

# Análise de modelos adversariais aplicados à classificação de sinais de Sonar Passivo

Júlio de Castro Vargas Fernandes, Natanael Nunes de Moura Junior, José Manoel de Seixas

**Resumo**—Nas operações de guerra naval, várias técnicas foram desenvolvidas para detecção e classificação de navios. Dada a natureza sigilosa dos dados é extremamente difícil obter grandes quantidades dos mesmos, o que torna difícil o uso de técnicas que dependam de dados abundantes, especialmente o aprendizado profundo. Este trabalho propõe a aplicação de uma técnica de treinamento adversarial para classificação de sinais de sonar passivo. Como a técnica proposta treina um gerador de imagens sintéticas, que depois serão utilizadas para o treinamento de um classificador, a qualidade daquele será avaliada. O método proposto atinge uma acurácia de  $98,4 \pm 0,6$  %.

**Palavras-Chave**—Sistemas Sonar, Redes Neurais, Redes Neurais Generativas, Aprendizado Profundo.

**Abstract**—In naval warfare, several techniques have been developed for the detection and classification of ships. Given the confidential nature of the data it is extremely difficult to get a hold of large quantities of data which makes it extremely hard to use techniques that rely on abundant data, like deep learning. This paper proposes the use of generative adversarial neural networks for the classification of passive sonar signals. Since the proposed technique trains a synthetic image generator, which will later be used for the training of a classifier, its quality will be evaluated. The proposed method achieved an accuracy of  $98,4 \pm 0,6$  %.

**Keywords**—Sonar Systems, Neural Networks, Generative Adversarial Neural Networks (GAN), Deep Learning.

## I. INTRODUÇÃO

Um sistema **SO**und **NA**avigation **R**anging (SONAR) é um conjunto de técnicas que utiliza o som que se propaga através da água para comunicação, navegação e detecção de objetos [1]. Desde sua invenção, os sonares encontraram um vasto número de aplicações, tanto para uso civil como, por exemplo, reconhecimento de marcos subaquáticos, rastreamento e estudo de animais, detecção de minas, aplicações em petróleo e gás, como também militares. Os submarinos dependem, consideravelmente, dos sistemas de sonares para navegar e adquirir informações sobre as condições e profundidade do mar, bem como para a detecção e classificação de navios [2], chamados, neste contexto, de contatos.

Existem dois tipos de sistemas sonar usados em navios: o ativo e o passivo [21]. No sistema de sonar ativo, o navio envia um sinal e utiliza seus ecos para detectar, classificar e localizar um possível alvo [3]. Enquanto isso, em um sistema de sonar passivo, não há sinal, emitido obrigando o sonar a trabalhar analisando os sons emitidos pelos possíveis

alvos para a detecção e classificação destes [3]. Este trabalho foi desenvolvido dentro do contexto do sonar passivo, de uso militar. Deste modo, sinais, de sonares passivos são os ruídos emitidos pelas diversas fontes sonoras presentes no ambiente subaquático, sendo que, para este trabalho, os sinais de interesse são aqueles provenientes de navios, enquanto que o ruído de fundo é composto por todos os outros ruídos do ambiente subaquático, que podem incluir os ruídos emitidos pela vida marinha, bem como ruídos emitidos por plataformas de petróleo, por exemplo.

Dada a grande quantidade de ruídos de fundo presente no ambiente marítimo, a classificação de alvos se torna um problema de alta complexidade, pois a assinatura de um sinal de interesse pode ser mascarada pelo ruído de fundo. Os efeitos do ruído de fundo têm forte impacto no problema de classificação de contatos, pois diversas fontes sonoras ocupam a mesma banda de frequência. Não obstante, a quantidade de dados disponíveis para o treinamento de um modelo pode ser bastante diminuta, especialmente em se tratando de dados com possíveis usos militares.

Diversas técnicas visando a detecção e classificação de contatos em um ambiente sonar foram desenvolvidas ao longo dos anos. No contexto de sonar ativo, redes neurais artificiais foram utilizadas com resultados interessantes [8]. No contexto de sonar passivo, o método *Support Vector Machines* (SVM) foi aplicado em [9]. Também existem trabalhos que fazem uso de redes do tipo stacked auto encoder [5] tanto para a classificação de sinais quanto para a detecção de novidades em sonar passivo [11]. Trabalhos que exploram as características tonais [13] e espectrais [12] em conjunto com redes neurais também atingiram ótimos resultados. Em [20], tanto classificadores quanto detectores de novidades foram criados utilizando-se redes perceptron multicamadas (do inglês, *Multi Layer Perceptron*) e SVM atingindo bons resultados nos dois objetivos.

O aprendizado profundo é um ramo de algoritmos na área de aprendizado de máquina que visa o treinamento de modelos complexos através do aprendizado automático de representações de alto nível dos dados [5], sendo isso possível através, principalmente, do grande número de parâmetros que compõem tais modelos. Além disso, o aprendizado profundo tem recebido muita atenção de pesquisadores nos últimos anos [6], visto que as diversas topologias de redes neurais profundas demonstraram grande capacidade de alcançar resultados melhores do que o estado da arte em diversos campos de pesquisa [6]. Esses algoritmos encontram estruturas complexas em grandes conjuntos de dados e alteram seus parâmetros internos para que a representação em cada camada seja construída sobre

Júlio de Castro Vargas Fernandes, Laboratório de Processamento de Sinais - UFRJ - COPPE, e-mail: juliocvf@poli.ufrj.br; Natanael Nunes de Moura Junior, Laboratório de Processamento de Sinais - UFRJ - COPPE, e-mail: natmourajr@lps.ufrj.br; José Manoel de Seixas, Laboratório de Processamento de Sinais - UFRJ - COPPE, e-mail: seixas@lps.ufrj.br.

a representação da camada anterior [7]. Em sonar, temos as técnicas de aprendizado profundo aplicadas na extração de características em [10], aonde imagens obtidas por veículos subaquáticos não tripulados são processadas utilizando-se o aprendizado profundo. Os resultados obtidos fornecem melhor desempenho em comparação com outras técnicas de extração de características clássicas. Neste artigo exploramos técnicas de aprendizado profundo para a geração de amostras sintéticas e o uso dessas para o treinamento de classificadores profundos.

O artigo está organizado da seguinte forma: na Seção II, o processamento básico de um sinal de sonar passivo é descrito. As redes neurais adversariais são brevemente expostas na Seção III e o método desenvolvido é explicado na Seção IV. Os resultados experimentais são analisados na Seção V. Por fim, as conclusões são derivadas na Seção VI.

## II. SISTEMA DE SONAR PASSIVO

O primeiro passo na cadeia de processamento do ruído irradiado por navios, visando a classificação e identificação destes, é a transformação do sinal acústico em um sinal elétrico, através da conformação de feixe [2] [4]. Depois desta conformação o sinal resultante pode ser analisado tanto no domínio do tempo quanto na frequência, com destaque para a LOFAR [14] (*Low-Frequency Analysis and Recording*). A análise LOFAR é uma análise de banda larga que trabalha sobre o ruído de máquina de um dado navio.

### A. LOFAR

O fluxograma da análise LOFAR pode ser visto na Fig. 1. De maneira resumida, esta análise consiste nos seguintes passos: primeiro, o sinal é decimado e, em seguida, multiplicado por uma janela *Hanning* para enfatizar a faixa de frequência de interesse. Subsequentemente, a representação de frequência do sinal é obtida aplicando-se a transformada de Fourier (STFT - do inglês, *Short Time Fourier Transform*), utilizando-se janelas disjuntas de amostras. Deste ponto em diante, somente a magnitude do sinal é processada, a informação de fase é descartada. Na sequência, o algoritmo TPSW (do inglês, *Two Pass Split Window*) [20] é utilizado para atenuar o ruído de fundo e normalizar o sinal. O resultado dessa análise pode ser exibido em um lofargrama, como visto na Fig. 2. no qual, a informação de frequência processada de cada janela é exibida em cada linha do eixo x enquanto o eixo y representa o tempo.

Neste artigo imagens são criadas a partir da análise LOFAR. Esta forma de processar o lofargrama foi motivada pelos bons resultados obtidos usando-se redes neurais, especialmente CNN (CNN - do inglês, *Convolutional Neural Network*) [5], com processamento de imagens. A escolha por tratar o lofargrama como uma imagem, ao invés de processar cada linha deste individualmente se dá pelo fato de uma imagem (combinando várias janelas) conter mais informação do que apenas um espectro, além da habilidade de tratar as imagens de maneira eficiente com redes neurais.

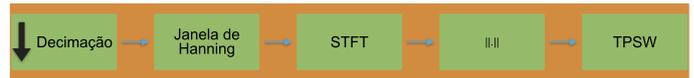


Fig. 1. Diagrama de blocos da análise LOFAR.

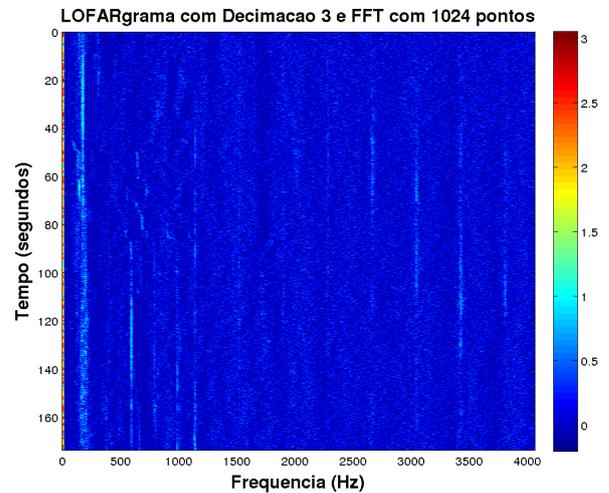


Fig. 2. Lofargrama. O espectro da janela sendo processada é representado no eixo x enquanto que o tempo é representado no eixo y.

## III. REDES NEURAIS ADVERSARIAIS GENERATIVAS

Como dito anteriormente, o treinamento de modelos profundos requer uma grande quantidade de dados, o que pode se tornar desafiador em um contexto de sonar. Para solucionar esse problema, este artigo propõe o uso de redes adversariais generativas (ou do inglês, *Generative Adversarial Neural Networks (GAN)*). Em [17], Goodfellow et al. introduziram o conceito de GAN, que são modelos não supervisionados que tentam gerar dados sintéticos que seguem a mesma distribuição de probabilidade dos dados que se quer reproduzir. Goodfellow propôs um novo paradigma de treinamento em que duas redes feedforward, chamadas de gerador e discriminador, participariam de um jogo de dois jogadores de soma zero, segundo a definição constando na teoria de jogos [16], que permite às redes aprender a distribuição de dados. O objetivo do modelo gerador G é gerar amostras de dados que se assemelhem ao máximo possível àquelas contidas no conjunto de treinamento em um nível tal que o modelo discriminador seja incapaz de distinguir as amostras sintéticas das reais. Enquanto isso, a tarefa do modelo discriminador é avaliar se uma determinada amostra é real ou não. Um esquema da estrutura do GAN pode ser vista na Fig. 3.

Neste artigo, uma das muitas topologias de GAN, a Wasserstein GAN [18] que tem este nome pois utiliza a distância de Wasserstein como função custo, é utilizada para aumentar a estatística presente no conjunto de dados. Depois disso, utilizando-se apenas estas amostras sintéticas geradas, um classificador, uma rede neural convolucional, foi treinada para identificar quatro classes de navios presentes no conjunto de dados. Esta abordagem nos permite criar tantos dados quanto forem necessários para o treinamento de modelos profundos

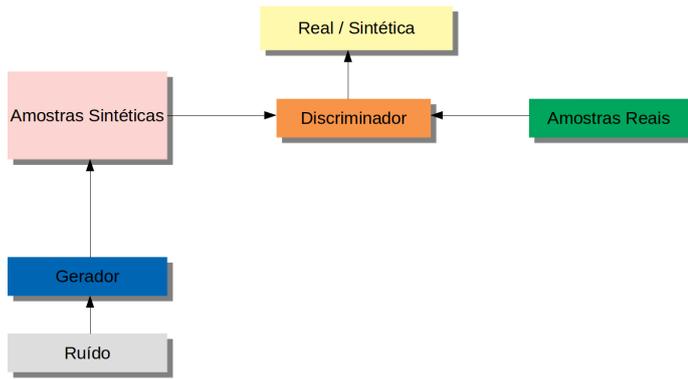


Fig. 3. Visão geral da estrutura da GAN, que consiste em dois modelos, o Gerador (G) e o Discriminador (D). G recebe ruído como entrada e gera uma amostra de dados. O modelo D é treinado com amostras do conjunto de treinamento e amostras geradas por G. Ele aprende como separar amostras reais de amostras geradas e é treinado para produzir 1 para o primeiro e 0 para o segundo. G é treinado como qualquer outra rede por backpropagation [5] vindo de D. Durante o treinamento, G gera amostras e D as julga. O objetivo aqui é usar as amostras que foram realistas o suficiente para enganar D para melhorar G. É importante notar que G nunca é mostrado qualquer amostra real sendo treinado apenas através do feedback que recebe de D.

efetivamente superando o problema da quantidade de dados em classificação de contatos.

#### IV. MÉTODO

##### A. Pré-Processamento

Este artigo opera sobre os lofargramas criados pela análise LOFAR. Essas imagens precisam ser normalizadas antes de treinarmos os nossos modelos. Para isso, os lofargramas são normalizados utilizando-se a norma- $\ell_2$ , i.e, para cada linha do lofargrama calculamos a norma- $\ell_2$  desta linha e depois dividimos toda a linha pela sua norma- $\ell_2$ . Esta normalização é o único pré-processamento aplicado nos dados.

##### B. Validação Cruzada

Objetivando-se estimar a flutuação estatística do método proposto, o método de validação cruzada *jackknife* foi utilizado. Primeiramente, os dados foram aleatoriamente particionados em  $k$  conjuntos disjuntos de mesmo tamanho. Em cada fold,  $k - 1$  desses conjuntos são utilizados como conjunto de treino enquanto o conjunto restante serve de conjunto de teste. A variação estatística foi medida através do cálculo das figuras de mérito nos  $k$  modelos provenientes dos  $k$  folds. Os resultados apresentados são os valores médios conjuntamente com os valores RMS das figuras de mérito.

Para reduzirmos a possibilidade de uma má inicialização dos pesos nas redes neurais (o que, em teoria, poderia levar até a um mínimo local), cada modelo foi inicializado um certo número de vezes e em cada fold apenas a melhor inicialização foi mantida.

##### C. Redes Neurais

Em cada fold de validação cruzada o mesmo procedimento é aplicado. Primeiramente, as redes GAN são treinadas, utilizando-se o conjunto de treino, para gerar amostras

sintéticas. Nesta etapa, treinamos uma GAN especialista para cada classe, ou seja, temos uma GAN que gera somente dados da classe A, outra para a classe B e assim por diante. Deste modo, GAN especialistas são treinadas, uma para cada classe de navio. Depois do treinamento das GAN ter sido concluído utilizamos estas para gerar imagens sintéticas de cada classe e estas são utilizadas para treinar um classificador para as quatro classes, i.e, uma CNN. Depois disso, utilizamos o conjunto de teste para validar o classificador, ou seja, o classificador é treinado apenas em dados sintéticos e é validado em dados reais que não participaram do treinamento da GAN.

Duas perguntas importantes devem ser respondidas concernentes ao treinamento da GAN: a GAN realmente aprendeu a pdf geradora dos dados? A GAN está criando dados novos ou está simplesmente reproduzindo os mesmos dados que esta viu ao longo do treinamento? A primeira pergunta concerne à convergência do treinamento, enquanto que a segunda trata de overfitting.

A primeira pergunta pode ser respondida em parte pela própria função custo de treinamento, a distância de Wasserstein, já que esta quantifica quanto trabalho deve ser feito para transformar uma pdf em outra e assim sendo quanto menor a distância de Wasserstein mais parecidas as duas pdf são. Embora a análise da função custo possa responder a primeira pergunta, neste trabalho também foi utilizada a divergência de Kullback-Leiber (KL) para medir se os dados sintéticos seguem a mesma pdf dos dados originais. A ideia é simples: já que cada classe possui diversas tomadas de dados, podemos medir a variância estatística de cada classe medindo a divergência KL entre todas estas tomadas (KL entre a tomada 1 e 2, 1 e 3, ..., n-1 e n) para todos os bins de frequência da análise LOFAR e averiguar se a divergência entre os dados reais e sintéticos fica dentro da variância medida. Uma outra característica importante dessa análise é que os pontos de maior variação da KL costumam ser os pontos em que há maior informação, então, esta técnica também permite averiguar os bins mais relevantes para cada classe.

A segunda pergunta foi respondida através do uso do erro  $\ell_2$  (definido como a soma dos erros ao quadrado entre os pixels correspondentes de duas imagens). Para cada imagem real no dataset, o erro  $\ell_2$  entre esta imagem e todas as imagens sintéticas (foi utilizado um conjunto de imagens sintéticas de mesmo tamanho das imagens reais) foi calculado; se alguma imagem gerada fosse idêntica ou parecida com as imagens sintéticas teríamos o erro  $\ell_2$  em zero ou na sua vizinhança.

#### V. RESULTADOS EXPERIMENTAIS

Os dados utilizados neste trabalho são gravações do ruído irradiado por navios de quatro classes diferentes (identificadas como Classe A, B, C e D) feitas em uma raia acústica pertencente à Marinha do Brasil localizada em Arraial do Cabo, Rio de Janeiro. Cada gravação foi feita utilizando-se um único hidrofone omnidirecional posicionado no fundo da raia. Para cada gravação, um navio operando em uma mesma condição de máquina durante toda a gravação é colocado na raia acústica e o ruído irradiado por este é gravado. Este processo de tomada de dados, chamado de corrida, é repetido por várias vezes com diferentes condições de mar e máquina.

A análise LOFAR utilizou uma taxa de decimação de 3 e janelas de 1024 amostras na STFT. No final do processo cada espectro tinha 400 bins de frequência, deste modo cada lofargrama de cada corrida possui  $n$  linhas  $\times$  400 bins, aonde  $n$  depende do tamanho de cada corrida.

Para cada corrida, o lofargrama correspondente é gerado e então, a partir deste, diversos sub-lofargramas são gerados. Assim como o lofargrama, esses sub-lofargramas também são imagens, porém, de tamanho  $L \times 400$ . O processo de geração destas sub-imagens começa na primeira linha do lofargrama,  $L$  linhas sequenciais são selecionadas para compor uma imagem. Em seguida, a cada  $R$  linhas, outras imagens de tamanho  $L \times 400$  são geradas da mesma forma. Os parâmetros  $L$  (altura da sub-imagem) e  $R$  (o tamanho do passo a ser dado entre uma imagem e outra) foram selecionados fazendo-se um grid search aonde se variou  $L$  de 5 até 40 de 5 em 5 e  $R$  de 1 até 10 de 1 em 1. Os valores que obtiveram os melhores resultados foram 20 e 5 para  $L$  e  $R$ , respectivamente. Para todas as corridas e todas as classes de navio, o mesmo processo de criação de imagens é repetido.

O algoritmo de treinamento da Wasserstein GAN utilizado neste trabalho é o WGAN-GP [19]. Tanto a arquitetura do gerador quanto a arquitetura do discriminador da GAN são inspirados na arquitetura da DC-GAN [22]. A arquitetura do classificador utilizado é a mesma utilizada no discriminador da GAN. A GAN foi treinada por 5000 épocas com bateladas de 32 amostras utilizando o algoritmo ADAM [5]. O classificador foi treinado por 100 épocas com o mesmo algoritmo.

#### A. Geração de Dados Sintéticos

Primeiramente, explicitaremos as arquiteturas usadas neste trabalho. O gerador recebe como entrada um vetor aleatório de dimensão 100 e gera uma imagem de tamanho  $20 \times 400$ . A arquitetura da rede consiste em uma camada totalmente conectada, três camadas convolucionais, todas com filtros de tamanho  $5 \times 5$ . O número de filtros usados em cada uma das camadas convolucionais é  $\{64, 32, 1\}$ , respectivamente. Batch-Normalization [23] é aplicada a cada camada na rede, exceto na camada de saída. A técnica de batch-normalization estabiliza o processo de aprendizagem e impede que o gerador colapse todas as amostras em um mesmo ponto conforme descrito em [23]. A função ReLU [5] é usada como função de ativação em todas as camadas, exceto na última camada, que usa a função tangente hiperbólica.

O discriminador também é uma CNN. Este recebe como entrada uma imagem de tamanho  $20 \times 400$  e tem como saída um escalar, aonde 1 significa que a imagem é real e zero significa que esta é sintética. A arquitetura é composta por três camadas convolucionais com filtros de tamanho  $5 \times 5$  e duas camadas totalmente conectadas. O número de filtros usados nas camadas convolucionais é  $\{16, 32, 32\}$ , respectivamente, e o número de neurônios nas camadas totalmente conectadas é  $\{50, 1\}$ . Mais uma vez, a técnica de batch-normalization é aplicada em todas as camadas, exceto na primeira e na última. Todas as camadas usam as funções LeakyReLU [5], exceto a última, que usa uma função linear.

Na Fig. 4 podemos ver as curvas de treinamento para a GAN especializada na classe A para todos os folds. Como

podemos ver o treinamento convergiu para todos os folds. O comportamento para as outras classes é idêntico.

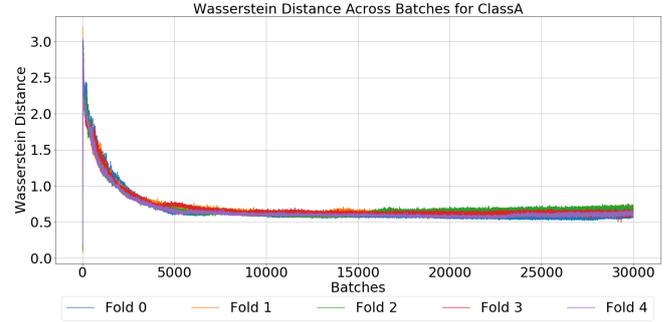


Fig. 4. Curva de treinamento da classe A.

Com respeito à geração das imagens, a divergência entre os dados reais e sintéticos fica dentro da variância esperada como pode ser visto na Fig. 5 para a maioria dos bins, sendo que em alguns bins, especialmente nos bins mais ao final do espectro, há divergências. Porém estes bins com erros mais proeminentes costumam representar ruído o que atenua este problema já que o mais importante é que a GAN gere os bins com sinal corretamente.

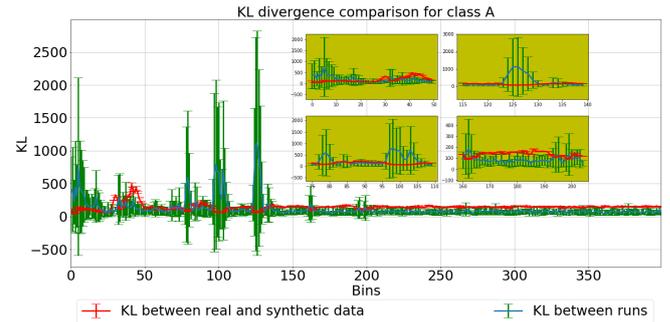


Fig. 5. Comparação entre as divergências de KL dentro da classe A e a divergência de KL entre amostras sintéticas e reais da classe A.

Na Fig. 6 podemos ver o erro  $\ell_2$  calculado entre as imagens reais e sintéticas e também entre as imagens reais. Como podemos ver as amostras sintéticas são tão diversas quanto as imagens reais, e também, não há dentro as amostras sintéticas nenhuma amostra sintética na vizinhança de uma amostra real.

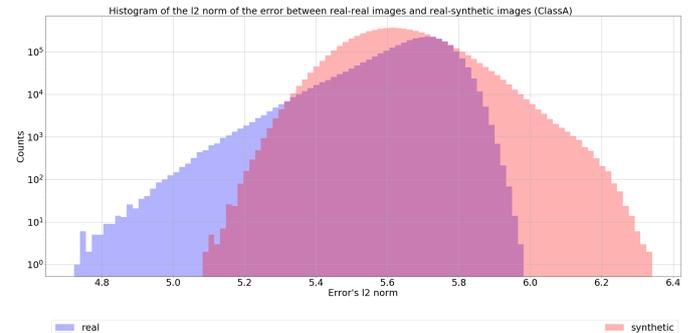


Fig. 6. Erro  $\ell_2$  calculado *pixelwise* entre as imagens sintéticas e as imagens reais.

## B. Classificação

Depois do treinamento das GAN podemos treinar os classificadores. Como dito anteriormente a arquitetura do classificador é idêntica a do discriminador utilizado na GAN, exceto que na última camada a tangente hiperbólica é utilizada como função de ativação em vez da função linear. Como prova de conceito do método proposto, e como dito anteriormente, o classificador preliminar foi treinado somente com amostras sintéticas e validado em amostras reais.

Na Fig 7 podemos ver que a acurácia obtida pelo método proposto fica excepcionalmente alta para todas as 4 classes, com resultados melhores do que na literatura [15] aonde no mesmo conjunto de dados uma rede neural mlp atingiu 91% de acurácia. Como o classificador foi treinado apenas em dados sintéticos mostramos aqui a acurácia no conjunto de treino (chamado aqui de corrida de treinamentos), que lembrando foi usado somente para treinar as GAN, e para o conjunto de validação (chamado aqui de corrida de validação), que foi usado somente para validar o classificador.

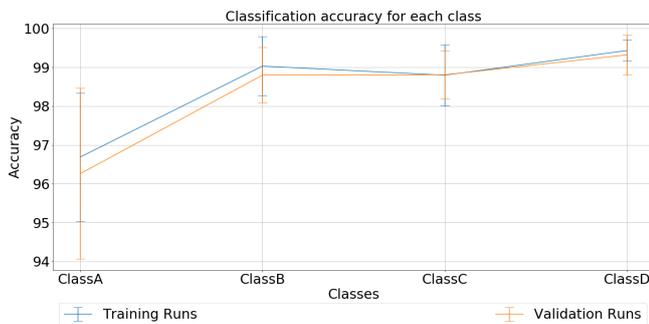


Fig. 7. Acurácia para todas as classes presentes no dataset para as corridas de treinamento e validação.

## VI. CONCLUSÕES

Sistemas de sonar passivo são importantes em muitas áreas da acústica submarina, tanto em aplicações civis quanto militares. Neste artigo, exploramos o uso de modelos generativos para produzir amostras sintéticas realistas para então usarmos estas amostras no treinamento de um sistema de classificação que utiliza redes neurais convolucionais. A necessidade de um sistema de classificação em um contexto militar é óbvia, dado que há uma necessidade de tomar decisões e essas decisões serão muito diferentes se estivermos lidando com um destróier ou um navio de pesca.

Neste trabalho, podemos ver que a geração de amostras sintéticas realistas é uma tarefa viável no contexto de sistemas de sonar passivo, o que é uma grande conquista em si. Nós também atestamos que essas amostras sintéticas podem ser usadas para treinar classificadores e os resultados preliminares até agora apontam que sim mas ainda é preciso tratar o problema de classificação de forma mais rigorosa.

Como possíveis trabalhos futuros, podemos explorar se o uso de técnicas como bagging e boosting poderiam evitar o problema de colapso de modo, quando a GAN não aprende certos modos contidos no dataset, presente na GAN.

## AGRADECIMENTOS

Os autores gostariam de agradecer à CAPES, ao CNPq e à FAPERJ pelo apoio a este trabalho. Um agradecimento especial ao Instituto de Pesquisa da Marinha do Brasil (IPqM) por fornecer o conjunto de dados e por discussões proveitosas sobre este trabalho. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES).

## REFERÊNCIAS

- [1] R. Urick, Principles of Underwater Sound for Engineers. McGraw-Hill, 1983.
- [2] R. P. Hodges, Underwater acoustics: Analysis, design and performance of sonar. John Wiley & Sons, 2011.
- [3] R. O. Nielsen, Sonar Signal Processing. Norwood, MA, USA: Artech House, Inc., 1991.
- [4] W. S. Burdick, Underwater Acoustic System Analysis. Peninsula Pub, 2003.
- [5] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.
- [6] J. Schmidhuber, Deep learning in neural networks: An overview, Neural Networks, vol. 61, no. Supplement C, pp. 85 - 117, 2015.
- [7] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, Nature, vol. 521, no. 7553, pp. 436 - 444, 2015.
- [8] C. H. Chen, Neural networks for active sonar classification, in Proceedings., 11th IAPR International Conference on Pattern Recognition. Vol.II. Conference B: Pattern Recognition Methodology and Systems, 1992, pp. 438 - 440.
- [9] F. B. TA and P. R. Support vector machine applied to underwater target classification, in 2014 Fourth International Conference on Advances in Computing and Communications, 2014, pp. 46-49.
- [10] P. Zhu, J. Isaacs, B. Fu, and S. Ferrari, Deep learning feature extraction for target recognition and classification in underwater sonar images, in 2017 IEEE 56th Annual Conference on Decision and Control (CDC), 2017, pp. 2724-2731.
- [11] 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.
- [12] X. JIANG, Q. WANG, X. ZENG, Cavitation Noise Classification Based on Spectral Statistic Features and PCA Algorithm, 3rd International Conference on Computer Science and Network Technology, pp. 438 - 441, 2013.
- [13] C. CHIN-HSING, L. JIANN-DER, L. MING-CHI, Classification of underwater signals using wavelet transforms and neural networks, Mathematical and Computer Modelling, v. 27, n. 2, pp. 47 - 60, 1998.
- [14] Q. Li, Digital Sonar Design in Underwater Acoustics. Springer-Verlag Berlin Heidelberg, 2012.
- [15] W. Soares-Filho, J. M. Seixas, and N. N. de Moura, Preprocessing passive sonar signals for neural classification, in 2010 in IET Radar, Sonar and Navigation.
- [16] M. Maschler, E. Solan, & S. Zamir, Game Theory. Cambridge: Cambridge University Press, 2013.
- [17] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative Adversarial Networks, 2014, *arXiv e-prints*, arXiv:1406.2661.
- [18] M. Arjovsky, S. Chintala, L. Bottou 2017. Wasserstein GAN. *arXiv e-prints arXiv:1701.07875*.
- [19] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville 2017. Improved Training of Wasserstein GANs. *arXiv e-prints arXiv:1704.00028*.
- [20] N. N. de Moura Junior, Detecção de Novidades para sistemas de sonar passivo. Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2018.
- [21] Q. Li, Digital sonar design in underwater acoustics: principles and applications. Springer Science & Business Media, 2012.
- [22] A. Radford, L. Metz, S. Chintala 2015. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv e-prints arXiv:1511.06434*.
- [23] S. IOFFE, C. SZEGEDY, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, 2015.