

# Rastreamento de Abelhas da Espécie *Melipona capixaba* utilizando Visão Computacional

Marcos Veniciu Sá Barbalho<sup>1</sup>, Tiago N. Matos<sup>1</sup>, Ingrid Gomes<sup>1</sup>,  
Ricardo S. Ferreira<sup>1</sup>, Helder Canto Resende<sup>1</sup>, José Augusto M. Nacif<sup>1</sup>

<sup>1</sup>Universidade Federal de Viçosa (UFV) – Viçosa, MG – Brasil

{marcos.barbalho, tiago.nunes, ricardo}@ufv.br,  
{helder.resende, jnacif}@ufv.br, ingomes00@gmail.com

**Abstract.** *Monitoring the foraging of *Melipona capixaba*, an endangered species, is vital for conservation, yet manual methods are invasive. In this paper, we propose a non-invasive, cloud-based computer vision system for continuous directional tracking at the colony entrance. Evaluating YOLO11n, YOLO26n, and RT-DETR with ByteTrack reveals RT-DETR excels in GPUs (74.29% IDF1), fully meeting the demands of biologists. In the feasibility assessment for edge computing using a Raspberry Pi 5, RT-DETR fails under INT8 quantization, whereas YOLO models remain resilient for future autonomous field analysis.*

**Resumo.** *Monitorar o forrageamento da *Melipona capixaba*, uma espécie ameaçada de extinção, é vital para sua conservação, mas métodos manuais são invasivos. Neste trabalho, propomos um sistema não invasivo de visão computacional em nuvem para rastreamento direcional contínuo na colônia. Avaliando YOLO11n, YOLO26n e RT-DETR com ByteTrack, o RT-DETR obteve precisão superior em GPU (IDF1 de 74,29%), o que viabiliza análises ecológicas precisas. Na avaliação de viabilidade para processamento embarcado na borda com Raspberry Pi 5, o RT-DETR falhou na inferência sob quantização INT8, enquanto os modelos YOLO mostraram-se resilientes para futuras análises autônomas em campo.*

## 1. Introdução

A polinização por abelhas nativas sem ferrão é vital tanto para a biodiversidade neotropical quanto para a agricultura, mas suas populações enfrentam um declínio alarmante devido à degradação ambiental [Nogueira et al. 2014]. Para formular estratégias de conservação para espécies endêmicas ameaçadas, como a *Melipona capixaba*, é essencial monitorar a atividade de forrageamento das colônias [Soudeep et al. 2024, Lei et al. 2024]. Historicamente, a coleta desses dados encontrou limitações decorrentes de métodos manuais e invasivos, tornando o desenvolvimento de sistemas de monitoramento remoto um desafio urgente na ecologia moderna [Odemer 2022, Lei et al. 2024, Sledevič et al. 2024].

O monitoramento contínuo das entradas e saídas no ninho é um bioindicador valioso da saúde da colônia e da integridade do ecossistema circundante [Odemer 2022], contudo, aplicá-lo à colônia em campo impõe severos desafios técnicos [Soudeep et al. 2024, Silva et al. 2023]. Embora arquiteturas de aprendizado profundo, como o YOLO, sejam eficazes na detecção isolada [Jocher and Qiu 2024], manter a identidade (IDs) de

múltiplos insetos no rastreamento contínuo (MOT) sob luz natural instável continua sendo um grande gargalo computacional [Zhang et al. 2022, Jocher and Qiu 2024].

Para superar esses desafios, propomos um *pipeline* de visão computacional e hardware embarcado (Raspberry Pi 5) para quantificar o forrageamento da *Melipona capixaba*. Diferenciando-se das soluções predominantemente focadas na *Apis mellifera* e em análises de quadros estáticos isolados [Lei et al. 2024, Silva et al. 2023, Zhang et al. 2022], este trabalho foca na consistência temporal (MOT), que é a capacidade do algoritmo de manter a identidade de um mesmo indivíduo entre quadros consecutivos. Isso é fundamental para a *Melipona capixaba*, pois as densas aglomerações geram severas oclusões visuais (sobreposição de corpos), o que frequentemente confunde detectores tradicionais e corrompe a contagem [Zhang et al. 2022, Jocher and Qiu 2024]. Essa abordagem viabiliza a extração contínua de métricas sob condições reais, preenchendo a lacuna entre a detecção tradicional e a análise ecológica de longo prazo.

Os resultados validam a viabilidade do *pipeline* para futuras aplicações em computação de borda. No processamento sem restrições de *hardware* (GPU), o modelo RT-DETR estabeleceu o estado da arte para a aplicação, atingindo um IDF1 de 74,29% no rastreamento contínuo. Esse índice comprova a eficácia do sistema em preservar a identidade do inseto sob oclusão mútua. Para a ecologia, isso se traduz na mitigação de contagens duplicadas, garantindo a integridade dos dados de longo prazo. Como um passo evolutivo rumo à implantação em computação de borda, as avaliações de viabilidade para hardware limitado (simulando inferência em CPU) demonstraram que, embora a arquitetura RT-DETR sofra um colapso sob quantização máxima (INT8), os modelos da família YOLO mantêm notável resiliência e alta precisão mesmo sob severa compressão de dados. Isso valida a viabilidade técnica da arquitetura convolucional para futuras implantações contínuas e locais diretamente no campo.

O artigo detalha a arquitetura e os algoritmos na seção Metodologia de Identificação e Rastreamento, apresenta os cenários multicondicionais na Metodologia de Avaliação e discute a precisão do sistema em Resultados Preliminares e Discussão. Por fim, a Conclusão sintetiza as descobertas e os próximos passos.

## 2. Trabalhos Relacionados

O monitoramento automatizado de colmeias evoluiu significativamente, substituindo antigos sensores físicos por processamento de imagens e vídeos [Odemer 2022]. Pesquisas recentes destacam o sucesso da Visão Computacional e do *Machine Learning* na análise de tráfego, detecção de pólen e identificação de parasitas [Bilik et al. 2022]. Contudo, essas tecnologias focam quase exclusivamente na *Apis mellifera*, havendo uma clara escassez de estudos voltados para abelhas nativas sem ferrão (*Meliponini*).

No âmbito do rastreamento direcional, abordagens utilizando Redes Neurais Convolucionais (CNNs) e Rastreamento de Múltiplos Objetos (MOT), como as arquiteturas da família YOLO, demonstram alta precisão na avaliação do forrageamento [Lei et al. 2024]. A literatura recente expandiu essas capacidades para estimar a orientação de voo via *key-points* [Sledevič et al. 2024], identificar comportamentos complexos na entrada do ninho [Sledevič et al. 2025] e monitorar fluxos de pólen em tempo real [Silva et al. 2023].

A integração desses modelos visuais com arquiteturas de Internet das Coisas (IoT) tem viabilizado redes de monitoramento contínuo [Dsouza and Hegde 2023,

Mahamud et al. 2019]. Entretanto, como os sistemas existentes estão fortemente otimizados para a morfologia e dinâmica da *Apis mellifera*, surge uma lacuna tecnológica significativa no estudo de polinizadores nativos. Este trabalho se diferencia ao adaptar e aplicar modelos estado da arte de detecção e rastreamento para solucionar os desafios visuais específicos da *Melipona capixaba*, fornecendo um método não invasivo e escalável para auxiliar na conservação desta espécie ameaçada.

**Tabela 1. Comparativo entre trabalhos correlatos de monitoramento por visão computacional e a abordagem proposta.**

Referência	Espécie-Alvo	Foco Principal da Análise	Abordagem de Visão	Embarcado na borda	Trata Oclusão
Bilik et al. (2022)	<i>Apis mellifera</i>	Tráfego, pólen e parasitas	Revisão (Múltiplas abordagens)	Variada	Não
Lei et al. (2024)	<i>Apis mellifera</i>	Forrageamento (In/Out)	Rastreamento (MOT)	Variada	Sim (via MOT)
Silva et al. (2023)	<i>Apis mellifera</i>	Fluxo de pólen em tempo real	Deteção de Objetos	Variada	Não
Sledevič et al. (2024, 2025)	<i>Apis mellifera</i>	Orientação de voo e comportamentos	Estimativa via <i>keypoints</i> (Pose)	Variada	Parcial
Dsouza & Hegde (2023)	<i>Apis mellifera</i>	Redes de monitoramento contínuo	Não se aplica (Sensores IoT)	Sim (IoT)	N/A
<b>Nosso Trabalho</b>	<i>Melipona capixaba</i>	<b>Fluxo direcional contínuo</b>	<b>Consistência temporal (MOT)</b>	<b>Simulação em CPU</b>	<b>Sim</b>

### 3. Metodologia de Identificação e Rastreamento

Para atender simultaneamente às restrições do ambiente de campo e à necessidade de precisão da pesquisa biológica, a arquitetura do sistema foi dividida em duas frentes. A primeira consiste em um nó de aquisição na borda, responsável pelo registro em condições reais. A segunda frente concentra a inferência computacional em nuvem, garantindo a confiabilidade dos dados ecológicos.

#### 3.1. Aquisição de Dados e Hardware

Para atuar como o nó central de aquisição de imagens em campo, o sistema empregou um microcomputador Raspberry Pi 5 (8 GB) acoplado a uma Camera Module 3, posicionada lateralmente à entrada da colônia. Para capturar o voo rápido, gravou-se em 1280x720 a 120 FPS, para reduzir o *motion blur*, foi necessário aumentar a velocidade do obturador (*shutter speed*) e a compensação de ganho analógico. A calibração de foco e posicionamento foi realizada via *USB tethering* e VNC, reduzindo a dependência de monitor e teclado. Os vídeos foram segmentados em blocos de 1 minuto e capturados apenas sob luz natural (04:00 às 19:00).

Durante a operação em campo, observou-se que o fluxo de exaustão unidirecional do Raspberry Pi Active Cooler concentrava o ar quente diretamente na base do equipamento. Devido a esse acúmulo térmico, que era severamente agravado pela incidência solar, foi necessário desenvolver uma *case* impressa em 3D (filamento ABS) com exaustão aprimorada para prevenir o superaquecimento e a deformação da estrutura.

#### 3.2. Detecção de Objetos (Dataset e Modelo)

A construção do conjunto de dados (5.036 imagens, divididas nas classes *bee* (6189 anotações) e *pollen* (330 anotações)) empregou heurísticas específicas de curadoria. Em cenários de baixa luminosidade, utilizou-se a “Densidade de Informação”, priorizando *frames* ativos; sob luz diurna, adotou-se a divisão por quadrantes para mitigar vieses espaciais. A anotação preservou a variabilidade real, englobando desde instâncias isoladas até densos agrupamentos de insetos na entrada da colmeia.

Após a partição estatística (85% treino, 9% validação, 6% teste) e o redimensionamento para 640x640 *pixels*, aplicou-se um robusto conjunto de técnicas de

*Data Augmentation* ao conjunto de treino. As transformações incluíram: *Horizontal Flip* para invariância geométrica direcional [Bilik et al. 2022], variações de brilho e exposição simulando o ciclo circadiano [Dsouza and Hegde 2023, Buslaev et al. 2020], e injeção de ruído gaussiano para simular a granulação de sensores de baixo custo [Hernández-Cámara et al. 2023].

Avaliamos três arquiteturas estado da arte para a detecção de pequenos objetos: YOLO11n [Jocher and Qiu 2024], YOLO26n [Jocher and Qiu 2026] e RT-DETR [Zhao et al. 2023]. Para garantir a acurácia exigida pelos parâmetros biológicos, tanto o treinamento quanto a inferência principal e a validação de todo o *pipeline* ocorreram em um ambiente de nuvem acelerado por GPU. O treinamento para a família YOLO utilizou o otimizador AdamW (taxa de aprendizado inicial de 0,001667, *momentum* de 0,9, *weight decay* de 0,0005 e 3 épocas de *warmup*), configurado para 200 épocas e *batch size* de 32. Para a arquitetura RT-DETR, adaptou-se a taxa de aprendizado inicial para 0,0001 sob o mesmo otimizador. Manteve-se a técnica de *Mosaic Augmentation*, essencial para a identificação de oclusões parciais e alvos diminutos [Bochkovskiy et al. 2020]. O conjunto de dados, códigos-fonte e os arquivos de configuração (YAML) detalhados estão publicamente disponíveis no GitHub<sup>1</sup>.

### 3.3. Algoritmo de Rastreamento e Lógica de Contagem

Para associar as detecções e reconstruir as trajetórias, implementou-se o algoritmo MOT ByteTrack [Zhang et al. 2022], capaz de mitigar a perda de identidades (IDs) ao não descartar predições de baixa confiança durante oclusões severas. O rastreador foi configurado com um limiar de confiança alto (*high threshold*) de 0,50, um limiar baixo (*low threshold*) de 0,10 e um limite de associação geométrica baseado em IoU (*match threshold*) de 0,95. O centroide da *bounding box* foi adotado como referencial espacial. Parametrizou-se um *track\_buffer* de 360 *frames* (equivalente a uma tolerância de 3 segundos a 120 FPS), prevenindo a fragmentação prematura das identidades nos agrupamentos.

A Região de Interesse (ROI) consistiu em uma área quadrada ao redor do orifício de entrada. Como a *Melipona capixaba* não dispõe de rampa de pouso externa, a ROI recebeu uma margem propositalmente superdimensionada para absorver oscilações mecânicas da câmera induzidas por ventos, garantindo a estabilidade do referencial de entrada.

A lógica de contagem do fluxo direcional baseou-se nas coordenadas do primeiro e último registro de cada ID rastreado:

- **Entrada (In):** O ID surge fora da ROI e seu rastreamento encerra no interior da ROI, inferindo a entrada na colmeia.
- **Saída (Out):** O ID surge no interior da ROI e desaparece no ambiente externo.
- **Voos Incertos:** A trajetória inicia e termina na mesma zona (ambos dentro ou ambos fora da ROI). Refletem dinâmicas de reconhecimento ou guarda, e não são computados como fluxo de forrageamento.

## 4. Metodologia de Avaliação e Cenários de Teste

Para garantir que o *pipeline* de monitoramento atenda às exigências do ambiente não controlado e viabilize a futura execução em hardware embarcado, o protocolo de testes

<sup>1</sup><https://github.com/lesc-ufv/wcama2026-melipona-tracking.git>

foi dividido em duas frentes de avaliação fundamentais: a análise estática de extração de características (Métricas de Detecção) e a análise dinâmica de consistência temporal (Métricas de Rastreamento).

Todos os experimentos de inferência e cálculo de métricas foram conduzidos no ambiente virtual Google Colab, alocando-se uma Unidade de Processamento Gráfico (GPU) NVIDIA Tesla T4 (15 GB de VRAM) e uma Unidade de Processamento Central (CPU) Intel Xeon @ 2.20 GHz com 12 GB de memória RAM.

#### 4.1. Métricas do Modelo de Detecção

A primeira fase da avaliação focou na capacidade intrínseca das arquiteturas (YOLO11n, YOLO26n e RT-DETR) de localizar e classificar os alvos (abelhas e pólen) em quadros isolados. Para esta etapa, utilizou-se o conjunto de teste isolado do dataset, composto por 504 imagens estáticas que não foram apresentadas aos modelos durante a fase de treinamento. A qualidade das predições foi mensurada utilizando as métricas clássicas de Visão Computacional:

- **Precisão (Precision) e Revocação (Recall):** Avaliam, respectivamente, a proporção de detecções corretas em relação ao total de predições e a capacidade do modelo de encontrar todas as abelhas reais presentes na imagem.
- **mAP@50 e mAP@50-95 (Mean Average Precision):** A precisão média global foi calculada sob os limiares de Intersection over Union (IoU) de 50% e no intervalo de 50% a 95% (mAP@50-95), avaliando não apenas a detecção do alvo, mas também a precisão do seu isolamento espacial.

O desempenho ideal, servindo como limite superior, foi estabelecido no ambiente utilizando a GPU Tesla T4. No entanto, para simular o cenário real de *deploy* em dispositivos com recursos limitados, como o microcomputador Raspberry Pi, os modelos também foram submetidos à avaliação em um regime de processamento exclusivo de CPU. Neste cenário de CPU, foram empregados formatos de exportação otimizados: ONNX (nas precisões de ponto flutuante FP32 e FP16) e OpenVINO quantizado em inteiros de 8 bits (INT8).

#### 4.2. Avaliação de Rastreamento (MOTA)

O principal diferencial desta pesquisa reside na transição da detecção de quadros isolados para o rastreamento contínuo em vídeo. Para validar a confiabilidade do algoritmo ByteTrack acoplado aos detectores, desenvolveu-se um cenário de teste focado na interação simultânea de múltiplos insetos..

Utilizou-se um videoclipe de 10 segundos extraído do ambiente de campo, selecionado especificamente por registrar um pico de forrageamento com pelo menos três abelhas interagindo simultaneamente na zona de entrada da colmeia. As posições reais de cada inseto foram anotadas manualmente, frame a frame (gerando o arquivo de *Ground Truth*), para formar a base de comparação.

Para avaliar a assertividade das trajetórias, adotou-se a biblioteca `motmetrics` para calcular métricas padronizadas pelo MOTChallenge, sendo as principais:

- **MOTA (Multiple Object Tracking Accuracy):** Métrica global que compila penalidades por Falsos Positivos (FP), Falsos Negativos (FN) e Trocas de Identidade (ID Switches).

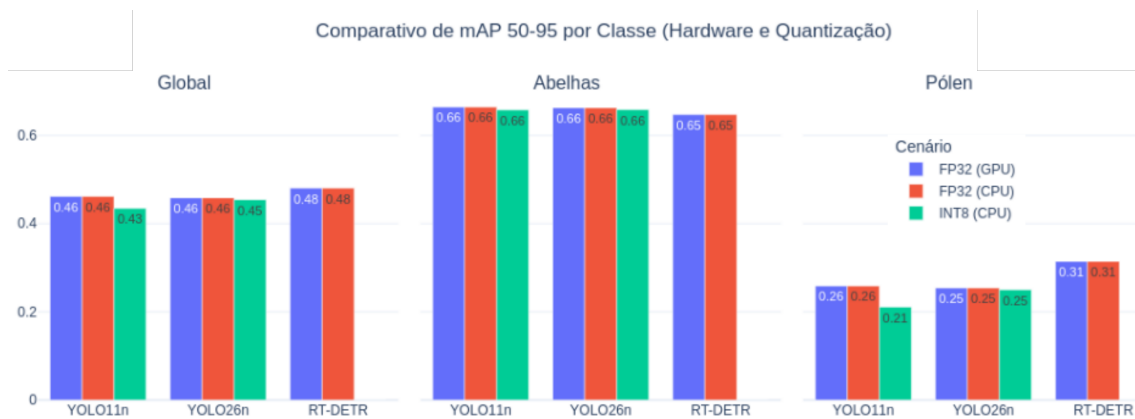
- **IDF1 (ID F1-Score):** Foca especificamente na capacidade do rastreador de manter a mesma identidade (ID) atrelada ao mesmo alvo ao longo de toda a sua trajetória.
- **ID Switches (IDSW):** Quantifica o número de vezes que uma trajetória é fragmentada e um novo ID é atribuído erroneamente a um inseto que já estava sendo rastreado.

No contexto da Apicultura de Precisão e do monitoramento de *Melipona capixaba*, o índice de ID Switches é o fator mais crítico do cenário de teste. A perda temporária de uma *bounding box* devido a oclusões pode ser tolerada; contudo, se o algoritmo trocar a identidade do inseto durante um cruzamento de voo, o sistema interpretará erroneamente que uma abelha entrou na colmeia e uma nova abelha saiu, comprometendo a contagem do bioindicador de forrageamento.

## 5. Resultados Preliminares e Discussão

Para refletir a arquitetura proposta e a demanda biológica, a discussão dos resultados está estruturada em duas frentes. Primeiro, avaliamos o desempenho do *pipeline* sem restrições de *hardware* (em nuvem com GPU), demonstrando a capacidade do sistema de atingir a precisão de identificação e rastreamento direcional exigida pela pesquisa ecológica. Em seguida, apresentamos uma avaliação de viabilidade técnica simulando restrições de processamento, visando o passo evolutivo de implantação do sistema na borda.

### 5.1. Desempenho de Detecção

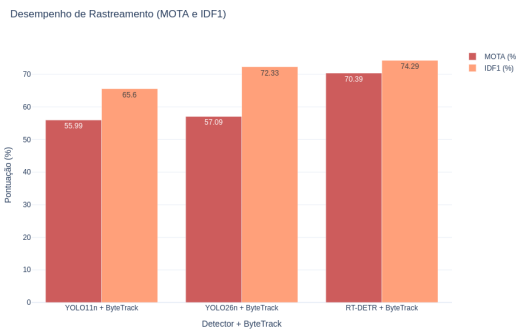


**Figura 1. Avaliação da métrica mAP 50-95 para os modelos YOLO11n, YOLO26n e RT-DETR nas categorias global, abelhas e pólen sob diferentes cenários de precisão computacional.**

A análise estática das arquiteturas, avaliada pela métrica mAP 50-95, revela diferenças relevantes na capacidade de generalização dos modelos sob diferentes níveis de precisão (FP32, FP16 e INT8). No cenário de avaliação global com pontos flutuantes (FP32 e FP16), o modelo baseado em Transformers, RT-DETR, apresenta uma leve superioridade, alcançando um mAP 50-95 de 0,48, em comparação com a marca de 0,46 obtida tanto pelo YOLO11n quanto pelo YOLO26n. Apesar do resultado promissor, a classe *pollen* possui forte desbalanceamento (330 anotações contra 6.189 de abelhas).

Neste cenário restrito, o pólen atua como um alvo secundário probatório, mantendo-se o fluxo direcional das abelhas como a métrica primária do sistema. Além das limitações do conjunto de dados, a aplicação da quantização extrema (INT8) expõe as limitações dessa precisão global, causando uma degradação no desempenho do YOLO11n para o pólen (queda para 0,21) e a falha de inferência da arquitetura RT-DETR, que não apresenta resultados viáveis nesse formato de compressão.

## 5.2. Consistência Temporal e Rastreamento de Múltiplos Alvos



**Figura 2.** Resultados das métricas de rastreamento (MOTA e IDF1) comparando as arquiteturas.



**Figura 3.** Análise dos erros empilhados e composição das penalidades no rastreamento.

O objetivo final do sistema não é a contagem estática de instâncias, mas o rastreamento direcional contínuo. A validação dinâmica, acoplando as detecções da GPU ao algoritmo ByteTrack, confirmou o RT-DETR como a arquitetura mais assertiva espacialmente para este fim (**Figuras 2 e 3**).

Na literatura de rastreamento de múltiplos objetos (MOT), o MOTA (*Multiple Object Tracking Accuracy*) é uma métrica reconhecidamente punitiva, pois aplica dupla penalidade a variações espaciais e contabiliza sucessivos falsos negativos caso um alvo sofra oclusão por breves frações de segundo [Milan et al. 2016]. No contexto biológico da *Melipona capixaba* caracterizado por alvos pequenos, visualmente uniformes, e severa oclusão mútua na entrada da colmeia, alcançar um MOTA de 70,39% (RT-DETR) e 57,09% (YOLO11n) evidencia um desempenho altamente promissor para a aplicação.

Para a Apicultura de Precisão, o indicador mais relevante da viabilidade do contador bidirecional é o IDF1 (*ID F1-Score*), que avalia a capacidade do modelo em preservar a identidade correta do inseto ao longo de sua trajetória. O RT-DETR atingiu 74,29% de IDF1, apresentando apenas 6 Trocas de Identidade (*ID Switches*) durante os cruzamentos de voo. Esta baixa taxa de fragmentação demonstra que o sistema é capaz de minimizar contagens duplicadas ou “viagens fantasmas”, garantindo a confiabilidade dos dados para análises ecológicas de longo prazo.

Na validação prática da contagem direcional (Tabela 2), comparou-se o fluxo do sistema com o *Ground Truth* (3 entradas, 2 saídas). O RT-DETR alcançou Erro Absoluto nulo ( $EA = 0$ ), contabilizando perfeitamente a dinâmica espacial. Em contraste, os modelos YOLO subestimaram as entradas ( $EA = 2$ ) devido à fragmentação por *ID Switches*.

Esse resultado consolida a eficácia empírica do RT-DETR em mitigar contagens fantasmas sob oclusão severa.

**Tabela 2. Validação prática da contagem direcional e Erro Absoluto comparado ao *Ground Truth* (3 In, 2 Out).**

Modelo	Entradas (Pred)	Erro (In)	Saídas (Pred)	Erro (Out)
YOLO11n	1	2	1	1
YOLO26n	1	2	2	0
RT-DETR	3	0	2	0

### 5.3. Viabilidade em Computação de Borda e os Efeitos da Quantização

**Tabela 3. Resultados de Detecção para a classe Abelhas**

Modelo	Formato	Precision_bee	Recall_bee	mAP50_bee	mAP50-95_bee
RT-DETR	FP32	0,960	0,974	0,982	0,647
YOLO11n	FP32	0,953	0,965	0,982	0,664
YOLO11n	FP16	0,953	0,965	0,982	0,664
YOLO11n	INT8	0,925	0,972	0,982	0,658
YOLO26n	INT8	0,942	0,957	0,979	0,659
YOLO26n	FP32	0,941	0,958	0,979	0,663
YOLO26n	FP16	0,940	0,957	0,977	0,663
RT-DETR	INT8	0,000	0,000	0,000	0,000

Uma vez atestada a eficácia do modelo em nuvem para suprir a demanda da biologia, a próxima etapa metodológica busca a autonomia da estação de campo. A futura implantação de sistemas de visão computacional diretamente no nó de aquisição, como o Raspberry Pi 5, exige redes neurais altamente eficientes. Para avaliar essa viabilidade de inferência em computação de borda, simulamos o processamento restrito à CPU e avaliamos a quantização dos pesos do modelo para inteiros de 8 bits (INT8), uma técnica padrão para a borda, mas que pode acarretar degradação de acurácia. A Tabela 3 detalha os impactos dessa degradação especificamente para a detecção da classe “Abelhas”.

As arquiteturas convolucionais demonstraram uma alta resiliência à compressão de dados. O YOLO11n, por exemplo, reduziu seu mAP 50-95 de forma marginal, caindo de 0,664 (FP32) para apenas 0,658 em INT8, enquanto manteve seu mAP 50 praticamente inalterado em 0,982. Embora sua precisão tenha sofrido uma leve queda (de 0,953 para 0,925 ao passar de FP32 para INT8), o modelo compensou com um ligeiro aumento na revocação (Recall de 0,965 para 0,972), indicando que ele se torna levemente mais permissivo, porém sem perder a capacidade de encontrar os alvos. O YOLO26n apresentou estabilidade semelhante, mantendo seu mAP 50-95 em 0,659 sob INT8.

Em contrapartida, o RT-DETR expôs uma vulnerabilidade arquitetural severa. Apesar de estabelecer o estado da arte no formato de ponto flutuante (FP32) com uma Precisão de 0,960 e mAP 50 de 0,982, o modelo sofre de uma perda massiva na capacidade de inferência ao ser submetido à quantização máxima em INT8, resultando em métricas nulas (0,0) em todas as frentes. Esses resultados empíricos evidenciam que, embora a arquitetura Transformer seja superior em ambientes com aceleração por GPU,

ela sofre de limitações críticas ao ser comprimida para execução em CPU. Consequentemente, a falha do RT-DETR exclui seu uso em cenários 100% embarcados, consolidando as redes da família YOLO como uma solução tecnológica mais resiliente e estável para inferência na borda em apicultura de precisão.

## 6. Conclusão

Este trabalho validou um *pipeline* não invasivo de visão computacional para monitorar a *Melipona capixaba*, utilizando hardware de baixo custo (Raspberry Pi 5 e Camera Module 3) para superar as limitações das inspeções manuais na ecologia. A principal contribuição reside na entrega de um sistema funcional, capaz de quantificar o fluxo direcional direcional das abelhas, minimizar a dupla contagem e preservar a integridade do ninho.

A análise revelou um importante *trade-off* tecnológico: em ambientes com GPU, o RT-DETR acoplado ao ByteTrack estabeleceu o estado da arte para a aplicação, lidando bem com oclusões severas e alcançando a maior precisão na detecção de pólen e consistência de rastreamento (IDF1 de 74,29%). Contudo, visando a autonomia da estação de campo na borda, o RT-DETR apresentou vulnerabilidades críticas sob quantização (FP16 e INT8), inviabilizando seu uso embarcado. Em contrapartida, as redes YOLO demonstraram alta resiliência à compressão, consolidando-se como a arquitetura ideal para a inferência 100% em CPU.

Como trabalhos futuros, a inferência será migrada para a borda com modelos YOLO quantizados e aceleradores dedicados (*AI HAT+*), integrando sensores IoT microclimáticos para a construção de modelos comportamentais preditivos. Adicionalmente, técnicas de *data augmentation* serão aplicadas para balancear a classe "Pólen", aprimorando essa detecção. Por fim, para mitigar a limitação da atual validação do rastreamento (MOT), o *ground truth* será expandido mediante anotação de múltiplos fragmentos de 10 segundos em horários-chave (7h, 10h, 12h, 15h, 17h e 18h), atestando a robustez do algoritmo frente às variações diárias de iluminação e densidade de voo.

## Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 e FUNBIO (Chamada de Projetos nº 02/2024). Agradecemos também ao Instituto Terra Brasilis de Desenvolvimento Socioambiental (ITB), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig) pelo apoio ao desenvolvimento deste trabalho.

## Referências

- Bilik, S., Zemcik, T., Kratochvila, L., Rikanek, D., Richter, M., Zambanini, S., and Horak, K. (2022). Machine learning and computer vision techniques in continuous beehive monitoring applications: A survey. *arXiv preprint arXiv:2208.00085*.
- Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
- Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., and Kalinin, A. A. (2020). Albuementations: Fast and flexible image augmentations. *Information*, 11(2):125.

- Dsouza, A. and Hegde, S. (2023). Hivelink, an iot based smart bee hive monitoring system. *arXiv preprint arXiv:2309.12054*.
- Hernández-Cámara, P. et al. (2023). Analysis of deep image quality models. *arXiv preprint arXiv:2302.13345*.
- Jocher, G. and Qiu, J. (2024). Ultralytics yolo11. <https://github.com/ultralytics/ultralytics>. Version 11.0.0.
- Jocher, G. and Qiu, J. (2026). Ultralytics yolo26. <https://github.com/ultralytics/ultralytics>. Version 26.0.0.
- Lei, C., Lu, Y., Xing, Z., Zhang, J., Li, S., Wu, W., and Liu, S. (2024). A honey bee in-and-out counting method based on multiple object tracking algorithm. *Insects*, 15:974.
- Mahamud, M. S. et al. (2019). Mouchak - an iot based smart beekeeping system using mqtt. In *4th International Conference on Robotics and Automation Engineering (ICRAE)*.
- Milan, A., Leal-Taixé, L., Reid, I., Roth, S., and Schindler, K. (2016). Mot16: A benchmark for multi-object tracking. *arXiv preprint arXiv:1603.00831*.
- Nogueira, J. et al. (2014). Conservation study of an endangered stingless bee (*Melipona capixaba*-hymenoptera: Apidae) with restricted distribution in Brazil. *Journal of Insect Conservation*, 18(3):317–326.
- Odemer, R. (2022). Approaches, challenges and recent advances in automated bee counting devices: A review. *Annals of Applied Biology*.
- Silva, D. A., Bomfim, I. G. A., Braga, A. R., and Gomes, D. G. (2023). Applying computer vision models to detect in real time the pollen flow at the input of honeybee hives (*Apis mellifera* L.). In *Anais do Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*.
- Sledevič, T., Serackis, A., Matuzevičius, D., Plonis, D., and Andriukaitis, D. (2024). Keypoint-based bee orientation estimation and ramp detection at the hive entrance for bee behavior identification system. *Agriculture*, 14:1890.
- Sledevič, T., Serackis, A., Matuzevičius, D., Plonis, D., and Vdoviak, G. (2025). Visual recognition of honeybee behavior patterns at the hive entrance. *PLoS ONE*, 20(2):e0318401.
- Soudeep, S. et al. (2024). Dgnn-yolo: Interpretable dynamic graph neural networks with yolo11 for small occluded object detection and tracking. *arXiv preprint arXiv:2411.17251*.
- Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W., and Wang, X. (2022). Bytetrack: Multi-object tracking by associating every detection box. In *European Conference on Computer Vision (ECCV)*.
- Zhao, Y., Lv, W., et al. (2023). Detsr beat yolos on real-time object detection. *arXiv preprint arXiv:2304.08069*.