

PLANO DE TRABALHO

Título do Projeto: Visão Computacional e Aprendizagem Automática para Aplicações em Agropecuária e Ciências Forenses.

Título do Plano de Trabalho: Desenvolvimento de Ferramenta para detecção e contagem de alevinos de pintado-real utilizando Faster R-CNN.

Orientador: Hemerson Pistori (pistori@ucdb.br)

Coorientador: Marcio Carneiro Brito Pache (marcio.pache@ifms.edu.br)

Orientado: Fabio Prestes Cesar Rezende **RA:** 169665

Curso: Engenharia de Computação **Semestre:** 5º semestre

Resumo

Um problema comum na aquicultura é a contagem de alevinos que exige a busca por soluções tecnológicas inovadoras, ambientalmente sustentáveis e usualmente amigáveis. Utilizando técnicas de Visão Computacional e Aprendizagem Automática, é possível providenciar dados estatísticos mais precisos e consistente ao produtor, auxiliando assim na homogeneização e barateamento de sua produção. O objetivo é testar a eficácia do método Faster R-CNN para detecção e contagem de alevinos de pintado-real, observando as diferentes performances ao mudar a quantidade de épocas usadas durante o treinamento, e comparando tais redes neurais treinadas a ferramenta atualmente em uso pelo INOVISÃO, o Contador de Alevinos criado por Vanir et al. (2015), e caso este método obtenha melhores resultados em relação a ferramenta atual, implementar esta rede neural em uma ferramenta para uso do INOVISÃO.

1. Antecedentes e Justificativa

Segundo o Anuário da Associação Brasileira de Piscicultura, a produção de peixes e alevinos teve crescimento de 4,5% sobre a produção de 2017 de 691.700 toneladas, totalizando 722.560 toneladas em 2018. Devido às adversidades enfrentadas tais como: dificuldade nos processos de regulamentação dos piscicultores; clima desfavorável para produção; problemas sanitários; modesto crescimento do PIB; preferência por produtos alimentícios mais baratos

devido a elevada taxa de desemprego, tal aumento não foi tão significativo quanto a anos anteriores. Entretanto, com uma expectativa otimista da profissionalização da área da piscicultura resultando em uma maior produtividade, será possível diminuir os impactos negativos criados por tais problemas (PEIXEBR, 2019).

Outro método para criar um impacto positivo no mercado seria o avanço tecnológico na área, visto que novas soluções tecnológicas não só automatizam o processo industrial, mas também melhoram a qualidade de vida do produto. Segundo pesquisas, o estresse causado pelo manuseio leva a diminuição da qualidade da carcaça final do peixe, com pontuações menores em quesito de textura e odor, assim como período de *rigor-mortis* mais curto após o abate (SIGHOLT et al., 1997), e antes do abate os peixes podem sofrer variações significativas em níveis de colesterol no sangue e em suas funções metabólicas, durando até 24 horas após o manuseio (WEDEMEYER 1972).

Um dos maiores problemas atualmente é a contagem e pesagem dos peixes, que necessita de grande quantidade de mão de obra humana e causa grandes prejuízos. (EMBRAPA, 2018). Devido às elevadas quantidades que devem ser processadas, utilizar-se de mão-de-obra humana está sujeito a medições errôneas durante as contagens, pesagens ou causar danos aos peixes, diminuindo o valor da carcaça final. Tais problemas se acumulam, implicando em prejuízos ao produtor e no aumento do preço final para o consumidor, levando-o a escolher produtos mais baratos.

Em julho de 2018, uma parceria foi formalizada entre várias instituições privadas e públicas para criar o BRS Aqua, um projeto de 4 anos que pretende estabelecer a infraestrutura e pesquisa científica necessária para atender as demandas do mercado de aquicultura (EMBRAPA, 2018). Com a automatização do processo de contagem, seria possível o produtor economizar em mão de obra, permitindo assim aumentar a escala de sua produção e vender o produto por um preço menor, atingindo assim mais consumidores de várias classes econômicas.

A Visão Computacional possui ótimas ferramentas e métodos capazes de solucionar tais problemas com interação mínima do usuário com a área de produção/medição. Vários testes já foram feitos utilizando tais métodos, como por exemplo experimentos com métodos de subtração de fundo para identificação de componentes conexas (*blobs*) que seriam peixes, ou usando um

conjunto de métodos de detecção de borda de Canny e análise de formas de Zarnike para classificação dos peixes via a forma das componentes conexas, ou utilizando pré-processamento das imagens dos peixes no ambiente utilizando Ruído Gaussiano (*Gaussian Blur*), e em seguida a subtração desta imagem pré-processada com a imagem original para facilitar o rastreamento dos peixes utilizando algoritmos de componentes conexas heurísticas (LING e PHOOI, 2019).

Assim, O Grupo de Pesquisa, Desenvolvimento e Inovação em Visão Computacional (INOVISÃO) tem explorado recursos científicos do ramo da visão computacional para criar soluções inovadoras para problemas locais em diversas frentes de pesquisa. O grupo é constituído por docentes, pesquisadores, estudantes de mestrado, doutorado, estudantes de graduação envolvidos em TCC (Trabalho de Conclusão de Curso) e iniciação científica e estudantes de ensino médio envolvidos em iniciação científica júnior, com projetos em quatro frentes de pesquisa: perícia forense/palinologia (PERICIA/PALINOVIC), piscicultura (FISHCV), pecuária de corte (PECVC) e agricultura de ponta (VANTAGRO). Este projeto em específico trabalha em junção ao FISHCV para desenvolver tecnologicamente a área da piscicultura aplicando métodos de visão computacional e inteligência artificial para automatizar a produção e torná-la mais homogênea e eficiente.

2. Objetivo

O objetivo deste trabalho é testar o uso do método Faster R-CNN em junção a técnicas de visão computacional para a detecção e contagem de alevinos de pintado real de modo autônomo.

2.1 Objetivos específicos

- Capturar vídeos de alevinos de pintado-real para formulação de um banco de dados (*dataset*).
- Implementar o algoritmo Faster R-CNN e analisar sua performance e acurácia utilizando análise estatística

3. Revisão de literatura

3.1 Piscicultura no Brasil

A piscicultura é uma área que tem recebido grande atenção mundial, e no Brasil tem tido um rápido crescimento, saltando de 172.000 toneladas de peixe de cativeiro em 2000 para 629.300 toneladas em 2011 (FAO, 2011), e segundo o Anuário da PEIXEBR (2019), tal produção vem aumentando, alcançando 722.560 toneladas em 2018, e conseguindo um total de exportações de 32.417 toneladas totalizando US\$ 136 milhões no mesmo ano, mesmo com uma suspensão sobre as exportações para a União Européia. Segundo este mesmo Anuário, entre os estados brasileiros, o Mato grosso do Sul também aumentou sua produção, conseguindo entrar no ranking dos 10 maiores produtores de peixe de cativeiro do Brasil, com 25.850 toneladas em 2018. (PEIXEBR, 2019).

Segundo a Embrapa (2012), o processo aquicultural se difere da pesca comum devido a homogeneidade de seus produtos, habilidade de rastreamento durante a totalidade do tempo de crescimento dos peixes, além de outros benefícios que permitem uma segurança alimentar e maior qualidade ao produto final com regularidade.

3.2 Iniciativa de inovação na piscicultura de Mato Grosso do Sul

Segundo a Embrapa (2012), o processo aquicultural se difere da pesca comum devido a homogeneidade de seus produtos, habilidade de rastreamento durante a totalidade do tempo de crescimento dos peixes, além de outros benefícios que permitem uma segurança alimentar e maior qualidade ao produto final com regularidade.

Todavia em Mato Grosso do Sul, o grupo INOVISAO estabeleceu parceria desde 2015 com o Projeto Pacu, localizado na cidade de Terenos, para desenvolverem um contador de alevinos com o objetivo de solucionar o problema de contagem manual, normalmente associado a estresse e danos aos alevinos, bem como, mão de obra dispendiosa que leva a erros e imprecisões.

Assim, Garcia et al. (2015) criaram um sistema de visão computacional que contém o hardware, responsável pela passagem dos alevinos e aquisição de vídeos, e da ferramenta responsável pelo processamento dos vídeos em tempo real de execução, realizando a contagem dos alevinos. Apesar dos avanços, melhorias no contador precisam ser implementadas a fim de

aferir o equipamento para se tornar um novo produto no mercado nacional e internacional acessível a pequenos e médios piscicultores.

3.3 Visão Computacional e Redes Neurais

Segundo G. Bradski (2016) e G. Weinstein (2018), Visão Computacional é a geração de novas representações ou decisões utilizando ferramentas computacionais a partir de uma imagem estática ou em movimento, utilizando os valores de seus pixels para inferir o conteúdo da imagem e criar identidades para certas sequências ou agrupamentos de pixels, podendo assim automatizar o processo desejado.

Esta área da computação é altamente usada para problemas onde é necessário transformar informações visuais em uma linguagem que o método computacional possa reconhecer e interpretá-la, por exemplo reconhecimento de padrões para detecção de objetos em imagens ou vídeos, onde se é obtido os valores dos pixels de interesse e se observa os valores de seus pixels vizinhos, criando assim agrupamentos semelhantes e permitindo que seja possível identificar objetos semelhantes a tal agrupamento (ROWLEY, 1999).

Esta junção de visão computacional e reconhecimento de padrões é muito utilizada em Redes Neurais, unidades computacionais paralelas e interconectadas parcial ou totalmente por nódulos chamados “neurônios artificiais”, onde cada *neurônio* realiza certas operações e transmite seu resultado para outras unidades conectadas a ela (*feed forward*). Através de um processo de treinamento, é possível para esta rede identificar e reconhecer padrões, mesmo que os dados em questão sejam contraditórios ou imprecisos, tornando-a eficaz em tarefas dispendiosas a especialistas humanos uma vez que há um conjunto de regras complexo (PERELMULTER, 1995).

3.4 Redes Neurais Convolucionais

Um dos métodos clássicos para classificação de vídeos antes do uso amplo de redes neurais envolve três estágios principais: extração de características visuais do vídeo, combinação

destas característica em descrições espaço-temporais do vídeo, e finalmente usar uma Máquina de Vetores de Suporte (SVM) para classificar as diferentes classes usando histogramas de palavras visuais (*Bag of Words*) resultante. Redes Neurais Convolucionais juntam estes três processos em uma única rede neural capaz de ser treinada desde a extração de características dos pixels até a classificação resultante, obtendo ótimos resultados devido ao avanço do *hardware* de GPUs (Placa Gráfica de Processamento), obtendo uma performance considerada estado-da-arte em banco de dados base como *ImageNET na época de sua criação*. (KARPATHY, 2014).

Após a apresentação dos resultados desta rede neural, chamada também de redes neurais convolucionais, *Convolutional Neural Networks, ConvNets* ou *CNNs*, Girshick et al. (2015) desenvolveram um método que possibilitasse a esta rede localizar objetos dentro de uma imagem com uma quantidade pequena de anotações feitas. Assim, utilizando um paradigma sobre “reconhecimento de regiões” para detecção de objetos dentro das imagens em junção a certas mudanças no método tradicional de treinamento supervisionado, foi possível criar uma rede que consegue detectar e classificar diferentes regiões de uma mesma imagem com acurácia semelhante a CNN original.

Porém, estas redes neurais convolucionais regionais ou *R-CNN* apresentam certa complexidade devido às várias localizações candidatas que possam possuir um objeto, chamadas propostas, que devem ser processadas, e em seguida refinadas para obter uma localização precisa, levando a soluções que comprometem a velocidade ou acurácia do algoritmo. Assim, foi proposto um método para simplificar e acelerar o processo de treinamento destas R-CNNs utilizando algoritmos para aprender a classificar propostas de objetos e melhorar suas localizações espaciais chamado Fast R-CNN, obtendo velocidades de treinamento nove vezes mais velozes que uma R-CNN comum (GIRSHICK, 2015).

As Redes Neurais Convolucionais mais Rápidas baseadas em propostas de região (Faster R-CNN) foram criadas para tentar resolver o problema das Fast R-CNNs, que ainda utilizam implementação em CPU (Unidade Central de Processamento) da proposta de regiões, pois o tempo gasto na criação das propostas de regiões é equivalente ao tempo demandado para a rede de detecção. Assim, ao utilizar uma segunda Fast R-CNN especializada na computação de regiões e que compartilha camadas de convolução com outra Fast R-CNN usada para detecção

de objetos, este método reduz o tempo para proposição de regiões marginalmente (REN et al., 2015).

4. Metodologia

Inicialmente, será necessário criar um banco de dados compostos por vídeos de alevinos de pintado-real em quantidades variadas para que possam ser realizados o devido treinamento e teste. As coletas de vídeos serão realizadas nas dependências da UCDB no bloco B sala B-116 onde estão localizados o aquário para armazenar os peixes e também está o dispositivo Contador de Alevinos criado por Garcia et al. (2015). Os animais serão adquiridos através de doações do Projeto Pacu e estarão sob os cuidados do grupo de pesquisa INOVISÃO, sob supervisão do professor Hemerson Pistori e com a aprovação de seu manuseio pela Comissão de Ética no Uso de Animais (CEUA) da UCDB até o dia 20/07/2020 com a finalidade de pesquisa científica.

Após criar o banco, ela será dividido em três grupos de modo aleatório: o primeiro grupo com 60% das amostras será utilizado para o treinamento da rede neural, o segundo grupo terá 20% dos vídeos para teste do método e ajuste dos hiperparâmetros para uma melhor classificação, e finalmente os 20% restantes serão usados como grupo de validação final do método após os ajustes necessários.

Para anotar as posições de cada peixe nos vídeos, será utilizado a ferramenta *LabelImg*¹ para criar arquivos XML e guardar regiões retangulares com informações relevantes. Este arquivo XML será usado pela Faster R-CNN para realizar a detecção e contagem dos alevinos do vídeo durante os processos de treinamento e validação.

Inicialmente, serão realizados quatro treinamentos e testes diferentes utilizando o mesmo grupo de vídeos (o primeiro grupo separado para os treinamentos, e o segundo para os testes), com variações na quantidade de épocas que serão necessárias para terminar o treinamento da rede neural. Isto será realizado para determinar a melhor quantidade para treinamento de modo a reduzir o sobreajuste, que ocorre quando a rede neural consegue identificar com grande sucesso o grupo usado para teste porem nao consegue utilizar os padrões aprendidos para identificar objetos fora deste grupo. Além das diferentes quantidades de épocas para cada conjunto de treino

¹ Disponível em: <https://github.com/tzutalin/labelimg>

e teste, também serão realizado ajustes em hiperparâmetros da rede neural afim de tentar evitar o sobreajuste enquanto melhora acurácia durante a validação.

Após cada treinamento e teste, será realizado o processo de validação da rede neural baseado no terceiro grupo de 20% separado anteriormente, utilizando o reconhecimento de padrões aprendido durante o treinamento, e em seguida será realizado então a análise estatística dos resultados obtidos nesta validação, onde serão observadas três métricas: Taxa de classificação correta (TCC), medida-F e área sob a curva ROC. Estas métricas serão usadas para analisar se a utilização da Faster R-CNN supera o método adotado por Garcia et al. (2015), assim como qual quantidade de épocas confere melhor precisão para detecção e contagem dos alevinos. Também serão realizados testes na sede da fazenda do Projeto Pacu, utilizando dados e condições reais e em tempo de execução, podendo assim resultar em um novo produto no mercado agropecuário brasileiro.

5. Cronograma de Execução

Atividades	Meses					
	2	4	6	8	10	12
Revisão de Literatura	x	x				
Coleta de videos		x	x			
Treinamento e teste			x	x		
Validação			x	x		
Análise estatística				x	x	
Implementação da ferramenta					x	x

Relatório Parcial			x			
Relatório Final						x

6. Referências Bibliográficas

PEIXEBR. Anuário PeixeBR da Piscicultura 2019. BRASIL. 2019. Disponível em: <<https://www.peixebr.com.br/Anuario2019/AnuarioPeixeBR2019.pdf>>. Acesso em: 15 mai. 2019.

LING, Yi Jun; LAU, Phooi Yee. Fish monitoring in complex environment. **In: International Workshop on Advanced Image Technology (IWAIT) 2019. International Society for Optics and Photonics**, 2019. p. 1104902. Disponível em: <<https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11049/1104902/Fish-monitoring-in-complex-environment/10.1117/12.2521251.full?SSO=1>>. Acessado em: 16 mai. 2019.

SIGHOLT, T. et al. Handling stress and storage temperature affect meat quality of farmed-raised Atlantic salmon (*Salmo salar*). **Journal of Food Science**, v. 62, n. 4, p. 898-905, 1997. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2621.1997.tb15482.x>>. Acessado em: 16 mai. 2019.

WEDEMEYER, Gary. Some physiological consequences of handling stress in the juvenile coho salmon (*Oncorhynchus kisutch*) and steelhead trout (*Salmo gairdneri*). **Journal of the Fisheries Board of Canada**, v. 29, n. 12, p. 1780-1783, 1972. Disponível em: <<https://www.nrcresearchpress.com/doi/10.1139/f72-284#.XN2sJXVKib8>>. Acessado em: 16 mai. 2019.

EMBRAPA. Brasil inicia o maior projeto de pesquisa já elaborado para desenvolver a aquicultura. BRASIL. 2018. Disponível em: <<https://www.embrapa.br/busca-de-noticias/-/noticia/35429495/brasil-inicia-o-maior-projeto-de-pesquisa-ja-elaborado-para-desenvolver-a-aquicultura>>. Acessado em: 16 mai. 2019.

Food and Agriculture Organization of the United Nations (FAO). Fishery and Aquaculture Country Profiles: The Federative Republic of Brazil. Disponível em: <<http://www.fao.org/fishery/facp/BRA/en>>. Acessado em: 18 mai. 2019.

EMBRAPA. Tema: Pesca e Aquicultura: Nota Técnica. BRASIL . 2015. Disponível em: <<https://www.embrapa.br/tema-pesca-e-aquicultura/nota-tecnica>>. Acessado em: 18 mai. 2019.

GARCIA, Vanir. Sistema para contagem automática de alevinos baseado em visão computacional. **UNIVERSIDADE CATOLICA DOM BOSCO** , 2015. Disponível em: <<http://www.gpec.ucdb.br/pistori/orientacoes/planos/vanir2015.pdf>>. Acessado em: 18 mai. 2019.

KAEHLER, Adrian; BRADSKI, Gary. Learning OpenCV 3: computer vision in C++ with the OpenCV library. " O'Reilly Media, Inc.", 2016.

ROWLEY, Henry A. Neural network-based face detection. **CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER SCIENCE**, 1999. Disponível em: <<https://apps.dtic.mil/dtic/tr/fulltext/u2/a366182.pdf>>. Acessado em: 20 mai. 2019.

PERELMUTER, Guy et al. Reconhecimento de imagens bidimensionais utilizando Redes Neurais Artificiais. **Anais do VIII Sibgrapi**, p. 197-203, 1995. Disponível em: <<http://sibgrapi.sid.inpe.br/col/sid.inpe.br/sibgrapi/2013/02.18.15.52/doc/25%20Reconheciment%20de%20imagens%20bidimensionais.pdf>>. Acessado em: 20 mai. 2019.

REN, Shaoqing et al. Faster r-cnn: Towards real-time object detection with region proposal networks. In: **Advances in neural information processing systems**. 2015. p. 91-99. Disponível em:

<<http://papers.nips.cc/paper/5638-faster-r-cnn-towards-real-time-object-detection-with-region-proposal-networks.pdf>>. Acessado em: 15 mai. 2019

GIRSHICK, Ross. Fast r-cnn. In: **Proceedings of the IEEE international conference on computer vision**. 2015. p. 1440-1448. Disponível em: <https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Girshick_Fast_R-CNN_ICCV_2015_paper.pdf>. Acessado em: 15 mai. 2019.

WEINSTEIN, Ben G. A computer vision for animal ecology. **Journal of Animal Ecology**, v. 87, n. 3, p. 533-545, 2018. Disponível em: <<https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/1365-2656.12780>>. Acessado em: 21 mai. 2019.

KARPATHY, Andrej et al. Large-scale video classification with convolutional neural networks. In: **Proceedings of the IEEE conference on Computer Vision and Pattern Recognition**. 2014. p. 1725-1732. Disponível em: <https://www.cv-foundation.org/openaccess/content_cvpr_2014/papers/Karpathy_Large-scale_Video_Classification_2014_CVPR_paper.pdf>. Acessado em: 21 mai. 2019.

GIRSHICK, Ross et al. Rich feature hierarchies for accurate object detection and semantic segmentation. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. 2014. p. 580-587. Disponível em: <http://openaccess.thecvf.com/content_cvpr_2014/papers/Girshick_Rich_Feature_Hierarchies_2014_CVPR_paper.pdf>. Acessado em: 21 mai. 2019.