

# DETECÇÃO E SEGMENTAÇÃO DE LARANJAS EM IMAGENS DA PLANTAÇÃO - UM ESTUDO COM LARANJAS NÃO MADURAS DO TIPO FOLHA MURCHA

Matheus Felipe Gremes<sup>1</sup>, Murilo Silva Colombo<sup>2</sup>, Rafael Krummenauer<sup>3</sup>, Franklin Cesar Flores<sup>4</sup>, Cid Marcos Gonçalves Andrade<sup>5</sup>, Oswaldo Curty da Motta Lima<sup>6</sup>

<sup>1</sup>Mestrando do Programa de Pós-graduação em Engenharia Química, Campus Maringá/PR, Universidade Estadual de Maringá – UEM. Bolsista CAPES. pg402992@uem.br.

<sup>2</sup>Doutorando, Mestre, Programa de Pós-Graduação em Engenharia Química, Campus Maringá/PR, Universidade Estadual de Maringá – UEM. pg401365@uem.br

<sup>3</sup>Doutor, Departamento de Engenharia Química, Campus Maringá/PR, Universidade Estadual de Maringá – UEM. rkrummenauer2@uem.br.

<sup>4</sup>Doutor, Departamento de Informática, Campus Maringá/PR, Universidade Estadual de Maringá – UEM. fcflores@din.uem.br.

<sup>5</sup>Coordenador, Doutor, Departamento de Engenharia Química, Campus Maringá/PR, Universidade Estadual de Maringá – UEM. cid@deq.uem.br.

<sup>6</sup>Orientador, Doutor, Departamento de Engenharia Química, Campus Maringá/PR, Universidade Estadual de Maringá – UEM. oswaldo@deq.uem.br.

## RESUMO

A agricultura vem se tornando cada vez mais importante para a humanidade. Sistemas capazes de realizar a colheita de uma safra de forma automática utilizando robôs ou de realizar uma boa estimativa de produção podem diminuir os custos envolvidos e aumentar a eficiência da produção. Com o avanço da visão computacional, métodos que utilizem processamento de imagem estão se tornando cada vez mais viáveis nas soluções de problemas da agricultura. Dessa forma, o objetivo desse trabalho é implementar uma solução integrando a rede *YoLov4* e o algoritmo de segmentação *watershed* para realizar a detecção e segmentação de laranjas a partir de imagens da plantação. A rede *YoLov4* foi competente em realizar a detecção das laranjas mesmo essas tendo uma coloração que se assemelham às folhagens aos seus arredores, enquanto o algoritmo *watershed* se beneficiou das informações fornecidas pela rede para posicionar os marcadores centralizados nas laranjas e realizar a segmentação apenas no interior das caixas delimitadoras. Para a quantificação dos erros de estimação foi proposto o emprego de um operador de diferenças simétricas entre o modelo do formato estimado da laranja dentro da caixa delimitadora fornecida pela rede e o resultado da segmentação fornecido pelo algoritmo *watershed*. Os resultados encontrados mostram que laranjas maiores, que estejam posicionadas na parte superior da árvore ou que não estejam obstruídas por folhagens ou por outras laranjas apresentam uma melhor segmentação e, portanto, uma diferença simétrica menor.

**PALAVRAS-CHAVE:** Segmentação; *Watershed*; Laranjas Pré-colheita; *YoLov4*.

## 1 INTRODUÇÃO

A agricultura se tornou uma das áreas mais essenciais para a humanidade. As frutas, grande fonte de nutrientes, estão presentes na dieta da maioria das pessoas. Dessa forma, a produção contínua é necessária para atingir as demandas globais (ABDULLAHI; SHERIFF; MAHIEDDINE, 2017). Uma das formas de aumentar a qualidade de produção e diminuir os custos é através de inovações tecnológicas como a visão computacional, que vem mostrando avanços significativos no processamento de imagens de frutos pré colheita (GREMES et al., 2021).

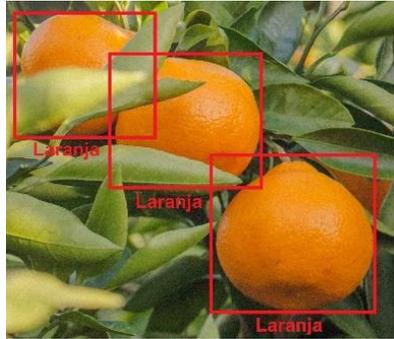
As duas aplicações de visão computacional que mais se desenvolveram nessa área recentemente foram no setor da estimativa de produção dos frutos e no setor da colheita robótica (GREMES et al., 2021).

Para realizar a estimativa de produção com precisão, têm se mostrado efetivos algoritmos de detecção de objetos da família *YoLo* (REDMON et al., 2016) para realizar a contagem das frutas diretamente das árvores (KOIRALA et al., 2019), (BRESILLA et al., 2019) e (XIONG et al., 2020).

Para realizar a colheita das frutas utilizando robôs de forma automática geralmente é necessária mais informação do que as fornecidas por algoritmos de detecção que

fornece apenas as caixas delimitadoras dos frutos. Nesse tipo de aplicação, é mais frequente a utilização de algoritmos de segmentação de objetos como a arquitetura de rede *Mask Region Neural Network* (Mask R-CNN) (HE et al., 2017), como demonstrado nos trabalhos (YU et al., 2019), (GE et al., 2019) e (KANG; CHEN, 2019).

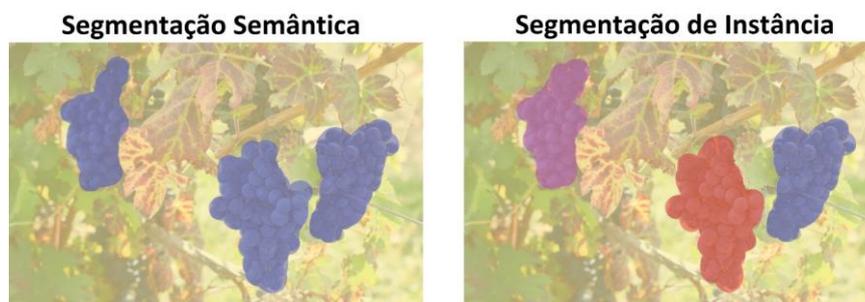
Detecção de objetos é a função que consiste em detectar e demarcar caixas delimitadoras em objetos em uma imagem (GOLLAPUDI, 2019). Veja na Figura 1 um exemplo de detecção.



**Figura 1:** Detecção de imagem em foto com 3 Laranjas

**Fonte:** (ORANGE FRUIT PHOTO – FREE FOOD IMAGE ON UNSPLASH, 2021) alterado pelo autor

Segmentação de imagem é a tarefa de subdividir uma imagem em regiões ou objetos que a compõem (GONZALEZ; WOODS, 2007). Existem duas formas principais de segmentação de imagem: segmentação semântica (GUO et al., 2018) e segmentação de instância (HAFIZ; BHAT, 2020). O objetivo da segmentação semântica é rotular cada pixel da imagem por classe de acordo com o objeto ou região dentro do qual está inserido (GARCIA-GARCIA et al., 2018). Enquanto a segmentação de instância fornece rótulos diferentes para entidades diferentes de objetos pertencentes à mesma classe de objeto. Dessa forma, a segmentação de instância pode ser definida como a tarefa de encontrar soluções simultâneas para a detecção de objetos e para a segmentação semântica (CHEN et al., 2018). Na Figura 2 é mostrado um exemplo de segmentação semântica e segmentação de instância.



**Figura 2:** Segmentação semântica e Segmentação de instância em cachos de uva

**Fonte:** (BLUE BERRIES ON BROWN STEM PHOTO – FREE NATURE IMAGE ON UNSPLASH, 2021) alterado pelo autor

*Watershed* é uma técnica de segmentação de imagem que pertence ao campo da morfologia matemática. *Watershed* é baseado na visualização da imagem como tendo três dimensões: duas coordenadas espaciais e uma de intensidade (VINCENT; MACHINE, 1991). Nessa interpretação topográfica, consideramos três tipos de pontos: a) pontos pertencentes a um mínimo regional; b) pontos nos quais uma gota d'água, se colocado no local de qualquer um desses pontos, deslizaria para um mínimo local; e c) pontos nos quais a água teria a mesma probabilidade de cair para mais de um mínimo. Para um mínimo regional específico, o conjunto de pontos que satisfaz a condição b) é chamado de

*watershed* desse mínimo. Os pontos que satisfazem a condição c) formam linhas de crista sobre a superfície e são denominados de linhas de *watershed* (GONZALEZ; WOODS, 2007). O principal objetivo dos algoritmos de segmentação baseados nesses conceitos é encontrar as linhas de *watershed*. Nesse trabalho, utilizamos o algoritmo *watershed* no gradiente da imagem de interesse.

A aplicação direta do algoritmo de *watershed* geralmente leva à segmentação excessiva devido a ruído e outras irregularidades locais do gradiente. Uma abordagem usada para controlar a segmentação excessiva é baseada no conceito de marcadores. Um marcador é um componente conectado pertencente a uma imagem. Temos marcadores internos, associados a objetos de interesse, e marcadores externos, associados ao *background* (GONZALEZ; WOODS, 2007). O problema é, portanto, reduzido a particionar a imagem em dois elementos: objetos de interesse e *background*. O ponto em utilizar marcadores é que eles trazem conhecimento a priori para lidar com o problema de segmentação. Dessa forma, o fato de que o algoritmo de segmentação *watershed* oferece uma estrutura que pode fazer uso eficaz deste tipo de conhecimento é uma vantagem significativa desse método.

O objetivo desse trabalho é utilizar a arquitetura de detecção de objetos *YoLov4* (BOCHKOVSKIY; WANG; LIAO, 2020) para realizar a detecção de laranjas diretamente das árvores e utilizar essa informação para aplicar marcadores e utilizá-los na segmentação das laranjas utilizando o algoritmo *watershed*. Este trabalho trata da variedade denominada "folha murcha", que é uma laranjeira tipo Valência (*Citrus sinensis* (L.) Osbeck) (OLIVEIRA; SCIVITTARO, 2019), e a época de maturação é de 7 meses antes da colheita. Este estudo integra um projeto maior, que engloba vários tipos de laranjeiras em variadas épocas de maturação. Por tratar de um estágio de maturação onde as laranjas estão verdes, o cenário para detecção e segmentação é desafiador e mostra o potencial da proposta.

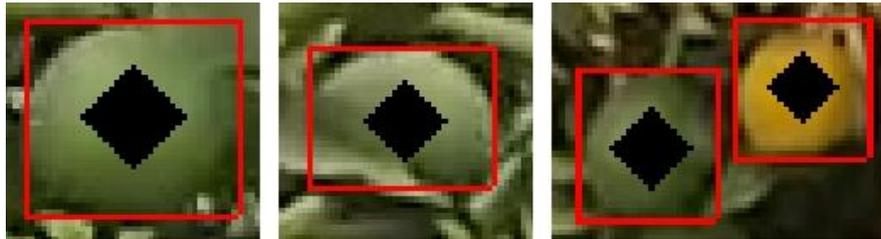
## 2 MATERIAIS E MÉTODOS

Para esse trabalho, foi utilizado a rede *YoLov4* (BOCHKOVSKIY; WANG; LIAO, 2020). Os dados de treinamento foram obtidos em campo utilizando-se um *Smartphone Xiaomi Redmi Note 9PRO*. As laranjeiras foram filmadas com resolução 1920x1080p a 60 quadros por segundo e as imagens para o treinamento da rede foram retiradas dos *frames* do vídeo em intervalos de 3 segundos.

As laranjas das imagens foram então devidamente anotadas utilizando a ferramenta CVAT (COMPUTER VISION ANNOTATION TOOL, 2018) (ROBERTS et al., 2019) totalizando um montante de 287 imagens. Das quais 236 foram separadas para treinamento e validação e 51 para teste. A rede *YoLov4* foi então treinada na plataforma *Google Colaboratory* (COLABORATORY – GOOGLE, 2017).

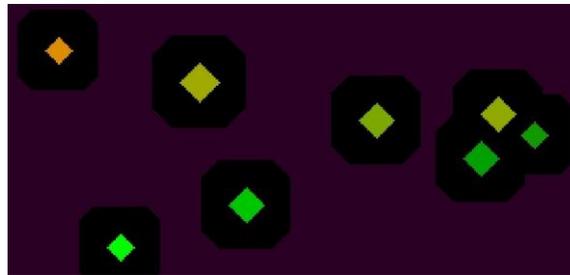
Após o treinamento da rede, foi possível utilizar a informação da localização e do tamanho das laranjas fornecidas pelas caixas delimitadoras da rede *YoLov4* para a segmentação do algoritmo *watershed* com marcadores.

Os marcadores das laranjas utilizados no algoritmo *watershed* foram selecionadas posicionando um marcador em forma de diamante no centro de cada caixa delimitadora como é ilustrado na Figura 3. Os tamanhos dos marcadores internos são calculados em função das caixas delimitadoras resultantes da aplicação da rede, caixas delimitadoras maiores resultam em marcadores maiores, assim como caixas delimitadoras menores resultam em marcadores menores. Apenas as caixas delimitadoras com confiança de classe superior a 50% foram utilizadas.



**Figura 3:** Marcadores das laranjas posicionados no meio das caixas delimitadoras  
**Fonte:** Dados da pesquisa

O marcador do *background* foi selecionado utilizando o conjunto de todos os *pixels* das imagens menos os *pixels* de todos os marcadores das laranjas após a dilatação deles com um elemento estruturante em forma de quadrado, como é ilustrado na Figura 4, em que os objetos coloridos em forma de diamante são os marcadores das laranjas, os *pixels* roxos representam o marcador do *background* e os *pixels* pretos são as regiões que não possuem nenhum marcador.



**Figura 4:** Marcadores coloridos das laranjas e o marcador roxo do *background*  
**Fonte:** Dados da pesquisa

Para calcular o gradiente da imagem, foi, primeiramente, calculado um gradiente para cada canal de cor (vermelho, verde e azul) utilizando-se vizinhança de distância 4 *pixels*. Então para cada posição na imagem, foi selecionada aquela com o maior valor entre os três gradientes anteriormente calculados na mesma posição. Depois disso, o algoritmo *watershed* foi utilizado no gradiente da imagem com os marcadores de laranjas e do *background*, já discutidos nos parágrafos anteriores. A segmentação do *watershed* para as laranjas foi restrita ao interior das caixas delimitadoras, isso significa que mesmo que haja segmentação de laranjas ou de partes delas dos *pixels* que estejam fora das caixas delimitadoras fornecidas pela rede *YoLov4*, esses *pixels* que não estão contidos nas caixas delimitadoras serão *segmentados* como sendo *background*.

Para estimar o erro de segmentação das laranjas resultante do emprego do algoritmo *watershed*, foi utilizada a diferença simétrica (HALMOS, 2017) entre um modelo do formato da laranja dentro da caixa delimitadora e a segmentação da laranja feita pelo algoritmo *watershed*. Uma elipse foi escolhida como modelo para representar o formato de uma laranja dentro de uma caixa delimitadora porque é o formato que a maioria das laranjas que não estão oclusas por outros objetos tem quando detectadas pela rede *YoLov4*. Primeiramente, uma elipse que se encaixa perfeitamente no formato da caixa delimitadora foi criada, como mostrado na imagem na esquerda da Figura 5. Então é aplicado o operador de diferença simétrica entre os *pixels* da elipse com os *pixels* segmentados da laranja, imagem do centro na Figura 5. Nesse caso, a diferença simétrica entre a elipse modelada e a área segmentada pelo *watershed* é o conjunto de todos os *pixels* que pertencem à união dos *pixels* da elipse com os *pixels* da área segmentada e não pertencem à interseção dos *pixels* da elipse com os *pixels* da área segmentada. Esta operação é mostrada na equação 1, assumindo que os *pixels* da elipse estão representados por  $A$  e os *pixels* da área segmentada estão representados por  $B$ .

$$A\Delta B = \{x: x \in (A \cup B) \text{ e } x \notin (A \cap B)\} \quad (1)$$

Dessa forma, quanto mais a segmentação do *watershed* se aproximar da elipse modelada, menor será a diferença simétrica e, em geral, melhor será a segmentação do *watershed*. De outra perspectiva, quanto menos a segmentação do *watershed* se aproximar da elipse estimada, maior será a diferença simétrica e, por conseguinte, pior será a segmentação do *watershed*.



**Figura 5:** Modelo do formato da laranja dentro da caixa delimitadora, segmentação da laranja feita pelo algoritmo *watershed* e sobreposição da estimativa do formato com a segmentação, respectivamente da esquerda para a direita

**Fonte:** Dados da pesquisa

O valor da diferença simétrica para cada laranja é então dividido pelo valor da soma das áreas da laranja segmentada e da elipse em pixels de modo a encontrar a diferença simétrica normalizada para cada laranja. Se a diferença simétrica normalizada de uma laranja for maior que 0,07, é feito uma marcação em formato de cruz no centro da laranja, como ilustrado na Figura 6. O valor 0,07 foi escolhido com base em testes visuais qualitativos e a marcação em formato de cruz foi escolhida para facilitar a análise visual qualitativa dos resultados.



**Figura 6:** Laranja com diferença simétrica normalizada superior ao valor limite e com marcação em formato de cruz no seu centro

**Fonte:** Dados da pesquisa

### 3 RESULTADOS E DISCUSSÕES

As imagens no canto superior esquerdo das Figuras 8, 9 e 10 demonstram as laranjas com as caixas delimitadoras em vermelho feitas pela rede *YoLov4*. Os diamantes pretos posicionados no meio de cada caixa delimitadora são os marcadores utilizados para realizar a segmentação utilizando o algoritmo *watershed*. As laranjas estão enumeradas de forma que aquelas com os menores índices possuem as maiores confianças de classe e aquelas com maiores índices possuem as menores confianças de classe.

As imagens no canto superior direito das Figuras 8, 9 e 10 são os gradientes das imagens originais. As cores no gradiente seguem o mapa de cor da Figura 7. As cores que estão mais a esquerda no mapa de cor, preto e roxo, representam valores menores do gradiente, enquanto que as cores que estão mais a direita no mapa de cor, vermelho e branco, representam valores maiores do gradiente. É importante salientar que as imagens do gradiente possuem apenas uma única banda de cor, em tons de cinza, em que os pixels são transições entre o preto e o branco. Para facilitar a análise visual das imagens do

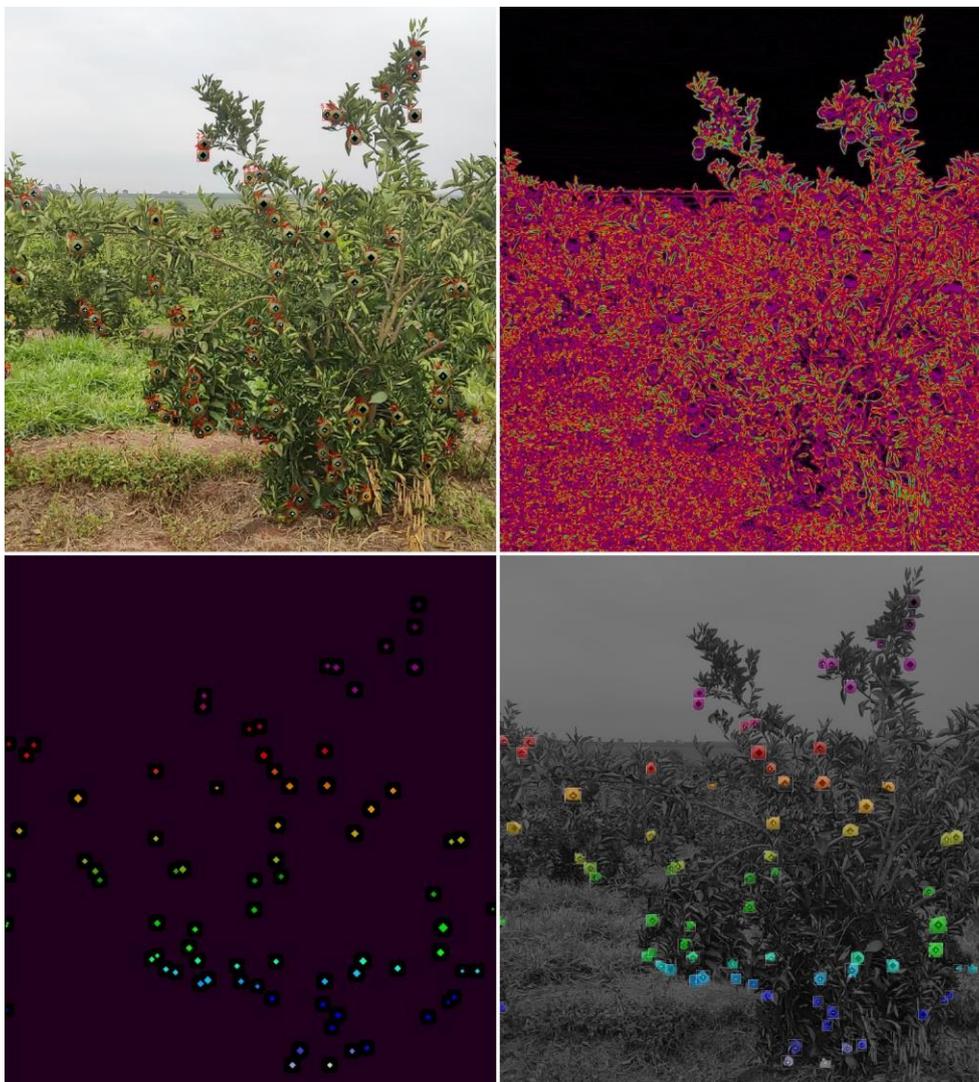
gradiente, foi escolhido o mapa de cor da Figura 7 para representar as intensidades dos pixels.



**Figura 7:** Mapa de cor utilizado para a imagem do gradiente  
**Fonte:** Dados da pesquisa

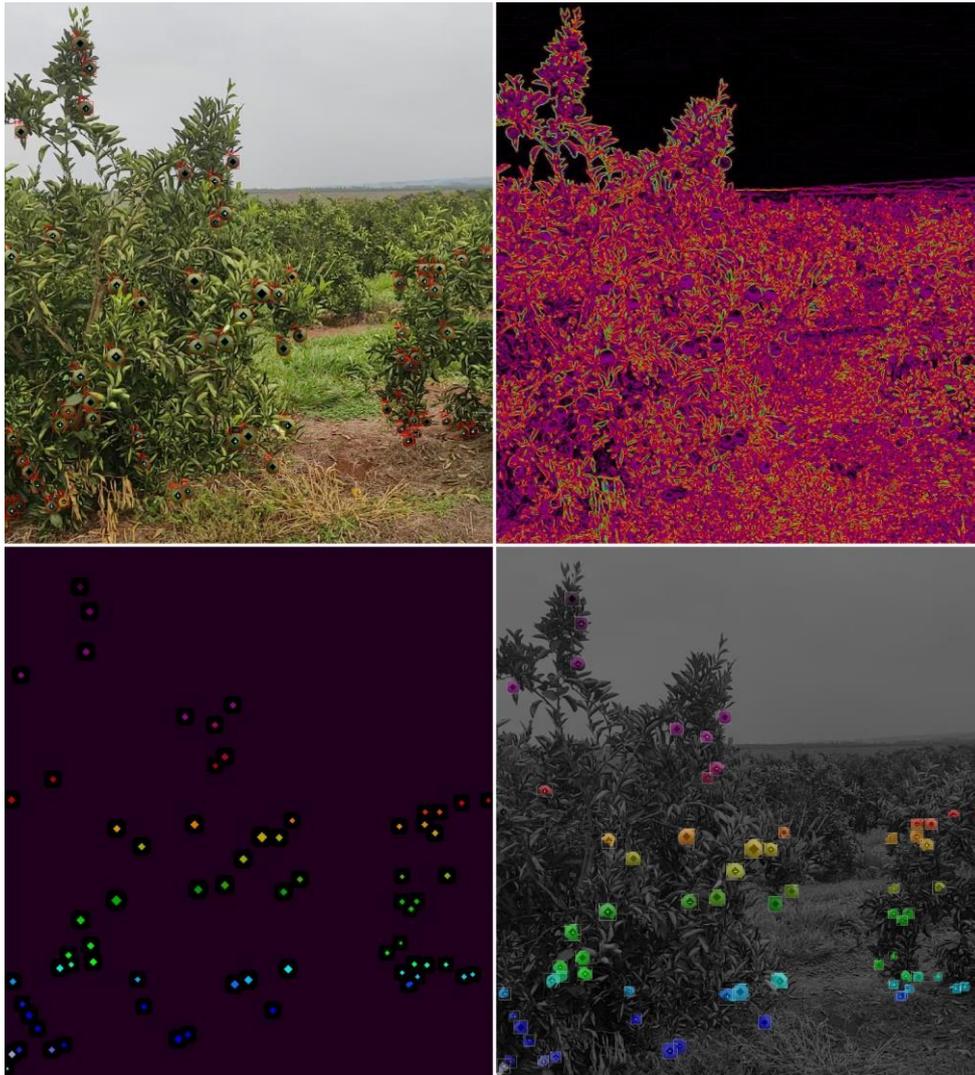
As imagens no canto inferior esquerdo das Figuras 8, 9 e 10 representam os marcadores utilizados no algoritmo *watershed*. Os diamantes coloridos são os marcadores utilizados para cada laranja, o fundo roxo é o marcador do *background* e as áreas em preto são as regiões que não possuem nenhum marcador.

As imagens no canto inferior direito das Figuras 8, 9 e 10 mostram as segmentações feitas pelo *watershed*. As laranjas segmentadas estão indicadas pelos pixels coloridos enquanto a segmentação do *background* está indicada pelos pixels cinzas. Além do *watershed* segmentar cada laranja como pertencente à classe laranja, ele também foi capaz de segmentá-las como sendo entidades diferentes, segmentação de instância, dessa forma é possível diferenciar uma laranja de outra.



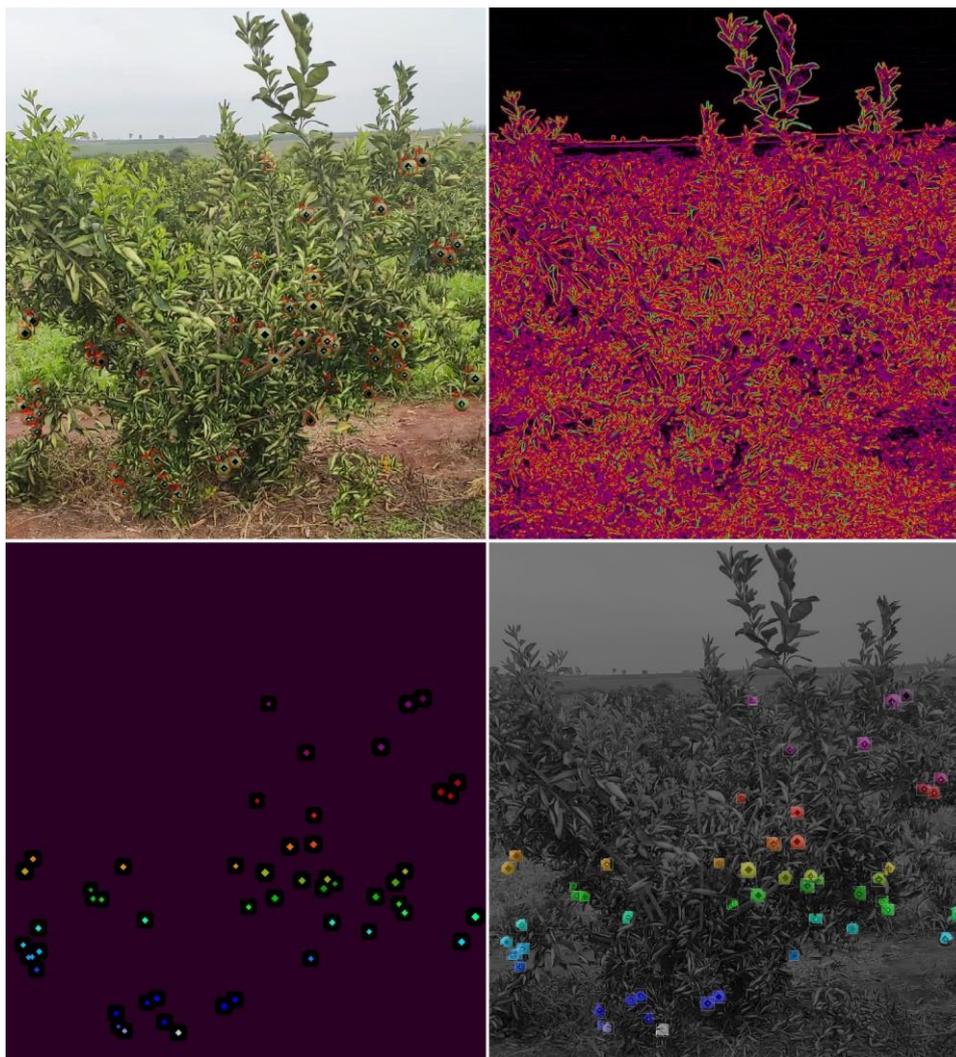
**Figura 8:** Foto 1; foto original com caixas delimitadoras e marcadores por laranja (canto superior esquerdo), gradiente da foto original (canto superior direito), marcadores para o algoritmo *watershed* (canto inferior esquerdo), foto com segmentação feita pelo algoritmo *watershed* (canto inferior direito)

Fonte: Dados da pesquisa



**Figura 9:** Foto 2; foto original com caixas delimitadoras e marcadores por laranja (canto superior esquerdo), gradiente da foto original (canto superior direito), marcadores para o algoritmo *watershed* (canto inferior esquerdo), foto com segmentação feita pelo algoritmo *watershed* (canto inferior direito)

Fonte: Dados da pesquisa



**Figura 10:** Foto 3; foto original com caixas delimitadoras e marcadores por laranja (canto superior esquerdo), gradiente da foto original (canto superior direito), marcadores para o algoritmo *watershed* (canto inferior esquerdo), foto com segmentação feita pelo algoritmo *watershed* (canto inferior direito)

**Fonte:** Dados da pesquisa

A Tabela 1 mostra o número de laranjas em cada foto com a respectiva confiança de classe fornecida pela rede *YoLov4*. Uma confiança de classe alta indica que a rede está convencida de que há uma laranja devidamente contida naquela caixa delimitadora, enquanto que uma confiança de classe mais baixa indica que a rede não está muito certa de que há uma laranja naquela caixa delimitadora e nem que o fruto está devidamente contido nela.

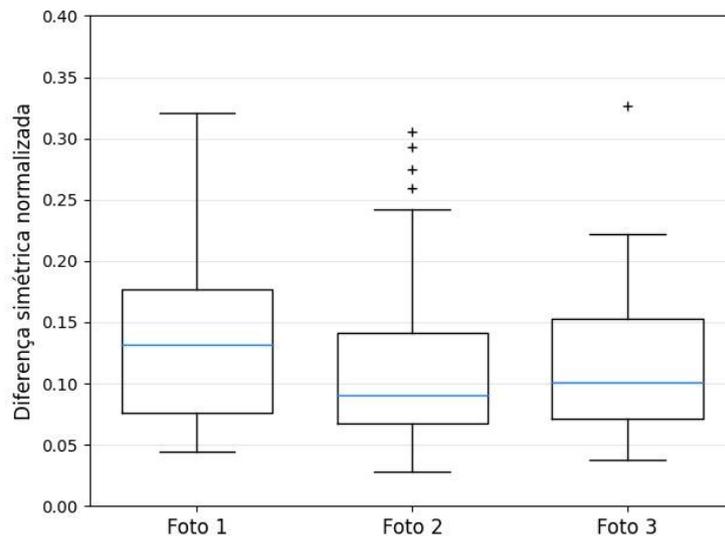
**Tabela 1:** Número de laranjas por categoria de confiança de classe por foto

Número de laranjas nessa categoria de confiança de classe			Confiança de classe
Foto 1	Foto 2	Foto 3	
52	50	41	100%
22	16	8	Entre 99% e 90%
5	3	3	Menor que 90%

**Fonte:** Dados da pesquisa

A diferença simétrica normalizada das laranjas por foto foi ilustrada no diagrama de caixa na Figura 11. Em que os riscos azuis são as medianas; os limites superior e inferior das caixas representam o primeiro e o terceiro quartil, respectivamente; os limites superior

e inferior do segmento de reta vertical representam os limites superior e inferior do diagrama de caixa; e os símbolos acima dos limites superiores do segmento de reta representam os valores discrepantes. Imagens que tiverem a maioria de suas laranjas com boas segmentações terão riscos azuis e limites superiores das caixas posicionados mais abaixo no diagrama, enquanto que imagens que tiverem a maioria de suas laranjas com segmentações ruins terão riscos azuis e limites superiores das caixas posicionados mais acima no diagrama.



**Figura 11:** Diagrama de caixa da diferença simétrica normalizada por foto  
**Fonte:** Dados da pesquisa

Laranjas que, em geral, não foram bem segmentadas pelo algoritmo *watershed* e obtiveram um valor de diferença simétrica normalizada superior ao limite imposto receberam uma marcação no centro delas, como é ilustrado nas 4 laranjas a esquerda na Figura 12. Por outro lado, laranjas que, em geral, tiveram boas segmentações pelo algoritmo *watershed* e assim obtiveram um valor de diferença simétrica normalizada baixa não receberam nenhuma marcação, como é ilustrado nas 4 laranjas a direita na Figura 12.



**Figura 12:** 4 laranjas com diferença simétrica normalizada superior ao limite marcadas a esquerda e 4 laranjas com diferença simétrica normalizada inferior ao limite a direita  
**Fonte:** Dados da pesquisa

É possível observar que laranjas maiores, ou que foram fotografadas de uma distancia menor, em geral apresentaram uma segmentação melhor, enquanto que laranjas menores, ou que foram fotografadas de uma distância maior, em geral apresentaram uma segmentação pior. Dessa maneira, uma das possíveis formas de obter melhores resultados, seria fotografar as laranjas de uma distância menor.

Laranjas que tinham objetos em suas frentes, como por exemplo um graveto ou uma folha, tiveram esses objetos ou partes deles segmentados juntos com as laranjas também trazendo erro à segmentação.

Outro fator que contribuiu para uma segmentação pior foi a iluminação. Laranjas que, em geral, estavam posicionadas de forma a receber mais iluminação natural tiveram uma segmentação melhor, enquanto que as laranjas que estavam mais obstruídas por folhagens, ou posicionadas na parte inferior da árvore, e dessa forma recebendo menos iluminação natural, tenderam a ter uma segmentação pior. Uma forma de contornar esses obstáculos seria utilizar alguma fonte de iluminação artificial no momento da aquisição da imagem, como por exemplo um flash externo.

Como o algoritmo *watershed* é utilizado diretamente no gradiente da imagem e não na imagem original. Outro possível caminho para obter resultados melhores, seria utilizar outra técnica para calcular o gradiente das imagens.

## 4 CONCLUSÃO

Este estudo tratou dos problemas de detecção e segmentação de laranjas a partir de imagens da plantação. Para o fluxo de processamento das imagens foi proposta uma solução integrando a rede *Yolov4* e o algoritmo *watershed*, a qual forneceu resultados promissores. A rede *YoLov4* demonstrou competência em detectar laranjas verdes com uma cor que se assemelhava muito com as folhagens aos seus arredores, enquanto o algoritmo *watershed* se beneficiou das informações fornecidas pela *YoLov4* para posicionar os marcadores centralizados nas laranjas e realizar a segmentação apenas no interior das caixas delimitadoras. Para a quantificação dos erros de estimação foi proposto o emprego de um operador de diferenças simétricas entre o modelo estimado da laranja e o resultado da segmentação. Os resultados encontrados mostram que laranjas maiores, que estejam posicionadas na parte superior da árvore ou que não estejam obstruídas por folhagens ou por outras laranjas são melhor segmentadas por conta da maior exposição a iluminação natural no momento da captura da imagem.

## AGRADECIMENTOS

Os autores gostariam de agradecer a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro.

## REFERÊNCIAS

ABDULLAHI, Halimatu Sadiyah; SHERIFF, Ray E.; MAHIEDDINE, Fatima. Convolution neural network in precision agriculture for plant image recognition and classification 5G Mobile Wireless Network and Technology View project Automation on farm field operations for yield optimization View project Convolution Neural Network in. **researchgate.net**, [S. l.], 2017. DOI: 10.1109/INTECH.2017.8102436. Disponível em: <https://www.researchgate.net/publication/320971643>. Acesso em: 22 mar. 2021.

**blue berries on brown stem photo – Free Nature Image on Unsplash**. 2021. Disponível em: <https://unsplash.com/photos/jzmZeQyEgiU>. Acesso em: 22 mar. 2021.

BOCHKOVSKIY, Alexey; WANG, Chien Yao; LIAO, Hong Yuan Mark. **YOLOv4: Optimal Speed and Accuracy of Object Detection** XivarXiv, , 2020. Disponível em: <https://github.com/AlexeyAB/darknet>. Acesso em: 23 mar. 2021.

BRESILLA, Kushtrim; PERULLI, Giulio Demetrio; BOINI, Alexandra; MORANDI, Brunella; CORELLI GRAPPADELLI, Luca; MANFRINI, Luigi. Single-shot convolution neural networks for real-time fruit detection within the tree. **Frontiers in Plant Science**, [S. l.], v. 10, 2019. DOI: 10.3389/fpls.2019.00611.

CHEN, LC; HERMANS, A.; RECOGNITION, G. Papandreou-....; 2018, Undefined. Masklab: Instance segmentation by refining object detection with semantic and direction features. **openaccess.thecvf.com**, [S. l.], 2018. Disponível em: [http://openaccess.thecvf.com/content\\_cvpr\\_2018/html/Chen\\_MaskLab\\_Instance\\_Segmentation\\_CVPR\\_2018\\_paper.html](http://openaccess.thecvf.com/content_cvpr_2018/html/Chen_MaskLab_Instance_Segmentation_CVPR_2018_paper.html). Acesso em: 7 ago. 2021.

**Colaboratory – Google**. 2017. Disponível em: <https://research.google.com/colaboratory/faq.html>. Acesso em: 2 ago. 2021.

**Computer Vision Annotation Tool**. 2018. Disponível em: <https://cvat.org/tasks>. Acesso em: 2 ago. 2021.

GARCIA-GARCIA, A.; ORTS-ESCOLANO, S.; ... S. Oprea-Applied Soft; 2018, Undefined. A survey on deep learning techniques for image and video semantic segmentation. **Elsevier**, [S. l.], 2018. DOI: 10.1016/j.asoc.2018.05.018. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1568494618302813>. Acesso em: 7 ago. 2021.

GE, Y.; XIONG, Y.; TENORIO, GL; ACCESS, PJ From-IEEE; 2019, Undefined. Fruit localization and environment perception for strawberry harvesting robots. **ieeexplore.ieee.org**, [S. l.], 2019. Disponível em: <https://ieeexplore.ieee.org/abstract/document/8863343/>. Acesso em: 23 mar. 2021.

GOLLAPUDI, S. Learn computer vision using OpenCV. [S. l.], 2019. Disponível em: <https://link.springer.com/content/pdf/10.1007/978-1-4842-4261-2.pdf>. Acesso em: 22 mar. 2021.

GONZALEZ, Rafael C.; WOODS, Richard E. Digital Image Processing (3rd Edition). **Prentice-Hall, Inc. Upper Saddle River, NJ, USA ©2006**, [S. l.], p. 976, 2007. Disponível em: <http://dl.acm.org/citation.cfm?id=1076432>. Acesso em: 2 ago. 2021.

GREMES, Matheus Felipe; KRUMMENAUER, Rafael; ANDRADE, Cid Marcos Gonçalves; LIMA, AndOswaldo Curty da Motta. Pre-Harvest Fruit Image Processing: A Brief Review. **Brazilian Journal of Experimental Design, Data Analysis and Inferential Statistics**, [S. l.], v. 1, n. 2, 2021. DOI: 10.29327/232092.1.2-13. Disponível em: <https://bjedis.org/bjedis%2312-n2>. Acesso em: 2 ago. 2021.

GUO, Yanming; LIU, Yu; GEORGIU, Theodoros; LEW, Michael S. A review of semantic segmentation using deep neural networks. **International Journal of Multimedia Information Retrieval**, [S. l.], v. 7, n. 2, p. 87–93, 2018. DOI: 10.1007/s13735-017-0141-z.

HAFIZ, Abdul Mueed; BHAT, Ghulam Mohiuddin. A survey on instance segmentation: state of the art. **International Journal of Multimedia Information Retrieval**, [S. l.], v. 9, n. 3, p. 171–189, 2020. DOI: 10.1007/s13735-020-00195-x.

HALMOS, PR. Naive set theory. [S. l.], 2017. Disponível em:

[https://books.google.com.br/books?hl=pt-BR&lr=&id=bWu9DgAAQBAJ&oi=fnd&pg=PA1&dq=Naive+set+theory.&ots=7FGGIAg77f&sig=AOM4vS72d9qpjC\\_tTjbD7Xle35U](https://books.google.com.br/books?hl=pt-BR&lr=&id=bWu9DgAAQBAJ&oi=fnd&pg=PA1&dq=Naive+set+theory.&ots=7FGGIAg77f&sig=AOM4vS72d9qpjC_tTjbD7Xle35U). Acesso em: 7 ago. 2021.

HE, Kaiming; GKIOXARI, Georgia; DOLLÁR, Piotr; GIRSHICK, Ross. **Mask R-CNNopenaccess.thecvf.com**. [s.l.: s.n.]. Disponível em: [http://openaccess.thecvf.com/content\\_iccv\\_2017/html/He\\_Mask\\_R-CNN\\_ICCV\\_2017\\_paper.html](http://openaccess.thecvf.com/content_iccv_2017/html/He_Mask_R-CNN_ICCV_2017_paper.html). Acesso em: 23 mar. 2021.

KANG, Hanwen; CHEN, Chao. **Real-Time Fruit Recognition and Grasping Estimation for Autonomous Apple Harvesting**[arxiv.org](https://arxiv.org/abs/2003.13298). [s.l.: s.n.]. Disponível em: <https://arxiv.org/abs/2003.13298>. Acesso em: 23 mar. 2021.

KOIRALA, A.; WALSH, K. B.; WANG, Z.; MCCARTHY, C. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO'. **Precision Agriculture**, [S. l.], v. 20, n. 6, p. 1107–1135, 2019. DOI: 10.1007/s11119-019-09642-0.

OLIVEIRA, Roberto; SCIVITTARO, Walkyria. FOLHA murcha: laranja tipo Valência de maturação extremamente tardia e tolerante ao cancro cítrico. [S. l.], 2019. Disponível em: <http://www.infoteca.cnptia.embrapa.br/handle/doc/746757>. Acesso em: 7 ago. 2021.

**orange fruit photo – Free Food Image on Unsplash**. 2021. Disponível em: <https://unsplash.com/photos/U1iYwZ8Dx7k>. Acesso em: 22 mar. 2021.

REDMON, Joseph; DIVVALA, Santosh; GIRSHICK, Ross; FARHADI, Ali. **You Only Look Once: Unified, Real-Time Object Detection**[cv-foundation.org](https://pjreddie.com/yolo/). [s.l.: s.n.]. Disponível em: <http://pjreddie.com/yolo/>. Acesso em: 23 mar. 2021.

ROBERTS, D.; WANG, M.; TORRES CALDERON, W.; GOLPARVAR-FARD, M. An Annotation Tool for Benchmarking Methods for Automated Construction Worker Pose Estimation and Activity Analysis. **International Conference on Smart Infrastructure and Construction 2019, ICSIC 2019: Driving Data-Informed Decision-Making**, [S. l.], p. 307–313, 2019. DOI: 10.1680/ICSIC.64669.307. Disponível em: <https://www.icevirtuallibrary.com/doi/abs/10.1680/icsic.64669.307>. Acesso em: 7 ago. 2021.

VINCENT, L.; MACHINE, P. Soille. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. **cse.msu.edu**, [S. l.], v. 13, n. 6, p. 583, 1991. Disponível em: <http://cse.msu.edu/~cse902/S03/watershed.pdf>. Acesso em: 3 ago. 2021.

XIONG, J.; LIU, Z.; CHEN, S.; LIU, B.; ZHENG, Z.; ... Z. Zhong-Biosystems; 2020, Undefined. Visual detection of green mangoes by an unmanned aerial vehicle in orchards based on a deep learning method. **Elsevier**, [S. l.], 2020. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1537511020300970>. Acesso em: 23 mar. 2021.

YU, Y.; ZHANG, K.; YANG, L.; AGRICULTURE, D. Zhang-Computers and Electronics in; 2019, Undefined. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN. **Elsevier**, [S. l.], 2019. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0168169919301103>. Acesso em: 22 mar. 2021.