

Contagem de ovos do *Aedes Aegypti* em palhetas de ovitrampas baseada em Deep Learning

Lucas M.de Sousa Silva¹, Lucas L. Silva², Veruska Barros³, Flavio H. D. Araujo^{1,2}

¹Programa de Pós Graduação em Engenharia Elétrica,
Universidade Federal do Piauí - Teresina, Brasil

²Departamento de Sistemas de Informação,
Universidade Federal do Piauí - Picos, Brasil

³Departamento de Biologia,
Universidade Federal do Piauí - Teresina, Brasil

{lucasm.desenvolvedor, lucaslopes, vcbbio, flavio86}@ufpi.edu.br

Abstract. *Aedes aegypti* is the vector of one of the most difficult public health problems to be tackled in the tropical world: the spread of the dengue epidemic. As there is no vaccine or specific treatment, and the eradication of its vector has become practically impossible, the best way to avoid the disease is to control the *Aedes aegypti* mosquito. With that, this work presents a computational methodology for the segmentation of eggs in ovitrap straws that aims to help specialists in this counting. For the development of this work we performed tests with the convolutional neural networks developed for segmentation: U-Net, Segnet and a pre-trained network. Then, we performed a post-processing step based on mathematical morphology. The results achieved were promising, and the U-Net network showed the best performance, with an accuracy of 98.65% in egg segmentation and mean square error of 4.25% in counting.

Resumo. O *Aedes aegypti* é o vetor de um dos problemas de saúde pública mais difíceis de serem combatidos no mundo tropical: a proliferação da epidemia de dengue. Como não há vacina ou tratamento específico, e a erradicação do seu vetor se tornou praticamente impossível, a melhor maneira de se evitar a doença é realizando o controle do mosquito *Aedes aegypti*. Com isso, este trabalho apresenta uma metodologia computacional para a segmentação de ovos nas palhetas ovitrampas que visa auxiliar especialistas nessa contagem. Para o desenvolvimento deste trabalho realizamos testes com as redes neurais convolucionais desenvolvidas para segmentação: U-Net, Segnet e uma rede pré-treinada. Em seguida, realizamos uma etapa de pós-processamento baseada em morfologia matemática. Os resultados alcançados foram promissores, e a rede U-Net foi a que apresentou melhor desempenho, com Acurácia de 98,65% na segmentação dos ovos e erro médio quadrado de 4,25% na contagem.

1. Introdução

As arboviroses são doenças infecciosas causadas por vírus que se desenvolvem e são transmitidas pelos insetos vetores. São consideradas agravos emergentes de grande impacto na saúde pública no Brasil. Dentre as arboviroses de importância epidemiológica no

Brasil e transmitidas por *Aedes aegypti*, pode-se citar Dengue, e seus 4 sorotipos, Febre Amarela Urbana, Zika e Chikungunya [Barros et al. 2019].

Uma das formas de realizar esse monitoramento é por meio do uso de ovitrampas. As ovitrampas são recipientes que simulam criadouros em potenciais que são facilmente reconhecidas pelas fêmeas grávidas de mosquitos como locais adequados para a postura de ovos. O uso deste tipo de armadilha pode fornecer importantes informações para o monitoramento destes vetores em uma área. Por meio da contagem de ovos, pode-se verificar a densidade populacional dos mosquitos, sua frequência e dispersão geográfica, a dinâmica populacional de acordo com a sazonalidade, além de servir como método avaliativo de ações de controle anteriormente aplicadas, como o uso de adulticidas. É uma técnica barata, segura e de fácil manuseio, pela qual o agente de endemia poderá verificar presença e quantidade de ovos [Glasser and Gomes 2000].

Neste trabalho é apresentado um método que utiliza visão computacional para a segmentação de ovos em imagens de armadilhas ovitrampas e que pode ser utilizada para auxiliar na contagem dos ovos de mosquitos. A primeira etapa desse método consiste na aquisição de imagens. Em seguida, é aplicado um pré-processamento, onde as imagens são recortadas e normalizadas. Após o pré-processamento, para a segmentação dos ovos, são avaliadas três técnicas baseadas em redes neurais convolucionais (CNN): U-Net, Segnet e uma rede pré treinada. Após a segmentação é feita uma estimativa da quantidade de ovos presentes na palheta de acordo com as regiões segmentadas.

2. Trabalhos Relacionados

Alguns trabalhos de contagem automática, total ou parcial, vem sendo estudados, e as principais técnicas utilizadas são baseadas em: limiarização ou agrupamento com filtragem por área [Silva et al. 2012, Feitosa 2015, Sousa and Paiva 2014], filtragem morfológica [Bandong and Joelianito 2019], Wavelet [Rocha et al. 2019, Yussof et al. 2018] e *Deep Learning* [Santana et al. 2019, Garcia et al. 2019]. Analisando esses trabalhos nota-se que é necessário realizar uma automação no processo de aquisição das imagens, pois a padronização da distância e do aumento do microscópio irá facilitar a etapa de detecção dos ovos. Além disso, a maioria dos trabalhos presentes na literatura necessita da configuração manual de parâmetros para a obtenção de um melhor desempenho, o que aumenta o tempo necessário para a contagem dos ovos e a seleção inadequada dos parâmetros afeta o desempenho da solução. Por fim, os experimentos de validação desses trabalhos muitas vezes não refletem as condições reais encontradas nas palhetas das ovitrampas, tais como: sobreposição e oclusão de ovos, imagens com fundo ruidosos e problemas de iluminação.

Vale destacar que embora existam diversos trabalhos que propõem técnicas para a contagem de ovos, encontramos na literatura somente uma ferramenta que realiza essa tarefa, chamada de ICount [Gaburro et al. 2016]. Foi realizada uma busca, no entanto não obtivemos acesso a essa ferramenta. Em Gaburro et al. (2016) [Gaburro et al. 2016] foi realizada a proposta e avaliação do desempenho dessa ferramenta, e de acordo com esse trabalho, ela utiliza limiarizações e filtros de área para detectar a presença de ovos nas imagens. De acordo com os autores essa ferramenta pode alcançar 98% de precisão sob configurações adequadas, no entanto, a taxa de erro aumenta a medida que a densidade de ovos na imagem aumenta. Em imagens com mais de 200 ovos, por exemplo, a taxa de erro

sobe para cerca de 12%. Além disso, outro problema dessa ferramenta é a necessidade de ajuste manual de vários parâmetros para alcançar resultados satisfatórios. Por fim, as imagens utilizadas no artigo apresentam alto contraste entre a cor dos ovos e a região de fundo, o que torna inviável sua aplicação no mundo real, pois as palhetas apresentam sujeira como areia, fragmentos de folhas secas e restos de pequenos insetos mortos.

3. Método proposto

A metodologia proposta está dividida em cinco etapas, como mostra a Figura 1. A primeira consiste na aquisição das imagens seguida por um pré-processamento. Em seguida, é realizada a segmentação utilizando uma CNN, nesta etapa avaliamos diferentes arquiteturas para o problema proposto. Logo após, é feito o pós-processamento com a finalidade de refinar as regiões segmentadas e remover ruídos. Por fim, na etapa de contagem é feita a estimativa da quantidade de ovos utilizando as imagens segmentadas.



Figura 1. Fluxograma do método proposto dividido em cinco etapas.

3.1. Aquisição das imagens

A aquisição de imagens foi realizada por meio de uma câmera digital com resolução de 7,2 megapixels, LCD 2,5”, zoom óptico de 4,5 vezes e lente Leica DC Vario Elmarit. As ovitrampas foram digitalizadas com 700 dpi de resolução e zoom óptico de 4 vezes. Esse processo gera uma imagem digital RGB de 8 bits com 1024 pixels versus 768 pixels.

A coleção de imagens foi capturada em um laboratório da Universidade Federal do Piauí (UFPI) que é responsável pelo mapeamento da incidência do mosquito *A. aegypti* no Campus da UFPI da cidade de Teresina. Foram digitalizadas 10 paletas, totalizando 30 imagens, pois cada uma gerava 3 imagens de tamanho 1024×768 após a digitalização. Para gerar uma máscara segmentada, um especialista segmentou os ovos manualmente. As Figuras 2 e 3 mostram um exemplo de imagem de uma palheta de otarampa contendo alguns ovos e sua máscara equivalente.

3.2. Pré-processamento

As imagens obtidas originalmente são grandes para serem utilizadas diretamente no treinamento das redes neurais convolucionais, dessa forma, seria necessário fazer o redimen-

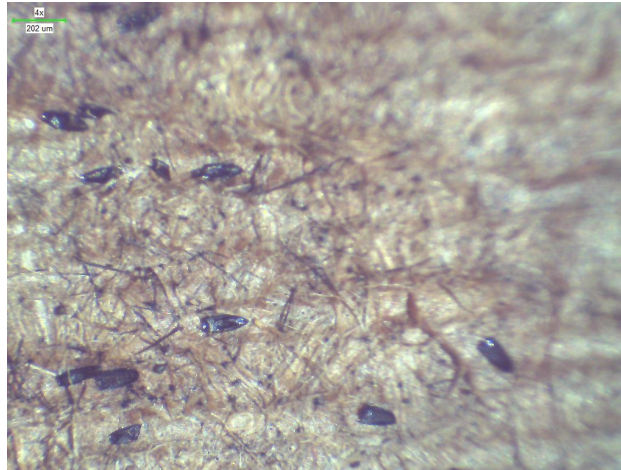


Figura 2. Imagem adquirida de uma palheta de ovitrampa

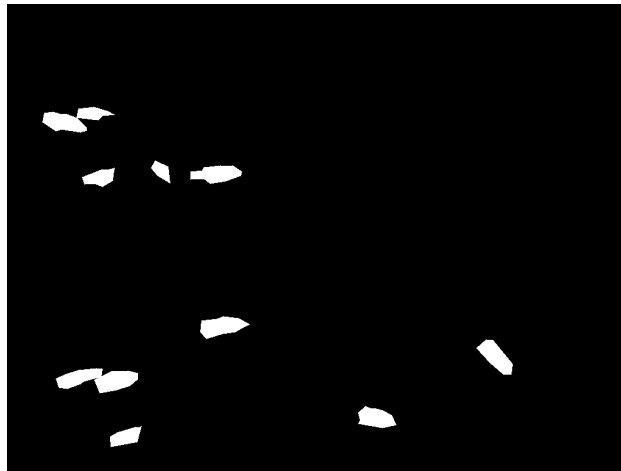


Figura 3. Imagem de máscara segmentada manualmente.

sionamento das mesmas. Além disso, como elas não são quadradas, esse redimensionamento poderia causar distorções. Por esse motivo, as imagens originalmente em tamanho 1024×768 são recortadas em 12 sub-imagens de dimensão 256×256 . e com isso com o corte a própria CNN identifica as cores dos ovos caso tenha cor/textura similar ela marca, depois os cortes são unidos, Assim as imagens ficam quadradas, e com um tamanho que torna possível o treinamento das redes neurais convolucionais. Com os recortes, a quantidade de imagens para treinamento aumenta, favorecendo o aprendizado das redes. Após o pré-processamento ficaram 228 imagens resultantes. Essas imagens foram normalizadas entre 0 e 1.

A Figura 4 apresenta o exemplo do recorte realizado para uma imagem da base, onde a imagem original de tamanho 1024×768 é dividida em 12 novas imagens de tamanho 256×256 .

3.3. Segmentação

Na etapa de segmentação avaliamos o desempenho de três arquiteturas: U-Net, SegNet e uma rede pré-treinada.

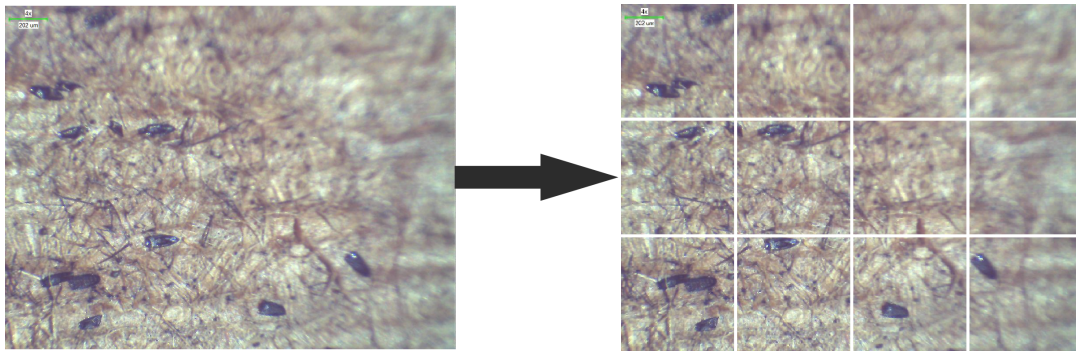


Figura 4. Imagem de uma palheta de ovitrampa (esquerda), divisão das imagens após o recorte (direita)

A U-Net possui duas etapas principais, a primeira é o caminho de contração, no qual a imagem de entrada passa por 2 convoluções de tamanho 3x3, stride 1, *padding* que preserva as bordas e ReLU como função de ativação. As camadas convolucionais são seguidas por uma operação de *max pooling* 2x2 com *stride* 2. Após cada operação de *max pooling*, a quantidade de mapa de características são aumentados por um fator de dois e o tamanho da entrada é reduzido pelo mesmo fator devido aos efeitos do *max pooling*. No caminho de contração, uma etapa é definida por duas convoluções e uma operação de *max pooling* [LONG 2015].

A rede SegNet, tem o objetivo de fazer a adoção de redes convolucionais originalmente aplicadas para classificação, sendo usadas para segmentação de pixels. Na SegNet, o encoder é topologicamente idêntico as camadas da rede VGG13 [Simonyan and Zisserman 2014] (treze camadas de convolução, salvo as que são totalmente conectadas. Cada operação de *max pooling* salva um índice, que é passado posteriormente às camadas apropriadas do encoder. As operações de *max pooling* e *up sampling* consistem basicamente em reduzir e aumentar o tamanho espacial da imagem, respectivamente [Andrade 2010].

Na terceira técnica avaliada, foi usada uma a rede pré-treinada como ponto de partida. Isso se deve ao fato de que as primeiras camadas das CNNs aprendem a reconhecer elementos básicos das imagens, tais como bordas, cantos, formatos arredondados e outras formas geométricas básicas, além de constatações básicas sobre cores. Quando utilizamos redes pré-treinadas para treinar em cima de novas categorias, essas camadas iniciais são pouco modificadas, pois representam elementos básicos presentes em qualquer tipo de imagem. Assim, o treinamento fica muito mais rápido [Silva et al. 2012]. Neste trabalho foi usado a Resnet50 que primeiro usa uma convolução 1x1 para redução da dimensionalidade, depois a convolução 3x3 e, finalmente, usando a dimensionalidade 1x1 para restaurar a dimensão original. Isso é equivalente a reduzir o número de parâmetros para o mesmo número de camadas, para que possa ser estendido para modelos mais profundos. Portanto a ResNet com 50 camadas, além de não ter problemas de degradação, a taxa de erro é bastante reduzida e a complexidade computacional fica muito baixa.[MENDES 2017]

3.4. Pós-processamento

É nesta etapa que os principais defeitos ou imperfeições da segmentação são devidamente corrigidos. Primeiro foi utilizada uma operação de fechamento morfológico com elemento estruturante em disco de raio 7. O objetivo dessa operação foi conectar os componentes das regiões segmentadas separadamente, em seguida, foi utilizada uma operação de abertura morfológica para a remoção de *pixels* isolados que foram segmentados. O elemento estruturante utilizado também foi um disco de raio 7. Por fim, foi utilizado uma técnica para fechar os buracos segmentados em regiões de ovos.

A Figura 5 ilustra o pós-processamento. Pode ser visto que ele remove *pixel* isolados e também conecta regiões de ovos que foram segmentadas separadamente.



Figura 5. Imagem original da palheta (esquerda), imagem da palheta segmentada pela U-Net (centro) e imagem com pós-processamento (direita).

3.5. Contagem

A contagem dos ovos se baseia no uso da média e do algoritmo de rotulação de componentes conectados. Na aquisição das imagens os ovos foram separados individualmente, formando várias imagens com apenas 1 ovo em cada imagem. O conjunto de treino foi utilizado para estimar o tamanho médio do ovo, resultando no tamanho de 1321 pixels.

Após a obtenção da média é possível aplicar a técnica de contagem proposta. No resultado do pós-processamento é aplicado um algoritmo de rotulação de componentes conectados. Se uma região tiver a área menor que 1,5 do tamanho médio do ovo, é considerado como somente um ovo. Caso a região seja maior que 1,5 do tamanho médio do ovo, é realizada a divisão do tamanho da região pelo tamanho médio do ovo, e o resultado inteiro é a quantidade de ovos contido naquela região. Esse processo é feito dessa forma porque ovos que estão localizados próximos são segmentados em uma única região.

3.6. Resultados e Discursões

Para a obtenção dos resultados, 80% das imagens foram utilizadas para treino e 20% para teste. Em ambas as arquiteturas foram usadas 50 épocas de treino, *batch size* de tamanho 4 e coeficiente de aprendizado com o valor 0.001.

Cada teste foi realizado 5 vezes, a Tabela 1 apresenta a média e o desvio padrão dos resultados obtidos antes e após o pós-processamento. A Tabela 2 apresenta a quantidade real de ovos e a estimada após a segmentação. Vale destacar que para cada uma das 4 imagens apresentadas na Tabela 2, são segmentadas 12 imagens de tamanho 256×256 .

Pela análise da Tabela 1 percebe-se que a arquitetura U-net produziu os melhores resultados. Pode ser visto também que o pós-processamento teve um impacto positivo no resultado final, causando aumento nos valores das métricas.

Tabela 1. Resultado e desvio padrão antes e após o pós-processamento

Sem pós-processamento					
Arquiteturas	IoU	Dice	Precisão	Recall	Acurácia
U-Net	63.85 ± 06.12	77.32 ± 04.87	83.35 ± 03.31	75.00 ± 09.58	98.64 ± 00.30
SegNet	58,62 ± 01,16	73,82 ± 00,90	76,07 ± 00,02	72,51 ± 00,02	97,11 ± 00,35
Pré-treinada	60,04 ± 04,07	60,04 ± 03,75	60,04 ± 08,80	76,12 ± 12,18	97,99 ± 00,32
Com pós-processamento					
U-Net	64,29 ± 05,62	77,66 ± 04,42	83,36 ± 03,62	75,54 ± 09,12	98,65 ± 00,30
SegNet	59,13 ± 01,04	74,23 ± 00,81	75,71 ± 00,12	73,57 ± 01,92	97,13 ± 00,33
Pré-treinada	60,24 ± 04,14	74,57 ± 03,82	76,29 ± 08,71	76,15 ± 12,13	98,00 ± 00,32

Tabela 2. Estimativa da quantidade de ovos das imagens de teste

Imagem	Quantidade de ovos	Quantidade estimada de ovos
1	10	20
2	18	14
3	32	31
4	11	13

O erro médio quadrado obtido na contagem dos ovos foi de 4,25, ou seja, para cada imagem a diferença entre a quantidade real e a estimada é em média 4 ovos. É uma média aceitável, porém como pode ser visto pela imagem 1 da Tabela 2, existem erros individuais que são maiores. Isso acontece devido a variação do tamanho dos ovos, já que a estimativa se baseia no tamanho médio do ovo. Na base de imagens existem ovos que podem ser até 8x maiores que outros. Provavelmente uma nova abordagem deva ser explorada para uma contagem mais satisfatória.

4. Conclusão

Neste trabalho foi apresentada uma metodologia computacional que visa auxiliar especialistas na contagem de ovos nas palhetas das ovitrampas. Os melhores resultados da metodologia foram alcançados com a rede U-Net, com Acurácia de segmentação de 98,65% e erro médio quadrado de 4,25%. Além disso, a etapa de pós-processamento melhorou os resultados da segmentação.

Um dos problemas obtido na estimativa da quantidade de ovos segmentados por imagem foi causado pela diferença entre o tamanho dos ovos, por tanto, como trabalhos futuro pretendemos investigar outras abordagens para tratar esse problema. Além disso, a etapa de segmentação também pode produzir melhores resultados com o aumento da base de imagens e com a descoberta dos melhores parâmetros das CNNs e do pós-processamento.

Referências

- Andrade, L. (2010). Identificação automática de áreas cafeeiras em imagens de satélite utilizando redes neurais artificiais. *Embrapa Cafe-Artigo em anais de congresso (ALICE)*, 28.
- Bandong, S. and Joelianto, E. (2019). Counting of aedes aegypti eggs using image processing with grid search parameter optimization. In *2019 International Conference on Sustainable Engineering and Creative Computing (ICSECC)*, pages 293–298.

- Barros, V. C., Pacheco, A. C. L., Feitosa, L. M., Figueredo, J. S., Batista, F., Lima, I. P., and Barbosa, O. A. A. (2019). *Produtos naturais no combate ao mosquito Aedes aegypti*. Campinas: Atomo.
- Feitosa, L. N. (2015). Sistema de contagem automática de ovos do aedes aegypti a partir de processamento de imagens das palhetas de ovitrampas. Monografia, Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte.
- Gaburro, J., Duchemin, J.-B., Paradkar, P., Nahavandi, S., and Bhatti, A. (2016). Assessment of icount software, a precise and fast egg counting tool for the mosquito vector aedes aegypti. *Parasites Vectors*, 9:1–9.
- Garcia, P. S. C., Martins, R., Coelho, G. L. L. M., and Cámara-Chávez, G. (2019). Acquisition of digital images and identification of aedes aegypti mosquito eggs using classification and deep learning. In *2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 47–53.
- Glasser, M. and Gomes, C. (2000). Infestação do estado de são paulo por aedes aegypti e aedes albopictus. *Revista de Saúde Pública*, 34:570–577.
- LONG, J.; SHELHAMER, E. D. (2015). Fully convolutional networks for semantic segmentation. *IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 28:3174 – 3179.
- MENDES, C. C. T.; FRÉMONT, V. W. D. F. (2017). Exploiting fully convolutional neural networks for fast road detected). *IEEE. Robotics and Automation (ICRA)*,, 28.
- Rocha, C., Bizerra, A., Coutinho, D., Silva, L., de Deus, M., and Souto, P. (2019). *Contagem automática de ovos do Aedes Aegypti em palhetas de ovitrampas: um sistema para aquisição e processamento de imagens*, pages 67–74.
- Santana, C. et al. (2019). A solution for counting aedes aegypti and aedes albopictus eggs in paddles from ovitraps using deep learning. (17):1987–1994.
- Silva, M. A. G. N. M. d., Rodrigues, M. A. A. B., and Araujo, R. E. d. (2012). System for acquisition and processing of ovitraps images to fight dengue. *Brazilian journal of biomedical engineering*, 28:364 – 374.
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 28.
- Sousa, J. A. and Paiva, A. C. (2014). Contagem automática de ovos de mosquito da dengue em imagem de ovitrampa. In *2014 - V Jornada de Informática do Maranhão (JIM)*, pages 1–4.
- Yussof, W., Man, M., Hitam, M., Abdul Hamid, A. A., Awalludin, E., and Wan Abu Bakar, W. A. (2018). Wavelet-based auto-counting tool of aedes eggs. pages 56–59.