

Segmentação e contagem automática de ovos do *Aedes Aegypti* em palhetas de ovitrampa baseada em aprendizado profundo

Automatic segmentation and counting of *Aedes Aegypti* eggs in ovitraps based on deep learning

Lucas M.de Sousa Silva
Programa de Pós Graduação
em Engenharia Elétrica,
Universidade Federal do Piauí
- Teresina, Brasil
lucasm.des@ufpi.edu.br

Lucas S. Lopes
Departamento de Sistemas de
Informação,
Universidade Federal do Piauí
- Picos, Brasil
lucaslopes092020@ufpi.edu.br

Veruska Barros
Departamento de Biologia,
Universidade Federal do Piauí
- Teresina, Brasil
vcbbio@ufpi.edu.br

Flavio H. D. Araujo
Programa de Pós Graduação
em Engenharia Elétrica,
Departamento de Sistemas de
Informação,
Universidade Federal do
Piauí, Brasil
flavio86@ufpi.edu.br

ABSTRACT

Dengue is an endemic disease, being one of the main health problems worldwide. The ovitraps simulate a perfect environment for breeding *Aedes aegypti*. It consists of a black pot filled with water, which remains stationary, attracting the mosquito. A wooden straw is inserted into it, which makes easier for the female *Aedes aegypti* to lay eggs. However, the major disadvantage of this technique is the counting of mosquito eggs that is performed manually, one by one, using a magnifying glass, making the counting process slow. Thus, this work presents a computational methodology to segment and counting eggs in ovitrap straws that aims to help specialists in this counting. For the development of this work, we carried out tests using different color spaces and U-net architectures with different numbers of layers and filters. Then, we searched for the best parameters of a post processing step based on mathematical morphology. The results achieved were promising with an accuracy of 98.65% in the segmentation of eggs and a mean square error of 4.47 in the counting of eggs.

Keywords

Aedes aegypti; U-net; Egg count; Parameter selection.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

RESUMO

A dengue é uma doença endêmica, sendo um dos principais problemas de saúde em todo o mundo. As ovitrampas simulam um ambiente perfeito para a reprodução do mosquito *Aedes aegypti*. Consiste em um pote preto cheio de água, que permanece parado, atraindo o mosquito. Nele é inserido um canudo de madeira, o que facilita a postura de ovos da fêmea do *Aedes aegypti*. No entanto, a grande desvantagem desta técnica é a contagem dos ovos do mosquito que é realizada manualmente, um a um, utilizando uma lupa, tornando o processo de contagem lento. Assim, este trabalho apresenta uma metodologia computacional para segmentação e contagem de ovos em canudos de ovitrampa que visa auxiliar especialistas nesta contagem. Para o desenvolvimento deste trabalho, realizamos testes utilizando diferentes espaços de cores e arquiteturas U-net com diferentes números de camadas e filtros. Em seguida, buscamos os melhores parâmetros de uma etapa de pós-processamento com base em morfologia matemática. Os resultados alcançados foram promissores com acurácia de 98,65% na segmentação de ovos e erro quadrático médio de 4,47 na contagem de ovos.

Palavras-Chave

Aedes aegypti; U-net; Contagem de ovos; Seleção de parâmetros.

CCS Concepts

• **AUC Area under the receiver operating characteristic curve;** • **CNN_Convolutional neural network;** • **DL, Deep Learning;**

1. INTRODUÇÃO

As arboviroses são doenças infecciosas causadas por vírus que se desenvolvem e são transmitidas pelos insetos vetores [18]. 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 [8].

A ovitrampa que foi inicialmente proposta por [35], se caracteriza como uma importante ferramenta para detecção e monitoramento do mosquito. É uma técnica segura, barata e que não agride o ambiente, permitindo sua fácil utilização em qualquer local. Ela determina a dispersão geográfica, densidade, frequência, ocupação, dominância e sazonalidade [21]. Sua utilização permite a contagem dos ovos depositados.

Nas Américas, a espécie *Aedes aegypti* é a responsável pela transmissão da dengue. Estes mosquitos proliferam-se dentro ou nas proximidades de habitações e em recipientes onde se acumula água limpa[43]. O ciclo do *Aedes aegypti* é composto por quatro fases: ovo, larva, pupa e adulto. Os ovos do mosquito são postos milímetros acima da superfície da água e podem ficar até um ano inativo a espera de água propícia a sua incubação[14]. Na fase do acasalamento, em que as fêmeas precisam de sangue para garantir o desenvolvimento dos ovos, ocorre a transmissão da doença.

O único modo possível de evitar a transmissão da dengue é a eliminação do mosquito transmissor[40]. Para se fazer o combate ao mosquito faz-se necessário o monitoramento do vetor de transmissão e assim traçar planos de controle da doença[23]. A análise estatística da população de *Aedes aegypti* é uma forma de fazer o monitoramento, e baseia-se no uso de ovitrampas para fazer a coleta e a contagem dos ovos[32]. As ovitrampas são armadilhas usadas por especialistas para coletar os ovos do mosquito, é composta por uma palheta de madeiras contendo um larvicida, e tem capacidade para armazenar mais de 1000 ovos. A contagem dos ovos presentes nas ovitrampas geralmente é feita de forma manual com o auxílio de lupa ou microscópio, configurando-se assim como um trabalho difícil[7], cansativo e passível de erros [20]. Dessa forma, um sistema que automatiza, ou auxilia nessa contagem, pode contribuir para o aumento da eficiência do monitoramento dos mosquitos.

Neste trabalho é proposto um método para a segmentação de ovos em imagens de armadilhas ovitrampas. Para a criação desse método foram realizadas seis etapas. A primeira etapa consiste na aquisição de imagens. Em seguida, foi realizado um pré-processamento, onde as imagens são recortadas e normalizadas. As imagens pré-processadas são utilizadas para a definição do melhor modelo de cor a ser utilizado. As duas etapas seguintes consiste na seleção da melhor arquitetura de U-net e na seleção dos melhores parâmetros de pós-processamento. Por fim, após a segmentação das regiões de ovos, é realizada a estimativa da quantidade de ovos presentes na palheta.

2. TRABALHOS RELACIONADOS

A literatura apresenta uma série de trabalhos que propõem soluções para o problema de contagem de ovos de *Aedes aegypti* e *Aedes albopictus* em imagens. Em [33] foi apresentada uma solução automatizada para contagem de ovos de *Aedes aegypti* em palhetas de ovitrampas. Nessa solução, o

profissional utiliza uma câmera digital para obter imagens da palheta. Essas imagens são processadas por um programa de computador que executa uma sequência de técnicas de processamento de imagem para detectar e contar o número de ovos. O primeiro processamento realizado pelo programa é a conversão das imagens de RGB (vermelho, verde, azul) para o espaço de cores YIQ (Luminance, In-phase, Quadrature). Em seguida, o componente “I” é binarizado e passa por um algoritmo de componentes conectados e uma série de filtros de redução de ruído. Finalmente, usando a informação da área média de um ovo, o programa estima o número de ovos nas áreas apresentadas na imagem.

Uma solução totalmente automatizada foi apresentada por [12]. Essa solução é controlada por um microprocessador e é capaz de capturar as imagens das palhetas e, em seguida, as envia para um computador conectado via porta serial para que seja feita a detecção e contagem dos ovos nas imagens [16]. O algoritmo de contagem de ovos proposto é baseado em filtros de segmentação, e a estimativa do número de ovos foi baseada no tamanho médio de um ovo. Após a execução dos experimentos em uma base de dados contendo 100 imagens de palhetas de ovitrampas[19], o método proposto apresentou uma taxa de erro médio de 2,67 % .

Em [17] foi avaliado o desempenho da ferramenta de contagem de ovos de *Aedes aegypti* em imagens: o ICount. Essa ferramenta usa filtros de tamanho e thresholds para detectar a presença de ovos nas imagens. O ICount também pode ser utilizada para estimar a densidade de larvas de *Aedes aegypti* em imagens[42]. De acordo com os autores desse estudo, com a configuração adequada, essa ferramenta pode alcançar 98% de precisão. No entanto, vale ressaltar que a taxa de erro aumenta a medida que o a densidade de ovos nas imagens cresce [13]. Por exemplo, para imagens com mais de 200 ovos, a taxa de erro sobe para 12% . Outro problema dessa ferramenta está na necessidade de se ajustar manualmente vários parâmetros para alcançar resultados satisfatório.[9]

Em [42] foi empregado Gabor Wavelets para separar os ovos do fundo (palheta) das imagem. O resultado dessa operação passa por um algoritmo de limiar global baseado no método de Otsu para binarizar a imagem [22]. Em seguida, eles aplicam uma operação de dilatação e depois consideram o primeiro plano como a área que contém os ovos. A estimativa do número de ovos na imagem é realizada a partir da divisão da área do primeiro plano pelo tamanho médio do ovo (definido como 100 pixels nesse estudo)[5]. Para reduzir a ocorrência de falsos positivos, são desconsiderados todos objetos no primeiro plano que não possuam sua a área descrita por uma elipse de excentricidade maior que 0,5.

Pela análise dos trabalhos foi observado a medida que a densidade de ovos aumenta nas imagens, a tarefa de segmentação fica cada vez mais desafiadora, principalmente devido a proximidade, ou mesmo a ovos que ficam se tocando na imagem [34]. Dessa forma, as ferramentas encontradas não possuem bom desempenho nessas situações. Outro ponto observado, é que as ferramentas encontradas não realizam um trabalho completamente automático [15], muitas delas necessitam de iteração com o usuário para definição de técnicas e configuração de parâmetros que é realizada manualmente pelo usuário.

3. MÉTODO PROPOSTO

A metodologia proposta está dividida em seis etapas, como

mostra a Figura 1. A primeira consiste na aquisição das imagens, em seguida é realizada a etapa de pré-processamento. As imagens resultantes do pré-processamento são utilizadas para a avaliação de alguns modelos de cor e assim definir o melhor a ser utilizado. Após a definição do melhor modelo de cor, é realizada a seleção da melhor arquitetura de U-net, nesta etapa avaliamos arquiteturas com diferentes valores de camadas e filtros. Logo após, é feito a seleção dos melhores parâmetros da etapa de pós-processamento. Por fim, na etapa de contagem é feita a estimação da quantidade de ovos utilizando as imagens segmentadas.

3.1 Aquisição das imagens

As imagens utilizadas neste trabalho foram capturadas no laboratório da Universidade Federal do Piauí (UFPI) - LAPES (Laboratório de Parasitologia e Entomologia Sanitária), que é responsável pelo mapeamento da incidência do mosquito *A. aegypti* no Campus da UFPI da cidade de Teresina. A aquisição das 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.

Foram digitalizadas 9 palhetas, totalizando 27 imagens, já que cada palheta digitalizada gerou 3 imagens. Com o propósito de treinar a CNN e calcular o tamanho médio do ovo, para cada imagem um especialista segmentou os ovos manualmente. A Figura 2 mostra um exemplo de imagem de uma palheta de ovitrampa 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 redimensionamento das mesmas. Além disso, como elas não são quadradas, esse redimensionamento poderia causar distorções. Assim, as imagens originalmente em tamanho 1024×768 são recortadas em 12 sub-imagens de dimensão 256×256 . Após esse pré-processamento, 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 da CNN. Após o pré-processamento ficaram 300 imagens resultantes. Essas imagens foram normalizadas entre 0 e 1.

A Figura 3 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 Escolha do modelo de cor

Um modelo ou espaço de cor é um modelo matemático abstrato que descreve um método de representação de cores por tuplas de números, cada posição da tupla possui intervalos de valores definidos pelo próprio modelo [6]. O modelo mais comum para a representação de imagens coloridas é o RGB, as imagens da base utilizada estão originalmente nesse formato, apesar de existirem outros. Visivelmente é possível observar diferenças entre imagens em variados espaços de cor geradas a partir de uma mesma imagem colorida, o que torna alguns espaços de cores mais adequados para a identificação de determinadas regiões da imagem. Os sis-

temas de cor avaliados neste trabalho foram o RGB, HSV e LAB, eles foram selecionados por serem os mais recorrentes na literatura.

O Hue Saturation Value (HSV) é um modelo de cores formado por três componentes (matiz, saturação, brilho), em que os valores de matiz variam de 0 a 360, e os componentes saturação e brilho variam entre 0 e 1 [37]. Esse modelo é comumente utilizado no processamento de imagens, no qual pode ser empregado, por exemplo, para facilitar a segmentação de imagens [25].

O modelo LAB, ou L^*A^*B consiste em um espaço de cor perceptualmente uniforme, ou seja, as distâncias numéricas do sistema são relacionadas às diferenças perceptuais humanas [5]. Essa propriedade é útil em aplicações onde há a necessidade de decisões sobre a aceitabilidade de cores [37]. No canal L, os valores estão no intervalo [0, 100], onde 0 especifica preto e 100 especifica branco e à medida que L aumenta, as cores ficam mais brilhantes. Embora não haja um intervalo único para os canais A e B, seus valores geralmente situam-se no intervalo [-100, 100] ou [-128, 127] [39].

3.4 Configuração da U-net

Diversos trabalhos na literatura utilizam técnicas baseadas em deep learning para problemas de visão computacional, pois essas técnicas produzem excelentes resultados [2, 3, 10, 31]. Em [38] foram avaliadas as redes neurais convolucionais U-net, Segnet e uma ResNet pré treinada, sendo a U-net a arquitetura que obteve melhor performance na segmentação de ovos em imagens de palhetas de ovitrampas. Dessa forma, nesse trabalho foi realizada uma busca pela melhor arquitetura de U-net, onde foram avaliadas arquiteturas com diferentes quantidades de camadas e filtros.

A U-net é uma arquitetura de rede codificador decodificador em forma de U, que consiste em blocos codificadores e decodificadores conectados por meio de uma ponte [11]. A U-net é simétrica e as conexões de salto entre o caminho de contração e expansão combinam as informações de localização do caminho de downsampling com as informações contextuais no upsampling. Além disso, não tem nenhuma camada Densa. O modelo U-Net pode ser usado em diferentes conjuntos de imagens e os resultados são bastante satisfatórios. Técnicas de aumento de dados como invariância de deslocamento e rotação podem ser muito úteis para ensinar à rede as propriedades de invariância e robustez desejadas quando temos apenas algumas amostras de treinamento [26]. A U-net é uma arquitetura de DL que segue a estrutura de uma CNN que foi desenvolvida primordialmente para a segmentação de imagens biomédicas. Nessas aplicações existe uma certa complexidade e dificuldade de obter um banco de dados suficientemente grande para treinamento da rede [36].

A U-Net possui duas etapas principais, a primeira é o caminho de contração, no qual a imagem de entrada passa por 2 convoluções com 8 kernels de tamanho 3×3 , stride 1, padding que preserva as bordas e ReLU como função de ativação. As camadas convolucionais são seguidas por uma operação de max pooling 2×2 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, como é representado pela Figura 4. No caminho de contração, uma etapa é definida por duas convoluções e uma operação de max pooling [30].

Neste trabalho foram avaliadas arquiteturas com 3, 4 e

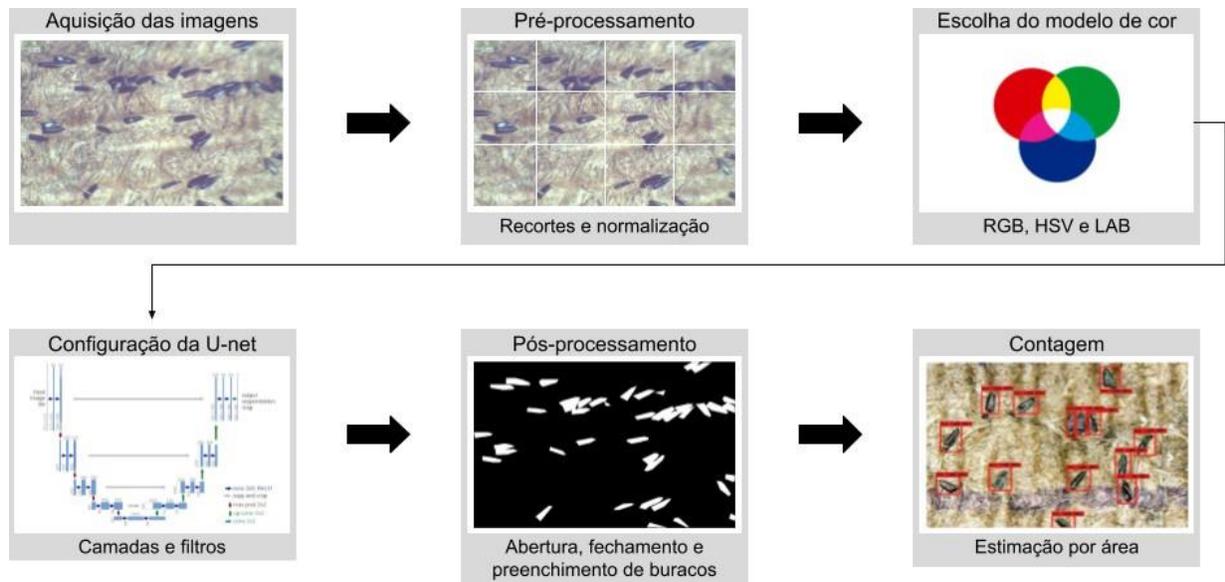


Figure 1: Fluxograma do método proposto dividido em seis etapas.

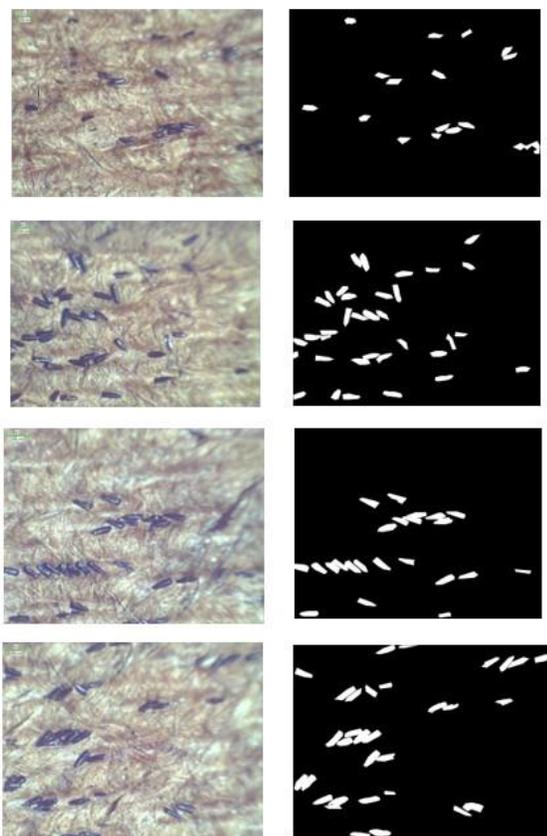


Figure 2: Imagens adquiridas de uma palheta de ovitrampa e suas máscaras equivalentes.

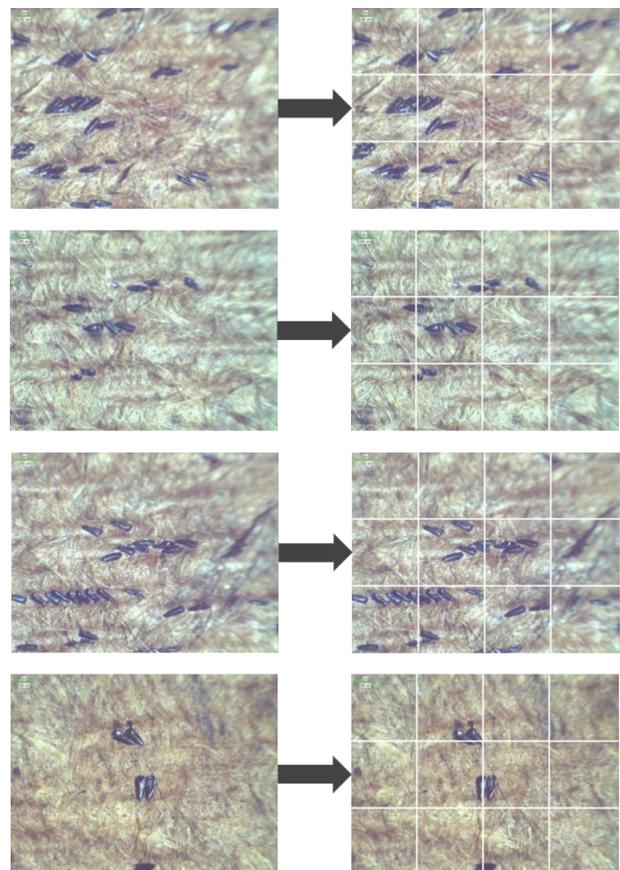


Figure 3: Imagens de palhetas ovitrampa (esquerda), divisão das imagens após o recorte (direita).

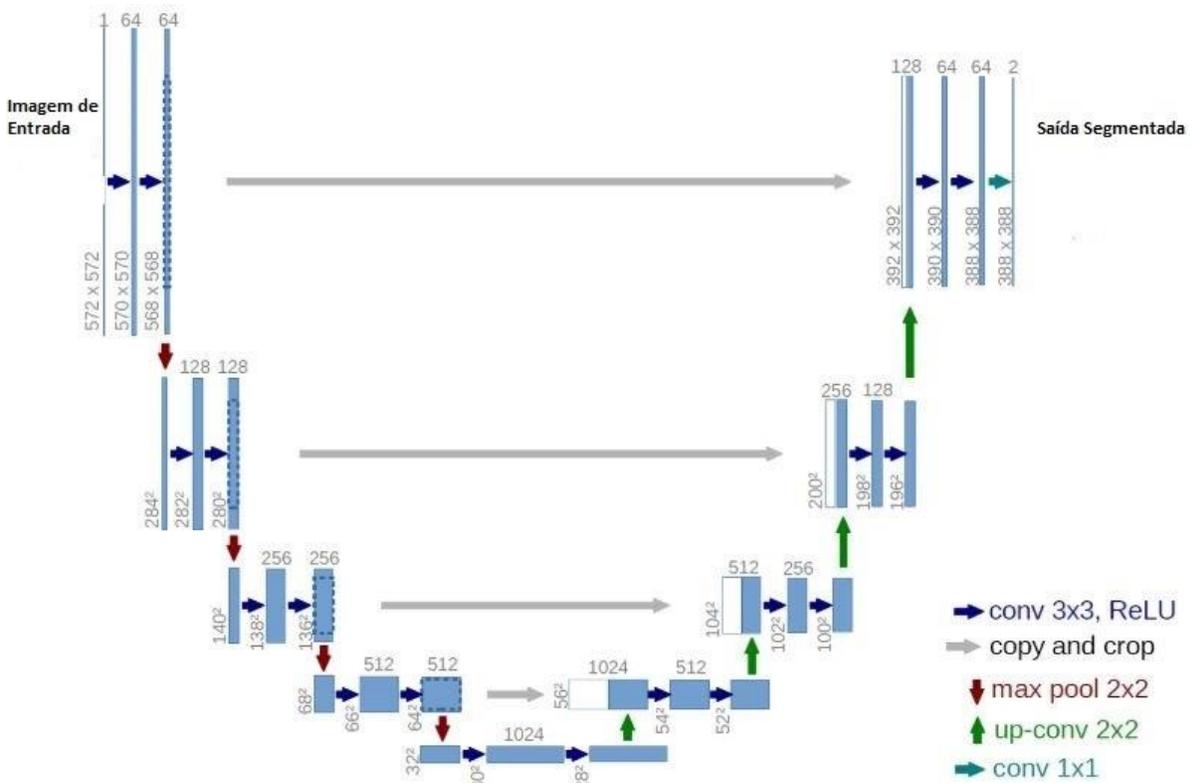


Figure 4: Representação da arquitetura da rede U-net.

5 camadas convolucionais, cada uma contendo 16, 32 e 64 filtros.

3.5 Pós-processamento

É nesta etapa que os principais defeitos ou imperfeições da segmentação realizada pela U-net são corrigidas, e também ocorre a remoção de pequenos ruídos [41]. Nesta etapa foram avaliadas as técnicas de fechamento e abertura morfológica, bem como foi feita a estimação do raio do elemento estruturante disco utilizado. Os valores avaliados foram 3, 5, 7, 9 e 11. O objetivo da operação de fechamento é conectar os componentes das regiões segmentadas separadamente, em seguida, a abertura morfológica é utilizada para a remoção de *pixels* isolados que foram segmentados. Após o uso dessas duas técnicas, também é realizado o preenchimento de buracos nas regiões segmentadas.

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.

3.6 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 1528 pixels.

Após a obtenção da média é possível aplicar a técnica de contagem proposta. Na imagem resultante após o pós-processamento é aplicado um algoritmo de rotulação de componentes conectados. Se uma região tiver a área menor que

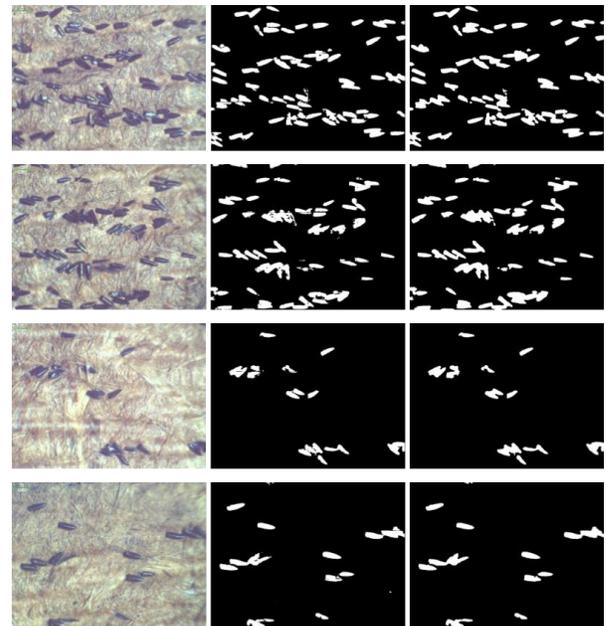


Figure 5: Imagens originais das palhetas (esquerda), imagens segmentadas pela U-net (centro) e imagens com pós-processamento (direita).

1,5 do tamanho médio do ovo, é considerado como somente um ovo. Caso a região seja maior que 1,5 do tamanho mé-

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.

3.7 Métricas de Avaliação

As métricas de avaliação são utilizadas para determinação o melhor modelo analisado para o problema. Nesse trabalho foram usadas as seguintes métricas: Intersection over Union (IoU), Dice, Precisão, Recall e Acurácia.

IoU, também conhecido como índice Jaccard, é a métrica de avaliação mais popular para tarefas como segmentação, detecção de objetos e rastreamento [24]. A detecção de objetos consiste em duas sub tarefas: Localização, que determina a localização de um objeto em uma imagem; e Classificação, que consiste em atribuir uma classe ao objeto. O IoU é calculado segundo a equação 1.

$$IoU = \frac{A \cap B}{A \cup B}, \quad (1)$$

onde A representa a máscara segmentada manualmente pelo especialista e B a região segmentada pelo método proposto.

O coeficiente Dice, também chamado de índice de sobreposição, é a métrica mais utilizada na validação de segmentações de volume médico [29]. A Equação 2 ilustra como calcular o coeficiente Dice.

$$Dice = \frac{2 * VP}{(VP + FP) + (VP + FN)} \quad (2)$$

De acordo com [11, 4], em problemas de classificação binária, predições podem ter quatro possíveis classes: Verdadeiro positivo (VP): quando o método diz que a classe é positiva e, ao verificar a resposta, vê-se que a classe era realmente positiva; Verdadeiro negativo (VN): quando o método diz que a classe é negativa e, ao verificar a resposta, vê-se que a classe era realmente negativa; Falso positivo (FP): quando o método diz que a classe é positiva, mas ao verificar a resposta, vê-se que a classe era negativa; Falso negativo (FN): quando o método diz que a classe é negativa, mas ao verificar a resposta, vê-se que a classe era positiva;

A precisão é uma métrica que avalia a quantidade de verdadeiros positivos sobre a soma de todos os valores positivos [27], como é mostrado na Equação 3

$$Preciso = \frac{VP}{VP + FP} \quad (3)$$

A métrica Recall calcula quantos dos Positivos Reais o modelo proposto acerta ao rotulá-los como Positivo (Verdadeiro Positivo) [9]. A Equação 4 ilustra o cálculo desse valor.

$$Recall = \frac{VP}{VP + FN} \quad (4)$$

A acurácia é considerada uma das métricas mais simples e importantes. Ela avalia o percentual de acertos [36], ou seja, ela pode ser obtida pela razão entre a quantidade de acertos e o total de entradas, como pode ser visto na Equação 5.

$$Acurcia = \frac{VP + VN}{VP + VN + FP + FN} \quad (5)$$

O erro médio quadrado (MSE) tem como principal objetivo encontrar a diferença média de um valor e o seu parâmetro inicial. De maneira mais prática, o seu uso é destinado a compreender um erro de previsão. Ele é calculado pela soma de todos os resultados tidos como erro em relação à previsão inicial e, posteriormente, dividi-los pela quantidade de valores somado, conforme mostra a Equação 6.

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (6)$$

3.8 Resultados e Discussões

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

Cada teste foi realizado 5 vezes e em cada um os conjuntos de treino e teste foram redistribuídos aleatoriamente. A Tabela 1 apresenta a média e o desvio padrão dos resultados obtidos pelos diferentes modelos de cor e arquiteturas avaliadas. Foi verificado que o modelo de cor RGB junto com a arquitetura da U-net que possui 5 camadas e 32 filtros produziram os melhores resultados, como pode ser observado na tabela 1. Acreditamos que o modelo de cor RGB foi o mais adequado nessa aplicação devido as regiões de ovos possuírem um maior contraste e por serem mais facilmente visíveis nas imagens [1], como pode ser observado na Figura 6. O melhor resultado utilizando esse modelo apresentou IoU de 81,10% e Dice de 89,50%, essas foram as duas métricas com maior peso nas escolhas da arquitetura e dos parâmetros do pós-processamento, pois elas medem a qualidade das regiões segmentadas.

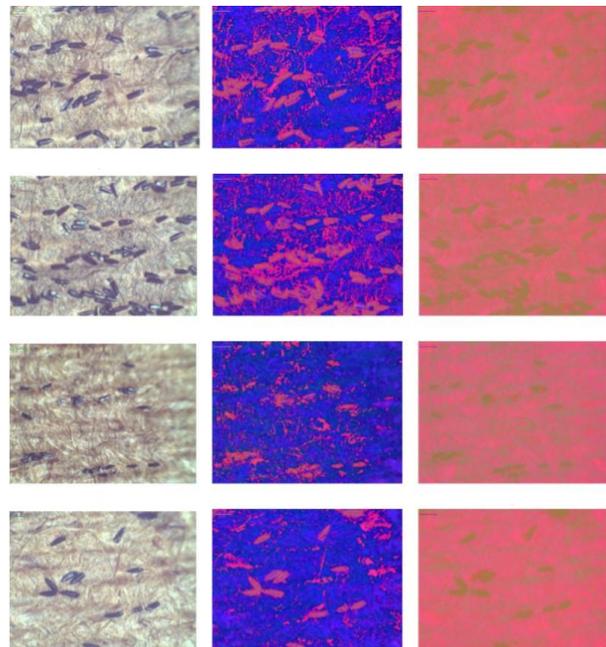


Figure 6: RGB (esquerda) HSV (centro) e LAB (direita).

Após a definição do melhor modelo de cor e da melhor arquitetura da U-net, é feita a seleção do melhor tamanho para

Table 1: Resultado e desvio padrão avaliando os diferentes modelos de cor e as configurações de arquitetura

Sistema de cor	Arquitetura da U-net	IoU	Dice	Precisão	Recall	Acurácia
RGB	3 camadas e 16 filtros	78,98 ± 02,48	88,12 ± 01,65	88,47 ± 03,60	88,33 ± 04,33	98,55 ± 00,35
	3 camadas e 32 filtros	78,05 ± 02,08	87,51 ± 01,44	85,99 ± 05,19	90,06 ± 04,29	98,95 ± 00,20
	3 camadas e 64 filtros	76,34 ± 09,14	86,09 ± 06,54	88,88 ± 03,36	85,36 ± 12,82	98,91 ± 00,32
	4 camadas e 16 filtros	77,00 ± 03,17	86,92 ± 02,01	89,58 ± 02,36	84,84 ± 04,23	98,64 ± 00,27
	4 camadas e 32 filtros	79,90 ± 01,77	88,75 ± 01,15	88,03 ± 03,65	90,01 ± 03,23	98,85 ± 00,17
	4 camadas e 64 filtros	79,31 ± 02,14	88,32 ± 01,42	89,15 ± 01,59	87,79 ± 03,22	98,71 ± 00,22
	5 camadas e 16 filtros	79,29 ± 01,29	88,35 ± 00,86	88,89 ± 03,13	88,18 ± 02,81	98,94 ± 00,12
	5 camadas e 32 filtros	81,10 ± 02,32	89,50 ± 01,49	91,34 ± 00,65	87,91 ± 02,79	98,54 ± 00,27
	5 camadas e 64 filtros	78,74 ± 02,39	87,99 ± 01,56	87,38 ± 03,09	89,05 ± 04,33	98,80 ± 00,35
HSV	3 camadas e 16 filtros	79,28 ± 02,97	88,07 ± 02,51	88,83 ± 02,80	88,42 ± 04,66	98,78 ± 00,33
	3 camadas e 32 filtros	80,61 ± 01,21	89,22 ± 00,79	89,21 ± 02,57	89,61 ± 02,36	98,83 ± 00,32
	3 camadas e 64 filtros	78,03 ± 02,29	87,33 ± 01,74	88,62 ± 03,20	86,93 ± 04,79	99,00 ± 00,23
	4 camadas e 16 filtros	76,05 ± 03,37	86,11 ± 02,42	87,30 ± 01,56	85,33 ± 03,72	98,63 ± 00,28
	4 camadas e 32 filtros	76,59 ± 03,87	86,41 ± 02,72	87,49 ± 03,80	86,12 ± 04,76	98,82 ± 00,13
	4 camadas e 64 filtros	80,53 ± 01,87	89,03 ± 01,46	89,27 ± 02,07	89,34 ± 04,05	98,80 ± 00,30
	5 camadas e 16 filtros	78,51 ± 02,18	87,84 ± 01,42	86,31 ± 02,40	89,86 ± 03,15	98,89 ± 00,33
	5 camadas e 32 filtros	79,84 ± 00,54	88,73 ± 00,34	86,76 ± 01,38	91,13 ± 01,35	98,76 ± 00,15
	5 camadas e 64 filtros	77,66 ± 02,99	87,14 ± 02,10	87,16 ± 01,87	88,02 ± 04,05	98,90 ± 00,26
LAB	3 camadas e 16 filtros	75,44 ± 08,17	85,50 ± 06,03	89,88 ± 03,48	83,07 ± 10,59	98,41 ± 00,54
	3 camadas e 32 filtros	79,31 ± 01,42	88,37 ± 00,95	89,76 ± 01,29	87,30 ± 02,37	98,65 ± 00,42
	3 camadas e 64 filtros	74,83 ± 08,29	84,45 ± 07,33	87,79 ± 01,81	84,07 ± 09,94	98,76 ± 00,28
	4 camadas e 16 filtros	77,35 ± 04,02	86,93 ± 03,06	91,09 ± 01,40	83,75 ± 05,39	98,88 ± 00,16
	4 camadas e 32 filtros	77,03 ± 05,36	86,37 ± 04,04	89,27 ± 02,12	85,27 ± 07,89	98,75 ± 00,35
	4 camadas e 64 filtros	78,78 ± 03,32	87,86 ± 02,33	88,96 ± 01,87	87,34 ± 04,41	98,83 ± 00,28
	5 camadas e 16 filtros	75,61 ± 04,83	85,67 ± 03,46	91,51 ± 02,22	81,69 ± 06,73	98,37 ± 00,15
	5 camadas e 32 filtros	79,12 ± 02,29	88,21 ± 01,55	89,50 ± 02,45	87,62 ± 03,98	98,90 ± 00,57
	5 camadas e 64 filtros	77,39 ± 02,51	86,91 ± 01,92	90,00 ± 02,33	84,88 ± 04,39	98,57 ± 00,27

o elemento estruturante do pós processamento. A Tabela 2 mostra a média e o desvio padrão dos resultados obtidos utilizando diferentes valores de raio para o elemento estruturante disco. O melhor resultado foi obtido com o tamanho de raio 7, pois para valores pequenos, os ruídos maiores que o elemento estruturante não são removidos, e também as regiões que ficaram separadas por uma distancia maior que o elemento estruturante não são conectadas. Já para valores grandes, podem ser removidos regiões de ovos que foram segmentadas separadamente, ou ovos menores que o elemento estruturante. O melhor resultado obtido após o pós-processamento foi com valor de raio 7 do elemento estruturante, com IoU de 81,90% e Dice de 89,96%. Percebe-se que a etapa de pós-processamento melhorou os resultados obtidos na etapa anterior.

A Tabela 3 apresenta três contagens, a manual realizada por um especialista, outra usando o método proposto para a contagem porém utilizando as máscaras que foram segmentadas manualmente e, por fim, usando as regiões segmentadas pelo método proposto. Vale destacar que para cada uma das 5 imagens apresentadas na Tabela 3, são segmentadas 12 imagens de tamanho 256×256 e depois elas são concatenadas para formar a imagem original segmentada.

Pela análise da Tabela 3 percebe-se que o erro médio obtido na contagem dos ovos foi de 4, ou seja, para cada imagem a diferença entre a quantidade real e a estimada é em média 4 ovos. É uma média aceitável, porém existem erros individuais por imagem que são bem maiores que essa média. Isso acontece devido a variação do tamanho dos

ovos, já que a estimação se baseia no tamanho médio do ovo. Na base de imagens existem ovos que podem ser até 8x maiores que outros. Esse resultado pode ser confirmado quando é analisado o erro médio obtido utilizando as máscaras segmentadas por um especialista, que foi de 4,8, sendo inclusive maior que a obtida utilizando a segmentação do método proposto. Embora o uso da área média das regiões segmentadas seja a estratégia mais utilizada na literatura para a estimação da contagem de ovos, os resultados apresentados neste trabalho confirmam que essa estratégia pode não ser tão eficiente para estimar a quantidade de ovos na palheta, principalmente quando existe uma grande variação no tamanho dos ovos.

3.9 Comparação com o estado da arte

A Tabela 4 mostra o resultado do método proposto em comparação com [38]. Como pode ser visto, o método proposto foi melhor em todas as métricas avaliadas. O principal motivo para esses melhores resultados se deve a busca de parâmetros realizada, tanto na escolha da melhor arquitetura, como nos melhores parâmetros do pós-processamento.

4. CONCLUSÃO

Neste trabalho foi apresentada uma metodologia computacional que visa auxiliar especialistas na contagem de ovos nas palhetas das ovitrampas. Como podemos observar pela análise dos resultados, o modelo de cor RGB foi superior ao HSV e LAB, provavelmente por apresentar um maior contraste entre as regiões de ovos e o fundo da imagem. Além

Table 2: Resultado e desvio padrão avaliando os diferentes valores de raio do elemento estruturante disco para o pós-processamento

Melhor arquitetura da U-net	Tamanho do raio do disco	IoU	Dice	Precisão	Recall	Acurácia
5 camadas e 32 filtros	3	80,94 ± 01,32	89,43 ± 00,82	91,45 ± 02,08	87,88 ± 03,14	99,07 ± 00,32
	5	78,87 ± 01,52	88,09 ± 00,93	90,32 ± 02,77	86,44 ± 03,31	98,53 ± 00,40
	7	81,90 ± 01,93	89,96 ± 01,20	90,01 ± 01,88	90,12 ± 03,71	98,80 ± 00,32
	9	77,91 ± 01,89	87,39 ± 01,33	90,95 ± 01,20	84,84 ± 02,49	98,81 ± 00,20
	11	78,21 ± 06,46	87,43 ± 04,47	85,45 ± 08,51	91,03 ± 04,87	98,75 ± 00,62

Table 3: Estimativa da quantidade de ovos das imagens de teste

Imagem	Contagem Manual	Estimativa utilizando as máscaras do especialista	Estimativa utilizando as máscaras preditas pelo método
1	9	9	12
2	45	55	52
3	5	8	10
4	11	12	10
5	68	78	72

Table 4: Comparação com o estado da arte

Trabalho	IoU	Dice	Precisão	Recall	Acurácia
[38]	64,29 ± 05,62	77,66 ± 04,42	83,36 ± 03,62	75,54 ± 09,12	98,65 ± 00,30
Método proposto	81,90 ± 01,93	89,96 ± 01,20	90,01 ± 01,88	90,12 ± 03,71	98,80 ± 00,32

disso, as técnicas de pós-processamento utilizadas melhoraram o desempenho do método proposto, que foi obtido com a rede U-net com 5 camadas e 32 filtros.

As técnicas de contagem baseadas em segmentação apresentam limitações devido a diferença de tamanho dos ovos, proximidade entre os mesmos [28], dentre outros fatores. Com isso, uma alternativa pode ser o uso de técnicas que estimem a quantidade de ovos na imagem sem a necessidade de segmentação. Dessa forma, como trabalho futuro iremos investigar técnicas de contagem que não necessitem da etapa de segmentação. Como exemplo, existem técnicas que estimam quantidade de objetos baseada em mapas de densidade criados por redes convolucionais. Outra possibilidade, consiste no uso de técnicas de detecção de regiões, nestas técnicas redes convolucionais são treinadas para detectar objetos ao invés de segmentar pixels na imagem.

5. REFERENCES

- [1] M. Angulakshmi and G. L. Priya. Brain tumour segmentation from mri using superpixels based spectral clustering. *Journal of King Saud University - Computer and Information Sciences*, 2018.
- [2] F. H. Araujo, R. R. Silva, F. N. Medeiros, D. D. Parkinson, A. Hexemer, C. M. Carneiro, and D. M. Ushizima. Reverse image search for scientific data within and beyond the visible spectrum. *Expert Systems with Applications*, 109:35–48, 2018.
- [3] F. H. Araujo, R. R. Silva, D. M. Ushizima, M. T. Rezende, C. M. Carneiro, A. G. Campos Bianchi, and F. N. Medeiros. Deep learning for cell image segmentation and ranking. *Computerized Medical Imaging and Graphics*, 72:13–21, 2019.
- [4] F. H. Araújo, A. M. Santana, and P. de A. Santos Neto. Using machine learning to support healthcare professionals in making preauthorisation decisions. *International Journal of Medical Informatics*, 94:1 – 7, 2016.
- [5] S. Armato III et al. The lung image database consortium (lidc) and image database resource initiative (idri): A completed reference database of lung nodules on ct scans. *Medical Physics*, 38:915–931, 01 2011.
- [6] A. Baazaoui, W. Barhoumi, A. Ahmed, and E. Zagrouba. Modeling clinician medical-knowledge in terms of med-level features for semantic content-based mammogram retrieval. *Expert Systems with Applications*, 94:11 – 20, 2018.
- [7] S. Bandong and E. Joelianto. 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, 2019.
- [8] V. C. Barros, A. C. L. Pacheco, L. M. Feitosa, J. S. Figueredo, F. Batista, I. P. Lima, and O. A. A. Barbosa. *Produtos naturais no combate ao mosquito Aedes aegypti*. Campinas: Atomo, 2019.
- [9] R. Boulic and O. Renault. 3d hierarchies for animation. In N. Magnenat-Thalmann and D. Thalmann, editors, *New Trends in Animation and Visualization*. John Wiley & Sons Ltd., 1991.
- [10] E. D. Carvalho, A. O. Filho, R. R. Silva, F. H. Araujo, J. O. Diniz, A. C. Silva, A. C. Paiva, and M. Gattass. Breast cancer diagnosis from histopathological images using textural features and cbir. *Artificial Intelligence in Medicine*, 105:101845, 2020.
- [11] M. Claro, L. Santos, W. Silva, F. AraUjo, A. Dalilia, and A. Santana. Automatic detection of glaucoma using disc optic segmentation and feature extraction. In *2015 Latin American Computing Conference*

- (CLEI), pages 1–7, Oct 2015.
- [12] M. A. B. D. A. R. E. DA SILVA, M. G. N. M.; RODRIGUES. Sistema de aquisição e processamento de imagens de ovitrapas para o combate a dengue. *Rev. Bras. Eng. Biom*, 28, 2012.
- [13] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893 vol. 1, June 2005.
- [14] J. A. de Sousa et al. Contagem automática de ovos de mosquito da dengue em imagem de ovitrapa. (22):1–7, 2014.
- [15] L. N. Feitosa. Sistema de contagem automática de ovos do aedes aegypti a partir de processamento de imagens das palhetas de ovitrapas., 2015. Monografia, Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte.
- [16] P. P. R. Filho, E. de S. Rebouças, L. B. Marinho, R. M. Sarmento, J. M. R. Tavares, and V. H. C. de Albuquerque. Analysis of human tissue densities: A new approach to extract features from medical images. *Pattern Recognition Letters*, 94:211 – 218, 2017.
- [17] J. Gaburro, J.-B. Duchemin, P. Paradar, S. Nahavandi, and A. Bhatti. Assessment of icount software, a precise and fast egg counting tool for the mosquito vector aedes aegypti. *Parasites Vectors*, 9:1–9, Dezembro 2016.
- [18] P. S. C. Garcia et al. Acquisition of digital images and identification of aedes aegypti mosquito eggs using classification and deep learning. (22):1–7, 2019.
- [19] P. S. C. Garcia, R. Martins, G. L. L. M. Coelho, and G. Cámara-Chávez. 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, 2019.
- [20] M. Glasser and C. Gomes. Infestação do estado de são paulo por aedes aegypti e aedes albopictus. *Revista de Saúde Pública*, 34:570–577, Dezembro 2000.
- [21] M. Glasser and C. Gomes. Infestação do estado de são paulo por aedes aegypti e aedes albopictus. *Revista de Saúde Pública*, 34:570–577, Dezembro 2018.
- [22] D. Gubler. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. *Trends in microbiology*, 10:100–3, Março 2018.
- [23] K. Hanbay and M. F. Talu. A novel active contour model for medical images via the hessian matrix and eigenvalues. *Computers and Mathematics with Applications*, 75(9):3081 – 3104, 2018.
- [24] M. Herlihy. A methodology for implementing highly concurrent data objects. *ACM Trans. Program. Lang. Syst.*, 15(5):745–770, November 1993.
- [25] I. K. Kallel, S. Almouahed, B. Solaiman, and E. Bosse. An iterative possibilistic knowledge diffusion approach for blind medical image segmentation. *Pattern Recognition*, 78:182 – 197, 2018.
- [26] A. Khatami, M. Babaie, A. Khosravi, H. Tizhoosh, and S. Nahavandi. Parallel deep solutions for image retrieval from imbalanced medical imaging archives. *Applied Soft Computing*, 63:197 – 205, 2018.
- [27] S. Kido, Y. Hirano, and N. Hashimoto. Detection and classification of lung abnormalities by use of convolutional neural network (cnn) and regions with cnn features (r-cnn). pages 1–4, 01 2018.
- [28] D. E. Knuth. *The T_EX Book*. Addison-Wesley, 15th edition, 1984.
- [29] Y. Lecun et al. Gradient-based learning applied to document recognition. In *Proc. of the IEEE*, pages 2278–2324, 1998.
- [30] E. D. LONG, J.; SHELHAMER. Fully convolutional networks for semantic segmentation. *IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 28:3174 – 3179, 2015.
- [31] D. S. Luz, T. J. Lima, R. R. Silva, D. M. Magalhaes, and F. H. Araujo. Automatic detection metastasis in breast histopathological images based on ensemble learning and color adjustment. *Biomedical Signal Processing and Control*, 75:103564, 2022.
- [32] G. M. Marques, G. Galante, J. V. Canabarro, and M. S. Oyamada. Método para contagem automática de ovos recolhidos em armadilhas do tipo ovitrapas, agosto 2017. Google Patents.
- [33] A. Masood et al. Image segmentation of ovitrapas for automatic counting of aedes aegypti eggs. *Journal of Biomedical Informatics*, 79:117 – 128, 2018.
- [34] P. N. Miranda. Contagem automática de ovos de aedes aegypti em imagens de ovitrapas, 2009. Dissertação, Escola Politécnica de Pernambuco - Universidade Federal de Pernambuco.
- [35] G. Nagarajan, R. Minu, B. Muthukumar, V. Vedanarayanan, and S. Sundarsingh. Hybrid genetic algorithm for medical image feature extraction and selection. *Procedia Computer Science*, 85:455 – 462, 2016. International Conference on Computational Modelling and Security (CMS 2016).
- [36] T. RONNEBERGER, O.; P.FISCHER; BROX. U-net: Convolutional networks for biomedical image segmentation. in: Medical image computing and computer-assisted intervention (miccai). 27:234–241, 2015.
- [37] N. Shrivastava and V. Tyagi. An efficient technique for retrieval of color images in large databases. *Computers and Electrical Engineering*, 46:314 – 327, 2015.
- [38] L. Silva, L. Silva, V. Barros, and F. Araujo. Contagem de ovos do aedes aegypti em palhetas de ovitrapas baseada em deep learning. In *Anais do XIV Encontro Unificado de Computação do Piauí e XI Simpósio de Sistemas de Informação*, pages 128–135, Porto Alegre, RS, Brasil, 2021. SBC.
- [39] A. Smith and B. Jones. On the complexity of computing. In A. B. Smith-Jones, editor, *Advances in Computer Science*, pages 555–566. Publishing Press, 1999.
- [40] A. K. O. Tiba, A. F. R. Araujo, H. F. Bassani, and T. I. Ren. Sistema autônomo de reconhecimento e contagem de ovos de aedes aegypti. In G. d. A. Barreto and J. A. F. Costa, editors, *Anais do 10 Congresso Brasileiro de Inteligência Computacional*, pages 1–8, Fortaleza, CE, 2011. SBIC.
- [41] D. Ushizima, C. Yang, S. Venkatakrishnan, F. Araujo, R. Silva, H. Tang, J. V. Mascarenhas, A. Hexemer, D. Parkinson, and J. Sethian. Convolutional neural

networks at the interface of physical and digital data. In *2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pages 1–12, Oct 2016.

- [42] W. Yussof, M. Man, M. Hitam, a. a. abdul hamid, E. Awalludin, and W. A. Wan Abu Bakar. Wavelet-based auto-counting tool of aedes eggs. pages 56–59, Outubro 2018.
- [43] L. Zheng, Y. Yang, and Q. Tian. SIFT meets CNN: A decade survey of instance retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(5):1224–1244, 2018.