

Aprendizagem Automática para Detecção de Embarcações em Sistemas Radar [★]

Tomás Pedro^{1,2}, João Cintra², Nuno Garcia¹, Alan Oliveira de Sá¹

¹LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal

²GMVIS SKYSOFT, 1990-392 Lisboa, Portugal

alan@di.fc.ul.pt, nrgarcia@ciencias.ulisboa.pt tpedro@gmv.com,
joao.cintra@gmv.com

Keywords: Autoencoder Convolutacional · YOLO · Radar de Superfície.

1 Introdução

A deteção e localização de embarcações é uma tarefa desafiadora, sobretudo devido ao ruído e interferência em sinais de radares de busca de superfície, originados por ondas do mar, chuva e presença de massas de terra. Estas perturbações geram ecos indesejados que podem ocultar embarcações, sendo particularmente proeminentes em condições climáticas adversas. A interferência marítima, causada pelas reflexões dos sinais radar nas ondas, aumenta em cenários com ondulações maiores, variando significativamente em intensidade, tempo e distribuição espacial. Tal complexidade apresenta um desafio na criação de um sistema robusto de aprendizagem automática para a deteção confiável de embarcações na presença dessas interferências.

Radares modernos geralmente possuem Sistema de Identificação Automática (*Automatic Identification System* - AIS), que permite às embarcações transmitir informações como posição, rumo e velocidade [6]. No entanto, mesmo o AIS apresenta limitações como as embarcações poderem desativar o AIS, o sistema não estar a funcionar corretamente, ou as embarcações podem não o terem equipado. Também são utilizados recursos como técnicas de *Moving Target Detection* (MTD) [7], que se baseia principalmente em métodos de processamentos de sinais [8], no entanto este não é capaz de eliminar eficazmente o ruído de mar.

Dadas as limitações dos sistemas radar, este trabalho sugere técnicas de aprendizagem automática para aprimorar a deteção de embarcações em cenários com ruído de mar. As contribuições deste trabalho são o uso do Autoencoder Convolutacional (*Convolutional Autoencoder* - CAE) para reduzir o ruído de mar nas imagens de radar, e a implementação e comparação de duas técnicas para detetar embarcações em imagens com ruído reduzido, pré-processadas pelo autoencoder. A primeira técnica estabelece um *threshold* para a deteção de embarcações, e a segunda aplica o algoritmo *You Only Look Once* versão 8 (YOLOv8) [3,5] às imagens pré-processadas pelo autoencoder.

[★] Este trabalho foi suportado pela FCT através do LASIGE Research Unit, ref. UIDB/00408/2020 and ref. UIDP/00408/2020.

2 Metodologia

A metodologia proposta neste trabalho para a detecção de embarcações envolve várias etapas: construção de um *dataset*, redução de ruído, e a aplicação de técnicas de detecção. O *dataset* é obtido da plataforma Shiplocus [1], consistindo em imagens de radar de banda X de uma zona costeira, com adicional informação das condições meteorológicas, marítimas e AIS para identificar embarcações. As imagens possuem uma resolução de 2048×2048 pixels e são anotadas com o software Label Studio [2,4].

Após a obtenção das imagens, realizou-se uma segmentação das mesmas, originando segmentos de 32×32 pixels. Esta etapa foi especialmente relevante devido à esparsidade dos dados, visto que grande parte dos pixels nas imagens possui uma intensidade nula ou próxima de zero. Através da segmentação, o Autoencoder Convolutivo (CAE) foi capaz de se concentrar mais eficazmente em detalhes locais, otimizando a extração de características.

Com um total de 194 imagens, o *dataset* foi dividido para treino (70%), teste (15%) e validação (15%). A criação deste *dataset* específico foi motivada pela falta de *datasets* públicos de radar de superfície. A informação AIS, um sistema automático de identificação usado pelas embarcações para transmitir dados relevantes para a segurança da navegação, serviu para confirmar a presença de embarcações nas imagens.

A etapa seguinte envolve a utilização de um Autoencoder Convolutivo (CAE) [9] para minimizar o ruído nas imagens de mar. O CAE é um modelo de aprendizagem profunda com duas partes: um codificador e o decodificador. O codificador contém camadas convolucionais seguidas por camadas de *maxpooling*. As camadas convolucionais extraem características locais e estruturas dos dados de entrada, enquanto as camadas de *maxpooling* reduzem as dimensões espaciais dessas características, resultando numa representação compacta dos dados de entrada. O decodificador consiste em camadas de *upsampling* seguidas por camadas convolucionais. As camadas de *upsampling* aumentam as dimensões espaciais da representação compacta, enquanto as camadas convolucionais são utilizadas para reconstruir a imagem original a partir desta representação expandida. O modelo é treinado com a métrica Erro Médio Quadrático (*Mean Squared Error* - MSE) para produzir imagens reconstruídas que se assemelham o máximo possível às originais.

O treino é feito exclusivamente com imagens de áreas marítimas sem embarcações, de forma a aprender a reconstruir imagens "limpas". Quando uma embarcação está presente numa imagem, surgem discrepâncias entre a imagem original e a imagem reconstruída pelo CAE, pois este não está adequadamente treinado para reconhecer embarcações como elementos "normais" das imagens. Essas discrepâncias, consideradas "anomalias" em relação ao padrão aprendido pelo modelo durante o treino, são isoladas pela subtração da imagem reconstruída da imagem original. Este processo de subtração destaca as embarcações nas imagens de radar, minimizando a presença de ruído e interferências. O resultado desta subtração serve como entrada para o métodos de detecção de embarcações com o algoritmo YOLO.

Na primeira abordagem, após o treino, o CAE é utilizado para gerar imagens reconstruídas dos sinais de radar de entrada. Nessas imagens, as embarcações são considerados anomalias. A detecção de embarcações é realizada nos segmentos anteriormente criados. Para cada um desses segmentos, é calculado o erro médio quadrático do segmento (MSE_S), que corresponde à diferença entre o segmento da imagem reconstruída pelo CAE e o mesmo segmento da imagem original. Com base no MSE_S calculado, utiliza-se, então, um *threshold* τ para concluir sobre a existência ou não de uma embarcação no segmento de imagem S analisado:

- Se $MSE_S < \tau$, então considera-se que não há embarcação em S ;
- Se $MSE_S \geq \tau$, então considera-se que há embarcação em S .

Para calcular τ , o *F1 score* é utilizado. Esta medida combina as métricas de *precision* (P) e *recall* (R), proporcionando uma visão mais equilibrada do desempenho do modelo. O *F1 score* é obtido a partir dessas duas métricas e o *threshold* τ é ajustado de forma a maximizar *F1* (1). Sendo P e R calculados de acordo com (2) e (3), com base nos verdadeiros positivos (VP), falsos positivos (FP) e falsos negativos (FN):

$$F1 = 2 \cdot \frac{P \cdot R}{P + R}. \quad (1) \quad P = \frac{VP}{VP + FP}, \quad (2) \quad R = \frac{VP}{VP + FN}. \quad (3)$$

O *threshold* τ é então ajustado de forma a maximizar *F1*. Este método permite que as embarcações sejam destacadas e distinguidas do ruído marítimo e de outros elementos presentes nas imagens de radar.

Na segunda abordagem, é utilizado o algoritmo YOLOv8 para detetar embarcações nas imagens radar pré-processadas pelo CAE. O YOLOv8 [3] é o mais recente modelo do conjunto de algoritmos de detecção de objetos YOLO [5]. Estes são reconhecidos pelo seu desempenho em tarefas de detecção de objetos, classificação de imagens e segmentação.

Para treinar o YOLOv8, as anotações fornecidas com o conjunto de dados, que marcam a localização e classificação das embarcações nas imagens, são adaptadas para o formato requerido pelo algoritmo. Posteriormente, o YOLOv8 é sujeito a um processo de *fine tuning*. Este processo envolve re-treinar o modelo já previamente treinado com um *dataset* maior e mais variado, com um conjunto de dados específico, permitindo assim ao YOLOv8 ajustar os seus pesos ao problema específico deste estudo.

Na fase de inferência, o YOLOv8 analisa as imagens de radar e retorna uma lista de detecções. Cada detecção é acompanhada por uma caixa delimitadora que neste caso indica a localização das embarcações nas imagens de radar.

3 Resultados

Os resultados obtidos indicam que a aplicação do CAE resultou numa redução do ruído nas imagens de radar. Após o processamento pelo CAE, verifica-se uma diminuição do ruído provocado pelas ondas do mar. Para quantificar a melhoria

proporcionada pelo CAE, a eficácia da redução de ruído é avaliada através da relação sinal-ruído (SNR - *Signal-to-Noise Ratio*), métrica amplamente usada para medir a relação entre a potência do sinal desejado (neste caso, embarcações) e a potência do ruído de fundo (resto da imagem). A SNR é calculada de acordo com a Equação (4):

$$SNR = \frac{P_{sinal}}{P_{ruído}}, \quad (4)$$

onde P_{sinal} é a potência do sinal (média dos valores dos pixels correspondentes às embarcações) e $P_{ruído}$ é a potência do ruído (média dos valores dos pixels do resto da imagem).

Nas imagens originais, a média de SNR é de 7.09, enquanto nas imagens com ruído reduzido a média de SNR é de 8.60. Isto confirma que, após a aplicação do CAE, há uma redução de ruído de mar e que esta redução é maior do que o desvanecimento das imagens de embarcações. Em outras palavras, os sinais de embarcações passam a ter mais destaque em relação ao ruído de mar.

A detecção de embarcações, efetuada em imagens pré-processadas pelo CAE, recorreu a dois métodos: um baseado em τ , e outro que emprega o algoritmo YOLOv8. Neste contexto, o *threshold* maximizando $F1$ resultou em $\tau = 4,3 \times 10^{-4}$. Para a avaliação de desempenho de cada método, as métricas de *precision* e *recall* foram utilizadas e estão detalhadas na tabela 1.

Método	Precison	Recall
CAE + <i>threshold</i>	0.6111	0.3503
CAE + YOLOv8	0.764	0.637

Tabela 1. Comparação do método baseado em limiar e método YOLOv8

Os resultados apresentados na tabela 1 mostram o melhor desempenho do algoritmo YOLOv8 em relação à detecção baseada no τ , tanto em termos de *precision* quanto de *recall*. Essa distinção indica que, na detecção de embarcações em imagens de radar, o YOLOv8 é a opção mais eficiente.

4 Conclusões

Este trabalho destaca a eficácia dos Autoencoders Convolucionais (CAE) para aprimorar a relação sinal-ruído em imagens de radar, facilitando a detecção de embarcações em ambientes marítimos. A combinação de CAE e YOLOv8 demonstrou ser uma abordagem promissora, superando um método baseado em *threshold* em termos de *precision* e *recall*. Futuramente, pretende-se testar estes modelos sob diferentes condições marítimas e explorar características temporais nos sinais de radar, considerando a implementação de uma LSTM para diferenciar mais eficazmente os ecos estáveis das embarcações dos ecos variáveis originados por condições climáticas e marítimas. Adicionalmente, planeia-se também aumentar o tamanho do conjunto de dados utilizado.

Referências

1. GMV: Shiplocus® (Nov 2022), <https://www.gmv.com/pt-pt/produtos/defesa-e-seguranca/shiplocus#shiplocus-gmv>, Acedido em 10 de junho de 2023
2. Heartex: Open source data labeling, <https://labelstud.io/>, Acedido em 10 de junho de 2023
3. Jocher, G., Chaurasia, A., Qiu, J.: YOLO by Ultralytics (Jan 2023), <https://github.com/ultralytics/ultralytics>
4. Labs, H.: Heartexlabs/label-studio: Label studio is a multi-type data labeling and annotation tool with standardized output format, <https://github.com/heartexlabs/label-studio>, Acedido em 10 de junho de 2023
5. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (June 2016)
6. Series, M.: Technical characteristics for an automatic identification system using time-division multiple access in the vhf maritime mobile band. Recommendation ITU: Geneva, Switzerland pp. 1371–1375 (2014)
7. Yan, H., Hou, Q., Zhang, J., Wang, L., Zhang, G., Zhu, D.: Scheme to implement moving target detection of coastal defense radar in complicated sea conditions. Journal of Applied Remote Sensing **16**(4), 046510 (2022)
8. Zhang, S., Zhou, Y., Sha, M., Zhang, L., Du, L.: Moving multitarget detection using a multisite radar system with widely separated stations. Remote Sensing **14**(11), 2660 (2022)
9. Zhang, Y.: A better autoencoder for image: Convolutional autoencoder. In: ICONIP17-DCEC. Available online: <http://users.cecs.anu.edu.au/Tom.Gedeon/conf/ABCs2018/paper/ABCs2018-paper.58.pdf> (accessed on 23 March 2017) (2018)