IX International Congress
of Management and Technology
COINTER PDVGT 2025

THRESHOLDING-BASED COMPUTER VISION TECHNIQUES ON AUTOMATIC COUNTING OF AEDES MOSQUITOES EGGS

TÉCNICAS DE VISÃO COMPUTACIONAL BASEADAS EM LIMIARIZAÇÃO NA CONTAGEM AUTOMÁTICA DE OVOS DO MOSQUITO AEDES

TÉCNICAS DE VISIÓN POR COMPUTADORA BASADAS EN UMBRALIZACIÓN PARA EL CONTEO AUTOMÁTICO DE HUEVOS DE MOSQUITOS AEDES

Presentation: Oral Communication

Fábio de L. F. Papais [1]; Mayanne Gomes da Silva[2]; Albert Alvin B. S[3];João Pedro P. B[4]; Advisor: Meuse Nogueira de O. Jr.

DOI:https://doi.org/10.31692/2596-0857.IXCOINTERPDVGT.0172

RESUMO

Os mosquitos Aedes são vetores reconhecidos de arbovírus como dengue, chikungunya, febre amarela e Zika, representando uma ameaça persistente à saúde pública em regiões tropicais e subtropicais. Programas de vigilância monitoram frequentemente a presença de fêmeas grávidas por meio da contagem de ovos depositados em paletes em ovitrampas. Tradicionalmente, esse processo de contagem é realizado manualmente sob um microscópio, o que consome tempo, exige muita mão de obra e é suscetível a erros humanos e inconsistências. Para lidar com essas limitações, vários sistemas automáticos de contagem de ovos foram propostos nos últimos anos, particularmente aqueles baseados em processamento de imagens e técnicas de limiarização. Neste estudo, conduzimos uma avaliação comparativa de 13 algoritmos distintos de contagem automática de ovos baseados em limiarização. Nossa análise considerou não apenas a contagem total de ovos previstos versus ovos reais, mas também empregou Precisão e Recall como métricas de desempenho primárias, fornecendo uma avaliação mais confiável da qualidade da detecção. O melhor desempenho foi alcançado utilizando o método de Limiarização de Histograma Balanceado em combinação com operações morfológicas extensivas; no entanto, todas as abordagens baseadas em limiarização testadas permanecem insuficientes para uma contagem automática consistentemente precisa em condições reais. Diversas características da imagem foram identificadas como fontes significativas de erro: sujeira e detritos em paletes devido à exposição ambiental; reflexos de luz causando ofuscamento na superfície dos ovos; e aberrações de foco resultantes da deformação ou empenamento de paletes após exposição à água. Essas descobertas destacam que, embora as técnicas de limiarização possam oferecer automação parcial, uma contagem de ovos robusta e precisa provavelmente exigirá abordagens híbridas ou baseadas em aprendizado capazes de compensar tais artefatos ambientais e de imagem. Os resultados e as limitações identificadas apresentadas aqui podem orientar o desenvolvimento futuro de algoritmos e aprimorar sistemas automatizados de vigilância entomológica.

[1] Curso Técnico em Eletrônica, Instituto Federal de Pernambuco, flfp@discente.ifpe.edu.br
[2] Curso Técnico em Eletrônica, Instituto Federal de Pernambuco, mgs@discente.ifpe.edu.br
[3] Tecnologia em Análise e Desenvolvimento de Sistemas, Instituto Federal de Pernambuco, aabs2@discente.ifpe.edu.br
[4] Tecnologia em Análise e Desenvolvimento de Sistemas, Instituto Federal de Pernambuco, jppb@discente.ifpe.edu.br
[5] Doutor em Ciência da Computação, Instituto Federal de Pernambuco, meusejunior@recife.ifpe.edu.br

Palavras-Chave: Contagem de ovos de mosquito, Aedes aegypti, Algoritmos de limiarização, Processamento de imagens.

RESUMEN
Los mosquitos Aedes son vectores reconocidos de arbovirus como el dengue, el chikunguña, la fiebre amarilla y el zika, lo que representa una amenaza persistente para la salud pública en las regiones tropicales y subtropicales. Los programas de vigilancia frecuentemente monitorean la presencia de hembras grávidas mediante el conteo de huevos depositados en palés en ovitrampas. Tradicionalmente, este proceso de conteo se realiza manualmente bajo un microscopio, lo cual es lento, laborioso y susceptible a errores humanos e inconsistencias. Para abordar estas limitaciones, en los últimos años se han propuesto diversos sistemas automáticos de conteo de huevos, en particular aquellos basados en el procesamiento de imágenes y técnicas de umbralización. En este estudio, realizamos una evaluación comparativa de 13 algoritmos distintos de conteo automático de huevos basados en umbralización. Nuestro análisis consideró no solo el conteo total de huevos predichos versus reales, sino que también empleó la Precisión y la Recall como métricas de rendimiento principales, proporcionando una evaluación más confiable de la calidad de la detección. El mejor rendimiento se logró utilizando el método de Umbralización de Histograma Balanceado en combinación con extensas operaciones morfológicas; sin embargo, todos los enfoques basados en umbralización probados siguen siendo insuficientes para un conteo automático consistente y preciso en condiciones reales. Se identificaron varias características de la imagen como fuentes significativas de error: suciedad y residuos en los palés debido a la exposición ambiental; reflejos de luz que causan deslumbramiento en la superficie de los huevos; y aberraciones de enfoque resultantes de la deformación o alabeo de los palés tras la exposición al agua. Estos hallazgos resaltan que, si bien las técnicas de umbralización pueden ofrecer una automatización parcial, un conteo de huevos robusto y preciso probablemente requerirá enfoques híbridos o basados en el aprendizaje capaces de compensar estos artefactos ambientales y de imagen. Los resultados y las limitaciones identificadas que se presentan aquí pueden guiar el desarrollo futuro de algoritmos y mejorar los sistemas automatizados de vigilancia entomológica.
Palabras Clave: Conteo de huevos de mosquito, Aedes aegypti, Algoritmos de umbralización, Procesamiento de imágenes.

ABSTRACT
Aedes mosquitoes are recognized vectors of arboviruses such as dengue, chikungunya, yellow fever, and Zika, representing a persistent public health threat in tropical and subtropical regions. Surveillance programs frequently monitor the presence of gravid females by counting eggs deposited on pallets in Ovitraps. Traditionally, this counting process is performed manually under a microscope, which is time-consuming, labor-intensive, and susceptible to human error and inconsistencies. To address these limitations, various automatic egg counting systems have been proposed in recent years, particularly those based on image processing and thresholding techniques. In this study, we conducted a comparative evaluation of 13 distinct automatic egg counting algorithms relying on thresholding. Our analysis considered not only the total counts of predicted versus actual eggs but also employed Precision and Recall as primary performance metrics, providing a more reliable assessment of detection quality. The best performance was achieved using the Balanced Histogram Thresholding method in combination with extensive morphological operations; however, all tested thresholding-based approaches remain insufficient for consistently accurate automatic counting under real-world conditions. Several image characteristics were identified as significant sources of error: dirt and debris on pallets due to environmental exposure; light reflections causing egg surface glare; and focus aberrations resulting from pallet warping or deformation after water exposure. These findings highlight that while thresholding techniques can offer partial automation, robust and accurate egg counting will likely require hybrid or learning-based approaches capable of compensating for such environmental and imaging artifacts. The results and identified limitations presented here can guide future algorithm development and improve automated entomological surveillance systems.
Keywords: Mosquito egg counting, Aedes aegypti, Thresholding algorithms, Image processing.

INTRODUCTION

Aedes mosquitoes can transmit a diversity of different arboviruses, such as dengue, chikungunya, yellow fever, and Zika (CDC, 2021). Aedes are also typical urban mosquitoes and can easily find breeding sites in domestic and surrounding areas, such as open tires and potted plants (Kraemer et al., 2019). Since their impact on the population, those mosquitoes, especially Aedes and Aedes albopictus are a relevant matter of public health studies and spread control campaigns in affected regions, usually tropical and subtropical countries.

Because of its great breeding potential, it is necessary to control the population of those mosquitoes to prevent outbreaks, endemic conditions, and even an epidemic. It is important for health authorities to estimate the population of mosquitoes in a certain area, and Ovitraps are the standard way of monitoring the presence of Aedes mosquitoes (Focks et al., 2003). Ovitraps are small, usually plastic and dark containers filled with water (and sometimes an infusion), that attract mosquitoes and can be used to detect the presence of gravid female mosquitoes (Reiter; Colon, 1991). The ovitrap also contains a small "palette" made of a substrate (e.g., wood or cloth) where the mosquitoes lay their eggs. After 5-7 days, the ovitraps are collected, and their pallets are gathered for the examination of the number of eggs within the sample using a microscope.

Egg's counting is done manually in the laboratory, and, since some pallets may have more than hundreds or thousands of eggs, the counting process is susceptible to many errors and inconsistencies, mainly while moving the microscope. Moreover, counting mosquitoes' eggs is a tedious and time-demanding task for the laboratory personnel. Technologies have been developed to automatize this process, usually by building image scanning hardware to get precise microscope images from the pallets and developing a computer vision algorithm to automatically count the eggs in the taken images (da Silva; Rodrigues; de Araújo, 2011) (Gaburro et al., 2016) (Brasil et al., 2015) ideally resulting in more precise and fast counting.

In this context, this article aims to approach the difficulties presented in the computer vision algorithm. For this, it analyzes methods, algorithms, and techniques for performing the image processing of the Ovitrap palette to, in addition to assisting in the counting of eggs, amplify the accuracy of the data obtained. This paper compares common thresholding-based approaches to the Aedes mosquitoes eggs counting problem, showing results of the variation of a simple base algorithm tested in more than 2000 images. Those results are based on the comparison of automatic counting with a gold standard reference of each pallet image, including metrics to evaluate the algorithm's accuracy, precision, and recall (not only the predicted quantity of eggs per pallet). Throughout this study, it will be possible to understand which directions may or may not be the most relevant to make thresholding-based algorithms more reliable and consistent, which will result in better integration of automated counting data and will allow control actions to be carried out on time to prevent mosquitoes from worsening. This task may seem simple and well-defined, but when dealing with real-world images, problems such as dirt and palette deformation may affect algorithms' assertiveness.

THEORETICAL FOUNDATION

The presence of the Aedes aegypti mosquito in Brazil is a persistent threat to public health, as it is the main vector of arboviruses such as dengue, chikungunya, and Zika. The reemergence of dengue in recent decades, with epidemics affecting all parts of the country, has intensified the need for more effective control strategies. To combat the spread of mosquitoes, various methods are used to monitor the vector, including larval collection, adult mosquito capture, and ovitraps, which are effective in detecting the presence of mosquitoes even in low-density areas. After being placed in strategic locations, such as homes, parks, and vacant lots, the ovitraps are collected and taken to a laboratory for analysis and counting of the eggs present in the traps. This strategy is mainly used for population control of arbovirus-transmitting vectors. By counting the eggs present in an ovitrap, it is possible to estimate the most vulnerable regions and study why these locations are more conducive to the proliferation of Aedes. Currently, the eggs present in the ovitraps are counted manually, making the work slow, susceptible to errors, and difficult to apply on a large scale.

Computer vision acts as an ally in optimizing this process. Computer vision, which models and replicates human vision using software and hardware, uses artificial intelligence and machine learning to make sense of images. These capabilities allow the machine to perform tasks such as image classification (identifying what the image represents), object detection (locating and detecting specific objects), and instance segmentation (isolating each object individually). But before applying computer vision techniques, there is the image processing stage (known as image preprocessing), acting as a preparatory tool for the inference performed by computer vision. Image processing is the manipulation of images using algorithms that alter and optimize them for analysis. This includes converting a color image to grayscale, applying filters to remove noise, adjusting brightness and contrast, and enhancing

ISSN: 2596-0857
PREFIXO DOI: 10.31692/2596-0857

sharpness or softening the image. This makes the image clearer and more suitable for the next stage of analysis. In short, both techniques work together, complementing each other in the field of machine image analysis. Machine processing modifies the image, and computer vision analyzes and extracts information. This combination allows automated systems to process large volumes of data accurately and efficiently, overcoming the limitations of manual processing.

One of the most widely used techniques is thresholding, a computer vision and image processing process that groups pixels into two distinct classes: objects and background, using a color or intensity threshold. This process typically converts grayscale images to binary (black and white) images, where 1 can represent white and black is represented by 0 (or vice versa, depending on the application). The main challenge in thresholding is automatically defining the ideal threshold for each application. To overcome this limitation, several different techniques have been developed, with the Otsu algorithm being one of the most widely used for global thresholding. This algorithm automatically determines a threshold based on the image histogram. This technique is used to convert the original image into a binary image, preprocess the region of interest to remove smaller components that may impair pixel recognition, and, from the binary image, calculate the image histogram to segment the region of the image containing the objects of interest.

Morphological operations are also used to process images. They are a class of nonlinear operations applied to binary or grayscale images, which aim to modify the image structure to remove noise and extract the shape of objects. The essence of these operations is to perform a logical operation on the pixels, using an element that defines the object's structure. This element functions as a mask that defines which pixels will be considered in the operation. The two most fundamental morphological operations are dilation and erosion. Dilation enhances the edges of objects and helps fill in small holes. It uses the logical OR operator, expanding the foreground areas of the image. Erosion, on the other hand, removes fine details and shrinks the foreground areas. It uses the rule that pixel color changes only if all pixels in the analysis area correspond to the structuring element. From these two basic operations, it is possible to create more complex operations, such as opening, which consists of erosion followed by dilation, and closing, which is dilation followed by erosion. Morphological Gradient—Is the difference between the dilated image and the eroded image. Morphological operations are essential in the field of computer vision for image preprocessing, allowing the

ISSN: 2596-0857
PREFIXO DOI: 10.31692/2596-0857

extraction of useful information for preparing data for machine learning algorithms.

METHODOLOGY

This study adopts a quantitative and experimental approach to investigate the effectiveness of thresholding-based computer vision techniques in the automatic counting of Aedes mosquito eggs. The research is based on the analysis of real-world entomological data, using a controlled laboratory environment to test and compare various algorithmic variations. This section details the procedures, instruments, and evaluation criteria employed throughout the research, providing a comprehensive basis for the interpretation and discussion of the results.

1. Pallets image capture and dataset

   In order to retrieve substantial data, this study used 8 different pallets previously used in real entomology studies at the city of Recife - Pernambuco, Brazil; these samples were kindly provided by Fiocruz-PE (Oswaldo Cruz Foundation), a recognized institution of science and technology in Latin America (Fiocruz, 2021).

   We used a treadmill system that performs automatic image captures of the ovitraps pallets through a microscope. For this, the robotic prototype uses a structure of servo mechanisms from old Inkjet printers, which also enables the recycling of electronic waste. The image capturing process is performed by a low-cost digital microscope, featuring a gain of x1000. This digital microscope also has a built-in LED light around its sensor. pallets have a dimension of [15 x 4.5] cm, and the printhead is used to tour the microscope along the X and Y axes of the palette, forming a microscopic array of 730 elements. In this process, the generated images have a dimension of 640 x 480 pixels and a resolution of 96 DPI. The capture process generates a small set of distorted images without eggs from palette edges, which were manually filtered. In the end, 2527 images were used for data processing and analysis.

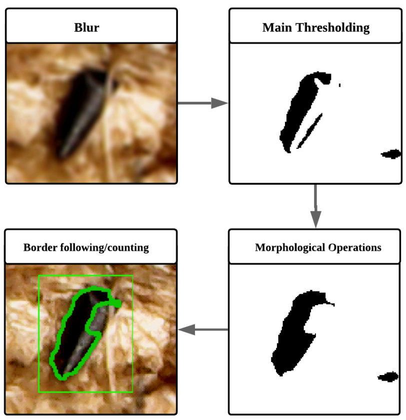
   In order to build a golden standard, each one of the retrieved images was analyzed and, in case of containing eggs, got all the eggs positions marked with rectangular bounding boxes. This information was stored in an XML file containing the position and size of each bounding box, per egg, per image. If the image contains eggs, this gold standard specifies the number of eggs, as well as the position of its rectangle bounding boxes at the image's Cartesian plane (upper-left x,y values plus width and height values). This classified XML file was later used to check the accuracy of the algorithms in each prediction of eggs.

2. Base algorithm and variations

Thresholding (often called binarization or image limiarization) is an image processing algorithm that can be used for image segmentation. This process takes a single-channeled image and generates a binary image (that contains only two exact colors) based on a threshold value, which defines the color of the pixel. Thresholding is a simple and established approach in the area of computer vision and image processing, being useful for images where there is enough contrast between the background and the regions of interest (Sezgin; Sankur, 2004). Since *Aedes Aegypti* and *Aedes albopictus* mosquitoes eggs' morphology is constant and easily distinguishable (black, cigar shaped, with slight dorso-ventral curvature (Suman et al., 2011), and Ovitraps pallets usually have lighter colors, thresholding is a clear and accessible option, especially considering the algorithm's cheap computational cost.

To enhance any image processing technique result, pre-processing and post-processing algorithms can also be used. This study's approach adopted blur pre-processing algorithms and morphological operations after main thresholding as post-processing techniques, including border following algorithms for structure (found eggs) counting. The whole process used in this study and its stages can also be visualized in Figure:

Figure 01: Stages of the base algorithm used in this study, visualized with a sample picture of an Aedes egg.



Source: Nogueira de O. et al. (2021).

ISSN: 2596-0857
PREFIXO DOI: 10.31692/2596-0857

Taking the described process as a base algorithm, this study tested several variations in each of its stages, including in the pre-processing and post-processing techniques. These variations included not only testing other recognized and suitable techniques but also varying actual stage parameters.

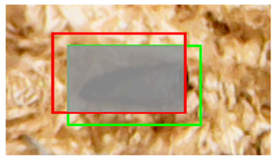
3. Algorithms variations evolutions and metrics

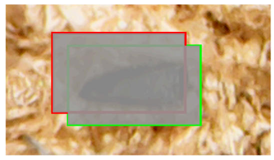
After image processing, another XML file containing the predictions of the respective algorithm is generated, allowing the comparison of the obtained and expected results by an automated software developed specifically for this task.

An initial analysis could lead to a comparison of results using as reference only the quantity of eggs found and the expected quantity per pallet, and, in fact, this is the case in many similar studies (Brun et al., 2020). Even so, it was observed that a lot of dirt and noise from the pallets were counted as actual eggs using this approach, affecting the results and subsequent comparison, since those false positives could be compared to real eggs that were not detected. As an example, an image with an expected number of three eggs can be measured as a correct prediction if the algorithm predicts two real eggs and one dirt in the image, which, manually analyzed, would be considered an error. As it will be further discussed, this kind of error may happen for a diversity of reasons; thus, for accurate measurement of false positives and negatives, it is necessary to verify the position of the eggs found by the algorithm and the expected position of these eggs according to the gold standard. In order to generate error measuring metrics for the variations of the algorithms, this study, in addition to comparing only the quantity of eggs in each image, also verified the correspondence of the positions of the eggs found. The metric used to analyze the correspondence between predicted and expected eggs was the Intersection over Union (IoU) also known as the Jaccard Index (Padilla; Netto; da Silva, 2020). For each correspondence comparison, the union area and intersection area (if any) of the eggs were measured. Based on this information, it was possible to calculate the IoU value of the two eggs , and predicted and expected eggs are considered matching if their $IoU \geq 0.4$.

Figure 02: Visual description of Intersection Over Union, metric used to assert expected and predicted egg's position matches

$$IoU = \frac{\textit{Intersection area}}{\textit{Union area}} =$$





Knowing whether an egg has been truly found or not, we can separate errors into False Positives and False Negatives, respectively excessive egg count and insufficient egg count. As a consequence, we can separate the matches and errors in a Confusion Matrix, as in table. It is important to note that since we are analyzing an object detection algorithm, True Negatives ("not objects") are not classified and can not be measured.

Table 01: Confusion Matrix describing main metrics; note that True Negatives are not considered in our scenario.

		Golden Standard	
Predicted	Egg	True Positives (TP)	False Positives (FP)
	Not Egg	False Negatives (FN)	True Negatives (FN)

Source: Own (2025).

From the values of this Confusion Matrix, we can extract other metrics that allow better evaluation of the algorithms. The metrics used in this study were Precision and Recall. Precision may be interpreted as the probability of predicted eggs being real eggs (with its $IoU > 0.4$), and Recall measures the probability of golden standard eggs being correctly predicted. High precision and low recall would mean that most predictions were correct, but there were many False Positives; similarly, low precision and high recall would mean that most predictions were correct, but there were many False Negatives.

ISSN: 2596-0857
PREFIXO DOI: 10.31692/2596-0857

$$PRC = TP / (TP + FP)$$
$$REC = TP / (TP + FN)$$

- PRC (precision): number of true positives in relation to the total number of predicted positives.
- REC (recall): number of true positives in relation to the total number of objects that should have been found.
- TP (True Positives): Number of correct positive predictions
- FP (False Positives): Number of incorrect positive predictions

Accuracy is a useful metric, and it is intuitive since it shows the correct predictions against all other guesses. But it might be misleading, especially when the data set is not balanced.

$$Accuracy = TP / (TP + FP + FN)$$

While analyzing the results, we found out that false positive and false negatives rates varied more than expected between each algorithm, making Accuracy not a reliable metric to compare each algorithm variation.

RESULTS AND DISCUSSION

Algorithm variations were made by taking the original base algorithm steps and characteristics and modifying them to the most common and available alternatives (e.g., standard thresholding to automatic thresholding methods). Even if those changes were made in isolation from others, the contrast in the algorithm's results helps to understand which changes may enhance results. Below, we describe each of the variations we tested:

• **Original**: as described in section methodology.
• **Otsu's Thresholding** (Otsu, 1979): changes standard thresholding to Otsu's thresholding algorithm.
• **Triangle Thresholding** (Zack; Rogers; Latt, 1977): changes standard thresholding to the histogram Triangle Thresholding technique.
• **Balanced Histogram Thresholding** (Dos Anjos; Shahbazkia, 2008): changes standard thresholding to the Balanced Histogram Thresholding (BHT) algorithm.
• **Median Blur** (Huang; Yang; Tang, 1979) (Mastin, 1985): changes Gaussian Blur to

Median Blur.

• **Bilateral Filtering** (Tomasi; Manduchi, 1998): changes Gaussian Blur to Bilateral Filtering, also known as Selective Blur.

• **No Blur**: removes any blur pre-processing.

• **Strong Morphology**: intensified opening and closing morphological operations.

• **Weak Morphology**: reduced opening and closing morphological operations.

• **No Morphology**: removes any morphological operation.

• **Contrast**: increase the image's contrast by 60%.

• **HSV**: converts image to HSV (Hue, Saturation, Value) color space, using V channel (brightness) as input to thresholding.

• **L*a*b***: converts image to L*a*b* color space (also referred to as CIELAB) using the L* channel as input to thresholding.

The goal of these algorithms is to find the correct total number of eggs in an egg palette or a group of palettes. Similarly to other studies, the EQ (estimated quantity) — the total number of eggs found — is also provided. In our dataset, the expected total number of eggs is 2941. Each algorithm was tested using the exact same set of images, and its metrics were extracted with the same evaluation software. In Table 02, we present the most promising results among the tested algorithm variations, based on the selected metrics (precision and recall) and Confusion Matrix values.

Table 02: Best performing results among the tested algorithm variations

Algorithm Variation	Precision	Recall	True positives	False Positives	False Negatives	EQ
Balanced Histogram Thresholding	0.48	0.12	346	378	2594	724
Otsu's Thresholding	0.47	0.15	451	507	2490	958
Median Blur	0.34	0.16	468	924	2473	1392
Weak Morphology	0.32	0.15	456	973	2485	1429
No Morphology	0.32	0.14	413	869	2528	1282

ISSN: 2596-0857
PREFIXO DOI: 10.31692/2596-0857

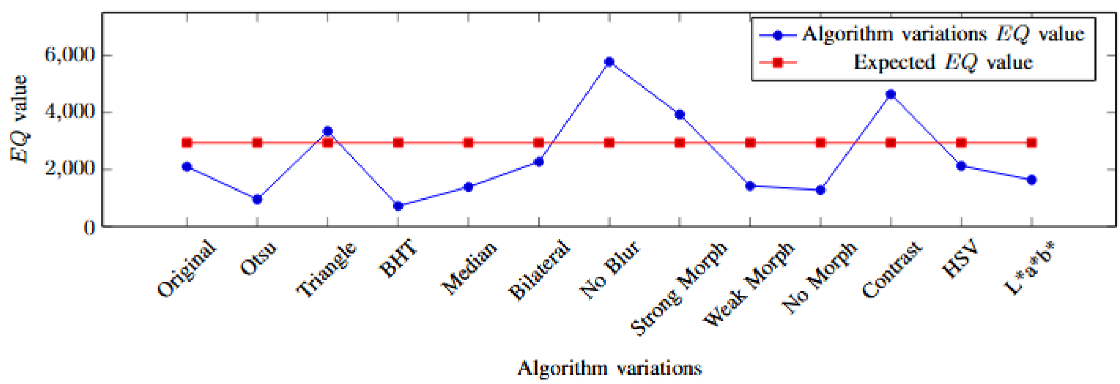
L*a*b	0.31	0.17	503	1136	2438	1639
HSV	0.30	0.21	647	1484	2294	2131
Original	0.26	0.19	554	1551	2387	2105
Strong Morphology	0.19	0.25	740	3194	2201	3934

Source: Own(2025).

Looking at *precision* and *recall* metrics at Table 02, we were able to observe that the proposed thresholding technique still presents results that are far below practical or reliable for their purpose. The highest precision was achieved using the BHT algorithm (*precision = 0.48*), and the highest Recall (*recall = 0.25*) was obtained using the Strong Morphology algorithm.

Similarly to other studies, the total amount of eggs found can also be used as an evaluation metric. Figure 03 shows how the *EQ* value varied in each algorithm, along with its expected result. Although it is intuitive, it is not possible to extract enough relevant information from such a metric by itself, as discussed in the methodology section.

Figure 03: The EQ values for each variation of the algorithm represented visually along with the expected total number of eggs in our dataset.



Source: Own(2025).

Among almost all algorithm variation metrics results, we can observe two main consistencies: *precision* > *recall* and *false negative* > *false positive*. If we calculate the standard deviation of precision and recall, we have that $\sigma precision = 0.125$ and $\sigma recall = 0.033$; in other words, the variation of the algorithms results occurred mainly in the decrease and increase of False Positives, while False Negatives related errors remained more balanced. It shows that the

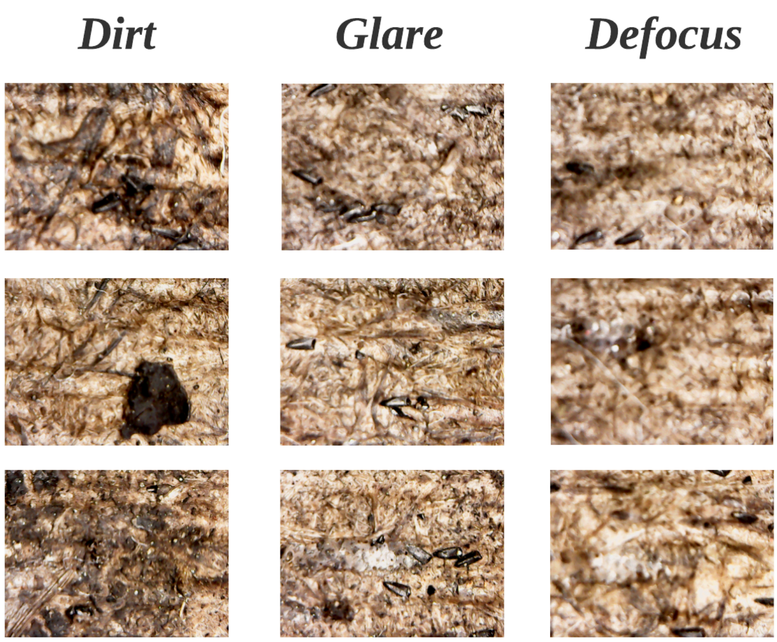
ISSN: 2596-0857
PREFIXO DOI: 10.31692/2596-0857

algorithms are constantly failing to identify certain eggs, producing False Negatives (FN). As if there exists some image feature that makes the algorithms blind to such eggs.

Looking at morphological approaches (No Morphology, Weak Morphology and Strong Morphology), we could note important aspects of the images in our dataset. While increasing the intensity, the morphological operations caused TP to increase, the cases of False Negatives and False Positives also increased. The role of these morphological operations is to eliminate noise and close larger structures that may be eggs. So they are involved in two main problems in our images dataset: dirt and reflected shine.

Dirt is very common on ovitrap pallets since they sit outdoors for days and are susceptible to any kind of environmental conditions and weather, including egg deposition from insects other than mosquitoes *Aedes*. This kind of dirt is usually dark and can confuse the thresholding algorithm, especially when the sensitivity is increased. The reflected shine on the eggs comes from the built-in LED light of the digital microscope; the surface of the eggs reflects this light, making the area of many of these (black) eggs to become partially white. This light interferes with the thresholding, causing only some parts of eggs to be recognized, which are usually ignored in the morphological operations and contouring stages.

Figure 04: Samples of images with common properties that caused errors on algorithm variations.



ISSN: 2596-0857
PREFIXO DOI: 10.31692/2596-0857

Dirt and shine are two hard problems to get around when it comes to thresholding-based algorithms, which have as their main parameter the chosen threshold (in the case of manually chosen thresholding). If we increase the threshold, the algorithm sensitivity increases, causing more eggs to be detected, but, in doing so, we cause more dirt and noise to be interpreted as eggs. The same happens with decreasing the threshold. As we decreased the algorithm's sensitivity in order to diminish the amount of dirt detected, fewer eggs were recognized due to the shine.

Another problem found in the image dataset that caused many errors in the algorithm was the blur on some images. The pallets used in this study were made of wooden material, which stays in contact with water and moisture for days, causing the pallet to expand and warp, even if little. The image capture system follows a straight horizontal line during the scanning of the palette, and the digital microscope used has its focus adjusted only at the beginning of the capture. Since the wooden pallet contains deformations, the distance between it and the microscope undergoes slight variations, which in some cases were enough to cause noticeable defocus aberrations while capturing the images.

Alternatively, there are studies that work with solutions based on Artificial Intelligence, using its learning capabilities to analyze images and capture patterns that enable detection in non-ideal scenarios. Machine Learning is a subfield of Artificial Intelligence that explores the construction and study of algorithms that can learn and make predictions from incoming data. Within the field of Machine Learning, Deep Learning is an approach that stands out for using multi-layer neural networks to learn hierarchical and complex data representations. These networks are capable of capturing intricate patterns and high-level abstractions, especially in tasks involving large volumes of data, such as image processing. Deep Learning has been widely adopted due to its ability to automate feature extraction. Convolutional Neural Networks (CNNs) are part of a specific Deep Learning architecture for image classification and object detection tasks. They are especially effective at processing data with a grid structure, such as images. These networks use convolutional layers to extract local features, such as edges and textures, and subsequent layers to combine these features into more complex representations, such as shapes and objects. Furthermore, CNNs are robust to variations such as translations, rotations, and scales, making them ideal for computer vision tasks.

ISSN: 2596-0857
PREFIXO DOI: 10.31692/2596-0857

CONCLUSIONS

This article aims to analyze and quantify the use of thresholding-based algorithms in the automatic counting of eggs from *Aedes* mosquito ovitraps pallets. Metrics other than the total number of eggs in the dataset were established. We adopted the concepts of True Positives, False Positives and False Negatives, which were individually measured from the results of each algorithm. In addition to these metrics, we calculated Precision and Recall, facilitating the analysis of algorithm variations. We also made available the total number of predicted eggs of each algorithm variation and the expected value, which proved to be an inaccurate way of analyzing the algorithm's results, even with simple computer vision tasks like the one discussed in this article. The chosen set of metrics proved to be useful and reliable for such analysis.

Results showed that the tested thresholding-based algorithms were not enough to develop a satisfactory algorithm, even applying pre-processing and other variations. Thus, such techniques are unsatisfactory for laboratories responsible for counting eggs on ovitraps pallets, for example. The best results among all the variations were the one that used the Balanced Histogram Thresholding technique and the one with intense morphology operations. It is important to note that the changes in the original algorithm were made isolated from each other, as well as its tests; so combining different variations may also increase the results. It was noticed that all the algorithms presented a high number of False Negatives, which may be the subject of new investigations about possible characteristics of the image or eggs causing that many errors. In addition, we were able to identify three main immediate issues in the images that were causing errors: dirt on the plates, reflected shine in the eggs by the light from the image capture tool, and defocus due to warping and physical distortions in the ovitrap pallets. The first two problems, dirt and shine, are especially difficult to get around with thresholding-based algorithms, due to their color contrast characteristic. Although shine can be improved through changes in the image capturing instrument, dirt and defocus are characteristic of pallets used in reality and very present in our image dataset. Approaching such issues is a challenge as well as a subject for future research on better materials choices for pallets and new image processing approaches.

REFERENCES

BRASIL, L. M.; GOMES, M. M. F.; MIOSSO, C. J.; DA SILVA, M. M.; AMVAME-NZE, G. D. Web platform using digital image processing and geographic information system tools: a

ISSN: 2596-0857
PREFIXO DOI: 10.31692/2596-0857

Brazilian case study on dengue. Biomedical Engineering Online, v. 14, n. 1, p. 1–14, 2015.

BRUN, A. L.; MORAES, P. L.; RIZZI, C. B.; RIZZI, R. L. 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.

CENTERS FOR DISEASE CONTROL AND PREVENTION – UNITED STATES DEPARTMENT OF HEALTH AND HUMAN SERVICES. Surveillance and control of Aedes aegypti and Aedes albopictus in the United States. Disponível em: https://www.cdc.gov/mosquitoes/pdfs/mosquito-control-508.pdf.. Acesso em: 08 dez. 2021.

DA SILVA, M. G. N. M.; RODRIGUES, M. A. B.; DE ARAUJO, R. E. Aedes aegypti egg counting system. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011. p. 6810–6812.

DOS ANJOS, A.; SHAHBAZKIA, H. R. Bi-level image thresholding. Biosignals, v. 2, p. 70–76, 2008.

FOCKS, D. A. et al. A review of entomological sampling methods and indicators for dengue vectors. Disponível em: https://apps.who.int/iris/bitstream/handle/10665/68575/TDR_IDE_DEN_03.1.pdf.. Acesso em: 08 dez. 2021.

FUNDAÇÃO OSWALDO CRUZ. The Foundation. 2021. Disponível em: https://portal.fiocruz.br/en/foundation.. Acesso em: 27 ago. 2021.

GABURRO, J.; DUCHEMIN, J. B.; PARADKAR, P. N.; NAHAVANDI, S.; BHATTI, A. Assessment of ICount software, a precise and fast egg counting tool for the mosquito vector Aedes aegypti. Parasites & Vectors, v. 9, n. 1, p. 1–9, 2016.

HUANG, T.; YANG, G. J.; TANG, G. A fast two-dimensional median filtering algorithm. IEEE Transactions on Acoustics, Speech, and Signal Processing, v. 27, n. 1, p. 13–18, 1979.

KRAEMER, M. U. G. et al. Past and future spread of the arbovirus vectors Aedes aegypti and Aedes albopictus. Nature Microbiology, v. 4, n. 5, p. 854–863, 2019.

MASTIN, G. A. Adaptive filters for digital image noise smoothing: An evaluation. Computer Vision, Graphics, and Image Processing, v. 31, n. 1, p. 103–121, 1985.

OTSU, N. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics, v. 9, n. 1, p. 62–66, 1979.

PADILLA, R.; NETTO, S. L.; DA SILVA, E. A. B. A survey on performance metrics for object-detection algorithms. In: International Conference on Systems, Signals and Image Processing (IWSSIP), 2020. p. 237–242.

REITER, P.; COLON, M. A. A. N. Enhancement of the CDC ovitrap with hay infusions for daily monitoring of Aedes aegypti populations. Journal of the American Mosquito Control Association, v. 7, n. 1, p. 52–55, 1991.

SEZGIN, M.; SANKUR, B. Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic Imaging, v. 13, n. 1, p. 146–165, 2004.

SILVA, M. G. N. M.; RODRIGUES, M. A. B.; ARAUJO, R. E. Aedes aegypti egg counting system. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011. p. 6810–6812.

SUMAN, D. S.; SHRIVASTAVA, A. R.; PANT, S. C.; PARASHAR, B. D. Differentiation of Aedes aegypti and Aedes albopictus (Diptera: Culicidae) with egg surface morphology and morphometrics using scanning electron microscopy. Arthropod Structure & Development, v. 40, n. 5, p. 479–483, 2011.

TOMASI, C.; MANDUCHI, R. Bilateral filtering for gray and color images. In: Proceedings of the 6th International Conference on Computer Vision, 1998. p. 839–846.

ZACK, G. W.; ROGERS, W. E.; LATT, S. A. Automatic measurement of sister chromatid exchange frequency. Journal of Histochemistry & Cytochemistry, v. 25, n. 7, p. 741–753, 1977.