



III MCSM

Modelagem Computacional no
Sertão Mineiro



PPGMCS
Programa de Pós-Graduação em
Modelagem Computacional e Sistemas

IDENTIFICAÇÃO DE CRIADOUROS DO *Aedes Aegypti* EM IMAGENS AÉREAS UTILIZANDO YOLO 12: APLICAÇÃO EM JANUÁRIA/MG

Mariane Silva Oliveira¹ - mso4@aluno.ifnmg.edu.br

Milena Beatriz Batista da Mota¹ - mbbm@aluno.ifnmg.edu.br

Hélder Seixas Lima¹ - helder.seixas@ifnmg.edu.br

¹Instituto Federal do Norte de Minas Gerais, IFNMG - Januária, MG, Brazil

Resumo. A expansão urbana desordenada e as deficiências no saneamento básico favorecem a proliferação do mosquito *Aedes aegypti*, principal vetor de arboviroses como dengue, zika e chikungunya. A inspeção convencional de potenciais criadouros é uma atividade onerosa e frequentemente limitada por restrições de acesso. Este trabalho utiliza a arquitetura YOLOv12 para a detecção automática de potenciais criadouros do *Aedes aegypti* em imagens aéreas adquiridas por Veículo Aéreo Não Tripulado (VANT) no município de Januária/MG. As imagens foram submetidas às etapas de particionamento, filtragem de fundo e balanceamento de classes antes do treinamento do modelo. O desempenho foi avaliado por meio das métricas precisão, revocação, curva F1–Confiança e mean Average Precision (mAP). O modelo alcançou desempenho global de 0,671 em precisão, 0,541 em revocação e 0,594 em mAP50, com melhores resultados para as classes Pneu (0,856) e Reservatório (0,797). Os resultados também evidenciam maior incidência de erros na classe Descarte, atribuída à elevada variabilidade visual dos objetos agrupados nessa categoria. Conclui-se que a abordagem apresentada demonstra potencial para apoiar ações de controle de arboviroses no município de Januária/MG.

Keywords: *Aedes aegypti*, Criadouros, Visão Computacional, Imagens Aéreas

1. INTRODUÇÃO

As arboviroses dengue, chikungunya e zika representam um problema de saúde pública no Brasil devido ocorrência frequente de surtos. São doenças transmitidas pelo mosquito *Aedes aegypti*, cuja proliferação é favorecida pelo descarte inadequado e pela ocupação desordenada do espaço urbano (Codeco et al., 2018). Recipientes capazes de reter água, como pneus, caixas d'água e piscinas, constituem potenciais criadouros (Scott and Takken, 2012). A dengue é uma arbovirose endêmica em Januária/MG. Com estimados 65.279 habitantes, o município registrou 178 casos em 2023, 7.812 em 2024 e 1.297 em 2025. A chikungunya também apresentou impacto expressivo no município, com 7.433 casos notificados desde 2022 (InfoDengue, 2026).

A identificação e o controle desses potenciais criadouros são fundamentais, dada a gravidade das doenças transmitidas. O método tradicional de investigação é a inspeção por agentes em áreas que podem conter possíveis locais de reprodução. O trabalho é habitualmente demorado, limitado e frequentemente impedido por imóveis com acesso não permitido. Segundo Soares et al. (2025), diante dessas dificuldades, o uso de Veículos Aéreos Não Tripulados (VANTs) tem sido considerado para a detecção desses focos. Aliado a isso, o avanço das técnicas de Redes Neurais Convolucionais (CNN) tem se mostrado uma solução eficiente para a identificação automatizada de criadouros (Sutanto and Ansharullah, 2025; Carrasco-Escobar et al., 2022).

Na literatura, a família de modelos YOLO (*You Only Look Once*) destaca-se pelo desempenho em tarefas de detecção de objetos. Nesse contexto, este trabalho investiga a aplicação da arquitetura YOLOv12 (Tian et al., 2026) para a detecção de potenciais criadouros do *Aedes aegypti* em imagens aéreas adquiridas por VANT no município de Januária/MG. Os resultados apresentados são parciais e integram a pesquisa *Desenvolvimento de Modelo de Visão Computacional para Detecção de Focos do Aedes aegypti Utilizando Imagens de Drone: uma aplicação na cidade de Januária/MG*, em desenvolvimento no IFNMG Campus Januária.

2. METODOLOGIA

2.1 Dados e pré-processamento

Aquisição das imagens. Um conjunto de dados composto por aproximadamente 6.000 imagens aéreas do município de Januária/MG foi fornecido pelo Departamento de Endemias e Vigilância Ambiental da Prefeitura Municipal de Januária. As imagens, com resolução de 8.192×5.460 pixels, foram adquiridas por meio de um VANT modelo DJI Matrice 350 Enterprise, equipado com uma câmera Zenmuse P1. Durante a aquisição, foi utilizada uma base RTK (*Real-Time Kinematic*) para aumento da precisão georreferencial das imagens. Os voos foram realizados a uma altitude de 120 m, cobrindo aproximadamente 600 hectares.

Anotações. Uma amostra de 300 imagens foi anotada usando a ferramenta *Computer Vision Annotation Tool* (CVAT) e exportada no formato YOLO. Cinco classes de potenciais criadouros do *Aedes aegypti* foram anotadas: *Piscina*, *Pneu*, *Lona*, *Reservatório* (caixa d'água destampada, tanque de água e tambor) e *Descarte* (sacolas plásticas, entulhos e sucatas).

Processo de slicing. Devido à alta resolução das imagens, elas foram recortadas em *tiles* de 640×640 pixels. Foi considerado sobreposição de 25% entre *tiles* adjacentes e preservação de anotação quando ao menos 40% de sua área original permanecia visível no *tile*.

Divisão do dataset. A divisão em treino, validação e teste foi realizada nas proporções de 70%, 15% e 15%, respectivamente, após o *slicing*. Para evitar vazamento de dados (*data leakage*), todos os *tiles* originados de uma mesma imagem permaneceram na mesma partição. A alocação utilizou um algoritmo de estratificação ponderado pelo número de *tiles* e pela frequência das classes, priorizando a cobertura de classes minoritárias em todas as partições.

Filtragem de fundo. A maioria dos recortes gerados não continha objetos de interesse, assim, reteu-se uma fração controlada de aproximadamente 10% de amostras de fundo em relação ao *dataset*, preservando a diversidade de cenários de fundo e mantendo a proporção original

entre treino, validação e teste. Essa estratégia visou mitigar falsos positivos em cenários de inferência real.

Balanceamento de classes. Para mitigar o alto desbalanceamento entre as classes, aplicou-se ao conjunto de treino a subamostragem via remoção controlada de imagens contendo apenas classes majoritárias, seguida de superamostragem com a duplicação de imagens de classes minoritárias. A seleção priorizou imagens mistas (contendo mais de uma classe), complementando com imagens puras da própria classe quando necessário para atingir a meta de balanceamento. As cópias foram submetidas a aumentos de dados (*data augmentation*) com espelhamento horizontal e ajustes de brilho/contraste visando diversificar e reduzir o risco de *overfitting*. Limites de segurança foram adotados através do fator máximo de triplicação por imagem-base e remoção de no máximo 50% das imagens puras por classe. Essa estratégia permitiu atingir integralmente a meta de razão máxima de 3:1 entre a classe majoritária e a minoritária no conjunto de treino. Os conjuntos de validação e teste não foram alterados.

Composição final do conjunto de dados. Após todas as etapas de pré-processamento, o conjunto final totalizou 4.573 imagens, distribuídas em 3.158 para treinamento, 694 para validação e 721 para teste. A Tabela 1 representa a quantidade de imagens por classe e partição.

Tabela 1: Distribuição do conjunto de dados por classe em cada partição

Classe	Imagens				Anotações			
	Treino	Validação	Teste	Total	Treino	Validação	Teste	Total
Piscina	388	22	20	430	404	22	22	448
Pneu	601	91	197	889	882	151	345	1.378
Lona	531	100	92	723	596	111	108	815
Reservatório	822	199	182	1.203	1.044	216	207	1.467
Descarte	917	256	249	1.422	1.211	327	321	1.859
Amostra de Fundo	321	81	70	472	-	-	-	-
Total*	3.580	749	810	5.139	4.137	827	1.003	5.967

*Nota: A soma das imagens excede o total real devido à presença de múltiplas classes por imagem.

2.2 Configurações e treinamento

A arquitetura YOLOv12 foi inicializada com pesos pré-treinados com o *dataset* COCO e treinada por 200 épocas com a partição de treino deste trabalho, utilizando *batch size* de 16 e *weight decay* de 0,0005. Os hiperparâmetros relacionados à taxa de aprendizado, ao otimizador e ao *momentum* foram ajustados automaticamente, com agendamento da taxa de aprendizado baseado em decaimento cosseno (*cosine decay*). Nas 15 últimas épocas, foi empregada a estratégia *close mosaic*. Para aumentar a capacidade de generalização, foram aplicadas técnicas de aumento de dados, incluindo *mosaic*, rotação livre e espelhamento vertical. A intensidade da perturbação de brilho no espaço de cores HSV foi reduzida e a técnica *mixup* foi desativada, a fim de evitar a diluição visual dos objetos pequenos presentes nas imagens.

2.3 Métricas de avaliação

O desempenho do modelo foi avaliado por meio das métricas precisão (*precision*), revocação (*recall*) e *mean Average Precision* (mAP), amplamente empregadas na avaliação de modelos

de detecção de objetos. A precisão corresponde à proporção de detecções positivas que são corretas, enquanto a revocação representa a proporção de objetos reais corretamente detectados. A qualidade da localização e da classificação das detecções foi avaliada utilizando as métricas mAP50 e mAP50-95. A métrica mAP50 considera como correta uma detecção cuja *Intersection over Union* (IoU) entre a caixa predita e a caixa de referência seja igual ou superior a 0,5. Já a métrica mAP50-95 calcula a mAP para múltiplos limiares de IoU, variando de 0,50 a 0,95, proporcionando uma avaliação mais rigorosa da capacidade do modelo de localizar e classificar corretamente os objetos. Adicionalmente, utilizou-se a curva F1–Confiança para identificar o limiar de confiança que maximiza o equilíbrio entre precisão e revocação, adotado na construção de uma matriz de confusão como ponto mais representativo do desempenho do modelo.

3. RESULTADOS

A Tabela 2 mostra os resultados obtidos pelo modelo na partição de testes, considerando as métricas de desempenho globais e por classe. Quanto ao resultado global, nota-se que o detector identifica corretamente a maioria das detecções realizadas, embora ainda deixe de detectar parte dos criadouros presentes nas imagens. A diferença entre mAP50 e mAP50-95 evidencia que o modelo apresenta boa capacidade de localização em critérios menos rigorosos de sobreposição, mas sua precisão espacial diminui à medida que o limiar de IoU aumenta.

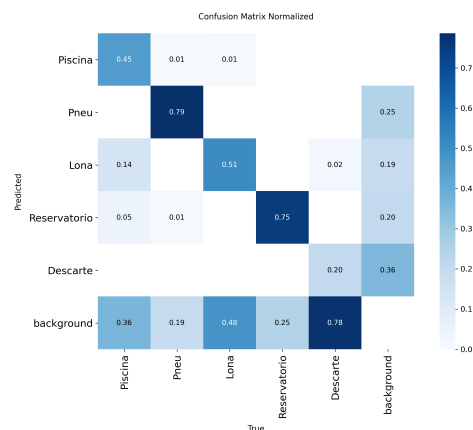
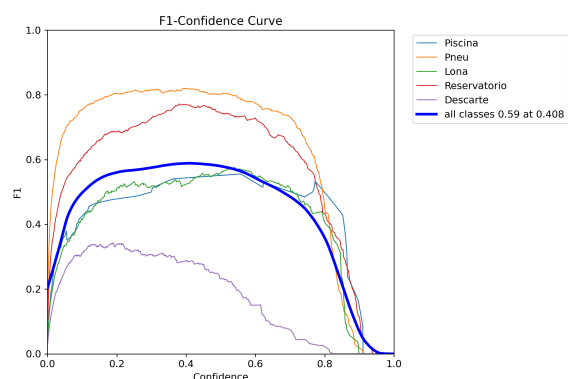
Tabela 2: Resultados de avaliação no conjunto de teste

Classe	<i>Precision</i>	<i>Recall</i>	mAP50	mAP50-95
Todas	0,671	0,541	0,594	0,404
Piscina	0,678	0,455	0,460	0,340
Pneu	0,852	0,789	0,856	0,604
Lona	0,545	0,510	0,578	0,391
Reservatório	0,788	0,750	0,797	0,553
Descarte	0,490	0,201	0,279	0,132

As classes *Pneu* e *Reservatório* apresentaram os melhores resultados, evidenciando maior capacidade de discriminação pelo modelo. As classes *Piscina* e *Lona* alcançaram desempenho intermediário. Em contrapartida, a classe *Descarte* apresentou o menor desempenho, embora tenha sido a mais representativa no conjunto de treinamento. Esse resultado indica que a maior disponibilidade de amostras não foi suficiente para capturar a elevada variabilidade intraclasses dessa categoria, composta por objetos visualmente heterogêneos, como sacolas plásticas, entulhos e sucatas, o que tornou sua detecção mais desafiadora.

A Figura 1a mostra que o limiar de confiança de 0,408 proporcionou o maior valor da métrica F1, indicando o melhor equilíbrio entre precisão e revocação para o modelo avaliado. Com base nesse ponto de operação, a Fig. 1b apresenta a matriz de confusão normalizada, a qual evidencia que a principal fonte de erro está associada à distinção entre as classes de criadouros e o *background*. Esse comportamento é refletido pela ocorrência de falsos positivos, quando regiões de fundo são classificadas como criadouros, e de falsos negativos, quando criadouros reais não são detectados pelo modelo.

A Figura 2 apresenta exemplos de detecções de criadouros de *Aedes aegypti* realizadas pelo modelo em imagens do conjunto de testes.



(a) Curvas F1–Confiança por classe e média global obtidas durante a avaliação do modelo para detecção de criadouros do *Aedes aegypti*. A legenda indica o limiar de confiança de 0,408 correspondente ao valor máximo da métrica F1.

(b) Matriz de confusão normalizada da avaliação do modelo para detecção de criadouros do *Aedes aegypti*, incluindo a classe *background*. A matriz foi construída considerando limiar de confiança de 0,408 e *Intersection over Union* (IoU) de 0,45.

Figura 1: Resultados da avaliação do modelo: (a) curva F1–Confiança e (b) matriz de confusão.

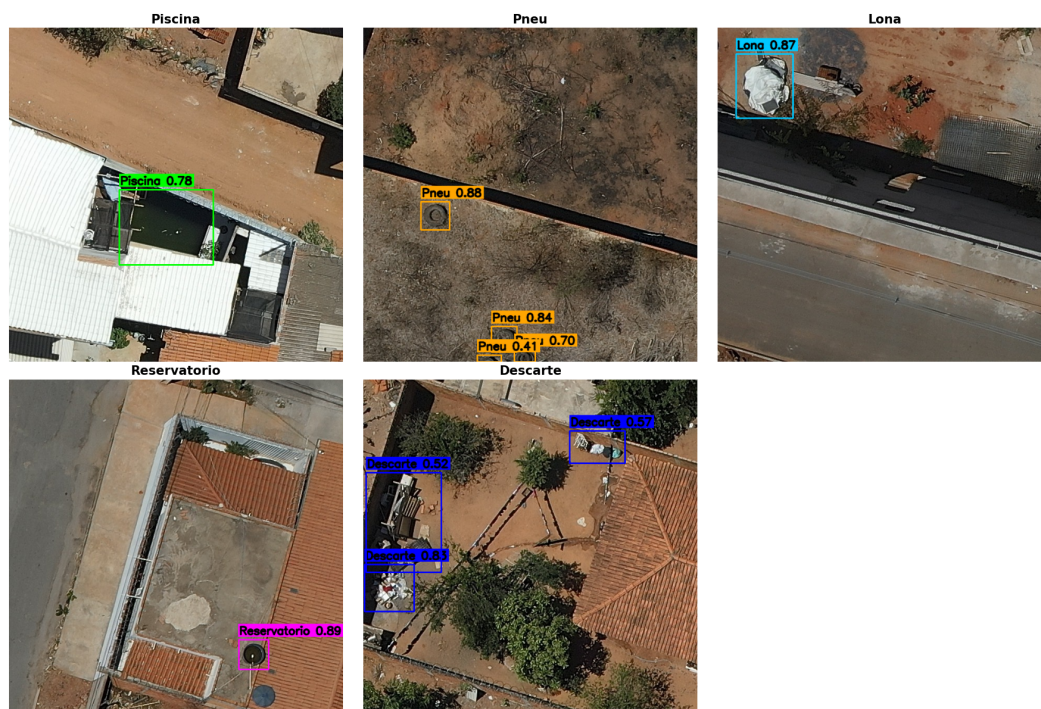


Figura 2: Exemplo de detecção do modelo em imagens de teste.

4. CONCLUSÃO

Este trabalho investigou a aplicação da arquitetura YOLOv12 para a detecção automática de potenciais criadouros do mosquito *Aedes aegypti* em imagens aéreas obtidas no município de Januária/MG. Os resultados obtidos demonstram o potencial da abordagem proposta para apoiar

ações de vigilância ambiental, alcançando desempenho global de 0,594 em mAP50, precisão de 0,671 e revocação de 0,541. A análise por classe evidenciou melhor desempenho na detecção de *Pneu* e *Reservatório*, enquanto a classe *Descarte* apresentou maior dificuldade de identificação, atribuída principalmente à elevada variabilidade visual dos objetos que a compõem. A matriz de confusão também indicou que parte significativa dos erros está associada à distinção entre os criadouros e o *background*, resultando em ocorrências de falsos positivos e falsos negativos.

Como trabalhos futuros, pretende-se empregar técnicas de fotogrametria para geração de ortomosaicos georreferenciados a partir das imagens originais, ampliar o conjunto de dados por meio de novas anotações e técnicas de pseudo-rotulagem, investigar estratégias de adaptação de domínio para aumentar a capacidade de generalização do modelo e explorar métodos de treinamento que priorizem regiões de fundo com elevada complexidade visual. Além disso, pretende-se avaliar a subdivisão da classe *Descarte* em categorias visualmente mais homogêneas, de modo a reduzir a variabilidade intraclasse e melhorar o desempenho da detecção.

Agradecimentos

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo apoio financeiro à realização deste evento (grant n OET-00243-26) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão de bolsa de iniciação científica (Processo nº 174114/2025-7). Agradecem, ainda, ao Instituto Federal do Norte de Minas Gerais (IFNMG), ao Laboratório de Pesquisa em Ciência de Dados e Inteligência Computacional *ciðic*, ao FutureLab da Universidade Federal de Minas Gerais (UFMG) e ao Departamento de Endemias e Vigilância Ambiental da Prefeitura Municipal de Januária pelo apoio institucional e técnico prestado durante o desenvolvimento da pesquisa.

Referências

- Carrasco-Escobar, G., Moreno, M., Fornace, K., Herrera-Varela, M., Manrique, E., and Conn, J. E. (2022). The use of drones for mosquito surveillance and control. *Parasites & vectors*, 15(1):473.
- Codeco, C., Coelho, F., Cruz, O., Oliveira, S., Castro, T., and Bastos, L. (2018). Infodengue: A nowcasting system for the surveillance of arboviruses in brazil. *Revue d'Épidémiologie et de Santé Publique*, 66:S386. European Congress of Epidemiology “Crises, epidemiological transitions and the role of epidemiologists”.
- InfoDengue (2026). Infodengue: Sistema de alerta para arboviroses. Disponível em: <https://info.dengue.mat.br>. Acesso em: 12 jul. 2026.
- Scott, T. W. and Takken, W. (2012). Feeding strategies of anthropophilic mosquitoes result in increased risk of pathogen transmission. *Trends in Parasitology*, 28(3):114–121.
- Soares, H., Pereira Jr., Z. K. G. D., and Guimarães, S. J. F. (2025). Visual identification system for aedes detection using drones. In *38th SIBGRAPI Workshop on Undergraduate Research*, pages 234–238, Salvador, BA.
- Sutanto, H. and Ansharullah, B. A. (2025). The role of artificial intelligence for dengue prevention, control, and management: A technical narrative review. *Acta Tropica*, 268:107741.
- Tian, Y., Ye, Q., and Doermann, D. (2026). Yolov12: Attention-centric real-time object detectors. *Advances in neural information processing systems*, 38:78433–78457.