

# Detecção Hierárquica de Classes Desconhecidas em Sonar por “Autoencoders” Convolucionais

Eduardo Sperle Honorato, Victor Hugo da Silva Muniz e João Baptista de Oliveira e Souza Filho

**Resumo**—As ondas acústicas captadas por sistemas de sonar passivo são analisadas por operadores humanos para a identificação de possíveis ameaças no ambiente submarino. Sistemas Automáticos de Classificação auxiliam o trabalho deste profissional, porém exigem mecanismos que lidem com a presença de classes desconhecidas. Este artigo propõe a utilização de um comitê hierárquico de *autoencoders* convolucionais para a construção destes sistemas, como uma alternativa mais robusta ao algoritmo dos  $k$ -vizinhos mais próximos, que representa o estado-da-arte para este problema. Dados reais pertencentes a 8 classes de navios em diferentes condições operacionais foram avaliados. Resultados sinalizam um desempenho competitivo da técnica proposta.

**Palavras-Chave**—Autoencoder, Redes Neurais Convolucionais, Aprendizado Profundo, Detecção de Novidades, Sistema de Sonar Passivo, Aprendizado de Máquina.

**Abstract**—Acoustic waves captured by passive sonar systems are analyzed by human operators, aiming to identify possible threats in the subsea environment. Automatic Classification Systems can aid in the work of this professional, however they require mechanisms to deal with the presence of unknown classes. This article proposes the use of a hierarchical committee of convolutional autoencoders to build such systems, as a more robust alternative to the  $k$ -nearest neighbors algorithm, which represents the state-of-the-art in this problem. Real data belonging to 8 classes of ships under different operational conditions were evaluated. Results signalize a competitive performance of the proposed technique.

**Keywords**—Autoencoder, Convolutional Neural Networks, Deep Learning, Novelty Detection, Passive Sonar System, Machine Learning.

## I. INTRODUÇÃO

Sistemas de sonar passivo são uma tecnologia presente em embarcações sigilosas de caráter militar. Tais sistemas, formados por um arranjo de hidrofones, captam as vibrações emitidas por outras embarcações, sem emitir qualquer sinal próprio durante este processo. De posse da assinatura acústica destas embarcações, operadores treinados classificam este contato como uma ameaça ou não.

A utilização de Sistemas Automáticos de Classificação (SAC) é de grande valia para o processo de tomada de decisão do operador de sonar, visto que esta ferramenta alivia o esforço de vigilância exigido na tarefa, ao passo que aumenta a confiabilidade desta ação. Idealmente, um SAC deve integrar informações sobre todas as classes de navios existentes, a fim de sempre obter êxito nas classificações. Entretanto, dado que

o número de classes disponíveis é limitado, é de suma importância que este sistema seja capaz de reconhecer novidades, neste caso, identificar novas classes de navios.

A literatura reporta resultados promissores quanto à detecção de novidades em aplicações de diferentes domínios [1], ao considerar técnicas de Aprendizado de Máquina baseadas em estratégias variadas, como por exemplo: reconstrução [2], distância [3] e descrição do domínio [4]. Tipicamente, o tratamento deste problema através de redes neurais adota uma estratégia baseada em reconstrução. Tal abordagem resulta em modelos mais robustos, desde que sejam treinados com extensas bases de dados, requisito necessário devido ao número elevado de parâmetros envolvidos. Dentro do contexto da detecção de novidades em sinais de sonar passivo, redes neurais do tipo *stacked autoencoder* foram propostas em estudos anteriores [5], produzindo sistemas de desempenho interessante.

Por se tratar de um problema multiclasse, a utilização de algoritmos baseados na estratégia “dividir para conquistar” costuma resultar em sistemas de melhor desempenho. Tais sistemas, que incluem múltiplos modelos classe-especialistas, frequentemente apresentam um melhor desempenho do que sistemas não-especializados [6]. Múltiplas estratégias e formas de fusão de modelos podem ser exploradas para este fim. Dentre elas, a solução hierárquica está associada aos melhores resultados reportados em [7]. Diante da excelência das redes convolucionais em várias aplicações, um aspecto relevante a se investigar é se uma arquitetura hierárquica de *autoencoders* convolucionais poderia beneficiar o problema em tela.

Este trabalho propõe a construção de um comitê hierárquico de *autoencoders* convolucionais para o problema da identificação de classes desconhecidas em sinais de sonar passivo. Tal proposta será comparada tanto com a adoção de um único *autoencoder*, mais típica na literatura, assim como com um sistema hierárquico desenvolvido por meio do algoritmo dos  $k$ -vizinhos mais próximos, o qual configura o atual estado-da-arte para este problema [7]. Adicionalmente, este estudo considera cenários de avaliação de variada complexidade quanto à tarefa de detecção de novidades, definidos pelo número de classes conhecidas assumidas durante o desenvolvimento do modelo.

A organização do artigo é a seguinte: na Seção II, é apresentada a estrutura adotada para a construção do sistema detector de novidades, juntamente com uma descrição dos algoritmos avaliados e das medidas de desempenho empregadas. Nas Seções III e IV, a base de dados e os resultados são reportados, respectivamente. A Seção V apresenta uma breve discussão. E na Seção VI, têm-se as conclusões e os trabalhos futuros.

## II. SISTEMA DETECTOR DE NOVIDADES

O detector de novidades aqui considerado busca identificar se um dado arbitrário ( $\mathbf{x}$ ) é proveniente de uma classe distinta daquelas consideradas no seu desenvolvimento (classes conhecidas).

### A. Topologias

A Figura 1 ilustra o diagrama em blocos do sistema detector de novidades considerando um único detector, aqui referido como Topologia “não-especializada” (NE). Neste caso, o bloco gerador do escore produz uma variável contínua  $s$ , que quantifica a dissimilaridade entre  $\mathbf{x}$  e as classes conhecidas. Após, no bloco decisor, seu valor é comparado a um limiar interno  $\lambda$ , gerando assim a variável binária  $d \in \{0, 1\}$ , onde 0 sinaliza que  $\mathbf{x}$  representa um sinal conhecido ( $s < \lambda$ ), enquanto 1 indica uma novidade ( $s \geq \lambda$ ).

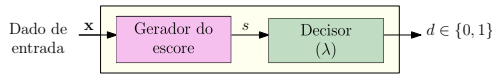


Fig. 1. Topologia “não-especializada”.

Sistemas detectores de novidades formados por múltiplos detectores classe-especialistas, isto é, constituídos por geradores de escore e decisores estritamente desenvolvidos para reconhecer as classes de sua especialidade e recusar as demais, devem apresentar um mecanismo específico para a fusão dos resultados fornecidos por cada detector. Neste trabalho, duas topologias foram consideradas para esta tarefa: a Unânime (UN) e a Hierárquica (HI).

Na Topologia UN, ilustrada na Figura 2, o dado é avaliado individualmente por cada  $i$ -ésimo detector classe-especialista ( $1 \leq i \leq N$ ). Ao final, é considerado um E-lógico sobre todas as variáveis de decisão  $d_i$ . Portanto, um dado somente será considerado novidade se houver uma concordância unânime entre todos os detectores.

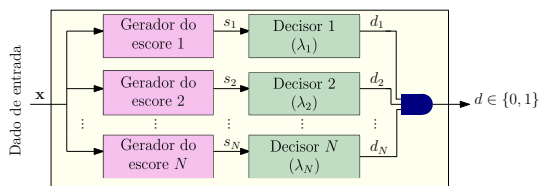


Fig. 2. Topologia Unânime.

Por sua vez, a Figura 3 exibe a Topologia HI. Nela, o dado sob análise é submetido a um bloco classificador, responsável por indicar a  $i$ -ésima classe mais provável de origem do sinal  $\mathbf{x}$  ( $1 \leq i \leq N$ ). Na sequência, apenas o  $i$ -ésimo detector classe-especialista é ativado, e sua respectiva saída  $d_i$  define a variável de decisão final  $d$ .

### B. Algoritmos

1) *Autoencoder Convolutacional (AEC)*: O autoencoder é um tipo de rede neural que busca codificar um dado de entrada sob uma representação de dimensionalidade reduzida, ao passo que

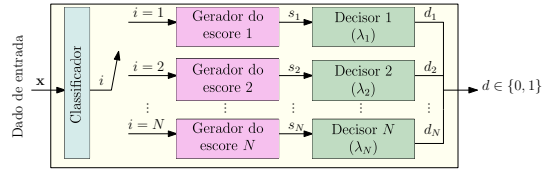


Fig. 3. Topologia Hierárquica.

utiliza esta mesma representação para gerar uma saída o mais similar possível à entrada original, por meio de mapeamentos a serem aprendidos dos dados. Em síntese, busca-se o menor erro de reconstrução possível [8]. Para a detecção de novidades por esta técnica, o escore de dissimilaridade mais comum é o erro quadrático de reconstrução do dado sob análise, visto que dados espúrios e de classes desconhecidas tendem a não ser adequadamente representados por tais mapeamentos, logo estariam associados a maiores erros [1].

Esta rede é formada por múltiplas camadas tipicamente simétricas, para uma maior simplicidade, as quais têm seu tamanho reduzido progressivamente até a camada latente. Tais dimensões são, de forma similar, aumentadas até a camada de saída, de mesma dimensão da entrada. Um AEC considera camadas convolucionais formadas por filtros convolucionais [9]. A Figura 4 ilustra um AEC arbitrário.

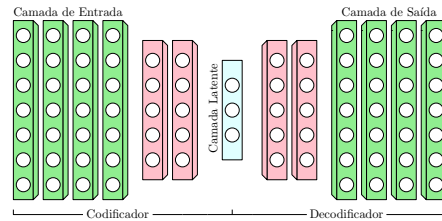


Fig. 4. Arquitetura de um Autoencoder Convolutacional.

Visto que a utilização de mecanismos de regularização costuma melhorar o desempenho dos AECs, neste trabalho foi adicionado um ruído gaussiano na entrada da rede durante o processo de treinamento para a redução de um possível *overfitting*. Esta estratégia, conhecida como *Denosing Autoencoder* [10], força uma reconstrução do dado original (sem ruído) através de sua versão ruidosa, o que leva a uma regularização da rede neural.

Quanto aos hiperparâmetros utilizados no modelo, têm-se: (i) o número de filtros em cada camada ( $n_f$ ); (ii) a função de ativação dos filtros de cada camada ( $a_f$ ); (iii) o comprimento dos filtros de cada camada ( $c_f$ ); e (iv) o desvio padrão do ruído gaussiano aplicado apenas na camada de entrada ( $\sigma$ ).

2) *k-vizinhos mais próximos (kNN)*: Consiste no atual estado-da-arte quanto ao problema da detecção de novidades em sinais de sonar passivo para a base de dados em análise [7]. Esta técnica se baseia em identificar os  $k$ -vizinhos mais próximos ao dado sob avaliação pertencentes ao conjunto de treinamento. Como escore de dissimilaridade, utiliza-se o valor da mediana da distância euclidiana entre o dado e os seus  $k$ -vizinhos.

### C. Índices de Desempenho

A sistemática para a produção e avaliação dos detectores é similar a [7]. De posse de uma base de dados  $\mathcal{D}$  constituída por diversas classes de navios, extraiu-se um conjunto de  $N$  classes:  $\mathcal{K} = \{K_1, K_2, K_i, \dots, K_N\}$ , a ser utilizado no treinamento do sistema. A taxa de detecção da  $i$ -ésima classe conhecida (TC<sub>*i*</sub>) será calculada pela média de acertos de detecção desta classe pelo sistema. Como índice global de desempenho, considerou-se a Taxa Média de Detecção de Eventos Conhecidos (TC), definida como a média dos valores de TC<sub>*i*</sub> [7].

Para aferir a capacidade do reconhecimento de novidades, deve-se apresentar ao sistema um conjunto de classes desconhecidas, isto é, diferentes daquelas exploradas no desenvolvimento do modelo. Assim, a partir da base de dados  $\mathcal{D}$ , reservou-se um conjunto de  $R$  classes:  $\mathcal{U} = \{U_1, U_2, U_i, \dots, U_R\}$ , de forma que:  $\mathcal{D} = \mathcal{K} \cup \mathcal{U}$  e  $\mathcal{K} \cap \mathcal{U} = \emptyset$ . Consequentemente, a taxa de detecção da  $i$ -ésima classe desconhecida (TN<sub>*i*</sub>) será dada pela média de acertos de detecção desta classe. Por fim, considera-se a Taxa Média de Detecção de Novidades (TN) como o índice global de desempenho, equivalente a média dos valores de TN<sub>*i*</sub> [7].

A escolha do limiar de decisão  $\lambda_i$  de cada detector classe-especialista define uma dada situação operativa, refletindo-se nos valores de TC e TN. Assim, pode-se traçar uma curva de compromisso para estas taxas, referida aqui como Curva de Operação do Detector (COD). Um critério possível para definir os limiares supracitados consiste em considerar, individualmente, diferentes valores de percentis ( $\alpha$ ) das medidas de dissimilaridade observadas para cada uma das classes, ou seja,  $\{\alpha_1, \alpha_2, \dots, \alpha_{p-1}, \alpha_p\} = \{0\%, 5\%, \dots, 95\%, 100\%\}$ , considerando o conjunto de validação de cada classe. Uma medida quantitativa sobre a COD é a área abaixo da curva (AUC) [11]. Portanto, quanto maior for o valor da AUC, mais eficaz será o sistema detector, visto que maiores serão os valores de TC relativos a cada TN.

### D. Sintonia de Hiperparâmetros

Cada detector classe-especialista tem seu desempenho influenciado pelo conjunto de hiperparâmetros com o qual foi desenvolvido. Sua definição considerou uma avaliação do tipo “um-contratodos” envolvendo as classes conhecidas. Portanto, hipoteticamente assumindo as classes A, B e C como conhecidas, o ajuste do detector da classe A assumiu as classes B e C como desconhecidas, e assim sucessivamente.

A figura de mérito adotada para esta sintonia foi a AUC relativa à COD produzida para cada detector, neste caso, através de um processo de validação cruzada, realizado por meio da técnica de reamostragem 10-fold [11]. Destaca-se que para o detector não-especializado, é impossível proceder esta escolha por validação cruzada sem a utilização de dados pertencentes a  $\mathcal{U}$ .

## III. EXPERIMENTOS

Os experimentos consideraram sinais acústicos irradiados por 8 classes de embarcações em 263 corridas de prova, realizadas na raia acústica da Marinha do Brasil, localizada

em Arraial do Cabo, Rio de Janeiro. O áudio emitido por cada embarcação foi coletado através de um arranjo de hidrofones situado no fundo da raia acústica. Em seguida, o sinal foi amostrado, pré-processado e normalizado, conforme procedimento descrito em [6]. Ao final, foram obtidas 29277 janelas espectrais de 557 pontos cobrindo a faixa de frequência de interesse (0 a 3 kHz). A Tabela I apresenta o total de janelas espectrais disponíveis para cada uma das 8 classes de navios.

TABELA I

DISTRIBUIÇÃO DE JANELAS ESPECTRAIS (JE) ENTRE AS CLASSES (CL).

CL	A	B	C	D	E	F	G	H
JE	2432	3432	4797	3072	7075	2934	2143	3392

As 8 classes de embarcações foram particionadas entre os conjuntos de classes conhecidas ( $\mathcal{K}$ ) e desconhecidas ( $\mathcal{U}$ ). Foram considerados três cenários de avaliação, assumindo 3, 5 e 7 classes conhecidas, que representam cenários de diferentes complexidades quanto à detecção de novidades. A escolha das classes integrantes de  $\mathcal{K}$  para cada caso foi similar à adotada em [12], considerando combinações de classes identificadas como mais críticas, resultando nos seguintes pares de classes conhecidas e desconhecidas: {A, C, G} / {B, D, E, F, H}; {A, B, C, F, G} / {D, E, H}; e {A, B, C, E, F, G, H} / {D}.

## IV. RESULTADOS

A avaliação dos modelos considerou os valores de AUC gerados através da técnica de reamostragem 10-fold. Por esta técnica, uma parcela de 10% do conjunto de dados contendo as classes conhecidas (supondo um determinado cenário de investigação) foi reservada como conjunto de teste a cada rodada. Os dados restantes foram divididos considerando as seguintes proporções: 10% para a validação e 90% para o treinamento. Ao final, foram obtidos 10 trios de conjuntos de treino, validação e teste.

Destaca-se que o classificador utilizado no primeiro estágio da Topologia HI foi baseado no algoritmo  $k$ NN, cujo número de vizinhos foi definido por validação cruzada, considerando a acurácia média como figura de mérito [11]. O experimento com 3 classes conhecidas adotou 9 vizinhos, enquanto aqueles com 5 e 7 classes conhecidas consideraram 11 vizinhos.

A Tabela II exhibe os valores dos hiperparâmetros avaliados para cada técnica considerada para a construção do sistema em tela.

TABELA II

VALORES INVESTIGADOS PARA OS HIPERPARÂMETROS DOS DETECTORES.

Algoritmo	Parâmetro	Valores
AEC	$n_f$	{55; 105; 155; ...; 555}
	$a_f$	{Selu (S); Relu (R); Elu (E); LeakyRelu (L)}
	$c_f$	{3; 5; 7}
	$\sigma$	{0; 0,1; 0,2; ...; 1}
$k$ NN	$k$	{1; 2; 3; ...; 30}

Tratando-se do algoritmo AEC, convém mencionar que: (i) o número de filtros deve sempre ser reduzido conforme se avança da camada inicial até a camada latente da rede neural; (ii) os hiperparâmetros das camadas de codificação

e decodificação simétricas são os mesmos; (iii) adotou-se uma busca *greedy* para os hiperparâmetros de cada camada, segundo a seguinte ordem: primeiramente definiu-se o par ótimo  $n_f \times c_f$ , seguido por  $a_f$ , e finalmente  $\sigma$ .

A Tabela III resume os hiperparâmetros ótimos do AEC  $\{n_f; c_f; a_f; \sigma\}$  para cada detector classe-especialista e também para o caso não-especializado (NE), considerando os diferentes cenários. Cabe destacar que as camadas identificadas são as de projeção (a arquitetura é simétrica). Adicionalmente, cada classe apresenta um número de camadas de projeção particular, variando de 1 até 4. Por sua vez, a Tabela IV exhibe os hiperparâmetros ótimos obtidos para o algoritmo  $k$ NN.

TABELA III  
HIPERPARÂMETROS ÓTIMOS DO AEC.

Cenário com 3 classes conhecidas				
Detector	1ª camada	2ª camada	3ª camada	4ª camada
A	{555;7;S;0,7}	{455;7;S}		
C	{555;5;L;0,7}	{455;5;L}	{205;5;L}	{55;5;L}
G	{555;7;S;0,0}			
NE	{505;5;R;0,5}	{405;5;R}		
Cenário com 5 classes conhecidas				
Detector	1ª camada	2ª camada	3ª camada	4ª camada
A	{505;3;L;0,6}	{55;5;L}		
B	{505;3;S;0,5}			
C	{505;5;L;0,8}	{405;5;L}	{105;5;L}	{55;5;L}
F	{555;3;L;0,6}			
G	{455;3;S;0,7}	{255;3;R}	{205;3;R}	
NE	{555;5;L;0,6}	{505;5;R}		
Cenário com 7 classes conhecidas				
Detector	1ª camada	2ª camada	3ª camada	4ª camada
A	{355;5;S;0,7}	{305;3;R}		
B	{505;3;L;0,5}	{205;5;L}		
C	{555;5;L;0,8}	{505;3;L}	{105;5;L}	
E	{555;5;L;0,6}	{505;5;L}	{105;5;L}	
F	{555;3;L;0,5}			
G	{555;5;R;0,1}	{505;3;R}		
H	{555;5;S;0,7}	{455;3;R}	{405;3;L}	{255;3;S}
NE	{555;5;S;0,4}	{505;5;R}		

TABELA IV  
HIPERPARÂMETROS ÓTIMOS DO  $k$ NN.

Cenário	Detector								
	A	B	C	D	E	F	G	H	NE
ACG	6	-	2	-	-	-	15	-	8
ABCFG	5	2	2	-	-	2	15	-	5
ABCEFGH	5	2	2	-	2	2	15	14	2

Aplicou-se o teste de *Friedman* com um nível de significância de 5% [11] para averiguar se os valores do índice AUC obtidos para os diferentes métodos eram estatisticamente distintos. Quando tal diferença foi identificada, adotou-se o teste *HSD* [11] para identificar entre quais pares de métodos esta se verificava, isto é, quando  $p \leq 0,05$ .

Na Figura 5 é exibido o diagrama de caixas da AUC para o cenário de avaliação contendo 3 classes conhecidas. Observa-se que ambas as técnicas apresentam medianas superiores ao se considerar a Topologia HI. Os testes de hipóteses ( $\chi^2(5) = 31,6; p < 0,001$ ) sinalizam que a técnica AEC-HI é superior à AEC-NE ( $p < 0,001$ ) e estatisticamente semelhante à  $k$ NN-HI ( $p = 0,891$ ), à  $k$ NN-UN ( $p = 0,261$ ) e à AEC-UN ( $p =$

0,206), porém a AEC-HI detém o maior valor de mediana da AUC.

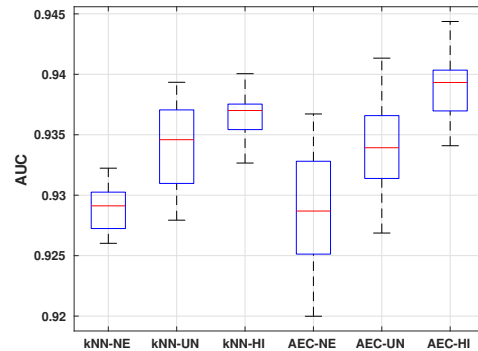


Fig. 5. Diagrama de caixas da AUC do experimento considerando 3 classes conhecidas.

Os resultados para o cenário de avaliação contendo 5 classes conhecidas são reportados na Figura 6. Verifica-se novamente que a Topologia HI detém as maiores medianas. Adicionalmente, observa-se que o acréscimo no número de classes conhecidas consideradas gerou um maior destaque da técnica AEC perante a  $k$ NN, ampliando a diferença entre as medianas, quando comparado ao experimento anterior. Através de testes estatísticos ( $\chi^2(5) = 46,4; p < 0,001$ ), constata-se que a técnica AEC-HI é superior à AEC-NE ( $p < 10^{-5}$ ) e semelhante estatisticamente à AEC-UN ( $p = 0,839$ ) e à  $k$ NN-HI ( $p = 0,121$ ), apresentando uma maior mediana.

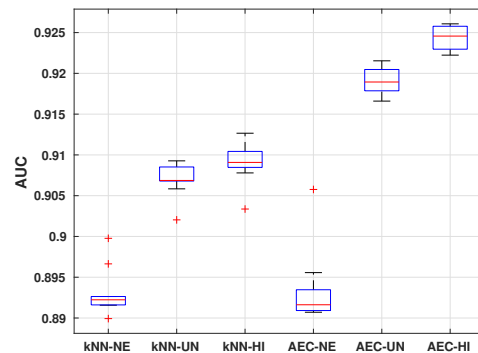


Fig. 6. Diagrama de caixas da AUC do experimento considerando 5 classes conhecidas.

A Figura 7 resume o cenário de 7 classes conhecidas. Verifica-se, novamente, que a especialização em classes é benéfica para as duas técnicas, especialmente para a AEC. Pelos testes estatísticos ( $\chi^2(5) = 45,5; p < 0,001$ ), observa-se que a técnica AEC-HI é superior à AEC-NE ( $p < 10^{-7}$ ), porém estatisticamente semelhante à  $k$ NN-HI ( $p = 0,394$ ) e à AEC-UN ( $p = 0,394$ ). Além disso, também apresenta a maior mediana.

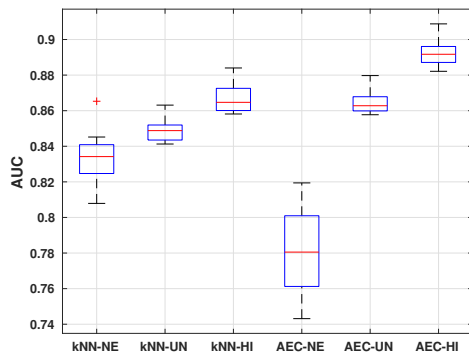


Fig. 7. Diagrama de caixas da AUC do experimento considerando 7 classes conhecidas.

## V. DISCUSSÃO

Relativamente ao método AEC-NE, o AEC-HI sempre apresentou resultados superiores, confirmados pela análise estatística. Na comparação entre as topologias HI e UN, para as técnicas AEC e  $k$ NN, a combinação AEC-HI obteve maiores valores de mediana e quartis, ainda que as diferenças observadas neste escopo de avaliação não tenham se mostrado estatisticamente significativas. Cabe ressaltar que tal resultado pode estar associado ao reduzido número de modelos avaliados. Trabalhos futuros consideram a inclusão de mais modelos nesta análise, através da técnica de  $k$ -fold aninhado, que permitirá o ganho de uma ordem de grandeza quanto à quantidade de modelos avaliada.

Nota-se que o acréscimo no número de classes conhecidas reduz o desempenho do sistema quanto à detecção de novidades. Este comportamento era esperado, visto que quanto maior for este número, maior será o número de fronteiras de classe incluídas, portanto mais complexo será o problema da decisão.

Avaliando-se os modelos associados à mediana dos diferentes ensaios de cada cenário, foram obtidos os pares de índices de desempenho TC e TN apresentados na Tabela V. Tendo em vista as limitações de uma comparação direta, face as diferenças entre bases de dados e índices de desempenho, para o método baseado em *stacked autoencoder* descrito em [5] é reportada uma detecção de sinais conhecidos de 35,6% considerando uma taxa de detecção de novidades de 87,0%, valores inferiores aos reportados na Tabela V. Em relação ao método  $k$ NN hierárquico com detectores navio-especialistas [7], é reportado um reconhecimento de sinais conhecidos de 95,8% para uma detecção de novidades de 75,0%, ao serem consideradas 4 classes conhecidas, valores que são comparáveis aos relatados na Tabela V.

TABELA V

VALORES DE PARES {TC - TN} OBTIDOS PELAS MELHORES TÉCNICAS DE CADA CENÁRIO INVESTIGADO.

Cenário	Técnica	{TC - TN}
ACG	AEC-HI	{76,4% - 93,0%}, {95,0% - 81,6%}
ABCFG	AEC-HI	{75,5% - 91,9%}, {93,8% - 75,9%}
ABCEFGH	AEC-HI	{74,1% - 88,0%}, {94,3% - 55,8%}

## VI. CONCLUSÃO

Este trabalho propôs a construção de um comitê hierárquico de *autoencoders* convolucionais para detectar classes de navios desconhecidas, a ser utilizado em sistemas de sonar passivo.

A topologia hierárquica se mostrou melhor que a tipicamente utilizada neste problema, baseada em um único detector para todas as classes. Na comparação com o estado-da-arte, a proposta obteve valores superiores de mediana, ainda que estatisticamente equivalentes aos obtidos pela técnica  $k$ NN, mostrando que a técnica de *autoencoders* é promissora para o problema. Destaca-se que o incremento do número de classes conhecidas ocasionou um aumento do ganho de eficiência da técnica proposta em relação a soluções alternativas baseadas em  $k$ NN.

Como trabalhos futuros, espera-se explorar outras arquiteturas de redes neurais na construção dos *autoencoders* aplicados na topologia proposta, bem como ampliar o número de modelos considerados em sua avaliação experimental.

## AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Agradecemos também ao Instituto de Pesquisas da Marinha pela disponibilização da base de dados e pela cooperação técnico-científica.

## REFERÊNCIAS

- [1] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," *Signal Processing*, vol. 99, pp. 215 – 249, 2014.
- [2] Y. Kawaguchi, "Anomaly detection based on feature reconstruction from subsampled audio signals," in *2018 26th European Signal Processing Conference (EUSIPCO)*, pp. 2524–2528, 2018.
- [3] Lili Yin, Huangang Wang, and Wenhui Fan, "The active learning-based nearest neighbor mean distance novelty detection for large data set," in *2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, pp. 150–155, 2017.
- [4] J. A. Carino, D. Zurita, A. Picot, M. Delgado, J. A. Ortega, and R. J. Romero-Troncoso, "Novelty detection methodology based on multi-modal one-class support vector machine," in *2015 IEEE 10th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*, pp. 184–190, 2015.
- [5] V. d. S. Mello, N. N. d. Moura, and J. M. d. Seixas, "Novelty detection in passive sonar systems using stacked autoencoders," in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, 2018.
- [6] G. M. R. Rocha and J. B. O. Souza Filho, "Sistemas de classificação de contatos de sonar passivo com rejeição à classes desconhecidas," in *XXI Congresso Brasileiro de Automática*, (Vitória, Brasil), 2016.
- [7] V. H. S. Muniz and J. B. O. Souza Filho, "Detecção especializada de novidades em sonar passivo," in *Anais do 14º Simpósio Brasileiro de Automação Inteligente (SBAI)*, 2019. Anais eletrônicos... Campinas, Galoá, 2020. Disponível em: <<https://proceedings.science/sbai-2019/papers/deteccao-especializada-de-novidades-em-sonar-passivo>>. Acesso em: 30 jun. 2020.
- [8] J. Chen, S. Sathe, C. Aggarwal, and D. Turaga, "Outlier detection with autoencoder ensembles," in *Proceedings of the 2017 SIAM International Conference on Data Mining*, pp. 90–98, 2017.
- [9] S. Boutarfass and B. Besserer, "Convolutional autoencoder for discriminating handwriting styles," in *2019 8th European Workshop on Visual Information Processing (EUVIP)*, pp. 199–204, 2019.
- [10] L. Gondara, "Medical image denoising using convolutional denoising autoencoders," in *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pp. 241–246, 2016.
- [11] N. Japkowicz and M. Shah, *Evaluating learning algorithms: a classification perspective*. Cambridge University Press, 2011.
- [12] V. H. S. Muniz, E. S. Honorato, and J. B. O. Souza Filho, "Aprendizado por instância para a identificação de classes desconhecidas em sonares passivos," in *Anais do 14 Congresso Brasileiro de Inteligência Computacional (CBIC)*, pp. 1–6, ABRICOM, 2019.