Colorado Potato Beetle Dataset and Detection for Monitoring and Management in Potato Fields

Thomas Leopold
Institute for Computer Technology
TU Wien
Gußhausstraße 27-29/E384
thomas.leopold@tuwien.ac.at

Axel Jantsch
Institute for Computer Technology
TU Wien
Gußhausstraße 27-29/E384
axel.jantsch@tuwien.ac.at

Abstract

The Colorado potato beetle (Leptinotarsa decemlineata) remains a significant threat to potato crops worldwide, imposing substantial economic losses and challenging sustainable agricultural practices. Manual pest control methods are labor-intensive, inefficient, and often insufficient to prevent widespread infestations. To address this challenge, we propose automated pest detection, have developed a labeled dataset (POBED), and studied several object detection models. Particularly YOLOv6 and CO-DETR, demonstrated promising performance in identifying Colorado potato beetle (CPB) stages, with $AP^{IoU=.50} = 72.7\%$ for YOLOv6 and $AP^{IoU=.50} = 79.2\%$ for CO-DETR. Despite challenges with background elements and labeling inconsistencies, this research highlights the potential of such models for generating detailed infestation maps and guiding targeted pest control strategies. Further refinement and exploration, including integration with autonomous removal mechanisms, offer exciting avenues for enhancing pest management efficiency and sustainability in agriculture.

1 Introduction

The CPB poses a significant threat to potato cultivation, with its infestations causing substantial reductions in crop yield and quality Alyokhin [2009]. Conventional pest control methods, such as pesticides and mechanical removal, are often ineffective and can have detrimental environmental consequences Casagrande [1987], Alyokhin et al. [2022].

Traditional agricultural practices, reliant on large machinery, raise environmental concerns like soil erosion, biodiversity loss, and energy consumption Batey [2009], Nawaz et al. [2013]. In contrast, smart farming offers a more sustainable alternative. Combining AI and automation, lightweight autonomous robots minimize soil disturbance, reduce pesticide use, and enhance efficiency Pearson et al. [2022].

By integrating AI-powered autonomous robots into smart farming practices, CPB infestations can be effectively controlled while minimizing environmental impact and enhancing agricultural sustainability. This approach can potentially revolutionize pest control and pave the way for a more environmentally friendly and productive agricultural future.

Therefore, this paper presents a CPB dataset called POtato BEetle Dataset (POBED) and employs object detection models to address these challenges.

The First Austrian Symposium on AI, Robotics, and Vision (AIROV24).

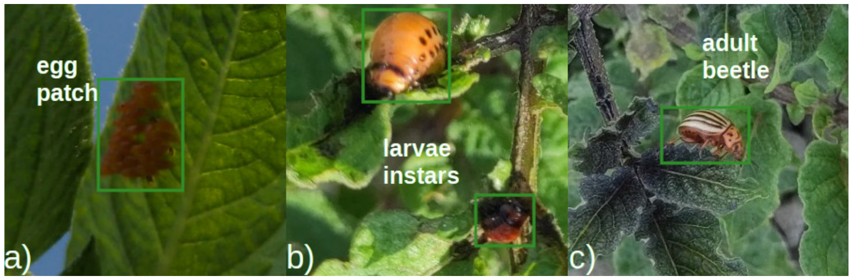

Figure 1: This concise visual illustrates the three distinct stages of the Colorado potato beetle's life cycle: a) eggs, b) larvae, and c) adults.

Key findings:

- All models achieved moderate performance on the stricter IoU metric (AP: 36.9%-44.1%). Using a less strict metric ($AP^{IoU=.50}$), performance reached 66%-79.2%, approaching the theoretical maximum based on the data quality of the dataset.
- YOLOv6 and Co-DETR models outperformed others, having potential for applications.
- Co-DETR has class-specific limitations in detecting larvae, possibly due to architecture.
- Small objects (first-instar larvae) pose significant challenges due to size and coloration.

Contributions:

1. POBED Dataset: We introduce POBED, a unique dataset specifically designed for training and evaluating object detection models for CPB stage identification, published on `https://zenodo.org/records/10599211`.
2. Benchmarking Object Detection Models: We benchmark state-of-the-art object detection methodologies, encompassing R-CNN, YOLO, and transformer architectures, on POBED.

2 Related work

2.1 Datasets

In agricultural technology, particularly in pest management, the detection of insects using advanced techniques has been a subject of ongoing research. A key component in developing these technologies is the availability of comprehensive and high-quality datasets. The IP102 dataset Wu et al. [2019] contains 75000 images of 102 species of pest insects. For object detection, 19000 images were annotated with bounding boxes. The rest are for classification only.

The dataset by Kuznetsov [2023] marks a significant advancement in pest detection within agricultural technology, focusing specifically on the Colorado potato beetle (CPB). It features 1810 images depicting the beetle in various stages, namely the adult and larvae stages, excluding the egg stage. An essential aspect of this dataset is its timestamp - from 2023, it offers contemporary insights into CPB detection.

The dataset provided by Sittinger [2023] Insect dataset for flying insects like wasps, bees, flies, hoverflies, and even shadows of the recorded insects. Different resolutions are provided.

To the best of our knowledge, no object detection dataset of this size encapsulates the whole growth process of potato beetles, i.e., covering the egg, larvae, and beetle stage at once. To detect all stages, especially the egg stage, we propose the new dataset POBED.

2.2 Insect detection

Recognizing and detecting pests on plants is essential in combatting the latter. In Ahmad et al. [2022], Verma et al. [2021], Lippi et al. [2021], Dai et al. [2023], Likith et al. [2021] you only look once (YOLO) based approaches are used to detect objects ranging from insects to bugs. The proposed

Figure 2: Illustration of leaves with eggs in the infrared spectrum. a) in controlled conditions, b) in direct sunlight from above, and c) in direct sunlight from below.

approaches use different advances in the YOLO detection classes. In Ahmad et al. [2022], the YOLOv5 model achieved the best results detecting 23% of insect pests with an $mAP@[0.5 : 0.95]$ of 79.8%. In Du et al. [2020], Choiński et al. [2023], faster region-based convolutional neural network (R-CNN) based models are used to detect insects.

3 Dataset

3.1 Dataset design

The POBED dataset (processed) can be found at Zenodo [1] and provides a comprehensive collection of 7192 images with 9640 labels for three classes. It features individual plants and plant sections, encompassing various stages of the CPB. These images capture the presence of multiple CPB instances, with most objects representing larvae, followed by beetles and a smaller number of egg clusters. The CPB exhibits distinct behavior patterns during its various developmental stages, necessitating a multi-pronged approach to effectively track and monitor its presence.

Detection of the different stages of the CPB

The fully grown CPB stage poses the most straightforward detection challenge due to its large size and recognizable pattern of yellow/orange and black stripes, allowing for precise object detection from above. These beetles typically reside on the upper surface of potato leaves, facilitating their identification from a vantage point. An illustration of the beetle captured from above can be found in Figure 1 c).

The larvae stage, encompassing four distinct substages, presents a more nuanced detection challenge. The last larval stage resembles the adult beetle in size but exhibits a different color pattern, featuring orange/yellow coloring with black dots along its sides. The second larval stage is smaller and darker, while the first is only a few millimeters long, partially orange, and partly black. Despite their smaller size, larvae are relatively mobile and consume a significant portion of potato plant material. While they can primarily be detected from above, their increased movement patterns may lead them to reside on stems or even the undersides of leaves. An illustration of larvae in two different instars can be found in Figure 1 b).

Early detection of egg patches is crucial for effective CPB control, as the first larval stage is challenging to identify due to its minuscule size. One approach investigated was near-infrared photography, utilizing filters attached to a smartphone camera. The higher transmission of infrared light compared to visible light was expected to enable differentiation between egg patches and leaves when exposed to infrared light from below. This method yielded promising results in controlled settings but proved less effective in field conditions. The interference from sunlight completely overpowered the infrared light source, and even with sun shading, scattered light interfered with the detection process. Consequently, an alternative approach is required. See Figure 2

The following approach employed angled cameras mounted on a stick to capture images of the undersides of potato leaves from below, allowing for a closer examination of the plant surface. This

[1] https://zenodo.org/records/10599211

Figure 3: Wheelbarrow and telescopic arm for capturing images in the potato fields.

method produced promising results, revealing several egg patches readily identifiable by hand. An illustration of egg patches on the underside of the leaves can be found in Figure 1 a).

3.2 Recording setup

The images were taken using multiple techniques. Four smartphones, including different resolution settings, were used to increase the diversity in image quality and improve the generalization by reducing the influence on one specific sensor and setting. Most images are taken over six days at three locations in Lower Austria in 2022 and 2023. The cameras used are the primary cameras of four different smartphones, including the Samsung Galaxy S20 Plus with a 12 MP, f/1.8, 26mm (wide), 1/1.76", 1.8µm, Dual Pixel phase detection auto focus (PDAF), OIS Sensor, the Google Pixel 4a 5G with a 12.2 MP, f/1.7, 27mm (wide), 1/2.55", 1.4µm, dual pixel PDAF, optical image stabilization (OIS) Sensor, the ZTE Axon 7 with a 20 MP, f/1.8, 1/2.6", 1.1µm, PDAF, OIS Sensor and the Fairphone 4 with a 48 MP, f/1.6, (wide), 1/2.0", 0.8µm, PDAF, OIS Sensor. Due to the necessity of covering different camera angles to detect the majority of the beetle stages, two mounts were constructed. A small wheelbarrow is modified to capture images while driving through the potato fields with mountings in the front, on each side, and at variations. The constructed parts for the top view camera holder are based on a telescopic arm mounted to the back of the wheelbarrow protruding from the vehicle on the front. The actual mount for the camera can be rotated 360° if needed to cover the tracks and the potato plants from the top left and right.

Table 1: Summary of the POBED Dataset

Category	Information	Additional
Images	7192	
Classes	3	CPB stages
Class 1	CPB Beetle	723 objects
Class 2	CPB Larvae	8730 objects
Class 3	CPB Eggs	187 objects
Number of labels	9,640	
Resolution processed	640x540	7192 Images
Resolutions raw	1080p, 2268p, 3024p, 3348p, 3000p, 2156p	3564 images
Capturing devices	smartphones	4 different sensors
Time	2022 and 2023	spring/summer
Data source	3 potato fields	
Field locations	Lower Austria	Wald-&Weinviertel

3.3 Ground truth generation

The open-source software CVAT CVAT.ai Corporation [2023] manages the labeling process. The captured images are split into tasks depending on the capturing device and the time of taking the picture. Jobs comprised of 20 to 30 images are distributed to ten annotators labeling the images. Each annotator was introduced to the task and was supervised by a CPB specialist during the first job. After the jobs are fulfilled, they are reviewed on a test basis. In addition to the review process, ground truth was generated by labeling 10% of the data a second time. A semi-automatic labeling setup was created to employ the trained detection models in the future. For the annotation of the dataset, three classes were defined to differentiate between the different stages of the CPB depicted in 1 showing ann egg patch, larvae instars, and the adult beetle. The instars are not differentiated due to the high resolution needed for differentiation.

3.4 Data quality

The integrated quality assessment of CVAT is used to assess the quality of the labeled data. This uses the ground truth labeled data and compares them with the labeled data from the jobs in each task. This produces different metrics that show the quality of the data. The primary metric is the mean annotation quality (MAQ). It consists of correct annotations, task annotations, Ground Truth annotations, accuracy, precision, and recall. The average of the MAQ of all the labeling tasks is around 80%. Most issues are due to low overlap, determined by the bounding box having less than 80% intersection over union (IoU). The other problems arise from additional or missing labels. In images with tiny objects, the chance is high that some objects are missed, whereas others not labeled in the task are found. Additionally, it is hard to determine whether the object should be labeled in some images due to low resolution and blur.

Table 2: Labeling quality analysis data

Subset	MAQ	Correct Ann.	Task Ann.	GT Ann.	Accuracy	Precision	Recall
Subset 1	**78.4%**	403	447	470	78.4%	85.7%	90.2%
Subset 2	**89.5%**	111	114	121	89.5%	91.7%	97.4%
Subset 3	**72.2%**	368	443	435	72.2%	84.6%	83.1%

4 Evaluation & results

For Comparison, three object detection approaches are used. A YOLO-based network approach uses a YOLOv6 Li et al. [2022] and a YOLOv7 Wang et al. [2022] model. The second type is an R-CNN based network, the Sparse-RCNN Sun [2024]. We use a transformer-based network called CO-DETR X-Lab [2024] for the last type.

For the training of the models, the dataset was split into training, validation, and test sets in the ratio of 7:2:1. The split was not taken at random to prevent images with similar backgrounds and objects from being placed in the same set. Furthermore, the difference in training and testing images is even higher as the days and conditions differ. The dataset is used in two ways for efficient use and training. The images were used with the recorded resolution for the first training batch. The models were trained with 640px train and test resolution and afterward with 1280px resolution as a comparison. For the final training data used in this paper, the images were converted to the same resolution by cropping 640x540px images out of the higher resolution images and discarding all the images with no objects. Discarding the empty images is necessary as the model training couldn't work with the large number of empty images, and the training time increased 10-fold. After preprocessing the images to the desired resolution and discarding the empty images the second training batch evaluates all the mentioned images on the processed dataset. The ratio of images changes as the objects in the images are not equally distributed over the raw images. For Comparison, the ratio stayed untouched.

4.1 Metrics used

The COCO evaluation metrics are employed to assess the performance of object detection models using the datasets mentioned. These metrics provide a robust framework for evaluating various aspects of model quality in object detection tasks. The most important metric is average precision (AP) defined in the coco metric. AP is averaged over ten different IoU values starting from 0.5 and going up to 0.95. The standard AP is also in the metric defined as $AP^{IoU=0.5}$. The second important metric is the average recall (AR). Both the AP and the AR are calculated for different object sizes, and AR is also given for other numbers of objects per image. Table 3 shows the performance of the models on the COCO evaluation tasks.

In addition to the coco metric, a confusion matrix is used to show the performance of the models in every predetermined class. It shows the percentage of a class detected as the actual or other classes, including the background. The default confusion matrix determines whether a positive detection of a model is a true or false detection and vice versa Narkhede [2021].

Table 3: COCO metrics coco [2024]

average precision (AP):	
AP	AP at IoU=.50:.05:.95 (primary challenge metric)
$AP^{IoU=.50}$	AP at IoU=.50 (PASCAL VOC metric)
$AP^{IoU=.75}$	AP at IoU=.75 (strict metric)
AP Across Scales	
AP^{small}	AP for small objects: area < 32^2
AP^{medium}	AP for medium objects: 32^2 < area < 96^2
AP^{large}	AP for large objects: area > 96^2
average recall (AR)	
$AR^{max=1}$	AR given 1 detection per image
$AR^{max=10}$	AR given 10 detection per image
$AR^{max=100}$	AR given 100 detection per image
AR across scales:	
AP^{small}	AR for small objects: area < 32^2
AP^{medium}	AP for medium objects: 32^2 < area < 96^2
AP^{large}	AP for large objects: area > 96^2

4.2 Model performance

The model performance is evaluated for implementations of R-CNN, YOLO, and Detection Transformer (DETR). For the training of the models, the MMDETECTION MMDetection Contributors [2018] framework is used for all the models. For the YOLO models, MMYOLO, a specialized branch of the MMDETECTION framework is used. Different model variants exist concerning size and computational effort, so the following list gives the details of the models.

Table 4: Models used with configured options

Model	Type	Scale	Epochs	Adaptations
Yolov6	yolov6_t_syncbn_fast	640	300	-
Yolov7	yolov7_tiny_syncbn_fast	640	300	-
Sparse-RCNN	sparse-rcnn_r50_fpn_ms-480-800	multiscale	36	-
Co-DETR	co_dino_5scale_swin_l	multiscale	36	lr = 2e-5

The detection results in Table 5 show detection of the objects in 66% to 79.2% of the cases with an IoU of 50%. When the stricter metric AP that uses ten different IoU values is used, the performance is reduced to the range of 36.9% up to 44.1%. This induces the general detection of the classes with unprecise bounding boxes. This could be due to mismatches in labeling and issues with identifying parts of the object, like legs and antennae, due to low resolution and matching background. According to the quality of the labeled data, there are issues with the intersection of the bounding boxes when labeled twice. Therefore, it reduces the AP. When high IoU is of no concern, the AP at an IoU of 50% almost reaches the precision of the labeled data. The models with the best performance are the YOLOv6 and the CO-DETR model. The chosen YOLOv7 model has the least parameters and seems to lack to achieve better results due to that which corresponds with the COCO detection quality in the COCO pre-trained models from MMDETECTION. The observation regarding the performance of the CO-DETR model, as evaluated through the confusion matrix, highlights a specific aspect of its effectiveness on a class-by-class basis. While the model demonstrates overall strong performance, this effectiveness is notably diminished when detecting the larvae class of pests. Since the other models do not demonstrate such differentiation in performance across different classes, the discrepancy observed in the CO-DETR model's handling of the larvae class is unlikely to originate from poor-quality training data. This indicates that the cause might be inherent to the specific architecture or processing techniques of the CO-DETR model itself. All the models show a high relative percentage of false positives for larvae. They are the smallest, especially in the first instars, hard to differentiate from the background, and the most overlooked while labeling.

Table 5: COCO metrics for POBED processed test set on all trained models

average precision (AP):	Yolov6	Yolov7	Sparse-RCNN	Co-DETR
AP	0.406	0.369	0.386	0.441
$AP^{IoU=.50}$	0.727	0.660	0.694	0.792
$AP^{IoU=.75}$	0.420	0.345	0.383	0.438
AP Across Scales				
AP^{small}	0.231	0.161	0.191	0.242
AP^{medium}	0.531	0.484	0.514	0.561
AP^{large}	0.570	0.429	0.529	0.615
average recall (AR)				
$AR^{max=1}$	0.401	0.370	0.412	0.436
$AR^{max=10}$	0.528	0.509	0.525	0.567
$AR^{max=100}$	0.549	0.520	0.558	0.613
AR Across Scales:				
AR^{small}	0.387	0.324	0.377	0.410
AR^{medium}	0.644	0.632	0.665	0.725
AR^{large}	0.637	0.645	0.544	0.713

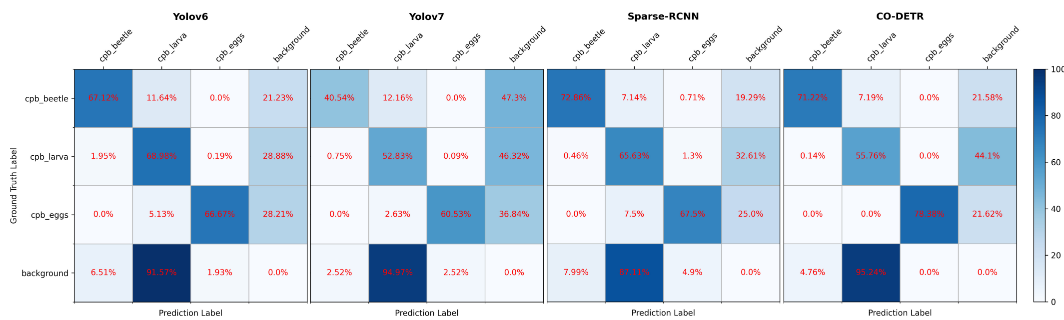

Figure 4: Confusion matrices for the trained models.

4.3 Qualitative evaluation/salience map ablation study

Figure 5 shows four different images with the corresponding gradcam visualization of the object detection of the YOLOv7 model. In image a), two objects are detected: a beetle and a larva. The beetle is detected by its stripes and correctly identified, whereas a consumed stem is wrongly identified as a larva. This could be due to the stem's coloring and the fact that larvae are typically found in that location. In the second image, b) both objects are detected, but the beetle is detected as a larva as the model looks at the front of the beetle, not at the stripes. The front of the beetle has patterns and colors similar to the larva and can, therefore, be misinterpreted. In the third image, c) both larvas are correctly identified by the orange and black variation at the head of the larva. In the last image, d) the egg patch is correctly identified, mainly looking at the center of the patch.

4.4 Size ablation study

The trained models perform differently concerning the size of the object. In the first approach to detecting the object, the raw dataset was used with a resolution of 640px. This rendered rather unsatisfying results as many objects were too small to be represented. The same model was fed with 1280px input resolution in the next step. This increased the results to an acceptable level, but the resources needed for training and interference are significantly higher. Therefore, the processed images with a resolution of 640px are used for the final models. When looking at the metrics of the models in Table 5, the data shows bad performance at small scales. This can be explained by the small area and pixels representing the object and the less colorful pattern the objects of this size exhibit. Objects of this size are mainly larvae of the first instars that are a few millimeters in size.

Figure 5: Comparison of original images with a gradcam visualization of the YOLOv7 model.

This is another indicator that the detection of the egg patches should be the primary goal to keep the infestation low.

5 Conclusion

This study presents a novel approach for detecting various CPB stages in potato fields, contributing significantly to precision agriculture and integrated pest management (IPM). The foundation of this research lies in creating the POBED dataset, meticulously designed to facilitate CPB stage identification using object detection models. This dataset is a valuable resource for advancing pest detection technologies.

We evaluated the suitability of various object detection models for CPB stage identification within the labeled POBED dataset. The analysis considered the dataset's labeling quality, particularly the observed 80% mean annotation quality and prevalent issues like bounding box misalignment and label inconsistencies.

The results suggest that YOLOv6 and CO-DETR models achieved performance levels nearing theoretical limits based on the dataset quality. Real-world field applications involve continuous imaging, capturing objects at varying scales, and potentially enhancing detection probabilities. This redundancy may compensate for individual model missed detections. However, challenges persist in distinguishing objects from background elements. Missed annotations during labeling, as evident from the data analysis, likely contribute to this issue. Discolored vegetation from diseased plants and soil hues can also confuse the algorithms. While plant discoloration might be significant only later in the growing season, soil interference can be addressed by adjusting camera angles or pre-processing images for soil removal before object detection.

This study explored diverse detection strategies, including multi-angle and near-infrared (NIR) approaches. While NIR detection showed potential, it did not produce the anticipated results, highlighting the complexities and challenges in developing effective pest detection systems for agricultural settings.

Looking ahead, the application of object detection models holds immense promise for generating detailed infestation maps of fields. Such maps could empower farmers with data-driven insights to make informed decisions regarding targeted pest control strategies tailored to the specific infestation level and nature. Furthermore, integrating these detection models with autonomous frameworks could revolutionize pest management, enabling the detection and autonomous removal of beetles. The potential synergy between autonomous detection and physical removal mechanisms opens exciting avenues for future research and development. This integrated approach could significantly enhance the efficiency and efficacy of pest control methods, reducing labor costs and minimizing environmental impact. Future work in this area will focus on refining these technologies, enhancing their accuracy and reliability, and exploring their potential to transform agricultural practices by providing more sustainable and effective solutions for pest management.

Acknowledgments and Disclosure of Funding

We would like to thank all the people helping to make this work possible, especially those collaborating in creating, labeling, and supervising the dataset.

References

Iftikhar Ahmad, Yayun Yang, Yi Yue, Chen Ye, Muhammad Hassan, Xi Cheng, Yunzhi Wu, and Youhua Zhang. Deep Learning Based Detector YOLOv5 for Identifying Insect Pests. *Applied Sciences*, 12(19):10167, January 2022. Number: 19 Publisher: Multidisciplinary Digital Publishing Institute.

Andrei Alyokhin. Colorado Potato Beetle Management on Potatoes: Current Challenges and Future Prospects. (1), 2009.

Andrei Alyokhin, Galina Benkovskaya, and Maxim Udalov. Chapter 4 - Colorado potato beetle. In Andrei Alyokhin, Silvia I. Rondon, and Yulin Gao, editors, *Insect Pests of Potato (Second Edition)*, pages 29–43. Academic Press, January 2022.

T. Batey. Soil compaction and soil management – a review. *Soil Use and Management*, 25(4):335–345, 2009. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-2743.2009.00236.x.

R. A. Casagrande. The Colorado Potato Beetle: 125 Years of Mismanagement. *Bulletin of the Entomological Society of America*, 33(3):142–150, September 1987.

Mateusz Choiński, Marcin Zegarek, Zuzanna Hałat, Tomasz Borowik, Jenna Kohles, Melina Dietzer, Katrine Eldegard, Reed April McKay, Sarah E. Johns, and Ireneusz Ruczyński. Insect Detection on High-Resolution Images Using Deep Learning. In Khalid Saeed, Jiří Dvorský, Nobuyuki Nishiuchi, and Makoto Fukumoto, editors, *Computer Information Systems and Industrial Management*, Lecture Notes in Computer Science, pages 225–239, Cham, 2023. Springer Nature Switzerland.

coco. COCO - Common Objects in Context, 2024.

CVAT.ai Corporation. Computer Vision Annotation Tool (CVAT), November 2023. original-date: 2018-06-29T14:02:45Z.

Min Dai, Md Mehedi Hassan Dorjoy, Hong Miao, and Shanwen Zhang. A New Pest Detection Method Based on Improved YOLOv5m. *Insects*, 14(1):54, January 2023. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

Yunpan Du, Yang Liu, and Nianqiang Li. Insect Detection Research in Natural Environment Based on Faster-R-CNN Model. In *Proceedings of the 2020 5th International Conference on Mathematics and Artificial Intelligence*, ICMAI '20, pages 182–186, New York, NY, USA, May 2020. Association for Computing Machinery.

Andrey Kuznetsov. PotatoBeetleAI, 2023.

Chuyi Li, Lulu Li, Hongliang Jiang, Kaiheng Weng, Yifei Geng, Liang Li, Zaidan Ke, Qingyuan Li, Meng Cheng, Weiqiang Nie, Yiduo Li, Bo Zhang, Yufei Liang, Linyuan Zhou, Xiaoming Xu, Xiangxiang Chu, Xiaoming Wei, and Xiaolin Wei. YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications, September 2022. arXiv:2209.02976 [cs].

Sarapu Likith, B. Rishidhar Reddy, and K. Sripal Reddy. A Smart System for Detection and Classification of Pests Using YOLO AND CNN Techniques. In *2021 International Conference on Computational Performance Evaluation (ComPE)*, pages 049–052, December 2021.

Martina Lippi, Niccolò Bonucci, Renzo Fabrizio Carpio, Mario Contarini, Stefano Speranza, and Andrea Gasparri. A YOLO-Based Pest Detection System for Precision Agriculture. In *2021 29th Mediterranean Conference on Control and Automation (MED)*, pages 342–347, June 2021. ISSN: 2473-3504.

MMDetection Contributors. OpenMMLab Detection Toolbox and Benchmark, August 2018. original-date: 2018-08-22T07:06:06Z.

Sarang Narkhede. Understanding Confusion Matrix, June 2021.

Muhammad Farrakh Nawaz, Guilhem Bourrié, and Fabienne Trolard. Soil compaction impact and modelling. A review. *Agronomy for Sustainable Development*, 33(2):291–309, April 2013.

Simon Pearson, Tania Carolina Camacho-Villa, Ravi Valluru, Oorbessy Gaju, Mini C. Rai, Iain Gould, Steve Brewer, and Elizabeth Sklar. Robotics and Autonomous Systems for Net Zero Agriculture. *Current Robotics Reports*, 3(2):57–64, June 2022.

Maximilian Sittinger. Image dataset for training of an insect detection model for the Insect Detect DIY camera trap, 2023.

Peize Sun. PeizeSun/SparseR-CNN, January 2024. original-date: 2020-11-19T07:37:38Z.

Shani Verma, Shrivishal Tripathi, Anurag Singh, Muneendra Ojha, and Ravi R Saxena. Insect Detection and Identification using YOLO Algorithms on Soybean Crop. In *TENCON 2021 - 2021 IEEE Region 10 Conference (TENCON)*, pages 272–277, December 2021. ISSN: 2159-3450.

Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, July 2022. arXiv:2207.02696 [cs].

Xiaoping Wu, Chi Zhan, Yu-Kun Lai, Ming-Ming Cheng, and Jufeng Yang. IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8779–8788, Long Beach, CA, USA, June 2019. IEEE.

SenseTime X-Lab. Sense-X/Co-DETR, February 2024. original-date: 2022-11-19T14:39:11Z.