INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Automatic monitoring of diseases and pests in tomato crops

Tiago André Raposo Domingues

Master in Computer Engineering

Supervisor:
PhD Tomás Gomes Silva Serpa Brandão, Assistant Professor
Iscte - Instituto Universitário de Lisboa

Co-Supervisor:
PhD João Carlos Amaro Ferreira, Assistant Professor with habilitation
Iscte - Instituto Universitário de Lisboa

October, 2022

iscte
TECHNOLOGY
AND ARCHITECTURE

Department of Information Science and Technology

Automatic monitoring of diseases and pests in tomato crops

Tiago André Raposo Domingues

Master in Computer Engineering

Supervisor:
Doctor Tomás Gomes Silva Serpa Brandão, Assistant Professor
Iscte - Instituto Universitário de Lisboa

Supervisor:
Doctor João Carlos Amaro Ferreira, Assistant Professor with habilitation
Iscte - Instituto Universitário de Lisboa

October, 2022

Acknowledgements

I would want to thank my family, particularly my parents and sister, for the education they provided for me as well as their support and encouragement during my academic journey. Without them it would be impossible for me to achieve what I have been achieving.

I would also like to thank my girlfriend Maria for being by my side and facilitating everything with her encouragement, enthusiasm, and unwavering companionship throughout my academic path. Her presence was essential at every stage of my journey.

To my supervisors, Dr. Tomás Brandão and Dr. João Ferreira, I have to express my deepest thanks for all the support, time allocated and mentorship provided. Their knowledge, guidance and monitoring were fundamental for this dissertation to be developed.

Finally, I would like to thank INOV for all the conditions it provided for the development of the project proposed in this dissertation, in particular to the researcher Ricardo Ribeiro for all the support and knowledge provided. This assistance was critical for the proposed work to be developed.

Tiago Domingues

Resumo

Considerando o crescimento populacional, prevê-se a necessidade do aumento da produtividade das culturas. Contudo pragas e doenças são um grande obstáculo deste aumento. Por conseguinte, é importante desenvolver métodos tecnológicos que ultrapassem este obstáculo em culturas como as do tomate, que são valiosas fontes de vitaminas e minerais.

Neste sentido, esta dissertação associou-se ao projeto europeu ANDANTE cujo objetivo é automaticamente prever o aparecimento de doenças e pestes com base em dados meteorológicos e de imagens de armadilhas de insetos, multiespectrais aéreas e de plantas ao nível do solo de plantações de tomate. Como tal, esta dissertação teve o objetivo de desenvolver quatro módulos que criam métricas, relativas aos dados, que potenciam a predição. Estabeleceu-se ainda o objetivo de realizar uma revisão de literatura para a sistematização do conhecimento existente, sendo publicado um artigo. Parte dos dados utilizados nesta dissertação foram providenciados por parceiros do ANDANTE.

O primeiro módulo consistiu num sistema de informação *web* que otimiza o acesso e análise dos dados e a aplicabilidade dos restantes módulos. O segundo consistiu num modelo de deteção e contagem de insetos nas armadilhas cujo melhor resultado foi 94.4% de *mAP_0.5* no *YoloV5x*. O terceiro consistiu na verificação, através de processamento de imagem, investigação e dados providenciados, dos índices de vegetação apropriados e adquiríveis. No quarto, devido às condições dos dados das imagens ao nível do solo utilizaram-se dados do PlantVillage para classificar a saúde de folhas de tomate utilizando-se transferência de conhecimento, sendo 98% de precisão no *ResNet152V2* o melhor resultado.

Palavras-chave: Agricultura inteligente; agricultura de precisão; aprendizagem de máquinas; inteligência artificial; doenças e pragas de plantas; classificão; deteção; previsão; plantações de tomate.

Abstract

Considering population growth, the need for increased crop productivity is expected. However diseases and pests are a major obstacle to this increase. Therefore, it is important to develop technological methods that overcome this obstacle in crops such as tomatoes, which are valuable sources of vitamins and minerals.

This dissertation was associated with the European project ANDANTE whose objective is to automatically predict the appearance of diseases and pests based on meteorological data and insect traps images, aerial multispectral images and ground-level plant images of tomato crops. As such, this dissertation aimed to develop four modules that create metrics, relative to the data, that enhance the prediction. It was also established the objective of carrying out a literature review for the systematization of the existing knowledge, resulting in the publication of an article. Part of the data used in this dissertation was provided by ANDANTE partners.

The first module consisted in an web information system that optimises data access and analysis and the applicability of the remaining modules. The second consisted of a model of detection and counting of insects in traps whose best result was 94.4% of *mAP_0.5* on YoloV5x. The third consisted of verification, through image processing, investigation and provided data, of the appropriate and acquireable vegetation indices. In the fourth, due to the conditions of the ground level image data the PlantVillage data-set was used to classify the health of tomato leaves using transfer learning, with 98% accuracy in ResNet152V2 representing the best result.

Keywords: Smart farming; precision agriculture; machine learning; artificial intelligence; plant diseases and pests; tomato crops; classification; detection; forecasting.

Contents

List of Figures

List of Tables

Glossary

AI - Artificial Intelligence

API - Application Programming Interface

ANDANTE - Ai for New Devices ANd Technologies at the Edge

ANN - Artificial Neural Network

BPNN - Back-Propagation Neural Network

BNDVI - Blue Normalized Difference Vegetation Index

CNN - Convolutional Neural Network

CVAT - Computer Vision Annotation Tool

DaP - Diseases and Pests

DSRM - Design Science Research Methodology

DT - Decision Tree

EXIF - Exchangeable Image File Format

GAs - Genetic Algorithms

GBNDVI - Green and Blue Normalized Difference Vegetation

GNDVI - Green Normalized Difference Vegetation

GRNDVI - Green and Red Normalized Difference Vegetation

GRNN - Generalized Regression Neural Network

H2020 - European Union's Horizon 2020

INOV - INOV-INESC Innovation

IoM - Intersection-over-Minimum

IoT - Internet of Things

IoU - Intersection-over-Union

JU - ECSEL Joint Undertaking

KNN - k-nearest neighbor

LSTM - Long Short Term Memory

MAE - Mean Absolute Error

mAP - mean Average Precision

ML - Machine Learning

NB - Naïve Bayes

NDRE - Normalized Difference Red Edge Index

NDVI - Normalized Difference Vegetation Index

NIR - Near Infra-Red

NMS - Non-Maximum Suppression

NN - Neural Network

OSS - Overlapping with Same Size

ODS - Overlapping with Different Size

PS - Pure Split

PSB - Pure Split with Borders

PRISMA - Preferred Reporting Items for Systematic Reviews and Meta-Analysis

RBNDVI - Red and Blue Normalized Difference Vegetation

R-CNN - Region-based Convolutional Neural Network

RF - Random Forest

SGD - Stochastic Gradient Descent

SVM - Support Vector Machine

SVR - Support Vector Regression

UML - Unified Modeling Language

VI - Vegetation Indices

CHAPTER 1

Introduction

In this chapter, a motivation associated to the theme of this dissertation is carried out in order to introduce the topic. The next section provides an overview of what this thesis is about, as well as an explanation of the scope of this work inside the infrastructure where it was inserted. Following this, the objectives of this dissertation are presented. Next, the methodology adopted in this thesis is addressed and explained. Finally, a brief explanation of the sequence and meaning of the remaining chapters of this dissertation is depicted.

1.1. Motivation

Due to extremely high infant mortality, the planet's human population slowly increased until the year 1700. The first billion was reached around 1800, followed by the second billion in 1928, the third billion in 1960 and its seventh billion in 2017. The rapid population growth in recent decades is mainly due to better healthcare. The United Nations forecasts that the world population will reach 9.7 billion people in 2050, and 10.9 billion people in 2100 [1].

Rapid population growth in recent decades has increased demand for agricultural goods, resulting in a significant rise in cultivation area [2]. Crop productivity will need to double by 2050 to fulfill the expanding population demands for food, bio-fuels, and animal production. To achieve this target, main crop yields need increase by 2.4% each year, but currently this growth is only about 1.3% per year [3]. However, this condition harms the environment by reducing biodiversity and increasing greenhouse gas emissions. It is thus important to make efficient use of resources, such as water and soil, to enable high yield crops, as traditional agricultural production is not sustainable from an economic or environmental point of view [2], as there is less caution, due to less knowledge, in the choice and application of chemicals.

Furthermore, crop production is constantly endangered by insect pests. It is estimated that, globally, the food supplement decreases annually with an average of 40% due to plant diseases and insect attacks [4]. Plant diseases and invasive insects cost the world economy roughly $220 billion and $70 billion, respectively, each year [5].

The rise in global temperature, induced by climate change, has an impact on pest damage and development. Insects metabolic rate increases when the temperature rises, forcing them to eat more food and cause more harm. Temperature also affects the pace of population expansion in numerous insect species. Global crop losses due to insect pests are predicted to grow between 10% and 25% with every degree of average global warming of the earth's surface [6].

Tomato is a fruit vegetable that has great potential to be cultivated since it is a source of vitamins and minerals. In terms of improving yields and fruit quality, tomatoes rank among the horticultural commodities with high economic value that still require careful handling [7]. It is critical to preserve this type of plantations against Diseases and Pests (DaP), in order to improve the quality and quantity of the crop [8]. According to data from the Food and Agriculture Organization of the United Nations, tomato production in Western Europe has increased considerably from at least 2000 to 2019 [9].

(a) *Tomato leaf affected by the mosaic virus disease.*

(b) *Tomato leaf affected by the late blight disease.*

Figure 1. *Examples of tomato leaves affected by diseases from PlantVillage data-set [10].*

Numerous fungal, bacterial, and viral diseases have severely afflicted this plant, with symptoms appearing in various areas of the plant, such as the leaf, stem, fruit, etc. Wilt, rot, stains on fruit, browning of foliage, and stunted development are some of the symptoms [11]. An example of these symptoms on plant leaves is shown in Figure 1.

The traditional method of detecting and identifying plant diseases involves an observation by experts. This takes time and talent, and it is not practical for monitoring huge farms. Therefore, to overcome the limitations of manual detection, automated methods for detecting and forecasting pests and illnesses are required [12]. A system with these capabilities is addressed in this dissertation. A structure capable of performing these tasks will prevent huge losses and the excessive use of pesticides and chemicals, reducing its associated costs as well as the damage done to the environment [12].

The increasing availability of big data analysis approaches, provide the capability to boost the research and development towards smarter farming, contributing to overcome the challenge of producing high yield crops in a larger scale and in a more sustainable way. The development towards smarter farming may support farmers in different tasks such as plant DaP detection and forecasting, or water and soil management, while safeguarding natural resources and protecting physical ecosystems [13].

Artificial Intelligence (AI) and Machine Learning (ML) approaches have been successfully utilised in a variety of areas, including the medical sector for illness detection from medical images [14], image classification in big data-sets [15], self-driving automobiles [16], and academic research fields such as physics [17].

ML in agriculture is still in its early stages, but it is showing promise. Disease classification can be done by using popular Convolutional Neural Networks (CNNs) architectures for different plants with different diseases [18]. Relationships between weather data and pest occurrence can be retrieved using Long Short Term Memory (LSTM) for forecasting future pest attacks [19]. Insect detection on leaves can be performed by using object segmentation and deep learning techniques [20].

1.2. Overview

With the recent technological advances that have been applied to agriculture using ML, this thesis seeks to contribute to the continuous development in this direction. Therefore this dissertation was associated to a project that aims to help monitoring tomato plantations, more specifically in forecasting the outbreak of DaP in the plantations, by developing a system to this effect. This project is called “Ai for New Devices ANd Technologies at the Edge” (ANDANTE) and it is funded and supported by the ECSEL Joint Undertaking (JU) and European Union’s Horizon 2020 (H2020) [21]. The development of the ANDANTE system and, consequently, parts of this dissertation, were performed at Iscte and at INOV-INESC Innovation (INOV). The ANDANTE consortium represents 30 partners from seven European countries, where four are from Portugal (INOV, CCTI, TerraPro - Technologies and Italagro). ANDANTE considers various use cases, one of which is “Use case 2.2: Tomato pests and diseases forecast” [22], that is the use case assigned to the Portuguese partners and explored on this dissertation.

ANDANTE proposes a system capable of acquiring data automatically, using Internet of Things (IoT) devices, making detections and classifications from that data and forecast the appearance of DaP in tomato crops based on all this data and metrics acquired from the detections and classifications. The data is captured by ANDANTE partners and is based on Unmanned Aerial Vehicle (UAV) and ground level images of tomato crops and images of insect traps in those plantations, as well as meteorological information.

In the context of this dissertation and ANDANTE, the results of the work associated with the detection and classification of objects in images are associated with the acquisition of metrics. However, the development to obtain the metrics is distinct and two aspects were considered. One associated with the object detection phase, i.e. object annotation, and other associated with the phase where object detection is not applied or has already been applied, i.e. metrics acquisition. In the context of this dissertation the first aspect is related to the automatic detection of insects in insect traps. The second is associated with the classification of tomato leaf disease and the acquisition of vegetation indices (VI) from the plantation. The first phase will therefore be referred to in this dissertation as object detection and the second as metrics acquisition.

In the context of ANDANTE the object detection and metrics acquisition of the images should be done in two ways: one, in a first phase, associated to devices such as a computer (that allows the use of any type of AI models, designated in this dissertation as “Normal Models”); other, at a later stage, associated to IoT devices (which causes a limitation of the AI models to be used, designated in this dissertation as “Tiny Models”) that need to consume as little energy as possible because will be used in the field to collect images, acquire metrics from them and send that information to the database.

This dissertation focused on the development of mechanisms associated to devices such as a computer, for the analysis, metrics acquisition and object detection of the data. The full system for forecasting the appearance of DaP using all possible data modalities is out of this dissertation’s scope. This is due to the fact that while writing this dissertation, data-sets from the partners with an extensive and varied record were not yet available with sufficient maturity, leading to the impossibility of forecasting. This dissertation ends up having a role in the development of the mechanisms that will allow the creation of these data-sets necessary for the forecast.

Additional contributions to the system as a whole were also performed, namely the development of a web information system, which arose out of a need, for easier access, management and storage of data collected in the tomato crop fields. In short, this dissertation is the first step towards the development of the system proposed by ANDANTE, by having developments concerning a web information system and mechanisms that allow the acquisition of metrics and object detection from the images. This will enable the forecast of the appearance of DaP at a later project´s stage. Figure 2 depicts the dissertation’s scope in the context of ANDANTE in Portugal.

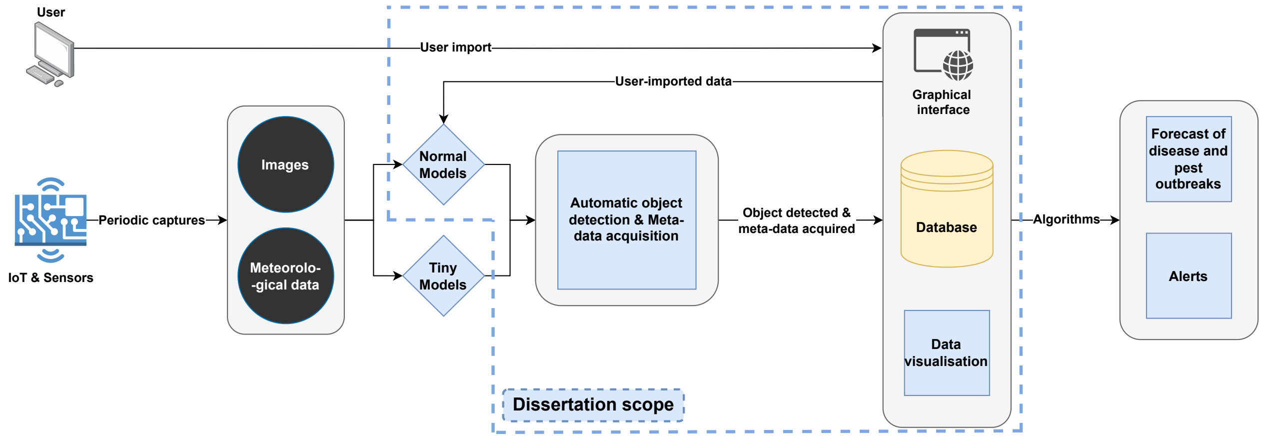


Figure 2*. Dissertation scope inside of the Portuguese context of ANDANTE in the "Use case 2.2: Tomato pests and diseases forecast".*

1.3. Objectives

The objectives of this dissertation emerged from the necessity to develop mechanisms to improve crop production, in particular tomatoes. This is important in order to combat DaP in time and as effectively as possible. In summary, this dissertation aimed to develop modules that will contribute to the development of an AI system for tomato plantations, denominated in the literature review as precision agriculture [23].

The main goal of this dissertation was to develop four modules using data provided by ANDANTE and image processing and ML techniques, namely:

- **Web information system** : The acquired data must be stored and organised, therefore an information system was developed for enabling an easier submission, storage, annotation and analysis of the project data;
- **Insect trap images** : Development of a model for the detection of insects present in insect trap images;
- **UAV images** : Literature review in order to gain knowledge of which VI could and would make sense to be tested in this context;
- **Ground level images** : Testing of different models to verify which is the best for classification of diseases present on tomato leaves.

Taking into account the association to ANDANTE, which is a project of more than 3 years, some data and information was not available at the time of the development of this dissertation. Therefore, there was also an aim to carry out a literature review in order to condense the knowledge that exists on the subject of applying AI to combat DaP in agriculture and enhance the future work of the project so that the forecasting of the appearance of DaP in tomato plantations could be done in the most effective and optimised way possible.

1.4. Methodology

Since the goal of this dissertation is the development of modules that enhance the development of a system, a methodology known as "Design Science Research Methodology" (DSRM) [24] was used. This choice was due to the nature of this dissertation and the characteristics of this methodology - DSRM is characterised by the direction and conceptual model for presenting the results of digital scientific research artifacts.

This dissertation started from the use of the "problem-centred approach", since the problem that this dissertation tries to solve has already been previously defined, as is demonstrated in section 1.2 and through the literature review in section 2.2. The motivation associated with the problem is present in section 1.1.

The second phase, related to the definition of objectives for the solution, was presented in section 1.3, where the objectives of this dissertation were depicted. The next step, which is the design and development of the artefact, represents the development of the solution to the problem mentioned and is present in sections 3 and 4. In this case, this phase involves the development of the necessary mechanisms to perform the object detection and metrics acquisition in tomato crops. In addition, the development of the web information system is also associated to this phase.

The fourth step involves putting into practice the mechanisms developed taking into account the real problem presented. In the case of the mechanisms associated with the images of insect traps and tomato plants at ground level this involves testing the mechanisms developed through data-sets called test sets, data never used in the developed mechanisms. In the other cases there were delays in the annotation of data and in the use of developed mechanisms which did not allow certain tests to be carried out.

The evaluation phase was performed through feedbacks from ANDANTE partners. Object detection and metrics acquisition from images were also evaluated by comparing their results with related work that performs similar tasks. At the time of the development of this dissertation all the data and information from ANDANTE was not yet available, such as annotations. Therefore, some evaluations, mainly those of the web information system and of the developments performed through the UAVs images, could not be carried out as planned. However, evaluations will be conducted in collaboration with the ANDANTE project next year, utilising the new tomato season.

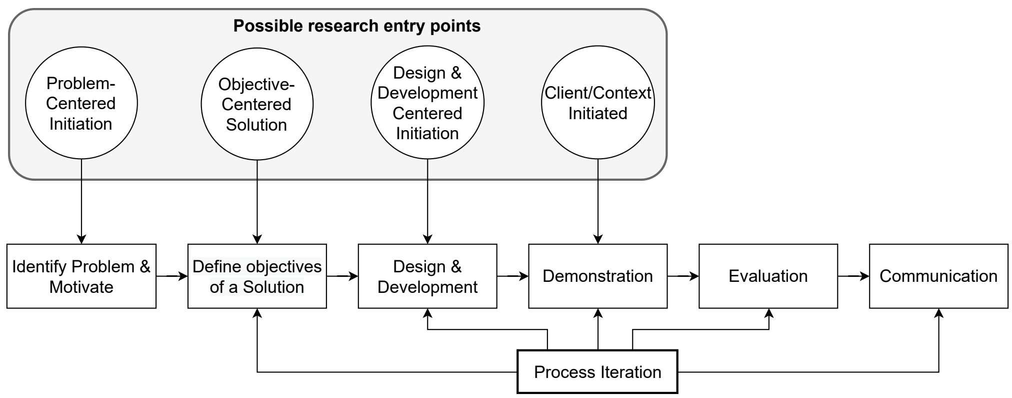


Figure 3. *Design science research methodology.*

Figure 3 summarises the iterative process followed by the DSRM, whose steps were described along this section.

1.5. Outline of the dissertation

After the introduction, the remaining part of the dissertation is organised as follows:

- **Chapter 2** : Presents the research methodology and literature review done;
- **Chapter 3** : Provides an overview of the web information system developed and the data-sets used;
- **Chapter 4** : Presents, evaluates and discusses the results obtained in image processing and analysis modules, including those of ML, developed in the context of the dissertation;
- **Chapter 5** : Presents the main conclusions of this dissertation, as well as its limitations and suggestions for future work.

CHAPTER 2

State of the art

In this section, a review of ML and image processing techniques applied to agriculture is presented, considering the objectives and context of this dissertation. The related work depicted along this section is focused on techniques that allow the object detection and metrics acquisition from images and the forecast of events associated with the emergence of DaP in crops. The methodology used for performing such literature review is also presented. At the end, a brief conclusion is shown in order to identify the literature gaps that can be addressed and explored by this dissertation.

2.1. Research methodology based on PRISMA

In order to find the relevant work related to the dissertation Scopus was used as the primary source of research and Google Scholar as a secondary one.

The type of search conducted in the two databases was different. While in Scopus a search was conducted through a query, in Google Scholar the search was conducted, when needed, through keywords that lead to works that would complement those found with Scopus. The search for documents in Scopus was done using three different queries, each aiming to addressed one of the following subjects: detection, metrics acquisition and agriculture events forecasting.

The query used for the detection subject was the following:

(Insect OR Insects OR Crop OR Crops OR Plantation OR Plantations OR Tomato OR Tomatoes) AND ("Automatic Annotation" OR "Automatic data annotation" OR "Automatic detection" OR "Bounding Box") AND ("Deep Learning" OR "Machine Learning" OR "Neural Network" OR "Artificial Intelligence" OR "Convolutional Neural Networks")

As for the metrics acquisition, the query used was the following:

(Tomato OR Tomatoes OR Insect OR Insects) AND ("Feature extraction" OR Identification) AND ("Deep Learning" OR "Machine Learning" OR "Neural Network" OR "Artificial Intelligence" OR "Convolutional Neural Networks")

And finally, for agriculture event forecasting, the query used was the following:

("Tomato disease" OR "Tomato diseases" OR "Crop disease" OR "Crop diseases" OR "Plantation disease" OR "Plantation diseases" OR "Tomato pest" OR "Tomato pests" OR "Crop pest" OR "Plant disease" OR "Plant diseases" OR "Plant pest" OR "Plant pests" OR "Crop pests" OR "Plantation pest" OR "Plantation pests") AND (Prediction OR Forecasting) AND ("Deep Learning" OR "Machine Learning" OR "Neural Network" OR "Artificial Intelligence" OR "Convolutional Neural Networks")

Besides this query-oriented search, a set of filters were also defined in order to highlight the most relevant works among the enormous amount of results retrieved. These filters are summarised in Table 1.

Including	**Exclusive**
Articles	Not being Articles
Written in English or Portuguese	Written in other than Portuguese or English
From 2015 or later	Older than 2015
In the area of Computer Science	Not in the area of Computer Science
In the area of Engineering	Not in the area of Engineering
In the area of Decision Science	Not in the area of Decision Science
Free or inside ISCTE's scientific license	Paid works

Table 1*. Exclusion and inclusion filters in the search for related work.*

After filtering, an analysis and selection procedure similar to the one defined in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) flow methodology [25] was adopted. Additional references discovered from the analysed works were added to the final selection when their contents were considered relevant.

At the beginning there were 1135 related works and then, after the removal of the duplicated ones and the application of the first phase of inclusion and exclusion criteria of Table 1, 135 related works were left. The next step was to analyse the title and abstract of the papers found. At the end of this step there were 120 papers left. As there were still too many documents out of the scope, a skimming read was performed on each work, and in the end, 124 papers were selected. This increase in the number of papers, compared to the previous step, was due to finding new relevant work referenced by the articles under analysis. The application of the described procedures resulted on a final selection consisting of 84 works. This process is summarised in the Table 2.

Filtering process phase	**Number of works**
Initial search with the inclusion and exclusion criteria	135
Analysis of the title and abstract of the documents	120
Skimming of documents and new relevant work found in references	124
Documents analysed and used	84

Table 2. *Filtering process of related works.*

The selected documents were grouped, according to the themes addressed in them, into three main groups: object detection, metrics acquisition and agriculture events forecasting. This grouping is what led to the organisation of the following sections as they are.

2.2. Related work

2.2.1. Object detection

To perform a wide variety of ML techniques associated with images is very important that the objects are properly detected. In this way, the metrics acquisition from the detected objects and the consequent forecast of events, based on the extracted information, will have a better performance [20]. For example, if the bounding box of a diseased leaf of a tomato plant is incorrect, when trying to classify the disease, i.e. to acquire metrics, the classification may be incorrect due to different possibilities, such as: an important area for the classification is cut off; the bounding box leaves large margins between the objects and the limits of the bounding box, which leads to the existence of more distractions for the model at the time of classification. For proper detection to happen there needs to be some manual annotation in order to properly train the AI models.

The work carried out in [26] aimed to detect, identify and count a specific pest species in insect traps using deep learning. Here a colour correction utilising a variation [27] of the "grey-world" approach [28] was used to mitigate the impact of lighting variability on detection performance. They suggest a sliding window-based detection pipeline that applies a CNN to image patches at various locations to calculate the probability that they contain a certain pest kind. Their work was inspired by algorithms proposed for pedestrian detection, analysed in [29]. The final detections were produced via Non-Maximum Suppression (NMS) [30] and thresholding of image patches based on their positions and related confidences. To evaluate the precision of the bounding boxes the Intersection-over-Minimum (IoM) was computed. It was concluded that many of the errors occur because the same moth could have various wing positions, occlusion levels, lighting circumstances, and decay patterns throughout time, indicating that the algorithm would improve in well-managed sites.

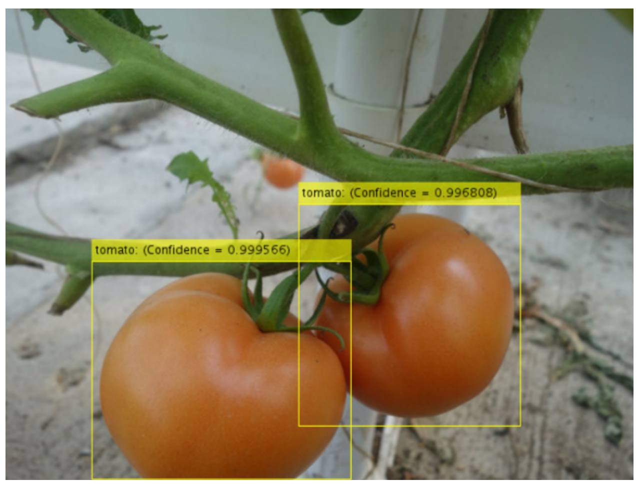


Figure 4. *Recognition of tomatoes [31].*

The work proposed in [31] uses Faster Region-based Convolutional Neural Network (R-CNN) [32] to recognise and locate tomatoes in images. This recognition is related to the creation of a bounding box in each tomato detected, Figure 4 shows an example for the outcomes of this recognition. In addition, the authors performed a more accurate detection by taking the resulting images from the bounding boxes and using gaussian density function of H and S, in the HSV color space, followed by erosion and dilatation on the tomato body to separate nearby tomatoes and eliminate periphery subpixels from all detected ripe tomatoes. They concluded that the Faster R-CNN classifier can quickly and accurately locate tomatoes.

In paper [33] the authors perform an automatic insect detection where they first use a spectral residual model and then extracted different colour features. In the end, whitefly and thrips were identified using a Support Vector Machine (SVM) classifier. The precision for the whitefly and thrips was 93.9% and 89.8%, respectively. As for the detection of the trap, a precision of 93.3% was obtained.

In [34] a pheromone trapping device was developed. In this work the original image was cropped into several sub-images without overlaps. These sub-images were then used to train the tested models. At the end, the image is reconstructed taking into account the detections made in each sub-image. The observed results showed a mean Average Precision (mAP) of 94.7%.

Using IoT and deep learning frameworks, the work in [35] provides a real-time remote insect traps monitoring system and insect identification algorithm. The authors used the Faster R-CNN ResNet 50 and an average accuracy using different databases of 94% was obtained.

To detect the location of a tomato leaf, in [36] the K-means algorithm was used to cluster the images of diseased tomato leaves and to improve the anchors based on the results, with the anchors representing initial guesses of the bounding boxes [37]. This paper indicates that the applied method has better detections than the Faster R-CNN alone. However, it must be taken into account that the data-set in question is lab-based and that only one leaf at a time subject to detect.

A geometric-based detection approach is proposed as part of the work in [38]. Here the goal was to obtain the cutting points of the peduncle based on the fruit bounding box in order to have a autonomous system that harvests most types of crops with peduncles. The Mask R-CNN [39] model was adapted and a geometric feature to detect the fruits and the peduncle cutting points was used. The results indicated that the cutting point can be recognised and the fruit cut at the proper peduncle location. Similarly, the work in [40] also used the Mask R-CNN to individually segment blueberries from an input image. In their work, the authors also suggest that the use of newer deep learning based segmentation models such as “YOLACT” [41], “SOLO” [42], “PolarMask” [43], “Blend-Mask” [44] and “SOLOv2” [45], may achieve better performance than the one used in their research.

In the case of [46], Faster R-CNN was used for the detection of rice seedling on images. In this case the authors tested, separately, three CNN models (“ZF Net” [47] “VGG_CNN_M_1024 Net” and “VGG_16 Net” [48]) in the Faster R-CNN. The best performance was achieved when using the “VGG_CNN_M_1024 Net” and “Approximate Joint Training” method, that is a method where the whole model is trained by setting a maximum number of iterations of the model [49]. Another work that used Faster R-CNN combined with other deep residual networks (“Resnet 50”, “Resnet 101”, and “Inception-Resnet-v2”) was proposed in [50]. Here, the best results were obtained using “Resnet 101”.

In [51], the authors main goal was to build a model that detects white-fly and thrips from sticky trap images in greenhouse conditions. They developed a model based on Faster R-CNN, calling it “TPest-RCNN”, and trained it using transfer learning with a public data-set in a first phase. In a second phase they used their data-set with the weights obtained from the first phase. The model was found to be reliable in detecting microscopic pests on images with varying pest concentrations and light reflections. In addition, it was shown that, for recognising insect species from images captured at sticky yellow traps, the best results were achieved by the proposed model, beating the Faster R-CNN architecture and techniques employing manual feature extraction (color, shape, texture).

Another type of method used to detect objects is the use of “YOLO”, whose concept is simplified in Figure 5 and addressed in [52]. In [53], the authors used “YoloV3” [54] and dealt with the problems of tomato detections under different lighting and occlusion conditions. The authors tested using circular bounding boxes instead of rectangular ones, which resulted in better Intersection-over-Union (IoU) values and a more accurate NMS. Data augmentation was also used.

In the referenced work [53], the “YOLO-Tomato” model was proposed. The proposed model is based on the “YoloV3” and incorporates the dense architecture proposed in

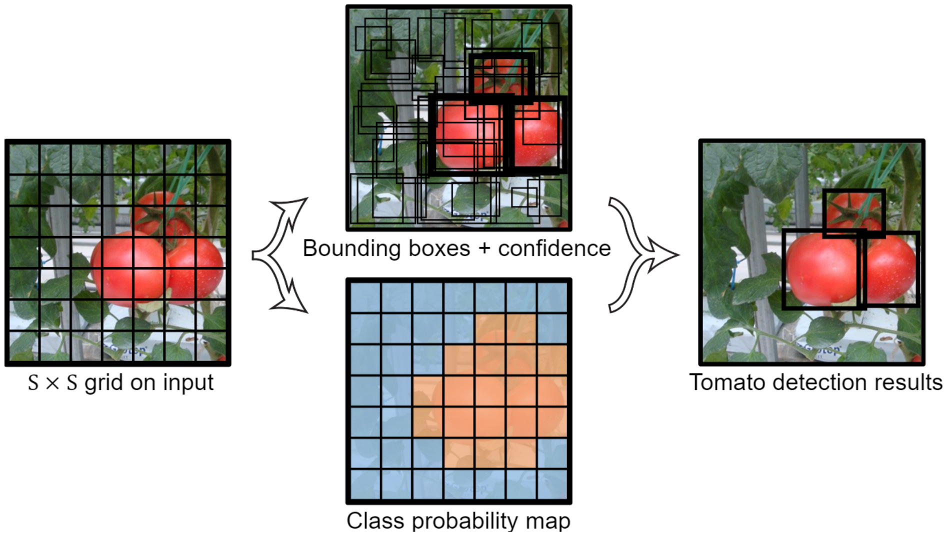


Figure 5. *YOLO model detection [53].*

[55]. To evaluate the detection performance a comparison was done between the use of different methods, data-set sizes and levels of illumination and occlusion in the images. It was concluded that the best method among the ones tested ("YoloV3", "YoloV2", Faster R-CNN and "YOLO-Tomato") was the one proposed by the authors. The results also showed that, for less than 450 images the F1-Score rises quickly as the number grows, but when the the amount of images on the data-set is above 450, performance increase is slower and begins to saturate. Finally, it was also verified that the proposed model is robust to variations in the illumination and occlusion of objects in the images.

One method not yet mentioned was the one proposed in [56]. This method uses the "CenterNet" algorithm [57] to detect vegetables and draw bounding boxes around them. It is proposed to remove the weeds from the background by determining and evaluating a color index using Genetic Algorithms (GAs) [58] according to bayesian classification error. This approach achieved 95.6% precision.

Taking into account the methods discussed, open source solutions can be used to help implement the detection process. In [59] this approach is performed, using the Computer Vision Annotation Tool (CVAT)[1] which contains a feature for automatic annotation/labeling. This software can also be powered by Nuclio[2], a serverless technology

[1]https://github.com/openvinotoolkit/cvat
[2]https://nuclio.io/

that allows to deploy trained models to CVAT. This tool was analysed and it was concluded that it could be interesting to use it given the infrastructure of the project, as CVAT allows to create and carry out annotation tasks and, with Nuclio, deploy trained models [60].

2.2.2. Metrics acquisition

It is possible to acquire a wide variety of information from images using ML algorithms and/or using non-ML computations performed on the image domain. For instance, an ML-based algorithm can be used for classifying the disease present on sick leaf, while a non-ML technique, as a simple python script, can be used for determining the image pixels within a given range of colors. The information can be acquired from the original image, from bounding boxes present on the image, which can be tiles of the original image focusing detected objects, or from the image with a type of processing, such as the transformation of the original image to gray-scale.

Neural Network (NN), LR, Decision Tree (DT), SVM, k-nearest neighbors (knn), Naïve Bayes (NB), and deep CNN are the most common AI methodologies for detecting and classifying plant diseases [61].

A common type of metrics acquisition is the classification of the object present in the image. Various works, performed this acquistion, performing the classification of insects [62]–[65] or diseases in tomato leaves [66]–[69]. To obtain this type of information one of the techniques used is transfer learning. [70] and [71] used this approach in their work. Transfer learning consists of taking the relevant parts of a pre-trained ML model and applying it to a new but similar problem.

Regarding the work done in [70], it is suggested “AlexNet”, “VGG-16 Net” and “SqueezeNet” as three pre-trained deep networks, where transfer learning is applied, in order to analyse and evaluate their performance in the categorisation of tomato leaf diseases. The authors compared the proposed methods with each other and with the state-of-the-art techniques. In addition they used two types of data-sets, a smaller and a larger one. It was concluded that, in general, the proposed method using “AlexNet” was the best among all the tests performed.

In [71], the authors used “AlexNet”, “VGG16”, “GoogLeNet”, “MobileNetv2”, and “SqueezeNet” and applied transfer learning to them to classify tomato plant leaves as healthy or diseased, classifying the type of disease on the illness cases. The results showed that “VGG16” achieved better performance.

This technique has also been applied in the area of insect classification. In [72] the authors used transfer learning to classify kissing bugs (*triatominae*) and their specific species. The results obtained show that “VGG16” achieved better performance among all testing configurations, reaching 96% accuracy for the classification of different species of kissing bugs.

Besides the simple classification of the objects of interest, other types of metrics can be collected. For example, [73] proposed a model to classify tomato maturity based on colours. Transfer learning was used and the obtained results outperformed other deep learning and ML techniques used in recent works on image classification in the context of tomato crops. Image pre-processing was applied by resizing, cutting and removing the background. The background removal was performed in four steps: the first refers to the transformation of the blue channel by setting the green and red channels to zero; the second to the gray-scale transformation; the third refers to the development of a binary mask using Otsu’s method [74]; the fourth concerns with the application of the binary mask to the colour image.

Another type of metrics possible to collect is the nutrient stress in the tomato plant [75][76]. In [75], transfer learning is applied to three pre-trained architectures “Inception-V3”, “ResNet50” and “VGG16” combined with two classifiers, Random Forest (RF) and SVM, to improve classification accuracy. The results showed that “Inception-V3” alone achieved the highest accuracy. Using the classifiers, the best accuracy was obtained combining “VGG16” with SVM. The proposed model is shown in Figure 6.

The work in [77] proposes an effective tomato experimental sorting method based on machine vision. The authors developed an algorithm to analyse the images. The parameters for sorting included shape, size, ripeness and defects, and these were the features that the algorithm collected. The techniques used to obtain these parameters were mostly based on image processing and mathematical methods. Image processing

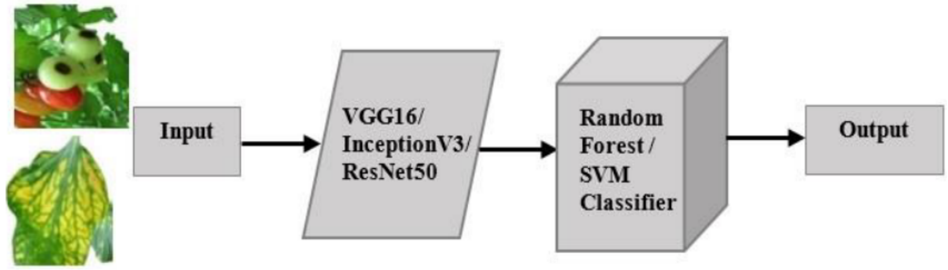


Figure 6. *Model proposed in [75].*

took place by converting the images to the HSI color space, and then, pixels whose colours were outside the specified HSI range were filtered and put in black. The resulting image was converted to RGB space and then its pixel colors were filtered in (0, 25) for red and (0, 64) for the blue and green range. Finally, the image was transformed to gray-scale and thresholded with Otsu's algorithm. The accuracy for defect identification, form, size, and the overall system, with all of the previous ones combined, was 84.4%, 90.9%, 94.5%, and 90%, respectively.

Due to differences in chlorophyll concentration, damaged leaves on diseased plants exhibit a different spectral reflectance than those in healthy plants. Sick plants absorb less visible light and more near-infrared light. Thus, reflectance information can be used for identifying diseased plants [78]–[80]. In [78], were was done a study targeting the late blight infection, a disease that affects different plant species including the tomato, it was discovered that spectral changes between healthy and sick plants are minimal in the visible spectrum, but substantial differences are observed in the Near Infra-Red (NIR).

The Normalized Difference Vegetation Index (NDVI) [81] is a standard indicator for measuring the level of vegetation in a region using leaf reflectance data. Satellite data or customised cameras can be used to calculate NDVI [78][80]. The combination of NVDI with the minimum temperature was shown to be more accurate than meteorological factors alone for forecasting the brown planthopper insect population. Relevance of NVDI for forecast pest alongside with temperature was found in [82], where it is suggested that NVDI may provide information regarding the relation between the crop growth stage and pest development.

Remote sensing may be used to get a variety of vegetation indexes [83]–[89]. Besides NDVI, many other indices have been developed taking into account the characteristics of the problem, developing NDVI variations such as Normalized Difference Red Edge Index (NDRE) [86] or Green Normalized Difference Vegetation Index (GNDVI) [87].

2.2.3. Forecast of diseases and pests

Besides the correct detection and classification of DaP in crops, it is very important to forecast their appearance.

For event forecast, it is important to understand which variables will have an impact on what is forecast, which was a subject analysed in [90] and [91]. In agriculture, a model introduced in [92] can be used to understand the impact of location and temperature on crops. In addition to these, it was verified that variables such as soil, humidity, rainfall and moisture can have an impact on the crop yield [93]. Long term analysis of meteorological data collected by unmanned observation can also be used for forecasting disease incidence [94].

A cloud-based platform that can handle the collection, analysis, and forecast of information about the agricultural environment, a smart farm service, was proposed in [94]. The suggested integrated system operates and monitors farms and manages related devices, data, and models to provide support for high-level application services. This system analyses environmental and growth data while registering, connecting, and managing IoT devices. Here, General Infection Model [95] was used to develop a model for forecasting *Botrytis cinerea* infection risk. It was concluded that an accurate plant disease forecast systems can be built if such integrated systems are implemented and various input data-sets necessary for system configurations and interrelationship analyses are collected [94]. This demonstrates the importance of having a lot of data when forecasting agriculture events.

A suggested model, proposed in [96], tried to identify a link between agrometeorological variables and the prevalence of four different types of rice illnesses. The authors used an Artificial Neural Network (ANN) to perform detection, classification and forecast of disease occurrence in rice crops. The data-set used was composed of weekly weather data from 1989 to 2019. Initially, the correlation between the variables was evaluated and pre-processing of the data-set was carried out. Then, using the proposed ANN model, the weather forecast was done. With the weather forecast done, disease forecast was performed for those weather forecasts. The authors compared different activation functions in the ANN model and came to the conclusion that Relu did the best job when it came to weather forecasting and Softmax when it came to classification. A Mean Absolute Error (MAE) of 0.46 was obtained for weather forecasting and an accuracy of 92.15% for the classification.

It is also common to use LSTM method to perform event forecasting, such as the work in [19] did. Initially, the authors used the Apriori algorithm to find the association rules between weather variables and the occurrence of cotton pests. Forecasting the presence of DaP was presented as a time series and an LSTM-based algorithm was devised to tackle it. Other traditional classification methods such as SVM, KNN and RT, were also implemented to compare with the use of LSTM. The model developed using LSTM has outperformed the other methods. The results suggest that the LSTM network has specific advantages in processing time-dependent problems and that model selection is critical.

In [97] an ANN was used to carry out the forecast, resulting in a classification according to three classes: no disease occurrence, low severity disease occurrence and high severity disease occurrence. A data-set with meteorological data from 2011 to 2015 was used and different activation functions and different splits of the data-set were tested. The results concluded that if a larger data-set is used, a forecast with a higher accuracy will occur. The maximum accuracy achieved was using the Sigmoid activation function and it was 90.909%.

Other works, such as [98] and [99], also proposed a method that uses ANN to perform a forecast. In [98] it is forecast the crop that can grow in a certain area based on soil and weather related parameters. In [99] a technique to forecast food quality using Back-Propagation Neural Network (BPNN) is developed. In [100] and [101] it is carried out a study where a forecast for crop yield is provided by an ANN. In [100] it is tested different types of NNs and came to the conclusion that increasing the number of neurons in the hidden layer does not contribute to increasing the accuracy of the system. The NN has the structure in the form of one input layer neuron, two neurons in the hidden layer and one neuron in the output layer. The results showed that ANN is better, in terms of accuracy, than the traditional regression methods used in this case. In [101] agricultural data collected over 30 years from a paddy field in Nepal was used. In addition climate and fertilizer use data were also used as input for the forecast. The authors concluded that the trained NN produced a minimal amount of error, indicating that the model is competent to forecast Nepali crop yield.

In the case of the work done in [102], SVM was used to perform disease forecast and it was concluded that SVM is better than conventional multiple regression and ML algorithms, BPNN and Generalized Regression Neural Network (GRNN) in this case. The authors based their forecast on six meteorological variables. As for the work in [103], the BLITE-SVR, which is a potato late blight forecast model, was evolved using Support Vector Regression (SVR) and 13 climatic variables. The procedure for building the potato late blight forecast model is shown in Figure 7. Regarding the work in [104], the authors used a linear model to forecast the crop growth period, but also used non-linear components to increase the accuracy of the system. On the other hand in [105] the non-linear Quasi-Newton multi-variate optimization method was used.

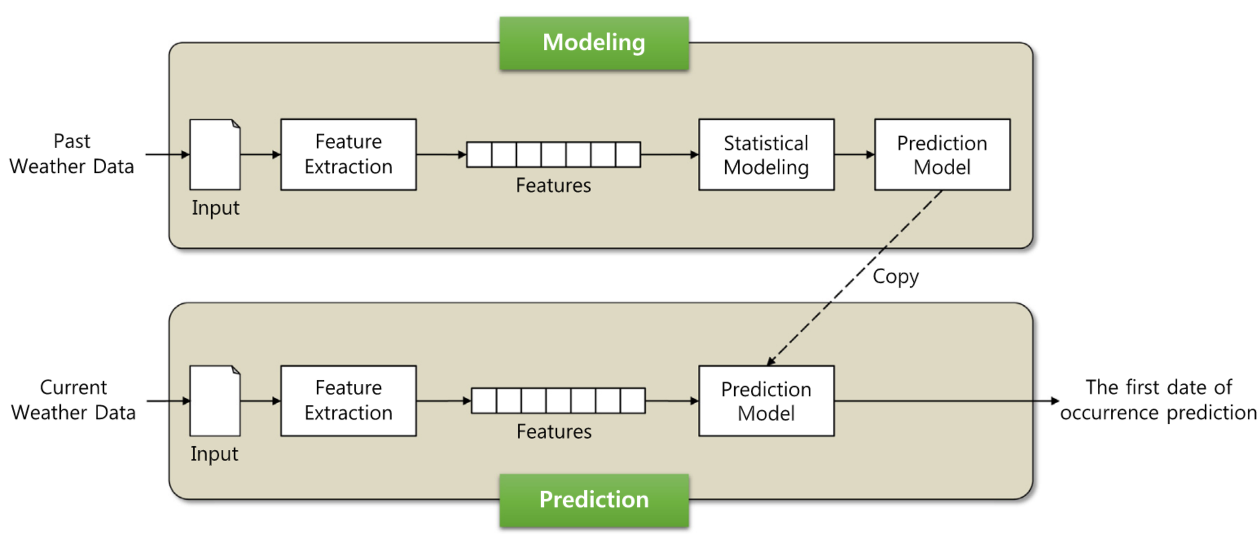


Figure 7*. Procedure for building a potato late blight forecast model by [103].*

Various surveys have been done in the area of forecasting events related to agriculture. For example, [106] has provided a survey article on the methodologies utilised in the field of agricultural wheat crop forecasting. Statistical, meteorological, simulation, agronomic, remote satellite sensed, synthetic and mathematical models were studied in the research. The study shows a compact combination of all these models that demonstrates why the NN model is important to other models for non-linear data behavior systems such as wheat crop yield forecast. In [107], the authors analysed and categorised research findings from the last ten years that forecast the start of disease at an early or pre-symptomatic stage, i.e. symptoms that are not visible to the unaided eye. They investigated the approaches and methods used, pre-processing techniques and data used, performance metrics, and expected results, highlighting the issues encountered. The study's findings show that this practice is still in its infancy and that many obstacles must be overcome. In the case of [108], the authors performed a review on studies that used ML methods applied to plant resistance genes discovery and plant diseases classification. The research showed that techniques such as SVM, NB and Markov clustering have been used to forecast plant diseases. It can be concluded from the survey that ML techniques have a great potential in disease forecast. The report in [109] outlines all of the models that have been developed for late blight of potato forecasting across the world. With these surveys, which addresses and summarise various works, it is possible to reach a wide variety of works related to the forecast of events in the context of agriculture.

2.2.4. Dissertation contribution

Taking into account that there is no work performing the forecast of the emergence of DaP in tomato plantations using data from meteorological stations, as well as aerial, ground level and insect traps images of tomato plantations, this dissertation seeks to explore this gap by providing foundations that will enable the development of this forecast.

CHAPTER 3

Data organisation and collection towards image analysis modules

At the first stage of the development of this dissertation there was a whole process of organising and selecting the data-sets. This process emerged from the quantity and condition of the data and from a need demonstrated by the end users and partners associated with the project, in order to optimise the project's progress. Taking this into account, an information system, that includes a web application to facilitate farmers inputs, was developed to help this process. This development aimed to facilitate the remaining work associated to this dissertation as well as future work associated with the ANDANTE project. In the following sections, details on the data-sets used and the web information system developed are provided. The open source software CVAT, its Application Programming Interface (API), and Nuclio (open source and managed serverless) were used in the developments described below and in Chapter 4, making model training, manual and automatic detection, and data management and selection easier.

3.1. Web information system

The development of a web information system that facilitates access to data, both for visualisation and analysis, as well as support for training, testing and use of ML-based mechanisms to predict the emergence of DaP in tomato crops, meets a need demonstrated by the ANDANTE partners.

The idea of developing a web information system such as this emerged since the data acquired was huge and was not organised and available for easy access and analysis. This was considered fundamental in order to develop more efficient and optimised techniques of object detection and metrics acquisition from the data provided. In addition, this system was designed to facilitate the sharing of information (annotations, new data, etc.) by end users.

This idea in combination with requirements surveys performed before and during the development of the other modules, with several meetings with the Portuguese partners of the ANDANTE project, led to the development of the web information system useful to those who were developing the project and to the end users. The meeting interactions demonstrated the need for an application where it would be possible to access information and conduct studies of metrics related to crop fields, such as the number of insects in different traps over time. In addition, a retrospection of the work previously done on the project was performed with INOV and it was concluded that it was necessary to reorganise and structure the data obtained so far in order to enable faster project development. This need was in line with the needs demonstrated by the partners.

The interface was developed on the basis of the results of the requirements survey, and as result, the following functionalities were developed:

- Creation/edition of new/old insect trap stations;
- Visualisation of the insect trap stations and their characteristics;
- Manual import of insect trap images;
- Display of collected/imported insect trap images and their characteristics;
- Analysis (line-chart) of the number of insects over time, with the possibility of filtering by date and trap station and/or cultivation field;
- Creation of tasks in CVAT of images that are not annotated and are in the system already;
- Synchronisation of the latest detection in CVAT with the respective images present in the system;
- Creation/edition of new/old weather stations;
- Visualisation of the weather stations and their characteristics;
- Manual import of weather data;
- Display of collected/imported weather data and their characteristics;
- Analysis (line chart) of the weather data over time, with the possibility of filtering by date, weather station, cultivation field and/or measurement;
- Manual import of ground level images;

- Display of collected/imported ground level images data and their characteristics;
- Analysis (line chart) of the plant images import rate over time, with the possibility of filtering by date and cultivation field;
- Manual import of UAV images;
- Display of collected/imported UAV images data and their characteristics;
- Creation/edition of new/old cultivation fields;
- Visualisation of the cultivation fields and their characteristics;
- Creation/edition of new/old diseases/pests;
- Form for the registration of diseases recorded in the crop field.

Some of the mechanisms developed within the scope of this dissertation, further detailed in Chapter 4, can be integrated into the website in order to put them to effective use, such as:

- Objects detection in imported insect trap images;
- Acquisition of Exchangeable Image File Format (EXIF) information from the different types of manually imported images;
- Automatic computation of VI when UAV images are imported.

This website was developed using Django [110] associated with SQLite [111] and Pyforms [112]. Django was used because it is a widely used technology today and facilitates communication with the database. The data models developed for the database, as well as their relations, are represented in Figure 8.

A model ("DiseaseEvent" table) representing DaP appearing in the crop field ("Crop" table) was created so events of the appearance of DaPs in areas of the crop field could be registered and DaP could be associated to UAV images ("DroneImage" table), insect traps images ("InsectsTrapImage" table) and plants images at ground level ("PlantsImage" table). The model corresponding to the acquired UAV images ("DroneImage" table) also registers the crop field ("Crop" table), the acquisition date, the location and the different spectra of the respective UAV image. The model for the ground level plant images ("PlantsImage" table) also registers information about the date, crop field ("Crop"

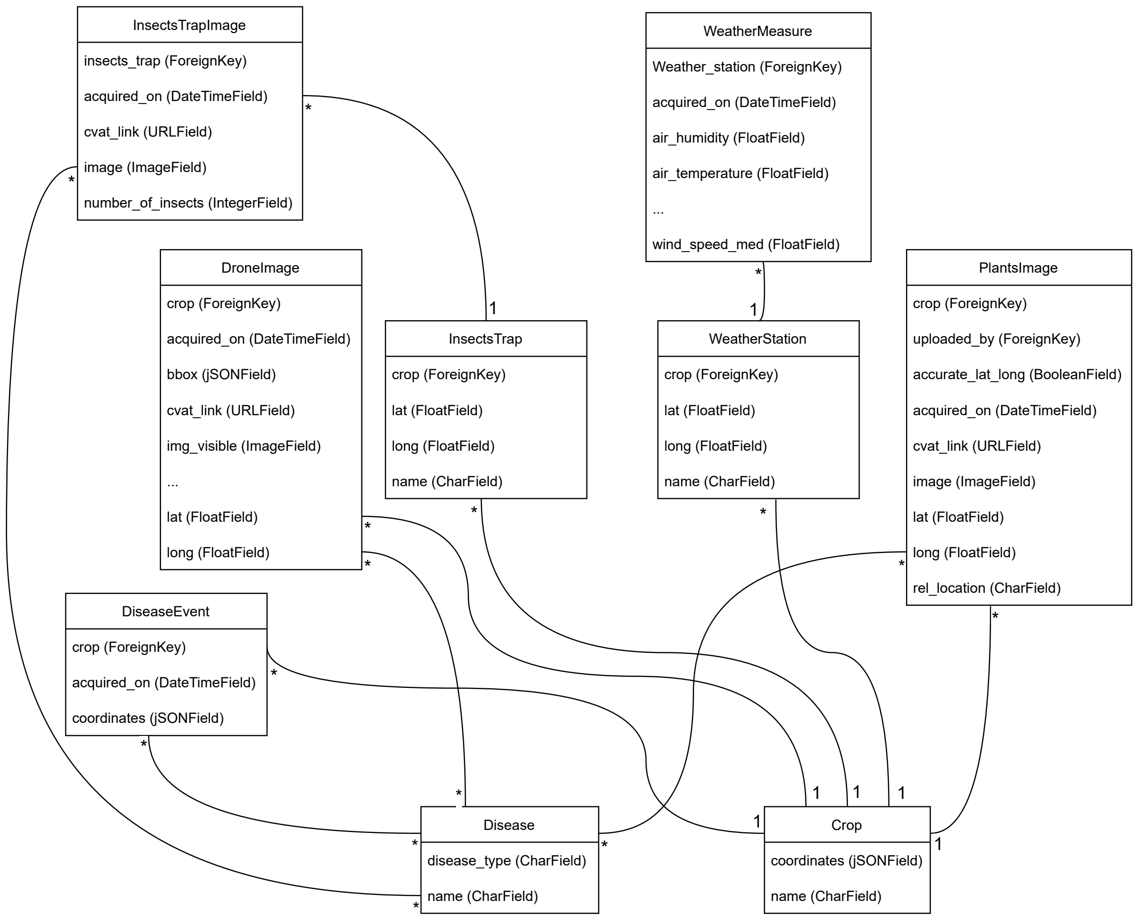


Figure 8. *Django models Unified Modeling Language (UML).*

table), location and acquisition date of the respective image. Regarding the models concerning the insect traps images ("InsectTrapImage" table), which records information about the number of insects and date of the respective image, and the meteorological data ("WeatherMeasure" table), which registers information about the measurements captured by the sensors, a model was created concerning the respective stations ("InsectsTrap" and "WeatherStation" table) where the data is acquired. The insect trap and weather station models ("InsectsTrap" and "WeatherStation" table) register information about the crop field ("Crop" table) and location of the respective station. These models emerged in this way taking into account the existing variables and the communications done with partners.

To help understand the developed web application and its functionalities, its interface is shown in Figure 9. In this case a random example (unrealistic) of the graph visualisation of the number of insects in the *trap 001* of the crop *Test_Crop_1* over time can be observed. Additional figures and information can be found on Appendix A.

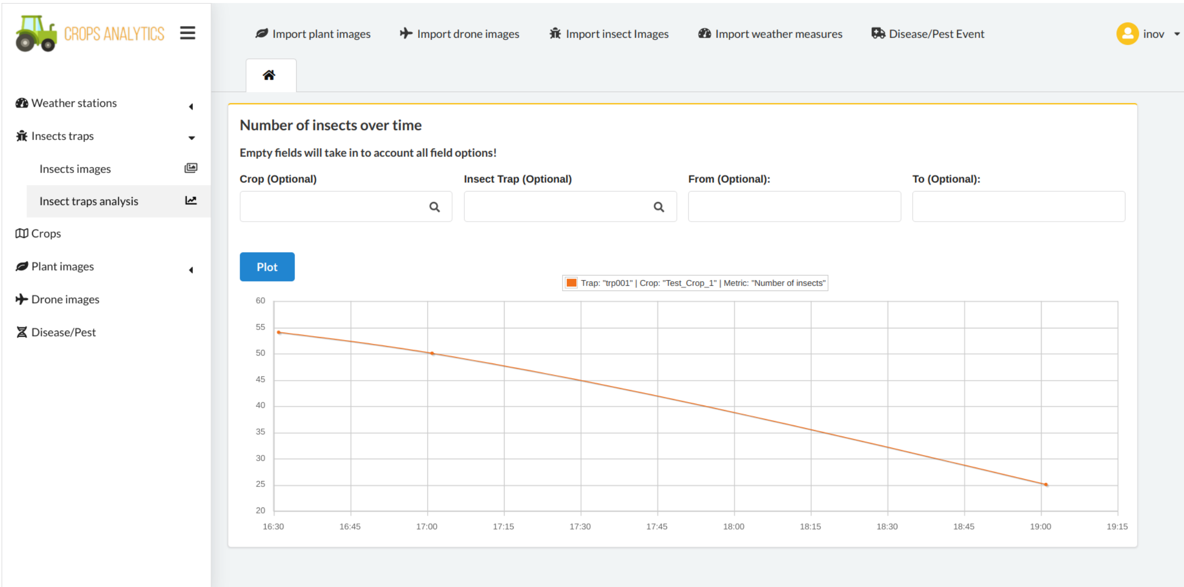


Figure 9. *Web application interface for the insect trap images data analysis.*

The web-based information system has been tested by importing new data concerning the tomato plantation in the Ribatejo area in Portugal in 2022 and data already acquired from the 2021 season. The tests carried out mainly showed that the system was prepared for the storage, access and analysis of imported data. For example, with the system developed it became possible to clearly access and link an weather station and its measurements or an insect trap and its images to a crop field. Taking into account the tests performed it can be stated that there has been a significant improvement, with needs that this system was intended to address being met.

3.2. Data-sets

The data-sets used are associated with Portuguese tomato plantations in the Ribatejo region, more precisely in Valada, Castanheira and Lezíria, where ANDANTE Portuguese partners performed the data acquisition. Information about these tomato crop fields can be found in Table 3. The acquired data was organized into four data-sets:

- **Insect Traps data-set :** Images of insect traps present in the plantations;
- **UAV images data-set :** Aerial images of the crop captured by a UAVs.

- **Ground level images data-set** : Images captured at ground level of the plantations;
- **Meteorological data-set** : Meteorological data provided from meteorological stations in the crops.

Location	Area (ha)	Planting date	Central GPS point
Castanheira	23	19/04/2021	38.982300, -8.954110
Lezíria	27	27/04/2021 and 10/05/2021	39.006537, -8.881018
Valada	20	07/05/2021	39.067730, -8.772214

Table 3. *Information on the tomato crop fields where data was acquired.*

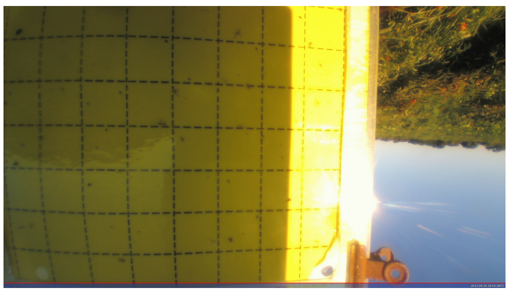
3.2.1. Insect traps data-set

This data-set consists of 5646 images of insect traps acquired by cameras located in front of the traps. However, only 4637 images were valid since some of them did not correspond to insect traps or were of insufficient quality to improve the models performance. These images were manually verified and marked as invalid. Table 4 shows the results of this filtering.

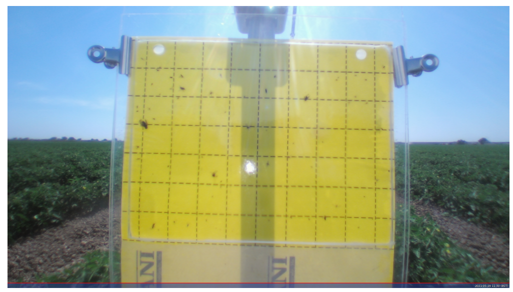
Figure 10 presents an example image for each of the six traps used and Table 4 shows additional information regarding the location and the image acquisition period on each installation. The images were captured every day between the dates indicated in Table 4. The acquisition was mostly done between 11 am and 8 pm at different times of the day (11 am, 11.30 am, 12 midday, 4 pm, 4.30 pm, 5 pm, 7 pm, 7.30 pm and 8 pm), usually nine images were captured per day. The ANDANTE partners defined this configuration based on their understanding of the insect's behaviour.

	Trap 001	Trap 002	Trap 003	Trap 004	Trap 005	Trap 006
Field	Valada	Castanheira	Valada	Lezíria	Lezíria	Castanheira
Period of operation	27/05/2021 to 03/09/2021	26/05/2021 to 08/09/2021	27/05/2021 to 08/09/2021	27/05/2021 to 23/09/2021	27/05/2021 to 24/09/2021	26/05/2021 to 06/09/2021
Total images	848	948	901	945	1071	933
Valid Images	733	756	784	763	845	756

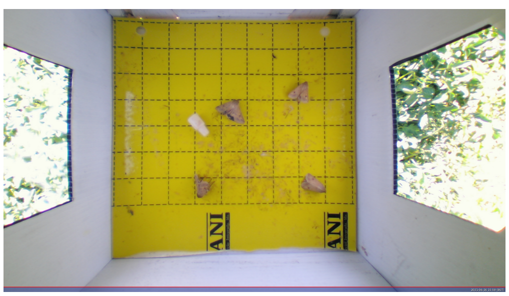
Table 4*. Numbers of the insect traps images.*



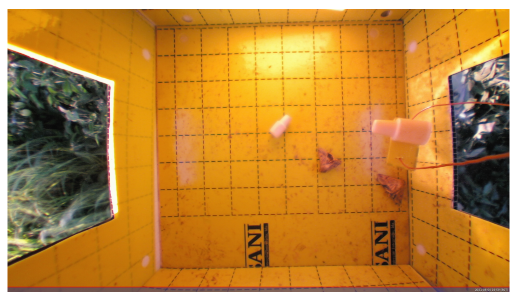
(a) *Insect Trap 001.*



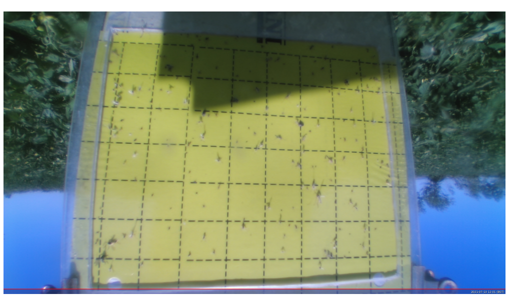
(b) *Insect Trap 002.*



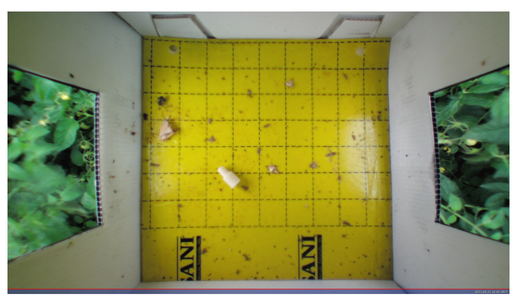
(c) *Insect Trap 003.*



(d) *Insect Trap 004.*



(e) *Insect Trap 005.*



(f) *Insect Trap 006.*

Figure 10*. Examples of insect traps images data-set.*

3.2.2. UAV images data-set

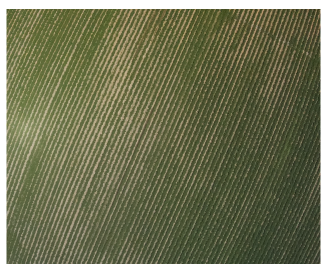
The aerial images collected by UAVs were captured on a weekly basis and at 120 and 30 metres altitude in the crop fields under study. The images have six wavelengths: visible, blue (450 nm ± 16 nm), green (560 nm ± 16 nm), red (650 nm ± 16 nm), red edge (730 nm ± 16 nm) and NIR (840 nm ± 26 nm). More information about this data-set is depicted in Table 5. Figure 11 depicts examples of UAV images of the different fields in the visible wavelength.

***(a)** Castanheira at 120m.*

***(b)** Castanheira at 30m.*

***(c)** Lezíria at 120m.*

***(d)** Lezíria at 30m.*



***(e)** Valada at 120m.*

***(f)** Valada at 30m.*

***Figure 11**. Examples of UAV images data-set.*

Location	First flight date	Last flight date	Number of flights	Number of images
Castanheira	13/05/2021	09/08/2021	14 (120m and 30m)	4055 (120m) , 5529 (30m)
Lezíria	31/05/2021	23/08/2021	11 (120m and 30m)	4548 (120m) , 3752 (30m)
Valada	20/06/2021	23/08/2021	6 (120m and 30m)	1878 (120m) , 2026 (30m)

***Table 5**. UAV images information.*

3.2.3. Ground level images data-set

This data-set consisted of images of plants of the crop fields under study captured manually by the ANDANTE partners. This was done in order to simulate a future implementation of a station equipped with a camera that captures images of plants in a regular and uniform manner. This image capture was done in order to detect and subsequently classify diseased leaves and fruits. Due to the fact that the image acquisition was done manually there was no frequency of capture as in the case of the insect traps images.

This data-set had a problem associated with the non-uniformity of the type of images, considering that there were images capturing multiple plants, others capturing only a leaf and others only a fruit. Examples of this non-uniform capture process can be observed in Figure 12. To increase the performance of object detection and metrics acquisition models, an assertive pre-selection of data, organising it into different categories, was required. This is due to the necessity of different types of detections depending on the image content. For instance, an image depicting a single leaf requires different detections than those for an image depicting a large crop field area.

Due to the scarcity and non-uniformity of the data, it was decided not to use this data-set and to use the PlantVillage [113] data-set to develop mechanisms that can be used and adapted to the real data-set of the project in the future.

From PlantVillage data-set, a balanced sub data-set was obtained considering only images associated to tomato plants. It consists of 11000 images of tomato leaves acquired in a laboratory environment (Figure 1 depicts two examples), i.e. leaves taken from the plant and placed on a background of mostly the same colour (grey), with 1100 images associated to each disease. In the data-set there are healthy tomato leaves and with the following diseases: *mosaic virus*, *target spot*, *bacterial spot*, *yellow leaf curl virus*, *late blight*, *leaf mold*, *early blight*, *spider mites twospottedspider_mite* and *septoria leaf spot*.

3.2.4. Meteorological data-set

The meteorological data collected in the study fields were collected using real and synthetic meteorological stations, with the real ones corresponding to three stations (one

(a) *Planting zone.*

(b) *Planting zone.*

(c) *Mildew disease.*

(d) *Bacterial disease.*

Figure 12*. Examples of data-set Ground level images.*

in each field) and the synthetic ones corresponding to four, whose values were obtained by ANDANTE partners combining the physical station values.

The data collected through the stations are: cloud percentage, atmospheric pressure, soil temperature, evapotranspiration amount, rainfall amount, air temperature, solar radiation density, air humidity, wind speed, wind direction, dew point, wet leaf percentage, plate temperature and the energy consumption of the station. The measurements for the real stations were taken every 30 minutes, whereas the measurements for the synthetic stations were taken every hour. Table 6 contains further information.

This data-set was not used in the scope of the dissertation due to delays in the provision of information by the ANDANTE partners which made it impossible to correlate the data and attempt to predict the onset of DaP using only meteorological data.

Location	Period of operation	Number of data of each measure
Castanheira	26/05/2021 to 07/12/2021	9479
Lezíria	31/05/2021 to 31/12/2021	11062
Valada	17/06/2021 to 31/12/2021	11647
Azambuja (synthetic)	05/03/2021 to 08/11/2021	5975
Benfica do Ribatejo (synthetic)	05/03/2021 to 08/11/2021	5975
Leziria (synthetic)	05/03/2021 to 31/12/2021	7224
Salvaterra (synthetic)	05/03/2021 to 31/12/2021	7224

Table 6. *Meteorological data information.*

CHAPTER 4

Image analysis modules

This chapter covers the work developed for the modules that perform image processing and analysis, either with or without the use of ML techniques.

4.1. Object detection in insect trap images

This section addresses the developed work associated with object detection. In the scope of this dissertation, this detection occurred in the images of insect traps, with the aim of counting the insects present in them in order to correlate this metric with other metrics gathered and contribute in the future development of the algorithm for forecasting the appearance of DaP. The objects to be detected were the traps (yellow sticky cards) and the insects present in them. Due to the fact that no annotation existed, some manual annotations were done at first. This was necessary to enable the training of the object detection models.

Due to the fact that insect traps are physically different between themselves and are subject to different illumination conditions during image acquisition, it was opted to use only AI models for object detection, discarding the use of manual image processing mechanisms for the detection of insects. The fact that the colour of the insects is usually the same as the colour of the lines present in the yellow sticky cards also led to the use of only AI models. Taking this into account and the literature review [31][38][40][46][50]–[53], where it was observed that AI models are being increasingly used, performing better and replacing more traditional methods that involved manual image processing, the manual image processing techniques were discarded.

The pipeline for insect detection was as follows: yellow sticky card present in the original image is detected; the resulting bounding box is split into tiles; the insects present on each tile are detected; the original image is reconstructed with all bounding boxes. For the sake of improving performance and results, the bounding box corresponding to the yellow sticky card, i.e. the result of the yellow sticky card detection model, was split into tiles which were used to train the insect detection models tested. Figure 13 depicts this pipeline divided into two phases, A and B.

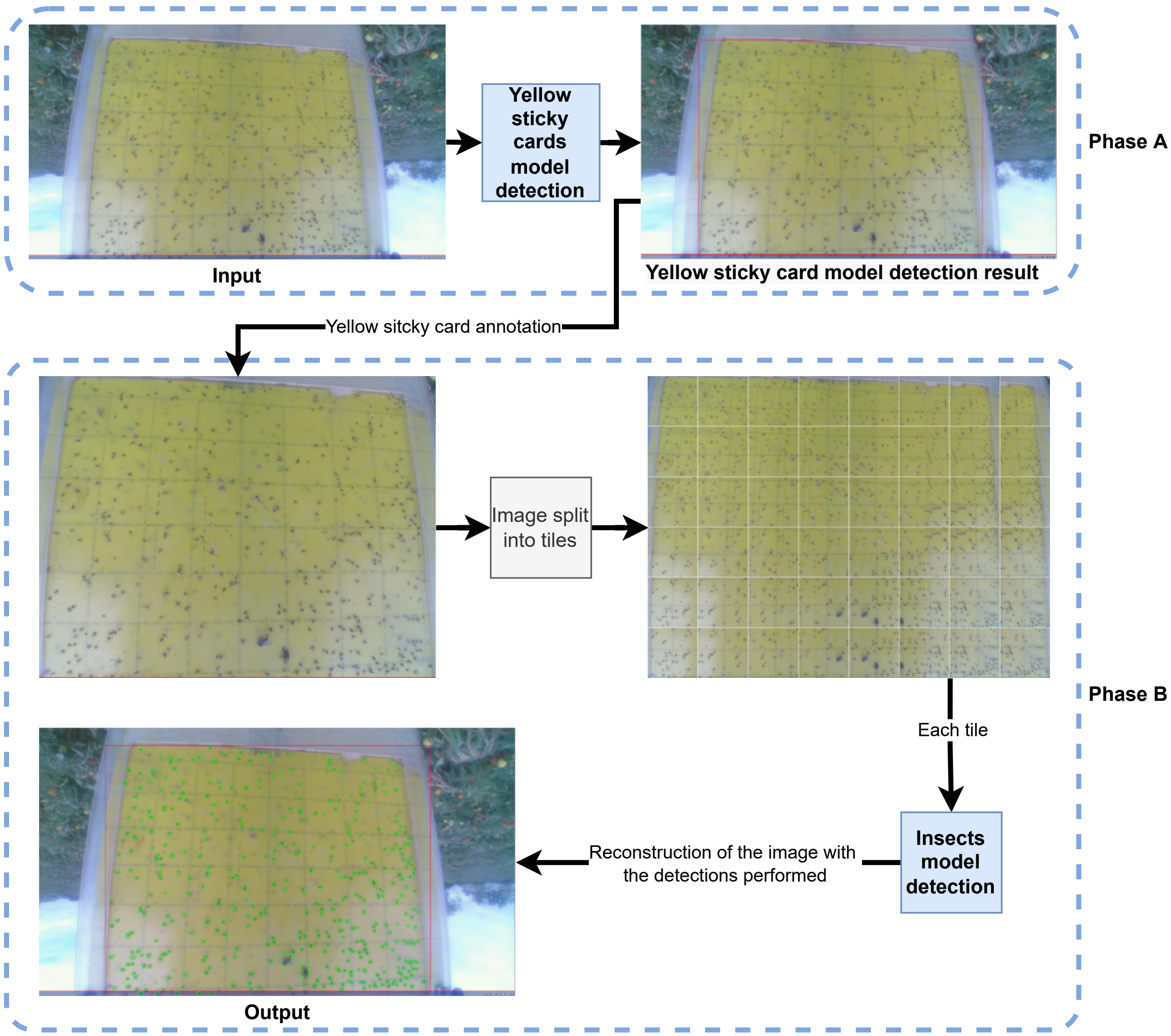


Figure 13. *Pipeline for insect detection.*

Considering the literature review performed [52]–[54], the YoloV5 object detection model was used to perform the necessary detections. Transfer learning was applied for training the model for this specific case of insects and yellow sticky cards detection.

The YoloV5 model as different versions (YoloV5s with small size, YoloV5m with medium size, YoloV5l with large size and YoloV5x with extra large size) and the basic structure of all these versions is the same. They differ only depending on the size of the model because there is a multiplier that influence the width and the length of the network. Generally, the larger the size of the model, the better the performance and the more processing time and memory will be used [114].

The parameters presented in Table 7 were used in all developments involving the use of YoloV5.

Epochs	Batch Size	Optimiser	Patience
300	16	Stochastic Gradient Descent (SGD)	100

Table 7. *YoloV5 Insect trap images parameters.*

4.1.1. Yellow sticky cards model detection

Phase A, concerning yellow sticky card detection, was developed with the intention of using the detection data to later detect the insects contained in the sticky cards.

From the valid images, explained in section 3.2.1, 1272 insect trap images were manually annotated concerning the yellow sticky trap. 80% of the data-set was used for training, 10% for validation and the remaining 10% for testing. The images were resized to 640 by 640 pixels in the training process.

The lightweight YOLO model YoloV5s was enough to achieve near-perfect results, as shown in Table 8. With the developed trap detection model getting good results, all the images that had not been manually annotated were passed through the developed model and it was verified the correct detection by the model.

Phase	mAP_0.5	mAP_0.5-0.95	Precision	Recall
Training	0.995	0.995	1	1
Testing	0.995	0.995	1	1

Table 8. *YoloV5s yellow sticky card model results.*

4.1.2. Insects model detection

Insect detection model was developed considering only the bounding box corresponding to the detection of the yellow sticky card. The YOLO model was again used, but in this case more powerful versions of YoloV5 were tested.

Initially, the tiles were obtained with an increment of the base tile size and in cases where these increment was not divisive of the width and/or length of the image, the tiles in the margins (right and/or bottom) were smaller than the remaining tiles (Figure 15c), this approach was denominated Pure Split (PS). In a second phase, in order to keep all tiles always with the same dimensions, black/yellow/white borders were added to the tiles with smaller dimensions (Figure 15d), this approach was denominated Pure Split with Borders (PSB). However, these approaches were discarded, since with these approaches it was possible for an insect to be split between tiles. This could lead to two detections representing the same object, one corresponding to the part of the object that was in a certain tile and the other to the part of the object that was in a tile in the vicinity of the previous one. This situation is illustrated in Figure 14.

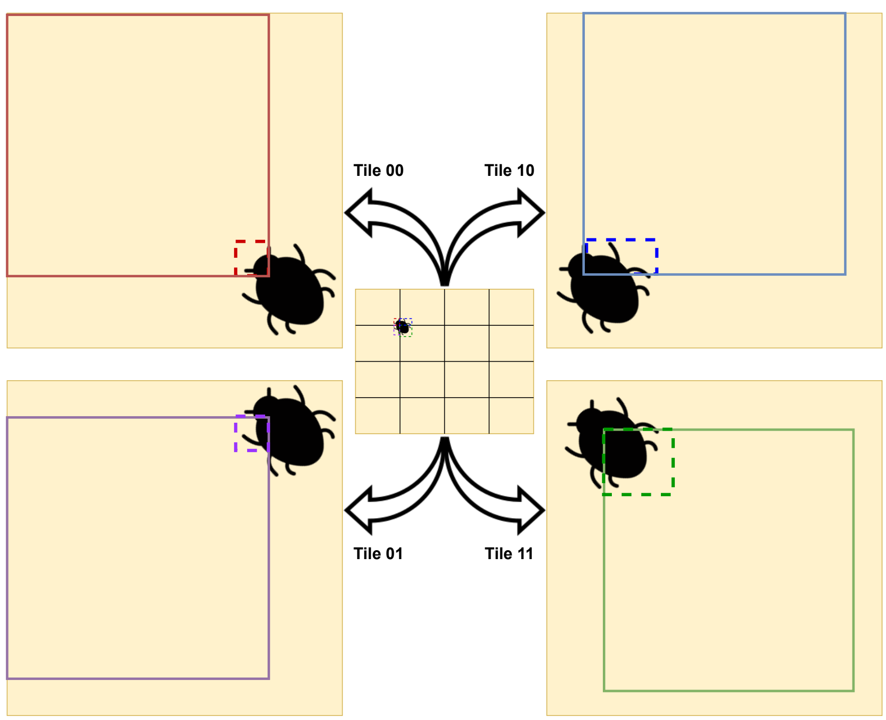
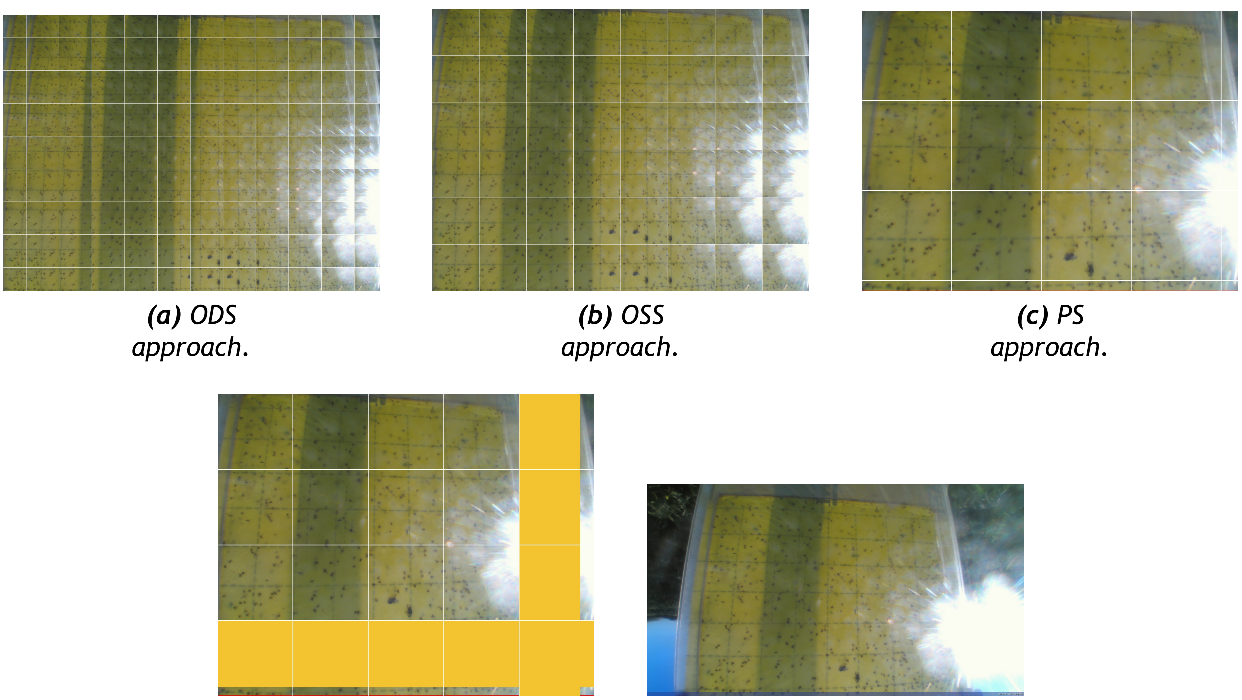


Figure 14. *Illustration of an insect cut by tiles obtained without overlap*

This situation would complicate the process of reconstructing the bounding boxes in the original image as the creation of the new bounding box based on the original ones would become complex and there would be a wide variety of possibilities when verifying which bounding boxes belong to the same object.

Arising this problem, the development concentrated on two new approaches, namely:

- **Overlapping with Different Size (ODS)** : Tiles get different dimensions depending on the position of the tiles relative to the image and overlapping occurs (Figure 15a);
- **Overlapping with Same Size (OSS)** : Tiles are all of the same dimensions (320x320px). Zones may have more overlapping areas than others (Figure 15b).



(a) *ODS approach.*

(b) *OSS approach.*

(c) *PS approach.*

(d) *PSB approach.*

(e) *Original image.*

Figure 15*. Yellow sticky card splitting approaches.*

For all the tests performed, the amount of images used was the same, 248 insect trap images. However, due to the different approaches to perform the splitting, the amount of tiles used to train the models was different for each approach. For ODS and OSS 11375 and 5092 tiles were used when training and testing the models, respectively.

In all approaches, 80% of the data-set was used for train, 10% for validation and the remaining 10% for test.

The overlapping of tiles was done with caution making sure that the overlapping zone occupied an area of 160x160px (Figure 16). This was because by analysing the images, the insects present in them and questioning experts in the area it was discovered that the maximum area that a bounding box could occupy would be below these values. In this way, the problem that arose was solved. This was due to the fact that if an insect is split between tiles it will be partially detected in some tiles but will always be fully detected on a neighbouring tile, this type of situation is illustrated in Figure 16. Thus, when reconstructing the image it became only necessary to understand which detections are overlapped, by checking and comparing each bounding box position, which ones have the largest area and confidence and remove the duplicated ones. This way only the bounding boxes detecting the whole object would remain.

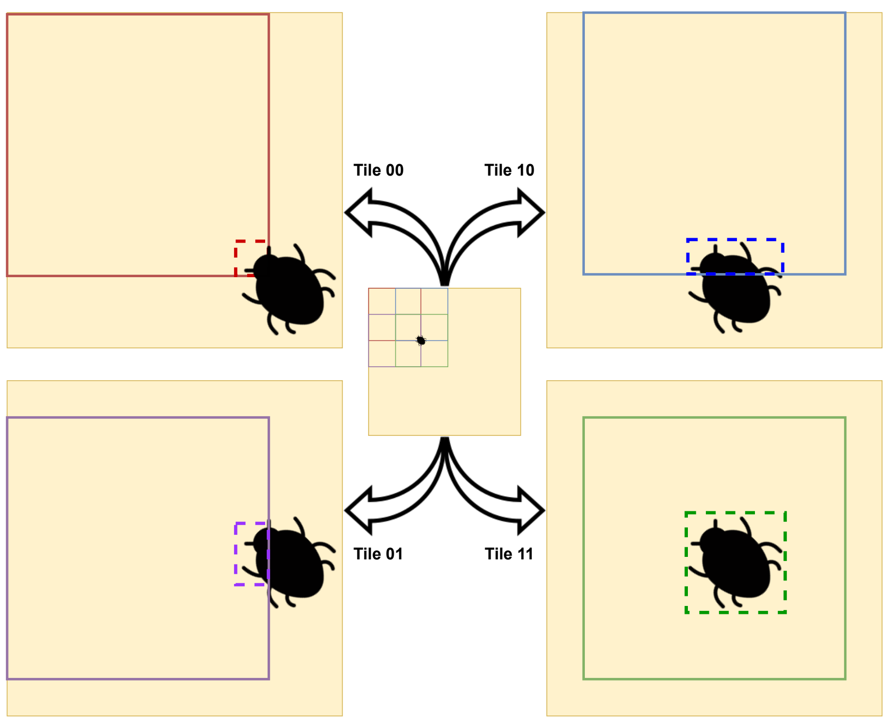


Figure 16. *Illustration of an insect cut by tiles obtained with overlap.*

From tests carried out, some incorrect detections and detections to be performed were observed, but they were in minority when compared to the accurate ones. These flaws can be suppressed when the values obtained in each image are associated with groups, for example, between 0 and 20 few insects, between 20 and 100 some insects, etc. This association is important when analysing the data and verifying the respective correlations. This type of failures is reflected in the *mAP_0.5-0.95* metric that is significantly lower than the *mAP_0.5* metric in all tests performed, this is depicted in Tables 9 and 10. These tables reflect that approach ODS and OSS had similar results with the YoloV5x having the best results in both cases. However, due to the uniformity that OSS provides to the dimensions of the tiles without the need for resizing, the OSS approach was the one taken in to account for the development of the remaining work.

Model	**Phase**	**mAP_0.5**	**mAP_0.5-0.95**	**Precision**	**Recall**	**F1-Score**
YoloV5s	Training	0.973	0.678	0.982	0.935	0.958
	Testing	0.945	0.539	0.937	0.89	0.913
YoloV5m	Training	0.975	0.7	0.976	0.94	0.958
	Testing	0.933	0.554	0.908	0.88	0.894
YoloV5l	Training	0.979	0.724	0.986	0.947	0.966
	Testing	0.952	0.567	**0.938**	**0.906**	**0.922**
YoloV5x	Training	0.98	0.733	0.982	0.951	0.966
	Testing	**0.952**	**0.573**	0.935	0.9	0.917

***Table 9**. YoloV5 Insect Model results for ODS.*

Model	**Phase**	**mAP_0.5**	**mAP_0.5-0.95**	**Precision**	**Recall**	**F1-Score**
YoloV5s	Training	0.964	0.632	0.963	0.940	0.951
	Testing	0.923	0.497	0.912	0.853	0.882
YoloV5m	Training	0.975	0.691	0.982	0.946	0.964
	Testing	0.946	0.542	0.946	0.874	0.909
YoloV5l	Training	0.973	0.694	0.981	0.939	0.960
	Testing	0.937	0.543	**0.951**	0.862	0.904
YoloV5x	Training	0.976	0.713	0.983	0.95	0.966
	Testing	**0.944**	**0.559**	0.942	**0.88**	**0.910**

***Table 10**. YoloV5 Insect Model results for OSS.*

The results observed were even better or similar to related work, such as work done in [33]–[35] that had a precision of 93.3%, a *mAP* of 94.7% and an accuracy of 94%, respectively.

An analysis of the context of the problem and communication with end users led to the conclusion that it was important to perform an analysis regarding false positives and false negatives. This is because if too many false detections (false positives) occur, it would mean a possible acquisition by end users of products in vain or a constant check in the field of the values reflected by the detections. On the other hand, if too many false negatives occur it would mean the possible appearance of pests without alerting the end user. Furthermore, the non-occurrence of false positives and negatives will always be the best situation to ensure that the correlations performed with other data acquired to predict the appearance of pests are not biased. Therefore, the precision-recall curve was analysed since precision and recall reflect the false negatives and false positives values, respectively. Figure 17 depicts the plot of the precision-recall curve.

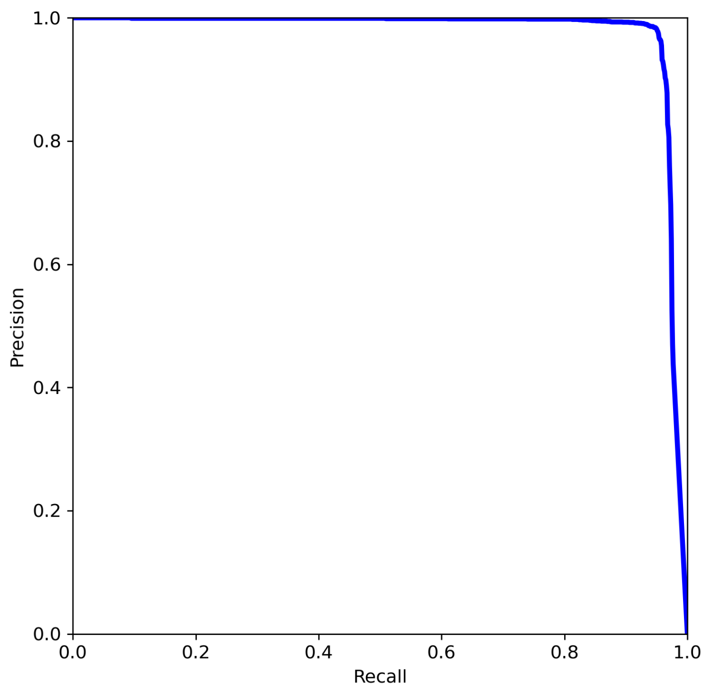


***Figure 17**. Precision-recall curve for YoloV5x OSS approach.*

An observation of the plot in Figure 17 suggests that it is possible to simultaneously achieve high values for the precision and recall by choosing a suitable decision threshold for the confidence associated to the bounding boxes outputted by the object detection model. This means that low values for both false positive and false negative rates can be kept, confirming the model's good performance [115].

4.2. Vegetation indices acquisition of UAV images

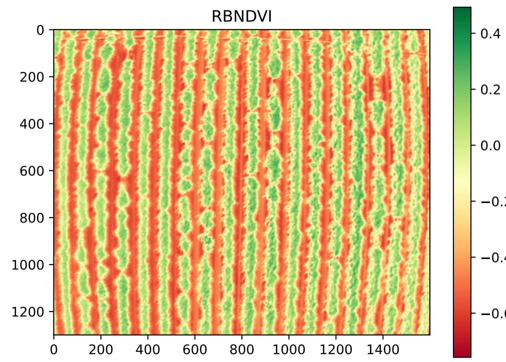
The acquisition of VI was performed in the UAV images since these technique is robust to the variation of the sun illumination [87]. Table 11 lists the indices considered.

Indice	Formula
NDVI	(NIR - RED)/(NIR + RED) [81]
NDRE	(NIR - RED_EDGE)/(NIR + RED_EDGE)[86]
GNDVI	(NIR - GREEN)/(NIR + GREEN) [87]
BNDVI (Blue Normalized Difference Vegetation Index)	(NIR - BLUE)/(NIR + BLUE) [88]
GRNDVI (Green and Red Normalized Difference Vegetation Index)	(NIR - (GREEN + RED))/(NIR + (GREEN + RED)) [89]
GBNDVI (Green and Blue Normalized Difference Vegetation Index)	(NIR - (GREEN + BLUE))/(NIR + (GREEN + BLUE)) [89]
RBNDVI (Red and Blue Normalized Difference Vegetation Index (RBNDVI)	(NIR - (RED + BLUE))/(NIR + (RED + BLUE)) [89]
CI_green (ChlorophyllIndex Green)	(NIR/GREEN) - 1 [84]
CVI (Chlorophyll Index Vegetation)	NIR * (GREEN/RED) - 1 [84]
ARI (Anthocyanin Reflectance Index)	(1/GREEN) - (1/RED_EDGE) [87]
MARI (Modified Anthocyanin Reflectance Index)	(1/GREEN) - (1/RED_EDGE) * NIR [87]
RGI (Redand Green Index)	RED/GREEN [87]
ACI (Anthocyanin Content Index)	GREEN/NIR [87]
MACI (Modified Anthocyanin Content Index)	NIR/GREEN [87]
DVI (Difference Vegetation Index)	NIR - RED [87]
RENDVI (Red-Edge-Normalized Difference Vegetation Index)	(RED_EDGE - RED)/(RED_EDGE + RED) [88]
ARVI (Atmospherically Resistant Vegetation Index)	(NIR – (2 * RED) + BLUE) / (NIR + (2 * RED) + BLUE) [85]
NDWI (Normalized Difference Water Index)	(GREEN – NIR) / (GREEN + NIR) [85]
OSAVI (Optimized Soil Adjusted Vegetation Index)	(NIR – RED) / (NIR + RED + 0.16) [85]
VARI (Visible Atmospherically Resistant Index)	(GREEN – RED) / (GREEN + RED – BLUE) [85]
SIPVI (Structure Intensive Pigment Vegetation Index)	(NIR – BLUE) / (NIR – RED) [85]
RECl (Red-Edge Chlorophyll Vegetation Index)	(NIR / RED) – 1 [85]

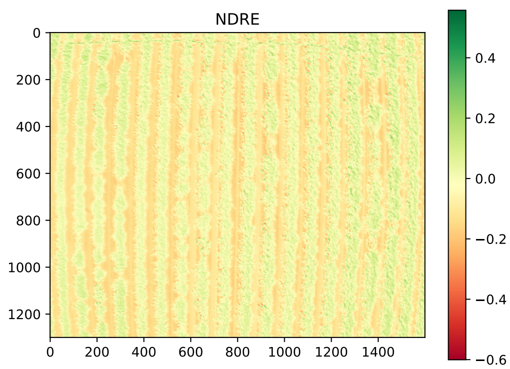
Table 11*. VI studied.*

The literature review and accessible spectra in UAV images were used to identify which indices to acquire since they are derived using mathematical formulas based on distinct spectrum bands. It was also sought the assistance of specialists in the field, ANDANTE partners, to determine which indices would be most useful to utilise. Some of the indices studies are depicted in Figure 18. This work was developed with the aim of understanding in the future which of the indices now studied are the ones that effectively have a stronger correlation with the other acquired data in order to optimise the algorithm for predicting the appearance of DaP.

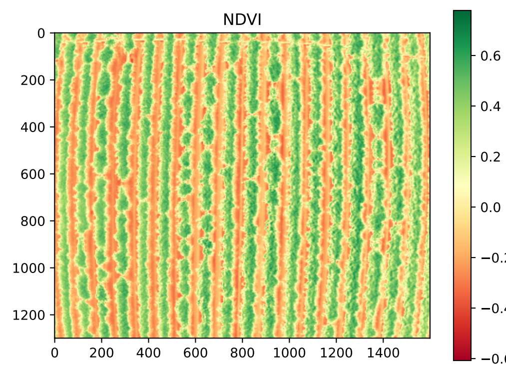
(a) *Original Image.*



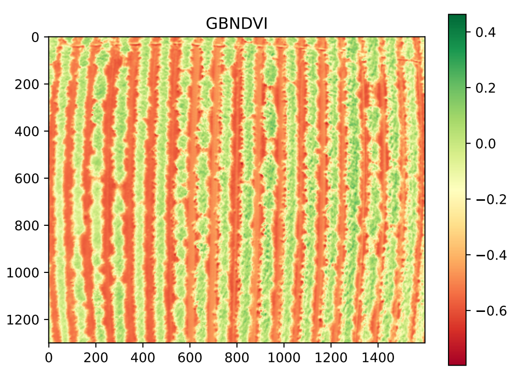
(b) *RBNDVI.*



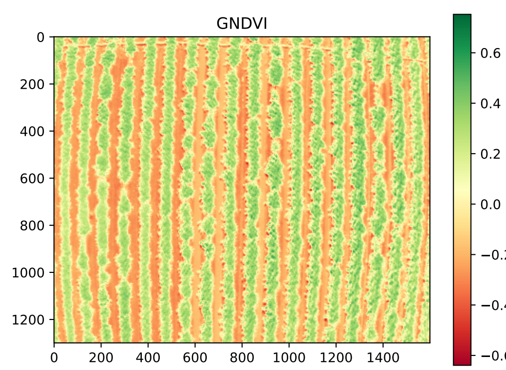
(c) *NDRE.*



(d) *NDVI.*



(e) *GBNDVI.*



(f) *GNDVI.*

Figure 18. *Examples of calculated VI.*

4.3. Disease classification on tomato leaves images

The development of mechanisms associated with the ground level images of tomato plantations was performed using PlantVillage data-set, since the data-set from ANDANTE did not have the necessary uniformity because it was collected by different people with different devices, as explained in section 3.2.3. This data-set is composed of laboratory images captured in a controlled environment with a uniform background. This was the approach adopted due to the insufficiency of uniform images acquired on the fields depicting tomato diseases. Nevertheless, training and testing the ML models using the images on the PlantVillage data-set should be similar to using another tomato leaf data-set. It is thus expected that, once sufficient uniform diseased tomato images are acquired on the fields, the AI mechanisms described in this dissertation can be easily adapted to the new data.

The disease classification problem associated to the tomato leaf images in the PlantVillage data-set contains 10 classes, where nine correspond to the diseases mentioned in section 3.2.3 and the remaining belongs to the healthy plants. To achieve the goal, the data was first organised into training, validation and testing sets. Then, different pre-trained CNN models were trained and tested, using a transfer learning approach. The network weights used for transfer learning were those associated to the ImageNet [116], a widely used large public image data-set. Additionally, the Keras Tuner [117] was used to build the most appropriate fully connected layer for the model in question. Finally, for each pre-trained model, the best dense-layer architecture resulting from Keras Tuner was used for training.

Keras Tuner is a library that helps choosing the ideal set of hyperparameters for the model under development. Hyperparameters are variables that remain constant throughout the training of the model and have direct impact on its performance. There are two types of hyperparameters: model hyperparameters, which influence the final model chosen and its structure, such as the number of hidden layers and their sizes; algorithm hyperparameters, which influence the speed and quality of the learning algorithm, such as the learning rate for the SGD. To perform the search for the best hyperparameters, it is necessary to configure those that are considered. Then, Keras Tuner randomly trains the model according to the possibilities for each hyperparameter (e.g. if the number of neurons in a dense layer is set as a hyperparameter, a range of values must be defined and, according to it, the Keras Tuner randomly chooses a value for the number of neurons in that layer). The number of times the Keras Tuner randomly chooses hyperparameters and trains based on them, corresponds to the number of trials defined. Each trial can have numerous runs, that is, a given trial/training associated to a set of hyperparameters can be run more than once - this can be useful since simply initialising the weights differently in the same model can lead to results with considerable differences. At the end, the model and respective hyperparameters that obtained the best results among all the trials carried out are obtained.

The data-set was divided into three sets, with 8000 images associated to train, 2000 to validation and 1000 to test. The data was always balanced during this procedure, with 1000 images of each class in the training set and 100 images in the validation set.

TensorFlow [118][119] library was used to make data augmentation and to generate the matrices of the images with the respective labels. The models used were deep learning models from state of the art, namely: MobileNetV2 [71]; VGG16 [71]; ResNet152V2 [120]; InceptionV3 [75].

All performed tests share a common setup configuration, namely: the percentage of images associated to test set, train set and validation set; the data augmentation characteristics (rotation range of 30 degrees, zoom range of 0.15, width and height shift range of 0.2, shear range of 0.15 and horizontal flip enabled); the characteristics of the Keras Tuner and the respective hyperparameters configurations (Figure 19), with Softmax used as the activation function of the last layer, since it is more suitable for multi-class problems for associating probabilities to each class, and Relu used in the hidden layers [121]; the callbacks used in the models training, with the validation set loss used as the metric monitored and patience set to six; the way that the models were compiled, with the Adam used as optimiser, because it is robust and suitable to a wide range of optimisation problems in the ML field [122], and the categorical cross entropy function used as loss function; batch size (32), target size (224,224), input shape (224,224,3) and number of epochs (50).

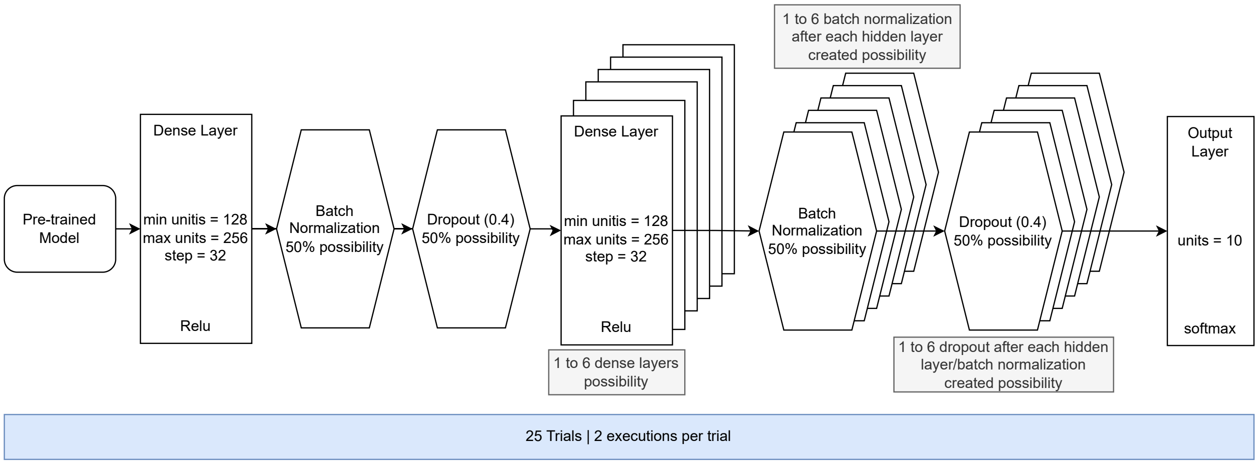


Figure 19*. Model architecture possibilities with Keras Tuner configurations.*

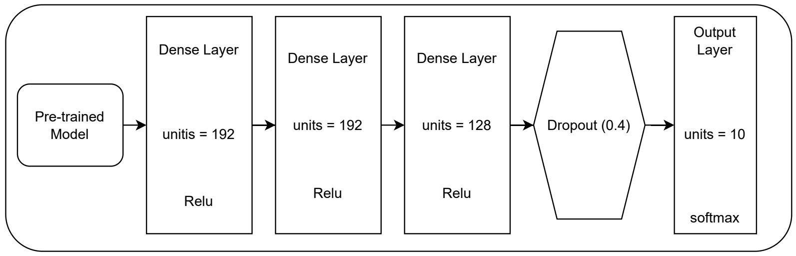
The best models obtained after the application of transfer learning and utilisation of Keras Tuner were those represented in Figure 20.

After obtaining the best models the next step was the training of the models. The best results from this training were the ones presented in Table 12, with additional information about the results in Appendix B.

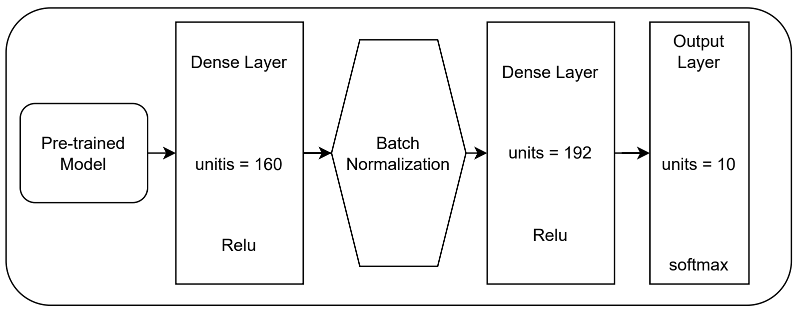
Model	Train set		Validation set		Test set	
	Accuracy	Loss	Accuracy	Loss	Accuracy	Loss
MobileNetV2	0.99	0.046	0.99	0.060	**0.97**	0.087
VGG16	0.89	0.330	0.90	0.330	0.86	0.498
ResNet152V2	0.98	0.070	0.99	0.040	**0.98**	0.046
InceptionV3	0.86	0.410	0.87	0.450	0.80	0.600

Table 12*. Results in ground level images.*

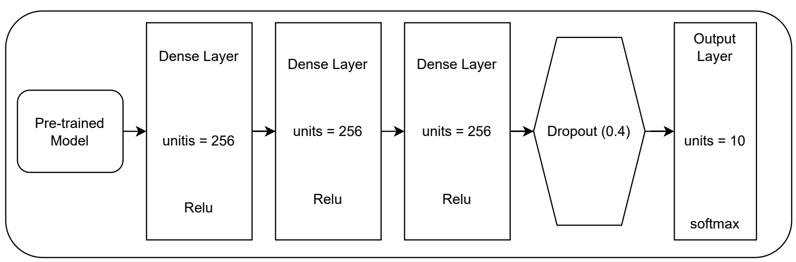
The model that showed the best performance for this problem was ResNet152V2 followed closely by MobileNetV2, with both models achieving higher accuracy values and lower loss values across all sets. Although it was not the best model, MobileNetV2 is a simpler model than ResNet152V2 and may be useful when computational power is limited. There was also verified some, but not very significant, overfitting in VGG16 and InceptionV3. In MobileNetV2 and ResNet152V2 there were realistic results, as there was no overfitting and/or underfitting, as the results from validation, training and test sets



***(a)** MobileNetV2 and VGG16.*



***(b)** ResNet152V2.*



***(c)** InceptionV3.*

***Figure 20**. Best final models for each pre-trained model.*

were always close. The best results obtained were quite similar to or surpassed similar works in the literature review [70][71].

It can be inferred from the confusion matrix of the two best models (Figure 21), with the values reflected in Tables 13 and 14, respectively, that MobileNetV2 had the lowest recall values, reaching 85% (highlighted in bold in Table 13) in the classification of disease *spider mites twospottedspider_mite*, while ResNet152V2 had the lowest value at 96% for diseases *mosaic_virus* and *late_light*. In this perspective it can be said that ResNet152V2 was the model that had a better performance classifying all classes.

	bacterial_spot ->	early_blight ->	late_blight ->	leaf_mold ->	septoria_leaf_spot ->	spider mites two-spotted spider mite ->	target spot ->	yellow leaf curl virus ->	mosais virus ->	healthy tomato->
bacterial_spot ->	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%
early_blight ->	0%	98%	1%	0%	1%	0%	0%	0%	0%	0%
late_blight ->	1%	0%	98%	0%	1%	0%	0%	0%	0%	0%
leaf_mold ->	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%
septoria_leaf_spot ->	0%	0%	2%	0%	98%	0%	0%	0%	0%	0%
spider mites two-spotted spider mite ->	0%	0%	0%	10%	0%	85%	1%	0%	0%	4%
target spot ->	1%	1%	0%	0%	0%	2%	96%	0%	0%	0%
yellow leaf curl virus ->	1%	0%	0%	0%	0%	1%	0%	98%	0%	0%
mosais virus ->	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%
healthy tomato ->	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%

***(a)** Confusion Matrix of MobileNetV2.*

	bacterial_spot ->	early_blight ->	late_blight ->	leaf_mold ->	septoria_leaf_spot ->	spider mites two-spotted spider mite ->	target spot ->	yellow leaf curl virus ->	mosais virus ->	healthy tomato->
bacterial_spot ->	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%
early_blight ->	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%
late_blight ->	1%	2%	96%	1%	0%	0%	0%	0%	0%	0%
leaf_mold ->	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%
septoria_leaf_spot ->	1%	0%	0%	0%	99%	0%	0%	0%	0%	0%
spider mites two-spotted spider mite ->	0%	0%	0%	0%	0%	97%	0%	0%	1%	2%
target spot ->	0%	1%	0%	0%	1%	1%	97%	0%	0%	0%
yellow leaf curl virus ->	1%	0%	0%	0%	0%	0%	0%	99%	0%	0%
mosais virus ->	0%	0%	0%	0%	4%	0%	0%	0%	96%	0%
healthy tomato ->	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%

***(b)** Confusion Matrix of ResNet152V2.*

***Figure 21**. Confusion Matrix's of the best leaf classification models*

	Precision	Recall	F1-Score
bacterial spot	0.97	1.00	0.99
early blight	0.99	0.98	0.98
late blight	0.97	0.98	0.98
leaf mold	0.91	1.00	0.95
septoria leaf spot	0.98	0.98	0.98
spider mites two-spotted spider_mite	0.97	**0.85**	0.90
target spot	0.99	0.96	0.97
yellow leaf curl virus	1.00	0.98	0.99
mosaic virus	1.00	1.00	1.00
healthy tomato	0.96	1.00	0.98
Accuracy	-	-	0.97
Macro average	0.97	0.97	0.97
Weighted average	0.97	0.97	0.97

Table 13. *MobileNetV2 additional metrics.*

	Precision	Recall	F1-Score
bacterial spot	0.97	1.00	0.99
early blight	0.97	1.00	0.99
late blight	1.00	0.96	0.98
leaf mold	0.99	1.00	1.00
septoria leaf spot	0.95	0.99	0.97
spider mites two-spotted spider_mite	0.99	0.97	0.98
target spot	1.00	0.97	0.98
yellow leaf curl virus	1.00	0.99	0.99
mosaic virus	0.99	0.96	0.97
healthy tomato	0.98	1.00	0.99
Accuracy	-	-	0.98
Macro average	0.98	0.98	0.98
Weighted average	0.98	0.98	0.98

Table 14. *ResNet152V2 additional metrics.*

CHAPTER 5

Conclusion

The main objective of this dissertation was to develop four modules to contribute to the development of a system capable to detect and acquire data from tomato crops and forecast the appearance of DaP in those crops based in the data acquired. The data was associated with weather stations, ground level and aerial images of tomato plantations and insect traps in these plantations. Furthermore, the realisation of a literature review that contributes to the future optimisation of the project was specified as an objective.

The initial aim was to perform DaP outbreak forecast through ML and image processing techniques using data from weather stations and aerial, ground level and insect trap images of tomato crops. However, this objective was discarded due to the quantity and condition of the data provided (limitations that are detailed in section 5.1). Therefore, the objectives in section 1.3 were defined and the work was thus developed in that direction.

The first module developed in this dissertation, which came from a need, was the development of an web information system, associated to a database where all data related to the project could be stored in a structured and organized way. Consequently, a web application was developed using the PyForms library and Django with the SQLite database. It can be concluded that an web information system that was able to suppress needs aroused was developed since with this application it became possible to access and analyse collected data, as the analysis of the number of insects detected in a trap on a certain field in a certain period.

Regarding object detection module, the second module developed, it was done in the insect traps images. The detection associated to the yellow sticky card and the subsequent training of AI models were performed in a first phase. In this phase optimal results were obtained using YoloV5s, and it was possible to perform the detection of yellow stick cards in all data-set.

The second phase was dependent on the first, as it was supposed to use the bounding box associated to the detection performed of the yellow sticky card in order to improve the accuracy of the detections of the insects in the traps. At this stage a problem was faced: how to perform the splits on the yellow sticky card bounding box image in a way that maximises the quality of the model and minimises its complexity while not causing insects to be lost during the process of splitting and reconstructing the bounding boxes on the original image? The defined approach contemplated the use of overlapping tiles. Within the tests carried out, the OSS approach ended up generally having the best results. In this approach the tiles had the same dimensions and were overlapped and the best results was 98.3% of precision with YoloV5x model.

It was possible to develop an insect detection model with the need for human supervision at times since the number and location of the bounding boxes performed may be inaccurate. However, these errors were never in substantial quantities and can end up mostly suppressed when associating the amount of detections performed in an image to a group. This association has advantages at the time of data treatment and analysis. This achievement was an important step as it provides fundamental future data for forecasting the appearance of DaP or for simple data analysis associated with the plantations.

Regarding the UAV images, the third module, only image processing was performed and metrics concerning VI was acquired. These indices were calculated based on the literature, the type of spectra of the images provided and the knowledge obtained from specialists associated to the ANDANTE project. This was very important since, as mentioned in the literature, VI are very important data for the detection, classification and forecast of DaP in crops. No evaluation could be done for this model due to delays by partners in providing information on UAV images.

Finally, in relation to ground level images, the fourth module, it was intended to perform a leaf and fruit detection and the subsequent classification of the type of diseases, in case of illness. However, due to the condition of the data provided this was not possible. Therefore, only leaf health classification was performed and in a public data-set. This happened since that the classification phase after the detection of the leaves in the original image has a similar process to the one performed in this work, serving this work as a basis for that phase.

The public data-set used was PlantVillage and positive results were obtained. Through the use of Keras Tuner, which allowed testing various options and weights for the constructed NN, the best models were obtained. It was concluded that the model with the best performance was ResNet152V2 closely followed by MobileNetV2. This is due to the fact that when the confusion matrix of each network was analysed, it was concluded that ResNet152V2 had a better performance in classifying all the diseases in question. However, if searching for the network with the lowest computational cost, MobileNetV2 would be the best option as it is smaller and less complex.

In the context of this dissertation a paper was published in MDPI Agriculture [123]. This paper depicts a literature review on ML techniques used in the agricultural sector, focusing on the classification, detection, and prediction of DaPs, with an emphasis on tomato crops. This survey aims to contribute to the advancement of smart farming and precision agriculture by encouraging the development of techniques that will allow farmers to use fewer pesticides and chemicals while maintaining and improving crop quality and productivity [124]. It was thus concluded that the objective of conducting a review that would contribute to the project was achieved. In addition, this dissertation resulted in another paper whose content is related to the work developed on insect trap images (present mainly in Section 3.2.1 and Section 4.1). This paper was published in MDPI Agriculture Special Issue "The Application of Machine Learning in Agriculture". More information about these papers is provided in Appendix C.

Taking into account the ANDANTE project, it can be concluded that this dissertation contributes to the development of the collaborative system that is intended to be developed in ANDANTE, since this dissertation is the first step towards this system development due to the modules and literature review performed. Thus, the primary objective of this dissertation was met by contributing to the final goal of ANDANTE, which is the prediction of the presence of DaP in tomato crops.

5.1. Limitations

The results and type of work carried out suggest that the quantity and quality of the data provided had a great influence on the objectives set and consequently on the results obtained. This was very much due to the fact that this dissertation was associated with an European project, suffering the problems and dependencies of a real project.

As far as quantity is concerned, this limited this dissertation as large and varied datasets are needed to train the developed AI models since in this way the models become robust and prepared for future situations when they are put on the field. When this does not happen, the results obtained may not reflect reality, and the model may be performing well only for the few data that is used to train, validate and test the model.

Regarding the module associated with the web information system there were limitations in the tests to be carried out due to the delay in using the application by the ANDANTE partners. Besides, all the other limitations mentioned about the other modules also limited the work developed, since the web information system should fully incorporate the remaining modules of the project, such as the detection and registration of sick zones from UAV images at the moment of their import.

Concerning the module associated with the images of insect traps, there were limitations due to the absence of manual annotations of insects made by the ANDANTE partners, which made it impossible to develop models for the detection and classification of insects trained with all the available images.

The UAV imagery module had limitations due to the lack of manual annotations of tomatoes and diseased areas, making it impossible to develop detection and classification models of tomatoes and/or areas affected by diseases and/or pests.

Regarding the module associated with the images at ground level of tomato crops, there were limitations due to the non-existence of manual annotations concerning tomatoes, leaves and diseased areas, making it impossible to develop models for the detection and classification of tomatoes, leaves and/or areas affected by diseases and/or pests. In addition the lack of quantity of diversified and uniformed images, due to the image acquisition being done manually by different people, also limited the development in this direction.

Considering all the limitations, developments were carried out with the aim of using the data-sets coming from ANDANTE whenever possible. In the case of the insect traps images, since there were no manual annotations and the partners were not able to perform the annotations, a large number of images were manually annotated, concerning the insects and the yellow sticky traps, in order to develop an object detection model for insects and yellow sticky cards. Regarding the images at ground level, because there were no annotations or uniformity to allow the implementation of what was intended to be developed, a public data-set of sufficient quality was used. In the case of the images from the UAVs, since the necessary annotations were not available, it was calculated the VI and verified the appropriate ones to be used in the development of this project. However, certain tests, which would serve to filter through correlations which indices actually make sense in this problem, could not be carried out due to the delay and lack of information already mentioned.

All these limitations meant that it was not possible to achieve the initial objective of forecasting the appearance of DaP using historical data from aerial images, insect traps images, plants at ground level images and weather stations in tomato plantations. This is due to the fact that the objective required a diversity, quantity and uniformity of data that did not exist in any of the data-sets, except for the meteorological data, which could not be used individually to make the forecasts because there was no historical record of the zones were DaP occurred in the 2021 season. That said, as it was not possible to achieve such an objective, the objectives already mentioned were established.

5.2. Future work

Regarding future work, a direct link can be established if the limitations associated with this dissertation are explored. That is, much of the future work will be focused on developing new data acquisition processes in order to collect enough high-quality data to realise the goal, which could not be defined, of forecast DaP onset.

Concerning insect traps images, future work may focus on performing a better manual and/or semi-automatic annotation, using the models developed in this dissertation, of at least the data-set associated with the 2021 season. This will allow the developed models to be more robust and accurate. This will also provide insight into whether other methods and techniques need to be adopted for insect detection if the results obtained are not as expected. Thus demonstrating that training with a larger amount of annotations with more quality was not the only improvement to be performed for the optimisation of results of the models. Part of the future work may also focus on the differentiation of the insects detected in the yellow sticky cards, in which case it will be necessary to have more images since a greater diversity of data is required in order to cover the various types of insects to be identified.

Regarding the UAV images, with the proper manual annotation by experts of the diseased areas, part of the future work may be to automatically identify diseased areas, and in a second stage to classify the identified diseases. In addition, future work may concern the acquisition of the fruit density present in the image using image processing and ML techniques, although this work may be hampered due to atmospheric conditions that create shadows and different luminosities in the plantation.

Concerning the images at ground level, part of the future work can include standardising the acquired images in order to have a pipeline applicable to all images from the crop field. The future work will involve the development of models that identify diseased leaves and fruits and, in a second phase, classify these diseases. The development of the classification part is already largely present in this dissertation.

A common future work can be applied to the different data-sets. This consists in testing different models applying transfer learning with different pre-trained weights, different combinations of hyperparameters and different layers set as trainable. It should also be noted that, taking into account the ANDANTE project, the use of tiny models in the object detection and metrics acquisition models that run on IoT devices present in the fields could be an interesting future work to be carried out.

Besides the future work applied individually to each set of data, there is also the future work that consists of combining all the data in order to analyse it, check its correlations and develop a model that enables the forecast of the onset of DaP.

Finally part of the future work may involve the improvement of the web application developed in this dissertation. This in order that it becomes robust enough for farmers to receive alerts, forecasts, and input new data, and for developers, for the analysis, access and management of the data collected so far.

Taking into account the knowledge acquired during this dissertation, it is expected that it is possible to develop a system that predicts the appearance of diseases and pests in tomato plantations, taking into account meteorological data and aerial, ground-level and insect trap images of these plantations. However this will not be possible in the short term, since this final goal depends on many other components, such as the modules in this dissertation developed and their optimisations. Moreover these modules also depend on other processes, such as the way the data is acquired. Therefore, it is expected that it will be possible to achieve this goal but in the long term due to the complexity and dependencies of the project.

References

[1] M. Roser and L. Rodés-Guirao, "Future population growth," *Our World in Data*, 2013, https://ourworldindata.org/future-population-growth.

[2] D. Fróna, J. Szenderák, and M. Harangi-Rákos, "The challenge of feeding the world," *Sustainability*, vol. 11, no. 20, p. 5816, 2019.

[3] D. K. Ray, N. D. Mueller, P. C. West, and J. A. Foley, "Yield trends are insufficient to double global crop production by 2050," *PloS one*, vol. 8, no. 6, e66428, 2013.

[4] R. Thangaraj, S. Anandamurugan, P. Pandiyan, and V. K. Kaliappan, "Artificial intelligence in tomato leaf disease detection: A comprehensive review and discussion," *Journal of Plant Diseases and Protection*, pp. 1–20, 2021.

[5] FAO, "The future of food and agriculture: Trends and challenges," 2017.

[6] C. A. Deutsch, J. J. Tewksbury, M. Tigchelaar, D. S. Battisti, S. C. Merrill, R. B. Huey, and R. L. Naylor, "Increase in crop losses to insect pests in a warming climate," *Science*, vol. 361, no. 6405, pp. 916–919, 2018.

[7] A. Anton, S. Rustad, G. F. Shidik, and A. Syukur, "Classification of tomato plant diseases through leaf using gray-level co-occurrence matrix and color moment with convolutional neural network methods," in *Smart Trends in Computing and Communications: Proceedings of SmartCom 2020*, Springer, 2021, pp. 291–299.

[8] M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep learning for tomato diseases: Classification and symptoms visualization," *Applied Artificial Intelligence*, vol. 31, no. 4, pp. 299–315, 2017.

[9] FAO, "FAOSTAT: FAO statistical databases," 2021. [Online]. Available: `https://www.fao.org/faostat/en/#data/QCL` (visited on 10/25/2021).

[10] *PlantVillage tomato | diseases and pests, description, uses, propagation*, `https://plantvillage.psu.edu/topics/tomato/infos`, Accessed: 2022-01-11.

[11] S. Verma, A. Chug, and A. P. Singh, “Prediction models for identification and diagnosis of tomato plant diseases,” in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, 2018, pp. 1557–1563.

[12] P. Kartikeyan and G. Shrivastava, “Review on emerging trends in detection of plant diseases using image processing with machine learning,” *International Journal of Computer Applications*, vol. 975, p. 8887, 2021.

[13] A. Kamilaris, A. Kartakoullis, and F. X. Prenafeta-Boldú, “A review on the practice of big data analysis in agriculture,” *Computers and Electronics in Agriculture*, vol. 143, pp. 23–37, 2017.

[14] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng, “CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning,” *arXiv e-prints*, arXiv:1711.05225, arXiv:1711.05225, Nov. 2017. arXiv: `1711.05225 [cs.CV]`.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.

[16] B. Huval, T. Wang, S. Tandon, J. Kiske, W. Song, J. Pazhayampallil, M. Andriluka, P. Rajpurkar, T. Migimatsu, R. Cheng-Yue, F. A. Mujica, A. Coates, and A. Y. Ng, “An empirical evaluation of deep learning on highway driving,” *CoRR*, vol. abs/1504.01716, 2015. arXiv: `1504.01716`. [Online]. Available: `http://arxiv.org/abs/1504.01716`.

[17] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, “Machine learning and the physical sciences,” *Reviews of Modern Physics*, vol. 91, no. 4, p. 045 002, 2019.

[18] S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using deep learning for image-based plant disease detection,” *Frontiers in plant science*, vol. 7, p. 1419, 2016.

[19] Q. Xiao, W. Li, Y. Kai, P. Chen, J. Zhang, and B. Wang, “Occurrence prediction of pests and diseases in cotton on the basis of weather factors by long short term memory network,” *BMC bioinformatics*, vol. 20, no. 25, pp. 1–15, 2019.

[20] A. Gutierrez, A. Ansuategi, L. Susperregi, C. Tubío, I. Rankić, and L. Lenža, “A benchmarking of learning strategies for pest detection and identification on tomato plants for autonomous scouting robots using internal databases,” *Journal of Sensors*, vol. 2019, 2019.

[21] *Andante*, https://www.andante-ai.eu/, Accessed: 2021-12-09.

[22] *Andante use case 2.2: Tomato pests and diseases forecast*, https://www.andante-ai.eu/project/use-case-2-2-tomato-pests-and-diseases-forecast/, Accessed: 2021-12-09.

[23] *Precision Agriculture, an international journal on advances in precision agriculture*, https://www.springer.com/journal/11119, Accessed: 2021-12-09.

[24] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, “A design science research methodology for information systems research,” *Journal of management information systems*, vol. 24, no. 3, pp. 45–77, 2007.

[25] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and P. Group, “Preferred reporting items for systematic reviews and meta-analyses: The prisma statement,” *PLoS medicine*, vol. 6, no. 7, e1000097, 2009.

[26] W. Ding and G. Taylor, “Automatic moth detection from trap images for pest management,” *Computers and Electronics in Agriculture*, vol. 123, pp. 17–28, 2016.

[27] D. Nikitenko, M. Wirth, and K. Trudel, “Applicability of white-balancing algorithms to restoring faded colour slides: An empirical evaluation.,” *Journal of Multimedia*, vol. 3, no. 5, 2008.

[28] G. Buchsbaum, “A spatial processor model for object colour perception,” *Journal of the Franklin institute*, vol. 310, no. 1, pp. 1–26, 1980.

[29] P. Dollar, C. Wojek, B. Schiele, and P. Perona, “Pedestrian detection: An evaluation of the state of the art,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 4, pp. 743–761, 2011.

[30] J. Hosang, R. Benenson, and B. Schiele, “Learning non-maximum suppression,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4507–4515.

[31] C. Hu, X. Liu, Z. Pan, and P. Li, “Automatic detection of single ripe tomato on plant combining faster r-cnn and intuitionistic fuzzy set,” *IEEE Access*, vol. 7, pp. 154 683–154 696, 2019.

[32] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” *Advances in neural information processing systems*, vol. 28, pp. 91–99, 2015.

[33] W. Li, Z. Yang, J. Lv, T. Zheng, M. Li, and C. Sun, “Detection of small-sized insects in sticky trapping images using spectral residual model and machine learning,” *Frontiers in Plant Science*, vol. 13, 2022.

[34] W. Yun, J. P. Kumar, S. Lee, D.-S. Kim, and B.-K. Cho, “Deep learning-based system development for black pine bast scale detection,” *Scientific reports*, vol. 12, no. 1, pp. 1–10, 2022.

[35] B. Ramalingam, R. E. Mohan, S. Pookkuttath, B. F. Gómez, C. S. C. Sairam Borusu, T. Wee Teng, and Y. K. Tamilselvam, “Remote insects trap monitoring system using deep learning framework and iot,” *Sensors*, vol. 20, no. 18, p. 5280, 2020.

[36] Y. Zhang, C. Song, and D. Zhang, “Deep learning-based object detection improvement for tomato disease,” *IEEE Access*, vol. 8, pp. 56 607–56 614, 2020.

[37] Y. Zhong, J. Wang, J. Peng, and L. Zhang, “Anchor box optimization for object detection,” in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2020, pp. 1286–1294.

[38] T. Zhang, Z. Huang, W. You, J. Lin, X. Tang, and H. Huang, “An autonomous fruit and vegetable harvester with a low-cost gripper using a 3d sensor,” *Sensors*, vol. 20, no. 1, p. 93, 2020.

[39] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.

[40] X. Ni, C. Li, H. Jiang, and F. Takeda, “Three-dimensional photogrammetry with deep learning instance segmentation to extract berry fruit harvestability traits,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 171, pp. 297–309, 2021.

[41] D. Bolya, C. Zhou, F. Xiao, and Y. J. Lee, “Yolact: Real-time instance segmentation,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 9157–9166.

[42] X. Wang, T. Kong, C. Shen, Y. Jiang, and L. Li, “Solo: Segmenting objects by locations,” in *European Conference on Computer Vision*, Springer, 2020, pp. 649–665.

[43] E. Xie, P. Sun, X. Song, W. Wang, X. Liu, D. Liang, C. Shen, and P. Luo, “Polarmask: Single shot instance segmentation with polar representation,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 12 193–12 202.

[44] H. Chen, K. Sun, Z. Tian, C. Shen, Y. Huang, and Y. Yan, “Blendmask: Top-down meets bottom-up for instance segmentation,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 8573–8581.

[45] X. Wang, R. Zhang, T. Kong, L. Li, and C. Shen, “Solov2: Dynamic, faster and stronger,” *arXiv e-prints*, arXiv–2003, 2020.

[46] S. Lin, Y. Jiang, X. Chen, A. Biswas, S. Li, Z. Yuan, H. Wang, and L. Qi, “Automatic detection of plant rows for a transplanter in paddy field using faster r-cnn,” *IEEE Access*, vol. 8, pp. 147 231–147 240, 2020.

[47] M. D. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” in *European conference on computer vision*, Springer, 2014, pp. 818–833.

[48] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.

[49] J. Fan, T. Huo, X. Li, T. Qu, B. Gao, and H. Chen, “Covered vehicle detection in autonomous driving based on faster rcnn,” in *2020 39th Chinese Control Conference (CCC)*, IEEE, 2020, pp. 7020–7025.

[50] Y. Mu, T.-S. Chen, S. Ninomiya, and W. Guo, “Intact detection of highly occluded immature tomatoes on plants using deep learning techniques,” *Sensors*, vol. 20, no. 10, p. 2984, 2020.

[51] W. Li, D. Wang, M. Li, Y. Gao, J. Wu, and X. Yang, “Field detection of tiny pests from sticky trap images using deep learning in agricultural greenhouse,” *Computers and Electronics in Agriculture*, vol. 183, p. 106 048, 2021.

[52] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, “A review of yolo algorithm developments,” *Procedia Computer Science*, vol. 199, pp. 1066–1073, 2022.

[53] G. Liu, J. C. Nouaze, P. L. Touko Mbouembe, and J. H. Kim, “Yolo-tomato: A robust algorithm for tomato detection based on yolov3,” *Sensors*, vol. 20, no. 7, p. 2145, 2020.

[54] J. Redmon and A. Farhadi, “Yolov3: An incremental improvement,” *arXiv preprint arXiv:1804.02767*, 2018.

[55] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[56] X. Jin, J. Che, and Y. Chen, “Weed identification using deep learning and image processing in vegetable plantation,” *IEEE Access*, vol. 9, pp. 10 940–10 950, 2021.

[57] X. Zhou, D. Wang, and P. Krähenbühl, “Objects as points,” *arXiv preprint arXiv:1904.07850*, 2019.

[58] L. Tang, L. Tian, and B. L. Steward, “Color image segmentation with genetic algorithm for in-field weed sensing,” *Transactions of the ASAE*, vol. 43, no. 4, p. 1019, 2000.

[59] C. Günther, N. Jansson, M. Liwicki, and F. Simistira-Liwicki, “Towards a machine learning framework for drill core analysis,” in *2021 Swedish Artificial Intelligence Society Workshop (SAIS)*, IEEE, 2021, pp. 1–6.

[60] M. Guillermo, R. K. Billones, A. Bandala, R. R. Vicerra, E. Sybingco, E. P. Dadios, and A. Fillone, “Implementation of automated annotation through mask rcnn object detection model in cvat using aws ec2 instance,” in *2020 IEEE region 10 conference (TENCON)*, IEEE, 2020, pp. 708–713.

[61] M. Agarwal, S. K. Gupta, and K. Biswas, “Development of efficient cnn model for tomato crop disease identification,” *Sustainable Computing: Informatics and Systems*, vol. 28, p. 100 407, 2020.

[62] T. Kasinathan, D. Singaraju, and S. R. Uyyala, “Insect classification and detection in field crops using modern machine learning techniques,” *Information Processing in Agriculture*, vol. 8, no. 3, pp. 446–457, 2021.

[63] L. Nanni, A. Manfè, G. Maguolo, A. Lumini, and S. Brahnam, “High performing ensemble of convolutional neural networks for insect pest image detection,” *Ecological Informatics*, vol. 67, p. 101 515, 2022.

[64] Q. Dai, X. Cheng, Y. Qiao, and Y. Zhang, “Agricultural pest super-resolution and identification with attention enhanced residual and dense fusion generative and adversarial network,” *IEEE Access*, vol. 8, pp. 81 943–81 959, 2020.

[65] D. J. Patel and N. Bhatt, “Insect identification among deep learning’s meta-architectures using tensorflow,” *Int. J. Eng. Adv. Technol*, vol. 9, no. 1, pp. 1910–1914, 2019.

[66] S. M. Hassan, A. K. Maji, M. Jasiński, Z. Leonowicz, and E. Jasińska, “Identification of plant-leaf diseases using cnn and transfer-learning approach,” *Electronics*, vol. 10, no. 12, p. 1388, 2021.

[67] S. U. Habiba and M. K. Islam, “Tomato plant diseases classification using deep learning based classifier from leaves images,” in *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, IEEE, 2021, pp. 82–86.

[68] H. Hong, J. Lin, and F. Huang, “Tomato disease detection and classification by deep learning,” in *2020 International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, IEEE, 2020, pp. 25–29.

[69] P. Tm, A. Pranathi, K. SaiAshritha, N. B. Chittaragi, and S. G. Koolagudi, “Tomato leaf disease detection using convolutional neural networks,” in *2018 eleventh international conference on contemporary computing (IC3)*, IEEE, 2018, pp. 1–5.

[70] S. M. Gharghory, “Performance analysis of efficient pre-trained networks based on transfer learning for tomato leaf diseases classification,”

[71] S. Wagle and H. Ramachandran, “A deep learning-based approach in classification and validation of tomato leaf disease,” *Traitement du Signal*, vol. 38, pp. 699–709, Jun. 2021. DOI: `10.18280/ts.380317`.

[72] B. A. Abdelghani, S. Banitaan, M. Maleki, and A. Mazen, "Kissing bugs identification using convolutional neural network," *IEEE Access*, vol. 9, pp. 140 539–140 548, 2021.

[73] P. Das, J. K. P. Singh Yadav, and A. K. Yadav, "An automated tomato maturity grading system using transfer learning based alexnet.," *Ingénierie des Systèmes d'Information*, vol. 26, no. 2, 2021.

[74] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[75] V. Kusanur and V. S. Chakravarthi, "Using Transfer Learning for Nutrient Deficiency Prediction and Classification in Tomato Plant," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 10, pp. 784–790, 2021, ISSN: 21565570. DOI: 10.14569/IJACSA.2021.0121087.

[76] L. Zhang, J. Jia, Y. Li, W. Gao, and M. Wang, "Deep learning based rapid diagnosis system for identifying tomato nutrition disorders," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 13, no. 4, pp. 2012–2027, 2019.

[77] O. O. Arjenaki, P. A. Moghaddam, and A. M. Motlagh, "Online tomato sorting based on shape, maturity, size, and surface defects using machine vision," *Turkish Journal of Agriculture and Forestry*, vol. 37, no. 1, pp. 62–68, 2013.

[78] J. M. Duarte-Carvajalino, D. F. Alzate, A. A. Ramirez, J. D. Santa-Sepulveda, A. E. Fajardo-Rojas, and M. Soto-Suárez, "Evaluating late blight severity in potato crops using unmanned aerial vehicles and machine learning algorithms," *Remote Sensing*, vol. 10, no. 10, p. 1513, 2018.

[79] W. Dake and M. Chengwei, "The support vector machine (svm) based near-infrared spectrum recognition of leaves infected by the leafminers," in *First International Conference on Innovative Computing, Information and Control-Volume I (ICICIC'06)*, IEEE, vol. 3, 2006, pp. 448–451.

[80] *Measuring Vegetation ndvi and evi*, https://earthobservatory.nasa.gov/features/MeasuringVegetation/measuring_vegetation_2.php, Accessed: 2022-01-11.

[81] J. Rouse Jr, R. Haas, J. Schell, and D. Deering, "Monitoring vegetation systems in the great plains with erts," in *Third Earth Resources Technology Satellite-1 Symposium: The Proceedings of a Symposium Held by Goddard Space Flight Center at Washington, DC on*, vol. 351, 1973, p. 309.

[82] S. Skawsang, M. Nagai, N. K Tripathi, and P. Soni, "Predicting rice pest population occurrence with satellite-derived crop phenology, ground meteorological observation, and machine learning: A case study for the central plain of thailand," *Applied Sciences*, vol. 9, no. 22, p. 4846, 2019.

[83] J. Xue and B. Su, "Significant remote sensing vegetation indices: A review of developments and applications," *Journal of sensors*, vol. 2017, 2017.

[84] J. Abdulridha, Y. Ampatzidis, S. C. Kakarla, and P. Roberts, "Detection of target spot and bacterial spot diseases in tomato using uav-based and benchtop-based hyperspectral imaging techniques," *Precision Agriculture*, vol. 21, no. 5, pp. 955–978, 2020.

[85] *Earth Observing System vegetation indices to drive digital agri solutions*, `https://eos.com/blog/vegetation-indices/`, Accessed: 2022-01-11.

[86] C. Evangelides and A. Nobajas, "Red-edge normalised difference vegetation index (ndvi705) from sentinel-2 imagery to assess post-fire regeneration," *Remote Sensing Applications: Society and Environment*, vol. 17, p. 100 283, 2020.

[87] J. Albetis, S. Duthoit, F. Guttler, A. Jacquin, M. Goulard, H. Poilvé, J.-B. Féret, and G. Dedieu, "Detection of flavescence dorée grapevine disease using unmanned aerial vehicle (uav) multispectral imagery," *Remote Sensing*, vol. 9, no. 4, p. 308, 2017.

[88] A. K. Chandel, L. R. Khot, and B. Sallato, "Apple powdery mildew infestation detection and mapping using high-resolution visible and multispectral aerial imaging technique," *Scientia Horticulturae*, vol. 287, p. 110 228, 2021.

[89] F.-M. Wang, J.-F. Huang, Y.-L. Tang, and X.-Z. Wang, "New vegetation index and its application in estimating leaf area index of rice," *Rice Science*, vol. 14, no. 3, pp. 195–203, 2007.

[90] D. Henderson, C. J. Williams, and J. S. Miller, "Forecasting late blight in potato crops of southern idaho using logistic regression analysis," *Plant disease*, vol. 91, no. 8, pp. 951–956, 2007.

[91] E. Lasso, D. C. Corrales, J. Avelino, E. de Melo Virginio Filho, and J. C. Corrales, "Discovering weather periods and crop properties favorable for coffee rust incidence from feature selection approaches," *Computers and Electronics in Agriculture*, vol. 176, p. 105 640, 2020.

[92] D. Diepeveen, L. Armstrong, and Y. Vagh, "Identifying key crop performance traits using data mining," 2008.

[93] N. N. Patil and M. A. M. Saiyyad, "Machine learning technique for crop recommendation in agriculture sector," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 1, pp. 1359–1363, Oct. 2019, ISSN: 22498958. DOI: 10.35940/ijeat.A1171.109119.

[94] S. Kim, M. Lee, and C. Shin, "Iot-based strawberry disease prediction system for smart farming," *Sensors*, vol. 18, no. 11, p. 4051, 2018.

[95] X. Yin, M. J. Kropff, G. McLaren, and R. M. Visperas, "A nonlinear model for crop development as a function of temperature," *Agricultural and Forest Meteorology*, vol. 77, no. 1-2, pp. 1–16, 1995.

[96] R. R. Patil and S. Kumar, "Predicting rice diseases across diverse agrometeorological conditions using an artificial intelligence approach," *PeerJ Computer Science*, vol. 7, e687, 2021.

[97] P. Sharma, B. Singh, and R. Singh, "Prediction of potato late blight disease based upon weather parameters using artificial neural network approach," in *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, IEEE, 2018, pp. 1–13.

[98] S. S. Dahikar and S. V. Rode, "Agricultural crop yield prediction using artificial neural network approach," *International journal of innovative research in electrical, electronics, instrumentation and control engineering*, vol. 2, no. 1, pp. 683–686, 2014.

[99] Z. Liu, L. Meng, W. Zhao, and F. Yu, "Application of ann in food safety early warning," in *2010 2nd International Conference on Future Computer and Communication*, IEEE, vol. 3, 2010, pp. V3–677.

[100] O. Trenz, J. Št'astnỳ, and V. Konečnỳ, "Agricultural data prediction by means of neural network," *Agricultural Economics*, vol. 57, no. 7, pp. 356–361, 2011.

[101] T. Ranjeet and L. Armstrong, "An artificial neural network for predicting crops yield in nepal," 2014.

[102] R. Kaundal, A. S. Kapoor, and G. P. Raghava, "Machine learning techniques in disease forecasting: A case study on rice blast prediction," *BMC bioinformatics*, vol. 7, no. 1, pp. 1–16, 2006.

[103] Y. Gu, S. Yoo, C. Park, Y. Kim, S. Park, J. Kim, and J. Lim, "Blite-svr: New forecasting model for late blight on potato using support-vector regression," *Computers and Electronics in Agriculture*, vol. 130, pp. 169–176, 2016.

[104] A. Murynin, K. Gorokhovskiy, and V. Ignatie, "Efficiency of crop yield forecasting depending on the moment of prediction based on large remote sensing data set," in *Proceedings of the International Conference on Data Science (ICDATA)*, The Steering Committee of The World Congress in Computer Science, Computer …, 2013, p. 1.

[105] A. K. Prasad, L. Chai, R. P. Singh, and M. Kafatos, "Crop yield estimation model for iowa using remote sensing and surface parameters," *International Journal of Applied earth observation and geoinformation*, vol. 8, no. 1, pp. 26–33, 2006.

[106] L. Vikas, V. Dhaka, *et al.*, "Wheat yield prediction using artificial neural network and crop prediction techniques (a survey).," *International Journal for Research in Applied Science and Engineering Technology*, vol. 2, no. 9, pp. 330–341, 2014.

[107] G. Fenu and F. M. Malloci, "Forecasting plant and crop disease: An explorative study on current algorithms," *Big Data and Cognitive Computing*, vol. 5, no. 1, p. 2, 2021.

[108] X. Yang and T. Guo, "Machine learning in plant disease research," *March*, vol. 31, p. 1, 2017.

[109] R. Arora, S. Sharma, and B. Singh, "Late blight disease of potato and its management," *Potato J*, vol. 41, no. 1, pp. 16–40, 2014.

[110] *Django*, https://www.djangoproject.com/, Accessed: 2022-01-10.

[111] *Sqlite*, https://www.sqlite.org/index.html, Accessed: 2021-12-09.

[112] *Pyforms*, https://pyforms.readthedocs.io/en/v3.0/, Accessed: 2022-01-10.

[113] *Plantvillage*, https://plantvillage.psu.edu/, Accessed: 2022-01-11.

[114] D. Dlužnevskij, P. Stefanovic, and S. Ramanauskaite, "Investigation of yolov5 efficiency in iphone supported systems," *Baltic Journal of Modern Computing*, vol. 9, no. 3, pp. 333–344, 2021.

[115] R. Padilla, S. L. Netto, and E. A. Da Silva, "A survey on performance metrics for object-detection algorithms," in *2020 international conference on systems, signals and image processing (IWSSIP)*, IEEE, 2020, pp. 237–242.

[116] *Imagenet*, https://www.image-net.org/index.php, Accessed: 2022-01-11.

[117] *Keras | keras tuner*, https://keras.io/keras_tuner/, Accessed: 2022-01-11.

[118] *Tensorflow*, https://www.tensorflow.org/, Accessed: 2022-01-11.

[119] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, *Tensorflow: Large-scale machine learning on heterogeneous distributed systems*, 2015. [Online]. Available: http://download.tensorflow.org/paper/whitepaper2015.pdf.

[120] V. K. Shrivastava, M. K. Pradhan, and M. P. Thakur, "Application of pre-trained deep convolutional neural networks for rice plant disease classification," in *2021 international conference on artificial intelligence and smart systems (ICAIS)*, IEEE, 2021, pp. 1023–1030.

[121] S. Sharma, S. Sharma, and A. Athaiya, "Activation functions in neural networks," *towards data science*, vol. 6, no. 12, pp. 310–316, 2017.

[122] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[123] *Mdpi agriculture*, https://www.mdpi.com/journal/agriculture, Accessed: 2022-01-11.

[124] T. Domingues, T. Brandão, and J. C. Ferreira, "Machine learning for detection and prediction of crop diseases and pests: A comprehensive survey," *Agriculture*, vol. 12, no. 9, 2022, ISSN: 2077-0472. DOI: 10.3390/agriculture12091350. [Online]. Available: https://www.mdpi.com/2077-0472/12/9/1350.

APPENDIX A

Web application interfaces

Figure 22 shows an example of a record in the web application of two diseases (*test1* and *test2*) that appeared in the crop field *Test_Crop_1* along with the marking on the map of the zone where the occurrence was found. As for Figure 23, it shows the layout and fields at the moment of importing ground level images and Figure 24 shows the interface where the number of this type of imported images over time can be analysed. Regarding Figure 25, it can be observed the interface when it is intended to analyse the meteorological data over time and the respective filtering options.

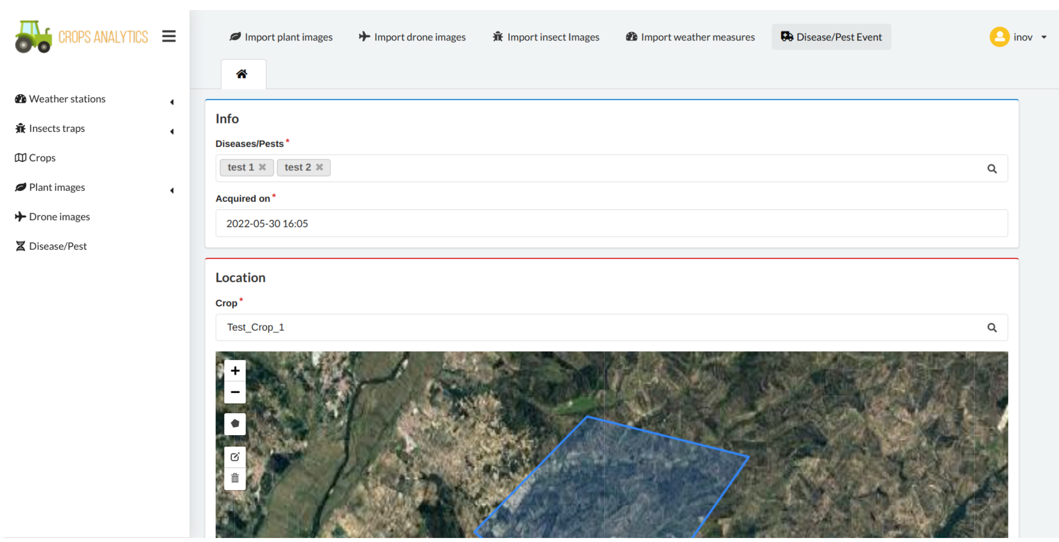


Figure 22. *Web application interface in the case of disease event.*

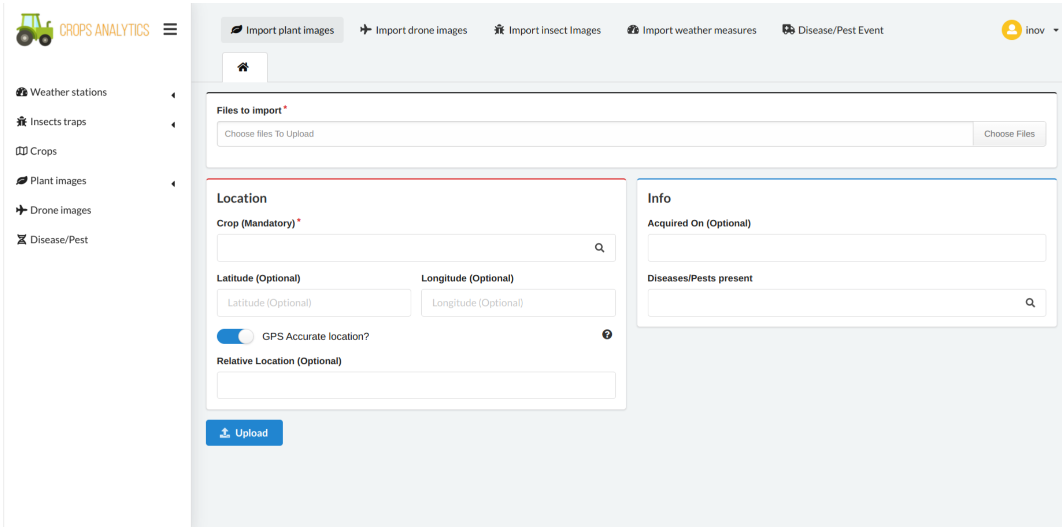


Figure 23. Web application interface in the case of plant image import.

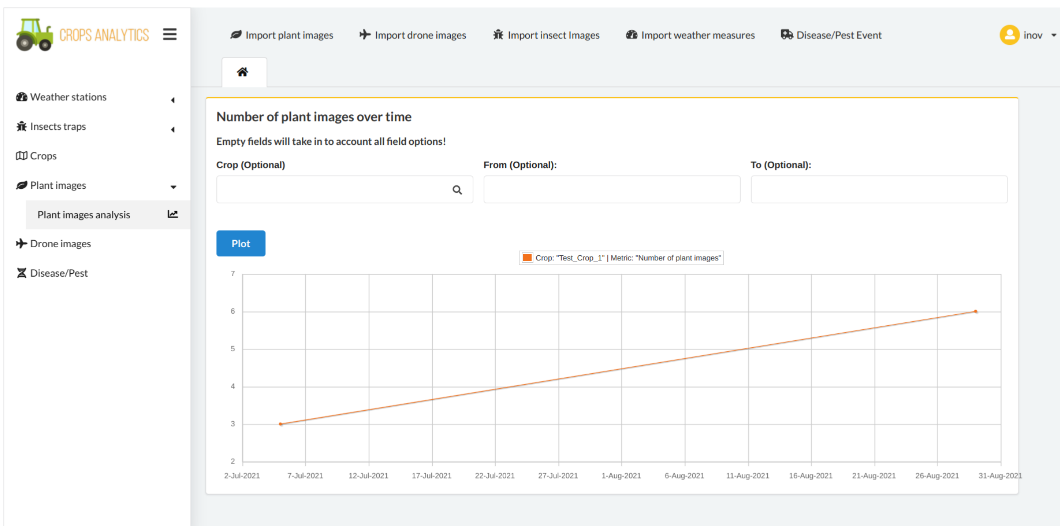


Figure 24. Web application interface in the case of plant analysis.

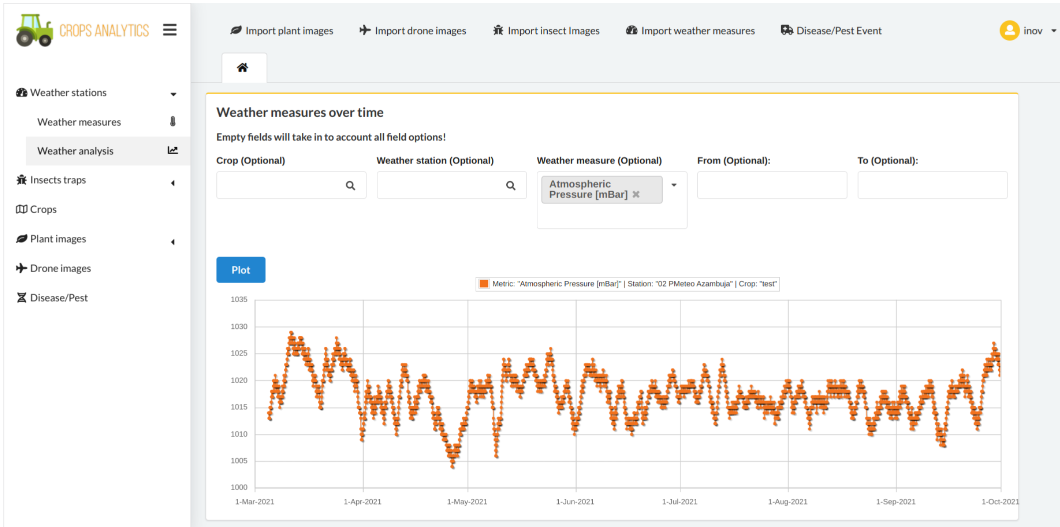


Figure 25*. Web application interface in the case of weather analysis.*

APPENDIX B

Results of disease classification on tomato leaves images

Figure 26 shows graphs that correspond to the data found in Tables 15 and 16. From the graph it is possible to verify the poor performance of these two models as the recall values are much lower than those of MobilteNetV2 and ResNet152V2.

	bacterial_spot ->	early_blight ->	late_blight ->	leaf_mold ->	septoria_leaf_spot ->	spider mites two-spotted spider mite ->	target spot ->	yellow leaf curl virus ->	mosais virus ->	healthy tomato->
bacterial_spot ->	97%	0%	1%	0%	2%	0%	0%	0%	0%	0%
early_blight ->	6%	70%	7%	1%	10%	3%	2%	0%	1%	0%
late_blight ->	0%	5%	89%	1%	5%	0%	0%	0%	0%	0%
leaf_mold ->	0%	2%	5%	75%	6%	5%	0%	3%	4%	0%
septoria_leaf_spot ->	1%	0%	6%	2%	88%	1%	1%	0%	1%	0%
spider mites two-spotted spider mite ->	0%	0%	2%	0%	0%	94%	0%	0%	4%	0%
target spot ->	1%	4%	4%	1%	9%	22%	55%	0%	3%	1%
yellow leaf curl virus ->	2%	0%	0%	0%	1%	1%	0%	92%	3%	1%
mosais virus ->	0%	0%	0%	0%	1%	1%	0%	0%	98%	0%
healthy tomato ->	0%	0%	0%	0%	2%	0%	0%	0%	1%	97%

***(a)** Confusion Matrix of VGG16.*

	bacterial_spot ->	early_blight ->	late_blight ->	leaf_mold ->	septoria_leaf_spot ->	spider mites two-spotted spider mite ->	target spot ->	yellow leaf curl virus ->	mosais virus ->	healthy tomato->
bacterial_spot ->	81%	1%	1%	2%	10%	1%	1%	1%	1%	1%
early_blight ->	4%	60%	16%	2%	6%	4%	6%	0%	2%	0%
late_blight ->	0%	1%	95%	0%	1%	0%	2%	1%	0%	0%
leaf_mold ->	0%	1%	9%	76%	2%	6%	0%	0%	5%	1%
septoria_leaf_spot ->	4%	4%	4%	3%	79%	1%	3%	0%	2%	0%
spider mites two-spotted spider mite ->	1%	0%	3%	0%	4%	82%	4%	0%	4%	2%
target spot ->	3%	1%	12%	2%	1%	12%	60%	0%	1%	8%
yellow leaf curl virus ->	3%	0%	1%	2%	1%	1%	0%	91%	1%	0%
mosais virus ->	0%	1%	4%	5%	3%	1%	0%	1%	85%	0%
healthy tomato ->	0%	0%	1%	0%	0%	3%	6%	0%	2%	88%

***(b)** Confusion Matrix of InceptionV3.*

***Figure 26**. Confusion Matrix's of the worst leaf classification models*

	Precision	Recall	F1-Score
bacterial spot	0.91	0.97	0.94
early blight	0.86	0.70	0.77
late blight	0.78	0.89	0.83
leaf mold	0.94	0.75	0.83
septoria leaf spot	0.71	0.88	0.79
spider mites two-spotted spider_mite	0.74	0.94	0.83
target spot	0.95	0.55	0.70
yellow leaf curl virus	0.97	0.92	0.94
mosaic virus	0.85	0.98	0.91
healthy tomato	0.98	0.97	0.97
Accuracy	-	-	0.85
Macro average	0.87	0.86	0.85
Weighted average	0.87	0.85	0.85

Table 15. *VGG16 additional metrics.*

	Precision	Recall	F1-Score
bacterial spot	0.84	0.81	0.83
early blight	0.87	0.60	0.71
late blight	0.65	0.95	0.77
leaf mold	0.83	0.76	0.79
septoria leaf spot	0.74	0.79	0.76
spider mites two-spotted spider_mite	0.74	0.82	0.78
target spot	0.73	0.60	0.66
yellow leaf curl virus	0.97	0.91	0.94
mosaic virus	0.83	0.85	0.84
healthy tomato	0.88	0.88	0.88
Accuracy	-	-	0.80
Macro average	0.81	0.80	0.80
Weighted average	0.81	0.80	0.80

Table 16. *InceptionV3 additional metrics.*

APPENDIX C

Articles

Article published to MDPI Agriculture:

- **Manuscript ID** : agriculture-1860035;
- **Digital Object Identifier** : 10.3390/agriculture12091350;
- **Type of manuscript** : Review;
- **Title** : Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey.

agriculture

Review

Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey

Tiago Domingues [1,2,*], Tomás Brandão [1] and João C. Ferreira [1,2]

1 Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR-IUL, 1649-026 Lisboa, Portugal
2 Inov Inesc Inovação, Instituto de Novas Tecnologias, 1000-029 Lisbon, Portugal
* Correspondence: tards@iscte-iul.pt

Abstract: Considering the population growth rate of recent years, a doubling of the current worldwide crop productivity is expected to be needed by 2050. Pests and diseases are a major obstacle to achieving this productivity outcome. Therefore, it is very important to develop efficient methods for the automatic detection, identification, and prediction of pests and diseases in agricultural crops. To perform such automation, Machine Learning (ML) techniques can be used to derive knowledge and relationships from the data that is being worked on. This paper presents a literature review on ML techniques used in the agricultural sector, focusing on the tasks of classification, detection, and prediction of diseases and pests, with an emphasis on tomato crops. This survey aims to contribute to the development of smart farming and precision agriculture by promoting the development of techniques that will allow farmers to decrease the use of pesticides and chemicals while preserving and improving their crop quality and production.

Keywords: plant diseases and pests; classification; detection; forecasting; precision farming; machine learning; smart farming

Citation: Domingues, T.; Brandão, T.; Ferreira, J.C. Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey. *Agriculture* **2022**, *12*, 1350. https://doi.org/10.3390/agriculture12091350

Academic Editors: Xiuliang Jin, Hao Yang, Zhenhai Li, Changping Huang and Dameng Yin

Received: 26 July 2022
Accepted: 23 August 2022
Published: 1 September 2022

1. Introduction

Due to extremely high infant mortality, the human population of the planet increased slowly until the year 1700. The first billion was reached in ca. 1800, followed by the second billion in 1928, the third billion in 1960. In 2017, the world's population reached its seventh billion. The fast population growth over recent decades is mainly due to better medical care. According to predictions from the United Nations, the world's population is expected to reach 9.7 billion in 2050, and 10.9 billion in 2100 [1].

Rapid population growth over recent decades has resulted in an increased demand for agricultural goods, which in turn has lead to a large expansion of cultivation [2]. To meet rising population demands for food, bio-fuels, and animal products, crop yield production must double its output by 2050. In order to achieve this goal, key crop yields must improve by 2.4% each year, but they are now only increasing by roughly 1.3% per year [3]. However, fulfilling this condition will have negative consequences for the ecosystem, including the loss of biodiversity and increased greenhouse gas emissions. Traditional agricultural production is not sustainable from an economic or environmental standpoint; hence, it is critical to optimize the use of resources such as water and soil to enable high yield crops [2].

Moreover, crop output is continually threatened by diseases and insect pests. It is estimated that between 20% to 40% of yearly crop production is lost due to plant diseases and insect assaults across the world, costing the global economy \$220 billion and \$70 billion, respectively. The amount of these losses varies across the globe and often occurs due to transboundary plant pests and diseases. For instance, the spread of crop pests and pathogens between 1950 and 2000 was greater in North America when compared with other world regions [4].

Pest damage and development are affected by the rise in global temperature brought by climate change. When the temperature rises, the metabolic rate of insects increases, driving them to consume more food and inflict more damage. Growth rates of several insect species are also affected by temperature. For each degree of average global warming of the earth's surface, worldwide agricultural losses due to insect pests are expected to increase by 10% to 25% [5].

Pesticides and chemical treatments have long been used by farmers to keep pests away. The use of pesticides for crop protection is on the rise [6], with negative consequences for human health and increased environmental damage to soil and groundwater. On the other hand, this also increases the risk of pests developing pesticide resistances [5].

The traditional method of detecting and identifying plant diseases involves naked eye observation by experts. This takes time and talent, and is not a practical solution for monitoring large farms. Therefore, to overcome the limitations of manual detection, automated methods for crop monitoring and forecasting are required [7]. A system capable of performing such tasks can play an important role in avoiding the excessive use of pesticides and chemicals, reducing both the damage caused to the environment and the production costs associated with the use of pesticides and chemicals [7].

The growing availability of big data analysis methods has the potential to spur even more research and development in smart farming. Besides promoting higher yield crops in a more sustainable manner, it also aims to contribute to event forecasting, detection of diseases, and management of water and soil. Big data is coming to the agriculture domain by collecting data from meteorological stations, remote sensors, historical data, and publicly available data-sets [8].

ML approaches have been successfully utilized in a variety of areas, including illness detection from medical images [9], image classification on large data-sets [10], self-driving automobiles [11], and academic research fields such as physics [12].

ML-based applications for agriculture are still young, but are already showing promise. For instance, disease classification from images can be done using popular Convolutional Neural Network (CNN) architectures for different plants with different diseases [13]; relationships between weather data and pest occurrence can be retrieved using Long Short Term Memory (LSTM) networks for forecasting future pest attacks [14]; insect detection on leaves can be performed using object segmentation and deep learning techniques [15].

Commercial tools and services for smart farming that make extensive use of machine learning are currently available to farmers. A few examples are as follows. *Plantix*, created by the German startup *Progressive Environmental and Agricultural Technologies* (PETA), is an android-based farming assistant tool that provides crop health information, helping with identification of plant diseases using computer vision and deep learning techniques [16]. Other examples of similar applications are *Agrio* [17] and *CropDiagnosis* [18]. *Gamaya* is a startup company based on Switzerland that offers a wide variety of smart farming services services based on the analysis of images images acquired by drones connected to IoT systems [19]. The asian *iFarmer* [20] is another company that offers IoT-based soil analysis and satellite imaging-based crop monitoring solutions. *See & Spray*, developed by California-based *Blue River Technology*, is a large tow-behind herbicide sprayer, that uses computer vision and deep learning-based algorithms to automatically locate and identify weeds (in real time), applying herbicides to the specific locations found rather than to the entire field [21].

Some related surveys can be found in the literature, but most of them are focused on traditional ML techniques: in [22], a comparison of ML algorithms for predicting the yield of soybean crops is presented; in [23] research papers from the last ten years for predicting the start of disease at an early or presymptomatic stage are analysed and categorised; in [24], the possibility of using different ML techniques in agriculture are discussed, but most of the present work is about statistical forecasting methods from weather data for predicting wheat yield. Since the mentioned surveys do not simultaneously cover forecasting, detection, and classification of diseases and pests, and do not fully explore recent deep learning-based techniques, the review performed in this paper aims to fill the gaps on these subjects.

The literature review presented in this paper also aims to provide guidance on the development of such ML-based tools, in order to provide farmers with data-driven decision making assistance systems. In this way, farmers can be assisted with lowering the need for pesticide application and the harm that comes with it, while also preserving and enhancing crop quality and yield. This contributes to the continued availability of food to meet global population demands while doing less damage to the planet.

The application of ML-based techniques has promoted the emergence of projects that have enriched the development and the evolution of smart farming [25]. With this in mind, this article also contributes to the progression, development, and success of such projects.

2. Literature Review

Data gathering, data pre-processing (i.e., data preparation that includes feature extraction), and ML classification models are the three basic steps of ML applications, represented in Figure 1. The following sections present and discuss different approaches used in these three stages.

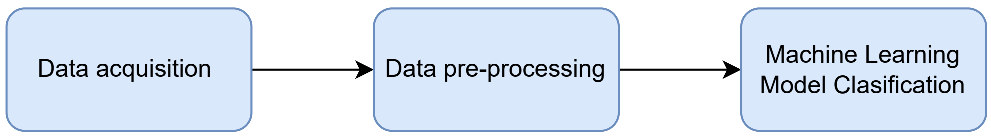


Figure 1. Simplification of the ML pipeline.

2.1. Data Acquisition

Data acquisition is the process of gathering data from various sources systems [26]. Previous studies gather their data various sources to be used for ML techniques. Some of them produce their own images by taking pictures of plants in greenhouses, such as in the studies from Gutierrez et al. [15] and Raza et al. [27]. However, image data acquisition using manual processes, as done by many, generally results in small image data-sets, which can compromise the development of effective ML-based models. Weather data collection is also proposed in the literature using for instance sensors in greenhouses, as done by Rustia and Lin [28]. Meteorological data can also be obtained from weather stations of regional areas, which typically store records for a longer period of time [14,29].

Images can be collected using search engines on their own [30,31]. This approach can get a large number of images, but ground truth must be checked by domain experts, and data cleaning is frequently used to filter out images that do not meet the requirements.

Remote sensing images from satellites and drones have the advantage of being able to retrieve image data for large agricultural areas. Remote sensing data from satellites typically consists of multi-temporal and hyper-spectral imagery data, which can be used to assess the development of the crops. This task can be performed by monitoring the evolution of vegetation indices [32], which provide important information about the development status of the crop fields. Spectral imagery can be used for computing different vegetation indexes, such as those proposed in [33–39], which are robust to variations on the sun illumination [37], an important advantage when compared to visible light spectrum imagery.

Images retrieved from drones can also be used, but have additional needs: to define the path of the device; to coordinate the drone position with the camera for image acquisition; and to correct geometric distortions on each acquired image in order to merge the different acquired images in order to reconstruct a larger image of the whole field [40].

Therefore, it can be stated that data consists of different modalities and variables. With ML-based and data analysis techniques it can be possible to understand their interaction and how they relate to a studied outcome. In the context of the cultivation fields, the questions are usually: which disease is affecting crops? What pest is causing damage? What is the relation between weather data and disease and pest occurrence? The most important variables related to the appearance of plant diseases and pests are reviewed in Section 2.1.1.

Freely available data-sets can also be utilised for the development of ML-based applications. This enables researchers to directly compare the performance of different ML techniques and approaches. In Section 2.1.2, a brief summary of the main available data-sets is presented.

The data conditions a significant impact on the performance of ML models. Section 2.1.3 addresses this issue. Data-sets should be representative and include enough records for the model to perform an effective generalization.

2.1.1. Variables Influencing Crop Diseases and Pests

It's crucial to be able to predict the arrival of diseases and pests in crops, in addition to correctly detecting and identifying them. Real-time meteorological data obtained by unmanned observation planes, as well as long-term data analysis from weather stations, have been used to create models capable of anticipating disease occurrence. In [41], the General Infection Model, proposed in [42], was used for assessing the prediction capabilities of the system. It was found that, if integrated systems such as this are implemented and various input data-sets essential for interrelationship analyses are collected, accurate plant disease prediction systems can be constructed.

When it comes to forecasting occurrences, it's crucial to know which variables will have an influence on what is being forecast. In the work by Henderson et al. [43] this was done by discovering which weather variables influence the forecast. On the other hand, Lasso et al. [44] determined the time period window for each weather variable and crop-related feature that is the most significant for the appearance of coffee leaf rust disease in coffee crops.

In [45], Small et al. used weather data, information on potato and tomato crops resistance to late blight (from published literature and field experiments), and management strategies, to create a web-based decision support system that allows the dynamic prediction of disease outbreaks, with an emphasis on the late blight disease on tomato and potato crops.

The work proposed by Ghaffari et al. [46] addresses the very early detection of diseases in tomato crops using atmospheric data and volatile organic compounds. Plants produce a wide spectrum of volatile organic compounds in reaction to physical and biotic stress, as well as infection [47]. In [46], the diseases under study were the powdery mildew and spider mites.

A model developed by Diepeveen et al. in [48] can be used in agriculture to understand the influence of location and temperature on crops. In addition, elements such as soil, humidity, rainfall, and moisture were found to have an influence on crop yield [49].

Plant diseases and pest development are greatly influenced by weather and environment conditions [50]. Humidity is a favorable condition for the development of fungus diseases. The humidity can be caused by the weather or by poor watering practices that cause a high wetness among the leaves, making tomatoes more susceptible to diseases, e.g., leaf mold or bacterial spot [51].

In addition, temperature is a primary driver of insect development, affecting their metabolic rate and population growth [5].

Plants absorb part of the radiation coming from the sun and reflect the rest. Depending on the health of the plant, the amount of radiation absorbed and reflected differs. This difference can be used to distinguish between healthy and diseased plants and to assess the severity of the damage [52]. The concept is illustrated in Figure 2.

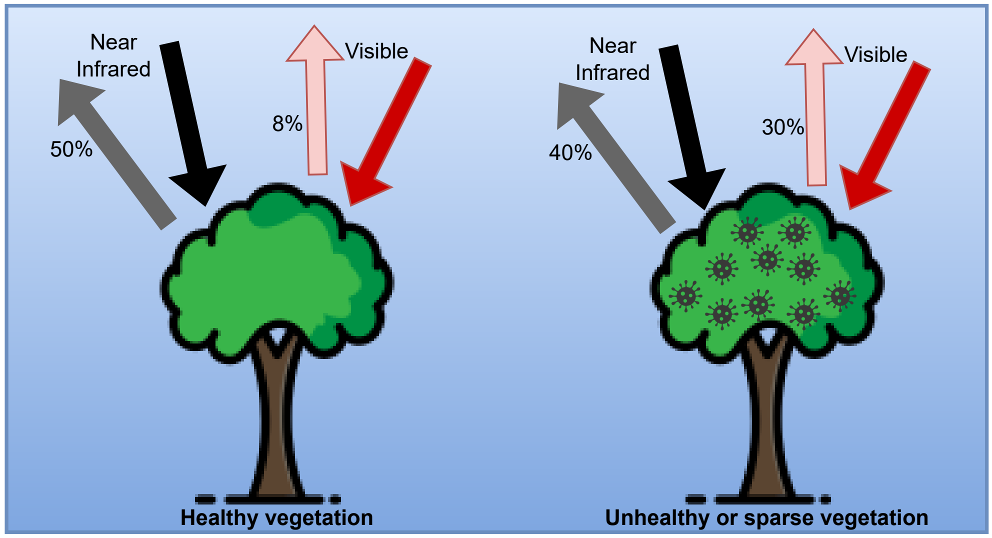


Figure 2. Absorbed and reflected radiation for plant's health estimation (adapted from [53]).

Temperature

Insects are ectothermic, meaning that they cannot regulate their internal temperature and have to rely on environmental heat sources. Temperature affects the population growth and metabolic rates of insects [5]. Thus, the duration of an insect's life cycle is highly influenced by the number of days where the temperature is suitable for its development. Two temperature thresholds can be define: an upper threshold, in which insect development slows down or stops and a lower one where there is no insect growth. These thresholds vary according to the specific insect species.

Degree day is a concept concerning the accumulation of heat by insects [54]. One degree day is a period of 24 h in which the temperature was one degree above a given baseline. Different models for determining the number of degree days associated to common pest species were proposed in [55]. For instance, tomato crops are susceptible to the greenhouse white fly (*Trialeurodes vaporariorum*), whose number of degree days from egg to adult is 380 DGG [56]. Depending on the temperature of the environment, this development time can be longer or shorter.

Biofix date is the date to start accumulating degree days associated with a given insect species [57]. This date can be determined by noticing specific insect species on traps or by detecting eggs on plant leaves. From this date, degree days can be used to estimate the period at which insects are reaching a given development stage suitable for pesticide application. Temperature and weather forecasts are nowadays sufficiently accurate to enable the estimation for the time required for an insect to reach a given development status [58].

In the context of ML-based applications, related work focused on studying the impact of weather in pest insect development found a higher correlation between the number of pest catches and temperature, when compared with other factors [28,32].

Some diseases affect the transpiration rate of the plant and, consequently, its temperature [27]. Therefore, plant leaf temperature can be used for disease detection. ML models can achieve higher accuracy for disease identification when combining thermal images with visible light images. The benefits are more useful for early detection when the plant has not yet developed symptoms recognizable by the naked eye.

Humidity

Diseases affecting plants are often caused by fungus or bacterial pathogens. High relative humidity environments favor the development of these microorganisms. Thus, humidity has to be managed by good watering practices, while avoiding excessive leaf moisture and soil moisture [59].

Different studies using regression models and weather data demonstrate the influence of humidity on disease and pest development [14,29]. Thus, the collection of humidity records in greenhouses using sensors can be helpful for disease forecasting.

Leaf Reflectance

Plants absorb solar radiation between 400 to 700 nm (photosynthetically active radiation) which corresponds approximately to the visible light region. For wavelengths greater than 700 nm (red) in the Near Infra-Red (NIR) region there is a sharp order-of-magnitude increase in leaf reflectance due to chlorophyll characteristics, a phenomenon known as *red edge* [60].

Diseased plants with damaged leaves have different leaf spectral reflectance compared to a healthy plant because of the different chlorophyll concentration and leaf tissue damage. Diseased plants end up absorbing less of the visible light and more of the NIR light. From this knowledge, disease detection can be done using leaf reflectance information [40,52,53]. In a study concerning late blight infection, a disease that tomatoes are also susceptible to, it was found that spectral differences in the visible region between healthy and diseased plants are small and more significant differences are noticeable in the NIR [40].

Various vegetation indices can be retrieved from remote sensing [33]. A common index is the Normalized Difference Vegetation Index (NDVI) (Figure 2) for assessing the degree of vegetation of an area by using leaf reflectance information. NDVI can be computed using satellite data or from modified cameras [40,53]. It was found that the combination of NDVI and temperature gives higher accuracy in predicting pests appearances than weather variables alone [32]. NDVI can also be used as input data for ML models to accurately evaluate disease severity.

Pest development varies depending on the development stage of the plants. NDVI can be used to monitor plant growth and establish relationships between the crop stage development and pest occurrence.

2.1.2. Agriculture Data-Sets

Many data-sets used in the context of agriculture include images of plant diseases or pests with the goal of classifying them. *PlantVillage*, *PlantDoc*, *IP102*, *Flavia* and, *MalayaKew Leaf* are some data-sets that are freely available. Here is a brief summary of each of these:

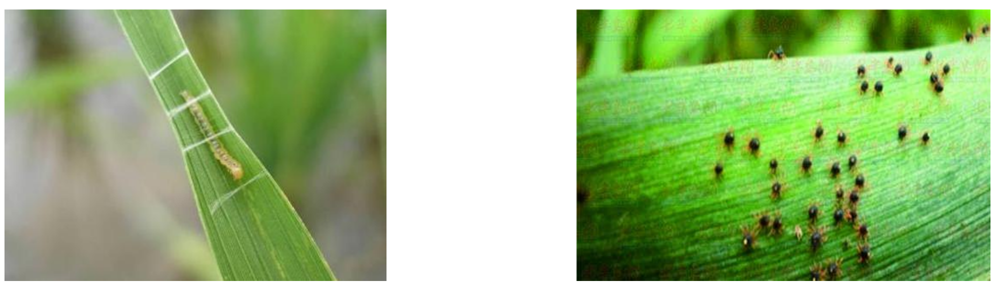
- *PlantVillage* [61]: popular data-set used for plant disease classification. Specifically for tomato, it contains 18,160 images representing leaves affected by bacterial spot, early blight, late blight, leaf mold, septoria leaf spot, spider mites, two-spotted spider mite, target spot and tomato yellow leaf curl virus. It also includes images of healthy leaves. Figure 3 depicts two sample images taken from this data-set.
- *IP102* [62]: data-set for pest classification with more than 75,000 images belonging to 102 categories. Part of the image set (19,000 images) also includes bounding box annotations. This is a very difficult data-set because of the variety of insects, their corresponding development stages (egg, larva, pupa, and adult) and image backgrounds. The data-set is also very imbalanced. Figure 4 presents two examples of images from this data-set.
- *PlantDoc* [63]: contains pictures representing tomato diseases which were acquired in the fields. Among the considered diseases are: tomato bacterial spot, tomato early blight, tomato late blight, tomato mold, tomato mosaic virus, tomato septoria leaf spot, tomato yellow virus and healthy tomatoes.
- *Flavia* [64]: contains photos of isolated plant leaves over a white background and in the absence of stems. This data-set covers 33 plant species.
- *MalayaKew Leaf* [65]: was gathered in England's Royal Botanic Gardens at Kew. It contains images of leaves from 44 different species. There are situations where leaves from

different species are very similar, presenting a greater challenge for the development of plant identification models.

Tomato Powdery Mildew Disease (TPMD) is a different type of data-set because it is related to meteorological data. It offers statistics on powdery mildew disease susceptibility depending on a variety of weather-related variables such as humidity, wind speed, temperature, global radiation, and leaf wetness [66].

(**a**) Tomato leaf affected by mosaic virus disease (**b**) Tomato leaf affected by late blight disease

Figure 3. Examples of tomato leaves affected by diseases taken from the *PlantVillage* data-set [61].



(**a**) Rice leaf roller (*Marasmia exigua*) (**b**) Winter grain mite (*Penthaleus major*)

Figure 4. Examples of insect images taken from the *IP102* data-set [62].

2.1.3. Field-Collected vs. Laboratory-Collected Data

ML models performance is influenced by the quality and type of input (image or other). Images acquired in a controlled laboratory environment and images acquired in the field can result in completely different processes and/or results. The difficulty for disease and pest classification is much higher for images acquired in the field than for images taken in a controlled environment.

Under a controlled laboratory environment, images typically contain a single leaf over a neutral artificial background [67]. The *PantVillage* data-set is an example of such situation [61]. It is possible to achieve great performance on these data-sets [13]. However, the creation of these types of data-sets is a time consuming and costly process.

When compared with images acquired in the laboratory, field images have much higher complexity, due to the presence of multiple leaves in the same image, presence of other plant parts, different shading, and lighting conditions, different ground textures, different backgrounds, etc. [63]. According to the studies in [63,68], training ML models using laboratory images provides poor outcomes when tested in the field, making them useless for the task. Training on field photographs and testing on laboratory photographs, on the other hand, produce reasonable outcomes [68]. The addition of field images in the training data has been shown to boost the results significantly, however testing on images from alternative data sources is advised [68]. *PlantDoc* demonstrates that cropping the leaves improves the accuracy of CNN architectures when dealing with in-field photos [63].

Table 1 shows the performance achieved on a few studies that analysed the impact of image acquisition conditions on the performance of disease classification models. In each table cell, "L" corresponds to lab images, "F" to field images and "L + F" for both types of images. In addition, the data-sets associated to the weights of the pre-trained models that were used for Transfer Learning (further explored in Section 2.3.4) are also shown.

Table 1. Performance comparison of field vs. laboratory data.

Study	Pretrained Weights	Training	Testing	Performance
[68]	-	L	F	33.0% acc.
		F	L	65.0% acc.
		L + F	L + F	99.0% acc.
[63]	ImageNet	L	F	15.0% acc.
	ImageNet + PlantVillage	F	F	30.0% acc.
	ImageNet + PlantVillage	F (cropped images)	F (cropped images)	70.0% acc.
[13]	ImageNet	L	L	99.0%+ acc.

2.2. Data Pre-Processing

Pre-processing data before feeding it to the model is common in most ML-based applications. Images are typically pre-processed using computer vision techniques to remove noise, to enhance the image contrast, to extract the regions of interest, to extract image features, etc. In general, image pre-processing steps usually lead to better model outcomes. The most common data pre-processing techniques are covered in the following sub-sections.

2.2.1. Noise Reduction

Different types of filters, such as Gaussian and median filters, are used to reduce noise to obtain smoother images. These filters have an effect of blurring and removing non relevant details of an image, at the expense of potentially losing relevant textures or edges [69].

Erosion and dilatation are two morphological image operations that can be applied to binary or grey-scaled images. Erosion removes islands and tiny items, leaving only larger objects. In other words, it shrinks the foreground objects. On the other hand, dilation increases the visibility of items and fills in tiny gaps, adding pixels to the boundaries of objects in an image [70]. These operations reduce details and enhance regions of interest. These methods are helpful, for instance, for pest detection against a neutral background, such as images of traps with captured insects [28,71].

Images are usually stored in the RGB format, which is an additive color model of red, green, and blue components. Due to the high correlation between these color components, it is usually not suitable to perform color segmentation in the RGB color space. Therefore it is important to bear in mind that there are others color spaces such as HSV or L*a*b*. In HSV the color components are: hue (pure color), saturation (shade or amount of grey), and value (brightness). In the L*a*b* color space, L* is the luminance (brightness), a* is the value along the red-green axis, and b* is the value along the blue-yellow axis. In these color spaces, the brightness of a color is decoupled from its chromaticity, allowing the images to be processed with different lighting conditions [69]. This is significant in the context of agricultural images acquired in the fields, since they can have been shot under various lighting circumstances or at different times of the day.

Histogram equalization is a technique for adjusting contrast. In low contrast images, the range of intensity values is smaller than in high contrast images. Equalization of the histogram spreads out the intensity levels throughout values in a wider range. Contrast enhancement is not directly applied in the RGB color space, because it applies to brightness values. Thus, images have to be converted either to grey-scale or to a color space that contains a brightness component, such as the HSV or L*a*b* color spaces [69].

2.2.2. Image Segmentation

Image segmentation is the process of grouping pixels into regions of interest. In the context of crop disease identification, these regions of interest can be, for instance, diseased areas on the plant leaves, for assessing the severity of the infection by the amount of the infected area, or for background removal, since the removal of the background allows highlighting of the regions of interest for further analysis. An example of background removal is shown in Figure 5.

(**a**) Strawberry leaf scorch (original) (**b**) Strawberry leaf scorch (segmented)

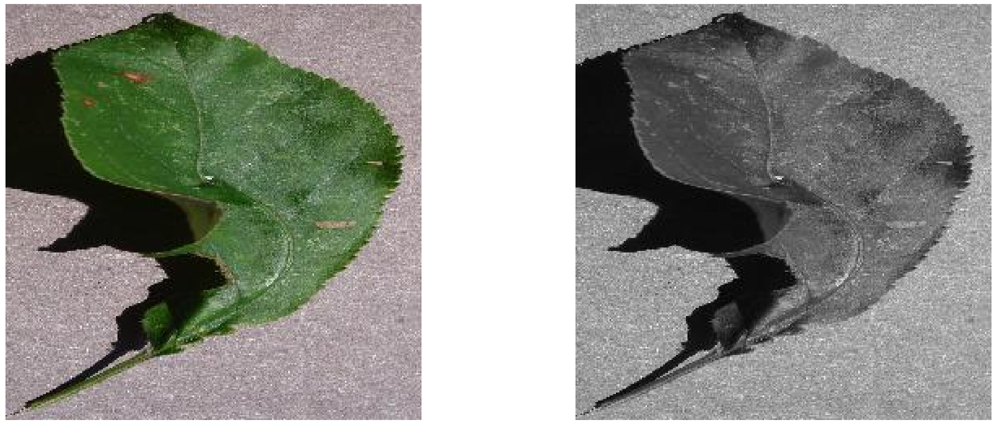
Figure 5. Example of background removal from the *PlantVillage* data-set [51].

Blob detection is a computer vision technique for getting regions of pixels that share common properties. The properties of these regions, such as color and brightness, differ greatly compared to their surroundings. This technique can be used, for instance, to detect and count insects in images [28,71].

The k-means clustering algorithm is a popular unsupervised ML algorithm that can be used for image segmentation. Pixels are grouped into clusters which have pixels with similar color and brightness values. This technique is helpful, for instance, to detect damaged regions on leaves [31,72]. Fuzzy c-means is a soft clustering technique where a pixel can be assigned to more than one group. This method was used by Sekulska-Nalewajko and Goclawski [73] and Zhou et al. [74] for plant disease classification.

Region growing is a region-based image segmentation technique used by Pang et al. in [75] to accurately define the image regions corresponding to the plant leaf parts affected by disease.

Intensity thresholding is a straightforward and simplified approach for image segmentation. According to the pixel value, that pixel is classified into a group (e.g., healthy or diseased). When using this technique, images are frequently converted to grey-scale first and then thresholded using a grey intensity value [76]. Figure 6 shows an example of an image converted to grey-scale.



(**a**) Black rot apple (original) (**b**) Black rot apple (grey)

Figure 6. Example of an image converted to grey-scale from the *PlantVillage* data-set [51].

2.2.3. Feature Extraction

Feature extraction is a common step in the pre-processing of images for shallow ML models. Common image feature extraction algorithms include the Histogram of oriented Gradient (HoG), Speeded Up Robust Features (SURF) and Scale Invariant Feature Transform (SIFT) [62,77]. Different feature extractors obtain different features that can be more or less suitable for the specific problem at hand. HoG focuses on the structure and shape of the image objects, by detecting edges on images oriented according to different directions. The distribution of gradients according to these directions are used as features. SIFT finds scale and rotation invariant local features through the whole image, obtaining a set of image locations referred to as the image's key-points. SURF is conceptually similar to SIFT, with the advantage of being much faster, which can be relevant for the implementation of real-time applications.

The distribution of image colors is represented by a color histogram. Since most diseases have symptoms that impact the color of the leaves, the histogram can also be used for distinguishing between healthy and unhealthy plants [77].

Some computer vision algorithms for feature extraction demand that pictures are converted to grey-scale, such as Haralick texture [78] or edge detection algorithms [79], etc. Haralick texture features are computed from a Grey Level Co-occurrence Matrix (GLCM), a matrix that counts the co-occurrence of neighboring grey-levels in the image. The GLCM acts as a counter for every combination of grey-level pairs in the image. Diseased and healthy leaves have different textures since a diseased leaf has a more irregular surface and a healthy leaf has a smoother one. These features allow differentiation of a healthy leaf from a diseased one.

Local Binary Pattern (LBP) [80] is another technique used for image texture features extraction robust to variations on lighting conditions. The LBP technique was used by Tan et al. in [81] for the extraction of information about diseases on tomato leaves.

Multi-spectral image data-sets can be exploited to create new data and improve the performance of models. For instance, in [40], originally, there were NIR pictures of the fields and from this data the authors created new images from spectral differences (between green and blue bands, and between NIR and green bands), band ratios and dimension reduction using principal component analysis. The authors also assess which type of data achieves best performance on the models.

2.2.4. Cropping and Resizing Images

Cropping and resizing images is used for decreasing the input image dimensions, to allow greater processing speed or to fit hardware requirements. It can also be used for creating more data to train the models, for example, from a low number of high resolution pictures, a much higher number of low resolution images can be retrieved [40].

2.2.5. Pre-Processing in Tabular Data

Tabular data consisting of weather records was commonly found in the literature analysed in the scope of this paper. When gathering data records with varying dates and locations, these records can be integrated in two ways: cross-year, where models are validated over the years at the same location, and cross-location, where models are validated across the various locations for the same year. The average coefficient of determination (r^2) was found to be higher for cross-year models for all ML algorithms tested [29].

Common procedures in pre-processing are scaling/standardization of data and missing values processing [14]. Most algorithms require that there are no missing values in data and others, such as neural networks, can benefit from the normalization of feature values to improve training and reduce the effects of vanishing gradients [29].

Down sampling is a useful way to process data when there is a high number of records. In [52], measurements of leaf reflectance were done, from 760 to 2500 nm with a 1 nm interval. The 1740 wavelengths measurements were compressed into 174, and afterwards 10 wavelengths were selected using the stepwise method. From the regression analysis, results showed a coefficient of determination $r^2 = 0.94$ for these wavelengths and leaf severity. Experiments showed that fewer than those 10 wavelengths would worsen performance.

2.2.6. Pre-Processing in Deep Learning

Deep learning pre-processing does not focus on feature extraction since one of the most essential and beneficial properties of deep learning is its ability to generate features autonomously. For this reason, pre-processing is focused mainly on creating more images through data augmentation and resizing the input images to fit the models input parameters.

Some studies have compared the manual selection of features with deep learning. When it comes to categorizing insects in the field, manually selected features were not able to capture all of the relevant information about insect infestations or to handle the noise of real-world photos. Manually selected features were also not able to capture subtle differences between different insect species that share similar appearance [62]. For insect detection, deep learning techniques achieved higher accuracy and took less time to process since they efficiently select regions of interest [15]. In the work done by Brahimi et al. in [67], tomato disease classification using deep learning achieves higher accuracy, with values above 98%, but the accuracy of models using feature extractors is not very far behind, reaching values above 94%.

When comparing the use of original color pictures with images converted to grey-scale or background segmentation, deep learning models performed better in the original color pictures [13]. These findings are also confirmed in [82], where the performance of color vs. grey-scale pictures is compared. This supports the idea of deep learning not requiring extensive pre-processing of images. Nevertheless cropping images achieve better performance on field images classification, by increasing the region of interest and reducing the varying background [63].

Data augmentation is a process to artificially expand and increase the diversity of the training data-set. This process benefits the performance of the models, by introducing variability in the data and allowing a better generalization of the domain [83]. Some common transformations are rotation, cropping, scaling, and flipping.

Data cleaning is the process of assessing the quality of the data and to either modify or delete it. It is usually applied in studies that retrieve their data-set images from search engines in an automatic way, removing pictures that do not correspond to the intended labels or that do not comply with minimum resolution requirements [30,62].

Image resizing is usually performed to fit the input parameters of the models. Studies have compared the performance of the models with different input image sizes, and concluded that with larger images the models achieve higher accuracy but require more time for each training epoch [68] and more powerful hardware [71].

Table 2 shows the pre-processing techniques applied to deep learning classification models analysed in the scope of this review. The 'type' column shows the data pre-processing technique used and the 'info' column contains additional details about it.

Table 2. Pre-processing when deep learning techniques were used.

Study	Type	Info
[13]	Greyscale	-
	Background Segmentation	Masks
	Resize	256 × 256
[30]	Data Augmentation	Affine, perspective, rotation
	Data Cleaning	-
	Resize	256 × 256
[71]	Resize	52 × 52, 112 × 112, 224 × 224
[62]	Data Cleaning	-
	Resize	224 × 224
[15]	Data Augmentation	Crop, rotation, Gaussian noise, scale, flip
	Resize	600 × 1024, 300 × 300
[67]	Resize	256 × 256
[68]	Resize	256 × 256
[82]	Greyscale	-
	Resize	60 × 60

From the table, it is noticeable that all analysed papers employing deep learning-based techniques used image resizing. It is also worth mentioning that the application of data augmentation was found in 25% of the depicted works, and the same goes for image color conversion to grey-scale and data cleaning.

2.3. Machine Learning Models

ML models enable researchers to get insight into data and existing correlations between various factors that influence occurrence of diseases and pests in crops. After data is processed and features are extracted, models can be used for classification, regression, among other goals. In classification, a new data sample is assigned a label according to the relations retrieved during the training process. In regression, a continuous output value is estimated from the input variables.

The following sub-sections contain a description about the ML models used, published work that have used them and the achieved performances. In addition, as a consequence of the conducted research, it was decided to include a sub-section about the use and potential of Transfer Learning (TF) in the research under consideration.

2.3.1. Support Vector Machine

SVM [84] is a model that creates a hyper-plane that separates two classes (can also be adapted and applied for multi-class problems). By maximizing the distance, or margin, between the nearest data points (support vectors) of each class to the hyper-plane, SVM chooses the optimum hyper-plane to segregate the data. SVM can also perform well in non-linear data by using the so called *kernel trick* technique. The SVM kernel is a function that transforms a low dimensional input space into a higher dimensional space that is linearly separable. For this reason, SVM can be very effective in high dimensional spaces.

SVM can also be used for regression problems [29,40,85]. Furthermore SVM can also be used in a hybrid way as Bhatia et al. did in [86], by using SVM together with logistic regression algorithm to predict powdery mildew disease in tomato plant.

A syntheses of agricultural studies using SVM as the ML model can be observed in Table 3. The type of SVM used, as well as its kernel and result can be observed. Linear, polynomial, and RBF kernels seem to be most commonly used on SVM-based classification and regression algorithms applied to agriculture contexts.

Table 3. SVM performance.

Study	Classification/ Regression	Kernels	
		Type	**Results**
[52]	Classification	Polynomial	90.0% acc.
		Radial Basis Function	97.4% acc.
[29]	Regression	Not specified	SVM outperformed
[40]	Regression	Linear	$r^2 = 0.45$
[27]	Classification	Linear	90.0%+ acc.
[31]	Classification	Radial Basis Function	90.5% acc.
		Quadratic	92.0% acc.
		Linear	91.0% acc.
		Multi-Layer Perceptron	
		Polynomial	
[67]	Classification	Not specified	94.6% acc., 93.1% f1

SVM can achieve better performance than other ML techniques such as ANNs and conventional regression approaches in forecasting plant diseases [29].

2.3.2. Random Forest

Random Forest (RF) is a widely known ensemble built from decision trees trained on different subsets of the training data. Also, when deciding which variable to split on a node, RF considers a random set of variables and not the whole set of features. During classification, each tree votes and the class most agreed upon is returned. As each tree is trained on a subset of data and of features, the computation is fast. A high number of trees and the diversity of each of them makes them robust to noise and outliers. Some studies that have employed Random Forest (RF) are shown in Table 4.

Table 4. Performance of Random Forests.

Study	Classification/Regression	Number of Trees	Performance
[40]	Regression	100	$r^2 = 0.75$
[32]	Regression	200	$r^2 = 0.75$
[77]	Classification	-	70.0% acc.
[67]	Classification	-	95.5% acc., 94.2% f1

RFs can achieve greater accuracy with less number of samples when compared to other ML techniques [77].

2.3.3. Artificial Neural Networks

Artificial Neural Networks (ANN) are models inspired by biological brains. ANN consists of neurons distributed in input, hidden, and output layers and can have multiple hidden layers and multiple units in each layer. With more hidden layers, an ANN is able to learn complex relations from the hierarchical combination of multiple features, and thus create high-order features, Figure 7 shows an illustration of an ANN. Deep learning is associated with ANNs that contain a large amount of layers.

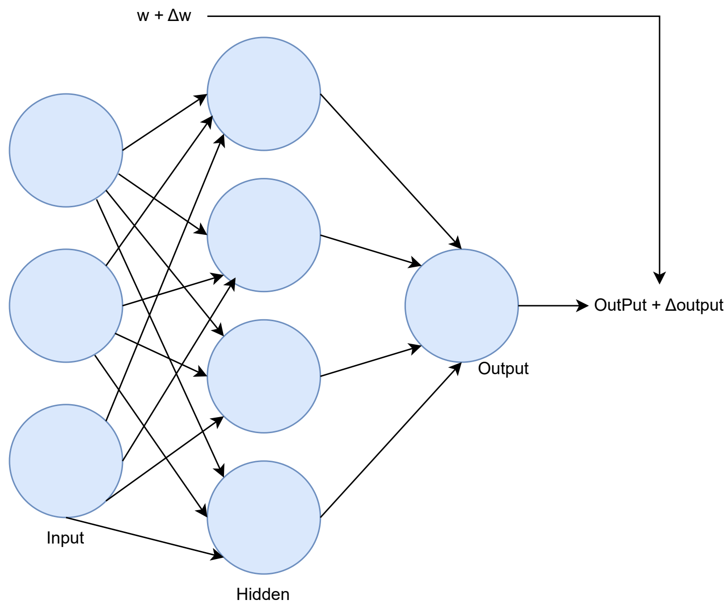


Figure 7. ANN example.

Learning occurs by a process called optimization, which is an iterative method for minimizing an error function, typically based the Gradient Descent algorithm. Instead of calculating the gradient from the entire data-set, the optimization process typically uses chunks of data records called batches. After the network processes the input, the output is compared to the expected output and the error is computed. The error is then propagated back through the network, one layer at a time, and the weights are updated according to the amount they contributed to the error. This updating process is called back-propagation. After all records in the data-set are processed once, a training epoch is completed. Training the network can require several epochs until desired results are achieved.

CNNs are a type of a deep learning network that commonly are applied on image classification tasks. In this type of network, the use of the so-called convolutional layers enables an hierarchical extraction of features, where simpler features such as edges are extracted in the first layers and more specific and complex features are extracted in deeper layers. The dimensionality of the input is decreased by the use of pooling layers. Fully connected neural networks are usually placed on top after the convolutional and pooling layers and act as classifiers using these high-level features.

Recurrent Neural Networks (RNN) are also a type of deep learning network, usually applied to time series data. RNNs extract features automatically from data and can capture temporal relationships. Because of the architecture of these networks, the gradients calculated to update the weights can become unstable, becoming too high (Exploding Gradient) or too low (Vanishing Gradient).

The recurrent layers can be structured in a wide variety of ways to produce distinct RNNs [87]. The LSTM cell was proposed by Hochreiter and Schmidhuber in [88]. Here, the remembering capacity for the standard recurrent cell was improved in order to deal with undesirable dependencies on the long-term.

Recently, Xiao et al. suggested in [14] that LSTM networks have specific advantages in processing time-dependent problems. LSTM networks can be used, for example, to retrieve relationships between meteorological data and pest occurrence in order to forecast future pest attacks.

In the context of agriculture, obtaining a large amount of annotated data for the training of ML-based algorithms can be a rather difficult task. Few-shot learning approaches have been trying to mitigate this problem by managing to learn with fewer data. The methods typically associated with this technique can be organized according to four groups [89]: data augmentation, metric learning, external memory, and parameter optimization. Yang et al. present a survey on the developments, application, and challenges of this approach.

When using ANN models, authors might use one of two methods. They either create their own model designs or adopt well-known architectures that have been shown to perform well in previous studies, particularly CNN architectures for image classification.

User-Defined Network Architectures

This sub-section presents studies where the authors defined their own neural network architectures.

In [40], Duarte-Carvajalino et al. built and compared the outcome of two different neural networks models. The first model was a Multi-Layer Perceptron (MLP) with 2 hidden layers, each having half the number of nodes of the previous layer. The authors used a learning rate of 0.01, the Adamax optimizer, batch normalization and dropout with probability of 0.2 in all layers, and ReLU as activation function. The other model was a CNN trained using the same hyperparameters used on the MLP. The CNN consists of two convolutional layers using 20 filter kernels of size 3×3, followed by a max pooling layer of size 2×2. The succeeding network layers are another two convolutional layers using 40 filter kernels of size 5×5 followed by a max pooling layer. After flattening, a dense layer is added before the output is computed. It was concluded that the CNN achieved better results than the MLP.

In [14], an LSTM network was used for processing time series data, i.e., winter and autumn data. The LSTM network consisted of two fully connected layers with five hidden units each. The results showed that the LSTM network achieved the best performance with 92% accuracy when compared to RF, SVM, and K-Nearest Neighbors (KNN). The Apriori algorithm [90] was applied for interpretability.

Disease prediction for different regions was also studied with the use of an ANN in [29]. In this case, the back-propagation neural network [91] and the generalized regression neural network [92] models were used.

A model suggested by Patil and Kumar in [93] attempted to identify the link between weather variables and the emergence of 4 types of rice diseases. In this work, the authors used an ANN to perform the detection, identification and prediction of the appearance of diseases in rice crops. The meteorological data-set referred to data between 1989 and 2019. The ANN consisted of 8 neurons in the input layer, 15 in the 2 hidden layers, and 5 in the output layer.

In [94], Sharma et al. performed a prediction of the potato late blight disease based on meteorological data only, using an ANN. In this case, data from 2011 and 2015 was used. Several tests with different network activation functions and data-set splits were done. It was concluded that the larger the data-set, the better was the performed prediction.

In addition, other algorithms relying on meteorological data and ANNs for performing predictions have been proposed. In [95], Dahikar and Rode present an ANN for predicting which crop will grow best in a certain area. The predictions were based on weather and soil data. Refs. [96,97] proposed ANN-based models for predicting crop yield.

Convolutional Neural Network Architectures

Image classification has achieved great results, with various model architectures being developed over the last 10 years. Most of these deep learning models were proposed in the context of the "Large Scale Visual Recognition Challenge" (ILSVRC). These models include well-known architectures such as AlexNet, GoogleNet, VGG, and ResNet, which have been widely used for image classification in different application domains.

Table 5 summarizes a set of studies that used pre-existing CNN architectures, depicting the architecture used in their work and the corresponding results.

Table 5. Performance of CNN architectures.

Study	Architecture	Results
[13]	GoogleNet	99.3%
	AlexNet	99.3%
[30]	CaffeNet	96.3%
[71]	VGG16	98.0% validation, 81.0% in new apple orchard
[62]	GoogleNet	43.5% acc., 32.7% f1
	FPN	54.9% mAP 0.5
	ResNet	49.4% acc., 40.1% f1
	VGGNet	48.2% acc., 38.7% f1
	AlexNet	41.8% acc., 34.1% f1
[67]	GoogleNet	98.7% acc., 97.1% f1
	AlexNet	99.2% acc., 98.5% f1
[68]	AlexNet	99.4% acc.
	VGG16	99.5% acc.
[63]	VGG16	60.4% acc., 60.0% f1
	InceptionResNet V2	70.5% acc., 70.0% f1
	Inception V3	62.1% acc., 61.0% f1
[82]	LeNet	98.6% acc., 98.6% f1

As can be observed from Table 5, several CNN architectures developed over the last decade have been successfully used, showing great potential for agriculture applications. From these, the use of older CNN architectures such as AlexNet (2012), VGG16 (2014), and GoogleNet (2014) were found on 44%, 33%, and 33% of the analysed papers, respectively. Although the use of the most recent CNN architectures is not expressed in the papers analysed in this review, we believe that, in the near future, the application of newer architectures to agriculture will be a reality.

2.3.4. Transfer Learning

TF makes use of already existing knowledge for some related task or domain in order and apply it to the problem under study. Models previously trained for image classification on large data-sets are usually used and adapted to the data-set under study. A common approach is to substitute the last network layers (i.e., the dense layers) of a pre-trained network, adapting it for a different classification task. The model is then trained but only the newly inserted layers are trainable—all network layers remain frozen during the training process. In extension of this approach, fine-tuning, is also commonly used. Besides training the newly inserted layers, fine-tuning allows the training of additional layers of the base model, typically the deeper convolutional layers of the network.

TF is usually done when the studied data-set is small, with insufficient samples for training a CNN model from scratch.

Table 6 synthesizes several deep learning-based studies where TF was applied. It presents details addressing: the data-set used for the base model training, the used TF method and the performance difference between using TF and training from scratch.

Table 6. TF analysis.

Study	Model	Dataset for Pretrain	Method	Performance Difference Compared to Training from Scratch
[13]	AlexNet, GoogleNet	ImageNet	All layers trainable	~−2% acc.
[30]	CaffeNet	ImageNet	Low learning rate for original layers (0.1), high for top layer (10)	~−0.50% acc.
[62]	AlexNet, GoogleNet, VGGNet, ResNet	ImageNet	Fine tune	~−14.0% acc. in best model (ResNet)
[15]	Faster RCNN (ResNet101, Inception V2, Inception ResNet V2)	COCO	Fine tune	No comparison
[67]	AlexNet, GoogleNet	ImageNet	Fine tune	~−2%
[63]	VGG16, Inception V3, Inception ResNet v2	ImageNet and/or PlantVillage	Fine tune	~−31.0% using ImageNet and PlantVillage

As can be observed from the table, the use of TF leads to lower performance when compared with training the full model from scratch. Nevertheless, there are many cases where such a difference is small, which means that TF can indeed be a useful possibility when the data-set is not sufficiently large.

3. Discussion

The studies collected in this review show that plant disease classification is a domain with promising results, with some studies achieving very high results [13,67]. Diverse data-sets have been employed, each with their own characteristics and associated difficulties: intraclass variability, background diversity, and different lighting and shading conditions during image acquisition. Due to these reasons, performance comparisons between the analyzed studies is not a straightforward task.

3.1. Data Acquisition

The data acquisition phase will have great influence on the quality of the ML model results, since the quantity and quality of data will influence the behaviour of ML models both in terms of good results and robustness.

Disease classification has shown promising results when the images of leaves are taken in laboratory conditions, with a single leaf against a neutral color background [13]. Laboratory image data is usually acquired using controlled lighting conditions using neutral color image backgrounds. On the other hand, images acquired in the cultivation fields are much harder to classify due to the presence of multiple leaves and plants, varying shading and lighting conditions, different ground textures and background objects [63]. Training ML models using laboratory images does not transfer well to testing in field conditions, achieving poor results. As for the opposite, reasonable results have been achieved [63,68]. The inclusion of images acquired in the field to the training processes can greatly improve performance of the models [68]. Nevertheless, the choice of data should always reflect the target objective of the application. For instance, if the model is intended to work in the laboratory environment, it isn't necessary to train it using data acquired on uncontrolled environments, since robustness to different acquisition conditions will not be necessary for the case. On the other hand, if the model is to be used in the field, images from the real context must be used. If the latter is not ensured using a large amount of images, the model will probably not perform well on the real conditions that exist in the crop fields.

Many studies applied ML-based algorithms to weather data records, since the weather conditions have an important role in the development of diseases and pests. Sensors for measuring weather data such as temperature, humidity, and rainfall fill data records that can be used to find correlation between the data and the development of pests or plant diseases in order to predict theses kind of developments [14,28,29]. Data records of meteorological measurements and pest occurrence can also be analyzed by deep learning with good results, as can be seen in [14]. Pest occurrence can be periodically manually analyzed by counting insect appearances [32] or by automated processes using computer vision techniques that detect insects on leaves or in traps [15,28,71].

Plants reflect NIR radiation differently depending on the disease damage [52]. If remote sensing data [32] or cameras equipped with filters [40] (allowing the capture of images at different bands of the light spectrum) are available, they can be used to compute NDVI. In [32], Skawsang et al. finds the relevance of NDVI and temperature for predicting pest occurrence. In [32], it is suggested that NDVI contains information about the relation between the crop growth stage and pest development. Duarte-Carvajalino et al. suggests in [40] that the NIR band is more suited for late blight detection than color imagery.

Thermal imagery combined with colored images provides higher accuracy when compared with using color features only, when detecting diseases that affect the temperature of the plant [27]. This is especially useful for detecting specific diseases in their early stages, where the plant has not yet developed visible symptoms. Knowledge about the variables that will influence the state of the plantation being worked on is very important when it comes to deciding which types of data should be acquired.

3.2. Data Pre-Processing

Data pre-processing techniques vary according to the used ML-based approach. In the case of image data, feature extraction can be done manually by applying computer vision algorithms, or automatically using deep learning.

Manual feature extraction processes typically demands pre-processing steps such as noise reduction or contrast enhancement. The researchers have to decide and select which feature extractors are more suitable for the problem at hand. When using deep learning, pre-processing is typically focused on data augmentation, enriching the training data-set in order to achieve a better model generalization. Deep learning shows better results when directly analyzing the originally acquired images when compared with the use of images converted to grey-scale [82] or subject to background removal [13]. This is a useful finding because background removal can be a complex and arduous task for images taken in field conditions, with complex and varying background [68]. When comparing the performance of ML models based on manual feature selection with models based on deep learning, the latter has shown better performance in studies that compared both approaches using the same input data [15,67,83].

Highlighting the region of interest of the leaves and reducing the background noise can increase the model's performance. This is valid for plant disease identification [63] as well as for insect classification [71].

3.3. Machine Learning Models

The studies presented along this paper have mostly used SVM, RF, or deep learning-based ML models. All of these have show promising results, highlighting the potential of using ML techniques for disease and pest classification, detection, and prediction. SVMs are robust and useful in high dimensional spaces due to their use of kernel trick. RF can avoid overfitting due to the high number of trees trained in different subsets of data. Deep learning usually achieves the best classification results due to its ability to create and extract hierarchical features from the inputs. Deep learning beats other ML models, particularly in image classification domains, especially when using pre-existing CNN architectures such as Inception and ResNet [62,63].

Despite deep learning models achieving higher accuracy values, SVM and RT can also achieve high values with accuracy above 94%, especially in disease classification on laboratory images [67]. SVM also achieves high accuracy, with values above 90%, in the detection of tomato diseases [30].

RNNs are capable of establishing relationships between weather data and pest occurrence, surpassing other models such as RF and SVM [14].

In scenarios where data is difficult to obtain, models trained with a lower amount of data can benefit from the use of TF, rather than having the models trained from scratch [13,52,72,76,77]. Most studies have their models pre-trained on large data-sets for image classification such as ImageNet or COCO. The inclusion of the PlantVillage data-set with ImageNet for pre-training helps to improve the accuracy of models for disease classification on images acquired in the field [62]. TF is typically applied by training some of the top layers of the pre-trained model jointly with the new classifier.

An alternative would be to address lack-of-data problems using few-shot learning approaches, as suggested in [89].

4. Conclusions

This survey presented an insight into existing research addressing the application of ML-based techniques for forecasting, detection, and classification of diseases and pests.

Data-sets containing weather, diseases, and pests data should keep records for long periods of time. Time-series ML models, such as RNN, can be employed to accurately forecast the occurrence of diseases and pests based on meteorological measurements series. NDVI measurements can also be helpful, since they provide additional information regarding the crop's development.

Detection and classification of pests and diseases can be performed using computer vision and deep-learning algorithms based on CNN models, which show better performance when compared with older image classification approaches based on "manual" features extraction. However, deep learning models require large amounts of data, which can be difficult to obtain. To tackle this issue, the use of transfer learning or few-shot learning methods can prove useful. Nonetheless, although the performance of deep learning-based methods is high for images acquired under controlled conditions, additional research is required regarding the analysis of images taken in the field, under real life conditions.

Since the literature does not yet include substantial work on pest and disease forecasting using combinations of different data modalities, this article also aimed to provide a general overview on the use of ML techniques over different types of data, in order to facilitate further developments that may help fulfill this gap.

Author Contributions: T.D.: research and writting; T.B.: writing and supervision; J.C.F.: supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 876925.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank Inov-Inesc Innovation for the conditions it provided for the realization of this article. This project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 876925. The JU receives support from the European Union's Horizon 2020 research and innovation programme and France, Belgium, Germany, Netherlands, Portugal, Spain and Switzerlan.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
CNN	Convolutional Neural Network
GLCM	Grey Level Co-occurrence Matrix
HoG	Histogram of oriented Gradient
ILSVRC	Large Scale Visual Recognition Challenge
JU	ECSEL Joint Undertaking
KNN	K-Nearest Neighbor
LPB	Local Binary Pattern
LSTM	Long Short Term Memory
ML	Machine Learning
MLP	Multi-Layer Perceptron
NDVI	Normalized Difference Vegetation Index
NIR	Near Infra-Red
PETA	Progressive Environmental and Agricultural Technologies
RF	Random Forest
RNN	Recurrent Neural Network
SGD	Stochastic Gradient Descent
SIFT	Scale Invariant Feature Transform
SURF	Speeded Up Robust Features
SVM	Support Vector Machine
TF	Transfer Learning
TPMD	Tomato Powdery Mildew Disease

References

1. Roser, M. Future population growth. In *Our World in Data*; University of Oxford: Oxford, UK, 2013.
2. Fróna, D.; Szenderák, J.; Harangi-Rákos, M. The challenge of feeding the world. *Sustainability* **2019**, *11*, 5816. [CrossRef]
3. Ray, D.K.; Mueller, N.D.; West, P.C.; Foley, J.A. Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* **2013**, *8*, e66428. [CrossRef]
4. FAO. *The Future of Food and Agriculture: Trends and Challenges*; FAO: Rome, Italy, 2017.
5. Deutsch, C.A.; Tewksbury, J.J.; Tigchelaar, M.; Battisti, D.S.; Merrill, S.C.; Huey, R.B.; Naylor, R.L. Increase in crop losses to insect pests in a warming climate. *Science* **2018**, *361*, 916–919. [CrossRef] [PubMed]
6. Food and Agriculture Organization of the United Nations FAOSTAT Pesticides Use. Available online: https://www.fao.org/faostat/en/#data/RP/visualize (accessed on 10 January 2022).
7. Kartikeyan, P.; Shrivastava, G. Review on emerging trends in detection of plant diseases using image processing with machine learning. *Int. J. Comput. Appl.* **2021**, *975*, 8887. [CrossRef]
8. Kamilaris, A.; Kartakoullis, A.; Prenafeta-Boldú, F.X. A review on the practice of big data analysis in agriculture. *Comput. Electron. Agric.* **2017**, *143*, 23–37. [CrossRef]
9. Rajpurkar, P.; Irvin, J.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.; Shpanskaya, K.; et al. Chexnet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv* **2017**, arXiv:1711.05225.
10. Wang, L.; Lee, C.Y.; Tu, Z.; Lazebnik, S. Training deeper convolutional networks with deep supervision. *arXiv* **2015**, arXiv:1505.02496.
11. Soni, A.; Dharmacharya, D.; Pal, A.; Srivastava, V.K.; Shaw, R.N.; Ghosh, A. Design of a machine learning-based self-driving car. In *Machine Learning for Robotics Applications*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 139–151.
12. Carleo, G.; Cirac, I.; Cranmer, K.; Daudet, L.; Schuld, M.; Tishby, N.; Vogt-Maranto, L.; Zdeborová, L. Machine learning and the physical sciences. *Rev. Mod. Phys.* **2019**, *91*, 045002. [CrossRef]
13. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using deep learning for image-based plant disease detection. *Front. Plant Sci.* **2016**, *7*, 1419. [CrossRef]
14. Xiao, Q.; Li, W.; Kai, Y.; Chen, P.; Zhang, J.; Wang, B. Occurrence prediction of pests and diseases in cotton on the basis of weather factors by long short term memory network. *BMC Bioinform.* **2019**, *20*, 688. [CrossRef]
15. Gutierrez, A.; Ansuategi, A.; Susperregi, L.; Tubío, C.; Rankić, I.; Lenža, L. A benchmarking of learning strategies for pest detection and identification on tomato plants for autonomous scouting robots using internal databases. *J. Sens.* **2019**, *2019*, 5219471. [CrossRef]
16. Plantix App. Available online: https://plantix.net/en/ (accessed on 11 January 2022).
17. Saillog. Available online: https://www.saillog.co/ (accessed on 11 January 2022).
18. CropDiagnosis. Available online: https://www.cropdiagnosis.com/portal/crops/en/home (accessed on 11 January 2022).

19. Gamaya. Available online: https://www.gamaya.com/ (accessed on 11 January 2022).
20. iFarmer. Available online: https://ifarmer.asia/ (accessed on 11 January 2022).
21. Chostner, B. See & Spray: The next generation of weed control. *Resour. Mag.* **2017**, *24*, 4–5.
22. Savla, A.; Israni, N.; Dhawan, P.; Mandholia, A.; Bhadada, H.; Bhardwaj, S. Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture. In Proceedings of the 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 19–20 March 2015; pp. 1–7.
23. Fenu, G.; Malloci, F.M. Forecasting plant and crop disease: An explorative study on current algorithms. *Big Data Cogn. Comput.* **2021**, *5*, 2. [CrossRef]
24. Vikas, L.; Dhaka, V. Wheat yield prediction using artificial neural network and crop prediction techniques (a survey). *Int. J. Res. Appl. Sci. Eng. Technol.* **2014**, *2*, 330–341.
25. Moysiadis, V.; Sarigiannidis, P.; Vitsas, V.; Khelifi, A. Smart farming in Europe. *Comput. Sci. Rev.* **2021**, *39*, 100345. [CrossRef]
26. Choudhuri, K.B.R.; Mangrulkar, R.S. Data Acquisition and Preparation for Artificial Intelligence and Machine Learning Applications. In *Design of Intelligent Applications Using Machine Learning and Deep Learning Techniques*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2021; pp. 1–11.
27. Raza, S.E.A.; Prince, G.; Clarkson, J.P.; Rajpoot, N.M. Automatic detection of diseased tomato plants using thermal and stereo visible light images. *PLoS ONE* **2015**, *10*, e0123262. [CrossRef]
28. Rustia, D.J.A.; Lin, T.T. An IoT-based wireless imaging and sensor node system for remote greenhouse pest monitoring. *Chem. Eng. Trans.* **2017**, *58*, 601–606.
29. Kaundal, R.; Kapoor, A.S.; Raghava, G.P. Machine learning techniques in disease forecasting: A case study on rice blast prediction. *BMC Bioinform.* **2006**, *7*, 485. [CrossRef]
30. Sladojevic, S.; Arsenovic, M.; Anderla, A.; Culibrk, D.; Stefanovic, D. Deep neural networks based recognition of plant diseases by leaf image classification. *Comput. Intell. Neurosci.* **2016**, *2016*, 3289801. [CrossRef]
31. Mokhtar, U.; Ali, M.A.; Hassanien, A.E.; Hefny, H. Identifying two of tomatoes leaf viruses using support vector machine. In *Information Systems Design and Intelligent Applications*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 771–782.
32. Skawsang, S.; Nagai, M.; K. Tripathi, N.; Soni, P. Predicting rice pest population occurrence with satellite-derived crop phenology, ground meteorological observation, and machine learning: A case study for the Central Plain of Thailand. *Appl. Sci.* **2019**, *9*, 4846. [CrossRef]
33. Significant Remote Sensing Vegetation Indices a Review of Developments and Applications. Available online: https://www.hindawi.com/journals/JS/2017/1353691/ (accessed on 11 January 2022).
34. Abdulridha, J.; Ampatzidis, Y.; Kakarla, S.C.; Roberts, P. Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques. *Precis. Agric.* **2020**, *21*, 955–978. [CrossRef]
35. Earth Observing System Vegetation Indices to Drive Digital Agri Solutions. Available online: https://eos.com/blog/vegetation-indices/ (accessed on 11 January 2022).
36. Evangelides, C.; Nobajas, A. Red-Edge Normalised Difference Vegetation Index (NDVI705) from Sentinel-2 imagery to assess post-fire regeneration. *Remote Sens. Appl. Soc. Environ.* **2020**, *17*, 100283. [CrossRef]
37. Albetis, J.; Duthoit, S.; Guttler, F.; Jacquin, A.; Goulard, M.; Poilvé, H.; Féret, J.B.; Dedieu, G. Detection of Flavescence dorée grapevine disease using unmanned aerial vehicle (UAV) multispectral imagery. *Remote Sens.* **2017**, *9*, 308. [CrossRef]
38. Chandel, A.K.; Khot, L.R.; Sallato, B. Apple powdery mildew infestation detection and mapping using high-resolution visible and multispectral aerial imaging technique. *Sci. Hortic.* **2021**, *287*, 110228. [CrossRef]
39. Wang, F.M.; Huang, J.F.; Tang, Y.L.; Wang, X.Z. New vegetation index and its application in estimating leaf area index of rice. *Rice Sci.* **2007**, *14*, 195–203. [CrossRef]
40. Duarte-Carvajalino, J.M.; Alzate, D.F.; Ramirez, A.A.; Santa-Sepulveda, J.D.; Fajardo-Rojas, A.E.; Soto-Suárez, M. Evaluating late blight severity in potato crops using unmanned aerial vehicles and machine learning algorithms. *Remote Sens.* **2018**, *10*, 1513. [CrossRef]
41. Kim, S.; Lee, M.; Shin, C. IoT-based strawberry disease prediction system for smart farming. *Sensors* **2018**, *18*, 4051. [CrossRef]
42. Yin, X.; Kropff, M.J.; McLaren, G.; Visperas, R.M. A nonlinear model for crop development as a function of temperature. *Agric. For. Meteorol.* **1995**, *77*, 1–16. [CrossRef]
43. Henderson, D.; Williams, C.J.; Miller, J.S. Forecasting late blight in potato crops of southern Idaho using logistic regression analysis. *Plant Dis.* **2007**, *91*, 951–956. [CrossRef]
44. Lasso, E.; Corrales, D.C.; Avelino, J.; de Melo Virginio Filho, E.; Corrales, J.C. Discovering weather periods and crop properties favorable for coffee rust incidence from feature selection approaches. *Comput. Electron. Agric.* **2020**, *176*, 105640. [CrossRef]
45. Small, I.M.; Joseph, L.; Fry, W.E. Development and implementation of the BlightPro decision support system for potato and tomato late blight management. *Comput. Electron. Agric.* **2015**, *115*, 57–65. [CrossRef]
46. Ghaffari, R.; Zhang, F.; Iliescu, D.; Hines, E.; Leeson, M.; Napier, R.; Clarkson, J. Early detection of diseases in tomato crops: An electronic nose and intelligent systems approach. In Proceedings of the 2010 International Joint Conference on Neural Networks (IJCNN), Barcelona, Spain, 18–23 July 2010; pp. 1–6.
47. Holopainen, J.K. Multiple functions of inducible plant volatiles. *Trends Plant Sci.* **2004**, *9*, 529–533. [CrossRef] [PubMed]
48. Diepeveen, D.; Armstrong, L.; Vagh, Y. Identifying key crop performance traits using data mining. In Proceedings of the IAALD-AFITA-WCCA Congress 2008 (World Conference on Agricultural Information and IT), Tokyo, Japan, 25–27 August 2008.

49. Patil, N.N.; Saiyyad, M.A.M. Machine learning technique for crop recommendation in agriculture sector. *Int. J. Eng. Adv. Technol.* **2019**, *9*, 1359–1363. [CrossRef]
50. Rosenzweig, C.; Iglesius, A.; Yang, X.B.; Epstein, P.R.; Chivian, E. *Climate Change and Extreme Weather Events-Implications for Food Production, Plant Diseases, and Pests*; NASA Publications: Nebraska, NU, USA, 2001.
51. PlantVillage Tomato | Diseases and Pests, Description, Uses, Propagation. Available online: https://plantvillage.psu.edu/topics/tomato/infos (accessed on 11 January 2022).
52. Dake, W.; Chengwei, M. The support vector machine (SVM) based near-infrared spectrum recognition of leaves infected by the leafminers. In Proceedings of the First International Conference on Innovative Computing, Information and Control-Volume I (ICICIC'06), Beijing, China, 30 August–1 September 2006; Volume 3, pp. 448–451.
53. Measuring Vegetation NDVI and EVI. Available online: https://earthobservatory.nasa.gov/features/MeasuringVegetation/measuring_vegetation_2.php (accessed on 11 January 2022).
54. Herbert, D.A.; Mack, T.; Reed, R.B.; Getz, R. *Degree-Day Maps for Management of Soybean Insect Pests in Alabama*; Auburn University: Auburn, AL, USA, 1988.
55. Research Models: Insects, Mites, Diseases, Plants, and Beneficials-from UC IPM. Available online: http://ipm.ucanr.edu/MODELS/models_scientific.html (accessed on 10 July 2014).
56. Entomology, L.O.E. Temperature-Dependent Development of Greenhouse Whitefly and Its Parasite Encarsia formosa. *Environ. Entomol.* ***1982***, 11*, 483–485.*
57. Miller, P.; Lanier, W.; Brandt, S. *Using Growing Degree Days to Predict Plant Stages*; Ag/Extension Communications Coordinator, Communications Services, Montana State University-Bozeman: Bozeman, MO, USA, 2001; Volume 59717, pp. 994–2721.
58. Calculating Degree Days. Available online: https://www.degreedays.net/calculation (accessed on 26 July 2022).
59. Tomato Diseases and Disorders | Home and Garden Information Center. Available online: https://hgic.clemson.edu/factsheet/tomato-diseases-disorders/ (accessed on 26 May 2021).
60. Seager, S.; Turner, E.L.; Schafer, J.; Ford, E.B. Vegetation's red edge: A possible spectroscopic biosignature of extraterrestrial plants. *Astrobiology* **2005**, *5*, 372–390. [CrossRef] [PubMed]
61. Hughes, D.; Salathé, M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv* **2015**, arXiv:1511.08060.
62. Wu, X.; Zhan, C.; Lai, Y.K.; Cheng, M.M.; Yang, J. Ip102: A large-scale benchmark dataset for insect pest recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 8787–8796.
63. Singh, D.; Jain, N.; Jain, P.; Kayal, P.; Kumawat, S.; Batra, N. PlantDoc: A dataset for visual plant disease detection. In Proceedings of the 7th ACM IKDD CoDS and 25th COMAD, Hyderabad, India, 5–7 January 2020; pp. 249–253.
64. Wu, S.G.; Bao, F.S.; Xu, E.Y.; Wang, Y.X.; Chang, Y.F.; Xiang, Q.L. A leaf recognition algorithm for plant classification using probabilistic neural network. In Proceedings of the 2007 IEEE International Symposium on Signal Processing and Information Technology, Giza, Egypt, 15–18 December 2007; pp. 11–16.
65. Lee, S.H.; Chan, C.S.; Mayo, S.J.; Remagnino, P. How deep learning extracts and learns leaf features for plant classification. *Pattern Recognit.* **2017**, *71*, 1–13. [CrossRef]
66. Bakeer, A.; Abdel-Latef, M.; Afifi, M.; Barakat, M. Validation of tomato powdery mildew forecasting model using meteorological data in Egypt. *Int. J. Agric. Sci.* **2013**, *5*, 372.
67. Brahimi, M.; Boukhalfa, K.; Moussaoui, A. Deep learning for tomato diseases: Classification and symptoms visualization. *Appl. Artif. Intell.* **2017**, *31*, 299–315. [CrossRef]
68. Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **2018**, *145*, 311–318. [CrossRef]
69. Garcia-Lamont, F.; Cervantes, J.; López, A.; Rodriguez, L. Segmentation of images by color features: A survey. *Neurocomputing* **2018**, *292*, 1–27. [CrossRef]
70. Types of Morphological Operations MATLAB and Simulink. Available online: https://www.mathworks.com/help/images/morphological-dilation-and-erosion.html. (accessed on 11 January 2022).
71. Albanese, A.; d'Acunto, D.; Brunelli, D. Pest detection for precision agriculture based on iot machine learning. In *International Conference on Applications in Electronics Pervading Industry, Environment and Society*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 65–72.
72. Sannakki, S.S.; Rajpurohit, V.S.; Nargund, V.; Kumar, A.; Yallur, P.S. Leaf disease grading by machine vision and fuzzy logic. *Int. J.* **2011**, *2*, 1709–1716.
73. Sekulska-Nalewajko, J.; Goclawski, J. A semi-automatic method for the discrimination of diseased regions in detached leaf images using fuzzy c-means clustering. In Proceedings of the Perspective Technologies and Methods in MEMS Design, Polyana, Ukraine, 1–14 May 2011; pp. 172–175.
74. Zhou, Z.; Zang, Y.; Li, Y.; Zhang, Y.; Wang, P.; Luo, X. Rice plant-hopper infestation detection and classification algorithms based on fractal dimension values and fuzzy C-means. *Math. Comput. Model.* **2013**, *58*, 701–709. [CrossRef]
75. Pang, J.; Bai, Z.y.; Lai, J.c.; Li, S.k. Automatic segmentation of crop leaf spot disease images by integrating local threshold and seeded region growing. In Proceedings of the 2011 International Conference on Image Analysis and Signal Processing, Wuhan, China, 21–23 October 2011; pp. 590–594.

76. Patil, S.B.; Bodhe, S.K. Leaf disease severity measurement using image processing. *Int. J. Eng. Technol.* **2011**, *3*, 297–301.
77. Ramesh, S.; Hebbar, R.; Niveditha, M.; Pooja, R.; Shashank, N.; Vinod, P. Plant disease detection using machine learning. In Proceedings of the 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), Bangalore, India, 25–28 April 2018; pp. 41–45.
78. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621. [CrossRef]
79. Heath, M.D.; Sarkar, S.; Sanocki, T.; Bowyer, K.W. A robust visual method for assessing the relative performance of edge-detection algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.* **1997**, *19*, 1338–1359. [CrossRef]
80. Ojala, T.; Pietikainen, M.; Harwood, D. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In Proceedings of the 12th International Conference on Pattern Recognition, Jerusalem, Israel, 9–13 October 1994; Volume 1, pp. 582–585.
81. Tan, L.; Lu, J.; Jiang, H. Tomato Leaf Diseases Classification Based on Leaf Images: A Comparison between Classical Machine Learning and Deep Learning Methods. *AgriEngineering* **2021**, *3*, 542–558. [CrossRef]
82. Amara, J.; Bouaziz, B.; Algergawy, A. A deep learning-based approach for banana leaf diseases classification. In *Datenbanksysteme für Business, Technologie und Web (BTW 2017)-Workshopband*; Gesellschaft für Informatik e.V.: Bonn, Germany, 2017.
83. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [CrossRef]
84. Suthaharan, S. Support vector machine. In *Machine Learning Models and Algorithms for Big Data Classification*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 207–235.
85. Gu, Y.; Yoo, S.; Park, C.; Kim, Y.; Park, S.; Kim, J.; Lim, J. BLITE-SVR: New forecasting model for late blight on potato using support-vector regression. *Comput. Electron. Agric.* **2016**, *130*, 169–176. [CrossRef]
86. Bhatia, A.; Chug, A.; Singh, A.P. Hybrid SVM-LR classifier for powdery mildew disease prediction in tomato plant. In Proceedings of the 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 27–28 February 2020; pp. 218–223.
87. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [CrossRef]
88. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
89. Yang, J.; Guo, X.; Li, Y.; Marinello, F.; Ercisli, S.; Zhang, Z. A survey of few-shot learning in smart agriculture: Developments, applications, and challenges. *Plant Methods* **2022**, *18*, 28. [CrossRef] [PubMed]
90. Agrawal, R.; Srikant, R. Fast algorithms for mining association rules. In Proceedings of the 20th International Conference on Very Large Data Bases, VLDB, Santiago, Chile, 12–15 September 1994; Volume 1215, pp. 487–499.
91. Buscema, M. Back propagation neural networks. *Subst. Use Misuse* **1998**, *33*, 233–270. [CrossRef]
92. Specht, D.F. A general regression neural network. *IEEE Trans. Neural Netw.* **1991**, *2*, 568–576. [CrossRef] [PubMed]
93. Patil, R.R.; Kumar, S. Predicting rice diseases across diverse agro-meteorological conditions using an artificial intelligence approach. *PeerJ Comput. Sci.* **2021**, *7*, e687. [CrossRef]
94. Sharma, P.; Singh, B.; Singh, R. Prediction of potato late blight disease based upon weather parameters using artificial neural network approach. In Proceedings of the 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 10–12 July 2018; pp. 1–13.
95. Dahikar, S.S.; Rode, S.V. Agricultural crop yield prediction using artificial neural network approach. *Int. J. Innov. Res. Electr. Electron. Instrum. Control Eng.* **2014**, *2*, 683–686.
96. Trenz, O.; Št'astnỳ, J.; Konečnỳ, V. Agricultural data prediction by means of neural network. *Agric. Econ.* **2011**, *57*, 356–361. [CrossRef]
97. Ranjeet, T.; Armstrong, L. An Artificial Neural Network for Predicting Crops Yield in Nepal. In Proceedings of the Asian Federation for Information Technology in Agriculture, Perth, Australia, 29 September–2 October 2014.

Article submitted to MDPI Agriculture Special Issue “The Application of Machine Learning in Agriculture”:

- **Manuscript ID** : agriculture-1971724;
- **Digital Object Identifier** : 10.3390/agriculture12111967;
- **Type of manuscript** : Article;
- **Title** : Insects detection in sticky traps images of tomato crops using machine learning.

agriculture

Article

Insect Detection in Sticky Trap Images of Tomato Crops Using Machine Learning

Tiago Domingues [1,2,*], Tomás Brandão [1], Ricardo Ribeiro [2] and João C. Ferreira [1,2]

[1] Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR-IUL, 1649-026 Lisbon, Portugal
[2] INOV-INESC Inovação—Instituto de Novas Tecnologias, 1000-029 Lisbon, Portugal
* Correspondence: tards@iscte-iul.pt

Abstract: As climate change, biodiversity loss, and biological invaders are all on the rise, the significance of conservation and pest management initiatives cannot be stressed. Insect traps are frequently used in projects to discover and monitor insect populations, assign management and conservation strategies, and assess the effectiveness of treatment. This paper assesses the application of YOLOv5 for detecting insects in yellow sticky traps using images collected from insect traps in Portuguese tomato plantations, acquired under open field conditions. Furthermore, a sliding window approach was used to minimize insect detection duplicates in a non-complex way. This article also contributes to event forecasting in agriculture fields, such as diseases and pests outbreak, by obtaining insect-related metrics that can be further analyzed and combined with other data extracted from the crop fields, contributing to smart farming and precision agriculture. The proposed method achieved good results when compared to related works, reaching 94.4% for *mAP_0.5*, with a precision and recall of 88% and 91%, respectively, using YOLOv5x.

Keywords: pests; insects; detection; identification; precision agriculture; machine learning; smart farming

check for updates

Citation: Domingues, T.; Brandão, T.; Ribeiro, R.; Ferreira, J.C. Insect Detection in Sticky Trap Images of Tomato Crops Using Machine Learning. *Agriculture* **2022**, *12*, 1967. https://doi.org/10.3390/agriculture12111967

Academic Editors: Nen Fu Huang and Ho-Hsien Chen

Received: 29 September 2022
Accepted: 17 November 2022
Published: 21 November 2022

1. Introduction

The world population has increased and is expected to continue to grow [1]. In recent decades, this growth has driven the demand for agricultural goods, resulting in an increase in crop areas [2]. However, traditional agricultural production is not economically or environmentally sustainable; hence, it is critical to make optimal use of resources to enable high-yield crops [2].

Furthermore, crop productivity is constantly threatened by insect pests. It is predicted that worldwide food supplements are declining by 40% on average every year owing to plant diseases and insect outbreaks [3]. Each year invasive insects cost the global economy around USD 70 billion [4].

Temperature influences the rate of population expansion in several insect species. In addition, the rise in global temperature caused by climate change influences insect damage and development. The metabolic rates of insects increase when the temperature rises, causing them to consume more food and inflict more harm. Crop losses due to insect pests are expected to increase by 10% to 25% for every degree of average global warming of the Earth's surface [5].

Tomato is a fruit–vegetable that has great potential to be cultivated since it is a source of vitamins and minerals. In terms of improving yields and fruit quality, tomatoes rank among the horticultural commodities with high economic value that still require careful handling [6]. It is critical to preserve these kinds of plantations against diseases and pests, in order to improve the quality and quantity of the crop [7]. According to data from the Food and Agriculture Organization of the United Nations, tomato production in Western Europe has increased considerably from at least 2000 to 2019 [8].

Numerous fungal, bacterial, and viral diseases have severely afflicted this plant, with symptoms appearing in various areas of the plant, such as the leaf, stem, fruit, etc.

Wilt, rot, stains on fruits, browning of foliage, and stunted development are some of the symptoms [9].

The advancements in information technology have allowed for the development of more precise farm management systems that overcome these invaders. Insect traps (ITs) are essential for keeping track of insect activity and are frequently used in pest detection and control programs, such as in [10], where trapping techniques for emerald ash borer and its introduced parasitoids were addressed. In [11], the authors address trapping, detection, control, and regulation of tephritid fruit flies, lures, area-wide programs, and trade implications associated with them. In [12], the authors address the use of pheromone traps to monitor the distribution and population trends of the gypsy moth; for further references, please also see [13–15]. ITs are also used to assess biodiversity, plan conservation [16–18], and evaluate pest activity and research initiatives, such as in [19], where over a two-year period, the association between female mating success and background male moth densities along the gypsy moth western front in northern Wisconsin, USA, was measured. In [20], the authors describe the usage of automated pheromone-baited traps, utilizing recording sensors and data loggers to collect male unique date–time stamps when they entered the trap; for further references, please also see [21–23].

As a result of the use of IT, a lot of research has been conducted to determine the effectiveness of traps, such as reference [24], where attraction and trapping capabilities of bucket- and delta-style traps with different pheromone emission rates for gypsy moths were compared. In [25], the performances of pheromone-baited traps to monitor the seasonal abundance of tortrix moths in chestnut groves were analyzed. In [26], the authors evaluated gravid traps for the collection of *culex quinquefasciatus*; for further references, please also see [27–30]. The research was also carried out to estimate the range of attraction, such as in reference [31], where the authors presented a novel method for estimating a pheromone trap attraction range to the pine sawyer beetle *monochamus galloprovincialis*. In [32], the range of attraction of pheromone traps to *agriotes lineatus* and *agriotes obscurus* was assessed. In [33], the authors assessed the attraction range of sex pheromone traps to *agriotes* male click beetles in South-Eastern Europe. In [34], the authors addressed the space of pheromone plume and its relationship with the effective attraction radius in applied models; for further references, please also see [35–38]. Work is also being conducted around the probabilities associated with insects, such as in [39,40]. Regarding the work in [39], the probability of detecting Caribbean fruit flies was addressed. Concerning the work in [40], the regional gypsy population trends (in an expanding population using a pheromone trap catch and spatial analysis) were predicted. This work on the probabilities associated with insects was conducted to better understand trap catches and to relate them to the absolute population density [41–47]. Regarding reference [41], the gypsy moth was used as the simulation model to interpret the capture of moths in pheromone-baited traps used for the surveillance of invasive species. Regarding the work in [44], the European pine sawfly was monitored with pheromone traps in maturing Scots pine stands. As for the work in [45], the autumn gum moth was monitored regarding relationships between pheromone and light trap catches and oviposition in eucalypt plantations.

For several insect trap systems, a relationship was found between trap catches and subsequent egg mass [44,45,48,49] and larval density [50–52]. However, translating trap catches into absolute population density and, in particular, interpreting zero catches, remains challenging at the quantitative level [12,24,41,53].

By gathering data on the target pest's existence, abundance, and dispersion, insect pest monitoring is often carried out in agriculture and forestry to evaluate the pest status in specific sites (such as a greenhouse, field, orchard/vineyard, or forest). The ultimate objective of insect pest monitoring within integrated pest management programs in agriculture is to give growers a useful decision-making tool. For instance, the intervention thresholds are crucial for optimizing the control method and grower inputs for a given insect pest infestation in a particular field at the ideal time. Insect population outbreaks can be predicted using monitoring data to develop prediction phenological models, providing extra knowledge to enhance control methods and maximize the use of insecticides [54]. Similarly, forestry relies heavily on the detection and monitoring of both native insect pests and invasive species to set up effective management programs. This is because forest insect species can have a serious negative influence on the biodiversity, ecology, and economy of the afflicted area [55].

The impetus for this work stemmed from the necessity to monitor insects that invade crops. The monitoring of insect populations potentiates an increased crop yield as the use of pesticides can be more efficient. Therefore, this work can contribute to precision agriculture [56]. On the other hand, the proposed technique for the detection and subsequent counting of insects, which corresponds to the number of bounding boxes retrieved, contributes to smart farming. To this end, use was made of YOLOv5 and a tiled image-splitting technique in order to optimize the model's performance.

Images from insect traps acquired in the open fields are subject to a wide variety of illumination conditions due to weather conditions, day-cycle light, landscape elements that cast shadows (e.g., trees, buildings, mountains), etc. The camera trap setup is also subject to oscillations due to the wind, which may result in lesser image quality due to motion blur. Trap imagery acquired in the open fields may also contain objects other than insects, such as leaves that stick to the traps. Machine learning models that use images acquired under these conditions tend to achieve worse results since they need to deal with such variability. On the other hand, images acquired in the laboratory are usually captured under fully controlled conditions (constant illumination, no wind, etc.), while images captured in greenhouses may also be subject to some uncontrolled environmental conditions (e.g., illumination variability), but not as adverse as on images captured in the fields.

This paper considers the much less controlled scenario of images acquired on the tomato crop fields, aiming to evaluate the applicability of YOLOv5 for the detection of insects in yellow sticky traps.

2. State-of-the-Art

Insect populations that exceed the economic threshold can cause significant harm to plants and, hence, diminish yields. The quantity of pests at an observed location is frequently determined by visually inspecting sticky surfaces in IT and counting the captured insects and this is a time-consuming job [57]. To overcome this problem, there has been much development of Internet of Things (IoT) systems with the support of machine learning for monitoring IT. This paper was developed in this direction, using images of IT captured by an IoT system to detect the number of insects present in the traps in the agricultural field through machine learning. This section will discuss some of the work that has been done in this area.

Deep learning was used to detect, identify, and count specific pest species in ITs in [58]. To reduce the impact of illumination variations on detection performance, a color correction variation [59] of the "gray-world" technique [60] was adopted. The authors suggested a sliding window-based detection pipeline that applies a convolutional neural network (CNN) to image patches at various locations to calculate the probability that they contain certain pests. Their work was inspired by algorithms proposed for pedestrian detection, analyzed in [61]. The final detections were produced via non-maximum suppression (NMS) [62] and thresholding of image patches based on their positions and related confidences. To evaluate the precision of the bounding boxes, the intersection-over-minimum (IoM) was computed. It was concluded that many of the errors occurred because the same moth could have various wing positions, occlusion levels, lighting circumstances, and decay patterns throughout time, indicating that the algorithm would improve in well-managed sites.

In [63], the authors' main objective was to create a model that detects whiteflies and thrips from sticky trap images in greenhouse settings. They developed a model based on faster region-based convolutional neural network (R-CNN), the "TPest-RCNN", and trained it using transfer learning with a public data set in the first phase. They utilized their data set with the weights from the first phase to the second phase. The model was proven to be accurate in detecting microscopic pests in images with varied pest concentrations and light reflections. It was also concluded that for recognizing insect species from images captured in sticky yellow traps, the best results were achieved by the proposed model, beating the faster R-CNN architecture and techniques employing manual feature extraction (color, shape, texture).

The research in [64] focuses on a four-layer deep neural network based on light traps with a search and rescue optimization strategy for identifying leaf folders and yellow stemborers. The search and rescue optimization approach was employed in the deep neural network to find the ideal weights to enhance the convergence rate, reduce the complexity of learning, and increase detection accuracy. The proposed method achieved 98.29% pest detection accuracy.

The proposed work in [65] studies the monitoring of spotted wing drosophila IT using image-based object detection with deep learning. The authors trained the ResNet-18 deep CNNs to detect and count the insect in question. From an image captured from a static position, an area under the precision–recall curve (AUC) of 0.506 was obtained for the female and 0.603 for the male. From the observed results, it was concluded that it is possible to use deep learning and object detection to monitor the insects.

In [66], the authors performed automatic insect detection where they first used a spectral residual model; different color features were then extracted. In the end, whiteflies and thrips were identified using a support vector machine classifier. The classification accuracies for the whiteflies and thrips were 93.9% and 89.8%, respectively. As for the detection of the trap, a precision of 93.3% was obtained.

To identify whiteflies and thrips, researchers in [67] presented an image-processing approach that included object segmentation and morphological processing of color features combined with classical neural networks. The images were acquired under controlled conditions, in a laboratory environment, from sticky traps moved from greenhouses. The proposed algorithms achieved 96% and 92% precision, respectively.

In [68], a pheromone-trapping device was developed. In this work, the original image was cropped into several sub-images with 30% overlap. These sub-images were then used to train the tested models, which were the images reconstructed with the detections performed. The results showed a mean average precision (mAP) of 94.7%.

Using IoT and deep learning frameworks, the work in [69] provided a real-time remote IT monitoring system and insect identification algorithm. The authors used the faster R-CNN ResNet 50 and an average accuracy (using different databases) of 94% was obtained.

The study in [70] used machine vision and deep learning to detect and count *Aphis glycines* automatically. To detect the insect, the authors used a sliding windows approach with a size of 400×400 pixels to slide over the acquired images with a stride of 400 pixels. Each image framed by the sliding windows in each step was fed into the faster R-CNN developed by the authors. The results demonstrate the high potential of the method proposed.

In [71], the authors proposed using low-cost cameras to capture and upload images of insect traps to the cloud. The authors used R-CNN and YOLO models to detect the insects, whitefly in this case, in yellow sticky traps. They used a public data set [72] for training the models. However, the images used for training were acquired under controlled illumination conditions. The authors do not explicitly state whether the images were split or used as a whole. The model with the best mAP was YOLOv5x, with a mAP of 89.70%.

The technique proposed in [73] combines high-tech deep learning with low-tech sticky insect traps. The authors propose a high-throughput cost-effective approach for monitoring flying insects as an enabling step towards "big data" entomology. In this work, the traps were captured a few days after being composed of a high number of insects, and images of them were only obtained after that capture, under laboratory and field conditions. The images were split into segments of 500×500 pixels. The authors concluded that the model was more likely to miss important images than it was to incorporate irrelevant ones.

Regarding the work in [74], the authors used yellow insect traps for the detection of *Trioza erytreae* and *Scaphoideus Titanus Ball* using image-processing techniques and the

FASTER R-CNN and YOLOv4 models. In order to promote the robustness of the models, images of the traps were taken by a 12-megapixel camera under different light conditions, backgrounds, and distortions. The authors did not perform splits on the images in order to train the models with tiles of the images instead of the images as wholes. The authors concluded that the models performed poorly with and without image processing.

Considering the methodologies stated, open-source solutions may be employed to aid in the detection process's implementation. In [75], this approach is followed, using the Computer Vision Annotation Tool (CVAT) (https://github.com/openvinotoolkit/cvat, accessed on 9 December 2021), which contains a feature for automatic annotation/labeling. This software can also be powered by Nuclio (https://nuclio.io/, accessed on 9 December 2021), a serverless technology that allows deploying trained models to CVAT. This tool was analyzed and it was concluded that it could be interesting to use it given the infrastructure of the project, as CVAT allows to create and carry out annotation tasks and, with Nuclio, deploy trained models [76].

From the state-of-the-art, it is not always clear that the approach used to split the image into tiles will feed the trained model. This is important, because in the case of splitting the image, in order to optimize the model performance, duplicated detections can arise. This problem is addressed in this paper and an approach to solve it is demonstrated. Furthermore, the main contribution of this paper was to test the application of YOLOv5 in detecting insects in traps (tomato plantations in this case). From the reviewed works, using YOLOv5, images acquired under controlled conditions (laboratory or greenhouses) were usually used. Thus, this paper contributes to the future developments of insect detection in images that are split using YOLOv5 and an approach that optimizes the performance of the trained model and the non-appearance of duplicate detections. Furthermore, this paper contributes to the monitoring and detection of insects in crop traps and, consequently, to the prediction of events in the agricultural field, by providing a new metric to be analyzed and correlated with other data from the crop.

3. Materials and Methods

In this article, a method was developed to detect insects in IT, yellow sticky cards, placed in agricultural fields. The work carried out in this article arose in the context of AI for new devices and technologies at the edge (ANDANTE) [77] project and, consequently, the data used in this work were provided by project partners. To carry out this work, first, the image was prepared to feed the artificial intelligence (AI) model, then the model was trained, and the results were evaluated and analyzed. This section presents the data set used and the pipeline of the method developed.

Given that there was no manual annotation on the images provided, the first stage of development was to manually annotate some yellow sticky cards and insects in the images. The open-source software CVAT, its application programming interface (API), and Nuclio (open source and managed serverless) were used in the developments described, making model training, manual and automatic detection, data management, and selection easier.

CVAT and its API allowed the creation of a website where all images were available and could be annotated manually and automatically. It was through CVAT that the bounding boxes of the yellow sticky cards and insects were manually created in the first phase. Through its API, it was possible to select images and access those same bounding boxes in the desired formats. With this access, everything was ready to start the development and training of the models with manual annotations. After the training, Nuclio was used to put the developed models into practice in CVAT, i.e., it became possible on the website to select a set of images in CVAT and apply the developed models to them with the immediate output of the results, in this case, the automatic bounding boxes of the yellow sticky cards and insects. This is because Nuclio allowed incorporating the developed models with the extra processing done, such as the splitting of the images into tiles and their consequent reconstruction, already with the respective automatic bounding boxes resulting from the annotations made by the model, thus providing CVAT with the coordinates of the bounding boxes to be placed on the image concerned. From CVAT API, it is thus possible to obtain the bounding boxes presented in each image and, consequently, the number of insects on the image in question.

3.1. Data Set

The data set used was related to Portuguese tomato plantations in the Ribatejo region, namely Valada, Castanheira, and Lezria, where ANDANTE Portuguese partners collected the data. Information about the tomato crop fields can be found in Table 1.

Table 1. Information on the tomato crop fields where data were acquired.

Location	Area (ha)	Planting Date	Central GPS Point
Castanheira	23	19 April 2021	38.982300, −8.954110
Lezíria	27	27 April 2021 and 10 May 2021	39.006537, −8.881018
Valada	20	07 May 2021	39.067730, −8.772214

The tomato cultivation fields where data were collected were fully mechanized, from planting to harvesting. The crop consists of natural tomato varieties, obtained from cross-pollination, without any kind of genetic modification. Sowing was in a greenhouse, starting at the end of January. Seedling production lasted about one-and-a-half to two months. The crop was staggered with a cycle of about 120 days, depending on the tomato varieties, and the start of planting took place between the end of March and the beginning of June. Planting was in 1.52 m wide ridges. Planting density was about 33,000 plants per hectare with drip irrigation.

The data set used contains 5646 images of IT captured by cameras placed in front of the traps. These were webcams with 12 megapixels. The traps were composed of chromotropic cards, yellow cards in this case, with glue, yellow in order to attract insects, such as *bemisia tabaci*. In addition, pheromones were placed in delta-type traps in order to attract the male insects so that they did not create offspring, such as *helicoverpa armigera*. The chromotropic leaves and pheromones were used in the biotechnical fight. In the whole data set, only 4637 images were considered legitimate since several did not correspond to IT or were not adequate to improve the model's performance. These images were considered invalid. This filtering is shown in Table 2.

Table 2. Data on the insect trap images where data were acquired.

	Trap 001	Trap 002	Trap 003	Trap 004	Trap 005	Trap 006
Field	Valada	Castanheira	Valada	Lezíria	Lezíria	Castanheira
Period of operation	27 May 2021 to 3 September 2021	26 May 2021 to 8 September 2021	27 May 2021 to 8 September 2021	27 May 2021 to 23 September 2021	27 May 2021 to 24 September 2021	26 May 2021 to 6 September 2021
Total images	848	948	901	945	1071	933
Valid Images	733	756	784	763	845	756

The images were captured every day, between the dates shown in Table 2. Furthermore, the acquisition was mostly done between 11 a.m. and 8 p.m. at different times of the day (11 a.m., 11.30 a.m., 12 midday, 4 p.m., 4.30 p.m., 5 p.m., 7 p.m., 7.30 p.m., and 8 p.m.), usually nine images were captured per day. The ANDANTE partners defined this configuration based on their understanding of the insect's behavior.

Figure 1 presents an example image for each of the six traps utilized.

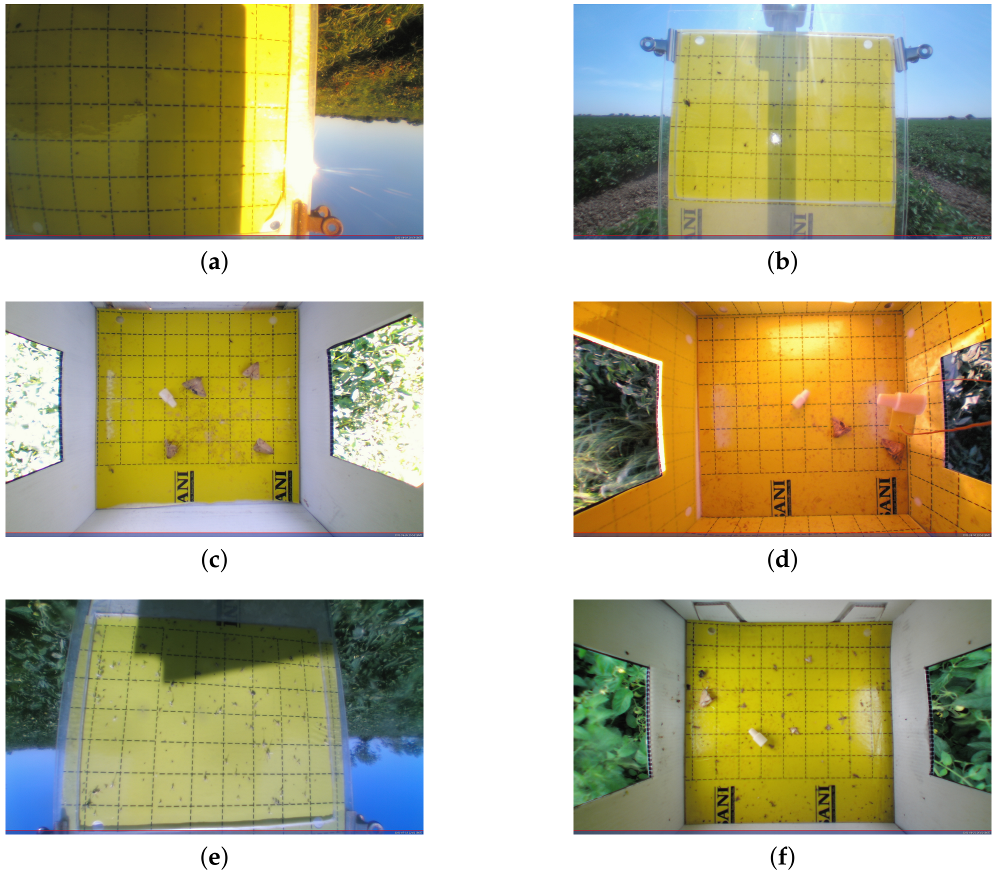


Figure 1. Examples of the data set. (**a**) Insect Trap 001. (**b**) Insect Trap 002. (**c**) Insect Trap 003. (**d**) Insect Trap 004. (**e**) Insect Trap 005. (**f**) Insect Trap 006.

3.2. Method Pipeline

An analysis of the images from the data set was carried out; a method was chosen in which the trap was first detected and then the insects presented in that trap through the bounding box resulting from the detection of the trap, the yellow sticky card.

Since ITs differ physically and are sensitive to varied lighting circumstances during image acquisition, we exclusively employed AI models for object detection, abandoning the usage of manual image-processing processes for insect detection. In addition, because the colors of the insects were generally the same as the colors of the lines on the yellow sticky cards, only AI models were used. Taking this into account, and the literature review [63,78–83], it was observed that AI models were increasingly being used, performing better and replacing more traditional methods that involved manual image processing; the manual image processing was discarded despite being considered at an early stage. Regarding the work in [79], it was verified that a YOLO model could perform better than the model used in the research for segmenting blueberries from an input image. In [63], the authors concluded that the faster R-CNN proposed had better results than techniques employing manual feature extraction for detecting whiteflies and thrips from sticky trap images in greenhouse conditions.

The insect detection process went as follows: the yellow sticky card in the original image was detected; the resultant bounding box was divided into tiles; the insects on each tile was detected; the original image was rebuilt with all bounding boxes. For the sake of improving the performance and results, cropping techniques were adopted [84]; the bounding box corresponding to the yellow sticky card, i.e. the result of the yellow sticky card detection model was split into tiles, and these tiles were used to train the insect models tested. From the performed detection, the number of insects presenting in each image can be directly inferred. Figure 2 depicts this pipeline split into two phases, A and B.

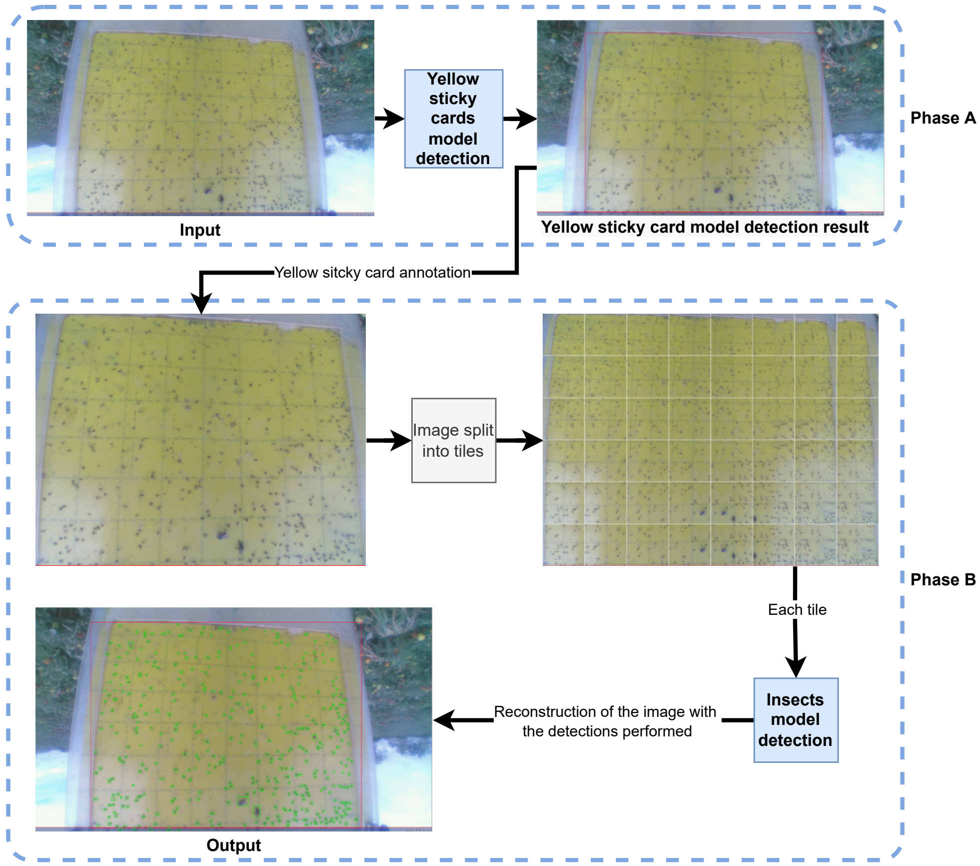


Figure 2. Pipeline for insect detection.

The YOLOv5 object detection model was used to perform the insect detection task. This choice is justified since YOLO is a widely used model that has been proposed for numerous object detection-based tasks and, its most recent version, the one used in this work, is showing an increasing usage trend [81]. Considering this trend and other works already mentioned in Section 2, it was decided to use YOLOv5 due to its potential performance in object detection tasks. Transfer learning was applied to train the model for insects and yellow sticky card detection.

The YOLOv5 model has different versions (YOLOv5s with a small size, YOLOv5m with a medium size, YOLOv5l with a large size, and YOLOv5x with an extra large size) and the basic structures of all these versions are the same. Their differences rely on the size of the model, with a multiplier that influences the width and the length (deepness) of the network. Generally, the larger the model, the better the performance at the expense of more processing time and required memory [85].

The parameters presented in Table 3 were used in all developments involving the use of YOLOv5.

Table 3. YOLOv5 insect trap image parameters.

Epochs	Batch Size	Optimizer	Patience
300	16	Stochastic Gradient Descent (SGD)	100

The results of YOLOv5 were obtained and analyzed through MLflow [86] integration. This integration made it possible to visualize the *mAP_0.5*, *mAP_0.5–0.95*, precision, recall, and loss during each training epoch. At the end of the training process, it was also possible to observe the F1 curve, as well as precision/recall curves. Of all the metrics obtained, due to the nature of the problem, the evaluation of the results was based on the *mAP_0.5*, *mAP_0.5–0.95*, precision, recall, F1 score, and F1 score curve.

The *mAP*, corresponds to the mean over classes, of the interpolated average of precision (AP), of each class (out of N classes), given by the area under the precision/recall curve [87], and is calculated as follows:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{1}$$

The precision measures the model's accuracy in classifying a sample as positive. It is calculated as the ratio between the number of positive samples correctly classified to the total number of samples classified as positive:

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives} \tag{2}$$

The recall of the model assesses its ability to recognize positive samples. The more positive samples identified, the larger the recall. The recall is computed as the ratio of positive samples that are properly categorized as positive to the total number of positive samples:

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives} \tag{3}$$

The F1 score combines the precision and recall of a classifier into a single metric by taking their harmonic mean. The F1 score formula is shown here:

$$F1\text{-}Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

4. Results

4.1. Yellow Sticky Card Model Detection

Phase A of the detection pipeline (see Figure 2), concerning yellow sticky card detection, was developed to use detection data to later detect the insects contained in the sticky cards.

From the valid images, explained in Section 3.1, 1272 insect trap images were manually annotated, which were the images of the data set used in this phase; 80% of the data set was used for training, 10% for validation, and the remaining 10% for testing. The images were resized to 640 by 640 pixels in the training process.

The lightweight YOLO model, YOLOv5s, was enough to achieve near-perfect results, as shown in Table 4, with the mAPs, precision, and recall reaching the maximum possible values or very close to them. With the developed trap detection model achieving good results, all of the images that had not been manually annotated were passed through the developed model and the correct detection was verified by the model.

Table 4. YOLOv5s yellow sticky card model results.

Phase	*mAP_0.5*	*mAP_0.5–0.95*	Precision	Recall
Training	0.995	0.995	1	1
Testing	0.995	0.995	1	1

4.2. Insect Model Detection

The insect detection model was developed considering only the bounding box corresponding to the detection of the yellow sticky card. The YOLO model was again used, but in this case, more powerful versions of YOLOv5 were tested.

Initially, the tiles were obtained with increments of the base tile sizes; in cases where these increments were not divisive of the widths and/or lengths of the images, the tiles in the margins (right and/or bottom) were smaller than the remaining tiles (Figure 3c); this approach was termed the pure split (PS). In the second phase, in order to keep all tiles with the same dimensions, black/yellow/white borders were added to the tiles with smaller dimensions (Figure 3d); this approach was termed pure split with border (PSB). However, these approaches were discarded since it was possible for an insect to be split between tiles in these approaches. This could lead to two detections representing the same object—one corresponding to the part of the object that was in a certain tile and the other to the part of the object that was in a tile in the vicinity of the previous one. This situation is illustrated in Figure 4.

This situation would complicate the process of reconstructing the bounding boxes in the original image as the creation of the new bounding boxes (based on the original ones) would become complex and there would be a wide variety of possibilities when verifying which bounding boxes belong to the same object.

Due to this potential problem, the development focused on two new alternative approaches, namely:

- **Overlapping with the different size(s) (ODS)**: Tiles with different dimensions depend on the positions of the tiles relative to the image and overlapping occurs (Figure 3a);
- **Overlapping with the same size(s) (OSS)**: The tiles are the same dimensions (320 × 320 px). Zones may have more overlapping areas than others (Figure 3b).

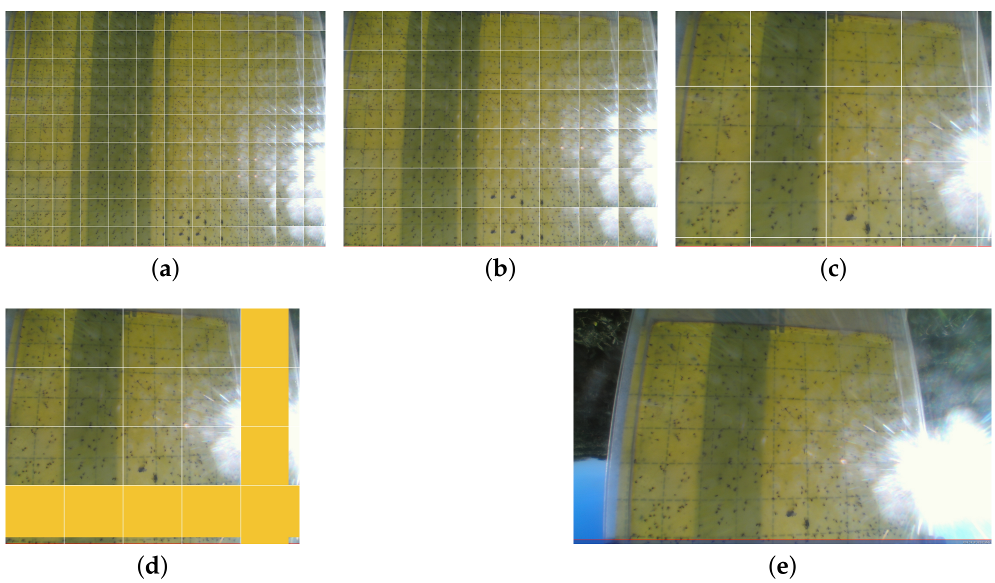


Figure 3. Yellow sticky card splitting approaches. (**a**) ODS approach. (**b**) OSS approach. (**c**) PS approach. (**d**) PSB approach. (**e**) Original image.

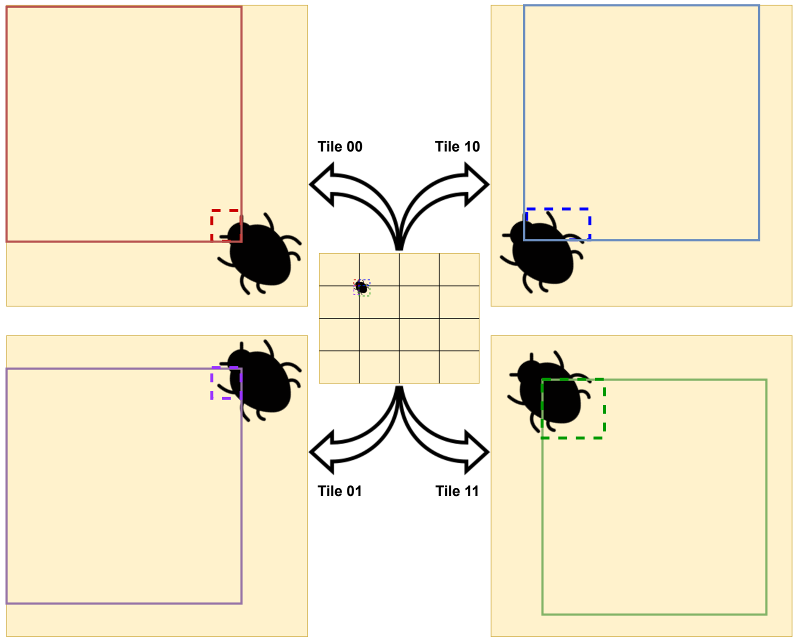


Figure 4. Illustration of splits without overlapping the split insects.

For all the tests performed, the number of images used was the same—248 insect trap images. However, due to the different approaches to performing the splitting, the tile numbers used to train the models were different for each approach. For ODS and OSS 11,375 and 5092 tiles were used when training and testing the models, respectively. In all approaches, 80% of the data set was used for train, 10% for validation, and the remaining 10% for the test.

The overlapping of tiles was done with caution making sure that the overlapping zone occupied an area of 160 × 160 px (Figure 5). By analyzing the images and the insects presenting in them, and questioning experts in the area, it was discovered that the maximum area that a bounding box could occupy is below these values. In this way, the problem that arose was solved. If an insect is split between tiles it will be partially detected in some tiles but will always be fully detected on a neighboring tile; this type of situation is illustrated in Figure 5. Thus, when reconstructing the image, it became only necessary to understand which detections overlapped, by checking and comparing each bounding box position, which ones had the largest area and confidence, and removing the duplicated ones. This way, only the bounding boxes detecting the whole object would remain.

From the tests carried out, a few incorrect detections or missing detections were observed, but they were in the minority when compared to the accurate ones. These flaws can be suppressed when the values obtained in each image are associated with groups, for example, between 0 and 20—few insects, between 20 and 100—some insects, etc. This association is important when analyzing the data and verifying the respective correlations with additional crop data (e.g., for performing event forecasting). These types of failures are reflected in the *mAP_0.5–0.95* metric, which is significantly lower than the *mAP_0.5* metric in all tests performed (these results are depicted in Tables 5 and 6). This can be expected since the *mAP_0.5–0.95* is computed over different intersection over union (IoU) [88] thresholds, from 0.5 to 0.95 with a step of 0.05, while *mAP_0.5* uses a fixed threshold at 0.5.

From the tables, it can be observed that the results achieved across all the tested models do not vary significantly. This means that, in cases where computational resources are limited, the lighter models can be used and still achieve good performance. By analyzing the precision, recall, and F1 score of all models, this situation becomes quite clear.

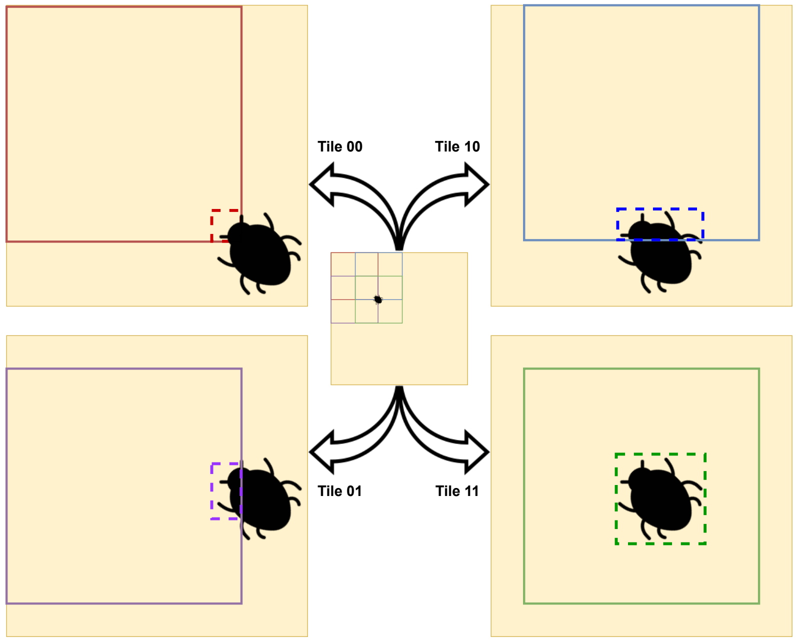


Figure 5. Illustration of overlapping tiles that split insects.

Tables 5 and 6 also reflect that ODS and OSS approaches achieve similar results with the YOLOv5x model, reaching the best results in both cases. However, due to the uniformity that OSS provides to the dimensions of the tiles without the need for resizing, the OSS approach was considered for the development of the remaining work.

Table 5. YOLOv5 insect model results for ODS.

Model	Phase	*mAP_0.5*	*mAP_0.5–0.95*	Precision	Recall	F1 Score
YOLOv5s	Training	0.973	0.678	0.982	0.935	0.958
	Testing	0.945	0.539	0.937	0.89	0.913
YOLOv5m	Training	0.975	0.7	0.976	0.94	0.958
	Testing	0.933	0.554	0.908	0.88	0.894
YOLOv5l	Training	0.979	0.724	0.986	0.947	0.966
	Testing	0.952	0.567	**0.938**	**0.906**	**0.922**
YOLOv5x	Training	0.98	0.733	0.982	0.951	0.966
	Testing	**0.952**	**0.573**	0.935	0.9	0.917

Table 6. YOLOv5 Insect Model results for OSS.

Model	Phase	*mAP_0.5*	*mAP_0.5–0.95*	Precision	Recall	F1 Score
YOLOv5s	Training	0.964	0.632	0.963	0.940	0.951
	Testing	0.923	0.497	0.912	0.853	0.882
YOLOv5m	Training	0.975	0.691	0.982	0.946	0.964
	Testing	0.946	0.542	0.946	0.874	0.909
YOLOv5l	Training	0.973	0.694	0.981	0.939	0.960
	Testing	0.937	0.543	**0.951**	0.862	0.904
YOLOv5x	Training	0.976	0.713	0.983	0.95	0.966
	Testing	**0.944**	**0.559**	0.942	**0.88**	**0.910**

An analysis of the applicability of this development and communication with the end users of ANDANTE led to the conclusion that it was preferable to have a balance between false positives and false negatives. If too many false detections (false positives) occur, it would mean a possible acquisition by end users of products, in vain, or a constant check in the field of values reflected by the detection. On the other hand, if too many false negatives occur, it would mean the possible appearance of pests without the perception of the end user. Furthermore, this balance will always be the best situation to ensure that the correlations performed with other data (acquired to make predictions regarding crop events) are not biased. Therefore, the F1 score was analyzed since it is adequate when both types of errors (false positives and false negatives) are not desired. Figure 6 depicts the graph of the F1 score curve.

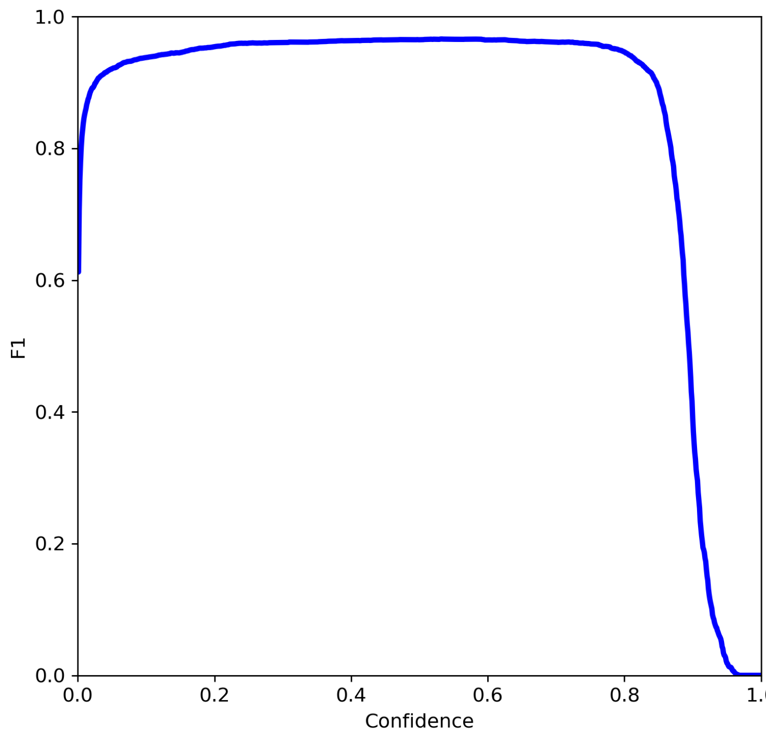


Figure 6. F1 score curve for the YOLOv5x model using the OSS approach.

By analyzing the plot of Figure 6, it is possible to have a significantly high confidence value that optimizes the F1 score at the same time; this value is between 0.7 and 0.8. Furthermore, *mAP_0.5* is a metric that is mostly used in object detection [89], and good results are obtained from it. Therefore, the analyses of the F1 score curve and *mAP_0.5* reflect the good performance of the model.

Although a comparison with other works cannot be directly performed, due to the use of different data sets and differences in the tasks performed by the object detection models, the results reported in the related work presented in Section 2 are summarized in Table 7.

Table 7. Comparison with other insect detection works.

Reference	Image Acquisition	Performance	Metric	Year
Proposed	Field, no controlled conditions	94.4% ∣ 94.2%	*mAP_0.5* ∣ Precision	2022
[63]	Greenhouse	95.2%	Accuracy	2021
[66]	Greenhouse	93.9% (whitefly) 89.8% (thrips)	Precision	2022
[67]	Laboratory	96% (whitefly) and 92% (thrips)	Precision	2016
[68]	Field, controlled conditions	94.7% (black pine bast scale)	*mAP_0.5*	2022

Using a faster R-CNN object detection model, the work in [63] achieved a mean F1 score of 94.4% and an accuracy of 95.2% in the detection of whiteflies and thrips as well as insect trap images acquired in greenhouses. In the approach followed in [66], automatic insect detection was conducted using a spectral residual model followed by the extraction of color features that were sent to a SVM classifier. The goal was to identify whiteflies and thrips; accuracies of 93.9% and 89.8% were achieved, respectively. As for the detection of the trap, a precision of 93.3% was obtained, which is less than the one achieved by the model proposed in this paper (100%). By comparing the results in both works, the approach using a deep learning-based object detection model in [63] seems to lead to better results than the approach in [66], which relies on image-processing techniques and classical machine learning models. As for [67], the images used for training and testing the system were acquired under controlled laboratory conditions, from sticky traps that were collected from greenhouses. They achieved precision rates of 96% and 92% for the detection of whiteflies and thrips. These results seem aligned with the ones achieved in [63]; however, since the images were acquired in a less adverse environment, the results may be biased when compared with those resulting from images acquired directly in the greenhouse. In [68], different object detection models were tested for detecting black pine bast scale pests Among the tested models, YOLOv5 achieved the best results, reaching an F1 score of 0.90 and mAP of 94.7%. The setup used for the image acquisition process (besides being used for a different task) was much more sophisticated than our own.

From Table 7, it can be seen that the approach presented in this paper is aligned with other works. It shows the potential of using the proposed image splitting approach together with YOLOv5 for detecting insects in sticky traps whose images are acquired in more adverse image acquisition conditions.

5. Conclusions

This paper presents the use and performance of YOLOv5 object detection models for insect detection in yellow sticky traps, using images acquired on tomato crop fields. The insect detection process uses a sliding window approach that minimizes the appearance of duplicate detections in yellow sticky card IT images. The presented YOLOv5 model demonstrated robustness and resilience for performing well under various illumination and adverse element exposure conditions. This work contributes to raising the bar for insect detection and monitoring. Furthermore, by creating another metric related to crop fields, this paper contributes to the development associated with forecasts of events regarding the agriculture field, such as the forecasting of disease and pest appearances.

There were limitations due to the absence of manual annotations of insects, which made it impossible to develop models for the detection and classification of insects trained with all available images.

The detection associated with the yellow sticky card and the subsequent training of AI models was performed in the first phase. In this phase, optimal results were obtained using YOLOv5s, and it was possible to perform the detection of yellow stick cards in all data sets.

The second phase was dependent on the first, as it was supposed to use the bounding box associated with the detection performed of the yellow sticky card in order to improve the accuracy of the detections of the insects in the traps. At this stage, a problem that this paper contributed to solving was faced: how does one perform the splits on the yellow sticky card bounding box image in a way that maximizes the quality of the model while not causing insects to be lost during the process of splitting and reconstructing the bounding boxes on the original image? The approach that ended up generally having the best results was OSS, where the tiles were the same sizes and overlapped, with 94.2% of precision in the test set with the YOLOv5x model. It can be concluded that the presented approach and the YOLOv5 models have potential in the detection of insects in insect traps scattered in an agricultural field.

It is possible to develop an insect detection model with the need for human supervision at times since the number and location of bounding boxes may be inaccurate. However, these errors are never in substantial quantities and can end up mostly suppressed when associating the number of detections performed in an image to a group. This association has advantages at the time of the data treatment and analysis.

6. Future Work

The annotation of all currently available images will be a part of future work, in order to build larger training and test sets. This annotation can either be manual or semi-automatic, assisted by the models presented in this paper. Larger data sets are expected to lead to more robust and accurate machine learning models.

Another topic for future work is the identification of specific insect species among those detected in the yellow sticky cards. For such a task, a larger number of images need to be acquired since greater diversities of data are required for covering the various species of insects to be identified.

It may also be valuable to evaluate the applications of other popular object detection networks, (e.g., faster R-CNN or single shot detector (SSD)) using the image splitting method proposed in this paper.

Future work will also involve testing the counting of insects themselves (in addition to their detection). Since the count is directly associated with the number of detections, and the detection model achieves high accuracy, we expect that the accuracies of insect counting will achieve results similar to the detection process. Nevertheless, this experiment will be put to the test and allow researchers to conclude its effectiveness in terms of considering the sliding window approach presented in this paper.

Author Contributions: T.D.: methodology, data curation, conceptualization, research, and writing; T.B.: conceptualization and supervision; R.R.: project administration, conceptualization and supervision; J.C.F.: conceptualization and supervision. All authors have read and agreed to the published version of the manuscript.

Funding: Part of this work was supported by Fundação para a Ciência e a Tecnologia (FCT), Portugal, under the Information Sciences, Technologies, and Architecture Research Center (ISTAR) projects UIDB/04466/2020 and UIDP/04466/2020.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Written informed consent was obtained from the patient(s) to publish this paper.

Data Availability Statement: Not applicable.

Acknowledgments: We thank Inov-Inesc Innovation for the conditions it provided for this article. We are also grateful to TerraPro Technologies for building the traps and acquiring the images. Finally, we would like to thank Italagro—Indústria de Transformação de Produtos Alimentares, S. A., for providing the fields to place the traps. This project received funding from the ECSEL Joint Undertaking (JU) under grant agreement no. 876925. This JU received support from the European Union's Horizon 2020 research and innovation programme, and France, Belgium, Germany, Netherlands, Portugal, Spain, and Switzerland.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	artificial intelligence
ANDANTE	AI for new devices and technologies at the edge
API	application programming interface
AUC	area under the precision–recall curve
CVAT	computer vision annotation tool
CNN	convolutional neural network
FCT	Fundação para a Ciência e a Tecnologia
ISTAR	Information Sciences, Technologies, and Architecture Research Center
IoT	Internet of Things
IT	insect traps
mAP	mean average precision
NMS	non-maximum suppression
ODS	overlapping with different size
OSS	overlapping with same size
PS	pure split
PSB	pure split with borders
R-CNN	region-based convolutional neural network
SGD	stochastic gradient descent
SSD	single shot detector

References

1. Roser, M. Future Population Growth. 2013. Our World in Data. Available online: https://ourworldindata.org/future-population-growth (accessed on 9 December 2021).
2. Fróna, D.; Szenderák, J.; Harangi-Rákos, M. The challenge of feeding the world. *Sustainability* **2019**, *11*, 5816. [CrossRef]
3. Thangaraj, R.; Anandamurugan, S.; Pandiyan, P.; Kaliappan, V.K. Artificial intelligence in tomato leaf disease detection: A comprehensive review and discussion. *J. Plant Dis. Prot.* **2021**, *129*, 469–488. [CrossRef]
4. FAO. *The Future of Food and Agriculture: Trends and Challenges*; FAO: Rome, Italy, 2017.
5. Deutsch, C.A.; Tewksbury, J.J.; Tigchelaar, M.; Battisti, D.S.; Merrill, S.C.; Huey, R.B.; Naylor, R.L. Increase in crop losses to insect pests in a warming climate. *Science* **2018**, *361*, 916–919. [CrossRef] [PubMed]
6. Anton, A.; Rustad, S.; Shidik, G.F.; Syukur, A. Classification of Tomato Plant Diseases Through Leaf Using Gray-Level Co-occurrence Matrix and Color Moment with Convolutional Neural Network Methods. In *Smart Trends in Computing and Communications: Proceedings of SmartCom 2020*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 291–299.
7. Brahimi, M.; Boukhalfa, K.; Moussaoui, A. Deep learning for tomato diseases: Classification and symptoms visualization. *Appl. Artif. Intell.* **2017**, *31*, 299–315. [CrossRef]
8. FAO. *FAOSTAT: FAO Statistical Databases*; FAO: Rome, Italy, 2021.
9. Verma, S.; Chug, A.; Singh, A.P. Prediction models for identification and diagnosis of tomato plant diseases. In Proceedings of the 2018 IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, India, 19–22 September 2018; pp. 1557–1563.
10. Abell, K.; Poland, T.M.; Cossé, A.; Bauer, L.S. Trapping techniques for emerald ash borer and its introduced parasitoids. In *Biology and Control of Emerald Ash Borer*; Van Driesche, R.G., Reardon, R.C., Eds.; FHTET-2014-09; US Department of Agriculture, Forest Service, Forest Health Technology Enterprise Team: Morgantown, WV, USA, 2015; Chapter 7, pp. 113–127.
11. Shelly, T.; Epsky, N.; Jang, E.B.; Reyes-Flores, J.; Vargas, R. *Trapping and the Detection, Control, and Regulation of Tephritid Fruit Flies: Lures, Area-Wide Programs, and Trade Implications*; Springer: Berlin/Heidelberg, Germany, 2014.
12. Elkinton, J.S.; Cardé, R.T. The use of pheromone traps to monitor distribution and population trends of the gypsy moth. In *Management of Insect Pests with Semiochemicals*; Springer: Berlin/Heidelberg, Germany, 1981; pp. 41–55.
13. Kuno, E. Verifying zero-infestation in pest control: A simple sequential test based on the succession of zero-samples. *Res. Popul. Ecol.* **1991**, *33*, 29–32. [CrossRef]
14. Tobin, P.C.; Onufrieva, K.S.; Thorpe, K.W. The relationship between male moth density and female mating success in invading populations of L ymantria dispar. *Entomol. Exp. Appl.* **2013**, *146*, 103–111. [CrossRef]
15. Tobin, P.C.; Sharov, A.A.; Liebhold, A.A.; Leonard, D.S.; Roberts, E.A.; Learn, M.R. Management of the gypsy moth through a decision algorithm under the STS project. *Am. Entomol.* **2004**, *50*, 200–209. [CrossRef]
16. Bossart, J.L.; Carlton, C.E. Insect Conservation in America: Status and Perspectives. *Am. Entomol.* **2002**, *48*, 82–92. [CrossRef]
17. Larsson, M.C. Pheromones and other semiochemicals for monitoring rare and endangered species. *J. Chem. Ecol.* **2016**, *42*, 853–868. [CrossRef]

18. New, T. Taxonomic focus and quality control in insect surveys for biodiversity conservation. *Aust. J. Entomol.* **1996**, *35*, 97–106. [CrossRef]
19. Contarini, M.; Onufrieva, K.S.; Thorpe, K.W.; Raffa, K.F.; Tobin, P.C. Mate-finding failure as an important cause of Allee effects along the leading edge of an invading insect population. *Entomol. Exp. Appl.* **2009**, *133*, 307–314. [CrossRef]
20. Tobin, P.C.; Klein, K.T.; Leonard, D.S. Gypsy moth (Lepidoptera: Lymantriidae) flight behavior and phenology based on field-deployed automated pheromone-baited traps. *Environ. Entomol.* **2009**, *38*, 1555–1562. [CrossRef] [PubMed]
21. Casado, D.; Cave, F.; Welter, S. Puffer®-CM Dispensers for mating disruption of codling moth: Area of influence and impacts on trap finding success by males. *IOBC Bull.* **2014**, *99*, 25–31.
22. Elkinton, J.; Cardé, R. Distribution, dispersal, and apparent survival of male gypsy moths as determined by capture in pheromone-baited traps. *Environ. Entomol.* **1980**, *9*, 729–737. [CrossRef]
23. Tcheslavskaia, K.; Brewster, C.C.; Sharov, A.A. Mating success of gypsy moth (Lepidoptera: Lymantriidae) females in southern Wisconsin. *Great Lakes Entomol.* **2002**, *35*, 1.
24. Cardé, R.T.; Bau, J.; Elkinton, J.S. Comparison of attraction and trapping capabilities of bucket-and delta-style traps with different pheromone emission rates for gypsy moths (Lepidoptera: Erebidae): Implications for understanding range of attraction and utility in surveillance. *Environ. Entomol.* **2018**, *47*, 107–113. [CrossRef]
25. Ferracini, C.; Pogolotti, C.; Lentini, G.; Saitta, V.; Busato, E.; Rama, F.; Alma, A. Performance of pheromone-baited traps to monitor the seasonal abundance of tortrix moths in chestnut groves. *Insects* **2020**, *11*, 807. [CrossRef]
26. Irish, S.R.; Moore, S.J.; Derua, Y.A.; Bruce, J.; Cameron, M.M. Evaluation of gravid traps for the collection of Culex quinquefasciatus, a vector of lymphatic filariasis in Tanzania. *Trans. R. Soc. Trop. Med. Hyg.* **2013**, *107*, 15–22. [CrossRef]
27. Elkinton, J.S.; Childs, R.D. Efficiency of two gypsy moth (Lepidoptera: Lymantriidae) pheromone-baited traps. *Environ. Entomol.* **1983**, *12*, 1519–1525. [CrossRef]
28. Hartstack, A., Jr.; Hollingsworth, J.; Ridgway, R.; Hunt, H. Determination of trap spacings required to control an insect population. *J. Econ. Entomol.* **1971**, *64*, 1090–1100. [CrossRef]
29. Hartstack, A.W., Jr.; Hollingsworth, J.; Lindquist, D. A technique for measuring trapping efficiency of electric insect traps. *J. Econ. Entomol.* **1968**, *61*, 546–552. [CrossRef]
30. Williams, C.B. Comparing the efficiency of insect traps. *Bull. Entomol. Res.* **1951**, *42*, 513–517. [CrossRef]
31. Jactel, H.; Bonifacio, L.; Van Halder, I.; Vétillard, F.; Robinet, C.; David, G. A novel, easy method for estimating pheromone trap attraction range: Application to the pine sawyer beetle Monochamus galloprovincialis. *Agric. For. Entomol.* **2019**, *21*, 8–14. [CrossRef]
32. Sufyan, M.; Neuhoff, D.; Furlan, L. Assessment of the range of attraction of pheromone traps to Agriotes lineatus and Agriotes obscurus. *Agric. For. Entomol.* **2011**, *13*, 313–319. [CrossRef]
33. Furlan, L.; Contiero, B.; Tóth, M. Assessment of the attraction range of sex pheromone traps to Agriotes (Coleoptera, Elateridae) male click beetles in South-Eastern Europe. *Insects* **2021**, *12*, 733. [CrossRef]
34. Byers, J.A. Active space of pheromone plume and its relationship to effective attraction radius in applied models. *J. Chem. Ecol.* **2008**, *34*, 1134–1145. [CrossRef]
35. Wall, C.; Perry, J. Range of action of moth sex-attractant sources. *Entomol. Exp. Appl.* **1987**, *44*, 5–14. [CrossRef]
36. Byers, J.A.; Anderbrant, O.; Löqvist, J. Effective attraction radius. *J. Chem. Ecol.* **1989**, *15*, 749–765. [CrossRef]
37. Dufourd, C.; Weldon, C.; Anguelov, R.; Dumont, Y. Parameter identification in population models for insects using trap data. *BioMath* **2013**, *2*, 1312061. [CrossRef]
38. Schlyter, F. Sampling range, attraction range, and effective attraction radius: Estimates of trap efficiency and communication distance in coleopteran pheromone and host attractant systems 1. *J. Appl. Entomol.* **1992**, *114*, 439–454. [CrossRef]
39. Calkins, C.; Schroeder, W.; Chambers, D. Probability of detecting Caribbean fruit fly, Anastrepha suspensa (Loew)(Diptera: Tephritidae), populations with McPhail traps. *J. Econ. Entomol.* **1984**, *77*, 198–201. [CrossRef]
40. Gage, S.H.; Wirth, T.M.; Simmons, G.A. Predicting regional gypsy moth (Lymantriidae) population trends in an expanding population using pheromone trap catch and spatial analysis. *Environ. Entomol.* **1990**, *19*, 370–377. [CrossRef]
41. Bau, J.; Cardé, R.T. Simulation modeling to interpret the captures of moths in pheromone-baited traps used for surveillance of invasive species: The gypsy moth as a model case. *J. Chem. Ecol.* **2016**, *42*, 877–887. [CrossRef]
42. Kirkpatrick, D.M.; Acebes-Doria, A.L.; Rice, K.B.; Short, B.D.; Adams, C.G.; Gut, L.J.; Leskey, T.C. Estimating monitoring trap plume reach and trapping area for nymphal and adult Halyomorpha halys (Hemiptera: Pentatomidae) in crop and non-crop habitats. *Environ. Entomol.* **2019**, *48*, 1104–1112. [CrossRef] [PubMed]
43. Kirkpatrick, D.M.; Gut, L.J.; Miller, J.R. Estimating monitoring trap plume reach and trapping area for Drosophila suzukii (Diptera: Drosophilidae) in Michigan tart cherry. *J. Econ. Entomol.* **2018**, *111*, 1285–1289. [CrossRef] [PubMed]
44. Lyytikäinen-Saarenmaa, P.; Varama, M.; Anderbrant, O.; Kukkola, M.; Kokkonen, A.M.; Hedenström, E.; Högberg, H.E. Monitoring the European pine sawfly with pheromone traps in maturing Scots pine stands. *Agric. For. Entomol.* **2006**, *8*, 7–15. [CrossRef]
45. Östrand, F.; Elek, J.A.; Steinbauer, M.J. Monitoring autumn gum moth (Mnesampela privata): Relationships between pheromone and light trap catches and oviposition in eucalypt plantations. *Aust. For.* **2007**, *70*, 185–191. [CrossRef]
46. Turchin, P.; Odendaal, F.J. Measuring the effective sampling area of a pheromone trap for monitoring population density of southern pine beetle (Coleoptera: Scolytidae). *Environ. Entomol.* **1996**, *25*, 582–588. [CrossRef]

47. Miller, J.R. Sharpening the precision of pest management decisions: Assessing variability inherent in catch number and absolute density estimates derived from pheromone-baited traps monitoring insects moving randomly. *J. Econ. Entomol.* **2020**, *113*, 2052–2060. [CrossRef]
48. Thorpe, K.W.; Ridgway, R.L.; Leonhardt, B.A. Relationship Between Gypsy Moth (Lepidoptera: Lymantriidae) Pheromone Trap Catch and Population Density: Comparison of Traps Baited with 1 and 500 < g (+)-Disparlure Lures. *J. Econ. Entomol.* **1993**, *86*, 86–92.
49. Evenden, M.; Borden, J.; Van Sickle, G. Predictive capabilities of a pheromone-based monitoring system for western hemlock looper (Lepidoptera: Geometridae). *Environ. Entomol.* **1995**, *24*, 933–943. [CrossRef]
50. Allen, D.; Abrahamson, L.; Eggen, D.; Lanier, G.; Swier, S.; Kelley, R.; Auger, M. Monitoring spruce budworm (Lepidoptera: Tortricidae) populations with pheromone-baited traps. *Environ. Entomol.* **1986**, *15*, 152–165. [CrossRef]
51. Sanders, C. Monitoring spruce budworm population density with sex pheromone TRAPS1. *Can. Entomol.* **1988**, *120*, 175–183. [CrossRef]
52. Sanders, C. *Pheromone Traps for Detecting Incipient Outbreaks of the Spruce Budworm, Choristoneura fumiferana (Clem.)*; NFP Technical Report TR-32; NODA: Peterborough, UK, 1996.
53. Östrand, F.; Anderbrant, O. From where are insects recruited? A new model to interpret catches of attractive traps. *Agric. For. Entomol.* **2003**, *5*, 163–171. [CrossRef]
54. Dent, D. Sampling, monitoring and forecasting. In *Insect Pest Management*; Springer: Berlin/Heidelberg, Germany, 2000; pp. 14–47.
55. Brockerhoff, E.G.; Liebhold, A.M.; Jactel, H. The ecology of forest insect invasions and advances in their management. *Can. J. For. Res.* **2006**, *36*, 263–268. [CrossRef]
56. Precision Agriculture. An International Journal on Advances in Precision Agriculture. Available online: https://www.springer.com/journal/11119 (accessed on 9 December 2021).
57. Marković, D.; Vujičić, D.; Tanasković, S.; Đorđević, B.; Ranđić, S.; Stamenković, Z. Prediction of pest insect appearance using sensors and machine learning. *Sensors* **2021**, *21*, 4846. [CrossRef]
58. Ding, W.; Taylor, G. Automatic moth detection from trap images for pest management. *Comput. Electron. Agric.* **2016**, *123*, 17–28. [CrossRef]
59. Nikitenko, D.; Wirth, M.; Trudel, K. Applicability Of White-Balancing Algorithms to Restoring Faded Colour Slides: An Empirical Evaluation. *J. Multimed.* **2008**, *3*, 9–18. [CrossRef]
60. Buchsbaum, G. A spatial processor model for object colour perception. *J. Frankl. Inst.* **1980**, *310*, 1–26. [CrossRef]
61. Dollar, P.; Wojek, C.; Schiele, B.; Perona, P. Pedestrian detection: An evaluation of the state of the art. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *34*, 743–761. [CrossRef]
62. Hosang, J.; Benenson, R.; Schiele, B. Learning non-maximum suppression. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4507–4515.
63. Li, W.; Wang, D.; Li, M.; Gao, Y.; Wu, J.; Yang, X. Field detection of tiny pests from sticky trap images using deep learning in agricultural greenhouse. *Comput. Electron. Agric.* **2021**, *183*, 106048. [CrossRef]
64. Muppala, C.; Guruviah, V. Detection of leaf folder and yellow stemborer moths in the paddy field using deep neural network with search and rescue optimization. *Inf. Process. Agric.* **2021**, *8*, 350–358. [CrossRef]
65. Roosjen, P.P.; Kellenberger, B.; Kooistra, L.; Green, D.R.; Fahrentrapp, J. Deep learning for automated detection of Drosophila suzukii: potential for UAV-based monitoring. *Pest Manag. Sci.* **2020**, *76*, 2994–3002. [CrossRef] [PubMed]
66. Li, W.; Yang, Z.; Lv, J.; Zheng, T.; Li, M.; Sun, C. Detection of Small-Sized Insects in Sticky Trapping Images Using Spectral Residual Model and Machine Learning. *Front. Plant Sci.* **2022**, *13*, 915543. [CrossRef]
67. Espinoza, K.; Valera, D.L.; Torres, J.A.; López, A.; Molina-Aiz, F.D. Combination of image processing and artificial neural networks as a novel approach for the identification of Bemisia tabaci and Frankliniella occidentalis on sticky traps in greenhouse agriculture. *Comput. Electron. Agric.* **2016**, *127*, 495–505. [CrossRef]
68. Yun, W.; Kumar, J.P.; Lee, S.; Kim, D.S.; Cho, B.K. Deep learning-based system development for black pine bast scale detection. *Sci. Rep.* **2022**, *12*, 606. [CrossRef]
69. Ramalingam, B.; Mohan, R.E.; Pookkuttath, S.; Gómez, B.F.; Sairam Borusu, C.S.C.; Wee Teng, T.; Tamilselvam, Y.K. Remote insects trap monitoring system using deep learning framework and IoT. *Sensors* **2020**, *20*, 5280. [CrossRef]
70. Hsieh, K.Y.; Kuo, Y.F.; Kuo, C.K. Detecting and Counting Soybean Aphids Using Convolutional Neural Network. In Proceedings of the 2018 ASABE Annual International Meeting. American Society of Agricultural and Biological Engineers, Detroit, MI, USA, 29 July–1 August 2018; p. 1.
71. Cardoso, B.; Silva, C.; Costa, J.; Ribeiro, B. Internet of Things Meets Computer Vision to Make an Intelligent Pest Monitoring Network. *Appl. Sci.* **2022**, *12*, 9397. [CrossRef]
72. Nieuwenhuizen, A.; Hemming, J.; Suh, H. Detection and classification of insects on stick-traps in a tomato crop using Faster R-CNN. In Proceedings of the The Netherlands Conference on Computer Vision, Eindhoven, The Netherlands, 26–27 September 2018.
73. Gerovichev, A.; Sadeh, A.; Winter, V.; Bar-Massada, A.; Keasar, T.; Keasar, C. High throughput data acquisition and deep learning for insect ecoinformatics. *Front. Ecol. Evol.* **2021**, *9*, 600931. [CrossRef]

74. da Silva Pinto Bessa, B.L. Automatic Processing of Images of Chromotropic Traps for Iden tification and Quantification of Trioza erytreae and Scaphoideus titanus. 2021. Available online: https://repositorio-aberto.up.pt/handle/10216/139335 (accessed on 9 December 2021).
75. Günther, C.; Jansson, N.; Liwicki, M.; Simistira-Liwicki, F. Towards a machine learning framework for drill core analysis. In Proceedings of the 2021 IEEE Swedish Artificial Intelligence Society Workshop (SAIS), Umea, Sweden, 14–15 June 2021; pp. 1–6.
76. Guillermo, M.; Billones, R.K.; Bandala, A.; Vicerra, R.R.; Sybingco, E.; Dadios, E.P.; Fillone, A. Implementation of Automated Annotation through Mask RCNN Object Detection model in CVAT using AWS EC2 Instance. In Proceedings of the 2020 IEEE Region 10 Conference (TENCON), Osaka, Japan, 16–19 November 2020; pp. 708–713.
77. Andante Use Case 2.2: Tomato Pests and Diseases Forecast. Available online: https://www.andante-ai.eu/project/use-case-2-2-tomato-pests-and-diseases-forecast/ (accessed on 9 December 2021).
78. Hu, C.; Liu, X.; Pan, Z.; Li, P. Automatic detection of single ripe tomato on plant combining faster R-CNN and intuitionistic fuzzy set. *IEEE Access* **2019**, *7*, 154683–154696. [CrossRef]
79. Ni, X.; Li, C.; Jiang, H.; Takeda, F. Three-dimensional photogrammetry with deep learning instance segmentation to extract berry fruit harvestability traits. *ISPRS J. Photogramm. Remote. Sens.* **2021**, *171*, 297–309. [CrossRef]
80. Lin, S.; Jiang, Y.; Chen, X.; Biswas, A.; Li, S.; Yuan, Z.; Wang, H.; Qi, L. Automatic Detection of Plant Rows for a Transplanter in Paddy Field Using Faster R-CNN. *IEEE Access* **2020**, *8*, 147231–147240. [CrossRef]
81. Jiang, P.; Ergu, D.; Liu, F.; Cai, Y.; Ma, B. A Review of Yolo algorithm developments. *Procedia Comput. Sci.* **2022**, *199*, 1066–1073. [CrossRef]
82. Liu, G.; Nouaze, J.C.; Touko Mbouembe, P.L.; Kim, J.H. YOLO-tomato: A robust algorithm for tomato detection based on YOLOv3. *Sensors* **2020**, *20*, 2145. [CrossRef] [PubMed]
83. Mu, Y.; Chen, T.S.; Ninomiya, S.; Guo, W. Intact detection of highly occluded immature tomatoes on plants using deep learning techniques. *Sensors* **2020**, *20*, 2984. [CrossRef] [PubMed]
84. Domingues, T.; Brandão, T.; Ferreira, J.C. Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey. *Agriculture* **2022**, *12*, 1350. [CrossRef]
85. Dlužnevskij, D.; Stefanovic, P.; Ramanauskaite, S. Investigation of YOLOv5 Efficiency in iPhone Supported Systems. *Balt. J. Mod. Comput.* **2021**, *9*, 333–344. [CrossRef]
86. MLflow. A Plataform for the Machine Learning Lifestyle. Available online: https://mlflow.org/ (accessed on 9 December 2021).
87. Henderson, P.; Ferrari, V. End-to-end training of object class detectors for mean average precision. In Proceedings of the Asian Conference on Computer Vision, Perth, WA, Australia, 2–6 December 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 198–213.
88. Rezatofighi, H.; Tsoi, N.; Gwak, J.; Sadeghian, A.; Reid, I.; Savarese, S. Generalized intersection over union: A metric and a loss for bounding box regression. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 658–666.
89. Everingham, M.; Van Gool, L.; Williams, C.K.; Winn, J.; Zisserman, A. The pascal visual object classes (voc) challenge. *Int. J. Comput. Vis.* **2010**, *88*, 303–338. [CrossRef]