



Comando & Controle e Defesa Cibernética: Sistemas Satelitais

KNN e Rede Neural Convolutacional para o Reconhecimento de Plataformas de Petróleo em Imagens SAR do Sentinel-1

KNN and Convolutional Neural Network for Oil Rig Recognition in Sentinel-1 SAR Images

Leonan Entringer Falqueto², Rafael Lemos Paes¹ e Angelo Passaro¹¹ Instituto de Estudos Avançados (IEAv), São José dos Campos/SP – Brasil² Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos/SP – Brasil

Article Info

Article History:

Received	11 May	2023
Revised	28 July	2023
Accepted	23 August	2023
Available online	22 September	2023

Palavras Chave:

**Synthetic Aperture Radar (SAR)
Reconhecimento Automático de Alvos (ATR)
Machine Learning**

Keywords:

**Synthetic Aperture Radar (SAR)
Automatic Target Recognition (ATR)
Machine Learning**

E-mail addresses:

falquetolef@fab.mil.br (L.Falqueto),
rafaelpaesrlp@fab.mil.br (R.L.Paes),
angeloap@fab.mil.br (A. Passaro).

Resumo

O reconhecimento automático de alvos (plataformas petrolíferas) por meio de imagens SAR de média resolução auxilia a vigilância de áreas extensas como o Atlântico Sul. Por isso, esse trabalho aprofundou o estudo do emprego da VGG-16 como extratora de atributos para alimentar algoritmos de Machine Learning, especificamente, o kNN. Variou-se o número de vizinhos para um conjunto de amostras de imagens SAR do Sentinel-1 contendo plataformas marítimas e falsos-alarmes, usando um experimento com 50 blocos de treinamento e teste. Demonstrou-se que o ajuste de parâmetros do classificador apresenta melhorias significativas, com um incremento de 6,46% no indicador AUC.

Abstract

The automatic recognition of targets (oil platforms) using medium-resolution SAR images aids in the surveillance of vast areas such as the South Atlantic. Therefore, this work delved into the study of employing VGG-16 as a feature extractor to feed Machine Learning algorithms, specifically, the kNN. The number of neighbors was varied for a set of SAR image samples from Sentinel-1 containing maritime platforms and false alarms, using an experiment with 50 training and testing blocks. It was demonstrated that the adjustment of classifier parameters results in significant improvements, with an increase of 6.46% in the AUC indicator.

I. INTRODUÇÃO

O Brasil, assim como os demais países com grandes dimensões territoriais, enfrenta o desafio de manter a vigilância sobre extensas áreas. Essa tarefa é sintetizada no conceito da Dimensão 22, veiculado pela Força Aérea Brasileira [1].

Os satélites imageadores proporcionam, rotineiramente, uma grande massa de dados distribuídas, inclusive, gratuitamente. Tais dados, se bem utilizados, podem contribuir para solucionar a questão da vigilância. Dentre esses sensores, destaca-se os satélites com SAR, que produzem amplas faixas de imageamento, operam diuturnamente e utilizam a banda espectral das micro-ondas, pouco afetada pelas condições meteorológicas da troposfera (atenuação do sinal por absorção de energia eletromagnética). Diversas constelações estão em operação nas principais bandas (X, C e L) e com diversas configurações, em termos de resolução e de polarização. Na verdade, a abundância de dados alçou o SAR à era do *Big Data* [2], [3].

A média de dados diária do Sentinel-1 (satélites SAR do Projeto Copernicus da ESA – *European Space Agency*), obtida no mês de novembro de 2017 foi de 3,79 Tb [3], [4]. A tendência é que essa a quantidade de dados aumente consideravelmente com a operação do Projeto NISAR, com a previsão de 85 Tb/dia [5]. Portanto, há uma necessidade premente de automatização no processo de vigilância por imagens orbitais, ou seja, são necessários métodos eficientes para tratar e sintetizar essa quantidade de dados em informações de apoio aos processos decisórios.

Dentre os alvos artificiais de interesse no oceano, a literatura é vasta para manchas de óleo e navios, porém pouco foi relatado sobre plataformas marítimas [6], [7]. Estudar a interação da onda eletromagnética com as plataformas marítimas tem relevância para as áreas de inteligência militar e econômica, para segurança da navegação naval e aérea de baixa altitude, para antecipação de providências em casos de desastres ambientais provocados por acidentes ou fenômenos naturais como furacões e para vigilância em casos de poluição ambiental. Além disso, uma das principais motivações é que a verdade de campo é obtida facilmente, por ser um alvo metálico

parado e, normalmente, de posição conhecida. Isso facilita a extrapolação do conhecimento para outros objetos de interesse como navios, sobre os quais não é tão simples obter a verdade de campo. O desafio para reconhecimento de plataformas petrolíferas é a dinamicidade do ambiente marítimo associada à presença de muitos alvos semelhantes, enquanto elas apresentam características geométricas diversas, não apresentando um comportamento tão homogêneo quanto ao retroespalhamento do sinal radar [6]–[8].

No âmbito das técnicas de reconhecimento automático de alvos por meio de *deep learning*, as Redes Neurais por Convolação (CNN) apresentam resultados promissores com imagens SAR de média e alta resolução. Uma taxa de acurácia de 79% foi alcançada para o reconhecimento de três tipos de navios (graneleiro, contêiner e petroleiro) usando imagens de média resolução Sentinel-1 (5m x 20m, SLC) [9].

O estado da arte no reconhecimento de alvos foi alcançado por meio de avanços na aprendizagem profunda [10]–[16], incentivando novos estudos em vigilância marítima com imagens coletadas pelos sensores espaciais atuais. Por exemplo, a CNN é usada para classificar quatro tipos de navios em imagens SAR de alta resolução (1,5m x 2,0m) do COSMO-SkyMed em [17]. Em [18], atingiu-se mais de 90% de acurácia na discriminação entre cinco classes de objetos metálicos no mar em imagens de alta resolução do TerraSAR-X (1,0m x 2,0m).

Em [7], foram comparadas duas CNN para extrair atributos de imagens SAR para o reconhecimento de plataformas de petróleo. Esses atributos foram submetidos a seis algoritmos de *Machine Learning* para finalizar o processo de distinção entre as plataformas de petróleo e os falsos-alarmes. A melhor performance foi obtida utilizando o algoritmo de Regressão Logística, extraído-se atributos por meio da VGG-16 de imagens SAR com a polarização VH, alcançando a acurácia de 86,4%, como média de 50 testes.

Foi apontado que não houve parametrização dos algoritmos de *Machine Learning* empregados em [7], sendo essa uma pesquisa futura para que o desempenho obtido com imagens de alta resolução também seja alcançado com

imagens de média resolução. Assim, esse trabalho procurou iniciar tal estudo avaliando o algoritmo de configuração mais simples: o kNN (*k-Nearest Neighbor*, em português, k-vizinhos mais próximos). Executou-se um experimento variando o fator k em quatro níveis (k = [3, 7, 15 e 31]), utilizando a seleção do conjunto de treinamento e de teste como bloco [19]. O objetivo principal é verificar se a acurácia do classificador sofre alterações significativas ao recorrer a um ajuste de parâmetros para a tarefa proposta em [7].

O artigo apresenta na próxima seção o método empregado por [7] para o reconhecimento de plataformas marítimas e o algoritmo kNN. A Seção III apresenta a metodologia, bem como os procedimentos de avaliação do experimento, seguida dos resultados e das discussões na Seção IV.

II. REDES NEURAS CONVOLUCIONAIS E O KNN

Em [7], duas redes neurais convolucionais, a VGG-16 e a VGG-19, foram utilizadas como extratoras de atributos de 800 amostras de imagens SAR, sendo 400 de cada polarização, 50% plataformas *offshore* e 50% falsos-alarmes. Esse trabalho utilizou a configuração de melhor desempenho, ou seja, a VGG-16 extraído atributos das 400 amostras da polarização VH.

A VGG-16 é uma rede neural convolucional profunda com arquitetura conforme Fig. 1 [20], [21]. Ela é composta por cinco grupos de camadas de Convolação e de *Max Pooling*, seguidas por três camadas *fully connected*. A rede usada foi aquela implementada em [22], que já é treinada no banco de dados ImageNet [23].

A última camada *fully connected* de 1000 neurônios (fc8) é substituída pelo algoritmo de classificação supervisionada kNN, sendo esse algoritmo alimentado pelo *output* da camada anterior (fc7), isto é, por 4096 valores que compõe o vetor de atributos.

O algoritmo kNN compara a amostra sob teste com k vizinhos mais próximos e classifica de acordo com a classe predominante, podendo ser uma comparação ponderada por algum fator (peso). Para determinar os vizinhos mais

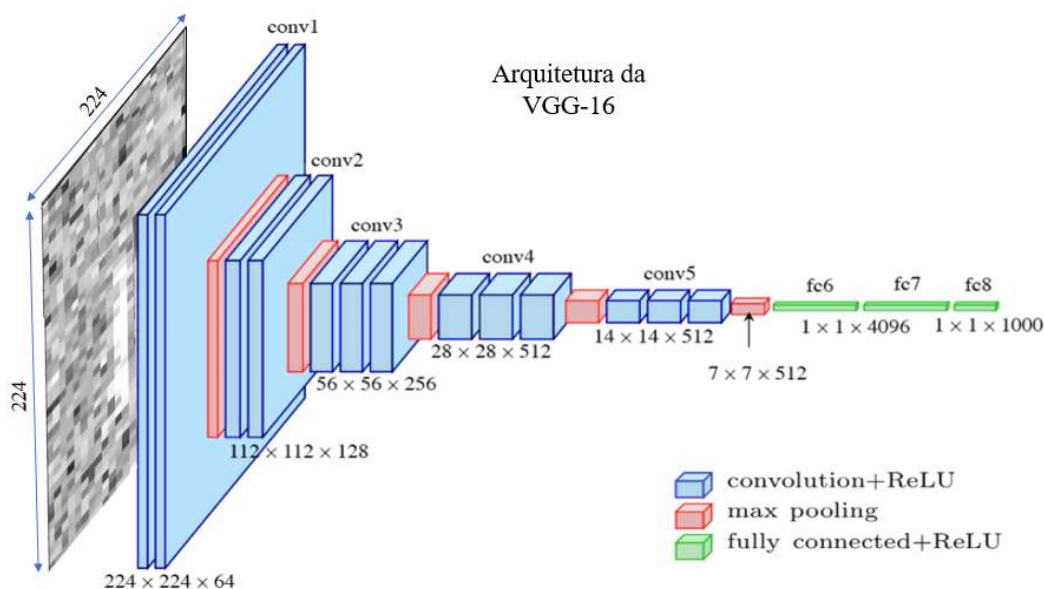


Fig. 1. Arquitetura da VGG-16. Adaptado de [26].

próximos, a amostra é posicionada no espaço de atributos e a distância para todas as amostras de treinamento é calculada. O espaço de atributos é organizado tendo em cada dimensão (eixo) um dos atributos. Os parâmetros são o tipo de cálculo de distância, o número de vizinhos (k) e peso de cada vizinho para comparação [24].

A Fig. 2 ilustra a classificação de uma amostra por meio de kNN com o uso de distância Euclidiana, número de vizinhos igual a 3 e peso uniforme.

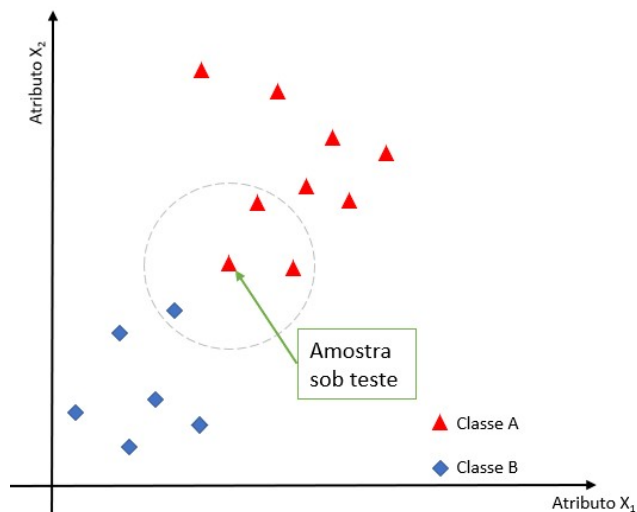


Fig. 2. Classificação utilizando o kNN, com k igual 3, distância Euclidiana e peso igual para todos vizinhos.

III. MATERIAIS E MÉTODOS

Em [7], foram utilizadas 400 amostras na polarização VH, obtidas de imagens do satélite SAR Sentinel-1 (produto GRD), com resolução de aproximadamente 20m em range x 22m em azimuth, pixel spacing de 10m x 10m, número de looks de 5 em range para 1 em azimuth e Número Equivalente de Looks (ENL) de 4,4. Alguns exemplos de amostras são apresentados na Fig. 3.

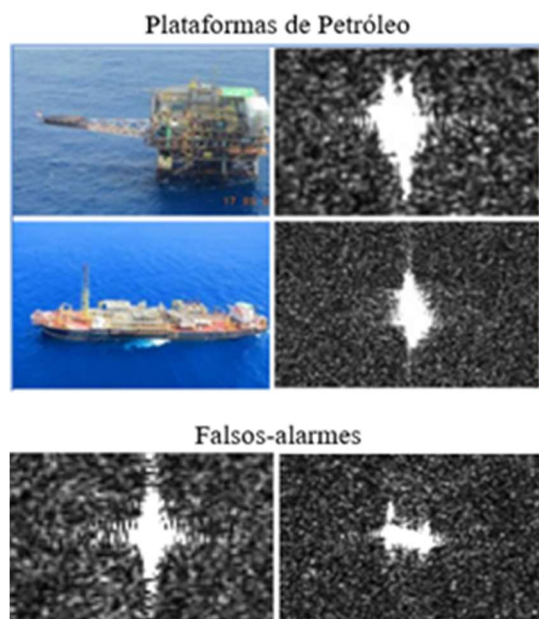


Fig. 3. Exemplos de amostras obtidas do banco de dados montado por [7].

Essas mesmas amostras, bem como os mesmos atributos extraídos pela VGG-16, foram utilizadas para compor o

conjunto de treinamento e teste para o algoritmo kNN. Nota-se na Fig. 3 que as imagens SAR são muito parecidas. Isso torna a tarefa determinar visualmente se a amostra é de uma plataforma ou não muito complexa para um especialista.

Em cada rodada, 80 amostras eram selecionadas aleatoriamente para compor o conjunto de teste balanceado entre plataformas e falsos-alarms, enquanto as outras 320 compunham o conjunto de treinamento também balanceado. Por balanceado, entende-se que os conjuntos são compostos por 50% de plataformas e 50% de falsos-alarms. Esse sorteio dos conjuntos de treino e teste foi repetido 50 vezes. Cada repetição é considerada um bloco de experimentação, de forma que a variância proveniente da seleção de amostras não afete a comparação dos tratamentos.

Cada bloco de treino e teste foi submetido ao kNN com quatro níveis para o fator número de vizinhos, ou seja, quatro tratamentos: 3 (parâmetros empregado em [7]), 7, 15 e 31. Para todos os tratamentos, foi utilizada a distância Euclidiana e o peso uniforme.

A análise foi feita com os indicadores AUC e Acurácia [25], inicialmente de forma gráfica por meio da visualização em *boxplots* da distribuição dos resultados. Posteriormente, foi feita a ANOVA, a análise de resíduos e o teste de Tukey.

Algumas informações sobre *Recall* foram extraídas para verificar se o algoritmo tende a ser conservador ou liberal, considerando as plataformas de petróleo como alvo de interesse [19].

IV. RESULTADOS E DISCUSSÃO

Para os indicadores AUC, Acurácia e *Recall*, os valores de média, mediana, variância, desvio-padrão e o *p-value* do Teste de Shapiro-Wilk são apresentados na Tabela I:

TABELA I. INFORMAÇÕES ESTATÍSTICAS PARA CADA INDICADOR.

AUC				
Número de k vizinhos	k=3	k=7	k=15	k=31
Média	0,821	0,866	0,874	0,864
Mediana	0,824	0,862	0,870	0,868
Variância	0,002	0,001	0,002	0,002
Desvio-Padrão	0,043	0,035	0,040	0,042
P-Value - Teste de Shapiro-Wilk	0,888	0,725	0,726	0,890

ACURÁCIA				
Número de k vizinhos	k=3	k=7	k=15	k=31
Média	0,767	0,784	0,786	0,762
Mediana	0,775	0,775	0,787	0,762
Variância	0,001	0,001	0,002	0,002
Desvio-Padrão	0,038	0,038	0,042	0,048
P-Value - Teste de Shapiro-Wilk	0,361	0,374	0,623	0,380

RECALL				
Número de k vizinhos	k=3	k=7	k=15	k=31
Média	0,800	0,831	0,840	0,827
Mediana	0,800	0,825	0,850	0,825
Variância	0,003	0,003	0,004	0,003
Desvio-Padrão	0,054	0,059	0,064	0,054
P-Value - Teste de Shapiro-Wilk	0,059	0,179	0,244	0,150

A Fig. 4 e a Fig. 5 apresentam graficamente as distribuições dos resultados para os indicadores AUC e Acurácia para cada um dos tratamentos:

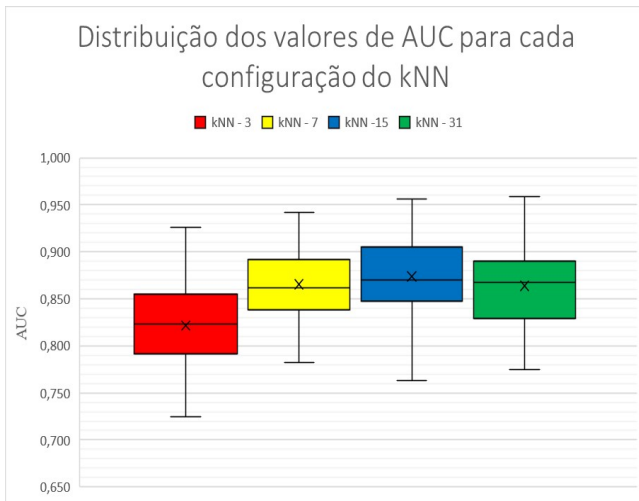


Fig. 4. Resultados do indicador AUC por número de vizinhos.

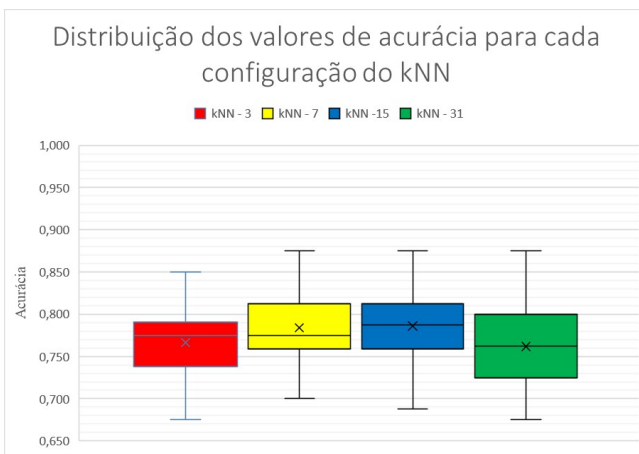


Fig. 5. Resultados do indicador Acurácia por número de vizinhos.

A Tabela II apresenta os resultados da ANOVA tanto para o indicador AUC, quanto para o indicador Acurácia, mostrando que há diferença estatística (significância de 5%) para as médias das distribuições das Fig. 4 e Fig. 5.

TABELA II. ANOVA PARA AUC E PARA ACURÁCIA.

ANOVA (AUC)					
Fonte da variação	SQ	gl	MQ	F	valor-P
Tratamentos	0,083	3	0,0278	68,34	2,00E-16
Blocos	0,258	49	0,0053		
Resíduos	0,06	147	0,0004		
Total	0,401	199			
ANOVA (Acurácia)					
Fonte da variação	SQ	gl	MQ	F	valor-P
Tratamentos	0,022	3	0,0073	10,07	4,47E-06
Blocos	0,234	49	0,0048		
Resíduos	0,107	147	0,0007		
Total	0,363	199			

A análise dos resíduos da ANOVA para o indicador AUC apresenta o valor-p para o teste de normalidade de 0,173 (Teste de Shapiro-Wilk - $W = 0,9899$), e para o teste de

homogeneidade das variâncias de $4,8E-05$ (Teste de Bartlett).

A análise dos resíduos da ANOVA para o indicador Acurácia apresenta o valor-p para o teste de normalidade de 0,9503 (Teste de Shapiro-Wilk - $W = 0,99672$), e para o teste de homogeneidade das variâncias de 0,0153 (Teste de Bartlett).

O Teste de Tukey (significância de 5%) de ambos os indicadores é apresentado na Tabela III.

TABELA III. TESTE DE TUKEY PARA AUC E ACURÁCIA.

Tratamento	IC-95% AUC	Grupos AUC	IC-95% Acurácia	Grupos Acurácia
kNN - 15	[0,867 - 0,881]	a	[0,776 - 0,796]	a
kNN - 7	[0,858 - 0,873]	a	[0,774 - 0,794]	a
kNN - 31	[0,857 - 0,871]	a	[0,752 - 0,772]	b
kNN - 3	[0,814 - 0,829]	b	[0,757 - 0,776]	b

A primeira constatação é de que os classificadores necessitam de uma parametrização adequada, pois a ANOVA tanto para Acurácia, quanto para AUC são significativas, demonstrando que há diferença na média dos tratamentos. Houve melhora de 6,46% na AUC, 5% no Recall e 3,14% na Acurácia, considerando o tratamento com pior e com o melhor desempenho.

O fato de o Recall apresentar uma melhora percentual de praticamente o dobro da Acurácia representa que a parametrização impactou mais o reconhecimento das plataformas de petróleo do que o dos falsos-alarmes. O Recall é a proporção de plataformas corretamente previstas com o total dessa classe existente no teste. A Acurácia considera os acertos de ambas as classes em seus cálculos.

Ressalta-se que o classificador tem um comportamento liberal, ou seja, acerta bastante a classe positiva (plataformas de petróleo), mas também tende a errar mais a classe negativa (falsos-alarmes). Ao comparar a tabela de Recall com a Acurácia, percebe-se que primeiro indicador é maior que o segundo em todos os tratamentos, uma característica para um classificador liberal, quando atuando em classificação binária balanceada.

Foram testadas as quantidades de vizinhos que variam de 0,93% da massa de treinamento ($k=3$) a 9,69% ($k=31$). A quantidade de vizinhos influencia como o classificador tende a perceber a topologia do espaço de atributos. Quanto maior, mais tende-se a classificar de acordo com um comportamento regional dentro do espaço de atributos, pesando menos os detalhes locais. Assim, quanto maior k , supõe-se que os ruídos afetem pouco a classificação, porém tende-se a suavizar as bordas dos limites entre as classes. Com isso, classes não-linearmente separáveis podem ser de difícil distinção ao utilizar um k muito grande em termos proporcionais ao conjunto de treinamento.

Para os resultados do experimento, percebe-se que $k=3$ apresenta o pior resultado, com média igual 0,0767. Portanto, o número muito pequeno de vizinhos provavelmente sofreu influência dos ruídos locais.

Ao utilizar o $k=31$, não se percebe melhora significativa. Na verdade, a média de todos os indicadores começa a cair. Pode-se inferir que as classes possuem limites complexos. Para um número de vizinhos muito grande, o classificador

deixa de ser sensível aos detalhes das fronteiras entre as classes. Além disso, um número maior de vizinhos exige um poder computacional maior, devido a maior quantidade de comparações necessárias.

Para essa base de dados, provavelmente o número de vizinhos ideal estaria entre 7 e 31, ou seja, em torno de 5% do conjunto de treinamento. Quanto aos resultados da referência [7], notou-se uma melhoria de 5,9% na AUC, passando de 0,825 para 0,874, confirmando que a parametrização do algoritmo oferece mudanças significativas em termos operacionais e estatísticos.

V. CONCLUSÃO

Nesse trabalho, continuou-se uma pesquisa que utilizou Redes Neurais Convolucionais para extrair os atributos de imagens SAR de média resolução de plataformas de petróleo (alvos de interesse) e falsos-alarmes. O objetivo era verificar se a parametrização do classificador que processava tais atributos tinha influência significativa no desempenho do reconhecimento automático desses alvos artificiais.

O classificador empregado foi o kNN e testou-se o parâmetro número de vizinhos em quatro níveis ($k = [3, 7, 15 \text{ e } 31]$). Chegou-se à conclusão de que há diferença na performance, sendo que os tratamentos com 7 e 15 vizinhos alcançaram as melhores médias para os três indicadores utilizados. Percebeu-se que o kNN tem um comportamento liberal para o tipo de alvo classificado.

Como trabalhos futuros, sugere-se um estudo mais detalhado da parametrização do kNN, bem como dos demais algoritmos empregados na pesquisa de base desse trabalho, para que se possa atingir o estado-da-arte em algoritmos ATR usando imagens de média resolução. Pode-se testar também o kNN associado à VGG-16 em outras bases de dados.

Contribuiu-se para a melhoria de sistemas de vigilância de grandes áreas, de forma que o controle da Amazônia Azul, por exemplo, seja beneficiado pelo emprego de técnicas ATR que usam a abundância de dados gerada pelos sensores espaciais.

REFERÊNCIAS

- [1] Comando da Aeronáutica, Ministério da Defesa, and Brasil, "Dimensão 22," 2018. [Online]. Available: <http://www.fab.mil.br/dimensao22>. [Accessed: 02-Apr-2018].
- [2] IEEE GRSS Beijing Chapter and RADL, "Proceedings of 2017 SAR in Big Data Era: Models, Methods and Applications, BIGSAR DATA 2017," *Proc. 2017 SAR Big Data Era Model. Methods Appl. BIGSAR DATA 2017*, vol. 2017-Janua, p. 2017, 2017.
- [3] F. Palazzo *et al.*, "RUS: A New Expert Service for Sentinel Users," *Proceedings*, vol. 2, no. 7, p. 369, 2018.
- [4] A. G. Castriotta and R. Knowelden, "COPE-SERCO-RP-17-0186: Sentinel Data Access 2017 Annual Report," Frascati, 2018.
- [5] J. Blumenfeld, "Getting Ready for NISAR — and for Managing Big Data using the Commercial Cloud," *ASF News*, 2018. [Online]. Available: <https://www.asf.alaska.edu/news-notes/winter-2017-1/getting-ready-for-nisar/>.
- [6] A. Marino, D. Velotto, and F. Nunziata, "Offshore Metallic Platforms Observation Using Dual-Polarimetric TS-X/TD-X Satellite Imagery: A Case Study in the Gulf of Mexico," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 10, no. 10, pp. 4376–4386, Oct. 2017.
- [7] L. E. Falqueto, J. A. S. Sa, R. L. Paes, and A. Passaro, "Oil Rig Recognition Using Convolutional Neural Network on Sentinel-1 SAR Images," *IEEE Geosci. Remote Sens. Lett.*, pp. 1–5, 2019.
- [8] PETROBRAS, "Tipos de plataformas," 2018. [Online]. Available: <http://www.petrobras.com.br/infograficos/tipos-de-plataformas/desktop/index.html>. [Accessed: 15-Jun-2018].
- [9] D. Zhang, J. Liu, W. Heng, K. Ren, and J. Song, "Transfer Learning with Convolutional Neural Networks for SAR Ship Recognition," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 322, no. 7, p. 072001, Mar. 2018.
- [10] S. Chen, H. Wang, F. Xu, and Y. Q. Jin, "Target Classification Using the Deep Convolutional Networks for SAR Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4806–4817, 2016.
- [11] M. Wilmanski, C. Kreucher, and J. Lauer, "Modern approaches in deep learning for SAR ATR," in *Modern approaches in deep learning for SAR ATR, Proc. SPIE 9843*, 2016, no. 98430N, p. 10.
- [12] J. Pei, Y. Huang, W. Huo, Y. Zhang, J. Yang, and T.-S. Yeo, "SAR Automatic Target Recognition Based on Multiview Deep Learning Framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 1–15, 2017.
- [13] H. Furukawa, "SAVERS: SAR ATR with Verification Support Based on Convolutional Neural Network," pp. 23–28, 2018.
- [14] F. Gao, T. Huang, J. Sun, J. Wang, A. Hussain, and E. Yang, "A New Algorithm of SAR Image Target Recognition Based on Improved Deep Convolutional Neural Network," *Cognit. Comput.*, vol. 18, no. 1, pp. 25–30, Jun. 2018.
- [15] F. Gao, Y. Yang, J. Wang, J. Sun, E. Yang, and H. Zhou, "A Deep Convolutional Generative Adversarial Networks (DCGANs)-Based Semi-Supervised Method for Object Recognition in Synthetic Aperture Radar (SAR) Images," *Remote Sens.*, vol. 10, no. 6, p. 846, May 2018.
- [16] E. G. Zelnio, M. Levy, R. D. Friedlander, and E. Sudkamp, "Deep learning model-based algorithm for SAR ATR," *Algorithms Synth. Aperture Radar Imag. XXV*, no. May, p. 10, 2018.
- [17] C. Wang, H. Zhang, F. Wu, B. Zhang, and S. Tian, "Ship classification with deep learning using COSMO-SkyMed SAR data," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 558–561.
- [18] C. Bentes, D. Velotto, and B. Tings, "Ship Classification in TerraSAR-X Images With Convolutional Neural Networks," *IEEE J. Ocean. Eng.*, vol. 43, pp. 258–266, Jan. 2017.
- [19] D. C. Montgomery, *Design and Analysis of Experiments*, Eighth. Arizona: John Wiley & Sons, Inc., 2013.
- [20] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *Inf. Softw. Technol.*, vol. 51, no. 4, pp. 769–784, Sep. 2014.
- [21] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [22] J. Demšar *et al.*, "Orange: Data Mining Toolbox in Python," *J. Mach. Learn. Res.*, vol. 14, pp. 2349–2353, 2013.
- [23] O. Russakovsky *et al.*, "ImageNet Large Scale Visual Recognition Challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
- [24] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Machine learning: a review of classification and combining techniques," *Artif. Intell. Rev.*, vol. 26, no. 3, pp. 159–190, Nov. 2006.
- [25] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006.
- [26] M. Ferguson, R. Ak, Y.-T. T. Lee, and K. H. Law, "Automatic localization of casting defects with convolutional neural networks," no. December, pp. 1726–1735, 2018.