

Sistema Inteligente de Visão Computacional para Detecção e Contagem de Peixes Utilizando YOLO11 e ByteTrack

Eduardo Mizael Clemente¹, Geraldo Andrade de Oliveira¹

¹Instituto Federal Goiano
Brasil

eduardo.mizael@gmail.com

Abstract. *This paper presents a computer vision system for fish detection, tracking, and counting in aquaculture videos using YOLO11 and ByteTrack. The system combines object detection, temporal tracking, polygonal regions of interest, geometric filtering, approximate size classification, and virtual line crossing. A preliminary proof of concept was evaluated with videos captured in a controlled aquaculture scenario and an initial annotated dataset. The detector achieved precision of 0.86, recall of 0.82, mAP@50 of 0.81, and mAP@50-95 of 0.41. The results indicate technical feasibility, although validation with larger datasets and real production environments is still required.*

Resumo. *Este artigo apresenta um sistema de visão computacional para detecção, rastreamento e contagem de peixes em vídeos de piscicultura utilizando YOLO11 e ByteTrack. O sistema combina detecção de objetos, rastreamento temporal, regiões poligonais de interesse, filtros geométricos, classificação aproximada de porte e contagem por cruzamento de linha virtual. Uma prova de conceito preliminar foi avaliada com vídeos capturados em cenário aquícola controlado e um conjunto inicial de imagens anotadas. O detector alcançou precisão de 0,86, recall de 0,82, mAP@50 de 0,81 e mAP@50-95 de 0,41. Os resultados indicam viabilidade técnica, embora ainda seja necessária validação com bases maiores e ambientes reais de produção.*

1. Introdução

A piscicultura demanda soluções capazes de apoiar o monitoramento contínuo dos organismos cultivados, incluindo estimativas populacionais, manejo alimentar e acompanhamento do desenvolvimento dos peixes. A contagem manual de indivíduos ainda é comum, mas pode ser demorada, dependente de operadores, suscetível a erros e potencialmente estressante aos animais. Revisões sobre contagem automática em aquicultura indicam que métodos baseados em sensores, visão computacional e aprendizado profundo podem reduzir trabalho manual e ampliar a frequência de monitoramento, desde que acompanhados por protocolos de coleta e validação bem definidos [Li et al. 2021].

Modelos da família YOLO (You Only Look Once) são amplamente utilizados em detecção de objetos por combinarem precisão e velocidade de inferência [Redmon et al. 2016]. No contexto de peixes, estudos recentes relatam desafios adicionais, como reflexos, turbidez, baixa iluminação, oclusões, densidade variável e movimentação rápida [Muksit et al. 2022, Kandimalla et al. 2022]. Esses fatores podem gerar falsos positivos, perdas temporárias de detecção e contagens duplicadas.

Neste contexto, este trabalho apresenta um sistema inteligente de visão computacional para detecção, rastreamento e contagem automática de peixes utilizando YOLO11 e ByteTrack. A proposta não se limita ao treinamento de um detector: ela integra aquisição de vídeo, região de interesse, detecção, rastreamento temporal, filtragem geométrica, classificação aproximada de porte e contagem por linha virtual. O ByteTrack foi adotado porque associa detecções de diferentes níveis de confiança, reduzindo fragmentações de trajetória em vídeos com detecções instáveis [Zhang et al. 2022].

O restante do artigo está organizado da seguinte forma. A Seção 2 descreve a arquitetura, o conjunto de dados, o treinamento e os parâmetros de reprodutibilidade. A Seção 3 apresenta os resultados preliminares e discute as métricas utilizadas. A Seção 4 sintetiza contribuições, limitações e trabalhos futuros.

2. Materiais e Métodos

Esta seção descreve o desenvolvimento da prova de conceito. Inicialmente, apresenta-se o fluxo de processamento do sistema. Em seguida, são descritos o conjunto de dados, o processo de anotação, o treinamento do modelo e os parâmetros necessários para reprodução do experimento.

2.1. Arquitetura do Sistema

O sistema foi desenvolvido em Python e estruturado em módulos independentes para aquisição de vídeo, definição de região de interesse (ROI), detecção de objetos, rastreamento temporal, filtragem, classificação aproximada de porte e contagem automática. Essa separação facilita ajustes em componentes específicos sem alterar o fluxo geral da aplicação.

Captura do vídeo → ROI → inferência com YOLO11 → rastreamento com ByteTrack → filtros geométricos e temporais → classificação aproximada de porte → cruzamento de linha virtual → exibição desktop ou web.

Figura 1. Pipeline geral do sistema proposto.

A detecção utiliza YOLO11 por meio da biblioteca Ultralytics [Ultralytics 2025], retornando caixas delimitadoras e níveis de confiança para os peixes identificados. Em seguida, o ByteTrack associa as detecções ao longo dos quadros e mantém identificadores persistentes para reduzir perdas temporárias e evitar múltiplas contagens do mesmo indivíduo [Zhang et al. 2022].

Para reduzir interferências, o sistema permite definir uma ROI poligonal. Essa região pode atuar como máscara espacial ou como recorte, enviando ao detector apenas a área útil da cena. Após a detecção e o rastreamento, filtros geométricos e temporais removem detecções incompatíveis com o padrão visual esperado dos peixes, considerando área mínima, proporção largura-altura e persistência mínima do objeto rastreado. A contagem ocorre quando o centro do objeto rastreado cruza uma linha virtual configurada na cena. A aplicação pode ser executada em ambiente desktop ou por interface web, com parâmetros definidos em arquivos externos.

2.2. Dataset, Anotação e Treinamento

O conjunto de dados utilizado nesta etapa foi composto por vídeos preliminares de peixes em ambiente aquícola controlado, gravados em resolução 720 x 1080 pixels. A decisão de iniciar em cenário controlado é compatível com trabalhos que constroem bases progressivas antes de avaliar cenários mais densos ou visualmente complexos [Muksit et al. 2022].

A partir dos vídeos, foram extraídos quadros individuais e anotados manualmente com a ferramenta Label Studio. Cada peixe visível foi delimitado por caixa retangular em formato compatível com YOLO. O conjunto resultante foi dividido em aproximadamente 80% para treinamento e 20% para validação. O treinamento foi conduzido com YOLO11 por 100 épocas, monitorando precisão, recall, mAP@50 e mAP@50-95.

2.3. Reprodutibilidade

Os principais parâmetros operacionais do sistema são: limiar mínimo de confiança do detector, limiar de confirmação temporal para IDs rastreados, coordenadas da ROI, área mínima da caixa delimitadora, proporção largura-altura permitida, posição da linha virtual de contagem e sentido de cruzamento considerado válido. O experimento pode ser reproduzido pelas seguintes etapas: gravar vídeos em ambiente controlado; extrair frames; anotar os peixes; dividir treino e validação; treinar o YOLO11; executar a aplicação sobre vídeos de validação; e comparar a contagem automática com observação manual. Quando os vídeos brutos não puderem ser publicados, recomenda-se disponibilizar ao menos frames anotados autorizados, arquivos de configuração e instruções de execução.

3. Resultados e Discussão

A avaliação preliminar demonstrou a viabilidade do uso conjunto de YOLO11 e ByteTrack para detecção, rastreamento e contagem automática de peixes. A Tabela 1 apresenta as métricas de validação do detector.

Tabela 1. Resultados preliminares do detector YOLO11.

Métrica	Valor
Precisão	0,86
Recall	0,82
mAP@50	0,81
mAP@50-95	0,41

A precisão indica a proporção de detecções positivas que correspondem a peixes, enquanto o recall indica a proporção de peixes presentes nos quadros que foram detectados. A mAP (*mean Average Precision*) resume a qualidade da detecção considerando a relação entre precisão e recall. A mAP@50 considera correta uma detecção cuja interseção sobre união (IoU) com a anotação de referência seja de pelo menos 0,50; a mAP@50-95 calcula a média em limiares de IoU entre 0,50 e 0,95, sendo mais rigorosa quanto à localização das caixas. Essas métricas são amplamente usadas em detecção de objetos e em estudos aplicados à detecção e contagem de peixes [Lin et al. 2014, Kandimalla et al. 2022, Muksit et al. 2022].

Os valores obtidos indicam boa capacidade inicial de identificação dos peixes nos quadros avaliados. Contudo, por se tratar de uma base preliminar e controlada, os resultados devem ser interpretados como evidência de viabilidade, e não como validação

definitiva para ambientes comerciais. A diferença entre $mAP@50$ e $mAP@50-95$ sugere que o modelo localiza os peixes de forma aceitável, mas ainda há margem para melhorar a precisão espacial das caixas delimitadoras.

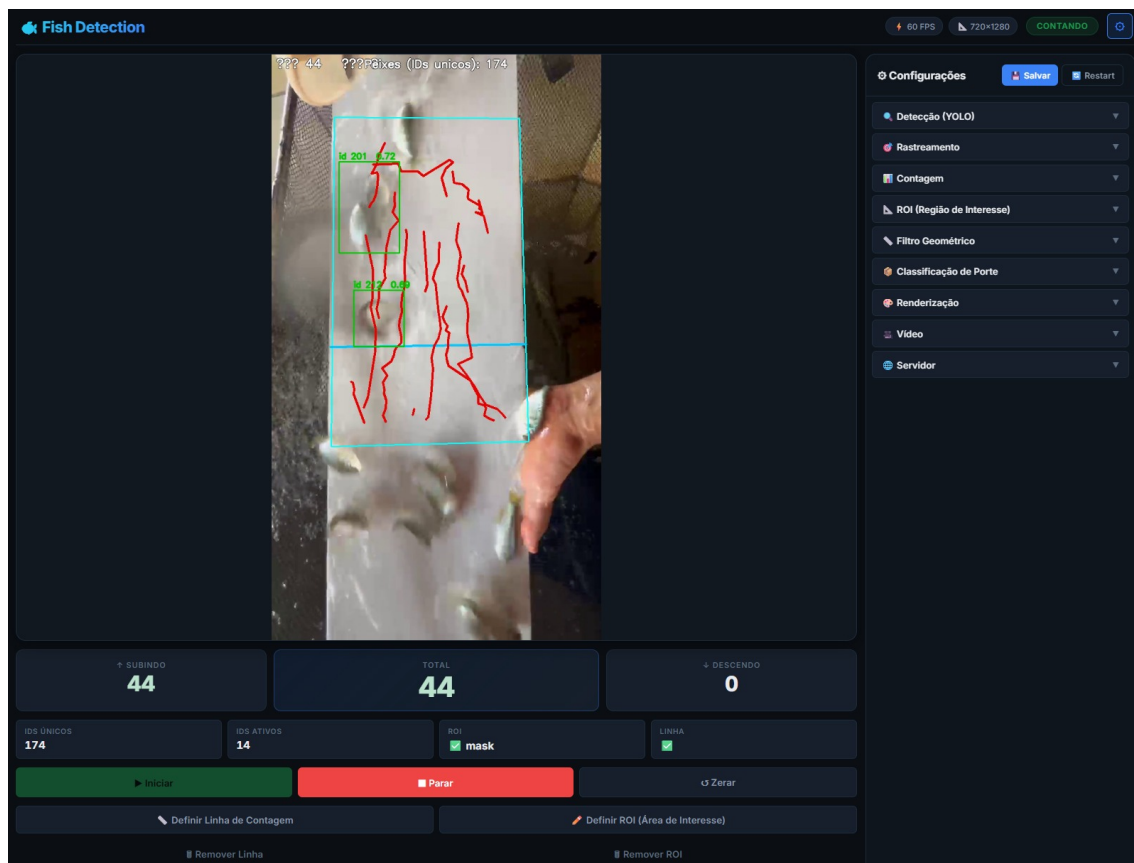


Figura 2. Exemplo de execução do sistema com detecção, rastreamento, região de interesse, linha virtual de contagem e contadores acumulados.

Na análise visual dos vídeos, a detecção quadro a quadro apresentou falsos positivos ocasionais em regiões com reflexos, variações de iluminação e pequenas movimentações do ambiente. A integração do ByteTrack contribuiu para preservar a continuidade das trajetórias, reduzindo perdas temporárias e múltiplas contagens do mesmo indivíduo. Os filtros geométricos também contribuíram para estabilizar o sistema ao descartar caixas incompatíveis com o padrão esperado dos peixes.

A utilização de regiões poligonais de interesse restringiu a análise às áreas úteis dos vídeos. O modo de recorte da ROI tende a favorecer aplicações em tempo real, pois reduz a quantidade de informação visual enviada ao detector. De modo geral, a combinação entre detecção por aprendizado profundo, rastreamento temporal e pós-processamento geométrico mostrou-se adequada para uma prova de conceito de contagem de peixes. Em cenários com alta densidade e forte sobreposição visual, métodos baseados em mapas de densidade podem ser investigados como complemento às abordagens por caixas delimitadoras [Li et al. 2024].

4. Conclusão

Este trabalho apresentou um sistema de visão computacional para detecção, rastreamento e contagem automática de peixes utilizando YOLO11 e ByteTrack. A principal contribuição está na integração dos componentes em uma arquitetura modular, configurável e extensível, com suporte a execução desktop e interface web.

Os resultados preliminares indicam viabilidade técnica da abordagem. O modelo treinado em uma base inicial apresentou métricas promissoras, e os mecanismos de rastreamento e pós-processamento contribuíram para reduzir falsos positivos e aumentar a estabilidade das contagens. Como limitações, destacam-se o tamanho reduzido da base, a coleta em condições controladas e a classificação de porte ainda baseada em estimativa relativa por área de caixa delimitadora.

Como trabalhos futuros, pretende-se ampliar o conjunto de dados com vídeos em ambientes reais, validar quantitativamente a contagem automática, avaliar o desempenho em hardware embarcado e disponibilizar os artefatos autorizados para reprodução independente.

Referências

- Kandimalla, V., Richard, M., Smith, F., Quirion, J., Torgo, L., and Whidden, C. (2022). Automated detection, classification and counting of fish in fish passages with deep learning. *Frontiers in Marine Science*, 8:823173.
- Li, D., Miao, Z., Peng, F., Wang, L., Hao, Y., Wang, Z., Chen, T., Li, H., and Zheng, Y. (2021). Automatic counting methods in aquaculture: A review. *Journal of the World Aquaculture Society*, 52(2):269–283.
- Li, X., Zhuang, Y., You, B., Wang, Z., Zhao, J., Gao, Y., and Xiao, D. (2024). LDNet: High accuracy fish counting framework using limited training samples with density map generation network. *Journal of King Saud University - Computer and Information Sciences*, 36:102143.
- 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 *European Conference on Computer Vision*, pages 740–755.
- Muksit, A. A., Hasan, F., Emon, M. F. H. B., Haque, M. R., Anwary, A. R., and Shatabda, S. (2022). YOLO-Fish: A robust fish detection model to detect fish in realistic underwater environment. *Ecological Informatics*, 72:101847.
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 779–788.
- Ultralytics (2025). Ultralytics yolo documentation.
- Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W., and Wang, X. (2022). ByteTrack: Multi-object tracking by associating every detection box. In *European Conference on Computer Vision*, pages 20–36.