



BACHAREL EM CIÊNCIA DA COMPUTAÇÃO

**AVALIAÇÃO DE MODELOS YOLO PARA INDENTIFICAÇÃO DE
ESPÉCIES DE ABELHAS GUAXUPÉ E IRAPUÃ**

LUIS EDUARDO BORGES DE SOUSA

Rio Verde, GO

2026



INSTITUTO FEDERAL GOIANO - CAMPUS RIO VERDE
BACHAREL EM CIÊNCIA DA COMPUTAÇÃO

AVALIAÇÃO DE MODELOS YOLO PARA IDENTIFICAÇÃO DE ESPÉCIES DE ABELHAS GUAXUPÉ E IRAPUÃ

LUIS EDUARDO BORGES DE SOUSA

Trabalho de Conclusão de Curso apresentado ao Instituto Federal Goiano - Campus Rio Verde, como requisito parcial para a obtenção do Grau de Bacharel em Ciência da Computação.

Orientadora: Profa. Dra. Heyde Francielle do Carmo França

Rio Verde, GO

03, 2026

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR PRODUÇÕES TÉCNICO-CIENTÍFICAS NO REPOSITÓRIO INSTITUCIONAL DO IF GOIANO

Com base no disposto na Lei Federal nº 9.610, de 19 de fevereiro de 1998, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia Goiano a disponibilizar gratuitamente o documento em formato digital no Repositório Institucional do IF Goiano (RIIF Goiano), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IF Goiano.

IDENTIFICAÇÃO DA PRODUÇÃO TÉCNICO-CIENTÍFICA

- ☐ Tese (doutorado)
☐ Dissertação (mestrado)
☐ Monografia (especialização)
☒ TCC (graduação)

- ☐ Artigo científico
☐ Capítulo de livro
☐ Livro
☐ Trabalho apresentado em evento

☐ Produto técnico e educacional - Tipo:

Nome completo do autor:

Luís Eduardo Borges de Sousa

Matrícula:

2020102201940148

Título do trabalho:

Avaliação de modelos YOLO para indentificação de espécies de abelhas Guaxupé e Irapuã

RESTRIÇÕES DE ACESSO AO DOCUMENTO

Documento confidencial: ☒ Não ☐ Sim, justifique:

Informe a data que poderá ser disponibilizado no RIIF Goiano: **19/03/2026**

O documento está sujeito a registro de patente? ☐ Sim ☒ Não

O documento pode vir a ser publicado como livro? ☐ Sim ☒ Não

DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA

O(a) referido(a) autor(a) declara:

- Que o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;
- Que obteve autorização de quaisquer materiais incluídos no documento do qual não detém os direitos de autoria, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia Goiano os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
- Que cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia Goiano.

Rio Verde, GO

Local

19/03/2026

Data

Luís Eduardo Borges de Sousa

Assinatura do autor e/ou detentor dos direitos autorais

Ciente e de acordo:



Documento assinado digitalmente

HEYDE FRANCESLE DO CARMO FRANCA

Data: 19/03/2026 15:36:46-0300

Verifique em <https://validar.it.gov.br>

**Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema Integrado de Bibliotecas do IF Goiano - SIBi**

B732a Borges de Sousa, Luis Eduardo
Avaliação de modelos YOLO para indentificação de espécies de
abelhas Guaxupé e Irapuã / Luis Eduardo Borges de Sousa. Rio
Verde 2026.

22f. il.

Orientadora: Profª. Dra. Heyde Francielle do Carmo França.
Tcc (Bacharel) - Instituto Federal Goiano, curso de 0219201 -
Bacharelado em Ciência da Computação - Integral - Rio Verde
(Campus Rio Verde).

1. CNN. 2. YOLO. 3. Abelhas Nativas do Cerrado. I. Título.

RESUMO

Borges de Sousa, Luis Eduardo. **Avaliação de modelos YOLO para indentificação de espécies de abelhas Guaxupé e Irapuã.** 03, 2026. 22 f. Trabalho de conclusão de curso – (Curso de Bacharel em Ciência da Computação), Instituto Federal Goiano - Campus Rio Verde. Rio Verde, GO.

As abelhas desempenham um papel fundamental como agentes polinizadores, sendo indispensáveis para a preservação da biodiversidade e o equilíbrio dos ecossistemas, em especial no bioma Cerrado. Entretanto, a identificação das espécies nativas é dificultada pelo chamado “impedimento taxonômico”, intensificado pela falta de especialistas e pela elevada semelhança morfológica entre as espécies. Este estudo utilizou redes neurais convolucionais YOLO para classificar de abelhas nativas do Cerrado *Trigona hyalinata* (Guaxupé) e *Trigona spinipes* (Irapuã) por serem especies que representam maior grau de dificuldade para sua identificação devido a sua semelhança “impedimento taxonômico”, visando minimizar essa limitação e contribuir para o monitoramento e a conservação desses polinizadores, a partir de uma base de imagens rotuladas manualmente. Nesse trabalho, foram testadas as versões YOLO v8, v9, v10, v11, v12 e 26, comparando desempenho em precisão, recall, mAP50 e F1.

Palavras-chave: CNN,YOLO,Abelhas Nativas do Cerrado.

ABSTRACT

Borges de Sousa, Luis Eduardo. **Avaliação de modelos YOLO para indentificação de espécies de abelhas Guaxupé e Irapuã.** 03, 2026. 22 f. Trabalho de conclusão de curso – (Curso de Bacharel em Ciência da Computação), Instituto Federal Goiano - Campus Rio Verde. Rio Verde, GO.

Bees develop a fundamental role as pollinating agents, being indispensable for the preservation and conservation of the biodiversity and the balance of the ecosystems, especially on the biome cerrado. However, the identification of native species is hindered by the so-called “taxonomic impediment,” intensified by the lack of specialists and the high morphological similarity among species. This study used YOLO convolutional neural networks to classify native Cerrado bees, *Trigona hyalinata* (Guaxupé) e *Trigona spinipes* (Irapuã) because these species represent a higher level of difficulty in identification due to their similarity and the resulting taxonomic impediment. The objective was to minimize this limitation and contribute to the monitoring and conservation of these pollinators using a dataset of manually labeled images. In this work, the versions YOLO v8, v9, v10, v11, v12, and 26 were tested, comparing their performance in terms of precision, recall, mAP50, and F1 score.

Keywords: Convolutional Neural Networks, YOLO, Bees native from cerrado.



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA GOIANO

Ata nº 7/2026 - GEPTNM-RV/DE-RV/CMPRV/IFGOIANO

ATA DE DEFESA DE TRABALHO DE CURSO

Ao(s) 12 dia(s) do mês de março de 2026, às 14 horas e 00 minutos, reuniu-se a banca examinadora composta pelos docentes: Heyde Francielle do Carmo (orientadora), Fábio Montanha Ramos(membro), Douglas Cedrim Oliveira(membro), para examinar o Trabalho de Curso intitulado “AVALIAÇÃO DE MODELOS YOLO PARA IDENTIFICAÇÃO DE ESPÉCIES DE ABELHAS GUAXUPÉ E IRAPUÃ ” do(a) estudante Luis Eduardo Borges de Sousa, Matrícula nº 2020102201940148, do Curso de Bacharelado em Ciência da Computação do IF Goiano – Campus Rio Verde. A palavra foi concedida ao(a) estudante para a apresentação oral do TC, houve arguição do(a) candidato pelos membros da banca examinadora. Após tal etapa, a banca examinadora decidiu pela APROVAÇÃO do estudante. Ao final da sessão pública de defesa foi lavrada a presente ata que segue assinada pelos membros da Banca Examinadora.

(Assinado Eletronicamente)

Heyde Francielle do Carmo

Orientador(a)

(Assinado Eletronicamente)

Fábio Montanha Ramos

Membro

(Assinado Eletronicamente)

Douglas Cedrim Oliveira

Membro

Observação:

() O(a) estudante não compareceu à defesa do TC.

Documento assinado eletronicamente por:

- **Heyde Francielle do Carmo Franca**, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 16/03/2026 15:18:14.
- **Douglas Cedrim Oliveira**, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 17/03/2026 08:23:12.
- **Fabio Montanha Ramos**, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 19/03/2026 14:46:52.

Este documento foi emitido pelo SUAP em 12/03/2026. Para comprovar sua autenticidade, faça a leitura do QRCode ao lado ou acesse <https://suap.ifgoiano.edu.br/autenticar-documento/> e forneça os dados abaixo:

Código Verificador: 799536

Código de Autenticação: 8c16b1a5b7



INSTITUTO FEDERAL GOIANO

Campus Rio Verde

Rodovia Sul Goiana, Km 01, Zona Rural, 01, Zona Rural, RIO VERDE / GO, CEP 75901-970

(64) 3624-1000

LISTA DE FIGURAS

Figura 1 – A esquerda uma guaxupé e a direita uma irapuã.	2
Figura 2 – A esquerda uma guaxupé coberta por polem colorindo seus membros inferiores de uma cor parecida com laranja e a direita uma irapuã onde em um ambiente que se confunde o fundo com a coloração de seus membros inferiores e asas.	3
Figura 3 – YOLO exemplo de como uma predição é feita pela arquitetura.	3
Figura 4 – Comparação de latência dos modelos YOLO com a base de dados COCO com diferentes valores de map50-95.	4
Figura 5 – Predições feitas S-YOLOv5m detectando e calssificando corretamente a esquerda (Sloth beetle), no meio (Red spider), a direita (Whitefly). . .	7
Figura 6 – Predições feitas pelo modelo treinado no trabalho de kumar corretamento diferenciando os insetos nas imagens.	8
Figura 7 – Comparação de métricas entre diferentes versões da YOLO treinadoas por Zhang.	9
Figura 8 – Predições feitas pelos modelos yolo treinados no trabalho de Zhang. . .	9
Figura 9 – Distribuição entre classes no trabalho de Zhang.	9
Figura 10 – Caixa delimitadora ao redor da abelha.	10
Figura 11 – Caixa delimitadora ao redor de multiplas abelha.	11
Figura 12 – Distribuição de classes.	12
Figura 13 – Progresso da métrica mAP50 por época.	13
Figura 14 – Predição feita pela YOLOv8s.	14
Figura 15 – Progressão das métricas para YOLOv8s.	15
Figura 16 – Progressão das métricas para YOLOv9s.	15
Figura 17 – Progressão das métricas para YOLOv10s.	16
Figura 18 – Progressão das métricas para YOLOv11s.	16
Figura 19 – Progressão das métricas para YOLOv12s.	17
Figura 20 – Progressão das métricas para YOLO26s.	17
Figura 21 – Predição errada feita pela YOLOv8s, a esquerda temos a caixas anotadas no dataset e a direita a predição realizada pela YOLOv8s.	18

LISTA DE TABELAS

Tabela 1	– Métricas de map50 obtidas pelos modelos avaliados no estudo de Tang.	7
Tabela 2	– Resultados do modelo treinado no trabalho de kumar contendo valores de precisão, recall, map50, e f1-score.	8
Tabela 3	– Parâmetros de treinamento do modelos YOLO.	11
Tabela 4	– Comparação de métricas entre diferentes versões da YOLO.	13

LISTA DE ABREVIATURAS E SIGLAS

CNN	Redes Neurais Convolucionais.
YOLO	You Only Look Once.
PGI	Informações de Gradiente Programáveis.
NMS	Non-Maximum Supprssion.
R-ELAN	Redes de Agregação de Camadas Eficientes.
DFL	Distribution Focal Loss.
COCO	Common Objects in Contex.
mAP	mean Average Precision.
IoU	Intersection over Union.

SUMÁRIO

1 – INTRODUÇÃO	1
2 – FUNDAMENTAÇÃO TEÓRICA	2
3 – REVISÃO DE LITERATURA	7
4 – MATERIAIS E MÉTODOS	10
5 – RESULTADOS E DISCUSSÕES	13
6 – CONCLUSÃO	19
Referências	20

1 INTRODUÇÃO

As abelhas exercem um papel indispensável na conservação da biodiversidade, atuando diretamente na reprodução de inúmeras espécies vegetais e na estabilidade dos ecossistemas. Estima-se que mais da metade das plantas presentes em florestas tropicais dependam da polinização realizada por esses insetos, percentual que ultrapassa 80% em biomas como o Cerrado brasileiro (FREITAS; SILVA, 2015). Em escala global, as abelhas são responsáveis pela polinização de cerca de 73% das culturas destinadas à alimentação humana e de 42% das 57 espécies agrícolas mais cultivadas (FREITAS; SILVA, 2015). No contexto brasileiro, aproximadamente 60% das 141 espécies de plantas cultivadas apresentam algum nível de dependência da polinização realizada por abelhas (FREITAS; IMPERATRIZ-FONSECA; MEDINA, 2009). O Cerrado, considerado um bioma estratégico para a conservação da biodiversidade e dos recursos hídricos (REIS, 2024), tem sua complexa diversidade ecológica sustentada, em grande parte, pela atuação de abelhas nativas como agentes polinizadores (TORRES; SOBREIRO; JUNIOR, 2023). Entretanto, ações humanas, como a expansão da fronteira agrícola, o desmatamento e a ocorrência de queimadas, têm provocado expressivas perdas de biodiversidade (MACEDO et al., 2015). Esse cenário, associado aos impactos das mudanças climáticas, intensifica a necessidade de identificar, classificar e monitorar espécies de forma eficiente. Este desafio é agravado pelo impedimento taxonômico, que envolve a limitação de especialistas e o difícil acesso a informações taxonômicas (CARVALHO; BOCKMANN; AMORIM, 2007). No caso específico das abelhas, fatores como elevada diversidade, similaridade morfológica entre espécies, dimorfismo sexual e a presença de espécies crípticas contribuem para o agravamento desse problema. Como solução para esse problema, é proposto neste trabalho o uso de redes neurais convolucionais (CNNs), que se destacam por seu desempenho na análise de imagens. Este trabalho tem como finalidade desenvolver e avaliar uma rede neural convolucional (CNN) para a classificação de espécies de abelhas nativas do bioma Cerrado: *Trigona hyalinata* (Guaxupé) e *Trigona spinipes* (Irapuã), empregando métricas de desempenho como precisão, revocação (recall) e F1-score. O modelo proposto busca apoiar pesquisadores ao minimizar o impedimento taxonômico, possibilitando a identificação rápida e eficiente de grandes volumes de imagens, inclusive com potencial de aplicação em tempo real.

2 FUNDAMENTAÇÃO TEÓRICA

As abelha irapuã e guaxupé são abelhas nativas do cerrado que apresentam um alto grau de semelhança possuindo tamanhos parecidos com uma media de 7mm e corpos pretos porem a principal diferença entre as duas esta na coloração as abelhas irapuã possuindo corbículas alaranjadas enquanto a guaxupé possui asas amarronzadas (Secretaria de Meio Ambiente, Infraestrutura e Logística do Estado de São Paulo, 2023). A figura 1 possui uma irapuã e uma guaxupé lado a lado:

Figura 1 – A esquerda uma guaxupé e a direita uma irapuã.



Fonte: próprio autor.

Essa semelhança denominada de dificuldade taxionômica dificulta a classificação dessas espécies entre si, especialmente em imagens onde as principais características que as diferenciam sendo essas a coloração das asas e dos membros inferiores não é clara ocorrendo principalmente em ambientes naturais onde essas características como demonstrado na figura 2:

O tipo de rede neural escolhida, a Rede Neural Convolucional. A arquitetura escolhida foi a YOLO (JOCHER; CHAURASIA; QIU, 2023). A escolha da arquitetura YOLO se deu por conta de ser o estado da arte e pela sua popularidade em problemas de detecção de objetos, como em detecção de tomates salada (YANG et al., 2023), detecção em imagens de sensoriamento remoto (SHEN; LANG; SONG, 2023), e até mesmo na detecção de comportamento de estudantes em sala de aula (CHEN; ZHOU; JIANG, 2023). De acordo com o criador da rede YOLO (REDMON et al., 2016), ela se destaca por sua simplicidade, pois utiliza apenas uma rede convolucional para prever simultaneamente várias caixas delimitadoras e as probabilidades de classe correspondentes. Diferentemente dos métodos tradicionais de detecção de objetos, a YOLO treina em imagens completas e otimiza diretamente o desempenho da detecção. Isso ocorre pois os métodos tradicionais de identificação e classificação de objetos precisam passar as imagens mais de uma vez pela rede: uma para identificar e separar os objetos da imagem e outra para classificá-los; já a YOLO faz ambos os processos em uma única passagem. Essa abordagem unificada apresenta várias vantagens, sendo a velocidade um dos principais benefícios. Ao formular a detecção como um problema de regressão, a YOLO evita a necessidade de pipelines

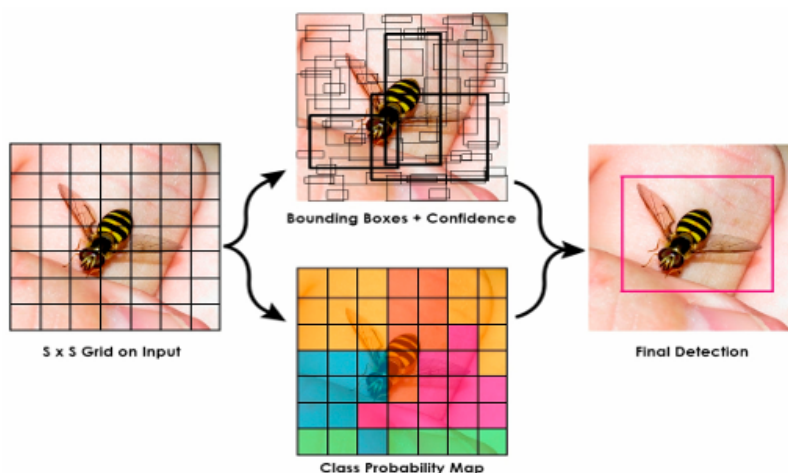
Figura 2 – A esquerda uma guaxupé coberta por pólen colorindo seus membros inferiores de uma cor parecida com laranja e a direita uma irapuã onde em um ambiente que se confunde o fundo com a coloração de seus membros inferiores e asas.



Fonte: próprio autor.

complexos. Durante o teste, a rede neural é aplicada diretamente a uma nova imagem para realizar as previsões de detecção exemplificado na figura 3:

Figura 3 – YOLO exemplo de como uma predição é feita pela arquitetura.

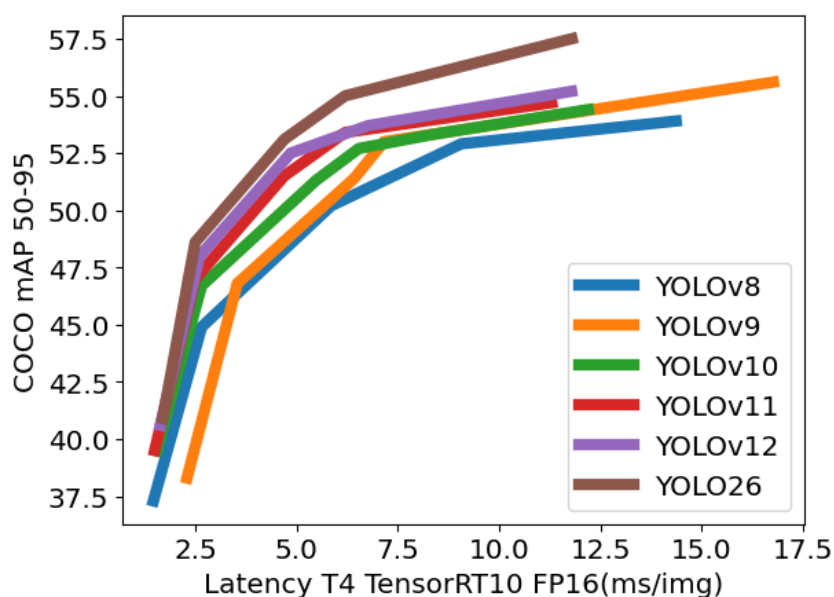


Fonte: Kumar, Nagarathna e Flammini (2023).

O YOLOv8 foi projetado para melhorar o desempenho da detecção de objetos em tempo real com recursos avançados. Ao contrário das versões anteriores, o YOLOv8 incorpora um cabeçalho Ultralytics dividido sem âncoras, backbone e arquiteturas de neck de última geração e oferece uma relação otimizada entre precisão e velocidade, tornando-o ideal para diversas aplicações (JOCHER; CHAURASIA; QIU, 2023). Já o YOLOv9 introduz métodos inovadores, como as Informações de Gradiente Programáveis (PGI) e a Rede de Agregação de Camadas Eficientes Generalizadas (GELAN), para resolver de maneira eficaz problemas relacionados à perda de informações e à eficiência computacional (WANG; LIAO, 2024). O YOLOv10 apresenta duas melhorias principais em relação às versões anteriores: um sistema de Atribuições Duplas Consistentes para Treinamento Sem NMS (Non-Maximum Suppression) é um algoritmo de pós-processamento usado para remover detecções duplicadas e um Design de Modelo Focado em Eficiência e Precisão,

visando otimizar o desempenho geral (WANG; CHEN; LIU, 2024). O YOLOv11 aprimora a extração de características com uma arquitetura avançada de backbone e neck, enquanto melhora a velocidade e meio de designs e pipelines de treinamento otimizados, equilibrando precisão e desempenho (JOCHER; QIU, 2024). O YOLOv12 incorpora várias inovações importantes para equilibrar velocidade e precisão. O mecanismo de Atenção de Área processa eficientemente grandes campos receptivos, reduzindo o custo computacional em comparação com a autoatenção padrão. As Redes de Agregação de Camadas Eficientes Residuais (R-ELAN) melhoram a agregação de recursos, abordando desafios de otimização em modelos maiores centrados na atenção. A Arquitetura de Atenção Otimizada, incluindo o uso de FlashAttention e a remoção da codificação posicional, aumenta ainda mais a eficiência (TIAN; YE; DOERMANN, 2025). E, por último, o YOLO26 tem como principais características a remoção de DFL (Distribution Focal Loss) é uma função de perda usada para melhorar a precisão da localização das bounding boxes, que simplifica a exportação e expande a compatibilidade de borda, a inferência End-to-End livre de NMS (Non-Maximum Suppression) para uma implementação mais rápida e simples (JOCHER; QIU, 2026). A figura 4 demonstra o progresso da métrica map50-95 dos modelos YOLO aqui analisados sendo treinados com a base de dados COCO (Common Objects in Context) (JOCHER; QIU, 2026):

Figura 4 – Comparação de latência dos modelos YOLO com a base de dados COCO com diferentes valores de map50-95.



Fonte: Jocher e Qiu (2026).

Nesse trabalho, são comparados as diferentes versões da YOLO. e para avaliar o desempenho da rede, foi utilizado a precisão, Recall, mAp50 e F1. A seguir, são apresentadas as definições e explicações dessas métricas. O treinamento de um modelo de aprendizagem requer validação de desempenho. Redes neurais utilizam dados rotulados como entrada para aprender padrões e prever resultados. Após o treinamento, aplica-se um conjunto inédito de dados para testar a precisão do modelo, cuja avaliação é feita por métricas específicas ao tipo de problema analisado. Assim, a precisão é uma métrica que valida a quantidade de verdadeiros positivos em razão da soma Verdadeiro positivo e falso positivo (PADILLA; NETTO; SILVA, 2020), calculada por:

$$Preciso = \frac{TP}{TP + FP} \quad (1)$$

onde:

$$\begin{cases} P &= \text{Precisão} \\ TP &= \text{Verdadeiro positivo} \\ FP &= \text{Falso positivo} \end{cases}$$

Já a sensibilidade ou recall avalia a capacidade do modelo de detectar com sucesso resultados classificados como positivos (PADILLA; NETTO; SILVA, 2020). E pode ser obtida pela equação:

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

onde:

$$\begin{cases} R &= \text{Recall} \\ TP &= \text{Verdadeiro positivo} \\ FN &= \text{Falso negativo} \end{cases}$$

A métrica mAP, ou mean Average Precision a um Intersection over Union (IoU) de 0.5, fornece uma medida global de desempenho do modelo de detecção de objetos (HENDERSON; FERRARI, 2016). O valor de mAP é calculado para diferentes limiares de confiança, e mAP50 refere-se ao valor médio da precisão considerando um limiar de 0.5 para o IoU, ou seja, se a sobreposição entre a caixa predita e a caixa real for superior a 50%, a predição é considerada correta. Quanto maior o valor de mAP50, melhor o desempenho global do modelo, com maior capacidade de detectar e classificar objetos com precisão, considerando um limiar de IoU de 0.5. A fórmula para o cálculo de mAP50 é:

$$M = \frac{1}{N} \sum_{i=1}^N AP_i \quad (3)$$

onde:

$$\begin{cases} M &= \text{mAP} \\ N &= \text{É o número total de classes} \\ AP_i &= \text{Precisão média para a classe } i \end{cases}$$

E a F-measure, F-score ou score F1 é uma média harmônica calculada com base na precisão e no recall (POWERS, 2020). Ela pode ser obtida com base na equação:

$$F = \frac{2 \times P \times R}{P + R} \quad (4)$$

onde:

$$\begin{cases} F &= \text{F1} \\ P &= \text{Precisão} \\ R &= \text{Recall} \end{cases}$$

Além dessas métricas mais importantes outras métricas que valem se citadas são as grandezas train/box loss, train/cls loss, train/dfn loss representam as perdas (losses) associadas à detecção das caixas delimitadoras, classificação da classe e perda focal, e a precisão do posicionamento das caixas delimitadoras. A tendência geral para essas métricas

é de decréscimo, isso indica que o modelo está aprendendo a realizar essas tarefas com mais precisão ao longo do treinamento. Flutuações podem ocorrer, especialmente no início do treinamento, à medida que o modelo se ajusta aos dados. Já a mAP50 (mean Average Precision) é uma métrica que combina precisão e recall, fornecendo uma avaliação mais completa do desempenho do modelo. A mAP50-95 considera um intervalo de IOUs de 0.5 a 0.95, oferecendo uma visão mais abrangente da performance do modelo em diferentes níveis de sobreposição entre as caixas previstas e as verdadeiras. O aumento da mAP50 indica que o modelo está melhorando na detecção de objetos.

O Roboflow serve como uma ferramenta para rotular bancos de dados, particularmente no contexto de detecção de objetos e aprendizado de máquina. Ele facilita a criação de diversos conjuntos de dados por meio de técnicas de rotulagem automatizadas e semiautomáticas, aumentando a eficiência e a precisão da preparação de dados para o treinamento de modelos (IWAMURA; TSUKADA; KISE, 2013).

3 REVISÃO DE LITERATURA

Durante os últimos anos, o sucesso das redes neurais convolucionais (CNNs) para classificação de imagens demonstra seu uso como um dos melhores métodos para o reconhecimento de imagem (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), e, em alguns casos, o uso de CNNs para classificação de imagens supera a capacidade humana (HE, 2015). Esses modelos vêm sendo aplicados, resultando em avanços significativos na assertividade de detecção. As CNNs se destacam por sua capacidade de reconhecer padrões visuais complexos, sendo amplamente empregadas na identificação de polinizadores em ambientes naturais (SPIESMAN, 2021). Há também o trabalho de França et al. (2025) que utiliza YOLOv8 para a identificação de espécies de abelha e obteve uma precisão de 79% na diferenciação entre 5 espécies nativas do cerrado sendo elas *Trigona spinipes* (Irapuã), *Tetragonisca angustula* (Jataí), *Paratrigona subnuda* (Jataí-da-terra), *Melipona quadrifasciata* (Mandaçaia), *Trigona pallens* (Olho de vidro) e a utilização da yolov8 para comparação das mesmas abelhas utilizadas neste trabalho obtendo, respectivamente, precisão e recall de 0,849 e 0,870. A média de precisão das duas classes (mAP@0.5) foi de 0,860. Dois estudos relevantes são o estudo que demonstra a identificação com o uso de uma versão modificada da YOLOv5 nomeada S-YOLOv5m de espécies de insetos que se proliferam em plantações de tomate (TANG et al., 2025), neste estudo foi utilizada uma base de dados de 3985 expandida para 18359 com técnicas de argumentação de dados, obtendo os resultados demonstrados pela tabela 1 e um exemplo de uma predição feita pela modelo S-YOLOv5m na figura 5 a baixa qualidade da imagem se da por ser a única imagem disponível no seu trabalho e em seu trabalho esta com esta qualidade:

Figura 5 – Predições feitas S-YOLOv5m detectando e classificando corretamente a esquerda (Sloth beetle), no meio (Red spider), a direita (Whitefly).



Fonte: Tang et al. (2025).

Tabela 1 – Métricas de map50 obtidas pelos modelos avaliados no estudo de Tang.

Modelo	mAP50	Modelo	mAP50
YOLOv3	0.721	YOLOv4	0.773
YOLOv4-tiny	0.686	YOLOv5	0.7961
YOLOv6-m	0.799	S-YOLOv5m	0.8434
YOLOv7	0.841	YOLOv7-tiny	0.681
YOLOx	0.824	SSD	0.755
Faster R-CNN	0.808		

Fonte: Tang et al. (2025).

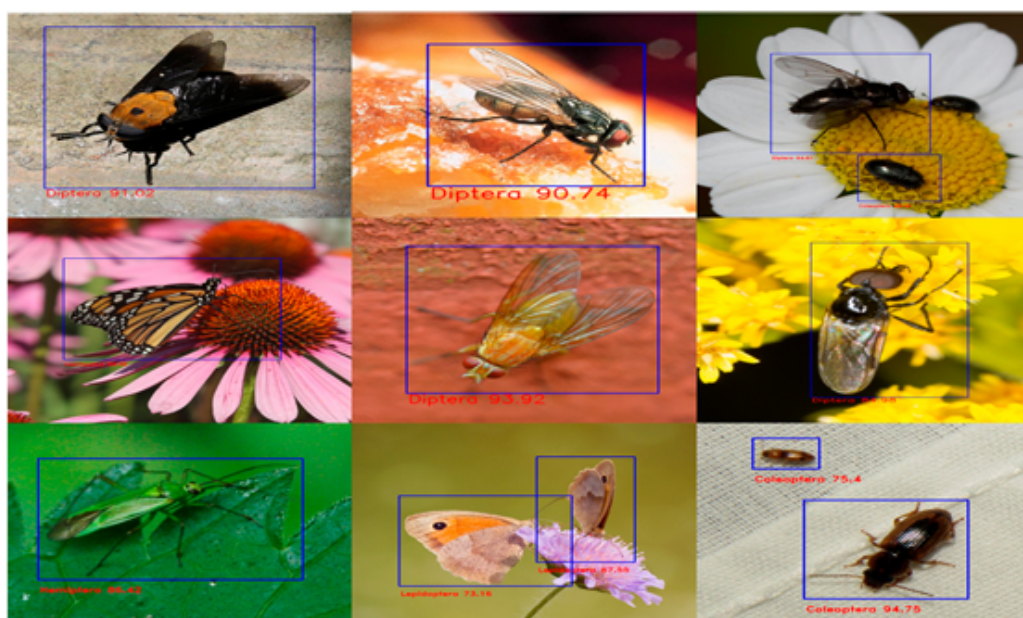
E o outro é Kumar, Nagarathna e Flammini (2023) que utiliza o modelo YOLOv5 para identificar insetos, com uma base de dados de 15.375 imagens separadas em 4 difentes classes de dados sendo essas nomeadas Diptera (mosca-da-fruta, mosca, abelha, mosquito), Odonata (Libélula), Lepidoptera (borboleta) e Araneae (aranha), obteve os resultados mostrados na tabela 2 e com exemplo de predições feitas pelo modelo na figura 6:

Tabela 2 – Resultados do modelo treinado no trabalho de kumar contendo valores de precisão, recall, map50, e f1-score.

Modelo	precisão	recall	mAP50	F1-score
YOLOv5	0.868	0.886	0.922	0.878

Fonte: Kumar, Nagarathna e Flammini (2023).

Figura 6 – Predições feitas pelo modelo treinado no trabalho de kumar corretamento diferenciando os insetos nas imagens.



Fonte: Kumar, Nagarathna e Flammini (2023).

Outro estudo e o mais relevante para este trabalho é sobre a aplicação da YOLO para diferenciar especies de abelhas de Zhang et al. (2024) diferenciando abelhas que polinizam a planta alfalfa , incluindo *Apis mellifera* (honeybee), *Bombus terrestris* (bumblebee) e *Megachile rotundata* (leafcutting bee) eles obtiveram excelentes resultados com os modelos YOLOv3, YOLOv5, YOLOv7 como pode ser observado na tabela 3 e um exemplo de predições realizadas no trabalho de Zhang et al. (2024) na figura 7:

Esses resultados foram obtidos com uma base dados 1000 imagens contendo 1 ou mais abelhas de cada espécie, com o numero de caixa limitadoras de cada espécie totalizando 1567 (honeybee), 2103 (bumblebee), and 1997 (leafcutting bee) essa divisão esta representada no figura 8.

]

Figura 7 – Comparação de métricas entre diferentes versões da YOLO treinadas por Zhang.

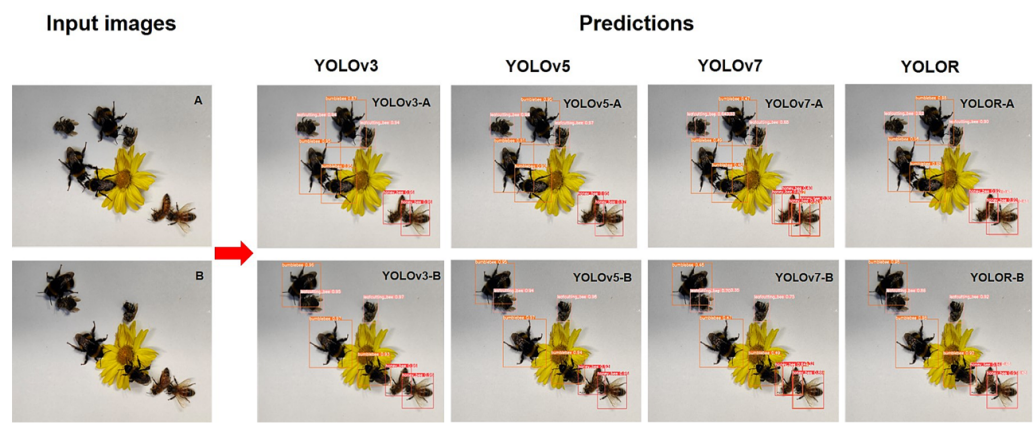


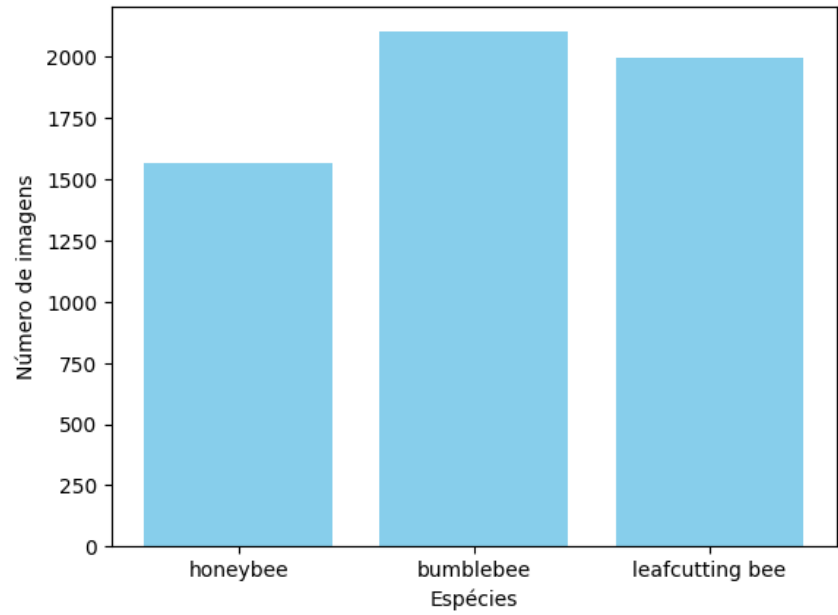
Figura 8 – Predições feitas pelos modelos yolo treinados no trabalho de Zhang.

Fonte: Zhang et al. (2024).

Modelo	precisão	recall	mAP50	F1-score
YOLOv3	0.99922	0.99783	0.99460	0.99852
YOLOv5	0.99922	0.99787	0.99477	0.99854
YOLOv7	0.85014	0.90289	0.94540	0.87572
YOLOR	0.99680	0.98721	0.99323	0.99198

Fonte: Zhang et al. (2024).

Figura 9 – Distribuição entre classes no trabalho de Zhang.



Fonte: Zhang et al. (2024)

4 MATERIAIS E MÉTODOS

Ter uma base de dados de alta qualidade é fundamental para treinar modelos eficazes de classificação. A baixa qualidade dos dados pode levar a previsões tendenciosas ou imprecisas do modelo (MAO; XU; SUAREZ, 2023). A base de dados utilizada disponibilizada pelo trabalho França et al. (2025) que possui 887 imagens que contem uma ou mais abelhas das espécies *Trigona hyalinata* (Guaxupé) e *Trigona spinipes* (Irapuã), criada a partir de imagens obtidas da plataforma iNaturalist (inaturalist.org) e Biodiversity4All (biodiversity4all.org), comunidades globais de ciência-cidadã que reúnem observações georreferenciadas de biodiversidade feitas por usuários voluntários, incluindo pesquisadores, fotógrafos, naturalistas e entusiastas da natureza. Além das imagens públicas, foram incluídas novas imagens de autoria própria, coletadas na cidade de Rio Verde, Goiás. Nesse trabalho, a base de dados foi administrada através do Roboflow (DWYER; NELSON; HANSEN T., 2026). Possuindo ao todo 287 imagens contendo abelhas *Trigona hyalinata* e 600 imagens contendo abelhas *Trigona spinipes*, não houve a utilização de nenhuma ferramenta de aumento de dados.

Figura 10 – Caixa delimitadora ao redor da abelha.

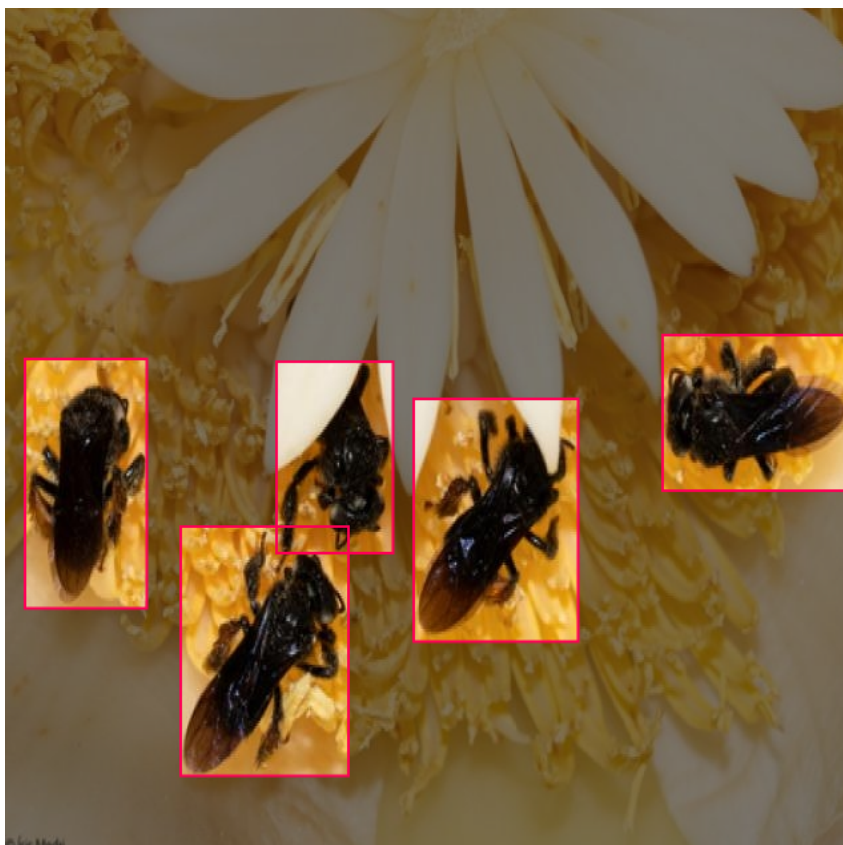


Fonte: próprio autor.

Essas imagens foram rotuladas com caixas delimitadoras com o auxílio do roboflow, semelhantes ao rótulo da Figura 9. Esse processo envolveu a abertura de cada imagem individualmente, onde se desenhou manualmente uma caixa delimitadora ao redor das abelhas presentes na imagem. As coordenadas dessas caixas foram convertidas em etiquetas

(labels) ou anotações, que representam as localizações e dimensões dos objetos (abelhas) de interesse. Junto à caixa delimitadora, foi atribuída a classe da abelha, no caso da Figura 9, Irapuã. Na marcação, foram utilizadas duas classes, uma para cada abelha, cada imagem pode ter um ou mais objetos (abelhas), conforme ilustrado na Figura 10.

Figura 11 – Caixa delimitadora ao redor de multiplas abelha.



Fonte: próprio autor.

A Figura 11 indica a quantidade absoluta de cada classe.

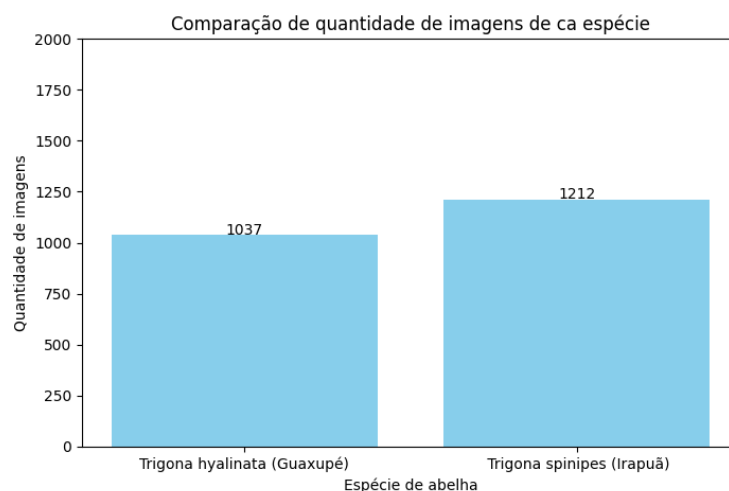
As imagens foram divididas em treino, validação, e teste com 70% em treinamento 15% em validação e 15% em teste. Para o processo de treinamento das distintas versões da YOLO, foram adotados os parâmetros apresentados na Tabela 4, valores amplamente utilizadas para o treinamento de redes YOLO na literatura e valores padrão fornecidos pela ultralytics (ZHANG et al., 2024)(JOCHER; QIU, 2026):

Tabela 3 – Parâmetros de treinamento do modelos YOLO.

Parâmetro	valor
Learning rate	0.01
Batch size	8
Número de iterações(Épocas)	1000
Momomentum	0.937
Weight decay	0.0005

Fonte: próprio autor.

Figura 12 – Distribuição de classes.



Fonte: próprio autor.

Esses valores foram definidos com o objetivo de maximizar o desempenho do modelo. A taxa de aprendizado foi fixada em 0.01, possibilitando atualizações controladas dos pesos ao longo do treinamento, o que contribui para a convergência do modelo e reduz o risco de oscilações excessivas. O tamanho do batch foi configurado em 8, buscando um compromisso entre o custo computacional e a capacidade de generalização. O treinamento foi realizado ao longo de 1000 épocas, assegurando que a rede dispusesse de tempo adequado para assimilar os padrões presentes nos dados. Adicionalmente, foi empregado um valor de momentum igual a 0.937, com a finalidade de acelerar a convergência ao incorporar parte das atualizações anteriores dos pesos, auxiliando o modelo a evitar mínimos locais. Para mitigar o problema de overfitting, aplicou-se uma penalização por decaimento de peso de 0,0005, desencorajando valores excessivos de pesos e favorecendo soluções mais simples e generalizáveis.

Os modelos utilizados foram YOLOv8x, YOLOv9x, YOLOv10x, YOLOv11x, YOLOv12x, e YOLO26x disponíveis em (Ultralytics, 2026) os códigos foram executados por meio de um ambiente python 3.10.12 no visual studio code com o uso das bibliotecas Ultralytics 8.4.9, torch-2.7.0+cu126 com uma placa de vídeo NVIDIA GeForce RTX 4080 SUPER, 15966MiB.

5 RESULTADOS E DISCUSSÕES

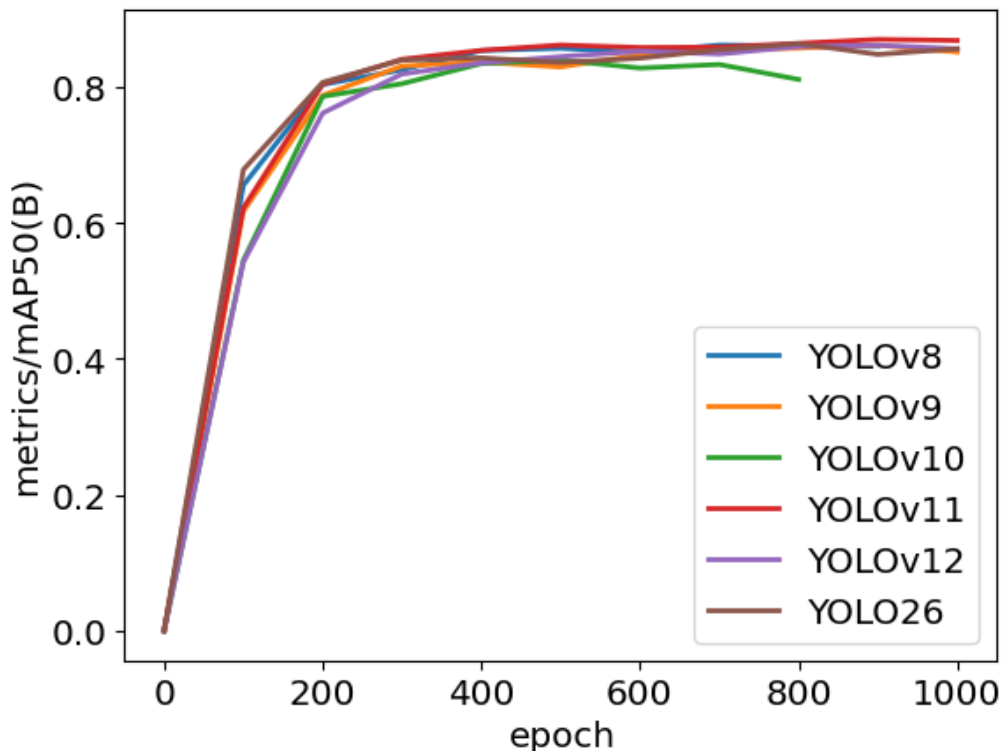
Para todos os treinamentos realizados, os resultados foram avaliados considerando as métricas precisão, recall, F1-score e mAP50, apresentadas na Tabela 5 e a figura 12 demonstra o progresso da métrica de map50 ao longo do treinamento:

Tabela 4 – Comparação de métricas entre diferentes versões da YOLO.

Modelo	precisão	recall	mAP50	F1-score	map50-95	tempo minutos(m)
YOLOv8s	0.87751	0.81478	0.90278	0.84498	0.57207	51.16 m
YOLOv9s	0.88085	0.79681	0.88311	0.83672	0.58782	101.46 m
YOLOv10s	0.82998	0.84766	0.89501	0.83872	0.55845	58.63 m
YOLOv11s	0.90236	0.77475	0.89412	0.83370	0.57667	61.85 m
YOLOv12s	0.83232	0.84117	0.89549	0.83672	0.57000	87.90 m
YOLO26s	0.90317	0.78316	0.90194	0.83889	0.61632	79.9 m

Fonte: próprio autor.

Figura 13 – Progresso da métrica mAP50 por época.



Fonte: próprio autor.

Com relação às métricas que avaliam o desempenho do modelo, a precisão é uma métrica que indica a capacidade do modelo de classificar corretamente os objetos nas imagens. A análise mostra que o modelo YOLO26s foi capaz de atingir um maior valor de precisão que os demais modelos, mas considerando a capacidade dos modelos em identificar todos os objetos relevantes em uma imagem o Recall o modelo YOLO26s deixa a desejar

sendo o melhor valor de recall alcançado pela YOLOv10s seguida da YOLOv12s, após analisar essas duas métricas, precisão e Recall, será analisada a mAP50, que é uma métrica que combina precisão e recall, fornecendo uma avaliação mais completa do desempenho de um modelo de detecção de objetos. Nessa métrica a versão YOLOv9s teve o pior desempenho se comparado com as demais versões. As demais versões da YOLO tiveram desempenho semelhante em mAP50. Contudo a YOLOv8s é a que apresenta o melhor mAP50. Portanto considerando todas as métricas de forma geral o modelo que se saiu melhor com essa base de dados foi a YOLOv8s, considerando isso a figura 13 mostra exemplos de predições corretas realizadas pela YOLOv8s e a figura 14 apresenta o progresso de todas as métricas com relação ao treinamento do modelo.

Figura 14 – Predição feita pela YOLOv8s.



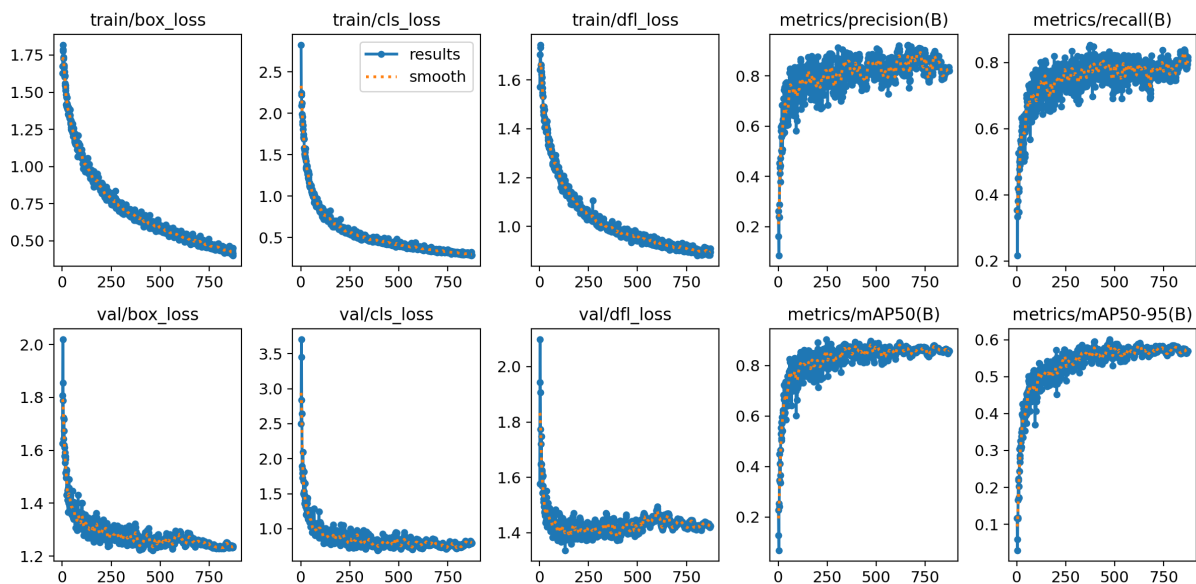
Fonte: próprio autor.

E as figuras de 15 a 19 mostram a progressão das métricas dos modelos yolov9s, yolov10s, yolov11s, yolov12s, yolo26s respectivamente:

O acréscimo das variáveis de perda nas últimas épocas em alguns dos modelos mais notavelmente no modelo YOLOv9s sugere que os modelos começaram a sofrer overfitting e considerando que esses acréscimos começaram em média na época 500 e o alto crescimento das métricas de precisão, recall, map50 e map50-95 nas primeiras épocas sugere que um treinamento de 1000 épocas não é necessário, podendo chegar em bons resultados com treinos mais rápidos.

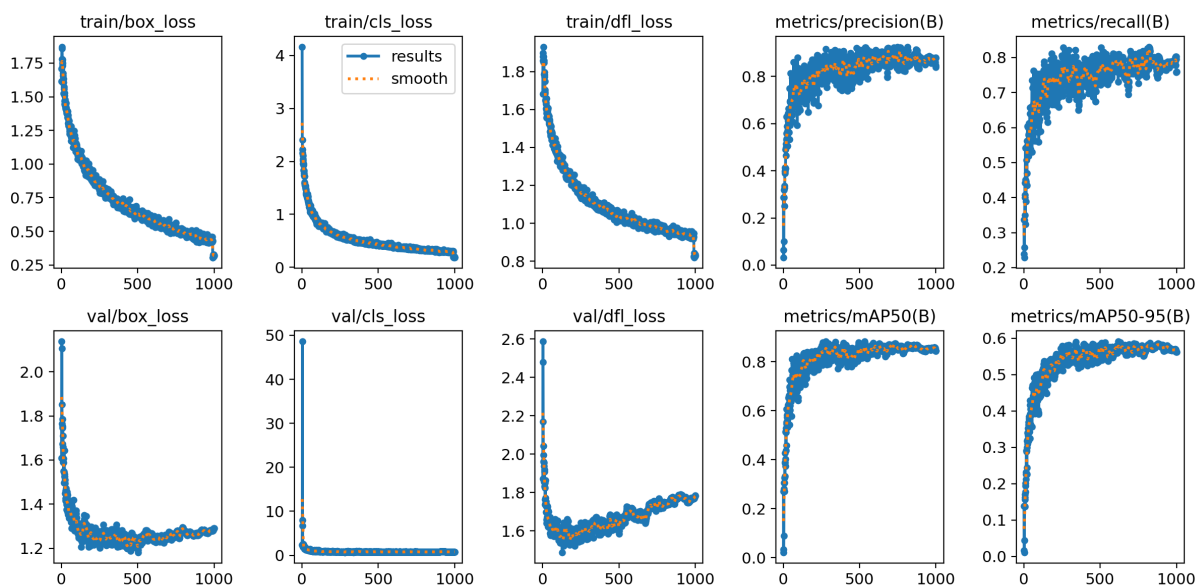
Um dos principais empecilhos para os modelos foi a diferenciação das abelhas do fundo da imagem, que muitas vezes acaba se camuflando no mesmo como mostrado na figura 20:

Figura 15 – Progressão das métricas para YOLOv8s.



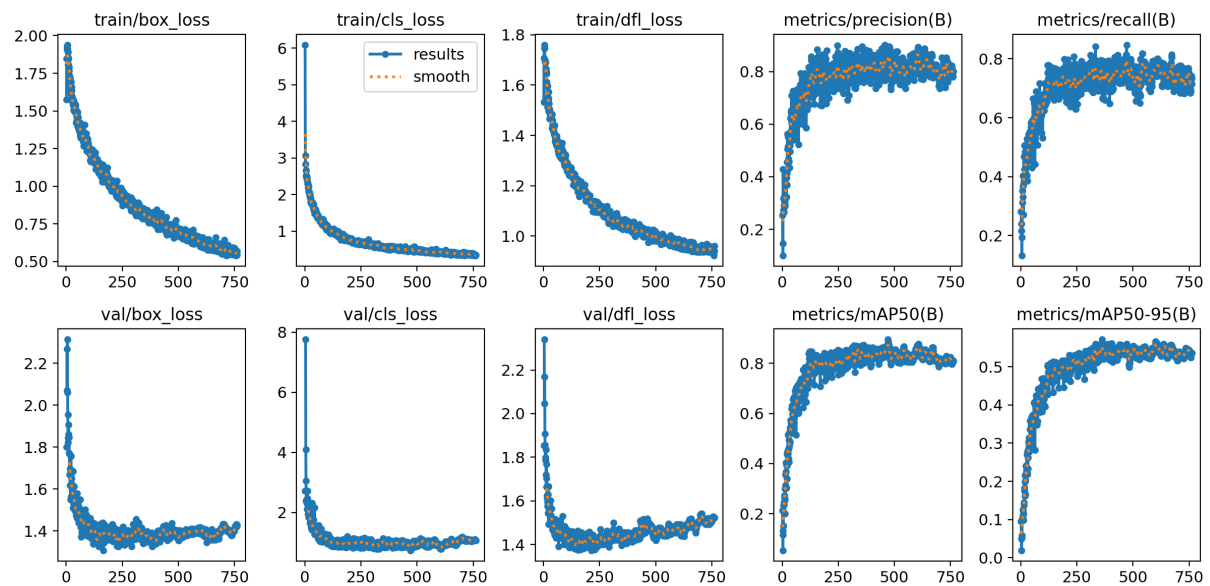
Fonte: próprio autor.

Figura 16 – Progressão das métricas para YOLOv9s.



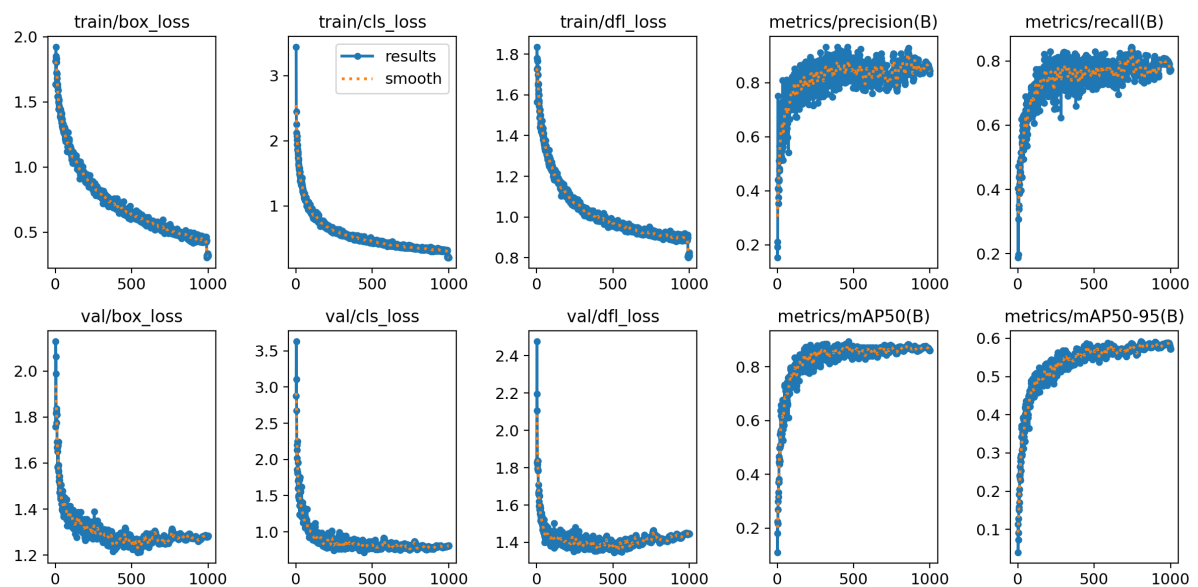
Fonte: próprio autor.

Figura 17 – Progressão das métricas para YOLOv10s.



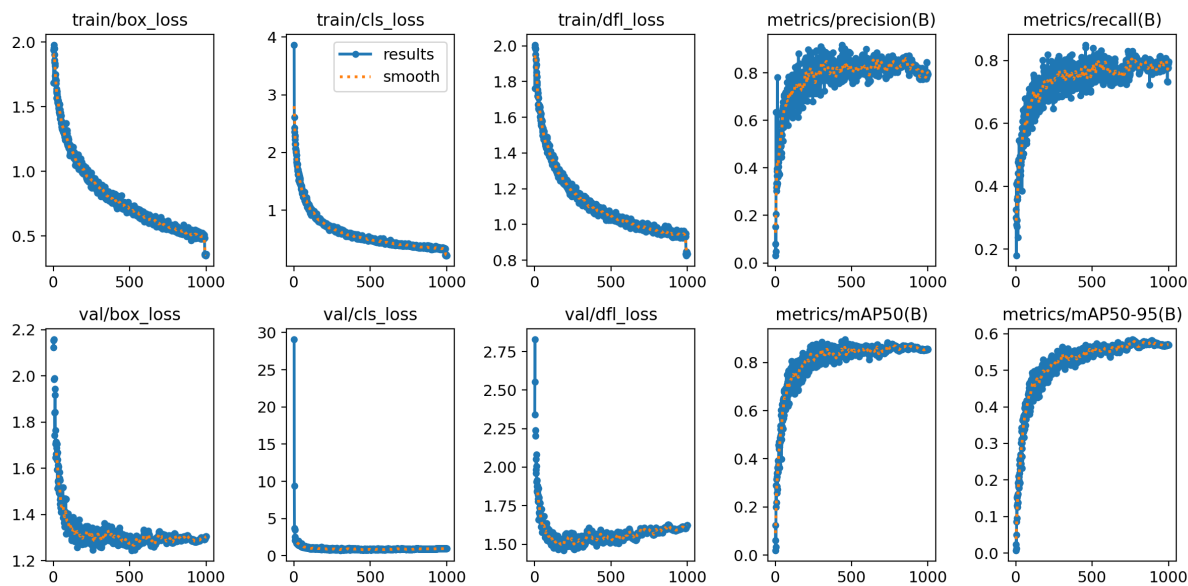
Fonte: próprio autor.

Figura 18 – Progressão das métricas para YOLOv11s.



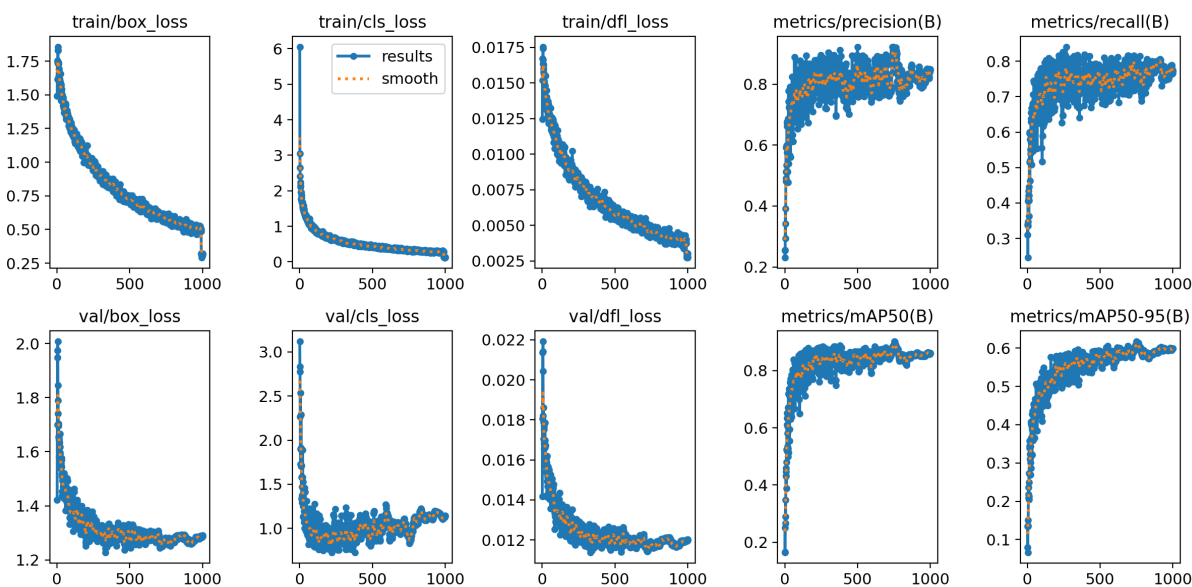
Fonte: próprio autor.

Figura 19 – Progressão das métricas para YOLOv12s.



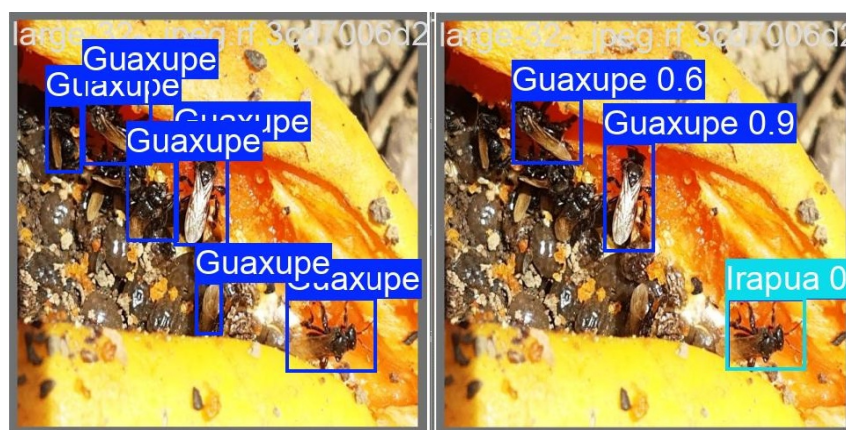
Fonte: próprio autor.

Figura 20 – Progressão das métricas para YOLO26s.



Fonte: próprio autor.

Figura 21 – Predição errada feita pela YOLOv8s, a esquerda temos a caixas anotadas no dataset e a direita a predição realizada pela YOLOv8s.



Fonte: próprio autor.

6 CONCLUSÃO

Este estudo avalia e compara o uso de 6 modelos YOLO, incluindo YOLOv8s, YOLOv9s, YOLOv10s, YOLOv11s, YOLOv12s e YOLO26s, para diferenciar entre duas espécies nativas do cerrado que apresentam grande impedimento taxonômico, sendo essas *Trigona hyalinata* (Guaxupé) e *Trigona spinipes* (Irapuã). Os resultados demonstram que a YOLOv8s apresenta desempenho superior, alcançando bons resultados mesmo levando em consideração que diferente de alguns estudos da literatura as imagens utilizadas neste dataset utiliza imagens de abelhas em ambientes diversos especialmente em ambientes naturais e não controlados, o que aumenta a variabilidade dos dados e dificulta o reconhecimento das espécies já que algumas de suas características podem ser camufladas pelo ambiente que dificulta a identificação e classificação das mesmas pelo modelo YOLO, mas ainda sim os resultados foram favoráveis provando mais uma vez a versatilidade e qualidade dos modelos YOLO para tarefas de identificação e categorização de objetos. Além disso, vale salientar que a qualidade da rede neural depende diretamente da base de dados utilizada (MAO; XU; SUAREZ, 2023), considerando que muitas vezes expandir a base de dados por meio da aquisição de novas imagens é uma tarefa mais complexa do que o próprio treinamento da rede neural convolucional (CNN), os modelos YOLO demonstra ser uma ferramenta eficaz para a automação do reconhecimento e classificação mesmo com bases de dados menores como a utilizada neste trabalho, oferecendo vantagens práticas em aplicações de detecção e reconhecimento de imagem de objetos aos quais não se despoem de uma alta disponibilidade de dados.

Referências

- CARVALHO, M.; BOCKMANN, F.; AMORIM, D. Taxonomic impediment or impediment to taxonomy? a commentary on systematics and the cybertaxonomic-automation paradigm. **Evolutionary Biology**, v. 34, p. 140–143, 2007. Citado na página 1.
- CHEN, H.; ZHOU, G.; JIANG, H. Student behavior detection in the classroom based on improved yolov8. **Sensors**, MDPI, v. 23, n. 20, p. 8385, 2023. Citado na página 2.
- DWYER, B.; NELSON, J.; HANSEN T., e. a. Roboflow (version 1.0) [software]. 2026. Available from <https://roboflow.com>. Citado na página 10.
- FRANÇA, E. L. d. C. et al. Detecção de abelhas nativas do cerrado brasileiro com rede neural convolucional yolov8. **Caderno Pedagógico**, v. 22, n. 11, p. e19970., 2025. Citado 2 vezes nas páginas 7 e 10.
- FREITAS, B.; IMPERATRIZ-FONSECA, V.; MEDINA, L. Diversity, threats and conservation of native bees in the neotropics. **Apidologie**, v. 40, p. 332–346, 2009. <https://doi.org/10.1051/apido/2009012>. Citado na página 1.
- FREITAS, B. M.; SILVA, C. I. **O papel dos polinizadores na produção agrícola no Brasil: A polinização. Agricultura e Polinizadores**. [S.l.]: São Paulo: Associação Brasileira de Estudos das Abelhas, 2015. Citado na página 1.
- HE, K. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: **Proceedings of the IEEE international conference on computer vision**, p. 1026–1034., 2015. Citado na página 7.
- HENDERSON, P.; FERRARI, V. End-to-end training of object class detectors for mean average precision. In: SPRINGER. **Asian conference on computer vision**. [S.l.], 2016. p. 198–213. Citado na página 5.
- IWAMURA, M.; TSUKADA, M.; KISE, K. Automatic labeling for scene text database. In: **2013 12th International Conference on Document Analysis and Recognition**. [S.l.: s.n.], 2013. p. 1365–1369. Citado na página 6.
- JOCHER, G.; CHAURASIA, A.; QIU, J. **Ultralytics YOLOv8**. 2023. Disponível em: <<https://github.com/ultralytics/ultralytics>>. Citado 2 vezes nas páginas 2 e 3.
- JOCHER, G.; QIU, J. **Ultralytics YOLO11**. 2024. Disponível em: <<https://github.com/ultralytics/ultralytics>>. Citado na página 4.
- JOCHER, G.; QIU, J. **Ultralytics YOLO26**. 2026. Disponível em: <<https://github.com/ultralytics/ultralytics>>. Citado 2 vezes nas páginas 4 e 11.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Advances in neural information processing systems**, v. 25, n. 11, 2012. Citado na página 7.
- KUMAR, N.; NAGARATHNA; FLAMMINI, F. Yolo-based light-weight deep learning models for insect detection system with field adaption. **Agriculture**, MDPI, v. 13, n. 3, p. 741, 2023. Citado 2 vezes nas páginas 3 e 8.

- MACEDO, H. et al. Composição florística em formações de cerrado com ação antrópica. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, v. 9, n. 5, p. 76–83, 2015. Disponível em: <https://www.gvaa.com.br/revista/index.php/RVADS/article/view/3148>. Acesso em: 25 nov. 2025. Citado na página 1.
- MAO, Z.; XU, Y.; SUAREZ, E. Dataset management platform for machine learning. **arXiv preprint arXiv:2303.08301**, 2023. Citado 2 vezes nas páginas 10 e 19.
- PADILLA, R.; NETTO, S. L.; SILVA, E. A. D. A survey on performance metrics for object-detection algorithms. In: IEEE. **2020 international conference on systems, signals and image processing (IWSSIP)**. [S.l.], 2020. p. 237–242. Citado 2 vezes nas páginas 4 e 5.
- POWERS, D. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. **arXiv preprint arXiv:2010.16061**, 2020. Citado na página 5.
- REDMON, J. et al. You only look once: Unified, real-time object detection. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2016. p. 779–788. Citado na página 2.
- REIS, A. C. B. C. Cerrado brasileiro: desafios e perspectivas na implementação de políticas públicas para conservação e desenvolvimento sustentável. **Contribuciones a las Ciencias Sociales**, v. 17, n. 7, p. e8501, 2024. Citado na página 1.
- Secretaria de Meio Ambiente, Infraestrutura e Logística do Estado de São Paulo. **Fichas de Identificação das Abelhas Paulistas**. 2023. Disponível em: <<https://semil.sp.gov.br/educacaoambiental/prateleira-ambiental/fichas-de-identificacao-das-abelhas-paulistas/>>. Citado na página 2.
- SHEN, L.; LANG, B.; SONG, Z. Ds-yolov8-based object detection method for remote sensing images. **Ieee Access**, IEEE, v. 11, p. 125122–125137, 2023. Citado na página 2.
- SPIESMAN, B. J. Assessing the potential for deep learning and computer vision to identify bumble bee species from images. **Scientific Reports**, v. 11, n. 1, p. 1–10, 2021. Citado na página 7.
- TANG, Y. et al. An improved yolo network for small target insects detection in tomato fields. **Computers and Electronics in Agriculture**, Elsevier, v. 239, p. 110915, 2025. Citado na página 7.
- TIAN, Y.; YE, Q.; DOERMANN, D. **YOLO12: Attention-Centric Real-Time Object Detectors**. 2025. Disponível em: <<https://github.com/sunsmarterjie/yolov12>>. Citado na página 4.
- TORRES, A. M.; SOBREIRO, A. I.; JUNIOR, V. V. A. Diversidade de abelhas em ambiente rural em área de transição entre os biomas de mata atlântica e cerrado. **Observatório de la Economía Latinoamericana**, v. 21, n. 10, p. 17327–17343, 2023. Citado na página 1.
- Ultralytics. **Ultralytics: AI Vision**. 2026. Disponível em: <<https://www.ultralytics.com/>>. Citado na página 12.
- WANG, A.; CHEN, H.; LIU, e. a. L. Yolov10: Real-time end-to-end object detection. **arXiv preprint arXiv:2405.14458**, 2024. Citado na página 4.

WANG, C.-Y.; LIAO, H.-Y. M. YOLOv9: Learning what you want to learn using programmable gradient information. 2024. Citado na página 3.

YANG, G. et al. A lightweight YOLOv8 tomato detection algorithm combining feature enhancement and attention. **Agronomy**, MDPI, v. 13, n. 7, p. 1824, 2023. Citado na página 2.

ZHANG, C.-J. et al. Evaluation of the YOLO models for discrimination of the alfalfa pollinating bee species. **Journal of Asia-Pacific Entomology**, v. 27, n. 1, p. 102195, 2024. ISSN 1226-8615. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1226861523001590>>. Citado 3 vezes nas páginas 8, 9 e 11.