.

3.2.1 Possibilidades Emergentes

Dependendo do modelo de armazenamento escolhido após a etapa de detecção, o Fluxo de Trabalho pode conter atividades extras que são consideradas possibilidades emergentes, dado que elas só são possíveis se o volume de dados for reduzido a uma quantidade viável para as tecnologias disponíveis atualmente.

A Fig. 7 mostra o Fluxo de Trabalho com três etapas extras além da detecção. As duas primeiras, distribuição e identificação podem ser executadas no equipamento caso desenvolvido para tal, já a última etapa de análise deve ser executada somente em laboratórios em terra.

Devido às dificuldades de prover recursos ótimos para dispositivos fixos autônomos, estes não disponibilizam outras formas de acesso aos dados que não seja a retirada do equipamento no mar. Considerando as possibilidades tecnológicas tais como Satélites, e

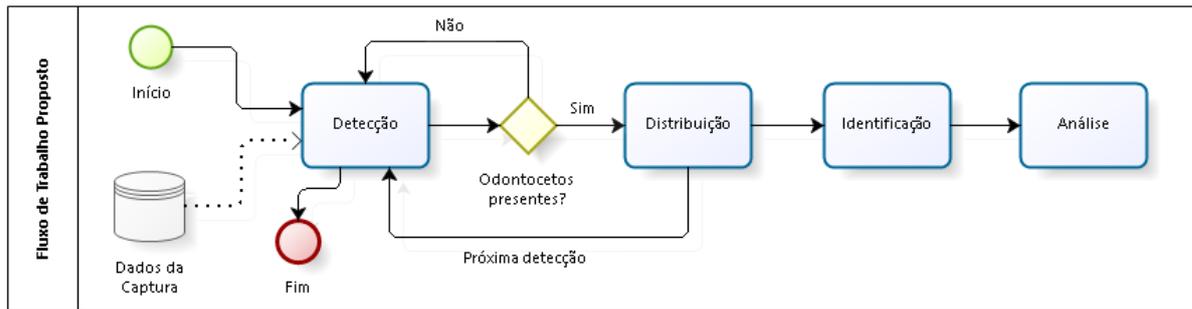


Figura 7 – Fluxo de Trabalho geral proposto com possibilidades emergentes.

projetos em desenvolvimento, como o caso do *Google Loon* [131], a retirada de equipamentos do mar para se ter acesso aos dados pode ser considerada uma abordagem obsoleta do ponto de vista tecnológico, apesar de ser ainda a mais utilizada. Satélites são tecnologias antigas, que tem seu custo reduzido cada vez mais, principalmente devido aos avanços recentes, onde mais satélites são enviados com capacidades de transmissão elevadas, como é o caso da rede Iridium [132]. O *Google Loon* é um projeto privado com intuito de prover o acesso a internet através de balões, inclusive em áreas remotas para pesquisa, como é o caso do oceano. Levando em consideração estes elementos, novos modelos de PAM em dispositivos fixos autônomos devem considerar a possibilidade de distribuição efetiva dos dados, assim como apresentado na Fig. 6. No entanto, aplicações como estas ainda carecem de uma infraestrutura que facilite a utilização de todo um aparato computacional de suporte.

Com intuito de possibilitar a transferência de informações relevantes detectadas, a etapa de distribuição foi inserida levando-se em conta potenciais tecnológicos de hardware e software. Para tal, a etapa de distribuição leva em consideração duas possibilidades. A primeira delas, restrita por limitações e custos tecnológicos, trata da possibilidade de transferir o segmento completo de dados, incluindo todas as amostras capturadas na parte do sinal de interesse. O problema desta abordagem é o custo necessário para transferir um volume de dados alto, como é o caso quando muitos cliques são detectados em um espaço de tempo pequeno. A segunda possibilidade leva em consideração a compactação dos dados antes que estes sejam transferidos, sendo neste caso, proposto a utilização de abordagens de mineração de dados [133] para uma distribuição eficiente, influenciando inclusive na etapa seguinte de identificação, como é o caso do *Locality Sensitive Hashing* (LSH) [134, 135].

Neste contexto, a *e-Science* [136] pode ser utilizada para caracterizar o importante papel das tecnologias computacionais na pesquisa, colaboração, compartilhamento de dados e documentos, e uso de recursos para automatizar a execução e análise de dados de experimentos científicos [137]. Assim, a área de *e-Science* está em busca de aprimorar e facilitar o uso deste aparato computacional através de uma infraestrutura que permita

projetar, reusar, anotar, validar, compartilhar e documentar artefatos gerados pela pesquisa científica. Apesar disso, atualmente existe uma grande lacuna entre os esforços empregados e a visão da *e-Science*, a qual tem um grau de facilidade de uso e automação muito mais elevados, e que propõe maneiras de colaboração e computação em escala global [138].

O termo *e-Science* foi introduzido no Reino Unido por Taylor, quando ele reconheceu a importância do uso de ferramentas computacionais na pesquisa científica colaborativa, multidisciplinar e com grande volume de dados [137], encapsulando as tecnologias necessárias ao seu suporte que emergiu em vários campos da ciência [136]. Portanto, o termo *e-Science* pode ser utilizado para descrever o desenvolvimento de infraestruturas de serviços de software capazes de prover acesso a facilidades remotas, recursos computacionais distribuídos, armazenamento de informações em bancos de dados dedicados, disseminação e compartilhamento de dados, resultados e conhecimento. A utilização de recursos computacionais no desenvolvimento da pesquisa beneficia o trabalho das comunidades científicas, facilitando o compartilhamento de dados e serviços computacionais, além de contribuir para a construção de uma infraestrutura de dados e de uma comunidade científica distribuída [137].

A terceira etapa, conhecida como identificação de padrões de espécies, possui diversos métodos já consolidados [139], sendo estes aplicados em diversos problemas que necessitam de reconhecimento de padrões. A identificação de padrões em dados segue uma lógica genérica, onde o processo básico pode ser descrito pela captura dos dados, seguido pela criação de um vetor de características com intuito de obter atributos relevantes que caracterizam individualmente cada espécie, e finalizado pela rotulação das espécies com base em dados históricos manualmente classificados por especialistas.

Normalmente, o que muda no processo completo de identificação é como o vetor de características é criado e quais são as características mais relevantes para o aumento da acurácia de classificação. Graças a grande variabilidade na estrutura de sons de mamíferos marinhos, talvez nenhum método seja capaz de ser efetivo para detecção automática e classificação acústica de todas as espécies e populações [11].

Com a aplicação de conceitos de Engenharia de Software, criando um Fluxo de Trabalho capaz de trabalhar com Componentes de Software em cada etapa [140], é possível que a etapa de identificação incorpore uma maior variedade de métodos de classificação disponíveis visando uma maior robustez no processo. Para tal, uma padronização no formato de entrada e saída de dados entre as etapas deve ser definida.

A saída da primeira etapa nada mais é que um segmento do sinal capturado em determinado tempo de interesse. O tamanho do segmento pode variar de acordo com configurações prévias do usuário. Normalmente, o tamanho do segmento é escolhido de forma que seja capaz de acomodar um clique, balanceando com a possibilidade de evitar mais de um clique dentro do mesmo segmento. A escolha do tamanho do segmento é um

passo crítico para a acurácia da etapa de detecção, mas também deve ser avaliado em relação à eficiência computacional. Isso porque existem algoritmos de transformação para o domínio da frequência computacionalmente otimizados para cálculos em segmentos com tamanhos que sejam potências de dois. Ou seja, caso a extração de características para a etapa de identificação seja feita no domínio da frequência, um segmento com 1.024 (2^{10}) amostras será calculado de maneira mais eficiente do que um com 1.025 amostras.

É importante ressaltar que, dado o modelo de Fluxo de Trabalho proposto pela Fig. 6, qualquer um dos métodos apresentados para criação do vetor de características [70, 96, 99, 100, 101, 102, 103, 104, 105, 106, 141, 142, 143], bem como os apresentados para classificação [81, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93] podem ser utilizados para identificação. Além disso, outros métodos mais eficientes de mineração de dados podem ser utilizados com intuito de possibilitar de maneira eficiente a execução em conjunto da etapa de distribuição e identificação. Em especial podemos citar o LSH.

Para finalizar, a etapa de análise propõe o efetivo uso dos dados de forma a usar Componentes de Software para realização da tarefa em questão. Este não é o foco deste trabalho, mas devido a padronização do Fluxo de Trabalho e formato de dados distribuído por cada etapa, algoritmos de análise podem ser implementados com facilidade como, por exemplo, algoritmos para estimação de densidade populacional.

4 DETECÇÃO EFICIENTE

Em experimentos científicos representados por Fluxos de Trabalho, além da especificação concisa de atividades e correlações entre as mesmas, outras variáveis do domínio podem ser exploradas para melhorar a eficiência da execução [58]. No caso do Monitoramento Acústico Passivo, os métodos mais comuns para detecção utilizam espectrogramas [22] que impõem a necessidade de conversão para o domínio da frequência. Com o intuito de evitar o gasto computacional [144] com essa conversão, o método proposto para detecção neste capítulo é executado inteiramente no domínio do tempo.

A criação de um espectrograma envolve a aplicação de uma Transformada de Fourier de Tempo Curto (*Short-Time Fourier Transform*, STFT) no sinal capturado. Na prática, uma STFT divide o sinal em segmentos menores normalmente de tamanho fixo, e aplica uma Transformada de Fourier Discreta (*Discrete Fourier Transform*, DFT) em cada um destes segmentos. O tamanho do segmento depende do balanceamento que se deseja alcançar entre resolução temporal e frequência, sendo este conhecido como princípio da incerteza. O princípio da incerteza trata da impossibilidade de garantir o tempo exato em que determinada frequência ocorreu no tempo, bem como precisar exatamente qual frequência foi capturada. Basicamente, um sinal não pode ser representado como um ponto exato no tempo e na frequência, dado que quando se aumenta a resolução de um, a do outro diminui [22].

Ao realizar operações de detecção no domínio do tempo, consegue-se aumentar a resolução temporal, conseqüentemente cria-se um alto nível de precisão nas próximas etapas para os métodos que dependam fortemente do tempo como, por exemplo, Diferença do Tempo de Chegada (*Time Difference of Arrival*, TDOA) para estimativas de densidade populacional. Caso aplicado recursivamente em segmentos menores, a precisão pode chegar a exatidão temporal baseada no tamanho do clique.

A Transformada Rápida de Fourier (*Fast Fourier Transform*, FFT) é o nome dado a algoritmos capazes de realizar a transformação do sinal para o domínio da frequência de maneira eficiente. Existem diversos algoritmos que se enquadram nesta categoria, cada um utilizando técnicas matemáticas diversas, variando de Aritmética de Números Complexos a Teoria dos Números e Grupos. O algoritmo mais utilizado para calcular a FFT é conhecido como algoritmo de Cooley–Tukey [145]. A variação mais utilizada do algoritmo divide os segmentos capturados recursivamente em unidades menores, sendo estas utilizadas para o cálculo da DFT. Ao se trabalhar com abordagens recursivas de divisão e conquista, a execução se torna mais eficiente quando analisado o custo assintótico.

A análise assintótica é um método utilizado para descrever o comportamento de limites, comumente utilizado para avaliar a complexidade de algoritmos através da notação *Big O*. Em computação esse tipo de notação é usado para avaliar como um determinado

algoritmo se comporta de acordo com o tamanho dos dados de entrada [146]. Na prática, o desempenho efetivo em computadores modernos é um tema complexo e requer uma análise extensa da metodologia utilizada [144].

Assim como apresentado no Capítulo 2, o volume de dados capturado demonstra um impacto notável na escolha do dispositivo, principalmente quando a captura em altas frequências é um processo contínuo no tempo, gerando dezenas ou até mesmo centenas de segmentos em um segundo. Apesar de existirem algoritmos eficientes capazes de realizar a conversão dos segmentos para o domínio da frequência utilizando um mínimo de operações [144], a diminuição das instruções executadas por segundo devido à execução no domínio do tempo pode constituir uma otimização relevante perante ao gasto energético total do processador.

Independente do domínio de trabalho, a criação de um vetor de características é um passo obrigatório a ser executado antes de qualquer algoritmo de Aprendizado de Máquina. Ao trabalhar utilizando dados brutos do sinal no domínio do tempo, importantes características únicas para classificação de Odontocetos são perdidas. Por exemplo, os cliques de ecolocalização das Baleia Bicuda de Baird (*Baird's Beaked Whales*) possuem frequência central entre 22 e 25 kHz , diferenciando-se das Baleia Bicuda de Cuvier (*Cuvier's Beaked Whale*), que possuem frequência central entre 30 e 48 kHz .

Diferente dos métodos mais comuns para detecção de mamíferos marinhos, que utilizam a energia em uma determinada faixa de frequência para verificar se um limiar definido pelo usuário é excedido [22], o método aqui proposto se baseia em características físicas da onda, tendo como referência o seu formato. Dessa forma, não é necessário um valor informado pelo usuário como limiar para considerar se é um sinal de interesse ou não, dado que é realizado um treinamento prévio do algoritmo utilizado.

Na Fig. 8 podemos ver claramente o padrão a ser detectado no domínio do tempo. Um clique quando capturado gera uma variação brusca de energia em um curto espaço de tempo, causando uma perturbação característica no sinal ambiente estacionário. A Fig. 8 (a) mostra um sinal contendo três cliques de Odontocetos capturados com uma taxa de amostragem de 200 mil amostras por segundo, sendo capaz de representar sinais de até 100 kHz . As frações (b), (c) e (d) da Fig. 8 mostram os três cliques em segmentos de 2.048 amostras cada. Esta característica é claramente visível em sinais capturados por dispositivos fixos autônomos, principalmente pela quantidade reduzida de ruídos.

Dada a intenção de detecção no domínio do tempo, abordagens utilizadas em outras áreas para detecção de picos e cliques neste domínio foram testadas. No caso de Nugent [147], a detecção de cliques funciona inteiramente no domínio do tempo, mas considerada que pra serem detectados os cliques devem possuir mais conteúdo em altas frequências do que o sinal desejado. Por causa disso, um filtro passa alta é executado antes de qualquer processamento. Um filtro passa alta é comumente utilizado para remover

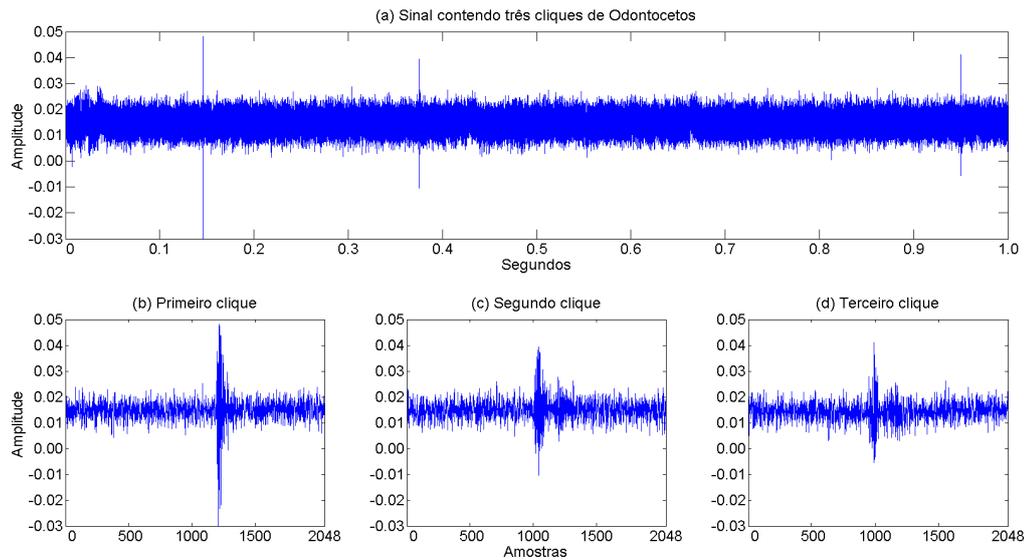


Figura 8 – Sinal de um segundo contendo três cliques de Odontocetos.

ruídos harmônicos simples como, por exemplo, os causados por estações de distribuição de energia elétrica. O problema de aplicar um filtro passa alta é que inevitavelmente o conteúdo de interesse pode ser perdido, não sendo idealmente aplicado em qualquer tipo de sinal. Em nenhum caso um filtro passa alta é adequado em sinal com harmônicos complexos [147].

A princípio o cenário de aplicação do algoritmo descrito é raro de ser encontrado em ambientes não controlados, como o ambiente marinho com ruídos antropogênicos e naturais. A proposta de Nugent funciona em segmentos de tamanho l previamente definidos, aplicando o filtro da Equação 4.1 em cada bloco gerado dentro do segmento.

$$h_i = \sqrt{\frac{1}{N-2} \left(\sum_{j=i-N}^{i+N} (s_{j-1} - 2s_j + s_{j+1}) \right)} \quad (4.1)$$

O tamanho N do bloco pode variar, mas o valor típico utilizado por Nugent foi 8. Este valor foi testado em nossos experimentos, mas devido ao tamanho do clique, testes foram realizados com valores distintos de 256, 512 e 1.024 amostras. A partir disso é calculado a média da saída produzida pelo filtro \bar{h} , a primeira derivada do conteúdo de alta frequência na Equação 4.2 para $j = (0 \dots l)$, bem como a média \bar{d} e desvio padrão d_σ .

$$d_j = \frac{s_{j+1} - s_{j-1}}{2} \quad (4.2)$$

A detecção do clique é realizada ao mover do fim até o início, testando se a condição

da Equação 4.3 é verdadeira para cada amostra em h .

$$d_j > \frac{3d_\sigma}{S} + \bar{d} \text{ e } h_j > \frac{3\bar{h}}{S} \quad (4.3)$$

Se a condição for verdadeira, é provável que o fim de um clique tenha sido encontrado. O processo continua movendo pelo segmento até que a condição da Equação 4.4 seja verdadeira, indicando que o provável início do clique foi encontrado.

$$h_j < \frac{3d_\sigma}{S} \quad (4.4)$$

Porém, devido a características únicas presentes nos sinais de Odontocetos, a abordagem apresentada não apresentou resultados satisfatórios, com uma taxa extremamente elevada de perda.

4.1 FLUXO DE DETECÇÃO

A etapa de detecção consiste de quatro fases específicas. Assim como mostrado na Fig. 9 e no Algoritmo 1, o sinal capturado é dividido em segmentos de dimensões previamente estabelecidas, baseando-se na taxa de amostragem e duração dos cliques de interesse. Através da Análise de Dados Simbólicos [148], cada segmento é transformado em um tipo de dado modal multi-valorado utilizando uma função de histograma [149]. Após o cálculo do dado multi-valorado, um algoritmo de ordenação de números inteiros é utilizado para criar um modelo de dado padrão a ser reconhecido. Para finalizar, um algoritmo de Aprendizado de Máquina é utilizado para classificar se aquele segmento multi-valorado ordenado contém alguma informação de interesse.

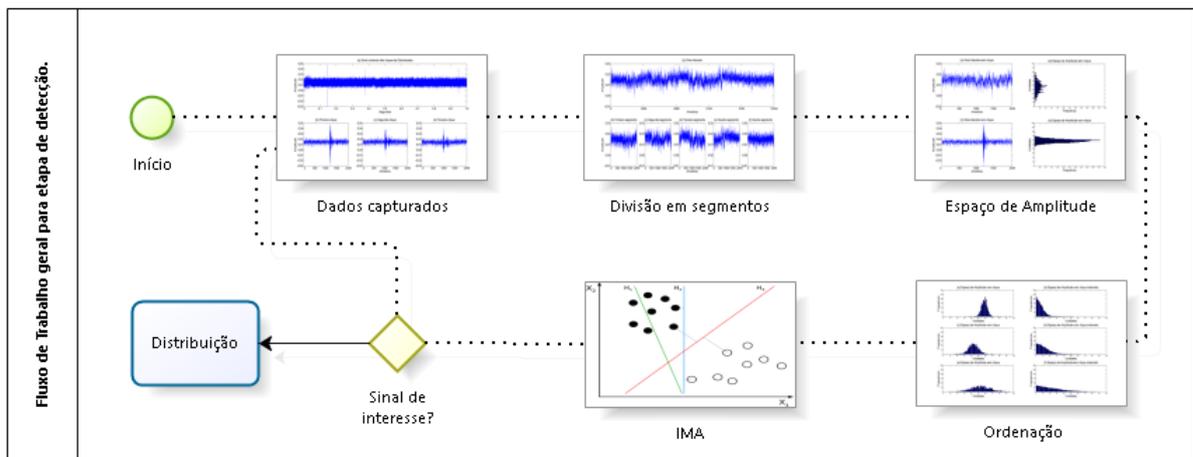


Figura 9 – Fluxo de Trabalho geral para etapa de detecção.

Algoritmo 1: Algoritmo para o fluxo de detecção.

Entrada: Sinal discreto capturado em tempo real

início

enquanto $SinalCapturado \neq \emptyset$ **faça**

 Divide $SinalCapturado$ em $Segmentos$;

para $Contador \leftarrow 0$ **até** $TamanhoSegmentos$ **faça**

 Aplica função de histograma no $Segmento$;

 Ordena $DadosSimbolicos$ com ordenação de inteiros;

 Classifica $DadosSimbolicos$ através do Aprendizado de Máquina;

fim

fim

fim

A divisão do sinal em segmentos é uma técnica comumente utilizada em processamento de sinais [5]. Um exemplo de divisão é apresentado na Fig. 10, onde um sinal discreto no tempo é dividido em 5 segmentos contendo 2.048 amostras cada. Para escolher o tamanho do segmento, alguns itens devem ser levados em consideração:

- A duração do clique é o item mais importante a ser considerado. O tamanho do segmento deve ser escolhido de modo que durante toda a sua duração, somente um clique seja capturado. Quando o mesmo segmento detecta mais de um clique, as etapas posteriores podem sofrer alterações indesejadas nos resultados. A etapa de identificação, por exemplo, poderia classificar somente um dos animais presentes naquele segmento. O mesmo caso poderia acontecer na etapa de análise, onde somente um clique seria levado em consideração pelo algoritmo de estimação de densidade populacional, ao invés dos dois capturados. Ressalta-se que nem todos os algoritmos das etapas futuras estão suscetíveis a isso, dependendo fortemente da sua implementação;
- A taxa de amostragem do sinal deve ser levado em consideração em conjunto com o item anterior. Quando o segmento é pequeno demais, podem não existir amostras suficientes para uma discriminação sucinta dos dados no domínio da frequência. O segmento também não pode ser muito grande, já que além de possibilitar mais de um clique dentro do mesmo segmento, a incerteza temporal é intensificada [135] prejudicando algoritmos dependentes do tempo;
- Por último relacionado com a eficiência computacional, quando o tamanho do segmento é uma potência de dois, variações do algoritmo de Cooley–Tukey podem ser utilizadas para um cálculo com complexidade computacional reduzida. A complexidade das implementações pode variar de $O(n^2)$ até $O(n \log n)$ [144].

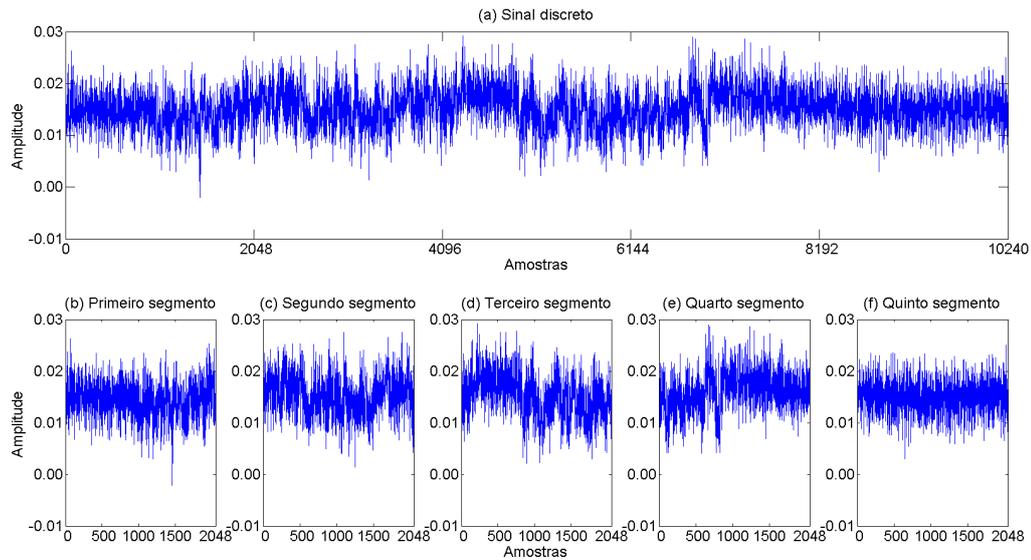


Figura 10 – Divisão do sinal em segmentos menores.

Assim como apresentado no Capítulo 5, este trabalho dividiu os sinais capturados com 200 mil amostras por segundo em segmentos com 512 amostras cada, gerando um espaçamento temporal de $2,56\text{ ms}$. Para os casos de captura com 320 mil amostras por segundo, cada segmento foi dividido nas mesmas 512 amostras, resultando em um espaçamento de $1,5\text{ ms}$. Os valores escolhidos proporcionaram uma boa resolução temporal, como consequência, os segmentos variam lentamente no tempo, proporcionando precisão à etapa de detecção bem como robustez à incerteza temporal na etapa de identificação.

4.2 ESPAÇO DE AMPLITUDE

O início do processo de criação do vetor de características pode ser considerado a partir desta fase da etapa de detecção. Infelizmente, a criação do vetor de características raramente é o foco de publicações na literatura, por isso os métodos utilizados são mencionados com poucos detalhes [11].

Para realizar a criação do vetor de características sem informações do espectrograma, uma representação mais rica dos dados deve ser utilizada. Para tal, a Análise de Dados Simbólicos [148] foi considerada, criando um novo tipo de dado modal multi-valorado através de uma função de histograma [149]. Uma das vantagens da utilização de Dados Simbólicos é o fato de reduzir a quantidade de dados a serem analisados, sem perda de generalidade [149]. Existem diversas formas de conversão de sinais no tempo para Dados Simbólicos [148], sendo algumas apresentadas abaixo:

- Valor quantitativo único, exemplo: altura(h) 1,75;
- Valores intervalados, exemplo: altura(h) [1,50 2,00];

- Multi-valorado categorizado, exemplo: alto, médio ou baixo;
- Modal multi-valorado, utilizado neste trabalho e detalhado a seguir.

A Fig. 11 mostra que quando uma função de histograma é aplicada no sinal capturado, o formato físico da onda é condensado em uma distribuição de frequências de amplitude espaçada. Como a representação gráfica do resultado gerado pela função de histograma não é utilizada, os dados resultantes são chamados de Espaço de Amplitude.

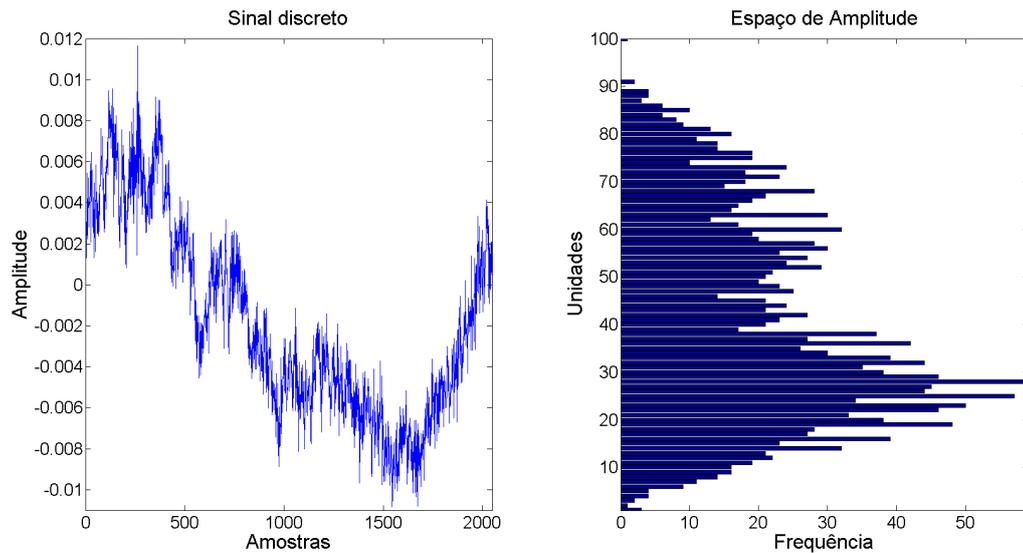


Figura 11 – Conversão do sinal para Espaço de Amplitude.

O exemplo da Fig. 11 constitui uma divisão entre o valor mínimo de $-0,01083$ e máximo de $0,01166$, sendo a diferença de $0,02249$ dividida em um Espaço de Amplitude com 100 unidades. A partir do Espaço de Amplitude gerado podemos inferir que a unidade do Espaço de Amplitude com maior frequência é a 28, além disso, uma maior atividade pode ser descrita entre as unidades 14 e 38.

O Algoritmo 2 mostra que inicialmente os valores mínimo e máximo da amplitude naquele segmento devem ser encontrados. A partir da diferença entre estes valores pode-se gerar um espaçamento linear dividido de acordo com a quantidade total de unidades desejadas. A quantidade total de unidades do Espaço de Amplitude não influencia demasiadamente na sua discriminação, entretanto, quando adiciona-se muitas unidades o desempenho da função de histograma e o treinamento do algoritmo de Aprendizado de Máquina podem ser prejudicados. Para finalizar a criação do Espaço de Amplitude o índice da amplitude é calculado de acordo com sua intensidade e armazenado em uma unidade dentro do Espaço de Amplitude. Neste momento é necessário uma iteração por todas as

unidades do Espaço de Amplitude, categorizando em uma complexidade assintótica de $O(n)$ em função do tamanho do segmento.

Algoritmo 2: Algoritmo para criação do Espaço de Amplitude

Entrada: Segmentos com dados capturados

início

 Armazena valor *Maximo* e *Minimo* de amplitude do *Segmento*;

 Gera espaçamento linear entre *Maximo* e *Minimo*;

para *Contador* $\leftarrow 0$ **até** *TamanhoSegmento* **faça**

 Calcula o *Indice* do elemento;

 Incrementa o valor atual do *Segmento* no *Indice* correto;

fim

fim

Normalmente o Espaço de Amplitude gerado a partir de um sinal ambiente (sem cliques capturados) possui uma forma bem distribuída e ordenada. Em contrapartida, quando o Espaço de Amplitude é gerado a partir de um sinal com pelo menos um clique, ele toma uma forma totalmente distinta, mantendo a maioria dos pontos discretos nas unidades centrais. O comportamento descrito pode ser visto na Fig. 12, onde os Espaços de Amplitude gerados podem ser facilmente distinguíveis um dos outros. No primeiro caso mostrado pela Fig. 12 (a) e (b), o sinal capturado não contém nenhum clique, gerando um Espaço de Amplitude similar a uma distribuição normal. Neste caso, os valores do Espaço de Amplitude variam de 0 até aproximadamente 45. A Fig. 12 (c) e (d) mostra o segundo caso onde um clique de Odontoceti foi capturado, gerando um Espaço de Amplitude totalmente distinto, com os valores do Espaço de Amplitude variando de 0 até aproximadamente 225.

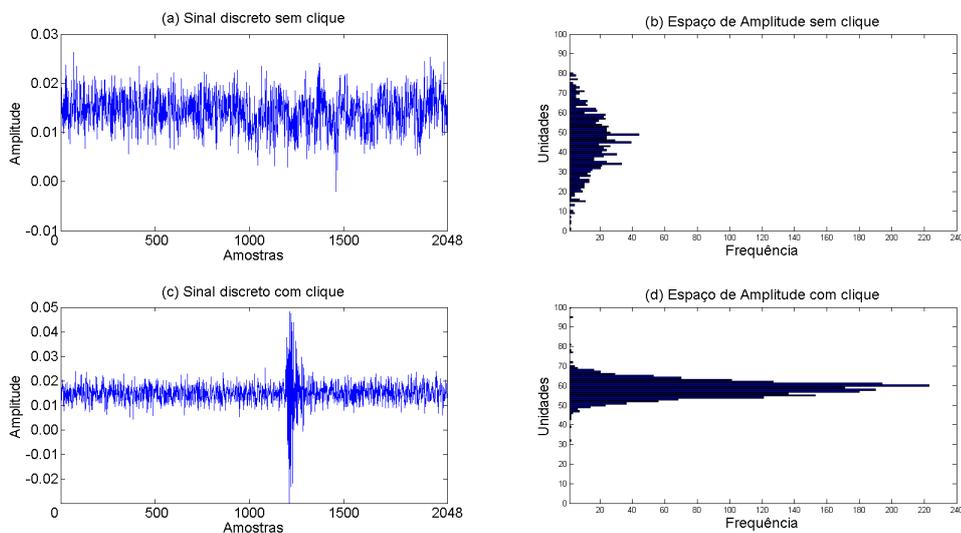


Figura 12 – Espaço de Amplitude com e sem clique de cetáceos.

É importante ressaltar que os segmentos não devem ser normalizados. O Espaço de Amplitude gerado só é facilmente distinguível quando os valores mínimo e máximo do segmento variam em função da intensidade do clique. A Fig. 13 mostra que quando normalizado, a maioria dos pontos discretos ficam concentrados no meio do Espaço de Amplitude. Neste caso, tanto o sinal sem clique mostrado pela Fig. 13 (a), quanto o sinal com clique mostrado pela Fig. 13 (c), ao serem convertidos para seus respectivos Espaços de Amplitude são difíceis de serem categorizados devido a sua similaridade.

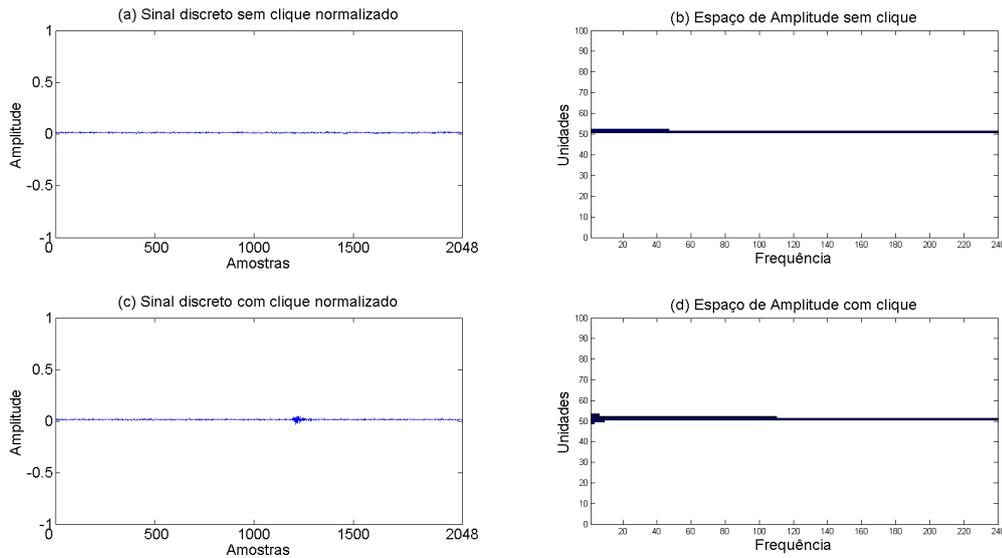


Figura 13 – Espaço de Amplitude gerado a partir do sinal normalizado.

4.2.1 Distâncias e Divergências

A partir do Espaço de Amplitude gerado, torna-se necessário definir uma maneira robusta de comparação entre eles, de forma que sejam rotulados corretamente como sendo um segmento de interesse ou não. Por se tratar de um processo estacionário [130], o modelo de detecção avaliado inicialmente procura por diferenças temporais entre os Espaços de Amplitude gerados [150], de forma a avaliar quando sinais de interesse estão presentes no sinal. Quando visualizado através de um espectrograma, o clique apresenta um formato característico, assim como visualizado no sexto segmento da Fig. 14. Comparando o segmento no tempo t com o anterior no $t - 1$, tem-se uma ocorrência que pode indicar que algum evento de interesse ocorreu ou não no sinal capturado. Apesar de todo o processo ser realizado no domínio do tempo, a Fig. 14 mostra o processo através de um espectrograma com intuito de facilitar o entendimento visual.

Em probabilidade e estatística, métricas, divergências e distâncias são quantificadores comuns para comparação entre distribuições de probabilidade e vêm exibindo boas propriedades discriminatórias em vários problemas [151]. Por ser criado a partir de uma

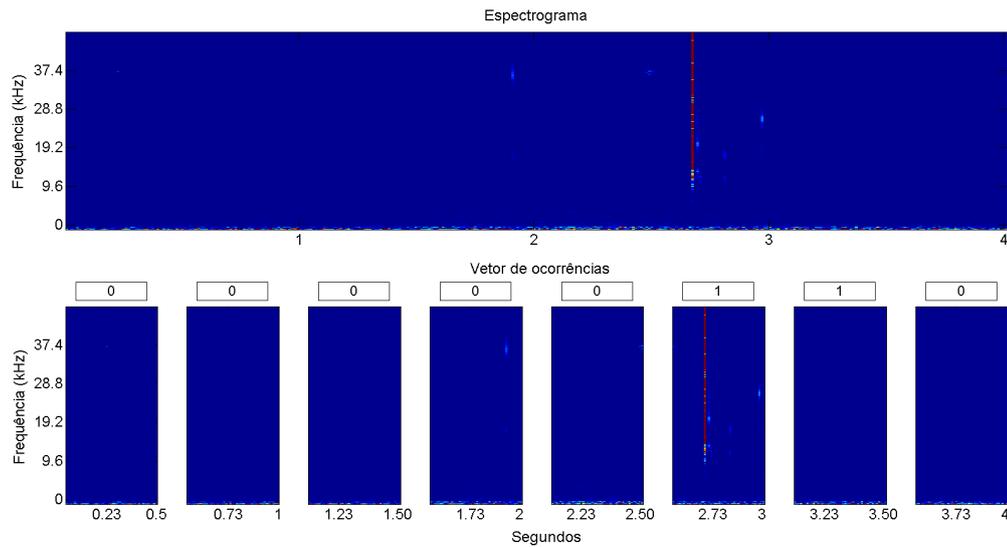


Figura 14 – Vetor de ocorrências gerados como exemplo em um espectrograma.

função de histograma, inicialmente alguns métodos para comparação entre histogramas foram avaliados.

4.2.1.1 *Histogram Intersection*

Histogram Intersection (HI) [152] é uma medida de similaridade útil para imagens e provou ser um *kernel* eficaz para Aprendizado de Máquina. Foi utilizada com sucesso como uma medida de similaridade para a tarefa de recuperação de imagem. O HI é definido como

$$K_{HI}(x, x') = \sum_{i=1}^n \min(x_i, x'_i) \quad (4.5)$$

onde $x = (x_1, \dots, x_n)$ and $x' = (x_1, \dots, x'_n)$ são dois histogramas com n bins cada.

Aplicações do *Histogram Intersection* podem ser encontradas em [152, 153].

4.2.1.2 *Histogram Match*

Histogram match (HM) [154], também conhecido como *Histogram Specification*, é uma operação de processamento de sinal básica. Consiste na transformação de um histograma em outro remapeando os valores do sinal para controlar a frequência relativa de sua ocorrência. O HM é a generalização da equalização de histograma, que estipula um histograma de saída uniforme, fazendo com que cada valor do sinal tenha probabilidade igual, a fim de maximizar a entropia do sinal e assim otimizar o conteúdo de informação do sinal quantificado [155].

O HM é usado para ajustar o perfil de estatística de uma gama dinâmica do sinal enquanto mantém a estrutura local e ordenação relativa dos valores do sinal. Por mais de 30 anos, o HM tem sido objeto de estudo com uma ampla variedade de aplicações, incluindo a síntese de imagem, fotografia digital, processamento de vídeo, sensoriamento remoto, imagens médicas, visão computacional e reconhecimento de fala [155].

O HM gera uma imagem de saída com base em um histograma especificado. É um processo em que uma série temporal, imagem ou dimensão mais elevada de dados escalar é modificada de tal modo que seu histograma corresponde a de um outro conjunto de dados. Uma aplicação comum do HM é combinar a imagem a partir de dois sensores com respostas ligeiramente diferentes, ou a partir de um sensor cuja resposta muda ao longo do tempo.

4.2.1.3 Distância Euclidiana

A distância euclidiana é uma métrica utilizada para medir diferentes pontos em um Espaço Euclidiano. Considere dois Espaços de Amplitude com três unidades cada, sendo eles $P = (x, y, z)$ e $Q = (a, b, c)$. Para calcular a distância euclidiana entre eles, tem-se:

$$d(P, Q) = \sqrt{(x - a)^2 + (y - b)^2 + (z - c)^2}. \quad (4.6)$$

É recomendável a padronização das variáveis antes de se obter o valor da distância euclidiana, visto que normalmente todos os dados não estão no mesmo padrão de medidas.

A distância euclidiana cresce à medida que cresce o número de variáveis. Uma maneira de eliminar o efeito do número de variáveis é dividir o valor da distância euclidiana pela raiz quadrada do número de variáveis, conforme Equação 4.7.

$$\bar{d}(P, Q) = \frac{1}{\sqrt{N_v}} \cdot d(P, Q) \quad (4.7)$$

onde $\bar{d}(P, Q)$ é a distância euclidiana média entre P e Q , N_v é o número de variáveis e $d(P, Q)$ é a distância euclidiana entre P e Q .

4.2.1.4 Distância Chi-quadrada

A distância chi-quadrada $d(x, y)$ [156] é a distância entre dois histogramas $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ com n bins cada. Além disso, ambos os histogramas são normalizados. A medida da distância d é geralmente definida como

$$d(x, y) = \sum \frac{\frac{(x_i - y_i)^2}{x_i + y_i}}{2}. \quad (4.8)$$

A distância chi-quadrada é amplamente utilizada em pesquisas ecológicas. A primeira premissa desta função é que ele é calculada em contagens relativas, e não sobre os originais, e a segunda é que padroniza pela média e não pela variância. O nome da distância é derivado do teste estatístico chi-quadrado de Pearson.

4.2.1.5 Divergência de Kullback-Leibler

A divergência de Kullback-Leibler (D_{KL}) foi originalmente introduzida por Kullback e Richard Leibler em 1951 [157] como a divergência dirigida entre duas distribuições. Ela é comumente usada em estatística com uma medida de semelhança entre duas distribuições de densidade.

A D_{KL} é não-simétrica e pode ser interpretada como o número de bits adicionais (se o logaritmo na base 2 for utilizado) necessários para codificar casos de verdadeira distribuição $p(x)$ utilizando o código com base na distribuição $q(x)$ [158]. Tipicamente, p é a distribuição inicial dos dados, observações ou uma distribuição teórica e q é uma aproximação de p . É também chamada de entropia relativa de P com relação a Q , e escrita como $H(P|Q)$. Para duas variáveis aleatórias contínuas, D_{KL} é dada por

$$D_{KL}(p, q) = \int p(x) \log \frac{p(x)}{q(x)} dx \quad (4.9)$$

A divergência Kullback-Leibler é um caso especial de uma classe mais ampla de divergências chamados f-divergências, bem como a classe de divergências Bregman. Ela é usada em vários aspectos de fala e reconhecimento de imagem como, por exemplo, determinar se dois modelos acústicos são semelhantes ao computar a melhor combinação utilizando modelos de histograma de imagens.

4.2.1.6 Divergência de Jensen-Shannon

A divergência de Jensen-Shannon (D_{JS}) [157] é uma versão resumida de todas as medidas de divergência importantes da teoria da informação. É baseada na divergência de Kullback-Leibler e é também conhecida como raio de informações ou divergência total da média.

Recentemente, descobriu-se que a divergência define uma verdadeira métrica - precisamente, é o quadrado de uma métrica [159]. Aparentemente, ela está ganhando em popularidade, especialmente entre os estatísticos. Ao contrário da divergência de Kullback-Leibler, ela é simétrica, sempre finita e semi-delimitada. A D_{JS} é definida como

$$D_{JS}(p, q) = \frac{1}{2}D_{KL}(p, m) + \frac{1}{2}D_{KL}(q, m). \quad (4.10)$$

onde $m(x) = \frac{1}{2}(p(x) + q(x))$. A estimativa de D_{JS} foi dada em [160], mas pode ser aproximada usando os estimadores de D_{KL} como em [158]

$$\begin{aligned}\hat{D}_{JS}(p, q) &= \frac{1}{2}\hat{D}_{KL}(p, m) + \frac{1}{2}\hat{D}_{KL}(q, m) \\ &= \frac{1}{2N_p} \sum_{i=1}^{N_p} [\log \hat{p}(x_{pi}) - \log \hat{m}(x_{pi})] + \frac{1}{2N_q} \sum_{i=1}^{N_q} [\log \hat{q}(x_{qi}) - \log \hat{m}(x_{qi})]\end{aligned}\quad (4.11)$$

onde $\hat{m}(x) = \frac{1}{2}(\hat{p}(q) + \hat{q}(x))$.

Algumas aplicações da divergência de Jensen-Shannon incluem [158]: agrupamento de característica para classificação de texto, detecção de *outlier* em dados de sensores, entre outros.

4.2.1.7 Divergência de Jeffrey

A divergência de Jeffrey D_J é uma derivação da D_{KL} , que se mostra numericamente estável, simétrica e robusta em relação ao ruído e ao tamanho dos dados de um histograma. Ela é definida como [158]

$$D_J(p, q) = D_{KL}(p, q) + D_{KL}(q, p). \quad (4.12)$$

Isso resolve o problema da não simetria de uma forma simples e intuitiva. Contudo, existe um outro problema: D_J não é definida se $p(x) = 0$ ou $q(x) = 0$, que é mais limitado do que no caso da D_{KL} . Uma aplicação da divergência de Jeffrey é na classificação de dados multimídia com SVM [161].

4.2.2 Variância

Abordagens mais simples utilizando a comparação entre Espaços de Amplitude se mostraram promissoras no início, mas após validações dos resultados obtidos, sua generalidade para todos os casos foi enfraquecida, de forma que a não normalização do Espaço de Amplitude inviabiliza a comparação entre Espaços de Amplitude gerados a partir de espaços lineares distintos.

Portando, é necessário definir as características do Espaço de Amplitude que o categorize bem. A primeira delas é conhecida como variância. Devido a centralização da maioria dos pontos discretos por causa do clique, a variância do Espaço de Amplitude é elevada a um patamar totalmente distinto quando nenhum clique é emitido. A Fig. 15 mostra o comportamento da variância no tempo dos Espaços de Amplitude quando comparada com um sinal temporal discreto.

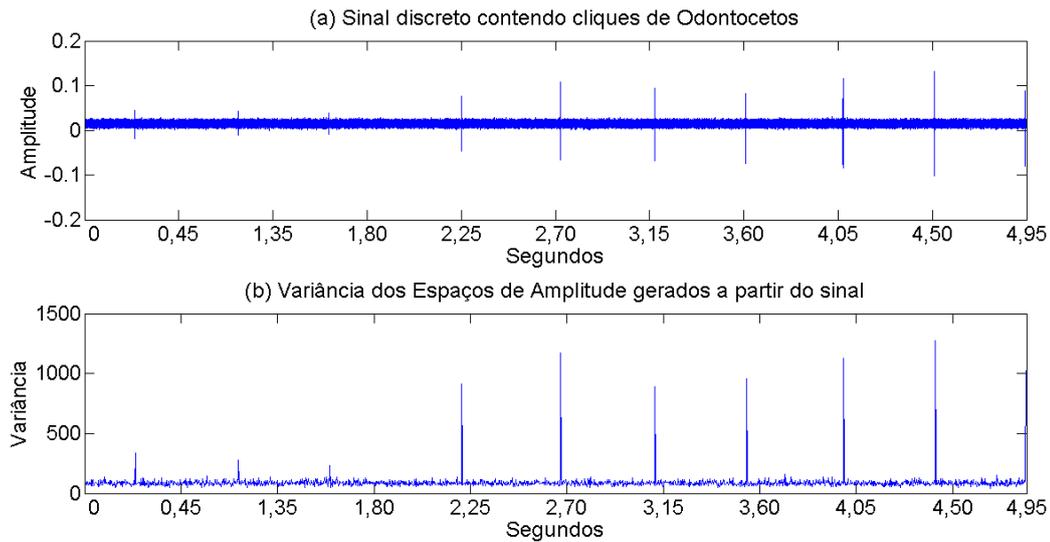


Figura 15 – Variância dos Espaços de Amplitude gerados no tempo.

Similar a abordagem utilizada para gerar ocorrências de possíveis segmentos de interesse, testes paramétricos e não paramétricos [162] foram utilizadas para comparação dos Espaços de Amplitude através da variância.

4.2.2.1 Testes Paramétricos

Testes estatísticos já consagrados podem ser utilizados para comparação da variância entre os Espaços de Amplitudes gerados. Quando se possui conhecimento suficiente sobre a distribuição de uma variável, podemos fazer predições sobre a mesma. Por exemplo, se pegarmos uma amostra com 100 pessoas adultas, e considerar a idade como a variável de interesse, é possível que a maioria das pessoas fiquem mais perto da média, decrescendo a quantidade de pessoas quanto mais distante da média. Isso representa uma distribuição normal, como ilustrado pela Fig. 16.

Em um segundo momento, utilizando-se pessoas de outra cidade, ao qual suspeitamos que são pessoas com altura menor que a média. Se a média de altura daquela população ficar fora dos 95% da nossa distribuição inicial, podemos assumir que temos duas populações distintas em relação a esta variável. Quando utilizamos testes paramétricos, só podemos assumir qualquer análise quando a distribuição dos dados for uma distribuição normal. Para um estudo mais aprofundado sobre distribuição normal e testes paramétricos, vide [162].

Para validar a capacidade de um teste paramétrico de distinguir os Espaços de Amplitudes gerados, é necessário avaliar se os Espaços de Amplitude gerados possuem uma distribuição normal. Para tal, o teste de Kolmogorov-Smirnov (KS) foi utilizado [163]. O teste KS compara a maior diferença vertical entre o Espaço de Amplitude e

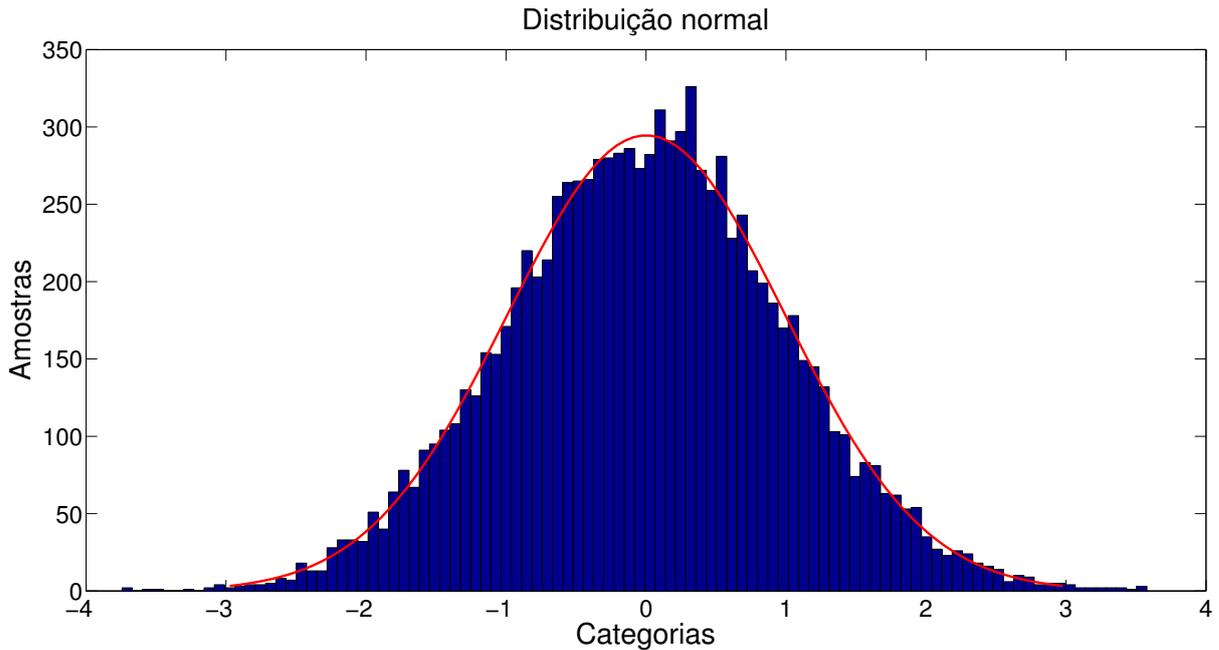


Figura 16 – Exemplo de distribuição normal.

uma distribuição de referência, no caso a distribuição normal. Ele é baseado na Função de Distribuição Empírica (*Empirical Distribution Function, ECDF*) dos dados e não faz qualquer suposição sobre a sua distribuição. O teste pode ser utilizado para avaliar as seguintes hipóteses:

- H_0 : Os dados seguem uma distribuição normal;
- H_1 : Os dados não seguem uma distribuição normal.

Para avaliar a normalidade do Espaço de Amplitude em dados reais, o conjunto de dados apresentado no Capítulo 5 foi utilizado. Do total de 1.404.062.490 de segmentos gerados, somente 14,67% apresentou distribuição normal não rejeitando a hipótese nula com um p-valor de 0,05. A partir da aplicação do teste KS, conclui-se que testes paramétricos não podem ser utilizados.

4.2.2.2 Testes não Paramétricos

Nos casos onde não se têm a certeza da distribuição normal, testes não paramétricos são utilizados. Basicamente, existe pelo menos um teste não paramétrico para cada teste paramétrico. Geralmente eles são divididos em três categorias, sendo elas: teste de diferença entre grupos (amostras independentes), teste de diferença entre variáveis (amostras dependentes) e teste de relação entre variáveis. Em nosso contexto, dado que determinado Espaço de Amplitude não tem relação alguma com o anterior, nos baseamos em amostras totalmente independentes.

Testes não paramétricos são mais apropriados quando utilizados com amostras pequenas. Quando se tem um volume maior de amostras (maior que 100), geralmente faz pouco sentido utilizar testes não paramétricos. Este comportamento é justificado pelo teorema do limite central, onde quanto mais amostras são utilizadas, existe uma tendência natural deste ser uma distribuição normal, mesmo que não exista a certeza da normalidade da variável utilizada [162]. Dessa forma, métodos paramétricos que são mais sensíveis na maioria dos casos são mais apropriados para grandes amostras. Considerando estas informações, e levando em consideração que a criação do Espaço de Amplitude não deve conter mais que 100 unidades, dado que sua representatividade é prejudicada com valores muito altos, ou muito baixos (abaixo de 50).

Em experimentos realizados, como a variável de interesse é a variância, utiliza-se um teste não paramétrico chamado de Brown–Forsythe, que é considerado bastante robusto para esse tipo de medição [164]. Outros testes para medição de variância existem, tais como *F-Test* [162] e o teste de Bartlett [162]. Entretanto, experimentos indicam que estes testes são muito sensíveis ao pressupostos da distribuição normal. Essencialmente, o Brown-Forsythe compara a igualdade de variância ao executar uma Análise de Variância (ANOVA) nos dados de entrada. Os resultados do modelo de ocorrência com teste não paramétrico de Brown–Forsythe foram adicionados no Apêndice A para referência futura.

O problema da utilização de testes não paramétricos aparece quando é analisada a mediana dos Espaços de Amplitudes. Dado que o Espaço de Amplitude foi gerado a partir de um sinal não normalizado, a mediana do Espaço de Amplitude varia de acordo com o sinal capturado. Neste caso, a alta oscilação da mediana inviabilizou a implementação de testes não paramétricos quando realizada visando detecções exatas da quantidade de cliques emitidos.

4.2.2.3 Lógica Difusa

A lógica *fuzzy* [165], também conhecida como lógica nebulosa ou difusa, se tornou conhecida a partir de 1965. Contudo, os princípios apresentados por esta lógica já existiam bem antes, por volta de 1920 quando um lógico chamado Jan Łukasiewicz utilizando-se do princípio da incerteza apresentou pela primeira vez as noções da lógica [166].

A lógica difusa é baseada na teoria dos conjuntos difusos e o raciocínio exato corresponde a um caso limite do raciocínio aproximado, sendo interpretado como um processo de composição de relações nebulosas. Ao contrário da lógica convencional, a lógica difusa utiliza da ideia de que todas as coisas admitem temperatura, velocidade, etc, ou seja, ela suporta os modos de raciocínio que são aproximados, ao invés de exatos. Tradicionalmente, uma proposição lógica tem dois extremos: ou completamente verdadeiro ou completamente falso. Entretanto, na lógica difusa uma premissa varia em grau de verdade de 0 a 1, o que leva a ser parcialmente verdadeira ou parcialmente falsa.

Uma variedade de aplicações é encontrada em vários campos que vão desde controle de processos industriais, diagnóstico médico, negociação de valores mobiliários, até sistemas de controle de processos em tempo real.

Para detecção de sinais de interesse utilizando a variância, o comportamento humano durante as detecções foi utilizado para criação das regras. A Fig. 17 mostra como as regras foram definidas. Durante o processo de detecção realizado por especialistas, existem duas certezas que podem ser considerados como cliques ou não, juntamente com uma área de incerteza com maior dificuldade de discriminação.

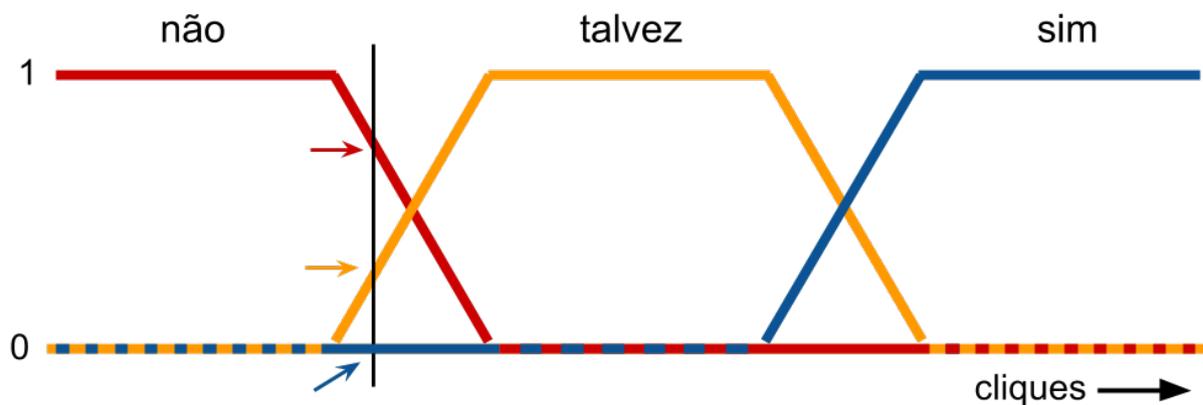


Figura 17 – Regras utilizadas pela lógica difusa.

O problema encontrado ao realizar testes com essa abordagem é a necessidade de criar um limiar definindo o corte entre as classes indicando se existe clique ou não. Outras variáveis foram consideradas para inserir informações extras na área de incerteza do modelo. A média do Espaço de Amplitude sempre será a mesma, já que a quantidade de amostras do sinal utilizada para criação do segmento nunca varia. A mediana como apresentado na Seção 4.3 possui características que inviabilizam sua utilização dada pela não normalização do sinal capturado.

4.3 ORDENAÇÃO

Como não é realizado uma normalização prévia, o Espaço de Amplitude perde seu padrão criado a partir da função de histograma. Quando se tem dois ou mais Espaços de Amplitude gerados a partir de segmentos distintos, as unidades de mesmo índice não correspondem ao mesmo espaçamento de amplitude, já que os valores mínimo e máximo variam entre os segmentos. O Espaço de Amplitude da Fig. 18 (b) e (c) foram gerados a partir de um sinal com clique (a). A Fig. 18 (e) e (f) também foram gerados a partir de um sinal com clique (d), entretanto, quando capturado, o clique estava em baixa intensidade. Os últimos Espaços de Amplitude na Fig. 18 (h) e (i) mostram o que acontece quando gerado a partir de um sinal (g) com nenhum clique presente. A maior dificuldade neste

caso é realizar a comparação entre os Espaços de Amplitude sem nenhuma correlação entre suas unidades de espaçamento, de forma que um algoritmo de Aprendizado de Máquina seja capaz de inferir e rotular cada um dos segmentos.

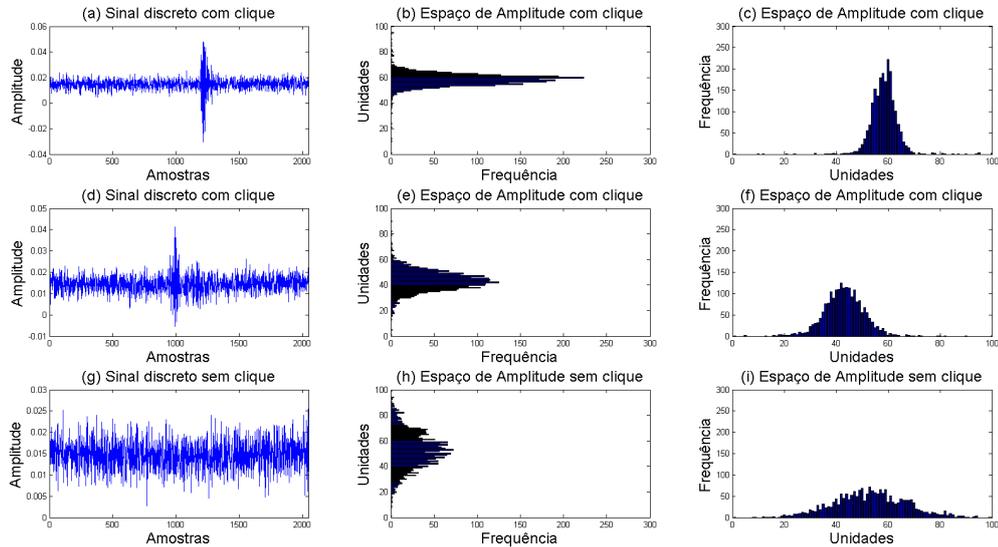


Figura 18 – Espaço de Amplitude sem ordenação.

Para resolver este problema, uma ordenação das unidades do Espaço de Amplitude é realizada, obtendo uma distribuição diferenciada do padrão a ser reconhecido. Ao invés de analisar a quantidade total de pontos discretos em determinada unidade, foi analisado a quantidade de pontos discretos nas unidade que possuem mais pontos. Como pode ser visto na Fig. 19 (c), (f) e (i), após ordenado, algoritmos de Aprendizado de Máquina podem ser utilizados para reconhecer o padrão de decaimento da quantidade total de pontos em determinada unidade do Espaço de Amplitude. Isso proporciona, por exemplo, que a unidade 1 de cada Espaço de Amplitude ordenado possa ser comparada com a unidade 1 de outros segmentos.

Desde os primórdios da computação, problemas de ordenação vêm atraindo atenção de pesquisadores, existindo registros datados desde 1956 [167]. Apesar de ser um processo familiar na vida das pessoas, para resolver um problema de ordenação de maneira eficiente através de meios computacionais, pode-se ter um aumento na complexidade do processo como um todo. A ordenação dos dados pode ser considerada a fase mais crucial da etapa de detecção. Caso o algoritmo de ordenação não seja escolhido corretamente, todo o ganho computacional esperado até o momento pode ser perdido. Até essa etapa, o ganho computacional esperado é dado em especial pela não utilização de algoritmos para a conversão do domínio da frequência. Para cálculo efetivo do ganho computacional do algoritmo proposto, uma análise extensa da metodologia utilizada deve ser avaliada [144]. Outros ganhos podem ser esperados quando os algoritmos apresentados forem

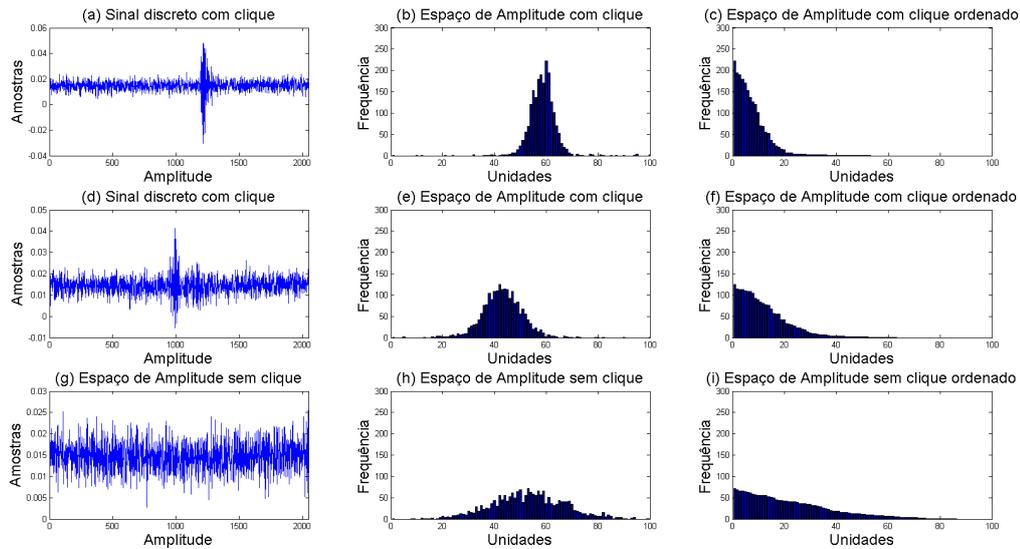


Figura 19 – Espaço de Amplitude com ordenação.

implementados em conjuntos de instruções otimizados para cada máquina, levando em consideração a arquitetura do processador em questão, bem como compilador utilizado para gerar as instruções. Comparações com outros algoritmos devem utilizar a mesma arquitetura computacional e compilador. Ressalta-se neste caso, que as implementações propostas para o fluxo de detecção são todas $O(n)$ quando avaliado seu custo assintótico.

Algoritmos de ordenação são categorizados através da sua complexidade computacional, podendo ser avaliado pelo melhor, médio e pior caso. O melhor caso, raramente encontrado em problemas reais, é quando um algoritmo de ordenação recebe um conjunto de dados totalmente ordenado, não sendo necessário realizar nenhuma operação extra. O caso médio possui maior relação com conjuntos de dados de problemas reais, onde uma parte dos dados já está ordenada. O pior caso é aquele onde todos os dados estão desordenados, sendo necessário a execução de diversas comparações e operações de troca para uma completa ordenação. Apesar do caso médio ser notoriamente o mais comum em problemas reais, o pior caso normalmente é considerado como base para comparação de algoritmos de ordenação através na notação *Big O* [168].

4.3.1 Ordenação Simples

Existem três algoritmos que são conhecidos como de ordenação simples: *Insertion Sort*, *Selection Sort* e *Bubble Sort*. Estes algoritmos baseiam-se em intuições comumente utilizadas por pessoas normais em tarefas de ordenações na vida real. Algoritmos de ordenação simples possuem um bom desempenho quando aplicado em pequenos conjuntos de dados, mas se tornam extremamente ineficientes quando a quantidade de dados aumenta.

Insertion Sort [169] é um algoritmo de ordenação simples, relativamente eficiente

para conjuntos de dados pequenos e quase ordenados [169]. Devido a sua eficiência nestes casos, normalmente ele é usado em conjunto com algoritmos de ordenação mais sofisticados. O algoritmo funciona da seguinte maneira: elemento por elemento é avaliado e colocado na sua posição ordenada. Isso requer que todos os dados sejam percorridos e comparados para saber onde o valor se encaixa melhor. Com isso, quando o conjunto de dados a ser ordenado é grande, e bastante desordenado, o *Insertion Sort* acaba se tornando um algoritmo computacionalmente caro com uma complexidade computacional de $O(n^2)$ para o pior caso.

Selection Sort [170] é um algoritmo de ordenação simples com complexidade alta para o pior caso. O algoritmo procura pelo menor valor e troca com o elemento na primeira posição. Feito isto, é realizada a mesma operação com o segundo elemento, e assim sucessivamente até que todos os valores sejam analisados. É um algoritmo pouco utilizado em grandes conjuntos de dados, e normalmente é mais lento que o *Insertion Sort*. Apesar disso, dependendo da implementação e considerando os dados de entrada, o *Selection Sort* pode ser executado mais rápido que o *Insertion Sort* [170]. O *Selection Sort* possui ainda algumas vantagens em certas situações de entrada de dados, reforçando a importância do conhecimento prévio dos dados na escolha de algoritmos de ordenação, mas com uma complexidade para o pior caso de $O(n^2)$.

Bubble Sort [171] também é considerado um algoritmo de ordenação simples. O algoritmo se inicia pela primeira posição do conjunto de dados e compara seu valor com a segunda posição e uma troca de posição é realizada, caso seja necessário. Em seguida, compara-se o segundo com o terceiro elemento, e assim sucessivamente até chegar no final do conjunto de dados. Ao atingir o final, o processo é reiniciado do primeiro elemento, realizando as mesmas comparações. O processo termina quando se passa por todo o conjunto de dados e nenhuma troca é realizada. O caso médio deste algoritmo é o mesmo do pior caso $O(n^2)$, por isso é raramente utilizado para ordenar grande conjunto de dados, bem como dados pouco ordenados. Em contrapartida, para dados de qualquer tamanho quase ordenados, este algoritmo pode ser utilizado com uma eficiência interessante. Neste caso, o conhecimento previsto do domínio aplicado é necessário.

Dependendo das configurações de criação do Espaço de Amplitude, a simplicidade do *Insertion Sort* poderia ser considerada como uma opção, devido a quantidade mínima de comparações e bom desempenho em dados quase ordenados. Em contrapartida, o *Selection sort* utiliza menos escrita para ordenar os dados, otimizando o acesso a memória, consequentemente, tornando uma boa opção quando os dados estão mais desordenados.

4.3.2 Ordenação Eficiente

Algoritmos de ordenação mais eficientes para um conjunto de dados grande, mais utilizados em problemas reais, normalmente são baseados na complexidade média. Os

mais comuns são: *Heap Sort*, *Merge Sort* e *QuickSort*.

Merge Sort [168] é baseado no paradigma de divisão e conquista. O algoritmo se inicia comparando dois elementos da lista. O primeiro elemento é comparado com o segundo, depois o terceiro com o quarto, e assim sucessivamente, trocando de posição, caso exista necessidade de ordenação. Na próxima iteração, é realizado a ordenação de quatro elementos, com duas listas previamente ordenadas na etapa anterior. Sendo assim, a quantidade de itens nas listas vai aumentando e os itens vão sendo ordenados até que se chegue ao final da lista. Pela sua abordagem de divisão e conquista, o algoritmo escala bem para listas muito grandes com uma complexidade do pior caso de $O(n \log n)$. Dado os conceitos de Orientação a Objetos, esse é um algoritmo que se aplica bem em listas, e não somente em *arrays*, visto que é requerido somente acesso sequencial e não randômico provido em *arrays*. Entretanto, uma complexidade de espaço adicional é necessária, bem como um grande número de cópias em implementações mais simples.

Heap Sort [168] é uma versão mais eficiente do *Selection Sort*. Ele também funciona por determinar o menor elemento na lista, colocando este elemento no início ou no fim da lista, continuando até o final. A grande diferença é que neste caso uma estrutura de dados chamada de *Heap* [168] é utilizada. Quando os dados da lista são adicionadas na *Heap*, é garantido que o nó principal é o maior elemento. Quando ele é removido e colocado novamente na lista, a *Heap* se atualiza, colocando novamente o maior elemento dos restantes no nó principal. Esse processo é contínuo até que todos os elementos sejam adicionados novamente na lista original. A complexidade assintótica *Heap Sort* é de $O(n \log n)$.

Quick Sort [168] é outra abordagem baseada no paradigma de divisão e conquista. Para realizar a divisão do *array*, um elemento chamado pivô é selecionado. Todos os elementos menores que o pivô são movidos para a esquerda, e todos os maiores movidos para a direita. O processo é então realizado recursivamente em cada uma das listas geradas à esquerda e à direita. Isso possibilita uma complexidade média boa, fazendo essa uma abordagem popular na prática. É importante ressaltar a importância da seleção do pivô, que pode levar a complexidade computacional para $O(n^2)$ mesmo em casos em que a lista já está totalmente ordenada, bem como possibilitar que dados totalmente ordenados não sejam considerados o pior caso. Por exemplo, quando o primeiro item é selecionado como o pivô em uma lista já ordenada, o *Quick Sort* será $O(n^2)$.

Assim como os algoritmos de ordenação simples, cada algoritmo de ordenação possui suas vantagens e desvantagens. Por exemplo, uma implementação simples do *Merge Sort* utiliza um espaço adicional do tamanho do conjunto de dados. Já implementações simples do *Quick Sort* possuem uma complexidade no pior caso elevada. Esses problemas podem ser evitados ou amenizados com a implementação de algoritmos mais complexos. Isso porque, apesar de serem considerados eficientes em dados aleatórios, para obter a

eficiência em dados reais diversas modificações podem ser implementadas.

O custo desses algoritmos em conjunto de dados menores são muito altos, por isso normalmente ordenações híbridas são utilizadas quando o conjunto de dados se torna pequeno. Por exemplo, a implementação recursiva do *Merge Sort* pode utilizar o critério de parada como sendo o tamanho de cada bloco, e quando alcançado, um algoritmo de ordenação simples é chamado para iniciar a volta da recursão.

4.3.3 Ordenação de Inteiros

Outros métodos de ordenação mais eficientes podem ser utilizados, dependendo do tipo de dados que se deseja ordenar. Algoritmos de ordenação de inteiros, por exemplo, são aqueles que ordenam um conjunto de valores pelas suas chaves numéricas inteiras. Algoritmos desse tipo também podem ser aplicados para outros tipos de dados como, por exemplo, ponto flutuante ou *string* [172]. A possibilidade de aplicação de aritmética inteira proporciona que estes tipos de algoritmos sejam mais rápidos que os algoritmos de comparação apresentados anteriormente. Mas a aplicação desse tipo ordenação depende da domínio aplicado, bem como do tamanho dos inteiros a serem ordenados.

Counting Sort e *Radix Sort* são exemplos de algoritmos de ordenação de inteiros extremamente eficientes quando as chaves são pequenas. Para casos onde os possíveis valores são grandes, esse tipo de método se torna lento, sendo neste caso melhor a utilização de algoritmos de ordenação de comparação. É justamente por essa característica que esse tipo de algoritmo é mais eficiente quando ordena-se o Espaço de Amplitude, em especial pelo fato dos valores máximos poderem ser controlados pelo tamanho do segmento. Algoritmos de ordenação de inteiros dependem basicamente de três parâmetros: o número de valores n a serem ordenados, a magnitude do maior valor k a ser ordenado, e o número de bits w que pode ser representado no computador em que o algoritmo será executado.

Algoritmos de ordenação de inteiros ainda possibilitam otimizações de baixo nível relacionado ao hardware aplicado. Esta é outra vantagem aparente em dispositivos fixos autônomos, onde o hardware a ser executado pode ser escolhido levando em consideração estes aspectos. É o caso de tabelas de mapeamento necessárias por alguns algoritmos, que podem ser substituídas por operações mais simples implementadas para manipulação de bits direto em hardware [173].

Radix Sort é um algoritmo de ordenação não comparativo de inteiros, que agrupa dígitos individuais que compartilham a mesma significância e valor com complexidade de $O(nk)$. Dado que inteiros podem representar outros tipos de dados através de *strings*, o *Radix Sort* pode ser utilizado para dados além de inteiros. O algoritmo pode ser dividido em duas categorias, os que utilizam os dígitos menos significantes (*Least Significant Digit*, LSD) e os que utilizam os dígitos mais significantes (*Most Significant Digit*, MSD). Ambos se iniciam pelo seu dígito significativo, e vai em direção ao outro. Por exemplo, o LSD

começa do dígito menos significativo e processa as representações inteiras até o mais significativo.

O *Counting Sort* é um algoritmo de ordenação otimizado para ordenar uma coleção de objetos de acordo com seus valores, que devem ser pequenos para realmente ser efetivo. Ele é executado contando a quantidade valores distintos, sendo o seu tempo de execução linear com o número de itens e a diferença entre o menor e maior valor. Dessa maneira, seu uso é recomendado somente em casos onde a variação entre valor mínimo e máximo não é significativamente maior que o número de itens. Pela sua característica, normalmente ele é utilizado em conjunto com *Radix Sort* para execução em grande volumes de dados mais eficientemente.

Basicamente, o algoritmo passa pelos itens calculando um histograma do número de vezes que cada valor ocorre nos dados de entrada. Após isto, ele faz uma segunda passagem, dessa vez no histograma para ver quantas vezes cada valor apareceu. Dessa forma, quando o item é encontrado no histograma ele é colocado de volta nos dados, que por sua vez, já estão ordenados.

A desvantagem desse tipo de ordenação é a utilização de um *array* que deve ser do tamanho da diferença entre o valor máximo e o valor mínimo possível. Esse espaço de memória é utilizado para criar um outro *array*, onde os índices são os valores do *array* a serem ordenados. Assim, essa abordagem não é muito adequada para valores muito grandes, pois acaba criando a necessidade de um espaço de memória maior e, conseqüentemente, uma pesquisa maior em todo o histograma.

Assim como apresentado pelo Algoritmo 3, a entrada do *Counting Sort* é um *array* de inteiros desordenados, onde o algoritmo realiza a ordenação e retorna como saída o mesmo conjunto de dados ordenados. Primeiramente, o *Histograma* é inicializado com zeros. O primeiro *loop* realiza a contagem de quantos itens nos *Dados* se encaixam em cada item do *Histograma*. O segundo *loop*, serve para pegar os valores contados no

Histograma, e colocar novamente nos *Dados*, mas dessa vez ordenados.

Algoritmo 3: *Counting Sort*

Entrada: Conjunto de dados desordenados

Saída: Conjunto de dados ordenados

início

 Inicializa o *Histograma* com zeros;

para *Contador* $\leftarrow 0$ **até** *TamanhoDados* **faça**

 Incrementa um ao *Histograma* de acordo com *Dados* no índice *Contador*;

fim

para *Contador* $\leftarrow 0$ **até** *TamanhoHistograma* **faça**

 Adiciona o valor *Contador* ao conjunto de *Dados* ordenados;

 Decrementa um do *Histograma* no índice *Contador*;

fim

fim

Dado que o algoritmo utiliza somente *loop* simples, sem recursão ou chamadas a sub-rotinas, sua análise é direta. O primeiro **para-faça**, responsável pela inicialização do histograma, realiza uma iteração de no máximo $O(n)$. O segundo **para-faça**, utilizado para colocar os valores de volta aos dados ordenados, deve percorrer o Histograma. Ele possui um pior caso de $O(k)$, onde k é o tamanho do histograma. Portanto, obtém-se uma complexidade final como sendo o somatório das duas etapas, ou seja, $O(n + k)$.

Como foi utilizado um *array* para representar o histograma, tem-se um gasto de memória a mais de $O(k)$. Assim, o algoritmo gasta no máximo $O(n + k)$ de memória. Para problemas que o valor máximo é significativamente menor que a quantidade de itens, o *Counting Sort* pode ser extremamente eficiente, dado que o tamanho necessário para o histograma é de no máximo o valor máximo do problema.

Apesar de qualquer método de ordenação apresentado poder ser utilizado, o mais indicado nas condições em questão é o *Counting Sort*. Seu uso é justificado por diversas questões. Primeiramente, temos o tamanho do vetor gerado. Assim como explicado anteriormente, a quantidade de unidades utilizadas pelo Espaço de Amplitude é de 100. Devido ao tamanho tão pequeno do vetor, algoritmos mais simples de ordenação são mais eficientes. O problema de uso deste tipo de ordenação é o tamanho dos valores do Espaço de Amplitude, que podem ficar muito grandes devido à quantidade de amostras utilizadas pela captura dos dados. Isso irá acontecer inevitavelmente quando cliques de alta frequência forem detectados. Apesar disso, para a maioria dos dispositivos disponíveis hoje em dia, isso não será um problema dado que o vetor de chaves ordenadas cabe na maioria das memórias *cache* dos processadores.

Sabe-se que para este trabalho não existe a necessidade de paralelização, dado a natureza sequencial de captura dos dados, e que a implementação apresentada pelo

Algoritmo 3 tende a manter o histograma zerado a cada execução. Assim, é possível alocar um espaço de memória global e ser utilizado durante toda a execução do algoritmo. Dado que o tamanho dos segmentos variam entre 1.024 amostras e 2.048, tem-se um espaço relativamente pequeno que pode ser mantido na *cache* do processador o tempo inteiro. Com isso, o tempo de execução é reduzido consideravelmente.

4.4 APRENDIZADO DE MÁQUINA

A última fase da etapa de detecção é realizada através de um processo de Aprendizado Supervisionado [174] construído com ferramentas específicas de Aprendizado de Máquina [175, 176]. No aprendizado supervisionado, um conjunto de dados previamente rotulado serve de referência para o processo de treinamento de um classificador que será utilizado para a predição de novas amostras através de uma hipótese de indução. Para resolver um problema de aprendizagem supervisionada, normalmente os seguintes passos são levados em consideração:

1. Determinar o tipo do conjunto de dados de treinamento. Nesta fase serão utilizados os Espaços de Amplitude ordenados que foram gerados a partir de uma função de histograma aplicada nos segmentos do sinal capturado;
2. Montar o conjunto de dados para treinamento a partir da representação do sinal já definida. Esta tarefa deve ser executada por especialistas com conhecimento prévio de cliques das espécies de interesse. Basicamente o especialista irá rotular os sinais baseando-se em sua experiência;
3. Determinar qual algoritmo será utilizado para gerar a hipótese de indução capaz de prever novos cliques. Diversas opções estão disponíveis, sendo Árvores de Decisão, Redes Neurais Artificiais e Máquinas de Vetores de Suporte alguns exemplos. Uma revisão da literatura sobre as possibilidades de aprendizado supervisionado pode ser vista em [174];
4. Avaliar a acurácia da hipótese gerada com a aplicação de uma validação cruzada. Assim como detalhado no Capítulo 5, o *k-fold* foi utilizado para validação da hipótese gerada. O conjunto de treinamento foi dividido em 10 (k) partes distintas, sendo 9 ($k-1$) utilizadas para treinamento e a última restante utilizada pra validação. Esse processo é executado 10 (k) vezes até que todas as partes tenham sido utilizadas para validação.

Existem diversos métodos para aprendizado supervisionado, cada um com suas vantagens e desvantagens. Usualmente nenhum método tem desempenho efetivo para diferentes tipos de problemas supervisionados. Como o objetivo desta etapa é detectar

a presença de um clique, ou seja, um rótulo binário indicando verdadeiro para cliques presentes e falso caso contrário, tem-se um problema de classificação binária. Um algoritmo baseado em margem foi escolhido, de forma a criar um hiperplano capaz de separar as duas classes de interesse maximizando a margem que as dividem.

Um dos problemas fundamentais na teoria de aprendizado é a construção de classificadores capazes de uma boa generalização. Neste sentido, o desenvolvimento de teorias baseadas em margem foi um passo importante na construção de classificadores eficientes. Como pode ser visto na Fig. 20, as amostras são rotuladas em duas classes, os círculos preenchidos e os círculos vazios. O hiperplano H_1 não consegue separar as duas classes. O hiperplano H_2 consegue separar, mas com uma margem pequena. O hiperplano H_3 , por sua vez, consegue separar as classes com a margem máxima possível entre elas, minimizando assim o possível erro de novas classificações.

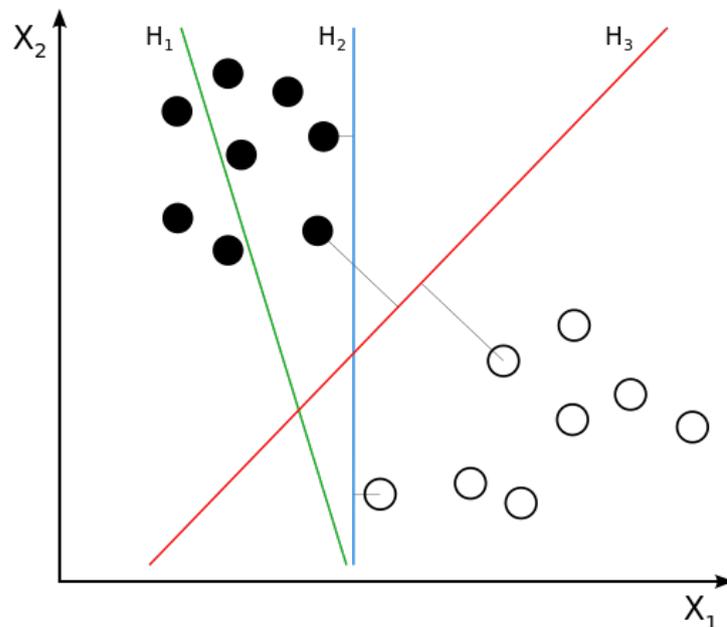


Figura 20 – Máxima margem, extraída de [177].

Visando a obtenção da máxima margem, as Máquinas de Vetor Suporte (SVM) foram desenvolvidas [178]. O SVM é um algoritmo baseado na ideia de que a generalização de um classificador linear pode ser melhorada quando encontrado um hiperplano que maximize a distância entre classes opostas. Para a construção deste hiperplano, um problema de programação quadrática deve ser resolvido.

Atenção considerável foi dada a procura de métodos computacionalmente mais eficientes para construção de classificadores de margem máxima, visando evitar a complexidade de problemas de programação quadrática. A maioria destes algoritmos são motivados pela busca da margem ótima, ou quase ótima, visando obter boa capacidade de generalização.

Neste contexto, o *Incremental Margin Algorithm* (IMA) [179] foi escolhido para realização de testes no Capítulo 5. O IMA é um algoritmo *online* e eficiente capaz de processar amostras isoladas por vez, construindo uma aproximação da margem máxima em cada iteração. Em nosso caso, determina-se uma margem inicial para evitar a utilização de um valor muito fora de um valor razoável de margem para os dados em questão, definida utilizando uma base previamente preparada com dados do início do monitoramento, discriminada por especialistas.

A metodologia proposta pelo IMA divide o processo em duas etapas iterativas. A primeira chamada de Perceptron de Margem Fixa (*Fixed Margin Perceptron*, FMP) é uma extensão do algoritmo Perceptron de Roseblatt [180] e mostrado pelo Algoritmo 4. Ele encontra uma solução linearmente separável do problema dado uma margem fixa.

Algoritmo 4: *Fixed Margin Perceptron.*

Entrada: $z_m, w_{init}, \gamma_f, \eta, T_MAX$

início

$w^0 \leftarrow w_{init}, t \leftarrow 0;$

repita

para $i \leftarrow 0$ **até** m **faça**

se $(y_i \langle x_i, w^t \rangle < \gamma_f \|w^t\|)$ **então**

$w^{t+1} \leftarrow w^t \lambda_t + \eta y_i x_i;$

$t \leftarrow t + 1;$

fim

fim

até (*não foram cometidos erros*) **ou** $(t > T_MAX);$

retorna $w^t;$

fim

A segunda etapa do método apresentada pelo Algoritmo 5, utiliza uma estratégia baseada nas margens de cada classe para alterar gradativamente a margem fixa utilizada pelo FMP com intuito de encontrar uma margem máxima similar a encontrada pelo SVM. Uma das vantagens dessa abordagem é que sempre é encontrada uma solução efetiva. Dessa forma, depois que a primeira iteração do FMP é finalizada, o algoritmo pode ser interrompido a qualquer momento com uma solução factível. Isso pode ser extremamente importante para aplicações dependente do tempo de execução, como e o caso de dispositivos

fixos autônomos.

Algoritmo 5: *Incremental Margin Algorithm.*

Entrada: z_m, η, δ, T_MAX

início

$w \leftarrow 0, \gamma_f \leftarrow 0;$

repita

$w \leftarrow \text{FMP}(z_m, w, \gamma_f, \eta, T_MAX);$

$\gamma_f \leftarrow \max((\gamma^+(w) \gamma^-(w))/2, (1 + \delta) \gamma_f);$

até convergência do FMP em T_MAX não for alcançada;

retorna último w factível;

fim

Os algoritmos apresentados para o IMA são utilizados durante seu processo de treinamento. Quando executado em ambiente marinho, a princípio será executado somente o algoritmo de classificação apresentado pelo Algoritmo 6.

Algoritmo 6: *Incremental Margin Algorithm Classify.*

Entrada: z_m, η, δ, T_MAX

início

para $k \leftarrow 0$ até *Dimensao* **faça**

$func \leftarrow func + w[k] * \text{sample} - > \text{points}[i].x[k];$

fim

se ($func < 0$) **então**

 sinal de interesse;

fim

senão

 não é sinal de interesse;

fim

fim

O Algoritmo 6 utilizado para classificação se é um sinal de interesse ou não tem a variância como primeiro atributo, e os primeiros dezenove atributos do Espaço de Amplitude ordenado. Ao utilizar somente a variância, não conseguimos poder discriminatório suficiente para casos com cliques em intensidade reduzida similar a ruídos, e caso utilize-se todas as unidades disponíveis no Espaço de Amplitude, um sobreajuste ocorre em função da similaridades de muitas unidades. Como pode ser visto pela Fig. 21, somente as unidades iniciais do Espaço de Amplitude possuem relevância discriminatória perante ao Espaço de Amplitude completo. No caso do Algoritmo 6 tem-se uma complexidade $O(d)$ em função da dimensão escolhida com os valores do Espaço de Amplitude, que no caso da Fig. 21 é de 19.

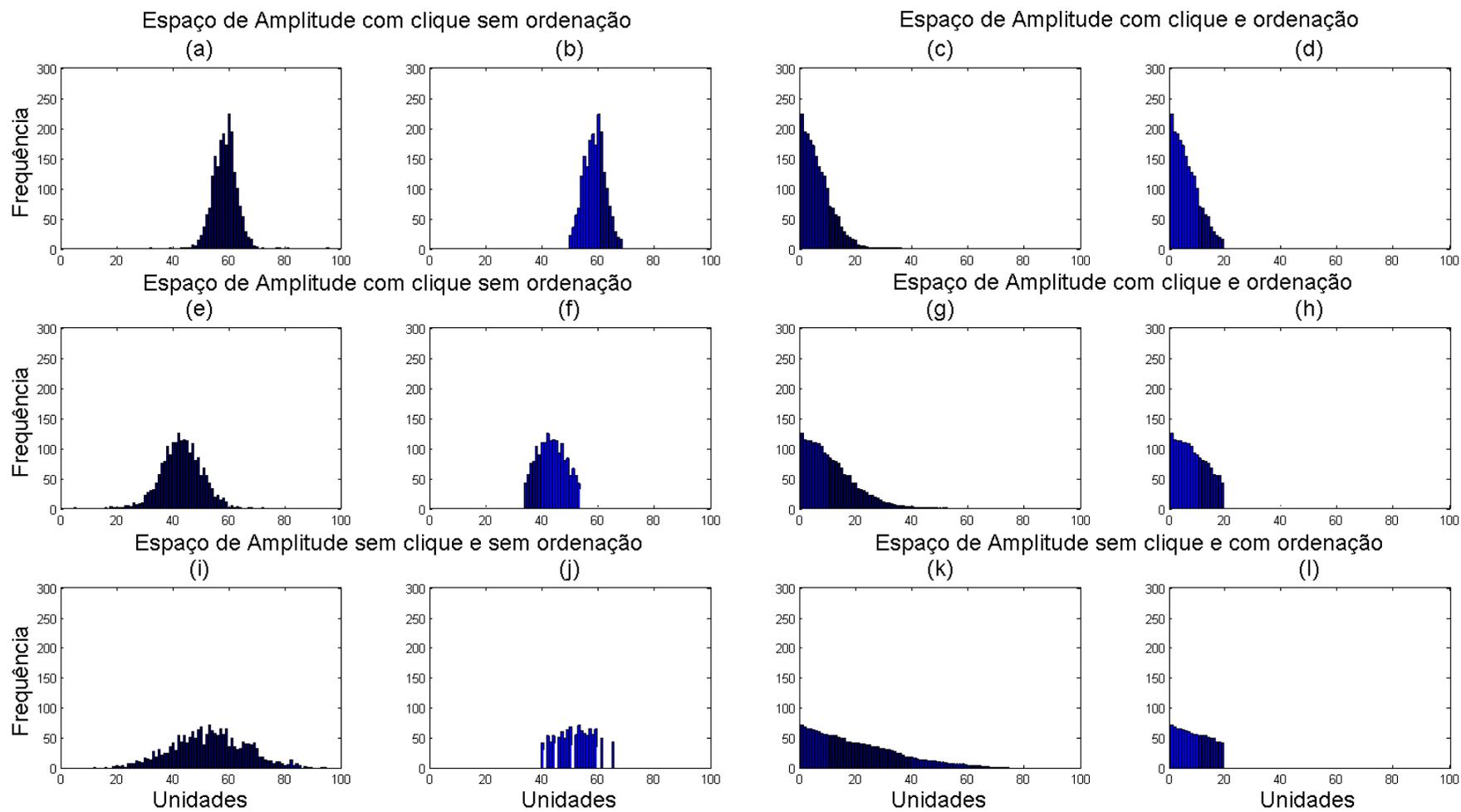


Figura 21 – Unidades relevantes do Espaço de Amplitude.

5 RESULTADOS EXPERIMENTAIS

O desenvolvimento de algoritmos eficientes, efetivos e padronizados para detecção de diversas espécies de mamíferos marinhos requer conjuntos de dados anotados e validados. A aquisição, manutenção e disponibilidade destes conjuntos necessita de esforço colaborativo entre diversos pesquisadores [11]. Para avaliar a acurácia e eficiência da etapa de detecção, de forma a reduzir o volume total de dados a ser processado, o algoritmo proposto foi executado em um conjunto de dados anotado por especialistas e disponibilizado pelo *Detection, Classification, Localization, and Density Estimation (DCLDE) Workshop 2015* [181].

Os dados consistem de sinais capturados pelo HARP em múltiplos locais e estações do ano, fornecidos como arquivos *Waveform Audio File Format (WAVE)*, onde o nome indica o local e o início temporal da captura [16]. O conjunto de dados foi dividido em sinais de Mysticetos de baixa frequência e Odontocetos de alta frequência. Os dados de baixa frequência consistem de anotações de duas espécies de Mysticetos, as Baleias Azuis (*Balaenoptera musculus*) e as Baleias Fin (*Balaenoptera physalus*), sendo armazenados com largura de banda de 1 e 1,6 *kHz*. Já os arquivos de alta frequência com sinais produzidos por Odontocetos - utilizados neste trabalho -, foram armazenados com largura de banda de 100 e 160 *kHz*, e consistem de múltiplas espécies de Odontocetos, como mostrado na Tabela 1.

Abreviação	Espécie
Bb	<i>Berardius bairdii</i> - Baleia Bicuda de Baird
Zc	<i>Ziphius cavirostris</i> - Baleia Bicuda de Cuvier
Pm	<i>Physeter macrorhynchus</i> - Baleia Cachalote
Lo	<i>Lagenorhynchus obliquidens</i> - Golfinho de Laterais Brancas do Pacífico
Gg	<i>Grampus griseus</i> - Golfinho de Risso
UPP	<i>Phocoenidae</i> - Golfinho não identificado
UO	<i>Odontoceti</i> não identificado

Tabela 1 – Espécies de Odontocetos capturadas.

A Fig. 22 mostra que o HARP foi implantado em sete locais distintos, proporcionando gravações entre 2009 e 2013. O objetivo deste conjunto de dados foi identificar encontros de espécies que estavam emitindo cliques de ecolocalização. O conjunto de dados foi examinado por especialistas, sendo anotado o início e fim aproximado do período de captura. Qualquer detecção separada por mais de 5 minutos foi anotada como um encontro separado.

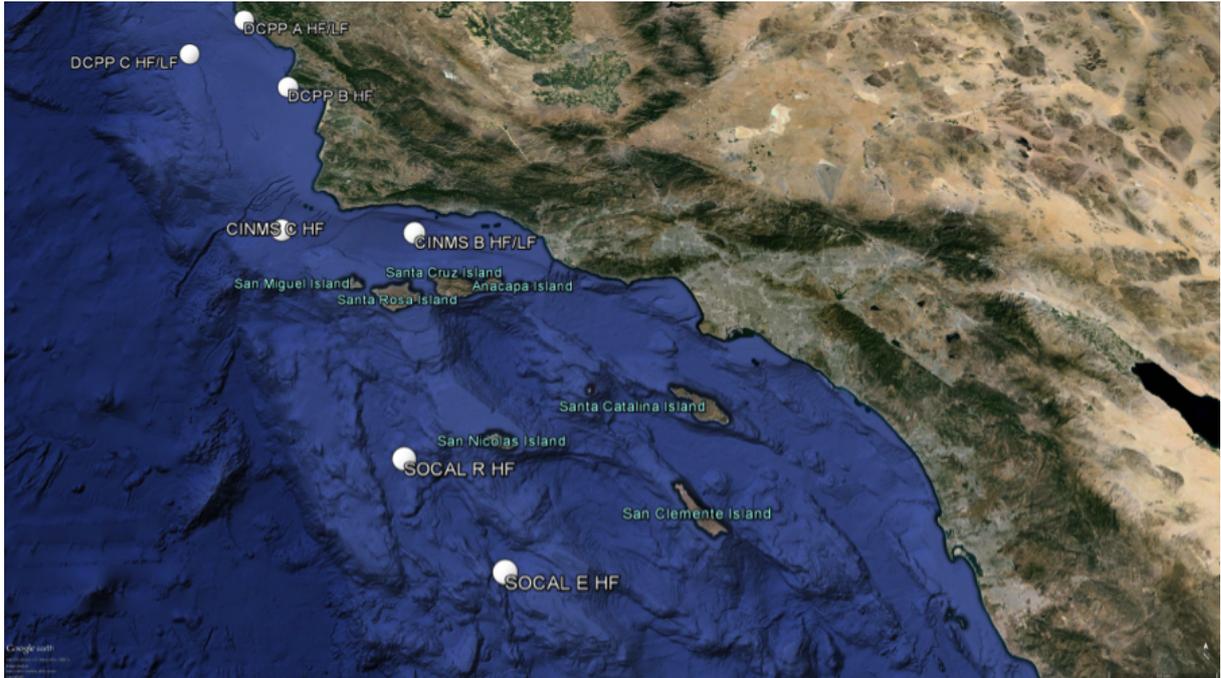


Figura 22 – Locais de implantação do HARP, extraída de [181].

Informações de localização exata da captura podem ser visualizadas na Tabela 2, onde são apresentados o nome do projeto, o sítio de captura, a profundidade em metros, a largura de banda em kHz e as coordenadas. Para referenciar os locais de captura, este trabalho utilizou uma combinação entre o nome do projeto e o sítio de captura. Por exemplo, CINMS B se refere ao projeto CINMS quando capturado no sítio B.

Projeto	Sítio	Profundidade	Largura de Banda	Latitude	Longitude
CINMS	B	600	100	34-17.0 N	120-01.7 W
CINMS	C	800	160	34-19.5 N	120-48.4 W
DCPP	A	65	160	35-36.7 N	121-14.5 W
DCPP	B	100	160	35-09.6 N	120-53.1 W
DCPP	C	1.000	100	35-24.0 N	121-33.8 W
SOCAL	E	1.300	100	32-39.4 N	119-28.4 W
SOCAL	R	1.200	100	33-09.6 N	120-00.6 W

Tabela 2 – Detalhes dos locais de implantação do HARP.

A parametrização da etapa de detecção considerou o tamanho dos segmentos para as capturas de $100 kHz$ como $2,56 ms$, sendo o total de 200 mil amostras por segundo divididas em aproximadamente 390 segmentos de 512 amostras. Para os casos com largura de banda de $160 kHz$, o segmento é de $1,6 ms$, dividindo o total de 320 mil amostras por segundo em 625 segmentos de 512 amostras. Neste caso, note-se a parametrização

do tamanho do segmento em função da quantidade de amostras independente da taxa de amostragem do sinal.

Para realizar o treinamento do IMA, um total de 5.600 amostras foram utilizadas, sendo 2.800 rotuladas como de interesse, e outras 2.800 rotuladas como de não interesse. As amostras rotuladas como de interesse possuem exemplos com cliques variando de baixa a alta intensidade, inclusive com ruído natural do ambiente implantado. Além disso, dentre estas amostras estão inclusos exemplares de quase todas as espécies (Bb, Zc, Pm, Lo, UO) apresentadas na Tabela 1. As amostras rotuladas como de não interesse possuem exemplares com diversos tipos de ruídos, sendo os mais notáveis movimentos bruscos no hidrofone e chuva. É importante ressaltar que o treinamento aqui apresentado pode ser utilizado para detectar uma vasta gama de espécies distintas, grande parte devido a similaridade de diferentes cliques no domínio do tempo.

Assim como mostrado pela Tabela 3, todas as amostras para treinamento foram retiradas do projeto SOCAL sítio R. Do total de aproximadamente 335 horas, somente 97 horas foram consideradas de interesse por anotações de especialistas. Dentre as 97 horas, um minuto foi extraído de cada dia de captura anotado e manualmente analisado a procura de 200 amostras com cliques. Outras 200 amostras sem cliques foram escolhidas de maneira aleatória dentro do minuto extraído, totalizando 400 amostras por dia em um total de 14 dias.

Para validação do treinamento, uma análise manual por especialistas foi realizada em três minutos distintos. Cada minuto foi extraído de projetos e sítios diferentes do utilizado pelo treinamento, categorizando amostras novas de capturas reais. Os minutos foram escolhidos levando em consideração algumas características importantes, comumente encontradas em ambientes reais:

- O minuto extraído do projeto CINMS sítio B foi escolhido pela alta taxa de cliques por minuto, onde são encontrados cliques bem próximos uns dos outros;
- O segundo minuto extraído do projeto DCPA A, foi selecionado pela alta taxa de ruído e baixo discernimento entre os cliques, boa parte devido a emissões de intensidade moderada a baixa; e
- Para finalizar, o terceiro minuto foi extraído do projeto SOCAL sítio E, que contém um volume considerável de cliques com intensidade moderada a alta.

Para anotação completa de cliques pelos especialistas, os dados foram analisados visualmente e acusticamente utilizando uma ferramenta computacional de processamento de sinais chamada de Raven [182]. Os cliques foram marcados temporalmente em um arquivo e salvos para comparação futura com o modelo proposto. A Tabela 4 mostra a

Data	Espécies	Captura	Tempo de Interesse	
			Total	Percentual
05/11/2009	Zc, UO	23:46:15	4:56:00	20,75%
06/11/2009	Zc, UO, Pm	24:00:00	5:01:16	20,92%
07/11/2009	Zc, UO, Pm	24:00:00	4:05:39	17,06%
08/11/2009	Zc, UO, Pm	24:00:00	7:17:38	30,39%
09/11/2009	Zc, UO	24:00:00	3:09:02	13,13%
10/11/2009	Zc, UO, Pm, Lo	24:00:00	9:52:35	41,15%
11/11/2009	Zc, UO	24:00:00	10:05:24	42,04%
13/04/2010	Zc, UO, Pm, Lo, Bb	23:47:30	3:48:28	16,00%
14/04/2010	Zc, UO	24:00:00	8:56:15	37,24%
15/04/2010	Zc, UO	24:00:00	5:25:00	22,57%
16/04/2010	Zc, UO, Bb	24:00:00	6:00:00	25,00%
17/04/2010	Zc, UO, Pm	24:00:00	5:41:01	23,68%
18/04/2010	Zc, UO, Pm, Lo	24:00:00	10:17:01	42,85%
19/04/2010	Zc, UO, Lo, Bb	23:48:44	12:50:09	53,90%
Total		335:22:29	97:25:28	29,05%

Tabela 3 – Dados do projeto SOCAL sítio R utilizados como treinamento do IMA.

contagem completa de cliques anotada por cada especialistas, bem como a união de cliques em suas detecções. A união representa o somatório de cliques únicos detectados pelos especialistas.

Local	Data	Hora	Especialista		União
			1º	2º	
CINMS B	26/11/2011	05:42:30	8.559	13.179	13.198
DCPP A	13/11/2012	01:59:10	302	1.318	1.318
SOCAL E	20/04/2009	00:03:45	135	111	135

Tabela 4 – Dados de análise por especialistas.

Um ponto importante na detecção manual realizada pelos especialistas e apresentado na Tabela 4 é a existência de uma grande variação na quantidade de detecções manuais, provavelmente relacionado com a experiência de cada especialista.

Os resultados comparativos das detecções indicadas pelos especialistas em relação

ao modelo proposto são apresentadas na Tabela 5. No primeiro local (CINMS B), do total de 23.437 segmentos, 17.084 (72,89%) destes foram detectados como de possível interesse com alta probabilidade de ser um clique. No segundo local (DCPP A), do total de 37.500 segmentos gerados, somente 4.048 (10,79%) segmentos foram considerados de interesse pelo algoritmo de detecção. No último local (SOCAL E), do total de 23.437 segmentos, somente 748 (3,19%) foram considerados de interesse. A alta variação percentual em cada um dos casos apresentados já era prevista por causa das características escolhidas em cada minuto. Todas as anotações temporais apresentadas pelos especialistas foram detectadas como segmentos pela etapa de detecção, ou seja, nenhuma informação relevante anotada por especialistas foi perdida. A coluna **Complemento** representa o complemento entre as detecções do modelo proposto com a união das detecções de especialistas.

Local	União	Modelo Proposto		Complemento
		Segmentos	Detecções	Detecções \ União
CINMS B	13.198	23.437	17.084	0
DCPP A	1.318	37.500	4.048	0
SOCAL E	135	23.437	748	0

Tabela 5 – Dados de análise por especialistas comparado com o modelo proposto.

Detectores automáticos podem cometer dois tipos de erros, sendo eles os falso negativos (perder um sinal de interesse), e os falso positivos (detectar incorretamente um sinal que não é de origem biológica). Para as detecções apresentadas anteriormente, o objetivo principal do treinamento é não perder nenhum sinal de interesse. Quando consideramos as próximas etapas do Fluxo de Trabalho, qualquer sinal detectado incorretamente pode ser validado novamente. O importante é conseguir reduzir o volume de dados de forma a facilitar a execução das etapas futuras.

Uma forma comum de avaliar a performance de um detector é chamado de matriz de confusão [11]. Uma matriz de confusão mostra a quantidade de verdadeiro-positivos (sinais classificados corretamente), verdadeiro-negativos (descartes corretos), falso-positivos e falso-negativos. Portanto, é necessário anotações detalhadas sobre a presença ou não de um sinal de interesse, normalmente sendo utilizado rotulação manual comparada com a automática [25].

A Tabela 6 mostra a matriz de confusão gerada a partir do treinamento realizado utilizando a captura do projeto SOCAL sítio R. Para tal, o conjunto de dados utilizado foi dividido em duas partes, uma para um novo treinamento e outra para teste. Do total de 5.600 amostras utilizadas anteriormente para o treinamento do IMA, 60% foi separado para o novo treinamento que será utilizado pela matriz de confusão, e os outros 40% utilizados como teste para validação do treinamento.

		Predição	
		Sim	Não
Rótulo	Sim	1.120	0
	Não	173	947

Tabela 6 – Matriz de confusão criada a partir dos dados de treinamento.

Como podemos ver, todas as 1.120 amostras de treinamento manualmente rotuladas como sinal de interesse foram preditas como tal. Em contrapartida, 173 amostras foram rotuladas incorretamente como de interesse, mas que deveriam ser rotuladas como sem interesse. A partir dessas informações, podemos avaliar a acurácia do treinamento como de 92,27%.

Para melhor perceber o desempenho do modelo proposto foi realizado posteriormente um teste utilizando a técnica de *10-fold cross-validation*. Nessa técnica, as 5.600 amostras foram divididas aleatoriamente em 10 partes de tamanho iguais. Do total de 10 partes, uma foi utilizada para o grupo de teste e as 9 partes restantes utilizadas para o grupo de treinamento. Em seguida, foi utilizada uma outra parte (diferente da primeira parte já utilizada) para o grupo de teste e as 9 demais partes para o grupo de treinamento. Esse processo é repetido até que todas as 10 partes sejam escolhidas para fazer parte do grupo de teste. A Fig. 23 exemplifica esse procedimento.

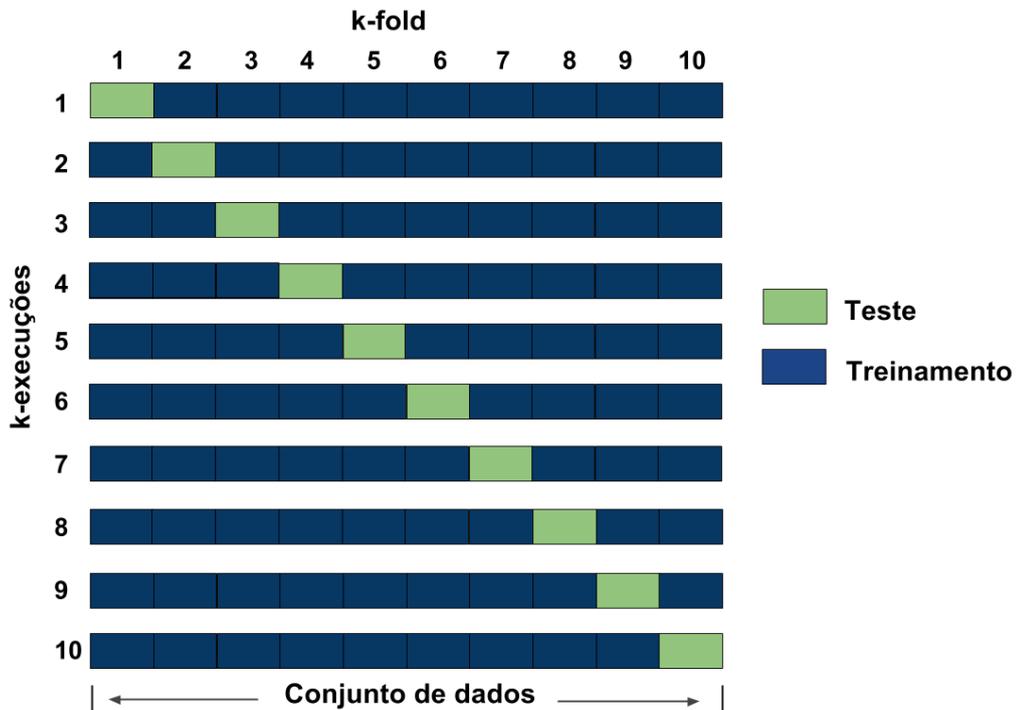


Figura 23 – Exemplo de um *10-fold*.

A Tabela 7 mostra a capacidade do modelo em prever o estado/fenômeno. Nela são apresentados os valores de especificidade, de sensibilidade e de acurácia de cada divisão com 560 amostras. A especificidade é a proporção de verdadeiros negativos, ou seja, a capacidade do sistema em prever corretamente a ausência da condição para casos que realmente não a têm. A sensibilidade é a proporção de verdadeiros positivos, indicando a capacidade do sistema em prever corretamente a condição para casos que realmente a têm. A acurácia é a proporção de predições corretas, sem levar em consideração o que é positivo e o que é negativo. Uma acurácia alta mostra que o modelo erra pouco.

Divisão	Especificidade (%)	Sensibilidade (%)	Acurácia (%)
1	94,32	100	97,14
2	92,40	100	96,60
3	97,80	100	98,57
4	70,71	100	92,67
5	88,39	100	95,35
6	97,44	100	98,21
7	78,57	100	95,71
8	96,42	100	98,21
9	94,58	100	96,96
10	99,09	100	99,28
Média	93,75	100	96,87

Tabela 7 – Avaliação dos modelos nos 10-fold do teste.

Como visto na Tabela 7, a sensibilidade alcançou 100% em todas as 10 divisões, o que mostra a capacidade do modelo em prever corretamente, uma vez que é preferível obter resultados falso-positivos do que falso-negativos, ou seja, após passar pelo modelo nenhuma informação relevante foi perdida.

A Tabela 8 apresenta os valores da média, mediana, desvio padrão, máximo e mínimo da acurácia para os dados dos grupos de teste. A mediana das acurácias foi de 97,05%, o que reflete a precisão do método proposto nesta tese. E o desvio padrão obtido foi de 1,93%, o que mostra a robustez do modelo.

Média	Mediana	Desvio Padrão	Máximo	Mínimo
96,87	97,05	1,93	99,28	92,67

Tabela 8 – Valores percentuais da acurácia para os 10 grupos de teste.

A Tabela 9 mostra uma visão geral dos resultados de execução do algoritmo proposto no conjunto de dados completo apresentado anteriormente. A primeira coluna mostra o local em que os dados foram capturados, utilizando em conjunto o projeto com o sítio da captura, assim como apresentado pela Fig. 22. A segunda coluna **Captura** indica o tempo total de sinais armazenados para aquele determinado local. A coluna **Tempo de Interesse**, dividida em **Total** e **Percentual**, indicam respectivamente, o tempo total considerado de interesse a partir das anotações de especialistas e o percentual total do tempo de interesse em relação ao tempo total de captura. A última coluna intitulada de **Detecções**, também dividida em outras duas colunas, a de **Segmentos** que indica a quantidade de segmentos que foram detectados dentro do tempo total de captura como clique, e seu respectivo **Percentual** em relação ao tempo total de captura. Para resultados mais detalhados por projeto considerando os dias de captura, veja o Apêndice A.

Local	Captura	Tempo de Interesse		Detecções	
		Total	Percentual	Segmentos	Percentual
CINMS B	334:59:58	54:07:26	16,16%	25.998.153	5,52%
CINMS C	311:21:57	28:20:56	9,10%	16.797.499	2,40%
DCPP A	311:36:03	22:21:14	7,17%	15.544.048	2,22%
DCPP B	287:28:10	17:44:55	6,17%	16.263.623	2,51%
DCPP C	310:39:58	79:46:17	25,68%	19.156.435	4,38%
SOCAL E	287:29:59	84:45:00	29,48%	17.017.159	4,21%
SOCAL R	335:22:29	97:25:28	29,05%	22.971.890	4,87%
Total	2178:58:34	384:31:06	17,64%	133.748.807	3,48%

Tabela 9 – Redução de dados por projeto e sítio.

Como podemos ver na Tabela 9, do total de aproximadamente 2.178 horas, somente 384 horas foram consideradas de interesse pelas anotações de especialistas, concretizando um total de apenas 17,64% do volume de dados. É importante ressaltar que, do total de dados de interesse quando considerado a nível de segmentos com cliques, esse valor pode ser reduzido bem mais. É justamente essa redução que o algoritmo proposto para detecção tenta alcançar, mostrando que do total de segmentos, somente 3,48% possui algum vestígio que possa conter cliques, possibilitando um possível interesse de identificação e análise.

6 CONCLUSÃO

O Monitoramento Acústico Passivo vêm se tornando cada vez mais essencial para pesquisas efetivas sobre Cetáceos. Com a utilização de sistemas computacionais, o PAM têm o potencial de prover um aparato operacional para os pesquisadores, otimizando além da utilização dos equipamentos, o tempo gasto com análises dos dados.

No caso de Odontocetos, a utilização de PAM se torna mais complexa, graças as características dos sinais emitidos por estes animais, que necessitam de uma atividade computacional mais intensa quando comparados às análises aplicadas a sinais de Mysticetos. A principal característica que encarece o processo é a alta frequência necessária para captura do sinais emitidos, que gera um grande volume de dados. Este grande volume de dados é prejudicial quando todos os dados devem ser salvos em discos rígidos instalados em dispositivos fixos autônomos, ou quando processados por algoritmos complexos [5, 13, 18], impossibilitando a detecção, distribuição e identificação de sinais *online*.

Alinhado aos objetivos apresentados por este trabalho, podemos citar o interesse de entidades públicas e privadas em maneiras robustas, com bom custo-benefício para coleta de dados de mamíferos marinhos, tais como sua distribuição, abundância, movimento e comportamento. Como é o caso do *Joint Industry Programme* (JIP), que considera a importância dessas informações para exploração e produção de gás e óleo [11].

Com intuito de reduzir o impacto do crescimento de volume de dados em um modelo de detecção atual, um Fluxo de Trabalho foi proposto nesta tese, focando principalmente na aplicação em dispositivos fixos autônomos. Com a utilização de um Fluxo de Trabalho padronizado, que considera diversas possibilidades do ambiente, bem como avanços e possibilidades computacionais, uma construção com componentes específicos para determinadas atividades é alcançado, otimizando todo o processo de planejamento e execução.

Quando um Fluxo de Trabalho é criado baseado em componentes de software, todos os algoritmos propostos podem ser utilizados de maneira similar ao seu uso original. Ou seja, mesmo que nem todas as etapas do Fluxo de Trabalho proposto estejam totalmente funcionais, abordagens comumente utilizadas na literatura podem ser alocadas na complementação do fluxo.

O Fluxo de Trabalho proposto é fortemente baseado em processos visando a redução de dados, como atestado em sua primeira etapa, ressaltando-se que este enfoque é raramente encontrado na literatura como objetivo principal. A importância da redução dos dados capturados é acentuada quando leva-se em consideração que os sons capturados por dispositivos fixos autônomos são na maioria das vezes processos estacionários, onde cerca de 70-90% dos dados são ruídos naturais do ambiente [16, 130].

Dessa maneira, foi proposto e desenvolvido um algoritmo baseado em dados simbólicos e seus padrões de representatividade que destaca-se pela sua simplicidade, desde a sua criação até execução, sendo capaz de reduzir o volume de dados a ser processado consideravelmente e sem perda de dados relevantes. O componente do Fluxo de Trabalho relacionado à redução dos dados apresentado passou por diversos refinamentos, com a versão final apresentada sendo capaz de reduzir em média 96,52% do volume total de dados a serem processados.

O modelo que viabilizou a redução foi baseado em um procedimento de Aprendizado de Máquina que evita a necessidade de definição prévia de um valor de referência que delimite se o sinal analisado é de interesse ou não. Um classificador baseado em margem *online* conhecido como IMA foi adaptado para discriminação dos sinais. O valor da margem mostrou-se efetivo como indicativo de referência para determinação dos sinais de interesse. Além disso, procedimentos adaptativos de correção da margem permite um melhor ajuste do indutor perante a variabilidade presente no ambiente. O treinamento realizado com uma pequena parte dos dados rotulados manualmente conseguiu, por meio da hipótese de indução gerada, a detecção de todos os cliques da captura, com a inclusão de uma pequena taxa de falso positivo.

Os dados utilizados nos experimentos estão disponibilizados pelo *Detection, Classification, Localization, and Density Estimation (DCLDE) Workshop 2015*. Para confirmar que o método não descarta informações relevantes, 3 minutos do sinal foram escolhidos e validados manualmente por especialistas a procura de cliques. Todos os cliques encontrados manualmente por especialistas foram encontrados na etapa de detecção.

Dada a redução substancial no volume total de dados, possibilidades emergentes foram vislumbradas com técnicas totalmente distintas das utilizadas atualmente para identificação de Odontocetos. Técnicas de mineração de dados vem ganhando espaço e sendo cada vez mais utilizadas por grandes empresas no meio corporativo, mas ainda não são exploradas para detecção e identificação de Odontocetos.

Apesar de ter apresentado excelentes resultados na etapa de detecção, ainda existe bastante coisa a ser avaliada. Como investigação futura, pretende-se:

- Realizar o estudo e implementação em uma arquitetura de hardware simplificada que considera as necessidades computacionais do Fluxo de Trabalho proposto;
- Avaliar a aplicação do algoritmo proposto para etapa de detecção em sinais capturados através de outros dispositivos e metodologias, em especial naqueles com incidência de ruídos elevada. Testes preliminares [150] mostraram bons resultados em dados capturados por arquiteturas de arrasto, mas uma avaliação pontual dos cliques ainda é necessária;

- Estudar uma melhor parametrização para do classificador baseado em margem adaptativa utilizado, bem como determinar um modelo de adaptação da margem que se ajuste à dinâmica de ambientes marinhos, usufruindo assim de todo seu potencial *online*;
- Avaliar a capacidade de identificação de algoritmos de similaridade para espécies distintas. O LSH mostra grande potencial nesse aspecto, sendo possível a compactação e distribuição *online* das informações identificadas, mas ainda é necessário um estudo maior avaliando as reais capacidades do LSH para identificação de diferentes espécies simultaneamente.

REFERÊNCIAS

- [1] Patrick JO Miller, Nicoletta Biassoni, Amy Samuels, and Peter L Tyack. Whale songs lengthen in response to sonar. *Nature*, 405(6789):903–903, 2000.
- [2] Mark F Baumgartner and Sarah E Mussoline. A generalized baleen whale call detection and classification system. *The Journal of the Acoustical Society of America*, 129(5):2889–2902, 2011.
- [3] Shannon Rankin, Julie N Oswald, and Jay Barlow. Acoustic behavior of dolphins in the pacific ocean: implications for using passive acoustic methods for population studies. *Canadian Acoustics*, 36(1):88–92, 2008.
- [4] David K Mellinger, Kathleen M Stafford, S Moore, Robert P Dziak, and Haru Matsumoto. Fixed passive acoustic observation methods for cetaceans. *Oceanography*, 20(4):36, 2007.
- [5] Tina M Yack, Jay Barlow, Marie A Roch, Holger Klinck, Steve Martin, David K Mellinger, and Douglas Gillespie. Comparison of beaked whale detection algorithms. *Applied Acoustics*, 71(11):1043–1049, 2010.
- [6] Gregory S Schorr, Erin A Falcone, David J Moretti, and Russel D Andrews. First long-term behavioral records from cuvier’s beaked whales (*ziphius cavirostris*) reveal record-breaking dives. *PLoS One*, 9(3):e92633, 2014.
- [7] Peter L Tyack, Mark Johnson, Natacha Aguilar Soto, Albert Sturlese, and Peter T Madsen. Extreme diving of beaked whales. *Journal of Experimental Biology*, 209(21):4238–4253, 2006.
- [8] John Hildebrand. Impacts of anthropogenic sound on cetaceans. *Unpublished paper submitted to the International Whaling Commission Scientific Committee SC/56 E*, 13, 2004.
- [9] Tara M Cox, TJ Ragen, AJ Read, E Vos, RW Baird, K Balcomb, J Barlow, J Caldwell, T Cranford, and L Crum. Understanding the impacts of anthropogenic sound on beaked whales. Technical report, DTIC Document, 2006.
- [10] PT Madsen. Marine mammals and noise: Problems with root mean square sound pressure levels for transients. *The Journal of the Acoustical Society of America*, 117(6):3952–3957, 2005.
- [11] Thomas F Norris, Julie N Oswald, and Renata S Sousa-Lima. A review and inventory of fixed installation passive acoustic monitoring methods and technologies. Final report to the International Association of Oil and Gas Producers Joint Industry Programme on E&P Sound and Marine Life, 2010. Contract reference # JIP2208-03. 214pp.
- [12] National Oceanic and Atmospheric Administration. Southwest Fisheries Science Center. <https://swfsc.noaa.gov>. [Online; acessado 5 de junho de 2014].
- [13] Tyler J Olmstead, Marie A Roch, Paul Hursky, Michael B Porter, Holger Klinck, David K Mellinger, Tyler Helble, Sean S Wiggins, LD Gerald, and John A Hildebrand.

- Autonomous underwater glider based embedded real-time marine mammal detection and classification. *The Journal of the Acoustical Society of America*, 127(3):1971–1971, 2010.
- [14] Holger Klinck and David K Mellinger. The energy ratio mapping algorithm: A tool to improve the energy-based detection of odontocete echolocation clicks. *The Journal of the Acoustical Society of America*, 129(4):1807–1812, 2011.
- [15] Julie N Oswald, Shannon Rankin, and Jay Barlow. The effect of recording and analysis bandwidth on acoustic identification of delphinid species. *The Journal of the Acoustical Society of America*, 116(5):3178–3185, 2004.
- [16] Sean M Wiggins and John A Hildebrand. High-frequency acoustic recording package (harp) for broad-band, long-term marine mammal monitoring. In *Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies*, pages 551–557. IEEE, 2007.
- [17] D Mellinger and J Barlow. *FUTURE DIRECTIONS FOR ACOUSTIC MARINE MAMMAL SURVEYS: STOCK ASSESSMENT AND HABITAT USE*. Pacific Marine Environment Laboratory (US), 2003.
- [18] Christopher W Clark, Thomas Calupca, Douglas Gillespie, Keith Von der Heydt, and John Kemp. A near-real-time acoustic detection and reporting system for endangered species in critical habitats. *The Journal of the Acoustical Society of America*, 117(4):2525–2525, 2005.
- [19] Suchita Nanaware, Rajveer Shastri, Yash Joshi, and Aruneema Das. Passive acoustic detection and classification of marine mammal vocalizations. In *Communications and Signal Processing (ICCSP), 2014 International Conference on*, pages 493–497. IEEE, 2014.
- [20] Michael Bittle and Alec Duncan. A review of current marine mammal detection and classification algorithms for use in automated passive acoustic monitoring. In *Proceedings of Acoustics*, volume 2013, 2013.
- [21] David K Mellinger and Christopher W Clark. Recognizing transient low-frequency whale sounds by spectrogram correlation. *The Journal of the Acoustical Society of America*, 107(6):3518–3529, 2000.
- [22] Renata S Sousa-Lima, Thomas F Norris, Julie N Oswald, and Deborah P Fernandes. A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals. *Aquatic Mammals*, 39(1):23–53, 2013.
- [23] Jacques Curie and Pierre Curie. Développement, par pression, de l’électricité polaire dans les cristaux hémihédres à faces inclinées. *Comptes rendus*, 91:294–295, 1880.
- [24] David L Bradley and Richard Stern. Underwater sound and the marine mammal acoustic environment: A guide to fundamental principles. *US Marine Mammal Commission*. (12 December 2011, 2008.
- [25] Elizabeth T Küsel, David K Mellinger, Len Thomas, Tiago A Marques, David Moretti, and Jessica Ward. Cetacean population density estimation from single fixed sensors using passive acoustics. *The Journal of the Acoustical Society of America*, 129(6):3610–3622, 2011.

- [26] Wikimedia commos. Humpback Whale. https://commons.wikimedia.org/wiki/File:Humpback_Whale_underwater_shot.jpg. [Online; acessado 1 de dezembro de 2015].
- [27] Wikimedia commos. Spinner Dolphin. https://commons.wikimedia.org/wiki/File:A_spinner_dolphin_in_the_Red_Sea.jpg. [Online; acessado 1 de dezembro de 2015].
- [28] Bia Hetzel and Liliane Lodi. *Baleias, botos e golfinhos: guia de identificação para o Brasil*. Editora Nova Fronteira, 1993.
- [29] Hal Whitehead. *Sperm whales: social evolution in the ocean*. University of Chicago Press, 2003.
- [30] Marcos Oliveira. Ecolocação. http://www.icnf.pt/portal/naturaclas/gestao-biodiv/roazes-do-sado/resource/img/ecolocaliz_1. [Online; acessado 11 de maio de 2016].
- [31] Renata S Sousa-Lima, Deborah P Fernandes, Thomas F Norris, and Julie N Oswald. A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals: 2013 state-of-the-industry. In *Acoustics in Underwater Geosciences Symposium (RIO Acoustics), 2013 IEEE/OES*, pages 1–9. IEEE, 2013.
- [32] Scripps Whale Acoustic Lab. HARP Packaging Configurations. http://cetuss.ucsd.edu/technologies_Mooring_LargeTraditional.html. [Online; acessado 1 de março de 2016].
- [33] David K Mellinger, Sharon L Nieukirk, Haru Matsumoto, Sara L Heimlich, Robert P Dziak, Joe Haxel, Matt Fowler, Christian Meinig, and Hendrick V Miller. Seasonal occurrence of north atlantic right whale (*eubalaena glacialis*) vocalizations at two sites on the scotian shelf. *Marine Mammal Science*, 23(4):856–867, 2007.
- [34] KM Stafford, SE Moore, M Spillane, and S Wiggins. Gray whale calls recorded near barrow, alaska, throughout the winter of 2003-04. *Arctic*, pages 167–172, 2007.
- [35] Ana Širović, John A Hildebrand, Sean M Wiggins, and Deborah Thiele. Blue and fin whale acoustic presence around antarctica during 2003 and 2004. *Marine Mammal Science*, 25(1):125–136, 2009.
- [36] Renata S Sousa-Lima and Christopher W Clark. Whale sound recording technology as a tool for assessing the effects of boat noise in a brazilian marine park. *Park Science*, 26(1):59–63, 2009.
- [37] David K Mellinger, Kathleen M Stafford, and Christopher G Fox. Seasonal occurrence of sperm whale (*physeter macrocephalus*) sounds in the gulf of alaska, 1999–2001. *Marine Mammal Science*, 20(1):48–62, 2004.
- [38] David K Mellinger, Kathleen M Stafford, SUEE MOORE, and Lisa Munger. Detection of north pacific right whale (*eubalaena japonica*) calls in the gulf of alaska. *Marine Mammal Science*, 20(4):872–879, 2004.
- [39] Rodney A Rountree, R Grant Gilmore, Clifford A Goudey, Anthony D Hawkins, Joseph J Luczkovich, and David A Mann. Listening to fish: applications of passive acoustics to fisheries science. *Fisheries*, 31(9):433–446, 2006.

- [40] Lisa M Munger, David K Mellinger, Sean M Wiggins, Sue E Moore, and John A Hildebrand. Performance of spectrogram cross-correlation in detecting right whale calls in long-term recordings from the bering sea. *Canadian Acoustics*, 33(2):25–34, 2005.
- [41] Sean Wiggins. Autonomous acoustic recording packages (arps) for long-term monitoring of whale sounds. *Marine Technology Society Journal*, 37(2):13–22, 2003.
- [42] Susanna B Blackwell and Charles R Greene Jr. Sounds from an oil production island in the beaufort sea in summer: characteristics and contribution of vessels. *The Journal of the Acoustical Society of America*, 119(1):182–196, 2006.
- [43] Mark A McDonald, John A Hildebrand, and Sean M Wiggins. Increases in deep ocean ambient noise in the northeast pacific west of san nicolas island, california. *The Journal of the Acoustical Society of America*, 120(2):711–718, 2006.
- [44] Christopher W Clark, William T Ellison, Brandon L Southall, Leila Hatch, Sofie M Van Parijs, Adam Frankel, and Dimitri Ponirakis. Acoustic masking in marine ecosystems: intuitions, analysis, and implication. *Marine Ecology Progress Series*, 395:201–222, 2009.
- [45] Teong Beng Koay, John R Potter, Torbjorn Johansson, and Pallayil Venugopalan. Panda; a self-recovering shallow water acoustic logger. In *OCEANS, 2001. MTS/IEEE Conference and Exhibition*, volume 3, pages 1879–1885. IEEE, 2001.
- [46] TB Koay, JR Potter, V Pallayil, and ET Tan. Panda: A rapidly deployable, self-recovering shallow water acquisition platform. *MTS Sea Technology Magazine*, 2002.
- [47] Faisal Hamady, Ayman Kayssi, Ali Chehab, and Mohammad Mansour. Evaluation of low-power computing when operating on subsets of multicore processors. *Journal of Signal Processing Systems*, 70(2):193–208, 2013.
- [48] ACPI. Advanced Configuration & Power Interface. <http://www.acpi.info>. [Online; acessado 26 de julho de 2014].
- [49] Wil Van Der Aalst and Kees Max Van Hee. *Workflow management: models, methods, and systems*. MIT press, 2004.
- [50] Wil MP van Der Aalst, Arthur HM Ter Hofstede, Bartek Kiepuszewski, and Alistair P Barros. Workflow patterns. *Distributed and parallel databases*, 14(1):5–51, 2003.
- [51] Stephen A White. Process modeling notations and workflow patterns. *Workflow handbook*, 2004:265–294, 2004.
- [52] Inc. Object Management Group. Unified Modeling Language (UML). <http://www.omg.org/spec/UML/>. [Online; acessado 30 de setembro de 2015].
- [53] Inc. Object Management Group. Business Process Model And Notation (BPMN) Version 2.0. <http://www.omg.org/spec/BPMN/2.0/>, 2011. [Online; acessado 29 de setembro de 2015].
- [54] Inc. Object Management Group. Object Management Group. <http://www.omg.org/>. [Online; acessado 30 de setembro de 2015].

- [55] Ewa Deelman, Dennis Gannon, Matthew Shields, and Ian Taylor. Workflows and e-science: An overview of workflow system features and capabilities. *Future Generation Computer Systems*, 25(5):528–540, 2009.
- [56] Amanda Mattos, F Silva, Nicolaas Ruberg, and M Cruz. Gerência de workflows científicos: uma análise crítica no contexto da bioinformática. *COPPE/UFRJ*, 2008.
- [57] Guilherme Horta Travassos and Márcio O Barros. Contributions of in virtuo and in silico experiments for the future of empirical studies in software engineering. In *2nd Workshop on Empirical Software Engineering the Future of Empirical Studies in Software Engineering*, pages 117–130, 2003.
- [58] Jonas Dias. *Exécution interactive pour expériences computationnelles à grande échelle*. PhD thesis, Universidade Federal de Rio de Janeiro, 2013.
- [59] Juliana Freire, David Koop, Emanuele Santos, and Cláudio T Silva. Provenance for computational tasks: A survey. *Computing in Science & Engineering*, 10(3):11–21, 2008.
- [60] Juliana Freire, Cláudio T Silva, Steven P Callahan, Emanuele Santos, Carlos E Scheidegger, and Huy T Vo. Managing rapidly-evolving scientific workflows. In *Provenance and Annotation of Data*, pages 10–18. Springer, 2006.
- [61] Bertram Ludäscher, Ilkay Altintas, Chad Berkley, Dan Higgins, Efrat Jaeger, Matthew B Jones, Edward A Lee, Jing Tao, and Yang Zhao. Scientific workflow management and the kepler system. *Concurrency and Computation: Practice and Experience*, 18(10):1039–1065, 2006.
- [62] Katherine Wolstencroft, Robert Haines, Donal Fellows, Alan Williams, David Withers, Stuart Owen, Stian Soiland-Reyes, Ian Dunlop, Aleksandra Nenadic, Paul Fisher, et al. The taverna workflow suite: designing and executing workflows of web services on the desktop, web or in the cloud. *Nucleic acids research*, page gkt328, 2013.
- [63] G Bruce Berriman, Ewa Deelman, John Good, Joseph C Jacob, Daniel S Katz, Anastasia C Laity, Thomas A Prince, Gurmeet Singh, and Mei-Hui Su. Generating complex astronomy workflows. In *Workflows for e-Science*, pages 19–38. Springer, 2007.
- [64] Alberto MR Davila, Pablo N Mendes, Glauber Wagner, Diogo A Tschoeke, Rafael RC Cuadrat, Felipe Liberman, Luciana Matos, Thiago Satake, Kary ACS Ocaña, Omar Triana, et al. Protozoadb: dynamic visualization and exploration of protozoan genomes. *Nucleic acids research*, 36(suppl 1):D547–D552, 2008.
- [65] Gabriel Guerra, Fernando A Rochinha, Renato Elias, Daniel De Oliveira, Eduardo Ogasawara, Jonas Furtado Dias, Marta Mattoso, and Alvaro LGA Coutinho. Uncertainty quantification in computational predictive models for fluid dynamics using a workflow management engine. *International Journal for Uncertainty Quantification*, 2(1), 2012.
- [66] Kary ACS Ocaña, Daniel de Oliveira, Eduardo Ogasawara, Alberto MR Dávila, Alexandre AB Lima, and Marta Mattoso. Sciphy: A cloud-based workflow for phylogenetic analysis of drug targets in protozoan genomes. In *Advances in Bioinformatics and Computational Biology*, pages 66–70. Springer, 2011.

- [67] Yogesh L Simmhan, Beth Plale, and Dennis Gannon. A survey of data provenance in e-science. *ACM Sigmod Record*, 34(3):31–36, 2005.
- [68] Samaran Flore, Olivier Adam, Jean-François Motsch, and Christophe Guinet. Definition of the antarctic and pygmy blue whale call templates. application to fast automatic detection. *Canadian Acoustics*, 36(1):93–103, 2008.
- [69] Kotaro Ichikawa, Chika Tsutsumi, Nobuaki Arai, Tomonari Akamatsu, Tomio Shinke, Takeshi Hara, and Kanjana Adulyanukosol. Dugong (dugong dugon) vocalization patterns recorded by automatic underwater sound monitoring systems. *The Journal of the Acoustical Society of America*, 119(6):3726–3733, 2006.
- [70] Susan Jarvis, Nancy DiMarzio, Ronald Morrissey, and David Morretti. *Automated classification of beaked whales and other small odontocetes in the tongue of the ocean, bahamas*. IEEE, 2006.
- [71] David K Mellinger. A neural network for classifying clicks of blainville’s beaked whales (*mesoplodon densirostris*). *Canadian Acoustics*, 36(1):55–59, 2008.
- [72] Tamaki Ura, Rajendar Bahl, Masao Sakata, Junichi Kojima, Tetsuo Fukuchi, Junya Ura, Yoshiaki Nose, Harumi Sugimatsu, Kyoichi Mori, Takeshi Nakatani, et al. Acoustic tracking of sperm whales using two sets of hydrophone array. In *Underwater Technology, 2004. UT’04. 2004 International Symposium on*, pages 103–107. IEEE, 2004.
- [73] DK Mellinger and JW Bradbury. Acoustic measurement of marine mammal sounds in noisy environments. In *Proceedings of the International Conference on Underwater Acoustical Measurements: Technologies and Results*, pages 273–280, 2007.
- [74] Ofer Tchernichovski, Fernando Nottebohm, Ching Elizabeth Ho, Bijan Pesaran, and Partha Pratim Mitra. A procedure for an automated measurement of song similarity. *Animal Behaviour*, 59(6):1167–1176, 2000.
- [75] Dan Valente, Haibin Wang, Peter Andrews, Partha P Mitra, Sigal Saar, Ofer Tchernichovski, Ilan Golani, and Yoav Benjamini. Characterizing animal behavior through audio and video signal processing. *MultiMedia, IEEE*, 14(4):32–41, 2007.
- [76] Odile Gerard, Craig Carthel, Stefano Coraluppi, and Peter Willett. Feature-aided tracking for marine mammal detection and classification. *Canadian Acoustics*, 36(1):13–19, 2008.
- [77] SP Van IJsselmuide and SP Beerens. Detection and classification of marine mammals using an lfas system. *Canadian Acoustics*, 32(2):93–106, 2004.
- [78] Brian R La Cour and Michael A Linford. Detection and classification of north atlantic right whales in the bay of fundy using independent component analysis. *Canadian Acoustics*, 32(2):48–54, 2004.
- [79] Maciej Lopatka, Olivier Adam, Cvstophe Laplanche, Jean-François Motsch, and Jan Zarzycki. New effective analytic representation based on the time-varying schur coefficients for underwater signals analysis. In *Oceans 2005-Europe*, volume 1, pages 160–165. IEEE, 2005.

- [80] Maciej Lopatka, Olivier Adam, Jean-François Motsch, and Jan Zarzycki. Attractive time-variant orthogonal schur-like representation for click-type signal recognition. *Canadian Acoustics*, 36(1):81–87, 2008.
- [81] Cornel Ioana, Andre Quinquis, and Yann Stephan. Feature extraction from underwater signals using time-frequency warping operators. *Oceanic Engineering, IEEE Journal of*, 31(3):628–645, 2006.
- [82] Christine Erbe and Andrew R King. Automatic detection of marine mammals using information entropy. *The Journal of the Acoustical Society of America*, 124(5):2833–2840, 2008.
- [83] C Sidney Burrus, Ramesh A Gopinath, and Haitao Guo. *Introduction to wavelets and wavelet transforms*. Prentice hall, 1997.
- [84] Denis Chabot. A quantitative technique to compare and classify humpback whale (*megaptera novaeangliae*) sounds. *Ethology*, 77(2):89–102, 1988.
- [85] Christopher W Clark. The acoustic repertoire of the southern right whale, a quantitative analysis. *Animal Behaviour*, 30(4):1060–1071, 1982.
- [86] Kurt M Fristrup and William A Watkins. Marine animal sound classification. Technical report, Woods Hole Oceanographic Institution, 1993.
- [87] T Scott Brandes. Feature vector selection and use with hidden markov models to identify frequency-modulated bioacoustic signals amidst noise. *Audio, Speech, and Language Processing, IEEE Transactions on*, 16(6):1173–1180, 2008.
- [88] SK Madhusudhana, Erin M Oleson, Melissa S Soldevilla, Marie Roch, John Hildebrand, et al. Frequency based algorithm for robust contour extraction of blue whale b and d calls. In *OCEANS 2008-MTS/IEEE Kobe Techno-Ocean*, pages 1–8. IEEE, 2008.
- [89] Julie N Oswald, Shannon Rankin, Jay Barlow, and Marc O Lammers. A tool for real-time acoustic species identification of delphinid whistles. *The Journal of the Acoustical Society of America*, 122(1):587–595, 2007.
- [90] Justin Matthews. Detection of frequency-modulated calls using a chirp model. *Canadian Acoustics*, 32(2):66–75, 2004.
- [91] Benoit Leprettre and Nadine Martin. Extraction of pertinent subsets from time-frequency representations for detection and recognition purposes. *Signal Processing*, 82(2):229–238, 2002.
- [92] PR White and ML Hadley. Introduction to particle filters for tracking applications in the passive acoustic monitoring of cetaceans. *Canadian Acoustics*, 36(1):146–152, 2008.
- [93] Patrick J Clemins and Michael T Johnson. Generalized perceptual linear prediction features for animal vocalization analysis. *The Journal of the Acoustical Society of America*, 120(1):527–534, 2006.
- [94] Xanadu C Halkias and Daniel PW Ellis. Call detection and extraction using bayesian inference. *Applied Acoustics*, 67(11):1164–1174, 2006.

- [95] Jack Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418, 1969.
- [96] Marie A Roch, Melissa S Soldevilla, Jessica C Burtenshaw, E Elizabeth Henderson, and John A Hildebrand. Gaussian mixture model classification of odontocetes in the southern california bight and the gulf of california. *The Journal of the Acoustical Society of America*, 121(3):1737–1748, 2007.
- [97] Hynek Hermansky. Perceptual linear predictive (plp) analysis of speech. *the Journal of the Acoustical Society of America*, 87(4):1738–1752, 1990.
- [98] John Makhoul. Spectral linear prediction: properties and applications. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 23(3):283–296, 1975.
- [99] WWL Au and PE Nachtigall. Artificial neural network modeling of dolphin echolocation, 1995.
- [100] HL Roitblat, PWB Moore, PE Nachtigall, RH Penner, and WWL Au. Natural echolocation with an artificial neural network. *Int. J. Neural Networks*, 1:239–248, 1989.
- [101] Nello Cristianini and John Shawe-Taylor. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.
- [102] Suleman Mazhar, Tamaki Ura, and Rajendar Bahl. Vocalization based individual classification of humpback whales using support vector machine. In *OCEANS 2007*, pages 1–9. IEEE, 2007.
- [103] Abdelmonem Affi, Susanne May, and Virginia A Clark. *Computer-aided multivariate analysis*. CRC Press, 2003.
- [104] Ronald A Fisher. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2):179–188, 1936.
- [105] Xuedong Huang, Alex Acero, Hsiao-Wuen Hon, et al. Spoken language processing, 2001.
- [106] Melissa S Soldevilla, E Elizabeth Henderson, Gregory S Campbell, Sean M Wiggins, John A Hildebrand, and Marie A Roch. Classification of risso’s and pacific white-sided dolphins using spectral properties of echolocation clicks. *The Journal of the Acoustical Society of America*, 124(1):609–624, 2008.
- [107] Alexander Waibel, Toshiyuki Hanazawa, Geoffrey Hinton, Kiyohiro Shikano, and Kevin J Lang. Phoneme recognition using time-delay neural networks. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 37(3):328–339, 1989.
- [108] Nikolai S. Belliustin, Sergei O. Kuznetsov, Irina V. Nuidel, and Vladimir G. Yakhno. Neural networks with close nonlocal coupling for analyzing composite image. *Neuro-computing*, 3(5):231–246, 1991.
- [109] Kunihiko Fukushima and Nobuaki Wake. Handwritten alphanumeric character recognition by the neocognitron. *Neural Networks, IEEE Transactions on*, 2(3):355–365, 1991.

- [110] Zezhen Kuang and Anthony Kuh. A combined self-organizing feature map and multi-layer perceptron for isolated word recognition. *Signal Processing, IEEE Transactions on*, 40(11):2651–2657, 1992.
- [111] M Van Der Schaar, E Delory, A Catala, and M André. Neural network-based sperm whale click classification. *Journal of the Marine Biological Association of the United Kingdom*, 87(01):35–38, 2007.
- [112] Sébastien Derégnaucourt, J-C Guyomarc’h, and V Richard. Classification of hybrid crows in quail using artificial neural networks. *Behavioural processes*, 56(2):103–112, 2001.
- [113] J Placer and CN Slobodchikoff. A fuzzy-neural system for identification of species-specific alarm calls of gunnison’s prairie dogs. *Behavioural Processes*, 52(1):1–9, 2000.
- [114] Peter-Christian Schön, Birger Puppe, and Gerhard Manteuffel. Linear prediction coding analysis and self-organizing feature map as tools to classify stress calls of domestic pigs (*sus scrofa*). *The Journal of the Acoustical Society of America*, 110(3):1425–1431, 2001.
- [115] Scott O Murray, Eduardo Mercado, and Herbert L Roitblat. The neural network classification of false killer whale (*pseudorca crassidens*) vocalizations. *The Journal of the Acoustical Society of America*, 104(6):3626–3633, 1998.
- [116] Joydeep Ghosh, Larry M Deuser, and Steven D Beck. A neural network based hybrid system for detection, characterization, and classification of short-duration oceanic signals. *Oceanic Engineering, IEEE Journal of*, 17(4):351–363, 1992.
- [117] John R Potter, David K Mellinger, and Christopher W Clark. Marine mammal call discrimination using artificial neural networks. *The Journal of the Acoustical society of America*, 96(3):1255–1262, 1994.
- [118] DS Houser, DA Helweg, and PW Moore. Classification of dolphin echolocation clicks by energy and frequency distributions. *The Journal of the Acoustical Society of America*, 106(3):1579–1585, 1999.
- [119] Jack McLaughlin, Nicolas Josso, and Cornel Ioana. Detection and classification of call types in the vocalizations of north-east pacific blue whales. *The Journal of the Acoustical Society of America*, 123(5):3102–3102, 2008.
- [120] J Luan, R Bahl, T Ura, T Akamatsu, M Yamaguchi, T Sakamaki, and K Mori. Model-based recognition of individual humpback whales from their vocalization features. In *Ocean Engineering Symposium, Tokyo*, 2003.
- [121] OGP E&P. Sound and Marine Life. <http://www.soundandmarinelife.org>. [Online; acessado 11 de julho de 2014].
- [122] Chelonia Limited. C-POD. <http://c-pod.com>. [Online; acessado 10 de julho de 2014].
- [123] Matlab. Matlab. <http://www.mathworks.com/products/matlab/>. [Online; acessado 22 de julho de 2014].

- [124] Karin Tubbert Clausen, Magnus Wahlberg, Kristian Beedholm, Stacy Deruiter, and Peter Teglberg Madsen. Click communication in harbour porpoises *phocoena phocoena*. *Bioacoustics*, 20(1):1–28, 2011.
- [125] Susan E Parks and Peter L Tyack. Sound production by north atlantic right whales (*eubalaena glacialis*) in surface active groups. *The Journal of the Acoustical Society of America*, 117(5):3297–3306, 2005.
- [126] Sirpa Nummela, JGM Thewissen, Sunil Bajpai, S Taseer Hussain, and Kishor Kumar. Eocene evolution of whale hearing. *Nature*, 430(7001):776–778, 2004.
- [127] P Madsen, M Wahlberg, and B Møhl. Male sperm whale (*physeter macrocephalus*) acoustics in a high-latitude habitat: implications for echolocation and communication. *Behavioral Ecology and Sociobiology*, 53(1):31–41, 2002.
- [128] Peter L Tyack and Christopher W Clark. Communication and acoustic behavior of dolphins and whales. In *Hearing by whales and dolphins*, pages 156–224. Springer, 2000.
- [129] Darlene R Ketten. The marine mammal ear: specializations for aquatic audition and echolocation. In *The evolutionary biology of hearing*, pages 717–750. Springer, 1992.
- [130] Ildar R Urazghildiiev and Christopher W Clark. Acoustic detection of north atlantic right whale contact calls using the generalized likelihood ratio test. *The Journal of the Acoustical Society of America*, 120(4):1956–1963, 2006.
- [131] Google Loon. Google Loon Project. <http://www.google.com/loon/>. [Online; acessado 15 de outubro de 2015].
- [132] Inc.n Iridium Communications. Iridium Networks. <https://www.iridium.com/>. [Online; acessado 15 de outubro de 2015].
- [133] Jure Leskovec, Anand Rajaraman, and Jeffrey David Ullman. *Mining of massive datasets*. Cambridge University Press, 2014.
- [134] Piotr Indyk and Rajeev Motwani. Approximate nearest neighbors: towards removing the curse of dimensionality. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 604–613. ACM, 1998.
- [135] Shumeet Baluja and Michele Covell. Waveprint: Efficient wavelet-based audio fingerprinting. *Pattern recognition*, 41(11):3467–3480, 2008.
- [136] Christine Hine. *New infrastructures for knowledge production: Understanding e-science*. IGI Global, 2006.
- [137] Savas Parastatidis. A platform for all that we know: creating a knowledge-driven research infrastructure., 2009.
- [138] Brahim Medjahed and Athman Bouguettaya. *Service composition for the Semantic Web*. Springer Science & Business Media, 2011.
- [139] Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.
- [140] Grant Larsen. Component-based enterprise frameworks. *Communications of the ACM*, 43(10):24–26, 2000.

- [141] LA Clark, D Pregibon, JM Chambers, and TJ Hastie. Statistical models in s. *chapter Tree-Based Models*, pages 377–419, 1992.
- [142] Douglas Gillespie and Marjolaine Caillat. Statistical classification of odontocete clicks. *Canadian Acoustics*, 36(1):20–26, 2008.
- [143] Julie N Oswald, Jay Barlow, and Thomas F Norris. Acoustic identification of nine delphinid species in the eastern tropical pacific ocean. *Marine Mammal Science*, 19(1):20–037, 2003.
- [144] Matteo Frigo and Steven G Johnson. The design and implementation of fftw3. *Proceedings of the IEEE*, 93(2):216–231, 2005.
- [145] James W Cooley and John W Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301, 1965.
- [146] Michael Sipser. *Introduction to the Theory of Computation*, volume 2. Thomson Course Technology Boston, 2006.
- [147] James Nugent. A comparison of techniques for detecting clicks on recordings taken from vinyl, 2008.
- [148] Edwin Diday, Monique Noirhomme-Fraiture, et al. *Symbolic data analysis and the SODAS software*. Wiley Online Library, 2008.
- [149] Alexandre Cury and Christian Crémona. Pattern recognition of structural behaviors based on learning algorithms and symbolic data concepts. *Structural Control and Health Monitoring*, 19(2):161–186, 2012.
- [150] Edson B Novais, Artur Andriolo, Carlos CH Borges, Fabrizzio C Oliveira, and Thiago OS Amorim. Statistical signal similarity check using symbolic data for power management on low capacity devices. In *Telecommunications and Signal Processing (TSP), 2015 38th International Conference on*, pages 1–5. IEEE, 2015.
- [151] Laura C Carpi, Osvaldo A Rosso, Patricia M Saco, and Martín Gómez Ravetti. Analyzing complex networks evolution through information theory quantifiers. *Physics Letters A*, 375(4):801–804, 2011.
- [152] Michael J Swain and Dana H Ballard. Color indexing. *International journal of computer vision*, 7(1):11–32, 1991.
- [153] Francesca Odone, Annalisa Barla, and Alessandro Verri. Building kernels from binary strings for image matching. *Image Processing, IEEE Transactions on*, 14(2):169–180, 2005.
- [154] EH Helmer and B Ruefenacht. Cloud-free satellite image mosaics with regression trees and histogram matching. *Photogrammetric Engineering & Remote Sensing*, 71(9):1079–1089, 2005.
- [155] Mark Grundland and Neil A Dodgson. Color histogram specification by histogram warping. In *Electronic Imaging 2005*, pages 610–621. International Society for Optics and Photonics, 2005.

- [156] Nong Ye, Connie M Borrer, and Darshit Parmar. Scalable chi-square distance versus conventional statistical distance for process monitoring with uncorrelated data variables. *Quality and Reliability Engineering International*, 19(6):505–515, 2003.
- [157] Solomon Kullback. *Information theory and statistics*. Courier Corporation, 1968.
- [158] Marcin Budka, Bogdan Gabrys, and Katarzyna Musial. On accuracy of pdf divergence estimators and their applicability to representative data sampling. *Entropy*, 13(7):1229–1266, 2011.
- [159] Flemming Topsøe. Jenson-shannon divergence and norm-based measures of discrimination and variation. *preprint*, 2003.
- [160] Jianhua Lin. Divergence measures based on the shannon entropy. *Information Theory, IEEE Transactions on*, 37(1):145–151, 1991.
- [161] Pedro J Moreno, Purdy P Ho, and Nuno Vasconcelos. A kullback-leibler divergence based kernel for svm classification in multimedia applications. In *Advances in neural information processing systems*, page None, 2003.
- [162] Jean Dickinson Gibbons and Subhabrata Chakraborti. *Nonparametric statistical inference*. Springer, 2011.
- [163] Frank J Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78, 1951.
- [164] Morton B Brown and Alan B Forsythe. Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69(346):364–367, 1974.
- [165] Kazuo Tanaka. An introduction to fuzzy logic for practical applications, 1997.
- [166] Earl Cox. The fuzzy systems handbook: a practitioner’s guide to building, using, maintaining fuzzy systems. boston: Ap professional. Technical report, ISBN 0-12-194270-8, 1994.
- [167] Howard B Demuth. Electronic data sorting. *Computers, IEEE Transactions on*, 100(4):296–310, 1985.
- [168] Thomas H Cormen. *Introduction to algorithms*. MIT press, 2009.
- [169] Michael A Bender, Martin Farach-Colton, and Miguel A Mosteiro. Insertion sort is $o(n \log n)$. *Theory of Computing Systems*, 39(3):391–397, 2006.
- [170] Sultanullah Jadoon, Salman Faiz Solehria, and Mubashir Qayum. Optimized selection sort algorithm is faster than insertion sort algorithm: a comparative study. *International Journal of Electrical & Computer Sciences IJECS-IJENS*, 11(02):19–24, 2011.
- [171] Owen Astrachan. Bubble sort: an archaeological algorithmic analysis. In *ACM SIGCSE Bulletin*, volume 35, pages 1–5. ACM, 2003.
- [172] Yijie Han and Mikkel Thorup. Integer sorting in $o(n \text{ radic}(\log \log n))$ expected time and linear space. In *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*, pages 135–144. IEEE, 2002.

- [173] Arne Andersson, Peter Bro Miltersen, and Mikkel Thorup. Fusion trees can be implemented with ac 0 instructions only. *Theoretical Computer Science*, 215(1):337–344, 1999.
- [174] Sotiris B Kotsiantis, I Zaharakis, and P Pintelas. Supervised machine learning: A review of classification techniques, 2007.
- [175] Tom M Mitchell. Machine learning. 1997. *Burr Ridge, IL: McGraw Hill*, 45, 1997.
- [176] Richard O Duda, Peter E Hart, and David G Stork. *Pattern classification*. John Wiley & Sons, 2012.
- [177] Cyc. Graphic showing 3 Hyperplanes in 2D. H3 doesn't separate the 2 classes. H1 does, with a small margin and H2 with the maximum margin. https://commons.wikimedia.org/wiki/File:Svm_separating_hyperplanes.png. [Online; acessado 15 de novembro de 2015].
- [178] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152. ACM, 1992.
- [179] Saul C Leite and Raul Fonseca Neto. Incremental margin algorithm for large margin classifiers. *Neurocomputing*, 71(7):1550–1560, 2008.
- [180] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.
- [181] DCLDE Workshop. Detection, Classification, Localization, and Density Estimation Workshop 2015. <http://www.cetus.ucsd.edu/dclde/>. [Online; acessado 4 de novembro de 2015].
- [182] Cornell Lab of Ornithology. Raven: Interactive Sound Analysis Software. <http://www.birds.cornell.edu/brp/raven/RavenOverview.html>. [Online; acessado 1 de dezembro de 2015].

APÊNDICE A – Resultados Completos de Detecção

Para as seguintes tabelas, a coluna **Data** indica o dia exato de captura. A coluna **Duração** mostra o tempo total da captura naquele dia. A coluna **Segmentos**, mostra a quantidade de segmentos gerados a partir do tempo total de captura. A coluna **Total** mostra o tempo total anotado como de interesse por especialistas. A coluna **Percentual** indica a relação entre o tempo total de captura quando comparado ao tempo total anotado como de interesse. A coluna **IMA** indica a quantidade de segmentos que foram detectados como de possível interesse pelo modelo proposto. A coluna **IMA%** indica a relação entre a quantidade total de segmentos gerados e os considerados de interesse pelo modelo proposto. A coluna **Brown–Forsythe** indica a quantidade de segmentos que foram detectados como ocorrência pelo modelo proposto utilizando teste não paramétrico de Brown–Forsythe. A coluna **Brown–Forsythe%** indica a relação entre a quantidade total de segmentos gerados e os considerados como ocorrência pelo teste não paramétrico de Brown–Forsythe.

Data	Captura		Tempo de Interesse		Detecções			
	Duração	Segmentos	Total	Percentual	IMA	IMA%	Brown–Forsythe	Brown–Forsythe%
26/12/2011	23:55:00	16.816.406	4:11:15	17,51%	2.469.332	7,34%	7.563.226	22,49%
27/12/2011	24:00:00	16.875.000	4:27:30	18,58%	731.834	2,17%	4.963.254	14,71%
28/12/2011	24:00:00	16.875.000	6:01:15	25,09%	1.621.063	4,80%	5.211.549	15,44%
29/12/2011	24:00:00	16.875.000	1:08:45	4,77%	1.849.558	5,48%	6.136.486	18,18%
30/12/2011	24:00:00	16.875.000	4:35:00	19,10%	1.672.913	4,96%	5.964.535	17,67%
31/12/2011	24:00:00	16.875.000	1:40:00	6,94%	1.620.491	4,80%	6.597.637	19,55%
01/01/2012	23:34:59	16.581.835	2:07:35	9,02%	2.519.610	7,60%	7.653.459	23,08%
15/04/2012	23:52:29	16.786.914	3:09:30	13,23%	1.650.284	4,92%	6.986.231	20,81%
16/04/2012	24:00:00	16.875.000	4:05:38	17,06%	1.335.567	3,96%	5.569.466	16,50%
17/04/2012	24:00:00	16.875.000	6:44:55	28,12%	1.812.541	5,37%	5.992.360	17,76%
18/04/2012	24:00:00	16.875.000	2:49:08	11,75%	2.159.593	6,40%	6.956.248	20,61%
19/04/2012	24:00:00	16.875.000	4:58:45	20,75%	1.885.317	5,59%	6.654.823	19,72%
20/04/2012	24:00:00	16.875.000	2:06:40	8,80%	2.401.751	7,12%	7.598.746	22,51%
21/04/2012	23:37:30	16.611.328	6:01:30	25,50%	2.268.299	6,83%	6.495.237	19,55%
Total	334:59:58	235.546.483	54:07:26	16,16%	25.998.153	5,52%	90.343.257	19,18%

Tabela 10 – Resultados detalhados por dia da captura no projeto CINMS sítio B.

Data	Captura		Tempo de Interesse		Detecções			
	Duração	Segmentos	Total	Percentual	IMA	IMA%	Brown-Forsythe	Brown-Forsythe%
07/04/2012	23:44:04	13.350.625	0:55:29	3,90%	1.931.233	3,62%	8.955.426	16,77%
08/04/2012	24:00:00	13.500.000	2:07:38	8,86%	941.602	1,74%	13.315.759	24,66%
09/04/2012	24:00:00	13.500.000	1:03:52	4,44%	570.995	1,06%	8.347.604	15,46%
10/04/2012	24:00:00	13.500.000	0:00:00	0,00%	1.672.100	3,10%	13.775.763	25,51%
11/04/2012	24:00:00	13.500.000	0:15:07	1,05%	1.730.267	3,20%	11.542.172	21,37%
12/04/2012	24:00:00	13.500.000	0:07:05	0,49%	862.806	1,60%	11.052.717	20,47%
13/04/2012	23:58:25	13.485.156	2:15:52	9,45%	714.967	1,33%	14.142.262	26,22%
01/11/2012	23:41:41	13.328.281	2:33:04	10,77%	1.595.843	2,99%	10.363.085	19,44%
02/11/2012	24:00:00	13.500.000	1:10:51	4,92%	1.779.591	3,30%	12.995.819	24,07%
03/11/2012	24:00:00	13.500.000	4:08:08	17,23%	1.669.294	3,09%	8.721.560	16,15%
04/11/2012	24:00:00	13.500.000	5:19:06	22,16%	936.547	1,73%	9.317.238	17,25%
05/11/2012	24:00:00	13.500.000	7:31:37	31,36%	1.138.137	2,11%	14.456.844	26,77%
06/11/2012	23:57:47	13.479.218	0:53:07	3,69%	1.254.117	2,33%	13.716.278	25,44%
Total	311:21:57	175.143.280	28:20:56	9,10%	16.797.499	2,40%	150.702.527	21,51%

Tabela 11 – Resultados detalhados por dia da captura no projeto CINMS sítio C.

Data	Captura		Tempo de Interesse		Detecções			
	Duração	Segmentos	Total	Percentual	IMA	IMA%	Brown-Forsythe	Brown-Forsythe%
13/11/2012	23:50:13	13.408.281	2:29:49	10,48%	1.677.388	3,13%	15.460.988	28,83%
14/11/2012	24:00:00	13.500.000	0:48:52	3,39%	634.453	1,17%	9.882.811	18,30%
15/11/2012	24:00:00	13.500.000	0:58:30	4,06%	1.784.070	3,30%	16.899.603	31,30%
16/11/2012	24:00:00	13.500.000	1:49:00	7,57%	1.393.892	2,58%	13.229.708	24,50%
17/11/2012	24:00:00	13.500.000	0:32:43	2,27%	781.526	1,45%	12.192.353	22,58%
18/11/2012	24:00:00	13.500.000	2:15:57	9,44%	1.294.407	2,40%	16.228.757	30,05%
19/11/2012	23:52:17	13.427.656	2:18:39	9,68%	995.303	1,85%	16.990.618	31,63%
01/03/2013	23:53:33	13.439.531	2:02:12	8,52%	1.165.676	2,17%	18.922.185	35,20%
02/03/2013	24:00:00	13.500.000	2:20:02	9,72%	732.837	1,36%	10.479.569	19,41%
03/03/2013	24:00:00	13.500.000	0:01:19	0,09%	723.329	1,34%	12.602.489	23,34%
04/03/2013	24:00:00	13.500.000	3:43:25	15,52%	771.994	1,43%	10.874.616	20,14%
05/03/2013	24:00:00	13.500.000	2:04:04	8,62%	2.009.032	3,72%	14.825.451	27,45%
06/03/2013	24:00:00	13.500.000	0:56:42	3,94%	1.580.141	2,93%	18.706.815	34,64%
Total	311:36:03	175.275.468	22:21:14	7,17%	15.544.048	2,22%	187.295.963	26,71%

Tabela 12 – Resultados detalhados por dia da captura no projeto DCPD sítio A.

Data	Captura		Tempo de Interesse		Detecções			
	Duração	Segmentos	Total	Percentual	IMA	IMA%	Brown-Forsythe	Brown-Forsythe%
11/12/2012	23:45:32	13.364.375	1:25:50	6,02%	842.906	1,58%	15.154.691	28,35%
12/12/2012	24:00:00	13.500.000	0:02:42	0,19%	1.325.095	2,45%	13.642.251	25,26%
13/12/2012	24:00:00	13.500.000	0:02:29	0,17%	1.392.942	2,58%	11.973.590	22,17%
14/12/2012	24:00:00	13.500.000	0:17:48	1,24%	1.897.689	3,51%	17.888.186	33,13%
15/12/2012	24:00:00	13.500.000	2:50:20	11,83%	1.827.091	3,38%	15.902.269	29,45%
16/12/2012	24:00:00	13.500.000	2:21:34	9,83%	754.277	1,40%	13.109.636	24,28%
17/12/2012	24:00:00	13.500.000	1:02:57	4,37%	1.919.873	3,56%	16.602.102	30,74%
06/03/2013	23:42:38	13.337.187	3:52:28	12,12%	838.176	1,57%	16.581.636	31,08%
07/03/2013	24:00:00	13.500.000	1:09:46	4,84%	1.739.788	3,22%	17.019.476	31,52%
08/03/2013	24:00:00	13.500.000	2:54:09	12,09%	560.463	1,04%	10.617.299	19,66%
09/03/2013	24:00:00	13.500.000	0:56:07	3,90%	1.597.525	2,96%	19.368.553	35,87%
10/03/2013	24:00:00	13.500.000	0:48:45	3,39%	1.567.798	2,90%	13.325.813	24,68%
Total	287:28:10	161.701.562	17:44:55	6,17%	16.263.623	2,51%	181.185.502	28,01%

Tabela 13 – Resultados detalhados por dia da captura no projeto DCPD sítio B.

Data	Captura		Tempo de Interesse		Detecções			
	Duração	Segmentos	Total	Percentual	IMA	IMA%	Brown-Forsythe	Brown-Forsythe%
09/12/2012	23:26:15	16.479.492	17:26:57	74,45%	2.090.601	6,34%	11.049.244	33,52%
10/12/2012	24:00:00	16.875.000	5:20:48	22,28%	758.678	2,25%	8.492.110	25,16%
11/12/2012	24:00:00	16.875.000	12:11:30	50,80%	1.353.205	4,01%	10.080.971	29,87%
12/12/2012	24:00:00	16.875.000	9:56:35	37,26%	2.255.147	6,68%	9.965.044	29,53%
13/12/2012	24:00:00	16.875.000	0:35:00	2,43%	1.463.729	4,34%	13.574.476	40,22%
14/12/2012	24:00:00	16.875.000	6:31:00	27,15%	2.131.287	6,31%	11.789.427	34,93%
15/12/2012	23:46:14	16.713.671	3:32:57	14,93%	1.146.685	3,43%	8.308.191	24,85%
05/02/2013	23:52:30	16.787.109	3:00:44	12,62%	1.069.402	3,19%	9.143.922	27,23%
06/02/2013	24:00:00	16.875.000	6:24:00	26,67%	939.189	2,78%	7.095.573	21,02%
07/02/2013	24:00:00	16.875.000	1:59:01	8,27%	2.412.150	7,15%	13.804.252	40,90%
08/02/2013	24:00:00	16.875.000	2:20:01	9,72%	703.159	2,08%	7.075.928	20,97%
09/02/2013	24:00:00	16.875.000	9:13:21	38,43%	1.981.645	5,87%	11.314.797	33,53%
10/02/2013	23:34:59	16.581.835	1:14:23	5,26%	851.558	2,57%	8.941.442	26,96%
Total	310:39:58	218.437.107	79:46:17	25,68%	19.156.435	4,38%	130.635.377	29,90%

Tabela 14 – Resultados detalhados por dia da captura no projeto DCPD sítio C.

Data	Captura		Tempo de Interesse		Detecções			
	Duração	Segmentos	Total	Percentual	IMA	IMA%	Brown-Forsythe	Brown-Forsythe%
20/04/2009	23:56:15	16.831.054	9:06:22	38,04%	699.440	2,08%	6.499.846	19,31%
21/04/2009	24:00:00	16.875.000	9:51:15	41,06%	1.831.799	5,43%	8.919.396	26,43%
22/04/2009	24:00:00	16.875.000	6:41:15	27,86%	2.108.960	6,25%	13.760.542	40,77%
23/04/2009	24:00:00	16.875.000	8:44:13	36,40%	835.415	2,48%	6.620.580	19,62%
24/04/2009	24:00:00	16.875.000	6:36:15	27,52%	1.613.436	4,78%	8.039.429	23,82%
25/04/2009	24:00:00	16.875.000	8:56:08	37,23%	1.980.749	5,87%	7.542.098	22,35%
26/04/2009	23:59:59	16.874.804	12:54:45	53,80%	852.925	2,53%	10.324.998	30,59%
05/06/2009	23:33:45	16.567.382	4:33:49	19,37%	1.046.288	3,16%	8.136.145	24,55%
06/06/2009	24:00:00	16.875.000	4:57:30	20,66%	1.585.434	4,70%	8.233.634	24,40%
07/06/2009	24:00:00	16.875.000	2:46:15	11,55%	1.861.807	5,52%	4.279.550	12,68%
08/06/2009	24:00:00	16.875.000	3:55:00	16,32%	1.623.414	4,81%	5.978.060	17,71%
09/06/2009	24:00:00	16.875.000	5:42:13	23,77%	977.492	2,90%	9.150.498	27,11%
Total	287:29:59	202.148.240	84:45:00	29,48%	17.017.159	4,21%	97.484.776	24,11%

Tabela 15 – Resultados detalhados por dia da captura no projeto SOCAL sítio E.

Data	Captura		Tempo de Interesse		Detecções			
	Duração	Segmentos	Total	Percentual	IMA	IMA%	Brown–Forsythe	Brown–Forsythe%
05/11/2009	23:46:15	16.713.867	4:56:00	20,75%	1.594.199	4,77%	8.373.530	25,05%
06/11/2009	24:00:00	16.875.000	5:01:16	20,92%	2.319.254	6,87%	7.956.617	23,58%
07/11/2009	24:00:00	16.875.000	4:05:39	17,06%	1.072.975	3,18%	10.128.348	30,01%
08/11/2009	24:00:00	16.875.000	7:17:38	30,39%	2.091.097	6,20%	7.723.011	22,88%
09/11/2009	24:00:00	16.875.000	3:09:02	13,13%	1.043.598	3,09%	6.046.920	17,92%
10/11/2009	24:00:00	16.875.000	9:52:35	41,15%	832.926	2,47%	7.596.895	22,51%
11/11/2009	24:00:00	16.875.000	10:05:24	42,04%	2.215.719	6,57%	8.775.842	26,00%
13/04/2010	23:47:30	16.728.515	3:48:28	16,00%	1.382.120	4,13%	6.110.767	18,26%
14/04/2010	24:00:00	16.875.000	8:56:15	37,24%	1.068.588	3,17%	9.313.492	27,60%
15/04/2010	24:00:00	16.875.000	5:25:00	22,57%	1.745.424	5,17%	8.074.506	23,92%
16/04/2010	24:00:00	16.875.000	6:00:00	25,00%	2.116.625	6,27%	11.532.008	34,17%
17/04/2010	24:00:00	16.875.000	5:41:01	23,68%	2.387.046	7,07%	8.016.417	23,75%
18/04/2010	24:00:00	16.875.000	10:17:01	42,85%	1.967.079	5,83%	5.878.025	17,42%
19/04/2010	23:48:44	16.742.968	12:50:09	53,90%	1.135.240	3,39%	10.980.609	32,79%
Total	335:22:29	235.810.350	97:25:28	29,05%	22.971.890	4,87%	116.506.987	24,70%

Tabela 16 – Resultados detalhados por dia da captura no projetoSOCAL sítio R.