

UNIVERSIDADE FEDERAL DO PARANÁ

LEONARDO JOSOÉ BIFFI

CONTRIBUIÇÕES NA DETECÇÃO DE FRUTOS DE MAÇÃ EM POMAR DE ALTA  
DENSIDADE UTILIZANDO IMAGENS OBTIDAS PARA FOTOGRAMETRIA  
TERRESTRE

CURITIBA  
2021

LEONARDO JOSOÉ BIFFI

CONTRIBUIÇÕES NA DETECÇÃO DE FRUTOS DE MAÇÃ EM POMAR DE ALTA  
DENSIDADE UTILIZANDO IMAGENS OBTIDAS PARA FOTOGRAMETRIA  
TERRESTRE

Tese apresentada ao Curso de Pós-Graduação em Ciências Geodésicas, Setor de Ciências da Terra da Universidade Federal do Paraná, como requisito parcial à obtenção do título de Doutor em Ciências Geodésicas.

Orientador: Prof. Dr. Edson Aparecido Mitishita

CURITIBA  
2021

Catálogo na Fonte: Sistema de Bibliotecas, UFPR  
Biblioteca de Ciência e Tecnologia

---

B591c Biffi, Leonardo Josué  
Contribuições na detecção de frutos de maçã em pomar de alta densidade utilizando imagens obtidas para fotogrametria terrestre [recurso eletrônico] / Leonardo Josué Biffi – Curitiba, 2021.

Tese - Universidade Federal do Paraná, Setor de Ciências da Terra, Programa de Pós-graduação em Ciências Geodésicas.

Orientador: Prof. Dr. Edson Aparecido Mitishita

1. Fotogrametria. 2. Agricultura. 3. Maçã - Poda I. Universidade Federal do Paraná. II. Mitishita, Edson Aparecido. III. Título.

. CDD: 526.982

---

Bibliotecária: Roseny Rivelini Morciani CRB-9/1585



MINISTÉRIO DA EDUCAÇÃO  
SETOR DE CIÊNCIAS DA TERRA  
UNIVERSIDADE FEDERAL DO PARANÁ  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO CIÊNCIAS  
GEODÉSICAS - 40001016002P6

## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em CIÊNCIAS GEODÉSICAS da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **LEONARDO JOSOE BIFFI** intitulada: **CONTRIBUIÇÕES NA DETECÇÃO DE FRUTOS DE MAÇÃ EM POMAR DE ALTA DENSIDADE UTILIZANDO IMAGENS OBTIDAS PARA FOTOGRAMETRIA TERRESTRE**, sob orientação do Prof. Dr. EDSON APARECIDO MITISHITA, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa. A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 29 de Julho de 2021.

Assinatura Eletrônica

30/07/2021 11:40:47.0

EDSON APARECIDO MITISHITA  
Presidente da Banca Examinadora

Assinatura Eletrônica

30/07/2021 13:21:44.0

JORGE ANTONIO SILVA CENTENO  
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

30/07/2021 13:16:07.0

VERALDO LIESENBERG  
Avaliador Externo (UNIVERSIDADE DO ESTADO DE SANTA  
CATARINA)

Assinatura Eletrônica

02/08/2021 10:45:42.0

LEO RUFATO  
Avaliador Externo (UNIVERSIDADE DO ESTADO DE SANTA  
CATARINA)

---

Centro Politécnico - Caixa Postal 19001 - CURITIBA - Paraná - Brasil

CEP 81531-980 - Tel: (41) 3361-3153 - E-mail: cpgcg@ufpr.br

Documento assinado eletronicamente de acordo com o disposto na legislação federal Decreto 8539 de 08 de outubro de 2015.

Gerado e autenticado pelo SIGA-UFPR, com a seguinte identificação única: 103688

Para autenticar este documento/assinatura, acesse <https://www.prppg.ufpr.br/siga/visitante/autenticacaoassinaturas.jsp>  
e insira o código 103688

Dedico à minha Esposa Claudia e minha Filha Isabela.

E aos meus Pais.

## **AGRADECIMENTOS**

Agradeço a DEUS por ser a luz presente em meu caminho e sempre me guiar para o esclarecer do autoconhecimento.

Agradeço do fundo do coração a minha filha Isabela e minha esposa Claudia pela compreensão, paciência e apoio ao longo desses anos, por estarem sempre ao meu lado, principalmente nas dificuldades.

A minha família, Pai, Mãe, Adriano, meu sogro e minha sogra, que muito me fiz ausente em eventos familiares, mas que todos compreenderam.

Ao meu Irmão Olavo e Milena e Miguel que me acolheram em Curitiba e até mudança de residência fizeram para que eu ficasse mais próximo a UFPR. Muito obrigado meu Irmão.

Ao meu Professor Orientador Edson Mitishita, pela orientação ao longo desses anos, apoio, compreensão, amizade e paciência.

Obrigado a Universidade do Estado de Santa Catarina – UDESC por conceder a licença para realizar o trabalho.

Ao Professor Centeno pelos auxílios nos primeiros trabalhos que originaram o primeiro artigo dessa tese.

Aos Professores da Universidade Federal do Paraná que tive aula, que contribuíram para minha formação.

Um abraço especial ao Professor Veraldo, que me acompanhou desde o meu ingresso até a finalização do trabalho e teve participação na elaboração dos artigos aqui apresentados, meu muito obrigado.

Obrigado pelas amizades que fiz nesse tempo no Labfoto: Marlo, Niarkios, Filipe, Júlio e Nathalia.

Agradeço ao técnico agrícola Renato que sempre me recebeu muito bem nas visitas de campo no pomar.

Agradecimento a equipe de trabalho dos Professores José Marcato Jr e Wesley Nunes Gonçalves, da Universidade Federal do Mato Grosso do Sul – UFMS, que proporcionaram a análise dos dados mesmo na decorrência da pandemia da Covid-19.

A equipe do Laboratório de Fruticultura e ao Professor Leo pelo apoio na pesagem dos frutos colhidos.

## RESUMO

A maleicultura brasileira faz uso de operações manuais para fazer avaliação de produção de pomar. Essas operações são morosas, despendem de esforço físico e necessitam no mínimo dois operadores para serem realizadas. Partindo desse problema propôs-se utilizar técnicas de Visão Computacionais e de Fotogrametria para obtenção de informações quantitativas e individualizadas de plantas de um pomar comercial de maçã. O estudo foi realizado em um pomar no município de Correia Pinto-SC, a cultivar analisada foi a Fuji Suprema, próximo ao período da colheita. Foram analisadas 8 linhas, identificadas de L1 a L8 com 10 plantas cada. A coleta de dados dividiu-se em duas etapas: 1) A Primeira etapa foi a aquisição das imagens que permitiram as análises realizadas por Fotogrametria; nesta etapa foram fotografadas 8 linhas em ambos os lados, e com sobreposição entre as fotos. 2) a Segunda foi a aquisição dos dados do pomar de modo convencional onde foram coletados: Diâmetro do Caule ( $Di$ ), Altura da planta ( $H$ ), Comprimento ( $Co$ ), Largura ( $Lg$ ), Número dos Frutos ( $NF$ ) e Peso Médio do Fruto ( $PMF$ ). Duas metodologias (A) e (B) de detecção de frutos de maçã foram desenvolvidas, empregando de forma independente técnicas de Processamento Digital de Imagem ( $PDI$ ) e de Aprendizado Profundo – *Deep Learning* ( $DL$ ). A metodologia (A) usou técnica do espaço de cores para detectar os frutos da maçã em uma Linha, com 20 imagens. Como avaliação calculou-se valores de comissão, omissão e acurácia para as detecções dos frutos. As bandas que apresentaram melhor desempenho foram *Lab-a*, Combinações C2 e C3, com valores aproximados de 0,06, 0,25 e 0,76 para comissão, omissão e acurácia, respectivamente. Na contagem de frutos proposta *Lab-a* apresentou um  $R^2$  de 0,73 e na comparação com o acerto na posição dos frutos o  $R^2$  foi de 0,96. A metodologia (B) explorou sete modelos diferentes de DL para realizar a detecção dos frutos. Os modelos de DL avaliados foram *ATSS*, *Faster-RCNN*, *Libra-RCNN*, *Cascade-RCNN*, *RetinaNet*, *FSAF* e *HRNet*. Das experimentações realizadas concluiu-se que os melhores resultados foram obtidos com *ATSS*, *HRNet* e *FSAF*, resultado Precisões Médias de 0,925, 0,922 e 0,922, respectivamente. Empregando o modelo *ATSS* foram efetuadas experimentações visando explorar diferentes tamanhos de caixas delimitadoras (BB), diferentes níveis de densidade de frutos por *patches*, e corrupção da imagem. Os melhores resultados foram: na BB o tamanho de 160 x 160 pixel, na densidade de frutos o valor de (20-29) frutos, e nas imagens corrompidas o método teve bom desempenho exceto para condições simuladas de neve, geada e névoa.

Palavras-chave: *Malus domestica* Borkh, detecção de objetos, agricultura de precisão.

## ABSTRACT

The Brazilian apple crop has used manual operations to evaluate its orchard production. These operations are time-consuming, require physical effort, and need at least two operators to be performed. Considering these difficulties, a study was performed to develop semi-automatic approaches to obtain quantitative and individualized information of plants from a commercial apple orchard using Computer Vision and Photogrammetry techniques. The study area was carried out in an orchard at Correia Pinto-SC; Fuji Suprema was the cultivar analyzed near to the harvest period. Eight rows with ten apple plants in each row, identified as L1 to L8, were analyzed. Data collection was divided into two steps: 1) The first step was the acquisition of images that allowed the analysis performed by Photogrammetry; In this step, 8 rows were photographed on both sides. A large overlap between the photos was used, resulting 20 images in each line. 2) In the second step, the basic dataset from the orchard was acquired using a conventional procedure; the following data were collected: Trunk Diameter ( $Di$ ), Plant Height ( $H$ ), Length ( $Co$ ), Width ( $Lg$ ), Number of Fruits ( $NF$ ) and Weight Fruit Medium ( $FMP$ ). Two methodologies (A) and (B) for apple fruit detection were developed using Digital Image Processing (PDI) and Deep Learning ( $DL$ ) techniques. Methodology (A) used the color space technique to detect apple fruits in a line. Commission, omission, and accuracy were the parameters' values calculated to measure the performance of the approaches used for apple detection. The bands that showed the best performance were *Lab-a*, Combinations C2, and C3; The obtained results from *Lab-a* were values near 0.06, 0.25, and 0.76 for commission, omission, and accuracy, respectively. Additionally, the obtained value of  $R^2$  was 0.73 and in the comparison with the correct position of the fruits, the  $R^2$  was 0.96. Methodology (B) explored seven different DL models to perform the process of fruit detection. The DL models evaluated were *ATSS*, *Faster-RCNN*, *Libra-RCNN*, *Cascade-RCNN*, *RetinaNet*, *FSAF*, and *HRNet*. From the experiments carried out, it can be concluded that the best results were obtained with *ATSS*, *HRNet*, and *FSAF*, resulting in Average Precision of 0.925, 0.922, and 0.922, respectively. Using the *ATSS* model, experiments were carried out to explore different sizes of bounding boxes ( $BB$ ), different levels of fruit density by patches, and image corruption. The best results were: in  $BB$  the size of 160 x 160 pixels, in the density of fruits the value of (20-29) fruits, and in the corrupted images the method performed well except for simulated conditions of snow, frost, and fog.

Keywords: *Malus domestica* Borkh, object detection, precision agriculture

## LISTA DE FIGURAS

Figura 1 – Ciclo anual da cultura da Macieira. ....	19
Figura 2 – Imagem da Cultivar Fuji Suprema (Correia Pinto, SC). ....	21
Figura 3 – Comparativo de exemplos de variáveis possíveis de serem analisadas com a referida escala da planta. ....	23
Figura 4 – Espectro Eletromagnético com destaque para espectro visível ....	26
Figura 5 – Representa o Cubo do espaço de cores RGB ....	27
Figura 6 – Representa o Hexágono do modelo HSV. ....	28
Figura 7 - Diagrama de Cromaticidade e Representação 3D do Espaço de Cor CIE XYZ. ....	30
Figura 8 – Definição de IoU na função de detectar objetos usando anotações de caixa delimitadora. ....	33
Figura 9 – Interpretação de precisão e recall para a função de detectar objetos. ....	34
Figura 10 – Princípio da dinâmica do SfM. ....	35
Figura 11 - (A) Brasil; (B) Estado de Santa Catarina; (C) Município de Correia Pinto e localização do pomar; (D) Vista aérea do local do estudo e identificação das linhas amostradas de Fuji Suprema (L1 a L8), e linhas de Gala (LA e LB); (E) perspectiva da aquisição de imagens no pomar durante o levantamento de campo. ....	38
Figura 12 – Modo ilustrado de obtenção das imagens no pomar: com a câmera perpendicular a linha de plantio (A) e com a câmera oblíqua a linha de plantio (B e C). As sobreposições das imagens também são apresentadas. ....	39
Figura 13 – Medindo a circunferência do caule para obtenção do diâmetro ( $D_i$ ) (A). Detalhe da operação (B). ....	41
Figura 14 – Exemplo de medição de altura de planta no pomar. Operador segurando a régua (A) e detalhe do desafio para fazer a leitura (B). ....	42
Figura 15 – Medidas de Largura (A) e Comprimento (B) da planta no pomar. ....	42
Figura 16 – Exemplo de contador de manual. ....	43
Figura 17 – Fluxograma da metodologia adotada para detecção de frutas de baixo custo. ....	46
Figura 18 – Esquema gráfico de vista superior da aquisição de imagens e movimento da câmera. ....	47

Figura 19 – Representação da remoção de ruído de imagem binária, aplicando operação morfológica de opening and closing, utilizando círculo como elemento estruturante. ....	51
Figura 20 – Imagem utilizada como exemplo para os resultados de realce para os frutos, imagem binária e remoção de ruídos. ....	53
Figura 21 – Quadro dos resultados das transformações do espaço de cores, imagem binária e remoção de ruídos. Bandas apresentadas: <i>rgb-r</i> , <i>HSV-h</i> e <i>L*a*b*-a</i> . ....	54
Figura 22 - Quadro dos resultados das transformações dos espaços de cores Otha, imagem binária e remoção de ruídos. Bandas apresentadas: $I''_2$ e $I''_3$ . ....	55
Figura 23 – Gráfico de comparação da média geral das imagens de Comissão, omissão e precisão por método de detecção. ....	57
Figura 24 – Contagem visual de frutos nas imagens RGB original (A) e resultados das detecções com o algoritmo em uma imagem resultante <i>Lab-a</i> (B). ....	57
Figura 25 – Regressões entre número de frutos contados visualmente na imagem RGB e: o número de objetos detectados no algoritmo nas imagens <i>Lab-a</i> (A); o número de acertos do algoritmo (B). ....	60
Figura 26 - Frutos aglomerados (A) e consideradas como um único objeto por terem as bordas unidas (B) ....	60
Figura 27 - Modo de obtenção das imagens no Pomar. As sobreposições das imagens também são apresentadas. ....	67
Figura 28 - O fluxo de trabalho para detecção de maçã está dividido em três etapas principais: (1) processo de anotação, (2) distribuição de <i>patches</i> nas imagens de treinamento, validação e teste, e comparação dos métodos aplicados e (3) processos de corrupção de imagem. ....	68
Figura 29 – Exemplo de duas imagens e sua divisão em <i>patches</i> sem sobreposição, com dimensões $w_p \times h_p$ . ....	72
Figura 30 – Subconjuntos de uma determinada imagem RGB mostrando alguns exemplos de rotulagem manual com pontos no centro dos frutos da maçã (a) e estimativa de caixa delimitadora com tamanho fixo para cada rótulo de ponto (b). ..	72
Figura 31 - Exemplos de detecção de maçã usando: (a) ATSS, (b) Cascade-RCNN, (c) Faster-RCNN, e (d) FSAF. ....	77
Figura 32 - Exemplos de detecção de maçã usando: (a) HRNet, (b) Libra-RCNN, e (c) RetinaNet. ....	78
Figura 33 – Exemplos de <i>patches</i> sem maçãs. ....	81

Figura 34 – Exemplo de dois patches diferentes (A e B) com névoa (ns = nível de severidade).....	84
Figura 35 - Apresentação de Nuvem de pontos referente as plantas da Linha L4, gerada por Fotogrametria.....	89
Figura 36 – Seção do caule para efetivar a leitura do diâmetro (Di). ....	89
Figura 37 – Caule individualizado em nuvem de pontos (A); Corte de faixa de transversal a 20 cm do ponto de enxertia (B); e medidas de diâmetro realizadas em duas direções (C).....	90
Figura 38 – Nuvem segmentada em perfil vertical com espaçamento de 0,8 m (A); e delimitação de uma borda simplificada da nuvem (B).....	90

## LISTA DE TABELAS

Tabela 1 – Informação de data, e horários de início e fim das aquisições das imagens nas linhas de plantio. ....	40
Tabela 2 – Valores das amostras após aquisição das imagens e contagem visual dos frutos.....	52
Tabela 3 – Valores de médias de contagem visual dos frutos, contagem automática, comissão, omissão e valores de acurácia para as 20 imagens e para as imagens dos lados W e E, separados. ....	56
Tabela 4 – Conjuntos das linhas de plantio para treinamento, validação e teste. ....	73
Tabela 5 – Hiperparâmetro de cada método de detecção. ....	74
Tabela 6 – Precisão Média ( <i>PM</i> ) dos métodos de detecção de frutos de maçã em três linhas de plantio do conjunto de teste. ....	76
Tabela 7 – Resultado da Precisão Média ( <i>PM</i> ) para diferentes tamanhos de caixas delimitadoras (BB) obtidas pelo ATSS. ....	80
Tabela 8 – Resultado do método ATSS em diferentes densidades de frutos de maçã por <i>patches</i> .....	81
Tabela 9 – Precisão Média para diferentes ruídos e borrões em três níveis de severidade obtidos por método ATSS.....	82
Tabela 10 – Precisão Média para diferentes condições de clima e processamento digital de imagem em três níveis de severidade obtidos por método ATSS.....	83
Tabela 11 – Valores médios por planta de dados coletados em campo de forma convencional, nas linhas amostradas do pomar. Diâmetro (m), Altura (m), Comprimento (m), Largura (m), Número de Frutos. ....	87
Tabela 12 - Valores médios por planta de variáveis determinadas com os dados coletados em campo de forma convencional, nas linhas amostradas do pomar. Altura corrigida (m), Volume (m <sup>3</sup> ), Peso médio por fruto (g), Produção por planta (kg), Produtividade (Mg.ha <sup>-1</sup> ), Área da Seção transversal do caule (cm <sup>2</sup> ) e Eficiência Produtiva (kg.cm <sup>-2</sup> ). ....	88
Tabela 13 – Dados de plantas da linha L4, obtidas pelo método convencional e por Fotogrametria. Os dados apresentados são: Número de Frutos, Diâmetro do caule (m), Volume (m <sup>3</sup> ), Área da Seção transversal do caule (cm <sup>2</sup> ) e Eficiência Produtiva (kg.cm <sup>-2</sup> ) .....	91

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>14</b>
1.1	JUSTIFICATIVA: .....	16
1.2	OBJETIVOS: .....	16
1.2.1	Geral:.....	17
1.2.2	Específicos: .....	17
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>18</b>
2.1	A CULTURA DA MAÇÃ ( <i>Malus domestica</i> BORKH.).....	18
2.1.1	Cultivar de Maçã.....	20
2.1.1.1	Cultivar Fuji.....	20
2.1.1.1.1	Cultivar Fuji Suprema:.....	21
2.1.2	Variáveis para avaliação de pomar.....	22
2.2	TÉCNICAS DETECÇÃO E LOCALIZAÇÃO DE FRUTOS .....	24
2.2.1	ESPAÇO DE CORES .....	25
2.2.1.1	Cor .....	25
2.2.1.2	Espaço RGB .....	26
2.2.1.3	Espaço HSV .....	27
2.2.1.4	Espaço CIELAB .....	29
2.2.1.5	Espaço OHTA .....	31
2.2.2	APRENDIZADO PROFUNDO ( <i>DEEP LEARNING</i> ).....	32
2.3	STRUCTURE FROM MOTION - SfM.....	34
<b>3</b>	<b>METODOLOGIA GERAL .....</b>	<b>37</b>
3.1	CARACTERIZAÇÃO E LOCAL DO ESTUDO REALIZADO.....	37
3.2	MODO DE AQUISIÇÃO DAS IMAGENS.....	39
3.3	AQUISIÇÃO DOS DADOS DE MODO CONVENCIONAL .....	40
3.3.1	Diâmetro do Caule ( <i>Di</i> ).....	40
3.3.2	Altura da planta ( <i>H</i> ).....	41
3.3.3	Volume da planta ( <i>Vol</i> ). .....	41
3.3.4	Contagem dos frutos .....	43
3.3.5	Produção e Eficiência Produtiva.....	43
<b>4</b>	<b>METODOLOGIA A:.....</b>	<b>45</b>
4.1	<b>ARTIGO 1: AVALIAÇÃO DO DESEMPENHO DE DETECÇÃO SEMI-AUTOMÁTICA DE FRUTOS DE MAÇÃ EM UM POMAR DE ALTA</b>	

	<b>DENSIDADE USANDO SENSOR DE IMAGEM DIGITAL RGB DE BAIXO CUSTO.....</b>	<b>45</b>
4.1.1	Método para detecção de frutos .....	45
4.1.1.1	Aquisição de imagens .....	46
4.1.1.2	Transformação de Espaço de Cores. ....	47
4.1.1.3	Segmentação da imagem e binarização. ....	49
4.1.1.4	Combinação de bandas .....	50
4.1.1.5	Remoção de ruídos da imagem .....	50
4.1.1.6	Análise estatística .....	51
4.1.2	Resultados.....	52
4.1.3	Discussão .....	58
4.1.4	Conclusão e Recomendações .....	60
<b>5</b>	<b>METODOLOGIA B:.....</b>	<b>62</b>
5.1	<b>ARTIGO 2: ABORDAGEM BASEADA EM DEEP LEARNING ATSS PARA DETECTAR FRUTOS DE MAÇÃ. ....</b>	<b>62</b>
5.1.1	Introdução.....	62
5.1.2	Material e Métodos .....	66
5.1.2.1	Aquisição de dados .....	66
5.1.2.2	Método de Detecção de Objetos .....	67
5.1.2.3	Proposta sugerida .....	71
5.1.2.4	Configurações do Experimento .....	73
5.1.3	Resultados e Discussão .....	75
5.1.3.1	Desempenho de detecção de maçã usando diferentes métodos de detecção de objetos. ....	75
5.1.3.2	Influência do tamanho da caixa delimitadora (Bounding Box) .....	79
5.1.3.3	Detecção em níveis de densidade .....	80
5.1.3.4	Robustez contra Corrupções das imagens .....	81
5.1.4	Perspectivas de pesquisas adicionais .....	84
5.1.5	Conclusão e Recomendações .....	86
<b>6</b>	<b>METODOLOGIA C.....</b>	<b>87</b>
6.1	<b>DADOS CONVENCIONAIS VS FOTOGRAMETRIA.....</b>	<b>87</b>
6.1.1	Dados Convencionais.....	87
6.1.2	Fotogrametria .....	88
6.1.2.1	Variáveis determinadas .....	89

6.1.2.1.1	Diâmetro do Caule .....	89
6.1.2.1.2	Volume .....	90
6.1.3	Comparação dos resultados .....	91
<b>7</b>	<b>CONSIDERAÇÕES FINAIS .....</b>	<b>93</b>
<b>8</b>	<b>REFERÊNCIAS.....</b>	<b>94</b>

## 1 INTRODUÇÃO

A produção agrícola sempre foi um requisito importante para a balança comercial brasileira, representou 6,8% do Produto Interno Bruto (PIB) em 2020 (IBGE, 2021), e vem crescendo em termos de produtividade graças a novas tecnologias voltadas para maximizar o rendimento do setor. Neste sentido, o conceito de automação de processos torna-se cada vez mais comum em todo setor agrícola, com a utilização de implementos e equipamentos agrícolas, com sensores embarcados. Estes sensores podem utilizar técnicas de processamento digital de imagens terrestres e aéreas.

O setor da fruticultura é um cliente em potencial de geotecnologias, pois as culturas possuem um alto valor agregado em seus produtos. Para isso investem em inovação e tecnologias para obter mais informações dos sítios de produção e tomarem as decisões corretas para maior eficiência nos manejos.

As informações qualitativas e quantitativas precisas, relacionadas à planta, que permitam ao técnico ter o diagnóstico correto do sítio de produção são fundamentais nos dias atuais. As geotecnologias de levantamentos vêm propiciando um ganho nesses setores, pois cada vez apresentam melhores resoluções e precisões nas informações coletadas. Um dos desafios da Agricultura de Precisão é coletar a informação por indivíduo, onde a decisão seria individualizada e não regionalizada, como era o conceito inicial (STAFFORD, 2000).

Para o setor frutícola se manter competitivo, a coleta de informações necessita de técnicas ágeis e de baixo custo, porém com confiança na precisão da informação coletada.

A Fotogrametria está entre as ciências que viabilizam técnicas de coleta de informações quantitativas existentes que recentemente, com as facilidades na obtenção de imagens digitais por sensores RGB simples ou hiperespectrais, tem sido empregada com maior frequência em diversas áreas da engenharia. O aumento da capacidade computacional em “hardware/software” tem permitido o uso de câmeras digitais de baixo custo para extração automática de informações métricas de objetos na superfície terrestre.

Aliado a técnicas de Processamento Digital de Imagens (PDI) e de modelagem tridimensional de objetos, alguns trabalhos têm apresentados avanços na coleta de dados na fruticultura. Tais como estimativa de produção e contagem de

frutos (DORJ et al., 2017). Gongal et al. (2015) citam em sua revisão o uso de sensores como câmeras preto e branco, coloridas e termais, e métodos de classificação na detecção de frutos. Herrero-Huerta et al. (2015) apresenta um método de reconstrução tridimensional de bagas de uva vinífera para cálculo de volume de cachos.

A automação na coleta dos dados traz benefícios para os produtores e gestores de áreas frutícolas, pois permite levantar dados amostrais extensos, mais rápidos e com menor custo, e com isso ter o diagnóstico com base em mais pontos observados, viabilizando analisar uma área maior em relação aos métodos tradicionais. Métodos estes, que além de morosos, em sua maioria necessitam de mais de uma pessoa *in loco* para executá-los.

Outros setores que prestam serviços para a fruticultura também podem se beneficiar da agilidade de coleta de dados para estimar a produção. Como o setor de Seguros Agrícolas, que em épocas de ocorrência de sinistros, (como granizos, geadas, ventos fortes e fogo) geralmente atingem áreas expressivas, e devem ser periciados com máxima urgência. Um diagnóstico correto da área atingida e estimativa da perda de produção no local deve ser obtido para o correto ressarcimento do beneficiário.

O uso de câmeras de baixo custo pode ser uma solução para obtenção de dados métricos de maneira rápida e barata na fruticultura. Trabalhos científicos recentes têm indicados esses benefícios nas áreas de visão robótica (AN et al., 2017 e GONGAL et al., 2016), detecção de frutos (FONT et al., 2014; LIU et al., 2016; WEI et al., 2014), colheitas automáticas (FONT, DAVINIA et al., 2014 e TIJSKENS et al., 2016) e previsão de produção (BARGOTI; UNDERWOOD, 2017 e DORJ et al., 2017).

Desde ao final do ano 2000, com o desenvolvimento acelerado de técnicas de correspondências na visão 3D e a evolução dos computadores fez com que nuvens de pontos oriundas de imagens se torne uma fonte de dados para a obtenção de informações métricas na fruticultura.

Desta forma, empregando as técnicas de PDI e da Fotogrametria, imagens digitais terrestres não métricas, obtidas com câmeras de baixo custo, associadas com levantamentos de campo tradicionais, esta pesquisa de doutorado delineou estudos e desenvolvimentos de metodologias automáticas de detecção de objetos que permitam a obtenção de informações individualizadas de plantas em um pomar de maçã, como métricas de produção (números de frutos, produção por planta, eficiência produtiva) e

de dados geométricos das plantas (altura, diâmetro do caule, área da seção transversal do caule, volume da copa).

### 1.1 JUSTIFICATIVA:

A aplicação de técnicas de baixo custo para detecção de frutos de maçã em pomares de alta densidade é uma necessidade nos sistemas de produção brasileiros. O uso de técnicas de detecção automática de frutos exige normalmente conhecimento técnico e recursos computacionais avançados, nem sempre disponíveis na realidade do produtor brasileiro. Tanto grandes produtores como pequenos produtores de maçãs adotam técnicas avançadas de cultivo que colocam Santa Catarina como o principal produtor no Brasil. No entanto, apesar dos avanços agronômicos neste campo, percebe-se uma carência de tecnologias avançadas na quantificação da produção ainda no pomar, como informação para o processo de gestão.

A situação econômica e social, especialmente dos pequenos e médios produtores, abre um campo de pesquisa e desenvolvimento de meios de baixo custo que facilitem o trabalho. Por outro lado, os grandes produtores e empresas agrícolas também carecem destas soluções. Esses produtores possuem grandes áreas cultivadas e o método convencional de levantar dados para estimativas de produção e da capacidade produtiva se tornam muito onerosos e sujeitos ao erro humano. Este método faz algumas amostragens, que após avaliação é extrapolado para a área.

Nesse sentido, a automação de processos agrícolas vem ao encontro desta realidade no campo, cujo conceito é um sistema de produção no qual os processos operacionais são monitorados, controlados e executados por meio de máquinas e ou dispositivos mecânicos, eletrônicos ou computacionais, para ampliar a capacidade de trabalho humano. Visam-se sempre o aumento da produtividade e do trabalho, otimizando tempo e recursos, reduzindo perdas na produção, aumentando a qualidade dos produtos e consequentemente beneficiando a qualidade de vida dos trabalhadores (EMBRAPA, 2018).

O uso de técnicas de Visão Computacional e da Fotogrametria é cada vez maior nas Ciências Agrárias, tais como o setor florestal (UZQUIANO et al., 2014), a agricultura de grãos (CHANG et al., 2017 e MESAS-CARRASCOSA et al. 2015) e cana de açúcar (MORIYA et al., 2017 e LUNA; LOBO, 2016). De forma geral, a fruticultura brasileira tem se envolvido pouco com essas tecnologias. Uma busca com

os termos técnicos em bases de dados nacionais indicam poucas publicações (LARA et al., 2021).

Neste sentido, este trabalho busca explorar técnicas de detecção de frutos de maçãs no pomar e informações individualizadas das plantas, a partir de técnicas já estabelecidas. A principal contribuição deste trabalho está em explorar a viabilidade de aplicação de tais técnicas na maleicultura catarinense.

## 1.2 OBJETIVOS:

### 1.2.1 Geral:

Desenvolver metodologias de obtenção de informações quantitativas e individualizadas de macieiras de um pomar, para estimativa da produção, utilizando recursos de Visão Computacional e da Fotogrametria.

### 1.2.2 Específicos:

- a) Desenvolver uma metodologia para detectar o fruto na macieira utilizando imagens terrestres de curto alcance e câmeras de baixo custo.
- b) Desenvolver e avaliar uma metodologia para detectar o fruto da maçã analisando a variação de cores dos pixels com métodos de processamento de imagens.
- c) Avaliar detectores de objetos de aprendizagem profunda na detecção do fruto de maçã.
- d) Mensurar as informações morfométricas e de produção da planta como número de frutos, diâmetro do caule, área da seção transversal do caule, altura e volume de copa, através das nuvens de pontos.
- e) Comparar os resultados com coleta de informações do pomar no modo convencional com as informações extraídas das imagens.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 A CULTURA DA MAÇÃ (*Malus domestica* BORKH.)

A macieira é uma espécie de fruteiras lenhosa, decídua de clima temperado, é muito adaptável a diferentes climas, podendo ser cultivada desde os trópicos até altas latitudes. Pertence à família das Rosaceas que abrange aproximadamente 100 gêneros e um número maior que 2000 espécies em todo o mundo (EPAGRI, 2002).

Ela possui origem na Ásia e seu cultivo doméstico é de conhecimento milenar. Se destaca como a terceira fruta mais cultivadas no mundo. A China é atualmente o maior produtor, concentrando aproximadamente 49% da produção mundial, em segundo lugar está os Estados Unidos com quase 6% da produção. O Brasil representa mundialmente uma produção de 1,4% da fruta (FAO, 2019).

Em 2017, a produção brasileira de maçã foi de  $1,254 \times 10^6$  Mg (Mega grama,  $1\text{Mg} = 10^6$  g), com Santa Catarina produzindo 638.351 Mg, que representa 50,9% da produção nacional. O Rio Grande do Sul produziu 577.774 Mg, 46%, o Paraná 29.700 Mg, 2,3% e os demais estados representam 0,8% (ABPM, 2018).

No Brasil o primeiro relato de cultivo comercial do fruto, é do ano de 1926, na região de Valinhos, São Paulo. Entre os anos de 1940 e 1960 essa mesma região já possuía aproximadamente de 500 mil a 1 milhão de plantas de macieira (PETRI et al., 2011).

A cultura foi introduzida no Estado de Santa Catarina na década de 1960, na região de Fraiburgo. E foi impulsionada através de um incentivo fiscal da Lei Federal nº 5.106 para abatimento de imposto de renda devido, implantando reflorestamento, que poderia ser feito com frutíferas (PETRI et al., 2011). Com o incentivo e condição climática da região favorecendo a cultura, os plantios se concentraram nas regiões de Fraiburgo e São Joaquim em Santa Catarina, e Vacaria no Rio Grande do Sul, formando o principal polo produtivo da cultura.

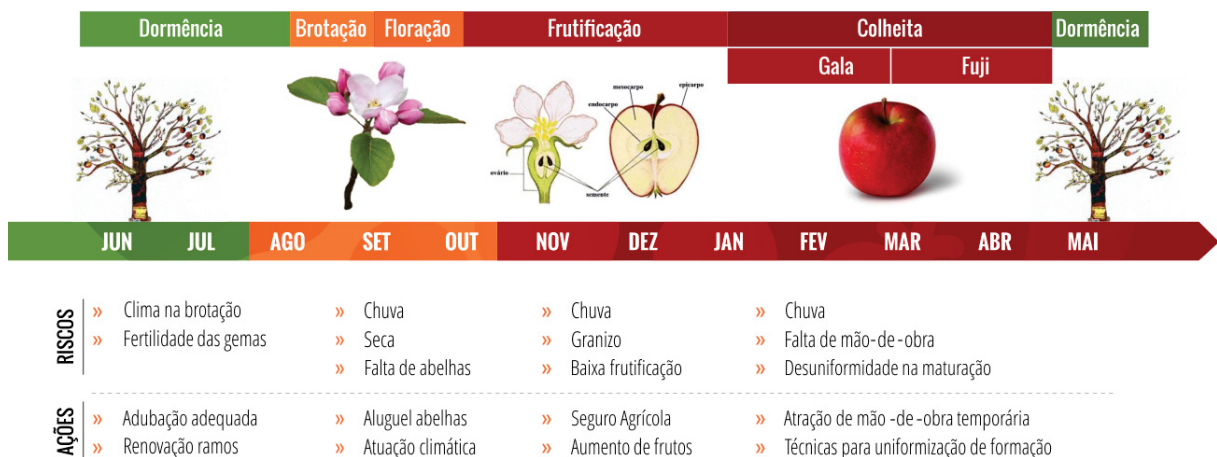
De acordo com Petri et al. (2011) a cultura é uma importante fonte de geração de empregos, sendo três empregos diretos e indiretos por ha, o que representa mais de 150 mil empregos na cadeia produtiva da maçã. Importantes tecnologias foram introduzidas ao longo dos anos que permitiram um aumento de qualidade e produtividade por unidade de área, onde em meados dos anos 80 produzia-se próximo de  $15 \text{ Mg} \cdot \text{ha}^{-1}$  e atualmente pode-se chegar a mais de  $50 \text{ Mg} \cdot \text{ha}^{-1}$ . A densidade de

plantas por ha passou de 500 a 800 para 2500 a 3000 plantas em média. Outras tecnologias foram desenvolvidas como a indução de brotação e floração, permitindo estabilidade de produção. Obteve-se também avanços nos sistemas de condução e poda, manejo de colheita, raleio químico, polinização, controle fitossanitário, conservação e armazenagem da fruta.

A cultura possui dois ciclos distintos, um de dormência e outro vegetativo. O ciclo de dormência vai de Junho a Setembro, e neste período ocorre no pomar a queda das folhas, são realizadas operações de limpeza do pomar como retirada de frutos caídos remanescentes e galhos danificados, efetua-se operações de podas e, ao fim do período começa os tratos culturais (Figura 1). Em regiões de baixa altitude inicia-se a brotação e a florada (ABPM, 2018).

O ciclo vegetativo ocorre de Outubro a Maio. No período de outubro a novembro ocorrem a florada plena, raleio químico, raleio manual (ambos os raleios pretendem eliminar frutos indesejáveis e diminuir a concorrência). De dezembro a janeiro ocorre o desenvolvimento do fruto, controle de ervas daninhas, limpeza e manutenção do pomar (Figura 1). Coincide também com o início da colheita da Gala em regiões de baixa altitudes. Em Fevereiro/Março e Abril/Maio ocorrem as colheitas da Gala e da Fuji respectivamente (ABPM, 2018).

Figura 1 – Ciclo anual da cultura da Macieira.



FONTE: [www.pomifrutas.com.br](http://www.pomifrutas.com.br), (2018).

### 2.1.1 Cultivar de Maçã

Segundo Petri et al. (2011), o sucesso do desenvolvimento da cultura da macieira está ligado ao avanço tecnológico que acompanharam a cultura ao longo do tempo. Um desses avanços tecnológicos é o desenvolvimento e uso de diferentes cultivares. Ainda segundo os autores, as primeiras cultivares utilizadas no Brasil foram Golden Delicious, Starkimson, Blackjon e Melrose. Porém, as cultivares Gala e Fuji rapidamente dominaram o mercado interno. Em 2011 a produção mundial dessas cultivares representavam 14,9% e 7,2%, respectivamente.

No Brasil os Grupos Gala e Fuji representa algo próximo de 95% da produção nacional, com uma participação relativa a área plantada dessas cultivares de 60% e 30%, respectivamente (Fioravanço et al., 2010).

Todos os grupos de cultivares possuem clones (mutações somáticas surgidas espontaneamente) com melhores características que a cultivar original, sendo assim de interesse para a maleicultura. Fioravanço et al. (2010) comentam, que as mutações mais comuns das cultivares estão relacionados a coloração da epiderme do fruto, assim como existem mutações relacionadas ao tamanho do fruto, vigor da planta, resistência a doenças e época de produção.

#### 2.1.1.1 Cultivar Fuji

A origem do grupo Fuji ocorreu no Japão, em 1939, com o cruzamento de 'Ralls Janet' x 'Delicious'. A denominação 'Fuji' ocorreu em 1962, em homenagem à cidade de Fujisaki, onde o cruzamento foi realizado (Fioravanço et al., 2010).

Caracterizada por apresentar floração intensa e contribuindo com *fruit set* e produção elevadas. Possui frutos de tamanho médio a grande, redondo-oblongos ou oblongos, desuniformes e com presença de um pouco de *russetting*, na epiderme do fruto. O requerimento em horas de frio acumulado ( $T < 7,2^{\circ}\text{C}$ ), para quebra da dormência, é de 600 a 800 horas (Fioravanço et al., 2010).

Segundo os mesmos autores, dentre as tendências de cultivo, há procura por clones com maior coloração vermelha na epiderme, outras características buscadas são clones menos vigorosos, com menor tendência à alternância de produção e com frutos mais uniformes. Os principais clones desse grupo são: Fuji Suprema, Fuji Select e Mishima.

#### 2.1.1.1.1 Cultivar Fuji Suprema:

A cultivar Fuji Suprema é produto de mutação espontânea da cultivar Fuji. Foi descoberta no município de Curitibanos - SC em 1986 por um extensionista rural da Epagri. Foi introduzida na Estação Experimental da Epagri de Caçador e enxertada no porta-enxerto M-26. As plantas formadas produziram frutos vermelho compacto, sem estrias, totalmente contrastante com os frutos da planta original. Foi introduzido no mercado como nova cultivar em 1997 sob a denominação de Fuji Suprema. A exceção da epiderme bem mais colorida que na Fuji standard, todas as demais características se mantiveram inalteradas (PETRI et al., 1997). Um exemplo da cultivar pode ser observado na Figura 2.

Figura 2 – Imagem da Cultivar Fuji Suprema (Correia Pinto, SC).



FONTE: O autor (2019).




### 2.1.2 Variáveis para avaliação de pomar

Nesse item será descrito algumas variáveis que são utilizadas em avaliações de desempenho de pomar, mas que dependem muito do objetivo final que se quer avaliar e da cultura estudada. Existem variáveis quantitativas e qualitativas, e nas quantitativas as variáveis se apresentam em diferentes níveis de escalas.

A figura descrita a seguir correspondem a variáveis analisadas em trabalhos realizados por Lima (2016) e Martins (2014), onde essas variáveis foram observadas e mensuradas nos pomares de estudo de seus trabalhos (Figura 3). A intenção aqui é mostrar um exemplo dos tipos de variáveis possíveis de serem avaliadas em cada nível de escala da planta.

Todas essas mensurações descritas pelas autoras, foram coletadas manualmente. Esses exemplos demonstram o trabalho e a dificuldade de coletar variáveis para avaliação de pomares.

Figura 3 – Comparativo de exemplos de variáveis possíveis de serem analisadas com a referida escala da planta.

Variáveis	Escala da planta
<p>Número de Anteras por flor (nº)            Número de pólen por antera (nº)            Germinação de pólen (%)            Comprimento e diâmetro de gema (mm)            Número de escama das gemas (nº)</p>	 <p>Fonte: (Francescato, 2014).</p>
<p>Comprimento do ramo do ano (cm)            Número de gemas do ramo do ano (nº)            Comprimento dos entrenós (cm)            Número de inflorescência por ramo do ano (nº)            Área Foliar (cm<sup>2</sup>)            Frutificação efetiva (%)            Crescimento dos frutos (mm)            Firmeza de Polpa dos frutos (Kg.cm<sup>2</sup>)</p>	 <p>Fonte: (Francescato, 2014).</p>
<p>Diâmetro do Caule (cm)            Área da Seção Transversal do Caule (cm<sup>2</sup>)            Peso médio por fruto (g)            Produção por planta (kg.planta<sup>-1</sup>)            Produtividade (t.ha<sup>-1</sup>)            Eficiência produtiva (kg.cm<sup>-2</sup>)            Volume (cm<sup>3</sup>)</p>	 <p>Fonte: O autor (2019)</p>

## 2.2 TÉCNICAS DETECÇÃO E LOCALIZAÇÃO DE FRUTOS

As técnicas de detecção e localização de frutos, além de outras estruturas de plantas tais como galhos, folhas, entre outros são usualmente retratados por segmentação do espaço de cores em imagens RGB, limiarização deste espaço de cores visando individualizar objetos de interesse na imagem ou Deep Learning (aprendizado profundo) pelo qual utilizam-se técnicas de aprendizado de máquina voltado também à identificação de objetos específicos.

Nas ciências florestais, agrícolas, ambientais e engenharias, o maior volume de trabalhos nas literaturas sobre detecção de estruturas ou partes de plantas estão ligados a detecção automática de troncos por meios de levantamentos a Laser Scanner ou de forma combinada deste com sensores de imagens (LIANG et al., 2012; SHALAL et al., 2015; OLOFSSON et al., 2014; LIANG et al., 2014; CABO et al., 2018; BERVEGLIERI et al., 2017).

Outra estrutura de planta de interesse diz respeito às copas das árvores. Autores tem estudado formas de detectar automaticamente as copas com uso de imagens obtidas por meio de plataformas pilotadas remotamente (ARP) (GUILLEN-CLIMENT et al., 2012; KALISPERAKIS et al., 2015) e Laser Scanner (SANZ-CORTIELLA et al., 2011; FERNÁNDEZ-SARRÍA et al., 2013; SANZ et al., 2018) a fim de estimarem volume de copa e Índice de Área Foliar (IAF).

Na fruticultura, o interesse está relacionado ao próprio fruto, sendo que técnicas para sua localização são desenvolvidas com uso de sensores de imagens, podendo ser associados a unidades robóticas autônomas (BARGOTI; UNDERWOOD, 2017 e STEIN et al., 2016; ZHAO et al., 2016) ou com aquisições manuais de imagens para localização e contagem (MEHTA et al., 2017; LIU et al., 2016; FONT, D. et al., 2014), em diferentes condições de iluminação, como luz natural ou a noite com luz artificial (LINKER, 2017, 2018; e LINKER; KELMAN; 2015 e LIU et al., 2016)

Na cultura da maçã existem alguns trabalhos relacionados à detecção do fruto como Gongal et al. (2016), que trabalharam com imagens tomadas de pontos de vistas opostos, mas convergentes, para resolver problemas de oclusão. Nguyen et al. (2016) fizeram uso de câmera RGB-D para reconhecimento automático de frutos de maçãs em nuvem 3D e cálculo dos diâmetros dos frutos.

Previsões de produção da maçã baseadas em imagens também são exploradas na literatura, relacionados com número de frutos por planta (ZHOU et al., 2012; BARGOTI; UNDERWOOD, 2017) e com base na detecção de flores no estágio de Floração Plena (AGGELOPOULOU et al., 2011 ). Estimativas essas, que só conseguem uma boa correlação com a produtividade, se não houver alterações bruscas do clima, ou seja, quando essa variável se comporta em condições ideais para o desenvolvimento da cultura durante todo seu ciclo (BRAGA, 1995).

Na situação de detectar frutos com o uso de câmeras de baixo custo, no processamento das imagens RGB, basicamente ocorre uma transformação para um ou mais espaços de cores, como HSI, HSV, CIELAB, CIELUV, OHTA e/ou RGBn, e em seguida a aplicação automática de um limiar, pelo método Otsu (JI et al., 2016; JIDONG et al., 2016; GONGAL et al., 2016; LIU et al., 2016; ZHAO, C. et al., 2016; ZHAO, Y. et al., 2016); ou pesquisadores calculam um limiar apropriado para o alvo de interesse (ZHOU et al., 2012) para segmentar a imagem. Em seguida ocorrem os processos visando filtrar a imagem por meio de filtros morfológicos ou componentes conexos e contar de objetos de interesse (frutos).

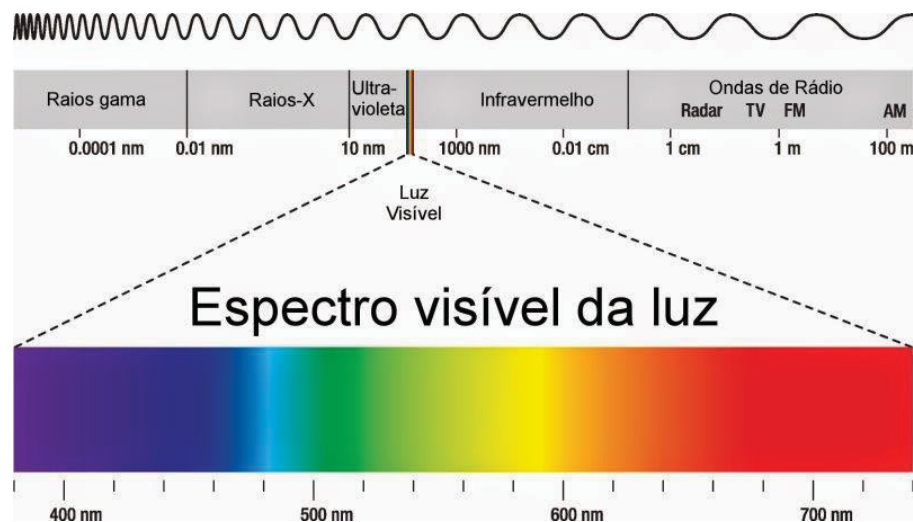
## 2.2.1 ESPAÇO DE CORES

### 2.2.1.1 Cor

A cor é um produto do processamento mental de uma fração da radiação eletromagnética detectada pelos olhos humanos. O que se percebe são respostas das cores apenas do espectro visível, que são refletidas pelas feições, porque uma parte da energia eletromagnética é absorvida e outra transmitida (NOGUEIRA, 2009).

A percepção da cor começa com a sensação visual, quando cones e bastonetes, localizados na retina dos olhos, são estimulados pela Radiação Eletromagnética (REM) de certos comprimentos de onda, que variam de aproximadamente 400nm a 700nm (nanômetro =  $1 \times 10^{-9}m$ ) (Figura 4).

Figura 4 – Espectro Eletromagnético com destaque para espectro visível



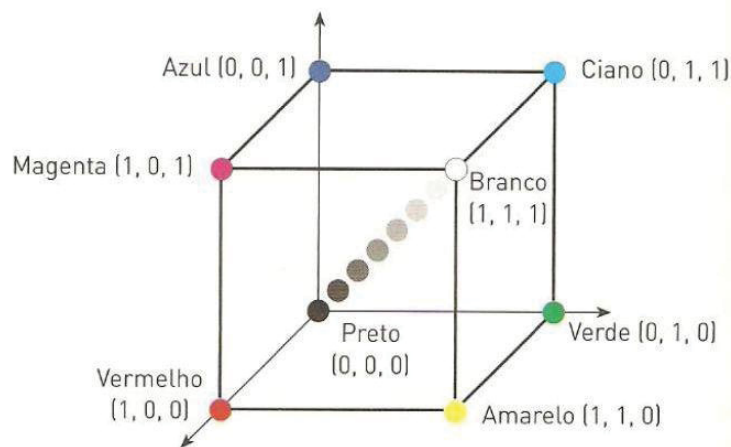
Fonte: Florenzano (2011).

### 2.2.1.2 Espaço RGB

O modelo RGB é muito utilizado em dispositivos como monitores e câmeras. Geralmente, assume-se que os valores R, G e B estão normalizados entre 0 e 1. Segundo Gonzalez & Woods (2000), o modelo RGB pode ser representado em um sistema Cartesiano ortogonal. Toda as possíveis combinações das componentes R, G e B ocupam um cubo dentro deste sistema. As cores primárias vermelho, verde e azul estão em três vértices do cubo que coincidem com os eixos do sistema Cartesiano ( $[1,0,0]$ ;  $[0,1,0]$  e  $[0,0,1]$ ), as cores secundárias ciano, magenta e amarelo estão nos outros três vértices. Na origem e localiza o preto e o vértice mais afastado da origem corresponde à cor branca. No modelo RGB, a variação de tons de cinza se estende através da diagonal do cubo, ou seja, a reta que une a origem (preto) até o vértice mais distante (branco) (Figura 5).

De acordo com Alves (2010), cada pixel tem o seu próprio valor RGB representado geralmente por três bytes, um para cada componente de cor: vermelho, verde e azul. Se cada componente for armazenado como um byte, cada um poderá representar 256 diferentes intensidades. Isso significa que é possível trabalhar com 16,7 milhões de cores.

Figura 5 – Representa o Cubo do espaço de cores RGB



Fonte: UFRGS (2018)

### 2.2.1.3 Espaço HSV

HSV é a sigla para o sistema de cores formadas pelas componentes *Hue* (matiz ou tonalidade), *Saturation* (Saturação) e *Value* (Valor). Suas componentes são descritas a seguir:

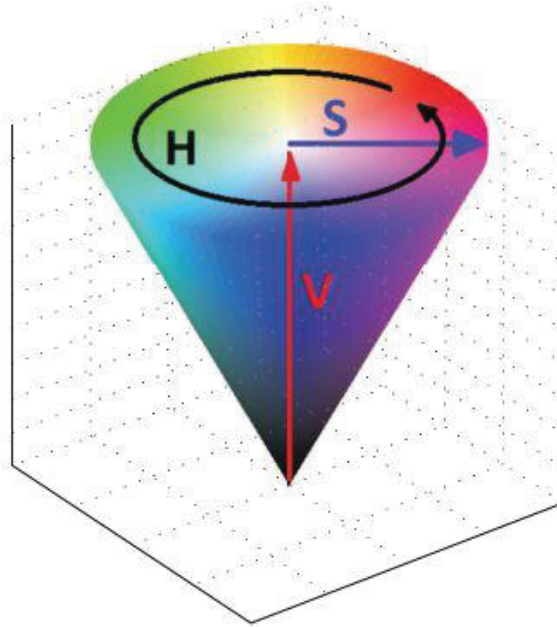
**Matiz:** é a medida do comprimento de onda médio da luz que o objeto reflete ou emite. Ela define a cor do objeto. Os valores variam de  $0^\circ$  a  $360^\circ$ , e pode ser normalizado para valer de 0 a 1. A componente *Hue* é também frequentemente traduzida como tonalidade.

**Saturação:** pode ser chamado de “pureza” da cor. A saturação mede a proporção de luz branca misturada na cor. Quanto menor seu valor, mais luz branca está presente e a cor do pixel de torna menos pura. Valores altos de saturação mostram que o matiz predomina sobre a luz branca. Possui valores de 0 a 1.

**Valor ou Brilho:** esta componente é também descrita como Brilho e é proporcional ao total de energia contida em todas as componentes RGB. Ela é expressa em valores de 0 (ausência de luz) a 1, brilho máximo (branco).

O sistema é caracterizado por ser uma transformação não-linear do sistema RGB. Ambos são parecidos geometricamente, pois a representação do sistema HSV se dá por um hexágono, onde altura do hexágono corresponde ao eixo acromático (do preto ao branco) do cubo RGB (Figura 6).

Figura 6 – Representa o Hexágono do modelo HSV.



Fonte:Kau; Lee (2013)

Para transformar o espaço de cores RGB para HSV, primeiro aplica-se uma rotação de corpo rígido ao espaço RGB, coincidindo um dos novos eixos com a diagonal principal do cubo RGB:

$$\begin{pmatrix} I \\ V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} \sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1)$$

A saturação é calculada como a distância da cor em relação à linha de cinza (ausência de saturação). O matiz é uma medida angular que descreve a tonalidade de cor predominante. Por ser uma variação contínua, no cálculo desta componente se fixa uma origem (por exemplo no vermelho) e uma direção de rotação para variar as tonalidades. Os valores de Matiz (H), Saturação (S) e Intensidade (V) são obtidos com as fórmulas abaixo:

$$H = \tan^{-1} \left( \frac{V_2}{V_1} \right) \quad (2)$$

$$S = \sqrt{(V_1)^2 + (V_2)^2} \quad (3)$$

$$V = I \quad (4)$$

#### 2.2.1.4 Espaço CIE-LAB

Segundo Scuri (2002) em monitor de computador as três cores primárias emitidas por cada um dos tubos de raios catódicos (vermelho a 700 nm, verde a 546 nm e azul a 436 nm) não correspondem às cores detectadas pelo olho humano. Deve-se então modificar as proporções de intensidade de cor aplicadas a cada uma das componentes primárias emitidas. Estas novas proporções podem assumir valores negativos em algumas gamas de comprimento, o que significa que, com um monitor, não é possível reproduzir todos os comprimentos de onda de luz visível, isto é, não é possível reproduzir todas as cores do espectro visível pela combinação ponderada de luzes vermelha, verde e azul. Existem, portanto cores que não podem ser reproduzidas em monitores a cores pela adição ponderada das cores vermelha, verde e azul.

Devido a esses problemas, em 1931 a CIE (*Commission Internationale de l'Éclairage*) resolveu adotar um novo modelo de representação padrão X, Y, Z, cujas cores primárias não correspondem a cores visíveis, mas suas componentes de cor são positivas sendo possível então reproduzir no monitor todos os comprimentos de ondas de luz visível (SCURI, 2002).

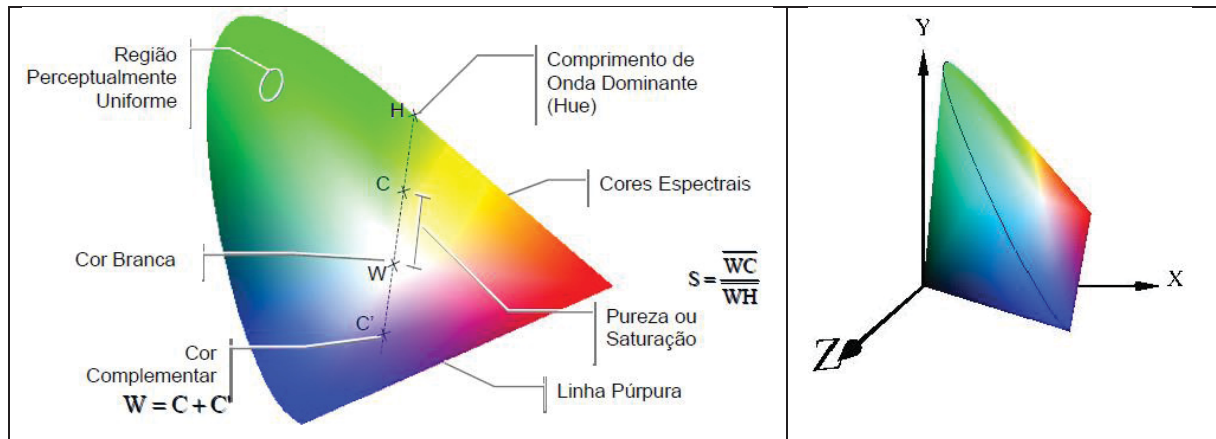
As coordenadas de cromaticidade dessas cores primárias são conhecidas, sendo possível a realização de cálculos que permitem não só obtenção de valores de grandezas no sistema XYZ a partir de grandezas do sistema RGB, assim como mudanças de coordenadas entre outros sistemas de cor (SCURI, 2002).

A conversão de sRGB para CIE XYZ é definida por:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0,4122 & 0,3576 & 0,1805 \\ 0,2125 & 0,7152 & 0,0722 \\ 0,0193 & 0,1192 & 0,9504 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (5)$$

A fim de representar esses valores criou-se o diagrama de cromaticidade. Luminância refere-se apenas a intensidade da cor, sendo assim se reproduziu uma visualização das cores visíveis que envolve somente as componentes X e Z. Mantendo a luminância constante Y (Figura 7).

Figura 7 - Diagrama de Cromaticidade e Representação 3D do Espaço de Cor CIE XYZ.



Fonte: Scuri (2002)

O espaço colorido tridimensional produzido pela plotagem dos valores tristimulares CIE XYZ em coordenadas retangulares não é visualmente uniforme, nem o espaço (x, y, Y) nem o bidimensional CIE (x, y) do diagrama de cromaticidade. Distâncias iguais nesses espaços não representam diferenças igualmente perceptíveis entre estímulos de cores. Por esta razão, em 1976, o CIE introduziu e recomendou dois novos espaços (conhecidos como CIE-Lab e CIE-Luv) cujas coordenadas são funções não-lineares de X, Y e Z. A recomendação foi apresentada na tentativa de unificar em espaços de cores. Ambos os espaços de cores se tornam mais uniformes sendo bem aceitos e amplamente utilizados.

As fórmulas para calcular os valores para CIE-LAB dependem dos valores convertidos para CIE XYZ, e dos valores  $X_w$ ,  $Y_w$  e  $Z_w$  que correspondem ao componente da cor branca de referência.

Variável de Luminosidade L:

$$L = 116 \left( \frac{Y}{Y_w} \right)^{1/3} - 16 \quad (6)$$

Coordenadas de cromaticidade A e B:

$$A = 500 \left[ \left( \frac{X}{X_w} \right)^{1/3} - \left( \frac{Y}{Y_w} \right)^{1/3} \right] \quad (7)$$

$$B = 200 \left[ \left( \frac{Y}{Y_w} \right)^{1/3} - \left( \frac{Z}{Z_w} \right)^{1/3} \right] \quad (8)$$

Caso os valores de  $(X/X_w)$ ,  $(Y/Y_w)$  ou  $(Z/Z_w)$  for  $< 0,008856$ , os valores nas equações acima são substituídas por:

$$\left( \frac{X}{X_w} \right)^{1/3} \Rightarrow 7,787 \left( \frac{X}{X_w} \right) + \frac{16}{116} \quad (9)$$

$$\left( \frac{Y}{Y_w} \right)^{1/3} \Rightarrow 7,787 \left( \frac{Y}{Y_w} \right) + \frac{16}{116} \quad (10)$$

$$\left( \frac{Z}{Z_w} \right)^{1/3} \Rightarrow 7,787 \left( \frac{Z}{Z_w} \right) + \frac{16}{116} \quad (11)$$

### 2.2.1.5 Espaço OHTA

O espaço de cor OHTA, uma conversão linear de RGB introduzida por Ohta et al. (1980) após uma análise colorimétrica de oito cores diferentes, proporcionando um melhor desempenho na segmentação de imagens coloridas. O espaço de cor OHTA pode ser determinado pela transformação linear dos componentes R, G e B, o que permite um processamento mais rápido.

Os componentes de cor ortogonais no espaço de cores OHTA ( $I_1$ ,  $I_2$ ,  $I_3$ ), são cálculos simples e os componentes são independentes. Há duas maneiras de expressar o espaço de cores OHTA usando o espaço de cores RGB, conforme mostrado nas equações abaixo:

$$\begin{cases} I_1 = (R + G + B)/3 \\ I_2 = (R - B)/2 \\ I_3 = (2G - R - B)/4 \end{cases} \quad (12)$$

$$\begin{cases} I'_2 = R - B \\ I'_3 = (2G - R - B) \end{cases} \quad (13)$$

Pode-se perceber que as equações (1) e (12) são semelhantes sendo assim sistemas paralelos.

Variações com relação as ordens das bandas dentro das fórmulas podem ser efetuadas, com o objetivo de realçar alguma cor específica de algum alvo de interesse.

### 2.2.2 APRENDIZADO PROFUNDO (*DEEP LEARNING*)

O aprendizado profundo ou *Deep Learning* (DL) é uma técnica de aprendizado de Máquina ou *Machine Learning* (ML) que nos últimos anos vem impulsionando os muitos aspectos da sociedade moderna, como pesquisas na web, recomendação de sites comerciais, e atualmente cada vez mais integrados a dispositivos que conseguem transcrever fala em textos, combinar itens de notícias e identificar objetos em imagens (LECUN et al., 2015).

O DL permite que modelos computacionais compostos com várias camadas de processamento aprendam a representar dados com vários níveis de abstração. O DL descobre a estrutura intrincada em elevados conjuntos de dados usando um algoritmo retropropagação para indicar como uma máquina deve alterar seus parâmetros internos, que são utilizados para calcular a representação em cada camada a partir da representação da camada anterior (LECUN et al., 2015).

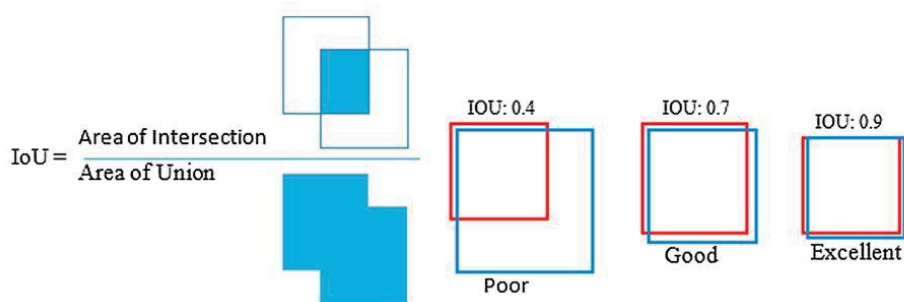
Essa é uma grande vantagem do DL, a extração automática de características de dados brutos de entrada, com características de um nível superior sendo formada pela composição de um nível inferior. É ideal para resolver problemas complexos, pois utiliza de modelos complexos. Kamilaris; Prenafeta-Boldú, (2018) afirmam que esses modelos de arquitetura complexos utilizados em DL aumentam a precisão de classificações ou diminui erros em problemas de regressão, desde que existam conjuntos de dados adequadamente grandes disponíveis descrevendo o problema.

#### 2.2.2.1 Avaliação e acurácia

Alguns parâmetros são utilizados para avaliar o desempenho dos resultados dos modelos adotados de arquitetura de DL. O primeiro parâmetro é o *Intersection over Union* (IoU).

Para fazer a avaliação de detecção de objetos em imagens o *IoU* é indicado. Ele é a razão entre a área de sobreposição e a área de união entre as caixas delimitadoras detectadas como verdadeiras (KOIRALA et al., 2019a). Os valores de IoU podem variar de 0 (sem sobreposição) a 1 (sobreposição total). A Figura 8 apresenta de forma gráfica como é determinado esse parâmetro. Algumas modelos adotam um valor de IoU como limiar para aceitar uma anotação como verdadeira.

Figura 8 – Definição de IoU na função de detectar objetos usando anotações de caixa delimitadora.

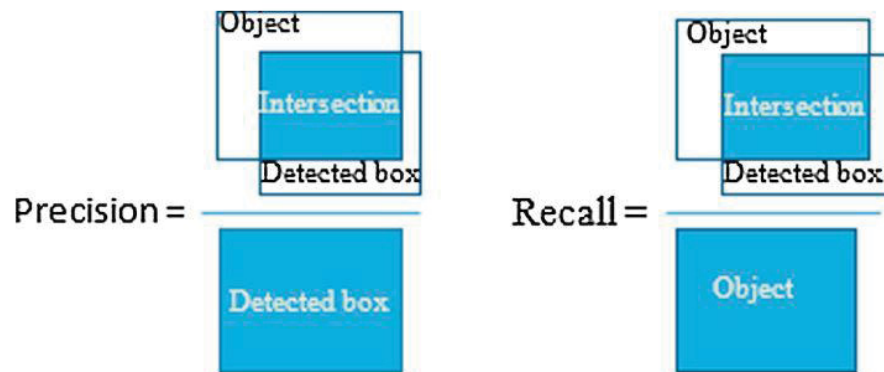


Fonte: Koirala et al. (2019a)

Em grande parte, o valor limite para IoU é 0,5 para aceitação de verdade para as anotações, valores acima desse limiar são aceitos como positivos. Já valores menores são considerados detecções falsas. Porém dependendo da necessidade de detecção esse limite pode ser mudado.

Outro parâmetro de avaliação de resultados de modelos adotados em DL são precisão e recall, que também podem ser avaliados pela sobreposição das caixas delimitadoras. O modo de determinação dos parâmetros é informado na Figura 9. A precisão (*precision*) é a proporção de caixas detectadas que correspondem as caixas de verdade do objeto. A recall é a proporção de caixas detectadas que correspondem às caixas de verdade do objeto em todas as caixas de verdade dos objetos (KOIRALA et al., 2019a).

Figura 9 – Interpretação de precisão e recall para a função de detectar objetos.



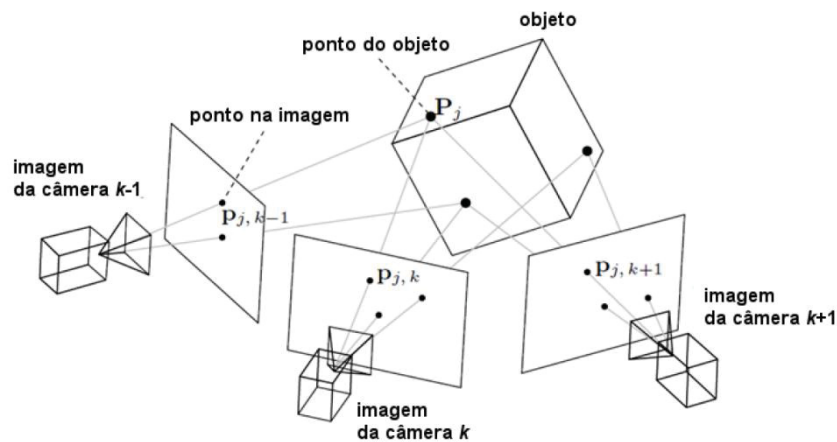
Fonte: Koirala et al. (2019a)

### 2.3 STRUCTURE FROM MOTION - SfM

Na última década, uma abordagem fotogramétrica que oferece aquisição flexível e econômica de dados combinados 3D e espectrais encontrou ampla aplicação e aceitação na geografia física: os algoritmos *Structure from Motion* (SfM), em conjunto com *multi-view stereo* (MVS) (SfM-MVS, comumente abreviado apenas para SfM) (IGLHAUT et al., 2019).

Trata-se de uma técnica estudada nos campos de visão de máquina e percepção visual, que se refere ao processo de obtenção de uma cena 3D a partir de uma série de imagens 2D adquiridas por um sensor em movimento. O processo é similar ao da visão estéreo de modo que nos dois casos a correspondência entre a reconstrução do objeto 3D e as imagens adquiridas precisa ser encontrada. Observando o SfM, para encontrar a correspondência entre as imagens, características como pontos de extremidades precisam ser rastreadas de uma imagem para outra. As trajetórias destas características entre as imagens são utilizadas para reconstruir sua posição no espaço tridimensional e o movimento da câmera, em outras palavras, suas posições relativas (VIANA, 2015). Para atender esse quesito, podem ser utilizadas imagens de uma câmera apenas, em movimento (*frames* extraídos de um vídeo ou fotos tiradas de diversos pontos) ou de várias câmeras em posições diferentes (Figura 10).

Figura 10 – Princípio da dinâmica do SfM.



Fonte: Adaptado de VIANA, (2015).

A SfM difere da fotogrametria tradicional principalmente em três aspectos: (i) as características podem ser automaticamente identificadas e combinadas em imagens em diferentes escalas, ângulos de visão e orientações, o que é particularmente benéfico quando pequenas plataformas instáveis são consideradas; (ii) as equações usadas no algoritmo podem ser resolvidas sem informações das posições das câmeras ou pontos de controle do solo, embora ambos possam ser adicionados e usados e (iii) a calibração da câmera pode ser resolvida ou refinada automaticamente durante o processo. SfM pode, portanto, fornecer modelos fotogramétricos automaticamente sem exigir homogeneidade rigorosa em imagens sobrepostas, posição de câmera e calibrações (IGLHAUT et al., 2019).

Segundo Viana (2015) estes dados são resolvidos simultaneamente através de um procedimento iterativo altamente redundante de ajuste de conjunto (*bundle adjustment*) baseado em uma fonte de dados extraída automaticamente de um conjunto de imagens sobrepostas. Para Westoby et al. (2012), o método tem melhor funcionamento em conjuntos de imagens com grandes áreas de sobreposição, que permitam capturar toda a estrutura 3D da cena vista de uma variada ordem de posições, ou seja, com o sensor em movimento.

Para agricultura a técnica SfM é vantajosa em coletas de dados tridimensionais (VÁZQUEZ-ARELLANO et al., 2016), por usar câmeras digitais de fácil acesso e economicamente viáveis, por existir *software open source* e comerciais para reconstrução 3-D e suportar aplicações aéreas e terrestres. Contudo, há a

necessidade de calibração da câmera e pontos de apoio para medições seguras. Além disso, atualmente, a geração de nuvem de pontos não é viável para uso em tempo real e requer experiência na obtenção dos dados brutos.

Maiores detalhes sobre o uso da técnica pode ser encontrada em Carrivick et al. (2016).

### 3 METODOLOGIA GERAL

Para ser possível detecção de frutos de maçã em um pomar comercial, através de imagens obtidas por câmeras de baixo custo, esse estudo apresenta o uso de duas metodologias: A primeira, identificada aqui como **METODOLOGIA A**, descreve uma sequência de etapas para efetuar a detecção de frutos por segmentação de cores nas imagens RGB. A segunda, identificada como **METODOLOGIA B**, descreve o processo de detecção com o estado da arte em detecção de objetos por *Deep Learning* (DL), onde sete métodos foram avaliados para detecção de frutos de maçã. A **METODOLOGIA B** foi elaborada e analisada em colaboração com a equipe e recursos dos Professores José Marcato Junior e Wesley Nunes Gonçalves da Universidade Federal de Mato Grosso do Sul (UFMS).

Uma análise com os dados obtidos de modo convencional e de modo automático foi feita ao final para avaliar os métodos (**METODOLOGIA C**).

A seguir neste capítulo, é descrito o local do estudo realizado, o método de aquisição das imagens e os métodos de aquisição dos dados do pomar de modo convencional. Estas informações são comuns para as metodologias A e B.

Cada metodologia (A e B) possui métodos específicos e estão apresentados de forma separadas nos **Capítulos 4 e 5**, pois cada capítulo corresponde a um artigo já publicado em periódico científico.

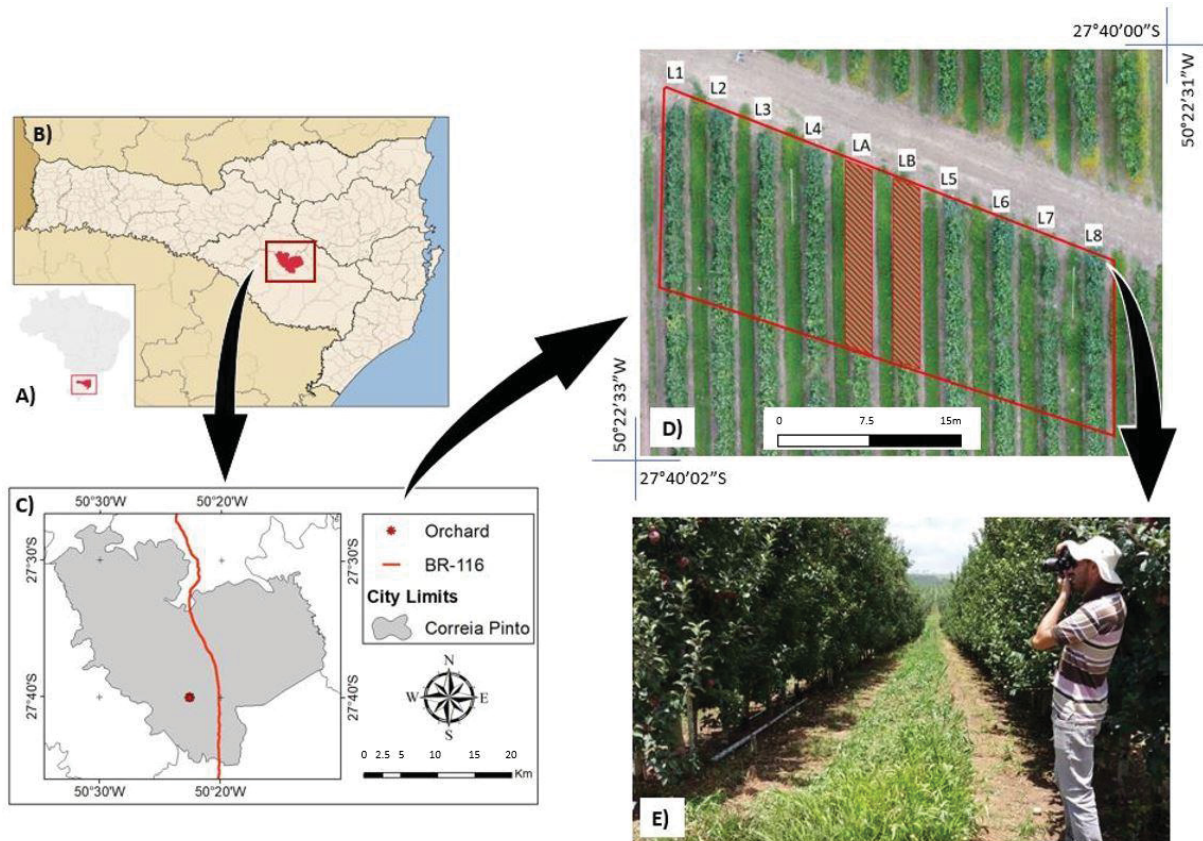
#### 3.1 CARACTERIZAÇÃO E LOCAL DO ESTUDO REALIZADO.

O estudo proposto foi conduzido no município de Correia Pinto, Estado de Santa Catarina, em pomar comercial (27°40'01"S, 50°22'32"W) (Figura 11 A-C). Segundo a classificação de Köppen, o clima é subtropical (Cfb) com verões temperados (ALVARES et al., 2013). O solo da região é classificado como Cambissolo Húmico Alumínico (CHa) no Sistema Brasileiro de Classificação de Solos (SANTOS et al., 2018). O relevo é relativamente plano e a altitude média local acima do nível do mar é de 880 m. A precipitação média anual acumulada é de 1560 mm, com base nos registros pluviométricos da estação de Lages (2750005 - INMET), localizada a 15 km do pomar, no período de 74 anos (1941-2014) (ANA, 2020). A precipitação é regularmente distribuída entre os meses de setembro a fevereiro, com uma média de

150 mm mensais. A precipitação média de abril a junho é inferior a 100 mm. Os eventos de geada geralmente ocorrem de abril a setembro.

O pomar selecionado para o estudo possui 46ha, e as macieiras foram plantadas em 2009, com espaçamento de 0,8m x 3,5 m. Resultando uma densidade de 3570 plantas.ha<sup>-1</sup>, com o alinhamento nas linhas de plantio seguindo a orientação N-S. Na área analisada, existem duas cultivares de maçãs plantadas: Fuji Suprema e Gala, e estão plantados em fileiras individuais, mantendo uma proporção de quatro fileiras de Fuji Suprema para duas fileiras de Gala, respectivamente. Esta prática é usada para proporcionar a polinização cruzada (DENARDI et al., 2019). A cultivar analisada foi a Fuji Suprema, as amostras foram divididas por fileiras identificadas de L1 a L8 (Figura 11 D). As duas fileiras de cultivares Gala não foram analisadas e foram identificadas como LA e LB na mesma. A área analisada no estudo não possui cobertura com tela anti-granizo.

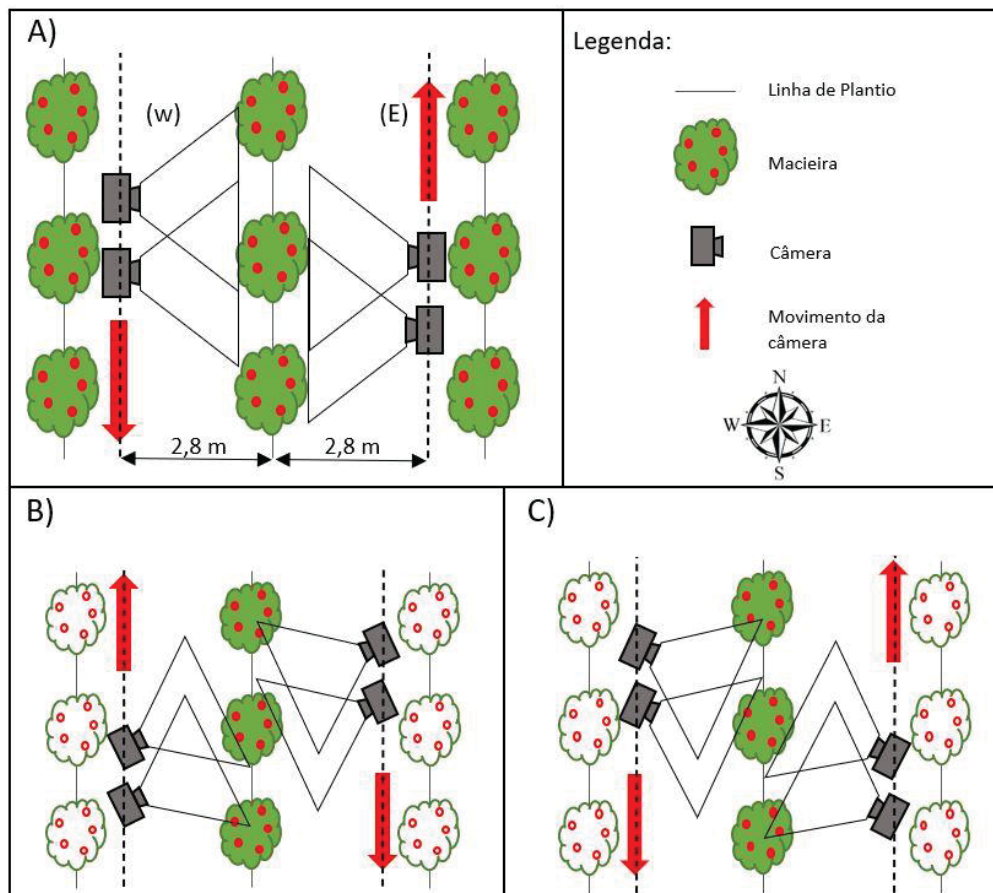
Figura 11 - (A) Brasil; (B) Estado de Santa Catarina; (C) Município de Correia Pinto e localização do pomar; (D) Vista aérea do local do estudo e identificação das linhas amostradas de Fuji Suprema (L1 a L8), e linhas de Gala (LA e LB); (E) perspectiva da aquisição de imagens no pomar durante o levantamento de campo.



### 3.2 MODO DE AQUISIÇÃO DAS IMAGENS

Para adquirir o conjunto de dados de imagens, uma câmera Canon EOS — T6 (Tóquio, Japão) com sensor CMOS de 5184x3456 pixels (17,9 Mp) e tamanho de pixel de 4,3  $\mu\text{m}$  (micrômetros), foi usada com uma distância focal nominal de 18 mm. A câmera foi operada manualmente, na posição retrato na altura do rosto do operador, com o flash auxiliar da câmera ativo (Figura 11 E). A Figura 12(A) mostra o modo de aquisição de imagem próximo à perpendicular. Imagens oblíquas das mesmas plantas também foram adquiridas (Figura 12 B e C). A cada captura de imagem, foi dado um pequeno passo paralelo à linha de plantio, e então feita uma nova aquisição de modo a manter uma sobreposição na imagem anterior, semelhante ao modo “*Stop and Go*” adotado por Liang et al. (2014). Ao total dez plantas por linhas de plantio foram fotografadas em ambos os lados.

Figura 12 – Modo ilustrado de obtenção das imagens no pomar: com a câmera perpendicular a linha de plantio (A) e com a câmera oblíqua a linha de plantio (B e C). As sobreposições das imagens também são apresentadas.



FONTE: O autor (2021).

As imagens foram obtidas em duas datas, a primeira em 29 de março e a segunda em 1º de abril de 2019, com variações nos horários de aquisição. Os tempos de aquisição para cada linha de plantio são mostrados na Tabela 1, e variaram aproximadamente entre 11h50 e 18h20. Nenhum fundo artificial foi utilizado, e a luz natural variou entre períodos nublados e sem nuvem.

Tabela 1 – Informação de data, e horários de início e fim das aquisições das imagens nas linhas de plantio.

Linhas de Plantio	Data	Hora de início	Hora de término
L1	29/03/19	12h50	13h20
L2	29/03/19	13h45	14h10
L3	29/03/19	14h30	15h05
L4	01/04/19	11h55	12h30
L5	29/03/19	16h10	16h40
L6	29/03/19	17h00	17h30
L7	29/03/19	17h45	18h20
L8	01/04/19	15h20	16h00

FONTE: O autor (2019).

### 3.3 AQUISIÇÃO DOS DADOS DE MODO CONVENCIONAL

Nesta seção demonstra-se brevemente como foi realizado a coleta dos dados do pomar de forma convencional. A coleta foi realizada nas linhas de L1 até L8 aproximadamente com um dia de diferença com as datas da aquisição das fotos.

Todas as operações foram realizadas com no mínimo duas pessoas.

#### 3.3.1 Diâmetro do Caule ( $D_i$ )

O Diâmetro do Caule foi obtido com auxílio de uma trena métrica, medindo a circunferência ( $C$ ) do Caule aproximadamente 20 cm acima do ponto de enxertia (Figura 13). Após a medida o valor da circunferência foi transformado para Diâmetro com o uso da equação (14).

$$D_i = \frac{C}{\pi} \quad (14)$$

Figura 13 – Medindo a circunferência do caule para obtenção do diâmetro ( $D_i$ ) (A). Detalhe da operação (B).



FONTE: O autor (2019).

### 3.3.2 Altura da planta ( $H$ )

A altura da planta foi obtida por comparação utilizando como referência uma régua topográfica, graduada em centímetros. A mesma foi colocada ao lado do caule da planta e observado o galho mais alto pertencente a referida planta. Um exemplo pode ser observado na Figura 14.

### 3.3.3 Volume da planta ( $Vol$ ).

Para o cálculo do volume da planta, foram usadas as medidas de largura, comprimento e altura corrigida da árvore. As medidas de largura ( $Lg$ ) e comprimento ( $Co$ ) foram obtidas com trena métrica, conforme apresentado na Figura 15, e a altura corrigida ( $Hc$ ) corresponde o ponto de inserção do primeiro ramo até o ápice da planta. Foi descontado na altura de planta ( $H$ ) o valor médio de 0,6m referente ao ponto de inserção do primeiro ramo. O Volume foi calculado com a multiplicação das três dimensões, conforme equação (15 **Erro! Fonte de referência não encontrada.**).

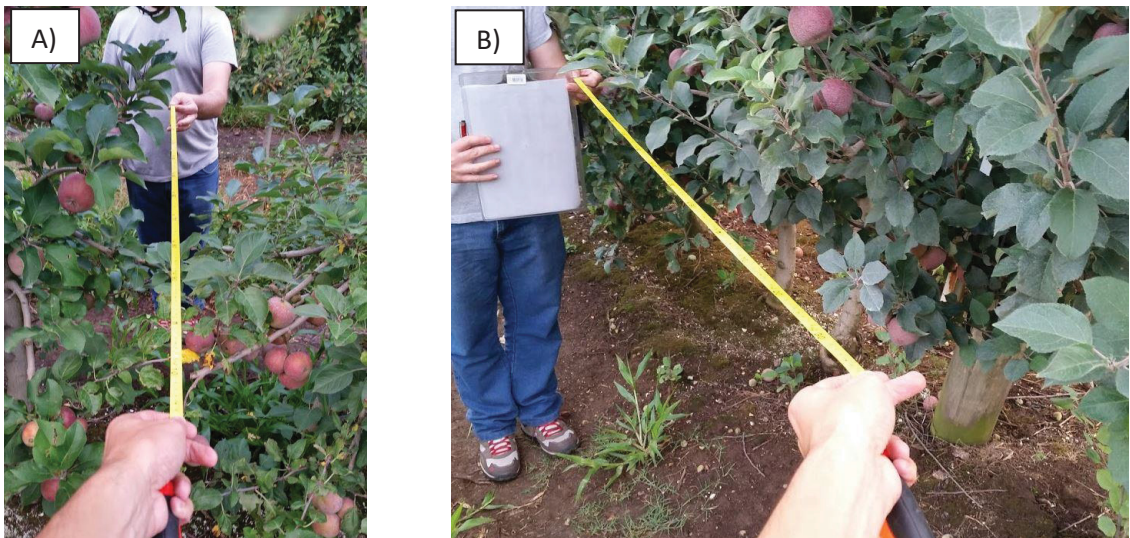
$$Vol = Lg. Co. Hc \quad (15)$$

Figura 14 – Exemplo de medição de altura de planta no pomar. Operador segurando a régua (A) e detalhe do desafio para fazer a leitura (B).



FONTE: O autor (2019).

Figura 15 – Medidas de Largura (A) e Comprimento (B) da planta no pomar.



FONTE: O autor (2019).

### 3.3.4 Contagem dos frutos

Essa operação foi realizada com auxílio de um contador manual para cada operador, que se posicionaram em lados opostos da planta. Ao final de cada planta a contagem dos dois operadores foi somada, e Número de Frutos por Planta (*NFP*) foi estabelecido.

Figura 16 – Exemplo de contador de manual



FONTE: [www.vonder.com.br](http://www.vonder.com.br) (2020).

### 3.3.5 Produção e Eficiência Produtiva

Para determinar a Produção (*Prod*), foram pesados amostras coletadas em 5 plantas aleatórias por linha analisada, contendo 20 frutos cada. Assim o Peso Médio por Fruto (*PMF*) foi determinada para cada linha. A Produção por planta ( $\text{kg.planta}^{-1}$ ) é o resultado da multiplicação entre o *NFP* e o *PMF* (equação ((16))).

$$Prod = NFP.PMF \quad (16)$$

A Eficiência Produtiva (*EP*) necessita do valor da Área da Seção Transversal do Caule (*ASTC*). A *ASTC* ( $\text{cm}^2$ ) foi determinada com o valor do *Di* e calculada conforme equação (17).

$$ASTC = \pi \left( \frac{Di}{2} \right)^2 \quad (17)$$

Para determinação da EP ( $\text{kg.cm}^{-2}$ ) adotou-se a seguinte equação (18).

$$EP = \frac{Prod}{ASTC} \quad (18)$$

## 4 METODOLOGIA A:

Este capítulo corresponde ao artigo publicado no periódico Boletim de Ciências Geodésicas, volume 27, número 2 de 2021 (BIFFI, et al. (2021a)<sup>1</sup>. O artigo original foi publicado em língua inglesa e para incentivar a exploração da técnica no país o artigo foi traduzido para língua portuguesa. A introdução do artigo foi suprimida, pois o assunto está abordado na seção 2.2 relativo a técnicas de espaço de cores. Assim como local de estudo e parte da metodologia que corresponde a metodologia geral descrita no capítulo 3. Permanecendo apenas a metodologia específica para este artigo.

Os objetivos da **metodologia A** presente no artigo são:

- a) Processamento digital de imagem RGB com um algoritmo simples de fácil uso;
- b) Detecção e contagem automática de frutas;
- c) Determinação da precisão da detecção do fruto de maçã em sensor de imagem RGB terrestre de curto alcance e baixo custo.

### 4.1 ARTIGO 1: AVALIAÇÃO DO DESEMPENHO DE DETECÇÃO SEMI-AUTOMÁTICA DE FRUTOS DE MAÇÃ EM UM POMAR DE ALTA DENSIDADE USANDO SENSOR DE IMAGEM DIGITAL RGB DE BAIXO CUSTO.

#### 4.1.1 Método para detecção de frutos

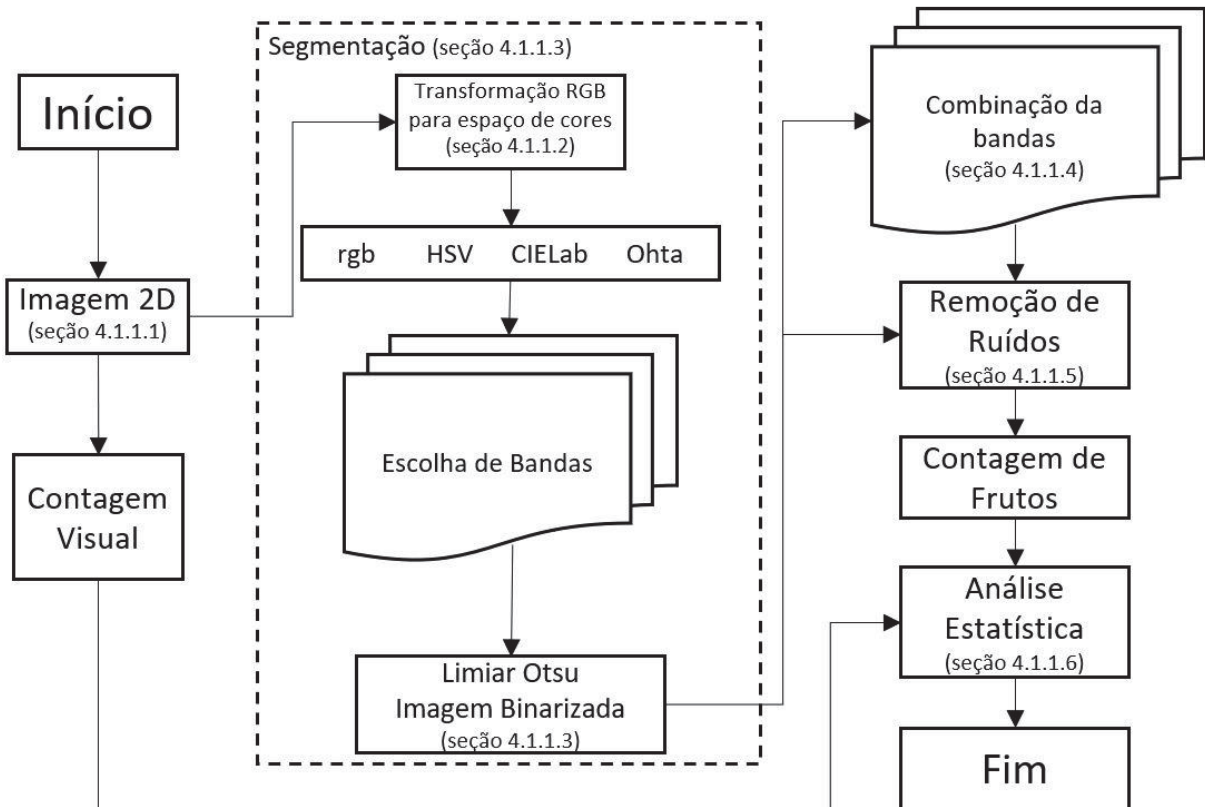
A **metodologia (A)** para contagem de maçãs, desenvolvida nesse estudo está fundamentada na Figura 17. Todos os algoritmos foram desenvolvidos utilizando o ambiente Matlab© instalado em um notebook montado com processador Intel Core i7-6700H, 2,60 GHz, memória de 16 GB e GPU NVIDIA GTX 960M com memória gráfica de 2 GB. A metodologia possui as seguintes etapas principais: 1) aquisição de imagens; 2) transformação do espaço de cores; 3) segmentação; 4) combinação de bandas; 5) remoção de ruído e 6) análise estatística.

---

<sup>1</sup> Artigo publicado no Boletim de Ciências Geodésicas (<https://doi.org/10.1590/s1982-21702021000200014>)

Os detalhes dos métodos utilizados em cada etapa são expostos e melhor explicados nas subseções 4.1.1.1 a 4.1.1.6.

Figura 17 – Fluxograma da metodologia adotada para detecção de frutas de baixo custo.

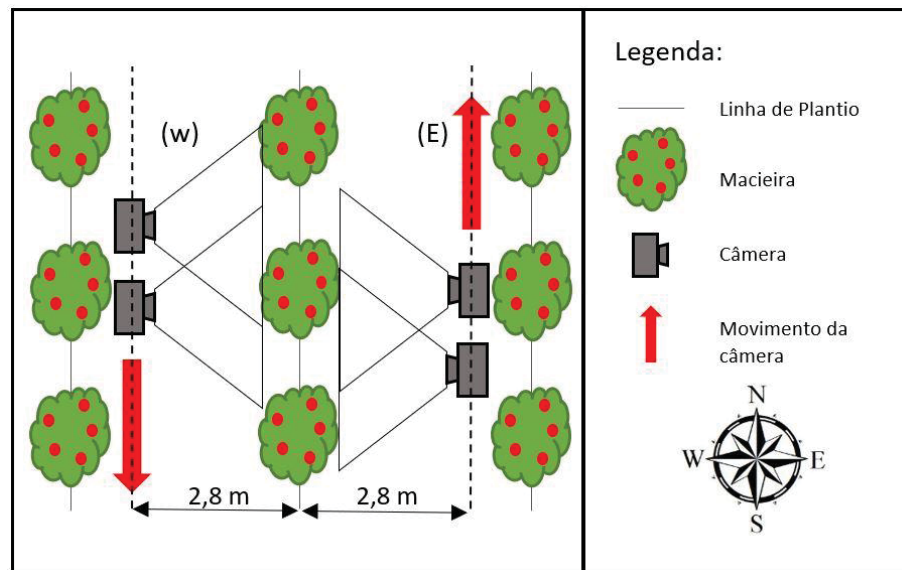


#### 4.1.1.1 Aquisição de imagens

A câmera e o método de aquisição de imagem estão descritos na seção 3.2. A diferença foi que só imagens perpendiculares as plantas da linha L1 foram analisadas. Uma série de imagens consecutivas foram capturadas, simulando um sensor fixado em um veículo em movimento. Ao todo dez imagens foram capturadas em cada lado da linha, resultando em vinte imagens selecionadas para este estudo. Embora o número de imagens seja baixo e considerando que é conduzido apenas em um pomar, a viabilidade de uma proposta de abordagem de baixo custo ainda é válida. A Figura 18 resume a estratégia empregada para tomada das imagens utilizadas neste estudo. A câmera foi acionada manualmente na altura do rosto do operador, e as imagens foram tiradas com a câmera na posição retrato, a uma distância próxima

de 2,8 m da linha de plantio. As plantas do pomar de alta densidade têm altura média de 3,5 – 4,0 m. É importante informar que às vezes o campo de visão da câmera não foi suficiente para fotografar o topo de todas as plantas. O dia estava nublado e o flash auxiliar da câmera estava ativo. As aquisições foram realizadas na data de 29 de março de 2019, entre as 12h00 e 14h00 hora local. Não foi adotado nenhum fundo artificial, de forma que a abordagem proposta represente condições reais de campo.

Figura 18 – Esquema gráfico de vista superior da aquisição de imagens e movimento da câmera.



Fonte: Biffi; Mitishita, E. A.; et al. (2021)

#### 4.1.1.2 Transformação de Espaço de Cores.

As Imagens originais RGB foram transformadas em alguns espaços (por exemplo: HVS, OTHA, CIE  $L^*a^*b^*$  e rgb). Para avaliar o uso de diferentes sistemas de cores e escolher o melhor sistema para detecção de frutos de maçã.

A primeira opção foi normalizar os valores RGB de acordo com a equação (19). Essa transformação permite separar as informações de cor da intensidade (Shaik et al., 2015):

$$\begin{cases} r = \frac{R}{R + G + B} \\ g = \frac{G}{R + G + B} \\ b = \frac{B}{R + G + B} \end{cases} \quad (19)$$

onde:  $r$ ,  $g$  e  $b$  são os valores normalizados das cores vermelha, verde e azul, respectivamente. E os valores  $R$ ,  $G$  e  $B$  são os valores da imagem original de vermelho, verde e azul, respectivamente.

Outro sistema comumente usado é o HSV, que significa os componentes Matiz, Saturação e Valor. É também conhecido como HSB (Matiz, Saturação e Brilho) e é definido da seguinte forma (Dorj et al., 2017):

i) Matiz: é a medida do comprimento de onda médio da luz que o objeto reflete ou emite. Os valores variam de  $0^\circ$  a  $360^\circ$ , representando todas as cores visíveis e podem ser normalizados de 0 a 1.

ii) Saturação: pode ser chamada de “pureza” da cor. Os valores mais baixos são encontrados em tons de cinza. Quanto maior a saturação, mais “pura” é a cor. Varia entre 0 (branco / cinza) a 1 (saturação total).

iii) Valor: representa o brilho da cor e varia entre 0 (preto) e 1 (branco).

O sistema é o resultado de uma transformação não linear, conforme descrito nas equações (20)-(22) (Shaik et al., 2015):

$$H = \arccos\left(\frac{\frac{1}{2}(2R - G - B)}{\sqrt{(R - G)^2 - (R - B)(G - B)}}\right) \quad (20)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (21)$$

$$V = \max(R, G, B) \quad (22)$$

Para a transformação do espaço de cores RGB para o espaço CIE  $L^*a^*b^*$ , foram usadas as fórmulas apresentadas por Jidong et al. (2016):

$$\begin{cases} L = 0.2126R + 0.7152G + 0.0722B \\ a = 1,4749(0.2213R - 0.339G + 0.1177B) + 128 \\ b = 0.6245(0.1949R + 0.6057G - 0.8006B) + 128 \end{cases} \quad (23)$$

Os componentes ortogonais da cor no espaço de cores OHTA ( $I_1$ ,  $I_2$ ,  $I_3$ ) são obtidos aplicando a transformação linear para  $R$ ,  $G$  e  $B$  os componentes são independentes, conforme mostrado na equação (24) (Ohta et al. 1980):

$$\begin{cases} I_1 = (R + G + B)/3 \\ I_2 = (R - B)/2 \\ I_3 = (2G - R - B)/4 \end{cases} \quad (24)$$

Pode-se notar que  $I_1$  na equação 6 é equivalente à intensidade ou brilho. Os outros dois componentes ( $I_2$  e  $I_3$ ) definem um plano perpendicular ao primeiro componente. Existem vários vetores possíveis para representar este plano e pode-se escolher um par que realce melhor uma determinada cor. Para destacar os frutos da maçã vermelha, os vetores da equação (25) foram adotados (adaptado de Ohta et al. 1980):

$$\begin{cases} I''_2 = (R - G) \\ I''_3 = (2R - G - B) \end{cases} \quad (25)$$

#### 4.1.1.3 Segmentação da imagem e binarização.

A segmentação de imagens é um processo computacional que permite dividir a imagem em regiões homogêneas, destacando áreas de interesse na imagem. Um método conhecido é o de limiarização. O limiar consiste em identificar o valor do nível de cinza que permite separar o objeto de interesse do restante da imagem. Como até imagens da mesma área podem ser obtidas em diferentes condições de iluminação, o limiar pode variar de imagem para imagem, o que não é desejado (Jidong et al. 2016). Uma possível solução para este problema é usar um método de limiar dinâmico, onde não é o operador que escolhe o valor de limiar.

O método de limiarização proposto por Otsu (1979) calcula o limiar ótimo “ $T$ ” analisando o histograma da imagem e selecionando o melhor nível de cinza que produz dois agrupamentos: um com valores baixos (plano de fundo) e outro com valores altos, o que permite binarizar a imagem.

Considerando que  $T$  pode assumir quaisquer valores de nível de cinza possíveis na imagem, todos os valores possíveis são avaliados com base nas probabilidades calculadas para os agrupamentos resultantes:  $w_0$ ,  $\mu_0$ ,  $w_1$ ,  $\mu_1$  e  $\sigma^2_B$ . Primeiro, para cada limite possível é calculado: a média dos valores de cinza no primeiro plano e no plano de fundo ( $\mu_0 \mu_1$ ) e a média global ( $\mu_T$ ).

Então, funções de probabilidade são calculadas para cada agrupamento:

- i)  $w_0$  é a probabilidade do agrupamento em primeiro plano, dada pela razão entre o número de pixels em primeiro plano e o número total de pixels.
- ii)  $w_1$  é a probabilidade do plano de fundo, dada pela razão entre o número de pixels do plano de fundo e o número total de pixels.

O valor de limiar ótimo “ $T$ ” será aquele que maximiza a variância entre as classes ( $\sigma_B^2$ ), conforme Equações (26)-(28 (Otsu, 1979).

$$\mu_T = w_0 \times \mu_0 + w_1 \times \mu_1 \quad (26)$$

$$\sigma_B^2 = w_0 \times (\mu_0 - \mu_T)^2 + w_1 \times (\mu_1 - \mu_T)^2 \quad (27)$$

$$\sigma_B^2 = w_0 \times w_1 \times (\mu_0 - \mu_1)^2 \quad (28)$$

#### 4.1.1.4 Combinação de bandas

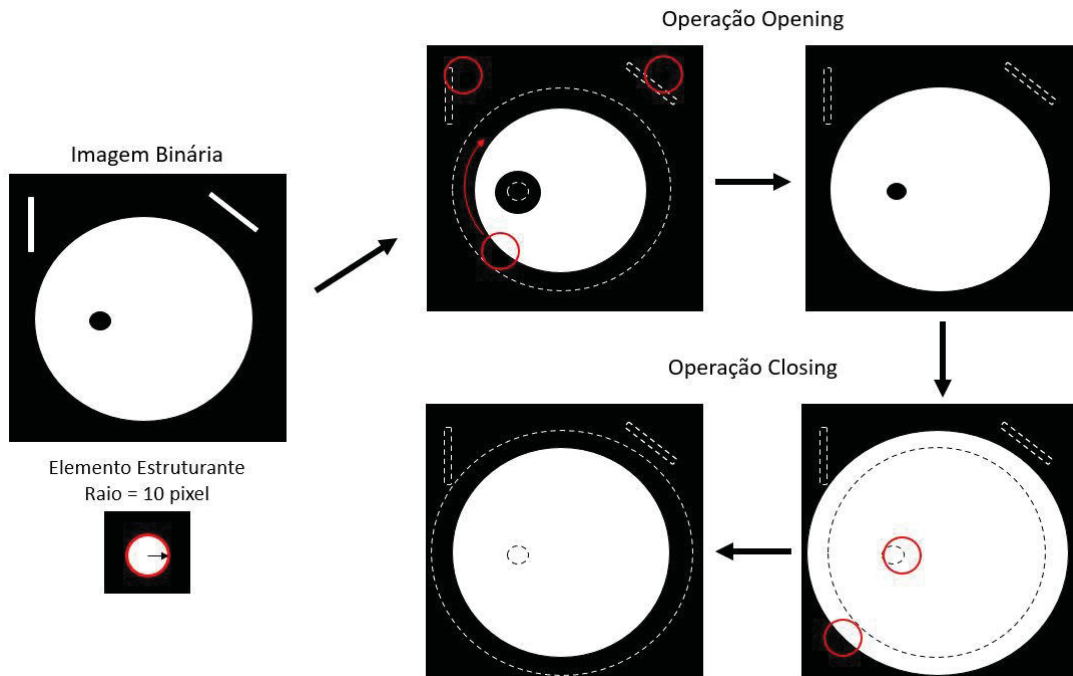
Visando obter melhores respostas para detecção de frutos de maçã, apresenta-se uma proposta de combinação de bandas das imagens, que tiveram respostas positiva dos frutos, após a etapa de segmentação. A ideia principal por trás deste procedimento é selecionar as imagens da seção 4.1.1.3 e combiná-las de forma aditiva. Como as imagens são binárias, elas possuem apenas os valores 0 e 1, correspondentes à identificação do plano de fundo e do fruto, respectivamente. Assim, a combinação de duas imagens (por exemplo) com áreas que correspondem ao mesmo fruto detectado resultaria em uma nova imagem com valor 2 nesta área do fruto.

As áreas de plano de fundo correspondentes em ambas as imagens permanecem com valor 0 e as áreas correspondentes que diferem em valores em ambas as imagens assumirão o valor 1. Assim, um limiar simples ( $T$ ) de  $n-1 < T < n$ , onde  $n$  é o número de imagens usado na combinação, ajudaria na remoção de ruídos discrepantes provenientes das bandas utilizadas.

#### 4.1.1.5 Remoção de ruídos da imagem

Após a binarização das imagens, pequenas regiões e espaços em branco (Ruídos) foram removidos aplicando operações morfológicas matemáticas, como *opening and closing*, utilizando como elemento estruturante um círculo de raio de 10 pixels (Figura 19). Os erros mais comuns são causados por bordas de galhos, pontas de folhas, frutos pequenos nas fileiras do fundo.

Figura 19 – Representação da remoção de ruído de imagem binária, aplicando operação morfológica de opening and closing, utilizando círculo como elemento estruturante.



#### 4.1.1.6 Análise estatística

Foram analisados os resultados do processo de detecção de frutos de cada imagem, em correspondência a contagem visual de frutos realizada em laboratório em cada imagem RGB. Portanto, cada fruta da maçã recebeu uma marca de localização na imagem RGB. Assim, todos os frutos da maçã visíveis ou parcialmente visíveis na planta, bem como os frutos no solo, foram marcados e contados.

Os frutos detectados que coincidiram com as marcações visuais foram contados como verdadeiro positivo (VP). Frutos detectados onde não ocorreu marcação visual foi considerado um falso positivo (FP). Todos os frutos restantes que não foram detectados pelo algoritmo foram contados como falso negativo (FN).

Gráficos foram gerados considerando a relação existente entre o número de objetos detectados versus o número de maçãs da contagem visual. Da mesma forma, o número de acertos de acordo com o algoritmo selecionado versus número de maçãs da contagem visual. Tal análise permite determinar o coeficiente de determinação ( $R^2$ ). Como análise dos dados, erros de omissão, comissão e acurácia foram

calculados para cada imagem usando as equações (29) a (31) (adaptado de Congalton, 1991).

$$Omissão = \frac{FN}{VP} \quad (29)$$

$$Comissão = \frac{FP}{VP} \quad (30)$$

$$Acurácia = \frac{VP}{VP + FN} \quad (31)$$

onde: *FN* - falso negativo; *VP* - verdadeiro positivo e *FP* - falso positivo

Os valores de erro de omissão são aqueles que não diferenciam a maçã do ambiente de fundo, deixando de marcá-los. E erros de comissão ocorrem ao detectar feições no ambiente que não correspondem aos frutos, e estes são contabilizados.

#### 4.1.2 Resultados

As imagens adquiridas em campo resultaram na amostragem de sete plantas inteiras. Considerando a sobreposição das cenas, foram necessárias vinte imagens, dez de cada lado da linha de plantio selecionada (Figura 18). O tempo médio de processamento da imagem para todas as cores do espaço até a contagem dos frutos é um processo que levou 12,4 segundos. A transformação da cor (por exemplo Lab-a) até a contagem dos frutos levou 4,2 segundos por imagem, o que resultou em um tempo total de 84 segundos. Na marcação visual, um total de 4.056 frutos visíveis foram contados e marcados (Tabela 2).

Tabela 2 – Valores das amostras após aquisição das imagens e contagem visual dos frutos.

<b>Amostras</b>	<b>Número Total</b>
Imagens	20
Macieiras	7
Frutos de maçã CV*	4.056

\* CV = Contagem Visual

Utilizou-se o subconjunto da imagem RGB terrestre mostrado na Figura 20 como um exemplo para destacar as diferentes etapas de processamento de imagem digital propostas neste experimento. As Figura 21 e Figura 22 mostram as sequência de segmentação e remoção de ruído, utilizando a metodologia de transformação do

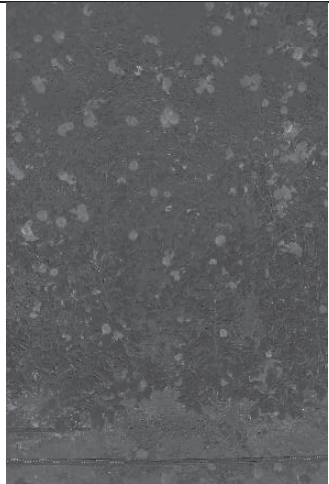
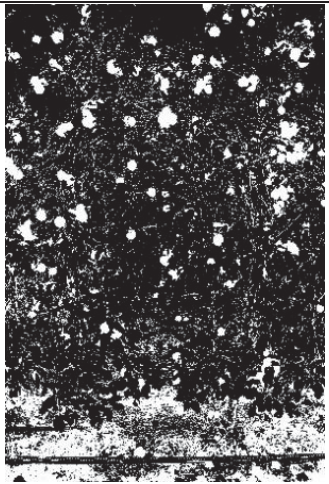
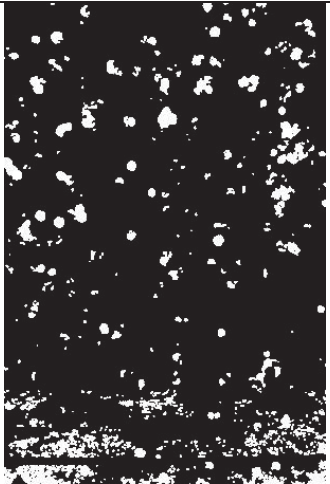
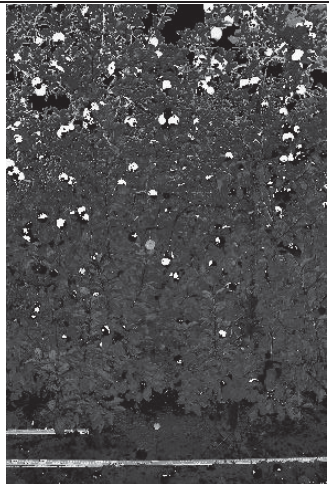
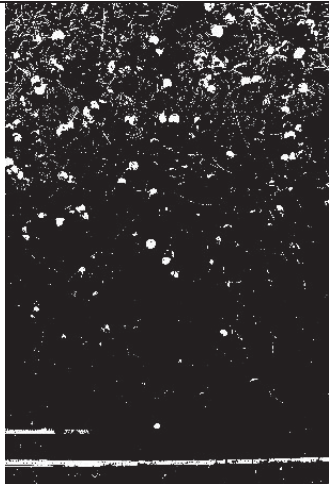

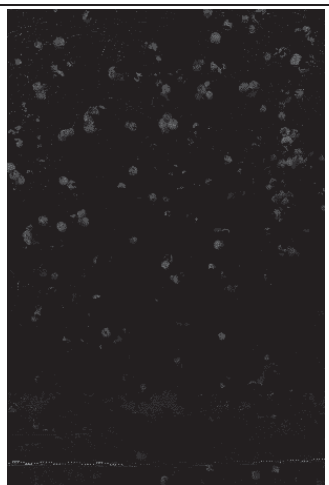
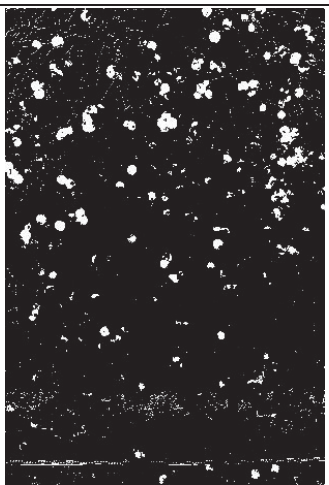
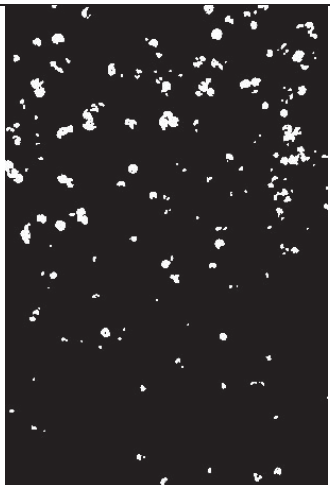
espaço de cores. Apenas as bandas que mostraram realce para os frutos foram selecionadas, ou seja,  $r$  do  $rgb$  normalizado ( $rgb-r$ ),  $h$  do espaço HSV ( $HSV-h$ ),  $a$  do espaço  $L^*a^*b^*$  ( $Lab-a$ ) e bandas  $l''_2$  e  $l''_3$  do espaço Ohta (Figura 21 (a) e Figura 22(a)).

Figura 20 –Imagem utilizada como exemplo para os resultados de realce para os frutos, imagem binária e remoção de ruídos.



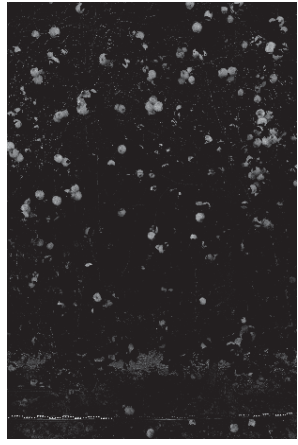
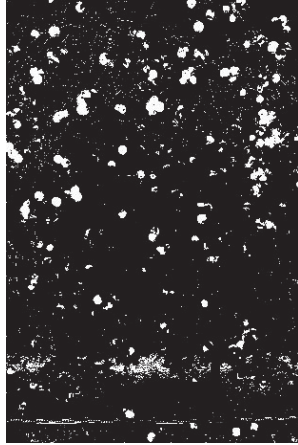
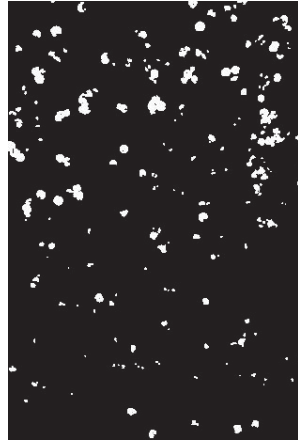
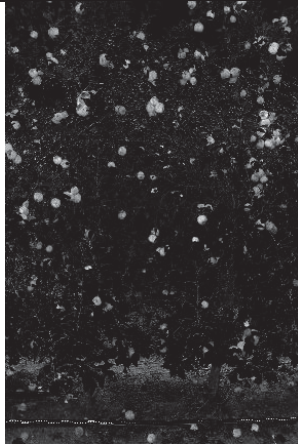
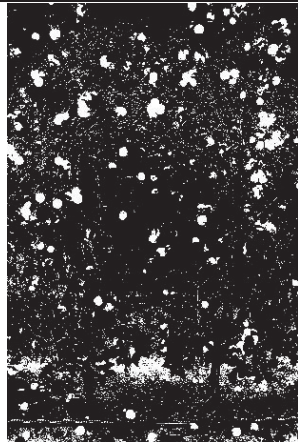
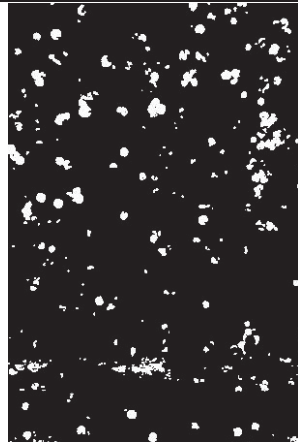
Fonte: O autor (2019)

Figura 21 – Quadro dos resultados das transformações do espaço de cores, imagem binária e remoção de ruídos. Bandas apresentadas: *rgb-r*, *HSV-h* e *L\*a\*b\*-a*.

	a) Realce para Frutos	b) Imagem Binária	c) Remoção de Ruídos
<i>rgb-r</i>			
<i>HSV-h</i>			
<i>L*a*b*-a</i>			

Fonte: O autor (2019)

Figura 22 - Quadro dos resultados das transformações dos espaços de cores Ohta, imagem binária e remoção de ruídos. Bandas apresentadas:  $I''_2$  e  $I''_3$ .

	a) Realce para frutos	b) Imagem Binária	c) Remoção de Ruídos
$Ohta - I''_2$			
$Ohta - I''_3$			

Fonte: O autor (2019)

A análise das imagens geradas após o processo de filtragem (Figura 21(c) e Figura 22(c)) mostra que algumas bandas (rgb-r, HSV-h e  $I''_3$ ) continuam a ter ruído na parte inferior da imagem, correspondendo ao solo e à escala de estádia de referência. Uma solução simples para esse problema foi fazer combinações entre bandas usando imagens binárias. As combinações realizadas foram:

$$\begin{cases} C1 = h + a + I''_3 \\ C2 = r + a + I''_3 \\ C3 = a + I''_2 + I''_3 \end{cases} \quad (32)$$

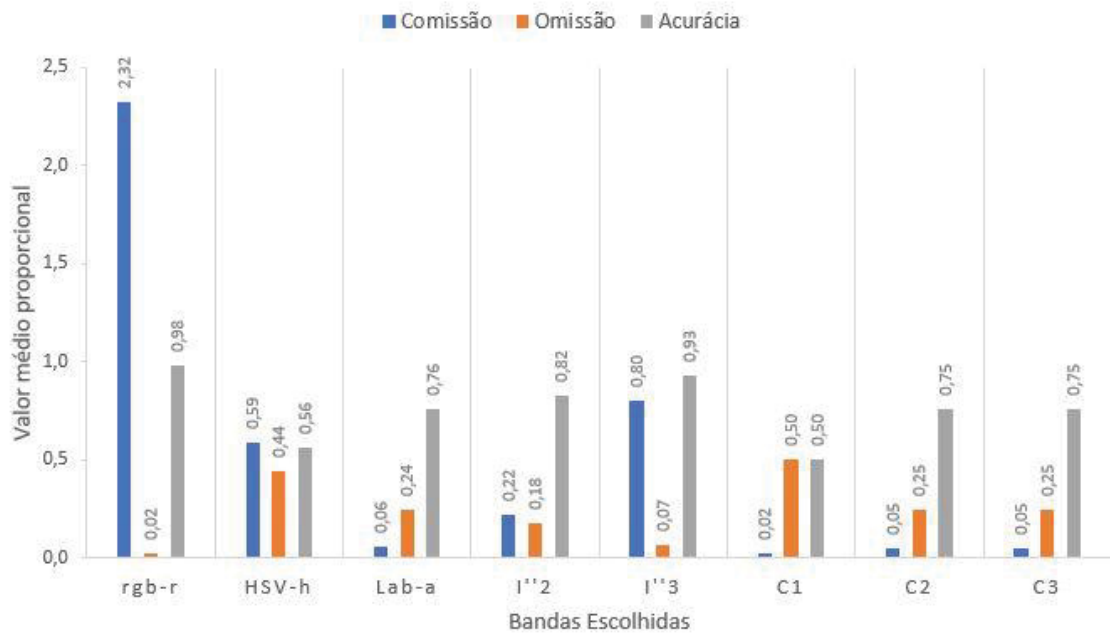
As novas três imagens geradas foram submetidas ao processo de erosão e dilatação. As bandas selecionadas e as imagens combinadas seguiram para a etapa de contagem das estruturas que o algoritmo identificou como frutos. Os valores médios das 20 imagens são apresentados na Tabela 3.

Tabela 3 – Valores de médias de contagem visual dos frutos, contagem automática, comissão, omissão e valores de acurácia para as 20 imagens e para as imagens dos lados W e E, separados.

		<b>Bandas</b>								
<b>N = 20 imagens</b>		<b>RGB</b>	<b>rgb-r</b>	<b>HSV-h</b>	<b>Lab-a</b>	<b><math>I''_2</math></b>	<b><math>I''_3</math></b>	<b>C1</b>	<b>C2</b>	<b>C3</b>
<b>CV</b>	$\mu$	203	-	-	-	-	-	-	-	-
	$\sigma$	45.5	-	-	-	-	-	-	-	-
<b>CA</b>	$\mu$	-	632	210	144	185	318	93	142	142
	$\sigma$	-	129.5	58.0	23.0	27.0	32.0	13.7	24.3	23.7
<b>Comissão</b>	$\mu$	-	2.32	0.59	0.06	0.22	0.80	0.02	0.05	0.05
	$\sigma$	-	1.18	0.42	0.03	0.18	0.31	0.01	0.02	0.02
<b>Omissão</b>	$\mu$	-	0.02	0.44	0.24	0.18	0.07	0.50	0.25	0.25
	$\sigma$	-	0.02	0.07	0.03	0.03	0.02	0.07	0.03	0.03
<b>Acurácia</b>	$\mu$	-	0.98	0.56	0.76	0.82	0.93	0.50	0.75	0.75
	$\sigma$	-	0.02	0.07	0.03	0.03	0.02	0.07	0.03	0.03
		<b>Lado W</b>								
<b>N = 10 imagens</b>		<b>RGB</b>	<b>rgb-r</b>	<b>HSV-h</b>	<b>Lab-a</b>	<b><math>I''_2</math></b>	<b><math>I''_3</math></b>	<b>C1</b>	<b>C2</b>	<b>C3</b>
<b>CV</b>	$\mu$	243 <i>a*</i>	-	-	-	-	-	-	-	-
	$\sigma$	24.9	-	-	-	-	-	-	-	-
<b>CA</b>	$\mu$	-	575	191	163	193	323	102	163	162
	$\sigma$	-	108.6	58.5	16.2	24.1	36.9	6.8	17.1	16.4
<b>Comissão</b>	$\mu$	-	1.53	0.33	0.04	0.11	0.55	0.01	0.04	0.04
	$\sigma$	-	0.53	0.22	0.01	0.07	0.18	0.01	0.01	0.01
<b>Omissão</b>	$\mu$	-	0.04	0.48	0.24	0.17	0.07	0.53	0.25	0.25
	$\sigma$	-	0.02	0.06	0.04	0.04	0.03	0.05	0.04	0.04
<b>Acurácia</b>	$\mu$	-	0.96	0.52	0.76	0.83	0.93	0.47	0.75	0.75
	$\sigma$	-	0.02	0.06	0.04	0.04	0.03	0.05	0.04	0.04
		<b>Lado E</b>								
<b>N = 10 imagens</b>		<b>RGB</b>	<b>rgb-r</b>	<b>HSV-h</b>	<b>Lab-a</b>	<b><math>I''_2</math></b>	<b><math>I''_3</math></b>	<b>C1</b>	<b>C2</b>	<b>C3</b>
<b>CV</b>	$\mu$	162 <i>b*</i>	-	-	-	-	-	-	-	-
	$\sigma$	13.6	-	-	-	-	-	-	-	-
<b>CA</b>	$\mu$	-	689	229	124	176	312	83	121	121
	$\sigma$	-	123.5	50.6	5.2	27	25	12.2	6	6
<b>Comissão</b>	$\mu$	-	3.11	0.85	0.07	0.34	1.05	0.03	0.06	0.06
	$\sigma$	-	1.12	0.41	0.03	0.18	0.17	0.02	0.03	0.03
<b>Omissão</b>	$\mu$	-	0.01	0.40	0.24	0.18	0.07	0.47	0.24	0.24
	$\sigma$	-	0.01	0.05	0.03	0.03	0.01	0.07	0.02	0.02
<b>Acurácia</b>	$\mu$	-	0.99	0.60	0.76	0.82	0.93	0.53	0.76	0.76
	$\sigma$	-	0.01	0.05	0.03	0.03	0.01	0.07	0.02	0.02

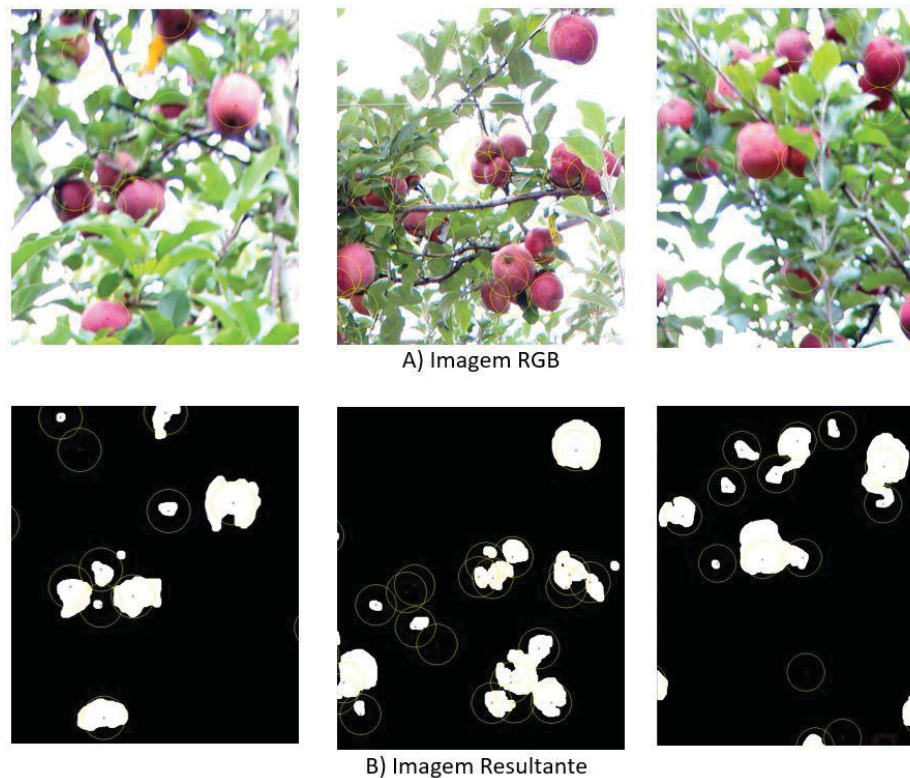
CV = Contagem Visual; CA = Contagem Automática; \*Médias seguidas pela mesma letra, não diferem significativamente entre si a 5% de probabilidade pelo teste de Tukey.

Figura 23 – Gráfico de comparação da média geral das imagens de Comissão, omissão e precisão por método de detecção.



Um exemplo dos resultados das detecções é apresentado na Figura 24. Com a visualização de sucesso e de falhas nas detecções.

Figura 24 – Contagem visual de frutos nas imagens RGB original (A) e resultados das detecções com o algoritmo em uma imagem resultante *Lab-a* (B)



Fonte: O autor (2019)

#### 4.1.3 Discussão

Como pode ser visto na Figura 23 as bandas que apresentaram erros de comissão com valores superiores a 20% (0,2) foram: *rgb-r*, *HSV-h*,  $I''_2$  e  $I''_3$ . Esses altos valores de erro mostram a impossibilidade de usar essas bandas como única banda para detecção de frutas de maçã. O erro de comissão traz mais problemas porque falsos positivos estão sendo contados como frutos de maçã. A Figura 23 mostra que a banda *rgb-r* apresentou maior número de detecções de objetos do que o número de frutos existentes nas imagens. A maior parte desse erro ocorreu devido à porção de solo exposta, apresentada na entre linha que aparece nas imagens, que possui uma coloração que não se diferenciou do fruto de maçã, neste espaço de cor. O mesmo problema foi detectado com o uso da banda  $I''_3$ .

As bandas com erros de omissão correspondentes a valores superiores a 20% (0,2) foram: *HSV-h*, *Lab-a*, C1, C2 e C3. Neste caso, o limiar dinâmico adotado falhou em diferenciar alguns frutos do ambiente de fundo complexo. Wei et al. (2014) também não foram capazes de identificar frutos verdes e frutos maduros. Esta omissão ocorreu principalmente nos frutos com alta taxa de oclusão que prejudicou a luminosidade no fruto. Este tipo de erro foi maior nos frutos que estavam na parte inferior da planta e dentro da camada interna da planta.

O grau de acurácia nas bandas também é apresentado (Figura 23). Os acertos em relação ao total de frutos de maçã nas imagens que ultrapassaram 75% (0,75) ocorreram com o uso das bandas: *rgb-r*, *Lab-a*,  $I''_2$ ,  $I''_3$ , C2 e C3. Os maiores valores de acerto foram *rgb-r*,  $I''_2$  e  $I''_3$ , todos acima de 80% (0,8). Porém, essas bandas também apresentaram altos valores de erro de comissão. Na prática, isso inviabiliza o uso isolado dessas bandas como método autônomo de detecção de frutos de maçã nas condições testadas neste estudo. As Figura 21 e Figura 22 mostram em detalhes a detecção dos frutos da maçã nessas bandas. No entanto, a maior fonte de erro ocorreu em detecção de elementos como frutos localizados no solo.

Assim, as únicas bandas que permaneceram na análise foram *Lab-a*, C2 e C3. Os três apresentaram valores praticamente idênticos de comissão (0,06; 0,05; 0,05), omissão (0,24; 0,25; 0,25) e acurácia (0,76; 0,75; 0,75), respectivamente. Esses valores corroboram o estudo realizado por Linker; Kelman (2015). Esses autores obtiveram valor de  $R^2$  de 0,76 na detecção de maçãs, com taxa de acerto de 95% quando comparada aos frutos contados na imagem. Seus valores de comissão,

omissão e precisão foram de 16%, 28% e 72%, respectivamente. Valores muito semelhantes aos encontrados para a banda *Lab-a* neste estudo.

Gongal et al. (2016) descreveram 78,9% na precisão dos acertos na detecção de frutos de maçã em avaliação completa da planta usando um conjunto de câmeras 2D e 3D. No entanto, a discreta melhoria implica um projeto de custo elevado e processamento de dados mais complexos. Zhou et al. (2012) apresentam coeficiente de correlação de 0,85 entre a contagem manual de frutos e a contagem pelo algoritmo de característica de cor antes da colheita.

Outros estudos, que não consideram a planta inteira, onde o único propósito de identificar o fruto por visão robótica, e analisar a frutificação de forma isolada, alcançaram taxas de acerto maiores, como 86-100% (Jidong et al., 2016) e 95 -100% (Wei et al., 2014). As bandas combinadas (C2 e C3) derivam de uma combinação com *Lab-a*, que se mostra a variável limitante nessas combinações. Portanto, por se originar de uma transformação direta e não apresentar diferenças das bandas combinadas (C2 e C3), o uso da banda *Lab-a* é uma alternativa para a detecção de frutos da maçã Fuji Suprema limitada às mesmas condições deste experimento.

O componente *a* do *Lab* representa valores em um eixo que variam entre vermelho (valores positivos) e verde (valores negativos), e no componente *b* os valores variam entre amarelo (valores positivos) e azul (valores negativos). Como os frutos da maçã enfatizam fortemente a cor vermelha e tanto as folhas quanto a cobertura do solo tendem ao verde, a diferenciação no mesmo eixo facilita a divisão desses objetos na imagem (Kahu et al., 2019).

A aplicação de uma análise de regressão entre o número de objetos detectados pelos Componentes Conectados e o número de frutos da contagem visual resultou em um coeficiente de determinação ( $R^2$ ) de 0,73, mas quando comparado com o cruzamento entre o número de acertos do algoritmo e o número de frutos da contagem visual,  $R^2$  é 0,96 (Figura 25). Isso ocorreu porque o algoritmo não conseguiu segmentar os agrupamentos de frutas em frutas individuais, e a contagem resultante dos objetos detectados era por meio do método de Componentes Conectados. Considerando este agrupamento apenas como um objeto detectado, onde havia dois ou mais frutos (Figura 26 A e B). Ainda em análise a mesma Figura 27 B, para calcular o número de acerto do algoritmo (Figura 25 B), aplicou-se uma detecção simples de modo inverso. Utilizou-se o objeto detectado cruzando com as marcações de contagem visual, nesse caso um elemento acerta a posição de quatro frutos.

Figura 25 – Regressões entre número de frutos contados visualmente na imagem RGB e: o número de objetos detectados no algoritmo nas imagens Lab-a (A); o número de acertos do algoritmo (B).

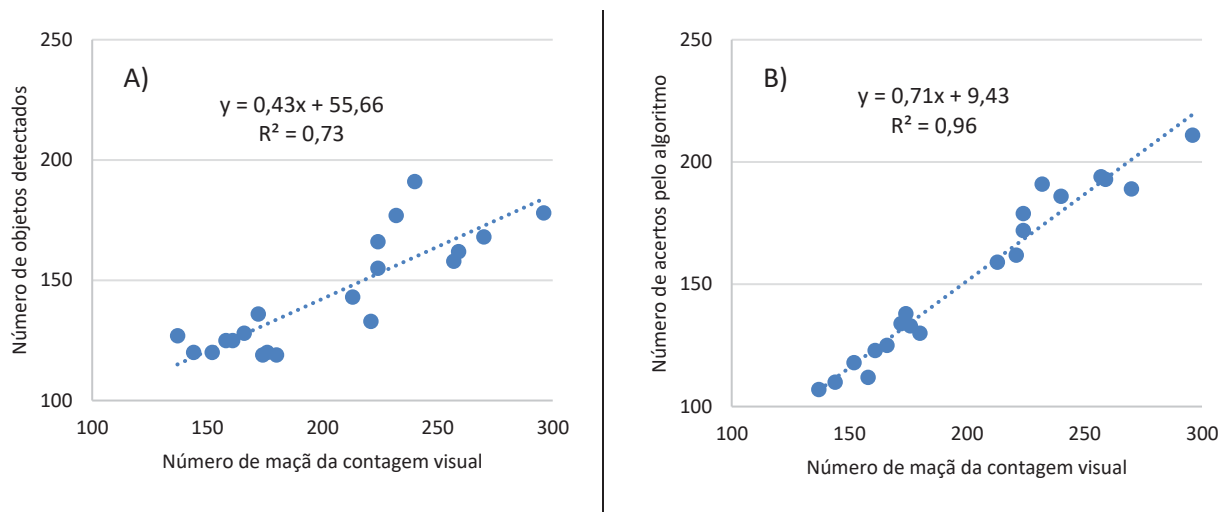


Figura 26 - Frutos aglomerados (A) e consideradas como um único objeto por terem as bordas unidas (B)



Fonte: O autor (2019)

#### 4.1.4 Conclusão e Recomendações

A detecção de frutos de maçã maduros antes da colheita, por um algoritmo semi-automático em imagem de câmera RGB comum, onde a imagem abrange toda a planta, foi testado e obteve um desempenho semelhante a outras metodologias que foram referenciadas nesse estudo, e que usaram algum anteparo como segundo plano.

Não utilizar uma tela como segundo plano pode representar grandes problemas de confusão de detecção, mas a transformação do espaço de cores *Lab-a* supriu essa adversidade. A taxa de detecção ( $R^2 = 0,73$ ) foi compatível com outros estudos, assim como a taxa de acurácia de reconhecimento de frutos ( $R^2 = 0,96$ ).

Frutos verdes com pouca iluminação não foram detectados na metodologia proposta, o que é uma limitação de uso neste caso. Técnicas como o aprendizado profundo que dependem de imagens RGB originais podem ser usadas para resolver o problema e também são sugeridas para projetos futuros.

Os índices de acertos podem melhorar sob sistemas de condução de plantas que deixam os frutos mais expostos. Isso pode contribuir para a redução do volume da copa das árvores, conferindo uma característica da forma bidimensional. Paralelamente contribui para um aumento na taxa de acertos dos frutos, diminuindo o erro de omissão.

Altos valores de detecção precoce relativos à colheita melhoram índices de previsão de rendimento. Este estudo contribuiu assim para um método de baixa intervenção do usuário e processamento simples para detectar frutos de maçã. Mas para melhorar os índices de número de detecção de frutos em imagens, projetos futuros devem focar na individualização e separação aglomerados de frutos, utilizando técnicas que, além de explorar espaços de cores, considerem outros atributos da fruta como forma e textura. Estudos em outros pomares e sob diferentes condições de iluminação também são recomendados para confirmar os resultados alcançados neste experimento.

Cálculos de estimativa de produção também devem ser explorados em estudos futuros para auxiliar na tomada de decisão por gestores de pomares, bem como explorar o uso da segmentação de cores na fruticultura, seja na detecção de frutos, folhas, troncos e galhos ou lesões causadas por doenças e pragas. Testes com outros espaços de cores combinados com técnicas de aprendizado profundo são alternativas promissoras para a pesquisa com frutas. Finalmente, as melhorias na apresentação da abordagem proposta dentro de ferramentas amigáveis que permitem seu uso prático em PCs ou smartphones são recomendações para estudos futuros.

## 5 METODOLOGIA B:

Este capítulo corresponde ao artigo publicado no periódico *Remote Sensing*, volume 13, número 1 de 2021 (BIFFI et al., 2021b)<sup>2</sup>. O artigo foi elaborado e analisado em conjunto com a equipe dos pesquisadores Dr. José Marcato Junior e Dr. Wesley Nunes Gonçalves, assim como uso de recursos da Universidade Federal do Mato Grosso do Sul. O artigo original foi publicado em língua inglesa e para facilitar a compreensão e leitura da técnica no país o artigo foi traduzido para língua portuguesa. Parte do Material e Métodos do artigo foi suprimida, como local de estudo e parte da metodologia que corresponde a metodologia geral descrita no capítulo 3. A metodologia específica para este artigo segue descrita na seção 5.1.2.

### 5.1 ARTIGO 2: ABORDAGEM BASEADA EM DEEP LEARNING ATSS PARA DETECTAR FRUTOS DE MAÇÃ.

#### 5.1.1 Introdução

Os sistemas de Sensoriamento Remoto têm ajudado com eficácia em diferentes áreas de aplicações por um longo período de tempo. No entanto, nos últimos anos, a integração desses sistemas com técnicas de aprendizado de máquina tem sido considerada o estado da arte para atender diversas áreas do conhecimento, incluindo a agricultura de precisão (Dian Bah et al. 2018, e Hunt et al. 2019, e Kamilaris & Prenafeta-Boldú 2018). Um subgrupo dentro dos métodos de aprendizado de máquina refere-se a *Deep Learning* (DL), que geralmente é arquitetado para considerar uma rede mais profunda na solução de tarefas com muitas camadas de transformações não lineares (Ball et al., 2017). Um modelo DL tem três características principais: (1) pode extrair recursos diretamente do próprio conjunto de dados; (2) pode aprender características hierárquicas que aumentam em complexidade através da rede profunda, e (3) pode ser mais generalizado em comparação com uma abordagem de aprendizado de máquina mais rasa, como máquina de vetor de suporte, floresta aleatória, árvore de decisão, entre outros (Ball et al., 2017).

---

<sup>2</sup> Artigo publicado na *Remote Sensing* (<https://doi.org/10.3390/rs13010054>)

A extração de características com métodos baseados em aprendizado profundo é encontrada em várias aplicações com imagens de Sensoriamento Remoto (Deng et al., 2018, Osco et al. 2020, Zhang et al. 2020). Essas redes profundas são construídas com diferentes tipos de arquiteturas que seguem um tipo hierárquico de aprendizagem. Neste aspecto, as arquiteturas frequentemente adotadas nos últimos anos incluem Redes Pré-treinadas Não Supervisionadas (UPN), Redes Neurais Recorrentes (RNN) e Redes Neurais Convolucionais (CNN). CNN é uma classe DL, sendo a mais utilizada para análise de imagens. CNNs podem ser usados para problemas de segmentação, classificação e detecção de objetos. O uso de métodos de detecção de objetos em Sensoriamento Remoto tem aumentado nos últimos anos. Li et al. (2020) conduziram uma pesquisa de revisão e propuseram um grande benchmark para detecção de objetos. Os autores consideraram imagens orbitais e 20 classes de objetos. Eles investigaram 12 métodos e verificaram que o RetinaNet superou ligeiramente os outros no benchmark proposto. No entanto, novos métodos já foram propostos que ainda não foram investigados em vários contextos.

Em problemas de agricultura de precisão, a integração de métodos DL e dados de Sensoriamento Remoto também apresentou melhorias notáveis. Recentemente, Osco et al. (2020) propuseram um método baseado em CNN para contar e localizar geograficamente as árvores cítricas em um pomar altamente denso usando imagens adquiridas por uma aeronave pilotada remotamente (APR). Mais especificamente para contagem de frutos, Apolo-Apolo, et al. (2020a) realizaram uma técnica de Faster R-CNN para a estimativa do rendimento e tamanho de frutas cítricas usando imagens coletadas por APR. Da mesma forma, Apolo-Apolo, et al. (2020b) gerou mapas de estimativa de rendimento de pomares de maçã usando imagens APR e uma técnica R-CNN.

Muitos outros exemplos são apresentados na literatura sobre detecção e contagem de objetos na agricultura, como para morango (Habaragamuwa et al., 2018, e Kirk et al., 2020), laranja e maçã (Liu et al., 2018), maçã e manga (Bargoti; Underwood, 2017), manga (Koirala et al., 2019), entre outros. Uma recente revisão da literatura (Koirala et al., 2019) sobre o uso de métodos DL para detecção de frutas destacou a importância da criação de novos métodos para facilitar o processo de rotulagem, uma vez que a anotação manual para obter conjuntos de dados de treinamento é uma tarefa trabalhosa e demorada. Além disso, os autores argumentaram que os modelos DL geralmente superam as abordagens de

segmentação por pixel com base em métodos de aprendizado de máquina tradicionais na tarefa de detecção de frutas na planta.

Em pomares de maçã, Wu et al. (2020) desenvolveram uma abordagem baseada em YOLOv4 para a detecção em tempo real de flores de macieira. Com relação à análise da doença foliar da maçã, outra pesquisa também foi conduzida por Jiang et al. (2019), e uma abordagem DL para detecção de bordas foi proposta para o monitoramento do crescimento da maçã por Wang, D. et al. (2020). Portanto, na detecção de frutos de maçã, trabalhos recentes foram desenvolvidos com base na adoção da CNN. Tian et al. (2019), aprimoraram o método YOLOv3 para detectar maçãs em diferentes estágios de crescimento. Uma abordagem para a detecção em tempo real de frutos de macieira em pomares foi proposta por Kang & Chen (2020) mostrando uma acurácia que varia de 0,796 a 0,853. A abordagem Faster R-CNN mostrou-se ligeiramente superior aos métodos restantes. Gené-Mola et al. (2019) adaptaram o método Faster R-CNN para detectar frutos de maçã usando um sensor RGB-D. Nesse aspecto, esses trabalhos apontam que oclusões podem dificultar o processo de detecção de frutos de maçã. Para resolver isso, Gao et al. (2020) investigaram o uso do Faster R-CNN considerando quatro classes de frutas: não ocluída, ocluído por folha, ocluído por ramo/fio e ocluído por fruta.

Mesmo fornecendo resultados precisos em aplicações atuais, Faster R-CNN (Ren et al., 2017) e RetinaNet (Tsung-Yi Lin et al., 2017) são datados antes de 2018 e, recentemente, novos métodos foram propostos, mas ainda não avaliados na detecção de frutos de maçã. Um deles se refere à Seleção de Amostras de Treinamento Adaptativo (ATSS) (Zhang, S. et al., 2020). O ATSS, diferente de outros métodos, considera um limite individual de interseção sobre união (IoU) para cada caixa delimitadora verdadeira. Consequentemente, são esperadas melhorias na detecção de pequenos objetos ou objetos que ocupam uma pequena área da caixa delimitadora. Apesar de ser um método recente, poucos trabalhos o exploraram, principalmente em aplicações de Sensoriamento Remoto, para atender a problemas relacionados à agricultura de precisão. Além disso, ao considerar imagens RGB, o método pode ser útil para sistemas de detecção com sensores de baixo custo.

Normalmente, a previsão da produção de frutos é feita manualmente, contando os frutos nas árvores selecionadas, e então é feita uma generalização para todo o pomar. Esse procedimento ainda é uma realidade para a maioria (ou até na totalidade) dos pomares do Sul do Brasil. Evidentemente, há algumas previsões

imprecisas devido à variabilidade natural dos pomares. Portanto, técnicas que visam contar os frutos da maçã de forma mais eficiente para cada árvore, mesmo tendo alguns frutos oclusos, auxiliam na previsão do rendimento. O monitoramento dos frutos é uma etapa importante imediatamente após a fase de frutificação até a maturação. Isso porque ocorrem perdas de frutos que caem ao chão devido a condições naturais ou por ações mecânicas, até que aconteça o processo de colheita. É importante mencionar que os frutos da maçã só são visíveis após a fase de frutificação (Srivastava, 2002). Os frutos definem seu formato final, aproximadamente, um mês antes da colheita e podem ser monitorados para verificar perdas de produtividade. Este dado também é uma informação importante para o gestor do pomar e para a previsão da produção de frutas.

É importante destacar que existe o uso crescente de cobertura de tela anti-granizo em pomares, que ajuda a prevenir danos causados por tempestades de granizo. A instalação de telas de proteção sobre pomares de macieiras de alta densidade é fortemente recomendada em regiões onde ocorrem fortes tempestades para evitar perdas na produção e na qualidade dos frutos (Brglez Sever et al., 2020 e Mészáros et al., 2019). Portanto, a proteção reduz a porcentagem de lesões em macieiras durante os estágios críticos da produção, principalmente nas fases de floração e desenvolvimento dos frutos. Como resultado, telas de proteção foram implementadas em diferentes regiões do mundo e no Sul do Brasil (Bosco et al., 2018 e 2020). Porém, o uso de telas dificulta o uso de imagens de Sensoriamento Remoto adquiridas em sistemas orbitais, aéreos e até mesmo de APR para detecção e previsão de produção de frutas.

Portanto, isso destaca a importância do uso de imagens RGB terrestres de curto alcance e baixo custo adquiridas por câmeras profissionais ou semi-profissionais. Esses dispositivos podem ser operados manualmente ou mesmo fixados em máquinas agrícolas ou tratores.

O Sensoriamento Remoto terrestre também é importante neste estudo de caso específico por causa da escala. O Sensoriamento Remoto de curto alcance também permite a detecção de frutos de maçã parcialmente oclusos por folhas e pequenos ramos. Complementarmente, a alta resolução espacial também possibilita a detecção de eventuais injúrias causadas por doenças, pragas e até efeitos climáticos nos próprios frutos, além de folhas, galhos e troncos que não seriam possíveis em imagens adquiridas por sistemas aéreos, dependendo da escala.

Este trabalho propõe uma abordagem baseada no método DL ATSS aplicado a imagens RGB terrestres de curto alcance e baixo custo para detecção automática de frutos de maçã. A principal característica do ATSS é que apenas o ponto central dos objetos é rotulado. É uma abordagem muito interessante, uma vez que a anotação da caixa delimitadora, principalmente em áreas de frutas muito densas, é difícil de executar manualmente além de ser demorada.

Experimentos foram conduzidos comparando esta abordagem com outros métodos de detecção de objetos, considerados o estado da arte em propostas de detecção de objetos nos últimos anos. Os métodos são: *Faster-RCNN*, *Libra-RCNN*, *Cascade-RCNN*, *RetinaNet*, *Feature Selective Anchor-Free - FSAF* e *HRNet*. Além disso, avaliou-se a capacidade de generalização do método proposto aplicando corrupções simuladas com diferentes níveis de severidade, tais como: ruído (Gaussiano, *shot* e *impulse*), borrões (*defocus*, *glass* e *motion*), condições climáticas (neve, geada, neblina e brilho) e processamento digital (transformação elástica, pixelizada e compressão JPEG). Finalmente, avaliou-se a variação do tamanho da caixa delimitadora na aplicação do método ATSS, variando entre 80 x 80 a 180 x 180 pixels o tamanho. Outra contribuição importante deste artigo foi a disponibilidade do conjunto de dados usado para ser implementado em pesquisas futuras que comparem novas redes DL no mesmo contexto.

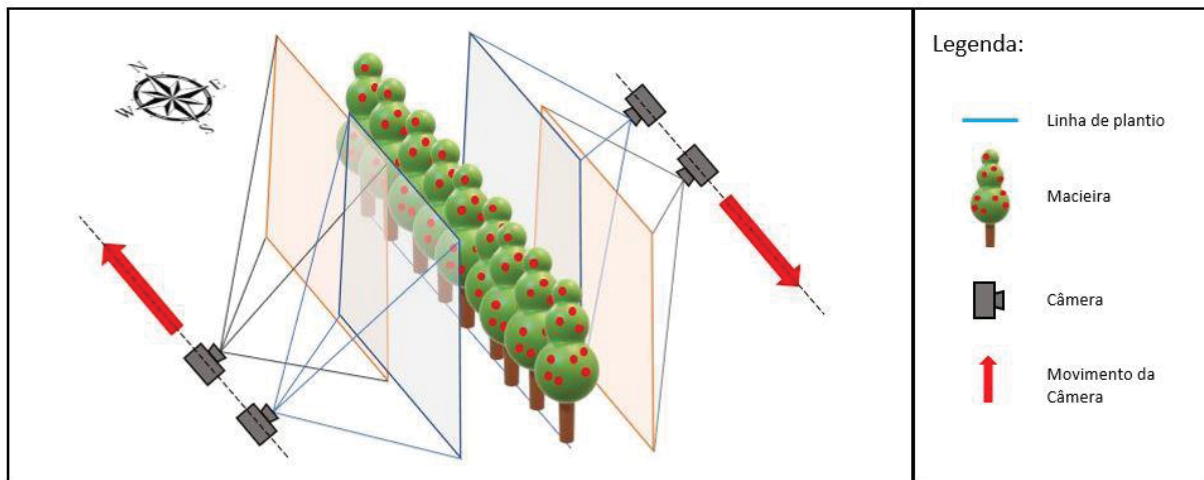
### 5.1.2 Material e Métodos

#### 5.1.2.1 Aquisição de dados

A aquisição de dados com a câmera e a descrição do local do estudo foram descritos nas sessões 3.1 e 3.2. As plantas avaliadas para esse estudo corresponde as 8 Linhas de plantio (L1 a L8, Figura 11D).

Em todas as 8 linhas de plantio, apenas as imagens adquiridas de modo perpendicular as linhas foram analisadas na **metodologia B**. A Figura 27 mostra o modo de aquisição de imagem para esse estudo.

Figura 27 - Modo de obtenção das imagens no Pomar. As sobreposições das imagens também são apresentadas.

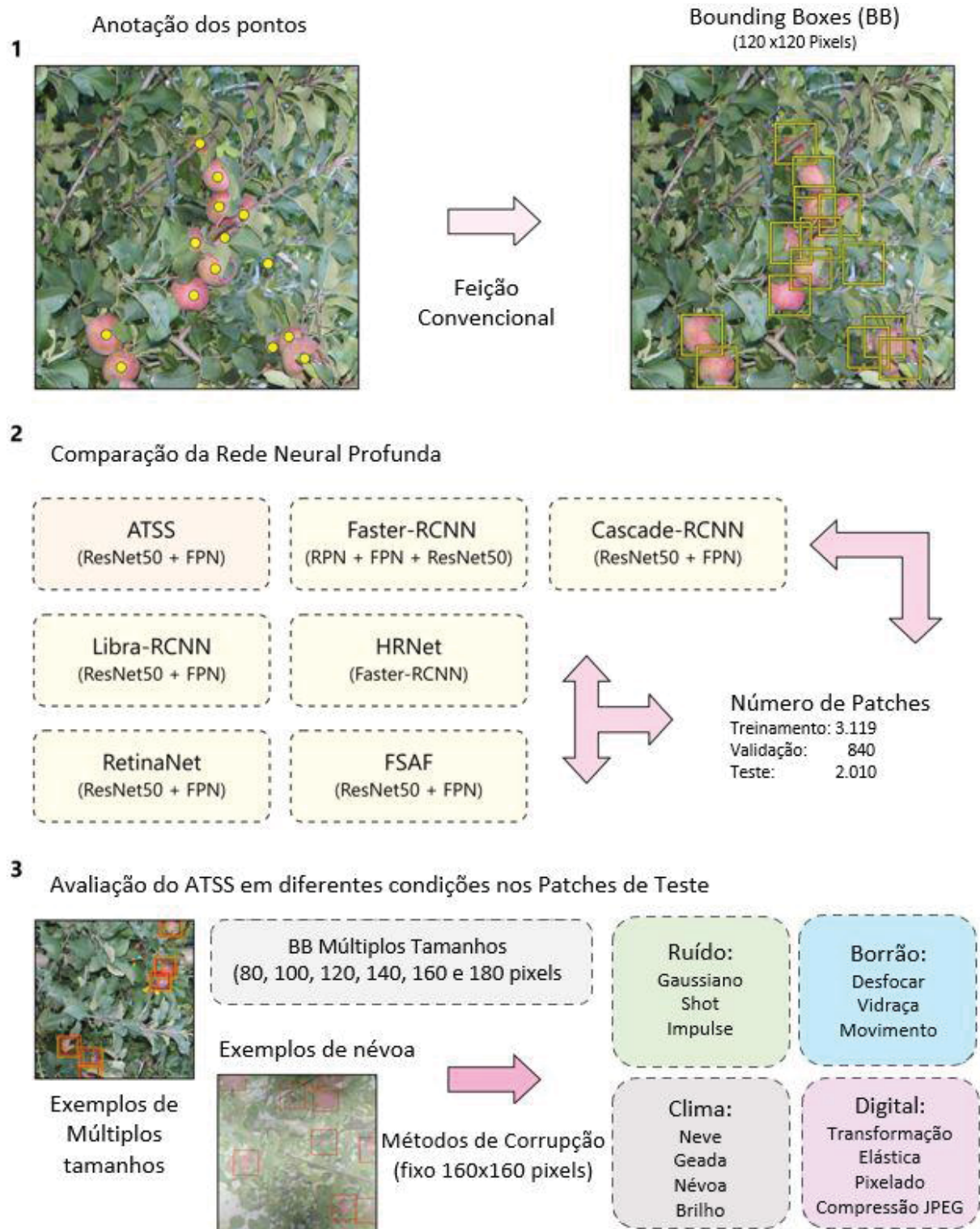


FONTE: O autor (2020).

#### 5.1.2.2 Método de Detecção de Objetos

Após a aquisição das imagens em campo, elas foram descarregadas em um computador com um cartão de memória. E iniciou-se a operação de marcações (anotações) para verdades de campo. As anotações dos frutos da maçã foram realizadas manualmente no aplicativo online *VGG Image Annotator - VIA* (DUTTA; ZISSERMAN 2019), por operadores. Neste procedimento, um ponto de marcação foi anotado no centro de cada fruta. A inspeção de cada ponto foi duplamente checada por um segundo operador. A abordagem resumida idealizada neste estudo é ilustrada na Figura 28 e pode ser resumida em três etapas principais. A primeira etapa (1) ilustra a anotação de ponto no centro de cada fruta da maçã e a configuração do tamanho da caixa delimitadora, com um valor fixo de 120 pixels. Na etapa (2), está destacado o estado da arte dos modelos de detecção de objetos comparados neste experimento, e a distribuição de *patches* nos conjuntos de imagens de treinamento, validação e de teste. A etapa (3) mostra as avaliações extras que o método ATSS foi submetido, como: a configuração do tamanho da caixa delimitadora, cujo valor varia de 80 a 180 pixels com intervalo fixo de 20 pixels, e quatro categorias dos métodos de corrupção de imagem (ruído, borrão, clima e processamento digital) aplicados no conjunto de teste para avaliar a robustez do modelo ATSS.

Figura 28 - O fluxo de trabalho para detecção de maçã está dividido em três etapas principais: (1) processo de anotação, (2) distribuição de *patches* nas imagens de treinamento, validação e teste, e comparação dos métodos aplicados e (3) processos de corrupção de imagem.



FONTE: (BIFFI et al., 2021).

A abordagem ATSS, proposta para detecção de frutos de maçã, foi comparada com seis métodos populares de detecção de objetos, são eles o Faster-RCNN, Libra-RCNN, Cascade-RCNN, RetinaNet, FSAF e HRNet. A escolha desses métodos ocorreu por dois motivos principais: (i) eles constituem o estado da arte em propostas de detecção de objetos nos últimos anos; (ii) cobrem as principais direções da pesquisa de detecção de objetos, tais como os detectores de um e de dois estágios, e detectores baseados em âncora (*anchor-based*) e livre de âncora (*anchor-free*). A seguir, descreve-se resumidamente as principais características de cada método adotado.

**ATSS (Adaptive Training Sample Selection):** Os resultados experimentais mostrados por Zhang, S. et al. (2020) concluíram que uma etapa importante no treinamento dos métodos de detecção de objetos é a escolha de amostras positivas e negativas. Portanto, ATSS propõe uma nova estratégia para selecionar amostras seguindo as características estatísticas dos objetos. Para cada verdade fundamental  $g$ ,  $k$  caixas de âncora, cujo centro está mais próximo do centro de  $g$ , são selecionadas como candidatas positivas. Em seguida, calcula-se a  $IoU$  e o  $g$  entre os candidatos, obtendo-se a média  $m_g$  e o desvio padrão  $\sigma_g$  das  $IoU$ . Um limiar é calculado como  $I_g = m_g + \sigma_g$  e usado para selecionar candidatos positivos finais cujo  $IoU$  é maior ou igual a esse limiar. As amostras restantes são usadas como negativos. Neste experimento, o ATSS tem o ResNet50 e o *Feature Pyramid Network* (FPN) como *backbone*, e  $k = 9$  caixas de âncora que são selecionadas como candidatas positivas.

**Faster-RCNN:** O Faster-RCNN é uma CNN de dois estágios que usa a *Region Proposal Network* (RPN) para gerar caixas delimitadoras candidatas para cada imagem. O primeiro estágio faz a regressão da localização de cada caixa delimitadora candidata e o segundo estágio prevê sua classe. Neste trabalho, construiu-se o FPN no topo da rede ResNet50, conforme proposto por Lin; Dollar; et al. (2017). Essa configuração traz algumas melhorias nos resultados para o Faster-RCNN, uma vez que o ResNet + FPN extrai recursos semânticos ricos nas imagens que alimentam o RPN para ajudar o detector a prever melhor a localização da caixa delimitadora.

**Cascade-RCNN:** Em treinamentos de métodos de detecção de objetos, um limiar  $IoU$  é definido para selecionar amostras positivas e negativas. Apesar dos resultados promissores, o Faster-RCNN usa um limiar fixo de  $IoU$  de 0,5. Um valor baixo para o limiar  $IoU$  durante o treinamento pode produzir detecções com ruídos, enquanto valores altos podem resultar em sobreajuste (*overfitting*) devido à falta de

amostras positivas. Para melhorar esse problema, Cai & Vasconcelos (2018) propõem uma sequência de detectores com aumento dos limiares *IoU* treinando estágio por estágio. A saída de um detector é usada para treinar um detector subsequente. Desta forma, o detector final recebe uma distribuição superior e produz detecções com mais qualidade. Seguindo os autores, usou-se o ResNet50 + FPN como *backbone* e três detectores em cascata com *IoUs* de 0,5, 0,6 e 0,7, respectivamente.

**Libra-RCNN:** Apesar da grande variedade de métodos de detecção de objetos, a maioria deles segue três etapas: (i) proposição de regiões, (ii) extração de características dessas regiões e (iii) classificação de categorias e refinamento de caixas delimitadoras. No entanto, o desequilíbrio nessas três etapas impede o aumento na acurácia geral. Assim, Pang et al. (2019) propõem três componentes para forçar o balanceamento nos três estágios. No primeiro estágio, a proposta de amostragem *IoU*-equilibrado extrai amostras sólidas de acordo com sua *IoU* com a anotação verdadeira, de maneira uniforme. Para reduzir o desequilíbrio na extração de características, no segundo estágio o método propõe a Pirâmide de Característica Equilibrada, que redimensiona os mapas de características do *FPN* para o mesmo tamanho, calcula a média e os redimensiona de volta ao tamanho original. Finalmente no terceiro estágio, a perda L1 balanceada reequilibra as tarefas de classificação e localização. Neste método, usou-se o ResNet50 + FPN como *backbone*.

**HRNet:** O objetivo principal do HRNet (WANG, J. et al., 2020) é manter a extração de recursos em alta resolução em todo o *backbone*. HRNet consiste em quatro estágios, onde o primeiro estágio usa convoluções de alta resolução. O segundo, terceiro e quarto estágios são compostos de blocos de resolução múltipla. Cada bloco de multi-resolução conecta convoluções de alta para baixa resolução em paralelo e fusiona representações de multi-escala através de convoluções paralelas. No final, as representações de baixa resolução são redimensionadas para alta resolução por meio de aumento amostral bilinear. Neste método, usou-se HRNet como *backbone* e Faster-RCNN *Heads* para detectar objetos.

**RetinaNet:** Para melhorar a precisão das detecções de objetos de estágio único com relação aos métodos de dois estágios, o RetinaNet foi proposto por Lin, T. et al. (2017) para resolver o desequilíbrio de classe extremo. O RetinaNet resolve o desequilíbrio de classe das amostras positivas (primeiro plano) e negativas (segundo plano), introduzindo uma nova função de perda chamada *Focal Loss*. Uma vez que a anotação mantém apenas as amostras positivas (caixas delimitadoras da verdade de

campo), as amostras negativas são obtidas durante o treinamento após a geração de caixas associadas aos candidatos que não correspondem à verdade de campo. Nesse sentido, dependendo do número de candidatas de caixas delimitadoras geradas, o número de amostras negativas pode ser super-representado. Para resolver esse problema, o RetinaNet reduz a perda atribuída a amostras bem classificadas (cuja probabilidade de ser verdadeira é acima de 0,5) e se concentra em amostras difíceis de classificar. A arquitetura RetinaNet é composta por uma única rede com ResNet + FPN como *backbone* e duas sub-redes independentes para regredir e classificar as caixas delimitadoras. Nos experimentos, combinou-se a rede Resnet50 com o FPN.

**FSAF:** Este método propõe o módulo Seleção de Características Livre de Âncora (*Feature Selection Anchor-Free, FSAF*) (Zhu et al., 2019) para ser aplicado a detectores *single-shot* com estrutura FPN, como o RetinaNet. Em tais métodos, escolher o melhor nível da pirâmide para o treinamento de cada objeto é uma tarefa complicada. Assim, o módulo FSAF anexa uma ramificação livre de âncora em cada nível da pirâmide que é treinada com a seleção de características *online*. Usou-se o RetinaNet com módulo FSAF, incluindo o ResNet50 + FPN como *backbone*.

### 5.1.2.3 Proposta sugerida

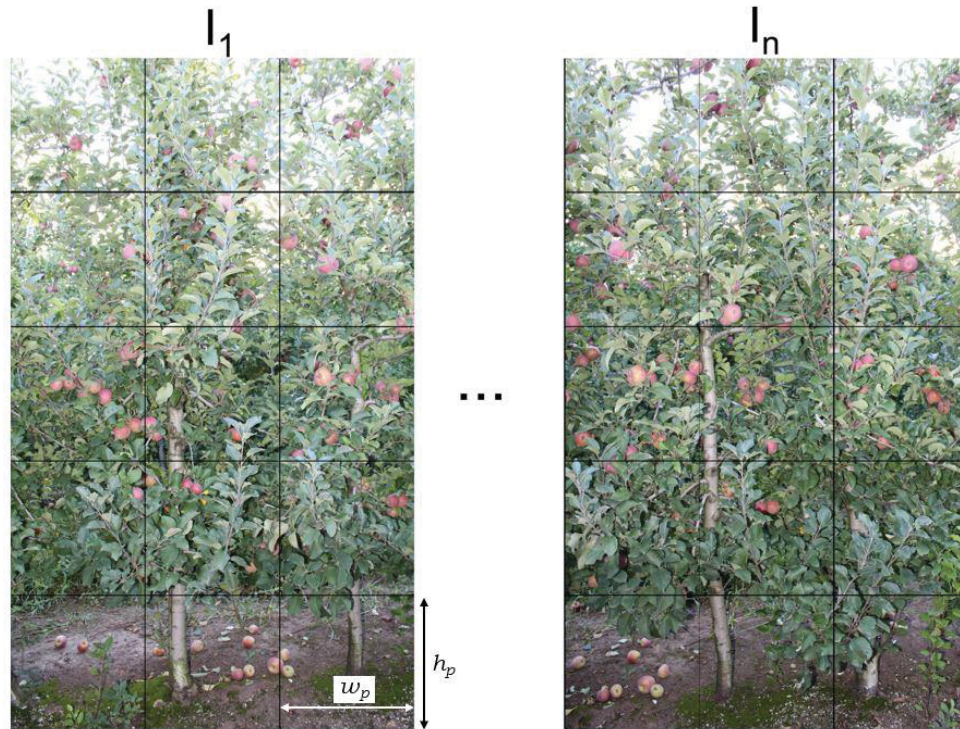
A abordagem proposta para treinar os métodos de detecção de objetos pode ser dividida em três etapas. O primeiro passo é dividir as imagens de alta resolução em retalhos de imagem (*patches*). Esta divisão visa auxiliar na detecção de maçãs pequenas e/ou oclusas. Dada uma imagem, ela é dividida em *patches* não sobrepostos com tamanho  $w_p \times h_p$  (largura x altura, em pixel), conforme mostrado na Figura 29.

Na segunda etapa, as maçãs foram rotuladas com um ponto no centro do fruto (Figura 30 a). Os frutos de maçã no chão também foram marcados. Este recurso de ponto acelera o processo de rotulagem e é essencial para rotular um grande número de imagens. No entanto, esses rótulos não podem ser usados diretamente em métodos de detecção de objetos que exigem uma caixa delimitadora.

Então, foi proposto estimar uma caixa delimitadora de tamanho fixo em torno de cada ponto rotulado (Figura 30 b). Finalmente, na última etapa, *patches* e os rótulos contendo as verdades de campo foram usados para treinar os métodos de detecção

de objetos. Em seguida, testou-se os métodos descritos na seção 5.1.2.2 Com método treinado, ele pode ser usado para detectar maçãs na etapa de teste.

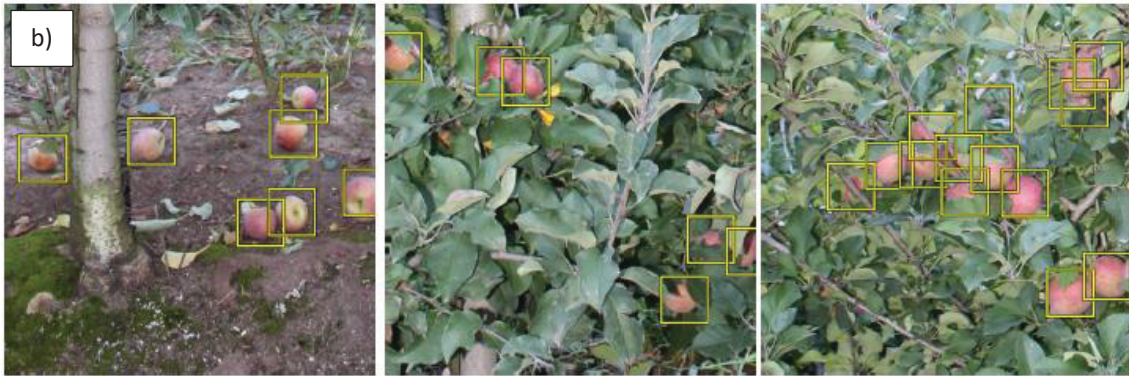
Figura 29 – Exemplo de duas imagens e sua divisão em patches sem sobreposição, com dimensões  $w_p \times h_p$ .



FONTE: (BIFFI et al. 2021b)

Figura 30 – Subconjuntos de uma determinada imagem RGB mostrando alguns exemplos de rotulagem manual com pontos no centro dos frutos da maçã (a) e estimativa de caixa delimitadora com tamanho fixo para cada rótulo de ponto (b).





FONTE: (BIFFI et al. 2021)

#### 5.1.2.4 Configurações do Experimento

As linhas de plantio foram divididas aleatoriamente em conjuntos de treinamento, validação e teste. Devido à alta resolução das imagens, elas foram divididas em *patches* de 1024 x 1024 pixels, conforme descrito na seção anterior. A Tabela 4 mostra a divisão do conjunto de dados com o número de imagens, *patches* e número de maçãs contadas e anotadas de forma visual por um operador, em cada conjunto.

Tabela 4 – Conjuntos das linhas de plantio para treinamento, validação e teste.

Conjunto	Linhas de Plantio	Nº de imagens	Nº de patches	Nº de maçãs (operador)
Treinamento	L1, L3, L6, L7	208	3.119	29.983
Validação	L4	56	840	7.010
Teste	L2, L5, L8	134	2.010	16.466
<b>Total</b>		<b>398</b>	<b>5.969</b>	<b>53.456</b>

Seguindo a metodologia de Zhang, S. et al. (2020), os *patches* de entrada foram redimensionados para 1333 x 800 pixels e normalizados usando valores médios de cor RGB de [123,675; 116,28; 103,53] e variâncias de [58,395; 57,12; 57,375]. Para aumento de dados, cada *patch* de entrada foi invertido aleatoriamente com uma probabilidade de 0,5. Com os centros dos frutos marcados e rotulados manualmente, as caixas delimitadoras com tamanho fixo de 120 x 120 pixels, foram avaliadas primeiramente (seção 5.1.3.1). Em seguida, testou-se diferentes tamanhos para as caixas, variando de 80 x 80 até 180 x 180 pixels (seção 5.1.3.2).

Para o treinamento, os suportes principais (*backbone*) dos métodos de detecção de objetos foram inicializados com os pesos pré-treinados no *ImageNet* em uma estratégia conhecida como *Transfer Learning*. Um conjunto de hiperparâmetros foi corrigido para todos os métodos após a validação. Faster-RCNN, Libra-RCNN, Cascade-RCNN e HRNet foram treinados usando o otimizador SGD para 17.500 iterações com uma taxa de aprendizagem de 0,02, um *momentum* de 0,9 e um declínio de peso de 0,0001.

RetinaNet, FSAF e ATSS seguiram os mesmos hiperparâmetros, exceto pela taxa de aprendizado inicial de 0,01 conforme mostrado na Tabela 5. Estes 3 métodos também foram treinados com uma taxa de aprendizado mais alta (0,02), mas não convergiram adequadamente durante o treinamento.

Tabela 5 – Hiperparâmetro de cada método de detecção.

Método	Taxa de aprendizagem	Otimizador	Nº de interações	Momentum
RetinaNet	0,01	SGD	17.500	0,9
Libra-RCNN	0,02	SGD	17.500	0,9
Cascade-RCNN	0,02	SGD	17.500	0,9
Faster-RCNN	0,02	SGD	17.500	0,9
FSAF	0,01	SGD	17.500	0,9
HRNet	0,02	SGD	17.500	0,9
ATSS	0,01	SGD	17.500	0,9

FONTE: (BIFFI et al. 2021b)

Como uma métrica de avaliação, usou-se a Precisão Média (PM) comumente usada na detecção de objetos com *IoU* de 0,5. Em resumo, esta métrica determina exemplos como verdadeiro-positivo (VP) quando o valor *IoU* é maior do que um limiar (0,5 adotado nos experimentos) e falso-positivo (FP) caso contrário. A *IoU* significa *Intersection Over Union*, que é calculada como a área de sobreposição entre as caixas delimitadoras preditas e verdadeiras divididas pela união entre elas. De acordo com esta quantidade de exemplos de *VP* e *FP*, o *PM* é calculado como a área sob a curva de recuperação de precisão e seu valor varia de 0 (baixo) a 1 (alto). A Precisão (P) e a *Recall* (R) são calculadas de acordo com as Equações (33) e (34), respectivamente.

$$P = \frac{VP}{VP + FP} \quad (33)$$

$$R = \frac{TP}{TP + FN} \quad (34)$$

A robustez do modelo foi medida aplicando 13 tipos de corrupção de imagem e três Níveis de Severidade (NS) em cada tipo de corrupção. Conforme indicado por Michaelis et al. (2019), essas corrupções de imagens não são recomendadas para serem usadas em dados de treinamento como uma *toolbox* de aumento de dados. Nesse sentido, as corrupções foram aplicadas apenas para as imagens do conjunto de teste. As corrupções são distribuídas em quatro categorias: (i) ruído (gaussiano, shot e impulse); (ii) borrões (*defocus*, *glass* e *motion*); (iii) condições climáticas que geralmente ocorrem na área de estudo (neve, geada, neblina e brilho); e (iv) processamento digital de imagens (transformação elástica, pixelado e compressão JPEG). Os experimentos foram realizados com apenas os 3 primeiros níveis de severidade, em vez dos 5 utilizados por Hendrycks & Dietterich (2019).

Todos os experimentos realizados na **metodologia B** deste estudo de doutorado, utilizaram-se de um computador desktop com Intel (R) Xeon (R) CPU E3-1270@3,80 GHz, 64 GB de memória e placa de vídeo NVIDIA Titan V (5120 Compute Unified Device Architecture - núcleos CUDA e memória gráfica de 12 GB). Os métodos foram implementados usando a toolbox *mmdetection* (<https://github.com/open-mmlab/mmdetection>) no sistema operacional Ubuntu 18.04. Computador esse pertencente a Universidade Federal do Mato Grosso do Sul – UFMS.

### 5.1.3 Resultados e Discussão

#### 5.1.3.1 Desempenho de detecção de maçã usando diferentes métodos de detecção de objetos.

Os resultados resumidos dos métodos de detecção de frutos de maçã são mostrados na Tabela 6. Para cada método de *Deep Learning* adotado, a *PM* de cada linha de plantio é apresentado além de sua Média e Desvio Padrão (std). Os resultados experimentais mostraram que o *ATSS* obteve o maior *PM* de 0,925 ( $\pm 0,011$ ), seguido por *HRNet* e *FSAF*. Os métodos tradicionais de detecção de objetos, como *Faster-RCNN* e *RetinaNet*, obtiveram resultados ligeiramente inferiores com *PM* de 0,918 ( $\pm 0,016$ ) e 0,903 ( $\pm 0,020$ ), respectivamente. Apesar disso, todos os métodos apresentaram resultados de *PM* acima de 0,9, o que indica que a detecção automática de frutos de maçãs, assim como a quantificação no número dessas detecções é

possível. Em comparação com a contagem manual efetuada por um operador, a contagem automática é muito mais rápida, com precisão semelhante.

Tabela 6 – Precisão Média (*PM*) dos métodos de detecção de frutos de maçã em três linhas de plantio do conjunto de teste.

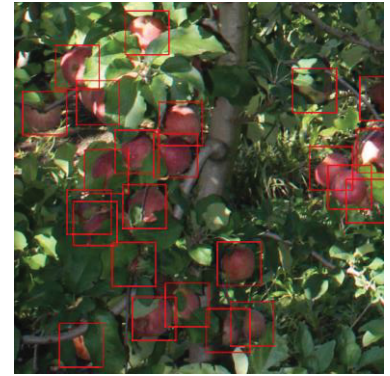
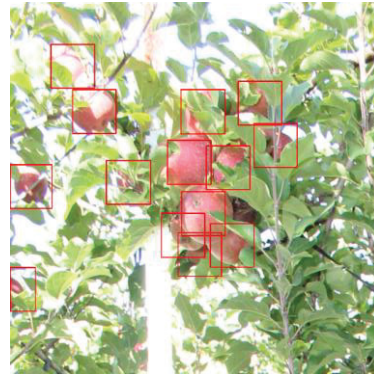
Métodos	Linhas de Plantio ( <i>Valores de PM</i> )				Nº de Maçãs (detectadas)
	L2	L5	L8	Média (std)	
RetinaNet	0,911	0,881	0,918	0,903(±0,020)	14.869
Libra-RCNN	0,910	0,896	0,922	0,909(±0,013)	14.968
Cascade-RCNN	0,920	0,892	0,923	0,912(±0,017)	15.017
Faster-RCNN	0,925	0,899	0,929	0,918(±0,016)	15.116
FSAF	0,929	0,906	0,932	0,922(±0,014)	15.182
HRNet	0,926	0,905	0,936	0,922(±0,016)	15.182
ATSS	0,925	0,914	0,935	0,925(±0,011)	15.231

FONTE: (BIFFI et al. 2021b)

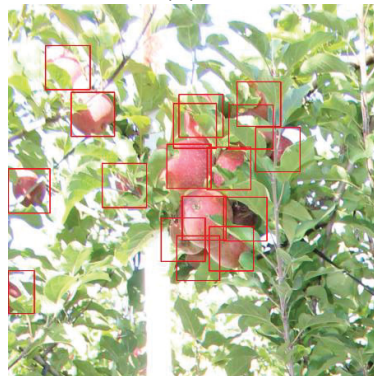
A última coluna da Tabela 6 mostra o número estimado de maçãs detectadas por cada método no conjunto de teste. Observa-se que o *ATSS* detectou aproximadamente 50 maçãs a mais do que o segundo melhor método (*HRNet* ou *FSAF*). É importante destacar que esse resultado foi obtido em três linhas de plantio, enquanto o impacto em todo o plantio seria muito maior. Além disso, o impacto financeiro em um sistema de colheita automática usando *ATSS*, por exemplo, seria considerável, apesar da pequena diferença na *PM*.

Exemplos de detecção de maçã são mostrados nas Figura 31 e Figura 32 para todos os métodos de detecção de objetos. Essas imagens apresentaram cenários desafiadores com alta oclusão de maçã e diferentes condições de iluminação. Oclusão e diferenças na iluminação são provavelmente os problemas mais comuns enfrentados por este tipo de detecção, conforme observado em aplicações semelhantes (Bargoti & Underwood, 2017; Kestur et al., 2019 e Liu et al., 2018). A precisão obtida pelo método *ATSS* no conjunto de dados, utilizado nesse estudo de doutorado, foi semelhante ou superior quando comparada com outros trabalhos de detecção de frutos na planta (Koirala, A. et al., 2019).

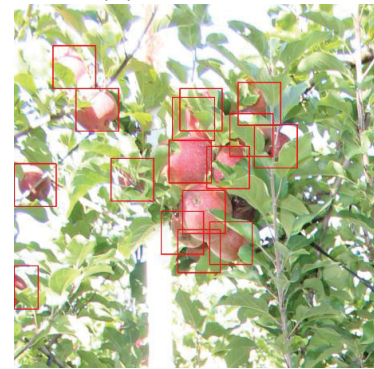
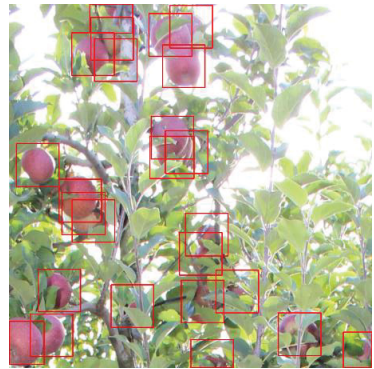
Figura 31 - Exemplos de detecção de maçã usando: (a) ATSS, (b) Cascade-RCNN, (c) Faster-RCNN, e (d) FSAF.



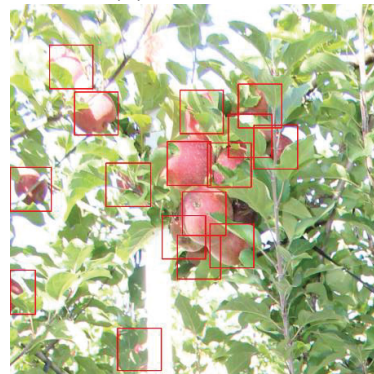
(a) ATSS



(b) Cascade-RCNN



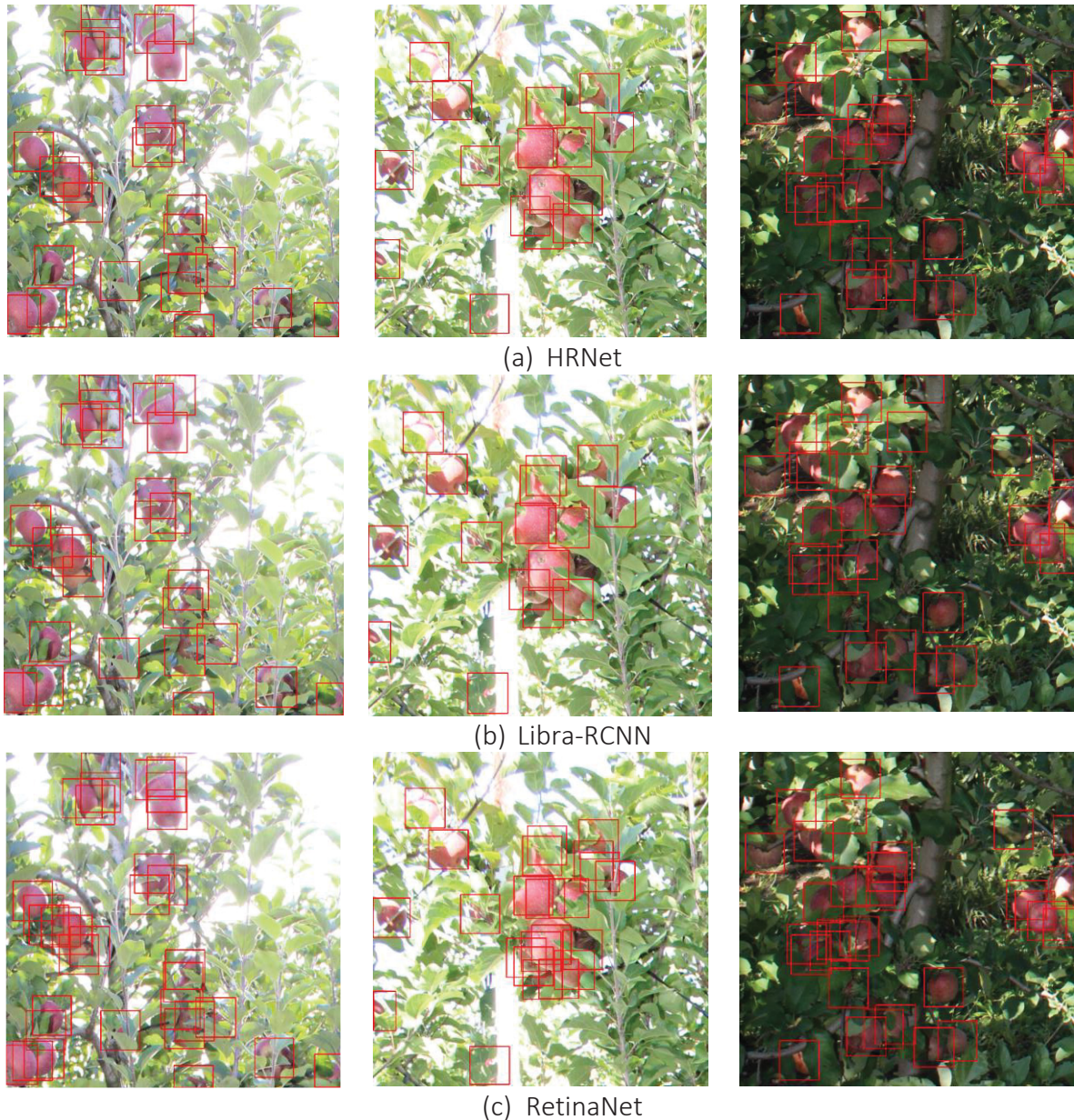
(c) Faster-RCNN



(d) FSAF

FONTE: (BIFFI et al. 2021b)

Figura 32 - Exemplos de detecção de maçã usando: (a) HRNet, (b) Libra-RCNN, e (c) RetinaNet.



FONTE: (BIFFI et al. 2021b)

Essas imagens são desafiadoras até mesmo em processos manuais, efetuados por operador, que necessitariam de um tempo elevado para realizar uma inspeção visual. Apesar das dificuldades existentes com esses exemplares de imagens, os métodos de detecção apresentaram bons resultados na identificação automática do fruto. Entretanto, o *PM* inferior do *RetinaNet* foi devido a falsos positivos, que ocorreram em regiões na imagem com reflexo pelo brilho do sol e principalmente quando se apresenta um agrupamento de frutas, onde um número maior que o real de frutos foi detectado conforme mostrado na Figura 32 (c). Pode-se

observar essa ocorrência ao comparar a esta figura com as Figura 32 (a) e (b), dos métodos HRNet e Libra-RCNN, respectivamente. Em geral quando o fruto aparece sozinho e isolado, mesmo com um pouco de oclusão, todos os métodos apresentarem a detecção correta.

É importante destacar que os métodos foram treinados com todas as caixas delimitadoras (*bounding box*) do mesmo tamanho 120 x 120 pixels. No entanto, as caixas delimitadoras estimadas ou resultantes da detecção, não eram todas do mesmo tamanho. Por exemplo, no *Faster-RCNN*, a largura e altura calculadas de todas as caixas delimitadoras estimadas no conjunto de teste apresentaram um tamanho médio em pixel de 114,60 ( $\pm 15,44$ ) x 114,96 ( $\pm 17,40$ ). O desvio padrão, portanto, indica que há uma variação entre as caixas delimitadoras estimadas.

#### 5.1.3.2 Influência do tamanho da caixa delimitadora (Bounding Box)

Os resultados da seção anterior foram obtidos com um tamanho de caixa delimitadora fixo de 120 x 120 pixels, no conjunto de treino. Nesta seção, avaliou-se a influência do tamanho da caixa delimitadora no ATSS, uma vez que obteve *PM* mais elevado entre os métodos de detecção de objetos. A Tabela 7 mostra os resultados ao usar tamanhos de caixas variando de 80 x 80 pixels até 180 x 180 pixels.

Caixas delimitadoras com menor tamanho (e.g., 80 e 100 pixels) alcançaram os menores valores de *PM* (0,842 e 0,868, respectivamente), pois não cobriram a maior parte das bordas dos frutos da maçã. O contraste e as bordas do objeto são importantes para a generalização adequada dos métodos. Por outro lado, tamanhos maiores (e.g., 180 x 180 pixels) fazem com que dois ou mais frutos de maçãs sejam incluídos em uma única caixa delimitadora. Portanto, o método apresentou dificuldade de detectar frutos de maçã individualmente, especialmente os pequenos ou em oclusão. Os experimentos mostraram que o maior valor de *PM* foi de 0,946, resultado obtido com um tamanho de caixa de 160 x 160 pixels, com uma pequena diferença no valor de *PM* para as dimensões de 140 e 180 pixels.

Apesar dos desafios impostos pela aplicação (e.g., frutos de maçãs com grande oclusão, escalas, iluminação, aglomerados), os resultados mostraram que a *PM* é elevado mesmo sem rotular uma caixa delimitadora para cada objeto. Uma vez que estudos referenciados de detecção de objetos de plantas frutíferas (Kestur et al., 2019 e Liu et al., 2018) usaram diferentes tamanhos de caixa como rótulo,

principalmente por causa das diferenças de perspectiva na posição dos frutos, este tipo de informação deve ser considerado, visto que a rotulagem manual é uma operação desgastante e trabalhosa.

Nesse experimento, demonstrou-se que no método ATSS, quando os objetos têm uma forma regular, o que geralmente é o caso das maçãs, o uso de uma caixa delimitadora de tamanho fixo é suficiente para obter *PM* aceitável. Isso, por si só, pode facilitar a rotulagem dessas frutas, uma vez que o operador pode utilizar uma abordagem do tipo ponto e, posteriormente, um tamanho fixo pode ser adotado pelos métodos baseados em caixa delimitadora, o que implica em uma redução significativa do tempo gasto em efetuar as anotações manuais nas imagens, já que a forma mais comum é a anotação por caixa delimitadora. Este processo necessita de no mínimo duas marcações na diagonal da caixa para finalizar a anotação ou rotulagem e a marcação por ponto requer uma marcação ao centro do objeto, como usado nesse estudo.

Tabela 7 – Resultado da Precisão Média (*PM*) para diferentes tamanhos de caixas delimitadoras (BB) obtidas pelo ATSS.

Tamanho (BB) (pixel)	Linhas de Plantio ( <i>Valores de PM</i> )			
	L2	L5	L8	Média (std)
80 x 80	0,843	0,827	0,855	0,842(±0,014)
100 x 100	0,892	0,810	0,901	0,868(±0,050)
120 x 120	0,925	0,914	0,935	0,925(±0,011)
140 x 140	0,940	0,926	0,950	0,939(±0,012)
160 x 160	0,946	0,936	0,956	0,946(±0,010)
180 x 180	0,933	0,929	0,947	0,936(±0,009)

FONTE: (BIFFI et al. 2021)

### 5.1.3.3 Detecção em níveis de densidade

Avaliação dos resultados do ATSS com uma caixa delimitadora de tamanho 160 x 160 pixels em *patches* com diferentes números de concentração de frutos maçãs (densidade). A Tabela 7 apresenta os resultados para *patches* com números de frutos de maçãs entre 0–9, 10–19, 20–29, 30–42. Quanto maior o número de maçãs em um patch, maior o desafio na detecção.

Pode-se ver que o método mantém a precisão em torno de 0,94–0,95 para densidades entre 0–29 frutos de maçãs. A Figura 33 mostra exemplos de detecção em *patches* sem frutos, por exemplo. Isso mostra que o método é consistente na

detecção de algumas maçãs até uma quantidade considerável de 29 maçãs em um único *patch*. Como esperado, com o aumento do desafio de detecção, com maior número de frutos em um *patche* e a possibilidade de problemas relacionados a agrupamento e oclusões de frutos, a precisão do método diminui no último nível de densidade (*patches* com 30-42 frutos). Porém, a precisão ainda é elevada mesmo com esse nível de densidade de frutos.

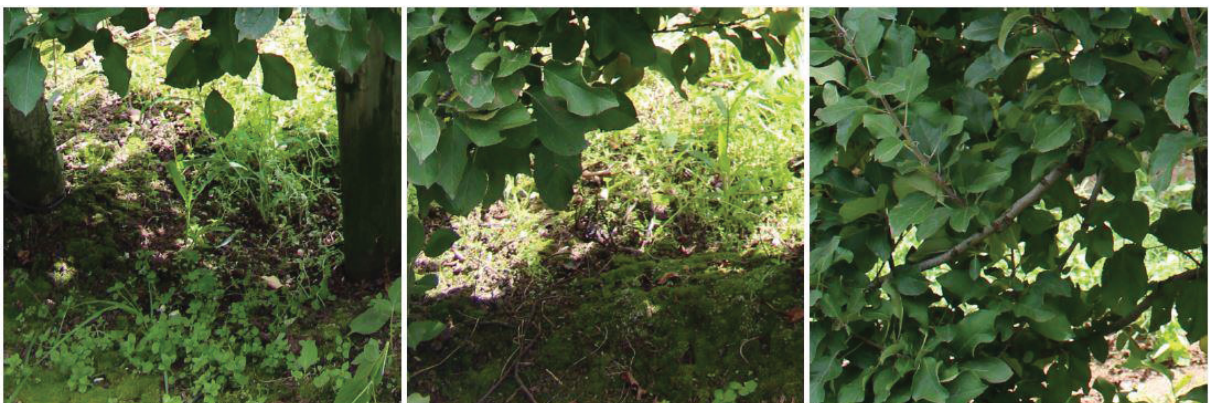
Os *patches* com maior densidade de frutos foram poucos, apenas 12, nota-se que a concentração de número de frutos média fica entre 10-19 frutos por *patches*. A quantidade de *patches* mais elevada foi na concentração 0-9 frutos. Isso se deve a posição e ao campo de visão da câmera, que captura parte do solo e também a parte mais alta da planta, onde na imagem aparece o céu. A Figura 33 apresenta exemplo de *patches* próximo ao solo que apresentaram frutos visíveis.

Tabela 8 – Resultado do método ATSS em diferentes densidades de frutos de maçã por *patches*.

Densidade de Frutos de Maçã	Nº de <i>Patches</i>	AP
0 – 9	1246	0,940
10 – 19	634	0,950
20 – 29	118	0,952
30 – 42	12	0,924

FONTE: (BIFFI et al. 2021)

Figura 33 – Exemplos de *patches* sem maçãs.



FONTE: (BIFFI et al. 2021)

#### 5.1.3.4 Robustez contra Corrupções das imagens

Nesta seção, avaliou-se a robustez do modelo do ATSS com uma caixa delimitadora de tamanho 160 x 160 pixels em simulações de diferentes corrupções, apenas para as imagens no conjunto de teste, pois as corrupções não foram utilizadas para treinamento e validação, para estes conjuntos foram utilizadas as imagens originais. Essas corrupções simulam possíveis condições de ocorrerem na aquisição de imagens desta cultura *in situ*, entre eles fatores ambientais adversos, atitude do sensor e degradação no registro digital dos dados.

As Tabela 9 e Tabela 10 mostram os resultados para diferentes corrupções e níveis de severidade (grau de degradação da imagem). Os resultados de ruído entre os três níveis de severidade (Tabela 9) indicaram uma redução na precisão média (*PM*) entre 2,7% e 11,1% em comparação com os valores normais sem corrupção. Ao considerar o ruído *gaussian* e o ruído *shot*, as precisões foram reduzidas na mesma proporção, enquanto a redução causada pelo ruído de *impulse* foi maior no primeiro e no segundo níveis. Considerando todos os níveis de severidade da corrupção de ruído, a precisão foi superior a 84%.

Tabela 9 – Precisão Média para diferentes ruídos e borrões em três níveis de severidade obtidos por método ATSS.

Nível de Severidade	Ruído			Borrão		
	Gaussian	Shot	Impulse	Defocus	Glass	Motion
SC*	0,947	0,947	0,947	0,947	0,947	0,947
1	0,919	0,921	0,908	0,941	0,944	0,945
2	0,892	0,896	0,874	0,931	0,932	0,938
3	0,844	0,848	0,842	0,902	0,910	0,921

SC\* = Sem Corrupção na imagem.

A redução na precisão para todos os componentes de borrão (Tabela 9) foi muito semelhante, mostrando uma ligeira diminuição em comparação com a condição de sem corrupção. Os valores de redução ficaram entre 0% e 4,8%. Em todos os três níveis de severidade, os resultados obtidos com borrão de *motion* impactaram menos quando comparados aos resultados obtidos com *defocus* e *glass*. A menor precisão foi devido ao *defocus*. Todos os resultados mostraram valores de precisão acima de 90%, mesmo considerando todos os níveis de severidade.

Tabela 10 – Precisão Média para diferentes condições de clima e processamento digital de imagem em três níveis de severidade obtidos por método ATSS.

Nível de Severidade	Clima				Digital		
	Neve	Geada	Névoa	Brilho	Transf. Elástica	Pixelado	Compressão JPEG
SC*	0,947	0,947	0,947	0,947	0,947	0,947	0,947
1	0,839	0,878	0,800	0,942	0,945	0,946	0,933
2	0,651	0,765	0,708	0,934	0,944	0,946	0,924
3	0,732	0,678	0,611	0,924	0,941	0,943	0,917

SC\* = Sem Corrupção na imagem.

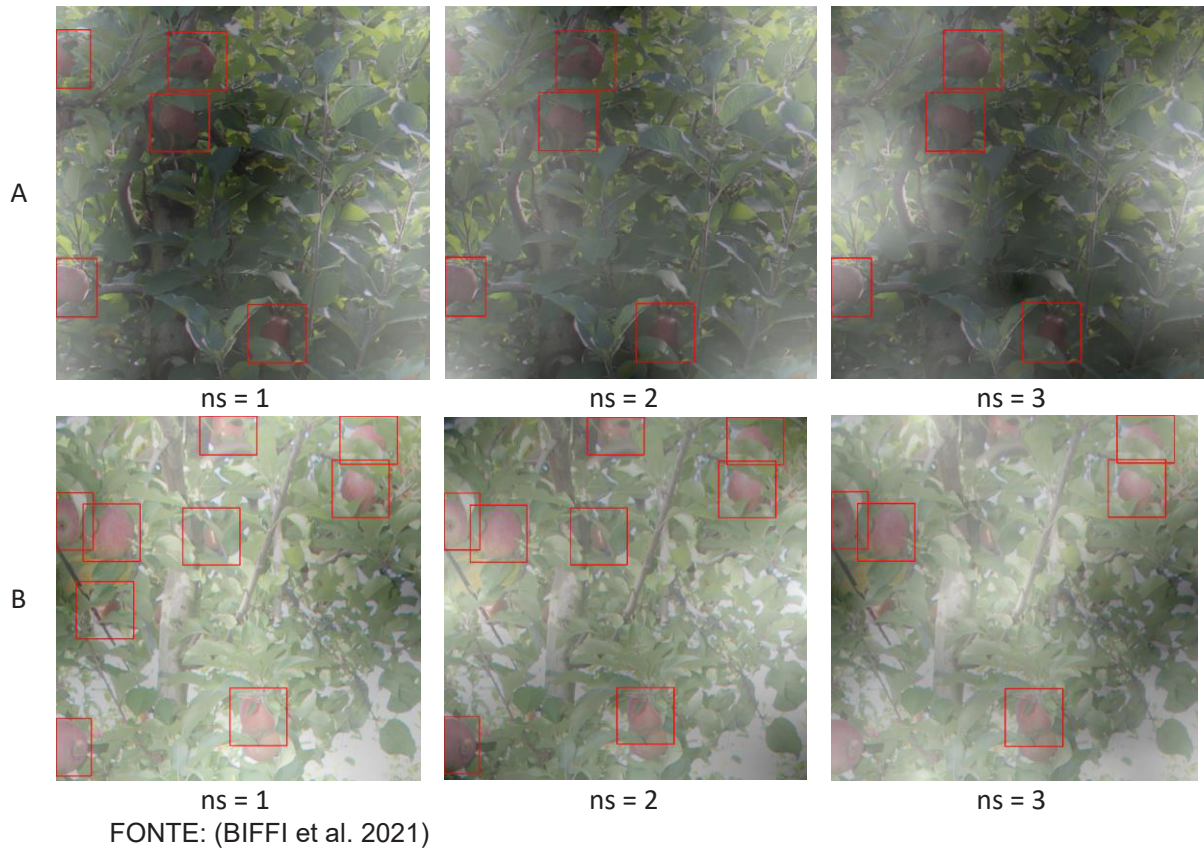
De acordo com os resultados apresentados na Tabela 10, a *névoa* resultou em menor precisão nas diferentes condições de clima simuladas. As reduções de precisão foram de aproximadamente 15,5% a 35,5% do primeiro ao terceiro nível de severidade, respectivamente. A Figura 34 mostra a detecção da maçã em duas imagens corrompidas pela *névoa*. Ao aumentar o nível de severidade da *névoa*, alguns frutos de maçã não são detectados. Essa perda de precisão é esperada principalmente quando há maçãs com grande oclusão foliar e/ou pouca iluminação.

O *brilho* implicou em uma pequena redução de precisão, entre 0,5% e 2,4%, em comparação com a imagem sem corrupção. As precisões obtidas no processamento Digital, para *transformação elástica* e *pixelado*, foram próximas a 95% para todos os níveis de severidade, e os resultados obtidos da *compressão JPEG* foram próximos a 92%. O *pixelado* apresentou os melhores resultados de precisão para todos os níveis de severidade entre as outras condições digitais. As reduções de precisão da *compressão JPEG* foram de 1,4% a 3,2% para os níveis de severidade 1 a 3, respectivamente, em comparação com a condição sem corrupção. Em geral, as condições climáticas afetaram mais a precisão do que o processamento digital.

Trabalhos referenciados Gao et al. (2020), Gené-Mola et al. (2019) e Tian et al. (2019) investigaram e adaptaram os métodos *Faster R-CNN* e *Yolo-v3*. Aqui nesse experimento, investigou-se um novo método baseado em ATSS e apresenta-se seu potencial para a detecção de frutos de maçã. Em um sentido geral, os resultados indicaram que o método baseado em ATSS superou ligeiramente os outros métodos de aprendizado profundo, com o melhor resultado obtido com um tamanho de caixa delimitadora de 160 x 160 pixels. O método proposto foi robusto em relação à maior parte da corrupção, exceto para condições climáticas de neve, geada e névoa. No entanto, os resultados ainda são promissores, o que implica bons registros de

precisão para detecção de frutos de maçã, mesmo quando as imagens RGB são adquiridas em períodos em que as condições meteorológicas não são favoráveis, como geada, nevoeiro e até eventos de neve que são frequentes durante o inverno em algumas áreas do planalto catarinense no sul do Brasil.

Figura 34 – Exemplo de dois patches diferentes (A e B) com névoa (ns = nível de severidade).



#### 5.1.4 Perspectivas de pesquisas adicionais

Esta pesquisa disponibilizou todas as imagens RGB terrestres de curto alcance e baixo custo, incluindo as anotações. Essas imagens foram adquiridas manualmente em campo. Porém, as câmeras também podem ser instaladas em implementos agrícolas, como pulverizadores, ou mesmo em tratores e veículos autônomos (Underwood et al., 2016). Em pesquisas futuras serão verificadas a aplicabilidade de tais imagens nas fases anteriores ou mesmo durante a colheita dos frutos da maçã. Pesquisas serão realizadas com a desenvolvimento de metodologias autônomas para a contagem de frutos que caem no solo e o impacto na estimativa da produção final de frutos de maçã. Outra tarefa importante é segmentar cada fruta, o

que pode ser executado usando uma estratégia de crescimento de região ou com base em métodos de segmentação de instância DL.

Além da identificação da fruta, outros atributos da fruta como forma, peso e cor também são informações muito relevantes direcionadas pelo mercado e são indicadas para futuros estudos. A adição de informações quantitativas, como forma, pode dar suporte à estimativa do peso médio do fruto. Essas informações forneceria estimativas de produção de frutas melhores e mais realistas. Porém, as imagens devem conter toda a planta da macieira, o que às vezes não acontecia devido à altura da árvore. Devido ao espaço limitado entre as linhas de plantio (ou seja, 3,5 m), às vezes foi impossível cobrir toda a macieira no quadro da imagem. Portanto, sugere-se pesquisas com dispositivos de câmera instalados em diferentes alturas de plataforma em implementos agrícolas ou tratores para superar esse problema. Alternativamente, novos estudos também poderiam explorar a possibilidade de aquisição de imagens oblíquas ou o uso de lentes olho-de-peixe.

Curiosamente, o uso de técnicas fotogramétricas com DL poderá localizar a posição referenciada dos frutos visíveis (Liu et al., 2018), bem como modelar sua superfície. Isso possibilitará a avaliação individual do tamanho do fruto, permitindo uma contagem mais realista dos frutos. Também elimina qualquer possibilidade de contar duas vezes uma única fruta, eventualmente em imagens sequenciais com sobreposição. Isso também pode ocorrer nas imagens na parte de trás da linha de plantio, que também é frequentemente computada mais de uma vez (Häni et al., 2020).

Os dados desse experimento encontram-se disponíveis no link do artigo na revista *Remote Sensing – Appendix A*. ( <https://doi.org/10.3390/rs13010054>).

### 5.1.5 Conclusão e Recomendações

Neste estudo, propôs-se uma abordagem baseada no método *ATSS DL* para detectar frutos de maçã usando uma estratégia de rotulagem de pontos centrais. O método foi comparado a outras redes baseadas em aprendizagem profunda de última geração, incluindo *RetinaNet*, *Libra R-CNN*, *Cascade R-CNN*, *Faster R-CNN*, *FSAF* e *HRNet*.

Também avaliou-se diferentes condições de oclusão e corrupção de ruído em diferentes conjuntos de imagens. A abordagem baseada em *ATSS* superou todos os outros métodos, alcançando uma precisão média máxima de 0,946 com um tamanho de caixa delimitadora de 160 x 160 pixels. Ao simular as condições de adversidade no momento da aquisição de dados, a abordagem forneceu uma resposta robusta para a maioria dos dados corrompidos, exceto para neve, geada e névoa.

Novos estudos são sugeridos com outras variedades de frutas em que a cor desempenha um papel importante na diferenciação das folhas. Atributos adicionais da fruta, como forma, peso e cor, também são informações importantes para determinar o preço de mercado e também são recomendados para investigações futuras.

## 6 METODOLOGIA C

### 6.1 DADOS CONVENCIONAIS VS FOTOGRAMETRIA

#### 6.1.1 Dados Convencionais

Nesta seção serão apresentados os dados levantados de forma convencional para avaliação de produção do pomar. Uma descrição do método de levantamento de campo foi apresentada na seção 3.3.

Os resultados destes dados estão apresentados na de Tabela 11 e trazem os valores de métricas geométricas que são usualmente aplicadas em pomar. Importante observar que se apresenta dados de produção, pois o NF produzido em cada planta, assim como, amostras para determinação de PMF foram coletadas. O Número de frutos por planta, que em média ficou em 111 frutos, pode parecer difere das análises das seções 4 e 5, mas esse é o valor real do número de frutos, e nas metodologias A e B o número de frutos foram quantos apareceram nas imagens, e estas tinham sobreposição, então o mesmo fruto foi detectado em mais de uma imagem.

Tabela 11 – Valores médios por planta de dados coletados em campo de forma convencional, nas linhas amostradas do pomar. Diâmetro (m), Altura (m), Comprimento (m), Largura (m), Número de Frutos.

Linhas	<i>Di</i> (m)	<i>H</i> (m)	<i>Co</i> (m)	<i>Lg</i> (m)	<i>NF</i>
L1	0,074	4,03	1,12	1,34	172,8
L2	0,074	3,97	0,95	1,24	103,0
L3	0,068	4,11	0,92	1,30	110,0
L4	0,078	4,26	0,95	1,30	127,9
L5	0,076	4,17	0,90	1,22	92,8
L6	0,088	4,24	0,97	1,41	83,7
L7	0,070	3,71	0,86	1,14	71,6
L8	0,077	4,08	0,91	1,27	123,6
<b>Médias</b>	<b>0,076</b>	<b>4,07</b>	<b>0,95</b>	<b>1,28</b>	<b>110,7</b>
DP	0,004	0,125	0,051	0,061	23,07

Di: Diâmetro; H: Altura; Co: Comprimento; Lg: Largura; NF: Número de frutos; DP: Desvio Padrão.

Após as coletadas em campo, outras variáveis de produção foram calculadas com estes dados. As variáveis calculadas foram Altura corrigida (*Hc*), Volume de Copa

(*Vol*), Peso médio dos Frutos (*PMF*), Produção por planta (*Prod*), Produtividade (*Produtiv*), Área da seção transversal do caule (*ASTC*) e Eficiência Produtiva (*EP*). Os valores obtidos para estas variáveis estão presentes na Tabela 12.

Tabela 12 - Valores médios por planta de variáveis determinadas com os dados coletados em campo de forma convencional, nas linhas amostradas do pomar. Altura corrigida (m), Volume (m<sup>3</sup>), Peso médio por fruto (g), Produção por planta (kg), Produtividade (Mg.ha<sup>-1</sup>), Área da Seção transversal do caule (cm<sup>2</sup>) e Eficiência Produtiva (kg.cm<sup>-2</sup>).

Linhas	Hc (m)	Vol (m <sup>3</sup> )	PMF (g)	Prod/pta (Kg)	Produtiv (M.ha <sup>-1</sup> )	ASTC (cm <sup>2</sup> )	EP (kg .cm <sup>-2</sup> )
L1	3,43	5,128	148	25,52	91,15	43,80	0,57
L2	3,37	3,980	164	16,88	60,29	52,66	0,43
L3	3,51	4,189	162	17,86	63,80	36,23	0,50
L4	3,66	4,562	172	21,93	78,34	49,20	0,45
L5	3,57	3,912	159	14,72	52,56	46,28	0,31
L6	3,64	4,967	156	13,08	46,72	65,08	0,23
L7	3,11	3,035	156	11,16	39,87	38,06	0,30
L8	3,48	4,009	150	18,51	66,09	49,60	0,43
<b>Médias</b>	<b>3,47</b>	<b>4,223</b>	<b>158,375</b>	<b>17,46</b>	<b>62,35</b>	<b>47,62</b>	<b>0,40</b>
DP	0,125	0,49	5,875	3,50	12,49	6,52	0,09

Hc: Altura corrigida; Vol: Volume; PMF: Peso médio por fruto; Prod/pta: Produção por planta; Produtiv: Produtividade; ASTC: Área da Seção transversal do caule; EP: Eficiência Produtiva, DP: Desvio Padrão

### 6.1.2 Fotogrametria

Com as fotos adquiridas para Fotogrametria, conforme descrito na seção 3.2 uma nuvem de pontos com o modelo das plantas foi gerada por técnica de SfM. Um exemplo dessa nuvem é apresentado na Figura 35.

Com a geração das nuvens de pontos, foi possível extrair informações métricas das plantas amostras. As mensurações foram efetuadas com os recursos de mensuração de distâncias do próprio software. O modo de operação para determinar as variáveis foi similar aos feitos em campo, porém em ambiente digital.

Figura 35 - Apresentação de Nuvem de pontos referente as plantas da Linha L4, gerada por Fotogrametria.



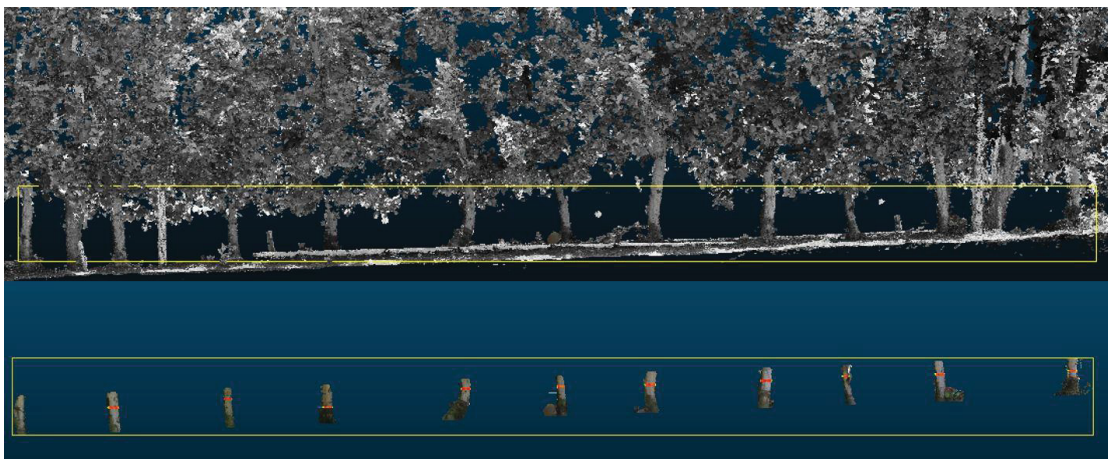
Fonte: O autor (2020).

#### 6.1.2.1 Variáveis determinadas

##### 6.1.2.1.1 Diâmetro do Caule

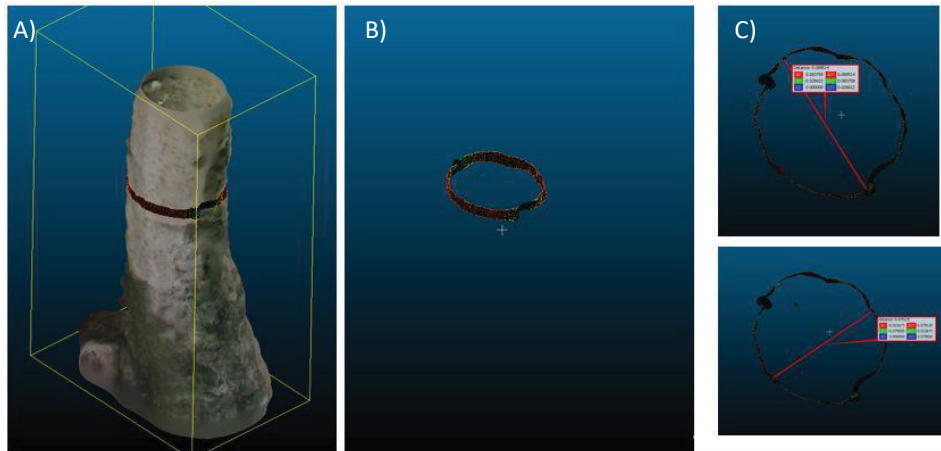
Para determinação dos diâmetros dos caules, primeiro realizou-se um corte na nuvem de pontos, em seguida cada caule foi separado em uma nuvem de pontos distinta (Figura 36). Em cada caule individualizado foi selecionado uma faixa transversal de corte a aproximadamente 20 cm do ponto de enxertia visível (Figura 37 A e B). Nesta faixa da seção transversal foram realizadas medidas de diâmetro em duas direções (Figura 37 C).

Figura 36 – Seção do caule para efetivar a leitura do diâmetro (Di).



Fonte: O autor (2020).

Figura 37 – Caule individualizado em nuvem de pontos (A); Corte de faixa de transversal a 20 cm do ponto de enxertia (B); e medidas de diâmetro realizadas em duas direções (C).

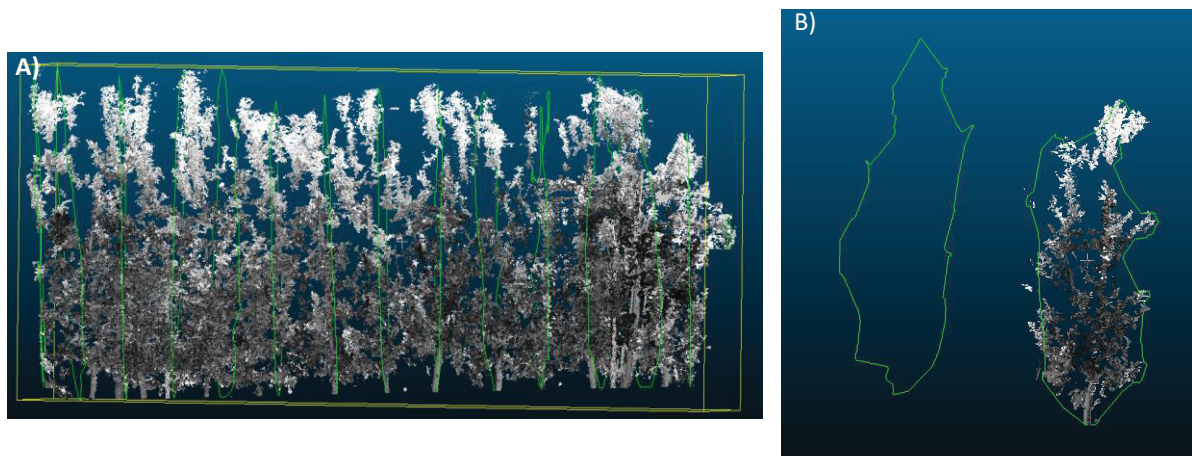


Fonte: O autor (2020).

#### 6.1.2.1.2 Volume

Para determinar o volume das plantas na nuvem de pontos, fez-se cortes verticais na nuvem com um espaçamento igual a distância entre plantas (0,8 m), e calculou-se a área da superfície formada pela borda simplificada da nuvem (Figura 38). Após os valores de área determinados, multiplicou-se a área média de suas seções, pela distância entre plantas.

Figura 38 – Nuvem segmentada em perfil vertical com espaçamento de 0,8 m (A); e delimitação de uma borda simplificada da nuvem (B).



Fonte: O autor (2020).

Com a obtenção desses valores as demais variáveis como ASTC e EP foram calculadas por planilha eletrônica.

### 6.1.3 Comparação dos resultados

Uma comparação entre os dados obtidos de modo convencional e por Fotogrametria é apresentada na Tabela 13. Nesta tabela os valores se referem a dados de plantas individuais da Linha L4. A intenção foi verificar se há diferenças entre os métodos de coleta de dados por planta. Um teste t de médias pareadas foi aplicado nos dados.

Tabela 13 – Dados de plantas da linha L4, obtidas pelo método convencional e por Fotogrametria. Os dados apresentados são: Número de Frutos, Diâmetro do caule (m), Volume (m<sup>3</sup>), Área da Seção transversal do caule (cm<sup>2</sup>) e Eficiência Produtiva (kg.cm<sup>-2</sup>)

Plantas	NF	Convencional				Fotogrametria			
		<i>Di</i>	<i>Vol</i>	<i>ASTC</i>	<i>EP</i>	<i>Di</i>	<i>Vol</i>	<i>ASTC</i>	<i>EP</i>
A1	206	0,088	4,900	60,62	0,58	0,080	3,560	49,64	0,71
A2	75	0,081	4,050	52,15	0,25	0,081	3,500	51,53	0,25
A3	81	0,064	3,312	31,83	0,44	0,059	4,036	26,88	0,52
A4	148	0,086	4,928	58,01	0,44	0,083	4,270	53,46	0,47
A5	71	0,087	4,810	59,74	0,20	0,085	3,848	56,41	0,22
A6	54	0,070	4,496	38,52	0,24	0,061	3,897	29,22	0,32
A7	155	0,074	4,862	42,83	0,62	0,068	3,817	36,32	0,73
A8	220	0,087	4,688	59,74	0,63	0,097	3,878	73,90	0,51
A9	128	0,065	4,680	33,12	0,66	0,062	4,331	30,19	0,73
A10	141	0,084	5,005	55,46	0,44	0,075	4,536	44,18	0,55
<b>Média</b>	<b>127,9</b>	<b>0,079<sup>ns</sup></b>	<b>4,57*</b>	<b>49,20<sup>ns</sup></b>	<b>0,45<sup>ns</sup></b>	<b>0,075</b>	<b>3,97</b>	<b>45,17</b>	<b>0,50</b>
DP	46,12	0,008	0,372	10,10	0,14	0,010	0,261	11,81	0,15

NF: Número de Frutos; Di: Diâmetro; Vol: Volume; ASTC: Área da Seção transversal do caule; EP: Eficiência Produtiva; DP: Desvio Padrão.

\* Valores significativos a 0,05 para teste t pareado.

Ao verificar os dados de *Di*, *ASTC* e *EP*, observa-se que não houve uma diferença significativa das médias entre modo convencional e por Fotogrametria. Ocorreu significância apenas na determinação do Volume das plantas. Isso se deve ao modo como o volume é calculado nos dois métodos. No convencional o cálculo é baseado em volume de caixa e no método por Fotogrametria é por área de superfície de corte vertical do modelo tridimensional da planta. O valor obtido no final é menor. Pode-se supor que o modo convencional esteja superestimando o volume da planta.

Como não houve diferença significativa nas outras variáveis, o modo como é levantado os dados não faz diferença para o resultado que se busca. Sendo assim, é viável o uso de produtos da Fotogrametria na coleta de dados demonstrados nesse estudo em pomar de alta densidade, com característica de pomares ditos bidimensionais.

O tempo de processamento das fotos para gerar a nuvem, ainda é um fator limitante para uso em tempo real. Porém, é possível através de seus produtos fazer medidas quantitativas com resultados semelhantes ao convencional.

## 7 CONSIDERAÇÕES FINAIS

Considerando a característica de pequeno produtor de maçã do Estado de Santa Catarina, foi possível propor uma metodologia de detecção de frutos de baixo custo e de simples processamento através do método de espaço de cores.

O uso da banda *Lab-a* apresentou  $R^2$  de 0,73 na detecção e contagem de frutos e  $R^2$  de 0,96 no reconhecimento dos frutos.

Assim como foi possível aplicar técnicas do estado da arte em detecção de objetos, na quantificação de frutos presentes em imagem de plantas de maçã, onde os métodos testados apresentaram alto grau de acerto e pequenas diferenças entre eles. Os métodos testados com melhores resultados foram ATSS, HRNet e FSAF, com valores de Precisões Médias de 0,925, 0,922 e 0,922, respectivamente. Foi possível estabelecer, para o método ATSS, uma configuração para anotação de objetos e dimensionamento de caixas delimitadoras para trabalhos semelhantes as condições dessa pesquisa. Os objetos podem ser anotados por um ponto no centro do objeto, mesmo que este esteja com oclusão parcial. A dimensão de caixa delimitadora com melhores resultados foi de 160 x 160 pixels. O método também apresentou bom desempenho na detecção de frutos em imagens corrompidas, exceto para condições simuladas de neve, geada e névoa.

Futuros estudos deverão considerar a alteração do ângulo do sensor para que o campo de visão não detecte a camada de solo reduzindo, portanto, os valores de erros apresentados nas duas metodologias.

O uso de imagens adquiridas para Fotogrametria de curto alcance pode ser utilizado para o propósito de detectar frutos em pomares comerciais, assim como fazer quantificações de dimensões das plantas, a proposta desta pesquisa pode servir de base para uso em outras culturas da fruticultura catarinense.

Para concluir, uma das contribuições relevantes do estudo está a disponibilidade do conjunto de dados de imagens para ser acessados pelo público. Isso pode ajudar outros a avaliar os métodos e suas respectivas abordagens para contar objetos, especificamente em frutas com condições altamente densas. Os dados podem ser acessados pelo endereço do artigo da metodologia B. Estão presentes os conjuntos das imagens de treinos, validação e teste, com as respectivas anotações dos objetos.

## 8 REFERÊNCIAS

- ABPM, A. B. D. P. D. M. Anuário Brasileiro de Maçã. , 2018. Santa Cruz do Sul.
- AGÊNCIA NACIONAL DE ÁGUAS (ANA). HIDROWEB V3.1.1 - Séries históricas de Estações. Disponível em: <<https://www.snirh.gov.br/hidroweb/serieshistoricas>>. Acesso em: 2/11/2020.
- AGGELOPOULOU, A. D.; BOCHTIS, D.; FOUNTAS, S.; et al. Yield prediction in apple orchards based on image processing. **Precision Agriculture**, v. 12, n. 3, p. 448–456, 2011.
- ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; DE MORAES GONÇALVES, J. L.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2013.
- ALVES, D. R. **Avaliação dos modelos de cores RGB e HSV na segmentação de curvas de nível em cartas topográficas coloridas**, 2010. Pontifícia Universidade Católica de Minas Gerais.
- AN, N.; WELCH, S. M.; MARKELZ, R. J. C.; et al. Quantifying time-series of leaf morphology using 2D and 3D photogrammetry methods for high-throughput plant phenotyping. **Computers and Electronics in Agriculture**, v. 135, p. 222–232, 2017.
- APOLO-APOLO, O. E.; MARTÍNEZ-GUANter, J.; EGEA, G.; RAJA, P.; PÉREZ-RUIZ, M. Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV. **European Journal of Agronomy**, v. 115, n. August 2019, p. 126030, 2020.
- APOLO-APOLO, ORLY ENRIQUE; PÉREZ-RUIZ, M.; MARTÍNEZ-GUANter, J.; VALENTE, J. A Cloud-Based Environment for Generating Yield Estimation Maps From Apple Orchards Using UAV Imagery and a Deep Learning Technique. **Frontiers in Plant Science**, v. 11, n. July, p. 1–15, 2020.
- BALL, J. E.; ANDERSON, D. T.; CHAN, C. S. Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. **Journal of Applied Remote Sensing**, v. 11, n. 04, p. 1, 2017. SPIE-Intl Soc Optical Eng.
- BARGOTI, S.; UNDERWOOD, J. P. Image Segmentation for Fruit Detection and Yield Estimation in Apple Orchards. **Journal of Field Robotics**, v. 34, n. 6, p. 1039–1060, 2017.
- BERVEGLIERI, A.; TOMMASSELLI, A. M. G.; LIANG, X.; HONKAVAARA, E. Vertical optical scanning with panoramic vision for tree trunk reconstruction. **Sensors (Switzerland)**, v. 17, n. 12, p. 1–13, 2017.
- BIFFI, L. J.; MITISHITA, E. A.; LIESENBERG, V.; et al. Evaluating the Performance of a Semi-Automatic Apple Fruit Detection in a High-Density Orchard System Using Low-Cost Digital Rgb Imaging Sensor. **Boletim de Ciências Geodésicas**, v. 27, n. 2, 2021.

BIFFI, L. J.; MITISHITA, E.; LIESENBERG, V.; et al. ATSS deep learning-based approach to detect apple fruits. **Remote Sensing**, v. 13, n. 1, p. 1–23, 2021.

BOSCO, L. C.; BERGAMASCHI, H.; CARDOSO, L. S.; et al. Microclimate alterations caused by agricultural hail net coverage and effects on apple tree yield in subtropical climate of southern Brazil. **Bragantia**, v. 77, n. 1, p. 181–192, 2018. Instituto Agronomico.

BOSCO, L. C.; BERGAMASCHI, H.; MARODIN, G. A. B. Solar radiation effects on growth, anatomy, and physiology of apple trees in a temperate climate of Brazil. **International Journal of Biometeorology**, p. 1969–1980, 2020. International Journal of Biometeorology.

BRAGA, H. J. **Previsão Agrícola: Uma Nova Abordagem - Uso de scanner aerotransportável e redes neurais**, Tese (Doutorado em Engenharia de Produção) - Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 1995.

BRGLEZ SEVER, M.; TOJNKO, S.; BREZNIKAR, A.; et al. The influence of differently coloured anti-hail nets and geomorphologic characteristics on microclimatic and light conditions in apple orchards. **Journal of Central European Agriculture**, v. 21, n. 2, p. 386–397, 2020.

CABO, C.; DEL POZO, S.; RODRÍGUEZ-GONZÁLVEZ, P.; ORDÓÑEZ, C.; GONZÁLEZ-AGUILERA, D. Comparing Terrestrial Laser Scanning (TLS) and Wearable Laser Scanning (WLS) for Individual Tree Modeling at Plot Level. **Remote Sensing**, v. 10, n. 4, p. 540, 2018.

CAI, Z.; VASCONCELOS, N. Cascade R-CNN: Delving into High Quality Object Detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. **Anais...** . p.6154–6162, 2018. IEEE Computer Society.

CARRIVICK, J.; SMITH, M.; QUINCEY, D. **Structure from Motion in the Geosciences**. 2016.

CHANG, A.; JUNG, J.; MAEDA, M. M.; LANDIVAR, J. Crop height monitoring with digital imagery from Unmanned Aerial System (UAS). **Computers and Electronics in Agriculture**, v. 141, p. 232–237, 2017.

DENARDI, F.; KVITSCHAL, M. V.; HAWERROTH, M. C. A brief history of the forty-five years of the Epagri apple breeding program in Brazil. **Crop Breeding and Applied Biotechnology**, v. 19, n. 3, p. 347–355, 2019.

DENG, L.; MAO, Z.; LI, X.; et al. UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 146, p. 124–136, 2018.

DIAN BAH, M.; HAFIANE, A.; CANALS, R. Deep learning with unsupervised data labeling for weed detection in line crops in UAV images. **Remote Sensing**, v. 10, n. 11, p. 1690, 2018.

DORJ, U. O.; LEE, M.; YUN, S. SEOK. An yield estimation in citrus orchards via fruit

detection and counting using image processing. **Computers and Electronics in Agriculture**, v. 140, p. 103–112, 2017. Elsevier B.V.

DUTTA, A.; ZISSERMAN, A. The VIA annotation software for images, audio and video. MM 2019 - Proceedings of the 27th ACM International Conference on Multimedia. **Anais...** . p.2276–2279, 2019. Association for Computing Machinery, Inc.

EMBRAPA, E. B. D. P. A. Automação e agricultura de precisão. Disponível em: <<https://www.embrapa.br/tema-mecanizacao-e-agricultura-de-precisao/nota-tecnica>>. Acesso em: 15/8/2018.

EPAGRI, (EMPRESA DE PESQUISA AGROPECUÁRIA E EXTENSÃO RURAL DE SANTA CATARINA). **A Cultura da Macieira**. Florianópolis: Epagri, 2002.

FAO, (FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS). FAOSTAT. Disponível em: <<http://www.fao.org/faostat/en/#data/QC>>. Acesso em: 22/1/2021.

FERNÁNDEZ-SARRÍA, A.; MARTÍNEZ, L.; VELÁZQUEZ-MARTÍ, B.; et al. Different methodologies for calculating crown volumes of *Platanus hispanica* trees using terrestrial laser scanner and a comparison with classical dendrometric measurements. **Computers and Electronics in Agriculture**, v. 90, p. 176–185, 2013.

FIORAVANÇO, J. C.; GIRARDI, C. L.; CZERMAINSKI, A. B. C.; et al. **Cultura da macieira no Rio Grande do Sul: análise situacional e descrição varietal**. 2010.

FLORENZANO, T. G. **Iniciação em sensoriamento remoto**. 3rd ed. São Paulo: Oficina de Textos, 2011.

FONT, D.; PALLEJÀ, T.; TRESANCHEZ, M.; et al. Counting red grapes in vineyards by detecting specular spherical reflection peaks in RGB images obtained at night with artificial illumination. **Computers and Electronics in Agriculture**, v. 108, p. 105–111, 2014.

FONT, DAVINIA; PALLEJÀ, T.; TRESANCHEZ, M.; et al. A proposal for automatic fruit harvesting by combining a low cost stereovision camera and a robotic arm. **Sensors (Switzerland)**, v. 14, n. 7, p. 11557–11579, 2014.

FRANCESCOTTO, P. **DESENVOLVIMENTO DAS ESTRUTURAS REPRODUTIVAS DA MACIEIRA (*Malus domestica* Borkh.) SOB DIFERENTES CONDIÇÕES CLIMÁTICAS – DA FORMAÇÃO DAS GEMAS À COLHEITA DOS FRUTOS**. Tese (Doutorado em Ciências) - Centro de Ciências Agrárias, Universidade Federal de Santa Catarina, Florianópolis, 2014.

GAO, F.; FU, L.; ZHANG, X.; et al. Multi-class fruit-on-plant detection for apple in SNAP system using Faster R-CNN. **Computers and Electronics in Agriculture**, v. 176, p. 105634, 2020.

GENÉ-MOLA, J.; VILAPLANA, V.; ROSELL-POLO, J. R.; et al. Multi-modal deep learning for Fuji apple detection using RGB-D cameras and their radiometric

capabilities. **Computers and Electronics in Agriculture**, v. 162, n. May, p. 689–698, 2019.

GONGAL, A.; AMATYA, S.; KARKEE, M.; ZHANG, Q.; LEWIS, K. Sensors and systems for fruit detection and localization: A review. **Computers and Electronics in Agriculture**, v. 116, p. 8–19, 2015.

GONGAL, A.; SILWAL, A.; AMATYA, S.; et al. Apple crop-load estimation with over-the-row machine vision system. **Computers and Electronics in Agriculture**, v. 120, p. 26–35, 2016.

GUILLEN-CLIMENT, M. L.; ZARCO-TEJADA, P. J.; BERNI, J. A. J.; NORTH, P. R. J.; VILLALOBOS, F. J. Mapping radiation interception in row-structured orchards using 3D simulation and high-resolution airborne imagery acquired from a UAV. **Precision Agriculture**, v. 13, n. 4, p. 473–500, 2012.

HABARAGAMUWA, H.; OGAWA, Y.; SUZUKI, T.; et al. Detecting greenhouse strawberries (mature and immature), using deep convolutional neural network. **Engineering in Agriculture, Environment and Food**, v. 11, n. 3, p. 127–138, 2018.

HÄNI, N.; ROY, P.; ISLER, V. A comparative study of fruit detection and counting methods for yield mapping in apple orchards. **Journal of Field Robotics**, v. 37, n. 2, p. 263–282, 2020.

HENDRYCKS, D.; DIETTERICH, T. Benchmarking neural network robustness to common corruptions and perturbations. 7th International Conference on Learning Representations, ICLR 2019. **Anais...** . p.1–16, 2019.

HERRERO-HUERTA, M.; GONZÁLEZ-AGUILERA, D.; RODRIGUEZ-GONZALVEZ, P.; HERNÁNDEZ-LÓPEZ, D. Vineyard yield estimation by automatic 3D bunch modelling in field conditions. **Computers and Electronics in Agriculture**, v. 110, p. 17–26, 2015.

HUNT, M. L.; BLACKBURN, G. A.; CARRASCO, L.; REDHEAD, J. W.; ROWLAND, C. S. High resolution wheat yield mapping using Sentinel-2. **Remote Sensing of Environment**, v. 233, p. 111410, 2019.

IBGE, I. B. DE G. E. E. PIB cai 4,1% em 2020 e fecha o ano em R\$ 7,4 trilhões. Disponível em: <<https://www.ibge.gov.br/>>. Acesso em: 29/3/2021.

IGLHAUT, J.; CABO, C.; PULITI, S.; et al. Structure from Motion Photogrammetry in Forestry: a Review. **Current Forestry Reports 2019 5:3**, v. 5, n. 3, p. 155–168, 2019.

JI, W.; QIAN, Z.; XU, B.; et al. Apple tree branch segmentation from images with small gray-level difference for agricultural harvesting robot. **Optik**, v. 127, n. 23, p. 11173–11182, 2016.

JIANG, P.; CHEN, Y.; LIU, B.; HE, D.; LIANG, C. Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks. **IEEE Access**, v. 7, p. 59069–59080, 2019. Institute of Electrical and Electronics Engineers Inc.

JIDONG, L.; DE-AN, Z.; WEI, J.; SHIHONG, D. Recognition of apple fruit in natural environment. **Optik**, v. 127, n. 3, p. 1354–1362, 2016.

KAHU, S. Y.; RAUT, R. B.; BHURCHANDI, K. M. Review and evaluation of color spaces for image/video compression. **Color Research and Application**, v. 44, n. 1, p. 8–33, 2019.

KALISPERAKIS, I.; STENTOUMIS, C.; GRAMMATIKOPOULOS, L.; KARANTZALOS, K. Leaf area index estimation in vineyards from UAV hyperspectral data, 2D image mosaics and 3D canopy surface models. **International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives**, v. 40, n. 1W4, p. 299–303, 2015.

KAMILARIS, A.; PRENAFETA-BOLDÚ, F. X. Deep learning in agriculture: A survey. **Computers and Electronics in Agriculture**, v. 147, n. July 2017, p. 70–90, 2018.

KANG, H.; CHEN, C. Fast implementation of real-time fruit detection in apple orchards using deep learning. **Computers and Electronics in Agriculture**, v. 168, p. 105108, 2020.

KAU, L. J.; LEE, T. L. An efficient and self-adapted approach to the sharpening of color images. **The Scientific World Journal**, v. 2013, 2013.

KESTUR, R.; MEDURI, A.; NARASIPURA, O. MangoNet: A deep semantic segmentation architecture for a method to detect and count mangoes in an open orchard. **Engineering Applications of Artificial Intelligence**, v. 77, n. April 2018, p. 59–69, 2019.

KIRK, R.; CIELNIAK, G.; MANGAN, M. L\*a\*b\*Fruits: A rapid and robust outdoor fruit detection system combining bio-inspired features with one-stage deep learning networks. **Sensors (Switzerland)**, v. 20, n. 1, p. 275, 2020.

KOIRALA, A.; WALSH, K. B.; WANG, Z.; MCCARTHY, C. Deep learning – Method overview and review of use for fruit detection and yield estimation. **Computers and Electronics in Agriculture**, v. 162, n. April, p. 219–234, 2019a.

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**, v. 20, n. 6, p. 1107–1135, 2019b.

LARA, L. M.; GEBLER, L.; LEITE JÚNIOR, M. J.; SOARES, A. L. Precision agriculture trends in fruit growing from 2016 to 2020. **Revista Brasileira de Fruticultura**, v. 43, n. 2, p. 1–14, 2021.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 2015.

LI, K.; WAN, G.; CHENG, G.; MENG, L.; HAN, J. Object detection in optical remote sensing images: A survey and a new benchmark. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 159, p. 296–307, 2020.

LIANG, X.; JAAKKOLA, A.; WANG, Y.; et al. The use of a hand-held camera for

individual tree 3D mapping in forest sample plots. **Remote Sensing**, v. 6, n. 7, p. 6587–6603, 2014.

LIANG, X.; LITKEY, P.; HYYPPÄ, J.; et al. Automatic stem mapping using single-scan terrestrial laser scanning. **IEEE Transactions on Geoscience and Remote Sensing**, v. 50, n. 2, p. 661–670, 2012.

LIMA, A. P. F. DE. **Fitorreguladores no retorno da floração e características agrônômicas de pereiras europeias**, Tese (Doutorado em Ciências) - Faculdade de Agronomia Eliseu Maciel, Universidade Federal de Pelotas, Pelotas, 2016.

LIN, T.-Y.; DOLLAR, P.; GIRSHICK, R.; et al. Feature Pyramid Networks for Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). **Anais...**, 2017.

LIN, T.-Y.; GOYAL, P.; GIRSHICK, R.; HE, K.; DOLLAR, P. Focal Loss for Dense Object Detection. IEEE International Conference on Computer Vision (ICCV). **Anais...** p.2980–2988, 2017.

LINKER, R. A procedure for estimating the number of green mature apples in night-time orchard images using light distribution and its application to yield estimation. **Precision Agriculture**, v. 18, n. 1, p. 59–75, 2017.

LINKER, R. Machine learning based analysis of night-time images for yield prediction in apple orchard. **Biosystems Engineering**, v. 167, p. 114–125, 2018.

LINKER, R.; KELMAN, E. Apple detection in nighttime tree images using the geometry of light patches around highlights. **Computers and Electronics in Agriculture**, v. 114, p. 154–162, 2015.

LIU, X.; CHEN, S. W.; ADITYA, S.; et al. Robust Fruit Counting: Combining Deep Learning, Tracking, and Structure from Motion. **IEEE International Conference on Intelligent Robots and Systems**, p. 1045–1052, 2018.

LIU, X.; ZHAO, D.; JIA, W.; et al. A method of segmenting apples at night based on color and position information. **Computers and Electronics in Agriculture**, v. 122, p. 118–123, 2016.

LUNA, I.; LOBO, A. Mapping crop planting quality in sugarcane from UAV imagery: A pilot study in Nicaragua. **Remote Sensing**, v. 8, n. 6, p. 1–18, 2016.

MARTINS, S. F. **Avaliação da eficiência em pomares de macieira na região de Carrazeda de Ansiães**, 2014.

MEHTA, S. S.; TON, C.; ASUNDI, S.; BURKS, T. F. Multiple camera fruit localization using a particle filter. **Computers and Electronics in Agriculture**, v. 142, p. 139–154, 2017.

MESAS-CARRASCOSA, F. J.; TORRES-SANCHEZ, J.; CLAVERO-RUMBAO, I.; et al. Assessing optimal flight parameters for generating accurate multispectral orthomosaics by uav to support site-specific crop management. **Remote Sensing**, v. 7, n. 10, p. 12793–12814, 2015.

MÉSZÁROS, M.; BĚLÍKOVÁ, H.; ČONKA, P.; NÁMĚSTEK, J. Effect of hail nets and fertilization management on the nutritional status, growth and production of apple trees. **Scientia Horticulturae**, v. 255, p. 134–144, 2019.

MICHAELIS, C.; MITZKUS, B.; GEIRHOS, R.; et al. Benchmarking Robustness in Object Detection: Autonomous driving when winter is coming. **arXiv**, , n. NeurIPS 2019, 2019.

MORIYA, E. A. S.; IMAI, N. N.; TOMMASELLI, A. M. G. Avaliação do potencial de índices de vegetação para detecção de doenças na cana-de-açúcar em imagens hiperespectrais adquiridas por VANT. Anais do XVIII Simpósio Brasileiro de Sensoriamento Remoto - SBSR. **Anais...** . p.886–893, 2017. Santos: INPE.

NGUYEN, T. T.; VANDEVOORDE, K.; WOUTERS, N.; et al. Detection of red and bicoloured apples on tree with an RGB-D camera. **Biosystems Engineering**, v. 146, p. 33–44, 2016.

NOGUEIRA, R. E. **Cartografia: Representação, Comunicação e Visualização de dados espaciais**. 3rd ed. Florianópolis: UFSC, 2009.

OHTA, Y.-I.; KANADE, T.; SAKAI, T. Color information for region segmentation. **Computer Graphics and Image Processing**, v. 13, n. 3, p. 222–241, 1980. Academic Press.

OLOFSSON, K.; HOLMGREN, J.; OLSSON, H. Tree stem and height measurements using terrestrial laser scanning and the RANSAC algorithm. **Remote Sensing**, v. 6, n. 5, p. 4323–4344, 2014.

OSCO, L. P.; DE ARRUDA, M. DOS S.; MARCATO JUNIOR, J.; et al. A convolutional neural network approach for counting and geolocating citrus-trees in UAV multispectral imagery. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 160, p. 97–106, 2020.

OTSU, N. A Threshold Selection Method from Gray-Level Histograms. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 9, n. 1, p. 62–66, 1979.

PANG, J.; CHEN, K.; SHI, J.; et al. Libra R-CNN: Towards balanced learning for object detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. **Anais...** . v. 2019-June, p.821–830, 2019. IEEE.

PAPAGEORGIU, E. I.; AGGELOPOULOU, K. D.; GEMTOS, T. A.; NANOS, G. D. Yield prediction in apples using Fuzzy Cognitive Map learning approach. **Computers and Electronics in Agriculture**, v. 91, p. 19–29, 2013.

PETRI, J. L.; DENARDI, F.; SUZUKI, A. EPAGRI. 405-Fuji Suprema: Nova cultivar de macieira. **Agropecuária Catarinense**, v. 10, n. 3, p. 48–50, 1997.

PETRI, J. L.; LEITE, G. B.; COUTO, M.; FRANCESCETTO, E. P. AVANÇOS NA CULTURA DA MACIEIRA NO BRASIL. **Rev. Bras. Frutic.**, v. Especial, p. 048–056, 2011. Jaboticabal-SP.

RAFAEL C. GONZALEZ, R. E. W. **Processamento de Imagens Digitais**. Editora Blucher, 2000.

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**, v. 39, n. 6, p. 1137–1149, 2017.

SANTOS, H. G. DOS; JACOMINE, P. K. T.; ANJOS, L. H. C. DOS; et al. **Sistema brasileiro de classificação de solos**. 5th ed. Brasília: Embrapa, 2018.

SANZ-CORTIELLA, R.; LLORENS-CALVERAS, J.; ESCOLÀ, A.; et al. Innovative LIDAR 3D dynamic measurement system to estimate fruit-tree leaf area. **Sensors**, v. 11, n. 6, p. 5769–5791, 2011.

SANZ, R.; LLORENS, J.; ESCOLÀ, A.; et al. LIDAR and non-LIDAR-based canopy parameters to estimate the leaf area in fruit trees and vineyard. **Agricultural and Forest Meteorology**, v. 260–261, n. May, p. 229–239, 2018.

SCURI, A. E. **Fundamentos da Imagem Digital**. Rio de Janeiro, 2002.

SHAIK, K. B.; GANESAN, P.; KALIST, V.; SATHISH, B. S.; JENITHA, J. M. M. Comparative Study of Skin Color Detection and Segmentation in HSV and YCbCr Color Space. **Procedia Computer Science**, v. 57, p. 41–48, 2015.

SHALAL, N.; LOW, T.; MCCARTHY, C.; HANCOCK, N. Orchard mapping and mobile robot localisation using on-board camera and laser scanner data fusion - Part A: Tree detection. **Computers and Electronics in Agriculture**, v. 119, p. 254–266, 2015.

SRIVASTAVA, L. M. Fruit Development and Ripening. **Plant Growth and Development**. p.413–429, 2002.

STAFFORD, J. V. Implementing precision agriculture in the 21st century. **Journal of Agricultural and Engineering Research**, v. 76, n. 3, p. 267–275, 2000.

STEIN, M.; BARGOTI, S.; UNDERWOOD, J. Image Based Mango Fruit Detection, Localisation and Yield Estimation Using Multiple View Geometry. **Sensors**, v. 16, n. 11, p. 1915, 2016.

TIAN, Y.; YANG, G.; WANG, Z.; et al. Apple detection during different growth stages in orchards using the improved YOLO-V3 model. **Computers and Electronics in Agriculture**, v. 157, p. 417–426, 2019.

TIJSKENS, L. M. M.; UNUK, T.; OKELLO, R. C. O.; et al. From fruitlet to harvest: Modelling and predicting size and its distributions for tomato, apple and pepper fruit. **Scientia Horticulturae**, v. 204, p. 54–64, 2016.

UNDERWOOD, J. P.; HUNG, C.; WHELAN, B.; SUKKARIEH, S. Mapping almond orchard canopy volume, flowers, fruit and yield using lidar and vision sensors. **Computers and Electronics in Agriculture**, v. 130, p. 83–96, 2016.

UZQUIANO, S.; MARTÍNEZ, J.; SAN MARTÍN, R.; BRAVO, F. Mediciones dendrométricas y dasométricas mediante técnicas LiDAR y fotogramétricas.

**Cuadernos de la Sociedad Española de Ciencias Forestales**, v. 40, p. 193–202, 2014.

VÁZQUEZ-ARELLANO, M.; GRIEPENTROG, H. W.; REISER, D.; PARAFOROS, D. S. 3-D imaging systems for agricultural applications—a review. **Sensors (Switzerland)**, v. 16, n. 5, 2016.

VIANA, C. D. **Análise Estrutural De Descontinuidades Baseada Em Técnicas De Structure From Motion : Aplicação Em Mina a Céu Aberto**Tese, 2015. USP.

WANG, D.; LI, C.; SONG, H.; et al. Deep Learning Approach for Apple Edge Detection to Remotely Monitor Apple Growth in Orchards. **IEEE Access**, v. 8, p. 26911–26925, 2020. Institute of Electrical and Electronics Engineers Inc.

WANG, J.; SUN, K.; CHENG, T.; et al. Deep High-Resolution Representation Learning for Visual Recognition. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, p. 1–1, 2020. Institute of Electrical and Electronics Engineers (IEEE).

WEI, X.; JIA, K.; LAN, J.; et al. Automatic method of fruit object extraction under complex agricultural background for vision system of fruit picking robot. **Optik**, v. 125, n. 19, p. 5684–5689, 2014.

WESTOBY, M. J.; BRASINGTON, J.; GLASSER, N. F.; HAMBREY, M. J.; REYNOLDS, J. M. Structure-from-Motion photogrammetry: A low-cost, effective tool for geoscience applications. **Geomorphology**, p. 300–314, 2012.

WU, D.; LV, S.; JIANG, M.; SONG, H. Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. **Computers and Electronics in Agriculture**, v. 178, p. 105742, 2020.

ZHANG, S.; CHI, C.; YAO, Y.; LEI, Z.; LI, S. Z. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection. **Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition**, p. 9756–9765, 2020.

ZHANG, X.; HAN, LIANGXIU; HAN, LIANGHAO; ZHU, L. How well do deep learning-based methods for land cover classification and object detection perform on high resolution remote sensing imagery? **Remote Sensing**, v. 12, n. 3, p. 417, 2020.

ZHAO, C.; LEE, W. S.; HE, D. Immature green citrus detection based on colour feature and sum of absolute transformed difference (SATD) using colour images in the citrus grove. **Computers and Electronics in Agriculture**, v. 124, p. 243–253, 2016.

ZHAO, Y.; GONG, L.; HUANG, Y.; LIU, C. A review of key techniques of vision-based control for harvesting robot. **Computers and Electronics in Agriculture**, v. 127, p. 311–323, 2016.

ZHOU, R.; DAMEROW, L.; SUN, Y.; BLANKE, M. M. Using colour features of cv. “Gala” apple fruits in an orchard in image processing to predict yield. **Precision Agriculture**, v. 13, n. 5, p. 568–580, 2012.

ZHU, C.; HE, Y.; SAVVIDES, M. Feature selective anchor-free module for single-shot object detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. **Anais...** . v. 2019-June, p.840–849, 2019. IEEE.