Computers and Electronics in Agriculture 181 (2021) 105922

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Review

Technological support for detection and prediction of plant diseases: A systematic mapping study

Vinicius Bischoff[a], Kleinner Farias[a], Juliano Paulo Menzen[a], Gustavo Pessin[b]

[a] Applied Computing Graduate Program (PPGCA), University of Vale do Rio dos Sinos (Unisinos), São Leopoldo, Rio Grande do Sul, Brazil
[b] Vale Institute of Technology (ITV), Ouro Preto, Minas Gerais, Brazil

ARTICLE INFO

Keywords:
Systematic review
Disease detection
Machine learning
Sensors

ABSTRACT

The field of plant disease diagnosis and epidemiology seeks to assess symptoms caused by pathogens. Different infectious and non-infectious agents can cause similar symptoms in plant organs. Diagnosing diseases is crucial, but it remains an inherently manual and error-prone task. Many works have been proposed to diagnose plant diseases, mainly using machine learning approaches. Even though this field affects agribusiness areas, little has been done to classify and map the current literature. This article presents a comprehensive overview of the current literature, and draw some research gaps, trends, and challenges that are worth investigating. A systematic mapping of the literature was carried out in pairs, following well-established practice guidelines. In total, 56 primary studies were carefully selected from a sample of 668 papers, which were retrieved from 9 widely recognized electronic databases. They were analyzed and categorized to answer seven research questions. The results show that 41% of primary studies applied machine learning techniques to detect diseases, 32% used image sensors to identify symptoms related to plant diseases, 30% focused on proposing new models of machine learning to detect diseases 34% were evaluation studies, and 71% were published in scientific journals. The association between computer vision and neural networks appears as a promising field of research for the detection of diseases. Finally, this article can serve as a starting point for upcoming studies, providing insights from a systematic map of the literature.

1. Introduction

The field of plant disease diagnosis and epidemiology seeks to assess symptoms caused by pathogens. Different infectious and non-infectious pathogenic agents can cause similar symptoms in plant organs in the context of agricultural environments, affecting the metabolism of plants (Moore, 2018). Plant diseases are the result of interactions between susceptible plants with pathogenic agents, in favorable environmental conditions, as illustrated in Fig. 1. These conditions are related to several factors, including stages of plant development, the availability of pathogenic to the climate, or other environmental conditions that can change over a given period (Raja et al., 2018; Garrett et al., 2016). The human activity may also affect the development of plant diseases (Agrios, 2005) through the application of pest and disease management tools (Bischoff and Farias, 2020).

To date, diagnosing diseases remains a pivotal task, requiring the detection and classification of diseases by farmers. Unfortunately, detecting and classifying diseases are still two inherently manual and error-prone tasks. To help control agricultural risks such as plant epidemics, the correct diagnosis and monitoring of plant diseases is essential, as farmers and technicians can decide (and measure) the proper management of plants and the protection of crops. Recognizing lesions caused by a specific pathogen in plant organs (i.e., stems, leaves, and roots) are key actions to guide the right choices and technical decisions in the field. Although the adoption of a single technique may be sufficient for a confident identification, two or more may be used to ensure the identification. For example, previous studies have invested some effort in such direction such as Shah et al. (2016), Prajapati et al. (2016), KKaur and Bhatiaaur and Bhatia (2020), Pantazi et al. (2019).

The use of images to measure diseases might be considered as an alternative for detection of diseases. However, image processing is not a trivial task, especially in an agricultural environment, requiring a labor-intensive and relatively costly procedure for farmers, and a large team of experts and continuous plant monitoring (Singh and Misra, 2017). Base on disease symptoms and visible appearances, researchers and agro-industry focus on the solution of software, computers and electronics in agriculture (Silva et al., 2016), coupled with deep machine learning techniques to detect and identify a wide range of diseases. Machine

E-mail address: viniciusbischof@edu.unisinos.br (V. Bischoff).

https://doi.org/10.1016/j.compag.2020.105922
Received 13 June 2019; Received in revised form 9 November 2020; Accepted 21 November 2020
Available online 7 January 2021

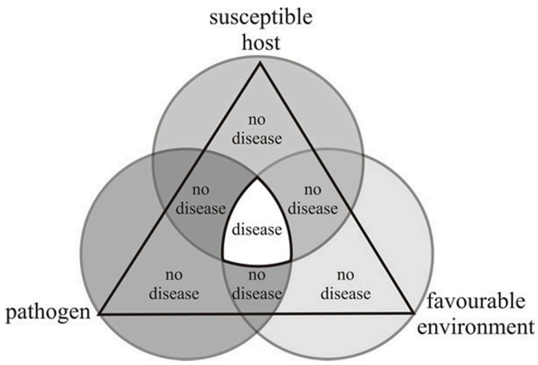

Fig. 1. Schematic diagram of the interrelationships of factors involved in plant disease epidemics Moore (2018).

learning models and techniques are often used in practice to detect a wide range of diseases (Pawara et al., 2018; Verma and Jain, 2019). The variety of diseases investigated with the use of these technologies has shown positive results, for example, the grapevine downy mildew on the vineyards (Abeledo et al., 2016), and mosaic viruses on the tomato crops (Fuentes et al., 2018). Recent studies show that the detection of plant diseases is still a challenging problem (Barbedo, 2019; Verma and Jain, 2019; Barbedo, 2016).

Over the decades, some works have been proposed to detect and classify plant diseases, mainly using machine learning approaches (Ferentinos, 2018; Mohanty et al., 2016). Even though plant diseases affect several agribusiness areas, little has been done to classify and map the current literature. Some literature reviews (Kamilaris and Prenafeta-Boldú, 2018; Kamilaris et al., 2017) on machine learning techniques have been published. The first explores deep learning techniques in agriculture (Kamilaris and Prenafeta-Boldú, 2018). The second investigates techniques for big data analysis in agriculture (Kamilaris et al., 2017). Although they are literature reviews, they have not explored issues related to disease detection in agricultural environments. Recent studies make use of deep learning models for plant disease detection and diagnosis (Ferentinos, 2018; Mohanty et al., 2016). However, their findings fall short of the required panoramic overview concerning the detection and classification of plant diseases. Even worse, these studies do not explore important dimensions, such as sensors and technologies commonly used to detect diseases, causing-disease factors, machine-learning techniques for detecting plant diseases, which diseases are currently detected, and the research methods used to investigate disease detection in agricultural environments.

The objective of this study is twofold: (1) to provide a comprehensive overview of the current literature; and (2) to draw some research gaps, trends, and challenges that are worth investigating in the context of detection and classification of plant diseases. A systematic mapping of the literature was carried out in pairs, following well-established practice guidelines (Petersen et al., 2015; Kitchenham et al., 2010). In total, 56 primary studies were carefully selected from a sample of 668 potentially representative studies, which were retrieved from 9 widely recognized electronic databases. The primary studies were analyzed and categorized to answer seven research questions. In particular, we try to understand little-known issues, mainly concerning (1) the technologies applied to detect plant diseases, (2) the sensors used to collect data from plants and agricultural environments, (3) the machine learning techniques applied for detecting diseases, (4) the main crops and their diseases explored in the literature, (5) the main contributions reported by the primary studies, (6) the research methods used by the primary studies, and (7) the venues of research where primary studies were published. Investigating these seven issues can be seen as an initial effort for a more ambitious agenda on how to characterize and improve techniques (and their tool support) for detecting plant diseases.

The results show that 41% of primary studies applied machine learning techniques to detect diseases, 32% used image sensors to identify symptoms related to plant diseases, 30% focused on proposing new models of machine learning to detect diseases, 34% were evaluation studies, and 71% were published in scientific journals. The association between computer vision and neural networks appears as a promising field of research for the detection of diseases.

The remainder of this paper is detailed in the following sections. Section 2 presents basic concepts to grasp this work. Section 3 presents the review protocol. Section 4 describes the filtering process of the potentially relevant studies. Section 5 shows the study results. Section 6 outlines additional discussions, taxonomy, trends, and challenges for future research. Section 7 presents the related works. Section 8 shows some actions performed to mitigate threats to validity of this work. Section 9 presents some concluding remarks and future works.

2. Background

This section provides an overview about the key concepts concerning the detection of plant diseases. Section 2.1 describes some factors that are responsible for causing plant diseases. Section 2.2 presents a case study on grapevine caused by Downy Mildew. Section 2.3 presents the concepts of machine learning and its applications to detect diseases in plants.

2.1. Plant diseases

Plant diseases are the result of interaction between hosts, causal agent, and environment that interrupts or modifies their vital functions (Francl, 2001). The alignment of these three factors must be synchronous for a particular disease to occur. The dependence of the climatic conditions for the occurrence of any disease and in any agricultural area may be illustrated in Fig. 1. Host, pathogen, and environment are represented by each side of a triangle, where the occurrence of the disease depends on the simultaneous combination of these three factors. An example of this is the lack of favorable weather (e.g., relative humidity, air temperature, and leaf wetness duration) conditions for the development of Mildew. For example, if the infectious agent is present in the growing area, as well as the susceptible host, but the weather is not favorable for the infection and the parasitism relationship, the interaction plant-pathogen lead to an incompatible interaction for a disease (Moore, 2018).

Plant diseases are classified based on the nature of their causal agents and their respective relationship of parasitism to produce related symptoms from the plant-pathogen interactions. Hence, plant diseases can be interpreted as a type peculiar of primary stress (the biotic stress) always imposed by pathogenic microorganisms such as fungi, bacteria, viruses, and nematodes (Buchanan et al., 2015). Diseases of non-infectious nature are a consequence of adverse climatic conditions — such as extreme temperature, the relationship between moisture and oxygen — which do not favor the cultivation of healthy crops, and the deficiency or even the excess of minerals in the soil (Garrett et al., 2016). Plant diseases affect the quality of fruits and the growth of their respective species (Francl, 2001). Therefore, early identification is very important. Many machine learning (ML) models have been used for the detection and classification of plant diseases (Singh and Misra, 2017; Kaur et al., 2018; Ferentinos, 2018).

Detection of diseases investigates whether a disease is present or not and if it is present, which part of the plant/leaf is infected. The classification in turn seeks to respond when a disease (which is an infected part of a plant) is shown to a model which disease is among the possible diseases in which the model has been trained. This way, machine learning techniques can be used to detect and classify diseases (Jordan and Mitchell, 2015), but they depend on the quality of data set for recognizing patterns and extracting meaning.

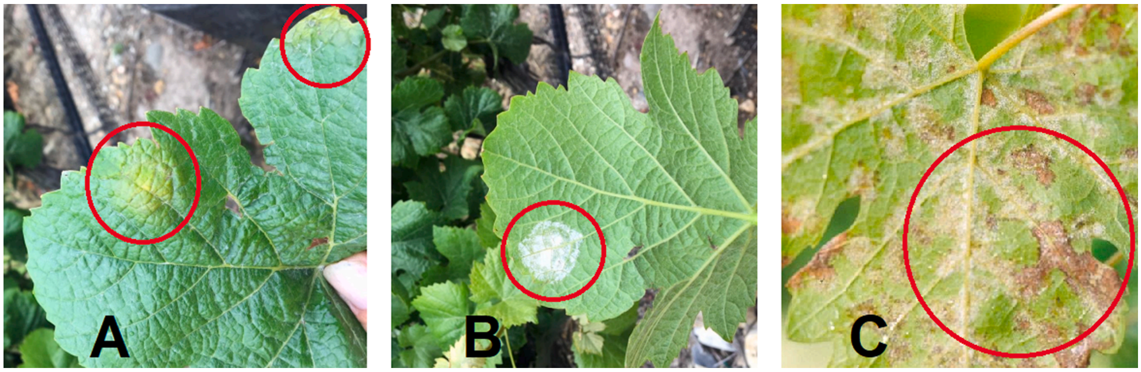

Fig. 2. Symptoms of mildew on vine leaf (A&M, 2018).

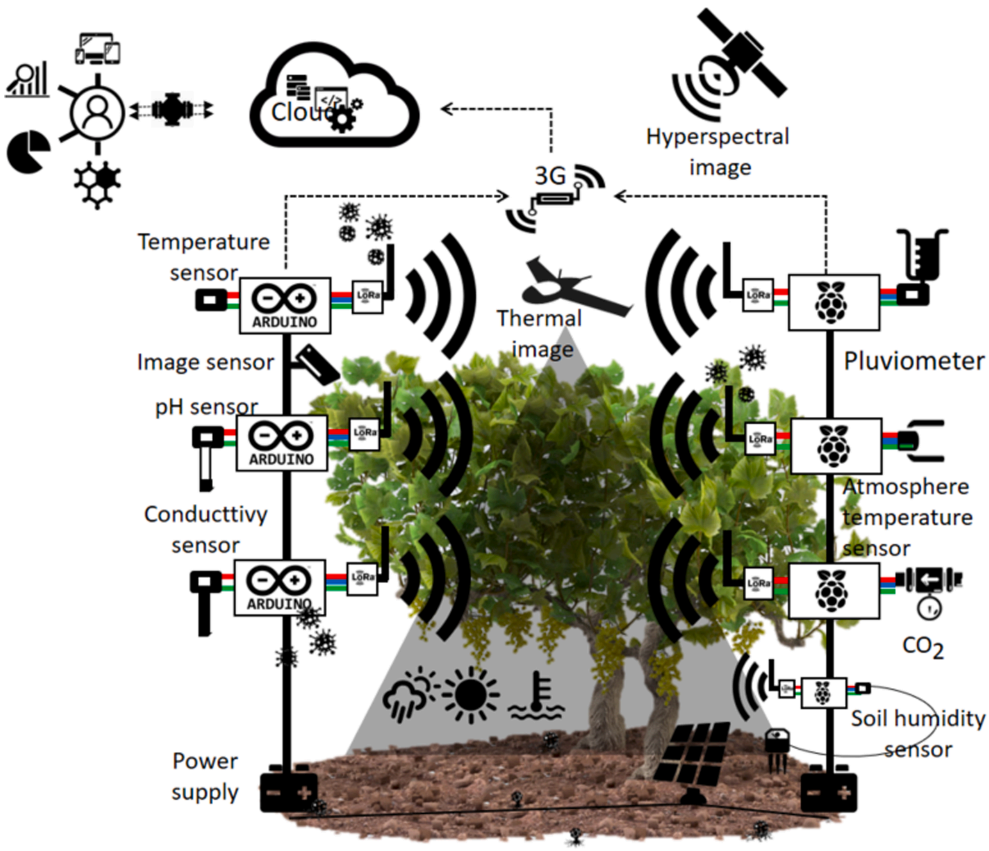

Fig. 3. An illustrative image of possible sensors that can be used in vineyards.

2.2. Grapevine downy mildew

Crops and plants are susceptible to a wide range of pathogenic microorganisms. For example, the cultivation of the grapevines (Vitis spp.) is subject to diseases worldwide, originated. Pathogens can be easily propagated thought vineyard because of the remarkable structures of dispersion (spores) lighter than clay particles. Wind and rain do favor the proliferation of diseases through the dispersion of spores by the whole crop fields, generally cultivated with a very restrict genetic diversity, leading to epidemics.

Even during favorable climatic conditions, the vine is subject to a number of fungal diseases which can lead to serious damage if they are not properly controlled. The main disease capable of completely impair grapevine yields is the Downy Mildew, caused by the oomycete (almost a fungus) *Plasmopara viticola*. The ideal temperature for downy mildew development is between 18°C and 25°C. The fungus requires free water in the tissues for at least 2 h for infection. The presence of free water, whether from rain, dew, or gutting, is indispensable for infection, with relative air humidity being above 98% necessary for sporulation. The fungus infection in leaves is due to the stomata present on the underside, stomata, and pedicels during flowering and beginning of fruiting and pedicels when the grape is already more developed (Gessler et al., 2011; Garrido et al., 2016).

Fig. 2 shows some symptoms caused by Mildew. In the leaves, at the top, there are yellowed spots on grapevine leaves, translucent against sunlight with a damp appearance, called “oil stains” (Fig. 2.A). At high relative humidity (above 95%), white sporulation appears in the lower part on the spot, since the signal refers to spore structures (zoosporangia) of the oomycete (Fig. 2.B). An affected area becomes necrotic and may cause the leaf to fall (Fig. 2.C).

The first method for identifying the grapevine downy mildew is usually done through the symptom disease recognizing and evaluation,

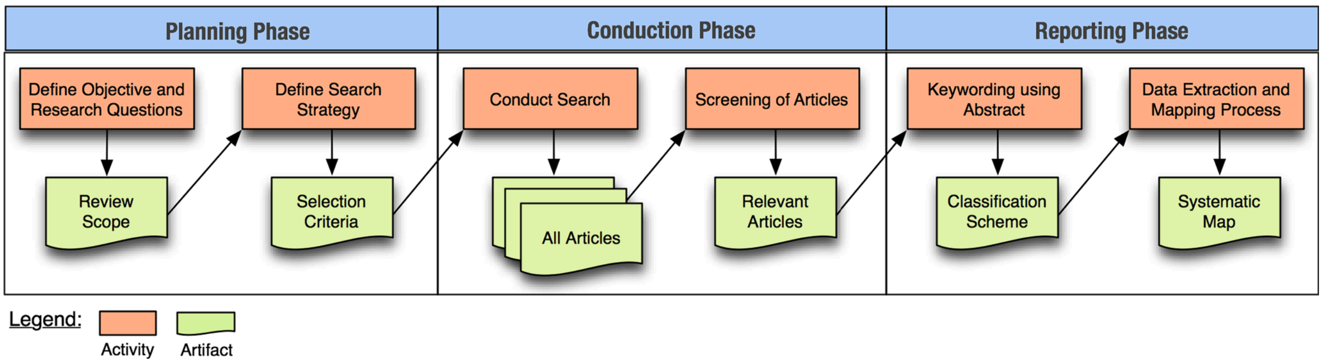

Fig. 4. Overview of the systematic mapping process (from Menzen et al. (2020)).

by the naked eye (Kaur et al., 2018). Plant diagnosis based on visual identification (symptomatology) may be a slow, arduous task that at the same time may lead to errors, because other diseases may display similar symptoms, that lead to misunderstandings. For example, the powdery mildew of grapevine (caused by the fungus *Erysiphe necator* also produces spots, but not too yellowed instead (Singh and Misra, 2017; Patrício and Rieder, 2018). This leads to production losses and also raises the costs of agricultural production (Singh and Misra, 2017). For this reason, it is essential to implement effective strategies for disease detection (Rimbaud et al., 2019) and monitoring (Sankaran et al., 2010).

2.3. *Technologies and machine learning techniques for the detection of diseases*

Fig. 3 outlines a set of electronic and digital devices for environmental monitoring of a plant disease pathosystem. This technological set includes a range of devices and sensors, as well as the use of complex approaches and machine learning techniques. It illustrates the monitoring and detection of diseases in plants through the use of sensors.

The implementation of machine learning techniques in agriculture applications has been rising in recent years (Sladojevic et al., 2016; Amara et al., 2017; Pawara et al., 2018; Ferentinos, 2018; Verma and Jain, 2019) to meet the growing demand for fast and accurate methods for precision cropping, agriculture, and food security. Machine learning deals with methods and techniques for computational applications capable of modifying or adapting their actions in order to make them more precise. It gives systems the ability to learn and automatically improve from experience without being explicitly programmed (Bottou et al., 2018).

Such techniques, combined with pattern recognition from data obtained from sensors, enable the detection of diseases in plants (Liscano et al., 2011). For example, from image sensors, it is possible to monitor the plant and generate images in real-time related to a specific disease, or to predict disease symptoms as discussed in (Fuentes et al., 2018), working with tomato root. In fact, it has been evidenced an increased number of works regarding the identification of plant disease and monitoring epidemic that applies a set of computing tools and technologies. Computer vision techniques can be also used to identify and classify the affected area by diseases (Kaur et al., 2018).

3. Systematic Mapping Study Planning

This section describes the review protocol. As previously mentioned in Section 1, we have chosen a Systematic Mapping Study (SMS) as a research method. According to (Cooper, 2016), this methodology tends to produce more reliable findings by reducing bias through a rigorous review process. This study protocol is based on well-established practice guidelines (Petersen et al., 2015; Kitchenham et al., 2010), and our experience in performing literature reviews (Menzen et al., 2020; Carbonera et al., 2020; Luz and Farias, 2020; Souza et al., 2020; Vieira and Farias, 2020; Gonçales et al., 2019; Gonçales et al., 2019; Bischoff et al., 2018; Gonçales et al., 2015). Fig. 4 shows an overview of the adopted systematic review process. Composed of three phases, in which activities are performed to create artifacts, this process serves as a guide on how to advance with the review (Menzen et al., 2020).

Table 1
Description of the investigated research questions.

Research Question	Motivation	Variable
RQ1: What technologies are being used to support disease detection and prediction?	Analyze the technologies used for detecting and predicting plant diseases.	Technological support
RQ2: What are the sensors used to collect data of plant diseases?	Compare applicability of sensors used, according to type of information collection.	Sensor types
RQ3: What machine learning (ML) techniques have been used to detect plant diseases?	Identify the techniques used for classification, localization, object detection, and segmentation in images.	Machine Learning
RQ4: What diseases and crops have been explored?	Discover and understand the diseases and crops that are most explored in the literature.	Crops and diseases
RQ5: What is the main contribution of the current studies related to detection of diseases?	Analyze the contributions of the studies.	Contributions
RQ6: What research methods have been used?	Identify research methods often applied in the current studies.	Research method
RQ7: Where have the articles been published?	Reveal the target venue used to publish the research results.	Research venue

The planning phase (Section 3) covers all the procedures for review design. The conduction phase (Section 4) details step by step how a sample of potentially relevant works was obtained, and how this sample was filtered to identify a set of representative works. The reporting phase (Section 5 and Section 6) focuses on reporting the results, drawing some trends, and highlighting some challenges that can be explored by the scientific community.

3.1. *Goal and Research Questions*

The article seeks to provide a comprehensive overview of the current literature, and to draw some research gaps, trends, and challenges that are worth investigating in the context of detection and classification of plant diseases (Section 5). Moreover, it proposes a taxonomy related to detecting and predicting plant diseases (Section 6). In this sense, we formulate seven research questions (RQs) to explore each facet of these goals properly. Table 1 shows a general view of our RQs, their motivations, and the explored variables.

Table 2
Description of main terms and their synonyms.

Main terms	Synonym
Plants	Vegetables OR Leafs OR Fruits
Sensors	Temperature, Humidity, Image
Detection	Prediction, Monitoring, Artificial Intelligence
Diseases	Pathologies, Infections, Contamination

Table 3
List of electronic databases.

Data Sources	Eletronic Address
1 - ACM Digital Library	https://dl.acm.org
2 - IEEE Xplore	https://ieeexplore.ieee.org
3 - Google Scholar	https://scholar.google.com.br
4 - Springer Link	https://link.springer.com/
5 - Science Direct	https://www.sciencedirect.com
6 - Frontiers Media	https://www.frontiersin.org/
7 - PLOS ONE	http://journals.plos.org/
8 - Hindawi Publishing Corporation	https://www.hindawi.com
9 - Taylor and Francis Group	https://www.tandfonline.com

3.2. Search Strategy

The next step was to define a search strategy to retrieve a representative sample of studies that could be used to answer the specified RQs. The search strategy was performed in two steps: (1) the definition of main terms or keywords, and (2) the construction of a search string (SS). To do this, we followed well-known guidelines (Barn et al., 2017; Petersen et al., 2015; Kitchenham et al., 2010).

We organized the search string according to the PIO method, i.e., Population, Intervention, and Outcome. The applied categories are specific and used for the set of keywords or derivatives. The term population refers to the target that is being investigated, i.e., plants, vegetables, leaves, and fruits. Intervention concerns to the means in which the population is treated. Two categories of intervention were identified in this work, such as the type of monitoring sensors (temperature, humidity, or image.), and dimensions (monitoring, artificial intelligence, reconnaissance by image) that enable detection. Outcome concerns to the result expected of the intervention over the population, i.e., the deceases, pathologies, infections, and contamination.

Table 2 introduces the major terms and their synonyms. The main keywords investigated in this mapping study are *plants*, *sensors*, *detection*, and *diseases*. Next, synonyms related to each major term were described. The Boolean operators AND and OR were used to compose our Search String. In particular, synonyms were grouped using the Boolean "OR", and then the major terms were group using the Boolean "AND". Thus, the obtained Search String is introduced as follows:

(plants OR vegetables OR leafs OR fruits) AND
(sensors OR temperature OR humidity OR image) AND
(monitoring OR prediction OR detection
OR artificial intelligence) AND
(diseases OR pathologies OR infections OR contamination)

After defining the search questions and the search string, the next step was to determine a search strategy.

Definition of search engines. The search engines refer to the sources of data used to find studies through the search string. For this, we used several combinations of such search terms that were formulated and applied in nine electronic databases, which have been widely adopted in previous systematic reviews of the literature, such as Shadroo and Rahmani (2018) and Patrício and Rieder (2018). In particular, search engines must be able to retrieve works related to the detection of plant diseases. Table 3 shows a list of nine electronic databases.

3.3. Inclusion and Exclusion Criteria

We present the inclusion criteria (IC) and exclusion criteria (EC) used to filter potentially relevant articles retrieved from the search engines shown in Table 3. The IC defines what criteria should be considered to include a particular work in our sample of representative articles. In contrast, the EC establishes the requirements to support the removal of works deemed inadequate. The IC sought to select studies that were:

- **IC1**: Academic works (i.e., articles, surveys, papers, master and doctoral thesis) aimed to propose techniques for detecting plant diseases, report empirical results or survey.;
- **IC2**: Works written, published or disseminated in English;
- **IC3**: Works found in scientific journals, conferences, research groups'web page or educational institutions; and
- **IC4**: Studies published from January 2008 until December 2018.

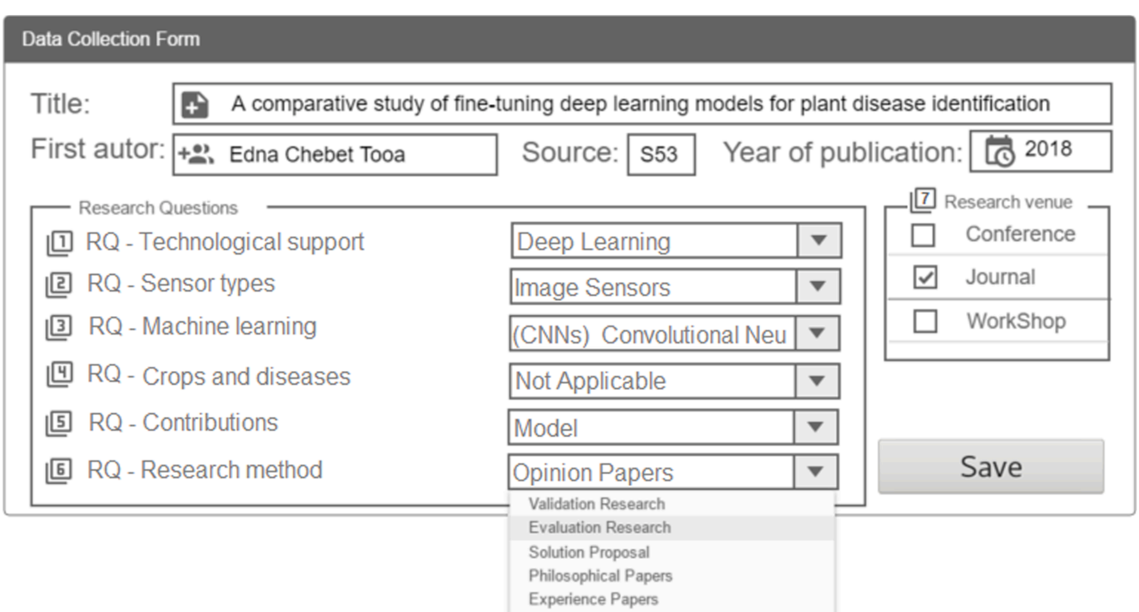

Fig. 5. An illustrative form to extract data from the selected studies.

Table 4
The classification scheme used to extract data from the studies.

Question	Variable	Answers
RQ1	Technological support	Machine Learning, Deep Learning, Statistical Inference, Genetic Testing, Generic Algorithms.
RQ2	Types of sensor	Multiple Sensors, Image Sensors.
RQ3	Machine Learning	Techniques, Categories.
RQ4	Crop and diseases	Vegetables, Fruits and Flowers.
RQ5	Contributions	Metric, Tool, Model, Method, Process.
RQ6	Research method	Validation Research, Evaluation Research, Solution Proposal, Philosophical Papers, Opinion Papers, Experience Papers.
RQ7	Research venue	Journal, Conference, Workshop.

Table 5
Category of technology.

Category	Description
Machine Learning	It uses algorithms to collect data, learn from them, and then to make a determination or prediction.
Deep Learning	It allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction.
Statistical Inference	It is the utilization of mathematical methods, models or algorithms to perform an analysis from the data abstraction.
Genetic Testing	It consists of techniques for extracting and classifying DNA molecules from plant cells for disease detection.
Generic Algorithms	Process of partial or total development of a software, through a set of activities for building a data structure. For example, forecasting temporal data, mathematical or statistical models.

Next, the EC, in turn, sought to throw away works that:

- **EC1**: The title, abstract or even their content was closely related to our search string, however without any semantic interplay;
- **EC2**: Works Were not published in English, were patent, or might be considered as an initial stage, typically represented by abstract and summary;
- **EC3**: No similarity with the research theme, or even the focal aim was completely contrary to the purpose of the issues addressed in the research questions;
- **EC4**: No aspect of the research questions was found in the abstract;
- **EC5**: Study was a duplicate;
- **EC6**: Studyt did not address issues about detection of diseases.

3.4. Data Extraction Procedures

After establishing the inclusion and exclusion criteria, the next step was to define how to extract data from potentially relevant articles. For this, we defined a classification scheme and a data extraction form (Fig. 5). Both were used to gather data so that the RQs might be properly answered.

Classification scheme. The proposed classification scheme, shown in Table 4, relates the RQs, the investigated variables, and the possible answers to the RQs. The answers are the possible values assumed by the explored variables (shown in Table 1).

This scheme was obtained based on the consensual understanding of the authors and considering potentially relevant studies. To reach this consensus, two meeting cycles were held. To mitigate threats to validity related to the elaborated scheme, we looked for previously validated schemes to take them as a basis, such as one proposed in Barn et al. (2017). We briefly describe each answer presented in Table 4.

Technological support (RQ1). For a better understanding of the technology used, we proposed a classification scheme, shown in Table 5. The categories are subdivided into six parts: (1) Machine Learning, (2) Deep Learning, (3) Statistical Inference, (4) Genetic Testing, and (5) Genetic Algorithms. In Table 5, we present the categories and their description.

Table 6
Definition of contributions.

Category	Description
Metric	It covers studies on general statistics and performance measures. Metrics can be direct and indirect and is tracked through attributes.
Tool	It is intended to develop a software, program or application.
Model	They describe the behavior of algorithms, allowing a better understanding and analysis of the produced data.
Process	They are studies that determine a set of actions and operations to be followed to reach an objective goal. That is, a method defines a process.
Method	It is planning rules and practices used to develop a technique.

Sensor types (RQ2). The sensors are responsible for monitoring, measuring, and collecting date of plants and climate environments. We define two categories of sensors: (1) Multiple sensors are devices composed by a range of single sensors, which are able to measure environmental agroclimatic features. A sensor converts a physical parameter (e.g., air temperature, relative humidity (%), air velocity, etc) into a signal that can be measured electronically; and (2) Image sensors are devices capable of capturing plant information from images. They typically capture light and electromagnetic waves, observing their variations in order to associate them with specific features or symptoms of plants.

Machine Learning (RQ3). We classify the categories and techniques as follows: (1) supervised learning techniques are performed from a previously defined set of labeled data where you want to find a function that is able to predict unknown labels. The labels are classified into "regression" and "classification" problems. In a regression problem, it is used to predict the results in a continuous output, which means that it seeks to map input variables to a continuous function (Hoo-Chang et al., 2016). For example, given a male/female image, we must predict the age based on the image data. The classification seeks to predict the results in a discrete output, being used to map input variables into different categories (Szegedy et al., 2016). For example, given a cancerous tumor, we try to predict whether it is benign or malignant through its size and age; and (2) in unsupervised learning techniques the data set used does not have any type of label. The purpose of this type of learning is to discover similarities between the analyzed objects in order to detect similarities and anomalies (DeGroff and Neelakanta, 2018; Nakasima-López et al., 2018).

In the computer vision domain, these techniques (Detection, Recognition, and Classification) help machines to understand and identify objects and environments in real-time with the help of digital images as inputs. Detection: The process of discovering specific objects in images, e.g. finding a specific disease on a leaf from the image that contains it. Recognition: The process of taking an image and recognizing the resources or structures of the image, e.g. if we have apple and potato, we can recognize their features or characteristics, and after that, we can classify them. Classification: The image classification process based on recognized features or characteristics is known as classification, e.g. after knowing or recognizing the features or characteristics of the apple and potato, we can classify them in different classes, that is, fruit and stem.

Crops and diseases (RQ4). We seek to identify which is the target of detection of the selected works with greater relevance. That is, what are the main crops and the recognition of which diseases affect the plantations. From this verified data, it will be possible to visualize the scenario that presents greater relevance in the agricultural sector. For this, we classified the crops into three categories: fruits, flowers, and vegetables (Kaur et al., 2018).

Main contributions (RQ5). Table 6 presents a definition that can be found in the explored studies. The contributions of the works are classified in metrics, tools, models, processes, and methods, as proposed in Petersen et al. (2008), Petersen et al. (2015). Thus, we try to categorize in a simple way the main scientific contribution of the selected primary studies.

Table 7
Research method Kitchenham et al. (2010).

Research method	Description
Validation Research	Techniques researched are new and have not yet been performed in practice. Techniques used are for example experimental, i.e., work done in the laboratory.
Evaluation Research	Techniques are executed in practice and an evaluation of the technique is fulfilled. That means, it is shown how the technique is implemented in practice (solution implementation) and what are the consequences of the execution in terms of advantages and disadvantages, (implementation evaluation). This also covers to identify problems in the industry.
Solution Proposal	A solution for a problem is proposed, the solution can be either new or an extension of an existing technique. The potential advantages and applicability of the solution are shown by a small example or a line of argumentation.
Philosophical Papers	These papers demonstrate a new mode of seeing at existing things by structuring the field in the form of taxonomy or conceptual framework.
Opinion Papers	These papers manifest the personal opinion of somebody whether a certain technique is positive or negative. They do not rely on related work and research methodologies.
Experience Papers	Experience papers explain on what and how something has been done in practice (Know-how). It has to be the personal experience of the author.

Research method (RQ6). This issue provides an overview of the direction of the current studies, i.e., the type of studies being produced. They seek to cover the whole process of research, from understanding the philosophical theory that supports the selection of a method, through the choice of methods used to answer problem solutions, research questions, data collection, and an analysis of these. We used the categories proposed in Kitchenham et al. (2010) to classify the selected works. Thus, we classify the primary studies according to Table 7.

Research venue (RQ7). Scientific production is presented to researchers, teachers, and the general public through newspapers, symposiums, seminars, conferences, workshops, etc. The evaluation of these articles occurs peer-review. Thus, giving total transparency and credibility to the information published.

We classified the selected primary studies into three categories: (1) Journal, which can be complete papers, published in scientific journals, and specialized in a particular area of study; and (2) conference, which can be annals, symposium, and seminars. In this format, full papers, short papers, or posters can be published. The main objective of the conference is to create a debate on topics presented. The main characteristic is the presentation of a speaker (researcher) followed by a discussion with the audience; and finally, (3) Workshop, which can be short papers and posters with training characters. Its purpose is to deepen the discussion on specific topics and, for this, presents practical cases.

Data extraction form. The data extraction process consists to perform a reading of each selected study. This procedure was performed in three review cycles, with all authors to avoid false positives or false negatives, and to cover important open questions. Finally, the data are stored extracted on a spreadsheet. For this, the data extraction form, shown in Fig. 5, was used. This form is based on Bischoff et al. (2018) and served as a template for easing the data synthesis, enabled us carefully to obtain data and generate qualitative indicators, as well as plot evidence about the formulated research questions in Table 4. The following section presents the procedures adopted to filter potential candidate studies.

4. Filtering Process

This filtering consists of nine stages, in which exclusion criteria are applied. Fig. 6 illustrates this process. Each step of the study filtering

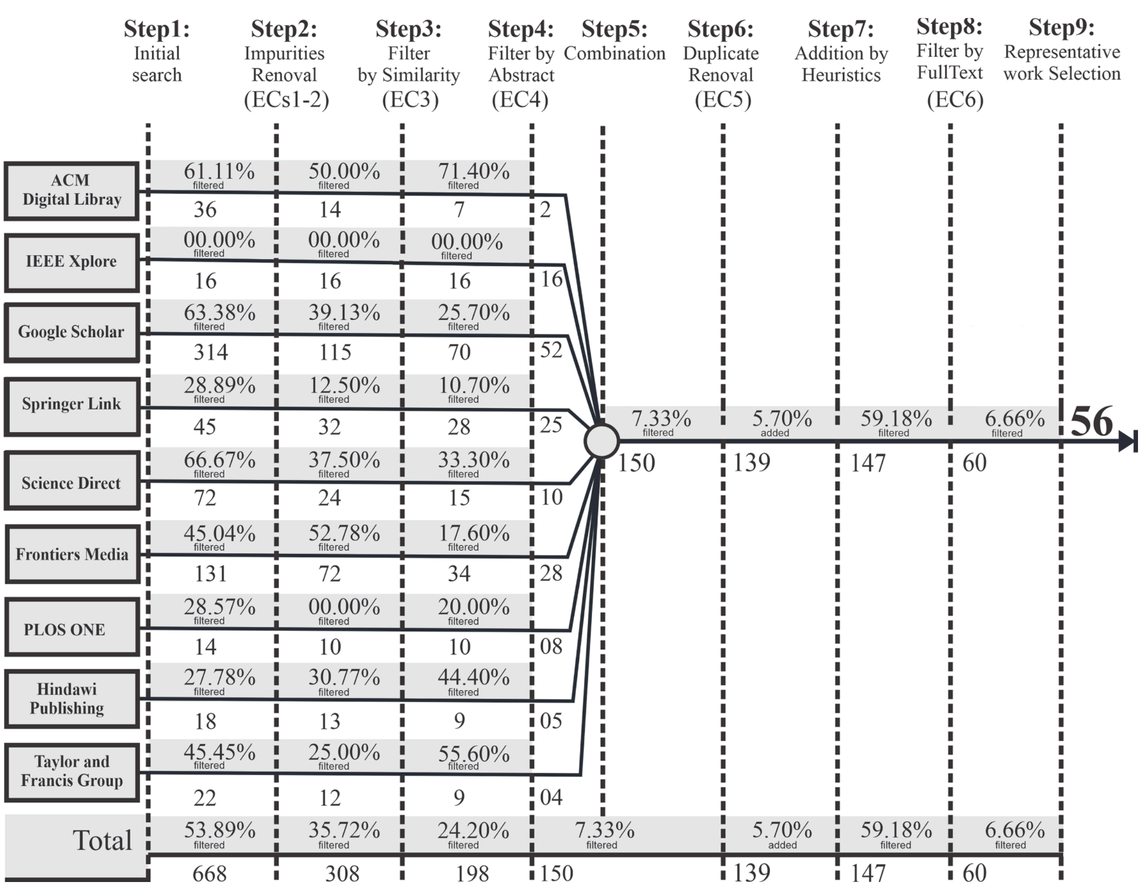

Fig. 6. The process of primary study selection.

Table 8
Classification of the primary studies based on their technological support (RQ1).

Technological Support	#Articles	Percentage	Article ID
Generic Algorithms	25	45%	[S01], [S02], [S04], [S05], [S07], [S08], [S09], [S20], [S21], [S22], [S23], [S24], [S26], [S27], [S32], [S39], [S43], [S44], [S45], [S46], [S49], [S50], [S51], [S52], [S54]
Deep Learning	14	25%	[S03], [S10], [S12], [S19], [S28], [S33], [S34], [S35], [S37], [S38], [S40], [S42], [S48], [S53]
Machine Learning	9	16%	[S06], [S15], [S17], [S18], [S25], [S41], [S47], [S55], [S56]
Statistical Inference	5	9%	[S11], [S14], [S16], [S31], [S36]
Genetic Testing	3	5%	[S13], [S29], [S30]
Total	56	100%	

process is described as follows:

- *Step 1: initial search.* It brings together the initial search results, after submitting the search string to the electronic databases in Table 3. In total, 668 candidate studies were retrieved.
- *Step 2: impurity removal.* We then applied two exclusion criteria, EC1 and EC2, to remove impurities. After applying EC1 and EC2, some were thrown away due to the absence of any semantic interplay of their title. In addition, studies that were not written in English were also discarded. In total, 53.86% (i.e., 360 works) were removed or filtered, while 308 studies remained for the next stage. Examples of these works that were retrieved would be a call for papers of conferences, special issues of journals, patent specifications.
- *Step 3: filter by similarity.* This step removed all studies that did not have similarity with our search string. For this, 35.72% (308/668) of the studies were filtered, 198 studies continued in the filtering process.
- *Step 4: filter by abstract.* This step examined 198 studies based on their abstract. In total, 24.20% of the articles were filtered. It was possible to remove studies whose content was not closely related to the key issues addressed by our research questions.
- *Step 5: combination.* The remaining studies were brought together to produce a sample of 150 candidate studies.
- *Step 6: duplicate removal.* Usually, a study can be found in several digital libraries. Thus, we applied EC5 to remove all duplicates, thereby ensuring the uniqueness of each study: 7.33% were removed, remaining 139 only.
- *Step 7: study addition by heuristic.* Although the search mechanisms in Table 3 are widely recognized, some works may not be retrieved. Thus, we inserted certain studies manually to our primary studies to mitigate this threat. We added 8 studies by applying heuristics and a snowballing process Wohlin et al. (2012), producing a sample of 147 studies, an increase of 5.70%. We reviewed the DBLP of some authors and the references and citations of the articles themselves.
- *Step 8: filter by full text.* After reading the full text of the remaining 147 studies, 59.18% were removed by applying the EC6, excluding studies whose contents were away from the expected issues concerning the detection of plant diseases. The following rules were applied to support our filtering process:

  Rule 1: Articles that were not directly aligned with the purpose. That is, we filtered all articles that were met by our search strings, but their content was not closely related to the purpose of this article;

  Rule 2: Articles whose size was small (up to 2 pages) were also filtered.

Table 9
Types of sensors used (RQ2).

Types of sensors	#Articles	Percentage	Article ID
Multiple sensors	33	59%	[S01], [S02], [S04], [S05], [S06], [S08], [S09], [S11], [S14], [S15], [S16], [S20], [S21], [S22], [S23], [S24], [S27], [S31], [S32], [S35], [S36], [S37], [S39], [S40], [S41], [S43], [S44], [S45], [S49], [S50], [S52], [S54], [S56]
Image sensors	18	32%	[S03], [S10], [S12], [S13], [S17], [S18], [S19], [S25], [S28], [S33], [S34], [S38], [S42], [S46], [S47], [S48], [S53], [S55]
Does not specify	5	9%	[S07], [S26], [S29], [S30], [S51]
Total	56	100%	

- *Step 9: representative work selection.* By exploring the remaining 60 studies, we observed that some were technically similar, i.e., studies produced based on previous ones, and their contributions were closely related. Thus, 6.66% were excluded. Finally, 56 works were selected as the most representative ones, hereinafter called *primary studies*, presented in Appendix A.

5. Systematic Mapping Studies Results

5.1. RQ1: What technologies are being used to support disease detection and prediction?

The RQ1 seeks to explore which technologies have been used to support the detection and prediction of plant diseases. After examining our primary studies, we noted that five technologies were commonly used for this purpose, including generic algorithms, machine learning, deep learning, statistical inference, and genetic testing. Table 8 shows the primary studies classified according to such technological support.

Genetic algorithms have received increasing attention from the scientific community, which was the most widely used technology (45%, 25/56). Next, technologies related to artificial intelligence, such as deep learning and machine learning, received high attention (41%, 23/56). Deep learning was explored in 25% of our sample (14/56), while machine learning in a slightly lower number (16%, 9/56). Most studies have explored neural networks in the field of computer vision. Statistical inference and genetic tests were the least explored technological supports, being observed only in 9% (5/56) and 5% (3/56) of our sample, respectively.

We could also complement that sensors and wireless networks have been used as a mechanism for collecting and distributing crop data, respectively. For example, [S27] employs sensors and wireless networks together to form sensor networks for crop monitoring. In addition, some studies have used neural network techniques to assist growers in detecting crop foliar problems. For example, [S18] explores the use of image processing for detecting plant diseases.

5.2. RQ2: What are the sensors used to collect data of plant diseases?

The RQ2 investigates how sensors have been used to detect plant diseases and which data sources they take into account. Appendix B shows the list of selected sensors. Table 9 shows the types of sensors classified as, multiple sensors and image sensors. Most studies used multiple sensors, representing (59%, 33/56) primary studies. The image sensors in turn represent a smaller portion (32%, 18/56) of the primary studies. Finally, (9%, 5/56) are not classified in the given question.

Table 10 shows where our primary studies applied sensors (i.e., data source) to collect data related to plant diseases. We have identified four data sources, namely Soil, Air, Water, Geolocation, and Cameras. In

Table 10
Description about where sensors were applied to collect data related to plant diseases (RQ2).

Data source	Article ID
Soil, air, water, geolocation	[S01], [S06], [S32], [S39], [S52]
Soil, water, geolocation	[S02]
Geolocation	[S04], [S05], [S09], [S54]
Air, geolocation	[S08]
Cameras (multiples, for example: optical, satellite, spectral, hyperspectral, multi-spectral)	[S11], [S14], [S15], [S24], [S35], [S36], [S37], [S44], [S56]
Soil, air	[S16]
Soil, water	[S20]
Geolocation, cameras	[S21], [S22]
Soil, air, water, cameras	[S23]
Soil, air, geolocation	[S27]
Air, cameras	[S31], [S40], [S45]
Air	[S41], [S43]
Soil, cameras	[S49]
Air, water, geolocation	[S50]

Table 11
Classification of the selected primary studies according to machine learning techniques used (RQ3).

Machine Learning	#Articles	Percentage	Article ID
Convolutional Neural Networks (CNNs)	11	48%	[S3], [S10], [S12], [S19], [S28], [S33], [S34], [S35], [S38], [S48], [S53]
Support Vector Machine (SVM)	3	13%	[S17], [S18], [S41]
Gradient Boosting Machine (GBM)	1	4,3%	[S6]
Neural Network (NN)	1	4,3%	[S15]
Multi-algorithms of Machine Learning	7	30,4%	[S25], [S37], [S40], [S42], [S47], [S55], [S56]
Total	23	100%	

particular, different types of cameras were used, from optical to satellite, spectral, hyper-spectral, and multi-spectral ones. Most studies (81.81%, 27/33) using multiple sensors made use of at least two data sources, while a smaller portion used Air (i.e., [S41], [S43]) or geolocation (i.e., [S04], [S05], [S09], [S54]) only.

A broad spectrum of sensors was used, not to satisfy generic purposes, but to support peculiar investigations in the contrasting sources of data. For example, [S01] applied sensors related to temperature, pH, and humidity to determine maturity in vineyard strains. [S28] used infrared spectroscopy sensors to identify diseases in plants. Finally, we realize that the variety and amount of data produced by sensors challenge current software technologies and architecture, particularly as it relates to their ability to generate accurate diagnoses of diseases in real-time.

5.3. RQ3: What machine learning (ML) techniques have been used to detect plant diseases?

The RQ3 aims to classify which machine learning techniques were used to detect plant diseases. Table 11 presents the classification of the primary studies concerning these questions. The collected data indicate that 41% (23/56) of our primary studies benefited from machine learning algorithms to classify plant diseases. This means that there is still a wide research avenue to be explored as the application of machine learning techniques to detect plant diseases.

Convolutional Neural Networks (CNNs) and Support Vector Machine (SVM) were by far the most adopted techniques, which were explored in 48% (11/23) and 13% (3/23) of the selected studies, respectively. Both Gradient Boosting Machine (GBM) and Neural Network (NN) were examined in just one work, i.e., [S6] and [S15], respectively. Moreover,

Table 12
List of machine learning techniques used by primary studies.

Studies	List of Machine Learning Algorithms
Kaur, Pandey, and Goel, 2018 [S25]	Convolutional Neural Networks (CNNs), Fully Connected Network (FCN), Support Vector Machine (SVM)
Park et al., 2018 [S37]	Convolutional Neural Networks (CNNs), Fully Connected Network (FCN), Deep Neural Network (DNN)
Picon et al., 2018 [S40]	Convolutional Neural Networks (CNNs), Fully Convolutional Networks (FCN)
Pukkela and Borra, 2018 [S42]	Support Vector Machine (SVM), Artificial Neural Networks (ANNs), Back Propagation Neural Network (BPNN), Generalized Regression Neural Network (GRNN)
Shanmuganathan, Sallis, and Narayanan, 2010 [S47]	Artificial Neural Network (ANN), Self-Organising Map (SOM)
Yalcin, and Razavi, 2018 [S55]	Support Vector Machine (SVM), Convolutional Neural Networks (CNNs)
Zhang et al., 2018 [S56]	Support Vector Machine (SVM), Artificial Neural Networks (ANNs), K-Means Clustering, Decision Trees

our collected data also indicated that some studies (34.4%, 8/23) have preferred to use multiple machine learning techniques. A list of the adopted techniques can be seen in Table 12.

We analyzed the studies to reveal the motivation behind the combined use of the techniques. Performance analysis between techniques was the main objective, such as [S25] and [S37]. For example, [S37] introduced a classifier based on two machine learning techniques. It extracted features of plant leaves using Convolutional Neural Networks, and classified the state of plant leaves using Fully Connected Network. [S25] implemented Support Vector Machine to examine classification quality through Precision-Recall metrics.

5.4. RQ4: What diseases and crops have been explored?

The RQ4 seeks to discover and understand the diseases and crops that have been most explored in the literature. We classified the explored crops in our primary studies into three categories, namely fruits, flowers, and vegetables (Tables 13 and 14).

Fruits, flowers, and diseases. Our results, shown in Table 13, indicate that a great diversity of fruits and flowers were examined in the literature, representing a total of 18 crops. Grape and Apple were the crops that received prominent attention. The other plants identified were Cherry, Peach, Blueberry, Raspberry, Strawberry, Banana, Lemon, Cantaloupe, Olive, Pear, Fig, Mango, and Watermelon that represent a total of 53.5% (30/56). Together, Grape and Apple were exploited in almost 40% of our sample — the first was explored by 23.2% (13/56), while the second by 14.2% (8/55). In addition, our primary studies did not prioritize the investigation of specific diseases in these two crops. Rather, a wide spectrum of them is exploited. A total of 23 diseases were surveyed in the context of Grape (12) and Apple. Another finding that deserves to be reported would be the high number of studies (78.57%, 44/56) that concomitantly explored several crops and pathogens. Flowers received less attention (3.5%, 2/56) when compared to fruit trees (96.4%, 54/56). We have learned from literature that Grape, Apple, and Orange are among the top 5 most-produced crops Statista (2017) and their economic importance is largely reconized Fowler et al. (2009). This may explain, in part, the great attention given to such crops in recent years.

Vegetables and diseases. Table 14 shows a list of vegetables and their explored diseases. In total, 16 vegetables were explored in 38 primary studies. Soybeans, Tomatoes, and Corn stood out as the first three most explored. Together, they represented 39.47% (15/38) of our sample. The other 13 crops totaled by 60.5% (23/38), including Bell Pepper, Pumpkin (squash), Rise, Wheat, Cucumber, Cabbage, Cassava, Celery, Cotton, Eggplant, Groundnut, Onion, and Potato. Again, there

Table 13
List of fruits, flowers and diseases.

Plant common name	Plant scientific name	Disease common name and scientific name	#Articles	Percentage	Article ID
Fruits and Flowers					
Grape	*Vitis vinifera* L.	Black rot (*Guignardia bidwellii*), Esca (*Phaeomoniella chlamydospora*), Leaf blight (*Pseudocercospora vitis*), Powdrey mildew (*Plasmopara viticola*), Downy Mildew (*Hyaloperonospora brassicae*), Greasy spot (*Mycosphaerella citri*), Melanose (*Phomopsis citri*), Scab (*Sphaceloma arachidis*), Rust (*Pucciniales*), Potassium deficiency, Neofusicoccum (*Neofusicoccum parvum*), and Botrytis (*Botrytis cinerea*)	13	23%	[S12*], [S13], [S16], [S21], [S25*], [S29], [S30], [S33*], [S34*], [S42*], [S48*], [S49], [S50]
Apple	*Malus* sp.	Powdery mildew (*Erysiphe cichoracearum*), Erwinia (*Erwinia amylovora*), Apple scab (*Venturia inaequalis*), Cedar apple rust (*Gymnosporangium juniperivirginianae*), Black rot (Botryosphaeria obtusa), Mosaic (*Cucumber mosaic virus*), Rust (*Pucciniales*), Brown spot (*Cercospora arachidicola*), Alternaria leaf spot (*Alternaria alternata*), Glomerella cingulata (*Colletotrichum gloeosporioide*), and Apple green dimple (*Marssonina blotch*).	8	14%	[S12*], [S28], [S33*], [S34*], [S35], [S37], [S42*],[S48*]
Citrus	*Citrus sinensis* L.	Huanglongbing (*Candidatus Liberibacter*), Canker (*Pseudomonas syringae* pv), Scab (*Sphaceloma arachidis*), Melanose (*Phomopsis citri*), and Black spot (*Colletotrichum acutatum*)	5	9%	[S12*], [S18*], [S25*], [S33*], [S34*]
Cherry	*Prunus avium* L.	Porosity, Powdery mildew (*Podosphaera* spp.)	5	9%	[S12*], [S18*], [S25*], [S33*], [S34*]
Peach	*Prunus persica* L.	Bacterial spot (*Xanthomonas campestris*), Powdery mildew (*Podosphaera* spp.), and Rust (*Taphrina deformans*)	4	8%	[S12*], [S33*], [S34*], [S48*]
Blueberry	*Vaccinium myrtillus* L.	Not applicable (healthy)	3	5%	[S12*], [S33*], [S34*]
Raspberry	*Rubus idaeus.* L.	Not applicable (healthy)	3	5%	[S12*], [S33*], [S34*]
Strawberry	*Fragaria* spp.	Leaf scorch (*Diplocarpon earlianum*)	3	5%	[S12*], [S33*], [S34*]
Banana	*Musa* spp.	Black sigatoka (*Mycosphaerella fijensis*), and Banana speckle (*Mycosphaerella musae*)	2	3%	[S03], [S12*]
Lemon	*Citrus limon* L.	Mycosphaerella (Greasy spot), Canker (*Pseudomonas syringae* pv), Scab (*Sphaceloma arachidis*), Melanose (*Phomopsis citri*)	2	3%	[S18*], [S25*]
Cantaloupe	*Cucumis melo* var.	not applicable (healthy)	1	2%	[S12*]
Olive	*Olea europaea* L.	Leaf scorch (*Xylella fastidiosa*)	1	2%	[S10]
Pear	*Pyrus* spp.	Porosity	1	2%	[S48*]
Fig.	*Ficus carica* L.	Alternaria (*Alternaria alternata*)	1	2%	[S17*]
Mango	*Mangifera indica*	Anthracnose (*Colletotrichum* spp.)	1	2%	[S17*]
Watermelon	*Citrullus lanatus*	not applicable (healthy)	1	2%	[S12*]
Oil Palm**	*Elaeis Guineenses* Jacq.	not applicable	1	2%	[S25*]
Maple**	*Acer palmatum*	Leaf spot (*Phyllosticta minima*) and scorch (*Xanthomonas campestris*)	1	2%	[S25*]
Total			56	100%	

Legend: *studies with multiple fruits and diseases, **flowers and diseases.

Table 14
List of vegetables and diseases.

Plant common name	Plant scientific name		#Articles	Percentage	Article ID
Vegetable					
Soybean	*Glycine max* L.	Downy mildew (*Hyaloperonospora brassicae*), Frogeye leaf spot (*Cercospora sojina*), Septoria leaf Blight (*Septoria glycines*), Septoria brown spot, (*Septoria lycopersici*), Myrothecium leaf blight (*Alternaria dauci*), Corynespora leaf spot (*Corynespora cassiicola*), Rust (*Pucciniales*), SDS (Sudden death syndrome), Stem canker (*Diaporthe phaseolorum*), and Bean pod mottle (*Bean pod mottle virus*)	6	16%	[S12*], [S18*], [S25*], [S33*], [S34*], [S42*]
Tomato	*Solanum lycopersicum.* L.	Leaf scorch (*Diplocarpon earlianum*), Bacterial spot (*Xanthomonas campestris*), Early blight (*Alternaria solani*), Late blight (*Phytophthora infestans*), Septoria leaf spot (*Septoria lycopersici*), Spider mites (*Tetranychus urticae*), Mosaic virus (*Tomato mosaic virus -ToMV*), Leaf Mold (*Fulvia fulva*), Target spot (*Corynespora cassiicola*), TYLCV (*Begomovirus Fam. Geminiviridae*), Leaf curl(*Tomato yellow leaf curl virus*), Fungal late blight (*Phytophthora infestans*), and Bacterial canker (*Clavibacter michiganensis subsp*)	5	13%	[S12*], [S25*], [S33*], [S34*], [S41]
Corn	*Zea mays* L.	Cercospora leaf spot (*Cercospora zeae-maydis*), Common rust (*Puccinia sorghi*), Northern Leaf Blight (*Exserohilum turcicum*), Leaf spot (*Septoria lycopersici*), Leaf blight (*Pseudocercospora vitis*), Sheath blight (*Septoria lycopersici*), Southern leaf blight (*Septoria lycopersici*), Powdrey mildew (*Plasmopara viticola*), Downy Mildew (*Hyaloperonospora brassicae*), Rust spots (*Pucciniales*), Gray leaf spot (*Septoria lycopersici*), Curvularia leaf spot (*Curvularia* sp.), Brown patch (*Rhizoctonia solani*) and Small spot (*Septoria lycopersici*)	4	11%	[S12*], [S25*],[S33*], [S34*]
Bell Pepper	*Capsicum annuum* L.	Bacterial spot (*Xanthomonas campestris*)	3	8%	[S12*],[S33*], [S34*]
Pumpkin (squash)	*Cucurbita* spp.	Cucumber mosaic (*Cucumber mosaic virus -CMV*) and Powdery mildew (*Erysiphe cichoracearum*),	3	8%	[S12*], [S33*], [S34*]
Rice	*Oryza sativa* L.	Bacterial blight (*Xanthomonas campestris*), Blast spots (*Oryza sativa*) Brown spots (*Cercospora arachidicola*), Sheath blight (*Rhizoctonia solani*), Marrow brown spot (*Cercospora arachidicola*), and Tungro (*Rice tungro bacilliform virus*)	3	8%	[S17*], [S25*], [S42*]
Wheat	*Triticum* spp.	Stripe rust (*Puccinia striiformis Westend*), Leaf rust (*Puccinia triticina*), Powdery mildew (*Blumeria graminis*), Septoria leaf (*Septoria lycopersici*), Tan spot (*Pyrenophora tritici-repentis*), and Snow mold (*Formally fusarium nivale*)	3	8%	[S22], [S25*], [S40]
Cucumber	*Cucumis sativus.* L.	Downy mildew (*Pseudoperonospora cubensis*), Powdery mildew (*Podosphaera xanthii*), Brown spot (*Corynespora cassiicola*), Angular leaf spot (*Pseudomonas syringae*), Blight (*Didymella bryoniae*), and Anthracnose (*Colletotrichum orbiculare*)	2	5%	[S12*], [S25*]
Cabbage	*Brassica oleracea* var. capitata	Black rot (*Xanthomonas campestris*)	1	3%	[S12*]
Cassava	Manihot esculenta Crantz	Brown leaf spot (*Cercosporidium henningsii*)	1	3%	[S12*]
Celery	*Apium graveolens* L.	Early blight (*Cercospora apii*)	1	3%	[S12*]
Cotton	*Gossypium hirsutum* L.	Red spot (*Passalora vaginae*), Powdery mildew (*Erysiphe cichoracearum*), Downy mildew (*Pseudoperonospora cubensis*), Leafminer (*Liriomyza trifolii*), Myrothecium (*Myrothecium roridum*), Bacterial blight (*Xanthomonas campestris*), and Alternaria (*Alternaria alternata*)	1	3%	[S25*]
Eggplant	*Solanum melongena.* L.	not applicable (healthy)	1	3%	[S12*]
Groundnut	*Arachis hypogaea* L.	Downy mildew (*Oidium arachidis*), Early leaf spots (*Arachis hypogaea*), Late leaf spots (*Mycosphaerella arachidis*) and Cercospora (*Cercospora arachidicola*)	1	3%	[S12*]
Onion	*Allium cepa* L.	not applicable (healthy)	1	3%	[S12*]
Potato	*Solanum tuberosum.* L.	Late blight (*Phytophthora infestans*), Early blight (*Alternaria solani*) and Cucumber mosaic (*Cucumber mosaic virus -CMV*)	1	3%	[S12*], [S42*]
Total			38	100%	

Legend: *studies with multiple vegetables and diseases.

Table 15
Main contributions of the primary studies (RQ5).

Main Contributions	#Articles	Percentage	Article ID
Model	17	30%	[S02], [S06], [S11], [S12], [S16], [S17], [S19], [S20], [S28], [S33], [S34], [S35], [S38], [S46], [S48], [S53], [S55]
Metric	12	21%	[S01], [S05], [S09], [S13], [S14], [S29], [S30], [S43], [S45], [S47], [S49], [S54]
Method	12	21%	[S03], [S04], [S10], [S18], [S21], [S24], [S36], [S37], [S40], [S41], [S44], [S56]
Tool	8	14%	[S07], [S22], [S23], [S26], [S32], [S39], [S50], [S51]
Process	7	13%	[S08], [S15], [S25], [S27], [S31], [S42], [S52]
Total	56	100%	

Table 16
Classification of primary studies according to the research methods (RQ6).

Research Method	#Articles	Percentage	Article ID
Evaluation Research	19	34%	[S02], [S05], [S08], [S09], [S11], [S14], [S16], [S21], [S22], [S26], [S27], [S31], [S32], [S36], [S39], [S41], [S44], [S50], [S56]
Validation Research	14	25%	[S17], [S19], [S20], [S24], [S28], [S29], [S30], [S33], [S34], [S35], [S43], [S45], [S46], [S48]
Solution Proposal	10	18%	[S03], [S06], [S10], [S12], [S23], [S37], [S40], [S47], [S49], [S55]
Opinion Papers	9	16%	[S07], [S15], [S18], [S25], [S38], [S42], [S51], [S53], [S54]
Experience Papers	4	7%	[S01], [S04], [S13], [S52]
Philosophical Papers	0	0%	
Total	56	100%	

was no prioritization, but rather a broad investigation of diseases that affect crops. We might highlight 27 diseases studied in the context of Soybeans (10 diseases), Tomatoes (13), and Corn (14). Recent studies Abhishek et al. (2019) highlight the strategic importance of such vegetables in the world food production for animal and human nutrition. Furthermore, the difference in the number of diseases investigated is easily perceived. However, we highlight the Grapevine Downy Mildew as the pathosystem that received the most attention, perhaps due to the severe effects caused to crops. This finding corroborates with some recent studies Bois et al. (2017), Soustre et al. (2018), Raja et al. (2018) that point out that the Mildew remains the major phytosanitary threat to plantations.

5.5. RQ5: What is the main contribution of the current studies related to the detection and prediction of diseases?

The RQ5 aims to analyze the contributions of primary studies, which are classified into models, metrics, methods, tools, and processes (Petersen et al., 2008; Petersen et al., 2015). Table 15 pinpoints the contributions identified in our primary studies. The model was the main contribution presented by the primary studies (30%, 17/56). 30% of the studies (17/56) presented a model as their main contribution. Both metrics and methods were the main contributions in a slightly smaller number of studies (21%, 12/56). Tool and process were reported as a contribution in a limited number of studies, 14% (8/56) and 13% (7/56), respectively.

We would like to emphasize that papers that placed a great emphasis on the collection and analysis of data had their contributions classified as metrics. For example, Abeledo et al. (2016),Correia et al. (2017), and Fraga et al. (2014) implemented a sensor network to obtain a set of data (e.g., temperature, humidity, and others) to assist in the monitoring of diseases in crops. Planning rules and practices were often used to create methods. For example, Amara et al. (2017) proposed a method based on a deep learning approach to classify leaf diseases. Mobile applications were also used as a tool to detect plant diseases. Johannes et al. (2017) used mobile's camera to automatically diagnose diseases in plants. Tools for decision support (Pérez et al., 2017), data analysis (Kameoka et al., 2014), and data mining (Kukar et al., 2018) have been proposed to promote facets of precision agriculture, such as the connection of technological support with traditional farming practice.

5.6. RQ6: What research methods have been used?

The RQ6 is concerned with exploring the research methods that have been used in the primary studies. For this, we have classified the studies into six categories, including evaluation research, philosophical papers, experience papers, opinion papers, solution proposals, and validation research. These categories were used because their usefulness has been already demonstrated in previous studies (Petersen et al., 2008; Bischoff et al., 2018).

Table 16 presents the collected data related to the classification of the primary studies considering the adopted research methods. The results indicate that most studies (34%, 19/56) were concerned with implementing and conducting an evaluation of detection techniques, outlining advantages, and disadvantages. Most primary studies (25%, 14/56) present the investigation of new technologies, followed by laboratory tests. The solution proposals appeared in lower numbers (18%, 10/56).

New techniques to detect diseases have been proposed or extended already existing techniques, as well as it was possible to reveal the implementation of the union of a set of techniques. Another representative portion refers to the opinion articles (16%, 9/56) that express deeper insights about the state of the art, and indicate an analysis of the exposed themes. Finally, few studies (7%, 4/56) explain how the practice of their work (i.e., the authors' experience in the application of theory/practice) was carried out in practice. In this investigation, we did not find any philosophical papers, all articles present contributions in their areas.

This means that there is a wide range of work investigating the current techniques for detecting diseases (i.e., a great number of techniques coupled to sensors in the identification of diseases). However, the application of such technology in the field depends heavily on external factors, climate changes that induce the proliferation of bacteria and fungi in the early stage of diseases. Those factors have been selected to be recorded by electronic and computational devices to monitor, quantify, and model the plant diseases and their epidemiology.

5.7. RQ7: Where have the primary studies been published?

The RQ7 aims to classify the location where the primary studies were published over the years. Fig. 7 answers this research question. It presents a chronology of the publication of the selected primary articles between 2008 and 2018, highlights publication venues and publication index.

Publication Index. The publication index is represented by a grey dashed line in Fig. 7. Each published primary study is equivalent to a point in the publication index. Thus, it enables to represent the publication of studies per year. The publication index indicates that 2017 was the most productive year, whereas 2009 was the less productive year as no article was identified. We did not find studies before 2008.

Trends. On average, 28 articles were published between 2008 and 2014. This average is far below the average of publications after 2014. Note that there was a significant increase in publications after 2014. We observed three reasons that explain this growth trend: (1) The constant

Publication Type	Percentage
Conference	25.00% (14/56) Primary Studies
Journal	71.42% (40/56) Primary Studies
Workshop	3.57% (02/56) Primary Studies

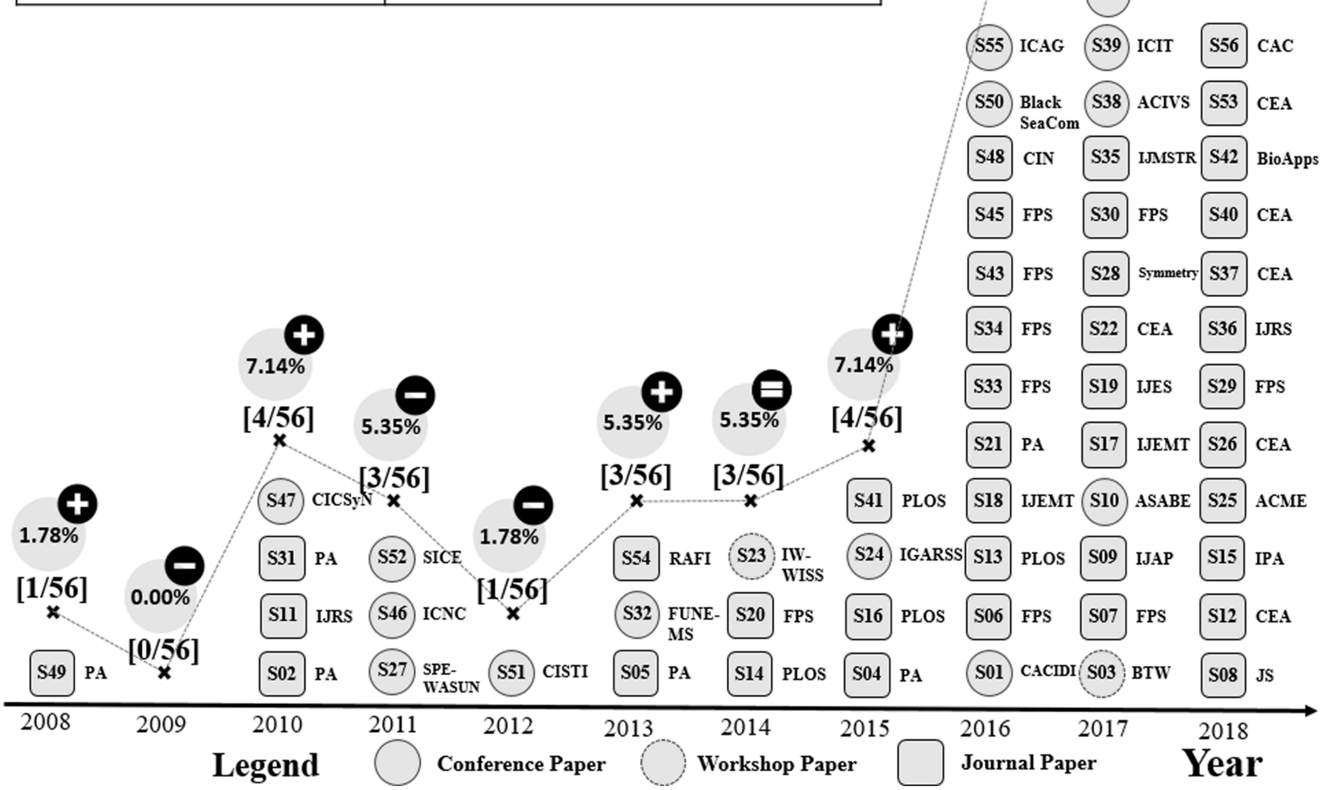

Fig. 7. Distribution of the primary studies over the past years (RQ7).

emerging problems in agriculture generated challenging research gaps in which academia and industry that are trying to address with the application of hardware and artificial intelligence; (2) The definition of a computational paradigm, named as the Internet of Things (IoT), that defined the application of technologies, such as RFID sensors, and microcomputers to analyze and collect data over all domains, including agriculture; and (3) The emerging of cheap microcomputers with enough computational capacity, such as Raspberry Pi (Pi, 2019), and Arduino (Arduino, 2019), and their respective sensors ecosystem that enables the monitoring in agriculture applications.

Venue of publication. Table 17 introduces the main research venue where the primary studies were published. We note that the Frontiers in Plant Science (16.07%, 9/56), Computers and Electronics in Agriculture (10.71%, 6/56), and Precision Agriculture (10.71%, 6/56) are the vehicles where the articles have been most commonly published. The journal PLOS ONE registered 7.14%, followed by the International Journal of Engineering & Management Technology and International Journal of Remote Sensing, each with 3.57%. The collected data suggest that there is no predominant vehicle in which researchers have prioritized the publication of their articles. In contrast, we perceive a great diversity in relation to the place of publication.

6. Discussion, Taxonomy and Challenges

6.1. Additional discussion

Fig. 8 presents a bubble chart that is formed by two x-y scatterplots. Each bubble represents the number of primary studies proportionally. Each scatterplot intersects three categories investigated, i.e., two research questions. The x-axis represents the RQ6, i.e., the research methods (right plot), and the RQ5 primary studies contributions (left). The y-axis represents the supported technology (RQ1). (Fig. 9).

Research facet. Evaluation research has been the most widely used research method (34%, 19/56), while the validation research method represents 25% (14/56), followed by the solution proposal with 18% of the cases (10/56). The remaining research methods together recorded 23% of cases (13/56), i.e., opinion paper (16%, 9/56) and experience paper (7%, 4/56). We did not locate philosophical papers in our investigation. We note that empirical studies have received more attention; however, none explores deep learning techniques.

Contributions facet. Primary studies focused on the development of models (30%, 17/56). The other contributions proposed the use of methods (21%, 12/56), application of metrics (21%, 12/56), tool development (8%, 14/56), and finally, they proposed processes (7%, 13/56) in the use of the technology investigated.

Technologies facet. Primary studies applied a wide range of algorithms (38%, 21/56) in the development of tools for detecting disease. In addition, the applications of models and methods are useful in the application of deep learning techniques (25%, 14/56) for validation research studies and solution proposals. The remaining technologies (37%, 21/56) is machine learning (16%, 9/56), statistical inference (9%, 5/56), genetic tests (5%, 3/56), and genetic algorithms (7%, 4/56).

Fig. 8 presents the proposed taxonomy. The color black represents the root of the taxonomy, while the gray color represents the

Table 17
Main research venue where the primary studies have been published.

Venue Description	Article ID
Frontiers in Plant Science	[S06], [S07], [S20], [S29], [S30], [S33], [S34], [S43], [S45]
Computers and Electronics in Agriculture	[S12], [S22], [S26], [S37], [S40], [S53]
Precision Agriculture	[S02], [S04], [S05], [S21], [S31], [S49]
PloS one	[S13], [S14], [S16], [S41]
International Journal of Engineering & Management Technology	[S17], [S18]
International Journal of Remote Sensing	[S11], [S36]
Congreso Ciencias de la Informática y Desarrollos de Investigación	[S01]
International Conference on Advanced Concepts for Intelligent Vision Systems	[S38]
International Conference Black Sea on Communications and Networking	[S50]
International Conference on Agro-Geoinformatics	[S55]
International Conference on Internet of Things	[S39]
International Conference on Natural Computation	[S46]
International Geoscience and Remote Sensing Symposium	[S24]
International Conference on Instrumentation, Control, Information Technology and System Integration	[S52]
International Conference on Advanced Video and Signal Based Surveillance	[S44]
Future Networks & Mobile Summit	[S32]
Conference Information Systems and Technologies	[S51]
International Conference on Computational Intelligence, Communication Systems and Networks	[S47]
Annual International Meeting - American Society of Agricultural and Biological Engineers	[S10]
Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, And Ubiquitous Networks	[S27]
Information Processing in Agriculture	[S15]
International Journal of Monitoring and Surveillance Technologies Research	[S35]
Archives of Computational Methods in Engineering	[S25]
Symmetry, Multidisciplinary Digital Publishing Institute	[S28]
International Journal of Antennas and Propagation	[S09]
Classification in BioApps	[S42]
Comprehensive Analytical Chemistry	[S56]
Computational Intelligence and Neuroscience	[S48]
International Journal of Engineering Science	[S19]
Journal of Sensors	[S08]
Robotics and Automation in the Food Industry	[S54]
International Workshop on Web Intelligence and Smart Sensing	[S23]
Business, Technologie und Web	[S03]

subcategories of the methodologies applied for the detection of diseases in plants. This taxonomy provides an overview of the field of research investigated through the distribution of the answers regarding our questions. The taxonomy is grouped into seven categories: (1) Supported technology, six categories were discovered for detection and prediction of plant diseases; (2) Types of sensors, the sensors used in data capture were divided into two categories. Multiple sensors and image sensors; (3) Architectures, 12 categories explored the use of hierarchical representations with neural networks; (4) Plant crops, disease detection is performed in three categories, vegetables, fruits, and flowers. There is a variety of crops, as well as a large number of diseases; (5) Contributions from the primary studies, are applied in seven categories; (6) Research methods, distributed in five categories; and (7) Research venue were classified into three categories.

6.2. Challenges for Future Research

Three main challenges were identified for further research, which were derived from our careful review process. These challenges are highlighted below.

(1) A quality model for evaluating techniques for detecting and classifying plant diseases. In recent years, many techniques for the detection of plant diseases were proposed. In particular, RQ6 pointed out that 18% of the primary studies proposed techniques for disease detection. These techniques were proposed without any validation or evaluation of their effectiveness and accuracy. Moreover, the evaluation and validation studies that were produced did not use common quality attributes to assess the techniques for disease detection. In other words, different aspects were evaluated in these approaches. Therefore, the evaluated studies do not converge to an understanding in which specific dimensions a detection technique evolved in relation to each other. The investigated studies demonstrate the need to conduct practical research, that is, to validate the models proposed by the literature. In this way, seeking to promote the results through software platforms attributed to these quality metrics to demonstrate the accuracy and effectiveness of the techniques used.

The solution for this is to propose a quality model for evaluating the techniques about the detection of plant diseases in empirical studies. However, there is a lack of a quality model in studies that applies technologies to the agriculture research field. In particular, we did not find any experimental study applied to detect diseases using a standard quality model for the evaluation of these techniques. Consequently, detection tools of plant diseases fall to evaluate their quality issues, mainly considering the precision, accuracy, recall, and f-measure metrics. Thus, proposing a quality model would be an interesting research topic that might be explored by the research community in upcoming studies.

In addition, some relevant issues should be considered as challenges for future research: (1) to propose an adaptable quality model to measure the techniques in the different agricultural scenarios; (2) to compare the results of single models in relation to multiple techniques applied together; and (3) to employ quality attributes (precision, accuracy, recall, and f-measure) to compose a model for evaluating plant disease detection techniques.

The development of low-cost machine learning models for future forecasting of microclimates in small rural properties is a challenge. The set of necessary data (labels) that infer the emergence of diseases (fungi) are precipitation (pluviometer), temperature (thermometer), and humidity (hygrometer), and some labels have constancies in their values, being predictable.

Therefore, we conclude that it will be necessary to apply different methods to each label to arrive at values closer to what has been achieved. It may not be possible to apply a single method to this data set, but multiple methods. For example, the label (precipitation), that is, the incidence of rain does not show consistency in its values, therefore it does not allow assertive predictions. Free water being a determining factor for the emergence of fungal diseases.

The set of climatic factors (labels) presented depends on an exact alignment for the development of the fungus in the plant. As a challenge to establish microclimate predictions of up to 12 h, to anticipate the application of pesticides is something of extreme importance. In order to be able to train a network of a specific fungal disease from future weather data to find the exact patterns that allow it to identify the disease and its mutations over time.

The second method refers to the classification, the use of field condition image data (that is, with changes in lighting and shading) for the creation of training data sets, as this will create robust models that can be incorporated into systems high-throughput (such as UAVs and other autonomous systems).

In addition, to implement training models under typical data storage restrictions in rural landscapes. For example, downy mildew, a fungal

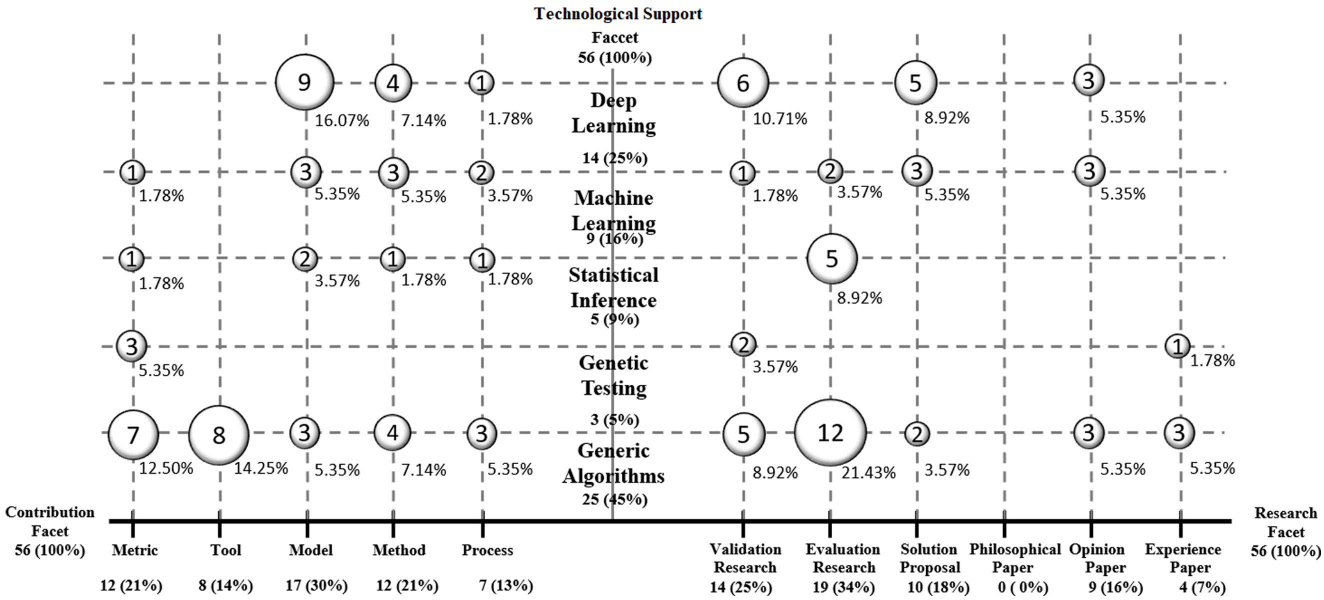

Fig. 8. A systematic map of our primary studies in the form of a bubble plot.

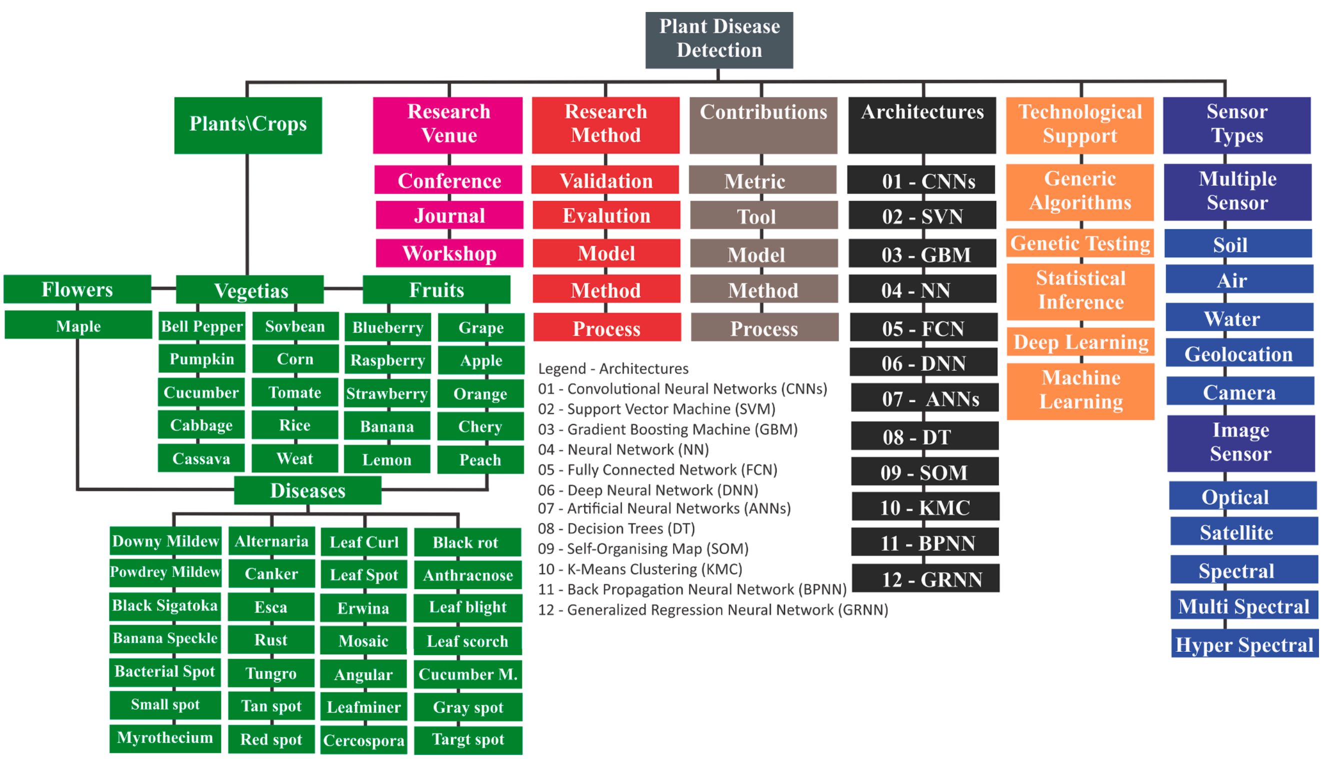

Fig. 9. Technological applications for the great areas of Plant Pathology.

disease of the vine that occurs at specific times during the flowering period. Even if it receives a protection cover against fungi (pesticide), the growth factor affects the branches and leaves branches every week on average between 10 to 15 (cm) (Garrido and Botton, 2017). Therefore, the classification systems must be designed for use in real time (monitoring the plantations through cameras), generating considerable increase in the volume of data to be processed.

(2) Empirical studies on detection of plant diseases. Most primary studies are empirical studies (59%, 33/56). In particular, many validations (25%, 14/56) and evaluation (34%, 19/56) studies were produced. This shows the agriculture research field that applies computing technologies to resolve the disease have concerned to produce empirical studies. Even so, studies are far from covering the detection of all problems related to diseases, and thus, conducting empirical studies to evaluate the effectiveness of detection techniques should still be a major focus for future studies.

It was observed that academia and the agribusiness industry are finding solutions to evaluate techniques in relation to various diseases, such as grapevine downy mildew, and rust. Specifically, among the most challenging diseases to detect are variants of mildew, downy mildew and powdery mildew (Bois et al., 2017; Soustre et al., 2018). In fact, it is important to affirm that the plant disease has an epidemiological aspect, since plants imply serious consequences such as the disposal of the infected repeated stem plant, and may even propagate throughout the

Table 18
Summary of related work.

Articles	Year	Number of Articles	Period	Objective
Mendes et al. (2020)	2020	38	2012–2020	Analyze on the use of applications (Apps) smartphones for the diagnosis and monitoring of plant diseases.
Roldán et al. (2018)	2018	91	1994–2017	Investigates the use of robotic technologies in the aid of agricultural activities.
Shadroo and Rahmani (2018)	2018	40	2010–2017	Analysis of deep learning techniques applied to agriculture.
Patrício and Rieder (2018)	2018	25	2014–2018	Application of computational vision in agriculture for the production of five types of grains.
Sibiya and Sumbwanyambe (2019)	2019	14	2010–2018	Identify the dominant frameworks in the literature for modeling disease detection systems.

Table 19
Comparative analysis of related works.

	Comparison Criteria				
Relatead Work	C1	C2	C3	C4	C5
Proposed SMS	●	●	●	●	●
Mendes et al. (2020)	◐	●	●	○	ø
Roldán et al. (2018)	◐	◐	○	○	○
Kamilaris and Prenafeta-Boldú (2018)	◐	○	●	◐	●
Patrício and Rieder (2018)	●	◐	◐	○	●
Sibiya and Sumbwanyambe (2019)	◐	○	◐	○	◐

● Meets Fully, ○ Does not answer.
◐ Meets partially, ø Not Applicable.

entire plantation causing damage and economic loss Agrios (2005). Current propagation practices favor cross-contamination by stem pathologies and cause physical-biological stress that affects the quality of the vines (Bidabadi et al., 2018).

The reduction in the yield of marketable seedlings over time, resulting from repeated cross-planting (graft) from previous harvests is common to most small farmers reusing stem seedlings coming from previous crops for a new planting (Basso et al., 2017). The recycling of planting material in this way can lead to "degeneration" considering the sequencing of the spread of diseases in the new generation of seedlings. Thus, the search for computational applications will provide data for replication of experiments, information based upon empirical data rather than relying on mere expert opinion, and specific points of threats to the validity of results.

Future empirical studies may focus on investigating the following research gaps: (1) Improve accuracy in detecting plant diseases, specifically downy mildew. Evaluate in practice the impact caused by different micro-climates in a plantation; (2) Investigate which variables are decisive for the detection of downy mildew. As well as validating the correlation of the severity caused by the disease in relation to the self-suffering suffered by these variables; (3) Another research focus to be explored refers to the combination of these variables and their connection to other types of data, using machine learning and data mining for the development of heuristics and the application of future forecasting to assist in the validation and evaluation of empirical studies.

(3) Application of edge computing in agriculture IoT systems. Edge computing is considered an evolution of cloud computing technology (Nastic et al., 2017; Shi and Dustdar, 2016). In cloud computing technology, sensors and devices constantly send data to cloud servers, where information processing is concentrated. In this way, the analyses are sent exclusively to the infrastructure that is part of the cloud for processing. Thus, this analysis is centralized in the infrastructure of cloud computing. In agriculture applications, this can be a problem because thousands of devices collect and send information to the cloud, implying high latency, increased bandwidth cost, privacy issues, and security (Nastic et al., 2017). In addition, the devices in the field are often in remote areas that hardly have limited internet access. Therefore, a device near data sources would be a natural solution to computing in agriculture. Such a device would be responsible for performing edge computing, i.e., performing computing at the edge of the Internet near the sensors. The devices that can perform edge computing are the farmer's smartphone, or microcomputer such as Arduino, or Raspberry Pi (Pi, 2019; Arduino, 2019).

A likely scenario is the monitoring of weather information for measure affordability for the incidence of plant diseases. A device near these data sources can receive real-time data from the soil, air, water, and biological sensors near the plant. This device is responsible for performing a pre-analysis and reporting the plant's vital condition. When a disease susceptibility signal is detected, both the farmer and the pesticide machines are notified by the application. Only in this situation, the data that implied on the alert is sent to the cloud for a more robust and critical analysis. This detailed analysis could help in better diagnosis of diseases, and treatment for diseases in plants. However, there are still some challenges to implement in the agricultural areas in practice on the use of information and communication technologies. The optimization of these resources and the planning of the operations required in data collection can be considered a hidden problem. This is due to the difficulty of storage and data transfer, considering a bottleneck due to the bandwidth (traffic) or its lack of certain coverage areas necessary for communication with the cloud. Thus, we expose some challenges to be investigated: (1) Which technological and infrastructure processes would be necessary to allow the effective use of computing, helping the farmer to detect diseases; (2) The use of edge computing would assist in this process; (3) It would be possible to prioritize the processing of plant data at the edge; (4) Which sensors or combination of sensors should be explored from the investigated culture and its disease; and (5) How the use of sensors can improve the accuracy of disease detection and classification.

7. Related Work

We examine and contrast the related works to the topic explored in this SMS. During the selection of primary studies, we identified similar studies, such as literature reviews and surveys. These studies were summarized in Table 18 which presents the dimensions in which related works were contrasted, i.e., the year in which articles were published, number of articles analyzed, a period of publication of surveyed articles, and the objective of the review.

7.1. Analysis of Selected Works

Mendes et al. (2020) presented a significant problem with agriculture. That is, the human being's perception (stress, experience, health, and age) in relation to the levels of precision in the diagnosis and monitoring of diseases in plants that may present different results among the same group of people. As a solution, they report the use of applications (Apps) to assist in the diagnosis of diseases. In this way, they conduct a systematic review of applications for smartphones intended for agricultural practices as a low-cost solution. The authors provide an overview of the state of the art on the type of existing application, what features they provide, and a comparison between them. They investigate 38 publications. In addition, they have some limitations in the development of mobile applications for use in precision agriculture. The processing capacity (operating time) of the mobile devices themselves is one of the main limitations. Deep Learning, Machine Learning or Artificial Intelligence algorithms and even some image processing techniques require considerable processing capacity.

Roldán et al. (2018) presented a collection of agricultural studies totaling 91 investigated studies, which addressed agricultural automation, precision agriculture, and greenhouse agriculture. The automation of tasks would be another issue also addressed, such as planting and harvesting, environmental monitoring, crop inspection, and treatment. They analyzed the use of terrestrial robots, applied to the precise treatment of plantations, and aerial used to construct field maps and detect weeds or irrigation deficits. The work investigates together greenhouse agriculture, which includes multitasking robots for monitoring environmental variables. Finally, the use of sowing and harvesting robots, which are already in place and a more advanced stage of development.

Kamilaris and Prenafeta-Boldú (2018) analyzed the technique of image processing and data analysis. The studies are applied to the challenges of agricultural production. The researchers performed a literature review, identifying 47 primary studies, of which 40 studies were done, employing deep learning techniques. The authors informed that deep-learning techniques provide high accuracy, surpassing the image-processing techniques commonly used in agricultural dominance. The authors focus on the (1) technical detail about the models used, (2) data sources used, (3) pre-processing tasks, and (4) the performance metrics used. The authors suggest the use of deep learning techniques in agricultural problems that involve classification or prediction, related to computer vision and image analysis. The authors report some limitations of the technique described in the literature. The models are not able to generalize/express beyond a certain set of data. For example, sorting individual sheets on both sides. Other factors refer to the classification of images of a disease, as it may be present on both sides of the leaf. In addition, the recognition of plants affected by environmental factors such as wrinkled surfaces, among others.

Patrício and Rieder (2018) presented a systematic review with the objective of identifying the applicability of computational vision in agriculture. The authors investigated several applications of the use of machine learning techniques, such as the processing of images and videos for the production of the five grains: corn, rice, wheat, soy, and barley. Concluding that the development of intelligent devices that use computer vision and artificial intelligence for the automation of tasks in the field is in the initial phase.

Sibiya and Sumbwanyambe (2019) presented a review of deep learning structures for the detection of foliar plant diseases. The authors identified software frameworks for supporting the modeling of detection systems of leaf diseases using machine learning techniques. In their discovery, the authors presented MATLAB as the dominant software framework in their studies.

7.2. Comparative of Selected Works

We present a comparative analysis of the related works. This comparison serves to identify similarities and differences between the proposed work and the selected literature. The comparison criteria (C) are presented below:

1. **Systematic review of the literature (C1):** It applies protocols of the systematic review of the literature.
2. **Research methods (C2):** It classifies the research methods and main contributions of the proposed works.
3. **Challenges (C3):** It discusses the future challenges of the literature.
4. **Taxonomy (C4):** It presents a schematization of the research area.
5. **Machine learning techniques (C5):** It presents a classification of the techniques of machine learning used in the detection of diseases in plants.

Table 19 presents the comparison considering these criteria. It is observed that only the proposed work fully meets the defined criteria, highlighting the contribution and the differential of this work.

8. Threats to Validity

We discuss the strategies used for mitigating some threats to validity.

Construct validity. Some works highlighted that the literature review may do not identify and include all the relevant works within the research field (Gómez et al., 2018; Pongnumkul et al., 2015). This can be caused by a mismatch between the search string and the keywords reported in the related works. As such, relevant works cannot be retrieved, or even irrelevant can be also identified. To mitigate this problematic, we carefully adopted rigorous procedures (reported in Section 3) for retrieving and filter primary studies. Search strings and their synonyms were defined according to well-established methods found in (Petersen et al., 2015; Kitchenham et al., 2010). In addition, we also included a considerable range of electronic databases, including ACM Digital Library, IEEE Xplore, Google Scholar, Springer Link, Science Direct, Frontiers Media, Plos One, Hindawi Publishing Corporation and Taylor and Francis Group.

Conclusion validity. This threat concerns on problems that can affect the reliability of our conclusions. To mitigate it, we have followed rigorously the steps provided by well disseminated systematic mapping study protocol (Gómez et al., 2018; Petersen et al., 2015; Pongnumkul et al., 2015; Kitchenham et al., 2010). As such, the conclusions outlined are derived from data produced from experimental procedures widely recognized for conducting the systematic mapping study. Finally, all the conclusions were made after collecting the results, then avoiding the fish-and-error measures.

Internal validity. Three major threats have been identified. First, it was the difficulty in establishing a relationship between the contributions surveyed, due to the various concepts (e.g., model, methods, tools, and processes and metrics). This thread was characterized as heterogeneous aspects, which exist among the evaluated contributions. We perform a careful analysis to identify common resources. The second refers to the techniques used to cover disease detection. We observe that it occurs in two distinct stages, through images applying computational vision algorithms and using different sensors to monitor the climatic conditions that interfere in the Triad of Disease (hosts, causal agent, and environment). In order to assure this process, we try to understand each technique and classify them, as described in Section 3.4. Next, it was difficult to identify the scope of each primary study. To mitigate these threats, we try to understand each technique and classify them. The filtering process was performed three times to avoid any bias. All primary studies are listed in Appendix A. In an attempt to ensure that the selection process of primary studies was as unbiased as possible, selection of primary studies was organized as a multiphasic activity, documenting the reasons for inclusion or exclusion of these studies and what were the selection criteria, as previously described in Section 3.

9. Conclusion and Future Perspectives

This work sought to understand, characterize, and synthesize current literature on the use of technology applied for the detection of diseases in plants. First, the imaging study refers to tools for diagnosis of plant disease (via symptoms) and in second, monitoring of climatic parameters serves to feed models of prediction, alert, and simulation of disease for studies at an epidemic level. To this, we performed a systematic mapping study to investigate seven research questions. We selected 56 primary studies applying a careful filtering process to a sample of 668 studies surveyed from 9 electronic databases.

We summarize the main results as follows. The primary studies 41% (23/56) applies machine learning techniques to detect diseases; 32% of studies (18/56) uses image sensors to collect information related to diseases of the plant; 48% (11/23) used the Convolutional Neural Networks (CNN) to classify diseases in plants; 23% (13/56) of the crop detect diseases in grape (Vitis vinifera L.); the most important disease in cultures\crops is downy mildew; 30% of primary studies (17/56) contributed mainly with new models of machine learning to detect diseases; 34% of primary studies (19/56) were evaluation studies, and 71% of studies (40/56) were published in scientific journals.

Based on the multiple issues listed and mapped, we understand that the key to meeting the challenge will come from a vast set of data. One of the challenges concerns the collection, organization, and preparation of this data to provide accurate predictions for complex contamination alert queries. Despite all the efforts of the software industry and the scientific community to provide accurate data on machine learning methods, this is not trivial. According to the results of this study, artificial neural networks are widely used. The training set of the networks consists of thousands of records, so as to provide accurate answers, which can be useless when new data is provided. In other words, (overfitting) is just one of the limits for current machine learning algorithms. The study points to a strong concentration on the investigation of vine culture. We believe that this fact is due to the entire productive chain involved. An important derivative of this productive chain is the wine industry. The wine industry has had a considerable drop in production over the past 2 years (OVI, 2018). This factor is due to the intense relationship between the environment and plant diseases, especially the downy mildew, as a strong threat to production. Predicting the limiting factors to climate variables, taking as a starting point the occurrence of different micro-climates in the same environment is a challenge in conjunction with new emerging technologies.

Thus, the design and implementation of a system to detect the particular stage of disease will be of great interest. Detection of early-stage disease, also known as disease prediction, can help farmers take the necessary precautions and accordingly reduce the percentage of damage to the crops.

We also observed that the selected studies identify plant diseases under specific conditions, i.e., in a controlled environment (greenhouse tests) (Santos et al., 2019). As far as we know, some studies carried out the identification of diseases in real-world scenarios with precision acceptable and economically viable to small farmers, e.g., (Fuentes et al., 2020 and Jiang et al., 2019). Thus, we propose as a future work the implementation in practice, through the development of a real-time system. Using concepts of continuous monitoring, edge computing, computer vision techniques, and prediction algorithms in a field area to identify diseases in plants.

Finally, we hope that the findings discussed throughout the paper may encourage researchers and practitioners to explore the findings. In addition, this study can be seen as the first step towards a more ambitious agenda on how to characterize and improve techniques for plant disease detection and prediction, in order to better protect crops and avoid as much as possible economic losses for small farmers and investors, especially those with restricted budgets.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was partially supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) under Grant 313285/2018–7.

Appendix A. List of Selected Primary Studies

The 56 articles selected as primary studies in the systematic mapping study are listed below.

S01 Abeledo, M. C., Bruschetti, F., Priano, D. A., Calbosa, D., Crubellier, R., Iriso, P., Abete, E., 2016. Application of wireless technology to determine optimum maturity in strains of malbec vineyards for argentine wine sectors. In: Ciencias de la Informatica y Desarrollos de Investigacio'n, IEEE Congreso Argentino de. IEEE, pp. 1–7.

S02 Acevedo-Opazo, C., Tisseyre, B., Taylor, J., Ojeda, H., Guillaume, S., 2010. A model for the spatial prediction of water status in vines (vitis vinifera l.) using high resolution ancillary information. Precision Agriculture 11 (4), 358–378.

S03 Amara, J., Bouaziz, B., Algergawy, A., et al., 2017. A deep learning based approach for banana leaf diseases classification. In: BTW (Workshops). pp. 79–88.

S04 Arnó, J., Masip, J., Rosell-Polo, J. R., et al., 2015. Influence of the scanned side of the row in terrestrial laser sensor applications in vineyards: practical consequences. Precision Agriculture 16 (2), 119–128.

S05 Arnó, J., Vallès, J. M., Llorens, J., Sanz, R., Masip, J., Palacín, J., Rosell-Polo, J. R., et al., 2013. Leaf area index estimation in vineyards using a ground-based lidar scanner. Precision agriculture 14 (3), 290–306.
S06 Brillante, L., Mathieu, O., Lévêque, J., Bois, B., 2016. Ecophysiological modeling of grapevine water stress in burgundy terroirs by a machinelearning approach. Frontiers in plant science 7, 796.
S07 Carvalho, L. C., Silva, M., Coito, J. L., Rocheta, M. P., Amâncio, S., 2017. Design of a custom RT-QPCR array for assignment of abiotic stress tolerance in traditional portuguese grapevine varieties. Frontiers in plant science 8, 1835.
S08 Castillo-Cara, M., Huaranga-Junco, E., Quispe-Montesinos, M., OrozcoBarbosa, L., Antúnez, E. A., 2018. Frog: A robust and green wireless sensor node for fog computing platforms. Journal of Sensors 2018.
S09 Correia, F. P., Alencar, M. S. d., Lopes, W. T. A., Assis, M. S. d., Leal, B. G., 2017. Propagation analysis for wireless sensor networks applied to viticulture. International Journal of Antennas and Propagation 2017.
S10 Cruz, A. C., Luvisi, A., De Bellis, L., Ampatzidis, Y., 2017. Visionbased plant disease detection system using transfer and deep learning. In: 2017 ASABE Annual International Meeting. American Society of Agricultural and Biological Engineers, p. 1.
S11 Cunha, M., Marcal, A. R., Silva, L., 2010. Very early prediction of wine yield based on satellite data from vegetation. International Journal of Remote Sensing 31 (12), 3125–3142.
S12 Ferentinos, K. P., 2018. Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture 145, 311–318.
S13 Fischer, J., Compant, S., Pierron, R. J., Gorfer, M., Jacques, A., Thines, E., Berger, H., 2016. Differing alterations of two esca associated fungi, phaeoacremonium aleophilum and phaeomoniella chlamydospora on transcriptomic level, to co-cultured vitis vinifera l. calli. PloS one 11 (9), 1–21.
S14 Fraga, H., Malheiro, A. C., Moutinho-Pereira, J., Cardoso, R. M., Soares, P. M., Cancela, J. J., Pinto, J. G., Santos, J. A., 2014. Integrated analysis of climate, soil, topography and vegetative growth in iberian viticultural regions. PloS one 9 (9).
S15 Golhani, K., Balasundram, S. K., Vadamalai, G., Pradhan, B., 2018. A review of neural networks in plant disease detection using hyperspectral data. Information Processing in Agriculture.
S16 González-Domínguez, E., Caffi, T., Ciliberti, N., Rossi, V., 2015. A mechanistic model of botrytis cinerea on grapevines that includes weather, vine growth stage, and the main infection pathways. PloS one 10 (10).
S17 Gupta, T., 2017. Plant leaf disease analysis using image processing technique with modified SVM-CS classifier. Int. J. Eng. Manag. Technol. 5, 11–17.
S18 Gupta, V. M. T., Tarun, 2016. An exploration on the identification of plant leaf diseases using image processing approach. International Journal of Engineering and Management Technology 4 (4), 47–52.
S19 Hanson, A. M. J., Joy, A., Francis, J., 2017. Plant leaf disease detection using deep learning and convolutional neural network. International Journal of Engineering Science 5324.
S20 Hofmann, M., Lux, R., Schultz, H. R., 2014. Constructing a framework for risk analyses of climate change effects on the water budget of differently sloped vineyards with a numeric simulation using the Monte Carlo method coupled to a water balance model. Frontiers in plant science 5, 645.
S21 Hou, J., Li, L., He, J., 2016. Detection of grapevine leafroll disease based on 11-index imagery and ant colony clustering algorithm. Precision agriculture 17 (4), 488–505.
S22 Johannes, A., Picon, A., Alvarez-Gila, A., Echazarra, J., RodriguezVaamonde, S., Navajas, A. D., Ortiz-Barredo, A., 2017. Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. Computers and Electronics in Agriculture 138, 200–209.
S23 Kameoka, T., Nishioka, K., Motonaga, Y., Kimura, Y., Hashimoto, A., Watanabe, N., 2014. Smart sensing in a vineyard for advanced viticultural management. In: Proceedings of the 2014 International Workshop on Web Intelligence and Smart Sensing. ACM, pp. 1–4.
S24 Karakizi, C., Karantzalos, K., 2015. Detecting and classifying vine varieties from very high resolution multispectral data. In: Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International. IEEE, pp. 3401–3404.
S25 Kaur, S., Pandey, S., Goel, S., 2018. Plants disease identification and classification through leaf images: A survey. Archives of Computational Methods in Engineering, 1–24.
S26 Kukar, M., Vračar, P.,Košir, D., Pevec, D., Bosnić, Z., et al., 2018. Agrodss: A decision support system for agriculture and farming. Computers and Electronics in Agriculture.
S27 Liscano, R., Jacoub, J. K., Dersingh, A., Zheng, J., Helmer, M., Elliott, C., Najafizadeh, A., 2011. Network performance of a wireless sensor network for temperature monitoring in vineyards. In: Proceedings of the 8th ACM Symposium on Performance evaluation of wireless ad hoc, sensor, and ubiquitous networks. ACM, pp. 125–130.
S28 Liu, B., Zhang, Y., He, D., Li, Y., 2017. Identification of apple leaf diseases based on deep convolutional neural networks. Symmetry 10 (1), 11.
S29 Liu, Y., Lan, X., Yin, L., Dry, I. B., Xiang, J., Lu, J., 2018. In planta functional analysis and subcellular localization of the oomycete pathogen plasmopara viticola candidate rxlr effector repertoire. Frontiers in plant science 9, 286.
S30 Massonnet, M., Figueroa-Balderas, R., Galarneau, E. R., Miki, S., Lawrence, D. P., Sun, Q., Wallis, C. M., Baumgartner, K., Cantu, D., 2017. Neofusicoccum parvum colonization of the grapevine woody stem triggers asynchronous host responses at the site of infection and in the leaves. Frontiers in plant science 8, 1117.
S31 Mazzetto, F., Calcante, A., Mena, A., Vercesi, A., 2010. Integration of optical and analogue sensors for monitoring canopy health and vigour in precision viticulture. Precision Agriculture 11 (6), 636–649.
S32 Medela, A., Cendón, B., Gonzalez, L., Crespo, R., Nevares, I., 2013. IOT multiplataform networking to monitor and control wineries and vineyards. In: Future Network and Mobile Summit, 2013. IEEE, pp. 1–10.
S33 Mohanty, S. P., Hughes, D., Salathe, M., 2016. Inference of plant diseases from leaf images through deep learning. Front. Plant Sci 7, 1419.

S34 Mohanty, S. P., Hughes, D. P., Salathe, M., 2016. Using deep learning for image-based plant disease detection. Frontiers in plant science 7, 1419.
S35 Nachtigall, L. G., Araujo, R. M., Nachtigall, G. R., 2017. Use of images of leaves and fruits of apple trees for automatic identification of symptoms of diseases and nutritional disorders. International Journal of Monitoring and Surveillance Technologies Research (IJMSTR) 5 (2), 1–14.
S36 Pádua, L., Marques, P., Hruška, J., Adão, T., Bessa, J., Sousa, A., Peres, E., Morais, R., Sousa, J. J., 2018. Vineyard properties extraction combining UAS-based RGB imagery with elevation data. International Journal of Remote Sensing, 1–25.
S37 Park, K., Ki Hong, Y., Hwan Kim, G., Lee, J., 2018. Classification of apple leaf conditions in hyper-spectral images for diagnosis of marssonina blotch using mRMR and deep neural network. Computers and Electronics in Agriculture 148, 179–187.
S38 Pawara, P., Okafor, E., Schomaker, L., Wiering, M., 2017. Data augmentation for plant classification. In: International Conference on Advanced Concepts for Intelligent Vision Systems. Springer, pp. 615–626.
S39 Pérez-Expósito, J. P., Fernández-Caramés, T. M., Fraga-Lamas, P., Castedo, L., 2017. An IOT monitoring system for precision viticulture. In: Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data (SmartData), 2017 IEEE International Conference on. IEEE, pp. 662–669.
S40. Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J., Johannes, A., 2018. Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. Computers and Electronics in Agriculture.
S41. Prince, G., Clarkson, J. P., Rajpoot, N. M., et al., 2015. Automatic detection of diseased tomato plants using thermal and stereo visible light images. PloS one 10 (4).
S42 Pukkela, P., Borra, S., 2018. Machine learning based plant leaf disease detection and severity assessment techniques: State-of-the-art. In: Classification in BioApps. Springer, pp. 199–226.
S43 Ranty, B., Aldon, D., Cotelle, V., Galaud, J.-P., Thuleau, P., Mazars, C., 2016. Calcium sensors as key hubs in plant responses to biotic and abiotic stresses. Frontiers in plant science 7, 327.
S44 Rilling, S., Nielsen, M., Milella, A., Jestel, C., Fröhlich, P., Reina, G., 2017. A multi-sensor platform for comprehensive detection of crop status: Results from two case studies. In: Advanced Video and Signal Based Surveillance (AVSS), 2017 14th IEEE International Conference on. IEEE, pp. 1–6.
S45 Sancho-Knapik, D., Medrano, H., Peguero-Pina, J. J., Mencuccini, M., Fariñas, M. D., Álvarez-Arenas, T. G., Gil-Pelegrìn, E., 2016. The application of leaf ultrasonic resonance to vitis vinifera l. suggests the existence of a diurnal osmotic adjustment subjected to photosynthesis. Frontiers in plant science 7, 1601.
S46 Shanmuganathan, S., Narayanan, A., Robison, N., 2011. A cellular automaton framework for within-field vineyard variance and grape production simulation. In: 2011 Seventh International Conference on Natural Computation (ICNC), Vol. 3, pp. 1430–1435.
S47 Shanmuganathan, S., Sallis, P., Narayanan, A., 2010. Data mining techniques for modelling the influence of daily extreme weather conditions on grapevine, wine quality and perennial crop yield. In: 2010 Second International Conference on Computational Intelligence, Communication Systems and Networks, pp. 90–95.
S48 Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., Stefanovic, D., 2016. Deep neural networks based recognition of plant diseases by leaf image classification. Computational intelligence and neuroscience 2016.
S49 Stoll, M., Schultz, H. R., Baecker, G., Berkelmann-Loehnertz, B., 2008. Early pathogen detection under different water status and the assessment of spray application in vineyards through the use of thermal imagery. Precision agriculture 9 (6), 407–417.
S50 Suciu, G., Fratu, O., Vulpe, A., Butca, C., Suciu, V., 2016. IOT agrometereology for viticulture disease warning. In: 2016 IEEE International Black Sea Conference on Communications and Networking. pp. 1–5.
S51 Tardío, M. A. M., Rosado, L. J. A., Bellot, G. O., 2012. Web-enabled decision support systems for precision viticulture. In: 2012 7th Iberian Conference on Information Systems and Technologies, pp. 1–6.
S52 Togami, T., Yamamoto, K., Hashimoto, A., Watanabe, N., Takata, K., Nagai, H., Kameoka, T., 2011. A wireless sensor network in a vineyard for smart viticultural management. In: SICE Annual Conference (SICE), 2011 Proceedings of. IEEE, pp. 2450–2454.
S53 Too, E. C., Yujian, L., Njuki, S., Yingchun, L., 2018. A comparative study of fine-tuning deep learning models for plant disease identification. Computers and Electronics in Agriculture.
S54 Wang, N., Li, Z., 2013. Wireless sensor networks (WSNS) in the agricultural and food industries. In: Robotics and Automation in the Food Industry, pp. 171–199.
S55 Yalcin, H., Razavi, S., 2016. Plant classification using convolutional neural networks. In: 2016 Fifth International Conference on Agro-Geoinformatics, pp. 1–5.
S56 Zhang, Y., Sun, X., Bajwa, S. G., Sivarajan, S., Nowatzki, J., Khan, M., 2018. Plant disease monitoring with vibrational spectroscopy. In: Comprehensive Analytical Chemistry. Vol. 80, pp. 227–251.

Appendix B. List of sensors found in our primary studies

Table B.20 we find a list of the main sensors presented in the articles. The Number sensor by Articles column displays a sensor totalizator per article. Finally, the last line shows the total of each sensor. The sensor legend (1 to 28) is shown below. If Applicable ●, Not Applicable ○, 1. Soil moisture, 2. Soil temperature, 3. Air humidity, 4. Air temperature, 5. Leaf moisture, 6. UV radiation, 7. Electrical conductivity of soil, 8. pH, 9. Pluviometer, 10. Inundation, 11. Pressure bomb, 12. CO2, 13. GPS, 14. Laser, 15. Network - CDMA - GSM - Wi-fi, 16. Optical image - satellite - spectral, 17. Hyperspectral image, 18. Flow of sap, 19. Multi-spectral images, 20. Cameras - smart phone, 21. Anemometer, 22. Wind speed, 23. Barometric pressure, 24. Ultrasonic sensor, 25. NFC-RFID, 26. RGB image, 27. Thermal image, 28. Vibrational spectral.

Table B.20
List of sensors.

Sensor ID																												Article ID	Number sensor per article
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28		
●	●	○	●	○	●	○	●	●	●	○	○	●	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	[S01]	9
○	○	○	○	○	○	●	○	○	○	●	○	●	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	[S02]	4
○	○	○	○	○	○	○	○	○	○	○	○	○	●	●	○	○	○	○	○	○	○	○	○	○	○	○	○	[S04]	2
○	○	○	○	○	○	○	○	○	○	○	○	○	●	●	○	○	○	○	○	○	○	○	○	○	○	○	○	[S05]	2
○	●	○	●	○	○	○	○	●	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	[S06]	4
○	○	○	○	○	○	○	○	○	○	○	●	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	[S08]	2
○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	[S09]	1
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	[S11]	1
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	[S14]	1
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	[S15]	1
●	●	●	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	[S16]	4
●	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	[S20]	3
○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	●	○	●	○	○	○	○	○	○	○	○	○	[S21]	3
○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	[S22]	2
●	●	●	●	○	●	○	○	●	○	○	○	○	○	○	○	●	●	○	○	●	●	●	○	○	○	○	○	[S23]	11
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	[S24]	1
●	●	●	●	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	[S27]	5
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	●	○	○	○	○	○	[S31]	2
●	●	●	●	●	●	○	○	●	●	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	[S32]	9
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	[S35]	1
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	[S36]	1
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	●	○	○	[S37]	2
●	●	●	●	○	○	○	○	●	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	○	○	○	○	[S39]	7
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	●	○	○	○	○	○	[S40]	2
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	[S41]	1
○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	[S43]	1
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	●	○	[S44]	2
○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	[S45]	2
●	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	[S49]	3
○	○	●	●	●	●	○	○	●	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	[S50]	6
●	●	●	●	○	○	○	○	●	○	○	○	○	○	●	○	○	○	○	○	●	○	○	○	○	○	○	○	[S52]	7
○	○	○	○	○	○	○	○	○	○	○	○	●	○	●	○	○	○	○	○	○	○	○	○	●	○	○	○	[S54]	3
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	[S56]	1
9	10	7	9	2	4	1	1	7	2	1	3	5	2	13	4	5	2	1	3	3	1	4	1	1	2	2	1	106	

References

Abeledo, C., Bruschetti, F., Priano, A., Calbosa, D., Crubellier, R., Iriso, P., Abete, E., 2016. Application of wireless technology to determine optimum maturity in strains of malbec vineyards for argentine wine sectors. In: Ciencias de la Informática y Desarrollos de Investigación (CACIDI). IEEE Congreso Argentino de. IEEE, pp. 1–7.

Abhishek, T., Richa, M., Kamlesh, M., Ram, S., Douglas, W., 2019. Estimates for world population and global food availability for global health. Academic Press, pp. 3–24.

Agrios, G., 2005. Plant pathology.

A&M, A.T., 2018. Diseases of the grapevine: Powdery mildew. https://aggie-horticulture.tamu.edu/vitwine/viticulture/viticulture-resources/viticulture-Articles/, [Accessed: 2018-10-18].

Amara, J., Bouaziz, B., Algergawy, A., 2017. A deep learning-based approach for banana leaf diseases classification. In: BTW (Workshops). pp. 79–88.

Arduino, 2019. Arduino uno rev3. https://www.arduino.cc/.

Barbedo, J., 2016. A review on the main challenges in automatic plant disease identification based on visible range images. Biosyst. Eng. 144, 52–60.

Barbedo, J., 2019. Plant disease identification from individual lesions and spots using deep learning. Biosyst. Eng. 180, 96–107.

Barn, B., Barat, S., Clark, T., 2017. Conducting systematic literature reviews and systematic mapping studies. In: Proceedings of the 10th Innovations in Software Engineering Conference. ACM, pp. 212–213.

Basso, M.F., Fajardo, T.V., Saldarelli, P., 2017. Grapevine virus diseases: economic impact and current advances in viral prospection and management. Revista Brasileira de Fruticultura 39 (1).

Bidabadi, S.S., Afazel, M., Sabbatini, P., 2018. Iranian grapevine rootstocks and hormonal effects on graft union, growth and antioxidant responses of asgari seedless grape. Horticult. Plant J. 4 (1), 16–23.

Bischoff, Vinicius, Farias, Kleinner, 2020. VitForecast: an IoT approach to predict diseases in vineyard. In: SBSI'20: XVI Brazilian Symposium on Information Systems, pp. 1–8, 6. https://doi.org/10.1145/3411564.3411584 (in press).

Bischoff, V., Farias, K., Gonçales, L., Barbosa, J., 2018. Integration of feature models: A systematic mapping study. Information and Software Technology.

Bois, B., Zito, S., Calonnec, A., 2017. Climate vs grapevine pests and diseases worldwide: the first results of a global survey. OENO ONE: Journal international des sciences de la vigne et du vin= International journal of vine and wine sciences 51 (2), 133–139.

Bottou, L., Curtis, F.E., Nocedal, J., 2018. Optimization methods for large-scale machine learning. SIAM Rev. 60 (2), 223–311.

Buchanan, B., Gruissem, W., Jones, R., 2015. Biochemistry and molecular biology of plants. John Wiley & Sons.

Carbonera, C.E., Farias, K., Bischoff, V., 2020. Software development effort estimation: a systematic mapping study. IET Software.

Cooper, I., 2016. What is a "mapping study?". J. Med. Library Assoc.: JMLA 104, 76–78.

Correia, F., Alencar, M., Lopes, W., Assis, M., Leal, B., 2017. Propagation analysis for wireless sensor networks applied to viticulture. Int. J.Antennas Propag., 2017.

DeGroff, D., Neelakanta, P.S., 2018. Neural network modeling: Statistical mechanics and cybernetic perspectives. CRC Press.

Ferentinos, K., 2018. Deep learning models for plant disease detection and diagnosis. Comput. Electron. Agric. 145, 311–318.

Fowler, G., Garrett, L., Neeley, A., Magarey, R., Borchert, D., Spears, B., 2009. Economic analysis: risk to us apple, grape, orange and pear production from the light brown apple moth, epiphyas postvittana (walker). United States Department of Agriculture: Raleigh, North Carolina.

Fraga, H., Malheiro, A., Moutinho, J., Cardoso, R., Soares, P., Cancela, J., Pinto, J., Santos, J., 2014. Integrated analysis of climate, soil, topography and vegetative growth in iberian viticultural regions. PLoS One 9 (9), e108078.

Francl, L., 2001. The disease triangle: a plant pathological paradigm revisited. The Plant Health. Instructor 10.

Fuentes, A., Yoon, S., Lee, J., Park, S., 2018. High-performance deep neural network-based tomato plant diseases and pests diagnosis system with refinement filter bank. Frontiers in plant science 9.

Fuentes, A., Yoon, S., Park, D.S., 2020. Deep learning-based techniques for plant diseases recognition in real-field scenarios. In: International Conference on Advanced Concepts for Intelligent Vision Systems. Springer, pp. 3–14.

Garrett, K., Nita, M., David, E.D.W.P., Gomez-Montano, L., Sparks, A., 2016. Plant pathogens as indicators of climate change. In: Clim. Change. Elsevier, pp. 325–338.

Garrido, L., Hoffmann, A., da Silveira, S., 2016. Produção integrada de uva para processamento: manejo de pragas e doenças. Embrapa Uva e Vinho-Livro técnico (INFOTECA-E).

Garrido, L. d. H., Botton, M., 2017. Recomendações técnicas para controlar as doenças e pragas da videira. https://ainfo.cnptia.embrapa.br/digital/bitstream/item/159114/1/Garrido-CampoNegocio-V22-N142-P68-71-2017.pdf.

Gessler, C., Pertot, I., Perazzolli, M., 2011. Plasmopara viticola: a review of knowledge on downy mildew of grapevine and effective disease management. Phytopathol. Mediterranea 50 (1), 3–44.

Gómez, M., Pratt, M.A., Molina, A., 2018. Wine tourism research: a systematic review of 20 vintages from 1995 to 2014. Curr. Issues Tour. 1–39.

Gonçales, L., Farias, K., da Silva, B., Fessler, J., 2019. Measuring the cognitive load of software developers: a systematic mapping study. In: 2019 IEEE/ACM 27th International Conference on Program Comprehension (ICPC). IEEE, pp. 42–52.

Gonçales, L.J., Farias, K., Oliveira, T.C.D., Scholl, M., 2019. Comparison of software design models: an extended systematic mapping study. ACM Comput. Surv. (CSUR) 52 (3), 1–41.

Gonçales, L.J., Farias, K., Scholl, M., Roberto Veronez, M., de Oliveira, T.C., 2015. Comparison of design models: A systematic mapping study. Int. J. Software Eng. Knowl. Eng. 25 (09n10), 1765–1769.

Hoo-Chang, S., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., Summers, R.M., 2016. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. IEEE Trans. Med. Imag. 35 (5), 1285.

Jiang, P., Chen, Y., Liu, B., He, D., Liang, C., 2019. Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. IEEE Access 7, 59069–59080.

Johannes, A., Picon, A., Alvarez, A., Echazarra, J., Rodriguez, S., Navajas, A.D., Ortiz, A., 2017. Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. Comput. Electron. Agric. 138, 200–209.

Jordan, M.I., Mitchell, T.M., 2015. Machine learning: Trends, perspectives, and prospects. Science 349 (6245), 255–260.

Kameoka, T., Nishioka, K., Motonaga, Y., Kimura, Y., Hashimoto, A., Watanabe, N., 2014. Smart sensing in a vineyard for advanced viticultural management. In: Proceedings of the 2014 International Workshop on Web Intelligence and Smart Sensing. ACM, pp. 1–4.

Kamilaris, A., Kartakoullis, A., Prenafeta-Boldú, F.X., 2017. A review on the practice of big data analysis in agriculture. Comput. Electron. Agric. 143, 23–37.

Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. Comput. Electron. Agric. 147, 70–90.

Kaur, M., Bhatia, R., 2020. Leaf disease detection and classification: A comprehensive survey. In: Proceedings of International Conference on IoT Inclusive Life (ICIIL 2019). Springer, NITTTR Chandigarh, India, pp. 291–304.

Kaur, S., Pandey, S., Goel, S., 2018. Plants disease identification and classification through leaf images: A survey. Arch. Comput. Methods Eng. 1–24.

Kitchenham, B., Pretorius, R., Budgen, D., Brereton, O.P., Turner, M., Niazi, M., Linkman, S., 2010. Systematic literature reviews in software engineering–a tertiary study. Inf. Softw. Technol. 52 (8), 792–805.

Kukar, M., Vračar, P., Košir, D., Pevec, D., Bosnić, Z., 2018. Agrodss: A decision support system for agriculture and farming. Computers and Electronics in Agriculture.

Liscano, R., Jacoub, J.K., Dersingh, A., Zheng, J., Helmer, M., Elliott, C., Najafizadeh, A., 2011. Network performance of a wireless sensor network for temperature monitoring in vineyards. In: Proceedings of the 8th ACM Symposium on Performance evaluation of wireless ad hoc, sensor, and ubiquitous networks. ACM, pp. 125–130.

Luz, M.A.d., Farias, K., 2020. The use of blockchain in financial area: A systematic mapping study. In: In: XVI Brazilian Symposium on Information Systems, pp. 1–8.

Mendes, J., Pinho, T.M., Neves dos Santos, F., Sousa, J.J., Peres, E., Boaventura-Cunha, J., Cunha, M., Morais, R., 2020. Smartphone applications targeting precision agriculture practices—a systematic review. Agronomy 10 (6), 855.

Menzen, J.P., Farias, K., Bischoff, V., 2020. Using biometric data in software engineering: a systematic mapping study. Behav. Inform. Technol. 1–23.

Mohanty, S.P., Hughes, D.P., Salathé, M., 2016. Using deep learning for image-based plant disease detection. Front. Plant Sci. 7, 1419.

Moore, D., 2018. Guidebook to Fungi. 21st century guidebook to fungi. http://www.davidmoore.org.uk/21st_Century_Guidebook_to_Fungi_PLATINUM/Ch14_09.htm, accessed: 2018-08-13.

Nakasima-López, S., Sanchez, M.A., Castro, J.R., 2018. Big data and computational intelligence: Background, trends, challenges, and opportunities. In: Computer Science and Engineering—Theory and Applications. Springer, pp. 183–196.

Nastic, S., Rausch, T., Scekic, O., Dustdar, S., Gusev, M., Koteska, B., Kostoska, M., Jakimovski, B., Ristov, S., Prodan, R., 2017. A serverless real-time data analytics platform for edge computing. IEEE Internet Comput. 21 (4), 64–71.

OVI, G.E.V., 2018. OIV. global economic vitiviniculture ovi. http://www.oiv.int/public/medias/5681/en-communiqu-depresse-octobre-2017.pdf, accessed: 2018-06-23.

Pantazi, X.E., Moshou, D., Tamouridou, A.A., 2019. Automated leaf disease detection in different crop species through image features analysis and one class classifiers. Comput. Electron. Agric. 156, 96–104.

Patrício, D.I., Rieder, R., 2018. Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. Comput. Electron. Agric. 153, 69–81.

Pawara, S., Nawale, D., Patil, K., Mahajan, R., 2018. Early detection of pomegranate disease using machine learning and internet of things. In: 2018 3rd International Conference for Convergence in Technology (I2CT). IEEE, pp. 1–4.

Pérez, J., Fernández, T., Fraga, P., Castedo, L., 2017. An iot monitoring system for precision viticulture. In: 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). IEEE, pp. 662–669.

Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M., 2008. Systematic mapping studies in software engineering. Ease. 8, 68–77.

Petersen, K., Vakkalanka, S., Kuzniarz, L., 2015. Guidelines for conducting systematic mapping studies in software engineering: An update. Inf. Softw. Technol. 64, 1–18.

Pi, R., 2019. Raspberry pi 3 model b+. https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus/.

Pongnumkul, S., Chaovalit, P., Surasvadi, N., 2015. Applications of smartphone-based sensors in agriculture: a systematic review of research. Journal of Sensors 2015.

Prajapati, B.S., Dabhi, V.K., Prajapati, H.B., 2016. A survey on detection and classification of cotton leaf diseases. In: 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT). IEEE, pp. 2499–2506.

Raja, M.U., Mukhtar, T., Shaheen, F.A., Bodlah, I., Jamal, A., Fatima, B., Ismail, M., Shah, I., 2018. Climate change and its impact on plant health: A Pakistan's prospective. Plant Protect. 2 (02), 51–56.

Rimbaud, L., Dallot, S., Bruchou, C., Thoyer, S., Jacquot, E., Soubeyrand, S., Thébaud, G., 2019. Improving management strategies of plant diseases using sequential sensitivity analyses. Phytopathology (ja).

Roldán, J.J., del Cerro, J., Garzón-Ramos, D., Garcia-Aunon, P., Garzón, M., de León, J., Barrientos, A., 2018. Robots in agriculture: State of art and practical experiences. In: Service Robots. InTech.

Sankaran, S., Mishra, A., Ehsani, R., Davis, C., 2010. A review of advanced techniques for detecting plant diseases. Comput. Electron. Agric. 72 (1), 1–13.

Santos, U.J.L., Pessin, G., da Costa, C.A., da Rosa Righi, R., 2019. Agriprediction: A proactive internet of things model to anticipate problems and improve production in agricultural crops. Comput. Electron. Agric. 161, 202–213.

Shadroo, S., Rahmani, A.M., 2018. Systematic survey of big data and data mining in internet of things. Comput. Netw. 139, 19–47.

Shah, J.P., Prajapati, H.B., Dabhi, V.K., 2016. A survey on detection and classification of rice plant diseases. In: 2016 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC). IEEE, pp. 1–8.

Shi, W., Dustdar, S., 2016. The promise of edge computing. Computer 49 (5), 78–81.

Sibiya, M., Sumbwanyambe, M., 2019. Deep learning models for the detection of plant leaf diseases: A systematic review.

Silva, Marcelino, Valente, Marco Tulio, Terra, Ricardo, 2016. Does Technical Debt Lead to the Rejection of Pull Requests?. In: SBSI 2016: Proceedings of the XII Brazilian Symposium on Information Systems on Brazilian Symposium on Information Systems: Information Systems in the Cloud Computing Era, vol. 1, pp. 248–254 (in press).

Singh, V., Misra, A., 2017. Detection of plant leaf diseases using image segmentation and soft computing techniques. Inform. Process. Agric. 4 (1), 41–49.

Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., Stefanovic, D., 2016. Deep neural networks based recognition of plant diseases by leaf image classification. In: Computational intelligence and neuroscience 2016.

Soustre, I., Lollier, M., Schmitt, C., Perrin, M., Buvens, E., Lallemand, J., Mermet, M., Henaux, M., Thibault, C., Dembelé, D., 2018. Responses to climatic and pathogen threats differ in biodynamic and conventional vines. Scient. Rep. 8 (1), 16857.

Souza, T.V. d., Farias, K., Bischoff, V., 2020. Big data analytics applied in supply chain management: A systematic mapping study. In: XVI Brazilian Symposium on Information Systems. pp. 1–8.

Statista, 2017. Global fruit production in 2017 by variety (in million tonnes). https://www.statista.com/statistics/264001/worldwide-production-of-fruit-by-variety/.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818–2826.

Verma, V.K., Jain, T., 2019. Soft-computing-based approaches for plant leaf disease detection: machine-learning-based study. In: Applications of Image Processing and Soft Computing Systems in Agriculture. IGI Global, pp. 100–113.

Vieira, R.D., Farias, K., 2020. Usage of psychophysiological data as an improvement in the context of software engineering: A systematic mapping study. In: XVI Brazilian Symposium on Information Systems, pp. 1–8.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M., Regnell, B., Wesslén, A., 2012. Experimentation in software engineering. Springer Science & Business Media.