

Aplicando o Modelo YOLOv11 para Detectar e Segmentar Viveiros de Camarão em Imagens Georreferenciadas

Alisson C. Ferreira¹, Luís J. R. Oliveira², Ana Carla C. da Silva¹,
Bruno S. Ursulino¹, Davidson A. Nunes¹, Danielo G. Gomes³, Raimundo V. C. F.¹

¹Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

²Universidade do Estado do Rio Grande do Norte (UERN)

³Programa de Pós-Graduação em Engenharia de Teleinformática,
Grupo de Redes, Engenharia de Software e Sistemas (GREat)
Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza - CE

{chavesalisson.a, luis.jeronimo044, anacarla162020}@gmail.com

danielo@ufc.br, {bruno.ursulino, nunes.davidson, valter.costa}@ifce.edu.br

Abstract. *The monitoring of shrimp farms is essential for tracking aquaculture production and identifying new cultivation areas. This study proposes a machine learning-based approach using the YOLOv11 architecture to detect and monitor farms in the states of Ceará and Rio Grande do Norte. A dataset of about 200 georeferenced images from the Copernicus[®] platform was created, covering the Red (R), Green (G), Blue (B), and Near-infrared (NIR) spectra, with annotations made in Roboflow[®]. The model was trained using the hold-out method (80% training, 20% validation) and evaluated using precision, recall, mAP50, and mAP50-95 metrics. The results confirm the effectiveness of the approach in identifying shrimp farms and detecting new cultivation areas, contributing to aquaculture monitoring and policy development.*

Resumo. *O monitoramento de viveiros de camarão é essencial para acompanhar a produção aquícola e identificar novas áreas de cultivo. Neste artigo, propomos uma abordagem baseada em aprendizado de máquina utilizando o modelo YOLOv11 para detectar e monitorar viveiros nos estados do Ceará e Rio Grande do Norte. Criamos um conjunto de aproximadamente 200 imagens georreferenciadas da plataforma Copernicus[®], abrangendo os espectros Red (R), Green (G), Blue (B), e Near-infrared (NIR), com anotações na plataforma Roboflow[®]. O modelo foi treinado com hold-out (80% treino, 20% validação) e avaliado pelas métricas precision, recall, mAP50 e mAP50-95. Os resultados ilustram a eficácia da abordagem na identificação de viveiros e na detecção de novas áreas de cultivo, contribuindo para o monitoramento aquícola e a formulação de políticas públicas.*

1. Introdução

A aquicultura tem se consolidado como um setor estratégico para a economia brasileira, impulsionando o crescimento da produção de pescado e camarão. Em 2023, a produção de camarão em cativeiro alcançou 127,5 mil toneladas, representando um crescimento de 13% em relação ao ano anterior e gerando uma receita de R\$ 2,63 bilhões. O Nordeste

lidera essa produção, sendo responsável por 99,6% do total nacional, com destaque para os estados do Ceará (57%) e Rio Grande do Norte (19,4%) [Igarashi 2022, IBGE 2023].

Diante desse cenário de crescimento, o monitoramento eficiente dos viveiros torna-se essencial para a sustentabilidade e planejamento do setor. Métodos tradicionais de fiscalização apresentam limitações devido à vasta extensão das áreas aquícolas, ao difícil acesso a determinadas regiões e à variabilidade ambiental. Como alternativa, imagens de satélite e técnicas de sensoriamento remoto têm se mostrado promissoras para o mapeamento e monitoramento sistemático dessas áreas. Além disso, ferramentas de aprendizado de máquina podem otimizar esse processo de monitoramento, proporcionando dados precisos para decisões estratégicas e formulação de políticas públicas [Werneck et al. 2023].

Modelos avançados de visão computacional, como os da família *You Only Look Once* (YOLO), têm mostrado eficiência na detecção e monitoramento de imagens de alta resolução [Gomes 2022, Redmon et al. 2016, Jocher and Qiu 2024]. É possível detectar áreas aquícolas avaliando a influência da resolução espacial [Dong et al. 2024], identificar gaiolas de cultivo de peixes [Quaade et al. 2024], segmentar e classificar viveiros aquícolas com variações temporais dessas áreas [Greenstreet et al. 2023].

Embora abordagens baseadas em YOLO tenham se mostrado eficazes na detecção de estruturas aquícolas, a qualidade das imagens e a necessidade de refinamento permanecem desafios [Dong et al. 2024, Quaade et al. 2024]. Alternativamente, técnicas de segmentação, aliadas a métodos de pré-processamento, oferecem maior detalhamento e adaptabilidade [Greenstreet et al. 2023].

Neste artigo, aplicamos o modelo YOLOv11 para detectar e segmentar de forma automática viveiros de camarão em imagens georreferenciadas da plataforma *Copernicus*[®].

2. Material e Método

A pesquisa envolveu a criação de um *dataset* segmentado contendo imagens de viveiros de camarão na região do Vale do Jaguaribe-CE, identificados nos espectros *Red* (R), *Green* (G), *Blue* (B) e *Near-infrared* (NIR). Para garantir uma abordagem estruturada, seguimos um *pipeline* sistemático voltado para a captação, organização e análise das informações necessárias à detecção e monitoramento dessas áreas aquícolas nos estados do Ceará e Rio Grande do Norte.

A metodologia adotada compreendeu cinco etapas: (i) obtenção e pré-processamento das imagens de sensoriamento remoto, (ii) construção do conjunto de dados para a tarefa de segmentação, (iii) adaptação do modelo *You Only Look Once* versão 11 (YOLOv11) para suportar um maior número de canais, (iv) treinamento do modelo com os dados processados e (v) avaliação do desempenho considerando métricas padronizadas.

As imagens das bandas R, G, B e NIR foram adquiridas a partir da plataforma *Copernicus*[®]. Para facilitar o processamento e garantir um conjunto balanceado entre imagens contendo e não contendo viveiros, essas imagens foram segmentadas em recortes de 200×200 pixels. Em seguida, os dados foram divididos em dois conjuntos: treinamento (78,9%) e validação (21,1%), sendo utilizados para treinar o algoritmo

YOLOv11. Durante o treinamento, monitoramos diversas métricas de desempenho para avaliar a eficácia do modelo. Entre elas, destacam-se *precision*, *recall*, *mean Average Precision* (mAP), considerando os cenários *mAP*50 (precisão média com *Intersection over Union* (IoU) mínimo de 50%) e *mAP*50-95 (média ponderada das precisões para IoUs variando entre 50% e 95% em intervalos de 5%), além do *F1-score*. Essas métricas permitiram uma análise detalhada da capacidade do modelo em detectar e segmentar corretamente os viveiros de camarão a partir das imagens georreferenciadas.

2.1. Área Estudada e Conjunto de Dados

O conjunto de dados deste estudo foi desenvolvido para detectar viveiros de camarão em imagens de sensoriamento remoto. A aquisição ocorreu via interpretação visual da cena T24MXV, capturada em 1º de outubro de 2024, às 12 : 53 : 09 *Universal Time Coordinated* (UTC), pela missão Sentinel-2B da *European Space Agency* (ESA), com o sensor *MultiSpectral Instrument* (MSI). A imagem, disponível na plataforma *Copernicus*®, está no formato JPEG2000, com resolução de 10 metros por pixel, profundidade de cor de 15 bits por canal e cobertura de nuvens de 3.55%. O produto utilizado é do nível de processamento L1C (Level 1C), órbita absoluta 39546 e relativa 52, com sistema de referência EPSG:32724 (*Universal Transverse Mercator* (UTM), zona 24S). A área analisada abrange as bacias dos rios Jaguaribe (médio e baixo) e Apodi-Mossoró (Figura 1). Extraída de uma região original de 110 × 110 km, a imagem¹ passou por ortorretificação baseada no modelo digital de elevação *Defence Gridded Elevation Data - 90m resolution* - (DEM GLO-90), resultando em um recorte trapezoidal de aproximadamente 106 km de base maior, 82,3 km de base menor e 110 km de altura, garantindo um ajuste preciso à projeção UTM.

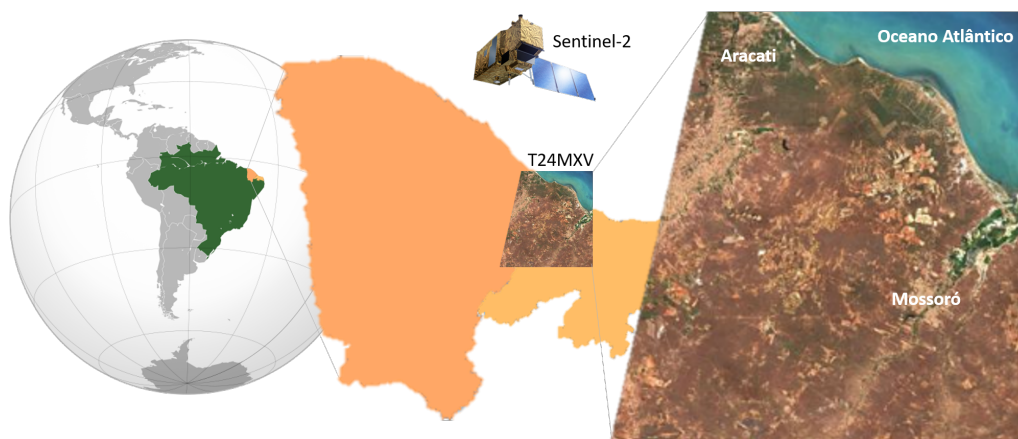


Figura 1. Localização da região analisada - Fonte: *Wikipedia*® e *Copernicus*®.

O *dataset* foi gerado a partir da imagem georreferenciada obtida na plataforma *Copernicus*® com quatro bandas espectrais: R (665 ηm), G (560 ηm), B (490 ηm) e NIR (842 ηm). Para garantir balanceamento, foram selecionadas 125 amostras com viveiros e 125 sem (*background*). As imagens contendo viveiros foram segmentadas manualmente na plataforma *Roboflow*®² conforme Figura 2 (a), gerando anotações no

¹[https://download.dataspace.copernicus.eu/odata/v1/Products\(b01abb74-210d-432e-8679-ee56efe66d41\)/\\$value](https://download.dataspace.copernicus.eu/odata/v1/Products(b01abb74-210d-432e-8679-ee56efe66d41)/$value)

²<https://roboflow.com/>

formato *annotation polygon* com coordenadas relativas a recortes de 200×200 pixels (Figura 2). O formato de exportação das anotações é compatível com a YOLOv11, permitindo sua integração direta ao *pipeline* de treinamento. Além disso, um arquivo auxiliar foi gerado para referenciar geograficamente cada recorte em relação à imagem raster original, assegurando a vinculação dos resultados do modelo às coordenadas reais do terreno.

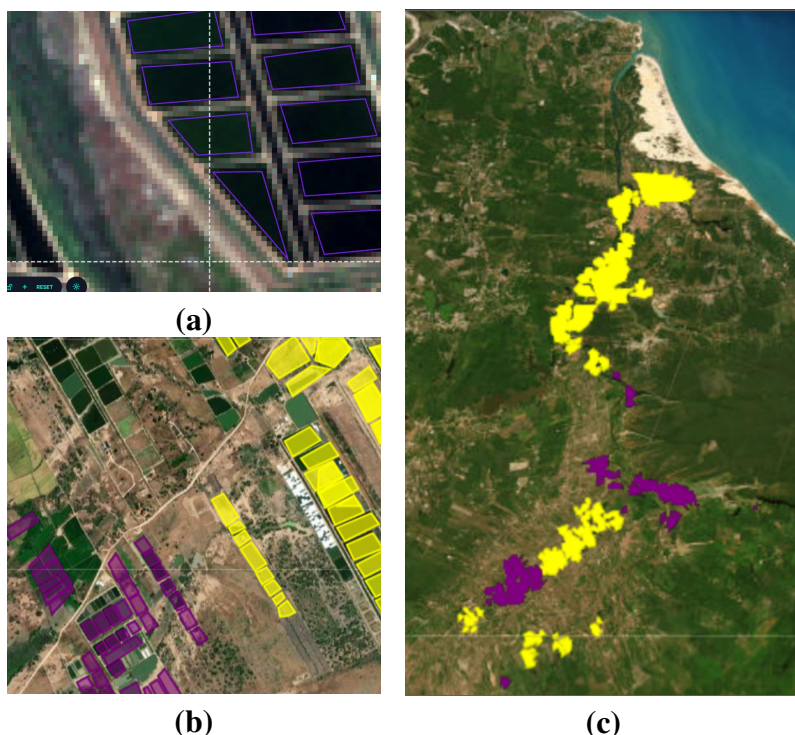


Figura 2. Exemplo de (a) segmentação de viveiros realizada por meio de anotação manual no *Roboflow*[®], (b) visualização detalhada dos conjuntos de treinamento e validação no *Google Earth*[®], e (c) distribuição dos conjuntos de treino (amarelo) e validação (roxo) nas regiões do médio e baixo Jaguaribe. Fontes: *Roboflow*[®], *Google Earth*[®].

Das 250 amostras aleatoriamente recortadas do vale do Jaguaribe (médio e baixo), Figura 2 (b) e (c), algumas foram descartadas devido à sobreposição entre treino e validação. As restantes, contendo um total de 5.491 viveiros anotados, foram manualmente divididas em 172 imagens para treino (92 com viveiros, 80 sem) e 46 para validação (23 com viveiros, 23 sem), mantendo uma proporção aproximada de 80.0% e 20.0%, respectivamente. As imagens foram armazenadas em formato *Tagged Image File Format* (TIFF) com quatro canais (R, G, B, NIR), sendo este último fundamental para distinguir vegetação e corpos d'água. O conjunto de dados poderá ser disponibilizado, mediante solicitação aos autores deste artigo, em conformidade com as regulamentações da plataforma *Copernicus*[®].

2.2. Adaptação do Modelo YOLOv11

O algoritmo original do YOLOv11 foi projetado para processar imagens com até três canais de entrada (RGB). Para permitir a incorporação de dados espectrais relevantes

para a segmentação de viveiros (NIR), foi necessária a modificação da camada de entrada do modelo para trabalhar com quatro canais.

Como modelo de partida, utilizamos a variante YOLOv11x-SEG [Jocher and Qiu 2024], contendo 62.051.411 parâmetros, previamente treinada com o *dataset Common Objects in Context - Segmentation* (COCO-SEG) [Lin et al. 2014]. Para minimizar a perda de conhecimento durante a adaptação, a primeira camada convolucional foi modificada para aceitar um quarto canal de entrada. Os pesos do novo canal foram inicializados como a média dos pesos dos três canais originais, garantindo uma adaptação eficiente e reduzindo a discrepância entre os filtros aprendidos anteriormente. Todas as demais camadas do modelo permaneceram inalteradas, preservando o conhecimento adquirido no treinamento prévio.

2.3. Pré-processamento

O *pipeline* de treinamento inicia-se com a normalização das imagens, convertendo os valores de cada canal do intervalo original $[0, 32, 767]$ (15 bits, inteiro) para $[0, 1]$ (32 bits, ponto flutuante). Além disso, foi aplicada uma estratégia de *data augmentation* para aumentar a diversidade do conjunto de dados e tornar o modelo mais robusto a diferentes orientações e condições visuais dos viveiros nas imagens. As transformações realizadas incluíram inversões horizontais e verticais, rotações fixas de 90° no sentido horário, anti-horário e 180° , rotações aleatórias com ângulos variando entre -15° e $+15^\circ$, desfoque com filtro gaussiano de até 2.5 *pixels* de intensidade, e inserção de ruído, com até 1.68% dos *pixels* alterados.

2.4. Treinamento do Modelo YOLOv11 Adaptado

O modelo adaptado foi treinado na plataforma *Google Colab*®, utilizando uma *Graphics Processing Unit* (GPU) *NVIDIA*®T4, com o conjunto de dados segmentado e processado conforme descrito anteriormente. A configuração inicial incluiu um tamanho de *batch* de 8, imagens redimensionadas para 640×640 pixels (a partir de recortes de 200×200 pixels), otimizador *Adaptive Moment Estimation* (ADAM) e taxa de aprendizado inicial de 0.01. O critério de parada (*early stopping*) interrompeu o treinamento após 30 épocas sem melhoria na métrica precisão, totalizando 274 épocas.

Durante o treinamento, foram monitoradas as seguintes métricas para avaliação de desempenho: precisão (*Precision*), revocação (*Recall*), *mAP*50, *mAP*50-95 e curvas de desempenho variando o *threshold* de confiança. Os resultados dessas métricas foram analisados a fim de validar a eficácia do modelo na detecção de viveiros de camarão, permitindo ajustes futuros na abordagem adotada.

3. Resultados e Discussão

A avaliação do modelo utilizou métricas padrões de detecção de objetos, incluindo *mAP*50 e *mAP*50-95, baseadas na interseção sobre união (IoU) das predições. Além disso, foi realizada uma análise visual detalhada para identificar erros, como *falsos positivos* e *falsos negativos*, complementada pelo uso de dados georreferenciados para correlacionar os resultados às características geográficas dos viveiros. Durante o treinamento, observou-se uma tendência ascendente das curvas de aprendizado, mesmo com flutuações pontuais, indicando que o modelo conseguiu aprender de maneira eficaz. A Figura 3 ilustra as

métricas do conjunto de validação: *Precisão*, *Recall*, *mAP50* e *mAP50-95*, as quais apresentam uma evolução consistente ao longo das 274 épocas.

Os resultados finais indicam uma **precisão de 87.64%** e um ***mAP50* de 88.25%**, sugerindo um alto nível de confiabilidade na identificação correta das classes. A métrica *mAP50-95* atingiu 52.87%, evidenciando uma capacidade considerável de generalização. As quedas abruptas de desempenho refletem a divisão do treinamento em múltiplas etapas devido às limitações da plataforma *Google Colab*®, resultando em reinicializações do otimizador e instabilidades momentâneas (Figura 3, épocas 105, 150 e 210). No entanto, o modelo demonstrou resiliência, recuperando seu desempenho após cada oscilação, o que sugere que a configuração dos hiperparâmetros e a estrutura do treinamento foram bem definidas, garantindo a manutenção da tendência positiva das métricas.

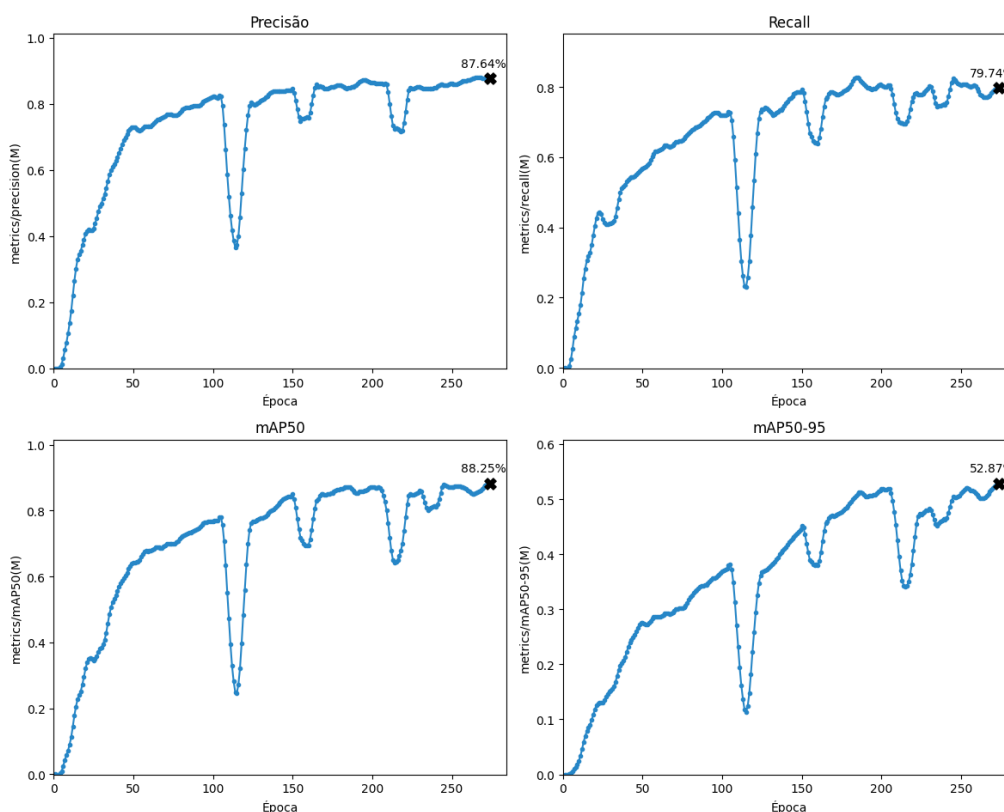


Figura 3. Evolução das métricas *Precisão*, *Recall*, *mAP50* e *mAP50-95* ao longo das 274 épocas para o conjunto de validação.

A curva *Precision-Confidence* e *Precision-Recall* (Figura 4) ilustram uma relação positiva entre a confiança das predições e a precisão obtida. A proximidade da curva *Precision-Recall* com o canto superior direito do gráfico sugere um desempenho de predição razoável, atingido um equilíbrio entre precisão e *recall* em 88.8% de confiança. Para um limiar de confiança de 86.1%, o modelo atinge máxima precisão, evidenciando alta confiabilidade em detecções com maior grau de certeza. No entanto, verifica-se uma queda abrupta na precisão em torno de 80.0% de confiança, causada por um aumento significativo de falsos positivos nesse intervalo. A análise do conjunto de dados revela que, nesse nível de confiança, o modelo atribui alta certeza a predições incorretas devido à presença de texturas semelhantes às dos viveiros de camarão em áreas não relacionadas.

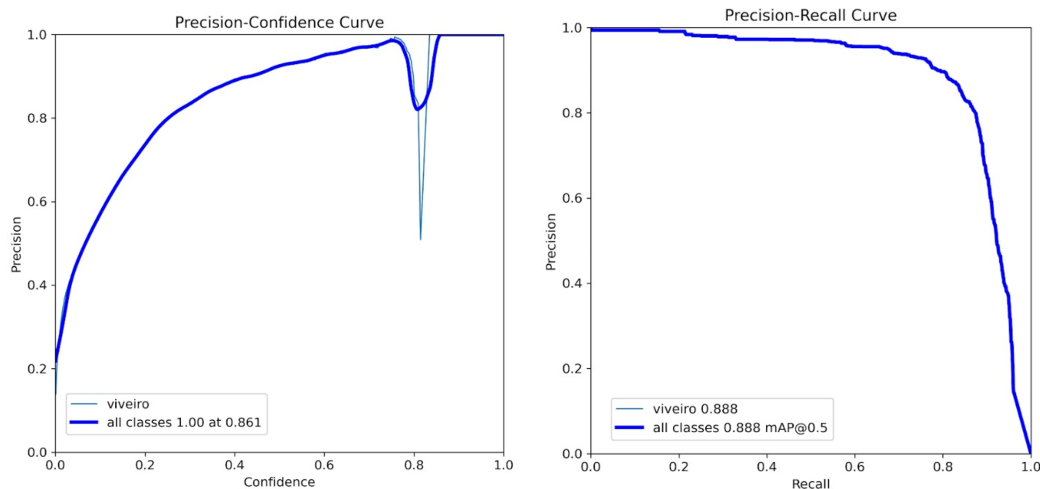


Figura 4. Pós-treinamento do modelo, apresentando as relações entre *Precision-Confidence* e *Precision-Recall*.

Ao aplicar o modelo à região de uma fazenda de produção de camarão presente no conjunto de validação, Figura 5 (a), observa-se que a maioria dos 58 viveiros presentes na imagem foi corretamente detectada. No entanto, o modelo não identificou 3 viveiros e gerou 1 detecção incorreta, apontando a existência de um viveiro inexistente. Ao aplicar o modelo à região da bacia Apodi-Mossoró, Figura 5 (b), observa-se muitas áreas com falsos positivos. Esses resultados indicam que, embora a abordagem proposta tenha alcançado um desempenho satisfatório, ainda há espaço para aprimoramentos, especialmente na minimização de *falsos positivos*.

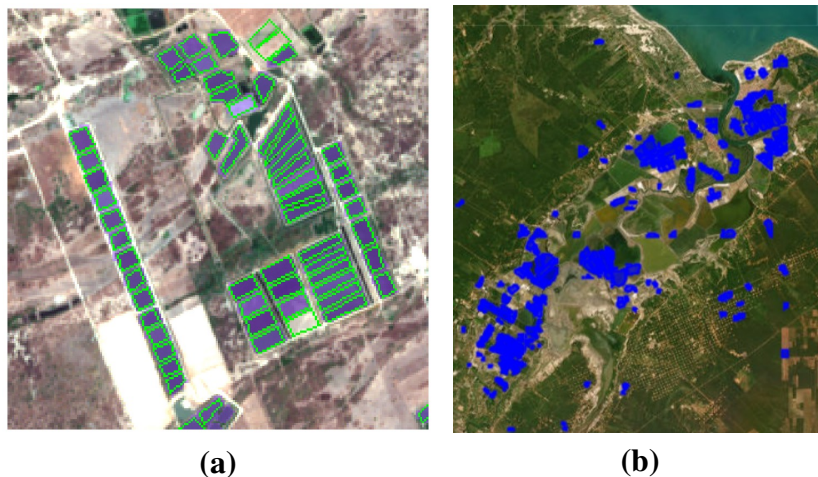


Figura 5. (a) Exemplo de segmentação de viveiros, com contornos em verde representando viveiros georreferenciados e, em roxo hachurado, viveiros detectados pelo modelo treinado. (b) Regiões prováveis de existência de viveiros (destacadas em azul) na bacia Apodi-Mossoró.

4. Conclusão

Neste artigo propomos aplicar o modelo YOLOv11 para detectar e segmentar imagens de viveiros de camarão por meio de sensoriamento remoto. Os resultados indicam alta precisão (87,64%) e um bom desempenho na identificação de áreas aquícolas. Além disso, o modelo identificou regiões que, embora não tenham sido confirmadas como viveiros por análise visual, apresentam características semelhantes às típicas de áreas de cultivo de camarão na bacia Apodi-Mossoró. Essa constatação abre caminho para novas investigações, com o objetivo de prospectar potenciais áreas de produção, reforçando a importância do modelo para o planejamento estratégico e a formulação de políticas públicas voltadas ao desenvolvimento sustentável da aquicultura. No futuro, essa abordagem pode ser aplicada a outros setores, como o monitoramento de recursos hídricos, a gestão de áreas agrícolas e a detecção de mudanças ambientais, ampliando seu impacto na preservação e uso eficiente dos recursos naturais.

Apesar dos resultados promissores, desafios permanecem. A ocorrência de falsos positivos e limitações na diferenciação entre viveiros e outras estruturas aquáticas indicam a necessidade de refinamentos, como a integração de informações geoespaciais, o uso de múltiplas séries temporais e a aplicação de índices espectrais. Além disso, a expansão do conjunto de dados para diferentes regiões permitirá testar a capacidade de generalização do modelo, tornando-o mais robusto para aplicações em larga escala. O aprimoramento contínuo dessas técnicas pode transformar a forma como a produção de camarão é monitorada, garantindo maior eficiência, sustentabilidade e suporte à tomada de decisão no setor.

Agradecimentos

Danielo G. Gomes agradece ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de produtividade (processo 311845/2022-3).

Referências

- Dong, D., Shi, Q., Hao, P., Huang, H., Yang, J., Guo, B., and Gao, Q. (2024). Intelligent detection of marine offshore aquaculture with high-resolution optical remote sensing images. *Journal of Marine Science and Engineering*, 12(1012).
- Gomes, J. V. E. (2022). Detecção de objetos com a arquitetura yolo. Master's thesis, Universidade Federal de Ouro Preto, João Monlevade, MG.
- Greenstreet, W., Bucini, G., Bowen, M., and Radeloff, V. C. (2023). Detecting aquaculture with deep learning in a low-data setting. *Remote Sensing in Ecology and Conservation*.
- IBGE (2023). Valor da produção da pecuária e aquicultura chega a R\$ 122,4 bilhões em 2023. Acesso em 29 de dezembro de 2024.
- Igarashi, M. A. (2022). Sinopse sobre os aspectos do potencial da produção e prospectos para o desenvolvimento do cultivo de camarão marinho no Brasil, especialmente no nordeste brasileiro. *Revista Semiárido De Visu*, 10(3).
- Jocher, G. and Qiu, J. (2024). Ultralytics yolo11.

- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *Computer vision–ECCV 2014: 13th European conference, zurich, Switzerland, September 6-12, 2014, proceedings, part v 13*, pages 740–755. Springer.
- Quaade, S., Vallebuena, A., Alcabes, O. D. N., Rodolfa, K. T., and Ho, D. E. (2024). Remote sensing and computer vision for marine aquaculture. *Science Advances*, 10(eadn4944).
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection.
- Werneck, P. R., Macedo, H. R., Morsoleto, F. M. d. S., Lira, K. C. d. S., and Feiden, A. (2023). Use of georeferencing as a sustainable planning tool for aquaculture production in small municipalities: a case study for ouro verde do oeste/pr. *Research, Society and Development*, 12(3):e17012338683.