



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE RUSSAS
CURSO DE GRADUAÇÃO EM ENGENHARIA DE SOFTWARE

NATHALIA CAVALCANTI GONÇALVES

**USO DE VISÃO COMPUTACIONAL PARA CLASSIFICAÇÃO DE FRUTOS DE
ACEROLA EM IMAGENS RGB**

RUSSAS

2021

NATHALIA CAVALCANTI GONÇALVES

USO DE VISÃO COMPUTACIONAL PARA CLASSIFICAÇÃO DE FRUTOS DE ACEROLA
EM IMAGENS RGB

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Engenharia de Software
do Campus de Russas da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Engenharia de Software.

Orientadora: Prof. Dra. Tatiane Fernan-
des Figueiredo

RUSSAS

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- G626u Gonçalves, Nathalia Cavalcanti.
Uso de visão computacional para classificação de frutos de acerola em imagens RGB. / Nathalia Cavalcanti Gonçalves. – 2021.
39 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Russas, Curso de Engenharia de Software, Russas, 2021.
Orientação: Prof. Dr. Tatiane Fernandes Figueiredo.
1. Agricultura 4.0.. 2. Smart Farming.. 3. Colheita Automatizada.. 4. Visão Computacional.. 5. Aprendizado de Máquina.. I. Título.

CDD 005.1

NATHALIA CAVALCANTI GONÇALVES

USO DE VISÃO COMPUTACIONAL PARA CLASSIFICAÇÃO DE FRUTOS DE ACEROLA
EM IMAGENS RGB

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Engenharia de Software
do Campus de Russas da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Engenharia de Software.

Aprovada em:

BANCA EXAMINADORA

Prof. Dra. Tatiane Fernandes
Figueiredo (Orientadora)
Universidade Federal do Ceará (UFC)

Prof. Dr. Bonfim Amaro Júnior
Universidade Federal do Ceará (UFC)

Prof. Ms. Eurinaldo Rodrigues Costa
Universidade Federal do Ceará (UFC)

AGRADECIMENTOS

A Deus, por não me desamparar nos momentos em que eu precisei de fé.

À minha família, por acreditar e confiar que eu sou capaz de alcançar todos os meus sonhos.

Ao NEMO - Laboratório de Pesquisa e Desenvolvimento em Aprendizado de Máquina e Otimização Núcleo e à todos que fazem parte dele, por ter tido a oportunidade de conhecer um mundo novo.

À Profa. Ms. Tatiane Fernandes Figueiredo por ter aceitado orientar meu trabalho de conclusão de curso.

Aos meus amigos, os que eu já tinha antes e os que fiz na universidade, sou grata por tê-los em minha vida.

Ao Prof. Dr. Tobias Rafael Fernandes Neto, coordenador do Laboratório de Sistemas Motrizes (LAMOTRIZ) onde este *template* foi desenvolvido.

Ao Doutorando em Engenharia Elétrica, Ednardo Moreira Rodrigues, e seu assistente, Alan Batista de Oliveira, aluno de graduação em Engenharia Elétrica, pela adequação do *template* utilizado neste trabalho para que o mesmo ficasse de acordo com as normas da biblioteca da Universidade Federal do Ceará (UFC).

Ao aluno Thiago Nascimento do curso de ciência da computação da Universidade Estadual do Ceará que elaborou o *template* do qual este trabalho foi adaptado para Universidade Federal do Ceará.

“Para alguém que tenha fé, nenhuma explicação é necessária. Para aquele sem fé, nenhuma explicação é possível.”

(São Tomás de Aquino)

RESUMO

Com o advento da Agricultura 4.0, a evolução da produção agrícola tem tornado cotidiano o uso da tecnologia aplicada ao campo. Dentre as inovações emergentes, tem se popularizado a colheita automatizada, técnica que reconhece e classifica padrões em imagens de cultivos. Por esse motivo, são necessárias abordagens efetivas, que tragam segurança nos seus resultados. A acerola é um fruto muito consumido no Brasil e no exterior. Porém ainda não existem estudos do uso de visão computacional aplicados a essa cultura. Este trabalho tem como objetivo apresentar um modelo inteligente, utilizando técnicas de visão computacional e aprendizado de máquina, capaz de classificar frutos de acerola em imagens RGB em duas classes: coloração avermelhada ou verde.

Palavras-chave: Agricultura 4.0. *Smart Farming*. Colheita Automatizada. Visão Computacional. Aprendizado de Máquina.

ABSTRACT

With the rising popularization of the Agriculture 4.0, the evolution of agricultural production has made customary the use of technology in the fields. Among the emerging innovations, automated harvesting, a technique that recognizes and classifies patterns in crop images, has become popular. For this reason, effective approaches to bring security to your results are needed. Acerola cherry is a fruit widely consumed in Brazil and abroad, however, there are still no studies on the use of computer vision applied to this crop. This study aims to present an intelligent model, using computer vision and machine learning techniques, capable of classifying acerola cherry fruits in RGB images in two classes: reddish or green coloring.

Keywords: Agriculture 4.0. Smart Farming. Automated harvesting. Computer vision. Machine Learnig

LISTA DE FIGURAS

Figura 1 – Processo de reconhecimento de imagens	17
Figura 2 – Processo de reconhecimento de imagens	18
Figura 3 – Processo de detecção de objetos	19
Figura 4 – Processo de procura seletiva	19
Figura 5 – Classificação com vetores de suporte	20
Figura 6 – Representação do conjunto de imagens criado por Leite (2020)	25
Figura 7 – Representação da segmentação adaptada da base de Leite (2020)	26
Figura 8 – Representação da avaliação do IoU	27
Figura 9 – Matriz de confusão de uma classificação	28
Figura 10 – Matriz de confusão do modelo 50-50	29
Figura 11 – Relatório do modelo 50-50	29
Figura 12 – Matriz de confusão do modelo 60-40	30
Figura 13 – Relatório do modelo 60-40	30
Figura 14 – Matriz de confusão do modelo 70-30	31
Figura 15 – Relatório do modelo 70-30	31
Figura 16 – Matriz de confusão do modelo 80-20	32
Figura 17 – Relatório do modelo 80-20	32
Figura 18 – Matriz de confusão do modelo 90-10	33
Figura 19 – Relatório do modelo 90-10	33
Figura 20 – Previsões de plano de fundo	34
Figura 21 – Previsões de frutos verdes	34
Figura 22 – Previsões de frutos vermelhos	34
Figura 23 – Região prevista pelo modelo 50-50	35
Figura 24 – Região prevista pelo modelo 60-40	35
Figura 25 – Região prevista pelo modelo 70-30	35
Figura 26 – Região prevista pelo modelo 80-20	36
Figura 27 – Região prevista pelo modelo 90-10	36

LISTA DE ABREVIATURAS E SIGLAS

AM	Aprendizado de Máquina ou <i>Machine Learning</i>
CNN	Rede Neural Convolutacional ou <i>Convolutional Neural Network</i>
DLNN	Rede Neural Profunda ou <i>Deep Learning Neural Network</i>
IoU	Interseccção sobre União ou <i>Intersection Over Union</i>
KNN	K-ésimo Vizinho mais Próximo ou <i>K-Nearest Neighbor</i>
MLP	Perceptron Multi-camada ou <i>Multilayer Perceptron</i>
PDI	Processamento Digital de Imagens ou <i>Digital Image Processing</i>
R-CNN	Rede Neural Convolutacional Baseada em Região ou <i>Region Based Convolutional Neural Network</i>
RPN	Rede de Proposta de Região ou <i>Region Proposal Network</i>
SVM	Máquina de Vetores de Suporte ou <i>Support Vector Machine</i>
VC	Visão Computacional ou <i>Computer Vision</i>

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Objetivos	13
<i>1.1.1</i>	<i>Objetivo Geral</i>	<i>13</i>
<i>1.1.2</i>	<i>Objetivos Específicos</i>	<i>13</i>
2	FUNDAMENTAÇÃO TEÓRICA	14
2.1	<i>Smart Farming</i>	<i>14</i>
2.2	Colheita automatizada	14
2.3	Visão Computacional	15
<i>2.3.1</i>	<i>Aprendizado de Máquina</i>	<i>15</i>
<i>2.3.2</i>	<i>Processamento Digital de Imagens</i>	<i>16</i>
<i>2.3.3</i>	<i>Etapas para reconhecimento de Imagens</i>	<i>16</i>
<i>2.3.4</i>	<i>Redes Neurais Convolucionais</i>	<i>18</i>
<i>2.3.4.1</i>	<i>Redes Neurais Convolucionais Baseadas em Região</i>	<i>19</i>
3	TRABALHOS RELACIONADOS	21
3.1	Segmentação de Imagens para Detecção de Frutas e Estimativa de Produção em Pomares de Maçã	21
3.2	Reconhecimento Automático de Frutas em Imagens Naturais usando Características de Cor e Textura	21
3.3	Reconhecimento de Frutas em Imagens usando Aprendizagem Profunda	22
3.4	Análise de Viabilidade do uso de Aprendizagem Profunda para Detecção de Frutos de Acerola em Imagens	22
3.5	Comparativo de atividades do estado da arte	23
4	PROCEDIMENTOS METODOLÓGICOS	24
4.1	Estudo Teórico	24
4.2	Estudo Prático	24
<i>4.2.1</i>	<i>Aquisição e Segmentação de Imagens</i>	<i>24</i>
<i>4.2.2</i>	<i>Geração do modelo de detecção e classificação</i>	<i>26</i>
5	RESULTADOS	28
5.1	Experimento prático	28
6	CONCLUSÃO	37

6.1	Considerações gerais	37
6.2	Trabalhos futuros	37
	REFERÊNCIAS	38

1 INTRODUÇÃO

Técnicas de visão computacional, aplicadas na criação de robôs autônomo, para análise de frutos em seu ambiente natural têm sido amplamente utilizadas como mecanismo de otimização produtiva em grandes fazendas, principalmente, em países do exterior. No entanto, não há registro do uso dessa tecnologia para otimização de colheita de frutos de acerola. Dentre os trabalhos apresentados na literatura que trazem o uso de Processamento Digital de Imagens ou *Digital Image Processing* (PDI) para reconhecimento e classificação de frutos, destacam-se os autores Arivazhagan *et al.* (2010), que trazem um método para reconhecimento de frutos por cor e textura, Dubey e Jalal (2012), apresentam uma abordagem para classificação de frutos e vegetais usando características de textura e Jana *et al.* (2017), que propõem classificar tipos diferentes de frutos em ambientes naturais. Há também autores que utilizam PDI para automatizar fazendas de fruticultura, como, Bargoti e Underwood (2017) que apresenta estimativas de colheita e Bhange e Hingoliwala (2015) que utiliza modelos inteligentes para detecção de doenças em frutos.

Ao considerar os cultivos de maior produção na fruticultura, a demanda da acerola orgânica, principalmente pela indústria farmacêutica, tem crescido a cada ano de forma grandiosa, gerando interesse de fruticultores por tecnologias que possam otimizar o processo de cultivo deste fruto. Segundo o Ipea - Instituto de Pesquisa Econômica Aplicada (2020), o Brasil se encontra em 12º lugar entre os países com as maiores áreas de produção orgânica do mundo. Porém, a capacidade de produzir em grande escala entra em pauta ao se tratar dos produtores, pois muitos não têm o treinamento e tecnologias necessárias para permitir melhor condução do negócio (MARTINS *et al.*, 2016).

Frutos de acerola possuem um alto valor de mercado, não só para consumo direto, mas principalmente pelo seu uso como insumo em diversos produtos medicinais, sendo uma das principais formas de extração de vitamina C (ASSIS *et al.*, 2008). De acordo com Diário do Nordeste (2014), o Ceará foi considerado o quarto maior produtor de frutas do país, dentre as fazendas produtoras de acerola existentes no estado, destaca-se a Fazenda Meri Pobo Agropecuária, atualmente considerada a maior fazenda de produção de acerola orgânica do mundo (SERPA, 2019), objeto de estudo deste trabalho.

Neste trabalho é proposto a criação de um modelo inteligente que classifica frutos de acerola contidas em imagens RGB em duas classes: coloração avermelhada ou verde. Para nosso melhor conhecimento, ainda não existem estudos aplicados a classificação de frutos de acerola em imagens RGB, e portanto uma análise de viabilidade do uso de algoritmos de visão computacional.

cional neste cenário pode trazer diversos benefícios para produtores. A base de dados utilizada para treinamento e teste do modelo foi gerada a partir de imagens de aceroleiras retiradas da Fazenda Meri Pobo Agropecuária, juntamente com imagens semelhantes encontradas na plataforma *Google* Imagens (LEITE, 2020). Buscando analisar e definir as melhores configurações, necessárias para obtenção de um modelo com taxas de acertos que possam ser consideradas de qualidade, os experimentos práticos foram definidos considerando vários cenários possíveis para treinamento e teste do modelo. Como conclusão, pode-se afirmar que a melhor configuração de treinamento é a do modelo com 60% de dados para treino e 40% para teste.

O restante deste documento está organizado conforme descrito a seguir. O Capítulo 2 apresenta a fundamentação teórica, na qual foi baseada esta pesquisa. No Capítulo 3 são apresentados os trabalhos relacionados a este. O Capítulo 4 descreve a metodologia utilizada para realização do trabalho. O Capítulo 5 descreve os resultados obtidos na pesquisa. E o Capítulo 6 apresenta uma conclusão sobre o estudo e possíveis trabalhos futuros.

1.1 Objetivos

1.1.1 Objetivo Geral

Apresentar um modelo inteligente, utilizando técnicas de visão computacional e aprendizado de máquina, capaz de classificar frutos de acerola em imagens RGB em duas classes: coloração avermelhada ou verde.

1.1.2 Objetivos Específicos

- Estudo do estado da arte atual relacionado a classificação de frutos semelhante ao fruto de acerola;
- Estudo de técnicas de visão computacional para reconhecimento e classificação de objetos em imagens RGB;
- Realização de testes com diferentes partições da base de imagens para serem utilizadas nas etapas de treinamento e teste do modelo;
- Análise estatística a cerca dos resultados obtidos.

2 FUNDAMENTAÇÃO TEÓRICA

Para entendimento do trabalho serão necessários conceitos de *smart farming*, colheita automatizada e visão computacional. Que serão apresentados nas seções 2.1, 2.2 e 2.3 respectivamente.

2.1 *Smart Farming*

O uso de tecnologias no setor agrícola num país como o Brasil, que recentemente se tornou um produtor e exportador estratégico de produtos básicos, têm contribuído para o aumento da produção de alimentos e minimização do uso de recursos naturais e mudanças climáticas (PIVOTO *et al.*, 2018).

A agricultura tem passado por muitas inovações com a chegada de ferramentas computacionais ao campo. Com o advento da Agricultura 4.0, mudanças estão acontecendo diariamente. Dentre os vários conceitos, técnicas e estudos científicos aplicados ao agronegócio, destaca-se o conceito de *Smart Farming* que tem como foco o uso de tecnologias em máquinas, equipamentos e sensores agrícolas, permitindo uma progressiva inserção de automação no processo de produção agrícola (PIVOTO *et al.*, 2018).

O conceito de Agricultura 4.0 engloba as inovações propostas pelo *smart farming* com o intuito de trazer tecnologias avançadas para o campo, gerando benefícios e maior produtividade em fazendas. Essa recente revolução, que incorpora conectividade e automação, com uso de veículos, *drones*, robôs e animais com sensores, tem sido um mecanismo gerador de informação e *insights* em tempo real, auxiliando agricultores em tomadas de decisão sobre seus cultivos (SANTOS *et al.*, 2019).

2.2 Colheita automatizada

A colheita automatizada tem sido um campo emergente na Agricultura 4.0. Dentre as diversas tecnologias utilizadas para automação da colheita, destaca-se um crescente uso de técnicas de Visão Computacional ou *Computer Vision* (VC) e Aprendizado de Máquina ou *Machine Learning* (AM) para reconhecimento e análise de imagens de cultivos. Os resultados obtidos, através de ferramentas computacionais nesta vertente, têm se apresentado promissores, sendo utilizados principalmente para análise de informações relevantes em diversos tipos de culturas. Como exemplo, o crescimento e maturação de frutas e legumes, dentre outros aspectos

de cultivos específicos (JANA *et al.*, 2017).

Através do uso de reconhecimento de padrões em imagens é possível mapear plantações e classificar informações sobre os padrões detectados, gerando informações relevantes referentes as culturas. Como exemplo, podemos citar o uso de VC para mensurar a quantidade média necessária de recursos durante uma etapa de plantio, ou até mesmo, automatizar a colheita através da análise de previsões relativas a possíveis melhores datas para sua realização. Para que se consiga alcançar resultados como os descritos anteriormente, técnicas robustas e precisas de PDI são necessárias (BARGOTI; UNDERWOOD, 2017). As próximas seções apresentam alguns conceitos importantes sobre Visão Computacional e reconhecimento de imagens, assim como as principais técnicas utilizadas para extrair informações de alto nível de imagens de cultivos, como localização, saúde, maturidade, carga da colheita (rendimento), e distribuição espacial.

2.3 Visão Computacional

A Visão Computacional (VC) é uma subárea de estudo da Inteligência Artificial que tem como foco simular em uma máquina a visão humana por meio de imagens digitais, sejam de câmeras, sensores ou outros dispositivos (MILANO; HONORATO, 2014). A VC consiste na distinção de objetos presentes em imagens, tendo como saída uma inferência sobre o mesmo, simulando a dedução humana ao manipular, pensar e reconhecer o conteúdo presente no arquivo digital (SCHALKOFF, 1989).

2.3.1 Aprendizado de Máquina

A área de estudo em Aprendizado de Máquina (AM), do inglês *Machine Learning*, é reponsável por desenvolver sistemas que sejam capazes de adquirir conhecimento, ou seja, programas de computador que são capazes de aprender baseados em experiências anteriores e a partir disso tomar decisões de forma autônoma (MONARD; BARANAUSKAS, 2003a). A literatura atual classifica os algoritmos de AM em quatro principais abordagens:

- **Aprendizado Supervisionado:** o algoritmo recebe um conjunto de dados já categorizados, ou seja, que foram rotulados baseados em suas categorias, e após a etapa de treinamento/aprendizagem, é capaz de rotular novas entradas (KHAN *et al.*, 2010). Normalmente essa técnica é utilizada na classificação automática de textos, por exemplo, determinar a área de estudo de um artigo de acordo com palavras-chave.

- **Aprendizado Não Supervisionado:** o algoritmo recebe dados de entrada não categorizados, tendo como objetivo buscar características em comum entre os dados de entrada para que possam ser agrupados e posteriormente reconhecidos em grupos similares (AYODELE, 2010).
- **Aprendizado Por Reforço:** frequentemente utilizado em aprendizado de robôs, o algoritmo aprende com seus erros e acertos. Quando uma ação é considerada satisfatória, esse algoritmo será induzido a repetir essa ação através de recompensas. Caso contrário, receberá punições para desincentivar o comportamento indesejado.
- **Aprendizado Profundo:** esse tipo de aprendizado é amplamente utilizado para reconhecimento de imagens. Funciona agregado à uma rede neural, onde os algoritmos aprendem a distinguir objetos baseado em dados rotulados previamente. Neste caso, é necessário que as imagens fornecidas para o treinamento já estejam classificadas para que o algoritmo possa aprender as categorizações desejadas e assim ser capaz de reconhecer padrões com autonomia (SOUZA *et al.*, 2020).

2.3.2 *Processamento Digital de Imagens*

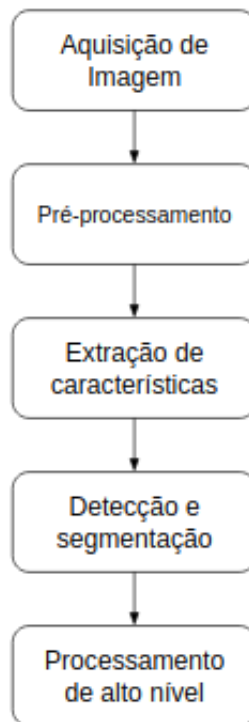
Técnicas de Processamento Digitais de Imagens (PDI) decorrem de duas principais aplicações: a) melhorar a informação visual de uma imagem para análise e interpretação humanas e b) extração de informações de uma imagem via processamento computacional (GONZALEZ; WOODS, 2000). Neste trabalho será utilizada a abordagem b).

O objetivo do PDI é fazer com que uma máquina seja capaz de diferenciar objetos em imagens assim como um humano é capaz. Então, com uma imagem de entrada, um algoritmo de PDI irá procurar padrões, que foram submetidos em uma base de dados, que serão reconhecidos ou não e retornará uma inferência baseada na análise feita.

2.3.3 *Etapas para reconhecimento de Imagens*

De maneira geral, os sistemas que fazem uso de técnicas de PDI tem as seguintes etapas em comum: reconhecer objetos em imagens e realizar inferências sobre os resultados desse reconhecimento (MILANO; HONORATO, 2014). Rehem e Trindade (2009) descrevem as etapas para reconhecimento de imagens da seguinte maneira, como apresentado na Figura 1:

Figura 1 – Processo de reconhecimento de imagens



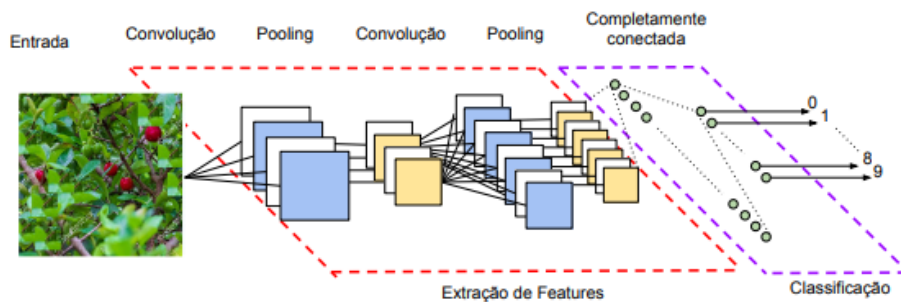
Fonte: Elaborado pelo Autor (2020).

- **Aquisição de imagens:** as imagens podem ser obtidas por máquinas como câmeras, radiografias, sensores, etc. Esse conjunto será a base de dados a ser utilizada.
- **Pré-processamento:** faz uso de técnicas que destacam detalhes importantes que poderão facilitar o reconhecimento do objeto.
- **Extração de características:** busca regiões que se diferenciam do restante da imagem, a fim de encontrar o objeto desejado.
- **Detecção e segmentação:** utilizando as regiões que foram encontradas na etapa anterior, nessa etapa as regiões serão destacadas do plano de fundo da imagem. É a etapa crítica do processo, pois se a segmentação for feita de maneira errônea, a detecção não encontrará o objeto.
- **Processamento de alto nível:** detecta e classifica as imagens segmentadas com intuito de reconhecer ou não o objeto. Alguns dos algoritmos mais utilizados são Máquina de Vetores de Suporte ou *Support Vector Machine* (SVM), K-ésimo Vizinho mais Próximo ou *K-Nearest Neighbor* (KNN) e Rede Neural Convolutiva ou *Convolutional Neural Network* (CNN) (MUKHERJEE *et al.*, 2017). As CNN estão sendo usadas progressivamente graças aos bons resultados e alta precisão no reconhecimento de imagens (DATA SCIENCE ACADEMY, 2018).

2.3.4 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (CNN) são redes de aprendizado profundo amplamente utilizadas na área de PDI. Esse algoritmo faz uso de camadas convolucionais para extração de características, ou seja, atribui pesos aos objetos que estão aparecendo na imagem. De forma geral, este tipo de rede é composta por camadas de convolução, camadas de agrupamento, do inglês *pooling*, e camadas completamente conectadas (FILHO, 2017) como apresentado na Figura 2.

Figura 2 – Processo de reconhecimento de imagens



Fonte: Adaptado de (VARGAS *et al.*, 2016).

Segundo Vargas *et al.* (2016) as camadas funcionam da seguinte maneira:

- **Extração de características:** a entrada passa pela camada de convolução, onde são aplicados filtros na imagem de estudo, para geração de um mapa de características, que representa os *pixels* que serão processados pelos neurônios. Depois, o mapa entra na camada de agrupamento, que é responsável por simplificar a saída diminuindo seu tamanho. Uma entrada pode passar por essas camadas inúmeras vezes dependendo da arquitetura utilizada.
- **Classificação:** é necessário uma camada completamente conectada que é responsável por aplicar um algoritmo classificador, por exemplo o método do gradiente descendente, que dado um valor de erro, obtido através do aprendizado do algoritmo, procura modificar o peso atribuído ao possível objeto para reduzir o erro na detecção de características, realizando desta forma o treinamento do modelo (FERREIRA *et al.*, 2016).

Para realizar o reconhecimento de frutos de acerola de coloração verde ou avermelhada, será utilizado o algoritmo pertencente a classe dos algoritmos de Redes Neurais Convolucionais

descrito a seguir.

2.3.4.1 Redes Neurais Convolucionais Baseadas em Região

O trabalho de Girshick *et al.* (2013) apresentou uma nova abordagem ao uso de CNNs, ao propor localizar objetos com uma rede neural profunda utilizando uma pequena quantidade de dados rotulados. A Figura 3 ilustra o processo de detecção e classificação de objetos com uma Rede Neural Convolutiva Baseada em Região ou *Region Based Convolutional Neural Network* (R-CNN).

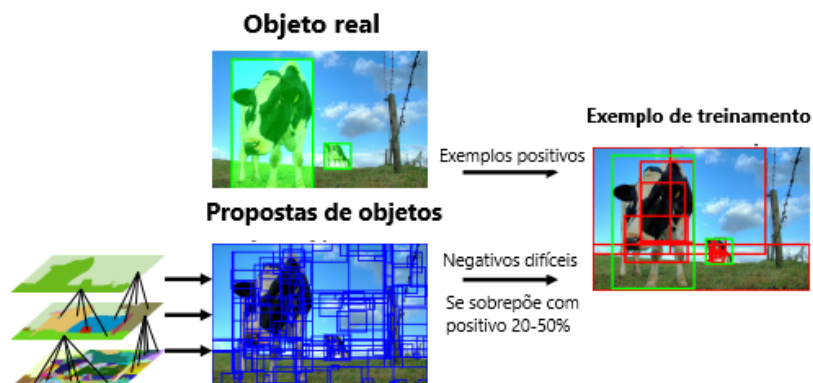
Figura 3 – Processo de detecção de objetos



Fonte: Adaptado de (GIRSHICK *et al.*, 2013)

O método R-CNN consiste em gerar 2.000 regiões propostas para cada imagem de entrada, usando o método de procura seletiva. Para gerar as regiões propostas, esse método combina estratégias de segmentação e busca exaustiva, usando as regiões segmentadas para gerar regiões de possíveis objetos, considerando todas as escalas do objeto, iluminação e planos de fundo (UIJLINGS *et al.*, 2013). Assim como ilustrado na Figura 4

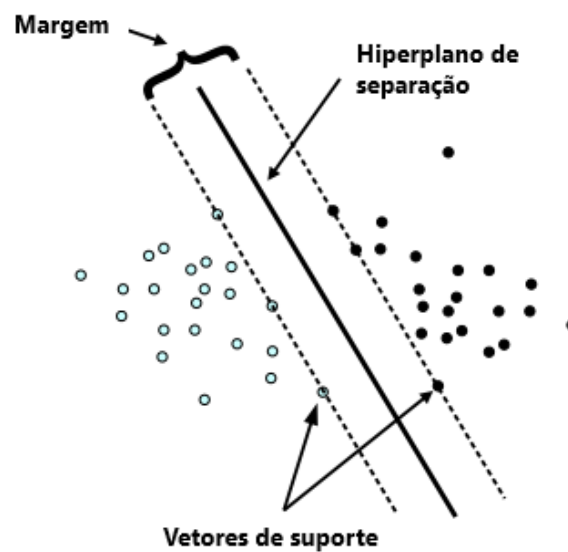
Figura 4 – Processo de procura seletiva



Fonte: Adaptado de (UIJLINGS *et al.*, 2013)

Em seguida, é feita a extração de características de cada região proposta usando uma CNN e para etapa de processamento de alto nível, uma SVM, vai analisar os dados e reconhecer se há ou não um objeto nessa região. Uma SVM é um classificador binário, onde os dados são separados em duas classes, buscando o hiperplano de separação ótimo, maximizando a distância da margem entre os vetores de suporte, que são os pontos mais próximos entre as classes (MEYER; WIEN, 2015). Assim como ilustrado na Figura 5.

Figura 5 – Classificação com vetores de suporte



Fonte: Adaptado de (MEYER; WIEN, 2015)

3 TRABALHOS RELACIONADOS

Nesta seção serão apresentados trabalhos que realizaram processamento de imagens digitais para reconhecimento e classificação de frutas.

3.1 Segmentação de Imagens para Detecção de Frutas e Estimativa de Produção em Pomares de Maçã

Bargoti e Underwood (2017) propõem um *framework* de processamento de imagens para detecção e contagem de frutos usando dados de imagens de pomares. Para tal, foram comparadas duas propostas arquiteturas de segmentação de imagem, Rede Neural Perceptron Multi-camada ou *Multilayer Perceptron* (MLP) e CNN. As imagens foram obtidas por câmeras acopladas em veículos terrestres não tripulados, que passavam entre os pomares na estação antes da colheita.

Foram comparados os resultados da MLP e da CNN com e sem uso de metadados, onde os metadados são informações de como as imagens foram capturadas, por exemplo, a distância do veículo para os pomares. Incluir essas informações pode melhorar o resultado da classificação. Foi observado que a CNN obteve uma melhor acurácia, ou seja, a porcentagem de quantas classificações corretas foram realizadas pelo algoritmo.

3.2 Reconhecimento Automático de Frutas em Imagens Naturais usando Características de Cor e Textura

Jana *et al.* (2017) tem como objetivo classificar diferentes tipos de frutos considerando como métrica de desempenho a acurácia obtida pelo modelo proposto. A abordagem definida pelos autores propõe uma diminuição no uso intensivo de recursos para processamento, sendo indicado para sistemas embarcados e computadores de placa única. De forma similar a monografia proposta, este trabalho traz o reconhecimento imagens de frutos em seu ambiente natural, considerado um desafio pelo autores, justificado pelos fato dos objetos do primeiro plano das imagens estudadas possuírem cores semelhantes as do plano de fundo.

A metodologia proposta pelos autores segue as seguintes etapas. Inicialmente, as imagens foram pré-processadas para separar o plano de fundo do primeiro plano. Logo após foram extraídas as características, são utilizadas textura e cor como *features*. O método obteve uma acurácia de 83.33% utilizando um modelo de classificação com uma SVM.

3.3 Reconhecimento de Frutas em Imagens usando Aprendizagem Profunda

Neste trabalho os autores apresentam uma base de dados contendo imagens de frutos populares com o intuito de testar a utilização de uma Rede Neural Profunda ou *Deep Learning Neural Network* (DLNN) para identificar frutas em realidade aumentada. As imagens foram obtidas a partir de filmagens, onde as frutas eram rotacionadas por um motor utilizando um plano de fundo branco. Posteriormente, foram aplicadas técnicas para aumentar a base de dados através de filtragens aleatórias para alterar a saturação das imagens.

Com o objetivo de melhorar a acurácia, anteriormente a aplicação do algoritmo de detecção, também foram concatenadas imagens em escala de cinza à base, extraíndo os objetos do plano de fundo, através de um algoritmo de inundação utilizando marcação de *pixels*. De início são marcados os *pixels* da borda das imagens, os *pixels* vizinhos serão marcados caso a diferença das cores seja menor que um dado valor, todo *pixel* marcado foi considerado plano de fundo e o restante foi considerado objeto. Mureşan e Oltean (2018) utilizaram uma CNN para o reconhecimento de imagens obtendo uma alta acurácia no reconhecimento de 60 classificações.

3.4 Análise de Viabilidade do uso de Aprendizagem Profunda para Detecção de Frutos de Acerola em Imagens

O trabalho de Leite (2020) tem como objetivo localizar frutos de acerola em imagens digitais. Foi criada uma base de dados que continha imagens de acerolas em seu ambiente natural. Para verificar a viabilidade de reconhecimento desses frutos foi utilizada uma adaptação de uma R-CNN, chamada *Mask-RCNN*, que utiliza uma *Faster-RCNN* para detecção de objetos.

A *Faster-RCNN* substitui a procura seletiva da R-CNN por uma Rede de Proposta de Região ou *Region Proposal Network* (RPN), que recebe uma imagem de entrada e usa uma CNN para extração de características e a saída será um conjunto regiões propostas (REN *et al.*, 2015). Originalmente, após a detecção dos objetos, a *Mask-RCNN* segmenta os objetos da imagem para realização da classificação, Leite (2020) modificou o algoritmo para que não houvesse a etapa de segmentação. Foi obtido uma alta acurácia na detecção de objetos, mesmo a base de dados sendo considerada pequena.

3.5 Comparativo de atividades do estado da arte

A Tabela 1 apresenta uma comparação entre as principais atividades realizadas pelos trabalhos relacionados e o presente trabalho.

Tabela 1 – Tabela comparativa de trabalhos relacionados

Lista de Atividades	Trabalhos Comparados				
	Bargoti e Underwood (2017)	Jana (2017)	Muresan e Oltean (2018)	Leite (2020)	TCC Nathalia (2020)
Classificação de frutos de acerolas usando um modelo R-CNN					X
Detecção de frutos de acerolas				X	X
Segmentação de imagens de frutas em seu ambiente natural	X	X		X	X
Detecção de frutas usando deep learning	X		X	X	X

Fonte: Elaborado pelo Autor (2020).

4 PROCEDIMENTOS METODOLÓGICOS

Nesta seção serão apresentados os procedimentos utilizados durante a pesquisa, que foram divididos em Estudo Teórico e Estudo Prático, 4.1, 4.2, respectivamente.

4.1 Estudo Teórico

Durante a etapa de estudo teórico foi realizada uma pesquisa exploratória através de revisões bibliográficas, em artigos, dissertações e livros, sobre visão computacional e processamento de imagens digitais. Posteriormente, foram avaliados quais algoritmos já utilizados no reconhecimento e classificação de frutos obtiveram as melhores taxas de acertos, considerando as métricas mais usuais para analisar a qualidade de classificadores inteligentes. Desta forma, definiu-se a utilização da rede neural R-CNN para geração do modelo inteligente proposto neste trabalho. A decisão foi fundamentada pela escolha do algoritmo considerado mais eficiente para resolução dos problemas de classificação de frutos apresentados na Seção 2.3.4.

4.2 Estudo Prático

Durante a etapa de estudo prático definiu-se as configurações e padronizações a serem utilizadas para geração do modelo de reconhecimento e classificação dos frutos de acerola em imagens RGB. Os modelos gerados utilizando a rede R-CNN foram criados utilizando a implementação de Thakur (2019), disponível em <<https://github.com/1297rohit/RCNN.git>>. O algoritmo foi implementado na linguagem de programação *Python* utilizando as bibliotecas de código aberto *TensorFlow*, *Keras*, e outras bibliotecas como *OpenCV* e *Numpy*. Para treinamento e testes dos modelos utilizou-se a plataforma *Google Colab*, um serviço em nuvem fornecido pelo *Google* que permite a execução de códigos *Python* através de um navegador web. A sua versão gratuita possui a seguinte configuração: processador da Intel(R) Xeon(R), com frequência base de 2.30GHz e um núcleo, 12.7 GB de memória RAM e GPU NVIDIA modelo Tesla K80 com memória de 12GB.

4.2.1 Aquisição e Segmentação de Imagens

A base de dados utilizada foi criada por Leite (2020), e contém 501 imagens do fruto de acerola *Malpighia emarginata DC*, que foram obtidas com aparelho celular com

câmera de 8 *megapixels* durante um estudo de campo na Fazenda Meripobo Agropecuária e por meio de pesquisa na plataforma *Google* Imagens. As imagens da base passaram por um pré-processamento, onde foram divididas em recortes menores e padronizadas no tamanho 512 *pixels* de altura e 512 *pixels* de largura. A Figura 6 ilustra exemplos de imagens da base.

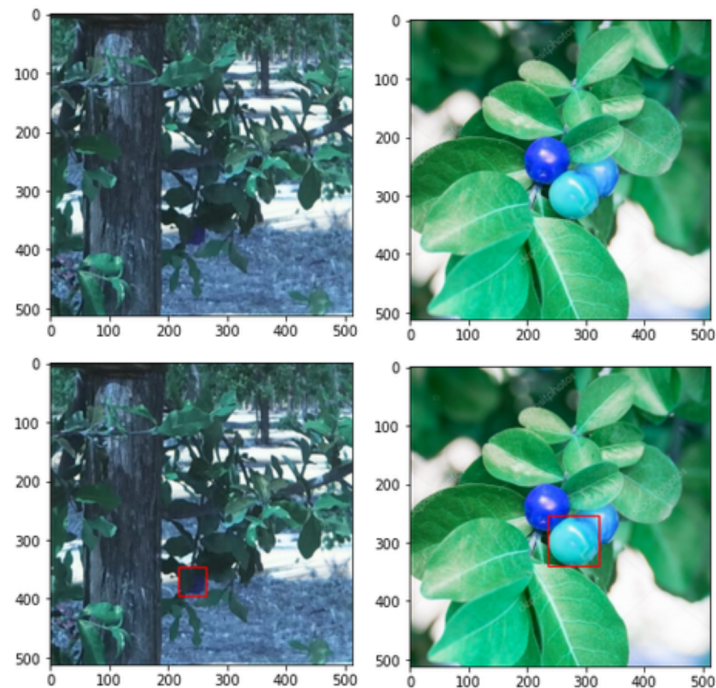
Figura 6 – Representação do conjunto de imagens criado por Leite (2020)



Fonte: Base de dados.

Utilizando a segmentação manual feita por Leite (2020), foram adaptadas as entradas para o modelo de (THAKUR, 2019). Para isso, foi necessário separar as imagens e segmentações classificando-as em acerolas de coloração avermelhada e verde. Para criação de categorias de cores, caso houvessem múltiplas acerolas das duas colorações em uma mesma imagem, foi necessário duplicar a imagem e marcar as acerolas separadamente, resultando em duas segmentações distintas, o que levou à ampliação da base de dados de 501 para 638 imagens. Por questões de restrições de hardware, para este trabalho, foram utilizadas apenas 240 imagens da base de dados, divididas em 120 de acerolas verdes e 120 de acerolas vermelhas. A Figura 7 ilustra a segmentação adaptada.

Figura 7 – Representação da segmentação adaptada da base de Leite (2020)



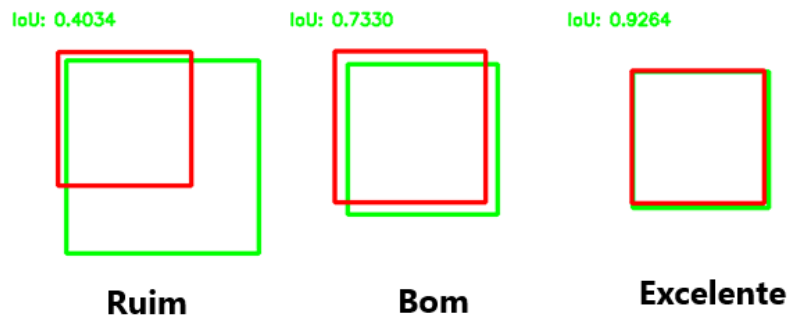
Fonte: Elaborado pelo Autor (2020).

4.2.2 Geração do modelo de detecção e classificação

Como mencionado anteriormente o modelo inteligente foi gerado utilizando a rede R-CNN de Girshick *et al.* (2013) adaptada, conforme a implementação de Thakur (2019). Para tal, efetuou-se uma procura seletiva na base padronizadas com 240 imagens, onde foram extraídas 2.000 regiões propostas de cada uma. Durante toda a fase de experimental foram definidas 3 classes de classificação, rotuladas como: 0 - plano de fundo, 1 - acerolas verdes e 2 - acerolas vermelhas.

Em seguida, as regiões propostas foram avaliadas através da função de cálculo de Intersecção sobre União ou *Intersection Over Union* (IoU), também conhecido como índice de *Jaccard*, que compara a localização da segmentação fornecida como entrada e a localização da segmentação feita com a procura seletiva (REZATOFIGHI *et al.*, 2019). Assim como Thakur (2019), foi considerado como uma boa predição todas as seleções com IoU acima de 0.7. A Figura 8 representa possíveis avaliações realizadas utilizando a métrica.

Figura 8 – Representação da avaliação do IoU



Fonte: Adaptado de Cowton *et al.* (2019).

Após a análise, as regiões com IoU considerados bons foram separadas em três conjuntos: plano de fundo, acerolas verdes e acerolas vermelhas. Cada amostra contém no máximo 30 regiões, formando assim o *dataset* completo utilizado neste trabalho. Na etapa de treinamento do modelo, foi utilizada a técnica *transfer learning*, com o algoritmo VGG16. Para a transferência de aprendizado, o algoritmo modifica os pesos das camadas de maneira similar ao aprendizado por reforço (FERREIRA *et al.*, 2016).

O algoritmo VGG16 foi utilizado para geração de um modelo de CNN pré-treinado, onde não há necessidade de aprendizado. Então, o valor dos pesos foi congelado para que o modelo obtivesse um aprendizado satisfatório mesmo com uma base de dados pequena. É importante mencionar que a definição de um *dataset* pequeno, se deve a quesitos de restrição de *hardware*.

A transferência de aprendizado do modelo pré-treinado foi realizada utilizando informações da base de dados *imageNet*. Como o modelo foi pré-treinado, a priori os pesos da rede já estão ajustados, e portanto após seus congelamentos o modelo realiza apenas pequenos ajustes, baseados no treinamento realizado posteriormente (PONTI; COSTA, 2018). Para o modelo proposto neste trabalho, foram congeladas 15 das 16 camadas geradas. Como existem 3 classes para serem previstas, foram adicionadas ao final 3 camadas de ativação densa *softmax*. O modelo foi compilado com o otimizador da biblioteca *Keras*, *Adam* com a taxa de aprendizado de 0,001. Por fim, foram definidos 10 lotes de amostras e 30 épocas, ou iterações, de treinamento.

Com o intuito de testar várias possíveis configurações de treinamento e teste, buscando compreender o efeito que os tamanhos das bases de imagens da etapa de treino teriam sobre a acurácia do modelo, foram gerados 5 experimentos computacionais distintos, cada uma com uma partição da base de dados de treino e teste diferente. Na primeira execução a base foi dividida em 50% treino e 50% teste, em seguida 60% treino e 40% teste, 70% treino e 30% teste, 80% treino e 20% teste e por último 90% treino e 10% teste.

5 RESULTADOS

Esta seção apresenta os resultados do experimento prático, onde foi aplicado o método proposto no Capítulo 4. Para cada ciclo de treinamento e teste, será apresentado seu desempenho e as métricas usadas para medi-lo.

5.1 Experimento prático

Os resultados apresentados são baseados nas matrizes de confusão gerada pelo modelo para cada combinação de particionamento de treinamento e teste. As matrizes de confusão representam as classes reais, as classes preditas e o número de erros e acertos para cada previsão feita pelo modelo. Para cada classe, os acertos se localizam na diagonal principal da matriz e os outros elementos representam erros nas previsões (MONARD; BARANAUSKAS, 2003b). A Figura 9 ilustra a representação de uma matriz de confusão. Para o modelo gerado, a classe 0 representa o fundo da imagem, e as classes 1 e 2 representam os frutos de acerola verde e vermelho, respectivamente. O treinamento e teste do modelo foi realizado com uma base dados de 240 imagens, resultando em 8.841 amostras para aprendizagem.

Figura 9 – Matriz de confusão de uma classificação

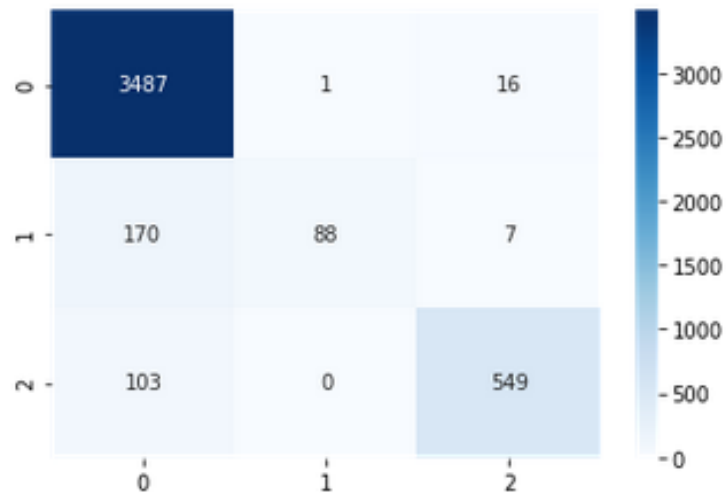
Classe	predita C_1	predita C_2	...	predita C_k
verdadeira C_1	$M(C_1, C_1)$	$M(C_1, C_2)$...	$M(C_1, C_k)$
verdadeira C_2	$M(C_2, C_1)$	$M(C_2, C_2)$...	$M(C_2, C_k)$
⋮	⋮	⋮	⋮	⋮
verdadeira C_k	$M(C_k, C_1)$	$M(C_k, C_2)$...	$M(C_k, C_k)$

Fonte: (MONARD; BARANAUSKAS, 2003b).

O primeiro ciclo de treinamento foi realizado com uma partição de 4.420 amostras, 50% do *dataset* final, para teste e os outros 50% do *dataset* para treino. A Figura 10 representa a matriz de confusão desse modelo. Analisando cada classe, podemos inferir que o modelo previu 3.487 amostras de plano de fundo corretamente, 17 amostras de plano de fundo como frutos e 273 de frutos de acerola como plano de fundo. Em relação aos frutos de cor verde, foram previstas 88 amostras corretamente, 7 amostras de acerola verde como vermelha e 1 amostra de plano de fundo como acerola verde. Para a classe de frutos de cor vermelha, o modelo previu 549 amostras corretamente, 103 amostras de acerola como plano de fundo e 23 amostras como

acerola vermelha incorretamente. Esse modelo atingiu acurácia, que é a porcentagem de acertos do modelo, de 93,28% e *f1-score* de 0.96 para plano de fundo, 0.50 para frutos verdes e 0.90 para frutos vermelhos. O *f1-score* é o valor que combina a precisão, quantidade de vezes que as classes foram previstas corretamente, e *recall*, quantidade de itens previstas como pertencentes a uma classe comparada a real quantidade de itens pertencentes as classes do modelo. A Figura 11 ilustra os valores atingidos pelo modelo.

Figura 10 – Matriz de confusão do modelo 50-50



Fonte: Resultado do modelo.

Figura 11 – Relatório do modelo 50-50

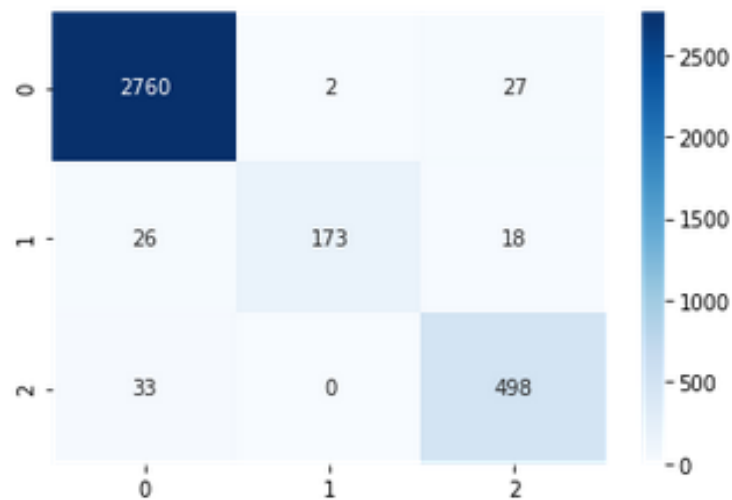
Classification Report				
	precision	recall	f1-score	support
Fundo	0.93	1.00	0.96	3504
Verde	0.99	0.33	0.50	265
Vermelho	0.96	0.84	0.90	652
accuracy			0.93	4421
macro avg	0.96	0.72	0.78	4421
weighted avg	0.94	0.93	0.92	4421

Fonte: Resultado do modelo.

O segundo ciclo de treinamento foi feito com uma partição de 5.304 amostras, 60% do *dataset* final, para teste e os outros 40% do *dataset* para treino, ou seja, 3.537 amostras. A Figura 12 representa a matriz de confusão desse modelo, foram previstas 2.760 amostras de plano de fundo corretamente, 29 amostras de plano de fundo como acerolas e 59 amostras de acerolas

como plano de fundo. Para os frutos de cor verde, foram previstas 173 amostras corretamente, 44 amostras de acerola classificadas incorretamente e 2 amostras de plano de fundo como acerola verde. Para os frutos de cor vermelha, o modelo previu 498 amostras corretamente, 33 amostras de acerola como plano de fundo e 45 amostras como acerola vermelha incorretamente. Esse modelo atingiu acurácia de 97,00% e *f1-score* de 0.98 para plano de fundo, 0.88 para frutos verdes e 0.93 para frutos vermelhos, ilustrados na Figura 13.

Figura 12 – Matriz de confusão do modelo 60-40



Fonte: Resultado do modelo.

Figura 13 – Relatório do modelo 60-40

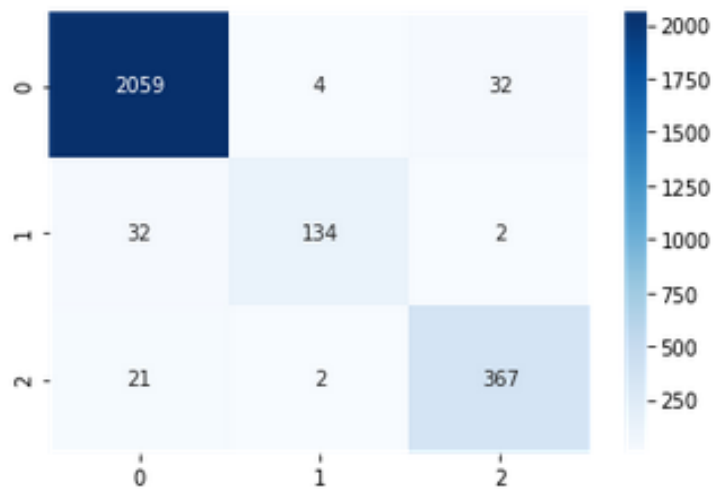
Classification Report				
	precision	recall	f1-score	support
Fundo	0.98	0.99	0.98	2789
Verde	0.99	0.80	0.88	217
Vermelho	0.92	0.94	0.93	531
accuracy			0.97	3537
macro avg	0.96	0.91	0.93	3537
weighted avg	0.97	0.97	0.97	3537

Fonte: Resultado do modelo.

O terceiro ciclo de treinamento foi feito com uma partição de 6.188 amostras, 70% do *dataset* final, para teste e os outros 30% do *dataset* para treino, ou seja, 2.653 amostras. A Figura 14 representa a matriz de confusão desse modelo, foram previstas 2.059 amostras de plano de fundo corretamente, 36 amostras de plano de fundo como acerolas e 53 amostras

de acerolas previstas como plano de fundo. Para os frutos de cor verde, foram previstas 134 amostras corretamente, 34 amostras de acerola classificadas incorretamente, 4 amostras de plano de fundo como acerola verde e 2 amostras de frutos vermelhos como verdes. Para os frutos de cor vermelha, foram previstas 367 amostras corretamente, 23 amostras de acerolas como plano de fundo, 32 amostras de plano de fundo e 2 amostras de acerola verde como acerola vermelha incorretamente. Esse modelo atingiu acurácia de 96,49% e *f1-score* de 0.98 para plano de fundo, 0.87 para frutos verdes e 0.93 para frutos vermelhos, ilustrados na Figura 15.

Figura 14 – Matriz de confusão do modelo 70-30



Fonte: Resultado do modelo.

Figura 15 – Relatório do modelo 70-30

```

Classification Report
              precision    recall  f1-score   support

   Fundo         0.97         0.98         0.98         2095
   Verde         0.96         0.80         0.87          168
   Vermelho      0.92         0.94         0.93          390

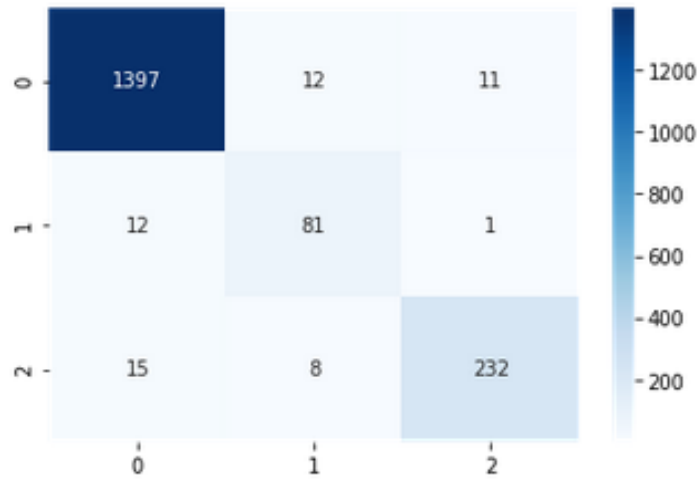
 accuracy              0.96         2653
 macro avg           0.95         0.91         0.93         2653
 weighted avg       0.97         0.96         0.96         2653
    
```

Fonte: Resultado do modelo.

O quarto ciclo de treinamento foi feito com uma partição de 7.072 amostras, 80% do *dataset* final, para teste e os outros 20% do *dataset* para treino, ou seja, 1.769 amostras. A Figura 16 representa a matriz de confusão desse modelo, foram previstas 1.397 amostras de plano de fundo corretamente, 23 amostras de plano de fundo como acerolas e 37 amostras de

acerolas previstas como plano de fundo. Para os frutos de cor verde, foram previstas 81 amostras corretamente, 13 amostras de acerola classificadas incorretamente, 12 amostras de plano de fundo como acerola verde e 8 amostras de frutos vermelhos como verdes. Para os frutos de cor vermelha, foram previstas 232 amostras corretamente, 15 amostras de acerolas como plano de fundo, 8 amostras de acerola vermelha como acerola verde, 11 amostras de plano de fundo e 1 amostra de acerola verde como acerola vermelha incorretamente. Esse modelo atingiu acurácia de 96,66% e *f1-score* de 0.98 para plano de fundo, 0.83 para frutos verdes e 0.93 para frutos vermelhos, ilustrados na Figura 17.

Figura 16 – Matriz de confusão do modelo 80-20



Fonte: Resultado do modelo.

Figura 17 – Relatório do modelo 80-20

```

Classification Report
              precision    recall  f1-score   support

   Fundo         0.98         0.98         0.98         1420
   Verde         0.80         0.86         0.83           94
   Vermelho      0.95         0.91         0.93          255

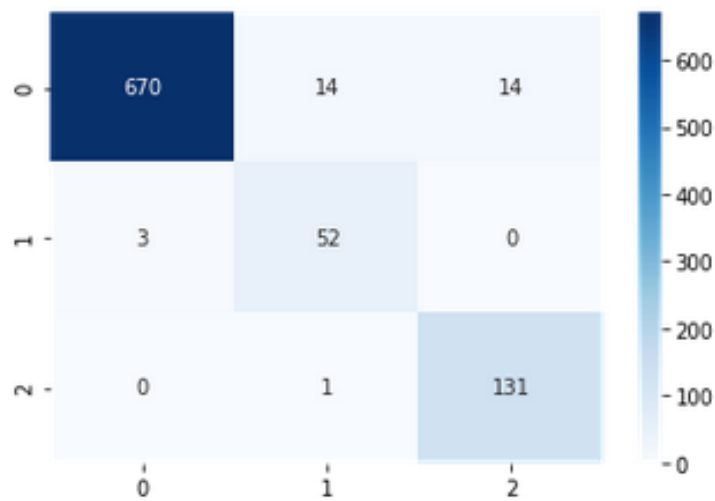
 accuracy              0.97         1769
 macro avg           0.91         0.92         0.91         1769
 weighted avg        0.97         0.97         0.97         1769
    
```

Fonte: Resultado do modelo.

Por fim, o último ciclo de treinamento foi feito com uma partição de 7.956 amostras, 90% do *dataset* final, para teste e os outros 10% do *dataset* para treino, ou seja, 885 amostras. A Figura 18 representa a matriz de confusão desse modelo, foram previstas 670 amostras de

plano de fundo corretamente, 28 amostras de plano de fundo como acerolas e 3 amostras de acerolas previstas como plano de fundo. Para os frutos de cor verde, foram previstas 52 amostras corretamente, 3 amostras de acerola classificadas como plano de fundo e 15 amostras de acerola verde classificadas incorretamente. Para os frutos de cor vermelha, foram previstas 131 amostras corretamente, 14 amostras de plano de fundo como acerola, e 1 amostra de acerola verde como vermelha. Esse modelo atingiu acurácia de 96,38% e *f1-score* de 0.98 para plano de fundo, 0.85 para frutos verdes e 0.95 para frutos vermelhos, ilustrados na Figura 19.

Figura 18 – Matriz de confusão do modelo 90-10



Fonte: Resultado do modelo.

Figura 19 – Relatório do modelo 90-10

```

Classification Report
precision    recall  f1-score   support

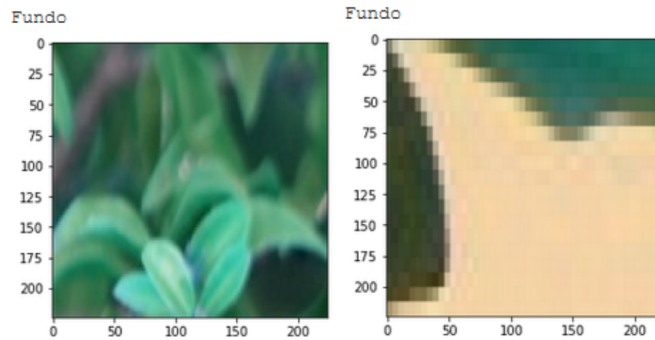
   Fundo      1.00    0.96    0.98     698
   Verde      0.78    0.95    0.85     55
 Vermelho     0.90    0.99    0.95    132

 accuracy          0.96     885
 macro avg         0.89     885
 weighted avg     0.97     885
    
```

Fonte: Resultado do modelo.

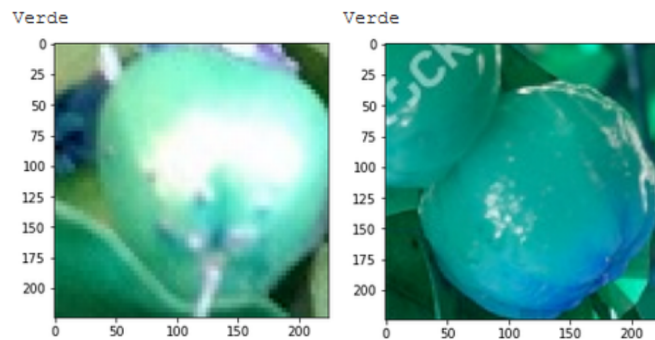
As Figuras 20, 21 e 22 exemplificam previsões feitas corretamente de regiões de plano de fund, frutos verdes e frutos vermelhos, respectivamente.

Figura 20 – Previsões de plano de fundo



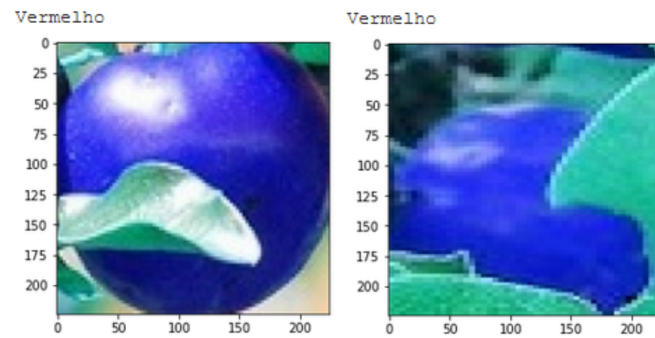
Fonte: Resultado do modelo.

Figura 21 – Previsões de frutos verdes



Fonte: Resultado do modelo.

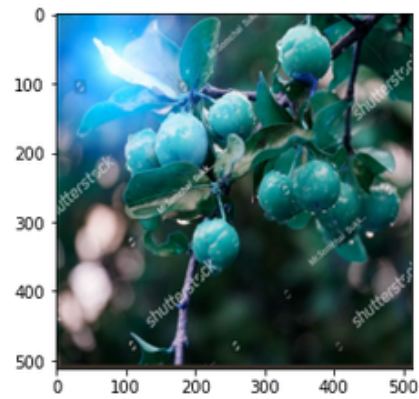
Figura 22 – Previsões de frutos vermelhos



Fonte: Resultado do modelo.

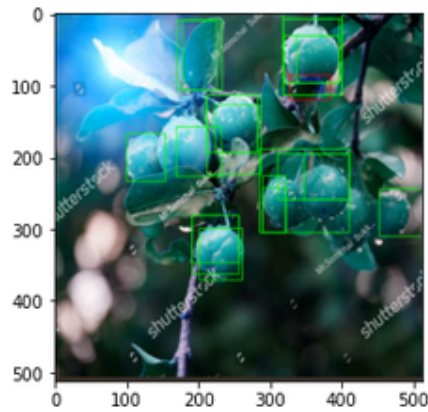
Após treinado e testado, o modelo é capaz de criar regiões de interesse baseado no que foi aprendido. A Figura 23 ilustra que não foram encontradas regiões de interesse na imagem usando o modelo 50-50. A Figura 24 ilustra as regiões encontradas com o modelo 60-40. A Figura 25 ilustra as regiões encontradas com o modelo 70-30. A Figura 26 ilustra as regiões encontradas com o modelo 80-20 e a Figura 27 ilustra as regiões encontradas com o modelo 90-10.

Figura 23 – Região prevista pelo modelo 50-50



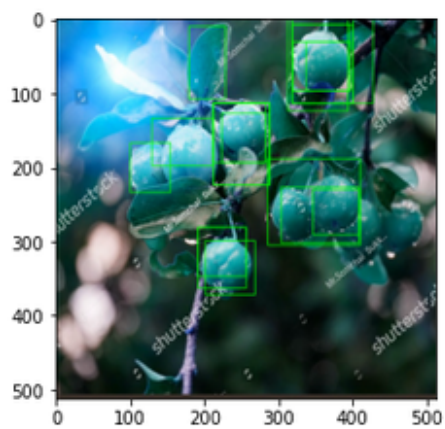
Fonte: Resultado do modelo.

Figura 24 – Região prevista pelo modelo 60-40



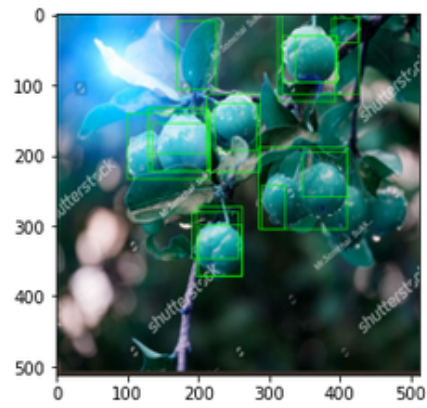
Fonte: Resultado do modelo.

Figura 25 – Região prevista pelo modelo 70-30



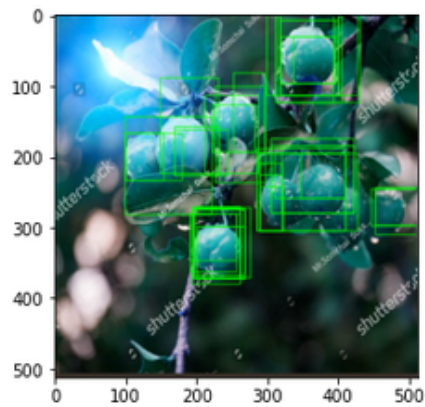
Fonte: Resultado do modelo.

Figura 26 – Região prevista pelo modelo 80-20



Fonte: Resultado do modelo.

Figura 27 – Região prevista pelo modelo 90-10



Fonte: Resultado do modelo.

6 CONCLUSÃO

Esta Seção apresenta as considerações e lições aprendidas sobre a metodologia para classificação de frutos de acerola, criada neste trabalho, assim como os resultados obtidos.

6.1 Considerações gerais

Neste trabalho foi elaborada uma metodologia para classificação de frutos de acerola em imagens RGB, retiradas em seu ambiente natural, usando a base criada por Leite (2020). O modelo foi gerado utilizando uma rede R-CNN, como descrita na Seção 2.3.4.1, adaptada da implementação de Thakur (2019), fazendo uso de técnicas de *augmentation* para aumentar as classes de treino e teste do modelo.

Devido às restrições de *hardware*, não foi possível utilizar a base de imagens completa, sendo utilizadas 240 imagens das 638 disponíveis. Portanto, foi comparado qual o melhor conjunto de teste e treino para uma base reduzida. O modelo que utilizou 60% do *dataset* final para treinamento e 40% para teste, retornou o melhor resultado de acurácia, obtendo o valor de 97%, e o melhor resultado *f1-score* para frutos verdes. A classe de frutos verdes pode ser considerada a mais complexa para classificações pois, a coloração dos frutos se assemelha à coloração do plano de fundo.

Como conclusões finais, pode-se afirmar que um dos principais desafios na realização desse trabalho foi a necessidade de um *hardware* de capacidade computacional que suportasse a complexidade da base de dados. É importante ressaltar que a escolha da melhor configuração de teste e treino influencia diretamente nos resultados obtidos, que será importante em atualizações da pesquisa.

6.2 Trabalhos futuros

Espera-se em trabalhos futuros aplicar a metodologia com a base completa em uma configuração de *hardware* que consiga suportar o esforço computacional que o modelo necessita. Por fim, espera-se também em trabalhos futuros desenvolver uma aplicação de análise de plantação para estimativa de colheita que utilize modelos de classificação de frutos semelhantes ao desenvolvido neste trabalho.

REFERÊNCIAS

- ARIVAZHAGAN, S.; SHEBIAH, R. N.; NIDHYANANDHAN, S. S.; GANESAN, L. Fruit recognition using color and texture features. **Journal of Emerging Trends in Computing and Information Sciences**, Citeseer, v. 1, n. 2, p. 90–94, 2010.
- ASSIS, S. A. de; FERNANDES, F. P.; MARTINS, A. B. G.; OLIVEIRA, O. M. M. de F. Acerola: importance, culture conditions, production and biochemical aspects. **Fruits**, EDP Sciences, v. 63, n. 2, p. 93–101, 2008.
- AYODELE, T. O. Types of machine learning algorithms. **New advances in machine learning**, InTech, v. 3, p. 19–48, 2010.
- BARGOTI, S.; UNDERWOOD, J. P. Image segmentation for fruit detection and yield estimation in apple orchards. **Journal of Field Robotics**, Wiley Online Library, v. 34, n. 6, p. 1039–1060, 2017.
- BHANGE, M.; HINGOLIWALA, H. Smart farming: Pomegranate disease detection using image processing. **Procedia Computer Science**, Elsevier, v. 58, p. 280–288, 2015.
- COWTON, J.; KYRIAZAKIS, I.; BACARDIT, J. Automated individual pig localisation, tracking and behaviour metric extraction using deep learning. **IEEE Access**, IEEE, v. 7, p. 108049–108060, 2019.
- DATA SCIENCE ACADEMY. **O que é visão computacional?** 2018. Disponível em: <<http://datascienceacademy.com.br/blog/o-que-e-visao-computacional/>>. Acesso em: 28 ago. 2020.
- DIARIO DO NORDESTE. **Fruticultura cearense é destaque em todo o País.** 2014. Disponível em: <<https://diariodonordeste.verdesmares.com.br/negocios/fruticultura-cearense-e-destaque-em-todo-o-pais-1.972429>>. Acesso em: 28 ago. 2020.
- DUBEY, S. R.; JALAL, A. Robust approach for fruit and vegetable classification. **Procedia Engineering**, Elsevier, v. 38, p. 3449–3453, 2012.
- FERREIRA, A.; FERREIRA, R. P.; SILVA, A. M. da; FERREIRA, A.; SASSI, R. J. Um estudo sobre previsão da demanda de encomendas utilizando uma rede neural artificial. **Blucher Marine Engineering Proceedings**, v. 2, n. 1, p. 353–364, 2016.
- FILHO, R. d. M. E. A study on deep convolutional neural networks for computer vision applications. Universidade Federal do Rio de Janeiro, 2017.
- GIRSHICK, R.; DONAHUE, J.; DARRELL, T.; MALIK, J. **Rich feature hierarchies for accurate object detection and semantic segmentation.** arXiv, 2013. Disponível em: <<https://arxiv.org/abs/1311.2524>>.
- GONZALEZ, R. C.; WOODS, R. E. **Processamento de imagens digitais.** [S.l.]: Editora Blucher, 2000.
- IPEA - INSTITUTO DE PESQUISA ECONOMICA APLICADA. **Demanda crescente estimula a produção orgânica no Brasil e no mundo.** 2020. Disponível em: <https://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=35326&Itemid=9>. Acesso em: 28 ago. 2020.

- JANA, S.; BASAK, S.; PAREKH, R. Automatic fruit recognition from natural images using color and texture features. In: IEEE. **2017 Devices for Integrated Circuit (DevIC)**. [S.l.], 2017. p. 620–624.
- KHAN, A.; BAHARUDIN, B.; LEE, L. H.; KHAN, K. A review of machine learning algorithms for text-documents classification. **Journal of advances in information technology**, Academy Publisher, PO Box 40 Oulu 90571 Finland, v. 1, n. 1, p. 4–20, 2010.
- LEITE, W. L. S. Análise de viabilidade do uso de aprendizagem profunda para detecção de frutos de acerola em imagens. 2020.
- MARTINS, É. d. A.; CAMPOS, R. T.; CAMPOS, K. C.; ALMEIDA, C. d. S. Rentabilidade da produção de acerola orgânica sob condição determinística e de risco: estudo do distrito de irrigação tabuleiro litorâneo do piauí. **Revista de Economia e Sociologia Rural**, SciELO Brasil, v. 54, n. 1, p. 9–28, 2016.
- MEYER, D.; WIEN, F. Support vector machines. **The Interface to libsvm in package e1071**, v. 28, p. 20, 2015.
- MILANO, D. de; HONORATO, L. B. **Visão computacional**. 2014.
- MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. **Sistemas inteligentes-Fundamentos e aplicações**, Manole Ltda, v. 1, n. 1, p. 32, 2003.
- MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. **Sistemas inteligentes-Fundamentos e aplicações**, Manole, v. 1, n. 1, p. 32, 2003.
- MUKHERJEE, S.; MUKHERJEE, S.; MUKHERJEE, D. P.; SIVASWAMY, J.; AWATE, S.; SETLUR, S.; NAMBOODIRI, A. M.; CHAUDHURY, S. *et al.* **Computer vision, graphics, and image processing**. [S.l.]: Springer., 2017.
- MUREŞAN, H.; OLTEAN, M. Fruit recognition from images using deep learning. **Acta Universitatis Sapientiae, Informatica**, Sciendo, v. 10, n. 1, p. 26–42, 2018.
- PIVOTO, D.; WAQUIL, P. D.; TALAMINI, E.; FINOCCHIO, C. P. S.; CORTE, V. F. D.; MORES, G. de V. Scientific development of smart farming technologies and their application in brazil. **Information processing in agriculture**, Elsevier, v. 5, n. 1, p. 21–32, 2018.
- PONTI, M. A.; COSTA, G. B. P. D. Como funciona o deep learning. **arXiv preprint arXiv:1806.07908**, 2018.
- REHEM, A.; TRINDADE, F. H. Técnicas de visão computacional para rastreamento de olhar em vídeos. **Publicado em**, v. 3, n. 02, 2009.
- REN, S.; HE, K.; GIRSHICK, R.; SUN, J. **Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks**. arXiv, 2015. Disponível em: <<https://arxiv.org/abs/1506.01497>>.
- REZATOFIGHI, H.; TSOI, N.; GWAK, J.; SADEGHIAN, A.; REID, I.; SAVARESE, S. Generalized intersection over union: A metric and a loss for bounding box regression. In: **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2019. p. 658–666.

SANTOS, T. C. dos; ESPERIDIÃO, T. L.; AMARANTE, M. dos S. Agricultura 4.0. **Revista Pesquisa e Ação**, v. 5, n. 4, p. 122–131, 2019.

SCHALKOFF, R. J. **Digital image processing and computer vision**. [S.l.]: Wiley New York, 1989. v. 286.

SERPA, E. **Ceará tem a maior fazenda de acerola orgânica do mundo**. 2019. Disponível em: <<https://blogdomanuelsales.com.br/2019/11/11/ceara-tem-a-maior-fazenda-de-acerola-organica-do-mundo/>>. Acesso em: 28 ago. 2020.

SOUZA, E. P. d.; GOMES, C. M.; BARROSO, D. H.; MIRANDA, V. L. d.; GURGEL-GONÇALVES, R. Aplicações do deep learning para diagnóstico de doenças e identificação de insetos vetores. **Saúde em Debate**, SciELO Public Health, v. 43, p. 147–154, 2020.

THAKUR, R. **Step-by-Step R-CNN Implementation From Scratch In Python**. 2019. Disponível em: <<https://towardsdatascience.com/step-by-step-r-cnn-implementation-from-scratch-in-python-e97101ccde55>>.

UIJLINGS, J.; SANDE, K.; GEVERS, T.; SMEULDERS, A. Selective search for object recognition. **International Journal of Computer Vision**, v. 104, p. 154–171, 09 2013.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: SN. **Proceedings of the xxix conference on graphics, patterns and images**. [S.l.], 2016. v. 1, n. 4.