

Classificação de Sinais Temporais de Sonar Passivo utilizando Redes Neurais Convolucionais

1st Marlon Jovenil de Souza

Laboratório de Processamento de Sinais, Coppe-UFRJ
Universidade Federal do Rio de Janeiro
Rio de Janeiro, Brasil
marlon.jovenil@coppe.ufrj.br

2nd José Manoel de Seixas

Laboratório de Processamento de Sinais, Coppe-UFRJ
Universidade Federal do Rio de Janeiro
Rio de Janeiro, Brasil
seixas@lps.ufrj.br

3rd Natanael Nunes de Moura Júnior

Dep. de Engenharia Eletrônica e de computação, Poli-UFRJ
Universidade Federal do Rio de Janeiro
Rio de Janeiro, Brasil
natmourajr@lps.ufrj.br

Resumo—O oceano é atualmente explorado em vários níveis. Devido ao interesse crescente pelo mundo no ambiente marinho, surge a necessidade de compreendê-lo melhor através da acústica submarina. Sinais irradiados no meio submarino podem ser detectados e classificados por diversos sistemas. As Redes Neurais Artificiais, em especial as chamadas Redes profundas, são ótimas ferramentas de apoio para o estudo desses sinais. Neste artigo são propostos dois modelos de Redes Neurais Artificiais, *Multilayer Perceptron* (MLP) e *Convolutional Neural Network* (CNN), a fim de classificar sinais de navios provenientes de um sistema de Sonar Passivo fornecidos pela Marinha do Brasil. Outros trabalhos desenvolvidos com esta base de dados analisam o sinal no domínio da frequência, entretanto este artigo visa uma análise no domínio do Tempo. Os resultados obtidos serão analisados utilizando algumas figuras de méritos já estabelecidas no meio acadêmico. Através delas é possível comparar o desempenho de classificação dos sinais do classificador baseado em modelo profundo com o classificador de rede rasa.

Palavras-chave—Sistemas de Sonar Passivo, Aprendizado de Máquina, Aprendizado Profundo, Redes Neurais Convolucionais.

I. INTRODUÇÃO

A identificação e classificação de sinais acústicos submarinos é uma das funções de um sistema SONAR (Sound Navigation and Ranging), o qual é um conjunto de técnicas que utiliza o som que se propaga através da água para extração de informações no ambiente marinho [1]. O Sistema Sonar é dividido em dois tipos: ativo e passivo. Em sistemas de sonar ativo há uma emissão de um pulso sonoro conhecido e através do eco deste pulso é realizado o processamento do sinal. Enquanto isso, em um sistema de sonar passivo, não há emissão de um pulso sonoro conhecido para análise do eco, o sinal processado é emitido pelo alvo de interesse [2]. Esses sistemas são comumente utilizados no meio militar e civil para diferentes aplicações. Como exemplo de utilização no meio civil temos sonar de pesca, ecobatímetros, exploração mineral, sonares de varredura, dentre outros. No meio militar, esses sistemas são empregados no monitoramento e vigilância de portos e também são as principais fontes de orientação para submarinos e outras embarcações [3].

Os sinais dos sistemas de sonar passivo são dados não-linearmente separáveis e possuem distribuição não gaussiana, tal contexto somado ao fato de que fenômenos sísmicos e outros ruídos marinhos podem emitir sinais com assinaturas acústicas similares ao sinal de interesse, torna desafiadora a tarefa de processar e classificar esses dados [4]. Em virtude desta complexidade, atualmente, diversas técnicas para classificação de contatos por meio de sinais adquiridos de sistemas sonar vêm sendo desenvolvidas [5]. As redes neurais artificiais (RNA) são especialmente indicadas para esta tarefa, tendo em vista que apresentam uma classificação eficiente, mesmo para um número expressivo de classes e para eventos de dimensionalidade elevada, além de possuírem facilidades quanto à adaptação, em especial, na incorporação de novas classes [6]. Dentre as RNA, as técnicas de aprendizado profundo se destacam, enquanto um ramo de aprendizado de máquina baseado em um conjunto de algoritmos que tentam modelar abstrações de alto nível de dados [7]. As duas arquiteturas de redes neurais utilizadas neste trabalho foram as redes perceptron multicamadas (*Multi Layer Perceptron*, MLP) que conseguem resolver computacionalmente problemas complexos graças à existência das camadas escondidas não-lineares, além de possuir como característica um alto grau de conectividade determinado pelas sinapses da rede [8] e as redes neurais convolucionais (*Convolutional Neural Networks*, CNN), pois mostraram em trabalhos anteriores serem uma alternativa viável para previsão de séries temporais [9].

Diversos assuntos já foram abordados dentro do contexto de sonar passivo, como por exemplo, técnicas de classificação baseadas em classificadores especialistas, do tipo classe-não-classe [10]. Foram abordados problemas de detecção de novidade para um classificador de sonar passivo utilizando modelos baseados em *Support Vector Machines* (SVM) [11] e redes MLP [7]. Neste contexto as chamadas componentes principais de discriminação foram aplicadas no pré-processamento para aumento da eficiência de detecção de novidade e constatou-se que quanto mais plásticos os classificadores (quanto menos

taxa de falso alarme), menos detecção de novidade é observada no detector de novidade [7] [11]. Um sistema classificador neural para sinais de sonar passivos foi desenvolvido com base em um método de pré-processamento que era baseado principalmente em informações espectrais médias e transformadas wavelet. Este sistema conseguiu classificar corretamente mais de 94% dos sinais de quatro classes diferentes [12]. Há trabalhos que fazem uso de redes do tipo *Stacked Autoencoders* (SAE) para inicializar os pesos de uma rede profunda de maneira não-supervisionada com o intuito de classificar sinais e detectar novidades em sonar passivo. Os resultados apresentados para as redes profundas mostraram resultados promissores se comparados aos de RNA rasas [4].

Este artigo tem por objetivo o desenvolvimento de classificadores no domínio do tempo para sinais de um sistema de sonar passivo, utilizando estruturas de redes MLP e redes CNN, tendo em vista que trabalhos como [7] abordam o processo de classificação desses sinais com *baselines* no domínio da frequência, os quais servirão de apoio para a construção dos *baselines* deste artigo. Os dados utilizados foram fornecidos pelo Instituto de Pesquisas da Marinha (IpqM) e são oriundos de navios da Marinha do Brasil (MB) pertencentes a 4 classes. Os resultados experimentais de classificação provenientes de cada modelo serão avaliados tendo como base as figuras de mérito especificadas na seção IV-C. Após avaliação, os resultados da análise no domínio do tempo serão comparados com os do domínio da frequência. Em suma, esperamos que esta nova abordagem para sistemas de sonar passivo possa agregar valor às abordagens já consolidadas no meio acadêmico.

O artigo está organizado da seguinte maneira: a Seção II descreve um sistema de sonar passivo e o processamento do sinal adquirido por ele, a Seção III apresenta brevemente as redes neurais convolucionais e na Seção IV é explicado o método desenvolvido. Os resultados são analisados na Seção V. Por fim, a conclusão dos resultados experimentais são discutidos na Seção VI.

II. SISTEMA DE SONAR PASSIVO

Os ruídos irradiados por navios ao longo do tempo podem ser modelados e representados como n sinais s_1, s_2, \dots, s_n , que são adquiridos por n transdutores, chamados de hidrofo-nes, os quais compõem o sistema de sonar passivo. Após a aquisição, esses sinais geralmente passam por um processo de conformação de feixes, que consiste em uma filtragem espacial dos sinais provenientes dos transdutores para a obtenção das direções de aproximação (DOA) [7]. Os sinais conformados podem ser observados nos domínios do tempo e da frequência [3]. Os dois domínios podem ser utilizados para detecção dos sinais, e, uma vez que os sinais forem detectados, o seu acompanhamento pode ser realizado [7]. Com os contatos sendo acompanhados, dá-se início ao processo de classificação do sinal através de uma análise no domínio do tempo, apoiado em modelos de aprendizado de máquina (*Machine Learning*) para decisão automatizada ou apoio à tomada de decisão por parte de um operador [10].

A. Análise no Domínio do Tempo

O fluxograma da análise no domínio do tempo pode ser visto na Fig.1. Os sinais de uma determinada direção de aproximação (DOA), na qual é estimada utilizando algoritmos de conformação de feixes, podem ser submetidos a uma extração de características visando a sua classificação [7]. Na análise temporal do sinal, utiliza-se uma decimação do sinal por um fator d , processo o qual reduz a frequência de amostragem do sinal, seguido de um filtro passa-baixa digital de ordem N , dado que a informação relevante para classificação de sinais de Sonar Passivo encontra-se nas baixas frequências [7]. Após este processo, blocos consecutivos de n pontos do sinal sofrem uma normalização na magnitude do sinal.

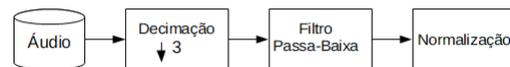


Figura 1. Diagrama em blocos da Análise no Domínio do Tempo.

A análise no domínio do tempo deriva de uma análise de banda larga que trabalha sobre o ruído de máquina, denominada LOFAR (*LOW Frequency Analysis and Recording*) [5] e seu processo está representado na Fig. 2. O início do processo da análise LOFAR é similar ao descrito na análise no domínio do tempo, porém ele difere após o processo de decimação do sinal. Enquanto o sinal no tempo é normalizado após a passagem por um filtro passa-baixa, o sinal na análise LOFAR passa por uma janela de *Hannig* seguida de uma transformada de Fourier de Tempo Curto (STFT). Em seguida, normaliza-se o módulo do espectro e, por fim, a informação passa por uma função TPSW (*two-pass split window*) com o propósito de estimar o ruído de fundo do ambiente de medição [7].

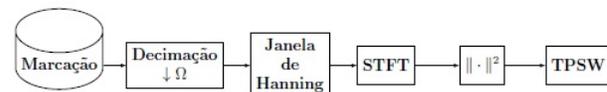


Figura 2. Diagrama Análise LOFAR.

III. ARQUITETURA DE REDES ARTIFICIAIS

Um estudo aprofundado sobre séries temporais foge ao escopo deste trabalho. De uma maneira mais formal, uma série temporal é uma realização de um processo estocástico. Ela difere de um processo *iid* (independente e identicamente distribuído) já que existe uma dependência entre seus valores. Os processos estocásticos baseados na premissa de estarem em um estado particular de equilíbrio estatístico são chamados de estacionários [13]. As RNAs, que são estruturas compostas por unidades de processamento não-lineares (neurônios) totalmente conectadas baseadas em redes neurais biológicas, vêm sendo amplamente utilizadas como ferramenta para análise, previsão e classificação de séries temporais [14].

As redes MLP são constituídas de uma camada que recebe os sinais de entrada, uma ou mais camadas intermediárias

e uma camada de saída. As camadas intermediárias são amplamente conhecidas como camadas escondidas [15]. As MLP são ainda caracterizadas pelas elevadas possibilidades de aplicações em diversos tipos de problemas relacionados com as mais diferentes áreas do conhecimento, sendo também consideradas uma das arquiteturas mais versáteis quanto à aplicabilidade [15].

Aprendizado profundo é o nome dado a uma classe de algoritmos de aprendizado de máquinas que utilizam várias camadas de processamento [16]. A ideia por trás da aprendizagem profunda é descobrir múltiplos níveis de representação com a expectativa de que recursos de alto nível representem uma semântica mais abstrata dos dados [17]. Dentre esses algoritmos, pode se destacar as CNN [9]. Criadas originalmente para tarefas envolvendo imagens, dados em duas dimensões, e são basicamente alimentadas por dados pré-processados através de convoluções na imagem de entrada por certos filtros (comumente chamados de *kernels*) [9]. Na análise de série temporais, os dados são unidimensionais, sendo assim, pode-se utilizar uma convolução unidimensional operando com vetores, como pode ser visto na Fig.3. Dessa forma, uma camada convolucional unidimensional é utilizada para extrair informações da série temporal por meio dos filtros e utiliza os dados filtrados como entrada de uma MLP, que tem em sua saída a previsão ou classificação desejada [9]. Como o operador convolução é aplicado em subsequências da série de entrada, um padrão aprendido em uma localização pode também ser identificado em outras posições [18].

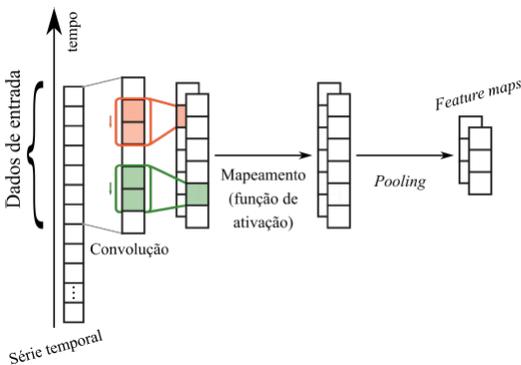


Figura 3. Diagrama de camada convolucional unidimensional.

IV. MÉTODO

A. Treinamento da Rede Neural

Após o pré-processamento do sinal pela análise no domínio do tempo descrita na seção II-A, o sinal serviu de entrada para duas arquiteturas de redes neurais, MLP e CNN, as quais foram projetadas e treinadas na linguagem *python*, utilizando a biblioteca *keras*.

O *baseline* da rede MLP possui nós de entrada equivalentes ao tamanho da janela definida no pré-processamento do sinal, uma camada intermediária com n neurônios e uma

camada de saída com a quantidade de neurônios correspondente a quantidade de classes do problema. Enquanto na camada intermediária foram utilizados neurônios com função de ativação do tipo ReLU, na camada de saída as funções dos neurônios são do tipo Softmax. A função Softmax tem a capacidade de indicar a probabilidade de uma saída ser um valor, dado o conhecimento da entrada, além de contribuir para um melhor desempenho da operação do classificador [16]. A rede é treinada usando o algoritmo *backpropagation*. Em suma, o método é composto por duas fases, *forward* e *backward*, que fazem com que os pesos sinápticos e limiares dos neurônios se ajustem automaticamente em cada iteração [15]. Seguindo as especificações citadas e utilizando o método da validação cruzada do item IV-B, alguns *hyperparâmetros* da rede foram variados, a fim de encontrar o melhor modelo de classificação. Em particular, a quantidade de neurônios da camada intermediária.

O *baseline* da rede CNN possui duas camadas convolucionais com n filtros e a saída das camadas convolucionais estão conectadas a entrada de duas camadas MLP com m neurônios. Os filtros e neurônios de todas as camadas (CNN e MLP) possuem funções de ativação ReLU, exeto a última com função Softmax, em virtude de ser a camada de saída. Assim como descrito no parágrafo acima para as redes MLP, também foi utilizada a validação cruzada aliada a variação dos *hyperparâmetros* para definir o melhor modelo de classificação para as redes CNN.

B. Validação Cruzada

Com o intuito de estimar a flutuação estatística dos modelos propostos durante o processo de treinamento, o método de validação cruzada *jackknife* [19] foi utilizado. O método consistiu em dividir o conjunto de dados aleatoriamente em K partes de mesmo tamanho (K -folds), a extração de características do modelo é feita utilizando $K - 1$ partes, e a parcela restante fica destinada a avaliar a generalização do modelo. Esse processo é repetido K vezes (em cada momento uma partição diferente é destinada a avaliar a generalização do modelo). Ao término, a incerteza estimada do processo pode ser feita através do cálculo das figuras de mérito nos modelos oriundos dos k folds.

C. Figuras de Mérito

Com o intuito de avaliar o desempenho do modelo, usamos as seguintes figuras de mérito clássicas: *Accuracy* (Accu), *Precision* (Pre), *Recall* (Rec) e F_1 score.

Para casos mais complexos de classificação, como o caso de 4 classes, precisaremos do auxílio de algumas figuras de mérito diferentes das figuras clássicas de classificação mencionadas acima. A Equação (1) apresenta a Eficiência de classificação que é determinada pelo número de eventos classificados corretamente dividido pelo número total de eventos. Esta Eficiência em conjunto com a Acurácia de classificação apresentada na Equação (2) servem como base para calcular outra figura de mérito importante a ser utilizada, o índice SP descrito na Equação (3), na qual tem sido adotado em

trabalhos anteriores mensurando a eficiência de classificação balanceada para um número grande de classes [4].

$$Eff_k(\theta) = \frac{\text{eventos}_k \text{ classificados corretamente}}{\text{total de eventos}_k} \quad (1)$$

$$A_{cc}(\theta) = \frac{\sum_{k=1}^{N \text{ classes}} Eff_k(\theta)}{N \text{ classes}} \quad (2)$$

$$SP(\theta) = \sqrt{\frac{\sum_{n=1}^N Eff_i(\theta)}{N}} \cdot \sqrt{\prod_{n=1}^N Eff_i(\theta)} \quad (3)$$

V. RESULTADOS EXPERIMENTAIS

A. Dados Experimentais

Os dados utilizados neste trabalho correspondem a gravações de ruídos irradiados por navios de quatro classes diferentes (Classe A, B, C e D) pertencentes à Marinha do Brasil e produzidos em uma raia acústica em Arraial do Cabo no Rio de Janeiro. Cada gravação foi feita com um único hidrofone omnidirecional posicionado a 45 metros de profundidade e a raia acústica é um ambiente controlado de 1500 metros de comprimento. O hidrofone foi acoplado a um sistema de aquisição com frequência de amostragem de $22.050Hz$ e 8 bits de resolução. Na raia acústica os navios realizavam corridas que iniciavam a 1000 metros do hidrofone e terminavam a 500 metros depois da posição do hidrofone. As classes B e D totalizaram 10 corridas diferentes cada uma, enquanto as classes A e C têm 5 e 9 corridas, respectivamente.

Os sinais de sonar passivo têm sua informação de interesse concentrada em frequências mais baixas [7], portanto, uma decimação de ordem baixa foi utilizada para a análise no domínio do tempo descrita na Seção II-A. Os sistemas da MB utilizam dois fatores de decimação, 3 e 4. Com base na análise descrita em [7], esses dois fatores são comparados, constatando-se uma diferença relativamente pequena nos resultados. Deste modo, este trabalho propõe a utilização de um fator igual a 3, o que resultou na redução da taxa de amostragem do sinal para $7.350Hz$. Em seguida, o sinal passou por um filtro passa-baixas de ordem 8, que foi desenvolvido com frequência de corte em $2.870Hz$ e atenuação de $60dB$ para frequências acima de $4.480Hz$. Além disso, outro fator importante na análise no domínio do tempo é o tamanho da janela de processamento. O tamanho da janela de processamento tem um impacto direto sobre a estacionaridade dos sinais processados. Uma das estratégias para atacar o problema da não-estacionaridade é reduzir o tamanho da janela de processamento. Quanto maior a largura de cada janela, maior a probabilidade de perda de estacionaridade. Nos resultados apresentados na Fig.4 foram avaliados a quantidade de pontos por janela, a qual foi variando, utilizando-se a variação em potências de 2 com os seguintes valores: 256, 512, 1.024, 2.048 e 4.096, a fim de definir o melhor tamanho de janela. As características do modelo utilizado serão definidas na seção

V-B. Por enquanto basta sabermos que a arquitetura utilizada é a CNN. A figura de mérito utilizada na avaliação do treinamento foi o índice SP, o qual teve o seu valor monitorado para o conjunto de validação. O conjunto de validação é mencionado com mais detalhes na seção V-B. Em suma, ele é uma subdivisão do conjunto de treinamento e fica responsável em verificar a eficiência da rede quanto a sua capacidade de generalização durante o treinamento. Considerando os resultados obtidos, a maximização do balanceamento da classificação ocorre para as janelas de 512 e 1024 pontos. Embora as janelas de 512 pontos tenham apresentado os melhores resultados (índice SP igual a $71,5 \pm 2,1$) para o conjunto de validação, os sistemas da MB utilizam janelas de 1024 pontos para a análise LOFAR. Desse modo, visando uma futura implementação da metodologia aplicada neste artigo, optamos por usar janelas de 1024 pontos, na qual obteve resultados (índice SP igual a $68,0 \pm 1,1$) próximos aos das janelas de 512 pontos. Por fim, os dados foram normalizados de forma que sua magnitude fique restrita a valores entre 0 e 1.

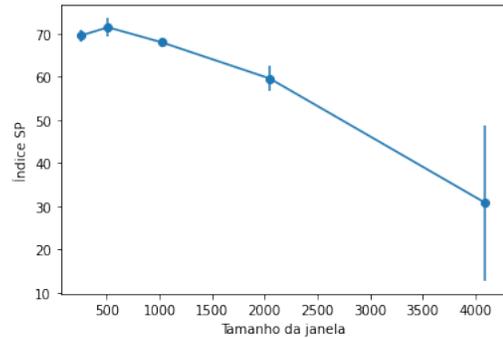


Figura 4. Variação do tamanho das janelas em comparação com o índice SP

A Tabela I apresenta a quantidade de amostras para cada classe obtida após esta etapa de processamento.

Tabela I
NÚMERO DE JANELAS POR CLASSE

Classes	Janelas de 1024 pontos sem <i>Overlap</i>
A	4.312
B	9.781
C	3.833
D	7.918
TOTAL	25.844

B. Treinamento e Definição da Topologia dos Modelos

Para encontrar o melhor modelo de classificação foram utilizados os métodos descritos na seção IV para treinamento e avaliação das arquiteturas mencionadas na seção III. Na topologia MLP as camadas de entrada e saída possuem respectivamente 1024 nós e 4 neurônios, o mesmo vale para a topologia CNN, dado o tamanho da janela do sinal de entrada e a quantidade de classes. Nas camadas intermediárias foram realizadas variações no número de neurônios e filtros, avaliando-se o índice SP nas diferentes topologias formadas a partir

desta variação. A validação cruzada descrita na seção IV-B foi utilizada durante o treinamento dos classificadores. Para tal, o conjunto de dados durante o processo de treinamento foi particionado em 5 *folds*, ou seja, para cada uma das 5 iterações, 80% do conjunto foi destinado ao treinamento do classificador neural e os 20% restantes foram utilizados para teste. O conjunto destinado ao treinamento foi subdividido em 5 partes para que dessa forma 4 partes efetuassem de fato o treinamento do modelo e a validação deste modelo ficasse com a parte restante, denominada conjunto de validação.

A função de perda de classificação usada em todas as redes foi a entropia cruzada e o método de otimização para calcular a taxa de aprendizado adaptativa de cada parâmetro foi o Adam [20], com uma taxa de aprendizado de 0,0001 e o parâmetro $\beta = 0,9$ para controlar a taxa de decaimento. Com o intuito de evitar *overfitting* do modelo, dois métodos foram aplicados. O primeiro é o critério de parada chamado, *Early Stopping*, o qual consiste em uma forma de regularização usada para evitar ajustes excessivos. Caso o treinamento não evolua em dez épocas, o treinamento é interrompido. O segundo é o método de regularização chamado de *Dropout*, que é amplamente difundido e utilizado no meio científico.

Para exemplificar a escolha dos hiperparâmetros, a figura 5 apresenta os resultados do índice SP aplicado ao conjunto de validação com base no modelo treinado para diferentes números de neurônios da camada intermediária da arquitetura MLP. Neste modelo apenas uma camada intermediária foi utilizada. Este resultado permite avaliar o melhor hiperparâmetro (neurônio), para o qual o desempenho de classificação dos sinais é maximizado. Analisando a figura podemos constatar que a partir de 70 neurônios, o índice SP não apresenta melhoras significativas. Sendo assim, com o intuito de obter o menor número de parâmetros para o modelo, o número de neurônios escolhido para a camada intermediária é de 70.

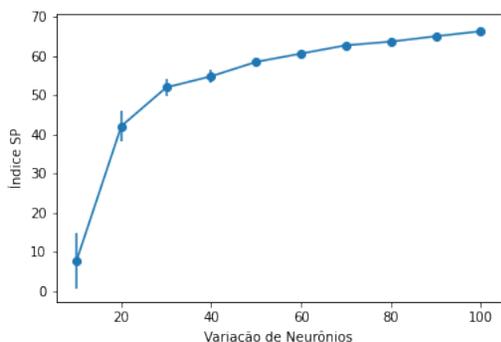


Figura 5. Escolha de neurônios do modelo MLP.

Vale ressaltar que a curva de perda de cada *fold* também é monitorada a fim de evitar o *overfitting*. A figura 6 mostra a curva de perda do conjunto de treinamento e do conjunto de validação para o melhor modelo de um *fold*.

A Tabela II apresenta os resultados da variação de filtros e neurônios para a arquitetura CNN em relação ao índice SP aplicado ao conjunto de validação com base no modelo

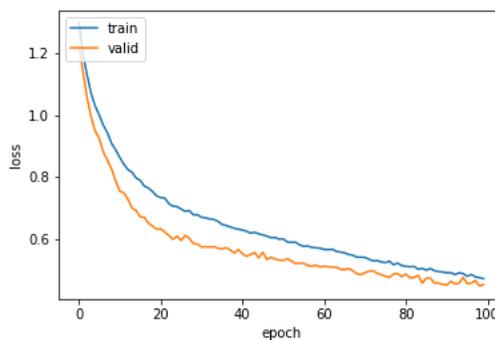


Figura 6. Curva de perda do melhor modelo de um *fold* de treinamento e validação para a arquitetura MLP.

treinado por *folds* do conjunto de treinamento. O modelo utilizado apresenta duas camadas convolucionais e duas camadas *fully Connected*, sendo uma delas a camada de saída. Com base nos resultados obtidos, o modelo com 32 e 64 filtros respectivamente na primeira e segunda camada convolucional e 50 neurônios na camada *fully Connected* (MLP), dado que a última camada possui 4 neurônios. Esta configuração apresentou o melhor desempenho na classificação dos sinais levando em consideração os valores do índice SP.

Tabela II
VARIAÇÃO DO NÚMERO DE FILTROS NA ARQUITETURA CNN

Filtros 1ª CNN	Filtros 2ª CNN	Neurônios MLP	SP(%)
32	64	32	$60,4 \pm 2,8$
32	64	50	$71,5 \pm 2,1$
32	32	50	$66,3 \pm 1,4$
64	64	50	$65,5 \pm 3,6$

C. Avaliação dos Classificadores

As Tabelas III e IV apresentam os resultados de Precisão, Recall e F1-score das duas topologias dos modelos de classificadores, MLP e CNN, descritos na seção V-B. Além disso, as figuras 7 e 8 também servem de apoio à comparação dos resultados obtidos pelos dois modelos.

Os resultados do índice SP aplicado ao conjunto de teste dos classificadores MLP e CNN desenvolvidos foram de $61,16 \pm 0,6$ e $71,0 \pm 2,6$ respectivamente. Analisando esses resultados podemos observar que o modelo CNN apresenta na classificação um melhor balanceamento das classes em comparação com o modelo MLP. Outro fator observado analisando os resultados de precisão e o valor de *recall* da classe C é que os classificadores tiveram um desempenho inferior de classificação para esta classe. O resultado já era esperando devido ao menor número de amostras desta classe em comparação com as outras do *dataset*. Embora o classificador CNN tenha apresentado uma pequena melhora na classificação da classe C. Apesar de o tamanho das amostras ser um fator importante para a classificação, é preciso ressaltar que as corridas realizadas pelos navios de mesma classe foram feitas sobre condições de máquinas e mar distintas. Sendo assim,

a correlação entre as corridas também é relevante para a composição do *dataset* das classes. Isso pode explicar o bom desempenho dos classificadores para a classe A, a qual possui um número reduzido de amostras.

Em [21] é proposto um sistema de suporte para identificação de alvos, especialmente em aplicações de múltiplos alvos. O sistema utiliza os dados oriundos do sistema de sonar passivo e apresenta resultados de classificação utilizando a análise LOFAR. Sendo assim, é possível comparar os resultados do classificador utilizando a análise no domínio do tempo com a do domínio da frequência. O índice SP encontrado pela análise LOFAR foi de aproximadamente 87%. Esse resultado é superior aos encontrados nos classificadores MLP e CNN no domínio do tempo. Logo, é possível constatar uma leve vantagem ao classificar os sinais no domínio frequência, embora o processamento no domínio do tempo tenha as vantagens de ser mais rápido e requer menos recurso computacional, o que é importante para as aplicações no meio naval.

Tabela III
RESULTADOS EXPERIMENTAIS PARA O MODELO MLP

Classes	Precisão(%)	Recall(%)	F1-score(%)
A	91,0 ± 0,7	73,8 ± 0,8	81,6 ± 0,5
B	65,2 ± 0,5	91,0 ± 0,7	76,0 ± 0,5
C	72,0 ± 2,2	16,6 ± 1,6	26,8 ± 2,3
D	87,8 ± 1,1	87,6 ± 0,5	87,6 ± 0,5

Tabela IV
RESULTADOS EXPERIMENTAIS PARA O MODELO CNN

Classes	Precisão(%)	Recall(%)	F1-score(%)
A	82,8 ± 1,8	77,4 ± 1,0	80,2 ± 1,0
B	74,8 ± 2,1	94,4 ± 1,5	83,6 ± 0,9
C	63,8 ± 2,5	35,2 ± 7,5	45,2 ± 5,7
D	94,0 ± 1,1	86,6 ± 0,9	90,0 ± 0,8

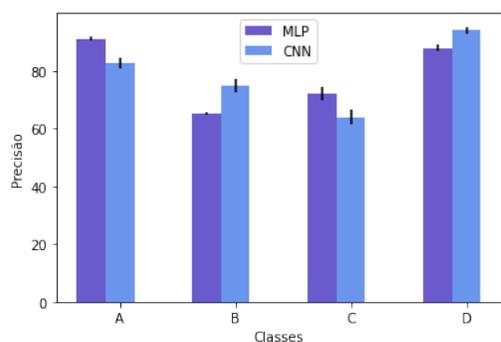


Figura 7. Gráfico de comparação da precisão dos modelos MLP e CNN.

VI. CONCLUSÃO

O sistema sonar passivo é importante em diversas áreas da acústica submarina. Neste artigo, exploramos o uso de modelos MLP e CNN para classificar sinais provenientes de um sistema de Sonar Passivo fornecidos pelo IpqM.

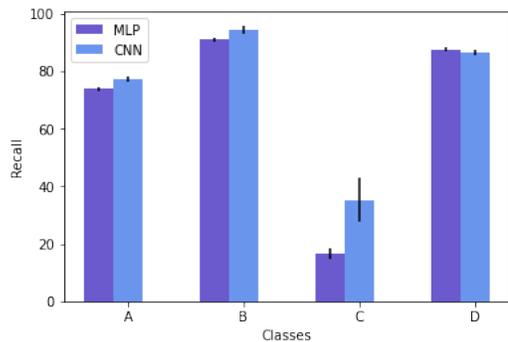


Figura 8. Gráfico de comparação do recall dos modelos MLP e CNN.

Para a avaliação dos resultados foi utilizado o índice SP e outras métricas de avaliação. Através dessas figuras de méritos, pode-se observar que quanto mais plásticos os classificadores, melhor é a classificação de novos sinais. Constatou-se também que o modelo CNN apresentou resultado superior para classificação dos sinais se comparado com o modelo MLP apesar de ambos os modelos apresentarem baixo desempenho na classificação da Classe C. Foi observado também um desempenho superior nos classificadores que utilizam o domínio da frequência, apesar de algumas vantagens operacionais que o domínio do tempo proporciona. Por fim, a proposta de classificação de sinais utilizando a Análise no Domínio do Tempo para sonar passivo mostrou um desempenho satisfatório.

Como trabalho futuro seria interessante a comparação desses resultados com um classificador baseado na arquitetura *Long short-term memory* (LSTM), o qual é amplamente difundido no meio científico na resolução de problemas que envolvem predições de séries temporais.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. Agradeço ao Programa de Engenharia Elétrica da COPPE/UFRJ, a FAPERJ e ao CNPq pela oportunidade de realizar esta pesquisa, bem como, ao IPqM (Instituto de Pesquisas da Marinha) pela disponibilização da base de dados.

REFERÊNCIAS

- [1] M. Ainslie, *Principles of Sonar Performance Modelling*. Springer-Verlag Berlin Heidelberg, 2010.
- [2] Q. Li, *Digital sonar design in underwater acoustics: principles and applications*. Springer Science & Business Media, 2012.
- [3] Goltz, Gustavo, *Classificação de Navios com Uso de Sinais de Sonar Passivo*. Tese de Doutorado, UFRJ/COPPE, Rio de Janeiro, RJ, 2019.
- [4] Vinicius dos Santos Mello et al. *Novelty Detection in Passive Sonar Systems using Stacked AutoEncoders*. 2018 International Joint Conference on Neural Networks (IJCNN) (2018): 1-7.
- [5] J. C. V. Fernandes, J. M. Seixas, M. Júnior, *Análise de modelos adversariais aplicados à classificação de sinais de Sonar Passivo*. XXXVII SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES E PROCESSAMENTO DE SINAIS - SBt2019, 29/09/2019–02/10/2019, PETRÓPOLIS, RJ.

- [6] Souza Filho, J. B. O., "Classificação Neural de Sinais de Sonar Passivo [Rio de Janeiro] 2007", Tese de doutorado, Programa de Engenharia Elétrica, Universidade Federal do Rio de Janeiro, 2007.
- [7] N. N. de Moura Junior, Detecção de Novidades para sistemas de sonar passivo. Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2018.
- [8] M. Júnior, J. M. Pires, Redes neurais dinâmicas para predição e modelagem não-linear de séries temporais. Dissertação de Mestrado, UFC, Ceará, CE, Brasil, 2006.
- [9] Barino, F. O, Santos, A. B., Rede Neural Convolutacional 1D aplicada à previsão da vazão no Rio Madeira. XXXVIII SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES E PROCESSAMENTO DE SINAIS - SBRt 2020, 22–25 DE NOVEMBRO DE 2020, FLORIANÓPOLIS, SC.
- [10] J. B. O. S. Filho and J. M. de Seixas, "Class-modular multi-layer perceptron networks for supporting passive sonar signal classification," IET Radar, Sonar Navigation, 2016.
- [11] N. N. de Moura Junior, J. M. Seixas, Avanços na Detecção de Novidade em Sinais de Sonar Passivo Utilizando Máquinas de Vector Suporte. XII Encontro de Tecnologia em Acústica Submarina, UFRJ, 2016.
- [12] J. M. Seixas, D. O. Damazio, P. S. R. Diniz, and W. Soares-Fillho, "Wavelet transform as a preprocessing method for neural classification of passive sonar signals," ICECS 2001. 8th IEEE International Conference on Electronics, Circuits and Systems (Cat. No.01EX483), 2001.
- [13] Villanueva, W.J.P., Comitê de máquinas em previsão em séries temporais. Mestrado em engenharia elétrica e computação; Faculdade de Engenharia Elétrica e de Computação; Campinas - São Paulo, 2006.
- [14] G. Dorffner (1996). "Neural networks for time series processing:" em Neural network world, 6, 447-468.
- [15] SILVA, Ivan Nunes da; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. Redes neurais artificiais para engenharia e ciências aplicadas. [S.l: s.n.], 2010.
- [16] GOODFELLOW, I., BENGIO, Y., COURVILLE, A. Deep Learning. MIT Press, 2016.
- [17] Guo, T., Dong, J., Li, H., & Gao, Y. (2017, March). Simple convolutional neural network on image classification. In 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(pp. 721-724). IEEE.
- [18] Q. Tao, F. Liu, Y. Li, & D. Sidorov (2019). "Air Pollution Forecasting Using a Deep Learning Model Based on 1D Convnets and Bidirectional GRU:" IEEE Access, 7, 76690-76698.
- [19] Efron, B., & Hastie, T. (2016). Computer age statistical inference: algorithms, evidence, and data science New York: Cambridge University Press.
- [20] Data Science (Institute of Mathematical Statistics Monographs). Cambridge University Press. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.
- [21] Moura, N. N. D. et al. "Passive Sonar Signal Detection and Classification Based on Independent Component Analysis." (2011).