Application of YOLOv7 for real-time detection of Aedes Aegypti

Danilo Oliveira, Samuel Mafra, Felipe Figueiredo, Guilherme Piedade, Mateus Cruz and Eduardo Teixeira .

***Abstract*— Public health faces significant challenges in combating the *Aedes aegypti* mosquito, which still threatens the Brazilian population. Despite awareness campaigns and control measures, the incidence of diseases such as dengue, zika and *chikungunya* is still high. However, technological advances have enabled the development of devices capable of detecting female mosquitoes *Aedes aegypti*, the main transmission vector of these diseases. The use of IoT systems (Internet of Things) and weather stations can assist researchers in controlling the population of these insects, enabling the monitoring of high-risk areas. This article presents an intelligent system that uses Computer Vision to detect *Aedes aegypti* mosquitoes. The objective of this article was to develop an IoT system architecture using the You Only Look Once v7 (YOLOv7) algorithm, which provides a superior solution compared to existing methods for real-time detection of the Aedes aegypti mosquito. By combining YOLOv7's real-time detection capabilities with the connectivity and intelligence of the IoT system, the proposed solution offers a significant advantage. Furthermore, the integration of the IoT architecture allows for continuous data collection and the implementation of advanced analytics, such as machine learning, enabling continuous improvements in the accuracy and efficiency of the detection system. This adaptive approach, coupled with real-time responsiveness, makes the proposed solution highly effective and promising in combating Aedes aegypti.**

***Keywords*— IoT; LoRa/LoRaWan; Dengue and Computer Vision**

I. INTRODUCTION

Dengue is an infectious viral disease transmitted by the *Aedes aegypti* mosquito, which is a public health problem in Brazil. Since the 1980s, the country has faced outbreaks and epidemics of the disease. In 2019, for example, more than 1.5 million cases were reported in the country [1] with more than 750 deaths. The problem is especially serious in the tropical and subtropical regions, where the climate is favorable for the proliferation of the transmitting mosquito [2].

In Brazil, the fight against Dengue has several available solutions, one of the main strategies being the elimination of possible breeding grounds of standing water. In addition, it is crucial to make the population aware of the importance of identifying and eliminating mosquito breeding sites in their own homes, as the mosquito has a flight range of approximately 800 meters [3]. Thus, residents must understand that mosquito control is not limited to the surroundings of their homes, but involves the collaboration of the entire community.

It is essential to adopt preventive measures throughout the region to reduce the spread of the dengue mosquito and prevent disease. It is a collective responsibility that involves the population and the government [4]. Medical treatment is based on supportive measures to control symptoms and prevent complications, as there is no specific cure for illness caused by the virus. Early diagnosis and adequate treatment are essential, including hospitalization in severe cases [5].

The scientific literature encompasses numerous studies addressing the application of IoT (Internet of Things) in monitoring and surveillance of urban and rural pests. These researches extensively explore smart devices equipped with internet-connected traps, including investigations of solutions and commercial prototypes [6]. By examining the published literature, it is possible to identify approaches that are based on the use of optoelectronics sensors to classify insects based on the frequency of wingbeats during flight [7], flying insect recognition and counting systems that are based on vision [8].

Computer Vision (CV) is a subfield of artificial intelligence that is currently used to automate various tasks that require the use of vision. In addition, several algorithms that make use of CV have already been proposed in the literature, including YOLO. Despite the development of the eighth version of the algorithm, YOLOv7 still stands out as an excellent solution for real-time detection when compared to other algorithms [9]. The YOLOv7 model is an improved version over previous versions like YOLOv4, incorporating a pre-processing technique from YOLOv5 to improve the identification of smaller objects. The architecture features the E-ELAN (Extended Efficient Layer Aggregation Network) block, which has enhanced the networks learning capability through cardinality expansion, scrambling, and merging. Group convolution is employed to increase the channel and cardinality of the computation block.

The objective of this study was to propose an update and

Danilo Machado de Oliveira, Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí-MG, e-mail: danilomachado@mtel.inatel.br; Samuel Baraldi Mafra, Departamento de Engenharia de Telecomunicações, Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí-MG, e-mail: samuelbmafra@inatel.br; Felipe Augusto Pereira de Figueiredo, Departamento de Engenharia de Telecomunicações, Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí-MG, e-mail: felipe.figueiredo@inatel.br;Guilherme Pires Piedade,Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí-MG,e-mail:guilherme.pires@mtel.inatel.br;Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí-MG, e-mail: mateusraimundo@mtel.inatel.br Eduardo Henrique Teixeira,Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí-MG, e-mail: eduardoteixeira@dtel.inatel.br; Eduardo Henrique Teixeira,Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí-MG, e-mail: eduardoteixeira@dtel.inatel.br. This work is partially supported RNP, with resources from MCTIC, Grant No. 01245.020548/2021-07, under the Brazil 6G project of the Radiocommunication Reference Center (Centro de Referência em Radiocomunicações - CRR) of the National Institute of Telecommunications (Instituto Nacional de Telecomunicações - Inatel), Brazil, Huawei, under the project Advanced Academic Education in Telecommunications Networks and Systems, contract No PPA6001BRA23032110257684, Brazil, the National Council for Scientific and Technological Development-CNPq (403827/2021-3), FAPESP (2021/06946-0),and by Minas Gerais State Agency for Research and Development (FAPEMIG) via Grant No. TEC - APQ-03283-17

improvement of existing insect surveillance systems, through the implementation of a solution based on disruptive technologies, such as the IoT, using the YOLOv7 algorithm for detection. This solution uses a system based on CV and IoT to automatically detect and count the number of *Aedes aegypti* mosquitoes, responsible for epidemic outbreaks. Through the localization and classification of mosquitoes through CV, data is sent remotely by LoRa to a cloud platform, allowing its visualization, storage and subsequent analysis. The objective is to mitigate the epidemiological outbreak of the *Aedes aegypti* mosquito, making detection more efficient and effective.

This article is organized into five sections: Section II reviews the literature on techniques for real-time detection of the *Aedes aegypti* mosquito; Section III describes the proposed system architecture in detail, including the model used; Section IV presents the experimental validation results; and Section V concludes the article and suggests future research directions.

II. Theoretical Background

In recent years, the use of CV and machine learning (ML) techniques for the identification, detection and classification of *Aedes aegypti* mosquitoes has been widely studied. IoT, intelligent traps and robotics, allowed the development of automated systems that aim to efficiently detect and control the infestation of these mosquitoes, contributing to epidemiological research.

In article[9], researchers investigate the use of sensors and unimodal classifiers for accurate identification of Aedes aegypti mosquitoes. In article [7], the authors evaluate the sensitivity and noise of an optoelectronic sensor used in mosquito monitoring, emphasizing the importance of an effective and reliable detection system. The work [10] proposes a deep convolutional neural network-based approach for the classification of Aedes albopictus mosquitoes, highlighting the potential of deep learning techniques in the identification and control of these disease vectors. The utilization of computer vision (CV) systems and deep learning techniques has shown great potential in the identification and classification of mosquito species, such as Aedes aegypti, which is a vector for diseases like dengue, Zika, and chikungunya.

The article "Computer vision system for automatic identification of potential Aedes aegypti mosquito breeding sites using drones"explores the use of drones to identify mosquito breeding sites [11], while articles [12], [13] highlight the power of convolutional neural networks in image processing and species classification. Furthermore, the article "Mapping the spatial distribution and predicting the abundance of dengue vectors using machine learning"presents a machine learning approach to map the spatial distribution and predict the abundance of dengue vectors, demonstrating the effectiveness of these techniques in combating mosquito-borne diseases [14].

Otherwise, there are still limitations and challenges in detecting *Aedes aegypti* mosquitoes through CV and ML, such as the variation in mosquito appearance at different stages of development and in different environments. In addition, it is important to highlight that these automated systems are complementary to traditional vector control methodologies and should not replace them [14]. It is possible to envisage several future applications of CV and machine learning for the control of *Aedes aegypti* mosquitoes, such as automated detection in smart traps and real-time monitoring of infestation areas. However, it is necessary to continue to invest in research and development of technologies in this field to overcome limitations and improve the effectiveness of these systems.

Basically, CNNs consist of three types of layers: convolutional, clustering, and fully connected [15]. Convolutional layers use filters to extract features from images, while clustering layers perform spatial sampling to generate lower-resolution versions of the convolutional layers. Finally, the fully connected layers act as classifiers, producing an n-dimensional matrix that indicates the probability of the input pattern belonging to a given class [15], [16], [12]. There are several CNN architectures, with YOLO being one of the most used for object recognition [17], [18], [19]. In the present work, YOLO was used to form the proposed CV system.

III. Materials and Methods

In this article, a solution based on IoT for real-time monitoring and detection of the Aedes aegypti mosquito using the YOLO v7 algorithm is discussed. The architecture, illustrated in Figure 1, consists of a Raspberry Pi 4 board with 8 GB of RAM, responsible for processing and executing the YOLOv7 algorithm. The images captured by the camera are sent to the Raspberry Pi board, where the CV algorithm identifies the insect being detected in real-time. The data is processed, stored, and then sent via serial communication to the ESP32 LoRa, which is handled and transmitted using the LoRaWAN protocol to the LoRaWAN gateway configured in the ChirpStack LoRaWAN® network server platform. LoRa technology is a low-power radio frequency (RF) communication method commonly used in situations that require low transmission rates and long distances. This technology is highly recommended for implementation in IoT devices and applications that utilize sensor networks [20].

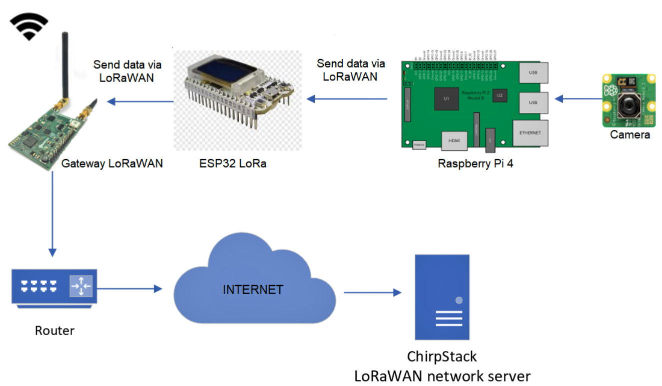

Fig. 1: IoT Trap System architecture

The Figure 2 illustrates the architecture of YOLOv7, which represents a significant evolution compared to previous versions [21]. This real-time object detection approach utilizes deep convolutional neural networks to divide the input image into a grid and then predicts the coordinates of bounding boxes and the classes of objects present in each grid cell.

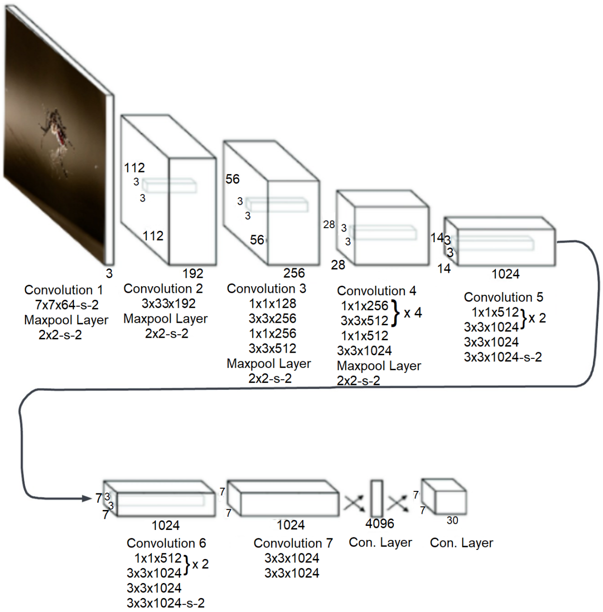

Fig. 2: YOLO architecture.

This architecture, characterized by its computational efficiency and high accuracy, incorporates improvements such as the utilization of CSPDarknet53 modules for feature extraction, enhancements in the loss function, and the implementation of attention mechanisms like Spatial Pyramid Pooling (SPP) and Path Aggregation Network (PANet) to enhance robustness and the model's generalization capacity in complex and diverse scenarios [22] The software development was divided into sections for microcontrollers and cloud storage. The ESP32 LoRa microcontrollers used Arduino IDE, while Raspberry Pi and the Radioenge gateway ran Rasbian OS. The open-source ChirpStack LoRaWAN® platform was chosen for reliable and secure cloud storage. The Node LoRa ESP32 served as a low-power wireless communication solution, enabling data transmission between sensor nodes and the LoRa gateway.

The LoRaWAN gateway is responsible for receiving the data transmitted by the ESP32 LoRa and forwarding them to the LoRaWAN server. This server, in turn, stores the received data and sends automatic alerts in case of detection of mosquitoes or infestation. This mosquito detection system is extremely efficient and can be used in different environments such as homes, schools, hospitals, parks, and other public areas. In addition, it is highly customizable and can be adapted to meet the specific needs of each monitored environment.

For the study, a dataset was built with images of different species of insects, including *Aedes aegypti*, bees and butterflies. This set of images was obtained by merging the datasets present in the Kaggle community. The resulting set has 7,673 images, 3,371 of the *Aedes aegypti*[23] class, 3,637 images of a class of bees [24] and 665 of the class of butterflies [25].

To assess the generalizability of the detection model, the data set was divided into training (70%), validation (20%) and testing (10%) for each type of insect. As the sizes of the images in the dataset were not uniform, an initial normalization phase was performed to standardize all the photos to an image with a resolution of 640×640 pixels. To complete the manual labeling of each insect class, image data annotation software was used [26]. After successfully labeling the images, text files corresponding to each of the images are generated containing the class and location information of the insect within the image. To ensure that the bounding box comprises as little of the background as possible, the images were labeled based on the smallest bounding box surrounding the insects. Sample images of the three types of insects are shown in Figure 3, including bee, butterflies, and aedes-aegypti.

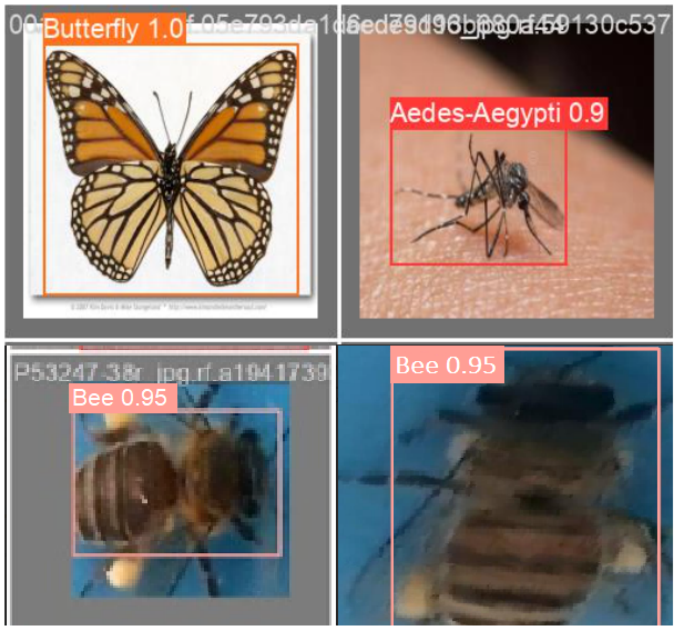

Fig. 3: Image annotation examples.

YOLO is trained in several steps, starting with collecting a large amount of training data, which is pre-processed and divided into training and validation sets.

There are three main matric's for evaluate the model: (I) box loss, (II) objectness loss, and (III) classification loss. The box loss reflects the algorithms ability to accurately locate the center of an object and predict its bounding box. Objectness measures the probability that an object is present in a particular region of interest, with high values indicating the likelihood of an object's existence. Classification loss indicates the model's ability to accurately predict the correct object class. The precision, recall, and mean average precision of the model improved quickly at the outset and reached a plateau after approximately 50 epochs. Similarly, the box, objectness, and classification losses of the validation data showed a sharp decrease until epoch 50. To choose the best weights, we utilized early stopping.

The neural network is initialized with random weights and then fed with the training images. During training, the network adjusts its weights to minimize the loss function, which measures the difference between the network's predictions and the true labels. This difference is calculated using a loss function such as the cross-entropy loss function. The training process is repeated several times, or epochs until the neural network

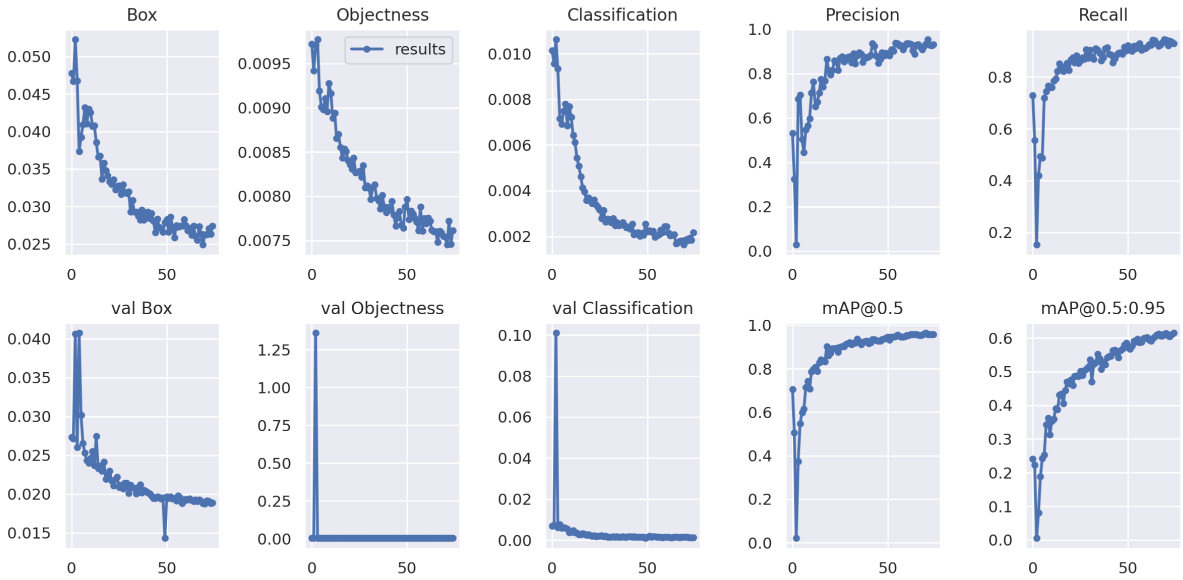

Fig. 4: Created graphs illustrating the changes in box loss, objectness loss, classification loss, precision, recall, and mean average precision (mAP) for the training and validation sets over the course of training epochs.

learns to recognize objects of interest with high accuracy. After training, the network is tested on a test set to assess its accuracy. If accuracy reaches a satisfactory level, the network is considered ready for use.

Epochs are an important concept in training neural networks, including YOLO. Each epoch represents a complete pass through the training data, during which the neural network updates its weights based on errors made in the predictions. The number of epochs needed to train a neural network depends on the size and complexity of the dataset, as well as the network architecture. In general, the larger the dataset and the more complex the network, the more epochs will be needed to achieve the desired accuracy. However, it is important to balance the number of epochs with the risk of overfitting, which occurs when the network fits too tightly to the training data and loses the ability to generalize to new data. Therefore, it is common to use techniques such as cross-validation and hyperparameter adjustment to determine the ideal number of epochs for neural network training.

IV. Results analysis

In this section, an analysis of the results obtained during the training of the computer vision model will be conducted with the aim of detecting the *Aedes aegypt*i mosquito in real-time. The YOLOv7 algorithm was employed, which stood out for its high efficiency and precision in image processing.The loss function shown in Figure 4 can be decomposed into three distinct components: the box loss measures the accuracy of the predicted bounding box parameters, such as width and height, as well as the offset from the center of the true object. The objectness loss is calculated based on the probability of an object being present in a given region of interest, with the aim of improving the detection of objects against the background. Finally, the classification loss evaluates the model's ability to accurately classify objects based on their features, such as shape, color, and texture. By minimizing these losses during training, the YOLOv7 algorithm can achieve high detection accuracy and real-time performance.

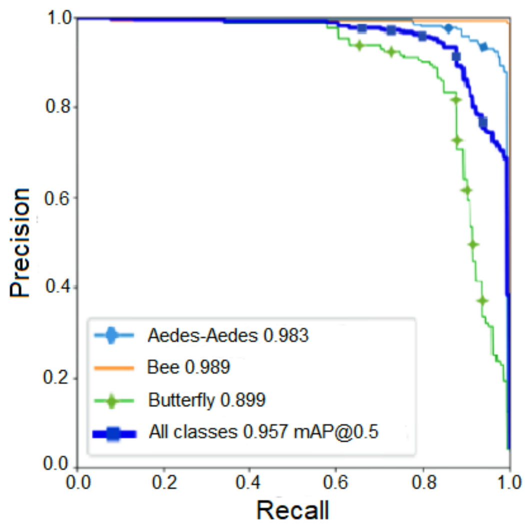

Fig. 5: R-Curve result

The objectivity metric quantifies the probability of an object being found in a given area, suggesting that an object is within the visible region of the image. Figure 5 presents a precision recovery curve that serves as a granular class-specific performance indicator. By observing the PR curve, we can conclude that the bee class had the best performance, reaching a precision of 98.9%, and following the bee class the *Aedes aegypti* class, which had a performance of 98.3%.

The butterfly classes showed a performance drop of 89.9%, respectively, due to the smaller number of images available in the dataset, which had only 665 images. In comparison, the image set for *Aedes aegypti* and bees contained over 3,000 images each. Despite this, the architecture had a satisfactory overall performance of 95.7%. Figure 5 shows the precision-recall curve of the global space criterion for mosquito detection using the YOLOv7 algorithm. To facilitate evaluation, the algorithm is able to detect almost all mosquitoes, even when they are partially occluded. This is also illustrated in Figure 5, where the maximum recall rate is above 98.9%.

V. Conclusion

An advanced Convolutional Neural Network structure was used to enable surveillance and population mapping of mosquitoes. After a careful evaluation, the general and specific objectives were fully achieved. An efficient *Aedes aegypti* and Bee detection algorithm was developed with YOLOv7. To assess the accuracy of the algorithm, several metrics were applied, including accuracy, recovery, F1 measures and precision. The results were extremely positive, demonstrating that YOLOV7 achieved an average accuracy of 95.7%. This modern technology presents new possibilities for monitoring and controlling the population of mosquitoes, in addition to helping to prevent diseases transmitted by vectors.

Future work includes expanding the insect monitoring system's database by adding more species. Furthermore, a prototype is being developed to analyze the behavior of different insects in the area, with emphasis on the *Aedes aegypti* mosquito. The prototype will be powered by rechargeable batteries, which will be charged through a system of photovoltaic modules that includes a charge controller, temperature and humidity sensors and an optoelectronic sensor capable of detecting the sex of the mosquito by the frequency of its wingbeats. The algorithm will also be updated to improve the effectiveness of the system. Several smart traps will be strategically placed around the city for field study. These systems will effectively control the spread of the *Aedes aegypti* mosquito, which is responsible for transmitting diseases such as dengue, zika and chikungunya. These initiatives represent a significant step forward in improving techniques for monitoring and studying the lives of insects, in addition to contributing to the control and reduction of diseases transmitted by insects. Our commitment is to continue developing innovative technological solutions to improve people's quality of life and prevent diseases transmitted by insects.

Referências

[1] S. N. S. Fonseca, "Changing epidemiology of dengue fever in children in south america," *Current Opinion in Pediatrics*, vol. 35, no. 2, pp. 147–154, 2023.

[2] T. F. Pinheiro, J. de Barros Alves, and Y. R. N. da Silva, "O impacto financeiro das arboviroses oriundas do aedes aegypti no brasil: uma projeção para 2019," *Brazilian Journal of Development*, vol. 6, no. 5, pp. 30 757–30 767, 2020.

[3] N. A. Honório, W. d. C. Silva, P. J. Leite, J. M. Gonçalves, L. P. Lounibos, and R. Lourenço-de Oliveira, "Dispersal of aedes aegypti and aedes albopictus (diptera: Culicidae) in an urban endemic dengue area in the state of rio de janeiro, brazil," *Memórias do Instituto Oswaldo Cruz*, vol. 98, pp. 191–198, 2003.

[4] R. Allan, S. Budge, and H. Sauskojus, "What sounds like aedes, acts like aedes, but is not aedes? lessons from dengue virus control for the management of invasive anopheles," *The Lancet Global Health*, vol. 11, no. 1, pp. e165–e169, 2023.

[5] T. N. Pereira, F. D. Carvalho, S. F. De Mendonça, M. N. Rocha, and L. A. Moreira, "Vector competence of aedes aegypti, aedes albopictus, and culex quinquefasciatus mosquitoes for mayaro virus," *PLOS Neglected Tropical Diseases*, vol. 14, no. 4, p. e0007518, 2020.

[6] J. Aira, T. Olivares, F. M. Delicado, and D. Vezzani, "Mosquiot: A system based on iot and machine learning for the monitoring of aedes aegypti (diptera: Culicidae)," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–13, 2023.

[7] D. A. Santos, L. E. Teixeira, A. M. Alberti, V. Furtado, and J. J. Rodrigues, "Sensitivity and noise evaluation of an optoelectronic sensor for mosquitoes monitoring," in *2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech)*. IEEE, 2018, pp. 1–5.

[8] Y. Zhong, J. Gao, Q. Lei, and Y. Zhou, "A vision-based counting and recognition system for flying insects in intelligent agriculture," *Sensors*, vol. 18, no. 5, p. 1489, 2018.

[9] V. M. Souza, "Identifying aedes aegypti mosquitoes by sensors and one-class classifiers," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications: 21st Iberoamerican Congress, CIARP 2016, Lima, Peru, November 8–11, 2016, Proceedings 21*. Springer, 2017, pp. 10–18.

[10] G. Adhane, M. M. Dehshibi, and D. Masip, "A deep convolutional neural network for classification of aedes albopictus mosquitoes," *IEEE Access*, vol. 9, pp. 72 681–72 690, 2021.

[11] G. A. Lima, R. O. Cotrin, P. A. Belan, S. A. de Araújo *et al.*, "Sistema de visão computacional para identificação automática de potenciais focos do mosquito aedes aegypti usando drones," *RISTI-Revista Ibérica de Sistemas e Tecnologias de Informação*, no. 43, pp. 93–109, 2021.

[12] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[13] K. Okayasu, K. Yoshida, M. Fuchida, and A. Nakamura, "Vision-based classification of mosquito species: Comparison of conventional and deep learning methods," *Applied sciences*, vol. 9, no. 18, p. 3935, 2019.

[14] M. S. Rahman, C. Pientong, S. Zafar, T. Ekalaksananan, R. E. Paul, U. Haque, J. Rocklöv, and H. J. Overgaard, "Mapping the spatial distribution of the dengue vector aedes aegypti and predicting its abundance in northeastern thailand using machine-learning approach," *One Health*, vol. 13, p. 100358, 2021.

[15] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[16] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 international conference on engineering and technology (ICET)*. Ieee, 2017, pp. 1–6.

[17] Z. Yi, S. Yongliang, and Z. Jun, "An improved tiny-yolov3 pedestrian detection algorithm," *Optik*, vol. 183, pp. 17–23, 2019.

[18] B. Benjdira, T. Khursheed, A. Koubaa, A. Ammar, and K. Ouni, "Car detection using unmanned aerial vehicles: Comparison between faster r-cnn and yolov3," in *2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS)*. IEEE, 2019, pp. 1–6.

[19] Z.-F. Xu, R.-S. Jia, H.-M. Sun, Q.-M. Liu, and Z. Cui, "Light-yolov3: fast method for detecting green mangoes in complex scenes using picking robots," *Applied Intelligence*, vol. 50, pp. 4670–4687, 2020.

[20] C. R. M. Silva and F. A. C. M. Silva, "An iot gateway for modbus and mqtt integration," in *2019 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC)*. IEEE, 2019, pp. 1–3.

[21] J. Terven and D. Cordova-Esparza, "A comprehensive review of yolo: From yolov1 to yolov8 and beyond," *arXiv preprint arXiv:2304.00501*, 2023.

[22] A. P. Rangari, A. R. Chouthmol, C. Kadadas, P. Pal, and S. Kumar Singh, "Deep learning based smart traffic light system using image processing with yolo v7," in *2022 4th International Conference on Circuits, Control, Communication and Computing (I4C)*, 2022, pp. 129–132.

[23] P. Isawasan, "Aedes mosquitos," https://www.kaggle.com/datasets/pradeepisawasan/aedes-mosquitos, 2021, acessado em 24 de Abril de 2023.

[24] I. Felizardo, "Honey bee pollen," https://www.kaggle.com/datasets/ivanfel/honey-bee-pollen, 2021, acesso em: 24 de abril de 2023.

[25] Veeralakrishna, "Butterfly dataset," https://www.kaggle.com/datasets/veeralakrishna/butterfly-dataset, 2021, acesso em: 24 de abril de 2023.

[26] "Roboflow - treinamento de modelos de visão computacional," https://app.roboflow.com/inatel-5zq1l, 2023, acessado em 24 de abril de 2023.