

A Novel Brazilian Dataset for Automatic Detection of Psyllid Attacks in Guava

Um Novo Conjunto de Dados Brasileiro para a Detecção Automática de Ataques de Psilídeos em Goiabas

Iago Magalhães de Mesquita^{1*}, Larissa Ferreira Rodrigues Moreira², Rhuan da Silva Nunes¹,
Iális Cavalcante de Paula Júnior¹

Abstract: Guava, a fruit native to America, plays a crucial role in the Brazilian economy, with production reaching 564,000 tons by 2022. It is widely used in culinary dishes and herbal remedies, and is exported to numerous countries. However, guava is also susceptible to pests and diseases. Among these, *Psyllid Triozoda Limbata* is particularly problematic, attacking young leaves and leading to reduced fruit yields and impaired plant development. To address this issue and assist producers in detecting pest infestations, this study introduces a novel dataset comprising images from Brazil. By leveraging Artificial Intelligence and Computer Vision techniques, this dataset provides a foundation for developing pest detection systems. We evaluated various object detection architectures including different versions of YOLO and Vision Transformer (ViT). The experiments indicated that YOLO V5 achieved promising results, with a precision of 93.14%. Moreover, the availability of this dataset opens new opportunities for enhancing pest management strategies, improving crop protection, and increasing the overall productivity in the guava industry.

Keywords: Guava Trees – Psyllids – Object Detection — Deep Learning

Resumo: A goiaba, uma fruta nativa do continente americano tem um impacto significativo na economia brasileira, que em 2022 chegou a produzir 564 mil toneladas do fruto. No entanto, a goiaba também é suscetível a pragas e doenças. Entre elas, o *Psilídeo Triozoda Limbata* é particularmente problemático, atacando as folhas jovens e levando à redução da produção de frutos e ao comprometimento do desenvolvimento da planta. Para resolver este problema e ajudar os produtores a detectar infestações de pragas, este estudo apresenta um novo conjunto de dados que inclui imagens do Brasil. Ao utilizar técnicas de Inteligência Artificial e Visão Computacional, esse conjunto de dados fornece uma base para o desenvolvimento de sistemas de detecção de pragas. Avaliamos várias arquiteturas de detecção de objetos, incluindo diferentes versões do YOLO e do Vision Transformer (ViT). Os experimentos indicaram que o YOLO V5 obteve resultados promissores, com uma precisão de 93,14%. Além disso, a disponibilidade deste conjunto de dados abre novas oportunidades para melhorar as estratégias de gestão de pragas, melhorar a proteção das culturas e aumentar a produtividade global na indústria da goiaba.

Palavras-Chave: Goiabeiras — Psilídeos — Detecção de Objetos — Aprendizado Profundo

¹Federal University of Ceará (UFC), Brazil

²Institute of Exacts and Technological Sciences, Federal University of Viçosa (UFV), Brazil

*Corresponding author: iagomagalhaes23@gmail.com

DOI: <http://dx.doi.org/10.22456/2175-2745.143544> • Received: 24/10/2024 • Accepted: 28/11/2024

CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introduction

Psidium guajava L. is the scientific name attributed to guava [1]. It is a fruit native to the American continent, which, owing to its adaptability, is cultivated in other tropical regions around the world [2]. According to the Brazilian Institute of Geography and Statistics (IBGE), in 2022, Brazil produces approximately 564 thousand tons of guava, highlighting the

importance of the fruit for the country's commerce and global consumption [3], [4].

Guava is widely used in various recipes as a phytotherapeutic agent. However, similar to many other fruit and vegetable crops, it is susceptible to several pests and diseases that directly affect fruit development [4]. Among these pests, the psyllid *Triozioida limbata* stands out as an insect that sucks

the sap from the leaves of the guava tree, causing necrosis in the leaves [1][4].

The proliferation of these pests can cause numerous problems in guava production, negatively affecting the economy of the fruit [4]. In this context, identifying the pests that affect plants can enable producers to act more quickly, applying insecticides at the correct time to reduce and eliminate the proliferation of these insects. With the advent of technology, this issue can be effectively tackled by leveraging Artificial Intelligence and Computer Vision, as demonstrated in various agricultural studies that have successfully identified the affected plants of different crops and distinguished them from healthy ones [5, 6].

This paper introduces a novel Brazilian dataset comprising images of guava plants afflicted by psyllid attacks. To the best of our knowledge, this is the first study to create a dataset specifically tailored to the Brazilian context to detect *Triozoida limbata* infestations. Our study represents a significant advancement in pest management in the guava industry, addressing a gap in the available datasets.

We trained and evaluated various object detection models based on deep learning using this new dataset, and applied metrics to determine the most effective approach. Our main goal was to provide producers with a robust tool that simplifies the detection of *Triozoida limbata* infestations, thereby enabling timely and informed intervention strategies.

The remainder of this paper is organized as follows. Section 2 presents the related work. Section 3 describes the materials and methods, including the novel dataset proposed in this paper. Section 4 presents the results and discussion. Section 5 concludes the paper.

2. Related work

This section presents studies that address the use of Artificial Intelligence and Computer Vision for disease identification in agriculture, both in guava crops and other cultures.

Barbedo [7] conducted a study to investigate how the size and variety of datasets can affect the effectiveness of deep learning techniques applied to phytopathology. They used a database of images of 12 plant species. The results indicated that although the automatic classification of plant diseases has made significant advances, the use of image datasets for training still has many undesirable consequences, hindering the effective dissemination of this technology.

Rajbongshi et al. [8] presented a dataset consisting of images of both healthy and diseased guava leaves and fruits. They proposed classification into six classes using machine learning techniques.

Pathmanaban et al. [9] developed a study focused on the early identification of diseases in guavas using image processing and computer vision. The study is composed of a database containing digital and thermal images of healthy, damaged, and diseased guavas categorized by maturity level and drop height.

Mattihalli et al. [10] proposed an approach was developed for the early detection of diseases in plant leaves and automation of the medication application process using a Raspberry Pi-based system. The system uses a database of leaf images that are compared with a reference database. If a disease is detected, the system notifies the user via the GSM and automatically activates the application of medications.

Shetty et al. [11] developed a system for the identification of diseases in plants with similar visual characteristics, using deep learning and computer vision methods, such as You Only Look Once (YOLO) and Faster Region Convolutional Neural Network with a ResNet-152 backbone. The system was trained with a customized dataset, in which the images were captured in an uncontrolled environment, considering guava and mango leaves.

In contrast to previous studies, our approach shows potential by employing a more comprehensive comparative analysis with different object detection techniques for the identification and classification of attacks on guava trees. Additionally, we developed a unique dataset with images of psyllid attacks.

3. Material and Methods

This section presents the research methodology and materials used for conducting the experiments, detailing the image dataset, architectures, and evaluation metrics employed for identifying psyllid attacks on guava leaves.

Figure 1 presents a flowchart representing the methodology used in this research. It outlines the models employed for object detection, such as YOLO V5, V8, V9, V10, NAS, and VIT. Additionally, the evaluation metrics used are indicated, including Precision, Recall, F1-Score, Average Precision, and Mean Average Precision.

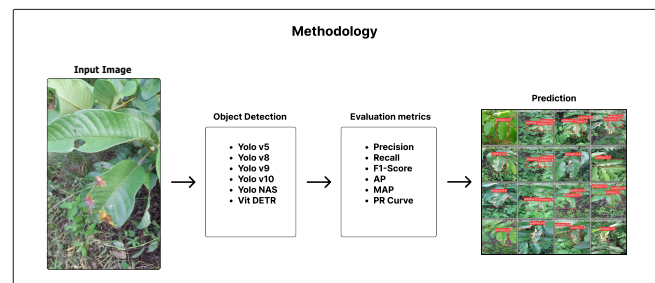


Figure 1. Steps of proposed method.

3.1 Dataset

The dataset used for the experiments was obtained from the city of Reriutaba, Ceará in the northeastern region of Brazil, specifically from an area affected by psyllid infestation. In total, 261 images were collected from 50 guava trees¹[12]. The dataset contains leaves with and without insect attacks, with variation in the number of affected areas per image, the smallest being a single positive sample. The plants used for data collection were three years old and displayed both mature

¹Available in: <https://doi.org/10.5281/zenodo.13953628>

and developing leaves. The data collection period was May 2024, during the rainy season in the region, which corresponds to the peak of insect attacks. For sample collection, mature leaves exhibiting the classic and most evident symptoms of insect presence, characterized by the curling of leaf edges, were selected. The image dimensions were 2250×4000 pixels, and the camera used for capturing the images was a Full HD with 48 megapixels. All images underwent a labeling process and were reviewed by experts. Figure 2 shows some samples included in the dataset.

3.2 Architectures

In this work, six object detection models were evaluated, selected for their ability to detect and classify in various scenarios, such as in agriculture, healthcare, and industry [13][14][15]. Another crucial factor for choosing these models was their use in market applications, with high accuracy and speed.

YOLO V5 is renowned for its capability to detect and classify objects in real-time. Among its advantages are precision, speed, and efficiency, making it popular in various applications across different niches. Its operation is based on dividing the input image into a grid, where each cell is responsible for predicting bounding boxes and object classes within that cell. Following this step, predictions are made, and the positions of the bounding boxes and probabilities are calculated. Finally, Intersection over Union (IoU) is calculated to measure the overlap between the predicted boxes and the ground truth boxes [16].

YOLO V8 is a superior version to V5, released in early 2023. This version aims to enhance both speed and precision for object detection by optimizing its architecture, which combines an efficient backbone with improvements in the detection head, as well as refined post-processing [17]. Compared to Yolo's V5, V8 shows superior results in evaluation metrics such as mAP and a shorter training time. In other words, it is faster to train and has better detection results. Also in this context, the V8 model adds to its portfolio a trend, which is to combine detection and tracking models with depth cameras, with the aim of tracking objects and people in space [18].

YOLO V9 was introduced in February 2024, surpassing previous models in terms of precision. YOLO V9 represents a significant advancement in object detection tasks, exceeding the results achieved by YOLO V8 through innovations such as Programmable Gradient Information (PGI) and the Generalized Efficient Layer Aggregation Network (GELAN) architecture. PGI allows for the preservation of input information, which is essential for the effective updating of network weights, while GELAN enhances parameter efficiency [19].

YOLO V10 was introduced in May 2024, aiming to reduce latency in response to input, improve object localization in images and videos, increase accuracy compared to previous versions, and decrease the number of model parameters [20]. This version also eliminates the use of non-maximum suppression during inference, employing consistent dual assignments

for training, which reduces inference time and improves accuracy. Consequently, YOLO V10 consolidates features from previous versions while adding significant innovations that enhance its performance.

YOLO NAS, or YOLO Neural Architecture Search, achieves up to 20% higher mAP performance compared to versions 5 and 8. Developed by a different company than the previously mentioned versions, it promises to further reduce inference time while offering greater accuracy [21].

ViT DETR is a combination of the Vision Transformer (ViT) and DEtection TRansformer (DETR) architectures designed to enhance object detection in images. ViT is a Transformer-based model initially created for image classification tasks, processing images by dividing them into patches and treating each patch as a sequence [22]. DETR, on the other hand, is a groundbreaking model for object detection that eliminates the need for intermediate steps such as region proposals and post-processing, utilizing a Transformer architecture to detect objects directly from global image representations [23].

3.3 Evaluation metrics

In the development of Deep Learning applications, it is necessary to use evaluation metrics to analyze the results obtained after training the networks, aiming to measure the performance of the models. Where the variables TP represents true positives, TN denotes true negatives, FP refers to false positives, and FN indicates false negatives [24].

- **Precision:** is a metric used to measure the accuracy with which classes are detected (Equation 1).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

- **Recall:** calculates the proportion of true positives among all actual positives, allowing the measurement of the model's ability to detect all instances of a class (Equation 2).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

- **F1-Score:** is the harmonic mean of precision and recall, providing a balanced assessment of a model's performance by considering both false positives and false negatives (Equation 3).

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

- **Average Precision (AP):** is a metric that summarizes the shape of the Precision-Recall curve into a single value by calculating the area under the curve. In other words, it uses multiple points from the curve to generate



Figure 2. Sample images from dataset.

the AP value, encompassing both Precision and Recall as defined by Equation 4.

$$AP = \int_0^1 \text{Precision}(\text{Recall}) dr \quad (4)$$

- **Mean Average Precision (MAP):** extends the concept of AP by calculating the average AP values across multiple object classes (Equation 5).

$$MAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (5)$$

4. Results and Discussion

In this section, the results obtained from the experiments conducted with the dataset created for the task of psyllid identification will be discussed. We used the resources from Google Colaboratory, configured with a Tesla T4 GPU with 16 GB of memory, 16 GB of RAM, and an Intel(R) Xeon(R) CPU @ 2.00GHz. The operating system used was Ubuntu 22.04.3 LTS.

To perform a comparative analysis between different object recognition models, we used the evaluation metrics described in the Section 3.3. These metrics served as key factors in selecting the model that demonstrated the highest efficiency in detecting psyllid attacks on guava leaves. Table 1 presents the results obtained during the training of the models. The experimental results demonstrate the varying efficacy of object detection models in identifying psyllid attacks on guava leaves, with YOLO V9 emerging as the most proficient model. Achieving a Precision of 90.90% and a recall of 86.40%, YOLO V9 not only excels in accurately identifying true positives, but also maintains a balance between precision and recall, reflected in its F1-Score of 88.59%. In contrast, both YOLO V10 and YOLO NAS exhibited perfect recall rates, yet their extremely low precision values (0.01% and 2.30%, respectively) indicate a propensity for false positives, undermining their utility in real-world scenarios. These findings

demonstrate the importance of balancing precision and recall in the development of reliable pest detection systems, positioning YOLO V9 as an alternative for future implementation in agricultural pest management strategies.

Table 1. Models results.

Model	Precision	Recall	F1-Score	AP	MAP
YOLO V5	93.14%	83.87%	88.27%	65.73%	65.73%
YOLO V8	90.27%	85.18%	87.65%	66.89%	66.89%
YOLO V9	90.90%	86.40%	88.59%	88.70%	95.00%
YOLO V10	0.01%	100.00%	1.03%	25.72%	25.72%
YOLO NAS	2.30%	100.00%	4.50%	2.30%	2.30%
ViT DETR	83.26%	88.89%	85.98%	65.77%	65.77%

Additionally, to provide a clearer analysis for future use of the detection model by small farmers and agricultural professionals, we compared main factors, such as model size, number of parameters, layers, and inference time. The results are listed in Table 2 and provide an overview of various object detection models, focusing on the key characteristics that affect their performance. Each model is listed along with its size in megabytes (MB), total number of parameters, number of layers, and inference time in milliseconds (ms).

- **Model Size:** indicates the amount of storage space the model occupies.
- **Parameters:** represents the complexity of the model; generally, a higher number of parameters suggests better performance, although it may lead to longer inference times.
- **Layers:** reflects the depth of the model, with more layers enabling more detailed feature extraction.
- **Inference Time:** shows the duration required for the model to process an image and deliver a prediction, which is critical for real-time applications.

This information helps farmers and agricultural professionals to choose a detection model by balancing accuracy with efficiency and practicality for field use.

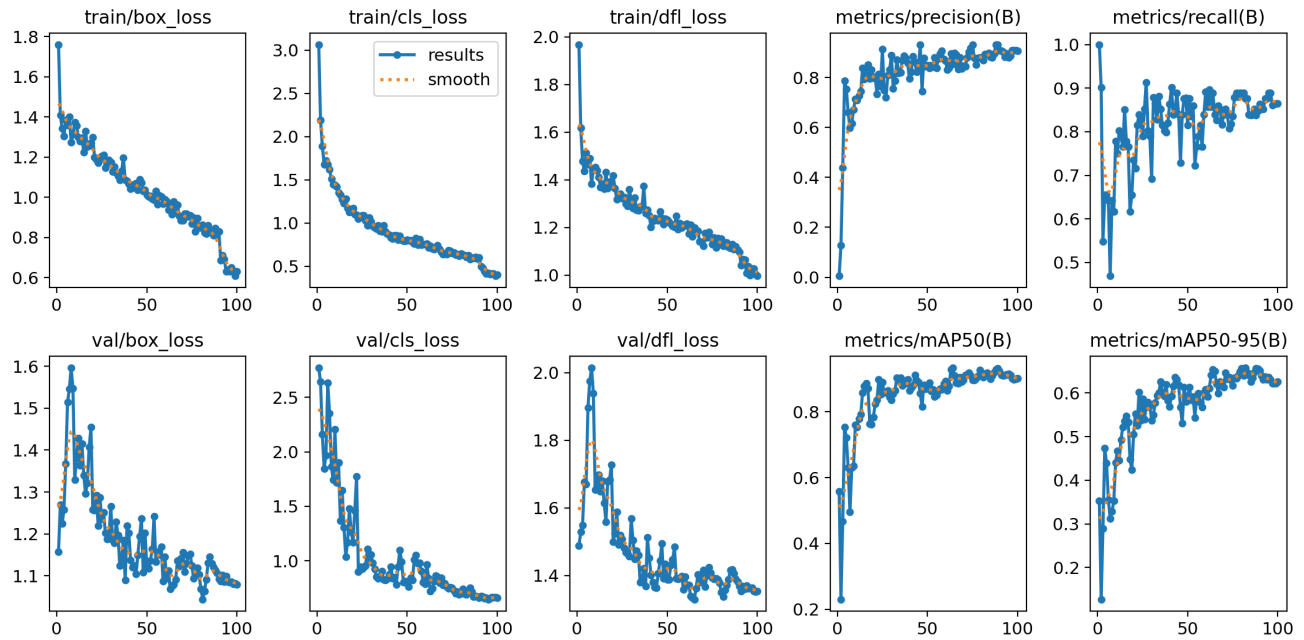


Figure 3. Training and validation performance metrics for YOLO V5.

Table 2. Summary of the characteristics of object detection models.

Model	Size Model	Parameters	Layers	Inference Time
YOLO V5	5.03Mb	2503139	193	24.1ms
YOLO V8	21.4Mb	11125971	168	28.7ms
YOLO V9	195.15Mb	25411731	467	44.1ms
YOLO V10	5.49Mb	2694806	285	16.0ms
YOLO NAS	255.86Mb	*	*	1.91s
ViT DETR	63Mb	31985795	498	46.8ms

We identify that, despite the results of the Recall, F1-Score, AP, and MAP metrics, YOLO V9 generates a model with very large and heavy weights compared to the other models and also has a high inference time. Among the versions that achieved better results, YOLO V5 not only had a smaller weight file, but also a lower number of trainable parameters and a shorter inference time. This meets the requirements for future work, where this study can serve as a reference for developing an application to assist farmers and professionals in the field.

We analyzed the behavior of YOLO V5 during the training models. As shown in Figure 3 the evaluation metrics obtained during the training epochs. The top row displays training metrics, while the bottom row shows validation metrics. The training and validation metrics for the YOLO V5 model exhibited an improvement over 100 epochs. Both the box and classification losses steadily decreased, indicating a better accuracy in object localization and classification. Precision remained high throughout, while recall improved gradually, reflecting enhanced object detection with fewer false positives and good generalization to the validation set.

Additionally, we considered the PR curve as a performance metric to evaluate the effectiveness of YOLO V5 in detecting psyllids on guava trees. Figure 4 shows the PR curve for the detection of psyllid infestations using the YOLOv5 model. The curve demonstrates a good performance, with a mAP of 0.930 at an IoU threshold of 0.5 for the psyllid class. This high precision is sustained for a wide range of recall values, indicating that the model maintains a low false positive rate even as recall increases. The model performed with minimal degradation in precision until the recall approached 0.8, where a slight drop was observed. The overall AUC indicates that the YOLO V5 model effectively balances precision and recall, making it well suited for detecting psyllid infestations across different scenarios.

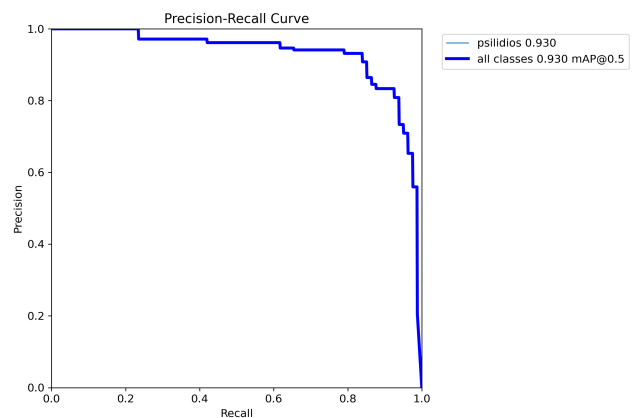


Figure 4. PR Curve.

Figure 5 shows examples of the predictions of the YOLO V5 model for detecting psyllid infestations in the guava tree

leaves. The model identified areas of infestation, with the majority of detections having a high confidence score of 0.9. Each red bounding box highlights a detected instance of psyllid damage, labeled with the class “psilidios” and the corresponding confidence score. The results demonstrate the capability of the model to detect psyllid infestations across various leaf samples, even in cases of partial damage or differing visual conditions.



Figure 5. YOLO V5 detection results of psyllid infestations on guava tree leaves

5. Conclusion

This study analyzed the use of object detection models on a dataset specifically developed for the identification of *Triozoida limbata* on guava trees. This pest affects various scales of cultivation and causes significant damage to producers. There remains a need for information technology to facilitate and automate the process of identifying infestations of these insects.

For the detection task, the methodology employed proved promising and yielded significant results. We conducted the tests using different versions of YOLO, namely V5, V8, V9, V10, and NAS, as well as a ViT. In the context of evaluation based on metrics and the analysis of feasibility for implementation in potential applications aimed at facilitating the use of the detector for farmers and industry professionals, YOLO V5 achieved the best results with Precision of 93.14%, Recall of 83.87%, F1-Score of 88.27%, and AP of 65.73%. Other versions of YOLO, such as V8, V9, and ViT DETR, also demonstrated excellent performance on the dataset used.

In addition to automating the process of identifying *Triozoida limbata* infestations as one of the contributions of this work, we can also highlight the development of the first Brazilian

dataset containing images of guava leaves infested by *Triozoida limbata*.

Despite the promising results obtained in this study, some limitations need to be addressed. The dataset size was small, which may limit the generalizability of our findings to other regions and cultivation conditions. In addition, the variability in images, such as different lighting conditions and capture angles, can affect the classification performance of the model.

In future work, we plan to expand the dataset to include a greater diversity of images, and explore other object detection architectures that may offer improvements in terms of classification performance and efficiency. Furthermore, developing a mobile application to facilitate field use can improve the practicality and adoption of this technology by farmers and agricultural professionals.

Acknowledgments

We thank the Federal University of Ceará (UFC), the Graduate Program in Electrical Engineering and Computing, and the Federal University of Viçosa (UFV) for providing accommodation, computational resources, and technical knowledge. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

Author contributions

IMM: Conceptualization, Methodology, Software, Investigation, Validation, Visualization, Writing - original draft. RSN: Investigation, Validation, Visualization, Writing - original draft. LFRM: Methodology, Validation, Visualization, Writing - Review & Editing. ICPJ: Supervision, Funding acquisition, Resources, Writing - Review & Editing.

References

- [1] SANTOS, T. P. D. et al. Natural parasitism in *triozoida limbata* (enderlein, 1918) (hemiptera: Triozidae) in a semi-arid region of brazil. *Revista Caatinga*, Universidade Federal Rural do Semi-Árido, v. 35, n. 1, p. 239–242, Jan 2022. ISSN 1983-2125. Disponível em: <https://doi.org/10.1590/1983-21252022v35n125rc>.
- [2] OLIVEIRA, G. S. d. et al. Participatory survey of citrus in the “recôncavo baiano region” with emphasis on the threat of huanglongbing. *Revista Brasileira de Fruticultura*, Sociedade Brasileira de Fruticultura, v. 43, n. 4, p. e–168, 2021. ISSN 0100-2945. Disponível em: <https://doi.org/10.1590/0100-29452021168>.
- [3] IBGE. *Produção Agropecuária*. 2023. Disponível em: <https://www.ibge.gov.br/explica/producao-agropecuaria/goiaba/br>.
- [4] RENAN, Q.; SAULO, F. Controle do psilídeo da goiabeira (*Triozoida limbata*) com o uso de medicamento

- homeopático. *Revista Científica Faesa*, v. 18, n. 1, p. 162–177, jul. 2022. ISSN 23167327. Disponível em: <http://revista.faesa.br/revista/index.php/Faesa/issue/archive>.
- [5] MOREIRA, R. et al. AgroLens: A low-cost and green-friendly Smart Farm Architecture to support real-time leaf disease diagnostics. *Internet of Things*, v. 19, p. 100570, 2022. ISSN 2542-6605. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2542660522000634>.
- [6] THAI, H.-T.; LE, K.-H.; NGUYEN, N. L.-T. Towards sustainable agriculture: A lightweight hybrid model and cloud-based collection of datasets for efficient leaf disease detection. *Future Generation Computer Systems*, v. 148, p. 488–500, 2023. ISSN 0167-739X. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0167739X23002388>.
- [7] BARBEDO, J. G. A. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Computers and Electronics in Agriculture*, v. 153, p. 46–53, 2018. ISSN 0168-1699. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0168169918304617>.
- [8] RAJBONGSHI, A. et al. A comprehensive guava leaves and fruits dataset for guava disease recognition. *Data in Brief*, v. 42, p. 108174, 2022. ISSN 2352-3409. Disponível em: <https://www.sciencedirect.com/science/article/pii/S235234092200378X>.
- [9] PATHMANABAN, P.; GNANAVEL, B.; ANANDAN, S. S. Comprehensive guava fruit data set: Digital and thermal images for analysis and classification. *Data in Brief*, v. 50, p. 109486, 2023. ISSN 2352-3409. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2352340923005863>.
- [10] MATTIHALLI, C. et al. Plant leaf diseases detection and auto-medicine. *Internet of Things*, v. 1-2, p. 67–73, 2018. ISSN 2542-6605. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2542660518300453>.
- [11] SHETTY, K. U. et al. Plant Disease Detection for Guava and Mango using YOLO and Faster R-CNN. In: *2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*. Gwalior, India: IEEE, 2024. v. 2, p. 1–6.
- [12] MESQUITA, I. M. de et al. *Psyllid Attacks in Guava*. Zenodo, 2024. Disponível em: <https://zenodo.org/doi/10.5281/zenodo.13953629>.
- [13] BADGUJAR, C. M.; POULOSE, A.; GAN, H. Agricultural object detection with You Only Look Once (YOLO) Algorithm: A bibliometric and systematic literature review. *Computers and Electronics in Agriculture*, v. 223, p. 109090, 2024. ISSN 0168-1699. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0168169924004812>.
- [14] BERGMANN, M. A. et al. An Approach Based on LiDAR and Spherical Images for Automated Vegetation Inspection in Urban Power Distribution Lines. *IEEE Access*, v. 12, p. 105119–105130, 2024.
- [15] VIJAYAKUMAR, A.; VAIRAVASUNDARAM, S. YOLO-based Object Detection Models: A Review and its Applications. *Multimedia Tools and Applications*, v. 83, n. 35, p. 83535–83574, Oct 2024. ISSN 1573-7721. Disponível em: <https://doi.org/10.1007/s11042-024-18872-y>.
- [16] AL., G. J. et. *YOLOv5 by Ultralytics*. 2020. Disponível em: <https://github.com/ultralytics/yolov5>.
- [17] AL., G. J. et. *Ultralytics YOLO*. 2023. Original-date: 2022-09-11T16:39:45Z. Disponível em: <https://github.com/ultralytics/ultralytics>.
- [18] SIRISHA, U. et al. Statistical Analysis of Design Aspects of Various YOLO-Based Deep Learning Models for Object Detection. *International Journal of Computational Intelligence Systems*, v. 16, n. 1, p. 126, Aug 2023. ISSN 1875-6883. Disponível em: <https://doi.org/10.1007/s44196-023-00302-w>.
- [19] WANG, C.-Y.; YEH, I.-H.; LIAO, H.-Y. M. *YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information*. 2024. Disponível em: <https://arxiv.org/abs/2402.13616>.
- [20] WANG, A. et al. *YOLOv10: Real-Time End-to-End Object Detection*. 2024. Disponível em: <https://arxiv.org/abs/2405.14458>.
- [21] AHARON, S. et al. *Deci-AI/super-gradients: 3.0.8*. Zenodo, 2023. Disponível em: <https://zenodo.org/records/7789328>.
- [22] DOSOVITSKIY, A. et al. *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. 2021. Disponível em: <https://arxiv.org/abs/2010.11929>.
- [23] CARION, N. et al. End-to-end object detection with transformers. In: VEDALDI, A. et al. (Ed.). *Computer Vision – ECCV 2020*. Cham: Springer International Publishing, 2020. p. 213–229.
- [24] DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification (2nd Edition)*. USA: Wiley-Interscience, 2000. ISBN 0471056693.