

# Classificação de sinais de Sonar Passivo utilizando *Stacked AutoEncoders*

Natanael Nunes de Moura Júnior<sup>1</sup>, Vinícius dos Santos Mello<sup>1</sup>, Lucas de  
Andrade Cerqueira<sup>1</sup> e José Manoel de Seixas<sup>1</sup>  
natmourajr@lps.ufrj.br, viniciusdsmello@poli.ufrj.br,  
lucas.cerqueira@poli.ufrj.br e seixas@lps.ufrj.br

<sup>1</sup>Laboratório de Processamento de Sinais, Universidade Federal do Rio de Janeiro,  
Av. Horácio Macedo 2030, 21941-914, Rio de Janeiro, Brasil

**Resumo** Em aplicações militares modernas, o ruído irradiado pelos navios de interesse tem se tornado cada vez menos intenso e, devido a isso, tarefas como detecção, acompanhamento e classificação, têm se tornado cada vez mais desafiadoras. Assim sendo, sistemas de apoio a decisão vem ganhando importância nesta área. De maneira geral, sistemas de apoio a decisão são treinados baseados nos dados disponíveis e, no caso de sistemas de sonar passivo, os dados são não-linearmente separáveis e possuem distribuição não-gaussiana. Neste tipo de dados, técnicas de aprendizado de máquina sofisticadas podem ser utilizadas para a tomada de decisão em níveis de abstração mais altos. Dentre estas técnicas, pode-se destacar as técnicas de *Deep Learning*. Uma das técnicas mais aplicadas de *Deep Learning* corresponde aos chamados *Stacked AutoEncoders* (SAE), que fazem o uso da informação dos dados para a inicialização dos pesos de uma rede neural profunda. Este trabalho propõe a utilização de um SAE para a classificação de sinais de sonar passivo.

**Keywords:** Sistemas de Sonar Passivo, Aprendizado de Máquina, Classificação, *Deep Learning*

## 1 Introdução

Em aplicações militares, comumente, se tem a utilização de sistemas especializados em processamento de sinais, como por exemplo: sistemas de radar e sistemas de sonar [1]. Como ondas eletromagnéticas são altamente atenuadas em meios aquáticos, aplicações navais submarinas utilizam ondas mecânicas, como o som, para sua operação [2].

Operações navais, tanto civis como militares, devem acessar informações sobre o ambiente ao redor como topologia do leito do mar, fenômenos sísmicos, movimentação de cardumes de peixes e outros. Como o som pode alcançar grandes distâncias no mar, geralmente, este é utilizado para os mais diversos fins e, devido a isso, sistemas específicos foram desenvolvidos para a análise, os chamados sistemas Sonar (**SO**und **N**avigation **AND** **R**anging).

Sistemas de sonar, geralmente, são divididos em sistemas de sonar ativo e sistemas de sonar passivo. No primeiro, ocorre a emissão de uma sinal conhecido e, a partir do seu eco, o processamento é realizado. No segundo, todo o processamento dos sinais é baseado na emissão dos navios de interesse.

Como navios, animais marinhos e fenômenos sísmicos submersos podem emitir sinais com assinaturas sonoras similares, a sua classificação se torna uma tarefa desafiadora. Assim sendo, a tomada de decisão sobre qual tipo de sinal foi adquirido pelo sistema pode ser feita por classificadores. Sistemas de classificação baseados em aprendizado de máquina têm sido utilizados para a melhora da eficiência de classificação em sistemas de sonar passivo.

Atualmente, algumas aplicações têm se baseado em técnicas de aprendizado de máquina e processamento de sinais para sistemas de sonar passivo. Em [3] temos a aplicação de redes neurais artificiais para a classificação de sinais sonar passivo. Em [4], técnicas de classificação baseadas em classificadores especialistas, do tipo classe-não-classe, foram aplicadas para sinais de sonar passivo. Uma outra forma de visualização de dados amplamente utilizada é a transformada Wavelet [5] e em [6] é apresentada a sua aplicação à classificação de sinais de sonar passivo. Além das técnicas descritas anteriormente, destacam-se também os artigos como [7] que apresentam outras implementações para classificação em sinais de sonar passivo. Em [7], a informação utilizada para o treinamento do classificador foi extraída utilizando os chamados *mel-frequency cepstrum* (MFC). Os MFC são uma representação do espectro de potência em uma curta janela de aquisição de dados. Para a sua extração, devemos fazer uma transformação cossenoidal do logaritmo do espectro de potência na escala mel (*mel scale*). Essa representação tem sido utilizada com sucesso para aplicações em áudio como reconhecimento de fala, extração de características sonoras e classificação de falantes [8], [9] e [10]. Em [11], uma técnica de *Deep Learning* chamada *Deep Belief Network* foi utilizada para fazer a classificação de sinais de sonar passivo. Neste trabalho, os dados utilizados foram processados com análises diferentes das utilizadas anteriormente e, as mesmas se mostraram promissoras para a aplicação em um ambiente submarino.

O objeto de estudo deste trabalho é um sistema de sonar passivo implementado pela Marinha do Brasil e a proposta é a utilização de classificadores baseados em SAE para a classificação de sinais de navios disponibilizados. Os resultados experimentais serão comparados com resultados de classificação de redes neurais e figuras de mérito específicas serão utilizadas para avaliar o desempenho de cada um dos modelos.

O texto se organiza da seguinte maneira: na seção 2 temos a descrição de um sistema de sonar passivo bem como a descrição detalhada da análise LOFAR, uma das análises mais utilizadas em sistemas de sonar passivo. Na seção 3 os *Stacked AutoEncoders* são apresentados e na seção 4 o método para obtenção dos resultados, bem como a descrição dos dados e das figuras de mérito são expostos. Por fim, os resultados experimentais são apresentados na seção 5, que é seguida dos agradecimentos e das referências bibliográficas.

## 2 Sistemas de Sonar Passivo

Os sinais de sonar passivo podem ser vistos como diversos sinais que se desenvolvem ao longo do tempo e podem ser modelados da seguinte maneira: os  $n$  sinais, representados por  $s_1, s_2, \dots, s_n$  são adquiridos por  $n$  hidrofones que compõem o sistema. Uma vez que os sinais são adquiridos, técnicas de conformação de feixe são aplicadas aos mesmos e os feixes são conformados. O processo de conformação de feixes pode ser visto como sendo uma filtragem espacial dos sinais provenientes dos transdutores para a obtenção das chamadas direções de aproximação, ou *DOA*.

Uma vez que os sinais são conformados, os mesmos podem ser observados nos domínios do tempo e da frequência. Os dois domínios podem ser utilizados para fazer a detecção dos sinais, e assim que uma das possíveis direções apresentar um ruído de interesse, o seu acompanhamento pode ser realizado. Com os contatos sendo acompanhados, o processo de classificação é iniciado. A classificação dos sinais se inicia com duas análises primárias: a análise DEMON (*DEtection MOdulation on Noise*) e a análise LOFAR (*LOw Frequency Analysis and Recording*). A primeira é uma análise de banda estreita que tem por objetivo a estimação da quantidade de pás e eixos que compõem o sistema de propulsão do contato de interesse que se encontra em uma determinada direção. A segunda será apresentada em detalhes abaixo.

### 2.1 Análise LOFAR

A análise LOFAR [12], que tem o seu diagrama em blocos mostrado na Figura 1, pode ser definida como uma análise de banda larga. Os sinais de uma determinada DOA são digitalizados por um ADC, neste caso específico, com 8 bits e frequência de amostragem de 22.050 Hz (período de amostragem de 45,3  $\mu$ s). A análise LOFAR consiste em um processo de decimação por 3 associado a um filtro FIR de décima ordem reduzindo assim a frequência de amostragem a 7,35 kHz [13]. Depois da decimação, uma janela de Hanning com 1.024 pontos é aplicada a blocos consecutivos de sinal sem *overlap*. Por fim, a informação no domínio das frequências é obtida após a aplicação de uma transformada de Fourier de tempo curto (STFT) e de uma modulação, fazendo assim com que o sinal se desenvolva de frequências próximas a DC até 3,67 kHz.

Para cada janela de sinal processada, os primeiros 400 *bins* de frequência são retidos, seguindo o método desenvolvido em [3]. Depois dessa filtragem espectral, o sinal processado contém frequências de 0 à 2,87 kHz, o que corresponde a uma janela de tempo de 139  $\mu$ s. Ao fim da cadeia de processamento, um algoritmo de estimação do ruído de fundo é aplicado visando a redução do mesmo e, o algoritmo escolhido foi o TPSW (*Two-Pass Split Window*) obtido de [2].

Depois de passar pela análise LOFAR, o sinal pode ser visto pelo chamado LOFARgrama (Figura 2). Neste gráfico, como pode ser observado, temos no eixo  $x$ , as frequências discretas e, no eixo  $y$  temos o tempo. Além disso, este gráfico é, geralmente, visto como um *waterfall display*, isto é, a medida que o tempo avança, mais e mais janelas são processadas, o que faz com que o mesmo seja



Figura 1: Diagrama em blocos da Análise LOFAR.

atualizado fazendo com que mais linhas sejam adicionadas a parte superior do gráfico, fazendo com que o mesmo se comporte como uma cascata.

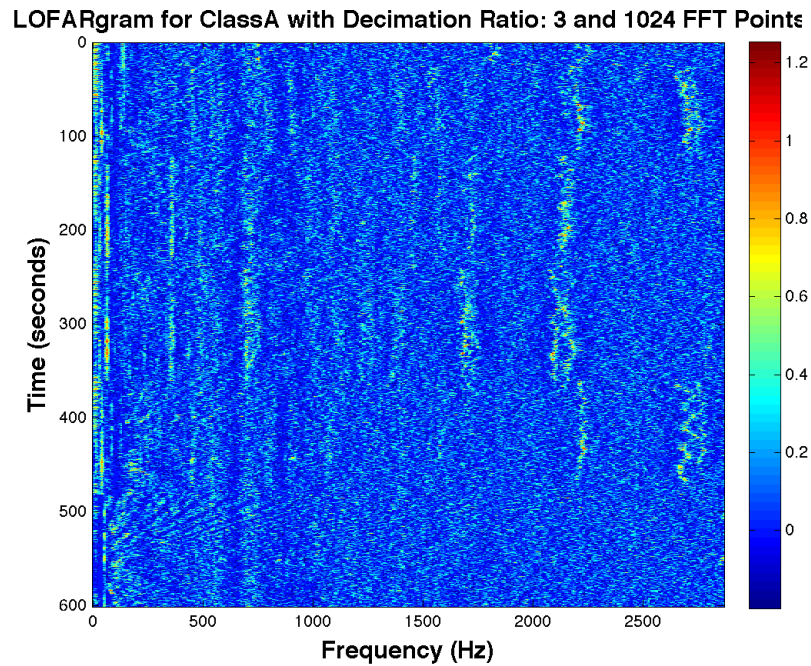


Figura 2: Gráfico de Frequência vs Tempo.

### 3 *Stacked AutoEncoders*

Redes neurais são estruturas compostas por unidades de processamento não-lineares (neurônios) totalmente conectadas baseadas em redes neurais biológicas. Redes neurais profundas são modelos baseados em redes neurais, embora apresentem uma maior quantidade de camadas escondidas [14]. Quando se trata de redes neurais profundas, pode-se encontrar um problema durante o processo de

treinamento devido à estrutura do modelo. Isto é, à medida que um treinamento, geralmente baseado em *backpropagation*, é realizado, as camadas escondidas mais próximas à saída são intensamente atualizadas enquanto as camadas mais próximas à entrada são menos atualizadas. Devido a este fato, pode ocorrer uma degradação na generalização do modelo.

Com o surgimento de técnicas de treinamento para redes neurais profundas, o chamado *Deep Learning* (DL), o problema de perda de generalização foi contornado. Uma das técnicas de DL mais utilizadas são os chamados *Stacked AutoEncoders* (SAE). Os SAE são uma maneira de se inicializar os pesos de uma rede neural profunda visando representar os dados nas camadas mais próximas à entrada como funções não-lineares. Assim sendo, a tomada de decisão é feita em níveis de abstração mais elevados e, assim, a classificação se torna mais eficiente.

Para se extrair os parâmetros desta representação, uma estrutura de *AutoEncoder* foi implementada. Nesta estrutura, o treinamento é realizado visando minimizar o erro de reconstrução dos dados de entradas após os mesmos passarem por uma rede neural com uma única camada escondida de funções não-lineares. A primeira camada de pesos da rede em questão é chamada de camada de codificação (*encoder*) enquanto que a segunda camada de neurônios é chamada de camada de decodificação (*decoder*). Na Figura 3 podemos ver um exemplo da estrutura de um *AutoEncoder* representada.

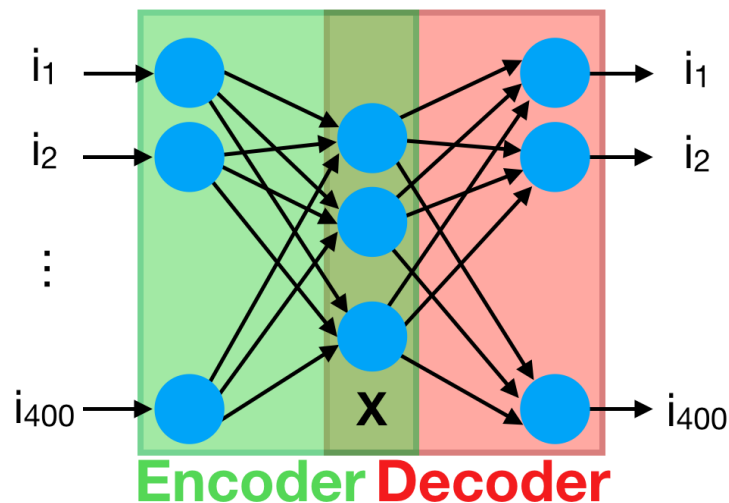


Figura 3: Exemplo de *AutoEncoders*.

Uma vez que as camadas de representação foram treinadas, as camadas de codificação das mesmas podem ser empilhadas para formar o SAE. Com o SAE formado, um processo de *fine tuning* é realizado visando a minimização do erro de classificação. O *fine tuning* é um treinamento da rede neural profunda ba-

seado, comumente, em *backpropagation*. Como as camadas mais próximas à entrada são pouco modificadas, a representação do dado, extraída anteriormente, é preservada.

## 4 Método

Os resultados foram obtidos a partir do processo de validação cruzada *jackknife*. Neste caso, se  $k$ -*folds* forem selecionados, o conjunto de dados é dividido aleatoriamente em  $k$  sub-conjuntos. A cada rodada de treinamentos,  $k - 1$  conjuntos são utilizados para a extração de características e 1 conjunto avalia a generalização do treinamento. Ao final, a estimativa da incerteza do processo pode ser feita utilizando a média e a variância dos resultados obtidos com os *folds*.

Para realizar a classificação dos sinais de sonar passivo, o diagrama de blocos da Figura 4 foi utilizado. Uma vez que os dados, no domínio do tempo, são obtidos, os mesmos passam pela análise LOFAR que os transforma do domínio do tempo para o domínio da frequência. Uma vez no domínio das frequências, os dados passam por uma etapa de pré-treinamento que realiza o treinamento dos *AutoEncoders*.

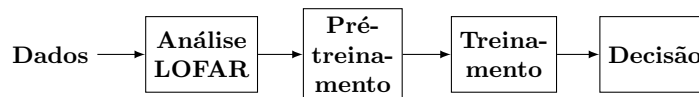


Figura 4: Diagrama em blocos para classificação de sinais de sonar passivo.

Nesta etapa, os parâmetros iniciais dos *AutoEncoders* são extraídos, conforme a Figura 5 e após esta extração inicial, um treinamento *backpropagation* clássico é feito, no chamado *fine tuning*. Enquanto o pré-treino tem como função objetivo o erro médio quadrático entre os dados de entradas e suas projeções, ou seja, visa a melhor representação do dado, o *fine tuning* tem por objetivo treinar os pesos para maximização da eficiência de classificação. A etapa de pré-treino é feita camada a camada enquanto o *fine tuning* opera em uma rede neural profunda, neste caso com 3 camadas escondidas.

Por fim, a tomada de decisão é feita na camada de saída. Como os alvos utilizados foram alvos maximamente esparsos, isto é, com uma dimensão dedicada a cada classe, a saída é definida com base no neurônio de maior saída (*winner takes all*).

### 4.1 Figuras de Mérito

Diversas medidas de desempenho são utilizadas para avaliar a eficiência de métodos de classificação. Em casos mais simples, onde a classificação se dá em

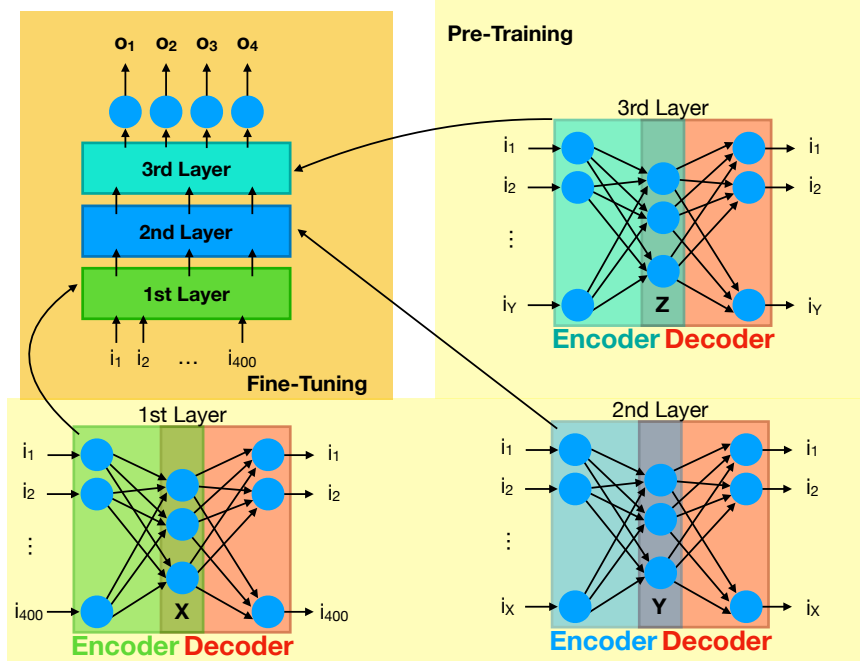


Figura 5: Estruturação do treinamento dos *Stacked AutoEncoders*.

duas classes, a curva ROC (*receiver operating characteristic*) [15] pode ser utilizada para a escolha do limiar de decisão entre as duas classes. Já para um caso mais complexo, como o caso de classificação para 4 classes, as figuras de mérito clássicas de classificação devem sofrer algumas modificações para serem aplicadas.

A eficiência individual de cada uma das classes envolvidas na classificação pode ser obtida pela equação 1. O parâmetro  $\theta$  define o limiar de classificação.

$$Eff_i(\theta) = 100\% \frac{\# \text{Eventos da classe } i \text{ classificados como classe } i}{\# \text{Total de eventos da classe } i} \quad (1)$$

Assim sendo, a primeira figura de mérito utilizada no trabalho foi a acurácia de classificação, mostrada na equação 2, onde  $N_{class}$  representa o número de classes. Esta figura de mérito, que nada mais é do que a média aritmética das eficiências de cada uma das classes envolvidas, tem sido bastante utilizada para classificação, embora seja pouco sensível a grandes variações nas eficiências individuais de cada classe.

$$Acc(\theta) = \frac{\sum_{i=1}^{N_{class}} Eff_i(\theta)}{N_{class}} \quad (2)$$

A segunda figura de mérito utilizada neste trabalho foi o índice SP [16]. Este índice foi desenvolvido para substituir a curva ROC para aplicações de classificação com mais de duas classes e o mesmo busca equalizar as eficiências individuais de cada uma das classes. Como vemos na equação 3, se uma das classes obtiver uma eficiência baixa, o valor obtido do índice SP será baixo, já se todas as classes obtiverem eficiências individuais altas, o índice SP obtido será alto.

$$SP(\theta) = \sqrt[2]{\left(\frac{\sum_{i=1}^{N_{class}} Eff_i(\theta)}{N_{class}}\right)} \cdot \left(N_{class} \sqrt[N_{class}]{\prod_{i=1}^{N_{class}} Eff_i(\theta)}\right) \quad (3)$$

## 5 Resultados Experimentais

Utilizando o método disposto anterior, os resultados experimentais foram obtidos com base nos dados disponibilizados pela Marinha do Brasil.

Os dados deste trabalho correspondem a ruídos irradiados por diferentes navios que foram adquiridos pela Marinha do Brasil em diferentes cenários. O conjunto de dados utilizado é composto por 4 classes de navios que foram adquiridas na raia acústica de Arraial do Cabo. Os experimentos na raia acústica consistem em navios se movendo em um ambiente controlado pela MB com 1500 metros de comprimento.

Geralmente, as gravações foram feitas com um único hidrofone que é acoplado a um sistema de aquisição que possui frequência de amostragem de 22.050 Hz e 8 bits de resolução. A gravação do experimento é feita quando o navio se encontra a 1000 metros antes da posição do hidrofone e termina quando o navio se encontra a 500 metros depois da posição do hidrofone. A profundidade do hidrofone é de 45 metros.

A estimação da incerteza estatística foi feita utilizando o processo de validação cruzada detalhado anteriormente. O número de *folds* utilizado foi de 10, o que significa que a cada rodada, um dos *folds* foi utilizado para teste e os outros 9 foram utilizados para a extração de características e treinamento dos modelos.

Para comparação dos resultados, redes neurais artificiais MLP foram treinadas para classificação das 4 classes utilizando como referências os resultados obtidos em [3]. Na tabela 1 são apresentados os resultados de classificação para redes neurais MLP.

Já para redes neurais profundas, os resultados foram obtidos com o aumento gradativo da complexidade das redes. Inicialmente, uma rede neural com uma camada escondida foi inicializada com os pesos obtidos através do treinamento de um *AutoEncoder*, após a extração dos pesos, uma rede neural MLP de uma camada foi treinada e os resultados encontram-se na tabela 2 (*1st Layer*). Nesta etapa, a camada de codificação utilizada possuía 300 neurônios. Em seguida,



Tabela 1: Resultados de classificação para redes neurais do tipo MLP.

	Eff.	Acc.	SP
Class 1	83,51% ± 5,75%		
Class 2	85,42% ± 3,41%	84,29% ±	84,27% ±
Class 3	86,78% ± 4,32%	4,56%	4,52%
Class 4	81,45% ± 4,76%		

uma outra camada escondida foi adicionada ao modelo, fazendo com que um outro *AutoEncoder* fosse treinado, desta vez com dimensão de codificação de 100 e com os pesos de representação extraídos. Um treinamento de classificação foi implementado obtendo a coluna *2nd Layer* na tabela 2. Por fim, um modelo com 3 camadas escondidas foi desenvolvido, seguindo os mesmos princípios dos modelos mais simples (*3rd Layer*), com 50 neurônios na camada de codificação.

Como pode ser visto, a maioria das redes neurais profundas obteve resultados melhores do que das redes neurais. No caso de classificadores baseados em redes neurais com 2 camadas escondidas, os resultados foram equivalentes aos obtidos com a rede neural padrão (no caso da acurácia e do índice SP). O mapeamento extraído, pode ser um mapeamento não-linear, pode ter gerado características confundidoras e devido a isso, diminuir a eficiência de classificação.

Além dos resultados quantitativos, um outro resultado interessante é mostrado na Figura 6. Nesta figura temos um LOFARgrama para a Classe 2 e o seu LOFARgrama após a projeção na primeira camada de um SAE. Como pode ser visto, depois do treinamento, o SAE evidenciou as frequências mais intensas e reduziu frequências com ruídos de fundo. Este resultado comprova a capacidade do modelo em aprender padrão que conduzem a uma maior eficiência de classificação.

## 6 Conclusões

Neste trabalho foi desenvolvido um classificador de sinais de sonar passivo de forma eficiente baseado em redes neurais profundas com inicialização dos pesos por *Stacked AutoEncoders*. Os dados experimentais foram processados pela

Tabela 2: Resultados de classificação para redes neurais profundas inicializadas com diferentes estruturas de SAE.

	1st Layer			2nd Layer			3rd Layer		
	Eff.	Acc.	SP	Eff.	Acc.	SP	Eff.	Acc.	SP
Class 1	86,71% ± 4,75%			85,34% ± 6,15%			90,12% ± 1,15%		
Class 2	87,12% ± 3,41%	86,39% ±	86,39% ±	84,10% ± 2,41%	82,07% ±	82,04% ±	91,05% ± 3,81%	88,98% ±	88,92% ±
Class 3	87,11% ± 3,32%	4,15%	4,11%	77,11% ± 3,32%	4,05%	3,93%	81,45% ± 2,98%	3,23%	3,03%
Class 4	84,65% ± 5,13%			81,74% ± 4,34%			93,31% ± 5,01%		

análise LOFAR e os resultados foram obtidos com diferentes figuras de mérito especialmente desenvolvidas para a classificação para mais de duas classes.

Nos resultados mostrados nas tabelas 1 e 2, podemos ver que os resultados melhoraram sensivelmente com a aplicação de redes neurais profundas. Os dados foram compactados de 400 dimensões na camada de entrada para 300 dimensões na primeira camada de codificação, 100 dimensões na segunda camada de codificação e 50 dimensões na terceira camada de codificação. A medida que os classificadores são formados por mais camadas não-lineares, maior o poder de processamento e assim, maior a eficiência de classificação. Esse comportamento possui uma exceção para o caso de uma rede neural profunda com 2 camadas inicializadas com os SAE.

Na Figura 6, podemos ver também um resultado qualitativo do treinamento de um SAE. Neste resultado, as frequências que continham alguma informação relevante para o modelo tiveram os seu valores acentuados enquanto frequências com ruído de fundo foram atenuadas. Este resultado pode ser de grande auxílio durante a operação de sistemas de sonar passivo, facilitando a visualização do operador sonar no que tange o LOFARgrama.

Como trabalhos futuros podem ser citados a análise de outras estruturas de *Deep Learning* para a classificação de sinais de sonar passivo. Por exemplo, as redes neurais convolucionais (CNN) que são modelos que dependem fortemente da informação espacial dos dados para a classificação. Os sinais de sonar passivo se desenvolvem ao longo do tempo de maneira dependente, ou seja, a estrutura temporal dos dados pode ser facilmente acessada com uma convolução no tempo. Além das CNN, um outro modelo que pode ser explorado são as redes neurais recursivas (RNN), que são modelos que acessam a informação temporal diretamente durante o seu treinamento.

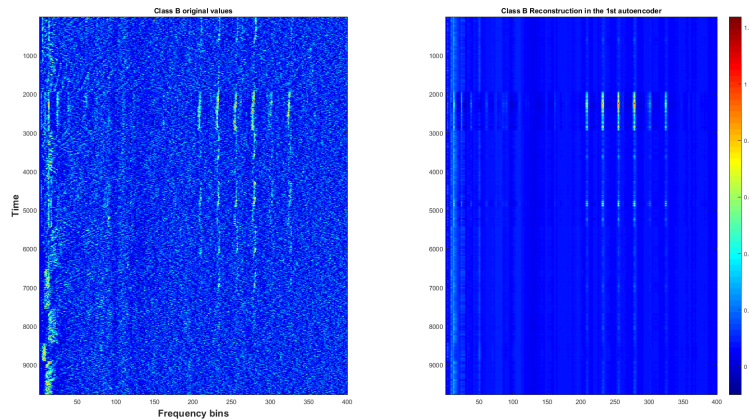


Figura 6: Reconstrução de um SAE: na esquerda é apresentada uma figura com o LOFARgrama original dos dados para a classe 2, enquanto que na direita, o LOFARgrama dos dados reconstruídos com uma camada de SAE foi apresentado para a mesma classe

## Agradecimentos

Os autores gostariam de agradecer à CAPES, CNPq e FAPERJ pelo suporte a este trabalho. Gostariam de agradecer também ao Instituto de Pesquisas da Marinha por disponibilizar o conjunto de dados, bem como à Natanael Nunes de Moura e Thiago Ciodaro pela colaboração com o trabalho.

## Referências

1. M. Ainslie, *Principles of Sonar Performance Modelling*. Springer-Verlag Berlin Heidelberg, 2010.
2. R. O. Nielsen, *Sonar Signal Processing*. Norwood, MA, USA: Artech House, Inc., 1991.
3. W. Soares-Filho, J. M. de Seixas, and L. P. Caloba, “Principal component analysis for classifying passive sonar signals,” in *ISCAS 2001. The 2001 IEEE International Symposium on Circuits and Systems (Cat. No.01CH37196)*, vol. 3, May 2001, pp. 592–595 vol. 2.
4. 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.
5. D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press, 2006.
6. J. M. Seixas, D. O. Damazio, P. S. R. Diniz, and W. Soares-Filho, “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.

7. A. Das, A. Kumar, and R. Bahl, "Marine vessel classification based on passive sonar data: the cepstrum-based approach," *IET Radar, Sonar Navigation*, 2013.
8. D. Chazan, R. Hoory, G. Cohen, and M. Zibulski, "Speech reconstruction from mel frequency cepstral coefficients and pitch frequency," in *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100)*, 2000.
9. X. Zhou, D. Garcia-Romero, R. Duraiswami, C. Espy-Wilson, and S. Shamma, "Linear versus mel frequency cepstral coefficients for speaker recognition," in *2011 IEEE Workshop on Automatic Speech Recognition Understanding*, 2011.
10. V. Tyagi, I. McCowan, H. Misra, and H. Bourlard, "Mel-cepstrum modulation spectrum (mcms) features for robust asr," in *2003 IEEE Workshop on Automatic Speech Recognition and Understanding (IEEE Cat. No.03EX721)*, 2003.
11. S. Kamal, S. K. Mohammed, P. R. S. Pillai, and M. H. Supriya, "Deep learning architectures for underwater target recognition," *2013 Ocean Electronics (SYM-POL)*, 2013.
12. Q. Li, *Digital Sonar Design in Underwater Acoustics*. Springer-Verlag Berlin Heidelberg, 2012.
13. P. S. R. Diniz, E. A. B. da Silva, and S. L. Netto, *Digital Signal Processing: System Analysis and Design*. Cambridge University Press, 2010.
14. Q. Xu, C. Zhang, L. Zhang, and Y. Song, "The learning effect of different hidden layers stacked autoencoder," in *2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2016.
15. H. L. V. Trees, *Detection, Estimation, and Modulation Theory: Radar-Sonar Signal Processing and Gaussian Signals in Noise*. Melbourne, FL, USA: Krieger Publishing Co., Inc., 1992.
16. A. dos Anjos, R. Torres, J. Seixas, B. Ferreira, and T. Xavier, "Neural triggering system operating on high resolution calorimetry information," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 559, no. 1, pp. 134 – 138, 2006, proceedings of the X International Workshop on Advanced Computing and Analysis Techniques in Physics Research ACAT 05X International Workshop on Advanced Computing and Analysis Techniques. [Online]. Available: [//www.sciencedirect.com/science/article/pii/S0168900205022552](http://www.sciencedirect.com/science/article/pii/S0168900205022552)