



Segmentação por Instâncias de Estágios de Desenvolvimento de Cochonilha de Escama com Mask R-CNN

Nikolaos Daskalakis¹, Josenalde B. Oliveira¹

¹Laboratório de Tecnologias Aplicadas às Ciências Agrárias – Escola Agrícola de Jundiá - UFRN
Caixa Postal 7 – 59.280-000 – Macaíba – RN – Brasil

nikolaos.ferreira.714@ufrn.edu.br, josenalde.oliveira@ufrn.br

Abstract. *This work presents preliminary results of a recognition, classification and counting system for scale cochineal (*Diaspis echinocacti*) in cactus pear samples, with the objective of predicting first and third instar stages of the cochineal, in addition to identifying males and females. Machine learning algorithms and the Mask R-CNN model, image processing techniques, and the use of technologies such as cloud computing were used. There are several challenges along the way, such as data labeling, demand for high computational processing power, with that there is room for improvement, but the model has learned the patterns well, and there is great progress.*

Resumo. *Este trabalho apresenta resultados preliminares de um sistema de reconhecimento, classificação e contagem para cochonilha de escama (*Diaspis echinocacti*) em amostras de palma forrageira, com objetivo de classificar fases da cochonilha de primeiro e terceiro instar, além de identificar machos e fêmeas. Foram utilizados algoritmos de machine learning e o modelo Mask R-CNN, técnicas de processamento de imagens, e uso de tecnologias como a cloud computing. Há vários desafios no caminho como a rotulagem dos dados, demanda de um alto poder de processamento computacional, com isso há possibilidade de melhorias mas o modelo aprendeu bem os padrões, e há grande avanço.*

1. Introdução

A palma forrageira é um suporte alimentar de grande valor para a pecuária do semiárido, com características como alta palatabilidade, produção de biomassa, potencial nutricional e resistência à seca. Além do nordeste brasileiro, países como Argélia, África do Sul, Israel, México, USA e Colômbia destacam-se em seu cultivo para alimentação de rebanhos. Contudo, é sujeita a ação de pragas, onde destaca-se a cochonilha de escama (*Diaspis echinocacti*). Esta pode enfraquecer a palma forrageira, deixar com tons amarelos e provocar queda das raquetes. Não havendo o manejo e tratamento adequado pode comprometer o palmar. Identificar em laboratório a ação de produtos naturais de controle como óleos vegetais tem sido alvo de pesquisas, como em Chagas et al. (2018). A análise é realizada em amostras de raquetes selecionadas e a contagem das cochonilhas realizadas por inspeção visual com auxílio de microscópio, embora ferramentas

computacionais também tenham sido propostas (OLIVEIRA et al., 2021), mas para contagem total dos insetos, independente do estágio de desenvolvimento. Um outro desafio consiste em identificar os insetos conforme o ciclo de vida, pelo impacto de maior dispersão nos primeiros sinais de aparecimento da praga. A cochonilha consiste em diferentes estágios de desenvolvimento. A observação de uma superfície pulverulenta nas faces da palma, pode indicar uma infestação de ninfas de 1º instar. Ao fixarem-se, as fêmeas começam a evoluir com aspecto circular ou elipsóide, passando pelo 2º e 3º instares. Já os machos possuem formato cilíndrico, alongado e possuem asas.



Figura 1. (a) Presença da cochonilha de escama (*Diaspis echinocacti*) na palma. (b) ciclo de vida com respectivos estágios

Este trabalho apresenta a detecção e segmentação dos estágios de desenvolvimento da cochonilha de escama (classes) por meio do aprendizado supervisionado profundo com rede neural convolucional do tipo Mask R-CNN (HE et al, 2017; REHMAN et al., 2021). A partir de imagens com amostras de raquetes de palma, o modelo proposto reconhece, classifica e quantifica a respectiva classe.

2. Problemática

Em visão computacional a detecção de objetos é um desafio computacional, sendo associada à etapa de segmentação. São considerados dois tipos de segmentação: a semântica, que rotula cada pixel de uma imagem com base em sua respectiva classe, e a segmentação por instâncias, onde cada instância (objeto) é tratada como unitária, independente de sua classe. Em resumo, a segmentação de instâncias detecta cada objeto individual dentro de um grupo de objetos semelhantes (MEIRA, 2020), conforme observado na Figura 2. Usualmente requer maior poder computacional.

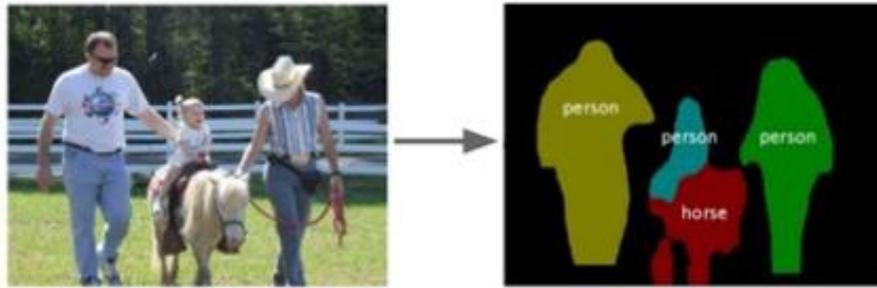


Figura 2. Exemplo da segmentação de instâncias. (MEIRA, 2020)

3. Metodologia

Este trabalho utiliza o aprendizado de máquina supervisionado, onde é necessário fornecer dados de treinamento (imagens de exemplo rotuladas) para o modelo aprender e identificar os padrões. O modelo utilizado neste trabalho foi o Mask R-CNN, uma arquitetura de rede neural convolucional baseada em regiões (R-CNN) bastante utilizada para a tarefa de segmentação de instâncias.

É uma extensão do algoritmo Faster R-CNN, utilizado para detecção de objetos. Produz como resultado os *bounding boxes* dos respectivos objetos com o percentual de acurácia na detecção. A técnica adiciona uma camada para prever máscaras de segmentação para cada região de interesse (RoI) em paralelo com as tradicionais camadas para classificação e regressão. Esta camada adicional é uma rede convolucional simples aplicada a cada RoI, com predição pixel a pixel. Internamente o modelo é composto por dois estágios, sendo o primeiro responsável por gerar regiões potenciais de existência de objetos (*proposals*, *RPN*) e o segundo prediz a classe, gerando os *bounding boxes* e máscaras, conforme a visão geral na Figura 3.

Foi utilizada a biblioteca Mask_RCNN¹ com a linguagem de programação python versão 3.7. Antes de criar o modelo é necessário elaborar o *imageset* para o problema em questão. O *imageset* é composto por 172 imagens utilizadas em Oliveira et. al. (2021), e cada imagem com sua anotação XML. As anotações foram feitas no Label Studio em formato pascal VOC (Visual Object Classes) e formato de caixas retangulares. O formato das imagens são .jpg (JPEG) na proporção quadrada de 300 pixels, e foram recortadas de imagens maiores (Figura 4); variam em resolução e

¹ Muitos exemplos e artigos utilizam a versão disponível em https://github.com/matterport/Mask_RCNN

escala dos objetos. Esta é uma das contribuições deste trabalho, visto não haver *imageset* similar para cochonilhas de escama disponíveis na web.

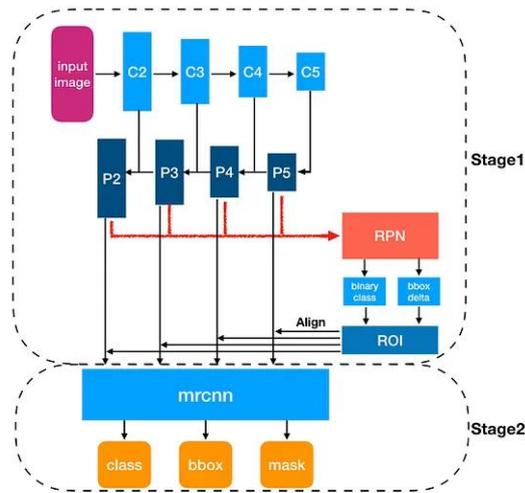


Figura 3. Estrutura das camadas internas Mask-rcnn (XIANG, 2018)



Figura 4. Rotulagem com caixas delimitadoras. Preto: 1º instar. Vermelho: fêmeas (2º e 3º instares). Azul: machos.

Para treinar o modelo, é recomendado uma máquina bastante robusta e com GPU. Para ganho de performance, foram utilizados recursos de Google Cloud onde se prepara uma máquina remota com especificações personalizadas. Inicialmente foram realizados treinos sem GPU, mas foi inviável com tempo de treino muito longo, cada passo levaria em média 2 minutos; no segundo momento os treinos foram executados em GPU uma Nvidia T4, onde foi observado uma redução drástica no tempo de treino, cada passo sendo executado em média 1 segundo, e um treino total durou em média 3 minutos.

4.1. Treino do modelo

A proporção utilizada foi de 80/20, sendo 80% dos dados para treino e 20% para teste. Os dados não foram embaralhados para evitar aleatoriedade e observar melhor as

variações dos hiperparâmetros (repetibilidade). Foi utilizado também *data augmentation* na proporção de +100% apenas com transformações geométricas. O *data augmentation* amplia o *dataset* no treinamento e evita o *overfit*, melhorando a previsão geral do modelo. Neste trabalho se utilizou taxa de 100%, ou seja, foram geradas mais 172 imagens com transformações geométricas de escala, rotação, e espelhamentos vertical e horizontal com os seguintes parâmetros: *Fliplr*(1), *Flipud*(1), *Affine*(rotate=(-45, 45)), *Affine*(rotate=(-90, 90)) e *Affine*(scale=(0.75, 1.2)); usando a biblioteca *imgaug*. Várias execuções foram aplicadas com variações de hiperparâmetros. Foi observado que um treino de camadas mais profundas propiciam resultados melhores que somente camadas iniciais ou ‘heads’. Um dos hiperparâmetros relevantes que foi observado foi a escala das âncoras (Figura 5) na Rede de Propostas de Região (RPN), chamado de *RPN_ANCHOR_SCALES*. Valores menores melhoraram a precisão do modelo.

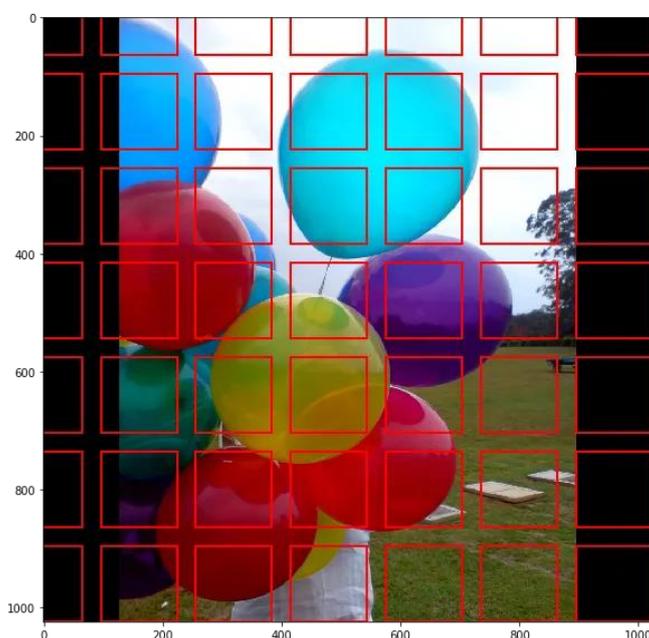


Figura 5. A ilustração representa caixas de âncoras, que podem ser em diferentes escalas.

O modelo que expressou melhores resultados foi teve seus parâmetros configurados para *batch_size* = 3, *steps_per_epoch* = 140, *validation_steps* = 40, *learning_rate* = 0.005, *learning_momentum* = 0.4, *rpn_anchor_ratios* = [0.5, 1, 1.5] e *rpn_anchor_scales* = [10,30,50,75,92]

5. Resultados preliminares

Foram salvos alguns outputs de treino para estimar o desempenho do modelo, a seguir métricas de previsão de classe, da época 2 a 4 com 140 iterações: Epoch 2/4, *mrcnn_class_loss*: 0.3959 e *val_mrcnn_class_loss*: 0.4007; Epoch 3/4, *mrcnn_class_loss*: 0.3585 e *val_mrcnn_class_loss*: 0.4043; Epoch 4/4, *mrcnn_class_loss*: 0.3315 e *val_mrcnn_class_loss*: 0.3834.

A Figura 6 apresenta alguns resultados. Em casos de aglomeração (Figura 6a), o modelo não foi bem em prever com acurácia todos os registros. O desejado, é o modelo reconhecer aglomerados do terceiro instar de cochonilhas fêmeas e não somente objetos

individuais. Isso deve ser resolvido, possibilitando assim avaliar a precisão do modelo; uma alternativa seria adicionar uma classe extra para representar estes aglomerados, isso facilitaria o reconhecimento para o modelo. O modelo teve melhor êxito em reconhecer diferentes escalas e tamanhos, graças ao Mask R-CNN, apresentando boa acurácia.

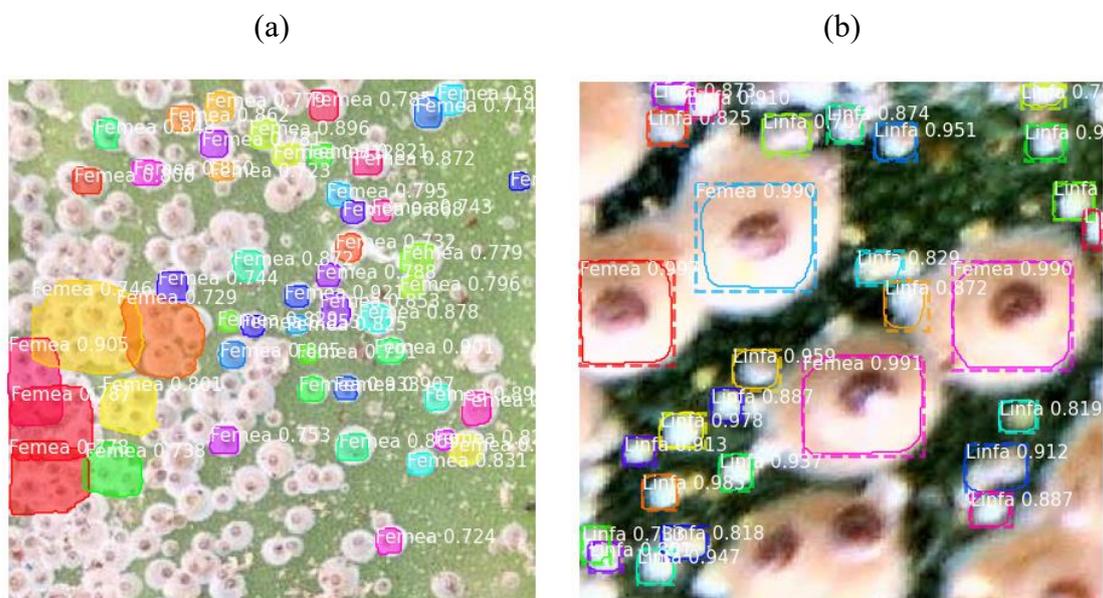


Figura 6 Saídas em diferentes tamanhos a) e escalas b). (reconhecimento de ninfa: 1 instar e fêmea (2. e 3 instares))

Referências

- Chagas, M.C.M., Silva, E.C.S., Nascimento, S.M., Lima, G.F.C., Lima, T.C.C. (2018). Cochonilha de escama na palma forrageira: aspectos biológicos e estratégias de manejo. Natal: EMPARN.
- He, K., Gkioxari, G., Dollár, P., Girshick, R. (2017). Mask R-CNN. In: IEEE Int. Conf. on Computer Vision (ICCV), p. 2980-2988.
- Meira, N. Edge AI: mask r-cnn e segmentação de instâncias. (2020). Disponível em: <<http://www2.decom.ufop.br/imobilis/segmentacao-instancias/>>. Acesso: 17/07/23.
- Rehman, Z.; Khan, M.A., Ahmed, F., Damasevicius, R., Naqvi, S.R., Nisar, W., Javed, K. (2021). Recognizing apple leaf diseases using a novel parallel real-time processing framework based on Mask R-CNN and transfer learning: an application for smart agriculture. IET Image Processing, p. 1-12.
- Xiang, Z. Simple understanding of Mask R-CNN. (2018). Disponível em: <https://alittlepain833.medium.com/simple-understanding-of-mask-rcnn-134b5b330e95>. Acesso: 17/07/23.
- Oliveira, J.B., Araújo, A.M.S., Teixeira, L.R.L., Silva, E.C.S., Santana, L.E.A.S., Rodrigues, T.C., Chagas, M.C.M. (2021). Uma proposta de sistema embarcado para contagem automatizada de cochonilhas de escama em laboratório. In: Anais do Simpósio Brasileiro de Automação Inteligente, p. 2063-2068.