

CAROLINA TAVARES DE OLIVEIRA

**MONITORAMENTO DA MATURAÇÃO DOS FRUTOS E DE DOENÇAS DO
CAFFEEIRO UTILIZANDO MODELOS DE DEEP LEARNING**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Agrícola, para obtenção do título de *Magister Scientiae*.

Orientador: Domingos Sarvio Magalhães
Valente

**VIÇOSA - MINAS GERAIS
2022**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

O48m
2021

Oliveira, Carolina Tavares de, 1990-
Monitoramento da maturação dos frutos e de doenças do
cafeeiro utilizando modelos de deep learning / Carolina Tavares
de Oliveira. – Viçosa, MG, 2021.

1 dissertação eletrônica (49 f.): il. (algumas color.).

Orientador: Domingos Sarvio Magalhaes Valente.
Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Engenharia Agrícola, 2021.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2022.136>

Modo de acesso: Worl Wide Web.

1. Café - Doenças e pragas. 2. Redes Neurais
(Computação). 3. Processamento de imagens. 4. Aprendizado
profundo (Aprendizado do computador). 5. Visão por
computador. I. Valente, Domingos Sarvio Magalhaes.
II. Universidade Federal de Viçosa. Departamento de Engenharia
Agrícola. Programa de Pós-Graduação em Engenharia Agrícola.
III. Título.

CDD 22. ed. 633.739

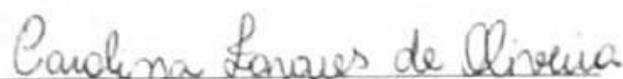
CAROLINA TAVARES DE OLIVEIRA

MONITORAMENTO DA MATURAÇÃO DOS FRUTOS E DE DOENÇAS DO
CAFFEEIRO UTILIZANDO MODELOS DE DEEP LEARNING

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Agrícola, para obtenção do título de *Magister Scientiae*.

APROVADA: 20 de dezembro de 2022.

Assentimento:



Carolina Tavares de Oliveira
Autor



Domingos Sarvio Magalhães Valente
Orientador

AGRADECIMENTOS

A Deus.

Aos meus pais.

À Universidade Federal de Viçosa, pela oportunidade de realizar a pós-graduação.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudos.

RESUMO

OLIVEIRA, Carolina Tavares de Oliveira, M.Sc., Universidade Federal de Viçosa, agosto de 2021. **Monitoramento da maturação dos frutos e de doenças do cafeeiro utilizando modelos de deep learning**. Orientador: Domingos Sarvio Magalhães Valente.

O valor da safra do café está relacionado à vários fatores entre eles: a oferta e a demanda, a quantidade produzida, o armazenamento, a qualidade dos frutos, dentre outros. A qualidade do café, por sua vez, é afetada por vários fatores, entre eles radiação solar, nutrição das plantas, altitude, presença ou ausência de pragas e doenças. Dentre as principais doenças que infestam a cultura do café estão: ferrugem, cercosporiose e mancha de phoma. A ferrugem (*Hemileia vastatrix* Berk. & Br.) ataca as folhas do cafeeiro o que gera queda prematura das mesmas e redução da fotossíntese. De forma semelhante, a cercosporiose causa danos, principalmente, em lavouras sem fertilização adequada, atingindo folhas e frutos. Já mancha de phoma é favorecida por diferentes eventos climáticos como: ventos fortes e frios, granizo e geada. Além das doenças, outro fator determinante que impacta diretamente na qualidade da bebida é o grau de maturação dos frutos no momento da colheita. Sabe-se que frutos maduros, no estágio denominado cereja, geram cafés de qualidade superior. Normalmente, a avaliação do grau de maturação é realizada por métodos destrutivos, com colheitas de algumas plantas no talhão, que podem ou não representar o talhão a ser colhido. Neste contexto, tecnologias computacionais, tais como: técnicas de inteligência artificial podem ser úteis para monitoramento do cafeeiro. A inteligência artificial tem tido um papel importante no desenvolvimento da agricultura, por meio dela é possível estimar a produtividade das safras, identificar pragas e doenças nas lavouras, definir de forma mais assertiva o momento adequado para realizar a colheita. Dessa forma, esse trabalho teve como objetivos: (1) Desenvolver modelo de detecção, classificação e segmentação de frutos a partir de imagens obtidas em ambientes não controlados (2) Desenvolver um classificador para classificar imagens entre: presença e ausência da doença no cafeeiro, e desenvolver um classificador para identificar as três espécies de doenças comuns que atacam o cafeeiro, a saber, cercosporiose, mancha de phoma e ferrugem. Para atender o primeiro objetivo foram coletadas 78 imagens com frutos em diferentes graus de maturação. As imagens foram rotuladas para identificação dos locais com frutos

verde, cereja e passa e fundo. A rotulagem foi realizada de forma manual/visual com auxílio de uma ferramenta desenvolvida por meio da linguagem de programação Python. A partir dos rótulos foi possível treinar um modelo de segmentação de instâncias com arquitetura Mask-RCNN. Para atender o segundo foram coletadas imagens de plantas contaminadas pelas doenças do cafeeiro, a saber, cercospora, mancha de phoma e ferrugem e imagens de plantas sem contaminação na região da Zona da Mata mineira. As mesmas imagens foram utilizadas para classificar somente as três doenças do cafeeiro. Em ambas as análises as imagens serviram de entradas para o treinamento dos modelos de classificação utilizando redes neurais convolucionais. Os resultados foram avaliados pelas métricas da matriz de confusão, a saber, a precisão, recall e f1-score. A precisão do modelo é a relação entre verdadeiros positivos (detecções corretas) sobre a soma de todas as detecções. As abordagens mostraram que o modelo de segmentação de frutos alcançou precisões de 0,897; 0,900; 0,891 para as classes: cereja, verde e passa, respectivamente. Os valores de recall para as mesmas classes respectivamente foram: 0,759; 0,700; 0,813, respectivamente. Já os valores de f1-score para as mesmas classes foram: 0,7336; 0,6802; 0,7692. Desta forma, o modelo foi mais eficiente na detecção, segmentação de classificação de frutos passas. Quando foi aplicado o mesmo modelo nas imagens geradas por janela deslizante os valores de precisão foram: 0,974; 0,906; 0,878; recall foram: 0,753; 0,740; 0,813; e f1-score foram: 0,844; 0,8105; 0,8427 para as classes: cereja, verde e passa, respectivamente. No modelo de classificação de doenças as precisões, recall e f1-score foram: 0,933 para ambas as classes: presença e ausência de doença. Já o modelo que classificou as três espécies de doenças apresentou valores de precisão: 0,900, 0,850 e 0,900, recall de 0,900, 0,850 e 0,900 e f1-score de 0,900, 0,850 e 0,900 para as classes ferrugem, cercospora e mancha de phoma, respectivamente.

Palavras-chave: Processamento de Imagens. Segmentação de instâncias. Deep Learning. Colheita do café. Inteligência artificial. Maturação do café.

ABSTRACT

OLIVEIRA, Carolina Tavares de Oliveira, M.Sc., Universidade Federal de Viçosa, August, 2021. **Monitoring of fruit maturation and coffee diseases using deep learning models**. Orientador: Domingos Sarvio Magalhães Valente.

The value of the coffee crop is related to several factors, including supply and demand, the quantity produced, storage, fruit quality, among others. Coffee quality, in turn, is affected by several factors, including solar radiation, plant nutrition, altitude, presence or absence of pests and diseases. Among the main diseases that infest the coffee crop are: rust, brown eye spot and phoma spot. Rust (*Hemileia vastatrix* Berk. & Br.) attacks coffee leaves, which causes premature leaf fall and reduced photosynthesis. Similarly, brown eye spot causes damage, mainly in crops without adequate fertilization, affecting leaves and fruits. Phoma stain is favored by different climatic events such as strong and cold winds, hail and frost. In addition to diseases, another determining factor that directly impacts the quality of the drink is the degree of maturation of the fruits at the time of harvest. It is known that ripe fruits, in the stage called cherry, generate coffees of superior quality. Normally, the assessment of the degree of maturation is carried out by destructive methods, with harvesting of some plants in the field, which may or may not represent the field to be harvested. In this context, computational technologies such as artificial intelligence techniques can be useful for monitoring the coffee tree. Artificial intelligence has played an important role in the development of agriculture, through which it is possible to estimate the productivity of crops, identify pests and diseases in crops, define in a more assertive way the appropriate time to carry out the harvest. Thus, this work aimed to: (1) Develop a model of detection, classification and segmentation of fruits from images obtained in uncontrolled environments (2) Develop a classifier to classify images between: presence and absence of the disease in coffee, and to develop a classifier to identify the three species of common diseases that attack the coffee tree, namely, brown eye spot, phoma spot and rust. To meet the first objective, 78 images were collected with fruits in different degrees of maturation. The images were labeled to identify the places with green, cherry and raisin fruits and background. Labeling was performed manually/visually with the aid of a tool developed using the Python programming language. From the labels it was possible to train an instance segmentation model with

Mask-RCNN architecture. To meet the second, images of plants contaminated by coffee diseases were collected, namely, cercospora, phoma stain and rust and images of plants without contamination in the Zona da Mata region of Minas Gerais. The same images were used to classify only the three coffee diseases. In both analyzes the images served as inputs for training the classification models using convolutional neural networks. The results were evaluated by the metrics of the confusion matrix, namely, precision, recall and f1-score. Model accuracy is the ratio of true positives (correct detections) over the sum of all detections. The approaches showed that the fruit segmentation model reached accuracies of 0.897; 0.900; 0.891 for the classes: cherry, green and raisin, respectively. The recall values for the same classes respectively were: 0.759; 0.700; 0.813, respectively. The f1-score values for the same classes were: 0.7336; 0.6802; 0.7692. In this way, the model was more efficient in the detection, segmentation and classification of raisin fruits. When the same model was applied to the images generated by the sliding window, the precision values were: 0.974; 0.906; 0.878; recall were: 0.753; 0.740; 0.813; and f1-score were: 0.844; 0.8105; 0.8427 for the classes: cherry, green and raisin, respectively. In the disease classification model, the precisions, recall and f1-score were: 0.933 for both classes: presence and absence of disease. The model that classified the three disease species, on the other hand, presented precision values: 0.900, 0.850 and 0.900, recall of 0.900, 0.850 and 0.900 and f1-score of 0.900, 0.850 and 0.900 for the rust, cercospora and phoma stain classes, respectively.

Keywords: Image Processing. Instance Segmentation. Deep Learning. Coffee Harvest. Artificial Intelligence. Coffee Maturation

SUMÁRIO

1. INTRODUÇÃO	12
2. CAPÍTULO 1	15
2.1. Introdução	17
2.2. Material e Métodos.....	19
2.2.1. Aquisição de imagens.....	20
2.2.2. Pré-processamento.....	20
2.2.3. Treinamento do modelo de segmentação de instâncias Mask-RCNN..	22
2.2.4. Avaliação da performance do modelo.....	25
2.2.5. Técnica de janelas deslizantes	25
2.3. Resultados e discussão	26
2.4. Conclusão	34
2.5. Referências	34
3. CAPÍTULO 2	37
3.1. Introdução	39
3.2. Material e Métodos.....	41
3.3. Resultados e Discussão.....	44
3.4. Conclusão	48
3.5. Referências	48
4. CONCLUSÃO GERAL.....	49

1. INTRODUÇÃO

O café se destaca entre as principais commodities comercializadas ao redor do mundo. Além disso, está entre as principais culturas do agronegócio brasileiro, visto que, além de ser um dos produtos mais importantes na pauta de exportações brasileiras, é a principal fonte de renda de milhares de pequenos agricultores (Maguire-Rajput et al., 2018; Hajjar et al., 2019). Apesar de sua importância econômica e social, produtores convivem com vários problemas que afetam a rentabilidade das lavouras. Tais problemas, estão relacionados a fatores climáticos, condições do solo, nutrição de plantas, relevo, presença de pragas dentre outros.

As pragas geram todos os anos prejuízos de safra para os agricultores devido a perda de produtividade no campo que causam (Savary & Willocquet, 2020). Nas condições da cafeicultura brasileira, a doença mais grave é a ferrugem, causada pelo fungo (*Hemileia vastatrix* Berk. Br.), em seguida tem-se a cercosporiose (*Cercospora coffeicola* Berk & Cook) e a mancha de phoma (*Phoma* spp.). As condições ambientais e de manejo dos cafezais favorecem a ocorrência das doenças, de forma que lavouras sombreadas e adensadas facilitam a infestação por ferrugem. Já lavouras abertas, excesso de insolação, em regiões mais quentes com alta umidade possibilita o desenvolvimento da cercosporiose enquanto áreas frias e úmidas, sujeitas a ventos frios favorece a infestação da mancha de phoma (Oliveira, et.al 2012). Os fungos da espécie *H. vastatrix* enfraquecem o cafeeiro tornando-o mais suscetível a outras doenças. Além disso, causam quedas precoces das folhas e redução de produtividade de safras futuras, além de reduzir a vida útil da planta (Zambolim et al., 1997). Já o ataque de *Phoma* spp. causa na planta, desfolha, seca de ramos e morte de brotos novos, comprometendo o desenvolvimento e a futura produção. Do mesmo modo, o fungo *C. coffeicola* ataca folhas e frutos, gerando prejuízos, principalmente em lavouras jovens plantadas no final do período chuvoso ou no início da seca. Tal fungo, promove a maturação e queda antecipada dos frutos. Os frutos que caem ao permanecerem em contato com o solo, resultam em cafés de qualidade inferior (Pozza, 2008). Além disso, geram o secamento da mucilagem, dificultam o descascamento e promovem fermentações indesejáveis.

Além das pragas, um dos fatores que mais tem influenciado a qualidade do café é o ponto de colheita (Fagan et al., 2011). A colheita do café em momentos inadequados prejudica a qualidade sensorial das bebidas, e gera desuniformidade de

cor dos grãos obtida após seu beneficiamento. De acordo Fagan et al. (2011), o café deve ser colhido em seu ponto ótimo de maturação, ao passo que quando colhido verde ou seco na planta pode proporcionar grande incidência de grãos verdes, ardidos e pretos que são considerados defeitos que afetam a qualidade da bebida. Os frutos colhidos verdes apresentam níveis inferiores de compostos fenólicos resultando em pior qualidade de bebida (Sittipod et al., 2019). Já a sacarose presente nos frutos maduros resulta em bebidas de qualidade superior.

Diante dos problemas causados pelas doenças e a necessidade de se colher o café no momento certo, o produtor precisa estar constantemente monitorando as condições da cultura. Geralmente esse monitoramento é feito de forma visual o que pode ser impreciso, uma vez que é um processo demorado, o que faz com que seja feito por amostragem, e a amostragem nem sempre é conduzida adequadamente. Alternativamente ao monitoramento visual, pode-se utilizar o monitoramento por imagens. As imagens podem ser capturadas por câmeras operadas manualmente ou obtidas por câmeras acopladas em tratores, implementos, robôs ou vants (UAV). Tais imagens, podem ser usadas em algoritmos para automatizar o processo de monitoramento, facilitando esse trabalho e trazendo maior exatidão para os diagnósticos.

Dentre as técnicas que podem ser utilizadas para o desenvolvimento desses algoritmos está o aprendizado de máquinas, uma subárea da inteligência artificial. O aprendizado de máquina envolve o desenvolvimento de modelos que geram previsões guiadas por dados. Tais algoritmos, aprendem visando reduzir o erro na previsão da variável de interesse à cada iteração e assim podem identificar padrões. A evolução dos hardwares e aumento do poder de processamento computacional vem fazendo com que o aprendizado de máquinas ganhe notoriedade em aplicações no mundo real (Skansi, 2018). Os algoritmos de aprendizado de máquina buscam aprender e melhorar seu desempenho com base nos dados históricos, e geram modelos matemáticos com base em dados amostrais (Mitchel, 1997). Os tipos de aprendizado existentes são: aprendizado supervisionado, não supervisionado e aprendizado por reforço.

Em trabalhos que envolvem a predição de uma variável de interesse predomina o uso de técnicas de aprendizado supervisionado. Nesta classe de algoritmos é necessário obter uma base de dados, que contém a variável alvo, ou seja, a variável dependente resultante das variáveis independentes observadas.

Nesse caso, diz-se que os dados são anotados com as respostas ou classes a serem previstas (Skiena, 2017). Para criação de modelos por meio destes algoritmos, o conjunto de dados deve ser dividido, usualmente, em três conjuntos: treinamento, validação e teste. O conjunto de dados de treinamento são utilizados para o aprendizado do modelo. O modelo é inicialmente ajustado em um conjunto de dados de treinamento. O modelo ajustado é usado para prever as respostas para as observações em um segundo conjunto de dados, a saber, conjunto de dados de validação. Tal conjunto, por sua vez, fornece uma avaliação imparcial do modelo no treinamento enquanto ajusta os hiperparâmetros do modelo. Já o conjunto de teste é usado para comprovar que o modelo ajustado realmente funciona. Os dados de teste são ignorados no processo de treinamento e no processo de escolha de hiperparâmetros.

Diante do exposto, acredita-se que o uso de imagens obtidas por câmeras digitais como entradas para modelos de aprendizado de máquina pode ser usado para quantificar e monitorar a maturação do café, a fim de determinar o ponto de colheita dos frutos. Além disso, tais técnicas podem possibilitar a identificação de pragas auxiliando nas tomadas de decisões dos agricultores. Dessa forma, esse trabalho foi realizado com nos seguintes objetivos: (1) Desenvolver um método para realizar a contagem de frutos do café em diferentes estágios de maturação; (2) Desenvolver um classificador para identificar quando há ou não a presença da doença na imagem e desenvolver um classificador para identificar as três espécies de doenças comuns que atacam o cafeeiro, a saber, ferrugem, cercosporiose, mancha de phoma.

A dissertação foi dividida em quatro capítulos. Nesse primeiro capítulo foi realizada uma introdução geral ao tema. No segundo capítulo foram relatadas as abordagens usadas na geração de uma ferramenta automática de rotulagem de frutos nos três diferentes estágios de maturação, a saber: cereja, verde e passa. Com base nos rótulos detectou-se frutos do café em diferentes estágios de maturação por meio de técnica de segmentação de instâncias envolvendo a arquitetura Mask-RCNN. Por fim, no terceiro capítulo foram desenvolvidos algoritmos de classificação supervisionada do café por meio de redes neurais convolucionais. Pretendeu-se com o estudo verificar a performance dos modelos em diferentes tarefas de monitoramento do cafeeiro, abordagens assim podem automatizar tarefas manuais e repetitivas das lavouras. E no quarto e último capítulo foram apresentadas as conclusões gerais do trabalho.

2. CAPÍTULO 1

SEGMENTAÇÃO DE FRUTOS DO CAFEIEIRO UTILIZANDO MASK-RCNN PARA ESTIMATIVA DA MATURAÇÃO DO CAFÉ

RESUMO

Um dos fatores associados à rentabilidade da lavoura é o grau de maturação dos frutos no momento da colheita. O agricultor precisa realizar a colheita quando o maior percentual de frutos atinge o estágio de maduros. Tal fato, é necessário pois a maturação afeta a qualidade do produto e esta, por sua vez, define o preço de comercialização das sacas. Todavia, estimar corretamente o ponto ideal da colheita não é uma tarefa trivial. Geralmente isso é feito com base nas experiências passadas do agricultor ou realizada por meio de amostragem destrutiva que nem sempre pode representar todo o talhão, assim, tais métodos podem gerar imprecisões nas estimativas. Sendo assim, as tecnologias de monitoramento automatizado das lavouras podem ser viáveis para a determinação do momento da colheita. A utilização de sistemas robóticos, em conjunto com tecnologias de visão computacional e algoritmos de inteligência artificial envolvendo modelos baseados em redes neurais, têm fornecido novos métodos para reconhecer e classificar imagens. Dessa forma, esse trabalho teve como objetivo desenvolver um para estimar o número de frutos nos diferentes estágios de maturação em imagens coletadas da lavoura. O modelo foi gerado por meio de deep learning com arquitetura de redes neurais do tipo Mask-RCNN. As imagens foram rotuladas de forma manual/visual com auxílio de um algoritmo desenvolvido na linguagem de programação python versão 3.8.6. Verificou-se que a Mask RCNN com uso de transferência de aprendizado foi eficiente na detecção e contagem de frutos de café nos diferentes estágios de maturação. Os valores de precisão foram 0,897; 0,900; 0,891 para os estágios: cereja, verde e passa, respectivamente. Os valores de recall foram: 0,759; 0,700; 0,813. Os valores de recall foram menores que os de precisão devido à falta de detecção de muitos frutos. Para a técnica de janela deslizante os valores de precisão foram: 0,974; 0,906; 0,878; e recall foram: 0,753; 0,74; 0,813 para os estágios: cereja, verde e passa, respectivamente. O método proposto permite estimar a maturação do café, o que pode auxiliar os produtores na tomada de decisão.

Palavras-chave: Cafeicultura, Inteligência Artificial, Maturação do café

SEGMENTATION OF COFFEE FRUITS USING MASK-RCNN TO ESTIMATE COFFEE MATURATION

ABSTRACT

One of the factors associated with the profitability of the crop is the degree of fruit maturation at the time of harvest. The farmer needs to carry out the harvest when the highest percentage of fruits reaches the ripe stage. This fact is necessary because maturation affects the quality of the product and this, in turn, defines the marketing price of the bags. However, correctly estimating the ideal harvest point is not a trivial task. This is usually done based on the farmer's past experiences or carried out through destructive sampling that cannot always represent the entire field, thus such methods can lead to inaccuracies in estimates. Therefore, automated crop monitoring technologies can be viable for determining the time of harvest. The use of robotic systems, together with computer vision technologies and artificial intelligence algorithms involving models based on neural networks, have provided new methods to recognize and classify images. Thus, this work aimed to develop a model based on to estimate the number of fruits at different stages of maturation in images collected from the field. The model was generated through deep learning with Mask-RCNN neural network architecture. The images were labeled manually/visually with the aid of an algorithm developed in the python programming language version 3.8.6. It was verified that the Mask RCNN with the use of transfer of learning was efficient in the detection and counting of coffee fruits in the different stages of maturation. Precision values were 0.897; 0.900; 0.891 for the stages: cherry, green and raisin, respectively. The recall values were: 0.759; 0.700; 0.813. Recall values were lower than precision values due to the lack of detection of many fruits. For the sliding window technique, the precision values were: 0.974; 0.906; 0.878; and recall were: 0.753; 0.74; 0.813 for the stages: cherry, green and raisin, respectively. The proposed method allows estimating coffee maturation, which can help producers in decision making.

Keywords: Coffee growing, Artificial Intelligence, Coffee maturation

2.1. Introdução

O café é uma das commodities que mais se destaca em transações internacionais em termos de quantidade e valor de sacas comercializadas. A produção mundial para a safra de 2020/2021 atingiu 169,64 milhões de sacas de 60 kg. O Brasil é o maior produtor e exportador de café e a produção de café arábica e conilon atingiu 45,6 milhões de sacas na safra 2020/2021 (International Coffee Organization, 2022).

Apesar da importância mundial do Brasil, o nível tecnológico empregado na cafeicultura é baixo em comparação às outras culturas como soja e milho (Ramos et al, 2017). Os produtores não dispõem de ferramentas eficientes para, por exemplo, antecipar informações sobre a estimativa de produtividade, determinação de maturação, e controle de pragas. Normalmente, cafeicultores realizam estimativas da produção baseadas na experiência, acompanhando os ciclos passados de produção e realizando inferências sobre a produção a partir da florada e frutificação do cafeeiro. De outra forma, as estimativas para determinar o ponto ideal da colheita baseada na maturação podem ser realizadas por meio da derrça manual do cafeeiro. No processo, coletam-se amostras da colheita e a partir delas são realizadas a contagem e a separação dos frutos manualmente em cada estágio de maturação (Kazama et al, 2020).

Além da produtividade do café, outro fator importante é a qualidade do café. Segundo Chagas (1998), a qualidade do café pode ser influenciada por diversos fatores, dentre eles, temperatura e umidade relativa do ar durante as fases de floração, amadurecimento e frutificação e precipitações pluviométricas. Estes fatores podem promover maturação heterogênea, além de fermentações indesejáveis nos frutos maduros, gerando perda de qualidade antes mesmo da realização da colheita. No café, para uma mesma planta há frutos em diferentes estágios de maturação. A maturação dos frutos no cafeeiro não é uniforme pois geralmente há de 2 a 4 períodos de floração por safra do café (Barros et al. 1978; Rena e Barros 2004).

Devido às variações de estágio de maturação, o café acaba sendo colhido verde, o que gera prejuízos relacionados ao peso do grão, rendimento da colheita e a classificação dos grãos, consequentemente, o valor final do produto é afetado (Alves et al, 2019). Ferrão et al. (2004) e Fonseca et al. (2007) sugerem que para dar início às atividades de colheita o percentual de grãos maduros deve alcançar 80%. Assim, considerando que estimar a maturação do café é importante para a definição do ponto

ideal da colheita e que estimativas a partir de técnicas atuais podem ser imprecisas e gerar operacionais e tempo adicional para o agricultor, uma alternativa seria utilizar novas tecnologias para automatizar a classificação dos frutos do café por planta e possibilitar a definição mais assertiva do momento da colheita. Essas tecnologias podem envolver dispositivos eletrônicos na captura de imagens (Kamilaris & Prenafeta-Boldú, 2018; Liu et al., 2018) e utilização de algoritmos de aprendizado de máquina para geração de modelos na detecção dos frutos (Kang & Chen, 2020; Kazama, 2020). Tais tecnologias, podem auxiliar no processo de tomada de decisão, promover a otimização de recursos consumidos na lavoura, e contribuir para a separação automática dos frutos.

Por meio de dispositivos móveis e visão computacional, os agricultores podem definir com maior precisão o ponto ideal da colheita, e delimitar as áreas onde há maior intensidade de frutos maduros. Assim, com uma imagem da planta pode-se contar o número de frutos e determinar a porcentagem de frutos nos diferentes estádios de maturação. Para melhorar a exatidão do processo várias plantas podem ser amostradas. A detecção de frutos em várias plantas por sua vez pode gerar estimativas da produção por meio da geração de mapas de produtividade.

A visão computacional tem sido utilizada nas tarefas como: classificação, segmentação semântica, detecção de objetos e segmentação de instâncias (Yang Yu, 2019). A classificação é uma técnica de aprendizado supervisionado que visa extrair classes de uma imagem visando identificar quais objetos são encontrados na imagem. A detecção de objetos é uma tarefa que identifica e classifica os objetos presentes nas imagens e usa caixas delimitadoras para delimitar o espaço onde o objeto se localiza. Posteriormente, os objetos são rotulados. A segmentação semântica vincula cada pixel à uma classe. Assim, as imagens são divididas em agrupamentos de pixels que podem ser classificados. Por fim, a segmentação de instâncias detecta os objetos individualmente na imagem vinculando os pixels do objeto à uma classe (Santaquiteria, 2020).

Na literatura, podem ser encontradas algumas abordagens para detecção e segmentação de frutos. Kazama (2020), utilizou deep learning para estimar a produtividade e maturação do café. Além disso, foram definidas zonas de manejo para determinar a melhor regulagem das máquinas para colher. No artigo, foi utilizado a arquitetura Fast-RCNN que realiza a detecção de objetos. Comparando a contagem real com a predita nas imagens obteve-se R^2 de 0,92. Ramos et al. (2017) propuseram

um algoritmo de processamento baseado em segmentação de regiões homogêneas e zonas com informações de textura, gerando contornos prováveis de serem arcos de frutas do café, obtendo valor de R^2 de 0,80 na relação contagem real e predita. Weikuan et al. (2020) propuseram um método aprimorado para reconhecimento e segmentação de maçãs, utilizando como rede backbone (entende-se por backbone a arquitetura principal de uma CNN orientada a classificação) a combinação de ResNet (Residual Neural Network) (Kaiming, 2015) e DenseNet (Huang, 2017) para extração de características de interesse nas imagens obtendo taxa de precisão de 97,31 % e recall de 95,75%. Hanwen et al. (2020) desenvolveram uma rede profunda DaSNet-v1 para realizar a detecção e segmentação de maçãs em galhos em ambientes de pomar obtendo precisão de 88,0 % e recall de 86,8%. Yang et al. (2019) introduziram a Mask-RCNN para detecção de morangos. No trabalho de Yang (2019), o Resnet50 foi adotado como rede backbone, combinado com a arquitetura Feature Pyramid Network (Lin, 2017) para extração de características. Liu (2020) propôs um modelo aprimorado baseado na YOLOv3 para realizar detecção de tomates e obteve precisão de 93,09 % e recall de 94,75%.

Diante do exposto, verifica-se que a determinação do ponto ideal de colheita é um problema enfrentado por produtores de café. As técnicas utilizadas pelos produtores consomem muito tempo e são imprecisas. Considerando que a combinação de técnicas de visão artificial com as de aprendizado de máquinas tem potencial para resolver esse problema, o presente trabalho foi realizado com o objetivo de desenvolver um sistema de segmentação de instâncias, que realiza a detecção, a contagem, a classificação dos frutos nos três estádios de maturação de frutos de café, a saber, cereja, verde e passa. Para isso, o sistema foi desenvolvido com base em imagens da planta do café e utilização da técnica de aprendizado profundo Mask R-CNN. Mask R-CNN é um método comum de detecção de alvos que utiliza a estrutura do Faster R-CNN incrementando uma ramificação extra para prever uma máscara de segmentação contendo os pixels do objeto de interesse, no caso, os frutos (He et al., 2019). Tal técnica, realiza a segmentação de instância para cada região proposta. Além disso, a segmentação é realizada em paralelo com a classificação e localização do alvo por meio das caixas delimitadoras (Kaiming, 2018).

2.2. Material e Métodos

O experimento foi conduzido em lavouras de café arábica em setembro de 2020 no município de Monte Carmelo, Minas Gerais. A lavoura está localizada na quadrícula delimitada entre as latitudes $-18,47217^{\circ}$ e $-18,4756^{\circ}$, e as longitudes $-47,21258^{\circ}$ e $-47,2190^{\circ}$. A cultivar de café arábica instalada na área do experimento era o catuaí vermelho, com espaçamento de 3,4 m na entrelinha e 0,8 m na linha de plantio, resultando em um estande de 3700 plantas por hectare. Na época, as plantas de café apresentavam ramos com frutos em variados estádios de maturação. Neste contexto, foram coletadas imagens para desenvolvimento do trabalho, o qual foi dividido nas seguintes etapas: aquisição de imagens, pré-processamento, rotulagem, redimensionamento, recorte, divisão dos dados e treinamento e avaliação do modelo.

2.2.1. Aquisição de imagens

As imagens foram coletadas em períodos diferentes do dia (manhã e tarde), três semanas antes da colheita do café, sob variadas condições de iluminação (dias ensolarados e nublados). As imagens foram obtidas no tamanho 1280 x 1280 pixels com smartphone Samsung Galaxy A51 com câmera de resolução de 32 MP. Buscou-se por obter diferentes angulações de visão na captura das imagens focando nos ramos produtivos do terço médio das plantas de café em 10 plantas escolhidas aleatoriamente. A coleta em situações reais de campo gerou fenômenos como coincidência e sobreposição de áreas, incluindo a sobreposição dos galhos e folhas sobre os frutos e a sobreposição de partes de um fruto pelo outro. Ao todo, foram obtidas 78 imagens, dentre elas, 72 foram coletadas para obter quantidades balanceadas de instâncias de cada classe de interesse, a saber: cereja, passa e verde. Assim, o conjunto das 72 imagens foi formado por 3 subconjuntos de 24 imagens com predominância de algumas das classes. As 6 imagens restantes foram utilizadas para predição a partir de técnicas de janelas deslizantes, após a construção do modelo final.

2.2.2. Pré-processamento

Após a coleta, as imagens foram pré-processadas para aumentar a quantidade de dados por meio da divisão delas, separar os conjuntos de treinamento, teste e validação e rotular os objetos de interesse presentes em cada imagem. Desta forma, na primeira etapa, as imagens foram redimensionadas para o tamanho de 1024 x 1024

pixels, em seguida cada uma foi repartida em 16 partes de 256 x 256 pixels, resultando em 1152 imagens (72 x 16). A partir das imagens 256 x 256, buscou-se por obter do conjunto resultante, imagens que continham as classes de interesse, ou seja, frutos cerejas, verdes e passas. Foi importante selecionar as imagens que continham frutos pois o modelo precisa das instâncias de interesse bem como seus rótulos. Sendo assim, aquelas que não apresentaram ao menos uma das classes foram descartadas. Tal processo, resultou em 800 imagens. Após essa etapa, selecionou-se imagens aproximadamente homogêneas em relação às classes presentes (com predominância de alguma delas). Assim, o conjunto final resultou em 300 imagens divididas em 3 conjuntos de 100 para abranger as três classes de frutos (passa, cereja, verde) conforme pode ser observado no exemplo da Figura 2.1. Nessa separação de classes não foi realizada distinção entre tonalidades de coloração de frutos, desta forma, quando o fruto em sua maior parte se enquadrava com coloração vermelha era classificado como cereja, o mesmo foi adotado para os demais, para as colorações verdes e pretas a classificação se enquadrava em verdes e passas respectivamente.

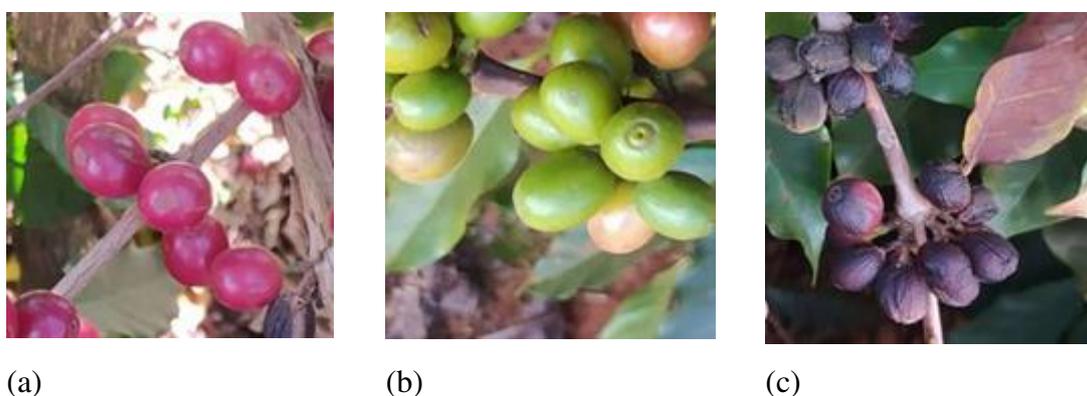


Figura 2.1: Imagens obtidas a partir do pré-processamento (a) frutos cerejas, (b) frutos verdes, (c) frutos passas.

Dentre as 300 imagens resultantes, os dados foram divididos em treinamento, validação e teste, correspondendo a 70%, 10% e 20% dos dados respectivamente. Desta forma, o conjunto de treinamento, validação e teste foi formado por 210, 30, e 60 imagens respectivamente. Para cada conjunto, foi realizada a rotulagem que consiste em demarcar manualmente os pixels dos frutos de café de acordo com suas classes. Desta forma, para auxiliar no processo de rotulagem, foi desenvolvida uma ferramenta para anotar as informações experimentais e gerar imagens com máscaras dos frutos. Para a rotulagem das imagens foram utilizados algoritmos de demarcação

de formas geométricas e armazenamento de informações utilizando as bibliotecas: cv2 (openCV) versão 3.4.15, numpy versão 1.21.2, skimage versão 0.18.3 do python 3.8. Assim, ao clicar nas coordenadas de centro, raio maior e raio menor das elipses representativas dos frutos, o algoritmo armazena automaticamente as informações em um arquivo do tipo json. As informações foram atualizadas a cada demarcação de um novo fruto. Com as coordenadas de cada fruto armazenadas foi possível obter as imagens com máscaras. Essas imagens com as máscaras foram usadas no treinamento e otimização dos parâmetros do modelo. Além disso, o desempenho do modelo treinado para segmentação de instância foi avaliado comparando as imagens segmentadas manualmente (com máscaras) com os resultados de segmentação do modelo (máscaras preditas). A Figura 2.2 mostra um exemplo de máscara gerada a partir da ferramenta de anotação.

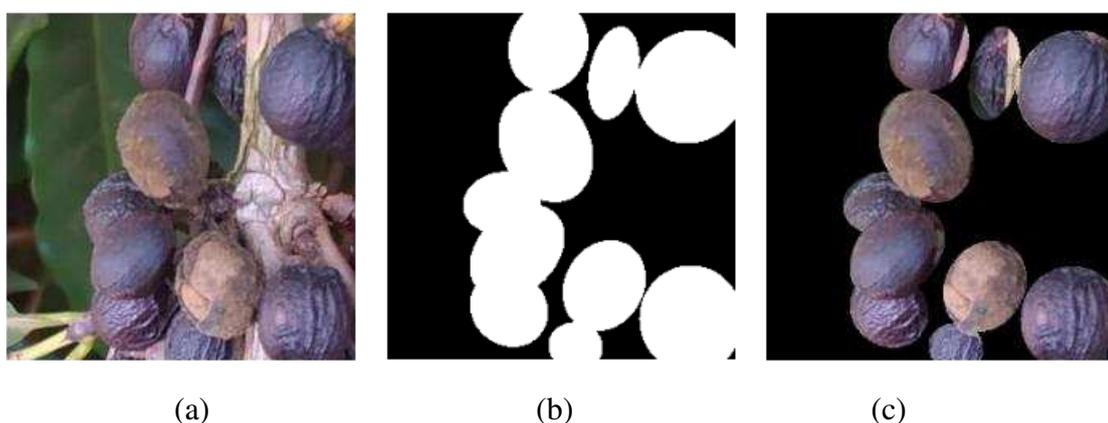


Figura 2.2 Máscara gerada a partir da ferramenta de rotulagem desenvolvida neste trabalho. a) imagem original, b) máscara binária gerada pela ferramenta de rotulagem desenvolvida, c) sobreposição da máscara em relação a imagem demarcando a área de interesse

2.2.3. Treinamento do modelo de segmentação de instâncias Mask-RCNN

Para desenvolvimento do sistema de detecção de frutos de café nos diferentes estágios de maturação utilizou-se a técnica Mask-RCNN. Mask-RCNN é uma rede neural profunda destinada a resolver problemas de segmentação de instâncias em visão computacional. Entende-se por segmentação de instâncias a tarefa que realiza a detecção do alvo (fruto) e em paralelo realiza a segmentação semântica em que cada pixel do alvo é rotulado. Mask-RCNN é uma extensão do Faster-R-CNN (Ren,

2017) utilizado em tarefas de detecção de objetos. O Faster R-CNN foi aprimorada a partir das redes RCNN (Region-based Convolutional Neural Network), que é uma arquitetura de rede neural que tem como objetivo agrupar e classificar regiões em uma imagem, indicando se cada uma dessas regiões contém um fruto do café por exemplo, comparando esta resposta com a resposta da rede neural em um dado momento. No caso de Faster-RCNN, ao invés de usar o método de busca seletiva do RCNN usa uma rede para geração de propostas de regiões conhecida como Regional Proposal Network (RPN), tal fato gerou um ganho de tempo de processamento (Ren, 2017). O algoritmo da RPN percorre a feature map (imagem processada por um filtro que tenta capturar os objetos da imagem) de cada imagem com janelas deslizantes, que por sua vez possuem 'âncoras' que encontram regiões de interesse em cada uma das instâncias das janelas. São aplicadas em cada janela diversas 'âncoras' de tamanho e proporção diferentes, para que objetos de tamanhos diferentes possam ser identificados (Ren, 2017). Mask-RCNN é uma estrutura de duas etapas. A primeira etapa escaneia a imagem e gera propostas (áreas que podem conter um objeto). A segunda etapa classifica as propostas e gera caixas delimitadoras e máscaras. Ambos os estágios são conectados à estrutura da rede backbone que gera as feature maps. A imagem de entrada é primeiro passada pelo backbone CNN para obter o mapa de características. A partir das saídas da rede backbone, rede RPN gera as Rols múltiplas (possíveis regiões com a localização dos frutos). Para isso, prevê as coordenadas e um rótulo para cada âncora. As âncoras são retângulos predefinidos de tamanhos diferentes, cobrindo toda a imagem de entrada. Em seguida, o método RoIAlign (alinhamento dos Rols) altera o tamanho dos Rols para um tamanho fixo. Finalmente, os Rols com tamanho fixo fluem para dois ramos diferentes: um que prevê frutos do café por regressão e classificação e outro ramo que realiza em paralelo a segmentação. Na regressão ocorrem as correções de coordenadas das caixas delimitadoras, na classificação os frutos serão classificados por tipo e na segmentação uma rede totalmente convolucional (FCN) realiza em paralelo a geração das máscaras. Mask-RCNN tem a capacidade de prever uma máscara de segmentação para cada objeto na imagem.

O algoritmo usado utilizou transferência de aprendizado na inicialização dos pesos da rede. A transferência de aprendizado visou reduzir o tempo que seria gasto com o treinamento do modelo do zero além de tentar suprir a falta de um grande conjunto de dados para o treinamento. Tal técnica, pode gerar ganhos de performance pois usa o

conhecimento de modelos previamente treinados com grande volume de dados (Mengola et al., 2017), neste caso os pesos foram pré treinados a partir de uma banco de imagens MS-COCO (Microsoft Common Objects in Context) . MS-COCO constitui um conjunto de dados com 328 k de imagens incluindo 91 categorias (Yang, 2019). A transferência de peso foi usada para prever os parâmetros de segmentação de instância para cada categoria (dos frutos do café).

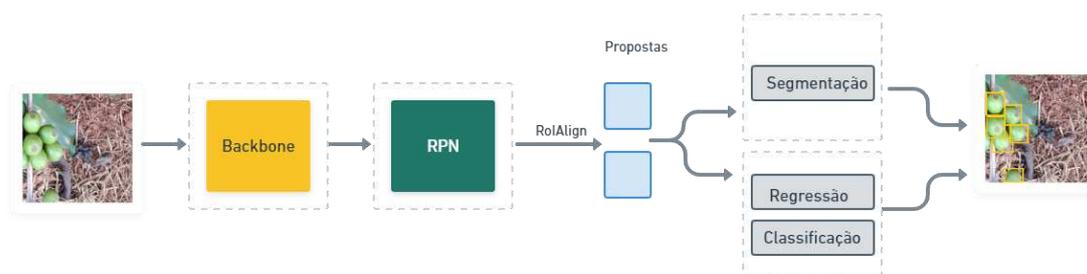


Figura 2.3. Arquitetura de aprendizado profundo adotada no trabalho Mask-RCNN. A rede backbone (ResNet-101) gera mapas de características que são entradas para redes RPN (Region Proposal Network). A RPN gera regiões de interesse com as possíveis localizações dos frutos. Posteriormente é aplicado o método de RoIAlign (manter os Rols em tamanhos fixos). Os Rols com tamanho fixo fluem para dois ramos diferentes para prever frutos de café por regressão e classificação e em paralelo realiza-se a predição das máscaras.

O experimento foi conduzido sob a estrutura de desenvolvimento de aprendizado profundo do TensorFlow e Keras na versão 1.15.3 utilizando o Google Colaboratory (GC). Google Colaboratory é um serviço em nuvem que permite prototipar modelos de aprendizado de máquina utilizando hardwares como GPUs e TPUs (Bisong, 2019). Neste trabalho, utilizou-se GPU do GC Tesla 12GB, processador Intel Xeon 2.2 GHz e 13GB de RAM no treinamento do modelo.

Para geração do modelo, utilizou-se Resnet-101 como rede residual backbone que é uma rede neural profunda do tipo FPN (Feature Pyramid Network). Resnet- 101 foi criada para classificação de mais de 1000 categorias (Ghosal, 2019) é uma rede neural residual profunda com 101 camadas de profundidade e foi escolhida neste trabalho por apresentar mais camadas de convolução do que outros backbones como: VGG-19, Resnet-50 (Liu, 2019).

O modelo construído a partir da arquitetura Mask RCNN visou detectar os alvos: frutos cereja, passa e verde nas imagens marcando a probabilidade da

classificação, caixas delimitadoras e máscaras correspondentes a cada fruto. A taxa de aprendizagem inicial foi definida como 0,001. O número de interações foi definido como 10 após verificações de curvas sobre a aprendizagem do modelo e considerando que como o método utiliza transferência de aprendizado o que dispensa muitas iterações para se alcançar bons desempenhos na predição. O limiar de IoU (definido pela área de intersecção sobre a área de união entre máscara real e predita) utilizado foi de 0,5. Se IoU da máscara foi maior que o limiar, classifica-se a máscara como TP (verdadeiro positivo), caso contrário, então é uma detecção errada e a máscara é classificada FP (falso positivo). O número de steps por época foi de 100.

2.2.4. Avaliação da performance do modelo

Para avaliar o desempenho do modelo, foram utilizadas as métricas advindas da matriz de confusão, a saber: precisão (P), recall (R) e F1-score (F1), em que TP é verdadeiro positivo, FP é falso positivo, FN é falso negativo. Tais métricas, são definidas pelas equações (2) a (4).

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = \frac{2 * P * R}{P + R} \quad (4)$$

2.2.5. Técnica de janelas deslizantes

A partir do modelo final desenvolvido e das 6 imagens restantes foi realizada a predição a partir de técnicas envolvendo janelas deslizantes. As imagens foram redimensionadas para tamanho 1024 e 1024 pixels. Foi criado um algoritmo que consiste em deslizar uma janela de 256 pixels por 256 pixels sobre a imagem com um intervalo entre uma janela e outra de 256 pixels na horizontal (na linha) e 256 pixels na vertical (coluna da matriz). Desta forma foi possível realizar a detecção dos alvos e predição da contagem dos frutos em uma imagem em situação real de campo. Seis imagens foram suficientes para avaliar o desempenho do modelo, pois 6 imagens resultaram em 96 janelas, pois nessa técnica a cada 256 pixel realizava-se a detecção das instâncias de forma a cobrir toda imagem.

2.3. Resultados e discussão

O modelo construído a partir da arquitetura Mask RCNN visou detectar os alvos: frutos cereja, passas e verde nas imagens marcando a probabilidade da classificação em cada instância, as caixas delimitadoras e máscaras correspondentes a cada fruto. O desempenho do treinamento é apresentado na Figura 2.4. Nesta Figura, as três perdas no treinamento do modelo foram mostradas para o conjunto de validação e treinamento. A função de perda no treinamento consiste em três componentes que são: a perda na classificação dos frutos de café, a perda de regressão das caixas delimitadoras e a perda das máscaras preditas. Os gráficos indicam que os valores de perdas diminuem, gradualmente de 1 para 9 iterações. Quando o número de iterações foi maior que 8 na análise do erro da máscara, os valores da função de perda tenderam a se estabilizar para o conjunto de validação e treinamento.

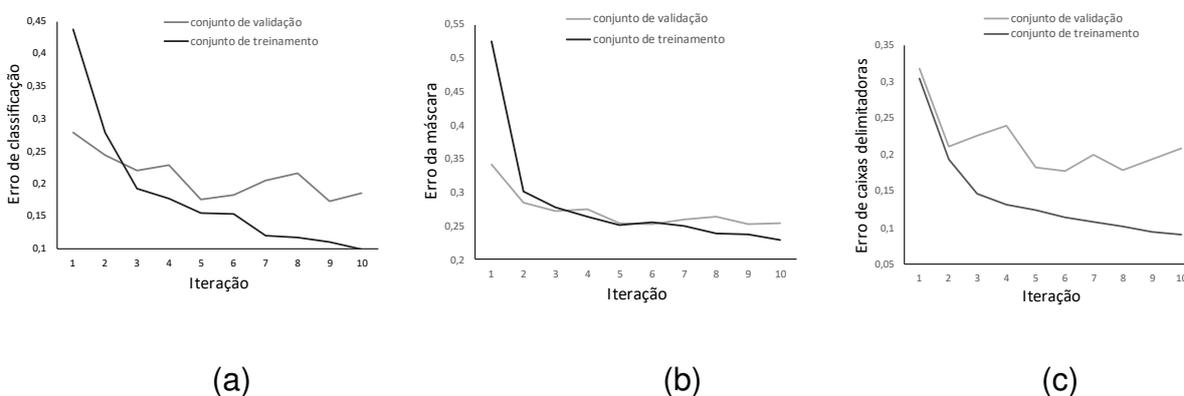


Figura 2.4. Curvas obtidas pelo treinamento do modelo MaskRCNN (função custo versus número de iterações) a) Erro na classificação dos objetos b) Erro na segmentação semântica c) Erro na demarcação das caixas delimitadoras.

Verificou-se que, como Mask-RCNN utilizou transferência de aprendizagem não foram necessárias muitas iterações para realizar as tarefas de segmentação de instâncias dos frutos. Na décima interação, os valores da função perda de máscaras (cross-entropy) (Bhuiyan, 2020) foram de 0,228 e 0,224 para conjunto de treinamento (CT) e do conjunto de validação (CV), respectivamente. Como foram valores semelhantes é um indicativo de que não ocorreu redução na capacidade de generalização da informação, visto que, nas amostras de validação, o desempenho foi satisfatório. Os valores de perda na classificação foram de 0,010 (CT) e 0,020 (CV)

e valores de perda na regressão de caixas delimitadoras foram de 0,087 e 0,237 para CT e CV, respectivamente. Desta forma, o resultado na regressão referente as caixas delimitadoras apresentaram um resultado inferior às outras tarefas. As curvas de validação para perda na classificação e máscara atingiram um estado de convergência. O resultado indicou que o treinamento do modelo funcionou bem para classificação e segmentação. Para as perdas regressão das caixas delimitadoras, a curva na validação tendeu a diminuir com o número de interações, mas não atingiu um platô. Erros assim podem ter sido gerados pela própria rotulagem manual, visto que, é sujeita a erros e imprecisões inerentemente humanas. Tais imprecisões, se devem ao fato que o processo de rotulagem é subjetivo e dependente da percepção do especialista que pode não demarcar todos os pixels do objeto de interesse.

Realizou-se uma comparação visual entre as imagens originais com segmentação de instâncias geradas pelo modelo. Contabilizou-se os frutos que o modelo classificou corretamente, os frutos não detectados e os frutos classificados erroneamente. O desempenho do modelo foi avaliado para as 60 imagens de teste a partir das métricas advindas da matriz de confusão.

Na Tabela 2.1 é mostrada a matriz de confusão. Por ela verifica-se que a contagem real de frutos das classes cereja, passa e verde foram: 375, 269, 216, respectivamente. Esses resultados evidenciam que mesmo realizando a divisão dos dados de forma a se obter conjuntos com balanceamento de classes houve uma predominância na contagem de frutos cerejas. O modelo realizou a detecção de 234, 185 e 117 frutos das classes cereja, preto e verde, respectivamente. Esses resultados são apresentados na matriz de confusão (Tabela 1), a acurácia global F1-score é mostrada na Tabela 2.2, o valor limiar de IoU é de 0,50.

Pela Tabela 2.1, verifica-se que houve muitas detecções errôneas em que o modelo classificou frutos cereja como fundo. Tal fato, pode ter acontecido pela similaridade das cores dos galhos e troncos com a coloração vermelha do fruto, a mesma explicação pode ser aplicada para detecções de fundo enquanto a classe verdadeira seria frutos passas. Detecções errôneas ocorreram também na classificação de partes do fundo como fruto verde, tal fato, se deve a semelhança de coloração verde com folhas do cafeeiro.

Tabela 2.1: Matriz de confusão, gerada a partir da contagem de frutos para as imagens de teste obtidas para o primeiro modelo.

Classe verdadeira	Classe Predita			
	Cereja	Passa	Verde	Fundo
Cereja	234	8	7	126
Passa	4	185	2	78
Verde	8	3	117	88
Fundo	0	16	10	-

Na Tabela 2.2 são mostradas as métricas: precisão, recall e f1-score para as três classes de interesse. Observa-se que o modelo para frutos cereja apresentou melhor desempenho de precisão de 95,10% em comparação aos demais estádios de maturação e segundo melhor desempenho de recall. Tal fato pode ter ocorrido pelo destaque que a coloração de frutos cerejas proporcionam gerando uma porcentagem menor de falsos positivos em relação às outras classes. Para os demais estádios de maturação (verde e passa), a precisão foi de 91,40% (verdes) e 87,26% (passa). Em relação ao recall, houve maior percentual de falsos negativos para os frutos verdes, principalmente, porque muitos frutos verdes não foram detectados. Tal fato pode ter ocorrido pela similaridade da coloração dos frutos e folhas. Além disso, 10 instâncias de fundo foram confundidas como frutos verdes. Tais fatos contribuíram para que os frutos verdes apresentassem o pior desempenho com valor de f1-score de 68,02 %.

Tabela 2.2: Métricas obtidas da matriz de confusão geradas para avaliar o primeiro modelo de detecção de frutos.

Parâmetro de avaliação	Cereja	Verde	Passa
Precisão %	95,12	91,4	87,26
Recall %	62,4	54,16	68,77
F1 %	75,36	68,02	76,92

Para frutos passa, houve 16 instâncias de fundo detectadas como sendo da classe frutos passa, pois, algumas folhas secas e galhos apresentaram coloração escura, além de ocorrer sombra, gerando detecções erradas. Problemas de sombreamento não foram evitados pois, buscou-se obter um imagens em ambientes não controlados, para verificar a aplicabilidade do modelo em situações reais de

campo. Verificou-se que frutos da classe cereja foram classificados erroneamente com maior frequência pelo modelo como verdes do que como passa. Além disso, frutos verdes foram confundidos mais frequentemente como frutos cerejas do que como passas. Tal resultado, pode ser esperado pois frutos verdes no processo de se tornarem maduros alteram sua coloração e podem se assemelhar a classes cerejas. Desta forma, erros assim, foram gerados pela coloração pouco avermelhada de alguns frutos em fase de amadurecimento. De uma forma geral, para todas as classes houve maior frequência de falsos negativos em detrimento a falsos positivos principalmente devido a grande quantidade de grãos não detectados. Isso pode ter relação com o fato de que imagens obtidas dos ramos apresentam o foco dos frutos à frente, e desfocando os frutos mais ao fundo. Tais problemas, poderiam ser contornados mudando a posição e foco das câmeras no processo de obtenção das imagens.

De uma forma geral, o modelo identifica os frutos de café de maneira satisfatória. Satisfatória pois o número de verdadeiros positivos superou o número de falsos positivos e falsos negativos para as três classes estudadas. No entanto, como os frutos do café se apresentam aglomerados, o modelo apresenta dificuldade na detecção de frutos parcialmente ocultos e frutos localizados no fundo na imagem, tal fato pode ser observado na Figura 2.5. O modelo foi capaz de identificar frutos em escalas variadas de tamanho para as três classes analisadas. Para regiões de proposta (boxes), o modelo também apresentou dificuldade em demarcar frutos que se apresentam aglomerados ou oclusos por outros frutos, folhas e ramos, muitas vezes marcando dois frutos como um único.

Na Figura 2.5 são mostradas algumas imagens de teste com as caixas delimitadoras nos frutos, as máscaras preditas e as classificações para cada alvo. Verifica-se que o modelo teve bom desempenho nas detecções dos frutos. Bom desempenho, pois, detectou a maioria dos frutos, não detectando frutos no fundo e frutos que não aparecerão por inteiro na imagem. Na imagem mais à esquerda evidencia que para frutos no fundo e sobrepostos o modelo teve dificuldade de detecção.



Figura 2.5. Exemplos de imagens com frutos detectados

Foi realizada uma comparação do número de frutos contados visualmente na imagem com o número de frutos detectados pelo modelo automático em cada estágio de maturação. Os resultados dessa comparação são apresentados nos gráficos de dispersão da Figura 2.6. O modelo reconheceu os frutos em estágio de maturação cereja com valores de R^2 de 0,80 e RMSE de 3,77. Para os demais estádios de maturação os valores do coeficiente de determinação (R^2) foram 0,73 e RMSE de 3,54 para frutos passas e R^2 de 0,73 e RMSE de 4,72 para frutos verdes.

Comparou-se as áreas preditas e reais de cada uma das classes. Sendo que a área predita é o número de pixels que o modelo identificou como pertencente à uma determinada classe e área real é a área rotulada gerada a partir das máscaras considerando as informações do arquivo *JSON* para as imagens de teste. O coeficiente de determinação para a relação área predita e área real foi de 0,80; 0,81; 0,66 para as classes cereja, passa e verde, respectivamente, e os valores de RMSE de 3166, 2936, e 3040 pixels para as classes cereja, passas e verdes, respectivamente. O pior desempenho foi para a classe verde devido a refletância desses frutos serem similares as das folhas.

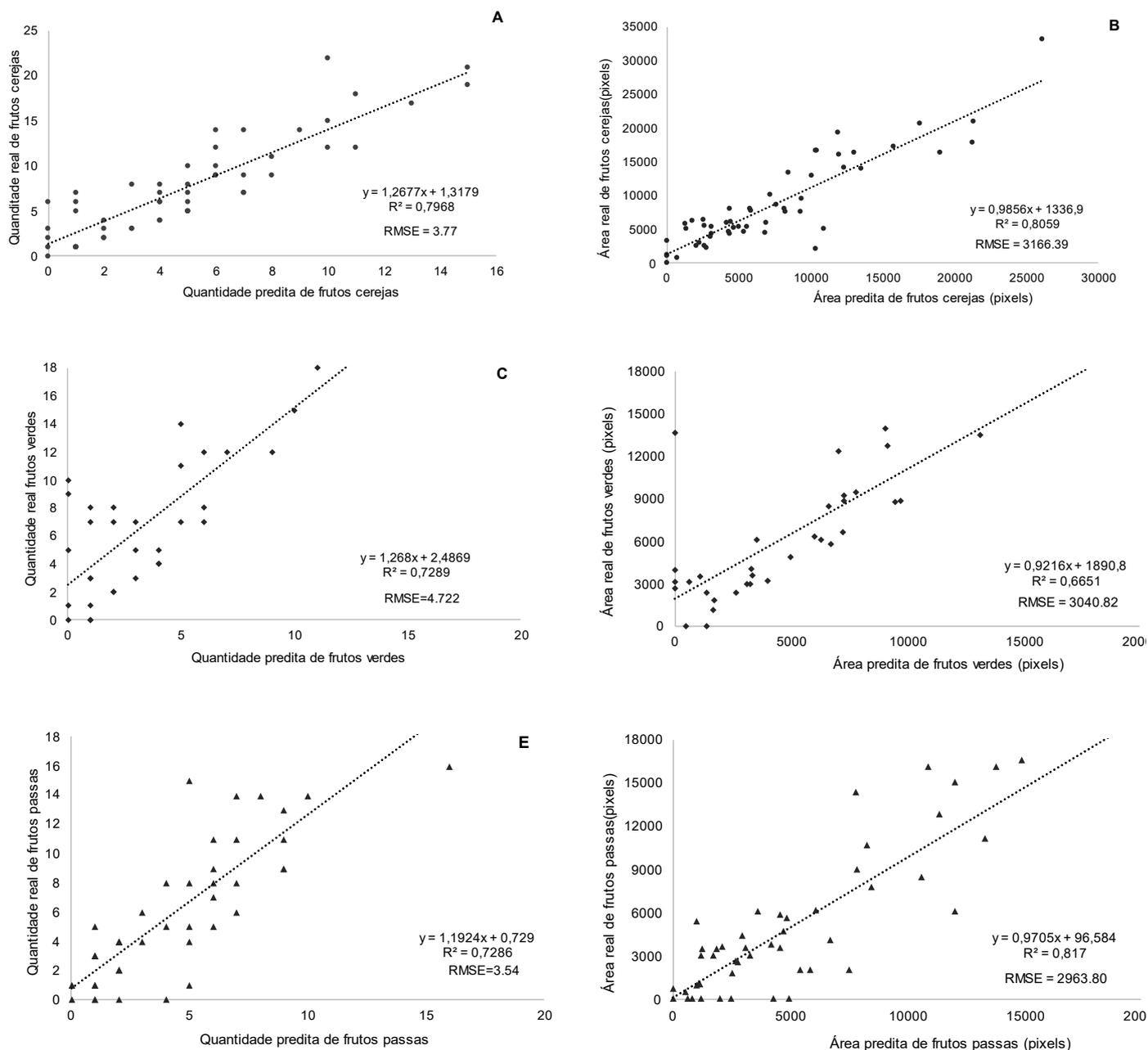


Figura 2.6. Eixo a esquerda regressão entre contagem de fruto real e contagem predita para os três estádios de maturação. No eixo à direita área relação entre área predita e área real

Com base nas técnicas abordadas foi possível gerar predição sobre imagens inteira (resolução original) para contagem de frutos nos diferentes estágios de maturação. Nessa abordagem, foram utilizadas técnicas de janelas deslizantes que consiste em deslizar uma janela de tamanho 256 x 256 com um intervalo de separação de 256 pixels na horizontal e vertical sobre uma imagem não processada a fim de realizar o processo de segmentação de instâncias em cada parte da imagem. O

processo começa no canto superior esquerdo da imagem, a janela se move da esquerda para a direita até o canto inferior direito. A cada movimento da janela é realizada a detecção e classificação dos frutos. O resultado do processo é armazenado em uma nova matriz com a mesma posição de coordenada.

Para classificar uma imagem inteira por janela deslizante, foram utilizadas as seis imagens não processadas durante o treinamento. Os resultados são apresentados nas Tabela 2.3 e Tabela 2.4. A acurácia global F1-score foi de 84,40 % para frutos cerejas, 81,05 % para frutos verdes e 84,27% para frutos passas. O número baixo de falsos positivos e falsos negativos torna o modelo viável para ser utilizado no monitoramento da maturação do cafeeiro e possível definição do ponto ideal da colheita. A classe cereja foi a que teve maior valor de precisão indicado pela menor quantidade de falsos positivos em relação as outras classes. No entanto, para a mesma classe teve o menor valor de recall devido à grande frequência de frutos não detectados. A classe verde teve menor valor de precisão. As imagens da Figura 2.7 mostram as detecções dos frutos. De modo geral, o uso de redes neurais profundas (Mask-RCNN com pesos pré-treinados) foram eficientes na detecção, segmentação e classificação dos frutos de café nos diferentes estádios de maturação. A técnica de janela deslizante associada a Mask RCNN se mostrou eficiente para a detecção dos frutos em imagens reais da cultura do café caracterizada por aglomeração de frutos em pequenas áreas.

Tabela 2.3. Matriz de confusão, gerada a partir da contagem de frutos obtidas das imagens não processadas aplicando-se o modelo em cada janela de 256 x 256 pelo método de janelas deslizantes.

Classe verdadeira	Classe Predita			
	Cereja	Passa	Verde	Fundo
Cereja	295	22	9	66
Passa	5	209	7	36
Verde	3	7	154	44
Fundo	4	1	2	-

Tabela 2.4. Métricas obtidas obtida para a abordagem que utilizou a técnica de janelas.

Parâmetro de avaliação	Cereja	Verde	Passa
Precisão %	96,09	89,53	87,44
Recall %	75,25	74,03	81,32
F1 %	84,4	81,05	84,27

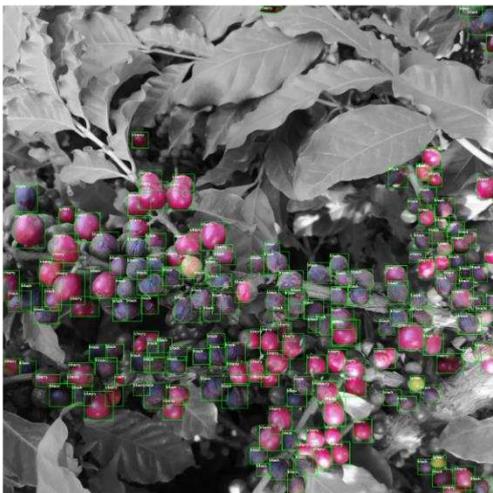
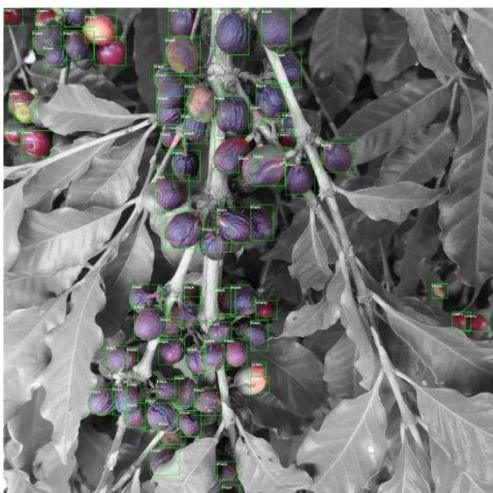


Figura 2.7. Imagem geradas da técnica de janelas deslizantes. A contagem dos frutos foi realizada na própria matriz de confusão.

2.4. Conclusão

Mask RCNN com uso de transferência de aprendizado foi eficiente na detecção e contagem de frutos de café nos diferentes estágios de maturação. Os valores de precisão foram 0,897; 0,900; 0,891 para as classes cereja, verde e preto, respectivamente. Os valores de recall das três classes foram: 0,759; 0,7; 0,813. Os valores de recall foram menores que os de precisão devido a muitos frutos não detectados. Para a técnica de janela deslizante os valores de precisão foram: 0,974; 0,906; 0,878; e recall foram: 0,753; 0,74; 0,813 para as classes cereja, verde e passa, respectivamente. Essa metodologia tem potencial para integrar a outras tecnologias e permitir a automatização de classificação de frutos na própria planta podendo ser útil na definição do ponto de colheita, além disso, pode ser adaptada para estimar a produtividade dos frutos com as adaptações necessárias conforme tipo de planta e local estudados. Com estimativas de produtividades, a metodologia proposta tem potencial para à partir de imagens georreferencias gerar mapas de produtividade das lavouras cafeeiras.

2.5. Referências

APARECIDO, L. E.O.; ROLIM, G. DE S.; MORAES, J. R. S. C.; VALERIANO, T. T. B.; LENSE, G. H. E. Maturation periods for *Coffea arabica* cultivars and their implications for yield and quality in Brazil. **Journal of the Science of Food and Agriculture**, v. 98, n. 10, p. 3880–3891, 2018.

ALVES, E. A.; COSTA, J. N. M.; SANTOS, J. C. F. Procedimentos de colheita do café. In: Marcolan, A. L.; Espídula, M. C. (ed.). **Café na Amazônia. Brasília, DF: Embrapa**, 2015. p. 347-358. Acesso em: 21 dez. 2019.

AMORIM, H. V. Estudo sobre a alimentação mineral do cafeeiro: efeito da adubação N, P e K, na composição química do solo, do fruto e na qualidade da bebida. **Anais da ESALQ, Piracicaba**, v. 22, p. 130-152, 1965.

BHUIYAN, M. A. E; WITHARANA, C; LILJEDAHN, A. Use of Very High Spatial Resolution Commercial Satellite Imagery and Deep Learning to Automatically Map Ice-Wedge Polygons across Tundra Vegetation Types. **Journal of Imaging**. 2020.

BISONG, E. "Google Colaboratory" in Building Machine Learning and Deep Learning Models on Google Cloud Platform, Berkeley, CA, **USA:Apress, pp. 59-64, 2019.**

CHAGAS, S. J. R.; LEITE, I. P. Alguns aspectos a serem observados na colheita para garantir a produção de um café de melhor qualidade. Lavras, MG: **EPAMIG. 2 p. EPAMIG. 1998.**

FERRÃO, R. G.; FONSECA, A. F. A.; FERRÃO, M. A. G.; VOLPI, P. S.; MARQUES, E. M. G.; ZUCATELI, F. Café conilon: técnicas de produção com variedades melhoradas. Vitória, ES: **Incaper, 2004.**

FUENTES, M. S.; ZELAYA, N. A. L.; AVILA, J. L. O. Coffee Fruit Recognition Using Artificial Vision and neural NETWORKS. 5th **International Conference on Control and Robotics Engineering 2020.**

GALEANO, E. V.; KROHLING, C. A. Análise de risco na cafeicultura de arábica no estado do espírito santo considerando colheita manual e semimecanizada. **X Simpósio de Pesquisa dos Cafés do Brasil, Vitória, 2019.**

GHOSAL, P; NANDANWA, L; KANCHAN, S; BHADRA, S; CHAKRABORTY, J; NANDI, D. Brain Tumor Classification Using ResNet-101 Based Squeeze and Excitation Deep Neural Network," 2019 Second International **Conference on Advanced Computational and Communication Paradigms (ICACCP)**, Gangtok, India, 2019.

HANWEN, K.; CHAO, C. Fruit detection, segmentation, and 3D visualisation of environments in apple orchards, **Computers and Electronics in Agriculture, 2019.**

HUANG, G.; LIU, Z.; MAATEN, V.; WEINBERGER, K. Densely Connected Convolutional Networks. IEEE. Conference on **Computer Vision and Pattern Recognition (CVPR), 2017.**

KAIMING, He; XIANGYU, Z; SHAOQING, R; SUN, J. Deep Residual Learning for Image Recognition. IEEE. Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

KAZAMA, E.H.; SILVA, R.P., TAVARES, T.O. et al. Methodology for selective coffee harvesting in management zones of yield and maturation. Precision Agric, 2020.

LIN, T; DOLLÁR. P; GIRSHICKi, K; BELONGIE, S "Feature Pyramid Networks for Object Detection," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

Historical Data on the Global Coffee Trade, **International Coffee Organization, 2022.**

LIU, G.; NOUAZE, J.C.; TOUKO, M.P.L.; KIM, J.H. YOLO-Tomato: A Robust Algorithm for Tomato Detection Based on YOLOv3. *Sensors* 2020.

LIU, L; OUYANG, W; WANG, X; FIEGUTH, P; CHEN, J; LIU, X; PIETIKAINEM, M. Deep Learning for Generic Object Detection: **A Survey. International Journal of Computer Vision, 2020.**

LIN, TY. Microsoft COCO: Common Objects in Context. In: Fleet D. *Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8693. Springer, Cham, 2014.*

MARTINS, D. R.; CAMARGO, O. A. de; BACTACLIA, O. C. Qualidade do grão e da bebida em cafeeiros tratados com lodo de esgoto. **Bragantia, Campinas, v. 64, n. 1, p. 115-126, 2005.**

RAMOS, P.J; PRIETO, F.A; MONTOYA, F.C; OLIVEROS, C.F. Automatic fruit count on coffee branches using computer vision, **Computers and Electronics in Agriculture, Volume 137, 2017.**

REN, S; He, K.; GIRSHICK, R.; SUN, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. **IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017.**

RODRÍGUEZ, J. P., CORRALES, D. C.; AUBERTOT, J.N. A computer vision system for automatic cherry beans detection on coffee trees. **Pattern Recognition Letters 2020.**

SANTAQUITERIA, J.; BUENO, G.; DENIZ, O.; VALLEZ, N.; CRISTOBAL, G.; Semantic versus instance segmentation in microscopic algae detection, **Engineering Applications of Artificial Intelligence, Volume 87, 2020.**

WEIKUAN, J.; YUYU, T.; RON, L, ZHONGHUA, Z.; JIAN, L.; YUANJIE, Z. Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot. **Computers and Electronics in Agriculture, Volume 72, 2020.**

YANG, Y.; KAILIANG, Z.; LI, Y; DONGXING, Z. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN, **Computers and Electronics in Agriculture, Volume 163, 2019.**

3. CAPÍTULO 2

USO DE REDES NEURAIIS PARA CLASSIFICAÇÃO DE DOENÇAS DO CAFEIEIRO

RESUMO

Um dos desafios da cafeicultura se refere a minimizar o impacto das pragas e doenças nas lavouras. As doenças afetam diretamente a segurança alimentar e resultam em perdas de safra todos os anos. Na cafeicultura, entre as principais doenças que afetam as plantações estão a cercospora, ferrugem e a mancha de phoma. Sistemas de classificação envolvendo aprendizado de máquina são abordagens promissoras para apoiar o monitoramento automático das pragas e doenças nas lavouras de café. Diante disso, com esse trabalho se objetivou gerar um sistema de classificação supervisionada de doenças a partir de imagens digitais utilizando deep learning. Verificou-se que o primeiro modelo proposto de classificação de três classes reconheceu efetivamente os três diferentes tipos de doenças da folha do café. O segundo modelo proposto diferenciou de forma satisfatória plantas com e sem presença de pragas. Comparando-se os dois modelos, o modelo que classificou as duas classes (presença e ausência da praga) obteve valores maiores de f1-score de 93,33 % para as duas classes. Já no modelo de classificação das 3 doenças do cafeeiro obteve f1-score de 90%, 85% e 90% para as classes cercospora, ferrugem e mancha de phoma, respectivamente. Os resultados experimentais indicam que a abordagem proposta pode apoiar significativamente a detecção doenças foliares do café com pouco esforço computacional.

Palavras Chaves: Monitoramento de Culturas. Inteligência artificial

USE OF DEEP LEARNIG FOR COFFEE DISEASES CLASSIFICATION

ABSTRACT

One of the challenges of coffee farming is to minimize the impact of pests and diseases on crops. Diseases directly affect food security and result in crop losses every year. In coffee farming, among the main diseases that affect plantations are brown eye spot, rust and phoma leaf spot. Classification systems involving machine learning are promising approaches to support automatic monitoring of pests and diseases in coffee plantations. Therefore, this work aimed to generate a supervised disease classification system from digital images using deep learning. It was found that the first proposed three-class classification model effectively recognized the three different types of coffee leaf diseases namely. The second proposed model satisfactorily differentiated healthy and diseased plants. Comparing the two models, the model that classified the two classes (diseased plants and healthy plants) obtained higher f1-score values of 93.33% for both classes. In the classification model of the 3 coffee diseases, it obtained f1-score of 90%, 85% and 90% for the cercospora, rust and phoma stain classes, respectively. The experimental results indicate that the proposed approach can significantly support the detection of foliar diseases in coffee with little computational effort.

Keywords: Cultures, monitoring, artificial intelligence.

3.1. Introdução

Pelo menos 10% da produção global de alimentos é perdida em função das pragas nas plantações (Strange & Scott, 2005). As pragas estão causando perdas de produção entre 25 a 41% para arroz, 10 a 28% para trigo, 20 a 41% para milho, 11 a 32% para soja, 8 a 21% para batata, 20 a 50% para o café (Savary, et al., 2019). O monitoramento das lavouras é essencial para detecção precoce de pragas e consequente adoção de medidas de controle.

Na cultura do café algumas doenças como: ferrugem, cercospora e mancha de phoma geram perdas nas lavouras todos os anos podendo comprometer em alguns casos 50% da produção (Zambolim et al., 1997). A ferrugem do cafeeiro (*Coffea arabica* L.) é causada por *Hemileia vastatrix* Berk. & Br. (Zambolim et al., 1997). Tal doença, gera danos indiretos ao cafeeiro pela indução de desfolha na colheita. A queda da folhagem resulta em menor vingamento dos frutos chumbinhos e da florada e na seca dos ramos plagiotrópicos. De forma semelhante, a doença causada pelo fungo *Phoma* spp. ocorre em folhas, ramos, flores e frutos do cafeeiro e os sintomas mais comuns da doença nas folhas são manchas irregulares de coloração escura. Além disso, o patógeno pode colonizar os ramos do cafeeiro causando queda de flores e frutos. Outra doença que acomete o cafeeiro é a cercosporiose causada pelo fungo *Cercospora coffeicola* Berk. & Cook. A doença é associada ao cafeeiro que recebe incidência direta do sol e são expostos a estresses nutricionais e hídricos (Boldini et al., 2000; Juliatti et al., 2000a; 2000b).

Tradicionalmente, a avaliação da incidência das pragas do cafeeiro é determinada visualmente por especialistas. Trata-se de um processo cognitivo que está sujeito a erros e imprecisões (Barbedo, 2018). Ademais, considerando a vastidão das áreas plantadas, a inspeção visual envolve maior demanda de tempo e dinheiro em relação a outros meios de detecção (Singh et.al, 2017).

Com o aumento da disponibilidade de processadores de alto desempenho, foi possível aumentar a velocidade das operações computacionais (Singh et.al, 2019). O desenvolvimento da inteligência artificial, associado à disponibilidade de dados gerados por sensores digitais, tem gerado avanços em diversas áreas da ciência. Dentre os sensores destacam-se as câmeras digitais. Técnicas de análise e processamento de imagens podem ser utilizadas para automatizar tarefas que são realizadas pelo sistema visual humano (Huang, 1996).

As imagens obtidas por câmeras digitais estão sendo utilizadas para identificação das pragas. As técnicas mais comuns para identificação de pragas são: aquisição, pré-processamento, segmentação, extração de características (features) e classificação supervisionada.

A aquisição de imagens pode ser feita por meio de câmeras obtidas de celulares. Além disso, existem conjuntos de dados públicos com uma diversidade de plantas e doenças disponíveis para uso, um deles é o Plant Disease Dataset. Por outro lado, o pré-processamento visa corrigir defeitos, ruídos, variações de iluminação da imagem, além de realçar os detalhes (Gomes, 2011). Uma etapa comum do pré-processamento é a conversão de imagens RGB para algum outro espaço de cor, por exemplo: CIELAB, CIELUV, YCbCr, HSI, HSV. Já o processo de segmentação de imagens visa dividir uma imagem em partes distintas e separar grupos de pixels com características parecidas (Gonzalez, 2010). Usos comuns para segmentação consistem em separar a folha do fundo, bem como separar a área afetada da área saudável. Por fim, os algoritmos de classificação têm por objetivo prever qual classe determinada imagem pertence. Um dos algoritmos utilizados para classificação de imagens na detecção de pragas são as ANN (Redes Neurais artificiais). As ANN geram modelos matemáticos a partir de amostras previamente classificadas. O processo de aprendizado ocorre na fase de treinamento em que gradualmente os pesos da rede neural são ajustados visando reduzir o erro entre a classe predita e real (Psaltis, 1988). Um tipo de rede comum usado nas classificações de pragas são as CNN (Redes neurais convolucionais) (Li, 2019). A classificação a nível de imagem feita pelos modelos de CNNs é automática, dispensando a necessidade da intervenção humana no processo, tornando o processo escalável pois permite avaliar um número grande de imagens em um período relativamente curto.

A arquitetura do modelo CNN é composta por camadas, sendo elas: camada convolucional, camada de unidade linear retificada (ReLU), camada de conjugação (pooling), totalmente conectadas (fully connected). Para validar os modelos é frequente o uso de validação cruzada e métricas advindas da matriz de confusão como precisão, acurácia, F1-score e Recall.

Geralmente, no processo de detecção de doenças foliares, a validação é feita na etapa de classificação. As redes neurais do tipo CNN são uma classe de redes neurais de aprendizado profundo. Primeiramente recebe-se uma imagem de entrada, realiza-se a convolução na imagem por meio de filtros, os valores do filtro são

multiplicados pelos valores originais dos pixels da imagem. A CNN aprende os valores dos filtros durante o processo de treinamento. Os pesos são inicializados aleatoriamente, atualizando-os a cada nova entrada durante o processo de backpropagation. Um filtro pode ser aplicado para aumentar a nitidez e o foco de uma imagem ou borrar uma imagem, ou detectar bordas. Um dos parâmetros utilizados é a stride que indica quantos pixels você movimenta o filtro por vez. Já o padding é um processo em que alguns pixels são adicionados ao redor da imagem antes da operação de convolução, de forma a manter a dimensionalidade na imagem resultante durante a operação. A camada de MaxPooling retira o maior elemento de determinada região da matriz é um processo simples de redução da dimensionalidade gerando as features maps (Sardogan, 2018).

O objetivo deste trabalho foi implementar um modelo de classificação envolvendo redes neurais convolucionais para classificar as principais doenças do cafeeiro, a saber: ferrugem, cercospora e mancha de phoma. Além disso, pretendeu-se gerar um modelo de classificação de presença e ausência de doença.

3.2. Material e Métodos

O experimento foi conduzido em lavouras de café arábica no município de Viçosa, Minas Gerais, em setembro de 2020. A lavoura está localizada na quadrícula delimitada entre as latitudes $-20,45217^{\circ}$ e $-20,4579^{\circ}$, e as longitudes $-42,52258^{\circ}$ e $-42,5390^{\circ}$. A cultivar de café arábica instalada na área do experimento foi o catuaí vermelho, com espaçamento de 3 m na entrelinha e 0,7 m na linha de plantio, resultando em um estande de 4761 plantas por hectare. No local, foram obtidas 800 imagens das plantas do café com câmera de celular smartphone Samsung Galaxy A51 com câmera de resolução de 32.0 MP. As imagens foram obtidas sob variadas intensidades de iluminação. As doenças foliares classificadas neste estudo foram: ferrugem, cercospora e mancha de phoma conforme Figura 3.1.



Figura 3.1. Doenças foliares do café estudadas no trabalho da esquerda para direita: ferrugem, phoma, cercospora.

A partir das 800 imagens obtidas, separou-se quatro grupos de imagens representando as quatro classes de interesse, a saber, ferrugem, cercospora, phoma e plantas saudáveis. As imagens foram então utilizadas como entradas para geração de modelos com deep learning. Os modelos foram treinados sob a estrutura de desenvolvimento de aprendizado profundo do TensorFlow e Keras utilizando o Google Colaboratory (GC). Para isso, utilizou-se GPU do GC Tesla 12GB, processador Intel Xeon 2.2 GHz e 13GB de RAM.

O trabalho foi dividido em duas abordagens. Na primeira abordagem, buscou-se por gerar um modelo de classificação de doenças foliares dos cafeeiros. Na segunda abordagem, as classes adotadas no modelo foram: ausência de doença e presença de doença. Na primeira abordagem, 540 imagens foram utilizadas para treinamento e 60 imagens foram utilizadas para teste. Na segunda abordagem, 340 imagens foram utilizadas para treinamento e 60 foram utilizadas para teste. Nas duas abordagens a planta foi classificada visualmente por um especialista afim de obter as classes verdadeiras tanto para o modelo que visou classificar as imagens com bases nas 3 classes de doenças quanto para o modelo que visou classificar as imagens de forma binária (presença ou ausência) da doença, a classificação visual visou obter os rótulos necessários para se trabalhar com os modelos de classificação supervisionada.

Nas duas abordagens, as imagens no conjunto de dados selecionado foram redimensionadas para o tamanho de 300x200. Aplicou-se argumentação nas imagens. Argumentação é uma técnica que pode ser usada para expandir

artificialmente o tamanho de um conjunto de dados de treinamento, criando versões modificadas de imagens no conjunto de dados.

As técnicas utilizadas para aumentar o conjunto de dados foram: deslocamento (horizontal e vertical), inversão horizontal, rotação horária de 90 graus, alteração do brilho e do zoom das imagens. O deslocamento significa mover todos os pixels da imagem em uma direção (horizontal ou verticalmente) mantendo as dimensões da imagem. Na inversão as linhas ou colunas de pixels mudam de direção. A rotação utilizada girou a imagem no sentido horário em 90°. O zoom ampliou aleatoriamente as imagens. O brilho da imagem foi aumentado. A variação do brilho foi adotada para permitir a generalização do modelo que usou imagens treinadas em diferentes níveis de iluminação. Um exemplo destas técnicas utilizadas aplicadas sobre as imagens pode ser observado na Figura 3.2. As técnicas de aumento dos conjuntos de dados visaram melhorar o desempenho dos modelos de classificação.



Figura 3.2. Exemplo de imagem aplicando-se técnicas de aumento do conjunto de dados como: deslocamento, inversão, rotação, alteração do brilho, e zoom.

Após o pré-processamento das imagens, foram utilizadas redes neurais convolucionais (CNN) para realização da modelagem. Na arquitetura adotada, inicialmente foi aplicado um processo de convolução com 16 filtros de tamanho 3x3 com padding e função de ativação de unidade linear retificada. Pode-se entender esse processo como a geração de 16 novas imagens a partir da imagem de entrada. Na camada subsequente, foi aplicada uma operação de MaxPooling reduzindo o tamanho

da imagem para 150x100. Na terceira etapa, foi aplicado uma camada convolucional com 32 filtros de tamanho 3x3. Na quarta camada foi aplicada uma MaxPooling reduzindo o tamanho da imagem para 75x50. Na quinta etapa, foi aplicado uma camada convolucional com 64 filtros de tamanho 3x3 com padding e função de ativação de unidade linear retificada (relu). A sexta camada foi uma MaxPooling que reduziu a dimensionalidade da imagem para 37x25. Por fim, foram utilizadas mais três camadas: uma Flatten e duas Dense. Na Figura 3.3 são mostradas algumas imagens de saída da primeira camada convolucional gerada a partir de uma imagem da folha sintomática para ferrugem. Nota-se que, em algumas imagens o dano causado a folha praga da ferrugem do café foi realçado.

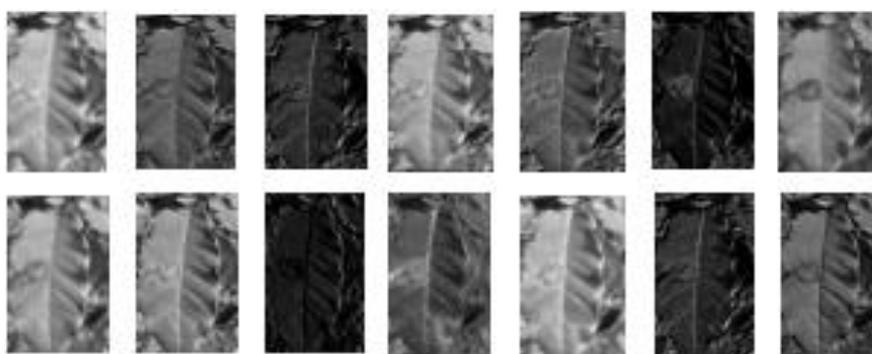


Figura 3.3. Saída da primeira camada convolucional, de forma que a convolução ocorreu sobre uma imagem de treino

Após o treinamento dos dois modelos, o desempenho nas imagens de teste foi avaliado, utilizando as métricas da matriz de confusão a saber, a precisão (P) e recall (R) e f1-score definidos nas equações (1) a (3).

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

$$F1 = \frac{2 * P * R}{P + R} \quad (3)$$

3.3. Resultados e Discussão

As curvas de acurácia obtidas no processo de treinamento do modelo são mostradas na Figura 3.4. O modelo convergiu para um acurácia de 99 % na curva de treinamento a partir da décima quinta iteração e atingiu erro de 0,2 na trigésima iteração. Os resultados da classificação das imagens de teste na primeira abordagem

considerando as classificações das doenças: cercospora, ferrugem e mancha de phoma são mostrados nas Tabelas 3.1

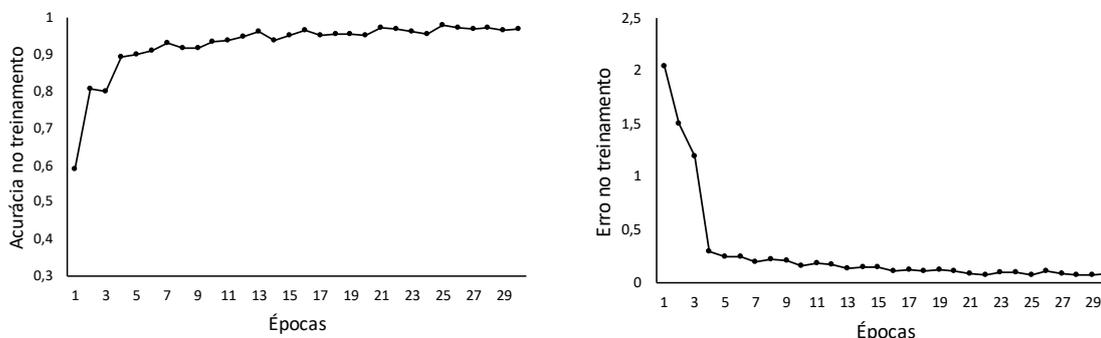


Figura 3.4. Gráfico obtido do treinamento das imagens na primeira abordagem, obtendo-se a curva de treinamento e validação com o passar do número de épocas Tabela 3.1. Matriz de confusão obtida na primeira abordagem para classificação das doenças mancha de phoma, ferrugem e cercospora.

Classe verdadeira	Classe Predita		
	Phoma	Ferrugem	Cercospora
Phoma	18	1	1
Ferrugem	2	17	1
Cercospora	0	2	18

Analisando-se os resultados apresentados na Tabela 3.1 verificou-se que a arquitetura proposta com redes neurais (CNN) alcançou uma precisão de 90%, 85% e 90% para as classes mancha de phoma, cercospora e ferrugem, respectivamente. Os valores de recall foram de 90 %, 85% e 90% para as classes mancha de phoma, ferrugem e cercospora, respectivamente. Problemas com falsos positivos e falsos negativos podem ter sido causados por reflexão especular da luz, pela coloração e formato semelhantes das manchas causadas pelas pragas, gerando interpretações errôneas pelo modelo. A classificação de pragas é um desafio pois espécies como cercospora e ferrugem são parecidas e sua classificação é difícil mesmo quando feita por especialistas. Contudo, a utilização de sistemas computacionais na resolução desse problema demonstrou-se confiável ao aplicar uma CNN especializada em problemas de classificação. De forma geral, a arquitetura proposta das redes neurais convolucionais classificou de forma satisfatória as pragas do cafeeiro estudadas mesmo considerando que as imagens

foram obtidas em ambientes não controlados, ou seja, em campo. Um fator que pode ter contribuído para o bom desempenho foi o uso de técnica para aumentar o conjunto de dados.

Para a classificação em que as classes foram: presença e ausência da doença os resultados são mostrados na Figura 3.5 e Tabelas 3.3. Na Figura 3.5 verificou-se que a curva de desempenho no treinamento convergiu na trigésima iteração alcançado na curva de treinamento. Verificou-se que o modelo foi eficiente na classificação das imagens em ambientes não controlados obtendo valores de 93,33 % de precisão para as classes saudáveis e doentes. Além disso, o modelo alcançou o mesmo valor de recall de 93,33% para plantas saudáveis e doentes. Portanto, o segundo modelo que classificou apenas a presença e a ausência da praga obteve melhor desempenho que o primeiro modelo do que o modelo que classificou os 3 tipos de pragas. Embora o trabalho tenha focado em um estudo para o cafeeiro, toda a metodologia pode ser aplicada de forma escalável a outras culturas para isso basta adaptar as saídas das redes para o número de classes desejadas.

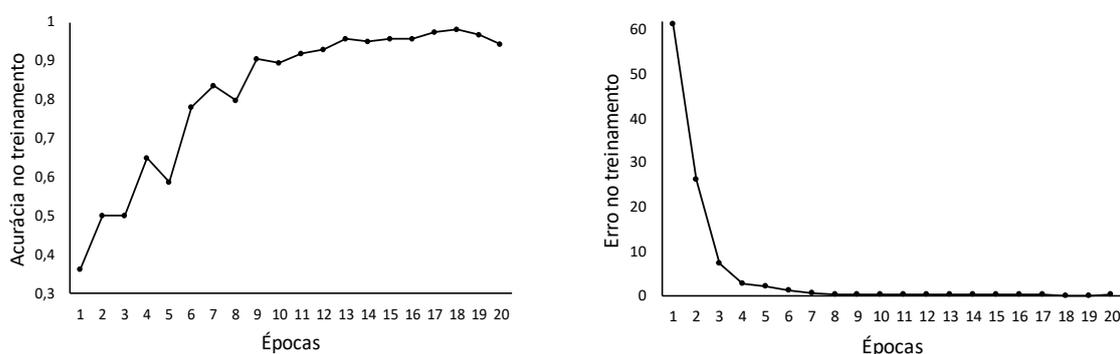


Figura 3.5. Gráfico obtido do treinamento das imagens na segunda abordagem, obtendo-se a curva de treinamento e validação com o passar do número de épocas.

Tabela 3.3. Matriz de confusão obtida na segunda abordagem para classificação de plantas doentes e saudáveis.

Classe verdadeira	Classe Predita	
	Ausência	Presença
Ausência	28	2
Presença	2	28

As duas abordagens de aprendizado profundo mostraram eficiência com a aplicação de técnicas de aumento do conjunto de dados. Na Figura 3.6 são mostrados alguns exemplos de predições que o modelo de classificação das três classes. O modelo confundiu ferrugem com mancha de phoma pela similaridade da coloração das folhas das plantas e confundiu cercospora com ferrugem pelos tons amarelados da doença. Já na Figura 3.7 são mostrados alguns exemplos de predição pelo modelo de classificação das duas classes. Tal modelo confundiu com maior frequência plantas doentes como saudáveis, provavelmente porque algumas plantas doentes apresentaram um tamanho pequeno de mancha característica em relação a área vegetativa. Estudos semelhantes abordaram outras arquiteturas de redes, outros métodos na identificação de pragas por exemplo por meio de segmentação semântica. A abordagem proposta aqui sugere que com um modelo mais simples baseado em aprendizado profundo pode sem muito esforço computacional identificar e classificar as pragas do café. Além disso, algoritmos assim podem ser utilizados em sistema para o monitoramento do cafeeiro auxiliando as tomadas de decisões dos agricultores.



Figura 3.6: Exemplos de predições realizadas pelo modelo de classificação das 3 classes (ferrugem, cercospora e phoma)

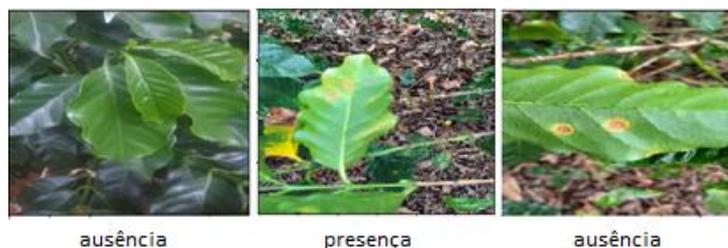


Figura 3.7: Exemplos de predições realizadas pelo modelo de classificação das 2 classes

3.4. Conclusão

Neste artigo, um método de classificação de doenças foliares do cafeeiro é apresentado com base em redes neurais convolucionais. Os experimentos foram realizados em imagens de plantas com e sem presença de praga no cafeeiro para realizar a classificação. Conclui-se que o método proposto reconhece efetivamente os três diferentes tipos de doenças da folha do café. E diferencia de forma satisfatória plantas com a presença e ausência de pragas. A classificação de pragas é um desafio pois muitas espécies são bastantes parecidas. Contudo, a utilização de sistemas computacionais na resolução desse problema demonstrou-se confiável ao aplicar uma CNN especializada em problemas de classificação. Comparando-se os dois modelos, o modelo que classificou as duas classes obteve valores maiores de f1-score. Obtendo f1-score de 93,3% para plantas com presença da praga e 93,3% para plantas com ausência da praga. Já no modelo de classificação das 3 pragas do cafeeiro obteve f1-score de 90%, 85% e 90% para as classes cercospora, ferrugem e mancha de phoma.

3.5. Referências

- BARBEDO, J.G. Factors influencing the use of deep learning for plant disease recognition. **Biosystems Engineering**, vol. 172, pp. 84 – 91,2018.
- DHAYGUDE. S.; KUMBHAR. N. "Agricultural plant Leaf Disease Detection Using Image Processing". **International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering** Vol. 2, Issue 1, January 2013
- GONZALEZ, R.E.; WOODS, R.C. **Digital image processing** (3rd edition), Mar. 2010.
- HUANG, T. E. C. Computer Vision: Evolution and Promise. 19th **CERN School of Computing**. Geneva: CERN. pp. 21–25 1996.
- KAMLESH, G.; BALASUNDRAM, S.; VADAMALAI, G.; BISWAJEET, P.A Review of neural networks in plant disease detection using hyperspectral data, **Information Processing in Agriculture**, Volume 5, 2018.
- PSALTIS, D.; SIDERIS, A.; YAMAMURA, A.A. A multilayered neural network controller, **IEEE Control Systems Magazine**, vol. 8, no. 2, pp.17–21, April 1988.

SARDOGAN, M.; TUNCER, A; OZEN, Y. Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm. 2018 3rd **International Conference on Computer Science and Engineering** (UBMK), 2018.

SAVARY, S.; WILLOCQUET, L.; PETHYBRIDGE, S.J. The global burden of pathogens and pests on major food crops, 3, 2019.

SINGH, U.P; CHOUHAN, S.S.; JAIN, S.; Multilayer convolutional neural network for the classification of mango leaves infected by anthrac nose disease, **IEEE Access**, vol. 7, pp. 43 721–43 729, 2019.

VIJAI, S.; Misra, M.A, Detection of plant leaf diseases using image segmentation and soft computing techniques, **Information Processing in Agriculture**, Volume 4, 2017.

4. CONCLUSÃO GERAL

De modo geral, pelos resultados obtidos neste trabalho, pôde-se concluir que:

Os produtos desenvolvidos nesse trabalho têm potencial para serem integrados à outras tecnologias como aplicativos de celular de forma a facilitar seu uso em campo. Tecnologias de baixo custo envolvendo dispositivos móveis e algoritmos gerados a partir de software livre são ferramentas acessíveis aos agricultores que podem auxiliar no monitoramento das lavouras. Os produtos desenvolvidos neste trabalho são promissores para promover um monitoramento automatizado das lavouras de café trazendo benefícios para a cafeicultura. Em pesquisas futuras, pode-se utilizar uma base de dados maior para que possamos melhorar os treinamentos das redes e um estudo mais detalhado sobre as diferentes arquiteturas de redes neurais convolucionais e até outros modelos e técnicas. É necessário um estudo mais aprofundado da arquitetura da planta e suas particularidades para que os modelos possam ser disponibilizados para uso em campo. A evolução deste trabalho pode envolver a disponibilização dos algoritmos em aplicativos para serem de fato usados em auxílio às tomadas de decisões nas lavouras.