

# CONTAGEM AUTOMATIZADA DE OVOS DE *Aedes Aegypti* EM PAPÉIS-FILTRO UTILIZANDO APRENDIZADO PROFUNDO

## *AUTOMATED COUNTING OF Aedes Aegypti* EGGS ON FILTER PAPERS USING DEEP LEARNING

João Marcelo Almeida de Macedo<sup>1</sup> 

Vinicius de Carvalho Rispoli<sup>2</sup> 

Jonas Lotufo Brant<sup>3</sup> 

**Resumo:** O mosquito *Aedes aegypti* é o principal transmissor da dengue e de outras arboviroses que afetam a saúde de milhares de brasileiros todos os anos. Entre as técnicas utilizadas para monitorar a infestação, destaca-se a contagem manual de ovos em papéis-filtro e ovitrampas — um procedimento demorado, sujeito a falhas humanas e que ainda carece de soluções modernas aplicadas de forma consistente. Este trabalho apresenta uma proposta de automação desse processo por meio do uso de redes neurais convolucionais, especificamente a arquitetura Mask R-CNN, implementada em *Python*. A metodologia adotada realiza a segmentação e a contagem automática dos ovos em imagens coletadas, alcançando métricas geralmente superiores às obtidas em estudos semelhantes que utilizam bases de dados menores (com menos de 100 amostras). Além disso, a pesquisa conta com um conjunto de dados autorais em papéis-filtro utilizados regularmente em pesquisas no Brasil. Assim, o estudo contribui para a aplicação do aprendizado profundo no enfrentamento da espécie *Aedes aegypti*, oferecendo uma alternativa inovadora para aprimorar a vigilância das arboviroses e direcionar os esforços de vigilância e controle de forma mais eficiente e econômica. Como perspectivas futuras, destaca-se a criação de um software acessível que incorpore a rede desenvolvida, além da ampliação do conjunto de dados de treinamento.

**Palavras-chave:** *Aedes Aegypti*. Aprendizado de Máquina. Contagem Automatizada. Redes Neurais.

**Abstract:** The *Aedes aegypti* mosquito is the main transmitter of dengue fever and other arboviruses that affect the health of thousands of Brazilians every year. Among the techniques used to monitor infestation, manual counting of eggs on filter paper and ovitraps stands out — a time-consuming procedure, subject to human error, and still lacking modern solutions applied consistently. This work presents a proposal for automating this process using convolutional neural networks, specifically the Mask R-CNN architecture, implemented in *Python*. The methodology adopted performs segmentation and automatic counting of eggs in collected images, achieving metrics that are generally superior to those obtained in similar studies using smaller databases (with less than 100 samples). In addition, the research relies on an authorial

---

<sup>1</sup>Mestrando em engenharia biomédica, Universidade de Brasília, 242107098@aluno.unb.br.

<sup>2</sup>Doutor em engenharia biomédica, Universidade de Brasília, vrispoli@unb.br

<sup>3</sup>Doutor em saúde coletiva, Universidade de Brasília, jonas.brant@unb.br

dataset of filter paper regularly used in research in Brazil. Thus, the study contributes to the application of deep learning in combating *Aedes aegypti*, offering an innovative alternative to improve arbovirus surveillance and direct control efforts more efficiently and economically. Prospects include the creation of accessible software that incorporates the developed network, as well as the expansion of the training data set.

**Keywords:** Aedes Aegypti. Machine Learning. Automated Counting. Neural Networks.

# 1 INTRODUÇÃO

*Aedes aegypti* é um mosquito vetor de inúmeras doenças, como a dengue, zika e chikungunya. É originário da África e de difícil e custoso controle populacional no ecossistema urbano moderno. Facilmente reconhecido pela coloração preta e branca em suas patas, este mosquito possui um ciclo de vida bem descrito e estudado pela literatura, e muitas tentativas de erradicá-lo foram realizadas no Brasil (SOPER, 1965).

A fêmea *Aedes aegypti* deposita seus ovos em recipientes naturais ou artificiais de água. Os ovos permanecem na borda dos recipientes, acima do nível da água. Eles podem sobreviver sem água por até oito meses e precisam de um baixo volume de água para eclodir. Quando ocorre uma chuva e o nível da água aumenta, os ovos eclodem e as larvas vivem na água. Após cerca de cinco dias, as larvas se desenvolvem em pupas que, por sua vez, se desenvolvem nos mosquitos adultos em cerca de dois a três dias (CDC, 2024).

Em seu ciclo de vida, o mosquito *Aedes aegypti* tende a voar não mais do que 100 a 200 metros de distância (CDC, 2024). Isso significa que o controle do mosquito implica a concentração de esforços em áreas de maior densidade de ovos, onde gerações de mosquito podem ter a oportunidade de se proliferar. Uma das formas de monitoramento da população da espécie *Aedes aegypti* mais utilizadas em virtude de seu baixo custo e simplicidade é a utilização de ovitrampas, armadilhas para a captura de ovos que consiste em uma palheta de madeira parcialmente imersa em um recipiente com água. Sua superfície rugosa é favorecida pela fêmea do mosquito, que então deposita seus ovos ali. É feita então a coleta da palheta e os ovos são contados manualmente pelo agente de saúde.

Com o resultado, é possível estimar a quantidade de fêmeas do mosquito presentes na região e então direcionar os esforços para os ambientes mais críticos em densidade de mosquitos. Porém, como a contagem é feita de forma manual, esta exige tempo e recursos humanos especializados. O Brasil, de forma generalizada, ainda carece de tecnologias do estado da arte para a contagem dos ovos de forma automatizada.

Tendo em vista tal contexto, é evidente a falta de trabalhos nesse tema que não envolvam imagens com alta amplificação ou montagem complexa, visto que se espera melhores métricas de objetos grandes comparados ao tamanho da imagem. Porém, há situações em que uma montagem complexa ou imagens amplificadas não são facilmente implementadas e, por isso, há uma necessidade de estudos que visem entregar métricas similares as que utilizam esses recursos. Isso possibilitaria a coleta de dados pela população geral, o que implica em um volume de dados muito maior para entidades que visam combater e prevenir epidemias.

O propósito deste trabalho é desenvolver um *software* que utilize redes neurais do estado da arte para realizar a contagem automatizada de ovos do mosquito *Aedes aegypti* evitando o uso de qualquer dispositivo que necessite de montagem como proposto por (SILVA, 2021). Tal contagem será feita em papéis-filtro. Dessa forma, planeja-se em trabalhos futuros incorporar tal rede neural em um aplicativo de celular para auxiliar profissionais de saúde na contagem de ovos e redução do esforço hoje direcionado a essa atividade.

Dessa forma, a inovação pretendida com esse trabalho envolve a automação de um processo que hoje exige tempo e recursos que poderiam ser direcionados para outras áreas de maior prioridade, bem como uma participação ativa da comunidade graças à ausência da montagem de qualquer tipo de equipamentos, o que geraria potencial para maior participação popular neste tipo de vigilância.

## **2 DESENVOLVIMENTO**

### **2.1 Revisão Bibliográfica**

Uma das maneiras de automatizar o processo de contagem é através do uso de redes neurais artificiais, em particular, redes neurais convolucionais. Há diversas topologias de tais redes descritas na literatura e para as mais diversas aplicações. As redes neurais convolucionais baseadas em região (do inglês, *region-based neural convolutional network*), também conhecidas como R-

CNNs, são arquiteturas que inicialmente propõem regiões que possam conter objetos de interesse e então extrai as características de cada região com uma CNN (GIRSHICK et al., 2013). Entretanto, seu lento processo de treinamento o torna proibitivo a depender da situação. Nesse contexto, a arquitetura Rede Neural Convolutiva Rápida Baseada em Região (do inglês, *fast region-based neural convolutional network*) ou Fast R-CNN utiliza técnicas de *pooling* globais para evitar o processamento de cada região com uma CNN individualmente (GIRSHICK, 2015). Outra evolução desta técnica é a Mais Rápida Rede Neural Convolutiva Baseada em Região (do inglês, *faster region-based neural convolutional network*) ou Faster-RCNN, que cria um tipo de rede capaz de realizar a proposição de regiões na GPU e compartilhar convoluções entre regiões (REN et al., 2015).

A última evolução principal na cadeia de redes baseadas em regiões, e que será utilizada neste trabalho, é a chamada Mask R-CNN (HE et al., 2017), que ramifica a Faster-RCNN calculando uma máscara de segmentação para cada objeto em paralelo com a classificação e detecção de cada objeto, sendo amplamente usada para casos de segmentação em instância. Como essa segmentação consegue lidar com a sobreposição de objetos, ela será uma base útil para a contagem de ovos automatizada.

Soluções no Brasil para a contagem automatizada raramente incluem redes neurais do estado da arte (BRUN et al., 2020) e nenhuma literatura foi encontrada utilizando a Mask R-CNN ou variações. Existem trabalhos que utilizam redes neurais para a detecção de objetos como a YOLOv4 (SILVA, 2021) com 93% de acurácia em testes de imagem, porém exigindo alguma competência para a construção/montagem de equipamentos de baixo custo para a aquisição das imagens das ovitrampas, o que pode ser inviável de forma generalizada no contexto nacional onde a contagem dos ovos é necessária.

No exterior, por outro lado, algumas tentativas de solução utilizando redes neurais foram feitas para este problema. Um software australiano foi desenvolvido chamado EggCountAI (JAVED et al., 2023) utilizando a mesma arquitetura, mas com uma ResNet101 como arquitetura base para resolver o mesmo problema de contagem de ovos, porém em fitas amareladas ao invés

das ovitrampas e papéis-filtro utilizados no Brasil. Outro software com a mesma intenção é o chamado ICount (GABURRO et al., 2016), que utiliza visão computacional para contar os ovos. É inconsistente quando os ovos estão se sobrepondo, mas possui altas métricas de avaliação no geral se comparado ao resto da literatura.

Também há estudos implementando arquiteturas similares, como no trabalho de Wang (WANG et al., 2024), onde uma Faster R-CNN é utilizada em conjunto de técnicas avançadas de segmentação para a construção de uma rede identificadora de ovos do mosquito *Aedes albopictus*, uma espécie próxima do mosquito *Aedes aegypti*. O conjunto de dados conta com 1729 imagens de microscópio e obtém uma métrica F1 de 0,977. Já no trabalho de Naranjo-Alcazar (NARANJO-ALCAZAR et al., 2024), foram utilizadas duas arquiteturas, Mask R-CNN e Cascade Mask R-CNN, e a precisão média de ambas foi calculada para o caso de contagem de ovos em paletas de madeira. O alvo também foi os ovos do mosquito *Aedes albopictus* e a aquisição do conjunto de dados envolveu uma montagem complexa de um microscópio que é automaticamente movimentado para capturar imagens aproximadas da paleta por toda a sua extensão. O estudo reportou uma precisão média de 0,91 e 0,85 para a Mask R-CNN e a Cascade Mask R-CNN, respectivamente, para uma métrica intersecção sobre união (IoU) de 0,5.

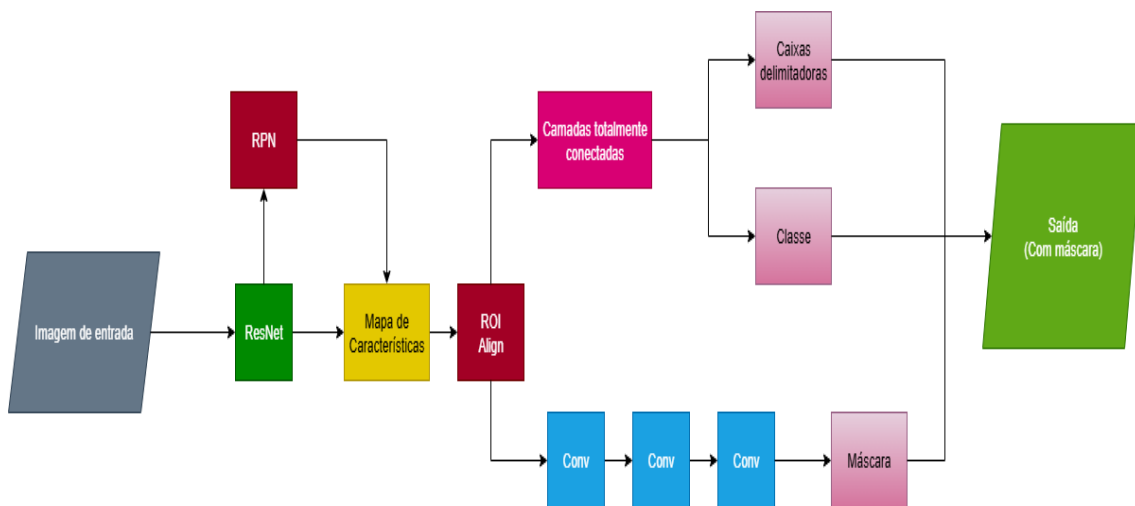
## 2.2 Mask R-CNN

A *Mask R-CNN* é um modelo de aprendizado profundo para segmentação de imagens em instâncias, estendendo a *Faster R-CNN* ao adicionar uma cabeça de segmentação que gera máscaras binárias para cada objeto detectado (HE et al., 2017). Ela opera em duas etapas principais: primeiro, uma RPN (do inglês, *region proposal network*) propõe regiões de interesse na imagem; depois, uma rede convolucional extrai características dessas regiões e as classifica, refina as caixas delimitadoras e gera máscaras pixel a pixel para cada objeto identificado.

A Figura 1 exibe a cadeia de processos internos da *Mask R-CNN*. A

rede divide a imagem em múltiplas potenciais regiões ao extrair seu mapa de características, usando uma ResNet50 como sua arquitetura base. A divisão de regiões não passa por nenhuma quantização graças à técnica *RoIAlign*, que utiliza interpolação bilinear para evitar a perda de informação da técnica mais clássica para a escolha de regiões, conhecida como *RoIPool*. Com as regiões, a *Mask R-CNN* atua em paralelo na definição das caixas delimitadoras, classes e máscaras, a principal diferença dessa arquitetura para a sua predecessora *Faster R-CNN* (HE et al., 2017).

**Figura 1** – esquema da arquitetura *Mask R-CNN*. A imagem original é inicialmente processada pela rede de proposição de regiões a fim de obter seções onde podem se encontrar objetos. O mapa de características é formado e, após o ROI Align, ocorre um processamento simultâneo para classificar máscara, classes e caixas, que irão compor a imagem final.



Fonte: dos autores (2025).

A atuação em paralelo da classe e máscara são apontados pelo artigo original como sendo essenciais para uma boa predição da máscara. A função de perda é então dada pela forma

$$L = L_{cl} + L_{box} + L_{mask},$$

onde  $L_{cl}$  é a perda associada à classe,  $L_{box}$  é a perda associada à caixa delimitadora e  $L_{mask}$  é a perda associada à máscara, todas definidas em

(GIRSHICK, 2015). Em adição à arquitetura do artigo original, (LI et al., 2021) utiliza normalização em lotes após as convoluções necessárias da rede de extração de características, quatro camadas convolucionais seguidas de normalização de lote e uma camada linear durante a determinação das regiões de interesse, e por último, duas camadas convolucionais na rede de proposição de regiões. Todas estas melhorias foram adicionadas ao código desse trabalho, mas optou-se por não explorar os transformadores como arquitetura base da rede, como proposto em (LI et al., 2021) pela complexidade da rede e tempo de treinamento, visto que o treino acontecerá em uma única GPU.

## 2.3 Métodos

A implementação da rede neural deste trabalho é feita em *Python* utilizando a biblioteca *Pytorch* (PAZSKE et al., 2019) como base. O treinamento foi feito localmente em um computador com processador Intel® Core™ i5 14600K, placa de vídeo NVIDIA RTX 4060 Ti 8 GB, memória RAM Corsair 32 GB 4800 MHz e placa-mãe Gigabyte Z790 UD AC. Pela excelente capacidade de generalização da arquitetura, a arquitetura base utilizado para a extração das características será uma ResNet50.

Acerca da obtenção do conjunto de dados, foi implementado o projeto das Ovitrapas numa parceria com a ASSBIO e o GERK-58<sup>a</sup> em um condomínio localizado na região dos lagos em Sobradinho, no Distrito Federal. As ovitrapas, como previamente mencionado, são armadilhas para capturas de ovos de mosquito para identificação de áreas de maior infestação, a fim de ajudar no combate a transmissão de doenças, principalmente da dengue. O objetivo do projeto é realizar o mapeamento e monitoramento da região atingindo resultados satisfatórios com o objetivo de definir melhor as áreas prioritárias para realização de ações a fim de controlar e prevenir a disseminação de doenças transmitidas pelo *Aedes aegypti* na comunidade. A proposta inicial era instalar 16 armadilhas, distribuídas no condomínio em raios de 200 metros, e a meta era selecionar imóveis para a instalação das armadilhas, onde os moradores colaboradores ajudariam a equipe a fazer o

monitoramento. A equipe CEA superou bastante as expectativas neste projeto, pois o monitoramento e implantação foi feito com sucesso e permanece até o presente momento no ano de 2025, tendo 36 armadilhas instaladas em áreas comuns e 6 armadilhas em casas de voluntários. O projeto vem superando expectativas no sentido do mapeamento e monitoramento da região do condomínio, atingindo resultados satisfatórios no campo de melhor definir as áreas prioritárias para realização de ações a fim de controlar e prevenir a disseminação de doenças transmitidas pelo *Aedes aegypti* na comunidade.

O conjunto de dados foi elaborado pelo autor com papéis-filtro contendo ovos do mosquito *Aedes aegypti* fornecidos pela Sala de Situação de Saúde da Faculdade de Ciências da Saúde da Universidade de Brasília (FS/UnB), provenientes do projeto das Ovitrampas. A Figura 2 mostra o equipamento utilizado para a captura dos ovos, onde o papel-filtro se encontra enrolado no interior da garrafa e os mosquitos depositam seus ovos sobre ela. Os papéis-filtro serão escaneados com 1200 dpi a partir de uma mesa digitalizadora comercial, vista a necessidade de manter a aquisição livre de qualquer montagem manual, uma abordagem diferente da observada em outras soluções no Brasil. Cada papel-filtro, em função do seu formato retangular, foi capaz de gerar quatro imagens para o conjunto de dados de diferentes tamanhos para uma melhor generalização da rede neural. O *software* utilizado para a anotação das imagens será o *labelme*, pela sua fácil utilização.

Foram realizados testes utilizando a biblioteca *MMdet* do *Python*, que possui versões da arquitetura Mask R-CNN prontas para o rápido treinamento utilizando um arquivo de configuração para que o usuário controle os hiperparâmetros da rede. Porém, o melhor resultado foi obtido utilizando a arquitetura já incluída na biblioteca *Pytorch*, que conta com algumas melhorias não incluídas no artigo original.

**Figura 2** – Armadilha utilizada para a coleta de ovos. O papel-filtro é colocado dentro da garrafa com água de modo a cobrir o seu raio interno. Os mosquitos depositarão seus ovos sobre ela, que serão aderidos ao papel-filtro para a coleta posterior do material.



**Fonte:** dos autores (2025).

No presente momento, o conjunto de dados conta com 200 imagens, onde 90% foram dedicadas para treino e o resto para validação. Os primeiros testes foram realizados com apenas 70 e, conforme mais imagens são obtidas, a rede é treinada e testada novamente. Embora seja um conjunto pequeno para o treinamento de uma rede neural, espera-se que o contraste entre o ovo e o fundo, aliado ao fato de que há apenas uma classe de interesse, faça com que a rede possa aprender rápido o suficiente para ter métricas de validação inferiores, mas dentro de 15% das métricas da literatura. Dada a heterogeneidade do conjunto, também não se espera que aconteça sobreajuste.

A arquitetura da rede é inspirada em (LI et al., 2021), com o número de épocas igual a 60 e o limite mínimo de confiança limiar igual a 85%. A taxa de aprendizado ótima após testes foi de  $5 \times 10^{-4}$ . O tamanho do lote foi de 2 imagens, visto que o treinamento é em apenas uma única placa de vídeo.

A Figura 3 ilustra um exemplo das imagens do conjunto de dados. A Figura 3a representa uma das entradas brutas do conjunto de dados, enquanto a Figura 3b exibe o resultado da anotação. O ruído presente na imagem deve ser filtrado pela rede neural no processo de treinamento, dado um conjunto de

imagens suficiente. Espera-se que a predição da rede neural se assemelhe ao resultado da imagem anotada.

**Figura 3** – Exemplo de uma imagem do conjunto de dados com e sem a anotação dos ovos.



(a) Imagem bruta do conjunto de dados (b) Imagem anotada através do software labelme.

Fonte: dos autores (2024).

Quanto às métricas de avaliação, propõe-se o uso do índice Jaccard (IoU), visto seu amplo uso para avaliação de questões de segmentação. O índice IoU é dado pela expressão

$$IoU = \frac{A \cap B}{A \cup B},$$

onde A representa a máscara de segmentação verdadeira e B representa a máscara predita pela rede neural. Em problemas de classificação binária, onde verdadeiros positivos (VP) e negativos (VN) significam, respectivamente, que a classe foi corretamente predita em sua presença e que a classe foi corretamente predita em sua ausência, e falsos positivos (FP) e negativos (FN) representam, respectivamente, que a classe foi incorretamente detectada quando não existia ali e que a classe foi incorretamente não detectada quando existia ali (SILVA et al., 2021).

Métricas como a acurácia, que nada mais é do que um percentual de acertos, definida como

$$Acc = \frac{VP+VN}{VP+VN+FP+FN},$$

é uma métrica simples para avaliar quanto a rede neural conseguiu acertar, mas que por si só não pode dizer o quão robusta ou poderosa a solução é. Outra métrica que costuma acompanhá-la é a chamada precisão, definida como

$$Pr = \frac{VP}{VP+FP},$$

e que analisa o percentual de acertos positivos tendo em vista o total de pixels preditos como positivos. Outra importante métrica é a chamada revocação (do inglês *recall*), que avalia a taxa de acertos positivos comparada com o total de *pixels* de fato positivos e é definida como

$$R = \frac{VP}{VP+FN}.$$

Para o problema em específico, visando a contagem de ovos, o erro absoluto de contagem será a mais importante métrica e será definido como  $Erro = |\theta - \hat{\theta}|$ , sendo  $\theta$  o número de ovos reais contidos na imagem e  $\hat{\theta}$  a predição da rede neural. Por último, o teste F1 é bastante usado para avaliar testes preditivos e é definido pela razão

$$F1 = \frac{2 \times Pr \times R}{Pr + R},$$

sendo então a média harmônica entre a revocação e a precisão.

## 2.4 Resultados

A Tabela 1 representa os resultados da análise de todas as imagens de validação durante o primeiro teste, onde havia apenas 70 imagens dedicadas para treino. A métrica de acurácia nos dá pouca informação sobre as predições da rede pois os ovos tomam um espaço muito pequeno da imagem. As métricas de precisão, *recall* e o índice F1 são mais adequadas para a avaliação. As imagens de validação eram compostas de no máximo 23 ovos.

A métrica de IoU apresenta-se como menor do que o esperado. Segundo testes posteriores, isso se dá pelo fato de que, com o baixo número de imagens, as predições da rede são conservadoras no geral. Isso significa que a máscara gerada pela rede, apesar de correta, foca-se mais em discernir o centro do objeto do que em delimitar toda a área do ovo. Dada a finalidade

principal de contar o total de ovos, isso não se torna um problema, mas espera-se que um aumento de imagens no conjunto de dados aumente a métrica IoU.

**Quadro 1** – Métricas das predições da rede neural nas imagens de validação (pixel por pixel) realizadas com 70 imagens de treino.

Imagens de validação	IoU	Acurácia	Precisão	Recall	F1	Erro absoluto de contagem	Total de ovos reais
1	0,6616	0,9998	1,000	0,6616	0,7963	0	6
2	0,7370	0,9992	0,8657	0,8332	0,8492	2	23
3	0,4895	0,9998	0,7954	0,5600	0,6572	1	5
4	0,6532	0,9998	0,8350	0,7500	0,7902	0	13
5	0,7993	0,9997	0,9494	0,8349	0,8885	0	15
6	0,6352	0,9999	0,8571	0,7105	0,7769	0	3
7	0,7406	0,9996	0,9040	0,8038	0,8509	0	18
Média	0,6732	0,9996	0,8866	0,7360	0,8011	0,428	11,25

**Fonte:** dos autores (2024).

Após a obtenção de mais imagens no ano de 2025, foram calculadas as métricas baseadas na contagem de ovos com acurácia espacial. O que significa que além das medidas pixel por pixel, há também medidas baseadas na contagem correta de ovos considerando uma sobreposição com as caixas delimitadoras (IoU) de pelo menos 0,3. Tal valor de confiança é geralmente abaixo dos usos gerais de redes neurais, porém uma particularidade deste trabalho é que a precisão da caixa delimitadora não necessita de ser alta, o que basta é que a rede encontre um ovo naquela região, visto que a contagem é a parte mais importante para a vigilância epidemiológica. A Tabela 2 contém tais valores. O conjunto utilizado aqui envolveu 19 imagens não antes vistas pela rede neural, fornecidas pela FS/UnB.

Tais imagens são ligeiramente diferentes das usadas no treino e nota-se métricas ainda elevadas se comparadas às métricas pixel a pixel, o que indica uma constante melhora da rede conforme o conjunto de dados se expande. Ainda que o contraste entre o fundo branco e o ovo preto ajude a rede em identificá-los, a alta presença de ruído nas imagens se tornou um entrave maior

do que esperado. Dividir a imagem em tamanhos menores auxilia o desempenho da rede, mas visto que as imagens são digitalizadas em 1200 dpi e não são de microscópio, há um limite no quanto o pré-processamento das imagens pode auxiliar.

**Quadro 2** – Métricas das predições da rede neural em contagem levando em conta a acurácia espacial com IoU  $\geq 0,3$ .

Verdadeiros positivos	Falsos verdadeiros	Falsos negativos	precisão	Recall	F1
107	16	26	86,99%	80,45%	83,59%

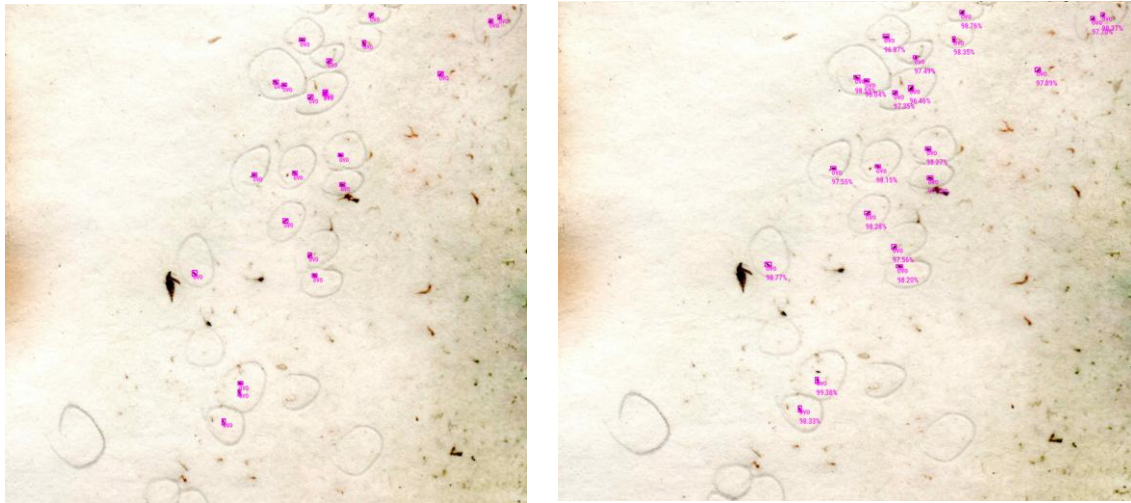
**Fonte:** dos autores (2025).

Em (SILVA, 2021), a rede YOLOv4 utilizada alcançou uma precisão de contagem em paletas de madeira de 94%, uma sensibilidade (*recall*) de 93% e um erro médio absoluto de 0,4 na média. Seu conjunto de dados, entretanto, é de cerca de 3500 imagens de microscópio. A disparidade no tamanho do conjunto de dados é mais do que suficiente para compensar o contraste do papel-filtro, visto que ambas as redes possuem altas capacidades de aprendizado devido a suas arquiteturas. Estima-se, portanto, que uma maior disponibilidade de dados deve melhorar as métricas de precisão e *recall* da rede para os mesmos níveis do trabalho citado.

Em (SILVA et al., 2021), os autores utilizam uma arquitetura U-Net para contar os ovos em ovitrampas. Suas métricas foram inferiores às métricas obtidas neste trabalho, onde ambos contam com um baixo número de imagens de conjunto de dados. As métricas de ambos os trabalhos estão a não mais do que 5% de distância, apontando para resultados similares.

A Figura 4 avalia a predição da rede neural. A Figura 4a mostra a segmentação real e a Figura 4b é a predição da rede. Observe a alta confiança da rede nos ovos. Hipotetiza-se que o alto contraste combinado com o formato singular do ovo faça com que a rede tenha altos níveis de confiança nas suas predições.

**Figura 4** – Exemplo da imagem anotada e da saída da rede neural. Nessa comparação, há 23 ovos reais e 21 ovos foram previstos pelo modelo, um erro de 8,69%.



**(a)** Uma das entradas do conjunto de dados devidamente anotada.

**(b)** Predição da rede neural das anotações da imagem real.

**Fonte:** dos autores (2024).

## 5 CONSIDERAÇÕES FINAIS

Neste trabalho, uma proposta de rede neural foi apresentada utilizando arquiteturas do estado da arte para a contagem de ovos do mosquito *Aedes aegypti* em papéis-filtro. A rede neural se mostrou comparável com outras redes da literatura, em especial (SILVA, 2021) e (SILVA et al., 2021). Houve uma surpresa no erro de contagem se comparado às expectativas passadas, visto que um conjunto de dados menor do que cem imagens produziram erros de contagem inferiores a 10% na média. Com o aumento do conjunto de dados, espera-se que o erro diminua e as métricas de precisão, *recall* e *IoU* cresçam para níveis similares ao de (SILVA, 2021).

A Mask R-CNN comprovou a hipótese de que consegue ter capacidades de generalização suficientes para esta tarefa, onde seu único impedimento será o tamanho do conjunto de dados para valores similares ao de (SILVA, 2021). Hipotetiza-se que, dados os avanços nas arquiteturas, a *Mask R-CNN* possa

superar os resultados dos trabalhos supracitados.

Como continuação deste trabalho, um *software* para a detecção de ovos está sendo produzido, considerando a rede proposta como o motor principal. O conjunto de dados ainda será expandido de acordo com a disponibilidade da Sala de Situação de Saúde da FS/UnB em fornecer os papéis-filtro para o refinamento da rede. Espera-se que com o aumento do conjunto de dados e a consequente melhoria dos resultados, somado ao *software* para encapsular a rede, o seu uso possa ser feito por agentes da saúde auxiliando a vigilância epidemiológica brasileira contra as arboviroses transmitidas pelo mosquito *Aedes aegypti*.

## REFERÊNCIAS

- BRUN, A. L. et al. Uma revisão das técnicas computacionais para contagem de ovos de *Aedes aegypti* em imagens de ovitrampas. **Revista Brasileira de Computação Aplicada**, v. 12, n. 3, p. 1–15, 2020. DOI: 10.5335/rbca.v12i3.10947.
- CDC – CENTERS FOR DISEASE CONTROL AND PREVENTION. **Life cycle of Aedes mosquitoes**. 2024. Disponível em: <https://www.cdc.gov/mosquitoes/about/life-cycle-of-aedes-mosquitoes.html>. Acesso em: 26 set. 2025.
- GABURRO, J. et al. Assessment of ICount software, a precise and fast egg counting tool for the mosquito vector *Aedes aegypti*. **Parasites & Vectors**, 2016. DOI: 10.1186/s13071-016-1870-1
- GIRSHICK, R. B. Fast R-CNN. **arXiv preprint**, 2015. DOI: 10.48550/arXiv.1504.08083.
- GIRSHICK, R. B. et al. Rich feature hierarchies for accurate object detection and semantic segmentation. **arXiv preprint**, 2013. DOI: 10.48550/arXiv.1311.2524.
- HE, K. et al. Mask R-CNN. **arXiv preprint**, 2017. DOI: 10.48550/arXiv.1703.06870.
- JAVED, N. et al. EggCountAI: a convolutional neural network-based software for counting of *Aedes aegypti* mosquito eggs. **Parasites & Vectors**, 2023. DOI: 10.1186/s13071-023-05956-1
- LI, Y. et al. Benchmarking detection transfer learning with vision transformers. **arXiv preprint**, 2021. DOI: 10.48550/arXiv.2111.11429.
- NARANJO-ALCAZAR, J. et al. Automatic counting and classification of mosquito eggs in field traps. **arXiv preprint**, 2024. DOI: 10.48550/arXiv.2405.20656.

PASZKE, A. et al. PyTorch: an imperative style, high-performance deep learning library. **arXiv preprint**, 2019. DOI: 10.48550/arXiv.1912.01703.

REN, S. et al. Faster R-CNN: towards real-time object detection with region proposal networks. **arXiv preprint**, 2015. DOI: 10.48550/arXiv.1506.01497.

SILVA, L. M. de S.; SILVA, L. L.; BARROS, V.; ARAUJO, F. H. D. **Contagem de ovos do *Aedes aegypti* em palhetas de ovitrampas baseada em deep learning**. In: ENCONTRO UNIFICADO DE COMPUTAÇÃO DO PIAUÍ (ENUCOMPI). 2021. p. 128–135. DOI: 10.5753/enucompi.2021.17763.

SILVA, R. E. V. **Contagem automática de ovos do mosquito *Aedes aegypti* utilizando métodos de aprendizagem profunda e dispositivo de baixo custo**. 2021. Dissertação (Mestrado) – Universidade Federal de Pernambuco, Recife, 2021.

SOPER, F. L. The 1964 status of *Aedes aegypti* eradication and yellow fever in the Americas. **The American Journal of Tropical Medicine and Hygiene**, 1965.

WANG, M. et al. Enhancing vector control: AI-based identification and counting of *Aedes albopictus* (Diptera: Culicidae) mosquito eggs. **Parasites & Vectors**, 2024. DOI: 10.1186/s13071-024-06587-w.