

INSTITUTO FEDERAL DE EDUCAÇÃO CIÊNCIA E TECNOLOGIA DE MINAS
GERAIS - *CAMPUS* BETIM
BACHARELADO EM ENGENHARIA DE CONTROLE E AUTOMAÇÃO

Rodrigo Rafael de Sousa

**REDES NEURAIS EM PROCESSAMENTO DE IMAGENS PARA DETECÇÃO DE
MARIPOSAS DA BROCA DA CANA-DE-AÇÚCAR EM ARMADILHAS**

Betim
2023

RODRIGO RAFAEL DE SOUSA

**REDES NEURAIS EM PROCESSAMENTO DE IMAGENS PARA DETECÇÃO DE
MARIPOSAS DA BROCA DA CANA-DE-AÇÚCAR EM ARMADILHAS**

Trabalho de Conclusão de Curso apresentado à banca examinadora do curso de Engenharia de Controle e Automação do Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais Campus Betim, como parte dos requisitos para obtenção do título de Bacharel em Engenharia de Controle e Automação.

Orientador: Virgil Del Duca Almeida

Betim

2023

FICHA CATALOGRÁFICA

S725r Sousa, Rodrigo Rafael de
Redes neurais em processamento de imagens para detecção de mariposas da broca da cana-de-açúcar em armadilhas / Rodrigo Rafael de Sousa. – 2023.
44 f. : il.

Trabalho de conclusão de curso (Bacharelado em Engenharia de Controle e Automação) - Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais, Câmpus Betim, 2023.

Orientador: Prof. Me. Virgil Del Duca Almeida

1. Redes neurais (Computação). 2. Processamento de imagens. 3. Mariposa. 4. Automação. I. Rodrigo Rafael de Sousa. II. Título.

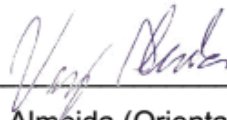
CDU: 681.5

Rodrigo Rafael de Sousa

**REDES NEURAIS EM PROCESSAMENTO DE IMAGENS PARA DETECÇÃO DE
MARIPOSAS DA BROCA DA CANA-DE-AÇÚCAR EM ARMADILHAS**

Trabalho de Conclusão de Curso apresentado à banca examinadora do curso de Engenharia de Controle e Automação do Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais *Campus* Betim, como parte dos requisitos para obtenção do título de Bacharel em Engenharia de Controle e Automação.

Aprovado em: 30/06/2023 pela banca examinadora:



Prof. Me. Virgil Del Duca Almeida (Orientador) – IFMG Campus Betim



Documento assinado digitalmente
LEANDRO FREITAS DE ABREU
Data: 11/07/2023 19:40:11-0300
Verifique em <https://validar.itl.gov.br>

Prof. Dr. Leandro Freitas de Abreu– IFMG Campus Betim



Documento assinado digitalmente
FELIPE AUGUSTO ROCHA DA SILVA
Data: 11/07/2023 23:52:33-0300
Verifique em <https://validar.itl.gov.br>

Prof. Me. Felipe Augusto Rocha da Silva– IFMG Campus Betim

RESUMO

Redes neurais para processamento de imagens é um campo de estudos com crescentes aplicações para a sociedade em geral, bem como em indústrias e no agronegócio.

Existem diversos desafios no agronegócio cuja utilização de técnicas de processamento de imagens pode contribuir para a solução parcial ou completa dos problemas, sendo um destes desafios o controle de pragas.

O objetivo central do trabalho é desenvolver um modelo de rede neural capaz de detectar mariposas presentes em armadilhas instaladas em fazendas de cana de açúcar. Foi utilizado um dataset com imagens de mariposas em armadilhas para treinar a rede, após isso os dados de desempenho da rede foram consolidados para viabilizar a análise do seu desempenho na detecção destes objetos.

A solução desenvolvida alcançou um percentil de detecção de 88%, além de mapear estratégias para tornar a rede neural ainda mais robusta e confiável, o que mostra a viabilidade da utilização de técnicas de processamento de imagens com redes neurais para este tipo de aplicação.

Palavras-chave: Rede neural; Processamento de imagens; Mariposas.

ABSTRACT

Neural networks for image processing is a field of study with increasing applications for society in general, as well as industries and agribusiness.

There are several challenges in agribusiness whose use of image processing techniques can contribute to the partial or complete solution of problems, one of these challenges being pest control.

The main objective of this work is to develop a neural network model capable of detecting moths present in traps installed in sugar cane farms. A dataset with images of moths in traps was used to train the network, after which the network performance data were consolidated to enable the analysis of its performance in detecting these objects.

The developed solution achieved a detection percentile of 88%, in addition to mapping strategies to make the neural network even more robust and reliable, which shows the viability of using image processing techniques with neural networks for this type of application.

Keywords: Neural network; Image processing; Moths.

LISTA DE ILUSTRAÇÕES

Figura 1 – Danos da broca da cana-de-açúcar.	16
Figura 2 – Armadilha de mariposas.	17
Figura 3 - Armadilha de mariposas com filtro de escala de cinza.	20
Figura 4 - Operador Sobel (fiveko, 2023).	20
Figura 5 – Filtros convolucionais (Rodrigues, 2022).	21
Figura 6 - Modelo matemático de um perceptron (Ferneda, 2006).	22
Figura 7 – Rede neural (Oliveira A. C., 2023).	23
Figura 8 – Filtro convolucional (Gonçalves, 2023).	24
Figura 9 - Rede neural convolucional (Sakurai, 2023).	25
Figura 10 – Detecção de objetos.	27
Figura 11 – Estratégia YOLO (Darknet, 2023).	29
Figura 12 - Cvat classificação de imagens.	32
Figura 13 - Cvat Objetos mapeados.	33
Figura 14 – Gráfico de precisão e função perda por épocas.	35
Figura 15 – YOLO v4 tiny.	37
Figura 16 - YOLO v7 tiny.	38
Figura 17 - YOLO v7 tiny customizada.	39
Figura 18 - YOLO v4 tiny customizada.	40
Figura 19- interface gráfica.	41
Figura 20 - Exemplo de detecção.	42
Figura 21 - Problemas de detecção.	43

LISTA DE TABELAS

Tabela 1 – Recursos.	31
---------------------------	----

LISTA DE ABREVIATURAS E SIGLAS

IFMG - Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais.

TCC - Trabalho de Conclusão de Curso.

YOLO – *You Only Look Once*.

CNNs - Convolutional Neural Network.

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Justificativa	15
1.2	Colocação do problema	16
1.2.1	<i>Diatraea saccharalis</i>	16
1.2.2	Monitoramento e controle da <i>Diatraea saccharalis</i>	17
1.2.3	Problemas no monitoramento	18
1.3	Objetivos	18
1.3.1	Objetivos específicos	18
1.4	Organização do trabalho	18
2	FUNDAMENTAÇÃO TEÓRICA	19
2.1	Processamento de imagens digitais por redes neurais	19
2.1.1	Imagem digital	19
2.1.2	Filtros	19
2.2	Redes Neurais Artificiais	21
2.2.1	Perceptron	21
2.2.2	Rede neural	22
2.2.2.1	<i>Redes Neurais Convolucionais</i>	23
2.2.3	Aprendizado de máquina	25
2.3	Arquitetura selecionada	26
2.3.1	Darknet	26
2.3.2	Detecção de objetos na rede neural artificial	27
2.3.3	YOLO	27
2.3.3.1	<i>Configuração e treinamento de uma rede YOLO</i>	29
2.3.3.2	<i>Modelos pré-treinados</i>	30
2.3.4	Dataset	30
3	METODOLOGIA	31
3.1	Recursos	31
3.2	Desenvolvimento	31
3.3	Coleta de imagens	31
3.4	Classificação	32
3.5	Configuração do projeto	34
3.6	Treinamento dos modelos	35
3.7	Critérios de avaliação	36
4	RESULTADOS E DISCUSSÃO	37
4.1	Desempenho dos modelos	37
4.1.1	Refinamento dos modelos	39
4.2	Modelo selecionado e seu desempenho	40
4.3	Interface gráfica	41
4.4	Dificuldades encontradas no dataset utilizado	43
4.5	Possíveis aplicações da rede	44
4.6	Futuros passos	44
5	CONCLUSÃO	45
6	BIBLIOGRAFIA	ERRO! INDICADOR NÃO DEFINIDO.

1 INTRODUÇÃO

No Brasil, o agronegócio é um dos setores mais relevantes da economia, tendo sua produtividade fortemente relacionada à economia brasileira, em 2021 representou 27,5% do PIB do país (USP-ESALQ, 2022). Dentre os principais desafios do setor para aumentar sua produtividade e garantir a segurança alimentar da população o controle de pragas é um fator determinante para o sucesso da safra.

No campo tecnológico existem diversas soluções que contribuem para o desenvolvimento sustentável e o aumento de produtividade no campo, dentre elas a visão computacional se destaca em diversas aplicações como leitura da qualidade do solo, separação de grãos e detecção de pragas.

O presente trabalho tem como foco a visão computacional para detecção da mariposa da cana-de-açúcar no setor sucroalcooleiro.

1.1 Justificativa

O Brasil é o maior produtor mundial do complexo sucroalcooleiro, exercendo a liderança em todos os segmentos: cana-de-açúcar, açúcar e álcool (CARVALHO e OLIVEIRA, 2006).

Empresas do setor estão constantemente em busca de métodos para aumentar sua produtividade protegendo a cultura da cana de diversas ameaças como pragas e doenças. Neste campo, o monitoramento de pragas se faz de grande importância.

Diversos fatores podem influenciar em melhores ou piores produtividades a cada safra. Dentre esses fatores destaca-se a ação de pragas e doenças. A cultura da cana-de-açúcar forma um ecossistema que abriga numerosas espécies de insetos, sendo que algumas delas, dependendo da época do ano e da região, podem ocasionar sérios prejuízos econômicos. Entre estas espécies destaca-se a *Diatraea saccharalis* popularmente conhecida como broca da cana-de-açúcar, uma das mais importantes pragas desta cultura devido à dimensão dos prejuízos que causa (FERREIRA, 2013, p. 13). O monitoramento contínuo da infestação de pragas é um dos elementos cruciais no sistema de controle de pragas (SUK-JU HONG, 2020).

1.2 Colocação do problema

Este capítulo discorre sobre os impactos causados pela broca da cana-de-açúcar em fazendas de cana e como é feito o monitoramento para o controle desta praga.

1.2.1 *Diatraea saccharalis*

Trata-se de uma praga extremamente nociva à cana-de-açúcar. É considerada a mais importante, devido a sua ampla distribuição geográfica e dos seus elevados prejuízos econômicos, causando redução da produtividade agrícola, perdas na qualidade da matéria-prima e custos de tratamento estimados em cerca de um bilhão de dólares anuais (TAVARES, 2019).

Em seu estágio larval, ela é comumente referida como broca da cana-de-açúcar, uma praga capaz de prejudicar canaviais inteiros se não tomadas às devidas medidas de controle (GARCIA, 2020).



Figura 1 – Danos da broca da cana-de-açúcar.

Conforme mostra a Figura 1, o impacto das brocas de mariposas é extremamente danoso ao cultivo da cana, uma vez que a praga consome o seu interior, o que causa impactos severos na qualidade da cana.

1.2.2 Monitoramento e controle da *Diatraea saccharalis*.

O controle e manejo de pragas é atualmente um dos maiores desafios do setor sucroalcooleiro.

Atualmente um dos processos de monitoramento da praga mais comum consiste na instalação de armadilhas, uma superfície de papel com cola e feromônio que atrai e capturam a *Diatraea saccharalis* nas proximidades de sua instalação, um exemplo deste processo pode ser visto na Figura 2.



Figura 2 – Armadilha de mariposas.

A Figura 2 mostra a *Diatraea saccharalis* em sua fase adulta, nesta fase a mariposa deposita seus ovos na cana, quando os mesmos eclodem suas larvas penetram no interior da cana, conforme demonstra a Figura 1.

Cerca de três dias após a instalação das armadilhas os técnicos de campo retornam ao local para contabilizar o montante de pragas capturadas. Uma vez feita a leitura de diversas armadilhas instaladas em toda a extensão da plantação, os técnicos são capazes de consolidar informações do estado de infestação do cultivo, possibilitando a tomada de decisão para aplicação de defensivos agrícolas ou outras medidas em cada região.

1.2.3 Problemas no monitoramento

Tendo em vista que o técnico responsável pela contagem de armadilhas precisa monitorar dezenas de armadilhas por dia, diversos problemas podem ocorrer como, por exemplo: grande esforço de tempo e grande variabilidade na contagem. Além de grande carga cognitiva exigida dos técnicos ao longo do dia.

1.3 Objetivos

Objetivo geral deste trabalho é desenvolver uma rede neural para a detecção de mariposas da broca de cana-de-açúcar a partir de imagens previamente coletadas.

1.3.1 Objetivos específicos

- Classificar banco de imagens localizando mariposas nas fotos.
- Construir dataset de imagens para treinamento da rede.
- Treinar rede com banco de dados.
- Consolidar dados com resultados de classificação da rede.
- Validar a viabilidade do processamento de imagens para a identificação de mariposas em armadilhas.

1.4 Organização do trabalho

O Capítulo 2 trata da fundamentação teórica, abordando temas relacionados a processamento de imagens e redes neurais, bem como as tecnologias utilizadas no desenvolvimento do projeto. O Capítulo 3 refere-se à metodologia, descrevendo o processo de desenvolvimento do projeto e os materiais e recursos necessários. O Capítulo 4 discorre sobre os resultados obtidos, possíveis aplicações da rede e futuros passos. Por fim, o Capítulo 5 trata da conclusão do trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

Para que a detecção de objetos em imagens ocorra em uma rede neural, é necessário que a imagem passe por tratamentos e processamentos específicos. As técnicas deste processamento envolvem a manipulação de imagens digitais como aplicações de filtros e máscaras.

2.1 Processamento de imagens digitais por redes neurais

As técnicas referentes à análise de dados multidimensionais, capturados por diversos tipos de sensores recebem o nome de processamento digital de imagens. Trata-se da manipulação de uma imagem por computador de modo onde a entrada e a saída do processo são imagens digitais (INPE, 2022). Essas técnicas podem incluir pré-processamento de imagens e aplicação de filtros.

2.1.1 Imagem digital

As imagens digitais são uma representação gráfica constituída por um conjunto de pontos denominados de pixels (ALMERON, 2010).

2.1.2 Filtros

Os filtros possuem o objetivo melhorar a qualidade da imagem, de forma a deixá-la mais adequada para o processamento. Entre os processos mais aplicados nesta etapa estão filtros de mediana, que têm como objetivo reduzir ruídos na imagem, filtros de passa alta que realçam contornos e bordas dos objetos na imagem e conversão para tons de cinza de forma a reduzir a quantidade de informações da imagem a ser processada (VALMORBIDA, 2018, p. 33).

Um exemplo comumente usado é a utilização do filtro em escala de cinza, para processamento de imagens cuja característica independe de sua cor. Desta forma reduzindo o custo de processamento da imagem, conforme ilustra a Figura 3.



Figura 3 - Armadilha de mariposas com filtro de escala de cinza.

Filtros também podem ser utilizados para destacar determinadas características de uma imagem, como o operador Sobel, no qual são utilizadas duas máscaras 3x3, uma para identificar bordas verticais e outra para bordas horizontais da imagem (KUZE, 2020).

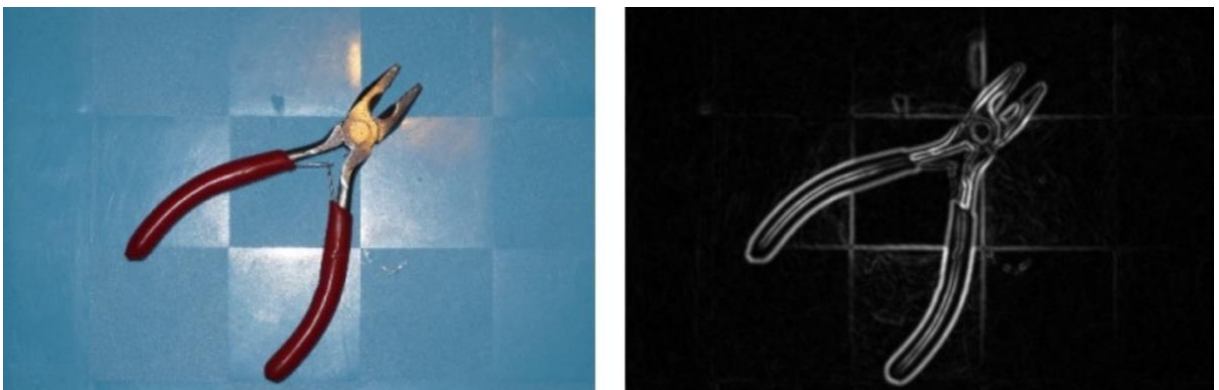


Figura 4 - Operador Sobel (Fiveko, 2023).

O operador Sobel demonstrado na Figura 4 é um filtro que tem como característica destacar os contornos das imagens.







Edge detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Box blur (normalized)	$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	
Gaussian blur 3 × 3 (approximation)	$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	

Figura 5 – Filtros convolucionais (RODRIGUES, 2022).

Na Figura 5 é possível observar como diferentes tipos de máscaras podem destacar diferentes aspectos de uma imagem.

Estes filtros podem ser aplicados na imagem para destacar determinadas características do objeto almejado para detecção, facilitando o reconhecimento do mesmo pela rede neural.

2.2 Redes Neurais Artificiais

2.2.1 Perceptron

O neurônio, em biologia, é a unidade básica de trabalho do cérebro, uma célula eletricamente excitável que se comunica com outras células por meio de conexões chamadas sinapses. Um neurônio é o elemento de uma rede neural. Um perceptron é o modelo matemático de um neurônio biológico onde os sinais elétricos são representados como valores numéricos. (AJALA, 2021).

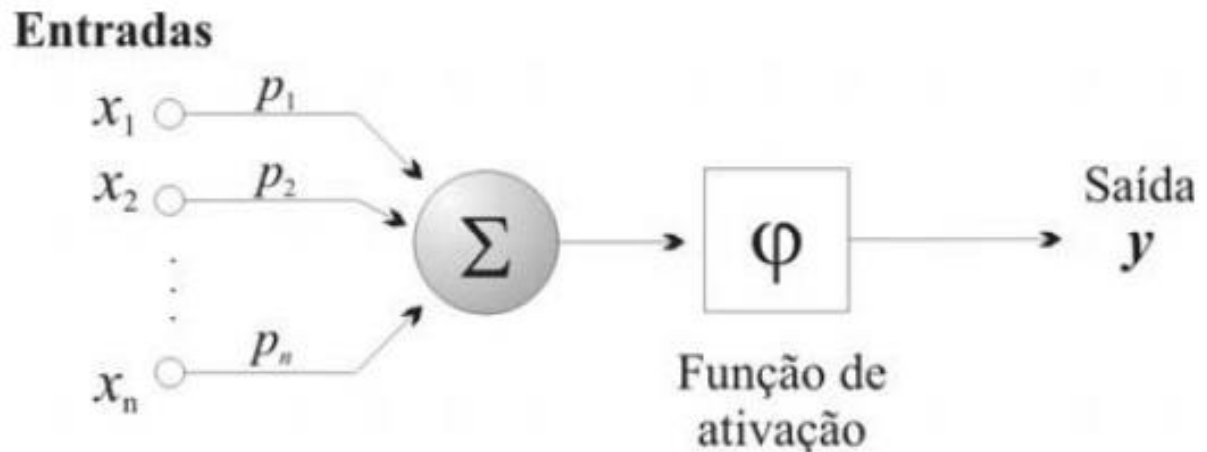


Figura 6 - Modelo matemático de um perceptron (FERNEDA, 2006).

A Figura 6 representa um perceptron com N entradas atribuídas pela letra X, dos quais estes valores de entrada são multiplicados pelos pesos atribuídos pela letra P, e somados para acumular os sinais de entrada. Após o somatório uma função de ativação φ que limita o intervalo permissível de amplitude do sinal de saída a um valor fixo, resultando no output representado pela letra y (FERNEDA, 2006).

Matematicamente o perceptron pode ser expresso da seguinte forma:

Equação 1 - Representação matemática do perceptron

$$y = \varphi\left(\sum_{1}^{N} X_n \cdot P_n\right)$$

2.2.2 Rede neural

Um conjunto de múltiplos neurônios em camadas forma uma rede neural artificial. As redes neurais artificiais são modelos que almejam simular o processamento de informação do cérebro. São formadas por neurônios, que se unem por meio de conexões sinápticas (FERNEDA, 2006).

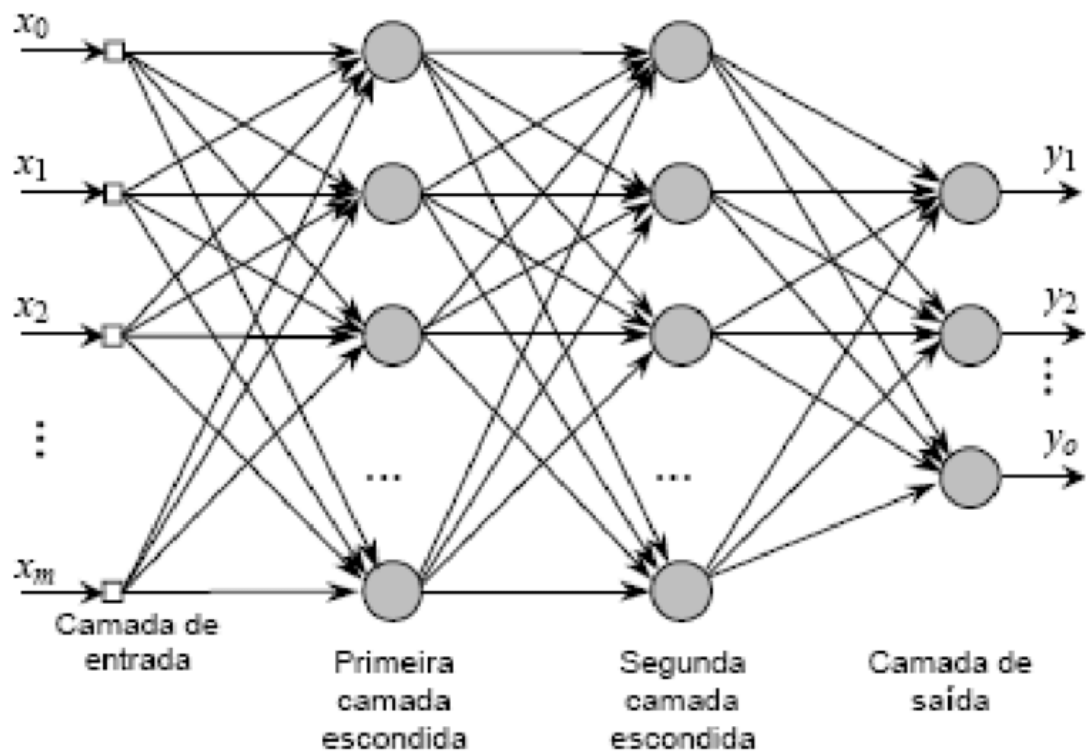


Figura 7 – Rede neural (OLIVEIRA, 2023).

A Figura 7 representa uma rede constituída por diversos perceptrons, e três camadas, representadas pelas colunas de perceptrons. Após receber os inputs, os dados são processados pelos perceptrons ao longo das camadas, multiplicando seus respectivos valores pelos pesos de cada perceptron e transferindo o resultado para as camadas subsequentes até que o resultado seja propagado para a camada de output.

Os neurônios, quando interligados, apresentam um enorme potencial para mapeamentos não lineares. O algoritmo de aprendizagem é o procedimento utilizado para realizar o processo de aprendizagem, cuja função é modificar os pesos sinápticos da rede para alcançar um objetivo de projeto desejado (NIED, 2007). Esta rede é constituída de um conjunto de neurônios que processam as entradas e possuem determinadas saídas como resposta.

2.2.2.1 Redes Neurais Convolucionais

Na área de reconhecimento e classificação de imagens, os melhores resultados são obtidos usando as redes neurais convolucionais (CNNs, do inglês Convolutional Neural Network), uma subclasse das redes neurais profundas (ROSA,

2019, p. 10). Uma rede neural convolucional difere de uma rede neural regular pelo uso de camadas convolucionais. Uma rede convolucional reconhece um objeto procurando por características de baixo nível, como bordas, linhas e curvas, e então segue construindo características mais abstratas através de uma série de camadas convolucionais (ROSA, 2019, p. 10).

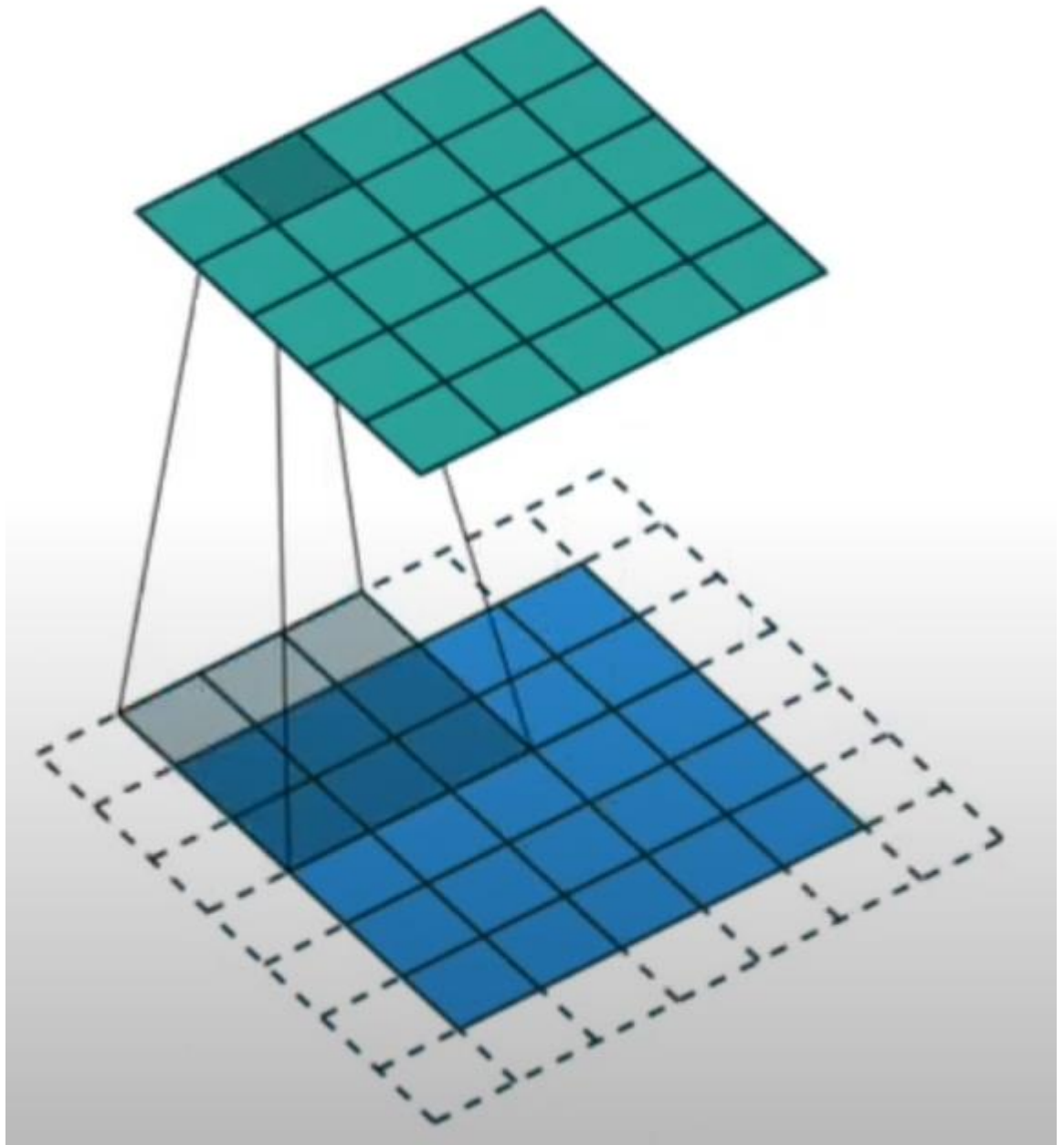


Figura 8 – Filtro convolucional (GONÇALVES, 2023).

O filtro convolucional demonstrado na Figura 8 percorre a imagem, representada pelos quadrados azuis, onde cada quadrado representa um pixel, e os quadrados pontilhados representam um padding a fim de permitir que o filtro 3x3 percorra toda a imagem fazendo um somatório destes pontos resultando em um

valor representado pelo quadrado verde em destaque. O resultado deste processo pode ser melhor compreendido através dos exemplos da Figura 5.

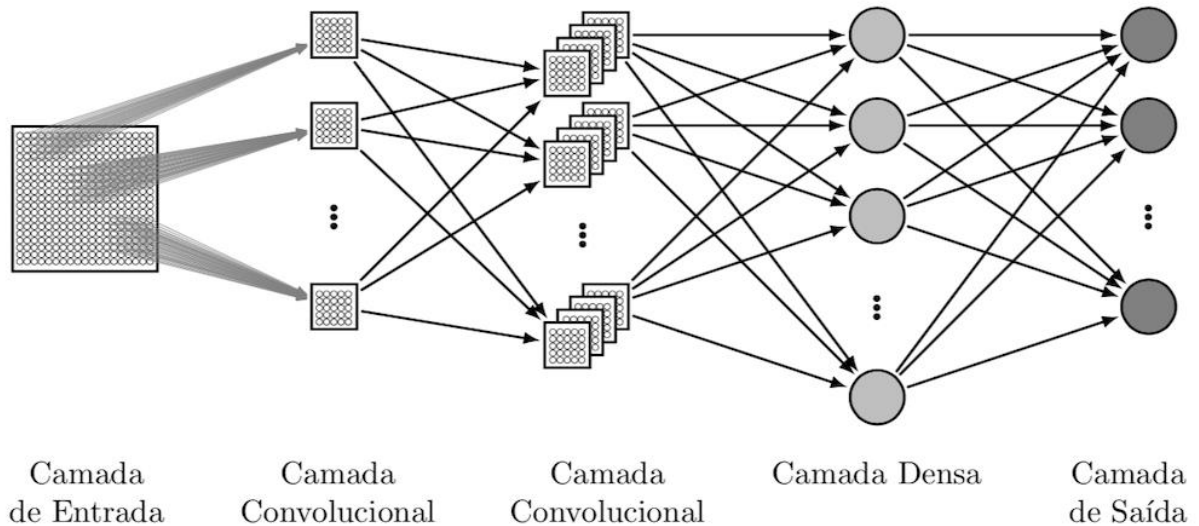


Figura 9 - Rede neural convolucional (SAKURAI, 2023).

Conforme ilustrado na Figura 9, podemos observar a aplicação de um conjunto de camadas convolucionais a um conjunto de camadas de perceptrons, o que constitui uma de Rede neural convolucional. Neste caso podemos observar os pixels de uma imagem digital sendo processados pela rede e como resultado tem múltiplos outputs. Para calibrar os pesos das camadas convolucionais e dos neurônios é utilizada a técnica de aprendizado de máquina.

2.2.3 Aprendizado de máquina

O aprendizado de máquina é o uso de métodos computacionais para coletar dados, aprender padrões, e então fazer uma previsão utilizando esses padrões. Isto é, a máquina consegue ser treinada usando uma quantidade de dados e métodos que dão a ela habilidades de aprender como executar a tarefa, sem ser explicitamente programada para isso (ROSA, 2019).

- I. Aprendizado supervisionado: são fornecidas entradas e saídas exemplos para o computador através do conjunto de dados, com o objetivo de ensiná-lo a mapear uma regra geral.
- II. Aprendizado não supervisionado: não são fornecidas saídas mapeadas e o objetivo da rede é identificar padrões, para que a parte de aprendizado seja feita pela própria máquina.
- III. Aprendizado por reforço: a máquina interage com o ambiente de forma dinâmica e recebe um feedback baseado em uma função de avaliação para ajustar seus pesos e melhorar a cada iteração (ROSA, 2019).

Para este projeto será utilizado o aprendizado supervisionado. A técnica de aprendizado de máquina vai permitir que a rede aprenda quais são os padrões da imagem que caracterizam uma mariposa.

O treinamento envolve o ajuste fino de uma rede neural de mau desempenho devido a pesos inicializados aleatoriamente em uma rede com alta precisão. O treinamento é equivalente a minimizar a função de perda. As redes neurais aprendem com os erros entre as previsões e as verdades básicas, atualizando seus parâmetros (LIU, 2011).

2.3 Arquitetura selecionada

Para o desenvolvimento deste projeto foi selecionada o framework Darknet, que oferece uma série de ferramentas para treinamento de redes neurais focadas em detecção de objetos em imagens e vídeos.

2.3.1 Darknet

Darknet é uma estrutura de rede neural de código aberto escrita em C e CUDA. É rápido, fácil de instalar e suporta computação de CPU e GPU (REDMON, 2016).

Este framework oferece uma série de ferramentas para treinamento e análise de redes neurais convolucionais.

2.3.2 Detecção de objetos na rede neural artificial

O reconhecimento de objetos é considerado o eixo central da visão computacional. Quando se trata da tarefa de reconhecimento de objetos na qual o que se busca é conhecido, esta técnica é denominada como detecção de objetos. (OLIVATTO, 2021, p. 42).

A detecção de objetos pode ser entendida como o procedimento para identificar um objeto, determinar a que classe o objeto pertence e estimar a sua localização exibindo a caixa delimitadora ao redor do objeto em questão. (OLIVATTO, 2021, p. 42).

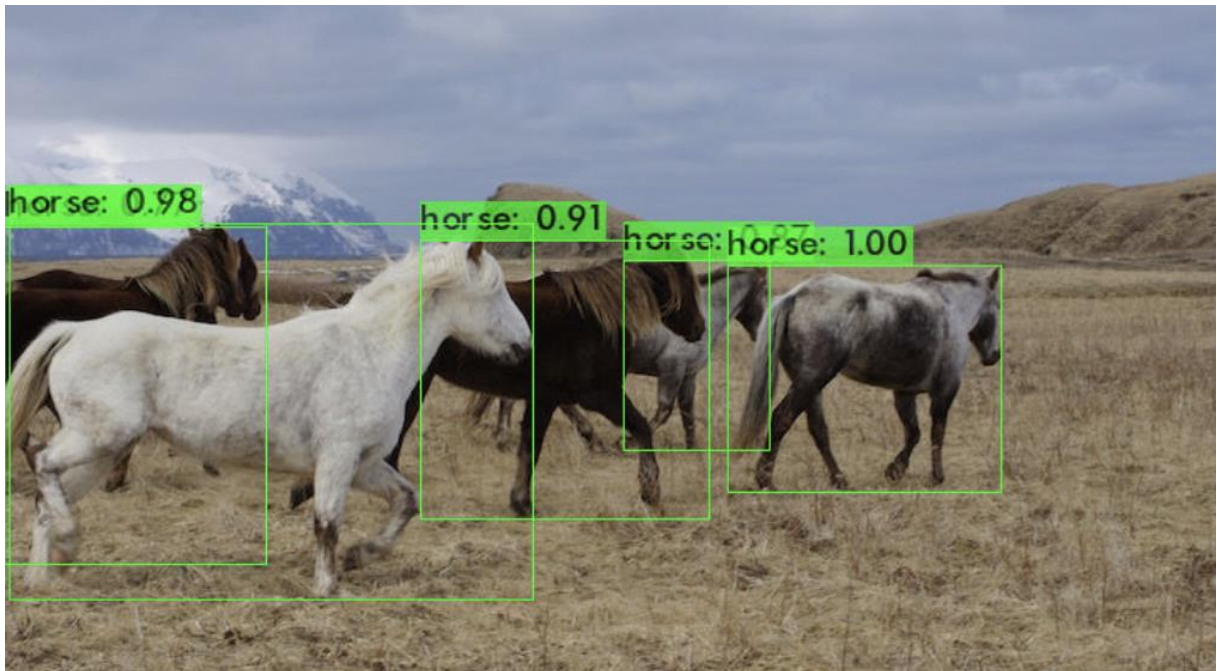


Figura 10 – Detecção de objetos.

Na Figura 10 é possível observar o objeto *horse* sendo detectado diversas vezes através das caixas verdes seguidas do valor de confiabilidade do reconhecimento do objeto que a rede definiu para cada detecção.

2.3.3 YOLO

Em 2016 foi apresentada uma estratégia para possibilitar a detecção de objetos em tempo real, a qual foi denominada YOLO (*You Only Look Once*).

Enquanto as estratégias anteriores utilizam uma rede neural projetada para classificação, na tarefa de detecção de objeto, a YOLO apresenta um método focado em detecção. O YOLO é um sistema de detecção de objetos em tempo real, capaz de identificar objetos em vídeos e em imagem. Para realizar essa tarefa ele faz uso de partes da imagem para identificar o objeto e sua posição (PIRES, 2018, p. 50).

O procedimento de predição pode ser compreendido com a definição dos seguintes passos (HOLLERWEGER, 2019, p. 18):

- I. A imagem é dividida em grid de 13 por 13 células.
- II. Cada uma das células é responsável pela predição de um grupo de caixas de seleção. As caixas podem ter dimensões e formatos variados, e correspondem à área onde a rede acredita que exista um objeto. Neste passo, também são obtidos valores de confiança para cada uma das caixas.
- III. Para cada uma das caixas de seleção, a rede faz a predição de uma classe. O valor de confiança de que a caixa contenha um objeto é combinado com um novo valor que representa a probabilidade de acerto no objeto classificado.
- IV. As caixas com valores de confiança abaixo de certo limiar (que normalmente é definido em torno de 30%) são eliminadas. Após tal operação, são obtidas as caixas com maior chance de conter um objeto. (HOLLERWEGER, 2019, p. 18)

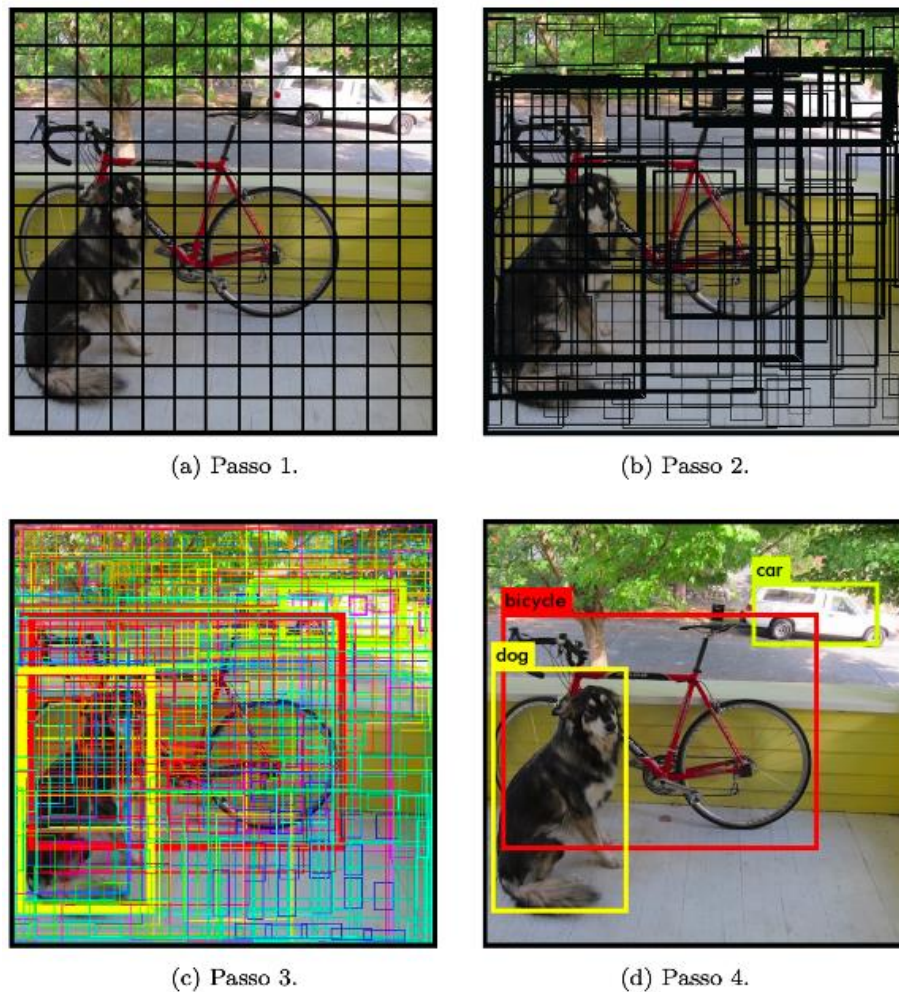


Figura 11 – Estratégia YOLO (Darknet, 2023).

Na Figura 11 ilustra as quatro principais etapas de detecção de objetos pela estratégia YOLO, onde como resultado é apresentado caixas indicando os objetos detectados na imagem e seus respectivos nomes.

2.3.3.1 Configuração e treinamento de uma rede YOLO

Para a configuração e treinamento do modelo YOLO através da Darknet são definidas as seguintes propriedades:

- Máscaras: Filtros aplicados com objetivo de destacar e identificar características dos objetos na imagem.
- Classes: Lista de objetos a serem detectados na imagem. Neste trabalho teremos apenas uma classe denominada MARIPOSA.
- Épocas: O número de vezes que o modelo é exposto ao conjunto de dados de treinamento. (PEREIRA, 2021)

2.3.3.2 Modelos pré-treinados

Modelos pré-treinados são facilitadores para o desenvolvimento de redes neurais convolucionais, estes modelos agilizam o treinamento da rede neural uma vez que estes já foram desenvolvidos com configurações e pesos específicos para determinadas tarefas, entre elas a detecção de objetos.

O projeto YOLO oferece diversos modelos pré-treinados, estes modelos estão disponibilizados no diretório do projeto Darknet.

2.3.4 Dataset

Banco de imagens com informações de classes e posições dos objetos presentes em cada imagem.

O dataset do presente trabalho é composto de um conjunto de imagens com seus respectivos arquivos .txt indicando a localização e dimensão dos objetos na imagem.

3 METODOLOGIA

Este capítulo aborda os recursos e técnicas utilizadas para treinamento e validação do desempenho da rede neural, abordando tecnologias utilizadas e a metodologia de desenvolvimento deste trabalho.

3.1 Recursos

Para o desenvolvimento da rede neural foi utilizado um computador e os respectivos softwares referenciados na Tabela 1.

Tabela 1 – Recursos.

Itens	Quantidade
Hardware	
Computador core i5, 500GB ssd, 8GB RAM, placa de vídeo GTX 1060 6GB.	1
Software	
Darknet	1
YOLO	1
VSCode	1
GitHub	1
Cvat	1

3.2 Desenvolvimento

Para o desenvolvimento de código foi utilizado o editor de texto VSCode. O framework Darknet para treinamento da rede, o software git para versionamento de código, o software Cvat para classificação das imagens e o pacote de arquivos YOLO com as configurações de arquitetura da rede.

3.3 Coleta de imagens

Para o desenvolvimento da rede foram utilizadas 674 contendo fotos de armadilhas com mariposas capturadas. Estas imagens foram triadas e selecionadas para o treinamento da rede. Os critérios para a seleção das imagens levam em conta, resolução, distância dos objetos, qualidade da imagem e formato da imagem.

3.4 Classificação

Para realizar o treinamento da rede é necessário que o arquivo da imagem acompanhe um arquivo .txt com informações de posicionamento e dimensões dos objetos presentes na fotografia.

Para o processo de classificação de imagens foi utilizado o software CVAT, através deste é possível localizar e delimitar objetos na imagem a fim de gerar o database requerido para o treinamento da rede. CVAT é uma ferramenta interativa de anotação de vídeo e imagem para visão computacional. É usado por dezenas de milhares de usuários e empresas em todo o mundo. O CVAT é gratuito e de código aberto (OPENCV, 2022).



Figura 12 - Cvat classificação de imagens.

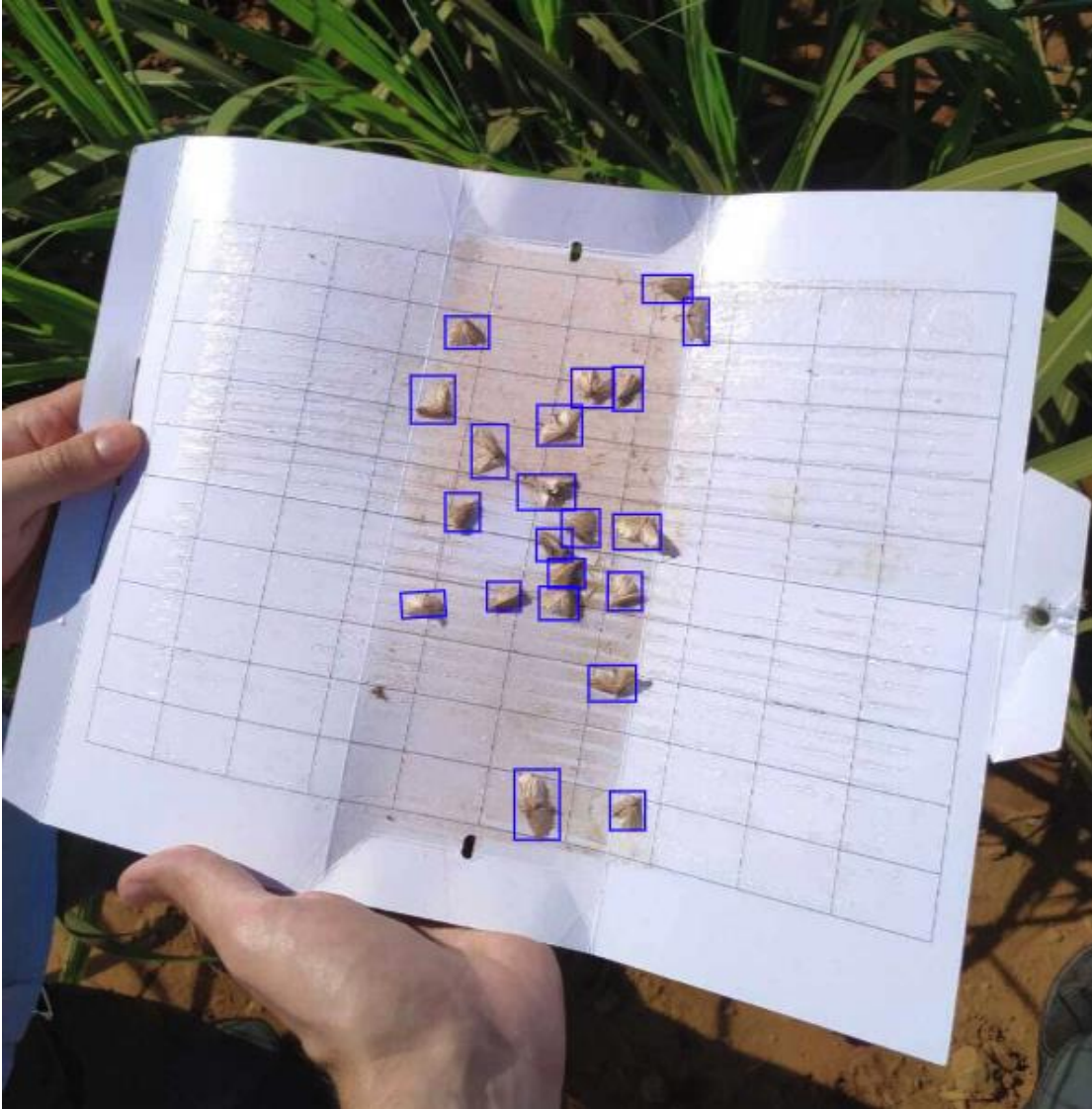


Figura 13 - Cvat Objetos mapeados.

As Figuras 12 e 13 demonstram a interface gráfica do Cvat, onde é possível destacar na imagem a localização e dimensão das classes de objetos determinados.

Após construir o dataset, as imagens foram divididas em três conjuntos distintos:

- I. Imagens de treinamento: Imagens utilizadas para o treinamento da rede (70% do banco de imagens).
- II. Imagens de validação: Imagens utilizadas para avaliar a capacidade de generalização da rede durante o treinamento (20% do banco de imagens).
- III. Imagens de teste: Imagens utilizadas para testar o modelo após a finalização do mesmo (10% do banco de imagens).

Para este projeto as imagens coletadas foram divididas seguindo a proporção referida acima, no qual foi coletado um total de 674 imagens, sendo, 472 estão no conjunto de treino, 135 no conjunto de validação e 67 para o conjunto de testes.

3.5 Configuração do projeto

As configurações dos modelos foram feitas com base na documentação do projeto (ALEXEY, 2022). Estas configurações levam em conta parâmetros como:

- Resolução: Padronização da resolução de todas as imagens, o framework se encarrega de ajustar a configuração das imagens.
- Épocas: O número de vezes que o modelo é exposto ao conjunto de dados de treinamento.
- Classes: Lista de objetos a serem detectados na imagem.
- Filtros: Quantidade de filtros a serem aplicados na imagem.
- Learning rate: taxa de ajuste de pesos na rede durante o treinamento.

O framework Darknet oferece diversas versões da rede neural YOLO, durante o desenvolvimento deste trabalho foram utilizadas as versões YOLO v4 tiny, YOLO v7 tiny.

As versões tiny das arquiteturas de rede foram selecionadas, pois são redes mais leves e com tempo reduzido de treinamento e processamento no reconhecimento das imagens (JIANG, ZHAO, *et al.*, 2013). Treinamentos nas redes tiny levam cerca de 5 horas para serem finalizados, enquanto nas redes tradicionais podem levar de 2 a 3 dias dependendo das configurações especificadas e do hardware escolhido. Apesar dos modelos mais robustos apresentarem maior precisão, foi ponderado o tempo de treinamento bem como o custo computacional do processo. Outro ponto importante para a decisão da escolha da arquitetura tiny se refere a sua maior facilidade de portabilidade para dispositivos móveis, o que no contexto do agronegócio pode ser vantajoso, uma vez que os técnicos de campo utilizam dispositivos móveis no trabalho de monitoramento de pragas.

3.6 Treinamento dos modelos

Após definida as configurações da arquitetura da rede desejada é gerado um arquivo .cfg o qual contém todas as especificações citadas como: resolução, quantidade total de épocas, quantidade e tipos de classes, quantidade de filtros e *learning rate*.

Uma vez configurada a rede, utilizamos a lista de imagens de treino para calibrar a rede no decorrer das épocas. Para este projeto a cada cem épocas a partir da época de número mil a rede foi validada com a lista de imagens de validação, gerando um percentil de precisão da rede. Estes valores são acompanhados durante o treinamento através de um gráfico que relaciona épocas e precisão, juntamente com o loss da rede.

Durante o treinamento da rede o framework Darknet fornece gráficos do desempenho dos modelos. Estes dados possibilitam o acompanhamento do treinamento da rede, bem como servem para comparar diferentes modelos treinados.

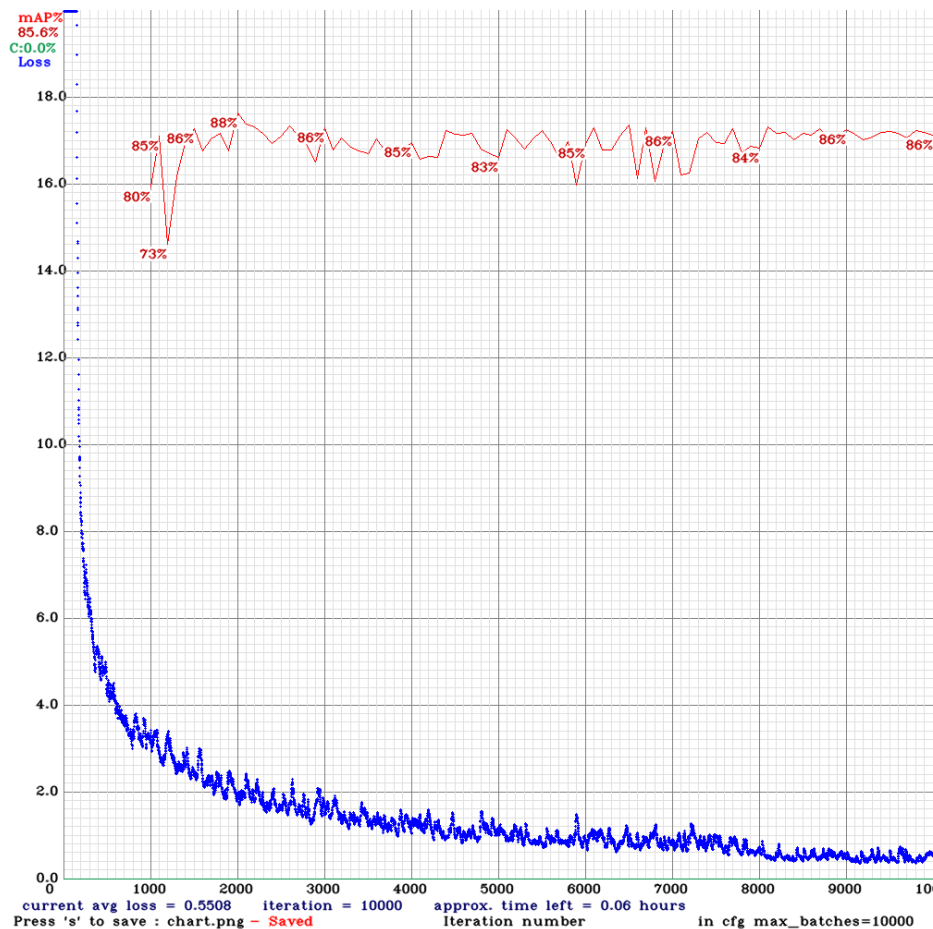


Figura 14 – Gráfico de precisão e função perda por épocas.

O loss apresentado na Figura 14 representa a média do erro entre as previsões da rede neural e os rótulos das imagens de treinamento. É uma métrica que indica o desempenho do modelo durante o treinamento. É esperado que o loss diminua ao longo das épocas de treinamento, indicando a melhora do modelo em relação a sua capacidade de fazer previsões corretas. O cálculo da função perda deste trabalho será feito através do método da soma do quadrado dos erros (REDMON e FARHADI, 2018). O mAP representa o percentil da média de acertos da rede referente ao grupo de imagens de testes ao longo das gerações. As iterações representam a quantidade de gerações de treinamento realizadas pela rede. Dado que as 1000 primeiras interações geralmente não apresentam bons resultados, o mAP passa a ser calculado a partir de iteração 1000, e é calculado a cada 100 interações. O que economiza tempo e recursos de processamento no início do treinamento.

3.7 Critérios de avaliação

Os critérios definidos para avaliação do modelo levam em conta o contexto de utilização almejado, sendo neste caso o mAP, loss e tempo de processamento. Tais métricas permitem validar a viabilidade da rede para futuras aplicações, bem como seu desempenho na detecção de mariposas.

Os dados de mAP e loss são fornecidos durante o treinamento das redes através do framework da Darknet. O tempo de processamento será medido realizando uma média do tempo de processamento de cinco imagens do banco de testes. Uma vez coletando estes dados é possível realizar um comparativo de desempenho e selecionar o modelo que mais se aproxima dos critérios estabelecidos.

4 RESULTADOS E DISCUSSÃO

Este capítulo aborda os resultados das redes neurais desenvolvidas, bem como a identificação de problemas encontrados durante o desenvolvimento e futuros passos para a melhoria da rede neural desenvolvida.

4.1 Desempenho dos modelos

A primeira rede a ser testada foi a YOLO v4 tiny, representada pela Figura 15. Nota-se que a rede apresentou um aumento do mAP ao longo das épocas, bem como uma redução do Loss. Tendo seu mAP em 71,7%.

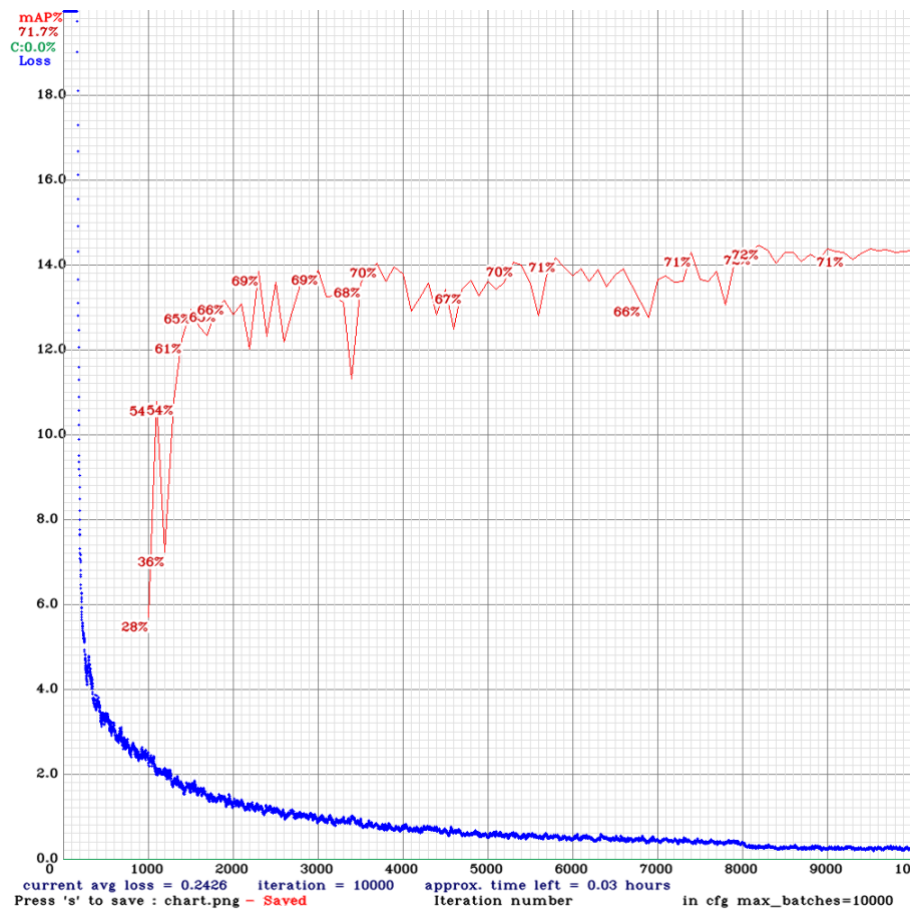


Figura 15 – YOLO v4 tiny.

A rede YOLO v7 tiny apresentou resultados próximos aos encontrados na rede YOLO v4 tiny, conforme mostra a Figura 16.

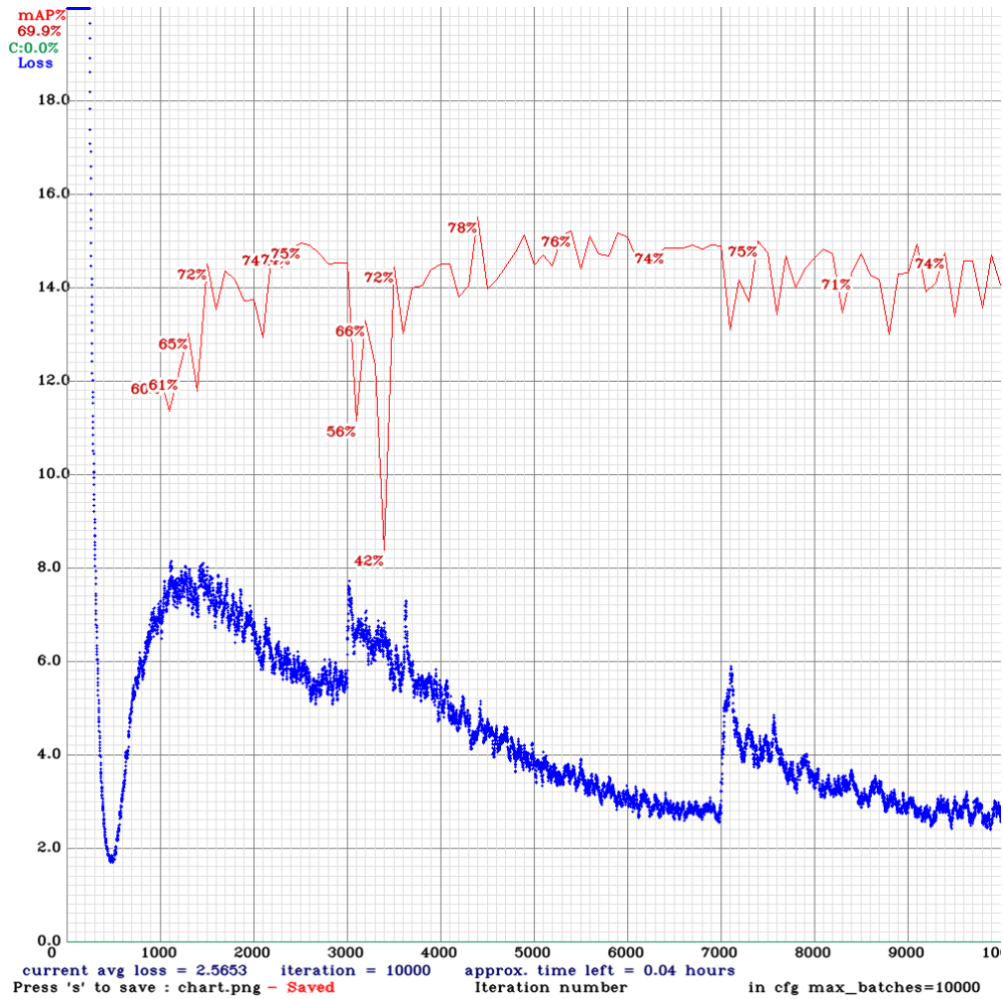


Figura 16 - YOLO v7 tiny.

Em relação à rede YOLO v4 tiny podemos observar um mAP próximo à rede YOLO v7 tiny. Por outro lado, o loss da rede apresenta maiores taxas de variações e um valor maior em comparação a rede YOLO v4 tiny.

Também se destaca o comportamento do loss nesta rede, apresentando picos nas épocas de número 3000 e 7000 acompanhados de quedas na taxa de assertividade. Estes picos são ocasionados devido à estratégia da rede, que em determinados pontos do treinamento pode aumentar sua taxa de aprendizado, para causar maior variabilidade no treinamento, ocasionalmente podendo aumentar ou diminuir a assertividade da rede.

4.1.1 Refinamento dos modelos

Após analisar os dados obtidos foram realizadas modificações na arquitetura da rede a fim de incrementar seu desempenho.

Conforme já descrito, o arquivo .cfg possui algumas configurações da rede para o treinamento como: resolução, quantidade total de épocas, quantidade e tipos de classes, quantidade de filtros e taxa de aprendizado.

Buscando refinar os modelos alguns desses parâmetros foram modificados almejando uma maior precisão da rede. Dentre eles a resolução foi o fator com maior predominância para a melhoria da precisão, sendo a resolução padrão das imagens de treinamento 416 pixels de altura por 416 pixels de largura, nos casos das redes customizadas as resoluções das imagens para treino foram aumentadas para 624 pixels de altura por 624 pixels de largura. Esta mudança foi aplicada em ambas as redes.

Em específico para a rede YOLO v7 tiny, foi removida a variação da taxa de aprendizado que ocorria nas épocas 3000 e 7000, tendo em vista que para este tipo de treino em específico estes picos de aumento na taxa de variação não causaram efeitos positivos, conforme descrito na explicação da Figura 16.

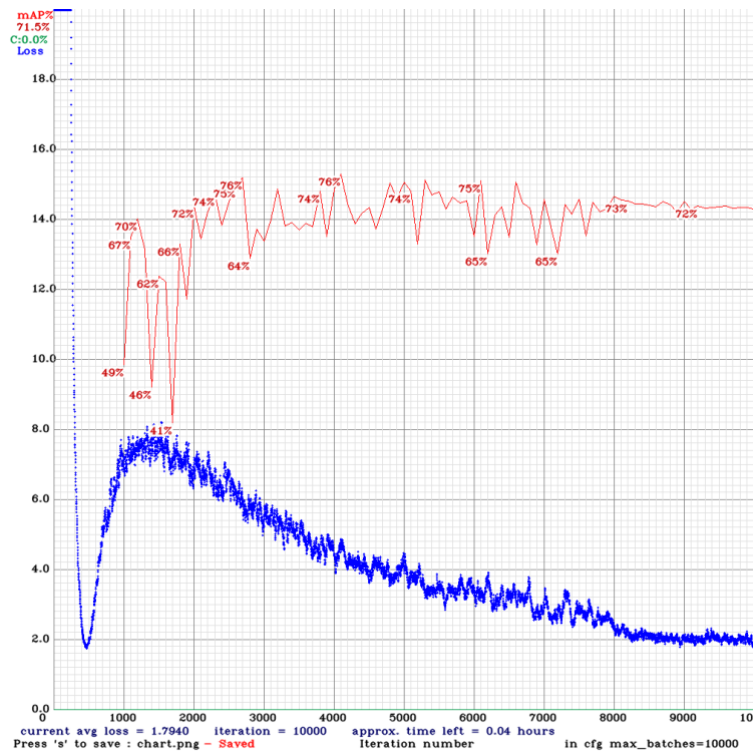


Figura 17 - YOLO v7 tiny customizada.

Analisando o resultado da Figura 17 em relação a sua versão anterior podemos perceber uma pequena melhora na taxa de assertividade.

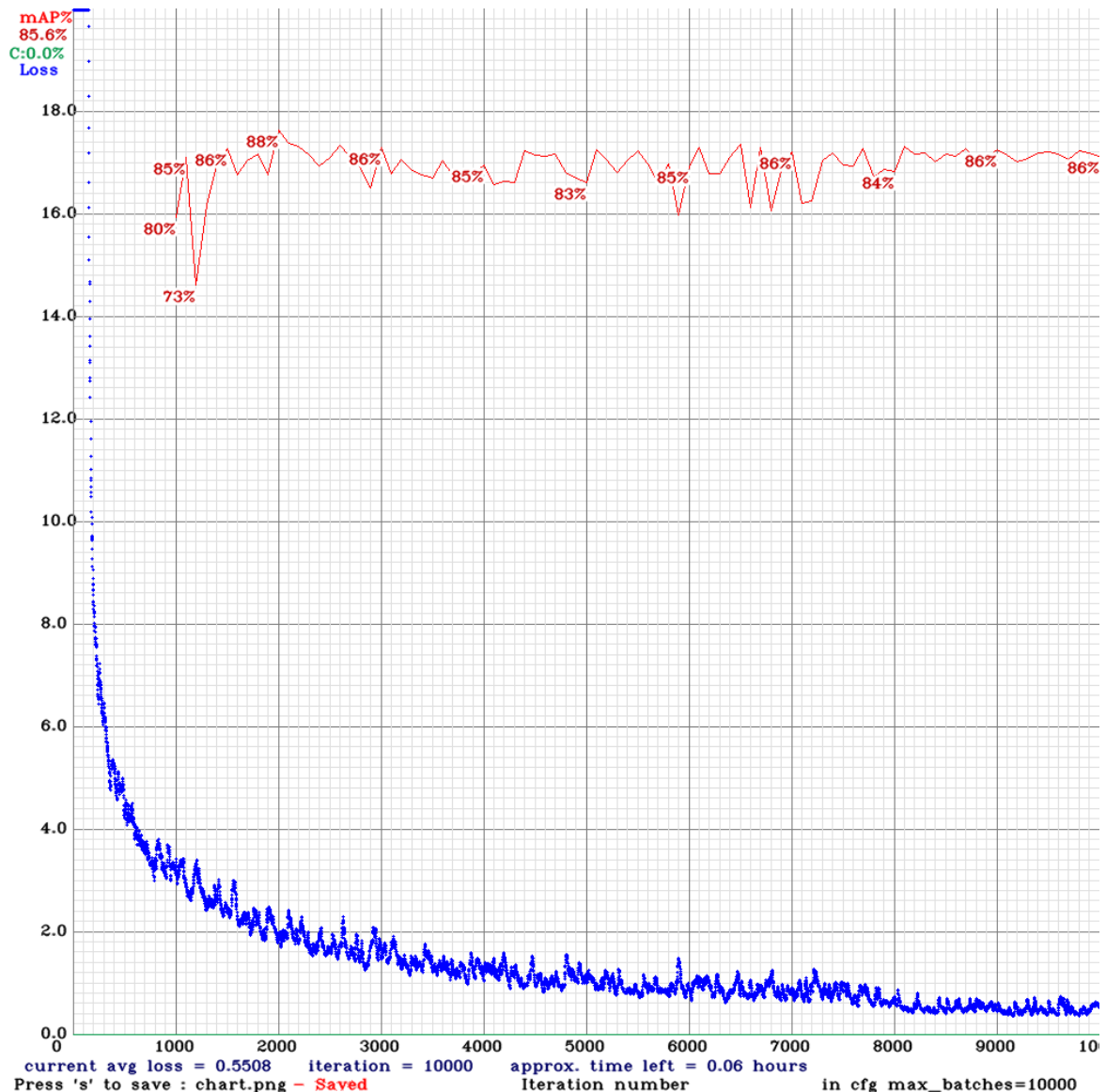


Figura 18 - YOLO v4 tiny customizada.

Na figura 18 é possível observar um aumento do mAP da rede YOLO v4 tiny de 71,7% para 85,6%. O que classifica a rede YOLO v4 tiny customizada como a mais performática dentre as testadas para a detecção de mariposas.

4.2 Modelo selecionado e seu desempenho

Após treinar os modelos citados acima, o modelo mais eficiente, levando em conta os critérios definidos foi à rede YOLO v4 tiny customizada. Indicado pela Figura 18 - YOLO v4 tiny customizada. Dado seu baixo loss e o mAP de 85,6%, tendo o seu maior valor de percentil em 88% em sua interação de número 2000. O tempo médio de processamento calculado de 350ms.

A fim de validar o seu desempenho em casos reais, foi utilizado o banco de imagens de testes, para testar a taxa de reconhecimento da rede, alcançando 83% de assertividade na detecção de mariposas.

4.3 Interface gráfica

A fim de validar visualmente o funcionamento da rede foi desenvolvido um serviço backend que recebe uma imagem, processa pela rede YOLO e retorna a imagem processada em uma interface gráfica Web.

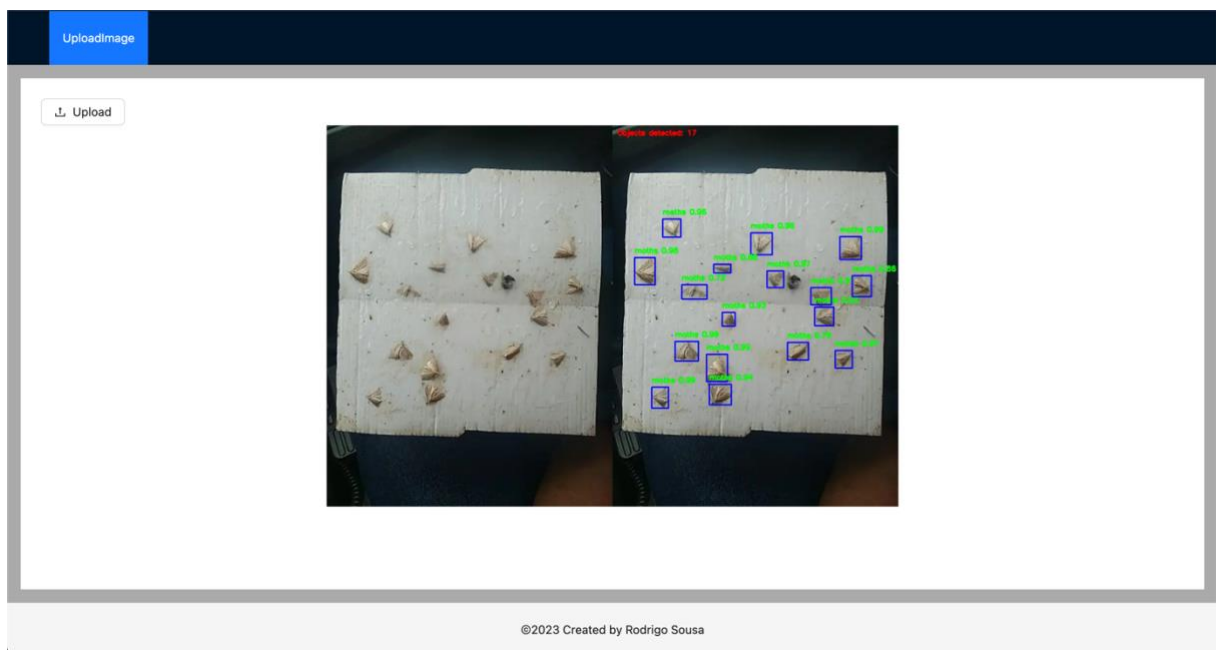


Figura 19- interface gráfica.

A Figura 19 apresenta o resultado da detecção de uma imagem selecionada ao lado da imagem original. Desta forma a interface gráfica contribui para a validação visual do funcionamento da rede, possibilitando identificar os pontos de melhoria e interpretar de forma qualitativa as falhas nas detecções de mariposas.

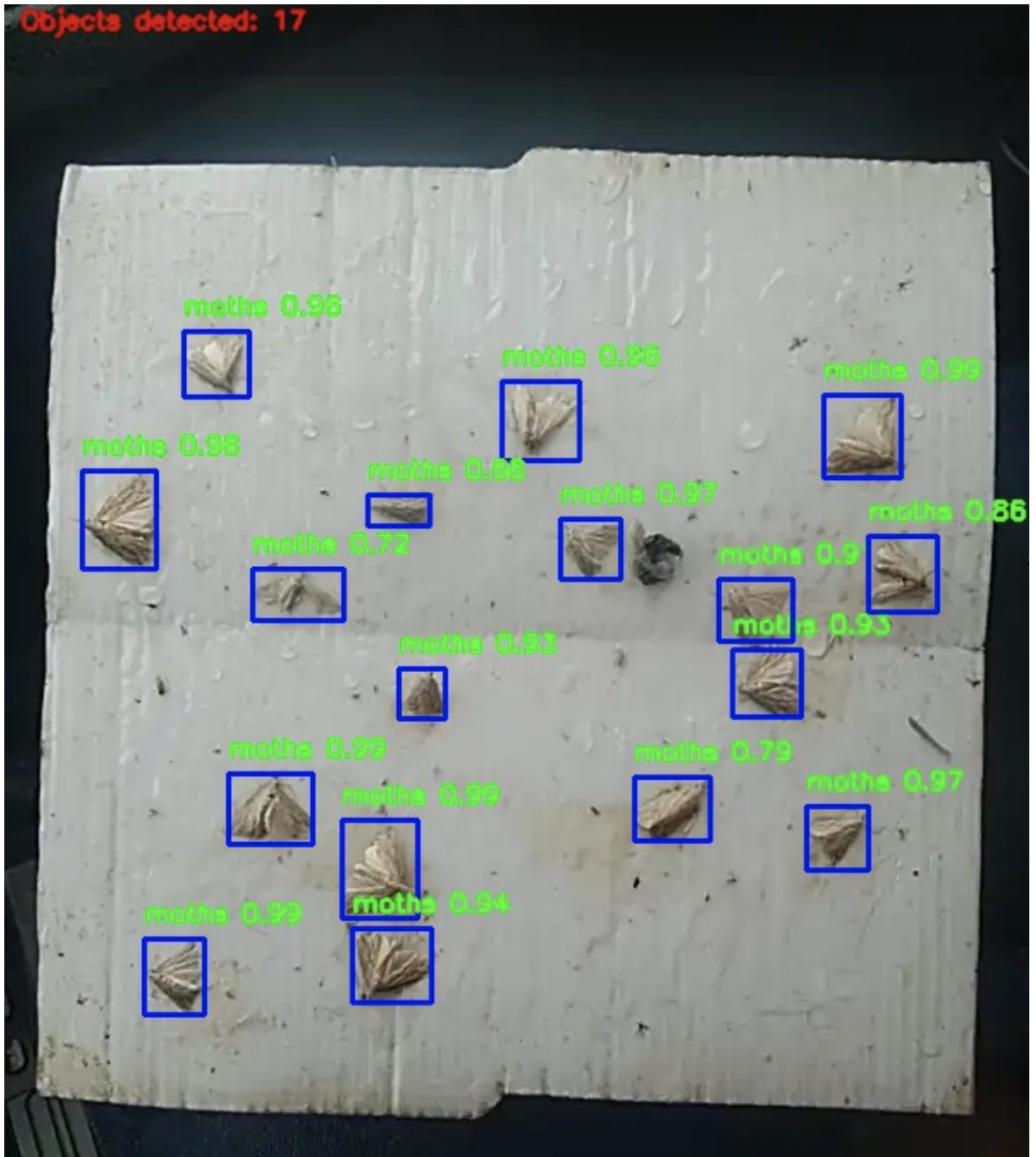


Figura 20 - Exemplo de detecção.

A Figura 20 destaca o output da rede, onde os objetos são destacados pelas caixas azuis seguidos do nome da classe (*moths*) e da confiabilidade de detecção da rede para cada objeto. A confiabilidade pode ser entendida um indicativo do quanto o objeto destacado em questão se aproxima do objeto entendido pela rede como uma mariposa. Sendo 1.00 100% de confiança e 0.00 0% de confiança.

4.4 Dificuldades encontradas no dataset utilizado

Durante os testes de desempenho da rede foi possível observar dois principais problemas de detecção, sendo estes reconhecer casos em que apenas partes de mariposas estão presentes na imagem ou quando ocorre a sobreposição de múltiplas mariposas aglomeradas em um único ponto.

Como a classificação de mariposas teve foco em classificar apenas mariposas inteiras, ou seja, descartando pedaços de asas e mariposas deformadas, era esperado que determinados casos a rede não tivessem treinamento o suficiente para detectar tais objetos. O mesmo ocorre para mariposas sobrepostas, nestes casos a rede detecta apenas um objeto ou considera um objeto diferente por não seguir o padrão de treino, sendo estes, mariposas inteiras e espaçadas.

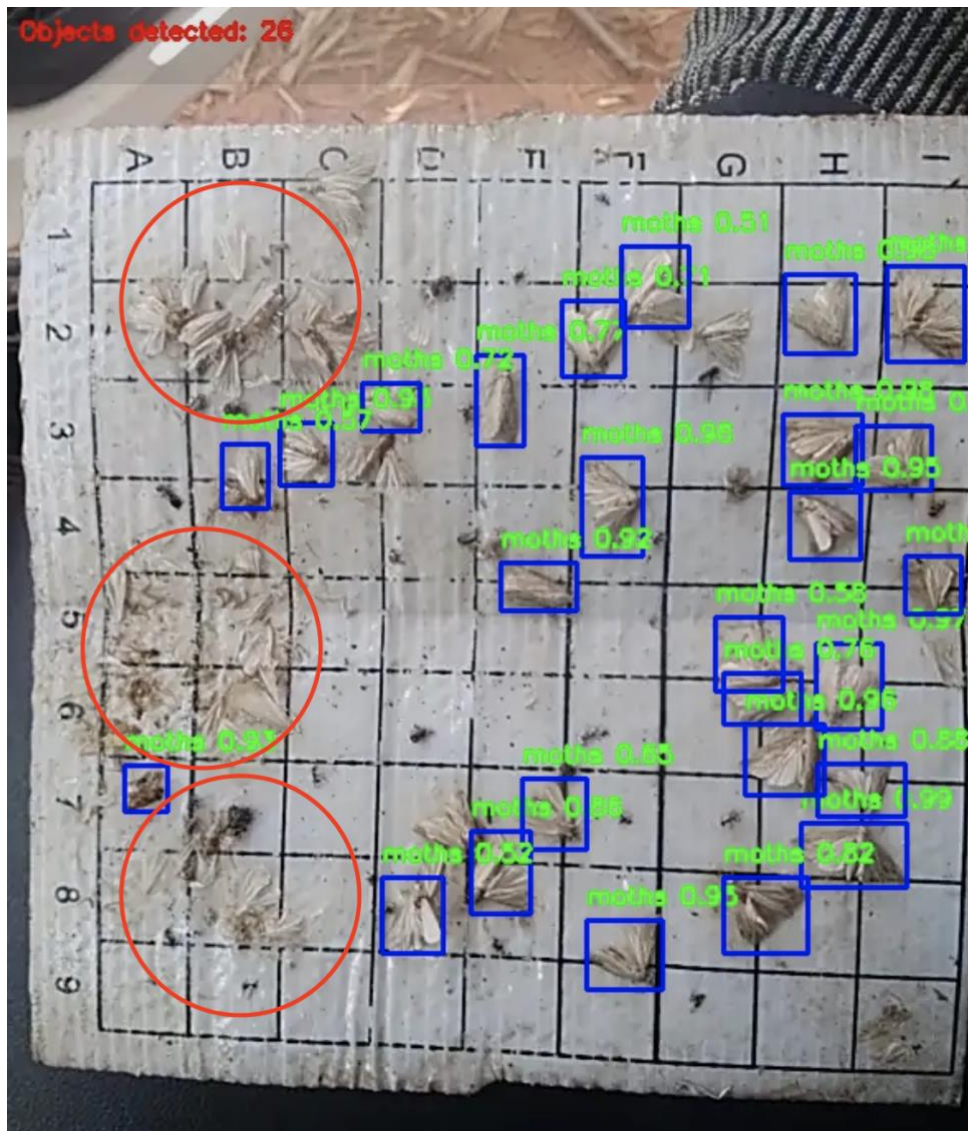


Figura 21 - Problemas de detecção.

A Figura 21 destaca com círculos vermelhos o problema discorrido neste tópico, onde partes de mariposas e mariposas aglomeradas não foram detectadas pela rede.

4.5 Possíveis aplicações da rede

Dentre as diversas aplicações podemos destacar a utilização da rede em armadilhas eletrônicas com um sistema de inteligência embarcado para detecção e contagem de mariposas de forma autônoma

Também é possível converter a rede para smartphones, onde podem se mostrar uma ferramenta eficaz para ajudar os técnicos de campo na contagem de mariposas reduzindo o esforço cognitivo dos mesmos.

A tecnologia também pode ser embarcada em modelos autônomos de armadilhas, coletando e enviando os dados processados sem a necessidade da presença humana no local de coleta, reduzindo custos de mão de obra e agilizando os processos de monitoramento.

4.6 Futuros passos

Analisando as dificuldades de detecção encontradas é possível tornar a rede mais robusta através da expansão do banco de imagens para treino do modelo, combinado com a adição dos casos de mariposas aglomeradas ou partes de mariposas no treinamento da rede.

5 CONCLUSÃO

Analisando os resultados obtidos foi possível constatar a viabilidade da utilização de redes neurais para a identificação de mariposas em armadilhas. O processamento de imagens para detecção de mariposas em armadilhas pode se tornar uma ferramenta de ganho de produtividade para o setor sucroalcooleiro.

Em conclusão, este trabalho alcançou os objetivos propostos no desenvolvimento de uma rede neural para a detecção de mariposas da broca de cana-de-açúcar. Alcançando uma taxa de detecção de mariposas de 83% com um tempo de processamento de 350ms.

Primeiramente, foi realizada a coleta de fotos, as quais foram usadas para o treinamento e validação da rede. Com esse conjunto de imagens, foi possível construir um banco de imagens, localizando as mariposas presentes nas fotos.

Com um conjunto diverso de exemplos, a rede foi capaz de aprender a reconhecer padrões e características específicas das mariposas da broca de cana-de-açúcar.

Após o treinamento da rede com o banco de dados, consolidou-se os dados com os resultados de classificação obtidos. A rede neural demonstrou uma capacidade de identificar corretamente as mariposas em diferentes contextos e situações, validando assim a viabilidade do processamento de imagens para a identificação dessas pragas em armadilhas.

Por fim, é importante ressaltar que este trabalho não apenas atingiu os objetivos propostos, mas também abre caminho para futuras pesquisas e aprimoramentos na detecção automatizada de mariposas e outras pragas, promovendo avanços significativos no manejo e preservação das culturas agrícolas. Uma vez que tal ferramenta pode agilizar o processo de contagem de mariposas reduzindo a carga cognitiva dos técnicos de campo. Ou até mesmo a construção de modelos de armadilhas autônomas, que realiza a coleta de dados e envia a contagem de forma automática para a central de manejo de pragas da fazenda.

6 BIBLIOGRAFIA

AJALA, J. **Object Detection and Recognition Using YOLO: Detect and**, p. 23-24, 2021.

ALEXEY. AlexeyAB/darknet#how-to-train-to-detect-your-custom-objects, 2022. Disponível em: <<https://github.com/AlexeyAB/darknet#how-to-train-to-detect-your-custom-objects>>. Acesso em: 2023.

ALMERON, C. P. **ESTUDO EM PROCESSAMENTO DE IMAGENS DIGITAIS APLICADO AO RECONHECIMENTO DE PLACAS DE VEÍCULOS AUTOMOTORES BRASILEIROS**, p. 14, 2010.

CARVALHO, G. R.; OLIVEIRA, C. D. O setor sucroalcooleiro em perspectiva. **Embrapa**, 2006. Disponível em: <<https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1008450/o-setor-sucroalcooleiro-em-perspectiva>>. Acesso em: 2023.

DARKNET , 2023. Disponível em: <<https://pjreddie.com/darknet/yolo/>>. Acesso em: 2023.

FERNEDA, E. **Redes neurais e sua aplicação em sistemas de recuperação de informação**, p. 2, 2006.

FERREIRA, C. A. D. S. **INTENSIDADE DE INFESTAÇÃO DE Diatraea saccharalis (FABRICIUS, 1794)**, p. 13, 2013.

FIVEKO , 2023. Disponível em: <<https://fiveko.com/sobel-filter/>>. Acesso em: 2023.

GARCIA, J. F. Medidas de manejo integrado no controle da broca da cana-de-açúcar. **Revista Cultivar**, 2020. Disponível em: <<https://revistacultivar.com.br/noticias/medidas-de-manejo-integrado-no-controle-da-broca-da-cana-de-acucar>>. Acesso em: 2023.

GONÇALVES, L. reconhecimento-de-escrita-manual-com-redes-neurais-convolucionais. **medium.com**, 2023. Disponível em: <<https://medium.com/luisfredgs/reconhecimento-de-escrita-manual-com-redes-neurais-convolucionais-6fca996af39e>>. Acesso em: 2023.

HOLLERWEGER, M. M. **Aplicação de Visão Computacional no Auxílio ao Levantamento de Defeitos em Pavimento Rodoviário**, p. 18, 2019.

INPE, 2022. Disponível em: <http://www.dpi.inpe.br/spring/portugues/tutorial/introducao_pro.html>. Acesso em: 2023.

JIANG, Z. et al. **Real-time object detection method for embedded devices**, p. 2-3, 2013.

JOSEPH REDMON, A. F. **YOLOv3: An Incremental Improvement**, p. 1, 2018.

KUZE, G. T. **AVALIAÇÃO DE MÉTODOS DE SEGMENTAÇÃO POR DETECÇÃO DE BORDA EM SEDIMENTOS DE AMOSTRAS DE URINA**, p. 5, 2020.

LIU, G. **REAL-TIME OBJECT DETECTION FOR AUTONOMOUS DRIVING BASED**, p. 13, 2011.

NIED, A. **TREINAMENTO DE REDES NEURAS ARTIFICIAIS BASEADO EM SISTEMAS DE ESTRUTURA VARIÁVEL COM TAXA DE APRENDIZADO ADAPTATIVA**, p. 21, 2007.

OLIVATTO, T. F. **IDENTIFICAÇÃO AUTOMÁTICA DE RAMPAS DE ACESSIBILIDADE APOIADA POR VISÃO COMPUTACIONAL A PARTIR DE IMAGENS PANORÂMICAS STREET-LEVEL**, p. 42, 2021.

OLIVEIRA, A. C. S. D. researchgate. **www.researchgate.net**, 2023. Disponível em: <https://www.researchgate.net/figure/Figura-1-Rede-Neural-Artificial-Multicamadas_fig2_240772105>. Acesso em: 2023.

OPENCV. **opencv/cvat**, 2022. Disponível em: <<https://github.com/opencv/cvat>>. Acesso em: 2023.

PEREIRA, J. K. O. **USO DE VISÃO COMPUTACIONAL PARA RECONHECIMENTO DE IMAGENS DE FRUTAS EM IMAGENS RGB**, p. 28, 2021.

PIRES, W. O. **RECONHECIMENTO DE ESPECIES FLORESTAIS PELA FOLHA, UTILIZANDO REDES NEURAS CONVOLUCIONAIS**, p. 50, 2018.

REDMON, J. **Darknet: Open Source Neural Networks in C**, 2016. Disponível em: <<http://pjreddie.com/darknet/>>. Acesso em: 2023.

REDMON, J.; FARHADI, A. **YOLOv3: An Incremental Improvement**, p. 1, 2018.

RODRIGUES, V. B. **vitorborbarodrigues.medium.com**, 2022. Disponível em: <<https://vitorborbarodrigues.medium.com/conhecendo-a-vis%C3%A3o-do-computador-redes-neurais-convolucionais-e1c2b14bf426>>. Acesso em: 2023.

ROSA, A. L. D. **Classificação de Imagens de Frutas utilizando Aprendizado de Máquina. Classificação de Imagens de Frutas utilizando Aprendizado de Máquina**, p. 8, 2019.

SAKURAI, R. sakurai, 2023. Disponível em: <<https://www.sakurai.dev.br/cnn-mapreduce/>>. Acesso em: 2023.

SUK-JU HONG, S.-Y. K. E. K. C.-H. L. J.-S. L. D.-S. L. J. B. G. K. **Moth Detection from Pheromone Trap Images Using Deep Learning Object Detectors**, p. 1, 2020.

TAVARES, D. M. **A broca da cana-de-açúcar, Diatraea saccharalis (Fabr.). promip**, 2019. Disponível em: <<https://promip.agr.br/broca-diatraea-saccharalis-cultura-cana-acucar-outras-culturas-agricolas-brasil/>>. Acesso em: 2023.

THENMOZHI KASINATHAN, D. S. S. R. U. **Insect classification and detection in field crops using modern machine learning techniques** , p. 1, 2020.

USP-ESALQ. <https://www.cepea.esalq.usp.br/br/pib-do-agronegocio-brasileiro.aspx>, 2022. Disponível em: <<https://www.cepea.esalq.usp.br/br/pib-do-agronegocio-brasileiro.aspx>>. Acesso em: 2023.

VALMORBIDA, A. D. **DESENVOLVIMENTO DE UM SISTEMA DE PROCESSAMENTO DE IMAGENS PARA INSPEÇÃO AUTOMATIZADA UTILIZANDO FPGA**, p. 33, 2018.

WANG, C.-Y. A. B. A. A. L. H.-Y. M. <https://github.com/AlexeyAB/darknet>, 2022. Acesso em: 2023.