Artificial Intelligence in Agriculture 9 (2023) 46–60

Contents lists available at ScienceDirect

Artificial Intelligence in Agriculture

journal homepage: http://www.keaipublishing.com/en/journals/artificial-intelligence-in-agriculture/

Deep learning methods for biotic and abiotic stresses detection and classification in fruits and vegetables: State of the art and perspectives

Sèton Calmette Ariane Houetohossou [a,*], Vinasetan Ratheil Houndji [a,b], Castro Gbêmêmali Hounmenou [a], Rachidatou Sikirou [c], Romain Lucas Glele Kakaï [a]

[a] Laboratoire de Biomathématiques et d'Estimations Forestières, Faculty of Agronomic Sciences, University of Abomey-Calavi, 04 BP 1525, Cotonou, Benin
[b] Institut de Formation et de Recherche en Informatique, University of Abomey-Calavi, 01 BP 526, Cotonou, Benin
[c] Laboratoire de Défense des Cultures, Centre de Recherches Agricoles d'Agonkanmey, Institut National des Recherches Agricoles du Bénin (INRAB), 01 BP 884, Cotonou, Benin

A R T I C L E   I N F O

Article history:
Received 1 September 2022
Received in revised form 21 July 2023
Accepted 8 August 2023
Available online 10 August 2023

Keywords:
Deep learning
Prediction
Fruits
Vegetables
Stress
Agricultural yield

A B S T R A C T

Deep Learning (DL), a type of Machine Learning, has gained significant interest in many fields, including agriculture. This paper aims to shed light on deep learning techniques used in agriculture for abiotic and biotic stress detection in fruits and vegetables, their benefits, and the challenges faced by users. Scientific papers were collected from Web of Science, Scopus, Google Scholar, Springer, and Directory of Open Access Journals (DOAJ) using combinations of specific keywords such as:'Deep Learning' OR'Artificial Intelligence' in combination with fruit disease', vegetable disease', 'fruit stress', OR 'vegetable stress' following PRISMA guidelines. From the initial 818 papers identified using the keywords, 132 were reviewed after excluding books, reviews, and the irrelevant. The recovered scientific papers were from 2003 to 2022; 93 % addressed biotic stress on fruits and vegetables. The most common biotic stresses on species are fungal diseases (grey spots, brown spots, black spots, downy mildew, powdery mildew, and anthracnose). Few studies were interested in abiotic stresses (nutrient deficiency, water stress, light intensity, and heavy metal contamination). Deep Learning and Convolutional Neural Networks were the most used keywords, with GoogleNet (18.28%), ResNet50 (16.67%), and VGG16 (16.67%) as the most used architectures. Fifty-two percent of the data used to compile these models come from the fields, followed by data obtained online. Precision problems due to unbalanced classes and the small size of some databases were also analyzed. We provided the research gaps and some perspectives from the reviewed papers. Further research works are required for a deep understanding of the use of machine learning techniques in fruit and vegetable studies: collection of large datasets according to different scenarios on fruit and vegetable diseases, evaluation of the effect of climatic variability on the fruit and vegetable yield using AI methods and more abiotic stress studies.

© 2023 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

* Corresponding author.
E-mail address: ariane.houetohossou@labef-uac.org (S.C.A. Houetohossou).

https://doi.org/10.1016/j.aiia.2023.08.001
2589-7217/© 2023 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Fruits and vegetables contain dietary fiber, which helps to lower the risk of cardiovascular disease and obesity (Slavin and Lloyd, 2012). They also provide vitamins, particularly C and A, and minerals essential for human well-being. For example, 100 g of fruit contains 2 to 14.8 g of dietary fiber, 61 to 89.1 g of water, and 90 to 646 Kcal of energy (Slavin and Lloyd, 2012). Despite their importance to human life, many stress factors induce a considerable loss of their productivity. Stressors include biotic factors caused by living organisms such as viruses, bacteria, fungi, parasites, and abiotic factors, caused by non-living organisms such as light intensity, water quantity, nutrient deficiency, and other environmental and climate factors (Goncalves et al., 2005; Hao et al., 2020). For instance, early blight is one of the most common diseases on tomatoes and can cause severe yield losses and many fruit lesions (Blancard, 2012; Brahimi et al., 2017). Likewise, wilting of Solanaceae, namely tomato, pepper, potato, and eggplant, is a genuine concern for farmers; tomato wilt can cause losses up to 100% (Sikirou et al., 2015). As abiotic stress, weak or high light can significantly affect crop physiology and morphology (Hao et al., 2020). Furthermore, when irrigation is minimal, and the plants are exposed to high water stress, they will die without any possibility of recovery (Wakamori et al., 2020). Likewise, nutrient deficiency can cause many diseases that significantly affect plant yield. For instance, deficiencies in three major mineral elements of plants can be observed through blossom end rot (BER) for Calcium, green or yellow shoulder and mottled maturation for Potassium, and paler green and uniform yellow or chlorosis on leaves for nitrogen (Tran et al., 2019). When diseases are not quickly identified and treated adequately, the production of hectares of fields can be destroyed. Therefore early identification of crop diseases and stress is essential for best productivity. Common stress and disease detection approaches focus mainly on visual recognition (Ferentinos, 2018). However, not all farmers have the experience to recognize disease symptoms to apply the appropriate treatment. Then, they call on experts who are not always available. In addition, visual recognition is a time-consuming and laborious task that often fails to meet recognition accuracy requirements (Liu et al., 2020; Dutot et al., 2013). These errors lead to the abusive use of pesticides, destroying the soil and harming consumers' health. Thus, some approaches have been proposed to automate the classification of diseases. Those approaches are based on machine learning and computer vision that use handcrafted features extracted from images by experts. Thus, the learning is not fully autonomous due to the dependence on handcrafted features (Breitenreiter et al., 2015; Brahimi et al., 2017). Deep Learning is an old concept based on artificial neural networks. It was first used in 1943 when Warren McCulloch and Walter Pitts published their first mathematical and computational model of the biological neuron (McCulloch and Pitts, 1943). Deep learning has become popular in recent years thanks to three main aspects: the ever-increasing power of computers, which allows the creation and training of neural networks with many hidden layers, the availability of big data and large datasets, and the possibility of using cloud computing. The functioning of biological neurons inspires ANN. ANN consists of an input layer, one or more hidden layers, and an output layer. Deep Learning consists of more than one hidden layer. Moreover, the hidden layers are of different types, such as convolutional and pooling layers. A DL system is self-teaching, learning by filtering information through many hidden layers. Due to its ability to extract raw data directly, DL is used to overcome the limitations of the methods based on handcrafted feature extraction. Since the first DL technique was introduced, many modifications have been made to achieve various architectures. Identifying the different DL techniques used so far, their limitations, and the users' main issues is important. A certain number of papers did similar works. For instance, Santos et al. (2020) provided a brief overview of DL applications in agriculture. They defined several DL architectures, including Deep Belief Networks (DBN), Fully Convolutional Networks (FCN), and Convolutional Neural Networks (CNN). They then discussed the various applications of DL in agriculture, including disease detection (focusing on 6 papers), crop identification (4 papers), land cover (3 papers), and weed identification (3 papers). Similarly, Paul et al. (2020) proposed a review of agricultural advancement based on machine learning and computer vision techniques. They focused on the details related to image acquisition and image processing techniques. They also presented the machine learning techniques commonly used, among which DL belongs. Singh et al. (2018) also reviewed recent works in which DL concepts were used for plant stress phenotyping from digital images. The review compared DL methods to other techniques in decision accuracy, data size requirements, and applicability in various scenarios. Despite all these works, the real challenges of using DL for stress detection in fruits and vegetables are missing in the literature. In addition, none of the authors used artificial intelligence and DL to address both abiotic stress and parasitic attacks in plants. The interest in focusing on biotic and abiotic stresses in plants is crucial since plants are designed to live in an environment where these two types of stress interact (Atkinson and Urwin, 2012). For example, drought stress can expose plants to pathogens. Therefore, focusing on the two types of stress would be an excellent asset for good monitoring to improve crop yield. This paper provided an analysis of the main stresses that face fruits and vegetables, the DL methods used to overcome them, their limitations, and some perspectives.

The rest of the paper is structured as follows. Section 2 clarifies the concepts and terminologies used in this paper. The methodology is described in Section 3. Section 4 then presents the findings and their discussion with the perspectives, and Section 5 concludes the work.

2. Clarification of concepts and terminologies

This section provides the history of the concept of Deep learning and a brief definition of the different terminologies considered.

2.1. History of deep learning

Deep Learning is a subset of Machine Learning (ML), an Artificial Intelligence (AI) field. ML is a field of study that allows computers to learn without being explicitly programmed (Samuel, 1959). Deep learning is built on artificial neural networks inspired by the human brain's functioning. The concept of artificial neurons was first used in 1943 by McCulloch and Pitts, who showed that neurons could be combined to build a Turing machine (McCulloch and Pitts, 1943). In a neural network, successive layers are connected to learn concepts. The simplest networks have only two layers: an input and an output layer, each of which can have several hundred, thousands, or even millions of neurons. As the number of neurons increases, so does the network's ability to learn more and more abstract representations. Rosenblatt (1958) invented the perceptron, a formal neuron, the smallest possible neural network, whose activation function is a step function called Linear Threshold Function. The perceptron has many inputs and weights for each input. The perceptron has an algorithm that allows it to learn the weights of the information from a set of data with labels (Rosenblatt, 1958). However, Minsky and Papert showed that perceptrons have some limitations (Minsky and Papert, 1969). Thus, research on neural networks was stopped for a decade. The following important milestone was multilayer perceptrons (MLPs) with the backpropagation algorithm introduced in 1985 by Ackley to revive the field. MLPs are neural networks that aim to classify more complex data than a perceptron. To do this, the MLP examines each piece of data and updates the weight of each neuron in each layer of its network to best classify that database. Later in 1988, Neocognitron, a hierarchical neural network capable of visual pattern recognition, was proposed (Fukushima, 1988). Then, Convolutional Neural Network with Backpropagation for document analysis was developed by LeCun (LeCun et al., 1998). A so-called deep neural network has at least two hidden layers (there can be as many as desired). Deep learning has been revolutionized since the 2010s with the creation of several architectures.

2.2. Deep learning architectures

Modifications of model architecture are an essential factor in improving the performance of models. The first Deep Learning has experienced considerable modifications up to the present day. Such modifications include structural reformulation, regularization, normalization, and parameter optimizations (Alzubaidi et al., 2021). These mainly occurred due to the reorganization of the processing unit and the development of novel blocks. In particular, the most novel results were performed using network depth. Some examples of Deep Learning architecture are AlexNet, ZefNet, Visual geometry group (VGG), GoogLeNet, ResNet, Inception-V3, Inception–V4, DenseNet, ResNet, Xception, Recurrent neural networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Autoencoders, Restricted Boltzmann Machines, Deep belief networks.

2.2.1. Importance of optimization techniques

In Machine Learning, optimization is done each time a model is trained: the learning algorithm optimizes the values of the model parameters to minimize the prediction error on the training dataset. However, the hyperparameters are not optimized during training. The number of layers in a neural network, the number of neurons per layer, etc., are hyperparameters that are not optimized during training. The choice of values for the hyperparameters influences the quality of the final model, sometimes strongly. On the other hand, the manual selection of the best hyperparameter values can be complex, even for the most experienced data scientists. Thus, there are simple search algorithms, such as Grid Search or Random Search, to explore the search space and retain the best-explored solution. These simple algorithms are beneficial if the number of possible combinations of hyperparameter values is quite limited and the search space is simple. Still, their use becomes prohibitive for larger and more complex search spaces, as is the case for the most sophisticated models. The problem of optimizing the hyperparameters of a Machine Learning model is generally related to the optimization with mixed variables (discrete/continuous). Indeed, some hyperparameters can take integer values, others real values, others strings of characters. Moreover, it is only possible to efficiently compute the gradients of the model performance (e.g., the accuracy score) concerning the hyperparameters. Finally, the search space can be pretty significant and complex if the number of hyperparameters is high enough. Based on this reflection, this is an ideal application for optimization algorithms.

2.3. Plants abiotic and biotic stresses

Stress can be defined as any state of discomfort for the plant. Many stressors, including abiotic and biotic factors, can slow or inhibit plant development and growth. Biotic factors include diseases, infections, and pest attacks. Infections can be bacterial, fungal, or viral. Viral diseases are caused by viruses which are tiny micro-organisms with a genome composed of Deoxyribonucleic Acid (DNA) and Ribonucleic Acid (RNA). They reproduce inside a living organism. Filamentous parasitic fungi cause fungal diseases. These fungi are disseminated by spores or mycelium present in the soil and penetrate the plant organism through stomata or roots. They then appear as spots on fruits, leaves, or seedlings and cause wilting, leaf dieback, and root rot. In contrast, phytopathogenic bacteria cause bacterial diseases. These bacterias live as parasites on the plants and cause cankers and soft rots. Abiotic stress factors include drought (water stress), excessive watering (waterlogging), extreme temperatures (cold, frost, and heat), salinity, mineral toxicity (Verma et al., 2013), nutrient deficiencies, and soil contamination by heavy metals such as cadmium and lead. In response to stress, plants slow down their metabolism and decrease energy expenditure. Stress reduces growth, photosynthesis, and therefore yield. When a plant is stressed, it prevents its normal development, and its yield can no longer be optimal.

3. Methodology

The methodology was organized into three subsections. The first described the research strategy, the second defined the research questions, and the last presented the literature synthesis and statistical analyses.

3.1. Description of the research strategy

Scientific articles were searched in databases such as Google Scholar, Science Direct, Web of Science, Scopus, Springer, and Directory Of Open Access Journals (DOAJ). The following keywords were used: 'Deep Learning' OR 'Artificial Intelligence' in combination: 'AND' with 'fruit disease', 'vegetable disease', 'fruit stress', OR 'vegetable stress'. The last papers were assessed on July 4, 2022. No time interval was specified during the search process since Deep Learning is a relatively new technique, and we want to retrieve as many articles as possible. Therefore, 818 research papers were identified, including journal articles (original and review articles), conference papers, and book chapters. Two hundred forty-six of the documents were duplicates. Identification was followed by a screening process that included reading the title, abstract, and content of the paper to ensure that it was related to the use of DL to detect stress on fruits and or vegetables.

By screening the titles and abstracts, 363 other papers were excluded. After the content screening, we removed 50 irrelevant articles, 12 literature reviews, and 15 book chapters. Among the irrelevant articles, some address stress on species other than fruits and vegetables using artificial intelligence or Deep Learning, while others refer to stress but do not use artificial intelligence or Deep Learning methods. Before excluding them from the final selection, we checked references cited in systematic reviews, books, and articles on related topics. Finally, 132 articles were included in the study. Thirteen of the 132 articles were not free of charge. We have therefore used only their abstracts. The methodology described was inspired by the Preferred Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021) and Fig. 1 summarizes it.

3.2. Definition of the research questions

Research questions help to structure better the systematic literature reviews. Our first concern was the objective behind using AI and DL methods to detect stress in fruits and vegetables. Then, we identify the species to which the techniques were applied and the type of stress involved. Data being an essential element for any model, we defined the data's type, size, and source. We also looked at the countries where the authors have installed experiments to collect the data. In addition, several algorithms and implementation frameworks in the literature are used for given tasks. Thus, we have defined them through the selected studies. Moreover, we sought to determine the models' hyperparameters, performances, and evaluation metrics. Finally,

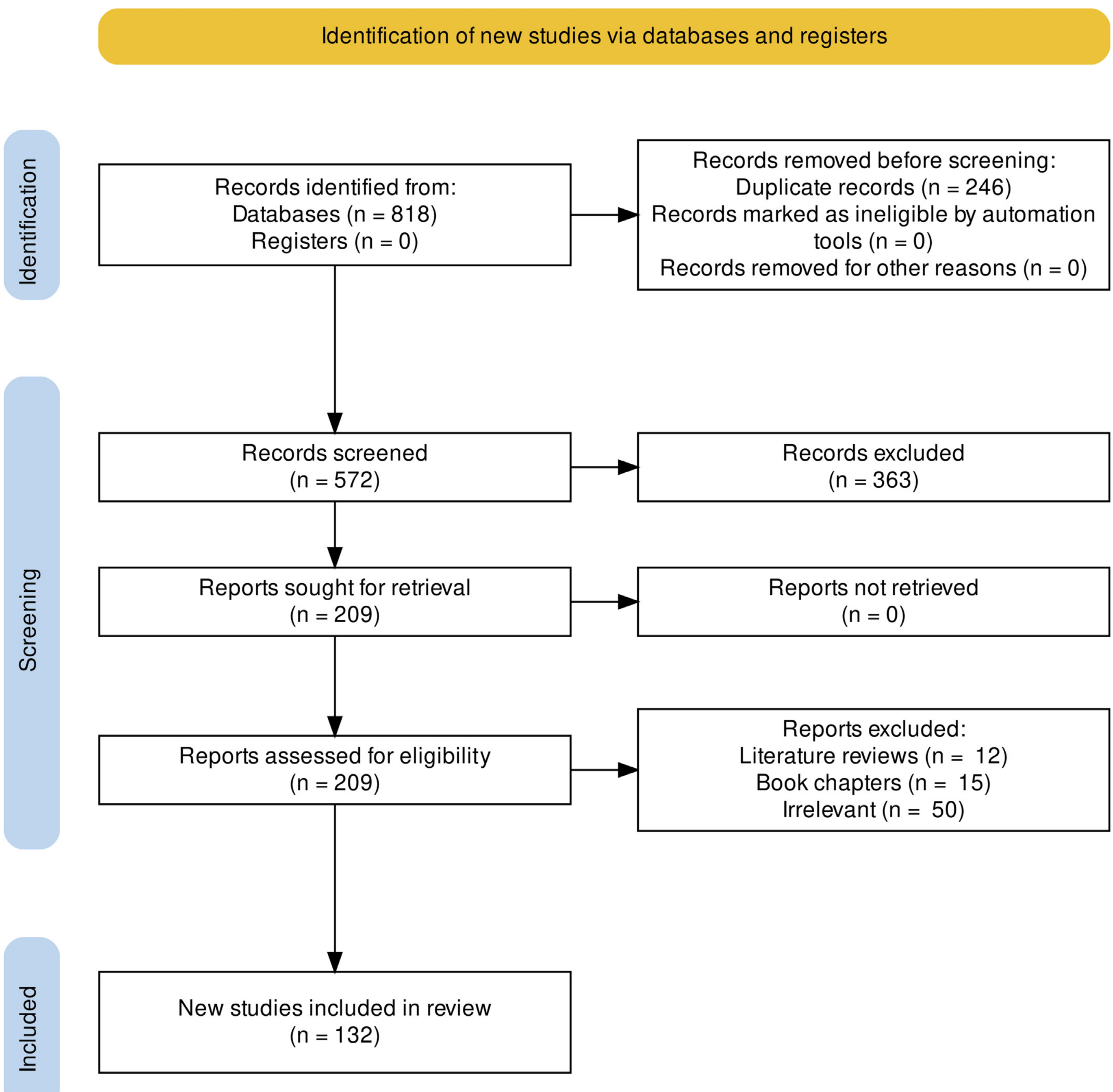

Fig. 1. PRISMA flow diagram of the selection process.

we have identified the challenges that researchers and practitioners face. An explicit reformulation of each research question is given in Table 1.

3.3. Literature synthesis and analysis

The literature synthesis was done in three parts. The first part concerns the bibliometric analysis and some general descriptive statistics. VOSviewer was used for the network of the keywords co-occurrence (van Eck and Waltman, 2022). VOSviewer provides a range of intuitive visualization, mainly for analyzing bibliometric maps. The second part focuses on the answers to the research questions. The occurrences of each response were used to calculate the frequencies. Similar terms were merged to make analysis easy. Some variables were analyzed together to get insight from their study. Results were presented in tables, pie charts, or bar charts. The data analysis was conducted in the Anaconda environment using the Spyder Notebook framework, with Python 3.9.12. The libraries used to synthesize data frames Figures were Pandas, Matplotlib and Numpy. We have also used the function: 'ggplot' of the package: 'ggplot2' (Wickham, 2016) of R software (R Core Team, 2021). Furthermore, the geographical location of the collection sites was represented using QGis. The last part was synthesizing the gaps and ranking them based on their relevance and making recommendations for new studies.

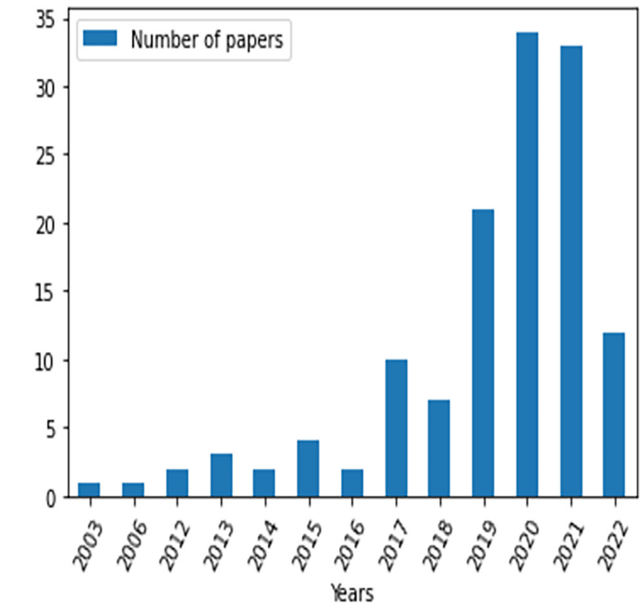

Fig. 2. Year-wise distribution of the articles.

4. Results and discussion

The use of Artificial intelligence in classifying and detecting biotic and abiotic stresses in fruits and vegetables is recent, as shown in Fig. 2. The first papers in this area appeared in 2003, followed by one paper in 2006. A rapid progression was observed from 2019 to 2021, with a peak of 34 papers in 2020. This increase is evidence of researchers' interest in using AI and DL methods for early disease detection, as precision agriculture has become a necessity to face the challenges related to food security. Twelve papers were published from January 2022 to July 2022. Most of the papers are journal articles. Conference proceedings account for about 36% of the total number of papers.

4.1. Bibliometric analysis of keywords

Keywords are the words or groups of words that inform the critical aspects addressed by a research paper. For more precision, we eliminated inconsistencies by merging similar keywords such as 'CNN', 'cnns', and 'convolutional neural networks'. Table 2 shows the top 10 most frequent keywords of the reviewed articles, their number of occurrences, weight links, and total link strength. The weight links are the number of connections of a keyword with other keywords indicating its importance. The total link strength represents the total strength of the co-occurrence links of a given keyword with other keywords. Fig. 3 completes Table 2 by illustrating the keyword co-occurrence network. According to this figure, we distinguish two main periods in the history of detecting and classifying diseases and stress in fruits and vegetables using artificial intelligence methods. The first period, the first cluster of the map, concerns the time when classical machine learning methods were employed through segmentation using k-means clustering. The K-means clustering algorithm groups pixels with common attributes that belong to a particular segment. The support vector machine (SVM) constitutes the core of this cluster. It has 13 occurrences and 22 connections (Table 2) with other words, including 'artificial neural networks (ANN)', 'k-means clustering', 'decision tree', 'disease detection', 'defect detection', etc. Indeed, k-means clustering was used to segment images. In contrast, SVM, ANN, random forest, and decision tree classified the diseases. The emergence of deep learning methods marks the second period represented by the second cluster. Deep learning constitutes the cluster's core with 36 occurrences, 29 connections, and 72 total link strengths (Table 2). The group of words 'deep learning' is linked to other keywords such as: 'Convolutional neural network', 'recurrent neural network', 'feature extraction', 'data augmentation', 'generative adversarial network', 'transfer learning', 'fine tuning', 'classification', 'agriculture' etc. Indeed, deep learning methods such as convolutional neural networks and recurrent neural networks were born during this period. Unlike classical machine learning methods, these methods aim to extract features directly from images. Their performance has improved thanks to the increased databases using algorithms such as generative adversarial networks. In addition, using pre-trained models through transfer learning, the models are fine-tuned and can classify images with much better accuracy for precision agriculture.

Table 1
Research questions.

Nubmer	Research questions
1	What is the objective behind the use of the DL or AI technique?
2	What species was concerned?
3	What type of stress was involved?
4	What are the types and source of data used?
5	What is the countries-wise distribution of the self made-data?
6	What models were used?
7	What are the hyper-parameters used to implement the models?
8	What are the evaluation metrics?
9	What are the performances achieved?
10	What are the gaps and perspectives?

Table 2
Top 10 most used keywords from the assessed papers.

Keywords	Occurrences	Weight Links	Total link strength
Deep Learning	36	29	72
Convolutional Neural Networks	28	29	56
Feature extraction	16	25	43
Support Vector Machines	13	22	33
Image processing	12	17	23
Transfer learning	10	15	26
Classification	9	11	17
Machine learning	8	12	17
k-means clustering	7	8	12
Artificial neural network	6	9	11

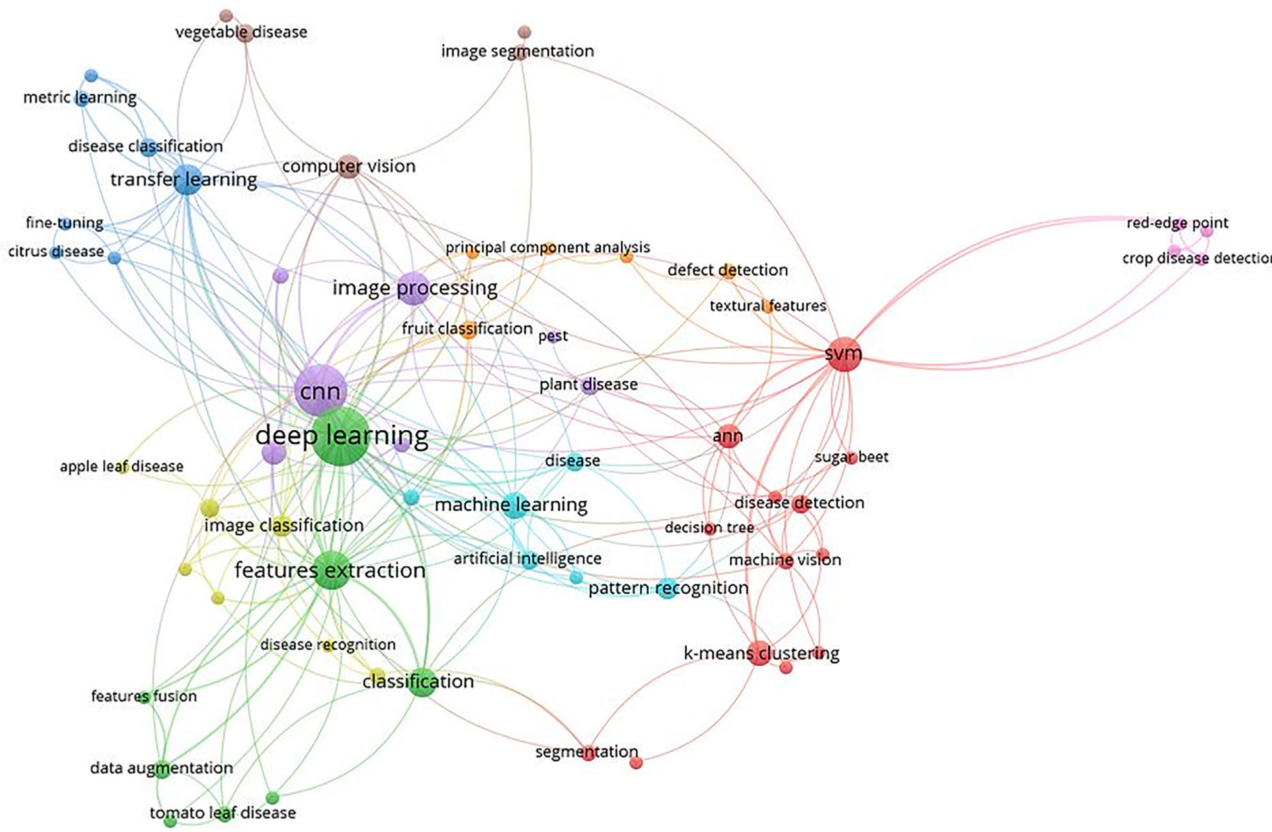

Fig. 3. Map of the most used keywords network.

4.2. RQ1:What is the objective behind the use of the DL or AI techniques?

According to the authors, the objectives behind using Deep Learning and AI techniques are diverse. After careful analysis of each paper, we have grouped these objectives into three main categories: the detection of biotic stress, the detection of abiotic stress, and the improvement of the robustness of the models.

In the first group, we have four sub-objectives. The first was the automatic detection of diseases in 46.46% of the collected papers. The second sub-objective of this group was the detection of vegetable diseases (9.45%). The third of this group concerned the detection of both fruit and vegetable diseases in 7.87% of the papers. The fourth is the fruits and vegetable disease segmentation in 3.94% of the papers. The second main objective of 5.5% of the articles was the automatic detection of abiotic stress, including nutrients deficiency, water stress, light stress, and metal contamination. The last main group includes three sub-objectives. The first is using complex backgrounds to improve the robustness of the models. It accounts for 11.02% of the papers. The second represents 8.66% of the documents that use hybrid models to improve algorithms' robustness. The use of transfer learning to improve the accuracy of prediction models and reduce the time and cost of calculations in 7.09% of the articles constitute this group's last sub-objective. Table 3 illustrates the objectives thus described and the percentage values given to each.

4.3. RQ2: What are the species of fruits and vegetables considered?

Papers considered in this review focused on 28 different species of fruits and vegetables, namely: apple (*Malus domestica* Borkh.), tomato (*Solanum lycopersicum* L.), grape (*Vitis vinifera* L.), lemon (*Citrus lemon* L.), peach (*Prunus persica* L.), orange (*Citrus sinensis* L.), cucumber (*Cucumis sativus* L.), strawberry (*Fragaria anassa* L.), cherry (*Prunus avium* L.), banana (*Musa paradisiaca* L.), mango (*Mangifera indica* L.), potato (*Solanum tuberosum* L.), pepper (*Capsicum annuum* L.), lettuce (*Lactuca sativa* L.), pear (*Pyrus communis L.*), guava (*Psidium guajava* L.), pomegranate (*Punica granatum* L.), cabbage (*Brassica oleracea* L.), sugar beet (*Beta vulgaris* L.), raspberry (*Rubus idaeus* L.), carrot (*Daucus carota* L.), pumpkin/Squash (*Cucurbita* spp.), jackfruit (*Artocarpus heterophyllus* Lam.), olive (*Olea europaea* L.), plum (*Prunus domestica* L.), avocado (*Persea americana* Mill), pineapple (*Ananas comosus* L.) and papaya (*Carica papaya* L.) (Table 4). Of all these species, apple and tomato are the most investigated, with 35 and 33 occurrences, respectively, representing 16.67% and 15.71%. This is easily explained because apple is the most consumed fruit in the world (Feng et al., 2021) while tomato is the second most consumed vegetable after the potato (Kabas et al., 2022). In 2019 the world production of apples reached 7,620,288 t (Oviedo-Mireles et al., 2021) while the world's annual tomato production was 160 million tonnes (Bazrafshan et al., 2022).

The worldwide interest in these two products justifies that they are the object of enough essential work to increase their productivity. After

Table 3
Objectives behind the use of DL or AI methods for stress detection.

Objective category	Objectives	Frequency (%)
Biotic stress detection	Fruits diseases	46.46
	Vegetables diseases	9.45
	Fruits and vegetables diseases	7.87
	Fruits and vegetable segmentation	3.94
Abiotic stress detection	Abiotic stress	5.51
Model improvement with	complex background	11.02
	hybrid models	8.66
	transfer learning	7.09

Table 4
Common and scientific names of species and occurrence frequency.

Common name	Scientific name	Occurrence	Frequency (%)
Apple	*Malus domestica*	35	16.67
Tomato	*Solanum lycopersicum*	33	15.71
Grape	*Vitis vinifera*	16	7.62
Lemon	*Citrus lemon*	14	6.67
Peach	*Prunus persica*	13	6.19
Orange	*Citrus sinensis*	12	5.71
Cucumber	*Cucumis sativus*	11	5.24
Strawberry	*Fragaria anassa*	10	4.76
Cherry	*Prunus avium*	9	4.29
Banana	*Musa paradisiaca*	8	3.81
Mango	*Mangifera indica*	8	3.81
Potato	*Solanum tuberosum*	6	2.86
Pepper	*Capsicum annuum*	5	2.38
Lettuce	*Lactuca sativa*	4	1.90
Pear	*Pyrus communis L.*	4	1.90
Guava	*Psidium guajava*	3	1.43
Pomegranate	*Punica granatum*	3	1.43
Cabbage	*Brassica oleracea*	2	0.95
Sugar beet	*Beta vulgaris*	2	0.95
Raspberry	*Rubus idaeus*	2	0.95
Carrot	*Daucus carota*	2	0.95
Pumpkin/Squash	*Cucurbita* spp.	2	0.95
Jackfruit	*Artocarpus heterophyllus*	1	0.48
Olive	*Olea europaea*	1	0.48
Plum	*Prunus domestica*	1	0.48
Avocado	*Persea americana*	1	0.48
Pineapple	*Ananas comosus*	1	0.48
Papaya	*Carica papaya*	1	0.48

apple and tomato, grape, lemon, peach, orange, cucumber, and strawberry came next with 16, 14, 13, 12, 11, and 10 occurrences, respectively. The least considered species are jackfruit, olive, plum, avocado, pineapple, and papaya, with one study for each.

4.4. RQ3: What type of stress was studied?

The species of fruits and vegetables identified face two types of stress: biotic and abiotic. Diseases constitute about 85% of the biotic stress factors, and pest attacks make up 7.93%. Viral, fungal, and bacterial infections are fruits and vegetables' main biotic stress factors. Fungal diseases are the most dominant. All species' most common fungal diseases are grey spots, brown spots, black spots, downy mildew, powdery mildew, and anthracnose. Regarding abiotic stress, only 6.34% of assessed papers deal with it and address aspects such as light intensity, nutrient deficiency, heavy metal contamination, and water stress. Three studies (Tran et al., 2019; Fuentes et al., 2017; Li et al., 2021) focused on nutrient deficiencies in respectively grape, tomato, and sugar beet. In addition, three other studies were about the cadmium concentration, the lead concentration, and the light intensity of lettuce (Xin et al., 2020; Zhou et al., 2020; Hao et al., 2020). We did not get any article that used DL to predict abiotic stress due to climate change. The low percentage of abiotic stress indicates the lack of interest in this stress detection on fruits and vegetables using artificial intelligence and DL.

4.5. RQ4: What are the type and the source of data used?

Training of the DL models requires large input databases. Because data is the most important input of models, it is necessary to ensure their quality. The data used are essentially images (92%). The other data types are climatic data: mean temperature, minimum temperature, maximum temperature, rainfall, wind speed, humidity, solar, and sunlight. The reviewed papers used about 50% of the data they collected in the field, while 28% relied on the plant village databases (www.kaggle.com/charuchaudhry/plantvillage-tomato-leaf-dataset). PlantVillage is a free downloadable web dataset containing images of 54,303 healthy and diseased leaves, divided into 38 categories by species and disease. Species include apple, blueberry, cherry, corn, grape, orange, peach, pepper, potato, raspberry, soybean, squash, strawberry, and tomato. Diseases are bacterial, fungal, and viral infections. The remaining 18% used data from AI Challenger 2018, ImageNet, Crowd AI, PlantDisease, Coffee leaf, etc. The sensors used to collect this data are cameras, smartphones, and drones.

4.6. RQ5: What is the countries-wise distribution of the self made-data?

Fig. 4 presents the map of the experimental site of data collection from the selected studies. It shows that most papers used data collected in Asia (14, 6, 4, and 3, respectively, from China, India, South Korea, and Bangladesh), followed by the United States of America, Europe (Italy, Germany, United Kingdom, Greece, and Latvia). In addition, there were some efforts from North Africa (Algeria and Egypt), Austral Africa (South Africa), and East Africa (Tanzania). From the selected studies, West Africa was not represented (Fig. 4).

4.7. RQ6: What models were used?

Deep Learning applications have accelerated exponentially over the last five years thanks to the advancement of powerful computing devices such as graphics processing units (GPU). Thus, this section considers the classical machine learning algorithms and ends with deep learning.

4.7.1. Classical machine learning algorithms for stress detection in fruits and vegetables

Classical Machine Learning methods are either supervised, semi-supervised, unsupervised, or reinforcement based. In supervised learning, the datasets are labeled, which means that the output parameters and expected results must be specified. Accuracy must also be adjusted during the learning process. Examples of algorithms in supervised learning are linear regressions, support vector machines (SVM), decision trees, etc. Semi-supervised learning methods combine labeled and unlabelled data. Algorithms of this type are fuelled by certain information through labeled categories, suggestions, and examples. Then they create their labels by exploring the data on their own, according to a rudimentary scheme or the guidance of data scientists. One example is the naive Bayes classifier that uses Bayes' theorem based on conditional probabilities. Researchers employ this algorithm to recognize classes of objects on labeled data sets. Then, the algorithm is trained on unlabelled data. Once this cycle is completed, the researchers associate the labels and restart the training. This technique is mainly used in the context of natural language processing. Unlike supervised algorithms, unsupervised algorithms are not trained. Unsupervised algorithms depend on extensive learning methods to identify patterns by combing unlabelled training data sets and observing correlations. Some examples are K-means clustering, Principal Component Analysis (PCA), and A priori. Reinforcement learning algorithms are built on reward and penalty systems. The algorithm is assigned a goal and tries to get closer to it to obtain maximum compensation. It relies on limited information and learns from its previous actions. These algorithms may depend on a pattern (a model); they must follow predefined steps, and the number of errors and trials is limited. Other algorithms do not rely on a scheme and interpret with each new attempt. Supervised and unsupervised methods are most commonly used for stress detection in fruits and vegetables.

4.7.1.1. Supervised machine learning. Samajpati and Degadwala (2016) propose the classification of apple fruit diseases from the extraction of features and colors. Random forest classifiers were used for disease classification. Still, on apple, Omrani et al. (2014) were interested in detecting three-leaf diseases (Alternaria, black spot, and apple leaf miner) using support vector regression based on the radial basis function. The authors compared ANN and SVM as a classifier and found that SVM performed better. Note that the data used are unbalanced between the

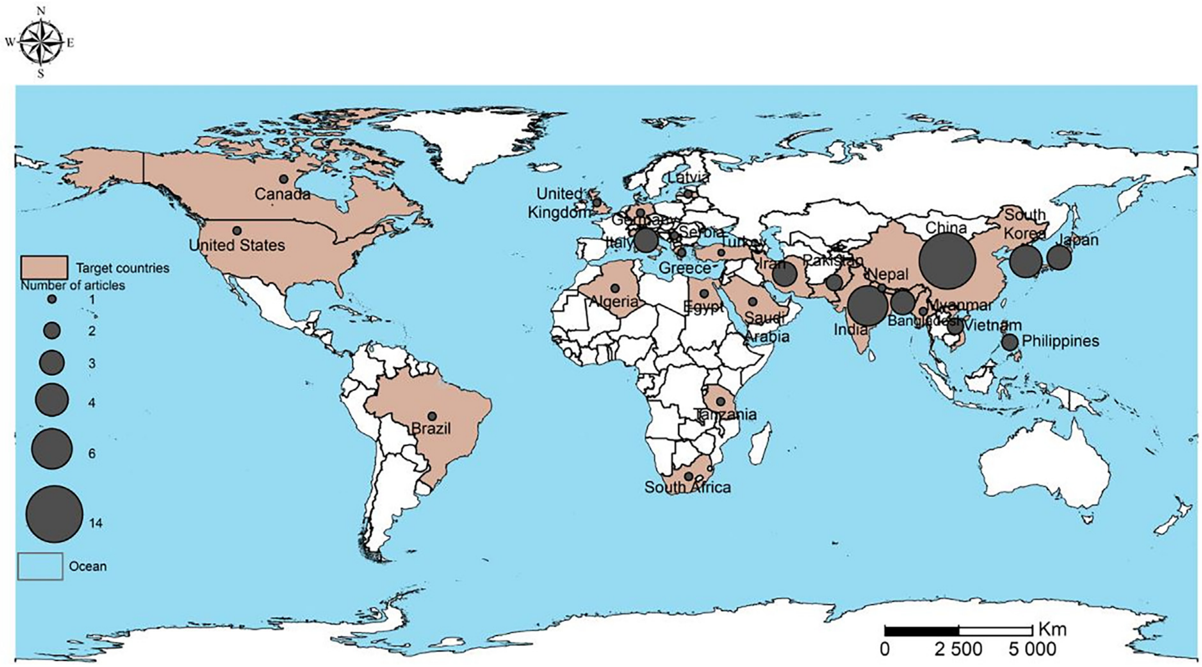

Fig. 4. Countries-wise distribution of self-data collected.

three classes considered. Nine classification algorithms are used for jackfruit disease classification in Bangladesh (Habib et al., 2022). Among the algorithms, Random Forest performed better than all other classifiers, with an accuracy of almost 90%. In addition, Vakilian and Massah (2013) implemented a device for detecting two fungal diseases of cucumbers. Inoculated plants are used as input for the ANN backpropagation algorithm, giving an acceptable performance. Zhang and Wu (2012) developed a classification method based on multiclass SVM in several steps. The first step consisted of the background removal of acquired images with a split-and-merge algorithm. The second step involved extracting each fruit image's color histogram, texture, and shape features for feature space composition. In the final step, the multiclass SVM was constructed. They come up with the Gaussian Radial Basis Kernel achieving the highest classification accuracy of 88.2%. In Razmjooy et al. (2012), they used SVM sequential minimal optimization (SMO) to reach 95% accuracy for the classification of potato defects.

4.7.1.2. Unsupervised machine learning. Dubey et al. (2013) proposed the defect segmentation of apple fruits using k-means clustering in two steps. The first step consisted of pixel clustering based on color and special features. The second step used the merged clustered block to a specific number of regions. Likewise, K-means clustering was used for segmentation in the case of any infection (Samajpati and Degadwala, 2016). Moreover, Omrani et al. (2014) employed K-means clustering to segment the region of interest in apple leaves. However, wavelet and grey-level co-occurrence matrix techniques were used to get the texture features. Also discussing segmentation of the diseased region, Rozario et al. (2016) suggested a computer vision-based approach in fruits and vegetables using K-means clustering, modified K-means clustering, and the Otsu method. In Bangladesh, an agromedical expert system for detecting and classifying papaya fruits has been developed (Habib et al., 2020). The system uses K-means clustering to detect the disease-infested region on a leaf. Still, in Bangladesh, another agromedical expert system for jackfruit detection and classification has been developed (Habib et al., 2022). The system also uses K-means clustering to detect the disease-infested region on a leaf. PCA was employed to reduce the dimensions of the feature space in Zhang and Wu (2012). Moreover, for potato defect detection and size sorting, Razmjooy et al. (2012) provided a system based on mathematical binarization and defect segmentation using color-based classifiers.

4.7.2. Deep learning approaches

Depending on the learning task, four deep learning approaches can be considered: deep supervised, deep semi-supervised, deep unsupervised, and deep reinforcement learning.

4.7.2.1. Deep supervised learning. The deep supervised learning technique deals with labeled data. It considers a collection of inputs and resultant outputs (x_i, y_i). For instance, we want to predict $\widehat{y}_i$ knowing the input x_i and will obtain a loss value $l(\widehat{y}_i, y_i)$. Then, the network parameters are repeatedly updated to get an improved estimate for the preferred outputs (Alzubaidi et al., 2021). There are several supervised DL techniques, and the most common are convolutional neural networks (CNNs) and recurrent neural networks (RNNs).

Convolutional neural networks, also known as CNN or ConvNet, are deep learning algorithms that take the input image, assign weights/biases to the components of the image, and then classify the entire picture. Based on the resolution and size of the image, it is transformed into an array. Each entry consists of a number between 0 and 255 for RGB (Red Green Blue) systems. This number represents the pixel intensity of the image. CNN's structure consists of three layers: the convolutional layer, the pooling layer, and the fully connected (FC) layer. The first layer makes extraction of high-level features. It convolved the image with a filter or kernel using the formula: $Z = X{*}f$, where X and f are the input image and the filter, respectively. The second layer intends to reduce the spatial size of the image representation and the computational and processing overhead of neural networks. It also extracts key positionally and rotationally invariant features. The last layer uses the flattened output of the previous pooling/convolution layer as input. Flattened means that a 3D matrix or array is expanded into a vector. Specific mathematical calculations are performed for each FC layer. After the vector has passed through all the fully connected layers, the softmax activation function is used in the last layer to calculate the probability that input will belong to a particular class. This process is repeated for other classes and individual images within those classes. It trains the network and teaches you, for instance, to distinguish between healthy and diseased. Fig. 5 shows that most papers used GoogleNet, ResNet50, and VGG16 for stress classification in fruits and vegetables. These are the architectures of convolutional neural networks (CNN). The CNN technique has performed well on tasks such as disease identification on apple fruit (Agarwal et al., 2019; Alharbi and Arif, 2020; Jiang et al., 2019; Sheikh et al., 2020), detection and diagnosis of damage to jackfruit by pests and diseases (Orano et al., 2019; Kukreja and Dhiman, 2020), tomato disease identification (Ashok et al., 2020; Mkonyi et al., 2020; Afifi et al., 2021; Brahimi et al., 2017), mango disease classification (Chouhan et al., 2019; Singh et al., 2018), guava disease detection (Farhan Al Haque et al., 2019; Howlader et al., 2019), strawberry disease diagnosis (Park et al., 2018; Sheikh et al., 2020; Afifi et al., 2021; Ferentinos, 2018), abiotic light stress grading in lettuce leaves (Hao et al., 2020), citrus disease identification (Pan et al., 2019), diagnosis of multiple cucumber infections (Tani et al., 2018), and grape diseases classification (Thet et al., 2020; Sheikh et al., 2020). One can attribute the widespread use of CNNs to their outstanding performance in solving a wide range of problems, particularly those involving images (Alzubaidi et al., 2021). In addition to performance, there are other advantages to using CNN models. For instance, the reduction of weight and convolutional neurons (Agarwal et al., 2019), the capability to transfer learning (Afifi et al., 2021; Thet et al., 2020), the possibility to deploy the models in a mobile application (Orano et al., 2019; Ferentinos, 2018; Sheikh et al., 2020; Pan et al., 2019), the alleviation of training time required (Atabay, 2017), the lower computational complexity (Ferentinos, 2018), and the simplicity of the model (Singh et al., 2018). Moreover, several open-source CNN software have been made available. However, unbalanced data often hurt CNN models performance (Hao et al., 2020). Apart from CNNs, there are other supervised deep learning models, including Recurrent Neural Networks (RNN).

The Recurrent Neural Networks (RNN) technique is specialized in sequences of data $x(t) = x(1), \ldots, x(\tau)$ with the time step index t ranging from 1 to τ. It uses the knowledge gained from its previous state as an input value for the current prediction. Therefore, it can help achieve short-term memory in a network and effectively manage time-based data systems. There are two different types of RNN: Long Short Term Memory (LSTM) and Gated Recurrent Units (GRUs). The LSTM is conceived to use memory to forecast data in temporal sequences. It has three gates: input, output, and oblivion. The GRUs also predict time sequences via memory but have two gates: Update and Reset. The RNN method works best for tasks where a single input is connected to a single output and a single input is associated with such an output sequence. It also works best with a series of inputs that produce a single output, such as sentiment analysis, and a set of inputs that create a set

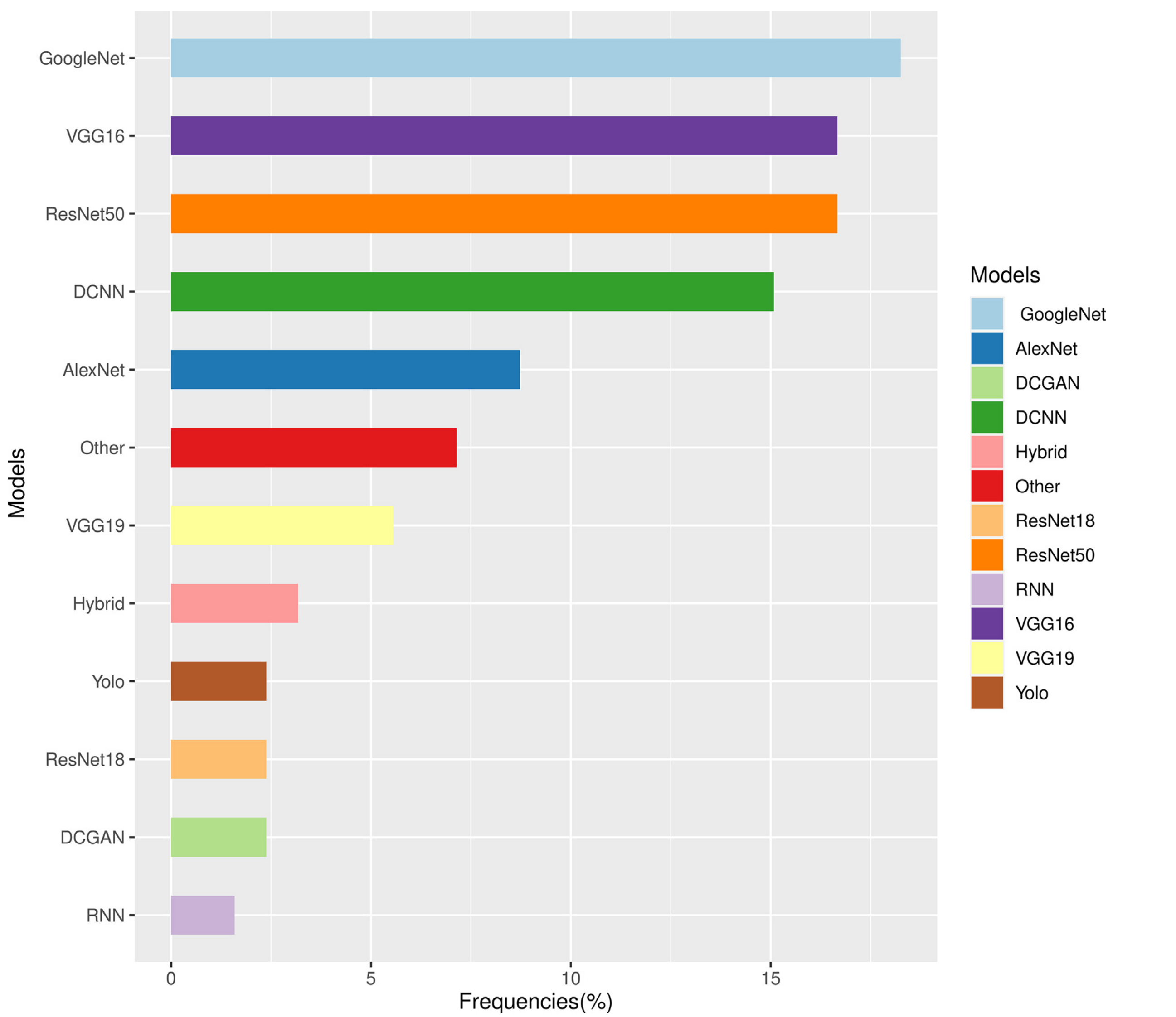

Fig. 5. Deep Learning models considered in the assessed papers.

of outcomes. The technique is used to measure water stress in tomato plants based on images, environmental data (temperature, relative humidity, vapor pressure difference, and scattered light), and stem diameter (Wakamori et al., 2020).

4.7.2.2. Deep semi-supervised learning. In Deep semi-supervised learning, the learning process is based on partially-labeled datasets. That is, we are in a situation where few labeled learning examples are available and many unlabelled samples (Ouali et al., 2020). The proportion of labeled examples is generally relatively small, ranging from 1 to 10% of the total number of instances (Ouali et al., 2020). Text document classifiers are among the most common examples of semi-supervised learning applications because it is challenging to retrieve large numbers of labeled text documents. Generative adversarial networks (GANs) are sometimes used in the same way as this technique. GANs are composed of two models: the Generator, which generates new samples based on the approximate distribution of the original dataset, and the Discriminator, which is used to distinguish the original dataset from the data generated by the Generator. Another variant of GANs techniques is Deep Convolutional Generative Adversarial Networks (DCGANs). DCGANs techniques helped in estimating the severity of citrus diseases, improving the model learning performance (Zeng et al., 2020). The advantage of these techniques is to minimize the amount of labeled data needed (Alzubaidi et al., 2021). One of the drawbacks of this technique is irrelevant input functions that exist in training data and can lead to incorrect decisions (Alzubaidi et al., 2021). GANs techniques are also known to be unstable to train (Zeng et al., 2020).

4.7.2.3. Deep unsupervised learning. This technique allows running the learning process without labeled data available. The model is supposed to organize the data on its own, based on the input data's features, to discover the data's unknown structure or relationship. This approach often includes Generation network technology, dimensionality reduction, and clustering. Some members of the DL family are working well on Nonlinear dimensionality reduction and clustering tasks. These include Restricted Boltzmann Machine (RBM) and Autoencoder.

The Restricted Boltzmann Machine (RBM) or Boltzmann Machine: is a generative unsupervised model that learns a probability distribution from the original dataset and infers data it has never seen before. The RBM has an input layer and one or more hidden layers. It uses a neural network with neurons connected to neurons in the same layer and other neurons in other layers. The nodes are connected in a circle. In contrast to all deterministic network models, the RBM model is called stochastic. It is ideal for system monitoring and handwritten digit recognition (eg: check verification and criminal evidence). The advantage of RBM is its possibility to encode any distribution due to its expressiveness and computational efficiency. In addition, using hidden layer activation as input to other models is a helpful feature to improve performance. This technique is more challenging to train.

Autoencoders are unsupervised deep learning techniques used to learn efficient data coding (Kunapuli and Bhallamudi, 2021). They consist of four or five flat, two symmetric deep belief networks. Half of the network encodes, and the other half decodes. The autoencoder learns essential functions in the data by minimizing reconstruction errors between input and output data (Abirami and Chitra, 2020). There are an equal number of neurons in the output layer and the input layer with Autoencoders. Autoencoders are flexible due to both linear and nonlinear transformations in encoding (Abirami and Chitra, 2020).

Autoencoder was used to predict and classify early and late phenomena resulting from macronutrient deficiencies in tomato plants (Tran et al., 2019) and for detection of lead concentration in lettuce (Xin et al., 2020). In general, the most crucial drawback of unsupervised learning methods is their inability to provide accurate information from the data sorting and computational effort (Singh et al., 2019).

4.7.2.4. Deep reinforcement learning. Developed in 2013 with Google DeepMind (Mnih et al., 2015; Alzubaidi et al., 2021), Deep reinforcement learning is the process by which an agent interacts with the environment to change its state. The agent takes action accordingly based on the observation made. Agents interact with the situation to help the network achieve its goals. This network model has an input layer, an output layer, and some hidden layers, and the state of the environment is the input layer itself. The model serves as a continuous attempt to predict future rewards for actions taken in a particular situation. This learning is much more challenging to perform because it cannot use the simple loss function (Alzubaidi et al., 2021) compared to the traditional deep supervised method. In addition, there are two crucial differences between deep supervised learning and deep reinforcement learning. First, one does not have full access to the function requiring optimization. That is, one needs to query through the interaction. Second, the state being interacted with is founded on an environment in which the input is based on previous actions (Arulkumaran et al., 2017; Alom et al., 2019; Alzubaidi et al., 2021).

4.8. RQ7: What are the hyper-parameters used to implement the models?

Fig. 6 shows four hyperparameters of DL models: the activation functions, the pooling methods, the optimizers, and the loss functions.

The activation functions are added at each layer of the DL model. They enable the algorithm to learn and understand something extremely complicated and non-linear between the inputs and response variables. Relu (Rectified Linear Unit) was preferred mainly (80%) over the other Sigmoid, Tanh, and LeakyRelu. Because of its mathematical simplicity: y = max(0,x), Relu uses less expensive operations. Therefore, its training rate is fast (Kukreja and Dhiman, 2020), and it improves convergence and rectifies vanishing gradient problems. LeakyRelu is an extension of the Relu function. The Sigmoid function is S-shaped and is used for binary classification problems such as 0 healthy and 1 diseased. It is expressed by: $y = 1/(1 + exp^{\,-x})$ The activation function Tanh is an extension of the Sigmoid function. It can be calculated using the expression $y = (1/(1 + exp^{\,-2x})) - 1$ Max pooling operators were used more than average pooling (Fig. 6). The objects of interest produce the highest pixel values, so taking the maximum value in a block will be better than the average. In addition, max-pooling acts as a noise canceller and provides better performance than average pooling.

The optimizers are algorithms used to minimize the loss function. The Stochastic Gradient Descent (SGD) and Adam(Adaptive Moment Estimation) were the most used, respectively, with 52% and 27%. The SGD frequently updates the model parameters and allows using large data sets with less memory. Adam optimizer proposes an adaptive learning rate for each parameter.

The loss function 'Categorical cross-entropy was used in more than 75% of the papers. The categorical cross-entropy function deals with probabilities. Each of the predicted probabilities is compared to the actual value of the class output (0 or 1), and a score is calculated. In contrast, the Contrastive loss function operates on the data points generated by the network and their positions relative to each other (Khosla et al., 2020). Then, the Categorical cross-entropy seems better for image classification tasks.

4.9. RQ8: What are the evaluation metrics?

The most frequently used metrics for assessing model performance are accuracy, precision, recall, and F1 score (Fig. 7). Accuracy or Classification Accuracy (CA) measures the effectiveness of a model as the percentage of correct predictions where the top class (the one with the highest probability) as indicated by the model, is the same as the target

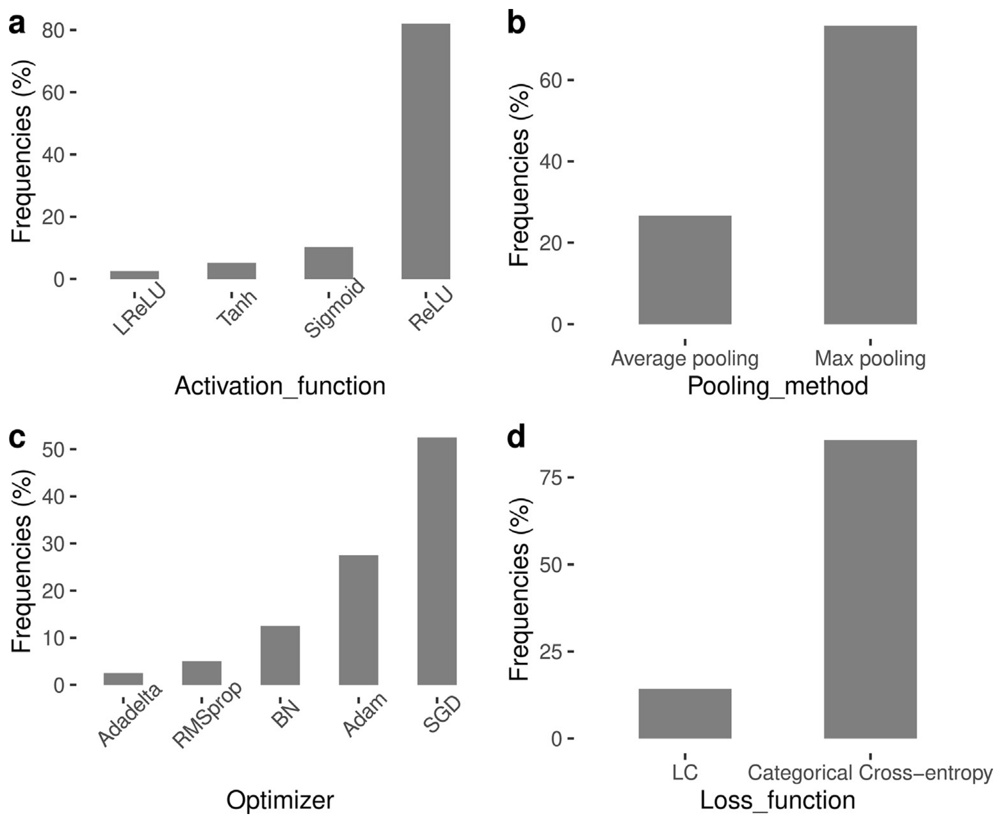

Fig. 6. Hyper parameters: (a) activation functions, (b) Pooling methods, (c) Optimizers, (d) Loss functions.

label as annotated beforehand by the author (Kamilaris and Prenafeta-Boldu, 2018). When precision (P) is the fraction of true positives (TP, correct predictions) from the total amount of relevant results, the Recall (R) is the fraction of TP from the total amount of TP and false negatives (FN). The F1 score is the harmonic average of precision (P) and recall (R). For multi-class classification problems, each metric gets averaged among all the classes. The following formulas are used to compute them (Kamilaris and Prenafeta-Boldu, 2018):

$$CA = (TP + TN)/(TP + TN + FN + FP) \tag{1}$$

$$P = TP/(TP + FP)$$

$$R = TP/(TP + FN) \tag{2}$$

$$F1 = 2 * (TP * FP)/(TP + FP) \tag{3}$$

The R2 and the root mean square error (RMSE) are the few observed regression metrics. The RMSE is the square root of the mean square error. It is used to measure the standard deviation of the residuals. The coefficient of determination or R-square is the proportion of the variance of the dependent variable that the linear regression model explains. The R2 is a scale-free score, which means that regardless of whether the values are small or large, the R-square value will be less than one. The R2 and RMSE are calculated using the following formulas.

$$RMSE = \sqrt{\left(\frac{1}{N}\right)\sum_{i=1}^{N}(y_i - \widehat{y})^2} \tag{4}$$

$$R^2 = 1 - \frac{\sum (y_i - \widehat{y})^2}{\sum (y_i - \overline{y})^2} \tag{5}$$

4.10. RQ9: What are the performances achieved?

We consider the most used evaluation metric (accuracy) to group the best-performing models reported by the authors into five classes. The five classes are defined as follows: class 1 from 95 to 99.99%, class 2 from 90 to 94.99%, class 3 from 80 to 89.99%, class 4 from 70 to 79.99%, and class 5 from 50 to 69.99%. We then calculated the occurrence of each model in a class and reported it as a frequency. This was done for two purposes. The first one is to identify the best models of all those written by the authors. The second purpose is to check if the best ones are also the most commonly used. We assume that a model would be best if it has been mentioned each time with at least a 90% accuracy value. Thus, a model must fulfill two criteria to be considered 'best'. The first criterion is that it has to be represented in the first class. The second criterion is that the model is neither in class 3 nor in class 4 or 5. According to Fig. 8, only ResNet 50 and AlexNet meet the criteria and can therefore be considered the 'best' models. Of the top three most used models, GoogleNet, VGG16, and ResNet50, only ResNet50 is among the best according to our criteria. One concludes that the most used models are not all included in the best ones.

4.11. RQ10: What are the gaps and perspectives

The automatic detection and classification of stress in fruits and vegetables is a useful approach that eases the task of both the farmer and the specialist of plant pathology. Indeed, it allows non-experienced farmers to monitor their plants and apply timely measures to optimize yield. The plant pathologist can no longer travel long distances to assist the farmer. However, there are five significant gaps linked to the use of these methods. The first one, the most important of the whole process, from training to testing the model, is related to the data. Data challenge is on the one hand due to the source and, on the other hand, the quantity

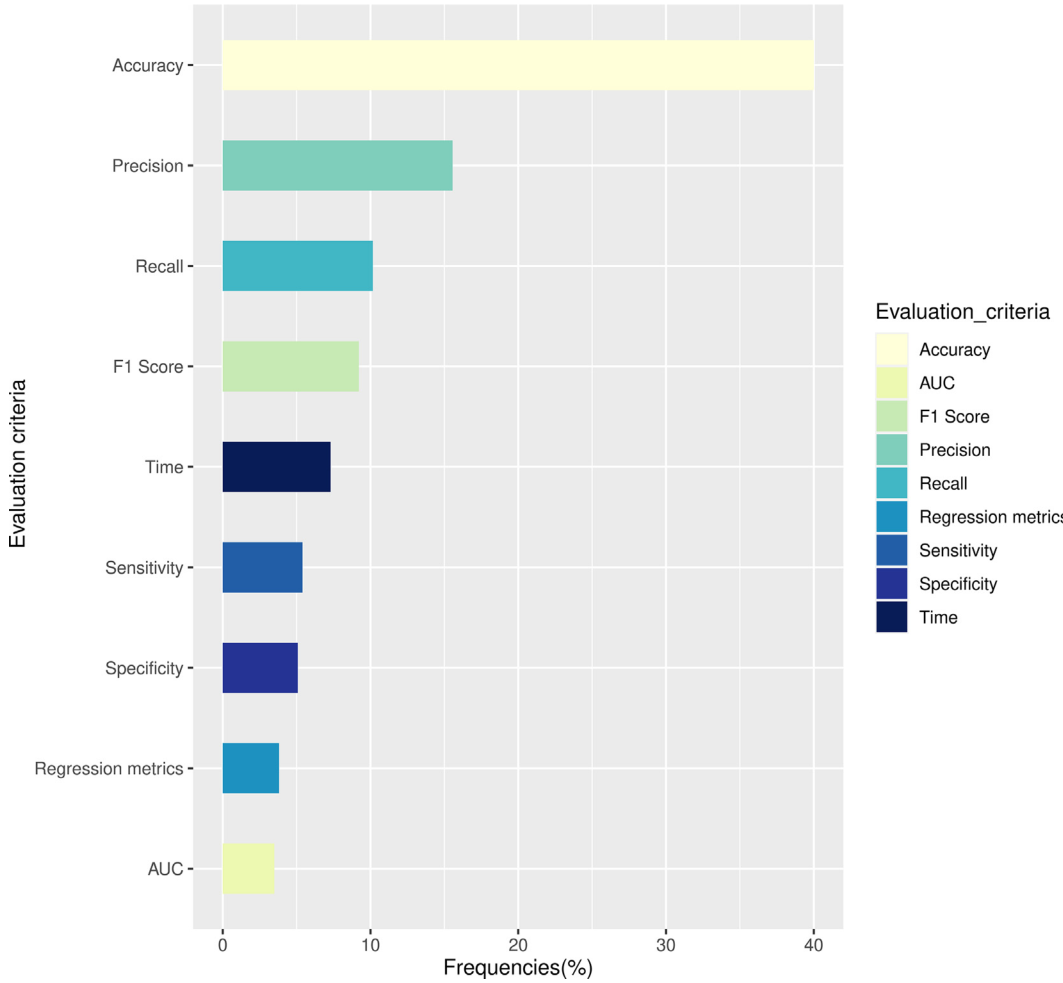

Fig. 7. Evaluation metrics used within the reviewed papers.

of data. Collecting data representing the characteristics of several diseases at different stages requires careful monitoring, financial means, time, and patience. It constitutes a constraint for researchers who, to remedy this, use online databases to train their models. For example, in more than 45% of the papers studied, the data used were obtained online (Hong et al., 2020; Zeng et al., 2020; Zhou et al., 2021; Adedoja et al., 2019; Agarwal et al., 2019; Pan et al., 2019; Zheng et al., 2019; Rangarajan et al., 2018). Plantvillage is the most popular of all the online databases and is found in 28% of the articles reviewed. They offer the possibility of having many images of several species. However, these are images taken in the laboratory. These pictures are mostly infested leaves with a homogeneous background (El-Kereamy et al., 2016). They do not show different illumination levels given the time of day the photo would be taken in the field, the crop substrate, or any other features observed in a natural field environment. As a result, these images do not reflect what is happening in the field. Another challenge related to the data is their small quantity which sometimes creates an imbalance in classes. Indeed, there are some diseases whose symptoms are not easily observed in the field. Thus, collecting data to balance these rare classes with other frequent classes is not always straightforward. From the reviewed papers, many authors faced unbalanced classes problem in their training data (Fenu and Malloci, 2021; Tani et al., 2018; Kodors et al., 2021; Fuentes et al., 2017; Liu et al., 2020). In addition, if the number of observations in some classes is minimal, it will be challenging to represent between the classes, which will cause problems in validation or creating test samples. For instance, Fuentes et al. (2017) considered 0.9% of their dataset for the minimum class and 43% of the data in the maximum class, while Kodors et al. (2021) used 6% of the observations in the minimum class while 65% represented the maximum class. The final consequence is the reduction of the global performance of the model. Of the studied articles, 20% used unbalanced data. Some augmentation methods, such as blur, rotation, contrast enhancement, etc., are used to overcome overfitting and class imbalance. However, these methods still do not reflect what is observed in the actual situation. Apart from augmentation methods, transfer learning is also used for small databases and unbalanced classes. However, even transfer learning does not improve the model's accuracy when the imbalance is enormous. The second gap is the consequence of the first. Indeed, the models deployed with data of homogeneous characteristics or unbalanced classes suffer from bias and overfitting. The unbalanced classes create problems because one doesn't get enough insight into the class on which the model is based. Then, optimized results were not obtained for imbalanced classes in real time. The models show good performances on training data but have difficulty generalizing. The final consequence is the reduction of the global robustness of the model. The third gap is the non-classification of the severity of some attacks. Although necessary, the infection's severity was not implemented in some models (Orano et al., 2019; Wang et al., 2020; De Luna et al., 2019). In particular, (Wang et al., 2020) and (De Luna et al., 2019) proposed to detect whether apple and tomato fruits were defective or not. But there are several levels of defectiveness in general. Thus, the merchant would be more profitable if the algorithm categorized the severity of defects. Therefore, the less attacked ones can be sold and the more attacked ones isolated quickly

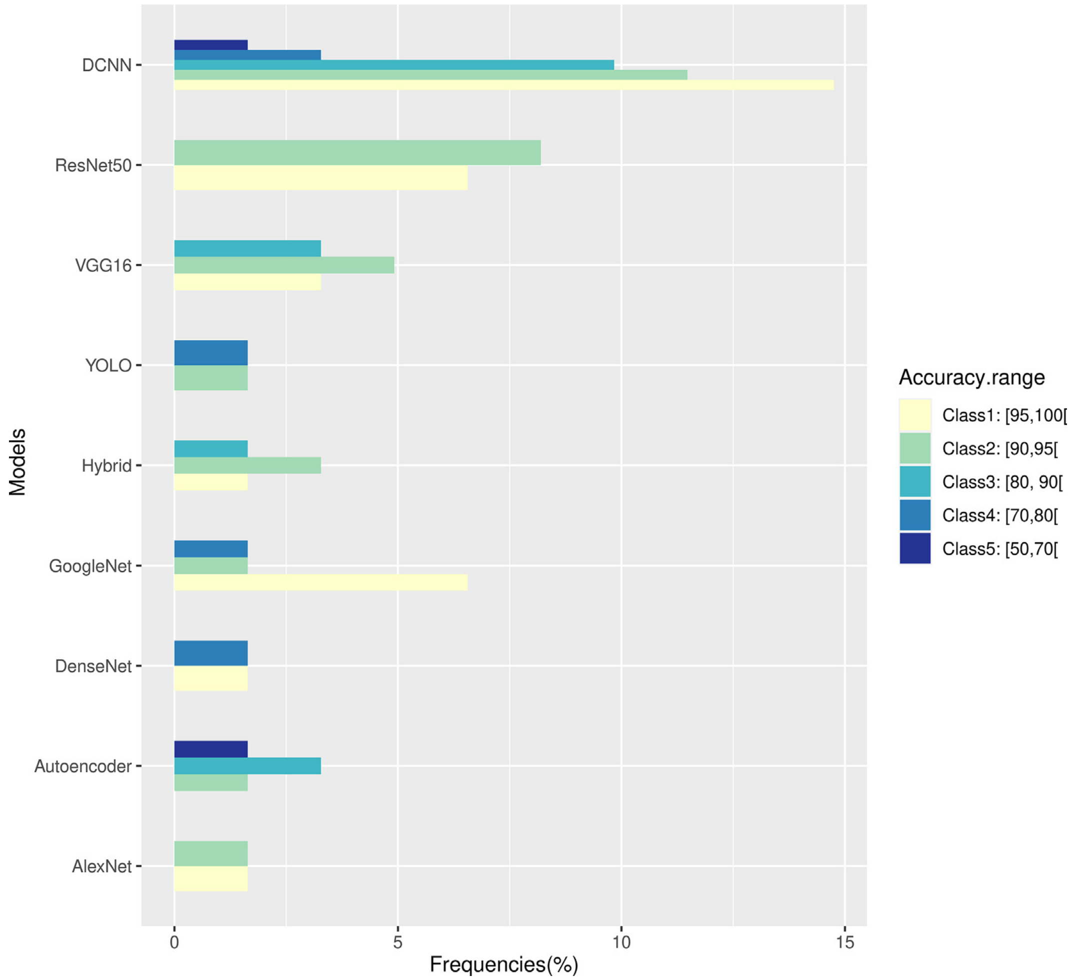

Fig. 8. Models classification based on the achieved performances.

to avoid losing everything. The fourth gap is related to the types of stress studied. Most of the work focuses on the early detection of biotic stress, such as diseases. Very little work has focused on abiotic stress due to nutrient deficiency, water (water stress), heavy metal contamination, and light quantity. At the same time, no work has been done on predicting stress due to climate change in fruits and vegetables. Indeed, climate change has many side effects in our lives. Thus, we recommend that AI-based models be developed to predict the impact of abiotic stress related to climate change on fruits and vegetables. It would allow us to start thinking about finding mitigation measures. Finally, the early detection of several diseases on the same leaf using AI methods remains a real challenge. Because, in the field, it is not uncommon to find several diseases or pest attacks on the same leaf. Experiments should be implemented to collect mega data with various characteristics on different species of fruits and vegetables. These data should be made available to researchers to have robust models that work well in real situations.

5. Conclusion

Applications of Deep Learning in agriculture have been steadily advancing in recent years. This review presented recent work in agriculture to identify and classify biotic and abiotic stresses in fruits and vegetables. The effectiveness of the models was evaluated based on the data sources, the models used, the hyper parameters adopted, and the evaluation metrics used. Finally, the limitations of the papers and some perspectives were presented. The finding is that most of the deep learning techniques used had a good performance but needed to be improved. This review also provides research guidelines for scientists who intend to work in this area.

Funding

This work was supported by the German Academic Exchange Service (DAAD) and the scholarship program" Artificial Intelligence for Development (AI4D) in Africa", funded by the International Development Research Centre (IDRC) and the Swedish International Development Cooperation Agency (SIDA), and managed by the African Centre for Technology Studies (ACTS).

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.

References

Abirami, S., Chitra, P., 2020. Chapter fourteen - energy-efficient edge based real-time healthcare support system. In: Raj, P., Evangeline, P. (Eds.), The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases. Elsevier, Advances in Computers, pp. 339–368 https://doi.org/10.1016/bs.adcom.2019.09.007.

Adedoja, A., Owolawi, P.A., Mapayi, T., 2019. Deep learning based on NASNet for plant disease recognition using leave images. icABCD 2019 - 2nd International Conference on Advances in Big Data, Computing and Data Communication Systems, pp. 1–5 https://doi.org/10.1109/ICABCD.2019.8851029.

Afifi, A., Alhumam, A., Abdelwahab, A., 2021. Convolutional neural network for automatic identification of plant diseases with limited data. Plants 10, 1–16. https://doi.org/10.3390/plants10010028.

Agarwal, M., Kaliyar, R.K., Singal, G., Gupta, S.K., 2019. Fcnn-Lda: a faster convolution neural network model for leaf disease identification on apple's leaf dataset. Proceedings of 2019 International Conference on Information and Communication Technology and Systems, ICTS 2019, pp. 246–251 https://doi.org/10.1109/ICTS.2019.8850964.

Alharbi, A.G., Arif, M., 2020. Detection and classification of apple diseases using convolutional neural networks. 2020 2nd International Conference on Computer and Information Sciences, ICCIS 2020, pp. 1–5 https://doi.org/10.1109/ICCIS49240.2020.9257640.

Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Hasan, M., Van Essen, B.C., Awwal, A.A.S., Asari, V.K., 2019. A state-of-the-art survey on deep learning theory and architectures. Electronics 8. https://doi.org/10.3390/electronics8030292.

Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L., 2021. Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. J. Big Data 8, 53. https://doi.org/10.1186/s40537-021-00444-8.

Arulkumaran, K., Deisenroth, M.P., Brundage, M., Bharath, A.A., 2017. Deep reinforcement learning: a brief survey. IEEE Signal Process. Mag. 34, 26–38. https://doi.org/10.1109/MSP.2017.2743240.

Ashok, S., Kishore, G., Rajesh, V., Suchitra, S., Sophia, S.G., Pavithra, B., 2020. Tomato leaf disease detection using deep learning techniques. 2020 5th International Conference on Communication and Electronics Systems (ICCES), pp. 979–983 https://doi.org/10.1109/ICCES48766.2020.9137986.

Atabay, H.A., 2017. Deep residual learning for tomato plant leaf disease identification. J. Theor. Appl. Inf. Technol. 95, 6800–6808.

Atkinson, N.J., Urwin, P.E., 2012. The interaction of plant biotic and abiotic stresses: from genes to the field. J. Exp. Bot. 63, 3523–3543. https://doi.org/10.1093/JXB/ERS100.

Bazrafshan, O., Ehteram, M., Latif, S.D., Huang, Y.F., Teo, F.Y., Ahmed, A.N., El-Shafie, A., 2022. Predicting crop yields using a new robust bayesian averaging model based on multiple hybrid anfis and mlp models. Ain Shams Eng. J. 13. https://doi.org/10.1016/j.asej.2022.101724.

Blancard, D., 2012. Tomato Diseases. Snd edition ed. Quae.

Brahimi, M., Boukhalfa, K., Moussaoui, A., 2017. Deep learning for tomato diseases: classification and symptoms visualization. Appl. Artif. Intell. 31, 299–315. https://doi.org/10.1080/08839514.2017.1315516.

Breitenreiter, A., Poppinga, H., Berlin, H., Technik, F., 2015. Deep learning. Nature 521, 2015.

Chouhan, S.S., Kaul, A., Singh, U.P., 2019. A deep learning approach for the classification of diseased plant leaf images. Proceedings of the 4th International Conference on Communication and Electronics Systems, ICCES 2019, pp. 1168–1172 https://doi.org/10.1109/ICCES45898.2019.9002201.

De Luna, R.G., Dadios, E.P., Bandala, A.A., Vicerra, R.R.P., 2019. Tomato fruit image dataset for deep transfer learning-based defect detection. Proceedings of the IEEE 2019 9th International Conference on Cybernetics and Intelligent Systems and Robotics, Automation and Mechatronics, CIS and RAM 2019, pp. 356–361 https://doi.org/10.1109/CIS-RAM47153.2019.9095778.

Dubey, S.R., Dixit, P., Singh, N., Gupta, J.P., 2013. Infected fruit part detection using k-means clustering segmentation technique. Int. J. Interact. Multimedia Artific. Intellig. 2, 65. https://doi.org/10.9781/IJIMAI.2013.229.

Dutot, M., Nelson, L.M., Tyson, R.C., 2013. Predicting the spread of postharvest disease in stored fruit, with application to apples. Postharvest Biol. Technol. 85, 45–56. https://doi.org/10.1016/J.POSTHARVBIO.2013.04.003.

El-Kereamy, A., Meitz-Hopkins, J.C., Testolin, A., Salathé, M., Mohanty, S.P., Hughes, D.P., 2016. Using deep learning for image-based plant disease detection. Front. Plant Sci. 7, 1419. https://doi.org/10.3389/fpls.2016.01419 — www.frontiersin.org.

Farhan Al Haque, A.S., Hafiz, R., Hakim, M.A., Rasiqul Islam, G.M., 2019. A computer vision system for guava disease detection and recommend curative solution using deep learning approach. 2019 22nd International Conference on Computer and Information Technology, ICCIT 2019, pp. 1–6. https://doi.org/10.1109/ICCIT48885.2019.9038598.

Feng, S., Yi, J., Li, X., Wu, X., Zhao, Y., Ma, Y., Bi, J., 2021. Systematic review of phenolic compounds in apple fruits: compositions, distribution, absorption, metabolism, and processing stability. J. Agric. Food Chem. 69, 7–27. https://doi.org/10.1021/acs.jafc.0c05481.

Fenu, G., Malloci, F.M., 2021. Using multioutput learning to diagnose plant disease and stress severity. Complexity 2021. https://doi.org/10.1155/2021/6663442.

Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. Comput. Electron. Agric. 145, 311–318. https://doi.org/10.1016/j.compag.2018.01.009.

Fuentes, A., Yoon, S., Kim, S.C., Park, D.S., 2017. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors (Switzerland) 17. https://doi.org/10.3390/s17092022.

Fukushima, K., 1988. Neocognitron: a hierarchical neural network capable of visual pattern recognition. Neural Netw. 119–130.

Goncalves, J.F.D.C., Barreto, D.C.D.S., Santo, U.M.D., Fernandes, A.V., Sampaio, P.D.T.B., Buckeridge, M.S., 2005. Growth, photosynthesis and stress indicators in young rosewood plants (aniba rosaeodora ducke) under different light intensities. Braz. J. Plant Physiol. 17, 325–334. https://doi.org/10.1590/S1677-04202005000300007.

Habib, M.T., Majumder, A., Jakaria, A.Z., Akter, M., Uddin, M.S., Ahmed, F., 2020. Machine vision based papaya disease recognition. J. King Saud Univ. Comp. Inform. Sci. 32, 300–309. https://doi.org/10.1016/J.JKSUCI.2018.06.006.

Habib, M.T., Mia, M.J., Uddin, M.S., Ahmed, F., 2022. An in-depth exploration of automated jackfruit disease recognition. J. King Saud Univ. Comp. Inform. Sci. 34, 1200–1209. https://doi.org/10.1016/J.JKSUCI.2020.04.018.

Hao, X., Jia, J., Gao, W., Guo, X., Zhang, W., Zheng, L., 2020. MFC-CNN : an automatic grading scheme for light stress levels of lettuce (*Lactuca sativa* L.) leaves. Comput. Electron. Agric. 179, 105847. https://doi.org/10.1016/j.compag.2020.105847.

Hong, H., Lin, J., Huang, F., 2020. Tomato disease detection and classification by deep learning. Proceedings - 2020 International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering, ICBAIE 2020, pp. 25–29 https://doi.org/10.1109/ICBAIE49996.2020.00012.

Howlader, R., Habiba, U., Faisal, R.H., Rahman, M., 2019. Automatic recognition of guava leaf diseases using deep convolution neural network. 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1–5.

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. https://doi.org/10.1109/ACCESS.2019.2914929.

Kabas, A., Fidan, H., Kucukaydin, H., Atan, H.N., 2022. Screening of wild tomato species and interspecific hybrids for resistance/tolerance to tomato brown rugose fruit virus (tobrfv). Chil. J. Agric. Res. 82, 189–196. https://doi.org/10.4067/S0718-58392022000100189.

Kamilaris, A., Prenafeta-Boldu, F.X., 2018. Deep learning in agriculture: a survey. Comput. Electron. Agric. 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016.

Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., Maschinot, A., Liu, C., Krishnan, D., 2020. Supervised contrastive learning. Advances in neural information processing systems 33, 18661–18673.

Kodors, S., Lacis, G., Sokolova, O., Zhukovs, V., Apeinans, I., Bartulsons, T., 2021. Apple scab detection using CNN and transfer learning. Agron. Res. 19, 507–519. https://doi.org/10.15159/AR.21.045.

Kukreja, V., Dhiman, P., 2020. A deep neural network based disease detection scheme for Citrus fruits. Proceedings - International Conference on Smart Electronics and Communication, ICOSEC 2020, 97–101. https://doi.org/10.1109/ICOSEC49089.2020.9215359.

Kunapuli, S.S., Bhallamudi, P.C., 2021. Chapter 22 - a review of deep learning models for medical diagnosis. In: Kumar, P., Kumar, Y., Tawhid, M.A. (Eds.), Machine Learning, Big Data, and IoT for Medical Informatics. Academic Press, pp. 389–404 Intelligent Data-Centric Systems.

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 521, 2278–2324.

Li, L., Zhang, S., Wang, B., 2021. Plant disease detection and classification by deep learning - a review. IEEE Access 9, 56683–56698. https://doi.org/10.1109/ACCESS.2021.3069646.

Liu, B., Ding, Z., Tian, L., He, D., Li, S., Wang, H., 2020. Grape leaf disease identification using improved deep convolutional neural networks. Front. Plant Sci. 11, 1–14. https://doi.org/10.3389/fpls.2020.01082.

McCulloch, W.S., Pitts, W.H., 1943. A Logical calculus of the Ideas Immanent in Nervous Activity. Pergamon Press PLC.

Minsky, M., Papert, S., 1969. Perceptron. MIT Press.

Mkonyi, L., Rubanga, D., Richard, M., Zekeya, N., Sawahiko, S., Maiseli, B., Machuve, D., 2020. Early identification of Tuta absoluta in tomato plants using deep learning. Sci. Afr. 10, e00590. https://doi.org/10.1016/j.sciaf.2020.e00590.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusi, A., Veness, J., Bellemare, M., Graves, A., Riedmiller, M., Fidjeland, A., Ostrovski, G., 2015. Human-level control through deep reinforcement learning. Nature 529, 33.

Omrani, E., Khoshnevisan, B., Shamshirband, S., Saboohi, H., Anuar, N.B., Nasir, M.H.N.M., 2014. Potential of radial basis function-based support vector regression for apple disease detection. Measurem. J. Int. Measurement Conf. 55, 512–519. https://doi.org/10.1016/J.MEASUREMENT.2014.05.033.

Orano, J.F.V., Maravillas, E.A., Aliac, C.J.G., 2019. Jackfruit fruit damage classification using convolutional neural network. 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management, HNICEM 2019 https://doi.org/10.1109/HNICEM48295.2019.9073341.

Ouali, Y., Hudelot, C., Tami, M., 2020. An Overview of Deep Semi-Supervised Learning. https://doi.org/10.48550/arXiv.2006.05278 arXiv:2006.05278 [cs, stat].

Oviedo-Mireles, J.C., Soto-Parra, J.M., Sanchez, E., Yanez-Munoz, R.M., Perez-Leal, R., Noperi-Mosqueda, L.C., 2021. Salicylic acid and nutrient immersion to maintain apple quality and bioactive compounds in postharvest. Notulae Botanicae Horti Agrobotanici Cluj-Napoca 49. https://doi.org/10.15835/nbha49312409.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372. https://doi.org/10.1136/bmj.n71.

Pan, W., Qin, J., Xiang, X., Wu, Y.A.N., Tan, Y.U.N., Xiang, L., 2019. A smart mobile diagnosis system for Citrus diseases based on densely connected convolutional networks. IEEE Access 7, 87534–87542. https://doi.org/10.1109/ACCESS.2019.2924973.

Park, H., Jeesook, E., Kim, S.H., 2018. Crops disease diagnosing using image-based deep learning mechanism. Proceedings of the 2nd International Conference on Computing and Network Communications, CoCoNet 2018, pp. 23–26 https://doi.org/10.1109/CoCoNet.2018.8476914.

Paul, A., Ghosh, S., Das, A.K., Goswami, S., Das Choudhury, S., Sen, S., 2020. A Review on Agricultural Advancement Based on Computer Vision and Machine Learning. vol. 937. Springer, Singapore. https://doi.org/10.1007/978-981-13-7403-6_50.

R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rangarajan, A.K., Purushothaman, R., Ramesh, A., 2018. Tomato crop disease classification using pre-trained deep learning algorithm. Procedia Comp. Sci. 133, 1040–1047. https://doi.org/10.1016/j.procs.2018.07.070.

Razmjooy, N., Mousavi, B.S., Soleymani, F., 2012. A real-time mathematical computer method for potato inspection using machine vision. Comput. Math. Appl. 63, 268–279. https://doi.org/10.1016/J.CAMWA.2011.11.019.

Rosenblatt, F., 1958. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. American Psychological Association.

Rozario, L.J., Rahman, T., Uddin, M.S., 2016. Segmentation of the region of defects in fruits and vegetables. Int. J. Comp. Sci. Inform. Security 14, 399–406. https://doi.org/10.1016/J.MEASUREMENT.2014.05.033.

Samajpati, B.J., Degadwala, S.D., 2016. Hybrid approach for apple fruit diseases detection and classification using random forest classifier. International Conference on Communication and Signal Processing, ICCSP 2016, pp. 1015–1019 https://doi.org/10.1109/ICCSP.2016.7754302.

Samuel, A.L., 1959. Machine learning. Technol. Rev. 62, 42–45.

Santos, L., Santos, F.N., Oliveira, P.M., Shinde, P., 2020. Deep learning applications in agriculture: a short review. In: Silva, M.F., Luıs Lima, J., Reis, L.P., Sanfeliu, A., Tardioli, D. (Eds.), Robot 2019: Fourth Iberian Robotics Conference. Springer International Publishing, Cham, pp. 139–151.

Sheikh, M.H., Mim, T.T., Reza, M.S., Hena, M.H., 2020. Leaf diseases detection for commercial cultivation of obsolete fruit in Bangladesh using image processing system. Proceedings of the 2019 8th International Conference on System Modeling and Advancement in Research Trends, SMART 2019, pp. 271–275 https://doi.org/10.1109/SMART46866.2019.9117505.

Sikirou, R., Ezin, V., Festonand, B., Afoha, E., Sètondji, A.P., Djindé, T.F., Ouessou, I., 2015. Geographical distribution and prevalence of the main tomato fungal wilt diseases in Benin. Int. J. Biol. Chem. Sci. 9, 603. https://doi.org/10.4314/ijbcs.v9i2.3.

Singh, A.K., Ganapathysubramanian, B., Sarkar, S., Singh, A., 2018. Deep learning for plant stress phenotyping: trends and future perspectives. Trends Plant Sci. 23, 883–898. https://doi.org/10.1016/j.tplants.2018.07.004.

Singh, U.P., Chouhan, S.S., Jain, S., Jain, S., 2019. Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease. IEEE Access 7, 43721–43729. https://doi.org/10.1109/ACCESS.2019.2907383.

Slavin, J.L., Lloyd, B., 2012. Health benefits of fruits and vegetables 1. Adv. Nutr., 506–516 https://doi.org/10.3945/an.112.002154.506.

Tani, H., Kotani, R., Kagiwada, S., Uga, H., Iyatomi, H., 2018. Diagnosis of multiple cucumber infections with convolutional neural networks. Proceedings - Applied Imagery Pattern Recognition Workshop 2018-Octob, 1–4 https://doi.org/10.1109/AIPR.2018.8707385.

Thet, K.Z., Htwe, K.K., Thein, M.M., 2020. Grape leaf diseases classification using convolutional neural network. Proceedings of the 4th International Conference on Advanced Information Technologies, ICAIT 2020, pp. 147–152 https://doi.org/10.1109/ICAIT51105.2020.9261801.

Tran, T.T., Choi, J.W., Le, T.T.H., Kim, J.W., 2019. A comparative study of deep CNN in forecasting and classifying the macronutrient deficiencies on development of tomato plant. Appl. Sci. (Switzerland) 9. https://doi.org/10.3390/app9081601.

Vakilian, K.A., Massah, J., 2013. An artificial neural network approach to identify fungal diseases of cucumber (cucumis sativus l.) plants using digital image processing. 46, pp. 1580–1588. https://doi.org/10.1080/03235408.2013.772321.

van Eck, N.J., Waltman, L., 2022. VOSviewer Manual. CWTS. p. 54.

Verma, S., Nizam, S., Verma, P., 2013. Biotic and Abiotic Stress Signaling in Plants. Springer, pp. 25–49 https://doi.org/10.1007/978-1-4614-6372-6_2 chapter 1.

Wakamori, K., Mizuno, R., Nakanishi, G., Mineno, H., 2020. Multimodal neural network with clustering-based drop for estimating plant water stress. Comput. Electron. Agric. 168, 105118. https://doi.org/10.1016/j.compag.2019.105118.

Wang, H.J., Mou, Q., Yue, Y., Zhao, H., 2020. Research on detection technology of various fruit disease spots based on mask R-CNN. 2020 IEEE International Conference on Mechatronics and Automation, ICMA 2020, pp. 1083–1087 https://doi.org/10.1109/ICMA49215.2020.9233575.

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.

Xin, Z., Jun, S., Yan, T., Quansheng, C., Xiaohong, W., Yingying, H., 2020. A deep learning based regression method on hyperspectral data for rapid prediction of cadmium residue in lettuce leaves. Chemom. Intell. Lab. Syst. 200, 103996. https://doi.org/10.1016/j.chemolab.2020.103996.

Zeng, Q., Ma, X., Cheng, B., Zhou, E., Pang, W., 2020. GANS-based data augmentation for citrus disease severity detection using deep learning. IEEE Access 8, 172882–172891. https://doi.org/10.1109/ACCESS.2020.3025196.

Zhang, Y., Wu, L., 2012. Classification of fruits using computer vision and a multiclass support vector machine. Sensors (Basel, Switzerland) 12, 12489–12505. https://doi.org/10.3390/S120912489.

Zheng, Z., Pan, S., Zhang, Y., 2019. Fruit tree disease recognition based on convolutional neural networks. Proceedings - 2019 IEEE International Conferences on Ubiquitous Computing and Communications and Data Science and Computational Intelligence and Smart Computing, Networking and Services, IUCC/DSCI/SmartCNS 2019, pp. 118–122 https://doi.org/10.1109/IUCC/DSCI/SmartCNS.2019.00048.

Zhou, X., Sun, J., Tian, Y., Lu, B., Hang, Y., Chen, Q., 2020. Development of deep learning method for lead content prediction of lettuce leaf using hyperspectral images. Int. J. Remote Sens. 41, 2263–2276. https://doi.org/10.1080/01431161.2019.1685721.

Zhou, C., Zhou, S., Xing, J., Song, J., 2021. Tomato leaf disease identification by restructured deep residual dense network. IEEE Access 9, 28822–28831. https://doi.org/10.1109/ACCESS.2021. 3058947.