

# Sistema de Visão Computacional para Identificação Automática de Potenciais Focos do Mosquito *Aedes aegypti* Usando Drones

Gustavo A. Lima<sup>1</sup>, Rafael O. Cotrin<sup>1</sup>, Peterson A. Belan<sup>1</sup>, Sidnei A. de Araújo<sup>1</sup>

[gustavo.araujo.lima94@gmail.com](mailto:gustavo.araujo.lima94@gmail.com); [rafa25.cotrin@uni9.edu.br](mailto:rafa25.cotrin@uni9.edu.br); [belan@uni9.pro.br](mailto:belan@uni9.pro.br); [saraujo@uni9.pro.br](mailto:saraujo@uni9.pro.br)

<sup>1</sup> Universidade Nove de Julho (UNINOVE), Rua Vergueiro, 235/249 – São Paulo/SP – Brasil

DOI: 10.17013/risti.43.93–109

**Resumo:** Os drones tem se tornado uma importante ferramenta tecnológica para auxiliar no combate aos focos de mosquitos. No entanto, as imagens adquiridas por eles são usualmente analisadas de forma manual, podendo consumir muito tempo nas atividades de inspeção. Neste trabalho é proposto um sistema de visão computacional (SVC) para identificação e geolocalização automática de potenciais criadouros do mosquito *Aedes aegypti* a partir de imagens aéreas adquiridas por drones. O SVC desenvolvido deu origem a um software, cujo núcleo é composto por uma rede neural convolucional (RNC) que apresentou taxas de acerto e mAP-50 (*mean average precision*) de 0,9294 e 0,9362 nos experimentos realizados com uma base composta por 500 imagens. Esses resultados, comparados com resultados recentes da literatura, corroboram a adequação da RNC para compor o SVC, o qual pode trazer melhorias para a utilização de drones em programas de prevenção e combate de fontes de reprodução de mosquitos.

**Palavras-chave:** Drone; Mosquito; Reconhecimento de Padrões; Visão Computacional; Redes Neurais Convolucionais.

## *Computer Vision System for Automatic Identification of Potential *Aedes aegypti* Mosquito Breeding Sites Using Drones*

**Abstract:** Drones have become an important technological tool to help fight mosquito breeding sites. However, the images acquired by them are usually analyzed manually, which can consume a lot of time in inspection activities. In this work, a computer vision system (SVC) is proposed for the automatic identification and geolocation of potential breeding sites of the *Aedes aegypti* mosquito from aerial images acquired by drones. The developed SVC gave rise to a software, whose core is composed of a convolutional neural network (CNN) that presented rates of recall and mAP-50 (*mean average precision*) of 0.9294 and 0.9362 in the experiments conducted with a database composed by 500 images. These results, compared with recent results from the literature, corroborate the adequacy of the CNN to compose the SVC, which can bring improvements to the use of drones in programs of prevention and combating mosquito breeding sources.

**Keywords:** Drone; Mosquito; Pattern Recognition; Computer Vision; Convolutional Neural Networks.

## 1. Introdução

As doenças causadas pelo mosquito *Aedes aegypti*, como dengue, chikungunya e zika, vêm preocupando as autoridades da Organização Mundial da Saúde - OMS, segundo a qual, o número de casos notificados de dengue aumentou rapidamente nas últimas décadas (OMS, 2017). Nas Américas, mais de 1,6 milhão de casos foram notificados apenas nos primeiros cinco meses de 2020, a maioria deles no Brasil (OPAS, 2020). Segundo o Ministério da Saúde do Brasil (MSB), em 2020, até novembro, foram registrados 971.136 casos de dengue, 78.808 casos prováveis de chikungunya e 7.006 casos prováveis zika. Ao todo, 528 pessoas morreram de dengue no Brasil em 2020 (MSB, 2020).

Atualmente a OMS considera o controle e eliminação de potenciais criadouros de mosquitos uma medida preventiva imprescindível para enfrentar as doenças causadas por eles (MS, 2020). Porém, as autoridades de saúde muitas vezes veem este desafio ampliado, especialmente em áreas onde uma parte considerável da população não tem acesso a estruturas habitacionais adequadas nem a serviços básicos bem geridos, como água, saneamento e eliminação de resíduos sólidos, que são fatores determinantes para o risco de doenças transmitidas por mosquitos (OMS, 2017; MSB, 2020).

O MSB, com objetivo de combater o mosquito *Aedes aegypti*, lança campanhas publicitárias que envolvem gestores estaduais, municipais e a população. No ano de 2019 o governo federal gastou cerca de R\$22 milhões com campanhas publicitárias de combate ao mosquito *Aedes aegypti*. Em adição, um levantamento feito por uma consultoria de pesquisa estimou que o prejuízo na economia brasileira devido à transmissão de dengue, chikungunya e zika, somente em 2016, ano em que o Brasil teve quase 2 milhões de casos de doenças relacionadas ao *Aedes aegypti*, atingiu R\$2,3 bilhões incluindo custos diretos e indiretos (Dantas, 2021).

As atividades de fiscalização e combate aos potenciais criadouros do mosquito no Brasil é realizada por agentes de saúde, a partir da identificação visual de situações de risco, caracterizadas pela existência de objetos e cenários suspeitos (BRASIL, 2009). Tais atividades normalmente são caras, demoradas, perigosas, além de serem temporal e espacialmente limitadas, resultando na falta de contabilização de grandes porções do espaço urbano (Grubestic et al., 2018). Além disso, é comum os agentes de saúde encontrarem impedimentos para o desempenho de suas atividades, como imóveis fechados, abandonados ou com acesso não permitido pelo proprietário, que afetam as ações de combate à proliferação do mosquito (BRASIL, 2009).

Assim, o combate ao mosquito *Aedes aegypti* tem demandado inúmeros outros esforços, pois nem sempre as atividades de fiscalização e as campanhas de informação e mobilização das pessoas são conduzidos de forma adequada. Neste sentido vários estudos vêm sendo feitos para agilizar a busca de possíveis focos do mosquito, principalmente em regiões com maior incidência de doenças, com o uso de veículos aéreos não tripulados - VANTs, popularmente conhecidos como drones (Passos et al., 2018; Diniz e Medeiros, 2018; Carrasco-Escobar et al., 2019).

Os drones possibilitam voos mais próximos ao solo obtendo imagens com altas resoluções espaciais e temporais, permitindo a detecção de pequenos objetos na superfície da terra e a percepção de mudanças em uma determinada região, em um curto espaço de tempo (Saura et al., 2019; Ponce-Corona et al., 2020; Vintimilla et al., 2020). Além disso, os drones facilitam a inspeção de locais de difícil acesso e requerem menos recursos humanos e financeiros do que as aeronaves tripuladas. Tais vantagens tornam os drones um instrumento poderoso para apoiar intervenções direcionadas à eliminação de criadouros de mosquitos, e assim prevenir doenças como dengue, chikungunya, zika e malária. No entanto, atualmente ainda são poucos os estudos propondo o uso de drones para identificar potenciais criadouros de mosquitos a partir da análise automática de imagens, sendo os principais apresentados na seção a seguir, na qual também estão identificados os principais desafios de pesquisa.

### 1.1. Trabalho Correlatos e Desafios de Pesquisa

Na literatura há estudos que, embora relacionados à temática aqui investigada, não podem ser diretamente comparados com o presente trabalho. Entre eles, podemos citar: Carrasco-Escobar et al. (2019), Haas-Stapleton et al. (2019), Minakshi et al. (2020), Stanton et al. (2020) e Suduwella et al. (2020), que investigaram a identificação de criadouros de mosquitos com base na análise das características de corpos d'água imageados por drones.

Entre os estudos que se assemelham ao presente trabalho, tanto no que tange aos objetos e cenários considerados quanto aos métodos computacionais propostos para a tarefa de identificação automática dos focos estão: Agarwal et al. (2014), Mehra et al. (2016), Dias et al. (2018), Passos et al. (2020), Rossi, Backes e Souza (2020) e Bravo et al. (2021).

Agarwal et al. (2014) apresentaram um método para detectar e visualizar possíveis criadouros de mosquitos, que emprega a técnica *Bag of Visual Words* (BoVW) combinada com o classificador *Support Vector Machine* (SVM) para classificação de imagens. Em experimentos envolvendo a classificação de 500 imagens, eles obtiveram uma acurácia em torno de 82%. Em Mehra et al. (2016), um *framework* foi proposto para detectar possíveis criadouros de mosquitos em imagens do Google e de vários outros dispositivos (câmeras digitais, smartphones e drones). Para a extração das características foi utilizada a técnica BoVW, sendo a classificação realizada por Redes Bayesianas. Nos experimentos conduzidos, os autores obtiveram uma precisão de 90%. Nesses dois trabalhos, as abordagens propostas analisam uma imagem e indicam se ela contém ou não um cenário suspeito, sem fornecer a sua localização na imagem.

Passos et al. (2020) exploraram a detecção de criadouros do mosquito *Aedes aegypti* a partir de vídeos aéreos adquiridos por drones em ambientes simulados. Contudo, em seus experimentos usando diferentes modelos de redes neurais convolucionais (RNC) eles obtiveram uma acurácia máxima de 78%. Embora os autores tenham composto um banco de vídeos para realização dos experimentos, infelizmente ele não foi disponibilizado na literatura.

Dias et al. (2018) propuseram um sistema inteligente para reconhecer e geolocalizar cenários suspeitos de serem criadouros do mosquito *Aedes aegypti* a partir de vídeos capturados por drones. Eles também consideraram a detecção de porções d'água nas

imagens. Segundo os autores, nos testes realizados foi alcançada uma acurácia geral acima de 99%. Contudo, assim como nos trabalhos de Agarwal et al. (2014) e Mehra et. al (2016), apenas um par de coordenadas (latitude e longitude) é associado a cada imagem, independentemente do número de objetos e/ou cenários detectados, dificultando a localização precisa dos mesmos. Em adição, a explicação dos experimentos sobre a detecção de porções d'água não possibilita saber se o algoritmo proposto é capaz de identificar a existência de água em objetos pequenos, que pode ser importante para melhor caracterização da suspeição do mesmo.

Rossi, Backes e Souza (2020) propuseram uma abordagem computacional para detecção automática de calhas a partir de imagens adquiridas por drones, para identificação de criadouros do mosquito *Aedes aegypti*. Além de considerar apenas calhas, desprezando outros objetos e cenários que representam potenciais criadouros, os autores não apresentaram dados quantitativos mensurando os resultados obtidos, bem como não disponibilizaram o banco de imagens coletadas.

Bravo et al. (2021) propuseram abordagens baseadas nas técnicas BoVW e RNC para detecção e localização automática de objetos e cenários suspeitos de serem criadouros do mosquito *Aedes aegypti* em imagens aéreas adquiridas por drones, as quais produziram as taxas de mAP-50 de 0,9651 e 0,9028 na detecção de objetos e cenários usando RNC, e de 0,6453 na detecção de cenários usando BoVW. No entanto, além dos autores terem empregado uma arquitetura específica de RNC para cada tipo de alvo (objetos/cenários), as abordagens propostas não resultaram em um software capaz de gerar mapas com indicações georreferenciadas. Neste sentido, os autores sugerem o desenvolvimento de um software de fácil utilização construído a partir de uma arquitetura única de RNC para detecção de objetos e cenários suspeitos. Não obstante, eles alertam para necessidade do sistema computacional produzir indicações georreferenciadas dos alvos detectados.

Fazendo uma análise dos trabalhos correlatos foram encontradas as seguintes limitações: dificuldade de localização de objetos pequenos nas imagens em virtude da resolução espacial; deficiência na geolocalização dos alvos detectados; falta de um banco de imagens disponível na literatura a partir do qual se possa avaliar e comparar os métodos computacionais propostos; e, principalmente, a falta de um software que possa ser operado por agentes de saúde através de uma interface interativa e de fácil uso, para viabilizar a aplicabilidade prática dos métodos computacionais propostos. Tais limitações constituem os desafios de pesquisa que foram considerados no desenvolvimento do SVC proposto neste trabalho, o qual deu origem a um software cujo núcleo é composto por uma rede neural convolucional (RNC) modelo YOLOv4.

## **2. Fundamentação Teórica**

### **2.1. Visão Computacional**

Visão computacional (VC) pode ser definida como uma subárea do processamento de imagens que estuda o desenvolvimento de métodos e técnicas que possibilitam um sistema computacional interpretar conteúdo de imagens digitais. Em outras palavras, um sistema de visão computacional (SVC) tem como objetivo dotar uma máquina com capacidade para descrever uma cena contida em uma imagem digital (Gonzalez e Woods, 2000).

De acordo Gonzalez e Woods (2000), um SVC típico executa os seguintes passos: i) Aquisição de imagens; ii) Pré-processamento - aplicação de técnicas de VC para aprimorar a qualidade da imagem de entrada, corrigindo iluminação, contraste e distorções, por exemplo; iii) Segmentação - isolar os objetos de interesse do fundo da imagem; iv) Extração de características - extração de vetores de atributos (textura, cor, forma, etc) que caracterizam os objetos de interesse; v) Reconhecimento e interpretação - definição das classes dos objetos segmentados de acordo com suas características.

## 2.2. Redes neurais convolucionais

As Redes neurais convolucionais (RNC), inicialmente propostas por (Le Cun et al., 1998), podem ser descritas como variações de uma rede neural perceptron multicamada, desenvolvida para exigir o menor pré-processamento possível. Isso ocorre porque a RNC tem a capacidade de extrair automaticamente recursos de padrões, uma tarefa que em um método tradicional de reconhecimento de padrões precisa necessariamente ser implementada separadamente, e que representa um dos principais problemas desses métodos (Le Cun et al., 1998). Essa capacidade da RNC é uma das principais vantagens para sua aplicação em tarefas de análise de imagens (Le Cun et al., 1998; Albawi et al., 2017; Schirrmester et al., 2017).

Basicamente, uma RNC consiste em três conjuntos de camadas: camadas convolucionais, camadas de agrupamento e as camadas totalmente conectadas (Le Cun et al., 1998). As camadas convolucionais são responsáveis por extrair as características das imagens. Eles empregam filtros que desencadeiam pequenas regiões ao longo de toda a imagem. As camadas de agrupamento, que são colocadas entre as camadas de convolução, realizam operações de amostragem espacial usando filtros que são aplicados pela imagem. Eles produzem versões de menor resolução das camadas de convolução e ajudam a tornar as representações invariantes às traduções (Goodfellow et al., 2016). As camadas totalmente conectadas funcionam de forma semelhante a uma rede neural perceptron multicamadas e agem como um classificador. Eles recebem entrada da camada anterior e produzem um vetor  $n$ -dimensional, onde  $n$  é o número de classes de saída. Assim, cada elemento vetorial é usado para indicar a probabilidade de que o padrão de entrada pertence a essa classe (Goodfellow et al., 2016; Schirrmester et al., 2017).

Há diversas arquiteturas de RNC, entre os quais está YOLO que tem sido amplamente utilizada na literatura para reconhecimento de objetos (Yi et al., 2019; Benjdira et al., 2019; Xu et al., 2020), e que foi empregada para compor o SVC proposto neste trabalho.

## 2.3. YOLOv4

YOLO é um framework composto de RNCs especialmente projetadas para detecção de objetos. O “YOLO – You Only Look Once”, proposto por (Redmon et al., 2016), tem essa denominação porque se refere ao fato de que a RNC “olha” a imagem apenas uma vez gerando as predições das classes dos objetos detectados. Em 2020 foi lançado o YOLOv4, cuja arquitetura é capaz de reconhecer, em tempo real, 80 objetos diferentes em imagens e vídeos, como demonstrado nos experimentos de Bochkovskiy et al. (2020), os quais pontuam ainda que, apesar de YOLOv4 usar parte da implementação da versão anterior (YOLOv3), seu desempenho é significativamente mais rápido do que YOLOv3 na detecção de objetos.

De acordo com Bohush et al. (2021), YOLOv4 representa o estado da arte na detecção de objetos, com uma das melhores taxas de precisão e velocidade. A abordagem é baseada na RNC Darknet53 que usa conexões *Cross-Stage-Partial-connections* (CSP), que consistem em pirâmides para extração de características. Para atingir altas taxas de velocidade e precisão na detecção, bem como a capacidade de usar o equipamento que estiver disponível para treinamento, YOLOv4 usa vários métodos para aumentar a eficiência sem, no entanto, aumentar o custo computacional. Entre eles, podemos citar: o uso de algoritmos genéticos para otimizar a taxa de aprendizagem, diferentes métodos de aumento do conjunto de dados e *Spatial Pyramid Pooling* (SPP). Este último aplica uma estratégia ligeiramente diferente na detecção de objetos com escalas diferentes. Ele substitui a última camada de pooling (após a última camada convolucional) por uma camada de pooling da pirâmide espacial, como mostrado na Figura 1.

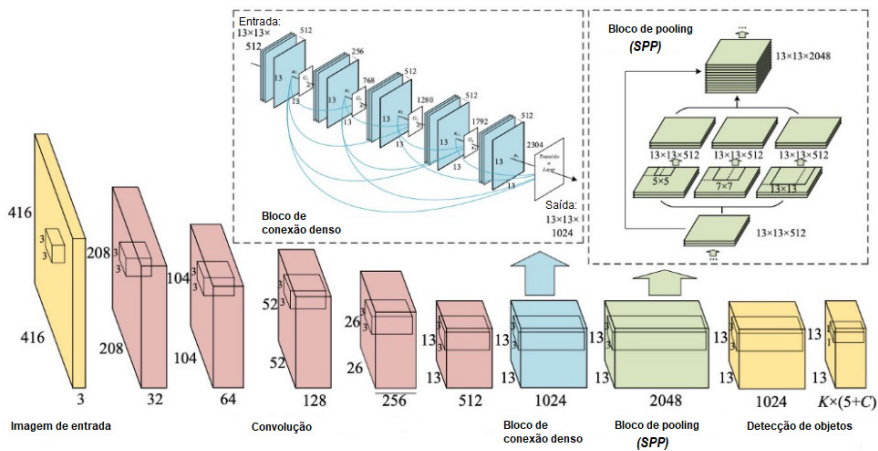


Figura 1 – Arquitetura YOLOv4. Fonte: adaptada de Huang et al. (2020)

Um bloco de conexão densa de YOLOv4 contempla várias camadas de convolução, sendo que cada uma delas produz quatro mapas de características. Ao invés de usar a saída da camada anterior apenas, cada uma das camadas leva em conta a saída de todas as camadas anteriores (Huang et al., 2020).

### 3. Materiais e Métodos

#### 3.1. Base de Imagens

Para a realização dos experimentos foi composta uma base contendo 500 imagens (vide Tabela 1) adquiridas em diferentes áreas da região metropolitana do estado de São Paulo (RMSP) - Brasil, as quais foram escolhidas por serem suscetíveis à proliferação do mosquito Aedes aegypti. Segundo o Ministério da Saúde do Brasil (MSB), são mais suscetíveis os locais onde há crescimento urbano desordenado com maior número de imóveis ocupados, que possuem menor renda per capita, com maior proporção de ruas sem pavimentação, existência de muitos depósitos de materiais de reciclagem e oficinas

mecânicas, com maior quantidade de caixas d’água parcialmente tampadas, lixos, garrafas, pneus e sucata a céu aberto. Além disso, tais regiões são distantes de áreas sensíveis como aeroportos e presídios, como instrui a Agência Nacional de Aviação Civil (ANAC), órgão responsável por regulamentar o registro e o uso de drones no Brasil.

Tipo	Conjunto	Núm. de Imagens	Resolução	Equipamento usado	Altura do solo	Ground sample distance (GSD)
Objetos	CONJ1	140	4000 × 3000	DJI Phantom 3	50	~ 2,2 cm/px
	CONJ2	100	4000 × 3000	DJI Phantom 3	70	~2,45 cm/px
Cenários	CONJ3	150	4000 × 3000	DJI Phantom 4	7, 10 e 13 m	~0,30 a 0,56 cm/px
	CONJ4	110	3000 × 2250	Câmera GoPro HERO4	3 a 5 m	~0,30 cm/px

Tabela 1 – Detalhamento da base de imagens

As imagens do CONJ1 e CONJ2 contemplam diversos tipos de recipientes para armazenamento de água para uso doméstico (caixas d’água) que podem estar cobertos ou descobertos e representam um dos principais criadouros do mosquito, enquanto as imagens do CONJ3 e CONJ4 contêm cenários suspeitos reais e simulados, os quais são caracterizados pela existência de lixo inorgânico a céu aberto, compreendendo objetos pequenos que podem acumular água como pneus velhos, garrafas pet, embalagens plásticas e de papel, entre outros. Alguns exemplos das imagens dos quatro conjuntos são ilustrados na Figura 2.

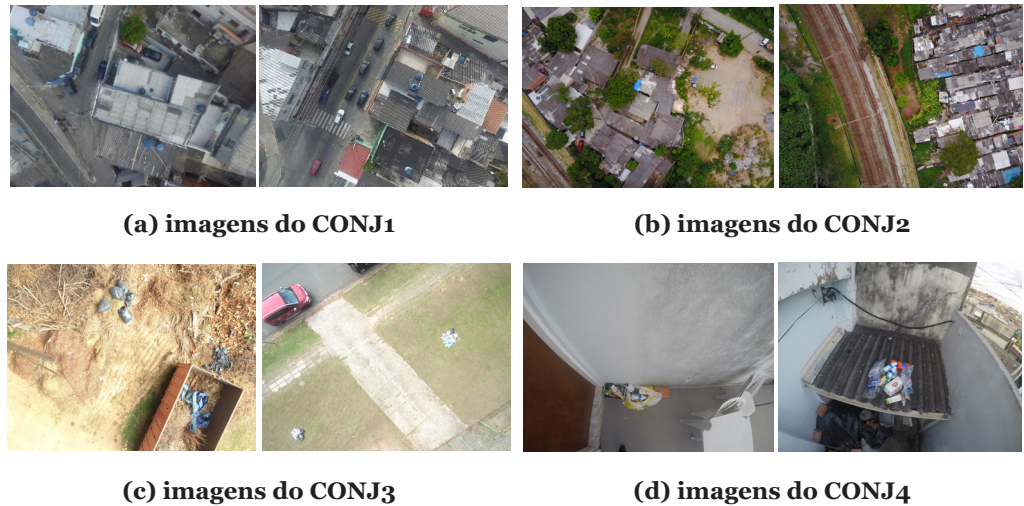


Figura 2 – Exemplos de imagens empregadas nos experimentos conduzidos

As imagens foram divididas em 2 partes: 70% para treinamento e 30% para avaliação da RNC empregada no SVC proposto. É válido ressaltar que, embora não seja uma regra, essas porcentagens são comuns na resolução problemas envolvendo métodos de aprendizagem de máquina e foram adotadas também no estudo de Bravo et al. (2021), o qual foi utilizado nas comparações dos resultados obtidos no presente trabalho. Durante a pesquisa realizada constatou-se que uma das principais dificuldades, também relatada em outros trabalhos encontrados na literatura, como em Passos et al. (2018), foi a aquisição de imagens contemplando os objetos e cenários que representam potenciais criadouros do mosquito. Por isso, a base de imagens composta pode ser considerada uma importante contribuição, uma vez que poderá ser disponibilizada, mediante solicitação, para outros pesquisadores avaliarem seus métodos.

### 3.2. Configuração e Treinamento de YOLOv4

Para treinamento e testes de classificação da RNC YOLOv4, desenvolvida em linguagem Python e denominada RNC\_Obj+Cen, foram utilizadas 168 imagens de objetos e 182 imagens de cenários, totalizando 350 imagens, enquanto as outras 150 imagens foram usadas para testes de classificação. Vale mencionar que durante a fase de treinamento é realizado o aumento de dados usando mecanismos automáticos providos pelo *framework* YOLOv4.

A RNC\_Obj+Cen foi configurada com os seguintes parâmetros: quantidade classes = 2, batches = 124, subdivisões = 16, taxa de aprendizagem = 0.001 e canais de cores considerados = 3 (RGB). Após 22 horas o treinamento foi finalizado com 1900 interações. Na Figura 3 é ilustrado o diagrama de funcionamento do método para mapeamento dos objetos e cenários usando RNC\_Obj+Cen.

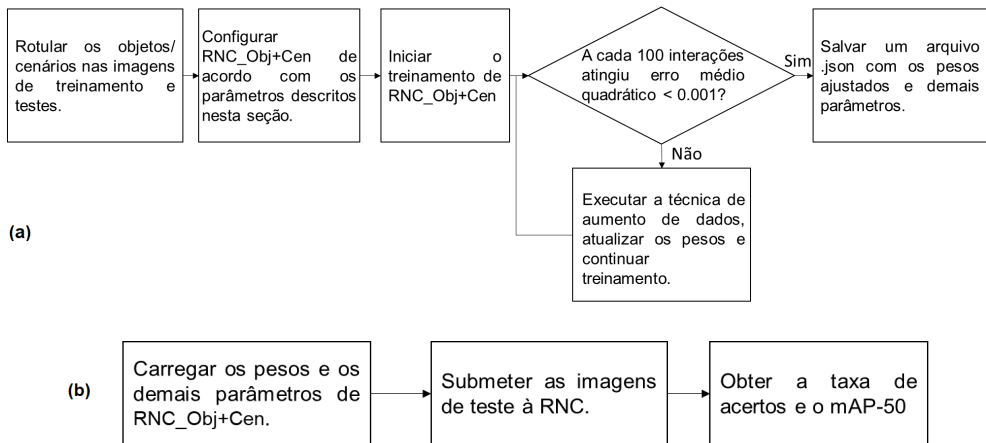


Figura 3 – Diagramas do funcionamento do método para mapeamento de objetos e cenários usando RNC\_Obj+Cen: (a) treinamento; (b) testes com a RNC treinada.

No diagrama detalhado na Figura 3a, que representa o treinamento da RNC, apenas as rotinas de rotulação dos objetos, que consiste em selecionar nas imagens as regiões

contendo objetos/cenários de interesse, e configuração dos parâmetros da RNC foram feitas de forma manual. Após isso, a RNC é responsável por todo o fluxo cujo passo final é salvar o arquivo com os pesos ajustados. Já a Figura 3b se refere aos testes de classificação com a RNC treinada. Para isso, o arquivo contendo os pesos é carregado e, em seguida, as imagens separadas para os testes são submetidas à classificação. Ao final desse processo são obtidas as métricas que indicam a qualidade da classificação para o conjunto de imagens de teste.

Para avaliar o desempenho da RNC foram consideradas as métricas taxas de acerto e mAP-50, as mesmas empregadas no trabalho de Bravo et al. (2021) considerado na análise comparativa, e que são calculadas a partir das quantidades de Verdadeiro Positivo (VP), Falso Positivo (FP), Verdadeiro Negativo (VN) e Falso Negativo (FN) obtidos na classificação.

## 4. Resultados e Discussão

### 4.1. Reconhecimento de Objetos e Cenários com RNC\_Obj+Cen

Para avaliar a precisão da rede RNC\_Obj+Cen, as 150 imagens separadas para os testes foram submetidas à tarefa de classificação, que consumiu aproximadamente 13 minutos. Alguns resultados estão ilustrados na Figura 4.

Das 150 imagens, 78 são de cenários contendo 185 criadouros de mosquito rotulados como *ground truth* por caixa delimitadoras, enquanto as outras 72 são imagens de objetos que apresentam 354 alvos rotulados como *ground truth*. Após a classificação das imagens foi possível obter as taxas de acerto e mAP-50 de 0,9011 e 0,9240 para objetos e de 0,9837 e 0,9485 para cenários.

Das 354 caixas delimitadoras de objetos rotuladas como *ground truth*, 319 foram classificadas corretamente por RNC\_Obj+Cen. Foram identificados 12 casos de FP e 35 casos de FN. Com relação aos cenários, das 185 caixas delimitadoras rotuladas como *ground truth*, 182 foram classificadas corretamente por RNC\_Obj+Cen. Nos experimentos realizados nos testes não houve casos de FP, entretanto houve 3 casos de FN. Figura 4 os casos de FP e FN estão indicados, respectivamente, por círculos vermelhos e círculos brancos.

Por fim, a Tabela 2 apresenta uma comparação dos resultados obtidos neste trabalho com aqueles apresentados em Bravo et. al. (2021), que empregaram duas diferentes arquiteturas oriundas do framework YOLOv3 para detecção de objetos e cenários.



Figura 4 – Exemplos de resultados da identificação de objetos e cenários por RNC\_Obj+Cen

Como pode ser visto na Tabela 2, as taxas médias de acerto e mAP-50 obtidas por RNC\_Obj+Cen, que contempla a detecção de objetos e cenários pela uma mesma arquitetura YOLOv4, superaram as médias das taxas obtidas pelas duas RNCs isoladas desenvolvidas por Bravo et al. (2021), mesmo considerando que o número de objetos e cenários para serem detectados (*ground truth*) foi maior nos experimentos com YOLOv4 realizados neste trabalho. É importante destacar ainda que houve redução significativa no número

de casos de FP. Unificar a detecção de objetos e cenários em uma mesma RNC era uma premissa para viabilizar a implementação do SVC proposto, apresentado na seção a seguir.

	RNC	Tipo de alvo	Qtd. de alvos (ground truth)	VP	FP	FN	Taxa de Acerto	mAP-50
Este trabalho	YOLOv4	Objetos	354	319	12	35	0,9011	0,9240
		Cenários	185	182	0	3	0,9838	0,9485
		<b>Totais</b>	<b>539</b>	<b>501</b>	<b>12</b>	<b>38</b>	<b>0,9294</b>	<b>0,9362</b>
Bravo et al. (2021)	YOLOv3	Objetos	152	140	16	12	0,9211	0,9651
	Tiny-YOLOv3	Cenários	111	96	11	15	0,8649	0,9028
		<b>Totais</b>	<b>263</b>	<b>236</b>	<b>27</b>	<b>37</b>	<b>0,8973</b>	<b>0,9340</b>

Tabela 2 – Sintetização dos resultados obtidos na avaliação das 3 RNCs desenvolvidas

#### 4.2. Sistema de Visão Computacional Desenvolvido

O sistema de visão computacional (SVC) deu origem a um software inteiramente desenvolvido em linguagem Python com uso das bibliotecas OpenCV<sup>1</sup> (composta por rotinas e algoritmos de processamento de imagens e visão computacional), Folium<sup>2</sup> (composta por algoritmos para visualização de dados em mapas), além do framework YOLOv4<sup>3</sup>. No desenvolvimento da interface do software, ilustrada na Figura 5, empregou-se a biblioteca TKinter<sup>4</sup>, a qual possui recursos que facilitam a criação de layouts responsivos. O software, cujo núcleo foi constituído por RNC\_Obj+Cen, possui 5 itens no menu principal, sendo o primeiro (analisar imagens) responsável por identificar (detectar) os objetos e cenários suspeitos em uma ou mais imagens, como mostrado na Figura 5.

O item “galeria de imagens” permite visualizar os objetos e cenários detectados em uma ou mais imagens. Além disso, como mostrado na Figura 6, é possível visualizar a partir desse item informações georreferenciadas sobre as imagens no *Open Street Map*<sup>5</sup> (OSM).

<sup>1</sup> <https://pypi.org/project/opencv-python/>

<sup>2</sup> <http://python-visualization.github.io/folium/>

<sup>3</sup> <https://github.com/AlexeyAB/darknet> 4

<sup>4</sup> <https://docs.python.org/3/library/tkinter.html>

<sup>5</sup> <https://www.openstreetmap.org/>

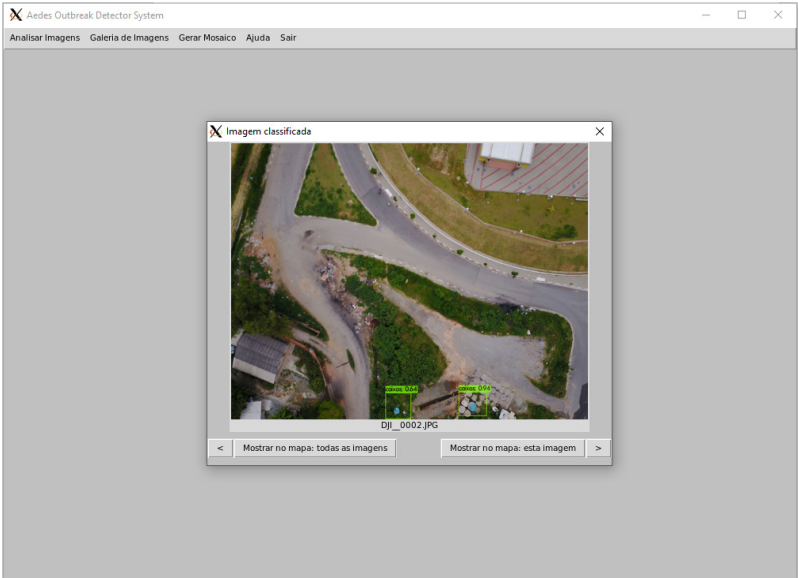


Figura 5 – Interface do software – Análise de imagens

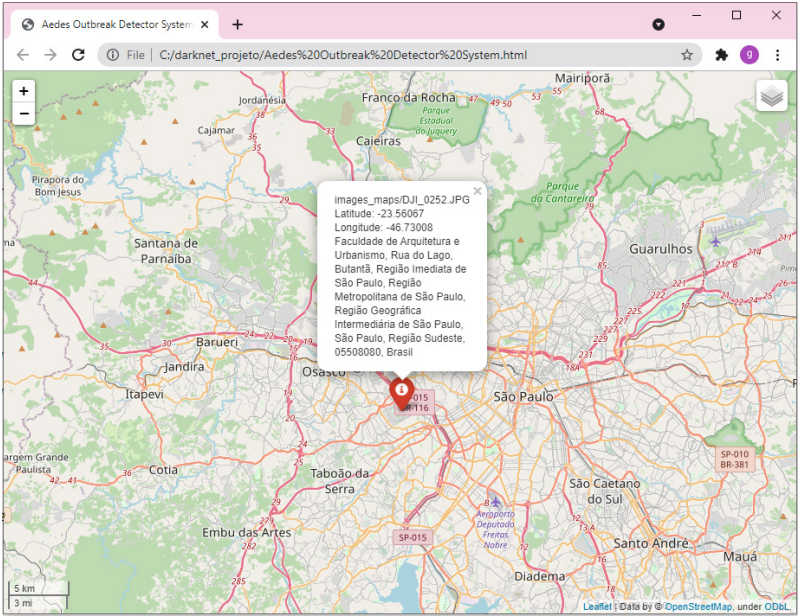


Figura 6 – Interface do software – Visualização de dados no Open Street Map

O item de menu “Gerar mosaico” permite criar um mosaico das imagens adquiridas em um voo (junção de duas ou mais imagens do local imageado). Esse módulo é importante porque permite gerar uma única imagem da área/região onde ocorreu a inspeção, facilitando o processo de análise. Por fim, o item “Ajuda” permite visualizar uma descrição detalhada de todas as funções oferecidas pelo software, enquanto o item “Sair” é utilizado para encerrar o software.

O software proposto neste trabalho, integrando uma única arquitetura de RNC para detecção de objetos e cenários suspeitos (RNC\_Obj+Cen), geração de mosaico de imagens e mapa com indicações georreferenciadas contempla não só as sugestões de (Bravo et al., 2021), mas também de outros trabalhos da literatura, apresentados na seção 1.1. Além do mais, ele poderá auxiliar agentes de saúde no combate ao *Aedes aegypti*, visto que foi projetado para identificar, por meio de uma interface de fácil uso, potenciais criadouros de mosquitos a partir de imagens adquiridas por drones. Obviamente, há ainda melhorias que precisam ser feitas para ampliar a sua aplicabilidade, entre as quais podemos citar: i) propiciar a geração de mosaicos de melhor qualidade (redução de distorções e efeitos de borda); ii) prover indicações georreferenciadas de todos os alvos detectados em uma ou mais imagens; iii) retreinamento de RNC\_Obj+Cen considerando imagens com mais tipos de objetos e cenários; iv) incorporação de um módulo para detecção de água acumulada nos cenários e objetos.

## 5. Conclusões e Trabalhos Futuros

Neste trabalho foi apresentado um SVC para identificação automática de potenciais criadouros do mosquito *Aedes aegypti* com o uso de drones. Para compor o núcleo do software desenvolvido uma RNC YOLOv4 (denominada RNC\_Obj+Cen) foi configurada e avaliada a partir de uma base com 500 imagens, adquiridas em áreas urbanas da Região Metropolitana de São Paulo (RMSP). Os resultados obtidos nos testes de classificação foram: taxa de acerto = 0,9294 e mAP-50 = 0,9362. A investigação da literatura conduzida neste trabalho, somada aos resultados alcançados, permitiu evidenciar a importância do uso de drones para a inspeção de áreas de risco de proliferação de criadouros de mosquitos, especialmente nas regiões periféricas dos grandes centros urbanos onde os serviços essenciais como coleta de lixo, infraestrutura e saneamento básico são muitas vezes escassos ou mesmo inexistentes, visto que tal instrumento de sensoriamento remoto permite a inspeção em áreas de difícil acesso, sem, contudo, colocar em risco a vida dos agentes de saúde responsáveis pelas inspeções de campo. Não obstante, de uma perspectiva transdisciplinar, o uso de drones em tarefas mapeamento automático, como a explorada neste trabalho, poder apoiar a sensibilização da comunidade incentivando sua participação em programas de redução das populações de mosquitos por meio de planos de ação de baixo para cima. Em trabalhos futuros pretende-se desenvolver um método para identificação de água acumulada nos objetos e cenários visando aumentar a aplicabilidade do software desenvolvido, além de outras melhorias incluindo a geração de ortomosaicos e geolocalização com maior precisão de todos os alvos detectados em uma imagem.

## Agradecimentos

Os autores agradecem à FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo pelo suporte financeiro (Processo nº. 2019/05748-0), e ao CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico pela bolsa DT concedida a Sidnei Alves de Araújo (Processo nº. 313765/2019-7).

## Referências

- Agarwal, A., Chaudhuri, U., Chaudhuri, S., Seetharaman, G. (2014). Detection of potential mosquito breeding sites based on community sourced geotagged images. In: Geospatial InfoFusion and Video Analytics IV and Motion Imagery for ISR and Situational Awareness II, p. 90890M. <https://doi.org/10.1117/12.2058121>
- Albawi, S., Mohammed, T. A., Al-Zawi, S. (2017). Understanding of a convolutional neural network. In: International Conference on Engineering and Technology (ICET). Ieee. 1-6. <https://doi.org/10.1109/ICEngTechnol.2017.8308186>
- Benjdira, B., Khursheed, T., Koubaa, A., Ammar, A., Ouni, K. (2019). Car Detection using Unmanned Aerial Vehicles: Comparison between Faster R-CNN and YOLOv3. In: Proceedings of the 1st International Conference on Unmanned Vehicle Systems (UVS), Muscat, Oman, 1-6. <https://doi.org/10.1109/UVS.2019.8658300>
- Bochkovskiy, A., Wang, C-Y., Liao, H-Y. M. (2020). YoloV4: Optimal speed and accuracy of object detection. arXiv preprint.
- Bohush, R., Ablameyko, S., Ihnatsyeva, S., & Adamovskiy, Y. (2021). Object Detection Algorithm for High Resolution Images Based on Convolutional Neural Network and Multiscale Processing. In CMIS (pp. 135-144).
- Brasil - Ministério DA Saúde, Secretaria De Vigilância em Saúde, & Departamento de Vigilância Epidemiológica. (2009). Diretrizes nacionais para prevenção e controle de epidemias de dengue. [https://bvsms.saude.gov.br/bvs/publicacoes/diretrizes\\_nacionais\\_prevencao\\_controle\\_dengue.pdf](https://bvsms.saude.gov.br/bvs/publicacoes/diretrizes_nacionais_prevencao_controle_dengue.pdf).
- Bravo, D. T., Lima, G. A., Alves, W. A. L., Colombo, V. P., Djogbénou, L., Pamboukian, S. V. D., & De Araujo, S. A. (2021). Automatic detection of potential mosquito breeding sites from aerial images acquired by unmanned aerial vehicles. Computers, Environment and Urban Systems, 90, 101692. <https://doi.org/10.1016/j.compenvurbsys.2021.101692>
- Carrasco-Escobar, G., Manrique, E., Ruiz-Cabrejos, J., Saavedra, M.; Alava, F., Bickersmith, S., & Gamboa, D. (2019). High-accuracy detection of malaria vector larval habitats using drone-based multispectral imagery. PLoS neglected tropical diseases, 13(1), e0007105. <https://doi.org/10.1371/journal.pntd.0007105>
- Dantas, C. G. (2021). O portal de notícias da Globo: Consultoria estima que Brasil perdeu R\$ 2,3 bilhões em um ano com doenças do Aedes. <https://g1.globo.com/bemestar/aedes-aegypti/noticia/consultoria-estima-que-brasil-perdeu-r-23-bilhoes-em-um-ano-com-doencas-do-aedes.ghtml>.

- Dias, T. M., Alves, V. C., Alves, H. M., Pinheiro, L. F., Pontes, R. S. G., Araujo, G. M., & Prego, T. M. (2018). Autonomous detection of mosquito-breeding habitats using an unmanned aerial vehicle. American Robotic Symposium, 2018 Brazilian Symposium on Robotics (SBR) and 2018 Workshop on Robotics in Education (WRE) (pp. 351-356). <https://doi.org/10.1109/LARS/SBR/WRE.2018.00070>
- Diniz, M. T. M., Medeiros, J. B. (2018). Mapeamento de focos de reprodução de *Aedes aegypti* na cidade de Caicó/RN com o auxílio de veículo aéreo não tripulado. Revista GeoNordeste, 2, 196–207.
- Gonzalez, R. C., Woods, R. E. (2000). Processamento de imagens digitais. Edgard Blucher.
- Goodfellow, I., Bengio, Y., Courville, A. (2016). Convolutional networks, in Deep Learning. MIT Press. <http://www.deeplearningbook.org>.
- Grubestic, T. H., Wallace, D., Chamberlain, A. W., & Nelson, J. R. (2018). Using unmanned aerial systems (UAS) for remotely sensing physical disorder in neighborhoods. Landscape and Urban Planning, 169, 148-159. <https://doi.org/10.1016/j.landurbplan.2017.09.001>
- Haas-Stapleton, E. J., Barretto, M. C., Castillo, E. B., Clausnitzer, R. J., & Ferdan, R. L. (2019). Assessing Mosquito Breeding Sites and Abundance Using An Unmanned Aircraft. Journal of the American Mosquito Control Association, 35(3), 228-232. <https://doi.org/10.2987/19-6835.1>
- Huang, Z., Wang, J., Fu, X., Yu, T., Guo, Y., & Wang, R. (2020). DC-SPP-YOLO: Dense connection and spatial pyramid pooling based YOLO for object detection. Information Sciences, 522, 241-258. <https://doi.org/10.1016/j.ins.2020.02.067>
- Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324. <https://doi.org/10.1109/5.726791>
- Mehra, M., Bagri, A., Jiang, X., Ortiz, J. (2016). Image analysis for identifying mosquito breeding grounds. International Conference on Communication and Networking (SECON Workshops), 1–6. <https://doi.org/10.1109/SECONW.2016.7746808>
- Minakshi, M., Bhuiyan, T., Kariev, S., Kaddumukasa, M., Loum, D., Stanley, N. B., & Jacob, B. G. (2020). High-accuracy detection of malaria mosquito habitats using drone-based multispectral imagery and Artificial Intelligence (AI) algorithms in an agro-village peri-urban pastureland intervention site (Akonyibedo) in Unyama SubCounty, Gulu District, Northern Uganda. Journal of Public Health and Epidemiology, 12(3), 202-217. <https://doi.org/10.5897/JPHE2020.1213>
- MSB – Ministério da Saúde do Brasil. (2020). Ministério da Saúde lança campanha de combate ao *Aedes aegypti*. <https://antigo.saude.gov.br/saude-de-a-z/combate-ao-aedes>. Acesso em: 4 jun 2021.
- OMS – Organização Mundial de Saúde. (2017). Keeping the vector out: housing improvements for vector control and sustainable development. <https://apps.who.int/iris/handle/10665/259404>. Acesso em: 4 jun 2021.

- OPAS - Organização Pan-Americana de Saúde. (2020). Casos de dengue nas Américas chegam a 1,6 milhão, o que destaca a necessidade do controle de mosquitos durante a pandemia. <https://www.paho.org/pt/noticias/23-6-2020-casos-dengue-nas-americas-chegam-16-milhao-que-destaca-necessidade-do-controle>. Acesso em: 19 fev 2021.
- Passos, W. L., Dias, T. M., Junior, H. M. A., Barros, B. D., Araujo, G. M., Lima, A. A., & Netto, S. L. (2018). Acerca da detecção automática de focos do mosquito *aedes aegypti*. In Brazilian Symposium on Telecommunications and Signal Processing, Campina Grande, Brazil (pp. 392-396).
- Passos, W. L., da Silva, E. A., Netto, S. L., Araujo, G. M., & de Lima, A. A. (2020). Spatio-temporal consistency to detect potential *Aedes aegypti* breeding grounds in aerial video sequences. arXiv preprint.
- Ponce-Corona, E., Sánchez, M. G., Fajardo-Delgado, D., Acevedo-Juárez, B., De-la-Torre, M., Avila-George, H., & Castro, W. (2020). Una revisión sistemática de la literatura enfocada al uso de vehículos aéreos no tripulados durante el proceso de detección de vegetación. RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação, (36), 82-101. <https://doi.org/10.17013/risti.36.82-101>
- Redmon, J., Divvala, S. K., Girshick, R. B., Farhadi, A. (2016). You only look once: Unified, real-time object detection. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition.
- Rossi, L., Backes, A. R., Souza, J. R. (2020). Rain Gutter Detection in Aerial Images for *Aedes aegypti* Mosquito Prevention. In: Anais do XVI Workshop de Visão Computacional, 1-5. <https://doi.org/10.5753/wvc.2020.13474>
- Saura, J. R., Reyes-Menendez, A., & Palos-Sanchez, P. (2019). Mapping multispectral digital images using a cloud computing software: applications from UAV images. Heliyon, 5(2), e01277. <https://doi.org/10.1016/j.heliyon.2019.e01277>
- Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggenberger, K., Tangermann, M., & Ball, T. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. Human brain mapping, 38(11), 5391-5420. <https://doi.org/10.1002/hbm.23730>
- Stanton, M. C., Kalonde, P., Zembere, K., Spaans, R. H., & Jones, C. M. (2021). The application of drones for mosquito larval habitat identification in rural environments: a practical approach for malaria control. Malaria journal, 20(1), 1-17. <https://doi.org/10.1186/s12936-021-03759-2>
- Suduwell, C., Amarasinghe, A., Niroshan, L., Elvitigala, C., De Zoysa, K., & Keppetiyagama, C. (2017). Identifying mosquito breeding sites via drone images. In Proceedings of the 3rd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications (pp. 27-30). <https://doi.org/10.1145/3086439.3086442>
- Vintimilla, A., Madera, P., Otero, N., Salinas, S., Duque, M., & Mullo, C. (2020). Método de estimación de velocidades operacionales en el diseño de UAVs de ala fija para fotografía aérea. RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação, (E29), 283-298.

- Xu, Z. F., Jia, R. S., Sun, H. M., Liu, Q. M., & Cui, Z. (2020). Light-YOLOv3: fast method for detecting green mangoes in complex scenes using picking robots. *Applied Intelligence*, 1-18. <https://doi.org/10.1007/s10489-020-01818-w>
- Yi, Z., Yongliang, S., & Jun, Z. (2019). An improved tiny-yolov3 pedestrian detection algorithm. *Optik–International Journal for Light and Electron Optics*, 183, 17–23. <https://doi.org/10.1016/j.ijleo.2019.02.038>