

**UNIVERSIDADE EVANGÉLICA DE GOIÁS – UNIEVANGÉLICA**

**CURSO DE ENGENHARIA DE SOFTWARE**

BRUNO SILVEIRA;

DOUGLAS PAULO DOS SANTOS;

KAROLINE NORONHA BARBOSA.

**Utilização de IA para Controle de Pragas na Agricultura**

ANÁPOLIS

2022

BRUNO SILVEIRA;  
DOUGLAS PAULO DOS SANTOS;  
KAROLINE NORONHA.

## **Utilização de IA para controle de pragas na Agricultura**

Trabalho de conclusão de curso apresentado como requisito parcial para a obtenção do título de Bacharel em Eng. de *Software* pela Universidade Evangélica de Goiás.

Orientador: Prof<sup>o</sup> a. Me. William Pereira dos Santos Júnior.

ANÁPOLIS

2022

## RESUMO

O uso de Inteligência Artificial para resolução de problemas complexos vem ganhando força no mundo inteiro. A empresa *Wittel*, em um dos seus artigos diz que a possibilidade de alcance de aplicação de algoritmos de *Machine Learning* é vasta. Com isso, é possível observar diversas áreas que começaram a usar esse tipo de tecnologia para se beneficiar de suas competências. Uma dessas áreas é a agricultura, onde a necessidade de resolução de problemas no cultivo é fundamental para o sucesso da colheita. Especialmente na cultura de soja, onde é comum o aparecimento de doenças e pragas ao longo do seu cultivo, a identificação de qualquer irregularidade na planta é importante para se utilizar o tratamento adequado e garantir produtividade na lavoura. Este trabalho visou o reconhecimento de duas das pragas mais comuns na soja, lagarta da soja e falsa-medideira, a partir da análise de imagens da folha utilizando algoritmos que implementam a arquitetura de *Machine Learning*. O trabalho irá envolver o desenvolvimento de um *software* para o treinamento de várias imagens e uma interface de usuário para poder classificar as imagens.

**Palavras-chave:** IA. *Machine Learning*. Soja.

## **ABSTRACT**

The use of Artificial Intelligence to solve complex problems has been gaining strength worldwide. The scope of application of Machine Learning algorithms is vast. That said, it is possible to observe several areas that have started to use this type of technology to benefit from their skills. One of these areas is agriculture, where the need to solve problems in cultivation is fundamental to the success of the harvest. Especially in soybean crops, where the appearance of diseases and pests is common throughout its cultivation, the identification of any irregularity in the plant is important to use the appropriate corrective and guarantee productivity in the crop. This work aimed at the recognition of two of the most common pests in soybean, from the analysis of leaf images, using algorithms that implement the Machine Learning architecture. The work will involve the development of software for training various images, and a user interface to classify the images.'

**Key-words:** IA. Machine Learning. Soybean.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Safra de 2020/2021 em Goiás .....	13
Figura 2 - Lagarta da soja.....	16
Figura 3 - Período de ocorrência da lagarta da soja .....	17
Figura 4 - Detalhe do dano causado pela lagarta da soja .....	17
Figura 5 - Lagarta falsa medideira .....	18
Figura 6 - Período de ocorrência da lagarta falsa medideira .....	18
Figura 7 - Dano causado pela falsa medideira .....	19
Figura 8 - Melhorando a performance do modelo.....	29
Figura 9 - Criação do modelo .....	29
Figura 10 - Compilação do modelo.....	30
Figura 11 - Treinando do modelo.....	30
Figura 12 - Gráfico de acurácia e perda do treino e validação antes das técnicas contra o overfitting .....	31
Figura 13 - Data augmentation .....	32
Figura 14 - Novo modelo com data augmentation a dropout aplicados.....	33
Figura 15 - Gráfico de acurácia e perda do treino e validação com as camadas data augmentation a dropout aplicadas .....	34
Figura 16 - Código para predizer uma imagem .....	35
Figura 17 - Exemplo de resposta depois do código executado .....	35

## SUMÁRIO

1. INTRODUÇÃO	7
1.1 DEFINIÇÃO DO PROBLEMA	8
1.2. OBJETIVOS	9
1.2.1 OBJETIVO GERAL	9
1.2.2 OBJETIVO ESPECÍFICO	9
2. REFERENCIAL TEÓRICO	10
2.1 AGRICULTURA	10
2.1.1 AGRICULTURA NO MUNDO	10
2.1.2 AGRICULTURA NO BRASIL	11
2.3. TIPOS DE PLANTAÇÕES	12
2.3.1 PLANTAÇÕES EM GOIÁS	12
2.3.2 PLANTAÇÕES DE SOJA	13
2.4 PRAGAS	14
2.4.1 PRINCIPAIS PRAGAS QUE ATACAM AS PLANTAÇÕES	14
2.5. LAGARTA DA SOJA ( <i>Anticarsia gemmatalis</i> )	15
2.5.1 DESENVOLVIMENTO DA LAGARTA	15
2.5.2 DANOS CAUSADOS	16
2.6 FALSA-MEDIDEIRA ( <i>CHRYSODEIXIS INCLUDENS</i> )	17
2.6.1 DESENVOLVIMENTO DA LAGARTA	17
2.6.2 DANOS CAUSADOS	18
2.7. <i>MACHINE LEARNING</i>	19
2.7.1 COMO FUNCIONA?	19
2.7.2 <i>MACHINE LEARNING</i> NA AGRICULTURA	20

2.7.3 O QUE JÁ TEM DE <i>MACHINE LEARNING</i> NO MUNDO	21
2.7.4 REDE NEURAL CONVOLUCIONAL	22
2.7.4.1 FUNCIONAMENTO CONVOLUCIONAL	23
2.8 OTIMIZADORES	24
2.9 FUNÇÕES DE PERDA	24
2.10 ACURÁCIA	24
2.11 FERRAMENTAS	25
2.11.1 <i>PYTHON</i>	25
2.11.2 <i>TENSORFLOW</i>	25
2.11.3 <i>KERAS</i>	26
3. METODOLOGIA	27
4. DESENVOLVIMENTO	28
4.1 CRIANDO E TREINANDO O MODELO	28
4.2 <i>OVERFITTING</i>	31
5. ANÁLISE E DISCUSSÃO DOS RESULTADOS	34
6. CONSIDERAÇÕES FINAIS	36
7.REFERÊNCIAS BIBLIOGRÁFICAS	37

## 1. INTRODUÇÃO

A aproximação entre homem e máquina sempre foi um assunto abordado pela ciência. Desde o início da Revolução Industrial, a partir da metade do século XVIII, começou a utilização de máquinas que substituíssem o trabalho humano e conseqüentemente aumentavam a velocidade de execução e a precisão das tarefas. A partir disso, existe um grande desafio tecnológico para construir um computador que se iguale ao cérebro humano. Mas, existem situações em que o computador sempre o vencerá, que foram circunstâncias que levaram à construção dele, situações que nada mais são do que a realização de tarefas.

Desde os anos 60, há uma necessidade de resolver problemas complexos, não mais uma tarefa simples, mas uma série de técnicas heurísticas permitirá que o software manipule informações e mostre-o na forma de conhecimento. Como característica a solução do problema limitado a uma área específica, acontecendo uma manipulação do conhecimento de um profissional humano. A partir disso, surgem os primeiros Sistemas especialistas, traçando uma linha tênue com a Inteligência Artificial moderna.

A ideia de que os computadores podem aprender e possuir inteligência autônoma não é nova, Alan Turing sugeriu essa possibilidade, não de forma explicativa, como se os computadores pudessem possuir inteligência semelhante à humana. Em seu artigo *Computing Machinery And Intelligence*, de 1950, ele propôs a questão se máquinas poderiam de fato pensar, através do Teste de Turing.

Os avanços que vieram com o crescimento do poder de computação tornaram possível resolver problemas extraindo grandes quantidades de conhecimento e grande quantidade de dados. O acúmulo dos dados por meio de algoritmos, visando conhecimento artificial, os pesquisadores apelidaram aprendizado de máquina.

As Redes Neurais Artificiais (RNAs) possuem a capacidade de aprender e se adaptar, e têm o potencial de extrair conhecimento, além de poderem se organizar por meio de grupos de conhecimento adquiridos, a partir da experiência no processamento de dados.



As redes neurais artificiais possuem aplicações em diversos setores econômicos, como indústria, agricultura e setor de serviços. São cada vez mais utilizados porque aliviam as pessoas de trabalhos complexos, reduzindo gastos e especialmente erros humanos.

A forma de aplicação que tornou as redes neurais artificiais amplamente conhecidas é o reconhecimento de imagens. Na medicina, estudou-se o uso do reconhecimento de imagens de RNA para identificar o câncer de pele (ESTEVA et al., 2017). Na área da agricultura já existem sistemas de identificação de doenças de plantas em lavouras inteiras, as imagens são adquiridas via satélite e processadas por meio de técnicas que geram imagens infravermelhas para uma inspeção completa do solo.

### 1.1 DEFINIÇÃO DO PROBLEMA

O uso de Inteligência Artificial (IA) para ajudar o trabalho humano, envolve em muitos casos um custo elevado e muito tempo para ser implementado. Na agricultura, um sistema que utiliza visão computacional para analisar uma plantação e saber o estado dela, requer um alto investimento. O mercado global de inteligência artificial na agricultura foi avaliado em cerca de US\$ 545 milhões, que agora está aumentando e previsto para chegar a quase US\$ 2.075 milhões até 2024. O mercado irá crescer a um CAGR crescente de 21% ao longo da previsão do período (REHAGRO, 2017). Isso mostra que a IA está disponível em sua maioria sob o domínio de grandes empresas do agronegócio. Com o intuito de analisar a presença de doenças e pragas, existem acompanhamentos de agrônomos para o pequeno e médio produtor. Esses acompanhamentos podem gerar altos custos para o produtor, bem como estarem sujeitos à disponibilidade do profissional.

Segundo o que foi divulgado pela Empresa Brasileira de Pesquisa Agropecuária nos números referentes à safra 2020/21 de soja no mundo. O Brasil atingiu o posto de maior produtor mundial do grão, uma cultura fomentada pelo grande, médio e pequeno produtor (EMPRAPA, 2021). Entre outras características, os campos de soja são facilmente atingidos por pragas, entre elas as mais comuns são *Anticarsia gemmatalis* mais conhecida como lagarta-da-soja e a *Chrysodeixis includens* conhecida como falsa-medideira. Como utilizar a inteligência artificial para detectar as pragas desenvolvidas na plantação?

## 1.2. OBJETIVOS

### 1.2.1 OBJETIVO GERAL

Utilizar soluções de *Machine Learning* para coletar e analisar imagens que mostram as condições das folhas da soja, e identificar se há ou não alguma praga a danificando.

### 1.2.2 OBJETIVO ESPECÍFICO

- Identificar corretamente os danos causados pelas pragas lagarta da soja e falsa-medideira.
- Pesquisar a ferramenta computacional que auxilie no processo de implementação de *Machine Learning*.
- Implementar um *software* que servirá de interface para os treinos, classificações, testes e análises, utilizando a biblioteca pesquisada.
- Treinar a IA com imagens suficientes para gerar uma base de dados consistente.
- Avaliar a eficácia do *software* criado a partir das análises dos resultados.

## 2. REFERENCIAL TEÓRICO

### 2.1 AGRICULTURA

Define-se “agricultura” como “o cultivo do solo por meio de procedimentos, métodos e técnicas próprias para produzir alimentos para consumo humano, como hortaliças, grãos, frutas e hortaliças, ou para serem utilizados como matéria-prima para a indústria” (AGRICULTURA, 2022). A palavra "agricultura" vem do latim e consiste em duas palavras "agru" e "colere" (cultura), "agru" significa "terra cultivada ou arável" e corresponde a "cultivo".

#### 2.1.1 AGRICULTURA NO MUNDO

Segundo Rodolfo F. Alves Pena, a agricultura é uma das principais atividades econômicas no mundo, trata-se de uma das formas principais de transformação do espaço geográfico, sendo uma das mais antigas práticas realizadas na história. Sua prática é estratégica para assegurar a disponibilidade de alimentos dos países, além de garantir uma importante fonte de renda. (PENA,2020). Os 4 maiores produtores agrícolas do mundo são China, Estados Unidos, Brasil e Índia (G1,2021).

De acordo com Luana Caroline Kunast, mestre em Geografia (KUNAST, 2020) a agricultura seja desenvolvida no mundo é necessário que os países mais desenvolvidos ofereçam incentivos governamentais impactando a inclusão dos demais países nas negociações. Isso acaba gerando algumas discussões entre os que defendem sob o argumento do desenvolvimento, e os que consideram injustos, pois acreditam que isso limita a ascensão dos países menos desenvolvidos economicamente.

Como funciona a agricultura de um país subdesenvolvido? Para um país onde a economia é baixa e com isso tornando o custo de vida das pessoas muito alto em relação ao seu ganho, a agricultura se torna a principal fonte de se manter, ou seja, usam como subsistência (Plantio para sobrevivência) com ferramentas e técnicas rudimentares.

Já em países desenvolvidos a agricultura se torna prioridade do governo onde recebem tais incentivos para a manutenção e seu desenvolvimento, uma de suas prioridades é manter sempre o abastecimento interno e não para exportações.

Hoje no mundo existe o protecionismo na agricultura que são medidas protecionistas que visam proteger os setores mais vulneráveis da economia de um país. As principais medidas protecionistas são: Subsídios; Barreiras alfandegárias; Barreiras Sanitárias e Fitossanitárias.

### 2.1.2 AGRICULTURA NO BRASIL

De acordo com a Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA,2020) a agricultura no Brasil é responsável por uma das principais fontes econômicas do País, este setor é que gera empregos em cerca de 22% da população ativa no Brasil e também é um dos setores que mais contribui para o crescimento do PIB nacional e que responde por 21% de todas as riquezas produzidas.

O Brasil já passou por várias transformações no setor agrícola onde no início sua produção era pautada principalmente na Cana-de-Açúcar onde segundo o Instituto Brasileiro de Geografia e Estatística (IBGE,2022) a produção por ano da safra e produto medidos em toneladas até março de 2022 chegou a produção de 600 milhões e logo atrás com média de 100 milhões temos a produção de Soja que começou a ser produzida por conta das transformações e expansões.

Graças ao avanço das indústrias, a aceleração do processo de urbanização colocou o campo economicamente subordinado à cidade, tornando-o dependente das técnicas e produções industriais, podendo assim dizer que a modernização da agricultura no país está diretamente ligada a esses fatores.

Rafaela Sousa (2020) mostra que no Brasil a agricultura tem suas características divididas por regiões, sendo elas: Região Sul; Região Sudeste; Região Nordeste; Região Centro-Oeste e Região Norte.

- Região Sul: Caracterizada pela ocupação de grupos imigrantes europeus, pela expansão da soja voltada para exportação e pela modernização agrícola. Além da soja, se cultivam o milho, a cana-de-açúcar e o algodão.
- Região Sudeste: É a região em que a agricultura se encontra mais subordinada à indústria. Os principais cultivos são o café, a cana-de-açúcar e a fruticultura.

- Região Nordeste: Na zona mais úmida predomina-se o cultivo de cana voltada para a produção de álcool e açúcar e nas áreas semiáridas predomina-se a agricultura familiar, e o principal cultivo é o de frutas.
- Região Centro-Oeste: É a área que mais se expande o cultivo pela produção mecanizada e seu principal produto é a soja, voltada para o mercado externo.
- Região Norte: É caracterizada por receber atualmente as principais frentes de expansão. As principais atividades praticadas são de caráter extensivo de de baixa tecnologia, com ênfase na pecuária primitiva, e na soja em expansão.

### 2.3. TIPOS DE PLANTAÇÕES

Rafaela Sousa (2020) ainda diz que, no mundo existem diversos tipos de formas de cultivo. O que é mais determinante para a escolha para o tipo de agricultura são características das regiões e a dimensão territorial e a quantidade de dinheiro a ser investido. Os tipos mais comuns de agricultura são:

- Agricultura Familiar: segundo o Censo Agropecuário realizado pelo IBGE (2017), 77% das propriedades agrícolas do Brasil são adeptas a esse sistema. O qual, como o próprio nome diz, é o tipo em que as pessoas da mesma família trabalham na terra. É responsável pelo abastecimento interno de alimentos do país.
- Agricultura Patronal: é voltada para a exportação. Requer trabalhadores qualificados e grande investimento para assegurar a rentabilidade do negócio.
- Agricultura Extensiva: utiliza técnicas tradicionais para produção, além do baixo investimento no maquinário e em tecnologias. O que ocasiona na baixa produtividade.
- Agricultura Moderna: é fundamentado na altíssima quantidade de tecnologias utilizadas. O exemplo disso é a agricultura de precisão e agricultura digital nas quais ganharam espaço desde 2010 e tendem a crescer ainda mais com o surgimento de conhecimentos em Inteligência Artificial (IA) e Internet das Coisas (IOT).

#### 2.3.1 PLANTAÇÕES EM GOIÁS

Embora tenhamos uma crescente industrialização, a agropecuária continua sendo uma das atividades econômicas mais importantes em Goiás, visto que a produção de

carnes e grãos impulsionou as exportações. Segundo a Companhia Nacional de Abastecimento (CONAB, 2021), no ano de 2020 Goiás passou a ser terceiro maior produtor de grãos do Brasil, ficando atrás somente do Mato do Grosso e Paraná.

O estado é o terceiro produtor nacional de grãos com uma produção em torno de 27,1 milhões de toneladas, o que representa 10,8% da produção de grãos brasileira. A pauta agrícola é bastante diversificada e composta principalmente por: soja, sorgo, milho, cana-de-açúcar, feijão, tomate, entre outros produtos (CONAB, 2020).

<b>SAFRA 2020/2021 GOIÁS (GRÃOS) - CONAB</b>				
<b>Produto</b>	<b>ÁREA (mil ha)</b>	<b>PRODUTIVIDADE (kg/ha)</b>	<b>PRODUÇÃO (mil ton)</b>	<b>Posição no ranking da produção nacional</b>
Grãos (geral)	6.123,5	4.456	27.285,3	4°
Soja	3.694,0	3.636	13.431,4	4°
Milho	1.815,8	6.560	11.911,7	3°
Sorgo	374,9	3.443	1.290,8	1°
Feijão	141,9	2.324	329,8	5°
Caroço de algodão	27,0	2.495	67,4	7°
Trigo	23,1	4.000	92,4	6°
Arroz	24,2	5.257	127,2	8°
Gergelim	3,0	500	1,5	2°
Girassol	19,6	1.689	33,1	2°

Nota: parte da área cultivada é sobreposta, ou seja, existem áreas que são cultivadas com mais de uma cultura por safra.

Figura 1

Fonte: CONAB.

Elaboração: Agricultura.Go.gov

### 2.3.2 PLANTAÇÕES DE SOJA

No 3° levantamento da Safra de Grãos no final do ano de 2020, mostrou que a soja é continua sendo principal grão produzido pelo Estado, em termos de volume, com aproximadamente de 13,4 milhões de toneladas produzidas no Estado. O número simboliza um aumento de 2,1% na produção em relação à safra 2019/2020, e era esperado em uma área plantada de 3,7 milhões de hectares por todo o Estado (aumento da área de 4,2% em relação à safra anterior) (CONAB, 2020). O Instituto para o Fortalecimento da Agropecuária

de Goiás (Ifag), reforçou o que foi dito pelo Conab dizendo que, a expectativa para o ano de 2021 era de crescimento de aproximadamente 4% na área plantada.

## 2.4 PRAGAS

As pragas consistem em uma população de organismos capazes de causar danos às plantas, seus produtos e subprodutos. Segundo o livro “Defesa Vegetal Fundamentos, Ferramentas, Políticas e Perspectivas” (2015), divulgado pela FAEMG (Federação da Agricultura e Pecuária do Estado de Minas Gerais), as perdas no Brasil podem chegar a R\$55 bilhões por ano em decorrência desse mal.

### 2.4.1 PRINCIPAIS PRAGAS QUE ATACAM AS PLANTAÇÕES

Segundo dados do Ministério da Agricultura, Pecuária e Abastecimento (MAPA,2020) e da Embrapa, as principais pragas que atingem a lavoura são:

- A lagarta da Soja (*Anticarsia gemmatilis*): é uma das pragas da soja de hábitos noturnos que, quando dia, fica em áreas sombreadas. Inicialmente as lagartas são de coloração verde-clara e possuem quatro pares de pernas no abdômen, sendo duas vestigiais. Os danos causados por essa praga são raspagens inicialmente em pequenas áreas das folhas. Quando as lagartas são maiores, alimentam-se da folha deixando grandes “buracos” ou mesmo se alimentando da folha inteira. A desfolha pode chegar a 100% se a lagarta-da-soja não for monitorada e controlada corretamente.
- A lagarta do cartucho (*Spodoptera frugiperda*): é uma espécie canibal e, por isso, geralmente são encontradas poucas lagartas por planta. Essa espécie penetra no colmo, criando galerias, e causa danos consumindo as folhas.
- Lagarta Elasm (*Elasmopalpus lignosellus*): Geralmente é cíclica, mas os surtos em soja têm sido frequentes, principalmente em solos arenosos e em anos com estiagem prolongada, na fase inicial das lavouras. Quando pequenas, as lagartas alimentam-se raspando o parênquima foliar.
- Lagarta-falsa-medideira (*Chrysodeixis includens*): Atacam as folhas, raspando-as enquanto são pequenas, ocasionando pequenas manchas claras; à medida que crescem, ficam vorazes e destroem completamente as folhas, podendo danificar até as hastes mais finas.
- Mosca-branca (*Bemisia sp.*): apesar do nome comum de mosca, trata-se de um inseto sugador comum em diversas culturas. Os insetos têm ação toxicológica, sendo que os maiores prejuízos são devido à transmissão de viroses.

- Percevejo-castanho (*Scaptocoris spp*): é um inseto que apesar de ter ocorrência em todo Brasil, no entanto, tem causado frequentes danos à lavoura da região dos cerrados. Seus danos são provocados por ninfas e adultos, que possuem hábito subterrâneo e fazem a sucção da seiva das raízes, causando atrofiamento das raízes e subdesenvolvimento das plantas.
- Bicudo-da-soja (*Sternechus subsignatus*): é considerado um inseto de difícil controle e vem ganhando importância pelos danos que tem causado às lavouras de vários municípios do Paraná, de Santa Catarina e do Rio Grande do Sul. O adulto raspa o caule e desfia os tecidos no local do ataque. Quando a população é alta e ocorre na fase inicial da cultura, o dano é irreversível e as plantas morrem podendo haver perda total de parte da lavoura.

## 2.5. LAGARTA DA SOJA (*Anticarsia gemmatalis*)

As larvas da lagarta da soja (*Anticarsia gemmatalis*) há muito são consideradas uma das principais pragas da soja, ocorrendo em todas as regiões produtoras do Brasil. Com a introdução da tecnologia Bt em 2010, *A. gemmatalis* tornou-se menos abundante nas lavouras, porém continua sendo uma praga economicamente importante devido ao seu potencial de causar perdas significativas de produtividade nas lavouras.

### 2.5.1 DESENVOLVIMENTO DA LAGARTA

As lagartas da soja são insetos desfolhadores que se desenvolvem em holometabolismo. Ocorre de novembro a março, atingindo seu pico populacional a partir de janeiro (Andrade et al., 2004). A fase adulta dura em média 15 dias. A mariposa tem envergadura de 30 a 38 mm e varia de cinza a marrom avermelhado ou amarelado, sem dimorfismo sexual. A posição dos ovos ocorre na parte inferior das folhas e os ovos eclodem em três dias. O estágio larval do inseto pode durar até 25 dias, com desenvolvimento de seis ínstars consecutivos.

O estágio larval do inseto pode durar até 25 dias, com desenvolvimento de seis ínstars consecutivos (Andrade et al., 2004). As lagartas da soja são verdes quando pequenas (até 1 cm) e possuem quatro pares de gastrópodes. Nesta fase, pode se movimentar como se estivesse “medindo palmos” confundindo-se com lagartas falsas-medideiras na hora do reconhecimento.



Lagartas maiores que 1,5 cm são verdes ou escuras. Possuem 3 faixas longitudinais brancas no dorso e 4 pares de gastrópodes, além de um par de patas anais, como mostrado na imagem 2.



Figura 2  
Fonte: Moscardi et al., 2012.  
Lagarta da soja, *Anticarsia gemmatilis*.

## 2.5.2 DANOS CAUSADOS

Os principais danos às lavouras de soja ocorrem durante a fase larval de *A. gemmatilis*, quando as lagartas inicialmente arranham as folhas da soja e causam danos consideráveis à medida que crescem. As lagartas se alimentam de folhas e até das veias das folhas e podem causar 100% de desfolha, como mostra na imagem 4. Cada lagarta pode comer uma média de 90 centímetros quadrados de folhas a cada 24 horas, ou 2,1 vezes sua própria massa, antes que o desenvolvimento larval esteja completo (Sosa-Gómez et al., 2010).

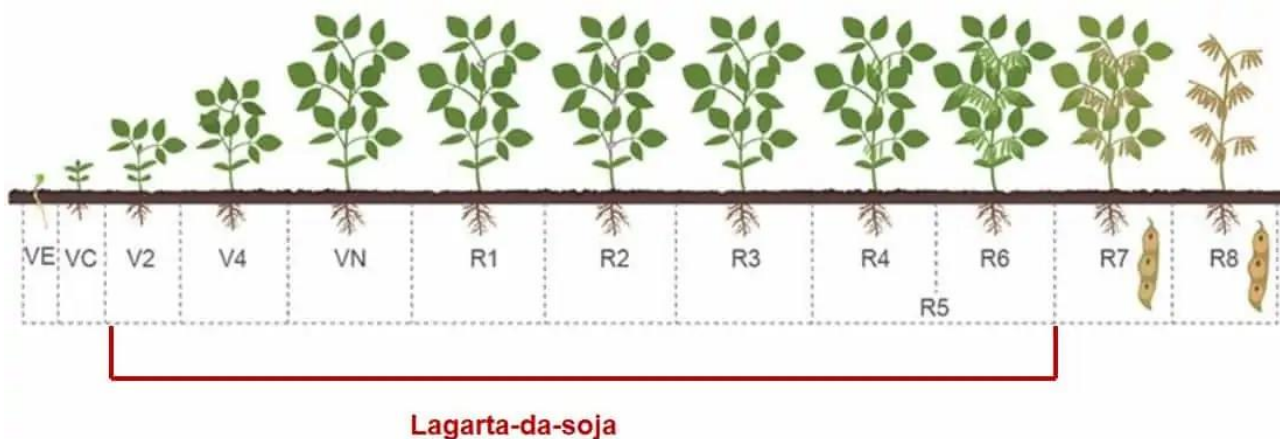


Figura 3

Fonte: Imagem adaptada de Instituto Phytus

Período de ocorrência da lagarta da soja, *Anticarsia gemmatalis*.

Elas, por serem desfolhadoras, começam a aparecer no estágio vegetativo V2, como mostra na imagem 3. A lagarta da soja se alimenta do terço superior das plantas, e o desfolho vai aumentando de acordo com o que vai se tornando maior. Dependendo da quantidade de indivíduos da espécie no local, podem se alimentar até mesmo de flores e vagens.



Figura 4

Fonte: Revista Cultivar

Dano causado pela lagarta da soja, *Anticarsia gemmatalis*.

## 2.6 FALSA-MEDIDEIRA (*CHRYSODEIXIS INCLUDENS*)

No Brasil, a falsa-medideira (*Chrysodeixis includens*) tornou-se um grave problema fitossanitário na cultura da soja, com vários surtos ocorrendo isoladamente ou em associação com lagartas da soja.

### 2.6.1 DESENVOLVIMENTO DA LAGARTA

Elas são assim chamadas porque têm apenas dois pares de pernas abdominais, então sua locomoção é como "medir uma mão" (Sosa-Gómez et al., 2010), como apresentado na imagem 5. Constituem um complexo de espécies associado à soja, sendo *Chrysodeixis* a espécie mais importante. As larvas recém-eclodidas são verdes pálidas com

listras verticais brancas e pontos pretos. Atingem de 40 a 45 mm de comprimento no estágio larval final.



Figura 5

Fonte: Pionner Sementes  
Lagarta falsa medideira, *Chrysodeixis includens*..

## 2.6.2 DANOS CAUSADOS

Da larva de primeiro ao terceiro ínstar, as lagartas selecionam como alimento as folhas jovens com baixo teor de fibra; por outro lado, as lagartas mais desenvolvidas tornam-se menos exigentes e passam a se alimentar de folhas mais velhas e fibrosas e geralmente são encontradas em rações na terço inferior da planta de soja e as folhas jovens na parte aérea secundária. Elas começam a aparecer no estágio vegetativo V4, como mostra na imagem 6. No primeiro e segundo ínstares apenas raspam as folhas, e a partir do terceiro ínstar conseguem perfurar as folhas.

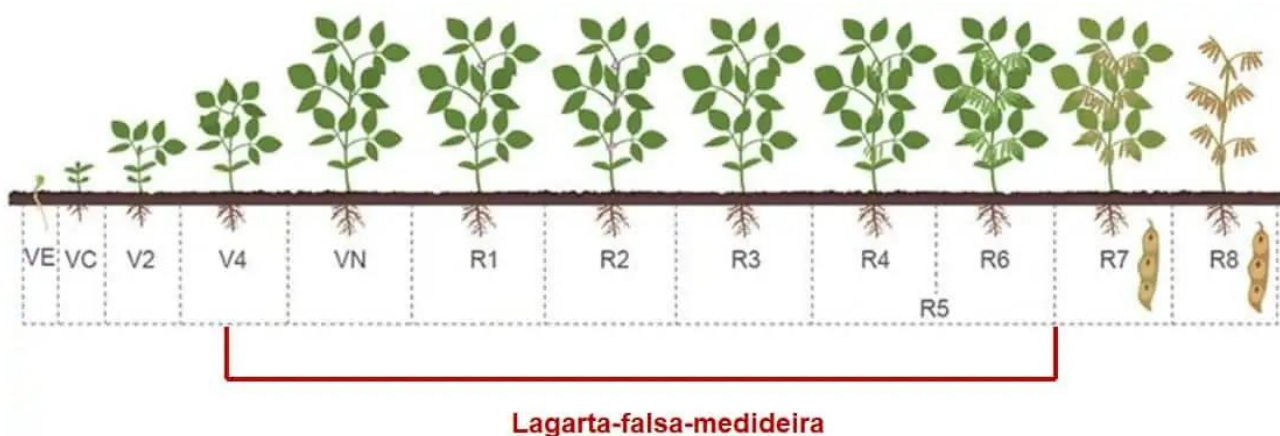


Figura 6

Fonte: Imagem adaptada de Instituto Phytus  
Período de ocorrência da lagarta falsa medideira, *Chrysodeixis includens*.

Ao contrário dos danos causados por outros desfolhadores, quando as lagartas de *C. includens* perfuram as folhas, deixam intactas as nervuras central e lateral, conferindo o aspecto característico de folhas rendilhadas, como apresentado na imagem 7.



Figura 7

Fonte: Moscardi et al., 2012.

Dano causado pela falsa medideira, *Chrysodeixis includens*.

## 2.7. MACHINE LEARNING

IBM Cloud Education (2020), explicou que *Machine Learning* é uma subcategoria da inteligência artificial, que se refere ao processo em que o computador desenvolve o reconhecimento de padrões, ou a capacidade de continuamente aprender e fazer previsões com base em dados, em seguida, depois desse treinamento consegue fazer ajustes de forma automatizada, eles aprendem com computações anteriores para produzir decisões e resultados confiáveis, passíveis de repetição.

### 2.7.1 COMO FUNCIONA?

Marcos Takana (2018), desenvolvedor de *software* na DB1 Global, mostra que Machine Learning pode ser complexa e como funciona dependerá de qual objetivo e do algoritmo usado para alcançá-lo. No entanto, o modelo continua o mesmo, o computador recolhe os dados e identifica padrões, e usando os resultados para melhor completar sua tarefa. Qualquer tarefa que dependa de um conjunto de pontos de dados ou regras que podem ser automatizadas usando o machine learning, até as mais complexas tarefas tais como responder a chamadas do Atendimento ao Cliente e revisar currículos.

Dependendo da situação, os algoritmos de machine learning funcionam usando muita ou pouca interação humana. Os quatro principais modelos do machine learning são, *supervised learning*, *unsupervised learning*, *semi-supervised learning* e *reinforcement learning*.

***Supervised Learning***: o computador é fornecido com um conjunto rotulado de dados que permite que ele aprenda como realizar uma tarefa humana. Este é o modelo menos complexo, pois tenta replicar o aprendizado humano.

***Unsupervised Learning***: é fornecido para o computador dados não rotulados e extrai dele padrões anteriormente desconhecidos. Existem diversas maneiras diferentes pelas quais os algoritmos de machine learning fazem isso, incluindo:

- *Clustering*, no qual o computador encontra pontos de dados semelhantes em um conjunto de dados e os agrupa de acordo.
- Estimaco de densidade, na qual o computador descobre insights observando como um conjunto de dados é distribuído.
- Deteco de anomalia, na qual o computador identifica pontos de dados em um conjunto que são significativamente diferentes do restante dos dados.
- Análise de componentes principais (PCA), na qual o computador analisa um conjunto de dados e os resume para que possa ser usado para fazer previsões precisas.

***Semi-supervised learning***: é fornecido um conjunto de dados parcialmente rotulados e executa sua tarefa usando os dados rotulados para entender os parâmetros para interpretar os dados não rotulados.

***Reinforcement learning***: o computador observa seu ambiente e usa esses dados para identificar o comportamento ideal que minimizará o risco e/ou maximizará a recompensa. Esta é uma abordagem interativa que requer algum tipo de sinal de reforço para ajudar o computador a identificar melhor sua melhor ação.

## 2.7.2 MACHINE LEARNING NA AGRICULTURA

No contexto da agricultura, estes algoritmos operam em grandes bancos de dados que contêm décadas de informações sobre o clima, produtividade, características do solo, sobre as pragas entre outras inúmeras informações, assim contribuindo para a área de prática agrícola fazendo com que o produtor rural consiga ser mais eficiente, podendo

auxiliar ele em todas as etapas do plantio e no bem-estar animal, sendo extremamente útil na gestão de espécies. Conseguindo, por meio de análise de banco de dados que contenha o desempenho de diferentes culturas em variados climas, construir um modelo que irá fornecer quais genes provavelmente irão contribuir com uma característica que será benéfica para uma cultura. Sendo o *machine learning* mais eficiente no reconhecimento de espécies do que um humano.

Auxiliando também no gerenciamento das condições do campo, por meio de algoritmos que analisam processos relacionados ao solo, como a umidade, temperatura e evaporação, fazendo com que o agricultor compreenda com muito mais assertividade como funciona a dinâmica natural do ecossistema na região de plantio ou de pasto. Também podendo investigar o impacto que a agricultura ou pecuária podem causar sobre o solo.

Assim, a partir dessas análises é possível produzir estimativas sobre melhor utilizar o sistema de irrigação. Destacando também que o *machine learning* é essencial para automatização da propriedade rural.

### 2.7.3 O QUE JÁ TEM DE *MACHINE LEARNING* NO MUNDO

O *machine learning* opera 24h por dia, sete dias por semana bem na palma de nossas mãos, temos vários exemplos dos mais variados tipos sendo alguns deles:

- **O que consumir:** *Netflix, Amazon Prime Video, Spotify, HBO Max*, entre outros serviços de *Streaming*, o *machine learning* é utilizado principalmente na personalização, aprimorando constantemente os algoritmos de recomendações, além de moldar o catálogo de filme, séries, podcasts e músicas de acordo com a resposta do usuário.
- **Melhor caminho:** *Uber, Google maps, Waze*, entre outros. Graças ao *machine learning* estes aplicativos sinalizam o tempo estimado de chegada, o horário de destino, qual melhor rota, se tem engarrafamento, se está interditado, entre outras coisas.
- **Ads em redes sociais:** os anúncios em navegadores e redes sociais funcionam na base de leilões. O vencedor terá seu anúncio exibido ao usuário. A estratégia torna

as campanhas mais lucrativas, por meio de taxas de CTR preditivas e estimativas de conversão com base no comportamento do usuário.

- **Google Tradutor:** máquina aprendendo outros idiomas, graças ao machine learning as traduções se tornam cada vez mais precisas, permitindo que a ferramenta aprendesse de acordo com as pesquisas dos usuários.
- **Assistentes virtuais:** sistema de *chatbot* principalmente desenvolvido para atender os clientes no pós-venda. Um exemplo é a Lu da Magalu.

#### 2.7.4 REDE NEURAL CONVOLUCIONAL

Segundo o capítulo 40 do livro Deep Learning Book (2022), uma Rede Neural Convolutiva (CNNs) ou RNC é um algoritmo de deep learning, utilizado para classificação e reconhecimento de imagens que é capaz de captar uma imagem de entrada e atribuir importâncias e diferenciar aspectos e objetos presentes nela, além disso as CNNs conseguem desempenhos de formas consistentes e eficaz em algumas tarefas visuais complexas. A estrutura de uma rede neural convolutiva é dividida em três objetivos principais. São elas:

- **Extração de Características:** É possível extrair as características locais ao receber os sinais de entrada de um campo receptivo.
- **Mapeamento de Características:** Cada camada computacional da rede neural convolutiva é composta por diferentes mapas de características, que dão robustez ao modelo, fazendo com que seja possível lidar com variações, como distorção, rotação ou translação das imagens.
- **Subamostragem:** A camada de subamostragem é aplicada para coletar amostras de cada mapa de características, produzindo uma sumarização.

O bloco de construção mais importante de uma CNN é a camada convolutiva, onde os neurônios na primeira camada convolutiva não estão conectados a cada pixel

de uma imagem de entrada, mas apenas aos pixels em seus campos receptivos. Já na segunda camada convolucional, está conectado apenas aos neurônios localizados dentro de um pequeno retângulo na primeira camada permitindo que a rede se concentre em recursos de baixo nível na primeira camada oculta, e em seguida monte-os em recursos de nível superior na próxima camada oculta, e assim por diante. Sendo uma estrutura hierárquica é comum em imagens do mundo real, que é uma das razões pelas quais as CNNs funcionam tão bem no reconhecimento de imagens.

#### 2.7.4.1 FUNCIONAMENTO CONVOLUCIONAL

Diferente de uma Rede Neural comum, as camadas das CNNs possuem neurônios organizados em 2 dimensões: Largura, Altura e Profundidade. Em um exemplo onde as imagens de entrada do nosso *dataset* é um volume de entrada de ativações e o volume tem dimensões 32 x 32 x 3 (Largura, altura e profundidade, respectivamente).

Logo uma CNN simples é composta por uma sequência de camadas, e cada camada de um CNN transforma um volume de ativações em outro através de uma função diferenciável. As principais camadas para criar arquiteturas CNN, empilharemos essas camadas para formar uma arquitetura CNN completa:

- **INPUT:** A camada de Input [32 x 32 x 3] manterá os valores brutos de pixel da imagem; nesse caso, uma imagem de largura 32, altura 32 e com três canais de cores R, G, B.
- **CONV:** A camada irá calcular a saída de neurônios conectados às regiões locais na entrada, cada um calculando um produto escalar entre seus pesos e uma pequena região à qual eles estão conectados no volume de entrada. Isso pode resultar em volumes como [32 x 32 x 12] se decidimos usar 12 filtros.
- **RELU:** A camada irá aplicar uma função de ativação elementar, com o limite máximo (0, X) em zero. Isso deixa o tamanho do volume inalterado [32 x 32 x 12]).



- **POOL:** A camada *pool layer* irá realizar uma operação de downsampling ao longo das dimensões espaciais (Largura, Altura), resultando em volume como [16x16x12].
- **FLATTEN:** O papel da camada *Flatten Layer* é realizar uma operação de nivelamento na saída da camada anterior, de forma que suas dimensões possuam a mesma forma da camada seguinte, ou seja, você precisaria "desempilhar" todo esse tensor multidimensional em um tensor 1D muito longo.
- **FULLY CONNECTED:** Irá calcular os resultados do processo de convolução/pool e em seguida irá classificar a imagem em um rótulo (Planta ou Inseto). A saída da camada anterior é achatada em um único vetor de valores, cada um representando uma probabilidade de que um determinado recurso pertença a um rótulo específico.

## 2.8 OTIMIZADORES

É importante trocar os pesos do modelo durante o processo de treinamento, para fazer as previsões mais corretas e otimizadas o possível. Os otimizadores conectam a função de perda e os parâmetros do modelo, atualizando o modelo em resposta à saída da função de perda. Ele molda o modelo da forma mais precisa possível lidando com os pesos.

## 2.9 FUNÇÕES DE PERDA

É um método de avaliação que envolve a função de previsão do modelo para dizer o quão bom o seu algoritmo modela o seu conjunto de dados, se ela retornar um valor alto é porque as suas previsões estão erradas, e se retornar um valor baixo elas estão corretas. A função de perda que guia para onde vai o código, dizendo se o algoritmo está indo na direção certa ou errada.

## 2.10 ACURÁCIA

A acurácia do modelo é a medida que determina se o modelo está bom em reconhecer os relacionamentos e padrões entre o conjunto de dados, com base no treinamento dos dados. Quanto melhor o modelo pode generalizar para dados não vistos antes, melhor serão as previsões que ele irá produzir, gerando um alto valor de acurácia.

## 2.11 FERRAMENTAS

Para o desenvolvimento estamos utilizando a linguagem de programação *Python* e a ferramenta *TensorFlow* que é uma biblioteca de código aberto para criação e treinamentos de redes neurais para detectar e decifrar padrões que também fornece a ferramenta *Keras* que roda por cima do *TensorFlow* que vai focar diretamente nas imagens.

### 2.11.1 PYTHON

*Python* é uma linguagem de programação de alto nível, interpretada por script, imperativa, orientada a objetos, funcional, de tipagem dinâmica e forte. Foi lançada por Guido van Rossum em 1991. Prioriza a legibilidade do código sobre a velocidade ou expressividade. Considerada pelo público a terceira linguagem “mais amada”, de acordo com uma pesquisa conduzida pelo *Stack Overflow* em 2018 e está entre as 5 mais populares, de acordo com uma pesquisa da *RedMonk*.

Desde o começo *Python* é utilizado para fins científicos e numéricos e *Machine Learning* é basicamente computação numérica *Python* se torna bastante útil para *Machine Learning*. Especialmente com a introdução do *TensorFlow*, a maior plataforma de *Machine Learning*, o *Python* se tornou a linguagem para aplicações no aprendizado de máquina.

### 2.11.2 TENSORFLOW

Segundo o Didática Tech (2022), *TensorFlow* é uma biblioteca de programação de código aberto que auxilia na construção de aprendizado artificial profundo, utilizável em computadores pessoais, dispositivos móveis, navegadores *web* e servidores. Oferece múltiplos níveis de abstração, abrangendo uma grande variedade de solução de problemas e flexibilizando o Aprendizado de Máquina, envolvendo áreas como reconhecimento de voz, visão computacional, robótica, recuperação de informação, processamento de linguagem, extração de informações geográficas e descoberta de drogas computacionais.

Inicialmente chamado de DistBelief, foi criado por uma equipe no segmento do Google voltado para pesquisa na área de inteligência artificial voltado para *deep learning* e redes neurais. Essa equipe de desenvolvimento promove pesquisas que aprimoram o estado da arte e as aplicam em campo, com políticas de boas práticas que visam a acessibilidade de IA para qualquer pessoa, atuando em áreas de grande potencial transformador e de

impacto na sociedade, como saúde, segurança, energia, transporte, manufatura e entretenimento.

*TensorFlow*, realiza computações em fluxo utilizando tensores. Um tensor é a generalização de matrizes e vetores para dimensões mais altas, ou seja, ele é um *array* multidimensional. Quando se utiliza essa biblioteca na programação, o tensor é o objeto principal do programa, representando uma computação parcialmente definida que retorna algum valor. Os programas desenvolvidos com essa ferramenta trabalham inicialmente construindo um grafo de objetos com os tensores, detalhando a forma que cada um é calculado, com base em tensores disponíveis, executando partes deste grafo para alcançar os resultados desejados.

O *TensorFlow* permite treinamento e inferência em grande escala, dispendo de centenas de servidores remotos para treinamento rápido, executando modelos treinados para inferência em diversas plataformas. Utiliza um grafo de fluxo de dados para retratar a computação de um algoritmo e o estado em que ele assume, inspirado por modelos de programação de alto nível. Em sistemas de fluxo de dados tradicionais, os vértices do grafo representam computação funcional e informações que não podem ser alteradas. No *TensorFlow*, o fluxo de dados permite que os vértices reproduzam cálculos que alteram o estado mutável do dado.

### 2.11.3 KERAS

O Didática Tech (2022) diz que *Keras* é uma biblioteca de código aberto criada para *deep learning* com *Python*. Ele é utilizado na criação de redes neurais para resolução de várias tarefas diferentes, como classificação de imagens, detecção de objetos e regressão. Para sua utilização é necessária a criação de códigos similares aos aplicados na linguagem *Python*, e por ser de código aberto qualquer pessoa pode utilizar gratuitamente.

*Keras* foi desenvolvido com base no *TensorFlow* e é de fácil uso, quando comparado com outros frameworks como o próprio *TensorFlow*. O *Keras* é largamente utilizado por pesquisadores e profissionais de Inteligência Artificial que desejam criar e experimentar modelos e precisam de uma interface de usuário. Muitas empresas utilizam o *Keras* em suas aplicações como: *Uber*, *Coca-Cola*, *PepsiCo*, entre outras.

Este também facilita a construção e o treinamento de redes neurais, fornecendo uma ampla variedade de redes neurais pré-treinadas que podem ser usadas para várias tarefas. Ele foi desenvolvido com foco em permitir a experimentação rápida.

### **3. METODOLOGIA**

Método de pesquisa exploratória que tem uma maior proximidade com o objeto de estudo pesquisado. Foi utilizada uma abordagem qualitativa pois o foco principal é o treinamento do algoritmo para o reconhecimento dos danos causados, e não focando nos números como resultados. Vale ressaltar que, iremos identificar a praga a partir do seu dano.

Foi utilizado pesquisa documental para coletar informações sobre a agricultura, as pragas, em sites especializados e revistas no setor da agricultura e das novas técnicas em agricultura de precisão, e o reconhecimento de imagem em sites especializados.

O banco de imagens para o sistema de treino e comparação da IA será coletado em repositórios de imagens da internet, e em sites especializados em agricultura como o site da Embrapa (Empresa Brasileira de Pesquisa Agropecuária). Nas plantações as imagens serão coletadas pelos drones.

A análise desses dados foi feita por um algoritmo de reconhecimento de imagem que foi desenvolvido e treinado, para atingir uma porcentagem de acerto que gere um alto nível de confiança nos resultados.

Foi utilizada a aprendizagem supervisionada para a solução escolhida, para Stuart e Peter (2013, p. 808): “Na aprendizagem supervisionada, o agente observa alguns exemplos de pares de entrada e saída, e aprende uma função que faz o mapeamento da entrada para a saída”. O modelo aprende a partir de resultados pré-definidos, estes valores servem como uma supervisão das previsões que são realizadas, permitindo assim o ajuste nas previsões com base nos erros. É supervisionado pois possui uma referência do que está certo e errado.

A última camada do modelo *fully connected* ou também conhecida como *dense*, de acordo com a documentação do *Keras* tem como objetivo calcular uma função de ativação em conjunto com nossos dados de entrada e pesos. Essa é uma camada profundamente conectada com a camada anterior, o que significa que os neurônios dessa camada são conectados com cada neurônio da camada anterior.

A função de ativação é usada para transformar os valores de entrada dos neurônios, basicamente decidem se um neurônio deve ser ativado ou não. Foi utilizado a função Unidade Linear Retificada (*Rectified Linear Unit*, ReLU). Segundo o capítulo 8 do livro *Deep Learning Book*, a função ReLU é não linear, quando houver um valor negativo na entrada, esse valor será convertido em zero, e o neurônio não será ativado. Sua função é definida como a parte positiva de seu argumento.

## 4. DESENVOLVIMENTO

A abordagem escolhida para o reconhecimento das imagens foi a rede neural convolucional, segundo o *Machine Learning Glossary* esse tipo de rede neural profunda é usado para processar estruturas de matrizes de dados. Comumente empregado em tarefas de visão computacional.

### 4.1 CRIANDO E TREINANDO O MODELO

Primeiramente é feito o carregamento das fotos, e a criação do modelo de dados, dividido entre conjunto de treinamento e de validação e é utilizado a regra 80/20, 80% das imagens para treinamento e 20% para validação. É necessário padronizar as imagens para trabalhar com a rede neural procurando tornar os valores de entrada pequenos, como são imagens coloridas os valores variam de 0 a 255, é aplicado uma camada de redimensionamento para colocar os valores entre 0 e 1.

Em seguida, o conjunto de dados é configurado para um melhor desempenho na hora de realizar o treinamento, como apresentado na imagem 8, para coletar os dados do disco sem bloquear a entrada e saída do disco. É feito um cache dos dados durante a primeira época do treinamento, mantendo as imagens carregadas na memória até o fim da

execução do modelo, e um prefetch dos dados para aumentar a velocidade de pré-processamento das imagens.

```
AUTOTUNE = tf.data.AUTOTUNE

treino = treino.cache().shuffle(1000).prefetch(buffer_size=AUTOTUNE)
validacao = validacao.cache().prefetch(buffer_size=AUTOTUNE)
```

Figura 8

Fonte: Autores.

Melhorando a performance do modelo.

Para a criação do modelo de dados foi utilizado um modelo sequencial, e adicionado as seguintes camadas, uma Rescaling para padronizar os valores de entrada, três Conv2D e MaxPooling2D para fazer a convolução e reduzir a resolução, uma *Flatten* para transformar o tensor 3D em 1D, e duas Dense para calcular os resultados e classificar a imagem em uma das classes, como apresentado na imagem 9.

```
model = Sequential([
    layers.Rescaling(1./255),
    layers.Conv2D(16, 3, padding='same', activation='relu'),
    layers.MaxPooling2D(),
    layers.Conv2D(32, 3, padding='same', activation='relu'),
    layers.MaxPooling2D(),
    layers.Conv2D(64, 3, padding='same', activation='relu'),
    layers.MaxPooling2D(),
    layers.Flatten(),
    layers.Dense(128, activation='relu'),
    layers.Dense(num_classes)
])
```

Figura 9

Fonte: Autores.

Criação do modelo.

Na compilação do modelo foi utilizado o otimizador Adam, a função de perda utilizando o *Crossentropy*, e a métrica de avaliação escolhida foi a acurácia, como apresentado na imagem 10.

```
model.compile(  
    optimizer='adam',  
    loss=tf.keras.losses.SparseCategoricalCrossentropy(from_logits=True),  
    metrics=['accuracy']  
)
```

Figura 10  
Fonte: Autores.  
Compilação do modelo.

Após compilado o modelo está pronto para o treino, onde é passado o conjunto de dados de treino e validação, e a quantidade de épocas, como apresentado na imagem 11, uma época é a iteração do modelo sobre o conjunto de dados.

```
epochs = 50  
  
history = model.fit(  
    treino,  
    validation_data=validacao,  
    epochs=epochs  
)
```

Figura 11  
Fonte: Autores.  
Treinamento do modelo.

Após as 50 épocas passadas temos as medidas de acurácia e da função de perda, como apresentado na imagem 12, e é visto que o modelo está sofrendo de *overfitting*, pois

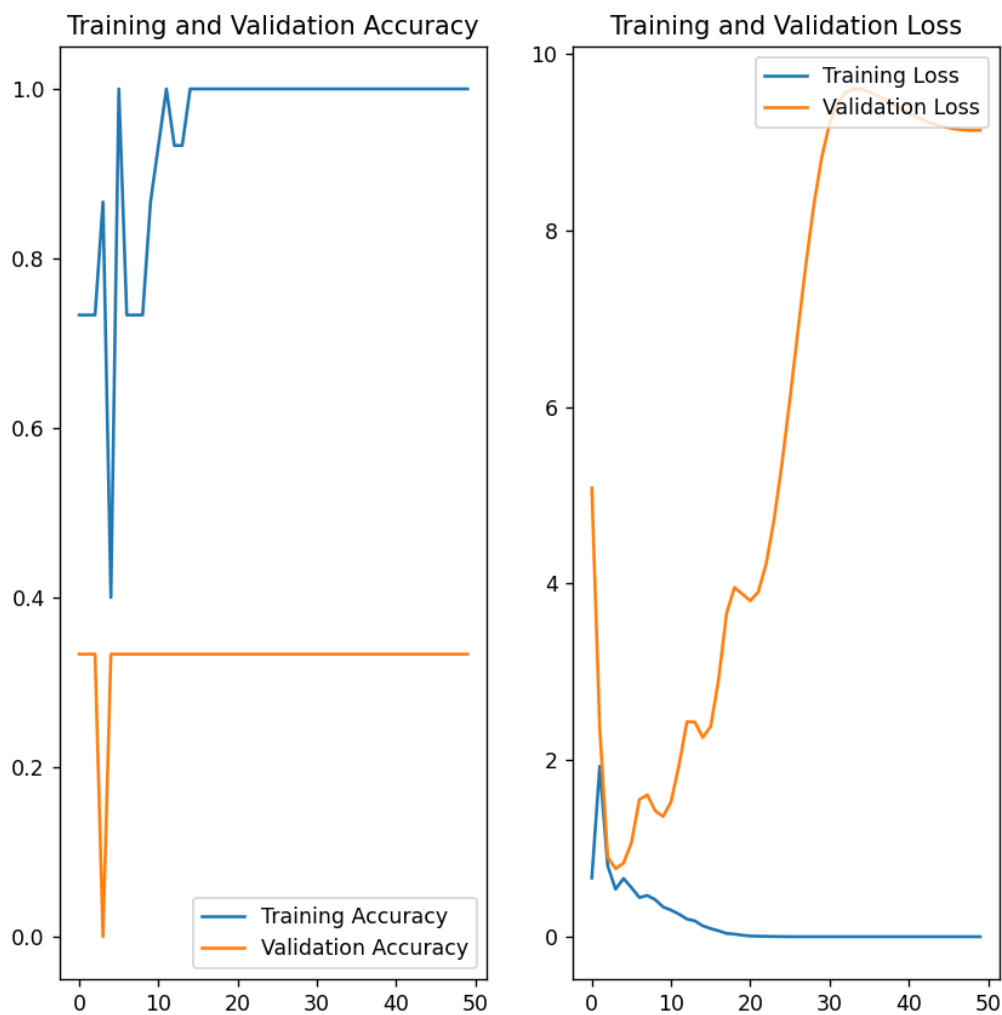


Figura 12  
 Fonte: Autores.

Gráfico de acurácia e perda do treino e validação antes das técnicas contra o overfitting.

a acurácia do treinamento aprendeu perfeitamente as imagens de teste, e quanto mais épocas, maior o valor da função de perda da validação.

#### 4.2 OVERFITTING

É a produção de uma análise de predição que corresponde muito perto ou exatamente a um conjunto de dados específico, mas que pode falhar ao fazer a predição de novos dados, deixando os resultados de validação e teste não confiáveis. O modelo



aprende de forma errada alguns sinais ou detalhes não importantes no conjunto de treinamento, em um ponto que isso impacta o desempenho do modelo em novos exemplos.

Algumas formas para diminuir o *overfitting* no modelo de treinamento e aumentar a acurácia na validação do modelo:

- *Data Augmentation*: aumenta a quantidade de imagens disponíveis aplicando transformações randômicas as imagens, criando imagens novas a partir de uma imagem, como apresentado na imagem 13.

```
data_augmentation = Sequential([
    layers.RandomFlip(
        "horizontal",
        input_shape=(img_height, img_width, 3)
    ),
    layers.RandomFlip(
        "vertical",
        input_shape=(img_height, img_width, 3)
    ),
    layers.RandomRotation(0.1),
    layers.RandomRotation(0.3),
    layers.RandomZoom(0.1),
    layers.RandomZoom(0.3),
    layers.RandomCrop(150, 150),
    layers.RandomTranslation(
        0.2, 0.2,
        fill_mode="reflect", interpolation="bilinear"
    )
])
```

Figura 13  
Fonte: Autores.  
Data augmentation

- *Dropout*: quando aplicado a camada *Dropout* ao modelo, ele de forma aleatória joga fora uma quantidade de imagens durante o processo de treinamento em cada época.

Agora adicionando as novas camadas ao modelo, para reduzir o *overfitting*, como apresentado na imagem 14.

```

model = Sequential([
    data_augmentation,
    layers.Rescaling(1./255),
    layers.Conv2D(16, 3, padding='same', activation='relu'),
    layers.MaxPooling2D(),
    layers.Conv2D(32, 3, padding='same', activation='relu'),
    layers.MaxPooling2D(),
    layers.Conv2D(64, 3, padding='same', activation='relu'),
    layers.MaxPooling2D(),
    layers.Dropout(0.2),
    layers.Flatten(),
    layers.Dense(128, activation='relu'),
    layers.Dense(num_classes, name="outputs")
])

```

Figura 14

Fonte: Autores.

Novo modelo com data augmentation e dropout aplicados.

Após a compilação do novo modelo, imagem 15, é visto que a acurácia do treinamento deixou de ser perfeita, e que as perdas da validação diminuíram significativamente. Mas a acurácia da validação continua baixa devido a pequena quantidade de imagens fornecidas para o modelo.

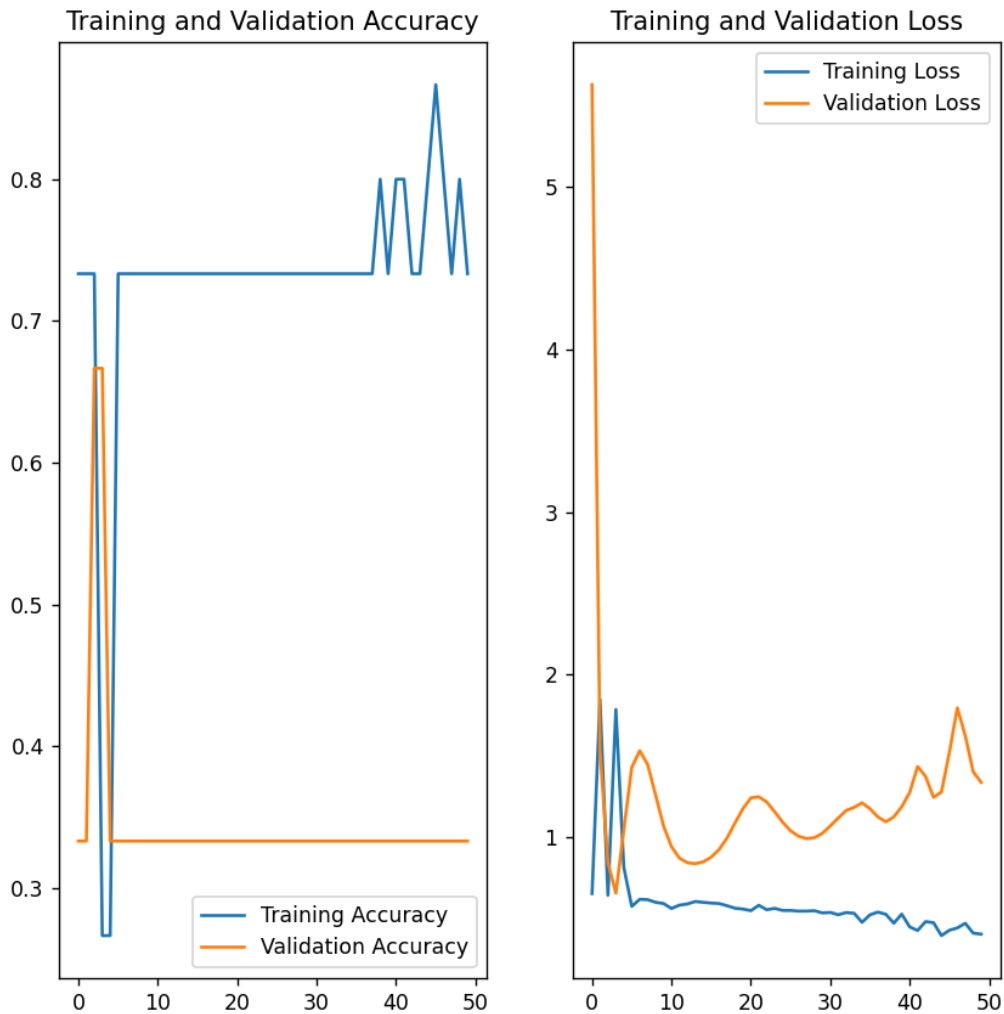


Figura 15

Fonte: Autores.

Gráfico de acurácia e perda do treino e validação com as camadas data augmentation e dropout aplicadas.

## 5. ANÁLISE E DISCUSSÃO DOS RESULTADOS

A partir das pesquisas já realizadas, obtivemos um resultado para contribuir com o gerenciamento do controle de pragas em plantações de soja visando utilizar soluções de *Machine Learning* com o auxílio de ferramentas, já apresentadas, para a sua implementação, no qual ele passou por um processo de treinamento com uma base de dados com imagens de danos para reconhecer os padrões nas folhas de soja, posteriormente identificando quais as pragas podem estar contaminando a plantação. E na

fase de teste foi colocada uma imagem nova para o algoritmo prever a qual classe pertence a folha de soja danificada, como apresentado na imagem 16.

```
image = load_img(image, target_size=(img_height, img_width))
image_array = img_to_array(image)
image_array = tf.expand_dims(image_array, 0)

predictions = model.predict(image_array)
score = tf.nn.softmax(predictions[0])

print(
    "Essa imagem provavelmente pertence a {} com {:.2f} por cento de confiança."
    .format(class_names[np.argmax(score)], 100 * np.max(score))
)
```

Figura 16

Fonte: Autores.

Código para prever uma imagem.

Não conseguimos ter um banco de dados com imagens suficientes para atingir uma grande assertividade. Com isso, a cada execução do código o valor vem diferente. Sendo assim, não foi atingido a acurácia esperada de 0.8 ou 80%, devido ao baixo número de imagens para treino e validação, portanto o classificador de imagens não é confiável para prever a qual classe pertence a imagem, como apresentado na imagem 17.

```
Essa imagem provavelmente pertence a lagarta-da-soja com 53.71 por cento de confiança.
```

Figura 17

Fonte: Autores.

Exemplo de resposta depois do código executado.

## 6. CONSIDERAÇÕES FINAIS

Este trabalho procurou entender a aplicação da inteligência artificial na detecção de pragas na agricultura para ajudar os produtores agrícolas na hora de aplicar o inseticida certo. Além disso, o foco principal foi o treinamento do algoritmo para o reconhecimento dos danos causados por cada praga nas folhas. Para que isso fosse realizado foi implementado um software que instancia uma rede neural, possibilitando adicionar conhecimento a ela através de aprendizagem supervisionada.

Para se atingir uma compreensão de como utilizar soluções de *Machine Learning* para coletar e analisar imagens que mostram as condições das folhas da soja, e identificar se há ou não alguma praga a danificando, definiu-se alguns objetivos específicos. O primeiro objetivo foi pesquisar a ferramenta computacional que auxilie no processo de implementação de *Machine Learning*. Verificou-se que as melhores ferramentas que nos auxiliaram nesse processo foram *Python*, *TensorFlow* e *Keras*. Depois, identificar corretamente os danos causados pelas pragas lagarta da soja e falsa-medideira. A análise permitiu concluir que há diferenças significativas que diferenciam cada dano, sendo a lagarta da soja comendo a folha por completo e a falsa medideira deixando o aspecto rendilhado. Em seguida, implementar um software que servirá de interface para os treinos, classificações, testes e análises. Certificamos que foi feita a implementação e foi realizada com algumas ressalvas.

A implementação da inteligência artificial para detectar as pragas foi executada, mas com o objetivo de melhora para trabalhos futuros será necessário: aumentar a base de imagens para treinamento; realizar uma inserção de imagens maior no banco de dados; com a inserção de imagens aumentar a acurácia para 80%; criar um aplicativo para execução nas lavouras.

Através do projeto pudemos apresentar resultados que poderão ter impacto através das ferramentas propostas, ratificando assim, a importância do gerenciamento de controle das pragas na agricultura.

## 7.REFERÊNCIAS BIBLIOGRÁFICAS

**ALVES, Rodolfo.** Agricultura no Brasil Atual. **Mundo Educação**, Disponível em: <https://mundoeducacao.uol.com.br/geografia/agricultura-no-brasil-atual.htm/>. Acesso em: 5 jun. 2022.

**Embrapa.** A agricultura Brasileira. **Portal Embrapa.** Disponível em: <https://www.embrapa.br/vii-plano-diretor/a-agricultura-brasileira>. Acesso em: 05 jun. 2022.

**Agronomia.são mateus.** A importância da agricultura no Brasil. **Agronomia.são mateus**, 2022. Disponível: <https://agronomia.saomateus.ufes.br/conteudo/import%C3%A2ncia-da-agricultura-no-brasil>. Acesso em: 6 mai. 2022.

**ANAC.** Drones. **Portal gov.br.** Disponível em: <https://www.gov.br/anac/pt-br/assuntos/drones>. Acesso em: 10 mai. 2022.

**BARSANTE, Christiano.** Tipos de Drones no Mercado e suas Aplicações. **Linkedin**, Disponível em: <https://pt.linkedin.com/pulse/tipos-de-drones-mercado-e-suas-aplica%C3%A7%C3%B5es-christiano-barsante>. Acesso em: 10 mai. 2022.

**AgroPós.** Drones agrícolas: Conheça essa Tecnologia. **AgroPós.** Disponível em: <https://agropos.com.br/drones-agricolas/>. Acesso em: 10 mai. 2022.

**PixForce.** Drones na Agricultura: tudo sobre a tecnologia que está mudando o setor. **PixForce.** Disponível em: <https://www.pixforce.com.br/post/drones-na-agricultura-tudo-sobre-a-tecnologia-que-est%C3%A1-mudando-o-setor>. Acesso em: 10 mai. 2022.

**IBGE.** Levantamento Sistemático da Produção Agrícola. **gov.br**, 2022. Disponível em: <https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9201-levantamento-sistematico-da-producao-agricola.html?=&t=destaques>. Acesso em: 06 mai. 2022.

**CAROLINE, Luana.** A Agricultura no Mundo Atual e as Políticas Agrícolas nos Países Desenvolvidos. **Brasil Escola.** Disponível em: <https://meuartigo.brasilecola.uol.com.br/geografia/a-agricultura-no-mundo-atual-as-politicas-agricolas-nos-paises-desenvolvidos.htm>. Acesso em : 7 mai. 2022.

**FIELDVIEW.** Maiores Produtores Agrícolas do Mundo e o que Eles Ensinam sobre a Agricultura Digital. 2021. Disponível em: <https://blog.climatefieldview.com.br/maiores-produtores-agricolas-mundo#:~:text=Os%205%20maiores%20produtores%20agr%C3%ADcolas%20do%20mundo%20s%C3%A3o%20China%2C%20Estados,do%20campo%20nas%20%C3%BAltimas%20d%C3%A9cadas>. Acesso em: 6 mai. 2022.

**FIELDVIEW.** Quais são os principais tipos de agriculturas praticadas no Brasil. 2021. Disponível em: <https://blog.climatefieldview.com.br/tipos-de-agricultura>. Acesso em: 6 mai. 2022.

**DSA.** 17 Casos de Uso de Machine Learning. **blog.dsacademy.com.br.** 2022. Disponível em: <https://blog.dsacademy.com.br/17-casos-de-uso-de-machine-learning/>. Acesso em: 8 mai. 2022.

**BIGTIRES.** Entenda a Aplicação de Machine Learning na Agricultura. **bigtires.com.br.** 2021. Disponível em: <https://www.bigtires.com.br/blog/post/entenda-a-aplicacao-do-machine-learning-na-agricultura>. Acesso em: 8 mai. 2022.

**IBM.** Aprendizado de Máquina Interativo. **ibm.com.** Disponível em: <https://www.ibm.com/br-pt/analytics/machine-learning#:~:text=A%20tecnologia%20Machine%20Learning%20permite,ele%20%C3%A9%20colocado%20para%20treinar>. Acesso em: 10 mai. 2022.

**HANASHIRO, Akira.** O que é e como funciona Machine Learning. **treinaweb.com.** 2019. Disponível em: <https://www.treinaweb.com.br/blog/o-que-e-e-como-funciona-machine-learning>. Acesso em: 10 mai. 2022.

**IMB.** Sobre Goiás. **imb.go.gov.br.** Disponível em: [https://www.imb.go.gov.br/index.php?option=com\\_content&view=article&id=79&Itemid=145](https://www.imb.go.gov.br/index.php?option=com_content&view=article&id=79&Itemid=145). Acesso em: 11 mai. 2022.

**FAEG.** Plantio de Soja chega a 35% no País. Goiás está entre os Estados que mais plantaram. 2020. Disponível em: <https://sistemafaeg.com.br/faeg/noticias/itr/plantio-da-soja-chega-a-35-no-pais-goias-esta-entre-os-estados-que-mais-plantaram>. Acesso em: 5 jun. 2022.

**AIRES, Rafaella.** Controle de Pragas Agrícolas: Tudo o que você precisa saber. **myfarm.com.** 2020. Disponível em: <https://www.mylfarm.com.br/controle-de-pragas/>. Acesso em: 15 mai. 2022.

**MORAES, Michelly.** Conheça as 9 principais Pragas e como combatê-las. **agropos.com.** Disponível em: <https://agropos.com.br/pragas-da-soja/>. Acesso em: 15 mai. 2022.

**G1.** Brasil é o 4º maior produtor de grãos, atrás de, China, EUA e Índia, diz estudos. **g1.globo.com** .2021. Disponível em: <https://g1.globo.com/economia/agronegocios/noticia/2021/06/01/brasil-e-o-4o-maior-produtor-de-graos-atras-da-china-eua-e-india-diz-estudo.ghtml>. Acesso em: 2 jun. 2022.

**ESTEVESTONI.** Disponível em: <https://estevestoni.medium.com/agrupando-conceitos-e-classificando-imagens-com-deep-learning-5b2674f99539>

**CHEATSHEET.** Disponível em: <https://ml-cheatsheet.readthedocs.io/en/latest/index.html>

**IMASTERS.** Disponível em: <https://imasters.com.br/back-end/classificacao-de-imagens-com-deep-learning-e-tensorflow>

**ANDRADE, F. G.; NEGREIRO, M. C. C.; FALLEIROS, A. M. F.** Aspectos dos mecanismos de defesa da lagarta da soja *Anticarsia gemmatalis* (Hübner 1818) relacionados ao controle biológico por *Baculovirus anticarsia* (AGMNPV). *Arquivos do Instituto Biológico*, v. 71, n. 3, p. 391-398, 2004.

**BERNARDI, Oderlei.** Avaliação do risco de resistência de lepidópteros-praga (Lepidoptera: Noctuidae) à proteína Cry1Ac expressa em soja MON 87701 x MON 89788 no Brasil. 2012. Tese de Doutorado. Universidade de São Paulo.

**GALLO, D.; NAKANO, O.; SILVEIRA NETO, S.; CARVALHO, R.P.L.; BATISTA, G.C.; BERTI FILHO, E.; PARRA, J.R.P.; ZUCCHI, R.A.; ALVES, S.B.; VENDRAMIN, J.D.; MARCHINI, L.C.; LOPES, J.R.S.; OMOTO, C.** Manual de entomologia agrícola. Piracicaba: FEALQ, 2002. 920p.

**LLOURENÇÃO<sup>1</sup>, ANDRÉ et al.** Produtividade de genótipos de soja sob infestação da lagarta-da-soja e de percevejos. *Neotropical Entomology*, v. 39, n. 2, p. 275-281, 2010.

**MOSCARDI, Flavio et al.** Artrópodes que atacam as folhas da soja. *Soja: manejo integrado de insetos e outros artrópodes-praga*, v. 4, p. 859, 2012.

**NEGREIRO, MCC de; ANDRADE, FG de; FALLEIROS, A. M. F.** Sistema imunológico de defesa em insetos: uma abordagem em lagartas da soja, *Anticarsia gemmatalis* Hübner (Lepidoptera: Noctuidae), resistentes ao AgMNPV. *Semina: Ciências Agrárias*, v. 25, n. 4, p. 293-308, 2004.

SOSA-GÓMEZ, D.R.; CORRÊA-FERREIRA, B.S.; HOFFMANN-CAMPO, C.B.; CORSO, I.C.; OLIVEIRA, L.J.; MOSCARDI, F.; PANIZZI, A.R.; BUENO, A. de F.; HIROSE, E. Manual de identificação de insetos e outros invertebrados da cultura da soja. Londrina: Embrapa-CNPSo, 2010. 90 p. (Embrapa – CNPSo. Documentos, 269).

**TensorFlow Core.** Image classification. **TensorFlow.** Disponível em: <https://www.tensorflow.org/tutorials/images/classification>. Acesso em: 04 nov. 2022.

**TensorFlow Core.** Convolutional Neural Network. **TensorFlow.** Disponível em: <https://www.tensorflow.org/tutorials/images/cnn>. Acesso em: 06 nov. 2022.

**TensorFlow Core.** Overfit and Underfit. **TensorFlow.** Disponível em: [https://www.tensorflow.org/tutorials/keras/overfit\\_and\\_underfit](https://www.tensorflow.org/tutorials/keras/overfit_and_underfit). Acesso em: 12 nov. 2022.

**ML Glossary.** Concepts Neural Networks. **ML Glossary.** Disponível em: [https://ml-cheatsheet.readthedocs.io/en/latest/nn\\_concepts.html](https://ml-cheatsheet.readthedocs.io/en/latest/nn_concepts.html). Acesso em: 04 nov. 2022.



**ML Glossary.** Loss Functions. **ML Glossary.** Disponível em: [https://ml-cheatsheet.readthedocs.io/en/latest/loss\\_functions.html](https://ml-cheatsheet.readthedocs.io/en/latest/loss_functions.html). Acesso em: 06 nov. 2022.

**ML Glossary.** Optimizers. **ML Glossary.** Disponível em: <https://ml-cheatsheet.readthedocs.io/en/latest/optimizers.html>. Acesso em: 06 nov. 2022.

MOREIRA, João Carlos; SENE, Eustáquio de. **Geografia:** ensino médio. São Paulo: Scipione, 2011.

VESENTINI, José William. **Geografia:** o mundo em transição: ensino médio. São Paulo: Ática, 2011.

NORVIG, Peter; RUSSELL, Stuart. **Inteligência Artificial.** 3.ed. Rio de Janeiro: Elsevier, 2013.

**Deep Learning Book.** Capítulo 40 - Introdução às Redes Neurais Convolucionais. **Deep Learning Book.** Disponível em: <https://www.deeplearningbook.com.br/introducao-as-redes-neurais-convolucionais/>. Acesso em: 19 dez. 2022.

**Deep Learning Book.** Capítulo 8 - Função de Ativação. **Deep Learning Book.** Disponível em: <https://www.deeplearningbook.com.br/funcao-de-ativacao/>. Acesso em: 21 dez. 2022.

**Keras.** Dense layer. **Keras.** Disponível em: [https://keras.io/api/layers/core\\_layers/dense/](https://keras.io/api/layers/core_layers/dense/). Acesso em: 21 dez. 2022.

**Keras.** Layer activation functions. **Keras.** Disponível em: <https://keras.io/api/layers/activations/>. Acesso em: 21 dez. 2022.