



UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE INFORMÁTICA

JOÃO GABRIEL JUNQUEIRA DA SILVA

Redes Neurais Profundas para Detecção e Contagem Automática de Moscas Brancas no Contexto Agrícola

Goiânia
2021



UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE INFORMÁTICA

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

☒ Dissertação ☐ Tese

2. Nome completo do autor

João Gabriel Junqueira da Silva

3. Título do trabalho

Redes neurais profundas para detecção e contagem automática de moscas brancas no contexto agrícola

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento ☒ SIM ☐ NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

a) consulta ao(a) autor(a) e ao(a) orientador(a);

b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.

O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **JOÃO GABRIEL JUNQUEIRA DA SILVA, Discente**, em 05/05/2021, às 10:18, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Cristhiane Gonçalves, Usuário Externo**, em 05/05/2021, às 10:38, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2047056** e o código CRC **5B855EE3**.

JOÃO GABRIEL JUNQUEIRA DA SILVA

Redes Neurais Profundas para Detecção e Contagem Automática de Moscas Brancas no Contexto Agrícola

Dissertação apresentada ao Programa de Pós-Graduação do Instituto de Informática da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Ciência da Computação.

Orientadora: Profa. Cristhiane Gonçalves

Goiânia
2021

Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Silva, João Gabriel Junqueira da
Redes neurais profundas para detecção e contagem automática de
moscas brancas no contexto agrícola [manuscrito] / João Gabriel
Junqueira da Silva. - 2021.
LXXI, 71 f.

Orientador: Profa. Dra. Cristhiane Gonçalves.
Dissertação (Mestrado) - Universidade Federal de Goiás, Instituto
de Informática (INF), Programa de Pós-Graduação em Ciência da
Computação, Goiânia, 2021.

Bibliografia. Apêndice.

Inclui fotografias, símbolos, gráfico, tabelas, lista de figuras, lista
de tabelas.

1. Redes neurais profundas. 2. Detecção de objetos. 3. Visão
computacional. 4. Agricultura de precisão. 5. Mosca-branca. I.
Gonçalves, Cristhiane, orient. II. Título.

CDU 004



UNIVERSIDADE FEDERAL DE GOIÁS

INSTITUTO DE INFORMÁTICA

ATA DE DEFESA DE DISSERTAÇÃO

Ata nº **09/2021** da sessão de Defesa de Dissertação de **João Gabriel Junqueira da Silva**, que confere o título de Mestre em Ciência da Computação, na área de concentração em Ciência da Computação.

Ao primeiro dia do mês de abril de dois mil e vinte e um, a partir das nove horas, via sistema de webconferência da RNP, realizou-se a sessão pública de Defesa de Dissertação intitulada “**Redes neurais profundas para detecção e contagem automática de moscas brancas no contexto agrícola**”. Os trabalhos foram instalados pela Orientadora, Professora Doutora Cristhiane Gonçalves (UTFPR) com a participação dos demais membros da Banca Examinadora: Doutor Marcelo Gonçalves Narciso (Embrapa), membro titular externo; Professor Doutor Anderson da Silva Soares (INF/UFG), membro titular interno. A realização da banca ocorreu por meio de videoconferência, em atendimento à recomendação de suspensão das atividades presenciais na UFG emitida pelo Comitê UFG para o Gerenciamento da Crise COVID-19, bem como à recomendação de isolamento social da Organização Mundial de Saúde e do Ministério da Saúde para enfrentamento da emergência de saúde pública decorrente do novo coronavírus. Durante a arguição os membros da banca não fizeram sugestão de alteração do título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Dissertação, tendo sido o candidato **aprovado** pelos seus membros. Proclamados os resultados pela Professora Doutora Cristhiane Gonçalves, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, ao primeiro dia do mês de abril de dois mil e vinte e um.

TÍTULO SUGERIDO PELA BANCA

Documento assinado eletronicamente por **Anderson Da Silva Soares, Professor do Magistério Superior**, em 01/04/2021, às 10:59, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Cristhiane Gonçalves, Usuário Externo**, em 01/04/2021, às 11:08, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **JOÃO GABRIEL JUNQUEIRA DA SILVA, Discente**, em 01/04/2021, às 11:59, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marcelo Gonçalves Narciso, Usuário Externo**, em 01/04/2021, às 12:16, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1938527** e o código CRC **99577B45**.

Referência: Processo nº 23070.011964/2021-98

SEI nº 1938527

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador(a).

João Gabriel Junqueira da Silva

Bacharel em Engenharia de Software pela Universidade Federal de Goiás - (UFG, 2018). Durante sua graduação, foi monitor no Instituto de Informática da UFG e atuante em projetos de desenvolvimento de software para o setor privado. Possui experiência em desenvolvimento de sistemas web e mobile, prestando serviços para empresas de porte nacional e internacional, além de empreender no setor de software sob demanda e plataformas de gamificação para recrutamento e seleção. Atualmente trabalha como pesquisador e desenvolvedor em projetos de inteligência artificial e redes neurais para visão computacional e processamento de linguagem natural no Centro de Excelência em Inteligência Artificial – CEIA.

Dedico este trabalho aos meus pais, irmão e a minha companheira, por acreditarem no sucesso de todos os meus projetos. Dedico esse trabalho a minha orientadora, professora Cristhiane Gonçalves, pela confiança e empenho em auxiliar na construção desse trabalho. Dedico esse trabalho a todas as pessoas e instituições que fomentaram este projeto. Não haveria êxito sem o apoio de todos vocês.

Agradecimentos

Agradeço primeiramente aos meus pais, irmão e demais familiares que mesmo distantes geograficamente, tanto me deram apoio nas decisões que eu tomei ao longo de minha vida. Agradeço à minha noiva e futura companheira Letícia Viana, por ser meu apoio diário e me incentivar e acompanhar nos momentos mais difíceis. Agradeço à minha orientadora Cristhiane Gonçalves, pela oportunidade, confiança e por ter ajudado no conduzir desse projeto, tutorando e me abrindo portas academicamente e profissionalmente. Agradeço ao Instituto de Informática da Universidade Federal de Goiás por dar apoio a nível de infra-estrutura e ensino, possibilitando não só a existência desse projeto, bem como provendo toda a base necessária desde a minha graduação, sinto-me acolhido nesta universidade que tenho como um segundo lar. Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de auxílio ao mestrado. Agradeço aos colegas do mestrado pelos ensinamentos compartilhados e estudos essenciais para provas e trabalhos. Agradeço a equipe da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), em especial a unidade Embrapa Arroz e Feijão, localizada em Santo Antônio de Goiás - GO e em especial o pesquisador Marcelo Narciso Gonçalves, pela disponibilidade dos dados fornecidos, aconselhamentos essenciais para o desenvolvimento deste projeto e a oportunidade de aplicar os conhecimentos adquiridos em problemas atuais da realidade, acredito na longevidade desta parceria em demais projetos profissionais e acadêmicos. Agradeço também ao Centro de Excelência em Inteligência Artificial (CEIA) por me propiciar crescimento profissional e novos desafios que sucederão a pós-graduação. Agradeço os meus colegas de mestrado e tantos outros nomes que foram importantes nessa trajetória de vida. Meu muito obrigado a todos vocês.

"The hardest thing in life is to get knocked down and to get back up constantly, but we do it because we love it and we know that deep down, if it's the right idea, it could be bigger than all of us."

Gordon Clark,
Halt and Catch Fire, 2015.

Resumo

Junqueira, João Gabriel. **Redes Neurais Profundas para Detecção e Contagem Automática de Moscas Brancas no Contexto Agrícola**. Goiânia, 2021. 70p. Dissertação de Mestrado. Instituto de Informática, Universidade Federal de Goiás.

A evolução da agricultura de precisão impulsionou o aumento da produção agrícola e proporcionou o surgimento de grandes desafios em termos de eficiência, segurança alimentar, sustentabilidade e impacto ambiental. Especificamente, o controle de pragas é de suma importância no que tange a saúde das plantações. Este trabalho tem como intuito propor uma arquitetura de rede neural para detecção e contagem da praga mosca branca em diferentes estágios ninfais e da sua vida adulta. Estão sendo considerados dados em ambientes controlados através de imagens microscópicas e ambientes abertos no campo com o auxílio de armadilhas adesivas. Com as etapas de revisão sistemática, aquisição de dados e análises comparativas de diferentes arquiteturas de redes neurais convolucionais finalizadas, os resultados da pesquisa indicam a viabilidade de uma aplicação de software e concepção de uma armadilha inteligente para ser utilizada no campo.

Palavras-chave

Visão computacional, Redes Neurais Profundas, Agricultura de Precisão, Mosca-branca.

Abstract

Junqueira, João Gabriel. **Deep Neural Networks for Automatic Detection and Counting of Whiteflies in Agricultural Context**. Goiânia, 2021. 70p. MSc. Dissertation. Instituto de Informática, Universidade Federal de Goiás.

The evolution of precision agriculture increased agricultural production and provided major challenges in terms of efficiency, food safety, sustainability and environmental impact. Specifically, the pest management is extremely important in healthy crop. This work proposes a neural network architecture for detecting and counting the whitefly pest at different stages of nymphs and its adult life. Data are being considered in controlled environments through microscopic images and open environments in the field with sticky traps. Systematic review, data acquisition and analysis steps are completed and results obtained so far point to a viability of a solution delivered in a software application and an intelligent trap to be used in the field.

Keywords

Computer Vision, Deep Neural Networks, Precision Agriculture, Whiteflies.

Sumário

Lista de Figuras	15
Lista de Tabelas	17
1 Introdução	18
1.1 Contexto	20
1.2 Objetivos	21
2 Revisão da literatura	22
2.1 Questões de pesquisa	22
2.2 Palavras-chave, sinônimos e estratégias de busca	23
2.3 Critérios de Inclusão/Exclusão	23
2.4 Condução da revisão	23
2.5 Estratégias, métodos e algoritmos	24
2.5.1 Imagens em armadilhas adesivas	25
2.5.2 Imagens em folhas de plantas	25
3 Conceitos básicos sobre detecção de objetos	27
3.1 Detecção de objetos	27
3.2 Redes neurais convolucionais	28
3.3 Limitações em pequenos objetos	29
4 Materiais e métodos	32
4.1 Materiais	32
4.2 Imagens de plantas (Embrapa)	32
4.3 Imagens de armadilhas adesivas (4TU Dataset)	34
4.4 Pré-processamento dos dados	35
4.5 As arquiteturas	36
4.5.1 You Only Look Once	36
4.6 YOLOv3-v5	37
4.6.1 Single Shot MultiBox Detector (SSD)	37
4.7 Equipamentos e programas utilizados	38
4.8 Métricas de avaliação	39
5 Resultados	42
5.1 Treinamentos	42
5.2 Análise dos testes	50
6 Conclusão e trabalhos futuros	52

Referências Bibliográficas	54
A Publicações	60

Lista de Figuras

1.1	Planta infestada com moscas-brancas	19
1.2	Sintomas de infestação pelo vírus do mosaico dourado	20
3.1	Classificação, classificação + localização e detecção de objetos respectivamente.	27
3.2	Processo de classificação de imagem através de uma CNN	28
3.3	Exemplo de camada <i>max pooling</i> para um modelo de redes neurais convolucionais	29
3.4	Resultados de acurácia em detecção de pequenos objetos para arquiteturas recentes de redes neurais profundas	30
4.1	Imagens obtidas com auxílio do Estereoscópio	33
4.2	Conjunto de imagens de ninfas obtido com apoio de pesquisadores da Embrapa Arroz e Feijão	33
4.3	Imagem obtida no dataset	34
4.4	Sequência de detecção de objeto para a arquitetura YOLO	37
4.5	Arquitetura da rede <i>Single Shot MultiBox detector</i>	38
4.6	Cálculo para a Intersecção sobre a União)	40
5.1	Curva de precisão média para o treinamento da rede YOLOv5m no 4tuDataset em resolução 640x640 com <i>threshold</i> de 0.5	43
5.2	Curva de precisão média para o treinamento da rede YOLOv5m no 4tuDataset em resolução 640x640 com <i>threshold</i> de 0.95	43
5.3	Curva de precisão no treinamento da rede YOLOv5m em resolução 640x640 (imagens de armadilhas)	44
5.4	Curva de recall no treinamento da rede YOLOv5m em resolução 640x640 (imagens de armadilhas)	44
5.5	Curva de precisão média para o treinamento da rede YOLOv5m no 4tuDataset em resolução 864x864 com <i>threshold</i> de 0.5	44
5.6	Curva de precisão média para o treinamento da rede YOLOv5m no 4tuDataset em resolução 864x864 com <i>threshold</i> de 0.95	44
5.7	<i>Loss curve</i> para o treinamento da rede SSD MobileNetV2 no conjunto de imagens de armadilhas adesivas	45
5.8	<i>Loss curve</i> para o treinamento da rede SSD MobileNetV2 no conjunto de imagens de ninfas em plantas.	45
5.9	Conjunto de treinamento para a rede YOLOv5m, destacando a habilidade da rede em detectar moscas-brancas ao redor e nas bordas da armadilha adesiva, problema recorrente em trabalhos similares citados anteriormente na etapa de revisão sistemática da literatura.	46

5.10	Exemplo 1 de inferência para o modelo YOLOv5m em imagens do conjunto de testes de armadilhas adesivas em resolução 640x640.	47
5.11	Exemplo 2 de inferência para o modelo YOLOv5m em imagens do conjunto de testes de armadilhas adesivas e resolução 640x640..	47
5.12	Conjunto de treinamento da rede YOLOv5m para as imagens de ninfas em resolução 640x640.	48
5.13	Exemplo 1 de inferência para o modelo YOLOv5m em imagens do conjunto de testes de ninfas em plantas em resolução 640x640.	49
5.14	Exemplo 2 de inferência para o modelo YOLOv5m em imagens do conjunto de testes de ninfas em plantas em resolução 640x640.	49
6.1	Raspberry Pi 3B+, um computador de placa única	52

Lista de Tabelas

2.1	Número de estudos na RSL.	24
4.1	Quantidade de amostras do conjunto de dados	36
5.1	Testes para o 4TUDataset de armadilhas adesivas	50
5.2	Testes para o conjunto de dados da Embrapa de imagens de ninfas	50

Introdução

Agricultura de precisão tem se tornado um fator de suma importância ao ser considerada um expoente do desenvolvimento futuro geral através da Internet das Coisas (IoT) e Big Data [Wolfert et al. 2017]. As tecnologias modernas deram à sociedade humana a capacidade de produzir alimentos suficientes para atender à demanda de mais de 7 bilhões de pessoas [Mohanty, Hughes e Salathé 2016]. Entretanto, com o aumento da capacidade produtiva, surgiram grandes desafios em termos de eficiência, segurança alimentar, sustentabilidade e impacto ambiental. Para enfrentar esses desafios, o complexo, multivariado e imprevisível ecossistema agrícola precisa ser melhor compreendido através do monitoramento, medição e análise contínua através de aspectos físicos e fenomenológicos [Kamilaris e Prenafeta-Boldú 2018]. As técnicas de processamento de imagem podem ser usadas para melhorar as práticas agrícolas, melhorando a precisão e a consistência dos processos e reduzindo o monitoramento manual dos agricultores. Frequentemente, oferece flexibilidade e efetivamente substitui a tomada de decisão visual dos agricultores [Saxena e Armstrong 2014]. Nos últimos anos, avanços significativos no campo de inteligência artificial, visão computacional e deep learning resultaram em performances notáveis em tarefas de classificação e visão de imagens [Schwartzman et al. 2016].

A mosca-branca (*Bemisia tabaci*) é uma das principais pragas da agricultura. Os espécimes adultos são capazes de alcançar longas distâncias ao serem carregados pelo vento e colonizar culturas básicas, ervas e plantas ornamentais, além de vetorizar vírus pertencentes a vários grupos taxonômicos importantes [Bar et al. 2019]. Afeta inúmeras culturas como abóbora, abobrinha, algodão, berinjela, brócolis, couve-flor, ervilha, feijão, mamão, mandioca, melancia, melão, pepino, pimenta, repolho, soja e muitos outros. A maneira mais fácil de controlar a infecção por pragas é o uso de pesticidas. Porém, quanto mais o seu uso se intensifica, além de matar as pragas nas plantas, também afeta a saúde de humanos, animais e plantas. Para superar esse problema, é necessário controlar o uso de pesticidas, e a detecção de pragas é o processo mais importante para um cultivo eficaz [Jige e Ratnaparkhe 19-2]. A figura 1.1 apresenta uma folha infestada com moscas-brancas.

Unindo a necessidade ao surgimento de novas técnicas, este projeto tem como

intuito explorar métodos e ferramentas de visão computacional, em específico redes neurais profundas, para coletar e monitorar conjuntos de dados em plantações de arroz e feijão para sistemas que possam apoiar a análise e reconhecimentos de padrões em campos de cultivo ao que tange a detecção de pestes e pragas, especificamente para a praga Bemisia Tabaci, popularmente conhecida como "mosca-branca".

Figura 1.1: Planta infestada com moscas-brancas



Fonte (Embrapa, 2019)

[[Barbedo 2014](#)] explica que a maneira mais direta de medir a infestação de mosca-branca é identificar e contar manualmente os insetos em uma região selecionada. Em geral, essa abordagem não requer aparato sofisticado e, mais importante, depende da notável capacidade humana de resolver ambiguidades e situações pouco claras, mesmo em condições longe das ideais. Por outro lado, os seres humanos são suscetíveis a fenômenos fisiológicos e psicológicos que podem ser importantes fontes de erro: fadiga, ilusões visuais, tédio, entre outros. Além disso, os humanos geralmente são muito mais lentos que as máquinas na execução de tarefas simples, como contagem.

A escolha do uso de redes neurais como ferramenta para esse projeto deve-se aos recentes avanços em benchmarks alcançados quando comparado com abordagens utilizando algoritmos clássicos de machine learning. Como exemplo é possível citar o problema de benchmark MNIST [[LeCun e Cortes 2010](#)], um dataset clássico de 70 mil números escritos a mão, apresenta em [[Ciregan, Meier e Schmidhuber 2012](#)] uma taxa de erro de apenas 0.23%, taxa próxima de 0.2% geralmente obtida por humanos executando a mesma tarefa.

No contexto específico de visão computacional para detecção e contagem de moscas-brancas, baseada na revisão sistemática da literatura, foi avaliado que: nos últimos dez anos, foram realizados alguns estudos sobre detecção da praga mosca-branca. Em média dois artigos por ano são publicados para esse fim, porém deixando aberto o campo para exploração de modelos de redes neurais convolucionais em datasets públicos e um campo pouco explorada para aplicação de tecnologias embarcadas. Ademais, este projeto segue pesquisas iniciais realizadas pela Embrapa Arroz e Feijão, unidade vinculada a

Empresa Brasileira de Pesquisa Agropecuária, onde trabalhos anteriores utilizaram-se de técnicas de segmentação e classificadores, pavimentando o estudo para o desenvolvimento de uma ferramenta para contagem de moscas-brancas em imagens obtidas em laboratórios e a construção de uma armadilha inteligente para contagem da praga no ambiente do campo, onde a infraestrutura de acesso a rede de dados é limitada e o uso de tecnologias embarcadas se faz necessário.

1.1 Contexto

Atualmente, 80% da produção agrícola mundial advém dos pequenos produtores [Meyer et al. 2014] e relatórios com índices de 50% de perda da produção desses produtores devido a pragas e doença de plantas são comuns [Harvey et al. 2014]. Na figura 1.2 é possível identificar sintomas de infestação pelo vírus do mosaico dourado no feijoeiro, sendo a mosca-branca um de seus principais vetores de transmissão.

Figura 1.2: Sintomas de infestação pelo vírus do mosaico dourado



Fonte (Embrapa, 2019)

“As plantas de feijão ficam anãs, amarelas, como que anêmicas; as folhas, encarquilhadas. O agricultor vê o problema no campo e sabe o futuro: se tiver sorte, irá lidar com perdas em torno de 40% da lavoura, mas corre o risco de perder tudo. O consumidor não toma conhecimento do problema, mas provavelmente pagará mais caro pelo feijão de qualidade que não foi afetado e chegará ao mercado” [Wendland 2002].

Tais números, além de afetarem a agricultura, ameaçam a segurança alimentar em escala global. Problema recorrente, a falta de monitoramento das plantas impossibilita a análise de informações referentes a detecção de doenças, infestações de pragas e a frequência desses problemas. Desta forma, o monitoramento contínuo tem como foco identificar e prevenir a infestação de doenças e pragas.

Técnicas de processamento de imagens foram provadas como um sistema eficaz de visão computacional para o domínio agrícola [Vibhute e Bodhe 2012]. No entanto, no que tange o desenvolvimento de novas técnicas de processamento de imagens para a agricultura, um fator que pode aumentar é a disponibilidade de conjuntos de dados

online. Não há bancos de dados de imagens públicos disponíveis para avaliação da qualidade dos alimentos, detecção de defeitos de frutas ou classificação de ervas daninhas / culturas [Saxena e Armstrong 2014] e seria interessante ver como novas técnicas, como deep learning, se comportariam em outros problemas relacionados a agricultura como a detecção de pestes e doenças nas comidas [Kamilaris e Prenafeta-Boldú 2018].

Resumindo, ao realizar coletas para imagens de pragas, será possível, através de análises, ter um entendimento em relação a como se dá a ocorrência dessas infestações. A proposta inclui o uso de conceitos de processamento digital de imagens e novas técnicas de processamento, tais como deep learning, como plataformas de armazenamento e análise dos dados referentes a infestações para a praga mosca-branca.

1.2 Objetivos

Em suma, pode-se definir que o objetivo geral deste projeto é explorar arquiteturas de redes neurais convolucionais capazes de detectar a praga mosca-branca em diferentes contextos agrícolas, pavimentando o início de uma solução para a construção de uma armadilha inteligente a ser utilizada no campo. Em específico, este trabalho se propõe a:

- Realizar uma revisão sistema da literatura de forma a estabelecer o estado da arte para a técnicas de detecção e contagem para a mosca-branca;
- Levantar conjuntos de dados para a praga mosca-branca em diferentes cenários, sendo estes laboratoriais e no campo;
- Efetuar treinamentos com diferentes algoritmos para detecção de objetos nos conjuntos de dados adquiridos;
- Avaliar e comparar os resultados, de forma a selecionar as melhores arquiteturas para o prosseguimento de futuros trabalhos;

Revisão da literatura

Com o intuito de analisar a viabilidade para o desenvolvimento do projeto, bem como o estado da arte para detecção e contagem automática da praga mosca-branca utilizando técnicas de processamento digital de imagens e visão computacional, foi efetuada uma Revisão Sistemática da Literatura (RSL), por oferecer um processo organizado e bem definido. [Mallett et al. 2012] define a RSL como uma forma rigorosa e transparente de revisão da literatura. Três etapas principais foram tomadas seguindo as diretrizes da RSL definidas por [Kitchenham e Charters 2007]: (I) definição de um protocolo de revisão; (II) realização da revisão dos estudos; e (III) relatar a revisão. A definição do protocolo de revisão inclui a definição das questões de pesquisa, palavras-chave e estratégia de pesquisa, critérios de inclusão e exclusão para a seleção dos estudos e uma lista de verificação para avaliação da qualidade do estudo. Esses elementos são detalhados a seguir. Para apoiar esse processo e fornecer uma maneira colaborativa entre o autor e sua orientadora, foi utilizada uma plataforma de software: Parsifal (<https://parsif.al/>).

Nesta revisão foram considerados artigos publicados entre 2008 e 2018 indexados por cinco bases de dados científicas. As palavras-chaves de busca, bem como os artigos selecionados, foram considerados na língua inglesa.

2.1 Questões de pesquisa

O principal objetivo desta RSL era responder à seguinte pergunta: “Qual é o estado da arte na aplicação de técnicas de processamento de imagem digitais para aquisição, detecção e contagem de pragas de mosca-branca em imagens?”.

Algumas perguntas mais específicas que se desenrolaram com a anterior foram formuladas: (Q.1) Quais são os métodos existentes para aquisição e pré-processamento de imagens para pragas de mosca-branca? (Q.2) Quais são as técnicas utilizadas para a detecção e classificação de pragas em plantas, especificamente para a praga da mosca-branca? (Q.3) Quais são os resultados obtidos até o momento para detecção e classificação da praga da mosca-branca? (Q.4) Que medidas são usadas para avaliar a eficácia dos métodos existentes?

2.2 Palavras-chave, sinônimos e estratégias de busca

As seguintes palavras-chave foram escolhidas para a condução da RSL, considerando os termos técnicos comumente utilizados e o nome científico e popular da praga mosca-branca: (K.1) computer vision; and (K.2) deep learning; (K.3) neural networks; (K.4) image processing; (K.5) image classification; (K.6) plague detection; (K.7) whiteflies; (K.8) bemisia tabaci; (K.9) plant disease.

A string de busca foi definida como ("Computer vision"OR "Deep learning"OR "Neural networks"OR "Image processing"OR "Image classification") AND ("Plague detection"OR "Whiteflies"OR "Bemisia tabaci"OR "Plant disease").

Essa string foi ajustada para todas as bases de dados eletrônicas utilizadas, com o intuito de atender suas particularidades em questões sintáticas de pesquisa. Cinco bases de dados científicos foram consultadas, sendo estas: (1) ACM Digital Library; (2) IEEE Digital Library; (3) ISI Web of Science; (4) Science Direct; and (5) Scopus, uma vez que reúnem a maioria das publicações na área de ciência da computação e aplicações.

2.3 Critérios de Inclusão/Exclusão

Critérios de inclusão e exclusão foram definidos para orientar a seleção dos estudos primários relevantes para a RSL. Um estudo foi incluído se: (1) disponível em texto completo; (2) publicado em uma revista ou nos anais de uma conferência; (3) um relatório técnico, incluindo pesquisas; ou (4) dissertação de mestrado ou tese de doutorado. Por outro lado, foi excluído se: (1) não estivesse escrito em inglês; (2) foi publicado antes de 2008; (3) apresentar avaliações sem descrever o método utilizado; (4) não usar técnicas de processamento de imagem digital; (5) não usar técnicas de processamento digital de imagens para a detecção de pragas de mosca-branca.

Um conjunto de perguntas de avaliação da qualidade Sim / Não / Parcialmente também foi definido com base no refinamento das principais perguntas da pesquisa. Por exemplo, se os métodos ou técnicas de classificação foram relatados objetivamente ou se há um aplicativo ou produto final, e se foi descrito em detalhes.

2.4 Condução da revisão

As strings de busca adequadas, com as condições de pesquisa corrigidas, foram aplicadas aos bancos de dados. Em seguida, como etapa de Seleção do Estudo, as referências bibliográficas foram avaliadas de acordo com os critérios de inclusão e exclusão (com a ajuda das plataformas de software para RSL). Quando necessário, o texto completo dos estudos foi consultado. A tabela [2.1](#) resume o número de artigos

recuperados ao pesquisar em cada banco de dados e a quantidade de estudos aceitos (selecionados), rejeitados e duplicados durante a etapa de seleção. No total, 25 estudos foram aceitos.

Tabela 2.1: Número de estudos na RSL.

Base de dados	Total	Aceito	Rejeitado	Duplicado
ACM DL	103	4	82	17
IEEE DL	94	3	89	2
WebOfScience	129	5	69	55
ScienceDir	81	9	64	8
Scopus	31	4	12	15

Os estudos foram incluídos de acordo com os critérios de inclusão e exclusão e, além disso, foram aplicados os seguintes critérios de qualidade: (1) Os métodos ou técnicas de classificação foram relatados objetivamente?; (2) Os métodos de aquisição e pré-processamento foram citados explicitamente?; (3) A técnica foi usada em comparação com outras técnicas existentes? ; (4) Existe uma avaliação dessa técnica?; (5) Se houver um aplicativo ou produto final, ele foi descrito em detalhes?

Esses critérios foram elaborados com base nas perguntas da pesquisa e, para cada critério, foi atribuído um valor em conjunto com um conceito: Sim (S) = 1 ponto, Não (N) = 0 e Parcialmente (P) = 0,5, onde a junção de todos os critérios formam um índice classificado da seguinte forma: trabalhos com índice entre 0 e 1 foram considerados muito ruins, índices entre 1,5 e 2 foram considerados ruins, trabalhos entre 2,5 e 3 foram considerados bons, entre 3,5 e 4 trabalhos foram considerados muito bons, e os trabalhos obtidos entre 4,5 e 5 foram considerados excelentes. O valor do índice indicou a prioridade na leitura dos artigos e também foi um fator de exclusão de artigos se o trabalho tiver um índice muito baixo e não houver contribuição específica, o que justificaria o baixo índice. Após a avaliação da qualidade, 21 estudos foram selecionados para leitura completa.

A seguir, é descrito os principais conceitos em alguns dos estudos selecionados pela RSL.

2.5 Estratégias, métodos e algoritmos

[[Barbedo 2014](#)] explica que existem duas estratégias principais para a contagem automática de moscas-brancas encontradas na literatura, uma usando armadilhas adesivas e a outra usando folhas de plantas diretamente. Em imagens de armadilhas adesivas, apenas espécimes capazes de voar são considerados; em imagens de folhas, ninfas em estágio inicial podem ser consideradas.

2.5.1 Imagens em armadilhas adesivas

As armadilhas adesivas consistem em cartões de cola onde os espécimes são atraídos pela cor do papel e capturados pela substância pegajosa. A superfície lisa e neutra facilita a contagem.

[[Cho et al. 2008](#)] propôs um método para explorar características de tamanho e cor para identificar e contar moscas-brancas, pulgões e tripes, usando transformações de cores, operações simples de morfologia matemática e limiarização. [[Boissard, Martin e Moisan 2008](#)] propôs um sistema de visão cognitiva que combina processamento de imagem, aprendizado neural e técnicas baseadas no conhecimento para detectar e contar pragas. [[Qiao et al. 2008](#)] descreveu um sistema de processamento de imagem para estimar a população de espécimes adultos de moscas-brancas coletadas em armadilhas adesivas com base no tamanho e na cor. [[Bechar et al. 23-2](#)] usa técnicas como Mistura de Gaussianas e Análise de Componentes Principais para extrair pequenos pontos presentes na cena e detectar moscas-brancas em armadilhas adesivas em um sistema em tempo real. [[Solis-Sánchez et al. 2011](#)] desenvolveu um novo algoritmo capaz de detectar e identificar cinco espécies, incluindo mosca-branca, usando características geométricas de insetos que afetam a estufa. [[Barbedo 2014](#)] apresentou um novo método para contar moscas-brancas em folhas de soja com base na imagem digital convencional, o algoritmo é capaz de detectar e quantificar não apenas moscas-brancas adultas, mas também espécimes ninfais usando delineamento de região de interesse (conhecido como ROI), aplicação de transformações de cores e segmentação baseada em limiares e detecção de ninfas jovens. [[Xia et al. 2015](#)] usou o algoritmo da bacia hidrográfica e a distância de Mahalanobis, no espaço de cores YCrCb, para propor um procedimento de identificação de pragas para imagens de baixa resolução em armadilhas adesivas usando dispositivos portáteis. [[Espinoza et al. 2016](#)] propôs uma nova abordagem para a detecção e monitoramento de mosca-branca em estágio adulto (*Bemisia tabaci*) e trip (*Frankliniella occidentalis*) em estufas com base na combinação de um algoritmo de processamento de imagem e redes neurais artificiais e descreveu uma nova gama de possibilidades com a melhoria das redes neurais nos últimos anos. [[Sun et al. 2017](#)] propôs uma técnica que, ao invés de diretamente contar as pragas capturadas nas armadilhas, o conceito é tratar as pragas aprisionadas como ruído em uma imagem bidimensional (2D) com 2DFT servindo como um coletor de ruído específico.

2.5.2 Imagens em folhas de plantas

Essa abordagem consiste em tirar fotos diretamente na superfície das folhas em estufa ou em campo aberto. As principais dificuldades são as interferências do ambiente e os distúrbios climáticos em campo aberto que dificultam a contagem.

[Huddar et al. 26-2] propôs um método capaz de contar insetos não apenas em estufas, mas também em fazendas abertas. O algoritmo possui quatro etapas: conversão de cores, segmentação baseada na diferença relativa nas intensidades de pixel, redução de ruído por erosão e contagem com base no rastreamento de vizinhos de Moore e na parada de Jacob. [Li, Xia e Lee 2015] propôs uma nova aplicação de análise multifractal para a detecção de pragas de pequeno porte (por exemplo, mosca-branca) a partir de imagens da superfície foliar *in situ*. A análise multifractal foi adotada para segmentação de imagens de mosca-branca com base na singularidade local e nos caracteres de imagem global com os mínimos regionais como estratégia de seleção. [Maharlooei et al. 2017] apresentou o uso da técnica de imagem digital para detectar e contar pulgões de soja nas folhas de soja. [Lino et al. 18-2] fez uma comparação entre os desempenhos dos recursos LBP e Haar em um Classificador Cascade, que visa automaticamente a contagem de moscas-brancas imaturas (ninfas). [Jige e Ratnaparkhe 19-2] apresentou um algoritmo para detecção e contagem automática de moscas-brancas em imagens de plantações de algodão. O algoritmo utiliza conversão de espaço de cores, subtração de fundo, operação de limiar, operação morfológica e contagem do número de mosca-branca na folha com algoritmo centróide e propriedades da região. [Fuentes et al. 2017] propôs um detector baseado em redes neurais para doenças em tomate em tempo real e reconhecimento de pragas, incluindo a praga de mosca-branca em imagens de plantas de tomate. O objetivo era encontrar a arquitetura da rede neural mais adequada para a tarefa e três tipos de detectores foram considerados: um usando a Rede Neural Convolutiva baseada em região mais rápida (Faster R-CNN), outro usando a rede totalmente convolutiva baseada em região (R-FCN) e o mais recente Single Shot Multibox Detector (SSD). [Chen et al. 2018] apresentou um método de processamento de imagem usando a CNN para segmentar e contar ninfas de pulgões nas folhas. [Wang et al. 2018] propôs uma abordagem de segmentação cognitiva para imagens de pragas. O método funciona da seguinte maneira: Primeiro, uma imagem de praga é dividida em blocos por meio de um método de processamento de bloco de imagem. Segundo, um algoritmo de aprendizado adaptativo é usado para selecionar com precisão os centros de cluster iniciais. Terceiro, os resultados preliminares da segmentação são alcançados usando o agrupamento K-means. Finalmente, três características morfológicas digitais de uma elipse são adotadas para remover as veias das folhas. [Gutierrez et al. 2019] apresentou o desenvolvimento e a comparação de duas abordagens diferentes para detecção e identificação automatizadas de pragas baseadas em visão, usando estratégias de aprendizado. Uma solução que combina visão computacional e aprendizado de máquina é comparada a uma solução de aprendizado profundo. O foco principal de seu trabalho está na seleção da melhor abordagem baseada na detecção de pragas e na precisão da identificação.

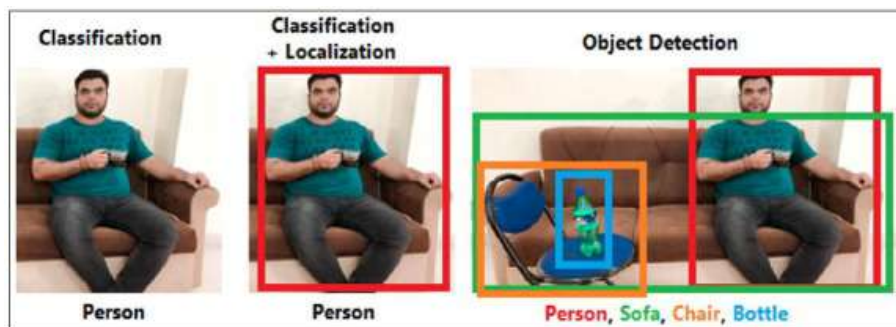
Conceitos básicos sobre detecção de objetos

De modo preliminar a descrição dos materiais e métodos empregados nesta pesquisa, é de suma importância estabelecer conceitos primordiais para o entendimento deste trabalho, sendo estes abordados no capítulo que se descreve a seguir.

3.1 Detecção de objetos

A tarefa de detecção de objetos está intrinsecamente ligada à etapa de classificação em visão computacional. Quando se deseja aplicar um algoritmo classificador em determinada imagem é necessário varrer toda imagem várias vezes com diferentes tamanhos de janelas de buscas. A janela de busca é a região que o classificador irá analisar para saber se a imagem possui o objeto que se deseja identificar. A partir deste processo é executada a detecção de objetos de uma determinada classe, onde diferente da classificação, o objetivo não é prever se determinada classe está presente na imagem. A detecção de objetos além de prever qual é a classe precisa identificar também a localização do objeto nessa imagem, de forma a caracterizar diferentes objetos para a mesma classe contidos na imagem.

Figura 3.1: Classificação, classificação + localização e detecção de objetos respectivamente.



Fonte: [Sharma e Mir 2019]

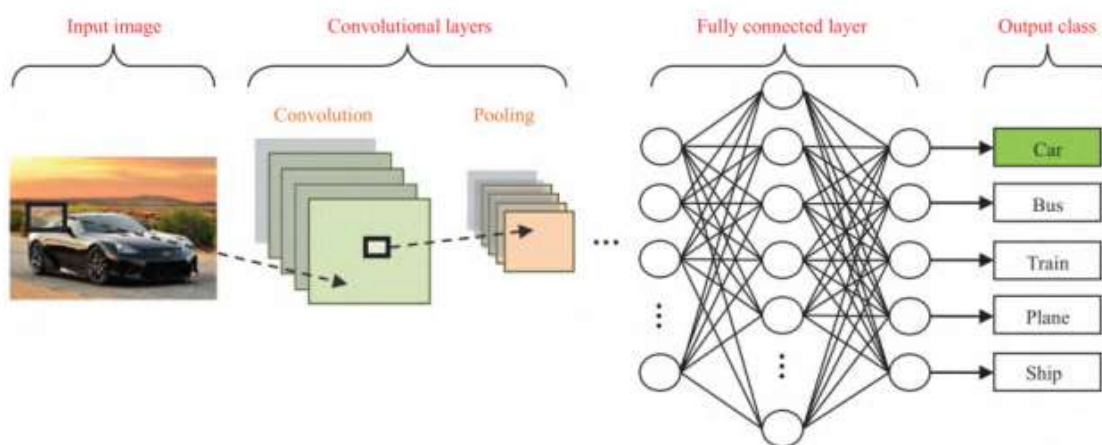
O processo de detectar objetos em uma imagem, visualizado na figura 3.1, tem se utilizado cada vez mais de estruturas de redes neurais convolucionais, sendo estas introduzidas na seção subsequente.

3.2 Redes neurais convolucionais

Apresentado em [Lecun et al. 1998] e inspirado no funcionamento do sistema nervoso dos seres vivos, as redes neurais convolucionais, ou *Convolutional Neural Networks* (CNN), são algoritmos computacionais que seguem o padrão *feed-forward*, onde todas as camadas se conectam a camada subsequente, seguindo o caminho da entrada para a saída da rede. Esta classe de rede neural artificial vem sendo aplicada com sucesso no processamento e análise de imagens digitais, onde ao converter imagens para o espaço de matrizes numéricas e atribuir pesos, torna-se possível simular sinapses, dando a habilidade ao algoritmo de manter o conhecimento adquirido durante o aprendizado.

Isto ocorre através dos ajustes dos pesos durante o treinamento, onde é realizado o processo de *back-propagation*, dividido em dois passos. No primeiro, uma amostra é fornecida à rede para que se obtenha uma saída. Esta saída é comparada em relação a saída desejada. A partir deste passo, o erro é computado e propagado da saída até a entrada e os pesos e limiares atualizados utilizando o gradiente descendente, de forma a se reduzir a somatória dos erros [Rumelhart, Hinton e Williams 1986].

Figura 3.2: Processo de classificação de imagem através de uma CNN



Fonte: [Rawat e Wang 2017]

Utilizando-se deste conceito, redes neurais convolucionais possuem a capacidade de extrair características, obtendo correlações entre as variáveis para detectar padrões e aprendizado por experiência através do ajuste automático de pesos e parâmetros para os exemplos selecionados. Tornando-se assim, o "estado da arte" em tarefas de classificação

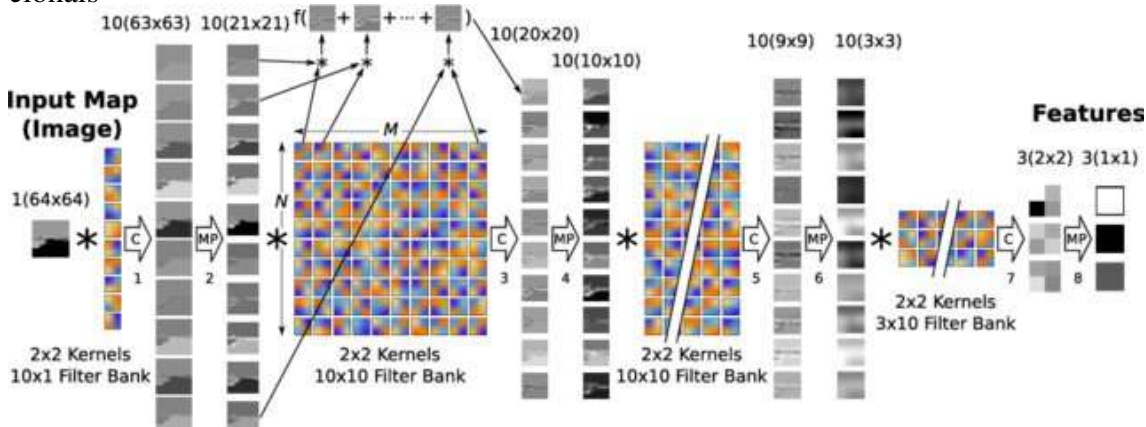
e detecção de objetos [Karpathy et al. 2014]. Atualmente diversas arquiteturas de redes neurais convolucionais são desenvolvidas e publicadas para aplicação em diferentes problemas, sendo a detecção de objetos uma de suas principais tarefas no âmbito de visão computacional.

3.3 Limitações em pequenos objetos

Um problema recorrente com os mapas de características de saída de um modelo de rede neural convolucional é a sensibilidade em relação à localização dos recursos na entrada. Uma abordagem para lidar com esta sensibilidade é reduzir a amostragem dos mapas de características através de uma camada de *pooling*, tendo esta como principal objetivo a redução dimensional da imagem e dos mapas de características [Sun et al. 2017]. Este processo de *downsampling* é necessário para que ocorra a invariância do modelo a pequenas alterações e distorções de imagem, bem como reduzir a quantidade de parâmetros treinados pela rede.

Através da figura 3.3 é possível observar uma *Max-Pooling Convolutional Neural Network (MPCNN)* com 8 camadas alternando entre convolução (C) e *downsampling* (MP: utilizando *max-pooling*). A primeira camada convolve a imagem de 64x64 pixels de entrada com filtros de tamanho 10x10 produzindo 10 mapas de tamanho 63x63, que são reduzidos para 10 amostras de 21x21 pela camada MP 2. A camada 3 convolve cada um desses 10 mapas com um filtro, soma os resultados e os passa pela função não linear f , produzindo 10 mapas de 20x20 pixels cada, e assim por diante até que a imagem de entrada seja transformada em apenas 3 recursos que são passados para as camadas subsequentes.

Figura 3.3: Exemplo de camada *max pooling* para um modelo de redes neurais convolucionais



Fonte: [Koutník, Schmidhuber e Gomez 2014]

A partir deste processo, caracteriza-se o problema para detecção de objetos de pequeno tamanho em modelos de redes neurais convolucionais, tendo em vista que a redução da resolução espacial das imagens nas camadas de *pooling* impacta na diminuição do conjunto de *pixels* dos objetos de interesse, reduzindo significativamente o desempenho do modelo.

Figura 3.4: Resultados de acurácia em detecção de pequenos objetos para arquiteturas recentes de redes neurais profundas

Method	Backbone	Size	FPS	AP	AP ₅₀	AP ₇₅	AP _S	AP _M	AP _L
YOLOv4: Optimal Speed and Accuracy of Object Detection									
YOLOv4	CSPDarknet-53	416	96 (V)	41.2%	62.8%	44.3%	20.4%	44.4%	56.0%
YOLOv4	CSPDarknet-53	512	83 (V)	43.0%	64.9%	46.5%	24.3%	46.1%	55.2%
YOLOv4	CSPDarknet-53	608	62 (V)	43.5%	65.7%	47.3%	26.7%	46.7%	53.3%
EfficientDet: Scalable and Efficient Object Detection [77]									
EfficientDet-D0	Efficient-B0	512	62.5 (V)	33.8%	52.2%	35.8%	12.0%	38.3%	51.2%
EfficientDet-D1	Efficient-B1	640	50.0 (V)	39.6%	58.6%	42.3%	17.9%	44.3%	56.0%
EfficientDet-D2	Efficient-B2	768	41.7 (V)	43.0%	62.3%	46.2%	22.5%	47.0%	58.4%
EfficientDet-D3	Efficient-B3	896	23.8 (V)	45.8%	65.0%	49.3%	26.6%	49.4%	59.8%
Learning Spatial Fusion for Single-Shot Object Detection [48]									
YOLOv3 + ASFP*	Darknet-53	320	60 (V)	38.1%	57.4%	42.1%	16.1%	41.6%	53.6%
YOLOv3 + ASFP*	Darknet-53	416	54 (V)	40.6%	60.6%	45.1%	20.3%	44.2%	54.1%
YOLOv3 + ASFP*	Darknet-53	608×	45.5 (V)	42.4%	63.0%	47.4%	25.5%	45.7%	52.3%
YOLOv3 + ASFP*	Darknet-53	800×	29.4 (V)	43.9%	64.1%	49.2%	27.0%	46.6%	53.4%

Fonte: [Bochkovski, Wang e Liao 2020]

Em [Bochkovski, Wang e Liao 2020] os autores demonstram na figura 3.4 a acurácia de diferentes modelos de redes neurais convolucionais em base de dados com diferentes tamanhos de objetos. Para as bases de dados de objetos de pequeno tamanho, é possível observar na coluna *average precision small* (APs) que o melhor desempenho de acurácia atinge apenas 26.7%, comparado a acurácias acima de no mínimo 40% para objetos em maior escala e 65.7% para o melhor resultado apresentado na coluna *average precision* (AP). No pior caso, para a rede *EfficientDet-D0* é alcançado ínfimos 12% de acurácia, demonstrando assim a extrema dificuldade em resolução de problemas de detecção de pequenos objetos em imagens.

[Solawetz 2020] define as principais estratégias para lidar na utilização de modelos de redes neurais profundas em tarefas de detecção de objeto de menor tamanho, sendo estas:

- **(a) Incrementar a resolução da imagem:** Objetos muito pequenos podem conter apenas alguns pixels dentro da caixa delimitadora, ressaltando a importância em aumentar a resolução espacial das imagens do conjuntos de dados, de forma a aumentar a riqueza de recursos que o modelo pode formar a partir dessa pequena caixa. Portanto, é de suma importância capturar imagens com a maior resolução espacial possível quando houver disponibilidade na etapa de aquisição de dados.

- **(b) Aumentar a resolução de entrada para o modelo:** Em complemento ao incremento da dimensão espacial das imagens, aumentar a resolução de entrada para as camadas iniciais da rede evita perdas ao decorrer do processo de *downsampling*.
- **(c) Repartir as imagens:** Efetuar durante a etapa de pré-processamento o processo de *tiling* das imagens, repartindo as figuras do conjunto de dados de forma a aumentar a caixa delimitadora dos objetos em relação ao espaço dimensional da imagem, diminuindo-se então as perdas de objetos de interesse nas camadas iniciais do modelo.

Os itens a e b referentes ao aumento de resolução da imagem, apesar de recomendáveis, resultam em aumento significativo do custo computacional durante os processos de treinamento e inferência dos modelos respectivamente, bem como elevam consideravelmente o tempo de execução para estas etapas, tornando essas estratégias inviáveis onde há baixa disponibilidade de recursos computacionais. Na seção 4 as estratégias adotadas para este trabalho são descritas.

Materiais e métodos

Neste capítulo são apresentados os dados e recursos utilizados neste trabalho. Além disso, são apresentados os métodos propostos e as métricas estatísticas utilizadas para avaliar os modelos executados.

4.1 Materiais

Nesta seção os conjuntos de dados utilizados são detalhados, bem como os processos de aquisição e pré-processamento para as respectivas bases. Foram consideradas imagens de plantas em contexto laboratorial e armadilhas adesivas, tendo em vista que o problema é dividido em dois contextos, como explicado nas seções [2.5.1](#) e [2.5.2](#).

4.2 Imagens de plantas (Embrapa)

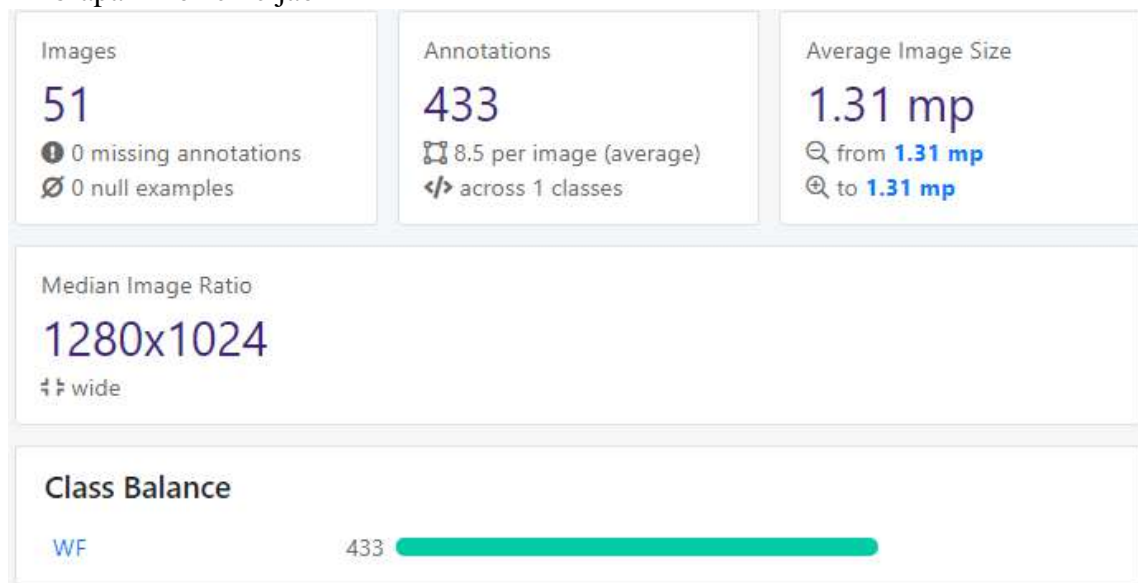
Tendo em vista que pesquisas iniciais no âmbito de detecção e contagens de ninfas foram executadas na unidade da Embrapa, um conjunto de imagens de plantas pré-adquirido foi levantado para a realização deste trabalho com o auxílio de estereoscópio (lupa). As imagens foram obtidas por pesquisadores durante suas atividades laboratoriais. Para não inviabilizar futuras aplicações e facilitar o processo de aquisição de imagens, não foram definidos padrões de distância focal e iluminação. Os pesquisadores usaram microscópio digital da marca *DinoLite*. O equipamento possui conexão usb, permitindo capturar imagens e salvá-las diretamente no computador. Para se acoplar com o estereoscópio é utilizado o modelo *Dino-Eye*. Este modelo transforma o estereoscópio em um dispositivo óptico digital. A figura [4.1](#) pertence ao conjunto de dados disponibilizado. Nela pode-se identificar facilmente as ninfas de mosca-branca que serão detectadas em etapa posterior. Ao todo foram capturadas 51 imagens contendo 433 amostras para a mosca-branca em estágio ninfal, As informações sobre este conjunto de dados podem ser visualizadas na [4.2](#).

Figura 4.1: Imagens obtidas com auxílio do Estereoscópio



Fonte (Embrapa, 2019)

Figura 4.2: Conjunto de imagens de ninfas obtido com apoio de pesquisadores da Embrapa Arroz e Feijão



Fonte: Elaborado pelo autor

4.3 Imagens de armadilhas adesivas (4TU Dataset)

Todas as imagens de armadilhas adesivas utilizadas neste estudo foram retiradas do conjunto de dados público em [Nieuwenhuizen et al. 2019] disponibilizado pela 4TU Research Data, instituto de pesquisa internacional administrado por um consórcio de universidades nos Países Baixos. As imagens foram captadas através de um Scoutbox [Agrocares], dispositivo com câmera e scanner acoplados, sob controle de condições climáticas em estufas localizadas na Bélgica, com uma resolução de 5184 x 3456 pixels. Das 6900 imagens obtidas pelo centro de pesquisa, um subconjunto de 284 imagens foi escolhido aleatoriamente para representar a variabilidade de populações esperadas, sendo este subconjunto publicado e utilizado neste trabalho. A figura 4.3 apresenta uma das imagens obtidas do dataset.

Essa base possui três classes de insetos, sendo 1312 amostras de *Macrolophus*, 510 para o *Nesidiocoris*, estas duas variações de espécie para o inseto denominado Tuta-Absoluta e 5591 amostras para o inseto mosca-branca. Tais classes estão anotadas no formato *PASCAL VOC*, contendo cada imagem seu arquivo em formato .jpg e outro arquivo em formato .xml com as anotações e coordenadas do inseto imagem. Para os experimentos, apenas a classe mosca-branca, objeto de interesse deste estudo, foi utilizada.

Figura 4.3: Imagem obtida no dataset



Fonte (4TU.Centre for Research Data, 2019)

Sendo previsto em trabalhos futuros o desenvolvimento de uma armadilha inteligente, integrando câmeras em tempo real e o uso de armadilhas adesivas, o conjunto de dados se mostrou viável para os experimentos realizados. A divisão dos dados se deu através de três conjuntos: treino, representando 70% do número total de imagens; validação, abrangendo 20% das imagens; e teste correspondendo à 10% da quantidade total de imagens. Esta divisão foi feita após a etapa de pré-processamento das imagens, descrita a seguir.

4.4 Pré-processamento dos dados

No geral, em sua fase adulta, as moscas-brancas medem entre 1 e 2 milímetros. Tais medidas caracterizam um problema de detecção de objetos de pequeno tamanho. Sendo assim as imagens possuem alta resolução com o intuito de ressaltar o inseto capturado pela armadilha. Inicialmente com imagens obtidas em resolução de 5184x3456 pixels, o processo de *data augmentation*, em outros termos, aumento de dados, foi aplicado no *dataset* através da técnica de *tiling*, abordagem a qual consiste em aumentar o conjunto de dados com o uso de blocos de subconjuntos que são cortados nas imagens originais de alta resolução, enquanto as anotações das coordenadas delimitadoras do objeto são mapeadas de acordo para um novo par de imagem e arquivo de anotação, isto é, dividindo a imagem de alta resolução em subpartes, consegue-se mapear pequenos objetos para uma área relativa maior, evitando perdas nas camadas iniciais da estrutura da rede neural convolucional durante a etapa de treino, que impossibilitariam o algoritmo de visualizar e identificar o objeto de interesse, além de manter o bom desempenho computacional do modelo, tendo em vista que as imagens em sua resolução original resultariam em um alto custo computacional para o processamento do algoritmo. Esta abordagem pode ser utilizada independentemente da arquitetura empregada, ressaltando que, ao realizar o recorte das imagens para a entrada da rede neural na fase de treino, deve-se realizar o mesmo recorte para imagens durante a fase de inferência, momento onde o modelo treinado detecta as moscas localizadas na imagem.

Em contrapartida, ao aplicar tal técnica, deve-se considerar possíveis perdas em amostras localizadas nas extremidades das subpartes da imagem, portanto, quanto maior a quantidade de recortes, aumenta-se a perda de moscas-brancas anotadas para o conjunto de dados. Através da tabela 4.1 é possível observar o número de imagens e amostras do inseto após a aplicação do algoritmo com uma grade de corte 6x4, ou seja, cada imagem de resolução original 5184x3456 pixels dividida horizontalmente em seis partes e verticalmente em quatro partes, totalizando vinte e quatro partes de 864x864. Tal resolução e grade foram estipulados considerando as arquiteturas de redes neurais selecionadas para o estudo.

Tabela 4.1: Quantidade de amostras do conjunto de dados

Base de dados	Imagens	Amostras
4TU Dataset Original	284	5591
4TU Dataset Recortado 6x4	6816	5049

É possível observar o aumento do conjunto em vinte quatro vezes o seu valor total, sendo a perda de amostras da mosca-branca em apenas 10%, tendo em vista que aumentar a grade resultaria em perdas significativas de amostras, bem como a diminuição ocasionaria imagens de alta resolução que inviabilizariam a solução devido ao alto custo computacional e o "sumiço" de amostras ao passar as imagens nas camadas iniciais dos modelos. O limiar de 6x4 se mostrou uma alternativa viável e satisfatória para a continuidade dos experimentos realizados neste trabalho.

4.5 As arquiteturas

Ao tratar-se de modelos de detecção de objetos, existem dois principais tipos: modelos de um ou dois estágios. Um modelo de um estágio é capaz de detectar objetos sem a necessidade de uma etapa preliminar. De maneira oposta, um modelo detector de dois estágios usa um estágio preliminar onde regiões de importância são detectadas e então classificadas para ver se um objeto foi detectado nessas áreas. A vantagem de um detector de um estágio é a velocidade com que é capaz de fazer previsões rapidamente, permitindo um uso em tempo real. Considerando a possibilidade de construção de uma armadilha inteligente em tempo real em estágios futuros desta pesquisa, foram selecionados modelos de um estágio para este trabalho.

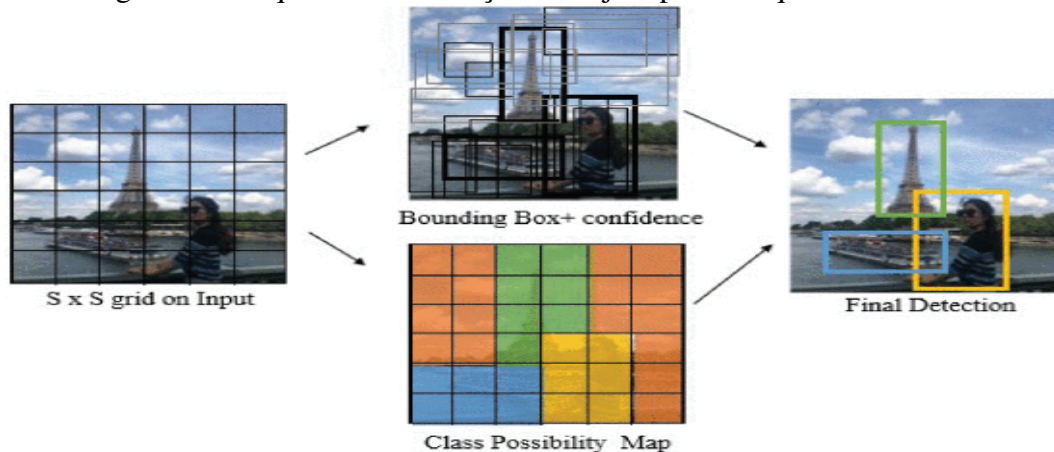
4.5.1 You Only Look Once

O YOLO (*You Only Look Once*), descrito por [Redmon et al. 2016], é um dos primeiros métodos para detecção de objetos em um estágio, onde através de uma rede convolucional é feita simultaneamente a predição das caixas delimitadoras e suas respectivas classes, permitindo uma única avaliação da imagem. Outra vantagem do método é dispor de informações do contexto, pois analisa a imagem como um todo, diminuindo erros de falsos positivos em relação ao plano de fundo. Além disso, esse tipo de abordagem oferece uma boa generalização, permitindo ser utilizado em domínios diferentes do qual foi treinado. No entanto, apresenta resultados inferiores ao se tratar de precisão e problemas na detecção de pequenos objetos.

Em suma, YOLO divide a imagem em regiões de grade $S \times S$ e prevê a caixa delimitadora de cada região de grade. Cada célula da grade prediz B caixas delimitadoras

com o grau de confiança para elas. Ao mesmo tempo, prevê a probabilidade da classe. Depois disso, a probabilidade da caixa e da classe são combinadas e o objeto é encontrado usando a supressão não máxima. Este processo é ilustrado na figura 4.4.

Figura 4.4: Sequência de detecção de objeto para a arquitetura YOLO



Fonte [Redmon et al. 2016]

4.6 YOLOv3-v5

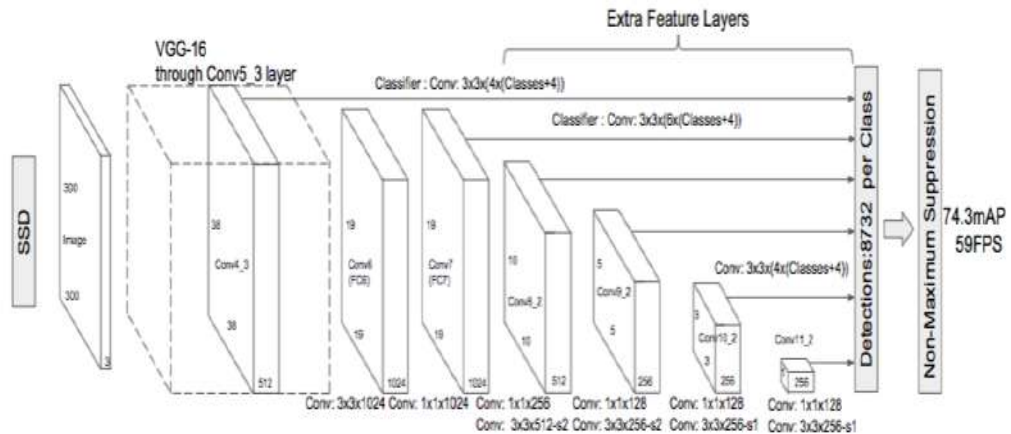
Foram selecionada para os experimentos a YOLO em suas versões 3, 4 e 5 devido ao fato de serem as mais recentes e representarem o estado da arte para o algoritmo. Estas versões se destacam pelo o uso de classificadores lógicos independentes para múltipla classificação em contextos mais complexos com sobreposição de classes. Três escalas diferentes foram propostas pelos autores para a predição das caixas delimitadoras, ao mesmo tempo que tensores de três dimensões são usados para codificar as classes e obter maior objetividade, bem como técnicas de data augmentation e seleção de hiperparâmetros por algoritmos genéticos.

4.6.1 Single Shot MultiBox Detector (SSD)

Single Shot Detector [Liu et al. 2016] é um método de detecção de objeto de estágio único que discretiza o espaço de saída das caixas delimitadoras em um conjunto de caixas padrão em diferentes proporções e escalas por localização do mapa de características. No momento da previsão, a rede gera pontuações para a presença de cada categoria de objeto em cada caixa padrão e produz ajustes na caixa para melhor corresponder à forma do objeto.

A SSD é dividida em duas etapas. A primeira é uma etapa de extração de *features* utilizando a arquitetura VGG-16 e após isso passa por filtros convolucionais extras para fazer a classificação e predição das *boundary box* considerando as diferentes escalas, aspectos e posicionamento. Tal arquitetura pode ser vista na figura 4.5.

Figura 4.5: Arquitetura da rede *Single Shot MultiBox detector*



Fonte [Liu et al. 2016]

A SSD demonstrou não se adaptar bem a pequenos objetos nas imagens, tendo em vista que seu processo de extração de característica pode diluir a informação do objeto para a detecção nas próximas camadas.

4.7 Equipamentos e programas utilizados

Os resultados obtidos foram alcançados em um computador com as seguintes configurações:

- Processador: Intel(R) Xeon(R) CPU E5-2698 v4 @ 2.20GHz
- Memória: 500 GB Ram
- GPU: NVIDIA Tesla V100 16GB

Os algoritmos de treino utilizarem-se das seguintes bibliotecas de código:

- CUDA and cuDNN: Desenvolvida pela empresa Nvidia, esta biblioteca permite a execução de algoritmos de alto custo computacional em dispositivos GPU;
- Darknet: Contém as implementações para as quatro primeiras versão da arquitetura YOLO, sendo este *framework* desenvolvido pelos autores da rede;
- LabelImg: *Software* de código-aberto desenvolvido em Python, utilizado para anotar as imagens dos conjuntos de dados e exportar os seus respectivos arquivos de anotação para os diferentes formatos de treinamento.

- Pytorch: Disponibilizada pela equipe de inteligência artificial do Facebook, contém a implementação da rede YOLO para a sua quinta versão.
- Roboflow: Plataforma online voltada para problemas de visão computacional deste a etapa de pré-processamento, anotação e treinamento dos modelos;
- TensorFlow 2: Biblioteca de código-aberto desenvolvida pela Google onde é possível configurar e executar treinamentos para diferentes arquiteturas de redes neurais, bem como uma ampla variedade de tarefas em Inteligência Artificial.

4.8 Métricas de avaliação

Para avaliação de resultados dos modelos utilizados, foi utilizada a *Average Precision*, principal métrica utilizada ao comparar diferentes detectores de objetos. Essa métrica é precedida por importantes conceitos:

Intersect Over Union (IoU): é uma medida baseada no índice Jaccard - estatística usada para medir a similaridade e diversidade de conjuntos de amostras - que avalia a sobreposição entre duas caixas delimitadoras. Requer uma caixa delimitadora com valores verdadeiros anotados e uma caixa delimitadora prevista pelo modelo, processo ilustrado na 4.6. Ao aplicar o IOU, podemos dizer se uma detecção é válida (Verdadeiro Positivo) ou não (Falso Positivo) com.

True Positive: Quando o modelo detecta e/ou classifica um objeto o qual ele deve detectar. Para que isto ocorra, é necessário que o IoU seja maior que um *threshold* pré-estabelecido para aquela classe de objeto. São valores para o threshold: 0.5, 0.75 e 0.95, sendo o primeiro representando 50% o mais utilizado.

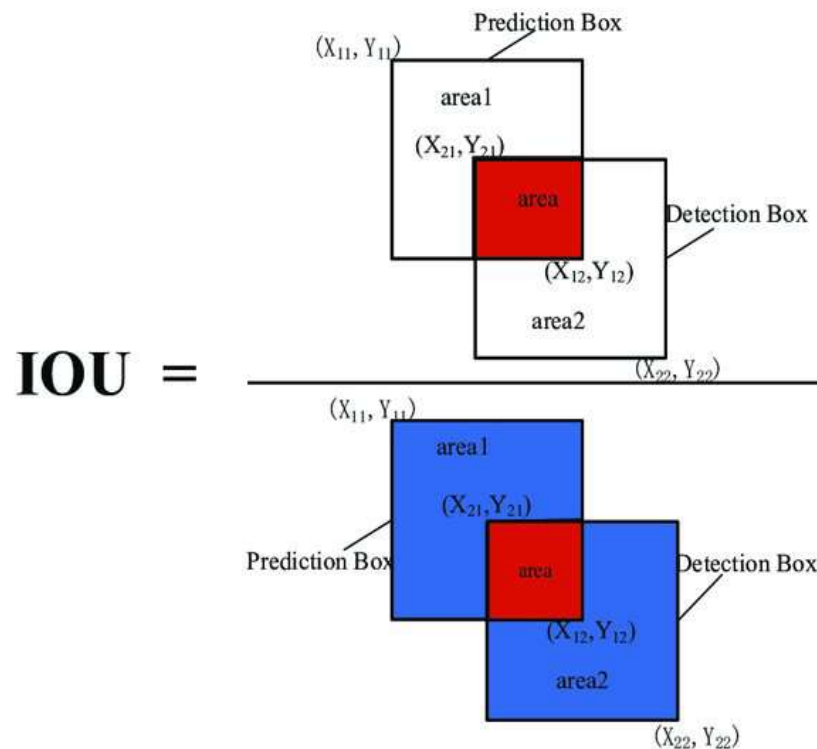
False positive: Ocorre ao se detectar um objeto o qual não deveria ter sido detectado para aquela classe. Em suma, caso uma caixa delimitadora prevista pelo modelo não represente o objeto de interesse e o valor de IoU seja menor que o *threshold*.

Precision: A precisão pode ser definida como a fração de elementos relevantes selecionados dentro de um grupo de elementos selecionados. Ou seja, em uma classificação binária, a precisão é a razão entre o número de verdadeiros positivos selecionados por um modelo e o número de elementos que foram classificados como positivos pelo modelo. Sendo assim, a precisão pode ser definida pela equação 4-1:

$$Precision = \frac{TP}{TP + FP} \quad (4-1)$$

Recall: é definido como a razão entre os elementos relevantes selecionados e todos os itens relevantes dentro do universo de elementos disponíveis. Em outras palavras, o recall é definido como a fração de verdadeiros positivos selecionados dentro de todos

Figura 4.6: Cálculo para a Intersecção sobre a União)



Fonte [Cheng, Zhao e Zhang 2019]

os positivos disponíveis no grupo de elementos. O recall pode ser definido pela equação 4-2:

$$Recall = \frac{TP}{TP + FN} \quad (4-2)$$

A curva Precision x Recall avalia o desempenho de um detector de objeto à medida que a confiança é alterada traçando uma curva para cada classe de objeto. Um detector de objeto de uma determinada classe deve permanecer com alta precisão à medida que a revocação (*recall*) aumenta, ou seja, variando o limite de confiança (*threshold*), a precisão e a revocação ainda serão altas. Outra maneira de se avaliar é observando a capacidade do modelo de identificar apenas objetos relevantes (0 Falsos Positivos = alta precisão), encontrando todos os objetos de verdade (0 Falsos Negativos = alta recuperação).

Um detector de objetos é considerado com baixo desempenho quando é necessário aumentar o número de objetos detectados (aumentando os falsos positivos = menor precisão) para recuperar todos os objetos verdadeiros (alta recuperação). Sendo assim, a curva Precision x Recall geralmente tende a iniciar com valores de alta precisão, diminuindo conforme a recall aumenta.

Average Precision (AP) : Outra métrica utilizada para comparar o desempenho dos detectores de objetos é calcular a área sob a curva (AUC) da curva Precision x Recall. Como as curvas de Precisão Média (AP) são frequentemente curvas em zigue-zague que sobem e descem, comparar curvas diferentes, ou seja, detectores distintos no mesmo gráfico geralmente não é uma tarefa compreensível, - tendo em vista que as curvas tendem a se cruzar com muita frequência. Sendo assim, a Precisão Média (AP), uma métrica numérica, é utilizada como método comparativo entre modelos detectores de objetos. Na prática, a precisão média AP é a precisão calculada em média em todos os valores de recall entre 0 e 1.

Caso haja mais de uma classe no detector de objetos, calcula-se a média das métricas de AP, formando a métrica *Mean Average Precision*. Para este trabalho especificamente, por se tratar de um detector de uma classe para o inseto mosca-branca, a métrica *Average Precision* (AP) representada pela equação 4-3 é suficiente.

$$AP = \sum_n (R_n - R_{n-1}) P_n \quad (4-3)$$

Resultados

Neste capítulo serão apresentados os principais resultados obtidos em cada etapa de testes nas bases de dados e arquiteturas avaliadas para este trabalho. Na seção 5.1 é descrito a etapa de treinamento, em seguida é apresentada a análise dos testes na seção 5.2.

5.1 Treinamentos

Para se treinar uma rede convolucional do início, é necessário alto custo computacional e uma vasta quantidade de dados, sendo estes itens limitantes e restringindo a execução dos treinos para projetos com baixo recurso disponível. Desta forma, um processo geralmente empregado é o *transfer learning*, onde utiliza-se um modelo inicial previamente treinado em um conjunto de dados maior, transferindo o aprendizado do modelo inicial através dos pesos e executando o algoritmo para o contexto do novo problema. Como descrito em [Pan e Yang 2010], o processo de transfer learning é necessário para criar modelos de alto desempenho treinados com dados obtidos mais facilmente em diferentes domínios. Foram efetuados treinamentos para as arquiteturas de rede YOLO e SSD MobileNet em suas diferentes versões e resoluções para os conjuntos de dados disponíveis neste trabalho. Para os testes do *dataset* de armadilhas adesivas, foram efetuados treinamentos em diferentes resoluções, sendo 864x864 a resolução de entrada original após o processo de *tiling*, descrito na seção 4.4, e testes com entrada 640x640, sugeridos pelos autores das implementações do modelo devido a diminuição do tempo de treinamento.

Para a arquitetura YOLO, em sua versão 3 e 4, é recomendável o treino em pelo menos 2.000 épocas para cada classe do *dataset*, porém não menos que a quantidade de imagens do conjunto e não menos que 6.000 iterações no total como descrito pelo autor em [Bochkovski, Wang e Liao 2020]. Sendo assim, foram executados o número máximo de 6.000 iterações para o conjunto de imagens de armadilhas adesivas. Para os modelos YOLOv5 foram realizados treinos de 300 épocas, como sugerido no repositório original do autor em [Jocher et al. 2020] para os dois conjuntos de dados utilizados neste trabalho.

Em relação à arquitetura SSD MobileNet, foi efetuado o treino em até 10 mil épocas, observando a diminuição dos valores de validação de forma a evitar o erro conhecido como *overfitting*, isto é, quando o modelo se adapta ao conjunto de dados por excesso de treinamento e perde a capacidade de generalização para conjunto de amostras não vistos anteriormente.

Primeiramente, analisando os testes para o conjunto de dados de armadilhas, observa-se, na figura 5.1 que a rede YOLOv5, em sua versão *medium* para resolução de imagens 640x640, apresenta boa evolução no treinamento, alcançando índice de acurácia mAP acima de 0.65 com um *threshold* de 50%, isto é, quando o *confidence score* considera apenas caixas delimitadoras previstas maiores que o limiar definido de 0.5.

Ao aumentar este delimitador para 0.95, é possível analisar considerável queda vide figura 5.2, atingindo aproximadamente 0.25 de mAP. Isto ocorre pela dificuldade do modelo em prever, com alto índice de confiança, os objetos detectados problema recorrente em detecção de pequenos objetos.

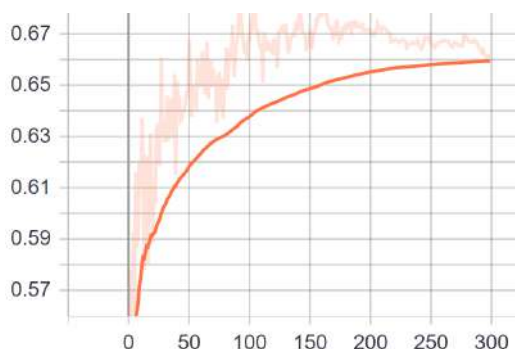


Figura 5.1: Curva de precisão média para o treinamento da rede YOLOv5m no 4tuDataset em resolução 640x640 com *threshold* de 0.5

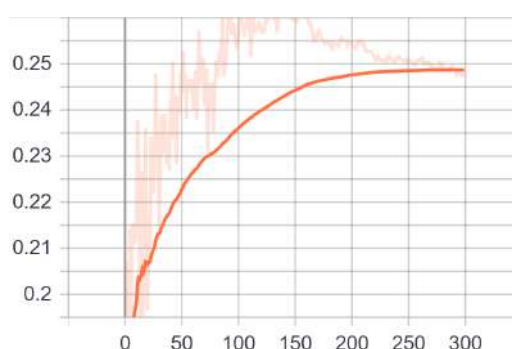


Figura 5.2: Curva de precisão média para o treinamento da rede YOLOv5m no 4tuDataset em resolução 640x640 com *threshold* de 0.95

Em seguida, os valores indicados para as curvas *precision* e *recall* apresentados nas figuras 5.3 e 5.4 para o mesmo treinamento efetuado.

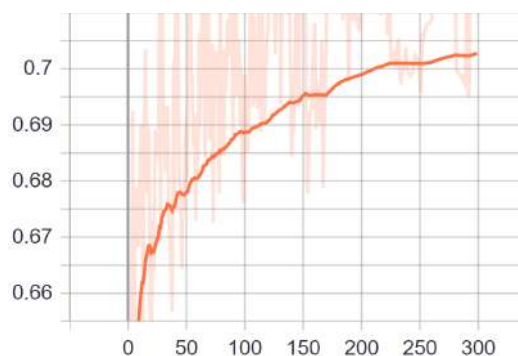


Figura 5.3: Curva de precisão no treinamento da rede YOLOv5m em resolução 640x640 (imagens de armadilhas)

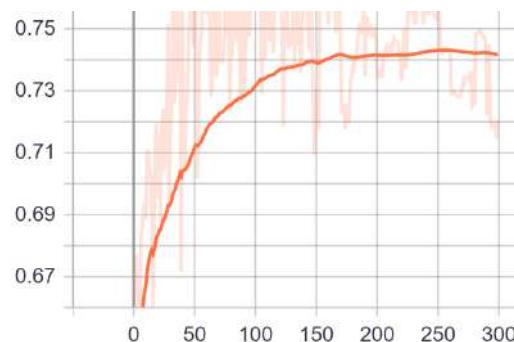


Figura 5.4: Curva de recall no treinamento da rede YOLOv5m em resolução 640x640 (imagens de armadilhas)

As curvas em direção crescente indicam que a rede não está cometendo muitos erros de falsas marcações (FP) e nem deixando de marcar muitos objetos necessários (FN) para o conjunto de dados analisados.

Em comparação ao treinamento para a rede em resolução 864x864 é perceptível o aumento da precisão acima de 0.7, se aproximando a acurácia de 0.8 observado nas figuras 5.5 e 5.6, demonstrando assim o aumento da acurácia do modelo ao realizar o treinamento para a escala original da imagem após a etapa de pré-processamento.

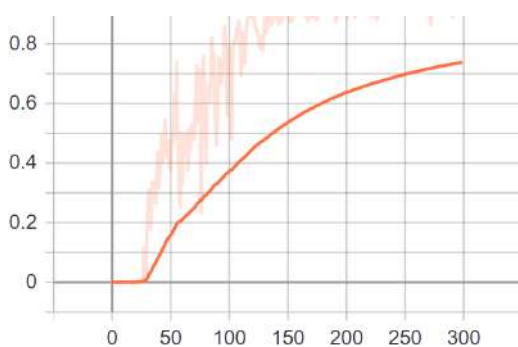


Figura 5.5: Curva de precisão média para o treinamento da rede YOLOv5m no 4tuDataset em resolução 864x864 com *threshold* de 0.5

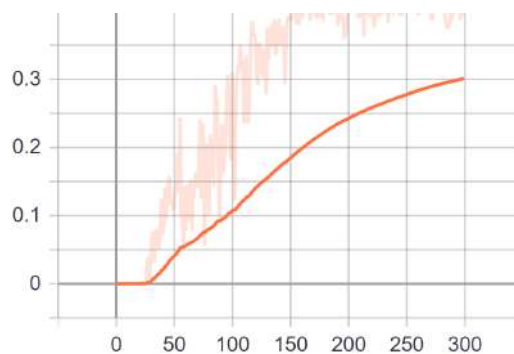
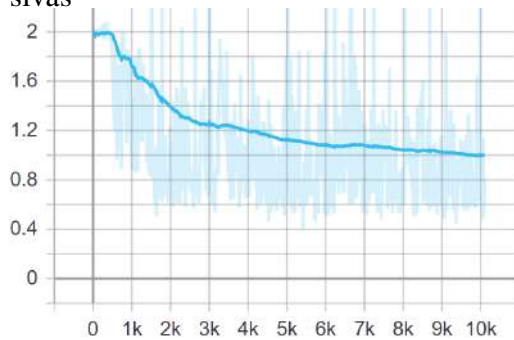


Figura 5.6: Curva de precisão média para o treinamento da rede YOLOv5m no 4tuDataset em resolução 864x864 com *threshold* de 0.95

Observando o treinamento para a rede SSD MobileNetV2, nos conjuntos de dados de armadilhas adesivas e imagens de ninfas, respectivamente, apresentados nas figuras 5.7 e 5.8, é possível identificar o melhor desempenho do modelo para o conjunto de imagens de ninfas com menores valores de perdas, sugerindo que, apesar de em menor

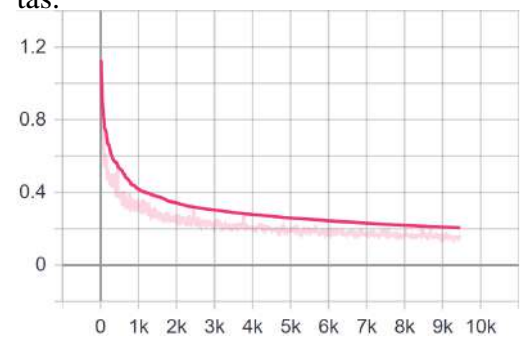
o quantidade de imagens, o modelo se adaptou melhor ao conjunto de dados anotados com alta resolução para as moscas em estágios ninfais.

Figura 5.7: *Loss curve* para o treinamento da rede SSD MobileNetV2 no conjunto de imagens de armadilhas adesivas



Fonte: Elaborado pelo autor

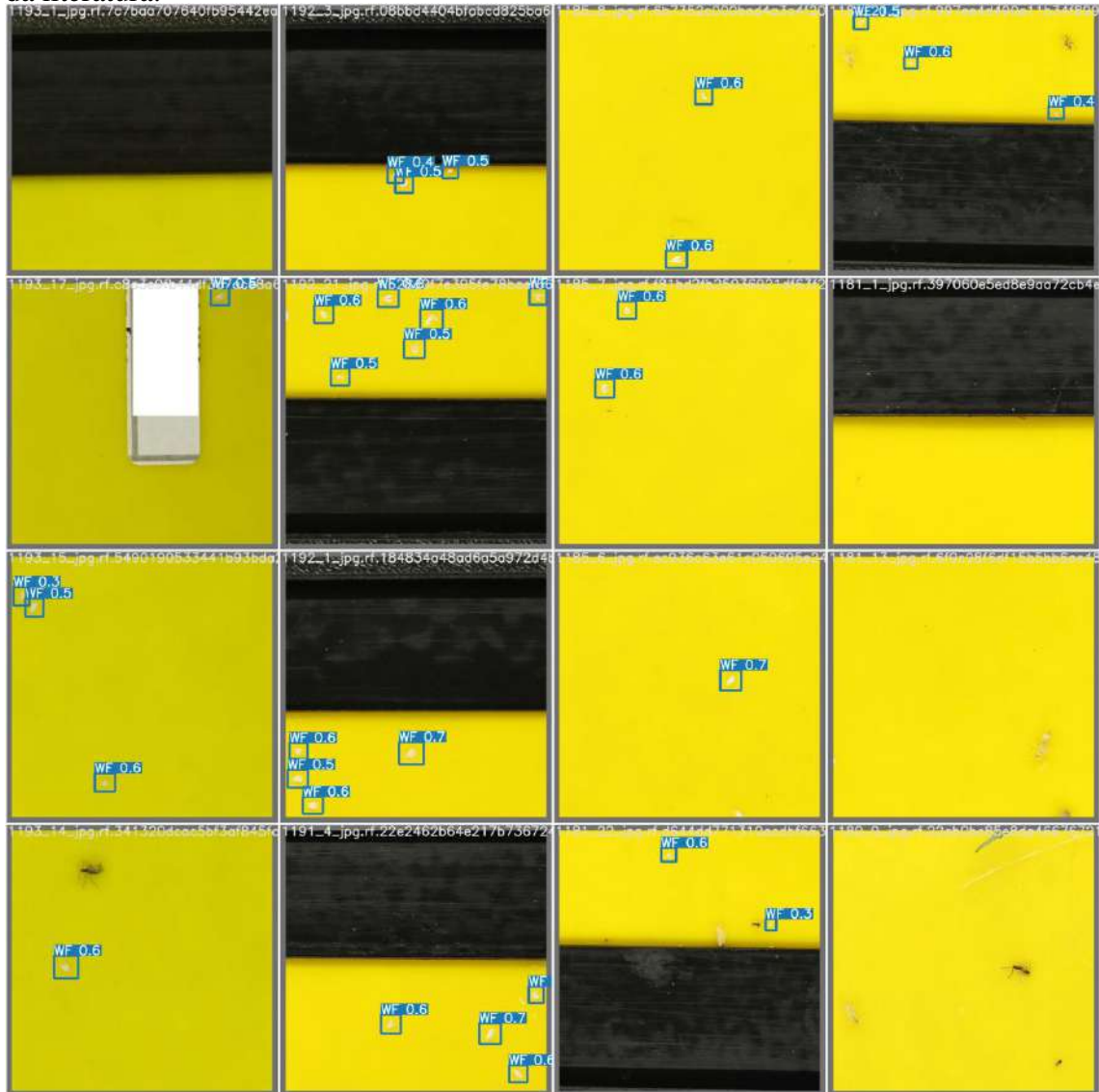
Figura 5.8: *Loss curve* para o treinamento da rede SSD MobileNetV2 no conjunto de imagens de ninfas em plantas.



Fonte: Elaborado pelo autor

Em seguida, é possível observar, nas figuras [5.9](#), [5.10](#), [5.11](#), [5.12](#), [5.13](#) e [5.14](#) alguns resultados de saídas em experimentos obtidos na etapa de treinamento e testes do sistema de detecção, destacando resultados promissores e a capacidade de generalização dos modelos treinados através do processo de *transfer learning*.

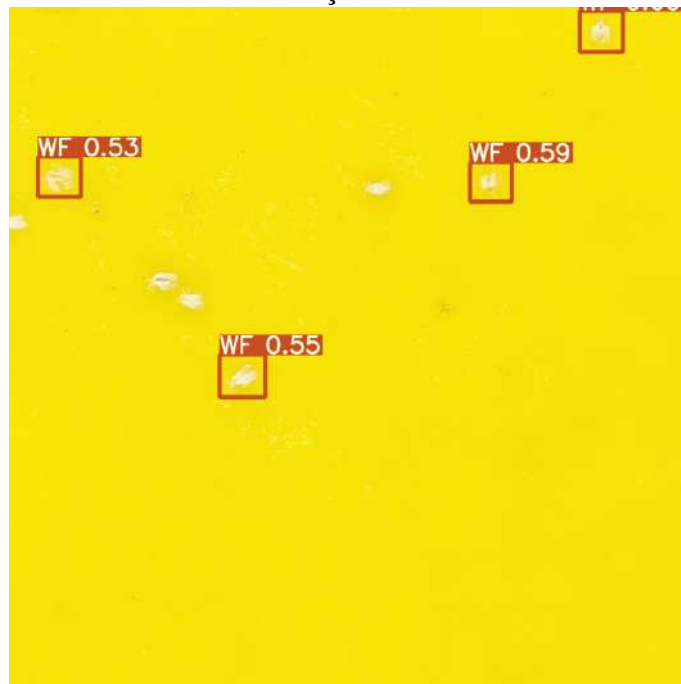
Figura 5.9: Conjunto de treinamento para a rede YOLOv5m, destacando a habilidade da rede em detectar moscas-brancas ao redor e nas bordas da armadilha adesiva, problema recorrente em trabalhos similares citados anteriormente na etapa de revisão sistemática da literatura.



Fonte: Elaborado pelo autor

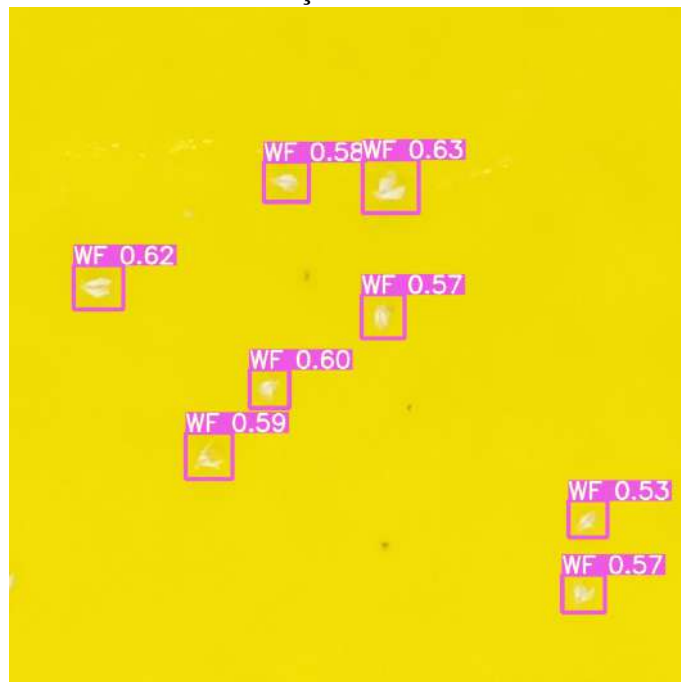
Durante a próxima seção, serão apresentados os valores de métricas de acurácia para os modelos de detecção de objetos treinados.

Figura 5.10: Exemplo 1 de inferência para o modelo YOLOv5m em imagens do conjunto de testes de armadilhas adesivas em resolução 640x640.



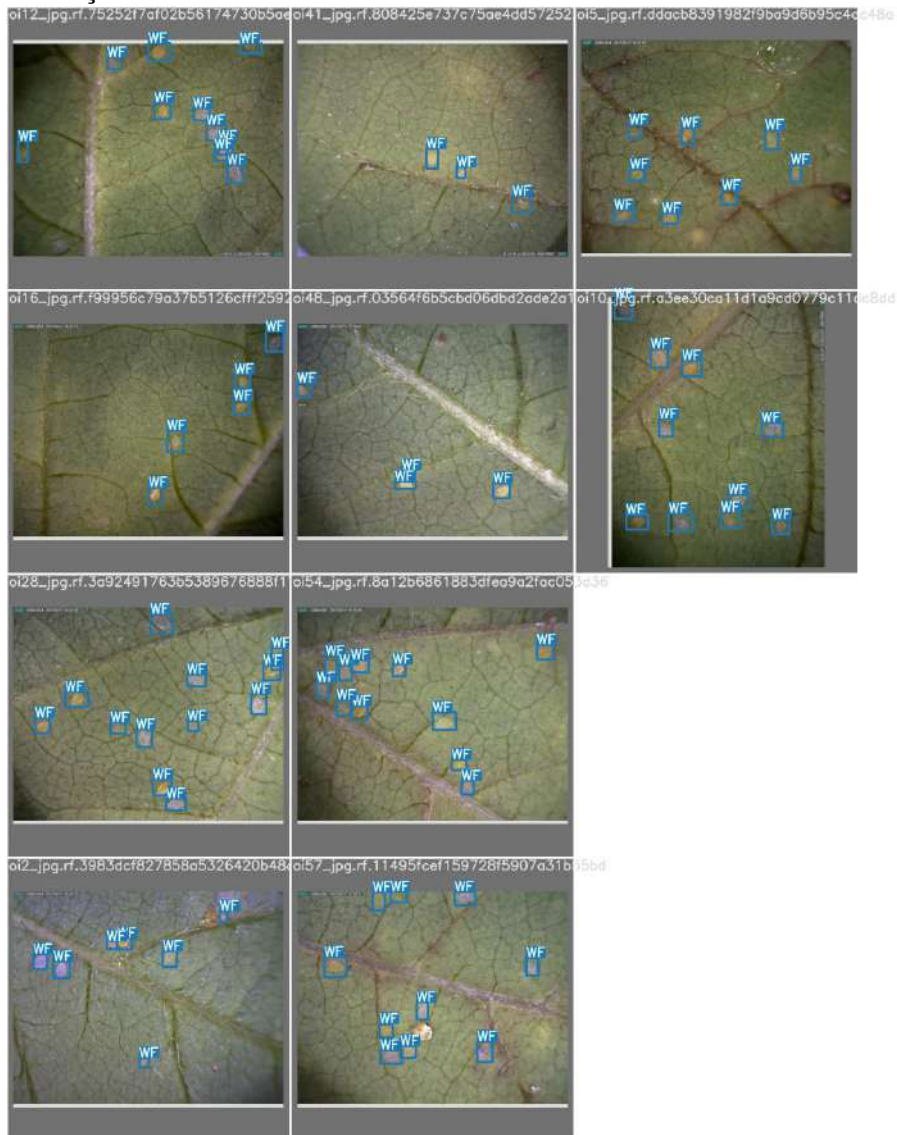
Fonte: Elaborado pelo autor

Figura 5.11: Exemplo 2 de inferência para o modelo YOLOv5m em imagens do conjunto de testes de armadilhas adesivas e resolução 640x640..



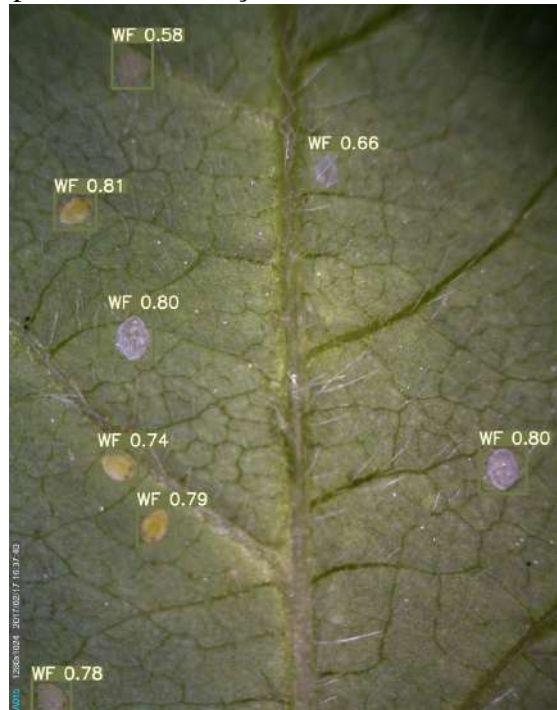
Fonte: Elaborado pelo autor

Figura 5.12: Conjunto de treinamento da rede YOLOv5m para as imagens de ninfas em resolução 640x640.



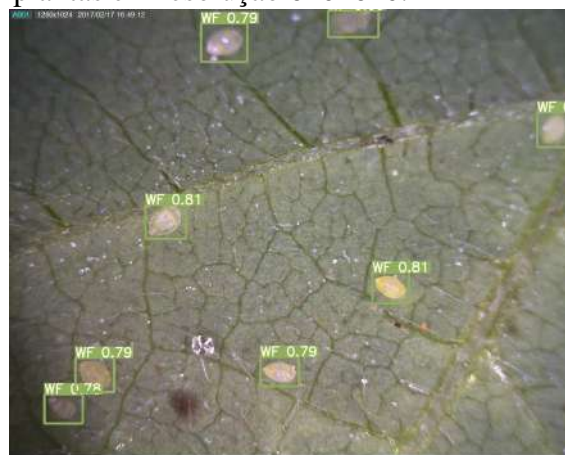
Fonte: Elaborado pelo autor

Figura 5.13: Exemplo 1 de inferência para o modelo YOLOv5m em imagens do conjunto de testes de ninfas em plantas em resolução 640x640.



Fonte: Elaborado pelo autor

Figura 5.14: Exemplo 2 de inferência para o modelo YOLOv5m em imagens do conjunto de testes de ninfas em plantas em resolução 640x640.



Fonte: Elaborado pelo autor

5.2 Análise dos testes

Utilizando os conjuntos de dados para testes, os resultados obtidos estão apresentados nas tabelas 5.2 e 5.2 respectivamente, em destaque (negrito) para os melhores valores observados para cada *dataset*. Os testes foram realizados seguindo o padrão de cada modelo, sendo o modelo YOLOv5m o de melhor desempenho em relação a métrica *average precision* nos dois casos, enquanto o YOLOv5s a melhor velocidade, isto é, menor tempo para processar uma imagem para as duas bases de dados. É possível observar uma baixa diferença entre os modelos YOLO, em sua versão *medium* e *small*, indicando o uso do segundo modelo em casos onde a queda de precisão não gere impactos significativos para a solução.

Os valores para AP foram obtidos com threshold de 50% e a velocidade de inferência avaliada em uma GPU Tesla V100 em ambos os casos. A tabela apresenta respectivamente: (a) o modelo treinado; (b) a velocidade de inferência para o modelo no dataset original dos autores; (c) A precisão média atingida na publicação dos modelos pelos autores; (d) O tamanho do modelo em megabytes; (e) A precisão média atingida para este trabalho nos datasets testados; (f) o tempo de inferência para o modelo em imagens executadas nos conjuntos de testes;

Modelo	Speed (ms)	COCO AP	Size (mb)	Author AP.50	Author Speed (ms)
YOLO v3 864x864	-	-	-	58.1	28
YOLO v4 864x864	-	-	-	59.2	21
YOLO v5m 640x640	3.4	44.3	41	66.0	10.7
YOLO v5m 864x864	3.4	44.3	41	67.2	13.2
YOLO v5s 640x640	2.4	37.0	14	63.1	5.2
YOLO v5s 864x864	2.4	37.0	14	64.3	6.1
SSD MobileNet V1 640x640	48	29.1	84	45.0	19
SSD MobileNet V2 640x640	39	28.2	20	37.9	13

Tabela 5.1: Testes para o 4TUDataset de armadilhas adesivas

Modelo	Speed (ms)	COCO AP	Size (mb)	Author AP.50	Author Speed (ms)
YOLO v5m 640x640	3.4	44.3	41	73.5	4.1
YOLO v5s 640x640	2.4	37.0	14	72.6	3.5
SSD MobileNet V2 640x640	39	28.2	20	47.2	6

Tabela 5.2: Testes para o conjunto de dados da Embrapa de imagens de ninfas

Um ponto interessante são os valores próximos de 66.0 e 64.3 para os modelos YOLOv5m, com imagens de entrada em resolução de 640x640, e o modelo YOLOv5s, para o tamanho de 864x864, respectivamente, indicando atingir desempenho similar de AP em menor tempo computacional apenas aumentando, relativamente, o tamanho da imagem. Reafirmando pesquisas antecedentes para o problema de detecção de pequenos

objetos, onde a resolução da imagem continua sendo a principal causa para diminuição dos resultados, tendo em vista a perda das informações que ocorrem ao passar nas camadas de diminuição da resolução. [Unel, Ozkalayci e Cigla 2019] explora o processo de *tiling*, reforçando a capacidade deste método e se mostra uma alternativa valiosa para problemas similares como o abordado nesta pesquisa.

Não obstante, ao analisar os resultados para a rede SSD Mobile Net, é possível observar que a detecção de pequenos objetos continua sendo um obstáculo para a utilização deste modelo, apesar do seu tamanho reduzido ser uma boa alternativa para dispositivos móveis. Portanto, as redes YOLOv5, nas versões *small* e *medium*, foram consideradas as melhores escolhas de arquitetura para este trabalho. Os maiores desempenhos de AP e velocidade de inferência foram obtidos nas versões deste modelo, recomendando a continuidade de futuros trabalhos para esta arquitetura de rede neural. Entre as resoluções para os modelos em destaque anteriormente, o ganho no aumento de resolução ocorre em menor escala, demonstrando a eficácia do processo de *tiling*, porém deixando a possibilidade em aberto da utilização de menores resoluções de entrada para o modelo com o intuito de incrementar o desempenho computacional da solução.

Conclusão e trabalhos futuros

Considerando os estudos efetuados anteriormente na área de visão computacional para detecção de pragas e doenças, especificamente para a mosca-branca, foi possível explorar em profundidade a utilização de modelos de redes neurais convolucionais, apresentando uma nova alternativa de abordagem para este problema. Não obstante, com a evolução de modelos como a YOLO e recentes trabalhos em otimização de desempenho para estas arquiteturas, bem como a chegada de tecnologias como a conexão 5G, a utilização de dispositivos embarcados para a inferência destes modelos em ambientes agrícolas se apresentar como alternativa futura para a integração de inteligência artificial e agricultura.

A partir deste projeto, foi iniciada pela equipe da Emprapa a construção de um dispositivo de armadilha inteligente, que tem como objetivo principal a capacidade de detectar moscas-brancas em diferentes pontos da lavoura, sendo possível levantar dados demográficos e efetuar o manejo de pesticidas de maneira responsável.

Figura 6.1: Raspberry Pi 3B+, um computador de placa única



Fonte (Sparkfun, 2020)

Este computador, como o ilustrado na 6.1, seria responsável pelo controle da câmera, obtendo e armazenando fotos da armadilha em um intervalo de tempo configurável. Não obstante, este dispositivo seria responsável por realizar uma análise

de cada imagem, fornecendo como saída a quantidade de insetos presos na armadilha através de um modelo de rede neural embarcado.

Para efeitos comparativos, seria interessante avaliar o desempenho de novas arquiteturas de redes neurais publicadas recentemente como a *EfficientDET* em suas diferentes versões, bem como detectores de objetos por região como a arquitetura *Faster R-CNN*.

O aumento do conjunto de dados de ninfas em imagens de plantas obtidas em contexto laboratorial se faz necessário, tendo em vista que a baixa quantidade de amostras pode afetar o desempenho da solução em grande escala. Para o *dataset de armadilhas* não foram avaliadas variações de luz e tempo, tendo em vista que variações meteorológicas e a incidência de raios de sol na lavoura é comum, a coleta de imagens neste cenário é necessária para a evolução dos experimentos.

Por fim, uma solução de *software* capaz de realizar inferências e contagem de moscas-brancas em diferentes estágios ninfaís e adultos poderia ser construída, utilizando os modelos explorados neste trabalho, gerando uma aplicação para pesquisadores e agricultores em diferentes contextos agrícolas.

Referências Bibliográficas

- [Agrocares]AGROCARES. *scoutbox | Smart Farming | Nutrient Testing – AgroCares*. <https://www.agrocares.com/products/scoutbox/>. (Accessed on 03/21/2021).
- [Bar et al. 2019]BAR, L. et al. Downregulation of dystrophin expression in pupae of the whitefly *bemisia tabaci* inhibits the emergence of adults. *Insect Mol Biol*, John Wiley & Sons, Ltd (10.1111), v. 0, n. 0, jul. 2019. ISSN 0962-1075. Disponível em: <<https://doi.org/10.1111/imb.12579>>.
- [Barbedo 2014]BARBEDO, J. G. A. Using digital image processing for counting whiteflies on soybean leaves. *Journal of Asia-Pacific Entomology*, v. 17, n. 4, p. 685–694, dez. 2014. ISSN 1226-8615. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1226861514000855>>.
- [Bechar et al. 23-2]BECHAR, I. et al. On-line video recognition and counting of harmful insects. In: *2010 20th International Conference on Pattern Recognition*. [S.l.: s.n.], 23–2. p. 4068–4071. ISSN 1051-4651.
- [Bochkovski, Wang e Liao 2020]BOCHKOVSKIY, A.; WANG, C.-Y.; LIAO, H.-Y. M. *YO-LOv4: Optimal Speed and Accuracy of Object Detection*. 2020.
- [Boissard, Martin e Moisan 2008]BOISSARD, P.; MARTIN, V.; MOISAN, S. A cognitive vision approach to early pest detection in greenhouse crops. *Computers and Electronics in Agriculture*, v. 62, n. 2, p. 81–93, jul. 2008. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169907002256>>.
- [Chen et al. 2018]CHEN, J. J. et al. Automatic segmentation and counting of aphid nymphs on leaves using convolutional neural networks. In: . [S.l.: s.n.], 2018.
- [Cheng, Zhao e Zhang 2019]CHENG, S.; ZHAO, K.; ZHANG, D. Abnormal water quality monitoring based on visual sensing of three-dimensional motion behavior of fish. *Symmetry*, v. 11, p. 1179, 09 2019.
- [Cho et al. 2008]CHO, J. et al. Automatic Identification of Tobacco Whiteflies, Aphids and Thrips in Greenhouse Using Image Processing Techniques. In: Anninos, P and

- Pham, T and Grebennikov, A (Ed.). *MABE'08: PROCEEDINGS OF THE 4TH WSEAS INTERNATIONAL CONFERENCE ON MATHEMATICAL BIOLOGY AND ECOLOGY*. [S.l.: s.n.], 2008. (Mathematics and Computers in Science and Engineering), p. 74+. ISBN 978-960-6766-32-9. ISSN 1792-4308. 4th WSEAS International Conference on Mathematical Biology and Ecology, Acapulco, MEXICO, JAN 25-27, 2008.
- [Ciregan, Meier e Schmidhuber 2012]CIREGAN, D.; MEIER, U.; SCHMIDHUBER, J. Multi-column deep neural networks for image classification. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2012. p. 3642–3649.
- [Espinoza et al. 2016]ESPINOZA, K. et al. Combination of image processing and artificial neural networks as a novel approach for the identification of *bemisia tabaci* and *frankliniella occidentalis* on sticky traps in greenhouse agriculture. *Computers and Electronics in Agriculture*, v. 127, p. 495–505, set. 2016. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169916304823>>.
- [Fuentes et al. 2017]FUENTES, A. et al. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. In: *Sensors*. [S.l.: s.n.], 2017.
- [Gutierrez et al. 2019]GUTIERREZ, A. et al. A benchmarking of learning strategies for pest detection and identification on tomato plants for autonomous scouting robots using internal databases. *Journal of Sensors*, Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University, Brno, Czech Republic, v. 2019, n. 5219471, 2019. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85066093058&doi=10.1155%2f2019%2f5219471&partnerID=40&md5=e756378745ae2105f6358ef46f>>.
- [Harvey et al. 2014]HARVEY, C. et al. Extreme vulnerability of smallholder farmers to agricultural risks and climate change in madagascar. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, v. 369, p. 20130089, 04 2014.
- [Huddar et al. 26-2]HUDDAR, S. R. et al. Novel algorithm for segmentation and automatic identification of pests on plants using image processing. In: *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*. [S.l.: s.n.], 26–2. p. 1–5.
- [Jige e Ratnaparkhe 19-2]JIGE, M. N.; RATNAPARKHE, V. R. Population estimation of whitefly for cotton plant using image processing approach. In: *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. [S.l.: s.n.], 19–2. p. 487–491.
- [Jocher et al. 2020]JOCHER, G. et al. *ultralytics/yolov5: v3.0*. 8 2020. Disponível em: <<https://doi.org/10.5281/zenodo.3983579>>.

- [Kamilaris e Prenafeta-Boldú 2018]KAMILARIS, A.; PRENAFETA-BOLDÚ, F. X. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, v. 147, p. 70–90, abr. 2018. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169917308803>>.
- [Karpathy et al. 2014]Karpathy, A. et al. Large-scale video classification with convolutional neural networks. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2014. p. 1725–1732.
- [Kitchenham e Charters 2007]KITCHENHAM, B.; CHARTERS, S. *Guidelines for performing Systematic Literature Reviews in Software Engineering*. 2007.
- [Koutník, Schmidhuber e Gomez 2014]KOUTNÍK, J.; SCHMIDHUBER, J.; GOMEZ, F. On-line evolution of deep convolutional network for vision-based reinforcement learning. In: . [S.l.: s.n.], 2014. p. 260–269.
- [Lecun et al. 1998]LECUN, Y. et al. Gradient-based learning applied to document recognition. In: *Proceedings of the IEEE*. [S.l.: s.n.], 1998. p. 2278–2324.
- [LeCun e Cortes 2010]LECUN, Y.; CORTES, C. MNIST handwritten digit database. 2010. Disponível em: <<http://yann.lecun.com/exdb/mnist/>>.
- [Li, Xia e Lee 2015]LI, Y.; XIA, C.; LEE, J. Detection of small-sized insect pest in greenhouses based on multifractal analysis. *Optik - International Journal for Light and Electron Optics*, v. 126, n. 19, p. 2138–2143, out. 2015. ISSN 0030-4026. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S003040261500409X>>.
- [Lino et al. 18-2]LINO, A. F. S. et al. Performance of haar and lbp features in cascade classifiers to whiteflies detection and counting. In: *2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*. [S.l.: s.n.], 18–2. p. 1–6.
- [Liu et al. 2016]LIU, W. et al. Ssd: Single shot multibox detector. *Lecture Notes in Computer Science*, Springer International Publishing, p. 21–37, 2016. ISSN 1611-3349. Disponível em: <http://dx.doi.org/10.1007/978-3-319-46448-0_2>.
- [Maharlooei et al. 2017]MAHARLOOEI, M. et al. Detection of soybean aphids in a greenhouse using an image processing technique. *Computers and Electronics in Agriculture*, v. 132, p. 63–70, jan. 2017. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169916310791>>.

[Mallett et al. 2012]MALLETT, R. et al. The benefits and challenges of using systematic reviews in international development research. *Journal of Development Effectiveness*, Routledge, v. 4, n. 3, p. 445–455, set. 2012. ISSN 1943-9342. Disponível em: <<https://doi.org/10.1080/19439342.2012.711342>>.

[Meyer et al. 2014]MEYER, J. de et al. The state of food and agriculture innovation in family farming. In: _____. [S.l.: s.n.], 2014. p. 138. ISBN 978-92-5-108537-0.

[Mohanty, Hughes e Salathé 2016]MOHANTY, S. P.; HUGHES, D. P.; SALATHÉ, M. Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, v. 7, p. 1419, 2016. ISSN 1664-462X. Disponível em: <<https://www.frontiersin.org/article/10.3389/fpls.2016.01419>>.

[Nieuwenhuizen et al. 2019]NIEUWENHUIZEN, A. A. et al. Raw data from Yellow Sticky Traps with insects for training of deep learning Convolutional Neural Network for object detection. 3 2019. Disponível em: <https://data.4tu.nl/articles/dataset/Raw_data_from_Yellow_Sticky_Traps_with_insects_for_training_of>.

[Pan e Yang 2010]PAN, S.; YANG, Q. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, v. 22, n. 10, p. 1345–1359, 2010.

[Qiao et al. 2008]QIAO, M. et al. Density estimation of bemisia tabaci (hemiptera: Aleyrodidae) in a greenhouse using sticky traps in conjunction with an image processing system. *Journal of Asia-Pacific Entomology*, v. 11, n. 1, p. 25–29, mar. 2008. ISSN 1226-8615. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1226861508000071>>.

[Rawat e Wang 2017]RAWAT, W.; WANG, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.*, MIT Press, Cambridge, MA, USA, v. 29, n. 9, p. 2352–2449, set. 2017. ISSN 0899-7667. Disponível em: <https://doi.org/10.1162/neco_a_00990>.

[Redmon et al. 2016]REDMON, J. et al. You only look once: Unified, real-time object detection. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.]: IEEE, 2016.

[Rumelhart, Hinton e Williams 1986]RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning internal representations by error propagation. In: _____. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*. Cambridge, MA, USA: MIT Press, 1986. p. 318–362. ISBN 026268053X.

[Saxena e Armstrong 2014]SAXENA, L. P.; ARMSTRONG, L. A survey of image processing techniques for agriculture. In: . [S.l.: s.n.], 2014.

[Schwartzman et al. 2016]SCHWARTZMAN, A. et al. Image processing, computer vision, and deep learning: new approaches to the analysis and physics interpretation of LHC events. *Journal of Physics: Conference Series*, IOP Publishing, v. 762, p. 012035, out. 2016. ISSN 1742-6596. Disponível em: <<http://dx.doi.org/10.1088/1742-6596/762/1/012035>>.

[Sharma e Mir 2019]SHARMA, V.; MIR, R. N. Saliency guided faster-rcnn (sgfr-rcnn) model for object detection and recognition. *Journal of King Saud University - Computer and Information Sciences*, 2019. ISSN 1319-1578. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1319157819306639>>.

[Solawetz 2020]SOLAWETZ, J. *Tackling the Small Object Problem in Object Detection*. 08 2020. <https://blog.roboflow.com/detect-small-objects/>. (Accessed on 03/24/2021).

[Solis-Sánchez et al. 2011]SOLIS-SÁNCHEZ, L. O. et al. Scale invariant feature approach for insect monitoring. *Computers and Electronics in Agriculture*, v. 75, n. 1, p. 92–99, jan. 2011. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S016816991000195X>>.

[Sun et al. 2017]SUN, M. et al. Learning pooling for convolutional neural network. *Neurocomputing*, v. 224, p. 96–104, 2017. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231216312905>>.

[Sun et al. 2017]SUN, Y. et al. A smart-vision algorithm for counting whiteflies and thrips on sticky traps using two-dimensional fourier transform spectrum. *Biosystems Engineering*, v. 153, p. 82–88, jan. 2017. ISSN 1537-5110. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1537511016303439>>.

[Unel, Ozkalayci e Cigla 2019]UNEL, F. O.; OZKALAYCI, B. O.; CIGLA, C. The power of tiling for small object detection. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. [S.l.]: IEEE, 2019.

[Vibhute e Bodhe 2012]VIBHUTE, A.; BODHE, S. K. Article: Applications of image processing in agriculture: A survey. *International Journal of Computer Applications*, v. 52, n. 2, p. 34–40, August 2012. Full text available.

[Wang et al. 2018]WANG, Z. et al. A cognitive vision method for insect pest image segmentation. *IFAC-PapersOnLine*, v. 51, n. 17, p. 85–89, jan. 2018. ISSN 2405-8963. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S2405896318311820>>.

[Wendland 2002]WENDLAND, A. *Vírus do mosaico comum do feijoeiro*. 2002. Disponível em: <https://www.agencia.cnptia.embrapa.br/gestor/feijao/arvore/CONTAG01_105_1311200215105.h>.

[Wolfert et al. 2017]WOLFERT, S. et al. Big data in smart farming – a review. *Agricultural Systems*, v. 153, p. 69 – 80, 2017. ISSN 0308-521X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0308521X16303754>>.

[Xia et al. 2015]XIA, C. et al. Automatic identification and counting of small size pests in greenhouse conditions with low computational cost. *Ecological Informatics*, v. 29, p. 139–146, set. 2015. ISSN 1574-9541. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1574954114001228>>.

Publicações

Com o intuito de validar os estudos levantados até o presente momento desta pesquisa, bem como reforçar o compromisso com a produção científica no ambiente acadêmico, foi submetido um artigo, em formato de revisão sistemática, para o Congresso Brasileiro de Agroinformática de 2019 (SBIAgro 2019), principal evento científico da área de informática aplicada à agricultura no Brasil, promovido pela Associação Brasileira de Agroinformática (SBIAgro 2019). O artigo intitulado como *Computer Vision and Image Processing for Detecting and Quantifying Whiteflies: A Systematic Review*", foi aceito e apresentado em novembro de 2019 pelo aluno João Gabriel Junqueira, e encontra-se publicados nos anais do congresso. Através da publicação e apresentação, o artigo pode ser compartilhado e avaliado por pesquisadores especializados em agroinformática, demonstrando a relevância acadêmica para a pesquisa efetuada.



Computer Vision and Image Processing for Detecting and Quantifying whiteflies: A Systematic Review

João Gabriel Junqueira da Silva¹, Marcelo Gonçalves Narciso², Crithiane Gonçalves¹

¹Instituto de Informática, Universidade Federal de Goiás
Goiânia, Goiás, Brasil
joaogabrieljunq@gmail.com, crithianega@gmail.com

²Embrapa Arroz e Feijão
Santo Antônio de Goiás, Goiás, Brasil
marcelo.narciso@embrapa.br

RESUMO

A evolução da agricultura de precisão impulsionou o aumento da produção agrícola e proporcionou o surgimento de grandes desafios em termos de eficiência, segurança alimentar, sustentabilidade e impacto ambiental. Especificamente, o controle de pragas é de suma importância no que tange a saúde das plantações. Este artigo apresenta uma revisão sistemática da literatura, abordando técnicas e métodos de processamento digital de imagens em plantas para a detecção e classificação de pragas, em específico para a praga Mosca Branca. Nesta revisão foram considerados artigos publicados entre 2008 e 2018 indexados por cinco bases de dados científicas. A revisão identificou que as técnicas são bem exploradas, diversificadas e possuem alto desempenho em termos de acurácia. No entanto há uma dificuldade no que tange a detecção de moscas em estágio ninfal, e principalmente técnicas aplicadas a campo aberto e dispositivos portáteis, o que daria maior autonomia à agricultores e cientistas. Ideias futuras são discutidas no final do artigo com base nas dificuldades relatadas e trabalhos futuros propostos.

PALAVRAS-CHAVE: Visão computacional, classificação de imagens, agricultura de precisão, mosca-branca.

ABSTRACT

The evolution of precision agriculture increased agricultural production and provided major challenges in terms of efficiency, food safety, sustainability and environmental impact. Specifically, the pest management is extremely important healthy crop. This paper presents a systematic review of literature, addressing techniques and methods of digital image processing in plants for the detection and classification of pests, specifically for whitefly pest. In this review,

articles published between 2008 and 2018 indexed by five scientific databases were considered. The review identified that the techniques are well explored and have high performance in terms of accuracy. However, there is a problem with the detection of whiteflies in the nymphal stage, mainly techniques applied to the open field and portable devices, which would give greater autonomy to farmers and scientists. Future ideas are discussed at the end of the article based on the difficulties reported and future works were proposed.

KEYWORDS: Computer vision, image classification, precision agriculture, whiteflies.

INTRODUCTION

Precision agriculture can be considered as an exponent of general future development through the Internet of Things (IoT) and Big Data (WOLFERT et al., 2017). Modern technologies have given human society the ability to produce enough supplies to feed the demand of more than 7 billion people (MOHANTY; HUGHES; SALATHÉ, 2016). However, with the increase in production capacity, great challenges have arisen in terms of efficiency, food safety, sustainability and environmental impact. To address these challenges, the complex, multivariate and unpredictable agricultural ecosystems need to be better understood by monitoring, measuring and analyzing continuously various physical aspects and phenomena (KAMILARIS; PRENAFETA-BOLDÚ, 2018).

Image processing techniques (IPTs) can be used to enhance agricultural practices, by improving accuracy and consistency of processes while reducing farmers manual monitoring. Often, it offers flexibility and effectively substitutes the farmers visual decision making (SAXENA; ARMSTRONG, 2014). During the last several years, spectacular advances in the fields of artificial intelligence, computer vision, and deep learning have resulted in remarkable performance in image classification and vision tasks. (SCHWARTZMAN et al., 2016).

Linking this need to the emergence of new techniques, this paper aims to explore methods and tools of IPTs through a systematic review of the bibliography, with the objective of presenting the state of the art in the use of IPTs to identify the whitefly pest in images of plants.

WHITEFLY

The whitefly (*Bemisia tabaci*) is a major pest to agriculture. Adults are able to fly for long distances and to colonize staple crops, herbs and ornamentals, and to vector viruses belonging to several important taxonomic groups (BAR et al., 2019). It affects several cultures like pumpkin, zucchini, cotton, eggplant, broccoli, cauliflower, peas, beans, papaya, cassava, watermelon, melon, cucumber, pepper, cabbage, soybean and many others. The easiest way to control the pest infection is the use of pesticides. But the more use of pesticide are hazardous and it not only kill pest in plant, but also affect the health of human, animal and plant. To overcome this problem it is necessary to control the use of pesticide, pest detection is the most important process for an effective cultivation (JIGE; RATNAPARKHE, 19-2).

Figure 1: Leaf infested with whiteflies



Fonte (Embrapa, 2019)

Barbedo (2014) explains the most direct way of measuring whitefly infestation is to manually identify and count the insects inside a selected region. In general, this approach does not require sophisticated apparatus and, more importantly, relies on the remarkable human ability to resolve ambiguities and unclear situations, even under far from ideal conditions. On the other hand, human beings are susceptible to physiological and psychological phenomena that may be important sources of error: fatigue, visual illusions, boredom, among others. Also, humans are usually much slower than machines in performing simple tasks like counting.

Image processing has been proved to be effective tool for analysis in various fields and applications (VIBHUTE; , 2012). In this paper we explore some of these techniques to detecting and counting the whitefly pest.

MATERIAL AND METHODS

In this research, we performed a Systematic Review of the Literature (SRL), since it offers a organized and well-defined process. Mallett et al. (2012) defines SRL as a rigorous and transparency form of literature review. Three main stages were taken following the SRL guideline defined by Kitchenham e Charters (2007): (I) defining a review protocol; (II) conducting the review of the studies; and (III) reporting the review.

The review protocol definition include the definition of the research questions, keywords and search strategy, inclusion and exclusion criteria for the selection of studies and a checklist for study quality assessment. These elements is detailed next. To support this process and provide a collaborative way between the authors a software platform were used: Parsifal (<https://parsif.al/>).

Research questions

The main objective of this SRL was to answer the following question: “What is the state of the art in the application of digital image processing techniques for acquisition, detection and counting of Whitefly pest in images?”.

Some more specific questions that unfolded the previous one were formulated: (RQ.1) What are the existing methods for acquisition and preprocessing images for Whitefly pest? (RQ.2) What are the techniques used for detection and classification of pests in plants, specifically for the Whitefly pest? (RQ.3) What are the results obtained so far for detection and classification for the white fly pest? (RQ.4) What measures are used to evaluate the effectiveness of existing methods?

Keywords, synonyms and search strategy

The following keywords were chosen for driving the SRL, considering the usually techniques terms and pest scientific and popular name: (K.1) computer vision; and (K.2) deep learning; (K.3) neural networks; (K.4) image processing; (K.5) image classification; (K.6) plague detection; (K.7) whiteflies; (K.8) *bemisia tabaci*; (K.9) plant disease.

The general logic search string was defined as ("Computer vision" OR "Deep learning" OR "Neural networks" OR "Image processing" OR "Image classification") AND ("Plague detection" OR "Whiteflies" OR "*Bemisia tabaci*" OR "Plant disease").

This string was adjusted for every electronic database in order to meet its search syntax. We searched on five databases: (1) ACM Digital Library; (2) IEEE Digital Library; (3) ISI Web of Science; (4) Science Direct; and (5) Scopus, since they gather most of the publications on the area of computer science and applications.

Inclusion/Exclusion criteria

Inclusion and exclusion criteria were defined for guiding the selection of primary studies relevant to the SRL. A study was included if it was: (1) available in full-text; (2) published in a journal or in the annals of a conference; (3) a technical report, including surveys; or (4) a master dissertation or a doctoral thesis. On the other hand, it was excluded if it: (1) was not written in English; (2) was published before 2008; (3) present evaluations without describing the used method; (4) do not use digital image processing techniques; (5) do not use digital image processing techniques for the whitefly pest detection.

A set of Yes/No/Partially quality assessment questions was also defined based on a refinement of the main research questions (RQs). For instance, they asked if the classification methods or techniques have been objectively reported or if there is an application or final product, has it been described in detail.

Conducting the review

The proper search strings, with the corrected search conditions, were applied to the databases. Next, as a Study Selection step, the bibliographic references were evaluated according to the inclusion and the exclusion criteria (with the help of the SRL software platforms). When necessary, the full text of the studies was consulted. Table 1 summarizes the number of papers

recovered when searching in each database, and the amount of accepted (selected), reject and duplicate studies during the Selection step. In Total, 25 studies were accepted.

Table 1: Number of studies in the SRL.

Database	Total	Accepted	Rejected	Duplicated
ACM DL	103	4	82	17
IEEE DL	94	3	89	2
WebOfScience	129	5	69	55
ScienceDir	81	9	64	8
Scopus	31	4	12	15

The studies were included according to the inclusion and exclusion criteria, and in addition, the following quality criteria were applied: (1) Have classification methods or techniques been objectively reported?; (2) Were the methods of acquisition and preprocessing explicitly cited?; (3) Was the technique used compared to other existing techniques?; (4) Is there an evaluation of this technique?; (5) If there is an application or final product, has it been described in detail?.

These criteria were elaborated based on the research questions and for each criterion a value was assigned in conjunction with a concept: Yes (S) = 1 point, No (N) = 0 and Partially (P) = 0.5, where the junction of all the criteria form an index that was classified as follows, works with index between 0 and 1 were considered very bad, indexes between 1.5 and 2 were considered bad, works between 2.5 and 3 were considered good, between 3.5 and 4 work was considered very good, and the works that obtained between 4.5 and 5 were considered excellent. The value of the index indicated the priority for reading the articles and was also a factor for the exclusion of articles if the work has a very low index and there is no specific contribution, which would justify the low index. After the quality assessment, 21 studies were selected for full read.

Next, we describe the main concepts in some of the studies selected by the SRL.

RESULTS AND DISCUSSION

Barbedo (2014) explains there are two main strategies for automatically counting whiteflies found in the literature, one using sticky traps, and the other using plant leaves directly. In sticky traps images, only specimens capable of flying are considered, in leaves images early stages like nymphs can be considered.

Sticky traps images

Sticky traps consist of glue cards where the specimens are attracted by the color of the paper and are caught by the sticky substance, the smooth and neutral surface make the counting easier.

Cho et al. (2008) proposed a method to explore size and color features to identify and count whiteflies, aphids and thrips, using color transformations, simple mathematical morphology operations, and thresholding. Boissard, Martin e Moisan (2008) proposed a cognitive vision system that combines image processing, neural learning and knowledge-based techniques to detect and count pests. Qiao et al. (2008) described a image processing system to estimate the population of adult specimens of whiteflies collected on sticky traps based on size and color. Bechar et al. (2012) uses techniques like Mixture-of-Gaussians and Principal Component Analysis to extract small spots present on the scene and detect whiteflies in sticky traps in a real-time system. Solis-Sánchez et al. (2011) developed a new algorithm capable to detect and identify five species, including whitefly, using geometrical characteristics of insects that affect greenhouse. Barbedo (2014) presented a new method for counting whiteflies on soybean leaves based on conventional digital image, the algorithm is capable of detecting and quantifying not only adult whiteflies, but also nymphal specimens using region-of-interest (ROI) delineation, application of color transformations, threshold-based segmentation, and detection of young nymphs. Xia et al. (2015) used watershed algorithm and Mahalanobis distance on YCrCb color space, to propose a pest identification procedure, for low resolution images in sticky traps using portable devices. Espinoza et al. (2016) proposed a novel approach for the detection and monitoring of adult-stage whitefly (*Bemisia tabaci*) and thrip (*Frankliniella occidentalis*) in greenhouses based on the combination of an image-processing algorithm and artificial neural networks and described a new range of possibilities with the improvement of neural networks in the last years. Sun et al. (2017) proposed a technique rather than directly counting the pests captured on the traps, the concept is to treat trapped pests as noise in a two-dimensional (2D) image with 2DFT serving as a specific noise collector.

Plant leaves images

This approach consists on take pictures directly on leaves surface in greenhouse or open fields, the main difficulty is that leaves imperfections and open field climatic disturbances make counting harder.

Huddar et al. (26-2) proposed a method capable of counting insects not only in greenhouses, but also in open farms, the algorithm has four steps: color conversion, segmentation based on the relative difference in pixel intensities, noise reduction by erosion, and counting based on Moore neighbor tracing and Jacob's stopping. Li, Xia e Lee (2015) proposed new application of multifractal analysis for the detection of small-sized pests (e.g., whitefly) from leaf surface images *in situ*. Multifractal analysis was adopted for segmentation of whitefly images based on the local singularity and global image characters with the regional minima selection strategy. Maharlooei et al. (2017) presented the use of digital imaging technique to detect and count soybean aphids on soybean leaves. Lino et al. (18-2) made a comparison between the performances of LBP and Haar features in a Cascade Classifier, which aims automatically to count immature whiteflies (nymphs). Jige e Ratnaparkhe (19-2) presented an algorithm for automatic detection and counting of whiteflies in images of cotton plantations, the algorithm uses color space conversion, background subtraction, threshold operation, morphological operation and counting of the whitefly number on the leaf with centroid algorithm and region properties. Fuentes et al. (2017) proposed a detector based on neural networks for real-time tomato diseases and pest recognition, including the whitefly pest in tomato plant images. The objective was to find the neural network architecture best suited for the task and three types of detectors were considered: one using Faster Region-based Convolutional Neural Network (Faster R-CNN), another using Region-based Fully Convolutional Network (R-FCN), and the latest Single Shot Multibox Detector (SSD). Chen et al. (2018) presented an image processing method using CNN to segment and count aphid nymphs on leaves. Wang et al. (2018) proposed a cognitive segmentation approach to pest images. The method works in the follow way. First, a pest image is divided into blocks via an image block processing method. Second, an adaptive learning algorithm is used to accurately select the initial cluster centers. Third, preliminary segmentation results are achieved using K-means clustering. Finally, three digital morphological features of an ellipse are adopted to remove leaf veins. Gutierrez et al. (2019) presented the development and comparison of two different approaches for vision based automated pest detection and identification, using learning strategies. A solution that combines computer vision and machine learning is compared against a deep learning solution. The main focus of his paper is on the selection of the best approach based on pest detection and identification accuracy.

CONCLUSIONS

Based on the SRL, we found that in the last ten years a few studies was conducted on whitefly pest detection, average two papers per year are published for this purpose. In the sticky traps approach the main difficulty related are detecting whiteflies caught at the edge of the trap, where it is relatively difficult for the image processing system to detect them. In the leaves images, the most related difficulty is detecting early stages nymphs and classify the different stages.

As future work, it is suggested the use of nymphs detection techniques used in other

species, such as aphids for the detection of whiteflies nymphs and could be interesting to see how these techniques could be improve in open field in images acquired in mobile devices. An automatic solution for use by farmers and field researchers could be proposed, deep learning techniques has become popular and is also an alternative in embedded technologies. There is no open dataset available and those reported in papers, show the scarcity of images and data for this field.

Despite these challenges, the field presents promising results and has followed the evolution of digital image processing techniques in agriculture.

ACKNOWLEDGEMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

REFERENCES

BAR, L. et al. Downregulation of dystrophin expression in pupae of the whitefly *bemisia tabaci* inhibits the emergence of adults. *Insect Mol Biol*, John Wiley & Sons, Ltd (10.1111), v. 0, n. 0, jul. 2019. ISSN 0962-1075. Disponível em: <<https://doi.org/10.1111/imb.12579>>.

BARBEDO, J. G. A. Using digital image processing for counting whiteflies on soybean leaves. *Journal of Asia-Pacific Entomology*, v. 17, n. 4, p. 685–694, dez. 2014. ISSN 1226-8615. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1226861514000855>>.

BECHAR, I. et al. On-line video recognition and counting of harmful insects. In: *2010 20th International Conference on Pattern Recognition*. [S.l.: s.n.], 23–2. p. 4068–4071. ISSN 1051-4651.

BOISSARD, P.; MARTIN, V.; MOISAN, S. A cognitive vision approach to early pest detection in greenhouse crops. *Computers and Electronics in Agriculture*, v. 62, n. 2, p. 81–93, jul. 2008. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169907002256>>.

CHEN, J. J. et al. Automatic segmentation and counting of aphid nymphs on leaves using convolutional neural networks. In: . [S.l.: s.n.], 2018.

CHO, J. et al. Automatic Identification of Tobacco Whiteflies, Aphids and Thrips in Greenhouse Using Image Processing Techniques. In: Anninos, P and Pham, T and Grebennikov, A (Ed.). *MABE'08: PROCEEDINGS OF THE 4TH WSEAS INTERNATIONAL CONFERENCE ON MATHEMATICAL BIOLOGY AND ECOLOGY*. [S.l.: s.n.], 2008. (Mathematics and Computers in Science and Engineering), p. 74+. ISBN 978-960-6766-32-9. ISSN 1792-4308. 4th WSEAS International Conference on Mathematical Biology and Ecology, Acapulco, MEXICO, JAN 25-27, 2008.

ESPINOZA, K. et al. Combination of image processing and artificial neural networks as a novel approach for the identification of *bemisia tabaci* and *frankliniella occidentalis* on sticky traps in greenhouse agriculture. *Computers and Electronics in Agriculture*, v. 127, p. 495–505, set. 2016. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169916304823>>.

FUENTES, A. et al. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. In: *Sensors*. [S.l.: s.n.], 2017.

GUTIERREZ, A. et al. A benchmarking of learning strategies for pest detection and identification on tomato plants for autonomous scouting robots using internal databases. *Journal of Sensors*, Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University, Brno, Czech Republic, v. 2019, n. 5219471, 2019. Disponível em: <<https://www.hindawi.com/journals/js/2019/5219471/>>.

HUDDAR, S. R. et al. Novel algorithm for segmentation and automatic identification of pests on plants using image processing. In: *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*. [S.l.: s.n.], 26–2. p. 1–5.

JIGE, M. N.; RATNAPARKHE, V. R. Population estimation of whitefly for cotton plant using image processing approach. In: *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. [S.l.: s.n.], 19–2. p. 487–491.

KAMILARIS, A.; PRENAFETA-BOLDÚ, F. X. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, v. 147, p. 70–90, abr. 2018. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169917308803>>.

KITCHENHAM, B.; CHARTERS, S. *Guidelines for performing Systematic Literature Reviews in Software Engineering*. 2007.

LI, Y.; XIA, C.; LEE, J. Detection of small-sized insect pest in greenhouses based on multifractal analysis. *Optik - International Journal for Light and Electron Optics*, v. 126, n. 19, p. 2138–2143, out. 2015. ISSN 0030-4026. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S003040261500409X>>.

LINO, A. F. S. et al. Performance of haar and lbp features in cascade classifiers to whiteflies detection and counting. In: *2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*. [S.l.: s.n.], 18–2. p. 1–6.

MAHARLOOEI, M. et al. Detection of soybean aphids in a greenhouse using an image processing technique. *Computers and Electronics in Agriculture*, v. 132, p. 63–70, jan. 2017. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169916310791>>.

- MALLET, R. et al. The benefits and challenges of using systematic reviews in international development research. *Journal of Development Effectiveness*, Routledge, v. 4, n. 3, p. 445–455, set. 2012. ISSN 1943-9342. Disponível em: <<https://doi.org/10.1080/19439342.2012.711342>>.
- MOHANTY, S. P.; HUGHES, D. P.; SALATHÉ, M. Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, v. 7, p. 1419, 2016. ISSN 1664-462X. Disponível em: <<https://www.frontiersin.org/article/10.3389/fpls.2016.01419>>.
- QIAO, M. et al. Density estimation of *bemisia tabaci* (hemiptera: Aleyrodidae) in a greenhouse using sticky traps in conjunction with an image processing system. *Journal of Asia-Pacific Entomology*, v. 11, n. 1, p. 25–29, mar. 2008. ISSN 1226-8615. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1226861508000071>>.
- SAXENA, L. P.; ARMSTRONG, L. A survey of image processing techniques for agriculture. In: . [S.l.: s.n.], 2014.
- SCHWARTZMAN, A. et al. Image processing, computer vision, and deep learning: new approaches to the analysis and physics interpretation of LHC events. *Journal of Physics: Conference Series*, IOP Publishing, v. 762, p. 012035, out. 2016. ISSN 1742-6596. Disponível em: <<http://dx.doi.org/10.1088/1742-6596/762/1/012035>>.
- SOLIS-SÁNCHEZ, L. O. et al. Scale invariant feature approach for insect monitoring. *Computers and Electronics in Agriculture*, v. 75, n. 1, p. 92–99, jan. 2011. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S016816991000195X>>.
- SUN, Y. et al. A smart-vision algorithm for counting whiteflies and thrips on sticky traps using two-dimensional Fourier transform spectrum. *Biosystems Engineering*, v. 153, p. 82–88, jan. 2017. ISSN 1537-5110. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1537511016303439>>.
- VIBHUTE, A.; , S. K. B. . *Application of Image Processing in Agriculture: A Survey*. [S.l.: s.n.], 2012. 34–40 p.
- WANG, Z. et al. A cognitive vision method for insect pest image segmentation. *IFAC-PapersOnLine*, v. 51, n. 17, p. 85–89, jan. 2018. ISSN 2405-8963. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S2405896318311820>>.
- WOLFERT, S. et al. Big data in smart farming - a review. *Agricultural Systems*, v. 153, p. 69–80, maio 2017. ISSN 0308-521X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0308521X16303754>>.
- XIA, C. et al. Automatic identification and counting of small size pests in greenhouse conditions with low computational cost. *Ecological Informatics*, v. 29, p. 139–146, set. 2015. ISSN 1574-9541. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1574954114001228>>.