A deep learning approach for detection small portions of water in images acquired by drones

Gustavo A. Lima, Rafael O. Cotrin, Gabriel R. Paz, Marcos Alexandruk, Sidnei A. de Araujo

Informatics and Knowledge Management Post-Graduation Program, Universidade Nove de Julho – UNINOVE, Vergueiro street 235/249, São Paulo, Brazil
gustavo.araujo.lima94@gmail.com, rafa25.cotrin@gmail.com, gabrielrpaz11@gmail.com, m.alexandruk@gmail.com, saraujo@uni9.pro.br

Abstract. Drones have been used to automatically identify objects and scenarios (normally water tanks, buckets, plant pots, and other containers contained in open-air trash) that characterize potential breeding sites of mosquito, such as the Aedes aegypt, from the acquired images. However, despite knowing that water stagnation is an essential condition for breeding mosquitoes, computer vision systems proposed in the literature for automatic image analysis do not include the detection of water in suspicious objects and scenarios, which constitutes a technical limitation for the effective use of drones in vector monitoring and control actions. In this work, an approach combining K-means with a convolutional neural network YOLOv8 is proposed to detect and locate small portions of water in images acquired by drones. To carry out the experiments, we composed two datasets of images acquired from simulated scenarios contemplating small containers with and without water in controlled environments. The overall results obtained in computational experiments (accuracy = 0.962; precision = 0.960; sensibility = 0.977) indicates the viability of adopting the proposed approach in existing computer vision systems for automatic identification of mosquito breeding sites, making them more effective for health surveillance actions.

Keywords: Drone, Mosquito, Aedes aegypt, Water, Deep Learning.

1 Introduction

The epidemics of dengue, Zika, chikungunya, and urban yellow fever, transmitted by mosquitoes, have caused significant concern among health authorities in Brazil and worldwide [1]. In 2024, dengue remains a serious public health issue in Brazil, with over 5 million cases reported and approximately 3,000 deaths recorded by May [2]. This significant increase in cases has overwhelmed the healthcare system, leading to high economic and social costs, including hospitalizations and loss of productivity due to absenteeism. Severe complications of dengue, transmitted by Aedes aegypt, such as shock and hemorrhages, increase mortality rates, particularly among the most vulnerable populations. In this context, ongoing efforts in prevention, control, and treatment are essential to mitigate the impacts of this disease in the country [3].

The Brazilian Ministry of Health conducts annual advertising campaigns to combat the Aedes aegypti mosquito, involving state and municipal managers as well as the public. In 2023, the federal government allocated approximately R$13 million for these campaigns [4]. Combating the Aedes aegypti mosquito also requires additional efforts, as informational campaigns and public mobilization do not always achieve the desired results. According to the Brazilian Ministry of Health, in addition to the dengue vaccine and the development of vaccines for other arboviruses, it is important to incorporate new technologies [2].

An alternative has been the use of drones to capture aerial images in regions more vulnerable to mosquito proliferation [5]. However, it is still common for the images captured by drones to be analyzed manually (visually), which is time-consuming for inspections.

On a global scale, there has been a growing number of scientific studies over the last decade proposing computer vision systems (CVS) for the automatic identification of objects and scenarios that characterize potential mosquito breeding sites in aerial images acquired by drones. Notable works in this area include [6-13]. These studies typically employ computer vision (CV) and artificial intelligence (AI) techniques to automatically identify key suspicious objects in the images, such as water tanks, tires, swimming pools, plant pots, gutters, and open inorganic waste. However, despite the understanding that stagnant water is a critical condition for mosquito breeding, the CVS proposed in the literature for automatic image analysis do not include water detection in the suspected objects and scenarios, which constitutes a technical limitation for the effective use of drones in monitoring and combating mosquitoes breeding sites.

2 Theoretical Background

2.1 Aedes aegypti

The Aedes aegypti mosquito is known for its diurnal habits. Inside residences, it tends to shelter in shaded and dark areas. While the male mosquito feeds on plant nectar, the female requires human blood for egg maturation. These eggs are laid separately on the internal walls of objects, usually near the water surface, providing ideal conditions for their survival [14].

According to information from the Ministry of Health, various containers can accumulate water and serve as breeding sites for mosquitoes, such as water tanks, stagnant water in plant pots, tires, and empty bottles. Therefore, it is important to be cautious with all locations that may collect water, as mosquito eggs are resistant to desiccation and can survive in the environment for up to 450 days. Even a small amount of water, such as a small puddle, is sufficient for the larvae to hatch [14]. In this study, the identification of stagnant water considers objects that may accumulate it, such as basins and other containers typically found in inorganic waste left exposed to the elements.

2.2 Water identification in aerial images

In the remote sensing (RS) literature, there are many studies addressing the detection of water bodies (any significant accumulation of water on the earth's surface, such as seas, rivers, and lakes) and other features like soil and vegetation, from aerial images acquired by satellites. There are indices calculated based on the spectral bands available on satellite imaging sensors to identify water, such as the NDWI – Normalized Difference Water Index [15] and IIA – Water Indicator Index [16], which are based on the combination of the visible and near-infrared (NIR) bands.

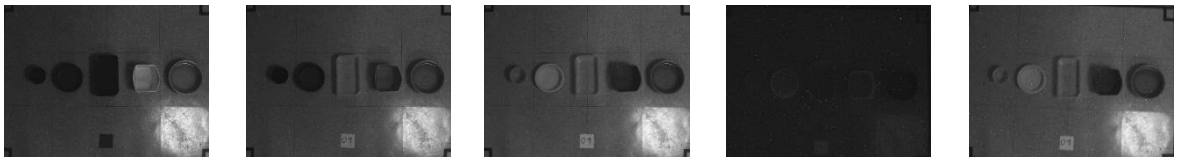
Although signatures (or indices) that characterize patterns associated with water bodies are a common task in remote sensing RS, detecting water contained in small objects from satellite images is challenging due to the low spatial resolution of these images. However, while drones provide high spatial and temporal resolution images, the sensors coupled on them are typically of low spectral resolution, making it difficult to utilize the spectral signatures proposed in the literature. This highlights a research gap that has motivated some recent studies.

Prasad [17] and Bravo et al. [10], for example, proposed approaches for identifying small lakes, swimming pools, fountains, and puddles. De Mesquita et al. [18] developed a drone with an embedded software capable of detecting both stagnant and moving water, whether clean or dirty, which sends notifications to the drone operator when water is detected, allowing for manual pesticide application. The study by Truong et al. [19] aimed to explore the feasibility of identifying temporary water bodies using real-time surveillance images obtained by drones, employing convolutional neural networks based on the YOLOv4 model. However, such approaches are not applicable for detecting water contained in tires, plant pots, gutters, and small containers typically found in open inorganic waste, as they are unable to identify such small volumes of water.

3 Material and methods

3.1 Image acquisition

Two image datasets were created for the experiments. The first dataset contains 108 images captured by a DJI Matrice 200 drone, equipped with a Micasense RedEdge-MX multispectral camera, positioned approximately 2 meters above the targets. Each capture provided an image with the spectral bands Blue (B), Green (G), Red (R), Infrared (NIR), and RedEdge (EDG), as illustrated in Figure 1. The second dataset consists of 51 RGB images captured by various smartphones, at a height of about 1 meter from the targets, as illustrated in Figure 2. Each image represents a scenario contemplating objects with different storage capacities (ranging from 0.1 to 12 liters), which may or may not contain water, and the water could be either clean or dirty.



(a) B (b) G (c) R (d) NIR (e) EDG

Figure 1. Spectral bands of an image of the dataset 1

Figure 2. Images from the dataset 2

3.2 Identification of the most important spectral bands for water detection

To identify the most suitable bands for the purpose investigated in this study, a clustering analysis of pixel intensities for each band was performed, with respect to the classes (Water − W / Non_Water − NW). For example, if the clusters for classes W and NW are very close for a particular band, then that band is not ideal as it will not provide good segmentation in the images. In other words, the greater the separation between clusters, the better the band's suitability for water detection. To achieve this, K-means algorithm [20] was configured with 2 clusters. The results of the K-means clustering of the data are presented in figures 3 and 4. Figure 3 shows the centroids representing the average gray intensities of the spectral bands for each class. Figure 4 presents graphs illustrating the separation of the classes for each band. As observed in figures 3 and 4, the G, B, and EDG bands are more suitable for segmentation, followed by the R and NIR bands.

	B	G	R	NIR	EDG
Water (W)	76.830508	72.593220	60.491525	28.254237	62.610169
No_Water (NW)	38.779661	27.966102	25.135593	25.474576	26.898305

Figure 3. Centroids calculated by K-means for each class

It is observed that for the NIR band, the centroids for classes W and NW are very close, making it difficult the segmentation (or classification) of the pixels. Based on this analysis, and considering the small difference between the EDG and Red bands, we decided to use all three bands of the visible spectrum (R, G and B) to train a convolutional neural network (CNN) [21] with the YOLOv8 [22] architecture for water identification in the images acquired by the drone. This allowed the combination of images from both datasets for CNN training. It is worth noting that if any of the non-visible spectrum bands (NIR and EDG) had been included among the selected bands, we would have had to train two separate CNNs, as dataset 2 contains only RGB images.

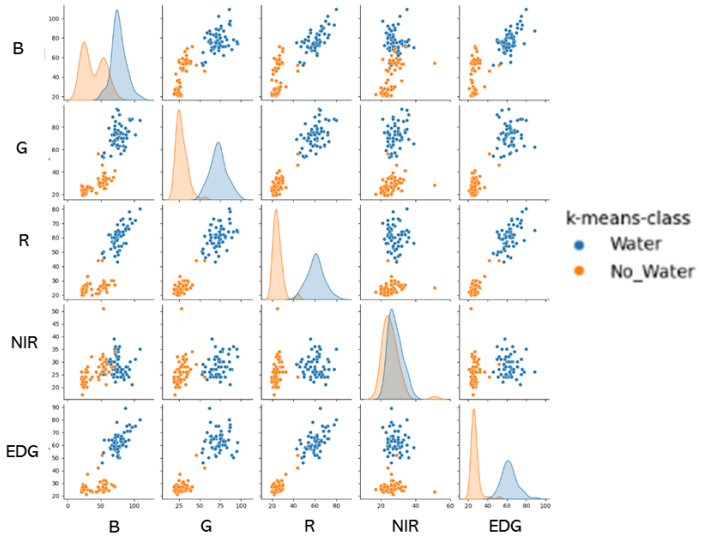


Figure 4. Graphical result of data clustering by K-means

3.3 Training and Predictions with the YOLOv8 CNN

Two experiments were conducted for water detection. In the first experiment, 76 images from dataset 1 were used for CNN training, while in the second experiment, images from both datasets were combined (76 images from dataset 1 and 36 images from dataset 2) for CNN training. The objective of these experiments was to analyze the CNN's predictions for images that were significantly different from those used in training. In both experiments, CNN was configured with the same parameters: a batch size of 16, an initial learning rate of 0.001, and three color channels (B, G, and R). Data augmentation was performed using a method integrated into the YOLO framework [22]. The online platform Roboflow [23] was used for annotating the training images (indicating regions of interest with bounding boxes – ground truth). Figure 5 illustrates the operational diagrams of the proposed approach for water detection in the images.

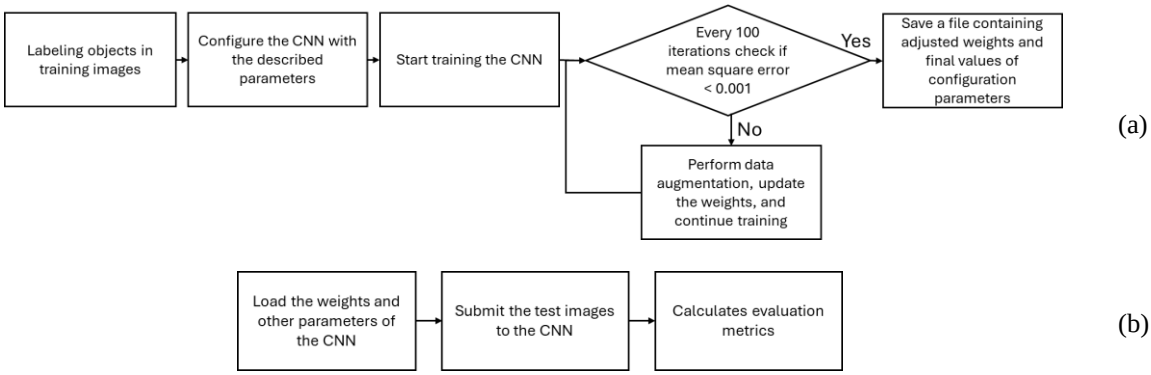


Figure 5. Operational diagrams of the proposed approach. (a) Training of the CNN and (b) Predictions with the images separated for testing

In the diagram of Figure 5a, representing the training of the CNN, only the first two blocks are configured manually. After that, the CNN manages the entire process, with the final step being to save the file containing the adjusted weights and other parameters, which represent the network's learning. The performance metric for the CNN during training was mAP50, the default in the YOLOv8 architecture, which evaluates how well the bounding boxes predicted by the CNN overlap with the ground truth boxes in the annotated images. Maximizing mAP50 indicates that the learning model is not only detecting objects but also accurately identifying their locations. Figure 5b refers to testing with the trained CNN. For this, the file containing the weights and other parameters is loaded, and then the images reserved for testing are processed by the CNN. At the end of this process, metrics are obtained that indicate the quality of water detection in the test images. In our experiments, we used accuracy, precision, and sensitivity metrics to evaluate the CNN's ability to successfully detect containers with water.

4 Results

During training, the following mAP-50 rates were achieved, respectively, in experiments 1 and 2: 0.980 and 0.950. The results obtained for the test images in both experiments are summarized in the confusion matrices presented in Figure 6. For the first experiment, 146 containers were detected by the trained CNN in the 32 images from dataset 1 that were analyzed, with 96 predicted as Water (W) and 50 predicted as Non_Water (NW). In the second experiment (Figure 6b), 47 images (32 from dataset 1 and 15 from dataset 2) were analyzed, and 215 containers were detected, with 131 predicted as W and 84 as NW. In terms of accuracy, precision, and sensitivity, the following results were obtained: Experiment 1 (accuracy = 0.993; precision = 0.989; sensitivity = 1.000) and Experiment 2 (accuracy = 0.930; precision = 0.931; sensitivity = 0.953).

	True Predicted	False Predicted
True	95	0
False	1	50

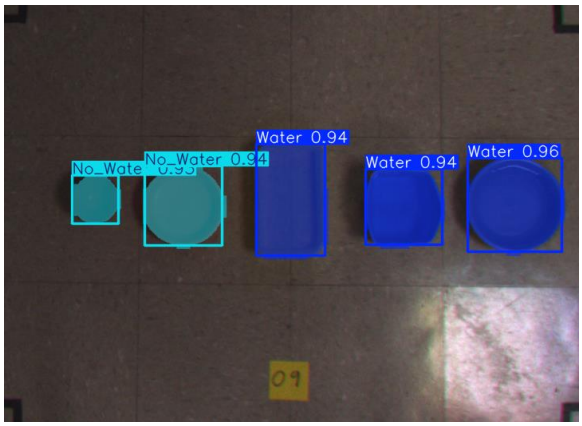
(a)

	True Predicted	False Predicted
True	122	6
False	9	78

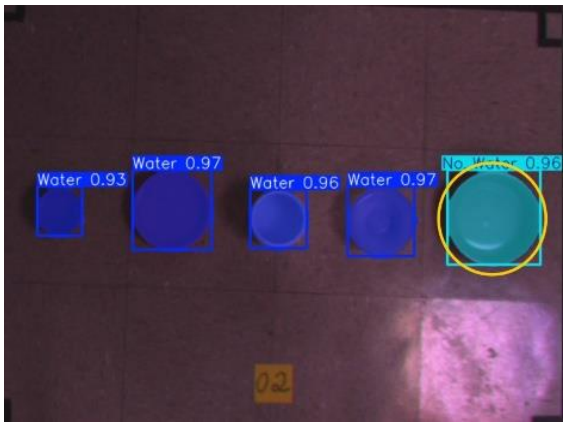
(b)

Figure 6. Confusion matrices obtained in the two experiments. (a) Experiment considering only images from dataset 1. (b) Experiment considering images from datasets 1 and 2

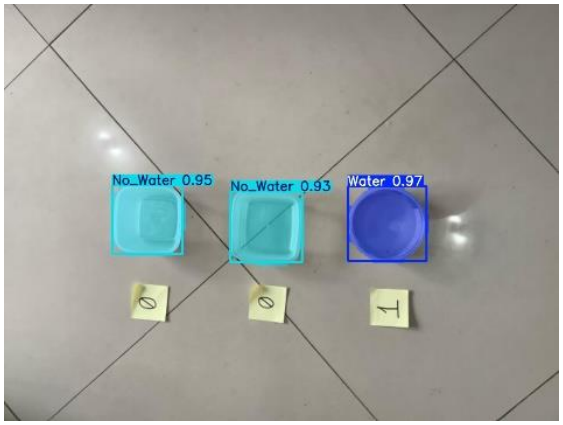
Figure 7 presents cases of water detection in containers from the test images. All containers with and without water shown in Figures 7a and 7c were correctly identified. In Figure 7b, out of the five containers containing water, four were correctly identified. One of them, highlighted with a yellow circle, was labeled as Non_Water, representing a false negative (FN). Finally, Figure 7d illustrates a false positive (FP), marked by a red circle, where a container that does not contain water was labeled as "Water".



(a)



(b)



(c)

(d)

Figure 7. Results obtained for four test images. (a) and (b) – images from dataset 1. (c) and (d) – images from dataset 2

In the conducted experiments, it was observed that the learning model encountered greater difficulty in detecting containers with water when they had lighter colors. Additionally, these errors occurred more frequently in areas with intense lighting, which made some regions of the images much brighter than others, hindering the learning process of the CNN. Another issue observed was the CNN's performance in detecting empty containers in the first experiment compared to its performance in the same task in the second experiment. This can be explained by the fact that there were more examples of containers without water in the second experiment, which contributed to a better interpretation of the pattern. Increasing the number of training images could likely help resolve these issues. Moreover, balancing the number of targets during training could improve the CNN's performance for both classes (W, NW). Although data augmentation was employed, brightness and contrast adjustments were not considered among the commonly used operations for this purpose, and no preprocessing operations were applied to the images.

Another problem observed was that when testing the model on images very different from those used in training, the CNN failed to detect water in a significant number of cases. This is probably due to the lack of similar examples in the training set. Therefore, it is suggested to include a variety of images for training, reflecting scenarios with different types of containers with and without water, taking into account the diversity of environments conducive to mosquito breeding. This can be seen as a limitation of the proposed approach, as no matter how comprehensive the training set is, there will always be distinct scenarios when the learning model is applied in practice. Thus, further studies are recommended to evaluate alternative solutions, one of which is the creation of indicators similar to those proposed in the works [15, 16].

Finally, it is important to note that although drone-acquired images offer significant advantages, such as high spatial and temporal resolutions, high-resolution images acquired by some satellite-borne sensors, such as WorldView-3, may also represent an important alternative for addressing the problem discussed in this work.

5 Conclusion

This study presented an approach for detecting small amounts of water in objects considered potential mosquito breeding sites. To achieve this, a CNN YOLOv8 was trained to segment the pixels of an image considering spectral bands chosen with the aid of the K-means algorithm. The conducted experiments demonstrated high sensitivity rates (1.000 and 0.953 − average of 0.977), indicating the feasibility of the proposed approach to identify containers with water. Nevertheless, some limitations were noted regarding the need for a large training dataset, which was expected since deep learning is being used. Finally, it is highlighted that the proposed approach, if properly parameterized and trained, can enhance the effectiveness of using drones in combating mosquito breeding sites. For future work, it is suggested to investigate more deeply the behavior of spectral bands in the identification of water, as the results were better when excluding the NIR band, which is not fully consistent with most literature involving water body identification. Additionally, it is recommended to explore the possibility of creating a water indicator, considering the specificities of the spectral bands available in cameras commonly used on drones.

References

[1] BVS − Biblioteca Virtual em Saúde.,Ministério da Saúde elabora plano para enfrentamento da dengue 2024/2025. Available at:https://bvsms.saude.gov.br/ministerio-da-saude-elabora-plano-para-enfrentamento-da-dengue-2024-2025/#:~:text=Mais%20de%20100%20pesquisadores%20e,laborat%C3%B3rios%20de%20refer%C3%AAncia%20do%20pa%C3%ADs. Accessed in: 10 Jul 2024.

[2] Portal G1 − SAÚDE. Brasil chega a 3 mil mortes confirmadas por dengue em 2024. Available at: https://g1.globo.com/saude/dengue/noticia/2024/05/24/brasil-chega-a-3-mil-mortes-confirmadas-por-dengue-em-2024.ghtml. Accessed in: 10 Jul 2024.

[3] MS – Ministério da Saúde. Dossiê de avaliação de tecnologias em saúde preparado para a CONITEC. Available at: https://www.gov.br/conitec/pt-br/midias/consultas/dossie/2023/DossietakedaVacinadengue.pdf. Accessed in: 10 Jul 2024.

[4] Portal Estadão – Saúde. Available at: https://www.estadao.com.br/saude/governo-lula-reduziu-em-58-gasto-com-campanhas-contra-dengue-em-2023-mesmo-com-alerta-de-epidemia/. Accessed in: 10 Jul 2024.

[5] Michelle, S., Kalonde, P., Zembere, K., Spaans, R. H., Jones, C., & Ghalayini, R. (2022, July). Combining Drone-Based Ultra-High-Resolution Earth Observation Data with AI for Mosquito Larval Habitat Identification: A Scalable Method in Malaria Vector Control. In IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium (pp. 4483-4485). IEEE.

[6] Agarwal, A., Chaudhuri, U., Chaudhuri, S., & Seetharaman, G. (2014, July). Detection of potential mosquito breeding sites based on community sourced geotagged images. In Geospatial InfoFusion and Video Analytics IV; and Motion Imagery for ISR and Situational Awareness II (Vol. 9089, pp. 175-182). SPIE.
[7] Mehra, M., Bagri, A., Jiang, X., & Ortiz, J. (2016, June). Image analysis for identifying mosquito breeding grounds. In 2016 IEEE International conference on sensing, communication and networking (SECON workshops) (pp. 1-6). IEEE.
[8] Dias, T. M., Alves, V. C., Alves, H. M., Pinheiro, L. F., Pontes, R. S. G., Araujo, G. M., ... & Prego, T. D. M. (2018, November). Autonomous detection of mosquito-breeding habitats using an unmanned aerial vehicle. In 2018 Latin American Robotic Symposium, 2018 Brazilian Symposium on Robotics (SBR) and 2018 Workshop on Robotics in Education (WRE) (pp. 351-356). IEEE.
[9] Passos, W. L., Dias, T. M., ALVES JUNIOR, M., BARROS, D., ARAUJO, M., LIMA, A., ... & LIMA, S. (2018). About automatic detection of aedes aegypti mosquito focuses. Anais do XXXVI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, 1-5.
[10] Bravo, D. T., de Araujo Lima, S. J., de Araujo, S. A., & Alves, W. A. L. (2019). Detection of Small Portions of Water in VIS-NIR Images Acquired by UAVs. In Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications: 23rd Iberoamerican Congress, CIARP 2018, Madrid, Spain, November 19-22, 2018, Proceedings 23 (pp. 168-176). Springer International Publishing..
[11] Lima, G. A., Cotrin, R. O., Belan, P. A., & de Araújo, S. A. (2021). Sistema de visão computacional para identificação automática de potenciais focos do mosquito aedes aegypti usando drones. Revista Ibérica de Sistemas e Tecnologias de Informação, (43), 93-109.
[12] Niluksha, T., Ananthajothy, G., Navaratnam, R., & Dissanayake, M. B. (2023, August). Weighted Ensemble Algorithm for Aerial Imaging Based Mosquito Breeding Sites Classification. In 2023 IEEE 17th International Conference on Industrial and Information Systems (ICIIS) (pp. 347-352). IEEE
[13] Sonali, B., & Patil, K. (2023). A Novel System for Potential Mosquito Breeding Hotspot Intimation and Monitoring Using MLOps and Improved YoloV3. Instrumentation, Mesure, Metrologie, 22(1), 35.
[14] MS – Ministério da Saúde Ministério da Saúde alerta para aumento de 149% dos casos de dengue no país. Available at: https://www.gov.br/saude/pt-br/assuntos/noticias/2019/fevereiro/ministerio-da-saude-alerta-para-aumento-de-149-dos-casos-de-dengue-no-pais. Accessed in: 10 Jul 2024..
[15] Gao, B. C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote sensing of environment, 58(3), 257-266.
[16] Polidorio, A. M., Imai, N. N., & Tommaselli, A. M. G. (2004). Índice indicador de corpos d'água para imagens multiespectrais. I Simpósio de Ciências Geodésicas e Tecnologias da Geoinformação, 9.
[16] Prasad, M. G., Chakraborty, A., Chalasani, R., & Chandran, S. (2015, December). Quadcopter-based stagnant water identification. In 2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG) (pp. 1-4). IEEE.
[17] de Mesquita, G. F., Santos, L. R., Dias, E. E. C., Lanzillotta, L. C., da Silva Pimentel, P. H. L., & de Toledo Quadros, J. R. (2021, June). Drone with anti-larvaes biomedicines, using deep learning to identify the reproduction zones of insects. In 2021 16th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.
[18] Truong, H. M., & Clavel, M. (2022, November). Identifying temporary water bodies from drone images at real-time using deep-learning techniques. In 2022 International Conference on Advanced Computing and Analytics (ACOMPA) (pp. 12-19). IEEE.
[19] Gonzalez, R. C.; Woods, R. E., ""Digital Image Processing. Massachusetts". São Paulo: Edgard Blucher, 2002.
[20] Yuan, C., & Yang, H. (2019). Research on K-value selection method of K-means clustering algorithm. J, 2 (2), 226–235.
[21] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.
[22] Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics yolov8. 2023. URL https://github. com/ultralytics/ultralytics.