

UTILIZAÇÃO DE IMAGENS SINTÉTICAS PARA IDENTIFICAÇÃO DE PROUTOS AGRICOLAS: ENFASE EM RECONHECIMENTO DE MAÇÃS

JOAB TAVARES FAGUNDES¹; FELIPE MACHADO²; FELIPE DIAS LOPES³;
MARCELO LEMOS ROSSI⁴

¹Universidade Federal de Pelotas – joabtavaresf02@gmail.com

²Universidade Federal de Pelotas – fmachado@ufpel.edu.br

³Universidade Federal de Pelotas – felopes_7@hotmail.com

⁴Universidade Federal de Pelotas – marcelo.rossi@ufpel.edu.br

1. INTRODUÇÃO

Cada vez mais técnicas de automação tem adentrado em tecnologias da área agrícola. Com a vinda da agricultura 4.0 mais e mais técnicas de inteligência artificial tem sido utilizada no auxílio de tomadas de decisão e de automação de processos, conforme pode ser visto em MORAES (2020) e REDDY (2021). A dificuldade da utilização de inteligência artificial é a necessidade de uma grande quantidade de dados para o treinamento adequado de forma que o sistema possa “aprender”. Além das análises de dados mensuráveis há também o crescente uso de imagens como apresentado em SHRESTHA et al. (2015) VREŠAK et al. (2016) e SHRESTHA et al. (2018). O uso de imagens pode ser utilizado na agricultura de precisão, automatizando o processo de análise e robotização da agricultura. Mais uma vez, o problema é a necessidade de uma grande quantidade de dados.

Sistemas de inteligência artificial podem ser utilizados na análise de imagens com três objetivos: Classificação de imagens; Reconhecimento de imagens; e Segmentação de imagens. Para o processo de classificação as imagens são separadas em grupos e o treinamento permite a identificação ou não de se um determinado elemento está na imagem. Já no processo de reconhecimento é feita uma anotação da área de cada imagem onde teria o elemento de interesse, o que exige um trabalho maior para preparar o *dataset* (conjunto de dados) maior do que o processo de classificação. Por fim, o processo de segmentação tem como objetivo de identificar quais os pixels da imagem pertencem ao elemento de interesse e, conseqüentemente, para preparar o *dataset* é necessário criar máscaras em que cada pixel identifica o elemento na imagem.

Para que uma inteligência artificial possa ser aproveitada na agricultura de precisão, com a utilização de robótica nas colheitas, o ideal é o processo de segmentação de imagem para a correta identificação e localização de objetos de interesse. A exemplo temos o trabalho de BARGOTI e UNDERWOOD (2017) que busca uma técnica de processamento de imagem para a segmentação para aplicação de robótica em pomares de maçãs.

Conforme apresentado, o processo para criar um *dataset* que permita a segmentação é muito trabalhoso e oneroso, já que o *dataset* deve ser constituído tanto das imagens quanto das máscaras necessárias para treinar a segmentação. Os custos envolvidos neste processo estão relacionados à obtenção das imagens e, também, realizar as marcações pixel-a-pixel para a construção das máscaras. Tal processo pode ser facilitado através de um processo automatizado para a criação das imagens e das máscaras, dando origem a imagens sintéticas que possam ser utilizadas em ambientes reais.

Assim, este trabalho tem como objetivo demonstrar que a utilização de imagens sintéticas criadas pela renderização de modelagens foto realistas podem

ser utilizadas. Para demonstrar isso foi gerado um *dataset* sintético que foi utilizado para treinar uma rede neural do tipo U-Net capaz de realizar segmentação de imagem e identificar pixels pertencentes a maçãs em imagens de pomares encontrados na internet. Escolheu-se maçãs pois a geometria desta fruta é de fácil modelagem 3D.

2. METODOLOGIA

Na primeira etapa buscou-se referências sobre a geometria de maçãs de tal forma a subsidiar o processo de modelagem 3D. De posse de algumas referências foi feito a modelagem 3D de elementos que possam representar as maçãs e buscou-se texturas foto realistas de maçãs de forma que a renderização das maçãs possa gerar resultados foto realistas. A Figura 1 apresenta dois modelos 3D de maçãs produzidos e renderizados no software Unity.



Figura 1 – Modelos de maçãs com textura foto realista

De posse dos modelos utilizou-se a biblioteca *Perception* da UNITY (2020) para gerar os dados sintéticos a serem utilizados no treinamento da rede U-Net. A Figura 2 apresenta o resultado do processo realizado pela biblioteca *Perception*, onde temos a imagem sintética e a máscara que será utilizada para treinar a segmentação.

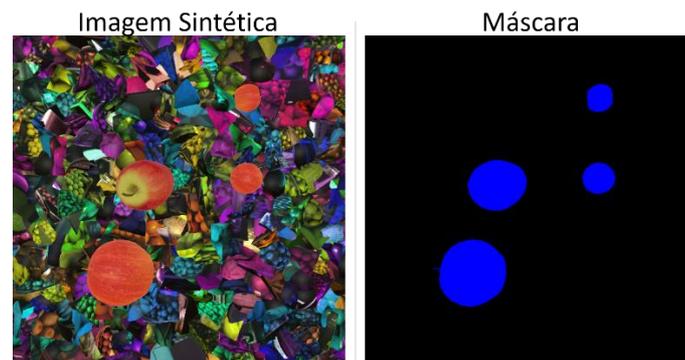


Figura 2 – Imagem sintética e a sua máscara

Na Figura 2 podemos verificar que na geração da imagem sintética foi utilizado um fundo aleatório com objetivo de “distrair” a rede neural, de forma que ela possa ter uma maior abstração e ser capaz de operar adequadamente em imagens que contenham elementos que possam atrapalhar no processo de reconhecimento das maçãs (galhos, folhas, etc.). Ainda na Figura 2 podemos ver que a máscara gerada consiste em uma imagem com todos os pixels com valor igual a 0 (zero), exceto se o pixel faz parte da maçã que o valor foi igual a 1 (um), que é representado de azul.

Após configurado o ambiente na Unity foi gerado 5.000 imagens com as maçãs tendo posição, rotação e tamanhos variados. Todo o processo levou aproximadamente 15 minutos utilizando um PC com processador i7-10700 e 16Mb

de memória RAM. Para cada imagem gerada foi gerado, também, a máscara para a imagem correspondente. Com posse deste *dataset* foi utilizada a plataforma Google Colab e a biblioteca TensorFlow para criar a rede U-Net que utiliza como codificador (*downsampler*) uma rede MobileNetV2 pré-treinada e como decodificador (*upsampler*) uma rede Pix2pix. A Figura 3 apresenta a estrutura da rede U-Net utilizada neste trabalho.

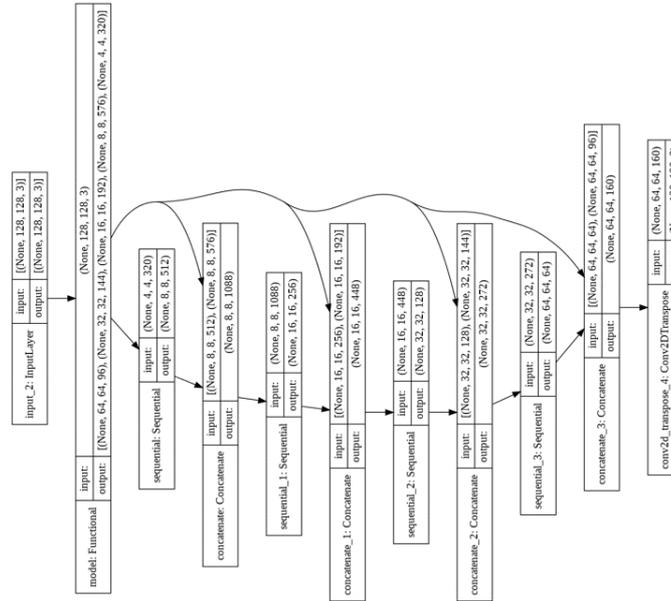


Figura 3 – Estrutura da rede U-Net

A rede neural apresentada na Figura 3 foi treinada por 20 épocas, utilizando lotes de com tamanho de 64 no ambiente do Colab utilizando GPU.

3. RESULTADOS E DISCUSSÃO

A metodologia apresentada indicou um excelente resultado com as imagens do *dataset*, conforme apresentado na Figura 4.

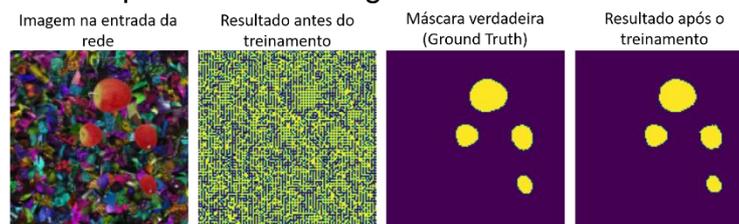


Figura 4 – Imagem sintética e a sua máscara

Observando os dados apresentados na Figura 4 podemos dizer que a rede U-Net utilizada foi capaz de aprender a realizar a segmentação da imagem.

De forma a validade o resultado buscou-se imagens de maçãs na internet e essas imagens foram utilizadas na rede U-Net criada para este trabalho. Um dos resultados encontra-se na Figura 5.



Figura 5 – Segmentação realizada pela rede U-Net utilizando imagens da internet

Na Figura 5 podemos verificar que a rede U-Net, treinada com apenas imagens sintéticas, foi capaz de realizar a segmentação de imagens obtidas através de fotografias do objeto real de forma satisfatória. Ainda observando a Figura 5 podemos verificar pequenos conjuntos de pixels claros isolados onde não correspondem o que era maçã, indicando a ocorrência alguns falsos positivos, e de alguns pixels escuros na região correspondente ao que deveria ser a maçã, indicando a ocorrência de alguns falsos negativos. A rede treinada encontra-se em <https://Colab.research.google.com/drive/1kc7TQVpC3ojvulsJIYis81Lebe0Rb9J2?u=sharing>

4. CONCLUSÕES

Conforme indicado na Figura 5 o resultado de todo o processo apresentou um resultado satisfatório, indicando que a rede treinada foi capaz de realizar a segmentação de imagens obtidas por fotografia mesmo sendo treinada puramente por imagens sintéticas. A grande vantagem deste processo é a agilidade para a criação do *dataset* que foi todo criado pelo computador eliminando, assim, a necessidade de um operador humano obter as imagens e gerar as máscaras para cada uma das 5.000 imagens utilizadas no *dataset*.

5. AGRADECIMENTOS

Os autores deste trabalho agradecem à Universidade Federal de Pelotas, ao CNPq, à Fapergs pelas bolsas de estudo concedidas e à Unity Technologies pela licença do software.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- REDDY, P. V. S. (2021) Data mining and fuzzy data mining using map reduce algorithms. **Data Mining: Methods, Applications and Systems**. DOI: 10.5772/intechopen.92232. Disponível em: <https://www.intechopen.com/chapters/72112>
- VREŠAK, M., OLESEN, M. H., GISLUM, R., BAVEC, F. JØRGENSEN, J.R. (2016) The use of image-spectroscopy technology as a diagnostic method for seed health testing and variety identification. **PLoS ONE**, v. 11, n. 3, p. 1-10, 2016. <https://doi.org/10.1371/journal.pone.0152011>
- SHRESTHA, S., DELEURAN, L. C., OLESEN, M. H., GISLUM, R. (2015) Use of multispectral imaging in varietal identification of tomato. **Sensors**, v. 15, n. 2, p. 4496–4512, 2018. <https://doi.org/10.3390/s150204496>
- BOELT, B., SHRESTHA, S., SALIMI, Z., JØRGENSEN, J., NICOLAISEN, M., CARSTENSEN, J.. Multispectral imaging - A new tool in seed quality assessment? **Seed Science Research**, v. 28, n. 3, p. 222-228. DOI: 10.1017/S0960258518000235.
- BARGOTI, S., UNDERWOOD, J. P.. Image Segmentation for Fruit Detection and Yield Estimation in Apple Orchards. **Journal of Field Robotics**, v. 34, n. 6, p. 1039-1060. 2017. DOI 10.1002/rob.21699
- Moraes, N. A. B. **Predição de ranqueamento de lotes de sementes de milho por inteligência artificial**. 2020. Dissertação de Mestrado. Programa de Pós-graduação em Ciência e Tecnologia de Sementes, Faculdade de Agronomia Eliseu Maciel, Universidade Federal de Pelotas.
- UNITY TECHNOLOGIES. Unity Perception Package, 2020. Disponível em <https://github.com/Unity-Technologies/com.unity.perception>