

DETECÇÃO DE PEIXES UTILIZANDO REDES NEURASIS CONVOLUCIONAIS

Iohan Torres Cabral¹, Yuri Silva Martins²

¹Instituto Federal Goiano (IFGoiano)
Rio Verde – GO – Brazil

iohan.torres@estudante.ifgoiano.edu.br, yuri.martins@estudante.ifgoiano.edu.br

Abstract. *Fish farming is an activity that needs proper management. At the time of sale, the fish is packaged and counted to meet customer demand, a time-consuming and manual process. To speed up this process, this project aims to use counting through image processing, a technique that reduces time consumption and reduces the fish's exposure to stressful situations. The presented algorithm guarantees an accurate and quick estimate of the amount of packed fish. For fish recognition, the Darknet neural network was used together with the YOLO detection system. Structure that is about three times faster than CNN's with other architectures.*

CNN was trained using the dataset of the fish culture laboratory of IFGoiano (Rio Verde), and it can be verified that despite the human mean of counting approaching the real value, the standard deviation presented is very large. Thus, with the accuracy and detection performance of the proposed method, it can be said that this method has a great potential to facilitate the identification and counting of fish.

At the end of this research, it is expected that the proposed algorithm will count different sizes of fingerlings with high precision.

Resumo. *O cultivo de peixes é uma atividade que necessita de manejo adequado. No momento de sua venda, o peixe é embalado e contado para atender a demanda do cliente, processo demorado e feito de forma manual. Para agilizar esse processo, esse projeto tem o intuito de utilizar a contagem pelo processamento de imagens, técnica que reduz o consumo de tempo e diminui a exposição do peixe a situações estressantes. O algoritmo apresentado garante uma estimativa precisa e rápida da quantidade de peixes embalados. Para reconhecimento dos peixes foi usada a rede neural Darknet juntamente ao sistema de detecção YOLO. Estrutura que é cerca de três vezes mais rápida que CNN's com outras arquiteturas.*

A CNN foi treinada utilizando o dataset do laboratório de piscicultura do IFGoiano (Rio Verde), podendo ser verificado que apesar da média humana de contagem se aproximar do valor real, o desvio padrão apresentado é muito grande. Sendo assim, com a precisão e o desempenho de detecção do método proposto, pode-se afirmar que esse método tem um potencial muito grande em facilitar a identificação e contagem dos peixes.

Ao final desta pesquisa, espera-se que o algoritmo proposto conte diferentes tamanhos de alevinos com alta precisão.

1. Introdução

É possível que a Aquicultura seja uma atividade rentável e sustentável, uma alternativa de fonte de alimentos e de diversificação produtiva. Como ramificação desta atividade a piscicultura pode ser definida como atividade que usa recursos hídricos para a criação, engorda e comércio de peixes [SALES 2009]. Cultivar peixes é uma atividade que envolve manejo adequado visando aumentar a produção do pescado.

Em fazendas e lojas de peixes, os alevinos são geralmente embalados em sacos plásticos para serem vendidos aos clientes. A embalagem envolve a contagem dos peixes para obter a quantidade certa exigida pelos clientes. No entanto, o processo de contagem é demorado e pode estar sujeito a erro humano. Neste artigo, exploramos como esta contagem de peixes alimentadores pode ser automatizada usando processamento de imagem.

Para a contagem de peixes alevinos, o problema é simplificado porque envolve apenas uma espécie de peixe que se parece em forma e tamanho entre si, de forma que a detecção das espécies deles não é necessária. [Toh et al. 2009]. Como o número de peixes no saco plástico é fixo, a intenção do projeto é encontrar um método simples de executar que possa fazer a operação de contagem rapidamente e com boa precisão.

Diante do exposto, a contagem pela técnica de processamento de imagens reduziria o consumo de tempo, minimizaria a exposição dos peixes a situações insalubres e garantiria a estimativa precisa dos alevinos a serem comercializados. Uma estimativa precisa e rápida irá melhorar a entrega de alevinos às fazendas, alimentação adequada e plano financeiro adequado para o crescimento deles [Aliyu et al. 2017].

2. Revisão Bibliográfica

Este trabalho está relacionado a inúmeros trabalhos que foram relatados nas áreas de aprendizado de máquina e visão computacional, especificamente em aprendizado profundo. Esta seção faz uma breve revisão dos trabalhos relacionados e apresenta alguns conceitos fundamentais necessários para a compreensão deste trabalho.

Em 2018 um grupo de pesquisadores demonstrou uma forma para detecção e rastreamento de peixes online, permitindo um sistema rápido e experimentação rápida. Foi desenvolvida uma nova detecção subaquática de peixes e estratégias de rastreamento combinando o algoritmo de detecção YOLOv3 mais recente na época (YOLO), algoritmo de detecção e filtro de correlação paralelo [Liu et al. 2018].

Um trabalho de reconhecimento e rastreamento de peixes muito bem sucedido também foi realizado na Universidade Nacional de Ciência e Tecnologia de Gyeongnam onde foi proposto um método para classificar objetos com precisão e contar seu número em imagens de vídeo sequenciais. Este teve uma alta probabilidade de classificação de 93,94% e 97,06% nos vídeos em tempo real [Park and Kang 2020].

Já em 2021, um método foi apresentado para melhorar a detecção de peixes e suas trajetórias onde as condições da água são desafiadoras. O algoritmo de detecção melhora a qualidade de imagens pouco nítidas e reconhece os peixes. No final, recursos como contagem de peixes e trajetórias são extraídos das coordenadas dos objetos detectados. O que trouxe um melhor rastreamento e detecção de peixes em tanques de peixes em fazendas de peixes [Wageeh et al. 2021].

No campo da medicina há uma grande necessidade de acurácia em resultados, e quando se

trata de contagem de células os olhos humanos podem ser muito falhos nos laboratórios. No estudo em questão, a estrutura YOLO foi treinada com um conjunto de dados BCCD de configuração modificada de imagens de esfregaço de sangue para identificar e contar automaticamente os glóbulos vermelhos, glóbulos brancos e plaquetas. Além disso, este estudo com arquiteturas de rede neural convolucional relatou precisão e tempo de execução excepcional. No geral, o sistema de detecção e contagem auxiliado por computador permite contar células sanguíneas a partir de imagens de esfregaço em menos de um segundo, o que é útil para diversas aplicações práticas [Alam and Islam 2019]. É interessante perceber que em todos os métodos citados a rede YOLO foi escolhida pelos pesquisadores para o treinamento dos seus sistemas, o que ressalta sua funcionalidade e importância no campo de detecção de objetos.

3. Redes Neurais Convolucionais

A Rede Neural Convolucional teve resultados inovadores na última década em uma variedade de campos relacionados ao reconhecimento de padrões; do processamento de imagem ao reconhecimento de voz. O aspecto mais benéfico das CNNs é a redução do número de parâmetros em Redes Neurais Artificiais (RNA). Essa conquista levou pesquisadores e desenvolvedores a abordar modelos maiores a fim de resolver tarefas complexas, o que não era possível com as RNAs clássicas.

A CNN tem várias camadas; incluindo camada convolucional, camada de não linearidade, camada de pool e camada totalmente conectada. As camadas convolucionais e totalmente conectadas têm parâmetros, mas as camadas de pool e não linearidade não têm parâmetros. A CNN tem excelente desempenho em problemas de aprendizado de máquina. Especialmente os aplicativos que lidam com dados de imagem, como o maior conjunto de dados de classificação de imagem (Image Net), visão computacional e em processamento de linguagem natural (PNL) [Albawi et al. 2017]

4. Metodologia

Para o reconhecimento dos peixes nas fotos o grupo procurou por diversas fontes para selecionar a mais adequada ao trabalho. Ao fim da pesquisa. Por fim, foi selecionada a rede Darknet juntamente com a YOLO, que são fáceis de manusear e uma das melhores e mais funcionais quando se trata do reconhecimento de seres em imagens.

Nosso objetivo é usar o algoritmo de detecção e classificação de objetos YOLO para detectar e contar os peixes diretamente das imagens disponibilizadas pelo laboratório. Precisamos treinar a estrutura YOLO com uma configuração modificada e imagens de treinamento de diversos peixes do dataset.

4.1. YOLO

Escolhemos a estrutura YOLO, que é cerca de três vezes mais rápida que CNN's com arquitetura VGG-16 [Redmon et al. 2016]. YOLO usa uma única rede neural para prever caixas delimitadoras e probabilidades de classe diretamente da imagem completa em uma ação de avaliação. Nós treinamos novamente a estrutura YOLO para identificar automaticamente e contar peixes alevinos presentes nas imagens.

A versão utilizada nesse projeto foi a YOLOv4 do AlexeyAB. Nela, são implementados novos recursos, como: WRC, CSP, CmBN, SAT, ativação Mish, aumento de dados Mo-saic, CmBN, regularização DropBlock e perda de CIoU, alguns deles combinados para

obter resultados de última geração: 43,5% AP (65,7% AP50) para o conjunto de dados MS COCO a uma velocidade em tempo real de em média 65 FPS em casos de vídeo [Bochkovskiy et al. 2020].

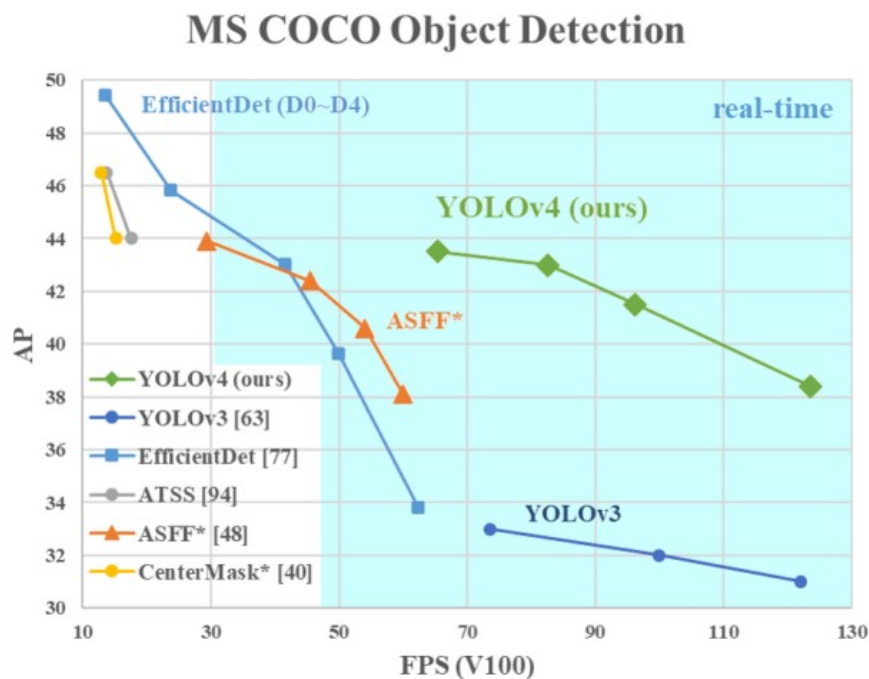


Figura 1. Gráfico de desempenho do YOLOv4 em comparação aos seus concorrentes.

Nosso método não requer nenhuma conversão de escala de cinza ou segmentação binária. Todo o processo é totalmente automatizado, rápido e preciso.

4.2. Darknet

Neste artigo, consideramos explicitamente a Darknet, que é uma Rede Neural Convolutiva (CNN) de código aberto usada para implementar o sistema de detecção de objetos YOLO. Darknet ficou famoso por suas atuações no Desafio PASCAL VOC 2007, pois era o único sistema em tempo real e tinha uma alta precisão em comparação com os outros sistemas [Everingham et al. 2009].

Darknet, como a maioria dos CNNs, realiza três operações para detecção de objetos, que são: (1) Convolução de camada. (2) Camadas de agrupamento máximas. (3) Classificação por meio de uma camada totalmente conectada [Redmon 2016]. Camadas convolucionais funcionam como extração de recursos no Darknet, deslizando um kernel convolutivo para a imagem de entrada. No Darknet, existem 24 camadas convolucionais e essas camadas produzem diferentes mapas de recursos a partir da entrada (ou seja, a saída de um kernel convolutivo na imagem). Os mapas de feições Darknet devem ser processados por camadas de pooling Max, o que reduz dimensionalmente o mapa de feições obtido anteriormente. O pool máximo divide o mapa de recursos em blocos e reduz cada bloco em um valor.

4.3. CUDA

A plataforma CUDA (Compute Unified Device Architecture) [Ryoo et al. 2008] é responsável por trazer o poder de processamento das GPUs (Graphics Processing Units) para as tarefas computacionais realizadas durante o processo de detecção de objetos, acelerando em torno de 500 vezes o tempo de detecção de objetos por frame, em comparação com tempo de detecção alcançado em CPUs (Central Processing Units), o que possibilita a detecção com o hardware adequado.

4.4. Dataset

Para criação do dataset foram usadas imagens capturadas por integrantes do laboratório de piscicultura do IFGoiano (Rio Verde). Estas imagens possuíam pastas contendo as quantidades de alevinos que ali estavam (10, 20, 30, 40 e 60 peixes) e videos deles em movimento. Isso proporcionou uma boa base de dados para o treinamento.

A fim de que a CNN pudesse identificar o que era um peixe na imagem a ser treinada, a equipe precisou demarcar manualmente a área de presença dos mesmos em toda a imagem. Esta ação foi realizada com o auxílio do programa "Labelimg", uma ferramenta gráfica de anotação de imagens escrita em Python e usa Qt para sua interface gráfica [tzutalin 2018]. A ferramenta foi adequada para a tarefa pois suportava a YOLO e facilitava o salvamento dos arquivos com as coordenadas dos animais.

Esse tipo de marcação é essencial para que os algoritmos de reconhecimento de imagens consigam entender o objeto que estão sendo treinados para encontrar na imagem.

Esta etapa pode determinar o sucesso do processo, pois depende da seleção correta das caixas dos objetos, já que é com ele que será treinado o reconhecimento.

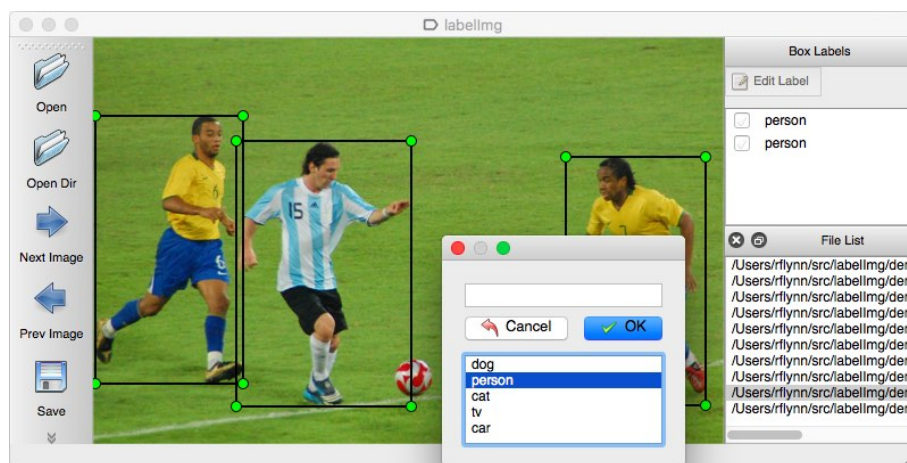


Figura 2. Exemplificação do funcionamento do Labelimg.

5. Processamento

Na etapa de processamento das imagens, a equipe conseguiu se limitar ao uso da biblioteca de visão computacional OpenCV, que foi invocada unicamente para mostrar as imagens dos resultados na tela. Todo resto foi tratado pela Darknet juntamente com a YOLO.

A vantagem do uso da Darknet sobre as demais foi o fato de que as imagens poderiam ser enviadas da forma que foram tiradas, sem precisar passar por qualquer pré-processamento antes do início do treinamento da rede.

6. Treinamento

O treinamento durou cerca de 7 horas e foi totalmente realizado pelo repositório Darknet/YOLOv4 do AlexeyAB no GitHub ¹.

As imagens marcadas com o LabelImg foram adicionadas a uma pasta "obj" do Google Drive, juntamente com uma segunda "test" para o teste da rede durante o treino.

6.1. Google Colab

A plataforma utilizada para realizar o treinamento da rede neural convolucional foi um ambiente virtual no Google Colab criado institucionalmente pelo estudante Iohan Torres Cabral², que funciona totalmente online e sem gastar recursos do computador.

Um problema detectado ao utilizar essa plataforma foi o tempo limite de execução do programa, o que acabou por encerrar o treinamento antes que tivesse uma acurácia considerável. Isso acabou por dificultar não só a obtenção de resultados, mas também a visualização das métricas de acompanhamento geradas

6.2. Métricas

Como descrito anteriormente, o encerramento do Google Colab dificultou a visualização em números do resultado. Porém foi possível perceber que ao início e ao fim do processo o Average Loss foi da casa dos 90.0 para 1.30 respectivamente, o que é considerado bom para a detecção de um único tipo objeto na imagem.

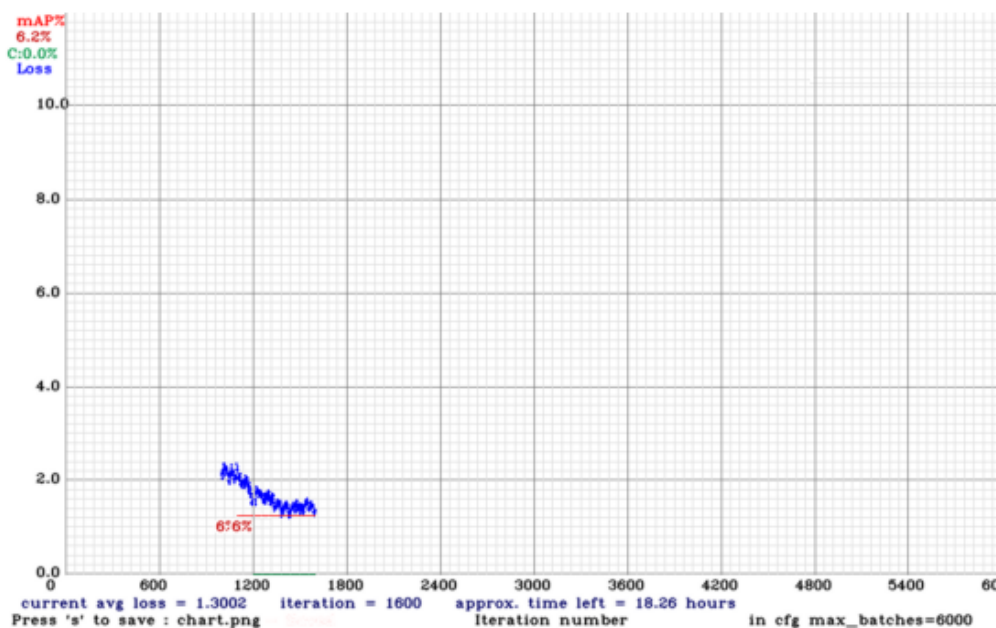


Figura 3. Parte do gráfico de resultados da rede neural.

Observando o gráfico é possível perceber uma certa consistência na curva a partir das 1200 interações, o que significa que ao fim do processo os resultados já estavam começando a se estabilizar, o que traz confiança ao teste realizado.

¹<https://github.com/AlexeyAB/darknet>

²<https://colab.research.google.com/drive/11xUW8ARnZhrPoiLoV2uajy-tLV2Yhyq?usp=sharing>

6.3. Avaliação dos Resultados

Para verificar a confiabilidade do algoritmo foi realizado um experimento de contagem para analisar se o valor obtido pelo contador se aproxima do valor real e o como o contador se sai quando comparado ao desempenho de seres humanos.

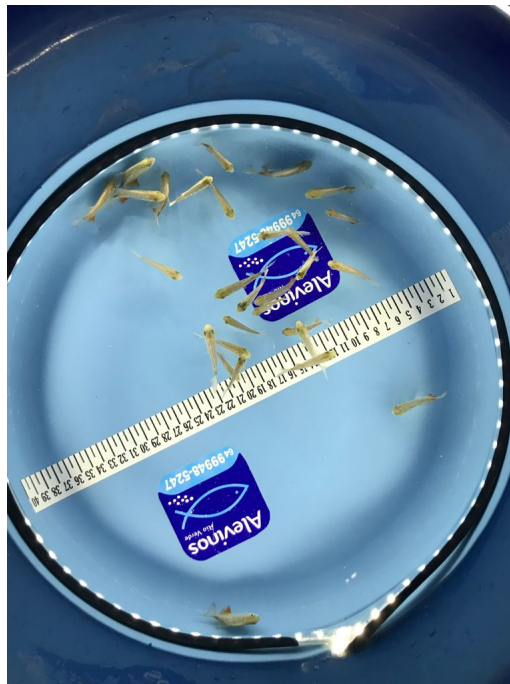


Figura 4. 29 peixes foram contabilizados manualmente nesta imagem.



Figura 5. 28 peixes foram contabilizados pelo programa nesta imagem.

Ao analisar os resultados obtidos pelo algoritmo percebe-se que a contagem apesar de precisa, ainda pode possuir alguns falsos negativos. Portanto a rede ainda requer mais tempo de treinamento e aperfeiçoamentos.

Além disso é interessante avaliar o quão acurado o algoritmo pode ser mesmo com variações e erros. Isto é, mesmo com as imprecisões, o quão próximo ao valor correto o algoritmo pode ser ao analisar figuras com quantidades de alevinos que seriam muito custosas para serem contabilizadas por seres humanos.

Na pesquisa foi notado que este método é muito útil para contar um pequeno número de peixes. Isso ocorre porque eles estão mais espaçados uns dos outros, então é mais provável que apareçam individualmente, reduzindo o número de vezes que temos que processar peixes que se cruzam. Quando os peixes se cruzam, a quantidade de sobreposição difere, portanto, os tamanhos das caixas de marcação podem ter uma grande variedade, mesmo para o mesmo número de peixes que se cruzam. Por exemplo, se um peixe se sobrepõe quase completamente a outro, é provável que a marcação seja apenas ligeiramente maior do que a de um único peixe. A gama de áreas que podem ser ocupadas por um peixe encontrado individualmente é maior do que para os peixes que se cruzam, então, se mais deles forem encontrados individualmente, a probabilidade de que mais peixes sejam categorizados no grupo correto é maior. Assim, à medida que o número de peixes aumenta, o número de peixes que se cruzam aumenta e consequentemente a precisão diminui.

7. Conclusões

Neste trabalho foi apresentado um sistema para a identificação e contagem de peixes de criatório em imagens digitais, geradas pela digitalização das amostras capturadas pelos integrantes de um laboratório do IFGoiano (Rio Verde). Para o treinamento do modelo inteligente utilizou-se uma rede neural convolucional profunda e um conjunto de dados com 70 imagens coloridas e bem iluminadas, contendo objetos (alevinos) rotulados manualmente pela equipe que realizou o trabalho.

Com relação as alternativas para aumentar a precisão do modelo, conclui-se que é necessário adicionar novas imagens no conjunto de imagens para treinamento, peixes em diferentes poses e posicionamento de cauda, bem como, executar mais épocas para a redução do valor de perda. Também é necessário revisar os procedimentos de campo para reduzir o envio de imagens com peixes muito juntos, visto que é possível que tanto o algoritmo quanto um humano os reconheça como um só, o que representa um ambiente altamente confuso para análise. Por fim, com a disponibilidade de uma GPU para a realização da inferência será possível processar imagens sem a necessidade da plataforma do Google Colab, que finaliza automaticamente a execução depois de passado certa quantidade de tempo.

Ainda que este algoritmo realize a contagem com alguns erros, conforme foi mencionado anteriormente, ele se aproxima com eficácia ao valor real verificado para os exemplos analisados.

No comparativo à qualidade humana de realizar esta mesma tarefa, fica claro que o algoritmo pode ser muito mais eficaz. Apesar da média da contagem humana se aproximar do valor real, o desvio padrão apresentado é muito grande, mesmo para uma pequena quantidade de peixes, isso significa que as chances de uma pessoa errar ainda continuam sendo maiores do que a máquina, mesmo que em seu estado de treinamento embrionário. Tal acontecimento não se limita apenas a contagem dos animais em questão, mas no geral quando se trata de contagem com imagens e vídeos. Isso é comprovado por um estudo

realizado em 2020 [de Andrade 2020] com contagem de veículos (Figura 6), onde a contagem humana, apesar de muitas vezes acurada, possuiu muito mais erros em relação a contagem realizada pela máquina, demonstrando uma vantagem em sua aplicação.

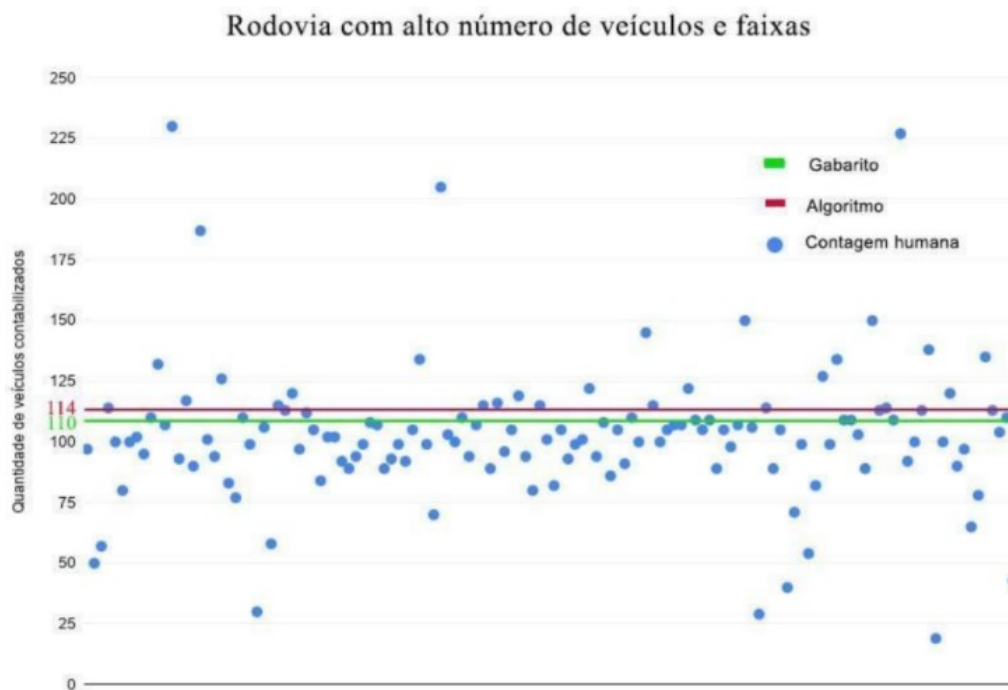


Figura 6. Contagem humana apresenta desvio padrão muito maior em comparação a máquina.

O método proposto também foi testado em um conjunto de dados diferentes das imagens inicialmente fornecidas pelo laboratório, onde mostrou um desempenho satisfatório. Com a precisão e o desempenho de detecção do método proposto, pode-se afirmar que o método tem potencial para facilitar o processo de identificação e contagem manual de peixes alevinos.

A precisão da técnica nas revisões precisa ser aprimorada, bem como para obter a precisão ideal. Para alcançar a contagem com alta precisão, um algoritmo de aprendizado de máquina para classificação e contagem de alevinos foi formulado. Ao final desta pesquisa, espera-se que o algoritmo proposto conte diferentes tamanhos de alevinos com alta precisão em comparação aos trabalhos existentes.

Este artigo foi baseado e utilizado dos mesmos conceitos analisados no artigo de detecção de peixes realizado por Liu, S., Li, X., Gao, M., Cai, Y., Nian, R., Li, P., Yan, T., and Lendasse, A. Embedded Online Fish Detection and Tracking System via YOLOv3 and Parallel Correlation Filter

Referências

- Alam, M. M. and Islam, M. T. (2019). Machine learning approach of automatic identification and counting of blood cells. *Healthcare technology letters*, 6(4):103–108.
- Albawi, S., Mohammed, T. A., and Al-Zawi, S. (2017). Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)*, pages 1–6. Ieee.
- Aliyu, I., Kolo, J. G., Aibinu, A. M., Agajo, J., Orire, A. M., Folorunso, T., and Adegboye, M. A. (2017). A proposed fish counting algorithm, using digital image processing techniques.
- Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection.
- de Andrade, M. M. (2020). *APLICAÇÃO DE VISÃO COMPUTACIONAL PARA RASTREAMENTO E CONTAGEM DE VEÍCULOS EM RODOVIAS*. PhD thesis, Universidade Federal do Rio de Janeiro.
- Everingham, M., Gool, L., Williams, C. K. I., Winn, J., and Zisserman, A. (2009). The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88:303–338.
- Liu, S., Li, X., Gao, M., Cai, Y., Nian, R., Li, P., Yan, T., and Lendasse, A. (2018). Embedded online fish detection and tracking system via yolov3 and parallel correlation filter. In *OCEANS 2018 MTS/IEEE Charleston*, pages 1–6. IEEE.
- Park, J.-H. and Kang, C. (2020). A study on enhancement of fish recognition using cumulative mean of yolo network in underwater video images. *Journal of Marine Science and Engineering*, 8(11):952.
- Redmon, J. (2013–2016). Darknet: Open source neural networks in c. <http://pjreddie.com/darknet/>.
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 779–788.
- Ryoo, S., Rodrigues, C. I., Baghsorkhi, S. S., Stone, S. S., Kirk, D. B., and Hwu, W.-m. W. (2008). Optimization principles and application performance evaluation of a multithreaded gpu using cuda. In *Proceedings of the 13th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, PPoPP '08*, page 73–82, New York, NY, USA. Association for Computing Machinery.
- SALES, E. B. d. (2009). Noções básicas de piscicultura. *Porto Velho: EMATER/RO*.
- Toh, Y., Ng, T., and Liew, B. (2009). Automated fish counting using image processing. In *2009 international conference on computational intelligence and software engineering*, pages 1–5. IEEE.
- tzutalin (2018). Labelimg.
- Wageeh, Y., Mohamed, H. E.-D., Fadl, A., Anas, O., ElMasry, N., Nabil, A., and Atia, A. (2021). Yolo fish detection with euclidean tracking in fish farms. *Journal of Ambient Intelligence and Humanized Computing*, 12(1):5–12.