



# Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture

Abhishek Upadhyay<sup>1</sup> · Narendra Singh Chandel<sup>2</sup> · Krishna Pratap Singh<sup>3</sup> · Subir Kumar Chakraborty<sup>4</sup> · Balaji M. Nandede<sup>5</sup> · Mohit Kumar<sup>6</sup> · A. Subeesh<sup>2</sup> · Konga Upendar<sup>7</sup> · Ali Salem<sup>8,9</sup> · Ahmed Elbeltagi<sup>10</sup>

Accepted: 29 December 2024 / Published online: 17 January 2025  
© The Author(s) 2025

## Abstract

Plant diseases cause significant damage to agriculture, leading to substantial yield losses and posing a major threat to food security. Detection, identification, quantification, and diagnosis of plant diseases are crucial parts of precision agriculture and crop protection. Modernizing agriculture and improving production efficiency are significantly affected by using computer vision technology for crop disease diagnosis. This technology is notable for its non-destructive nature, speed, real-time responsiveness, and precision. Deep learning (DL), a recent breakthrough in computer vision, has become a focal point in agricultural plant protection that can minimize the biases of manually selecting disease spot features. This study reviews the techniques and tools used for automatic disease identification, state-of-the-art DL models, and recent trends in DL-based image analysis. The techniques, performance, benefits, drawbacks, underlying frameworks, and reference datasets of more than 278 research articles were analyzed and subsequently highlighted in accordance with the architecture of computer vision and deep learning models. Key findings include the effectiveness of imaging techniques and sensors like RGB, multispectral, and hyperspectral cameras for early disease detection. Researchers also evaluated various DL architectures, such as convolutional neural networks, vision transformers, generative adversarial networks, vision language models, and foundation models. Moreover, the study connects academic research with practical agricultural applications, providing guidance on the suitability of these models for production environments. This comprehensive review offers valuable insights into the current state and future directions of deep learning in plant disease detection, making it a significant resource for researchers, academicians, and practitioners in precision agriculture.

**Keywords** Computer vision · Deep learning · Plant diseases detection · Vision transformers · Generative adversarial networks · Vision language models

## 1 Introduction

All cultures have their roots in agriculture. The agricultural sector employs approximately one billion individuals worldwide, which accounts for approximately 28% of the employed population (Anon.2018a). India, China and United States are the major cultivators globally, having the highest net cropped area (Anon. 2018b). Every aspect of crop cultivation must be considered to ensure a higher yield throughout the year. A global agricultural harvest is expected to be lost by 10–16% annually due to plant diseases, costing USD 220 billion. The global problem of food contamination induced by plant diseases is an ongoing issue that plant pathologists cannot afford to ignore (Yadav et al. 2021). Currently, fungi account for around 83% of plant-contagious diseases, phytoplasmas, and viruses for 9%, and bacteria for more than 7% (Pavlovskaya et al. 2018). Agricultural diseases are a particularly critical issue in crop production, affecting every field.

A plant disease refers to any modification that disrupts the inherent physiological processes of the plant. These pathogens can knowingly reduce the overall quantity and quality of harvest, affecting agricultural output. Several methods have been developed for diagnosing diseases to minimize serious harm. Molecular biology techniques, in particular, offer precise identification of pathogenic factors. Many farmers may not have access to these methods, though, and they are expensive and resource-intensive to obtain or execute (Sahu et al. 2021).

Therefore, it is preferable to detect diseases precisely and promptly to prevent such losses. The detection of plant diseases can be done manually or via computer-based technologies. The most evident symptoms of plant diseases to the human eye are spots on the leaves. On the other hand, some diseases do not manifest themselves on the leaves. In contrast, others manifest themselves in the latter stages and after the plants have already suffered significant damage.

Timely management of major crops necessitates careful oversight, including monitoring for diseases that could turn them into waste and addressing various issues promptly. Diseases can diminish plant productivity, and since each plant is susceptible to unique diseases, managing them with conventional disease control methods demands significant effort and skilled personnel. Farmers in most areas don't have enough resources and knowledge to consult professionals. In many situations, relying solely on visual observation for diagnosis is insufficient. The accuracy of diagnosis decreases when it is based solely on a single morphological characteristic. Computer vision (CV) and other methods should also be applied simultaneously to advance the accuracy of the diagnosis (Khakimov et al. 2022). The professional consultations are quite a costly and time-consuming process, which adds an extra financial strain to the farmer (Ferentinos 2018).

In such cases, it is advised that computerized systems would be the only way to quickly identify the disease utilizing various sophisticated algorithms and analytical tools, particularly with the aid of potent microscopes and other equipment. The main goal of this article is to give a comprehensive scenario of current findings in the field of plant disease detection using computer vision and deep learning. The scientific databases PubMed, Scopus, and Web of Science have been used to examine the research in the relevant fields in prior years. The implementation of digitalization in agriculture using deep learning and AI has progressed beyond the early conceptual stage. This study also highlights the challenges connected with implementing deep learning and computer vision techniques, along with their

technical specifics. This article will support the adoption of deep learning and computer vision-based systems on farms and help elucidate how these technologies can be integrated into agricultural operations.

In order to continue the conversation about certain topics that have made important contributions to the field of precision agriculture, this paper will focus on presenting and analyzing some recent and practical advancements in deep learning and computer vision for the purpose of detecting plant diseases. Understanding the future progress in this discipline requires analyzing the work that has already been done. By the time the conclusion of this review is reached, comprehension will be strengthened in the context of having a wide range of knowledge regarding the applications of deep learning in agriculture. The purpose of this article is to provide solutions to the following research questions to gain a better consideration of the developments in this field:

- Why is plant disease detection significant in precision agriculture?
- How far has deep learning and computer vision reached in this field, and what challenges remain?
- Which techniques and models show substantial development in plant disease detection?
- What are the main aspects that need to be considered for effective implementation in precision agriculture?

## 2 Review methodology and literature selection

### 2.1 Literature selection criteria and process

The criteria for inclusion and exclusion were developed to form the foundation of a rigorous literature review process, ensuring that the selected studies were relevant, current, and of high quality (Mahmud et al. 2023). In this comprehensive review of deep learning and computer vision in plant disease detection for precision agriculture, these criteria were carefully crafted to encapsulate the breadth and depth of existing research while maintaining a focus on practical and experimental implementations.

The inclusion criteria used to filter the literature for inclusion in the review are detailed below:

#### 2.1.1 Inclusion and exclusion criteria

The inclusion and exclusion criteria were designed to form the foundation of a rigorous literature review process, ensuring that the selected studies were relevant, current, and of high quality. In this comprehensive review of deep learning and computer vision in plant disease detection for precision agriculture, these criteria were carefully crafted to encapsulate the breadth and depth of existing research while maintaining a focus on practical and experimental implementations.

The inclusion criteria used to filter the literature for inclusion in the review are detailed below:

- **Relevance to topic:** The primary criterion for inclusion was the relevance of the study

to the topic of plant disease detection using deep learning and computer vision. Studies had to directly address one or more aspects of this topic, such as the development of DL models for disease identification, the application of computer vision techniques to analyze plant images, or the integration of these technologies in precision agriculture systems. Relevance was determined by examining the study's objectives, methodology, and findings to ensure alignment with the research focus.

- **Practical implementation and experimental evidence:** In order to provide a full overview of the current state of the art, studies offering practical implementations and experimental evidence of their findings were included. This includes papers describing the development, testing, and deployment of deep learning models and computer vision systems in real-world agricultural settings. Experimental evidence was crucial as it demonstrated the feasibility, effectiveness, and potential impact of the proposed methods. Studies including case studies, field trials, or extensive simulations were prioritized.
- **Methodological rigor:** The methodological rigor of a study was another essential inclusion criterion. The robustness of the data collection methods, research design, and analysis techniques used in each study was assessed. Studies employing well-established methodologies, large and diverse datasets, and thorough validation procedures were favored. This criterion ensured that the included studies provided reliable and generalizable results.
- **Contribution to knowledge:** Studies making significant contributions to the field were included. This includes papers introducing novel techniques, models, or frameworks, as well as those providing comprehensive reviews or meta-analyses of existing research. Studies identifying and addressing key challenges, gaps, and future directions in plant disease detection using deep learning and computer vision were also included.
- **Historical and foundational work:** While the review primarily focused on recent advancements, foundational studies published before the last five years were also included. These seminal papers provided the basic knowledge and theoretical underpinnings essential for understanding current developments. Foundational works from as far back as 1943 and 2005 were also considered, particularly if key concepts, methods, or technologies significantly influencing subsequent research were introduced.
- **Language and accessibility:** To ensure the review was accessible to a wide audience, only studies published in English were included. This criterion helped maintain consistency and comprehensibility across the reviewed literature. Additionally, studies published in open-access journals or repositories were favored to facilitate easy access for researchers and practitioners.

The exclusion criteria used to filter the literature for inclusion in the review are detailed below:

### 2.1.2 Exclusion criteria

- **Irrelevance to topic:** Studies not directly addressing plant disease detection using computer vision and deep learning were excluded. This included research focused on other aspects of agriculture or different applications of deep learning and computer vision

not pertaining to plant disease detection. Irrelevance was determined by reviewing the study's abstract, introduction, and conclusion sections.

- **Lack of practical implementation and experimental evidence:** Studies purely theoretical or conceptual, lacking practical implementation or experimental evidence, were excluded. While theoretical insights were valuable, the focus of this review was on applied research demonstrating real-world applicability and effectiveness. Papers without empirical validation or practical case studies were filtered out.
- **Methodological weaknesses:** Studies with significant methodological weaknesses were excluded. This included research with poorly designed experiments, small or biased datasets, inadequate validation techniques, or lack of reproducibility. Methodological weaknesses were identified through a critical appraisal of the study's data collection, research design, and analysis methods.
- **Limited contribution to knowledge:** Studies not making a substantial contribution to the field were excluded. This included papers reiterating well-known findings without offering new insights, failing to address significant challenges, or lacking originality in their approach. Additionally, studies not providing a comprehensive literature review or contextualizing their findings within the broader field were filtered out.
- **Non-English publications:** To maintain consistency and accessibility, studies published in languages other than English were excluded. This criterion ensured that the reviewed literature could be accurately interpreted and analyzed by the broader scientific community.

## 2.2 Selection process

### 2.2.1 Keyword identification

The first step in the literature selection process was the identification of relevant keywords. These keywords were derived from the research questions and objectives of the review and included terms such as “plant disease detection”, “deep learning in agriculture”, “computer vision”, “precision agriculture” and related synonyms.

### 2.2.2 Search strings formation

Next, search strings were generated by combining the identified keywords and their synonyms. Examples of search strings used include “(plant disease detection AND deep learning)”, “(plant disease detection OR computer vision)”, and “(deep learning AND agriculture)”. These search strings were designed to capture a broad range of relevant studies.

### 2.2.3 Database search

Extensive searches were conducted in multiple academic databases to ensure comprehensive coverage of the literature. The databases included PubMed (<https://pubmed.ncbi.nlm.nih.gov/>), Semantic Scholar (<https://www.semanticscholar.org/>), Scopus ([www.scopus.com](https://www.scopus.com)), Google Scholar (<https://scholar.google.com/>), IEEE Xplore (<https://ieeexplore.ieee.org/>), ScienceDirect (<https://www.sciencedirect.com/>) and Web of Science (<https://clarivate.com/>)

[products/web-of-science/](#)). The selection of these databases was based on their relevance to the following fields of computer science, engineering, and agricultural research.

### 2.2.4 Initial screening

An initial screening of the search results was conducted based on the title and abstract of each study. Studies that appeared irrelevant or did not meet the inclusion criteria were excluded at this stage. The preliminary screening procedure assisted in reducing the total number of studies to a more manageable number for detailed evaluation.

### 2.2.5 Full-text review

The remaining studies underwent an examination of the complete text to evaluate their relevance and the level of methodological rigor, and contribution to the field. During this stage, the criteria for both inclusion and exclusion were applied in detail, evaluating each study's objectives, methodology, findings, and overall quality.

### 2.2.6 Final selection

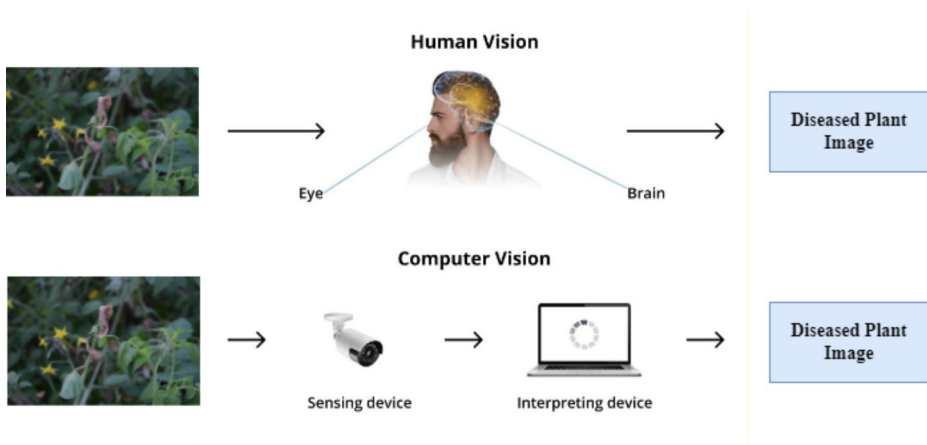
After the full-text review, 278 articles were chosen for final review and used in this work. This methodology ensures that a rigorous and systematic approach for selecting relevant literature was employed, thereby providing a comprehensive overview of recent developments and future directions in deep learning and computer vision for disease detection in precision agriculture.

## 3 Computer vision for image data acquisition

Computer vision (CV) aims to create a method that enables computers to “see” and comprehend the feature-based information in digital images and videos. It is a branch of science that enables computers to analyze, record, comprehend, and process visually traceable objects (Lawaniya 2020). CV algorithms leverage computers to comprehend and recognize the patterns from input images or video, as shown in Fig. 1. The successful implementation of deep learning models, particularly convolutional neural networks (CNNs), has been demonstrated in various CV applications in recent years. These applications include traffic detection (Yang et al. 2019a, b), recognition of medical images (Sundararajan et al. 2019), text identification (Melnik et al. 2019), facial recognition (Kumar and Singh 2020), and crop yield predictions (Damos 2015), etc.

Many deep learning-based methods have recently been used in agriculture to detect plant diseases and pests. Many domestic and international companies have built WeChat applets and object/disease identification apps. Additionally, it is made simpler and less expensive by automated disease diagnosis that requires the observation of symptoms on plant leaves. This section eases the use of machine vision to offer robot guidance and autonomous process control using images (Kumar and Singh 2020; Tsafaris et al. 2016; Yang et al. 2019a, b). In machine vision (MV) an imaging sensor is used to take plant photographs and assess if they have pests and disease attacks (Lee et al. 2017). Plant diseases and pest detection

## Human Vision vs Computer Vision



**Fig. 1** Perception of human vision and computer vision for disease detection

technologies based on machine vision have partially replaced visual identification (Martin et al. 2022).

Machine vision has been particularly beneficial in diagnosing the severity of diseases. An element of the MV, i.e., deep learning (DL), can be employed to ascertain the severity of diseases affecting plants and animals. MV plays a significant role in farm operations since it aids in the application of robotics for plant disease diagnosis, pest control, estimating yield and precision agriculture using visual information. The MV systems use imaging sensors such as RGB, hyperspectral, and thermal cameras to take images of the croplands and employ picture processing techniques such as pattern recognition algorithms, specifically trained neuronal networks like CNNs. Such automation assists the farmer in understanding disease development in crops and the use of effective control measures enhancing the management of crops and their yields. Despite the potential use of just-in-time crop monitoring using MV, several constraints that include equipment price, variability of the environment and incapability of acquiring enough data persist in hindering its practical use. It also categorizes diseases and prevents infections from being discovered belatedly (Lee et al. 2017). The technique might be used to examine and assess how much damage they have already done (Martin et al. 2022).

Therefore, it is essential to find the technology that is reasonable in price and can monitor plants to find and diagnose diseases, insect pests, etc. as farmers can take the appropriate procedures and safeguards after recognizing the diseases. The rapid development of software and hardware has significantly advanced image processing in agriculture. Image processing techniques (IPT) are used to manipulate and analyze images from visible light cameras, infrared imaging devices, and other electromagnetic spectrum sensors. Much fascinating research has been done on hyperspectral techniques for agricultural pest and disease identification (Oerke et al. 2016; Yu et al. 2013, 2018; Martinez-Martinez et al. 2018; Azadbakht et al. 2019). However, hyperspectral equipment is expensive and difficult for average farmers and extension workers to operate (Jiang et al. 2019).

Image processing methods used in pest and disease recognition utilizing RGB images evolve daily. In most situations, the affected plant's disease symptoms are visible (Atole and Park 2018; Zhang et al. 2018a, b, c). Image processing algorithms can be designed to operate accurately, quickly, and affordably to identify these diseases from common digital pictures (Ngugi et al. 2021).

- **Thresholding Technique:** The simplest technique is thresholding with the purpose of changing gray-scale pictures to black and white pictures. Regions of interest such as diseased sections of a leaf can be differentiated from the rest of the pixels by the thresholding technique, where an intensity threshold is set and only pixels above this threshold are included.
- **Edge Detection:** Edge detection algorithms are the ones that detect the edges of the infected areas in the images. Canny and Sobel are the two most popular edge detectors that are used to identify areas within an image that undergo sharp changes in pixel intensity for example leaf spots, lesions or discolouration.
- **K-means Clustering:** K-means, a widely used clustering method is an unsupervised learning technique that tends to classify image pixels by the nearest color. To help with that, algorithms are very useful when classifying affected areas of the plant in RGB images or hyperspectral images and healthy portions of plant imagery.
- **Region-based Segmentation:** This portion of the static approach is implemented to sub divide into parts of the image that appear to possess similar or same characteristics. For instance, when using the watershed algorithm one would be able to identify the individual disease spots in a leaf that are overlapping hence improving accuracy in disease diagnosis.
- **Convolutional Neural Networks (CNNs):** In deep learning, CNNs are a type of neural network that has completely transformed IPT in agriculture. They do most of the learning of the image data given that it involves shape, texture, and color of the leaves among other things to diagnose the diseases accurately and quickly. The use of CNNs in plant disease identification has become the norm.
- **Support Vector Machine (SVM):** SVMs, along with the HOG method of feature extraction, help in the discrimination of diseased features from healthy ones using the pixel patterns. It is a useful method for solving binary classifiers such as oriented towards classification of disease infested plants and their healthy counterparts.

The following benefits are specifically provided by IPT adoption:

- IPTs can accurately and efficiently identify agricultural diseases by utilizing photos of foliage, stems, fruits, and flowers.
- The size of the discoloured or deformed patch concerning the magnitude of the entire fruit, flower, or leaf can be used to evaluate the severity of the disease (Tiwari and Tarum 2017; Rothe and Kshirsagar 2014; Padol and Sawant 2016; Bierman et al. 2019; Dey et al. 2016).
- Monitoring the disease development in plants is important for seeing features like the contagion stage and spotting signs that are usually invisible to humans (Padol and Sawant 2016; Barbedo et al. 2016; Wang et al. 2017).
- IPTs will also assist researchers in evaluating the traits of novel crop cultivars under



laboratory evaluation for disease resistance (Bierman et al. 2019).

- People who live in rural places may easily and affordably access knowledge through IPTs (Prakash et al. 2017; Anand et al. 2016).
- The correct diagnosis will result in the more cost-effective usage of pesticides. In addition to lowering production costs, improved admittance to human professionals who can be accessed from distant locations rather than having physical visits to each farm will also preserve the environment and improve accessibility to heavily regulated markets (Anand et al. 2016; Krithika and Grace Selvarani 2017).

#### 4 Imaging techniques available for plant health status detection

Various imaging techniques, such as hyperspectral imaging, thermal imaging, and RGB imaging, have been utilized by several researchers to gather data for examining plant health status (Fig. 2). Fluorescence, thermal, hyperspectral, multispectral, visible, photo-acous-



**Fig. 2** Numerous imaging techniques utilized to detect plant diseases

tic, tomographic, thermographic, and MRT are useful imaging techniques (Singh et al. 2020). Additionally, 3D imaging techniques can also be utilized in combination with other techniques.

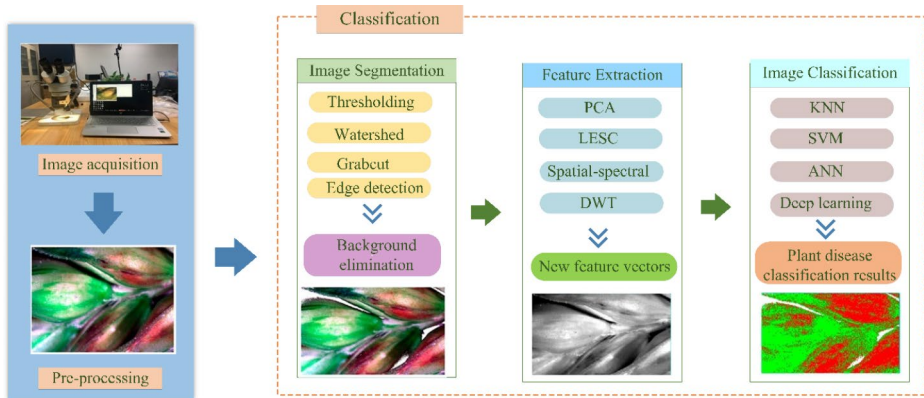
## 5 Imaging cameras and sensors for plant disease detection

Thermal and hyperspectral sensors are the most effective methods for detecting early-stage pathogen contagions in crops. However, later-stage infection severity may also be detected by RGB, multispectral, and hyperspectral sensors (Maes and Steppe 2019).

Among different imaging sensors, the digital RGB imaging sensors have been most commonly used for plant disease detection. RGB cameras are frequently less expensive and easily available. They can also be employed to take clear still images (MacPherson et al. 2022). Cameras use RGB (Red, Green, and Blue) sensors, which are often used to identify red, green, and blue values in pixels. These cameras provide images that depict the intensity of three different colours, allowing for the assessment of biomass in crops (Gruner et al. 2019; Roth and Streit 2017; Viljanen et al. 2018). When estimating biomass, RGB cameras are deployed in conjunction with multispectral and near-infrared cameras to improve precision (Roth and Streit 2017). Red filters are substituted for near-infrared filters in modified RGB cameras (Berra et al. 2017; Nijland et al. 2014). Commercial RGB cameras are inexpensive yet lackluster in terms of spectral resolution (Nijland et al. 2014). Not all wavelengths within 380 nm to 750 nm electromagnetic spectrum range of RGB cameras are appropriate for precise disease detection in crops (Bock et al. 2020). RGB colour information is transformed to other colour spaces like hue saturation value (HSV), LAB (lightness),  $YCbCr$  (blue difference, luma component, and red difference chroma components), and other colour spaces that are particularly helpful in diagnosing plant diseases. The lack of resolution in images from RGB cameras makes it impossible to distinguish between different severity levels of disease (Zhang et al. 2018a, b, c). RGB cameras, on the other hand, have the potential to acquire images with a high spatial resolution, which in turn provides better spatial features for plant disease monitoring and detection. This potential is one of the primary advantages of RGB cameras over multispectral systems. Using RGB cameras effectively ensures that the lighting and colour of the photos are consistent. Photos taken consistently will show less error in determining healthy and sick plants.

The fundamental hyperspectral image classification procedure for identifying plant diseases is shown in Fig. 3.

Agricultural analytics performed with multispectral cameras provide the best results, these can take photos with spatial resolution and measure near-infrared reflectance (Nhamo et al. 2020). Vegetative indices are produced using multispectral and NIR cameras using near-infrared or other light bands (Adao et al. 2017; Geipel et al. 2016; Iqbal et al. 2018). Various spectral bands, primarily red, red-edge, green, blue, and near-infrared, are used by multispectral cameras. Based on bandwidth, they can be divided into two groups: broadband and narrowband (Deng et al. 2018). Multispectral cameras are utilized for most aerial photos to monitor crop health concerns since they can calculate indices like NDVI, NIR and other metrics (Viljanen et al. 2018; Nhamo et al. 2020; Geipel et al. 2016; Zaman-Allah et al. 2015; Kalischuk et al. 2019). There are various other studies that have used multispectral



**Fig. 3** Fundamental procedure in classifying plant diseases using hyperspectral imaging

cameras for disease detection (Di Gennaro et al. 2016; Zhang et al. 2018a, b, c; Albetis et al. 2018; Calderon et al. 2014; Dash et al. 2018; Khot et al. 2015; Nebiker et al. 2016).

The primary distinction between hyper and multispectral cameras is that each pixel in the image produced by a hyperspectral camera collects light from various tiny size bands. In contrast, multispectral camera images are produced by a continuous spectrum (Lowe et al. 2017; Adao et al. 2017). Multispectral cameras can catch biomolecule-reflected light, but light bandwidth and location allow us to distinguish responses. These cameras are particularly good at detecting light emitted by biomolecules, including chlorophyll (Cilia et al. 2014; Gevaert et al. 2015), mesophyll (Lowe et al. 2017), xanthophyll (Proctor and He 2015), and carotenoids (Cilia et al. 2014; Gevaert et al. 2015). The expensive nature of the cameras (Adao et al. 2017; Deery et al. 2014) and the enormous amount of unusable data produced when they are not calibrated properly are the main drawbacks of employing a hyperspectral camera (Lowe et al. 2017; Saari et al. 2017). Hyperspectral cameras are mostly employed to improve upon the limitations of multispectral cameras. Identifying and differentiating target objects requires hyperspectral cameras, which can record details with fewer spectral variances. Hyperspectral cameras have significantly advanced image processing compared to conventional cameras (Thomas et al. 2018). They can identify plant stress with potential etiological factors (pathogen/disease).

The temperatures of the objects are shown as a thermal picture by thermal cameras, which collect infrared light between 0.75 and 1000  $\mu\text{m}$  (Costa et al. 2013). Thermal cameras are cheaper than other spectrum cameras, but the conversion of RGB cameras to thermal imaging is feasible (Mahajan and Bundel 2016). Thermal cameras were initially used to examine drought stress in crops (Deery et al. 2014; Mahajan and Bundel 2016; Gago et al. 2015; Granum et al. 2015). Thermal pictures have limited resolution compared to photos taken by other large cameras but contain the temperature of the nearby objects (Costa et al. 2013). According to Calderon et al. (2013) and Smigaj et al. (2015), thermal sensors are also used to monitor crops and detect agricultural diseases. Yang and coworkers created a technique that uses thermal photography for the early diagnosis of diseases in tea (Yang et al. 2019a, b). Thermal sensors perform better than multispectral and hyperspectral ones for observing drought stress (Ludovisi et al. 2017; Zhou et al. 2020). The studies that have reported disease detection using several sensors are listed below in Table 1.

**Table 1** Disease identification/detection/recognition/classification using various optical sensors

S. No.	Plant	Targets	Diseases	Optical sensors	References
1.	Ginseng	Identification	Water molds	LED	Gudkov et al. (2023)
2.	Olive	Identification	Verticillium wilt	LED	Calderon et al. (2013)
3.	Tobacco	Identification	Leaf mosaic	LED	Lei et al. (2017)
4.	Grapefruit	Detection	Citrus canker	Laser	Saleem et al. (2020a, b)
5.	Apple	Detection	Fruit rot	Laser	Matveyeva et al. (2022)
6.	Orange	Detection	Citrus canker	Laser	Marcassa et al. (2006)
7.	Chilli	Identification	Leaf curl	FTIR	Agustika et al. (2022)
8.	Wheat	Identification	Stripe rust	FTIR	Zhao et al. (2017)
9.	Cucumber	Identification	Powdery Mildew	Vis-NIR	Atanassova et al. (2019)
10.	Potato	Detection	Zebra Chip Disease	NIR	Liang et al. (2018)
11.	Pumpkin	Detection	Powdery Mildew	Vis-NIR	Rivera-Romero et al. (2020)
12.	Alfalfa	Identification	Phymatotrichopsis root rot	RGB	Mattupalli et al. (2018)
13.	Banana	Detection and classification	Bunchy top virus and Xanthomonas wilt	RGB	Gomez et al. (2020)
14.	Sugarbeet	Severity classification	Leaf spot ( <i>Cercospora beticola</i> Sacc.)	RGB	Ozguven (2018)
15.	Grapevine	Classification (Healthy or Diseased)	ESCA disease	RGB	Kerkech et al. (2018)
16.	Peanut	Identification	<i>Cercosporidium personatum</i> spot	RGB	Valasek et al. (2016)
17.	Potato	Detection	Late Blight	RGB	Sugiura et al. (2016)
18.	Wheat	Detection	Leaf Rust	RGB	Ashourloo et al. (2014)
19.	Rice	Severity classification	Sheath Blight	RGB	Zhang et al. (2018a, b, c)
20.	Tomato	Identification	Target spot and Bacterial spot	Multispectral	Abdulridha et al. (2019)
21.	Orange	Detection	Citrus Greening	RGB-D + Multispectral	Suproteem et al. (2016)
22.	Grapevine	Identification	Flavescence dorée	Multispectral	Al-Saddik et al. (2017)
23.	Grapevine	Identification	-	RGB + Infrared	Kerkech et al. (2020)
24.	Grapevine	Identification	ESCA complex and Flavescence dorée	Multispectral	Gallo et al. (2017)
25.	Pinewood	Identification	<i>Bursaphelenchus xylophilus</i>	Multispectral	Qin et al. (2021)

**Table 1** (continued)

S. No.	Plant	Targets	Diseases	Optical sensors	References
26.	Banana	Classification	Wilt	Multispectral	Ye et al. (2020)
27.	Tomato	Identification	Bacterial Spot	Hyperspectral	Abdulridha et al. (2019)
28.	Pears	Detection	Fire blight	Hyperspectral	Schoofs et al. (2020)
29.	Olive	Identification	-	Hyperspectral	Calderon et al. (2013)
30.	Forest Trees	Identification	-	Hyperspectral	Sandino et al. (2018)
31.	Rapeseed	Identification	Fungal Stress (Alternaria)	Thermal	Baranowski et al. (2015)
32.	Almonds	Detection	Red Leaf Blotch	Hyperspectral + Thermal	(Lopez-Lopez et al. 2016)
33.	Citrus	Detection	Huanglongbing	Thermal	Sankaran et al. (2013)
34.	Tobacco	Detection	Mosaic Virus	Thermal	Wang et al. (2013)
35.	Wheat	Detection	Fusarium Blight	Thermal	(Mahlein et al. 2019)
36.	Cucumber	Detection	Fusarium Wilt	Thermal	Wang et al. (2013)
37.	Banana	Detection	Wilt Spot	Thermal	Anasta et al. (2021)
38.	Tomato	Detection	Powdery Mildew	Thermal	Raza et al. (2015)
39.	Citrus	Detection	Greening disease	Depth	Suproteem et al. (2016)
40.	Tree	3D-Segmentation	-	Depth	Xia et al. (2015)
41.	Agricultural crop	Stress detection	-	Depth	Mahlein (2016)
42.	Wheat	Detection	Fusarium spp.	Depth	Paulus et al. (2013)

## 6 Traditional image processing techniques for plant disease detection

Machine vision approaches for plant disease and pest detection often use traditional image processing algorithms or manual feature development using classifiers (Lee et al. 2017). Plant diseases and pests' identification has several difficulties in a complex natural environment, including tiny differences between the lesion area and the backdrop, poor contrast, wide changes in the scale of the lesion area and different kinds, and significant noise in the lesion picture.

Colour, form, and texture are the three main characteristics of plant imagery. The shape is less useful for identifying plant diseases than colour and texture (Hlaing and Zaw 2018); combining texture and colour characteristics, Hlaing and Zaw classified tomato plant disease. They discovered the texture data, which included information on the shape, location and size, using the scale invariant feature transform (SIFT).

Traditional image-based disease detection and classification was done by digital image processing. Traditional image processing is divided into several steps, from image pre-processing to disease classification. Image pre-processing includes filtering, contrast limited adaptive histogram equalization (CLAHE) algorithm, etc., methods, while segmentation is

done by thresholding, clustering, histogram, compression, region growth, etc. After that, the feature of the disease was extracted by using local binary pattern (LBP), speeded-up robust features (SURF), histogram of oriented gradients (HOG), gray-level co-occurrence matrix (GLCM), histogram features, etc. methods. In the end, disease classification was done by using a SVM, naive bayes, decision tree, K-neural network (KNN), random forest, neural networks, fuzzy classifiers, etc. (Dhingra et al. 2018).

Automatic methods for disease detection in soybean crop were developed in 2015 (Dandawate and Kokare 2015) and 2023 (Kumar et al. 2023). In order to create an HSV colour space, they transformed the RGB image. Segmentation was done using colour and cluster-based techniques. Inferring the type of plant from its leaf form was done using the SIFT approach. Using colour texture traits and discriminant analysis, Pydipati and colleagues (2006) discovered citrus diseases. Additionally, they used the colour co-occurrence method (CCM) to see if the statistical classification methods and hue saturation and intensity (HSI) colour attributes could be used to distinguish the diseased leaves. An accuracy of greater than 0.95 was attained using this method (Pydipati et al. 2006). The following techniques may be employed to discern the presence of these pathogens, which have the potential to infiltrate various plant parts such as stems, vegetables, fruits, and more:

- Identifying and classifying the diseases.
- Determining the affected region.
- Retrieving the affected area's feature set.

This technique uses plant diseases and pests to construct the image scheme and choose the appropriate light source and shooting angle. This method ensures that images have uniform lighting. Designed imaging systems can reduce the complexity of constructing a standard algorithm but lead to an increase in application costs. It is sometimes difficult to expect standard algorithms to completely remove scene changes from recognized results to work in a natural environment (Dell' Aquila 2009).

Fairly consistent results were obtained across various feature extraction methods. In a nutshell, the standardization of the described methodologies has not yet been established and realized. The automatic detection of plant diseases has been a long-running research topic. Researchers reported highly satisfactory results using a relatively small number of images for training and testing. This study shows that discriminant analysis, especially linear discriminant analysis, and backpropagation neural networks outperform their competitors significantly. However, with recently introduced optimized deep neural networks, the overall recognition performance of the models has significantly improved. Deep convolutional neural networks can yield superior results when utilized more effectively, making them particularly beneficial for handling large data sets. Artificial neural networks (ANNs) are a popular deep learning method for image processing and categorization. ANNs are mathematical models that link one another similarly to how neurons and synapses do in the human brain (Ferentinos 2018). The neural networks are taught to operate on a comparable collection of data after being educated into a model using formerly recognized data. The operation of biological nerve systems significantly impacts ANNs, which are computational systems. ANNs primarily consist of several connected computational nodes, which merge dispersed to learn from the input and optimize the final output. The foundation of innumerable ANNs is that each neuron will continue to take in input and carry out an action. One

perceptive scoring function will relate the input raw picture matrices to the class score in the entire network (the weight). The last layer will include class-related loss functions, and normal ANN approaches apply.

Only the portion of the leaf damaged by the disease was removed by Al-bayati and Ustundag (Al-bayati and Üstündağ 2020). They additionally employed feature fusion, which aided in feature reduction. It should be ensured that the resources are available because image-based detection calls for many of them. The multilayer ANN serves as the basis for the underlying model. However, a convolutional layer executes kernel operations across various parts of the supplied image. The resultant representation is not affected by operations like rotation or translation. It has been demonstrated that these features perform better than the conventional features previously utilized in detecting plant diseases. Previous studies on using hyperspectral pictures in plant disease diagnosis have shown that categorization algorithms frequently use a correlation-based selection procedure despite the hyperspectral classification of plant diseases using the complete spectrum. Table 2 lists the research methods in hyperspectral image classification techniques for locating, identifying, and mapping plant diseases.

## 7 Deep learning-driven computer vision models

Deep learning, a subset of machine learning, excels at processing unstructured data. Deep learning outperforms standard machine learning. Computer models can gradually learn properties from input at different processing stages (Mathew et al. 2021). Deep Learning (DL) was first described in an article published by Hinton and Salakhutdinov (2006) in *Science*. Deep learning extracts data features using several hidden layers, each as a perceptron. Combining low-level features with abstract high-level features can significantly reduce the risk of getting stuck in local minima. Deep learning overcomes the limitation of traditional algorithms that rely on artificially engineered features, drawing increased interest from researchers. It is effectively used in recommendation systems, computer vision, pattern recognition, speech recognition, and natural language processing (NLP) (Liu et al. 2017).

In contrast to other image recognition techniques, deep learning-based image recognition technology does not require the extraction of specific features; instead, it finds the right features through iterative learning (Backpropagation), allowing it to acquire both global and contextual features from images while also being more robust and accurate at recognizing objects. For the analysis of multidimensional data, photographs, CNN, and DL in general have been developed (Martinelli et al. 2015). The underlying features of a picture can only be extracted using traditional manual image classification and identification methods, and it is challenging to extract the complex features information (Fergus 2012). Deep learning can eliminate this barrier. Unsupervised learning from the original image can reveal low-level, middle-level, and high-level semantic properties. Plant disease and pest image detection have great promise with deep learning. Recently developed deep neural network models include stack de-noising autoencoder (SDAE), deep belief network (DBN), deep Boltzmann machine (DBM), and deep convolutional neural network (CNN) (Bengio et al. 2013).

A computer model that uses the DL method to machine learning mimics the biological pathways of a human (McCulloch and Pitts 1943). In contrast to conventional neural networks, artificial neural networks used in deep learning have a variety of processing lay-

**Table 2** List significant contributions to plant disease diagnosis, identification, and mapping using hyperspectral image classification techniques

S. No.	Crop and diseases	Scales	Targets	Procedure and architecture	Performance metrics	References
1.	Strawberry, Anthracnose	Canopy	Early detection	BP, Machine learning, deep learning neural network, and CNN were chosen to develop a strawberry anthracnose recognition model using SPA, CARS, IRF, and TF combinations.	For healthy, asymptomatic, and symptomatic samples, the IRF + TF + BP model test set had 99.1%, 93.5%, and 94.5% accuracy, 100% recall, 94%, and 93% recall, and F1 values of 0.9955, 0.9375, and 0.9374, respectively.	Liu et al. (2023a, b)
2.	Soybean, frog eye leaf spot, anthracnose, rhizoctonia aerial blight, yellow mosaic virus, and soybean mosaic virus	Leaf	Disease classification	Mask R-CNN algorithm with Res Net 50 module was used for the classification of soybean leaf diseases.	Classification accuracy was determined at three levels of minimal detection confidence: 0.80, 0.85, and 0.90. The maximum detection accuracy was achieved at the 0.90 level of minimum detection confidence, which is greater than 85%.	Kumar et al. (2023)
3.	Apple, valsa canker	Branches	Disease detection	DC-CNN models were created using several spectral pre-processing algorithms. The models' classification mechanism was visualized using the gradient-weighted class activation mapping (Grad-CAM) approach with saliency maps and gradient histograms.	The DC-CNN model based on pictures and spectra after multiple scattering correction (MSC) pretreatment performed best for early apple valsa canker detection, with 98% accuracy.	Cui et al. (2022)



**Table 2** (continued)

S. No.	Crop and diseases	Scales	Targets	Procedure and architecture	Performance metrics	References
4.	Anthraco nose fungus in Olives	Fruit	Disease detection	Each olive's hyperspectral images (450–1050 nm) were taken until disease symptoms appeared. ResNet101 was developed to process two-class 61-band hyperspectral images.	Increasing the sample size with data augmentation yielded good accuracy (91.8%) and sensitivity (96.88%) but low specificity (73.08%) due to false negatives in the initial days after inoculation.	Fazari et al. (2021)
5.	Grapevine vein-clearing virus (GVCV)	Leaf/ Vine	Disease identification and classification	A statistical method was used to distinguish healthy and GVCV vine reflectance spectra. Pixel-wise and image-wise categorization were done simultaneously using deep learning architectures and traditional machine learning.	Temporal analysis revealed that GVCV vines' spectral signals were more distinct at VIS wavelengths (449–461 nm, $p < 0.1$ ). Achieving an accuracy of 71% and 75% for 5-fold cross-validation, respectively, and 50% for testing in both networks, the automated 2D-CNN and 3D-CNN feature extractors provided the image-based machine learning pipeline.	Nguyen et al. (2020)
6.	Citrus and citrus bacterial canker	Leaf/ fruit/ plant	Disease severity classification	Function-based on Neural network radial: KNN with spectral vegetation indices	92% canker detection at fruit scale and 100% at plant scale; 94%, 96%, and 100% by RBF and 94%, 95%, and 96% by KNN for three levels at the leaf scale	Supro-teem et al. (2016)
7.	Wheat, Fusarium head blight	Spike	Disease identification	SVM with reflectance and SVIs	Using SVIs and reflectance, 76.0% and 77.0% for multiclass classification, 95.0% and 99.0% for two classes of classification	Hlaing and Zaw (2018)
8.	Soybean, charcoal rot	Stem	Disease identification	3D convolutional neural network	95.73%	Nagasubramanian et al. (2019)

**Table 2** (continued)

S. No.	Crop and diseases	Scales	Targets	Procedure and architecture	Performance metrics	References
9.	Sugar beet, Cercospora leaf spot/ powdery mildew/sugar beet rust	Leaf	Disease identification	Spectral angle mapper (SAM)	97.23% for powdery mildew 14 days after inoculation (dai); 98.9% for Cercospora leaf spot at eight dai; 61.70% for sugar beet rust at 20 dai.	Mahlein et al. (2012)
10.	Wheat, yellow rust ( <i>Puccinia striiformis</i> )	Leaf	Disease detection	Self-organizing map /Quadratic Discriminant Analysis (QDA)	Around 99% by using SOM NN, 94.5% by using QDA	Moshou et al. (2005)

---

```

# Initialize parameters
initialize_transformer_weights()

# Split image into patches
patches = split_image_into_patches(image)

# Linear projection of flattened patches
projected_patches = [linear_projection(patch) for patch in patches]

# Add position embeddings
position_embeddings = create_position_embeddings(len(projected_patches))
encoded_patches = projected_patches + position_embeddings

# Transformer encoder layers
for each layer in transformer_layers:
    # Multi-head self-attention
    attention_output = multi_head_self_attention(encoded_patches)
    # Add & Normalize
    encoded_patches = layer_norm(attention_output + encoded_patches)
    # MLP
    mlp_output = mlp(layer_norm(encoded_patches))
    # Add & Normalize
    encoded_patches = layer_norm(mlp_output + encoded_patches)

# Classification head
class_output = classification_head(encoded_patches)

# Compute loss and backpropagation
loss = compute_loss(class_output, true_labels)
gradients = backward_propagation(loss)
update_transformer_weights(gradients)

# Function definitions (split_image_into_patches, linear_projection,
create_position_embeddings, multi_head_self_attention, layer_norm, mlp,
classification_head, compute_loss, backward_propagation, update_transformer_weights)
would be implemented here.

```

---

**Algorithm 1**

ers (Ferentinos 2018). It entails several phases: data gathering, picture categorization, and result interpretation. Artificial neural networks for image categorization come in a variety of forms: CNNs, generative adversarial networks (GANs), and recurrent neural networks (RNNs). Among these CNNs are the most often utilized for identifying and categorizing plant diseases.

Some researchers also used Hyperspectral Imaging (HSI) and DL models to more clearly observe plant disease signs in plant disease detection. A comprehensive assessment of the DL with the HSI approach was done (Signoroni et al. 2019). A thorough evaluation of many DL models, including 2D-CNN, LSTM/GRU, and a hybrid LSTM/GRU having 2D-CNN, was conducted to prevent overfitting and boost accuracy.

## 7.1 Convolutional neural networks (CNNs) based deep learning approach

Convolutional neural networks, akin to traditional ANNs, are composed of neurons that acquire the ability to optimize themselves. The fact that CNNs are primarily employed in feature representation inside pictures is the only distinguishing feature between CNNs and conventional ANNs. The network's appropriateness for image-focused activities is enhanced, and the number of setup parameters is decreased (O'Shea and Nash 2015). The GoogLeNet (Brahimi et al. 2018), Inceptionv3 (Ahmad et al. 2020), VGG19 (Ahmad et al. 2020), EfficientNet, ResNet50 (Ahmad et al. 2020), DenseNet (Tian et al. 2019), Xception (Verma et al. 2019), and MobileNet (Bi et al. 2019) are pre-trained network models. They have been used for various computer vision applications, including image classification, image production, anomaly detection, neural style transfer, image captioning, and more.

## 7.2 Plant disease identification using deep learning architectures

The application of DL models for disease recognition in crops is expanding quickly (Ferentinos 2018; Carranza-Rojas et al. 2017; Yang and Guo 2017). Using aerial images, CNNs are fundamental deep-learning methods for identifying plant diseases. They comprise potent modelling approaches that recognize intricate patterns in vast volumes of data (Ferentinos 2018). Studies that don't have enough data for neural networks to operate can nevertheless add data. CNNs replaced the earlier ANNs. ANNs were created for domains with recurring patterns, such as identifying unhealthy plant images. Numerous algorithms have been effectively employed to categorize plant diseases using CNNs, making crop health monitoring simpler.

According to previous studies, CNN's percentage of accurate predictions was 1 to 4% greater than that of SVM (Chen et al. 2014; Grinblat et al. 2016; Lee et al. 2015), and 6% greater than that of random forests (Kussul et al. 2017). Nevertheless, Song and colleagues (2016) revealed that the CNN model's correct prediction rates are 18% lower than ML models. Deep learning models for crop categorization aid in pest control, agricultural activity, yield prediction, and other tasks (Zhu et al. 2018). Deep learning models have made farmers' work easier, allowing them to take a photo in the field, click it, and then send it to a program to determine the disease. CNN models do not require feature engineering, a time-consuming procedure, because the key features are found during dataset training. They consist of layers that automatically learn to identify features from images. Key equations and pseudocode of these are as follows:

- Convolution operation:

$$(I * K)(i, j) = \sum_m \sum_n I(i + m, j + n) K(m, n) \quad (1)$$

Where  $I$  is the input image,  $K$  is the kernel, and  $(i, j)$  are the coordinates of the output.

- ReLU activation:

$$\text{ReLU}(x) = \max(0, x) \quad (2)$$

- Pooling operation (Max pooling):

$$P(i, j) = \max_{m, n} I(i + m, j + n) \quad (3)$$

Where  $P$  is the pooled feature map.

- Feature extraction:

$$F = \text{PretrainedModel}(I) \quad (4)$$

Where  $F$  is the feature map extracted from the input image  $I$ .

- Fully connected layer:

$$y = W.F + b \quad (5)$$

Where  $W$  and  $b$  are the weights and biases of the fully connected layer, and  $y$  is the output.

Through the convolution operation Eq. (1), an image  $i$  is treated with a small matrix, also known as kernel  $K$  to produce a feature map. The coordinates in the output feature map are referred to by indices  $i$  and  $j$ . While applying a kernel over an image, within the kernel window formed, the pixels are multiplied by the value in the corresponding position of the kernel and the sum of these weighted pixels is computed. This enables the model to learn features, like edges and corners and textures that are quite small, but very important in learning patterns contained in an image. The convolutional procedure serves to pinpoint these low-level patterns. These patterns are combined in a progressive way to learned higher-level patterns in the layers of the network that are deeper. In this part of the model, the projection after activation of the input layer applies the relu activation function (Eq. 2). It will simply remove the effect of neurons with negative activations in the feature map rather than their positive counterparts. This will help in complex learning as most networks are uninterested in learning non-linear relations from the data the task primarily deals with linear hypotheses. This is especially important for ReLU as it helps to keep the architectural design of the model simple while increasing its feature extraction capabilities from the input data significantly.

The next step is a pooling operation which is most times used as the max pooling which can be described by the Eq. (3). ‘Max pooling’ aids in scaling down the feature map, by utilizing small parts of the input feature map and finding only the extreme maximum value in

them. Means of down-sampling like this keep the essential characteristics and throw away the unimportant ones thus cutting down the processing costs and minimizing the chances of overfitting. In this way, max pooling enhances how well the network performs across data that is different from what it was trained on by retaining only the most important features. For feature extraction, it can be modified by Eq. (4) with accepted pretrained model. Most commonly, some models like ResNet or VGG are used as Pretrained models for transfer learning in which the filters trained on data are used for new tasks. They have learned how to build features within deep networks that have been exposed to copious amounts of data and have useful features for processing in the next stages of the network. The last part of this model; the fully connected layer illustrated in Eq. (5). It combines knowledge acquired by the model and serves for the final stage - prediction, i.e. discrimination of plant diseases in the image.

- Pseudo code for Simple CNN model for plant disease detection:

Due to the specific characteristics of each disease location, Barbedo (2018a) examined individual lesions and patches rather than the entire leaf. This approach can detect multiple diseases on one leaf and expand data by dividing the leaf image into many sub-images. In 2015, Lee et al. (2015) projected a new perspective on leaf disease detection that focused on identifying diseased area methods. Experiments showed that training the deep learning model with the general disease was more generic, regardless of crop type or lack of observation.

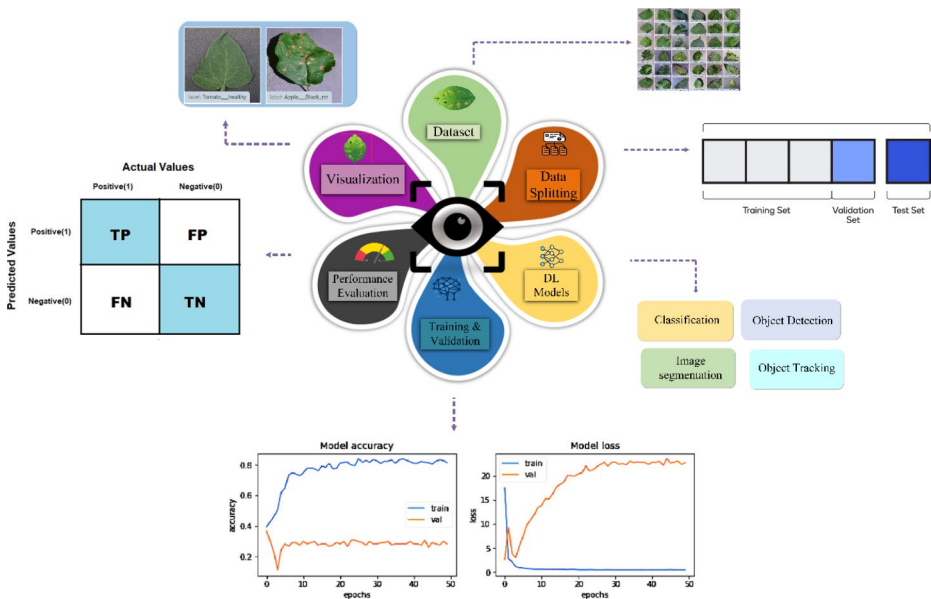
Customized deep learning models have proven to be highly effective in plant disease detection due to their ability to be tailored to specific datasets and issues. These models can be fine-tuned to deliver optimal performance for different types of plant diseases, varying image qualities, and diverse environmental conditions, providing a more accurate and efficient solution compared to generic models. A great example of this is the work by Mahmud et al. (2024) in their paper “Light-Weight Deep Learning Model for Accelerating the Classification of Mango-Leaf Disease,” where they developed a streamlined version of the DenseNet architecture specifically for mango-leaf disease classification. The custom DenseNet model was designed to be lightweight, reducing computational complexity while maintaining high accuracy. This makes it ideal for use in resource-limited environments, such as mobile devices or edge computing systems often used in agriculture. The model was fine-tuned using a dataset of mango-leaf images with various diseases, allowing it to learn disease-specific features more effectively and thus improve its classification accuracy. Moreover, the lightweight nature of this custom DenseNet model ensures faster processing times, which is crucial for real-time field applications. This combination of efficiency and accuracy makes it a valuable tool for farmers and agricultural experts who need quick and reliable disease diagnostics. The custom model achieved outstanding results in classification tasks, showcasing that tailored architectures can significantly enhance performance for specific agricultural challenges. The study highlighted improved accuracy and reduced computational overhead compared to standard deep learning models. The success of the custom DenseNet model by Mahmud et al. highlights the advantages of creating specialized deep learning architectures for plant disease detection. Custom models can be optimized for the unique characteristics of different plants and disease types, leading to better diagnostic tools that are both efficient and effective. Using customized deep learning models in plant

disease detection offers several benefits, including higher accuracy, faster processing times, and adaptability to various environments. As demonstrated by the custom DenseNet model for mango-leaf disease classification, these tailored approaches can significantly advance precision agriculture, contributing to healthier crops and more sustainable farming practices. Table 3 summarizes current studies that use the DL framework for disease detection and categorization directly. The flowchart of DL model implementation for disease detection is shown in Fig. 4.

Deep learning computer vision models, such as CNNs, offer several significant benefits. One of the primary advantages is their ability to automatically extract relevant features from raw image data, eliminating the need for extensive manual feature engineering (Chakraborty et al. 2022; Chandel et al. 2022). This capability allows these models to identify complex patterns and structures within images, making them highly effective for tasks like image classification, object detection, and semantic segmentation. These models have demonstrated impressive accuracy and performance in various real-world applications, par-

**Table 3** Disease identification/detection/recognition/classification using deep learning algorithms

S. No.	Plants	Targets	DL Algorithms/ Architectures	Accuracy	Reference
1.	Pines and Salinas	Identification	Modified CNN	95.40%	Hu et al. (2015)
2.	Rice	Identification	CNN, (CNN+segmentation), (CNN+SVM)	82.26%, 92.75% and 96.8%	Chen et al. (2016)
3.	To-mato	Identification	GoogLeNet	94%	Brahimi et al. (2018)
4.	Apple	Identification	ResNet152, Inception V3, and MobileNet	77.65%, 75.59%, and 73.50%	Bi et al. (2019)
5.	Wheat	Detection	Mask-RCNN (ResNet50 or ResNet101)	92.01%	Qiu et al. (2019)
6.	Corn	Classification	VGG16 (ImageNet)	95.33%	Li et al. (2019a, b)
7.	To-mato	Identification	VGG19, VGG16, ResNet, and Inception V3	-	Ahmad et al. (2020)
8.	Corn, potato, and tomato	Identification	RIC-Net	98.44%, 99.43% and 95.20%.	Zhao et al. (2022)
9.	To-mato	Classification	Modified CNN	95%	Shewale and Daruwala (2023)
10.	To-mato	Classification	EfficientNetB5	99.07%	Bhandari et al. (2023)



**Fig. 4** Plant disease detection flow diagram with DL implementation (Agarwal et al. 2020)

ticularly when ample labeled data is available for training. Another key strength of deep learning computer vision models is their scalability. Their performance improves with the availability of more labeled data and better hardware resources, such as powerful GPUs, making them suitable for large-scale image analysis projects. Moreover, these models are versatile and have been successfully applied across different domains, including medical imaging, autonomous driving, and agricultural monitoring, showcasing their wide-ranging applicability (Upadhyay et al. 2024).

However, there are notable limitations to these models. One significant challenge is their dependency on large amounts of labeled data for training. Without sufficient annotated datasets, the performance of these models can be limited. In comparison, advanced models often incorporate unsupervised or semi-supervised learning techniques to mitigate this dependency. Additionally, deep-learning computer vision models require substantial computational power for training and deployment, which can be a barrier for organizations with limited resources. Interpretability is another major concern. These models often act as “black boxes,” making it difficult to understand the reasoning behind their decisions and the features they have learned. While advanced models sometimes incorporate techniques to enhance interpretability, traditional deep learning models lag in this aspect. Overfitting is also a common issue, especially when these models are trained on limited datasets. Although regularization methods and data augmentation can help address this, advanced models often use more sophisticated techniques to prevent overfitting effectively. Lastly, deep learning computer vision models can be vulnerable to adversarial attacks, where minor perturbations in input data can lead to incorrect predictions. Advanced models may offer better defenses against such attacks, highlighting an area where traditional deep learning models need improvement.

In summary, while deep learning computer vision models provide automated feature extraction, high accuracy, scalability, and versatility, they also face challenges related to data dependency, computational complexity, interpretability, overfitting, and robustness to adversarial attacks. Advanced models often address these limitations to some extent, serving as a benchmark and pointing out areas for enhancement in traditional deep-learning approaches.

## 8 Advanced computer vision-based deep learning models

These methods start by using region proposal techniques to create several sparse candidate boxes, after which they use a detector employed with CNN to conduct regression with a bounding box for categorization. The second family of algorithms, known as single-stage detectors, are the single shot multi-box detector (SSD) (Liu et al. 2016) and you only look once (YOLO) series algorithms (Redmon et al. 2016; Redmon and Farhadi 2017, 2018). They simultaneously estimate bounding boxes and target class probabilities from entire pictures. These algorithms, which are based on CNN, have excelled in major competitions where objects were recognized in real-world images, including PASCAL VOC (Pattern Analysis), Statistical Modeling ImageNet (Deng et al. 2009), COCO (Common Objects in Context) (Lin et al. 2014) and Computational Learning Visual Object Classes (Everingham et al. 2010; Li et al. 2020a, b).

Two modules comprise Faster R-CNN: a Fast R-CNN detector and a region proposal network (RPN) (Everingham et al. 2010; Li et al. 2020a, b; Girshick et al. 2014; Girshick 2015; Ren et al. 2017). The RPN is a proposal generation network that is entirely convolutional. Each feature map site can generate nine anchors with three scales and different aspect ratios. Depending on the availability of targets, these anchors are then classified as positive or negative. In order to avoid bias, the positive and negative markers are chosen randomly in a 1:1 ratio as a minibatch. These anchors create candidate areas by comparing them to ground truth boxes of the objects during training. As a result, a batch of characteristic information that describes whether or not area proposals contain the candidate items may be created once the convolutional feature maps of any size are introduced into the RPN. The architecture of the Faster R-CNN detector and the RPN share is shown in Fig. 5. Better object detections can be obtained using a base network and then deployed into those convolutional layers to extract features (Li et al. 2020a, b).

Over YOLO v1 and v2, YOLO v3 combines both an improvement and an inheritance. Within the YOLO v1 algorithm, all input images are rescaled to a specified size and split into an  $S \times S$  grid. When a definite set  $B$  of bounding boxes and its related assurance score are explicitly predicted, a grid cell can only be connected with one item. A series of probabilities for each class item are generated simultaneously using the fully linked layer. The same dataset target, however, may have more than one box surrounding it. The detection boxes with the greatest confidence score are chosen using the non-maximum suppression (NMS) with an intersection over union (IoU) criterion to avoid redundant predictions. IoU calculates the amount of overlap between predicted and actual bounding boxes. NMS keeps the recognition box with the greatest confidence and abandons the rest. This result effectively reconstructs an object identification issue into an end-to-end regression job (Li et al. 2020a, b).



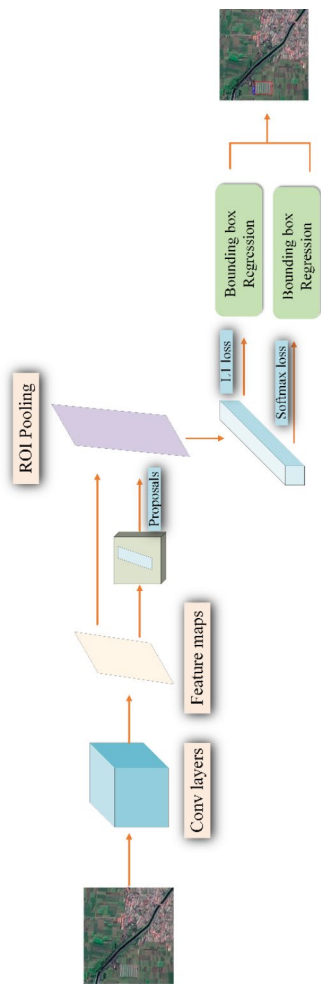


Fig. 5 Faster R-CNN architecture for Agricultural Greenhouses detection (Li et al. 2020a, b)

Recent versions of YOLO, including YOLOv4, YOLOv5, YOLOv7, and the latest YOLOv9, have brought substantial improvements in terms of accuracy, speed, and efficiency. YOLOv4 introduced advanced features like cross stage partial (CSP) connections, spatial pyramid pooling (SPP), and the Mish activation function, which together enhance detection performance and training stability, as shown in Fig. 6. YOLOv5, although not an official continuation by the original creators, has become popular for its ease of use and implementation improvements. YOLOv7 took this further by optimizing both speed and accuracy, making it ideal for embedded and mobile applications. YOLOv9 continues this trend with even more advanced optimizations and innovations, further boosting performance across various metrics. It features improved backbone networks, better handling of small objects, and enhanced real-time detection capabilities, making it one of the most efficient and accurate models available for object detection.

Faster R-anchor CNN's box technique and YOLO v1's regression concept are combined in the SSD method. In object identification models, the base network, which consists of the initial few layers, is a popular design, as shown in Fig. 7. Similar to Faster R-CNN, the same VGG-16 network was utilized. SSD employs a pyramid structure with feature maps of various dimensions following the base network. As the spatial resolution of feature maps decreases, 3-D detail information was continually lost, while abstract semantic characteristics continue to expand. Due to this, features of varying depths may identify both tiny and big objects simultaneously, which is crucial for resolving the issue of changing object sizes (Li et al. 2020a, b).

The advancements in models like YOLO demonstrate significant progress in addressing some of these challenges, particularly in object detection, by enhancing both speed and accuracy. Despite these advancements, there are still some challenges. One major issue is the need for large amounts of labeled data to train these models effectively. Without enough annotated datasets, their performance can suffer. In contrast, newer models often use unsupervised or semi-supervised learning techniques to address this dependency. Additionally, deep learning models for computer vision require significant computational power for training and deployment, which can be a hurdle for organizations with limited resources.

## 8.1 Vision-transformers (ViTs)

The Transformer architecture has established itself as the industry norm for jobs involving natural language processing, but its applicability to computer vision is yet very few. The vision, and attention are utilized to replace some of the convolutional networks' components while maintaining the overall structure of the network. Several researchers have demonstrated that there was no requirement for CNN dependency and that pure transformers can be used directly for collections of patches of pictures, which may get excellent results on image classification operation. Traditional CNNs process images by applying a series of convolutional layers to capture local features within the image. In contrast, ViTs divide an image into smaller patches and treat each patch like a word in a sentence. This method allows the model to examine the relationships between these patches, enabling a broader and more holistic analysis of the image rather than relying only on local pixel information. The transformer architecture stands out for its flexibility, making it suitable for a wide range of vision tasks beyond just classification, including object detection and segmentation. Unlike CNNs, which are often fine-tuned for specific tasks, ViTs can be more easily adapted

to different applications. One of the key strengths of ViTs is their ability to capture long-range dependencies, allowing them to grasp the overall context of an image, not just the local features. This holistic understanding is essential for tasks that require a comprehensive interpretation of the visual scene (Maurício et al. 2022). The architecture of ViT is mainly based on the original transformer, as shown in Fig. 8. ViT achieves good results when it is pre-trained on large quantities of data and transferred to various mid-sized or small image recognition platforms (CIFAR-100, ImageNet, VTAB, etc.) while substantially requiring very few CPU resources to train (Dosovitskiy et al. 2020).

After being proposed by Vaswani et al. (2017) for machine translation, transformers are now the most advanced method for many NLP tasks. If self-attention were unconditionally applied to images, each pixel would have to accompany every other pixel. This approach isn't scalable to actual input sizes because of the variations in the pixel density. Therefore, transformers have been used to process images in various approximate ways. Parmar et al. (2018) employed self-attention locally rather than globally for each query pixel. By using such local multi-head dot product self-attention blocks, convolutions may be completely replaced (Zhao et al. 2020; Ramachandran et al. 2019; Hu et al. 2019). The operations in the layers of ViTs were deduced from the following equations and pseudocode:

- Patch embedding:

$$x_p = Flatten(x_i) \quad (6)$$

Where,  $x_i$  is the  $i$ -th image patch.

- Linear projection:

$$z_i = x_p W_e + b_e \quad (7)$$

Where  $W_e$  and  $b_e$  are the projection weights and biases.

- Self-attention.

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (8)$$

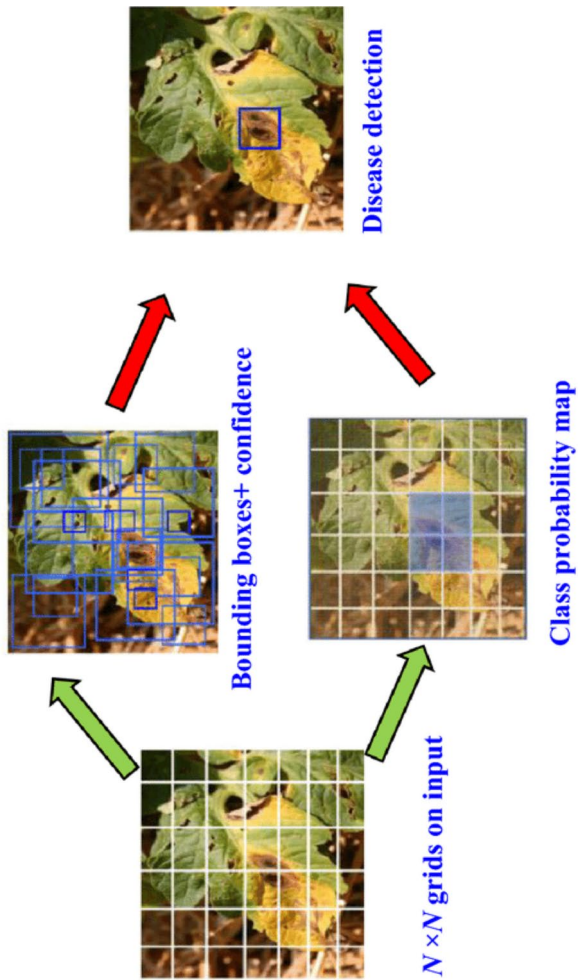
Where Q, K, and V are the query, key, and value matrices.

- Transformer encoder layer:

$$z'_l = MSA(LN(z_{l-1})) + z_{l-1} \quad (9)$$

$$z_l = MSA(LN(z'_l)) + z'_l \quad (10)$$

Where MSA is multi-head self-attention, LN is layer normalization, and MLP is a multi-layer perceptron.



**Fig. 6** Flowchart of YOLOv4 object identification method for disease (Roy et al. 2022)

- Pseudocode of a Vision Transformer (ViT) for plant disease detection:

## 8.2 Generative adversarial network (GAN)

Ian Goodfellow and colleagues (2014) introduced the idea of generative adversarial network (GAN), which have discriminator and generator networks, as shown in Fig. 9. The discriminator checks the generated material while the generator creates the content. The discriminator determines if a picture appears natural, whereas the generator makes images that appear natural. GAN is regarded as a two-player minimax method. Convolutional and feed-forward neural nets are used by GANs (Goodfellow et al. 2014).

Compared with explicit modelling, the implicit modelling approach of GAN may provide superb pictures and avoid complexities. Additionally, given that GAN may accommodate high-dimensional data dissemination and top-notch picture-generating performance, The generation model's current top approach is GAN. High-quality, fine-grained RGB plant pictures were produced by Zhu and others (2020) by modifying the different class tags with the desired class tags to the network architecture of CDCGAN. The photographs dramatically enhanced the recognition's classification performance after data augmentation; the F1 score increased by 0.23, supporting the hypothesis. The outcomes were comparable to those of the bigger training set without adding any info. A novel technique was presented, utilizing GANs to enhance the data for identifying leaf diseases with the pictures produced via generative adversarial networks with deep convolutions (DCGAN) (Wu et al. 2020). Using the actual pictures as Google Net's input, this model may attain a 94.33% average accuracy in identification. However, there is still much space for development in the precision and quality of the disease-related images produced by the techniques mentioned above for classification (Zhang et al. 2022). Additionally, in several related efforts, updated or upgraded DL architectures were used to produce better outcomes and create software for disease identification systems.

## 8.3 Vision-language models

Vision-language models (VLMs) first came into light in 2019, bridging the gap between computer vision and natural language processing. By combining image analysis with text comprehension, these models overcome the limitations of older object recognition systems. That year saw major strides in transformer architectures and dual-stream frameworks, which helped VLMs enable exciting new applications like image captioning and visual question answering. This progress has sparked continued innovation in artificial intelligence, pushing the field forward (Li et al. 2019a, b).

In 2020, there was a burst of innovation in VLMs. Researchers made great strides in refining how these models are pre-trained to better understand the relationship between images and text. For example, VL-BERT worked on creating flexible visual-linguistic representations that could adapt to various tasks (Su et al. 2019). XGPT focused on improving image captioning by enhancing cross-modal generative pre-training. Pixel-BERT took a creative approach by aligning individual image pixels with the right text components (Huang et al. 2020). The Multimodal Framework (MMF) provided researchers with a robust set of tools tailored for vision and language studies (Xing et al. 2021). OSCAR emphasized the importance of connecting visual elements with their corresponding text descriptions

(Li et al. 2020a, b). UNITER aimed to standardize the way images and text are formatted, reflecting a broader trend towards more adaptable VLM solutions (Chen et al. 2019a, b). These advancements collectively expanded what VLMs could do, paving the way for more advanced applications.

Recent advancements in vision-language models, such as CLIP (Contrastive Language-Image Pre-training) and ALIGN (A Larger-scale Image-Language Co-training), have significantly enhanced our ability to connect visual and textual information (Zhang et al. 2021). These models are designed to interpret and generate images from textual descriptions, enabling functions like zero-shot classification, image captioning, and visual question answering (Radford et al. 2021). By learning from extensive datasets that pair images with text, these models exhibit impressive versatility and adaptability, often excelling in a wide range of tasks without the need for specialized fine-tuning.

## 8.4 Foundation models

Large foundation models, like those from OpenAI and DeepMind, have become a major trend in the field. Trained on vast datasets and capable of handling various tasks, these models exemplify the move toward creating general-purpose AI systems. They utilize transfer learning to apply insights from one area to another, resulting in highly adaptable models that can tackle multiple tasks with minimal extra training. Their influence on computer vision has been significant, setting new performance standards and expanding the possibilities of AI (Li et al. 2023a, b, c).

A model like DALL-E could generate visual simulations of disease progression on crops, helping researchers understand how different diseases spread and evolve, leading to better preventive measures. GPT-3 is capable of creating detailed descriptions for images. For example, if you show it a picture of a busy city street, it can craft a vivid narrative that captures the essence of the scene, including the people, buildings, and overall atmosphere.

## 8.5 Advanced and hybrid computer-vision deep learning architectures

Major downsides of many existing deep neural network architectures include a huge number of parameters, a lengthy training period, expensive storage and processing costs, etc. New or modified DL structures are utilized to detect leaf disease, and a flow chart of their implementation is given in Fig. 4. Table 5 below summarizes current studies on enhancing DL in plant disease diagnosis.

## 9 Plant disease public datasets

An image dataset consists of digitized pictures that have been carefully selected for use in training, testing, and assessing the performance of computer vision algorithms. The data sets used to analyze leaves are built from primary data gathered in the fields. Because the statistics are based on visible characteristics of the leaves, they have a high degree of trustworthiness. Additionally, the data sets are separated into easily understandable portions.

For instance, the research by Atila et al. (2021) divides the study into sections according to diseases such as sheath blight (SB), rice blast (RB), and bacterial leaf blight (BLB).

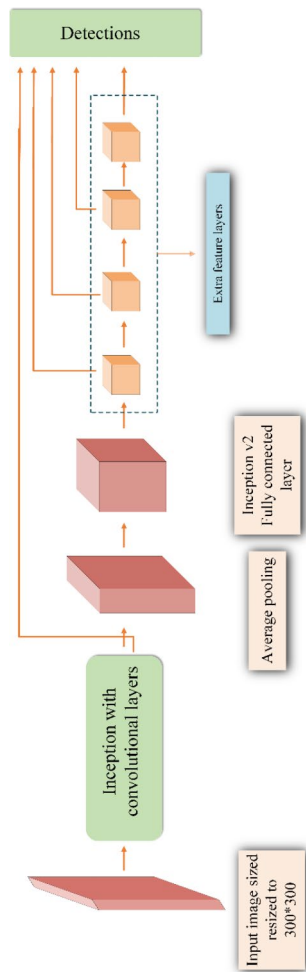


Fig. 7 Architecture of single shot detector (SSD) (Sivakumar et al. 2020)

In this study, another data set named as PlantVillage was used, having 54,306 photos of 14 distinct crops representing 26 plant diseases that make up the dataset. Different-coloured leaves were depicted in the photographs that were part of the data collection. Examples from the Plant-Village data set are shown in Fig. 10. The hues represent the areas of the leaves afflicted by the diseases being researched (Geetharamani and Pandian 2019).

Additionally, ImageNet data collection was employed in the research, and the interaction of merging different approaches produced high-quality research results (Atila et al. 2021). The requirement to demonstrate image-based detection algorithms for coffee leaf diseases promotes the utilization of coffee leaf data sets in research (Esgario et al. 2020). The researchers employed large numerical data sets and information-rich colour data sets to present their gathered data (Proctor and He 2015).

In PlantVillage, the descriptions of the leaves both before and during the disease's impact are included in the data sets. This data set displays healthy leaves with those that suffered dents from attacks by septoria leaf blight, frog eye leaf spot, and downy mildew. The data set is understandable and well-organized. The data set is unambiguous and shows how many leaves in total were examined and divided into four groups. The PlantVillage dataset, which Sharada P. Mohanty and others in 2016 compiled, comprises 87,000 RGB photos of healthy and unwell plant leaves divided into 38 groups. They have only chosen 25 classes to test their algorithm, and Table 6 displays these classes.

In the subsequent paragraphs, some findings from this investigation were summarized: finding leaf photos for particular plant diseases is challenging. The size of the accessible plant data sets is consequently quite modest. Only a few papers have submitted thousands of pictures for investigation (Barbedo et al. 2016; Sladojevic et al. 2016; Meunkaewjinda et al. 2008; Pires et al. 2016; Shrivastava et al. 2019; Schikora et al. 2012). The photographs in the database are taken under extremely restricted environmental settings. In order to make the algorithms more useful, photos must be collected under real-world circumstances. The current situation calls for effective picture acquisition of leaf images. The research community would appreciate such databases if these photos were recorded in real-time circumstances. Images captured using sophisticated mobile devices are becoming more common in most recently published works (Chandel et al. 2024). Although several single-click picture solutions are also presented, the researchers hope to significantly increase plant disease detection algorithms' automation. A severe problem with database size may be resolved with the move of image-capture systems to smart devices.

A smartphone-assisted disease detection system was developed by Mohanty et al. (2016). The prime motivation for the study was a combination of rising global smartphone usage and recent breakthroughs in computer vision enabled by deep learning. A deep convolutional neural network was trained to identify 14 crop species and 26 diseases using images of a public dataset of 54,306 photos of healthy and damaged plant leaves taken under controlled settings. These images were taken from the 'Plant Village' dataset. The dataset of images was resized to  $256 \times 256$  pixels. Different combinations of training and testing splits were tried. Two approaches were followed while developing deep learning models. The first approach was pre-trained transfer learning models used for classification model development. In the second approach, a CNN-based model for classification was developed from scratch. The effect of image type on classification accuracy was studied as input to network colour images, grayscale images, and background-removed images were used. The accuracy of the two models was expressed in terms of the F1 score. The average F1 score of



**Table 4** Comparison of multi box SSD, you look only once-v3, and faster R-CNN (Li et al. 2020a, b)

Characteristics	YOLO v3	Faster-RCNN	SSD
Network Type (Neural)	Convolutional (Fully)	Convolutional (Fully)	Convolutional (Fully)
Phases	Concurrent bounding-box classification and regression	RPN+Fast R-CNN detector	Concurrent bounding-box classification and regression
Location Detection	Based on Anchor	Based on Anchor	Default boxes/Prior boxes
Backbone Feature Extractor	Darknet-53 (53 convolutional layers)	VGG-16 or another feature extractors	VGG-16 or another feature extractors
Loss Function	Binary cross-entropy loss	Smooth L1 for regression; Softmax loss for classification	Smooth L1 Loss for localization and Softmax loss function for confidence
IoU Thresholds	One (at 0.5)	Two (at 0.3 and 0.7)	One (at 0.5)
Anchor Box	Nine anchor boxes with different sizes, K-means from VOC and coco	Nine standard boxes with varying aspect ratios and scales	Each feature map has a collection of bounding boxes in different sizes and aspect ratios.
Input Size	Random input multi-scale size	Keep the original image's aspect ratio; the enlarged dimensions vary from 500 to 1000.	Images resized to a defined size ( $300 \times 300$ or $512 \times 512$ ).

GoogleLeNet and AlexNet was 98.86% and 98.48%, respectively. The classification accuracy of the model reached up to 99.35%.

CNN-based deep learning network was developed to classify tubers into five classes, i.e., healthy and four disease classes (Oppenheim and Shani 2017). The dataset contains different shapes, sizes, and tones of potatoes acquired under normal conditions. The dataset was labeled manually by a subject matter specialist. The total dataset was made into different train and test split combinations to see the effect on model accuracy. The highest classification accuracy was 96% at a training dataset size of 90% and a testing dataset of 10%.

GoogleLeNet and Inception v3 models were deployed on TensorFlow to detect two types of pests and three diseases in cassava crops (Ramcharan et al. 2017). The dataset of 11,670 images was used during the training, validation, and testing stages. The confusion matrix was used as a performance metric. The classification accuracy ranged from 80 to 93%. Waheed et al. (2020) employed an optimized DenseNet model for three corn leaf disease detection and classification. The total trainable parameters were 77,612. The optimized DensNet classification accuracy was compared with the VGGNet, XceptionNet, EfficientNet, and NASNet. The DL models were competent on the 12,332 images dataset. The image

---

```

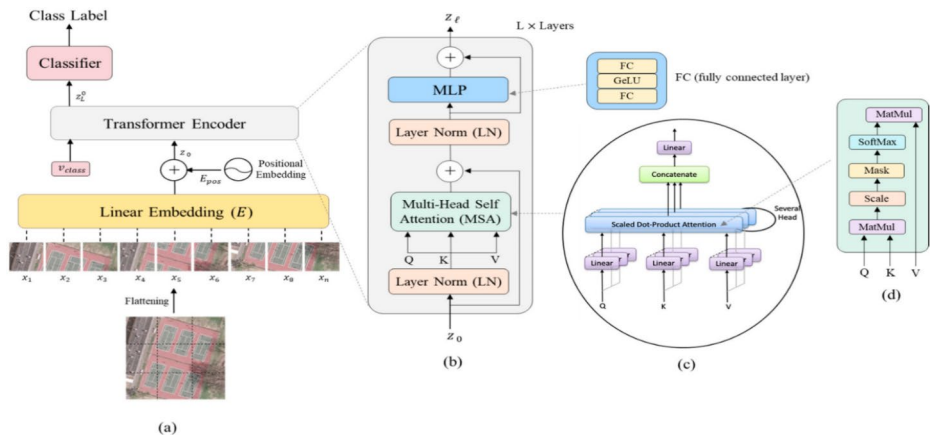
# Initialize parameters
initialize_weights()

# Forward pass
for each epoch:
    for each image in training_data:
        # Convolutional layer
        conv_output = convolution(image, kernel)
        # ReLU activation
        relu_output = relu(conv_output)
        # Pooling layer
        pooled_output = max_pooling(relu_output)
        # Fully connected layer
        fc_output = fully_connected(pooled_output)
        # Softmax layer
        predictions = softmax(fc_output)
        # Compute loss
        loss = compute_loss(predictions, true_labels)
        # Backward pass and update weights
        gradients = backward_propagation(loss)
        update_weights(gradients)

# Function definitions (convolution, relu, max_pooling, fully_connected, softmax,
compute_loss, backward_propagation, update_weights) would be implemented here

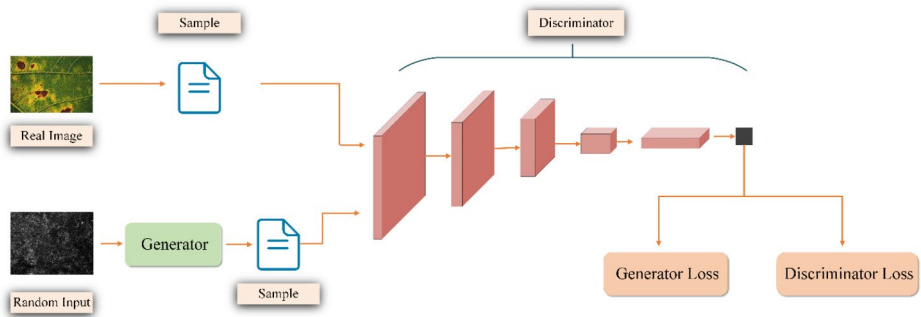
```

---

**Algorithm 2****Fig. 8** Vision Transformer architecture (Bazi et al. 2021)

resolution of each image was  $250 \times 250 \times 3$ . The data augmentation technique (cropping, padding, and horizontal flipping) increased relevant data. The accuracy of the DenseNet was found to be 98.06%. In the future, they intend to develop a mobile application for corn leaf disease detection and classification.

AlexNet, VGG16, and VGG19 CNN models were used for disease detection in rice fields (Sethy et al. 2020). Three models were trained to detect and classify four types of diseases



**Fig. 9** Block Diagram of Generative Adversarial Networks (GAN) (Aggarwal et al. 2021)

in paddy, i.e., Blast, Brown spot, Tungro, and Bacterial blight. The support vector machine classifier was used for four types of rice disease detection. In this study, CNN models have used a feature extractor from disease images 5932. The extracted features were used in the second stage to input the support vector machine classifier. AlexNet (feature layer: fc8) called resnet50 with support vector machine algorithm classified better than other algorithms. The F1 score was found to be 98.38%. CNN-based deep learning model was trained with 10,000 labeled images for cassava crop disease detection (Sambasivam and Opiyo 2021). The images were processed using the adaptive histogram equalization technique. The accuracy of the model varied from 76.9 to 99.30%. CCN and ANN models were used for plant disease detection (Shin et al. 2020). Feature extraction with CCN and ANN classification generated better results than other supervised machine learning models under different positions or leaf angles under field conditions.

It has been argued that deep learning-based image analysis techniques surpass traditional methods for visually assessing disease severity. However, these imaging systems are not without flaws. The quality of the training data significantly impacts the system's performance. In plant disease automation, the training images and a few extracted features greatly influence a system's performance. A well-trained system uses high-quality training data. However, a definite set of conditions must be met for most prevailing systems to function correctly. The method may produce false findings if requirements are not fulfilled which results in incorrect disease detection. Some generalized techniques that function in diverse situations need to be modified. In-depth knowledge of the techniques and appropriate tool use are also required to increase productivity. Table 7 compares and summarizes the most recent findings for diagnosing plant diseases using various data sets and techniques.

## 10 Performance metrics for classification and object detection models

Many measures have been introduced in research, each addressing specific aspects of an algorithm's performance. Consequently, for every machine learning problem, researchers need a suitable set of measures for performance assessment.

In this study, several common metrics were collected to obtain crucial information on the effectiveness of algorithms in categorization tasks and performed a side-by-side comparison. These metrics include precision, recall (Powers 2011), F1-score (Sasaki 2007),

**Table 5** Imaging-based disease identification/detection/recognition/classification using advanced and hybrid computer-vision deep learning architectures

S. No.	Plants	Targets	Modified DL Algorithms/Architectures	References
1.	Tomato	Leaf diseases	Transfer Learning with CNN and IncMB Module for early lesion detection	Radočaj et al. (2024)
2.	Various	Disease identification	Hybrid CNN-RNN architectures for sequential data analysis	Ait Nasser and Akhloufi (2022)
3.	Coffee	Plant diseases	Hybrid CNN models with attention mechanisms	Karthik et al. (2023)
4.	Corn	Detection	Multiple CNN	DeChant et al. (2017)
5.	Apple	Identification	AlexNet-precursor Cascade- Inception network	Liu et al. (2018)
6.	Apple	Identification	Multi-scale ResNet	Wang et al. (2017)
7.	Tomato	Detection	feature ranking-KNN	Xie et al. (2017)
8.	Apple	Detection	Orthogonal Subspace Projection	Shuaibu et al. (2018)
9.	Peanuts	Detection	Hyperspectral Vegetation Index	Chen et al. (2019a, b)
10.	Wheat	Detection	Convolutional and Bidirectional Gated Recurrent Network	Jin et al. (2018)
11.	Tomato	Detection	Generative Adversarial Nets	Wang et al. (2019)
12.	Potato	Detection	Generative Adversarial Nets	Polder et al. (2019)
13.	Wheat	Detection	Multiple Inception-Resnet	Zhang et al. (2019)
14.	Tomato	Identification	Deconvolution- Guided VGGNet	Ren et al. (2017)
15.	Tomato	Identification	Multi-layer CNN	Agarwal et al. (2020)
16.	Cassava	Classification	Vision Transformers	Thai et al. (2021)
17.	Tomato	Classification	MMD GAN, B-ARNet model	Zhang et al. (2022)
18.	Rice	Identification	PlantViT (CNN + Vision Transformers)	Thakur et al. (2022)
19.	Olive	Identification and classification	CNN + Vision Transformers	Alshammari et al. (2022)
20.	Grape	Recognition	GrapeGAN	Jin et al. (2022)
21.	Tomato	Identification	FWDGAN	Li et al. (2022)
22.	Cucumber	Recognition	SwinT-based and Attention-guided GAN	Wang et al. (2022)
23.	Pepper bell, potato, and tomato leaf	Classification	Transfer Learning Model and Vision Transformer (TLMVIT)	Tabbakh and Barpanda (2023)
24.	Bean	Identification	Inception convolutional vision transformers	Yu et al. (2023)
25.	Sugarcane leaf	Identification	Shuffle-convolution-based lightweight Vision transformer (SLViT)	Li et al. (2023a, b, c)

accuracy, ROC-AUC score, IOU (Breton and Eng 2019), mAP, and the confusion matrix (Fawcett 2006; Brown and Davis 2006).

### 10.1 Confusion matrix

This matrix is the most useful and clearest criterion for defining a machine learning algorithm's accuracy and correctness. It is mostly used for classification problems where the output might have two or more categories of various classes. It consists of true and false negatives (TN & FN) and true and false positives (TP & FP), as shown in Fig. 11. Precision and recall performance metrics were derived from the confusion matrix.



**Fig. 10** Plant-Village data set: illustrations of diverse plant phenotypes (Hughes and Salathe 2015).

## 10.2 Precision

It just indicates what number of selected data items are pertinent. In other words, the quantity of positive observations that a machine learning system predicted. *Formula* (11) specifies that accuracy is calculated by dividing the total number of true positives by the sum of erroneous and true positive results (Powers 2011):

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

## 10.3 Recall

It displays the fraction of relevant data items chosen. In reality, it remarks the number of truly encouraging data that have the algorithm's predictions. *Formula* (12) states that Recall is calculated by dividing the count of true positives by the sum of erroneous negatives and true positives (Powers 2011).

**Table 6** Dataset specifications (Hughes and Salathe 2015)

Disease Name	Plant	No. of Images
Healthy	Corn	1859
Cercospora leaf spot		1642
Common rust		1907
Northern Leaf Blight		1908
Healthy	Apple	2008
Scab		2016
Black rot		1987
Cedar apple rust		1760
Healthy	Tomato	1926
Bacterial spot		1702
Early blight		1920
Late blight		1851
Leaf Mold		1882
Septoria leaf spot		1745
Two-spotted spider mite		1741
Target Spot		1827
Yellow Leaf Curl Virus		1961
Tomato mosaic virus		1790
Healthy	Potato	1824
Early blight		1939
Late blight		1939
Healthy	Grapes	1692
Black rot		1888
Esca (Black Measles)		1920
Leaf blight (Isariopsis)		1722

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

#### 10.4 F1-score

It measures algorithm performance, often known as f-measure or f-score, which also considers recall and precision. It is the harmonic mean of precision and recall, written as follows in mathematics (Sasaki 2007):

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (13)$$

#### 10.5 Accuracy

Probably the most common and first way to evaluate an algorithm's categorization performance. It measures the percentage of successfully predicted data points to all observations (Formula (14)). In spite of being extensively applicable in various fields, accuracy may not be the best performance measure when the dataset's intended parameter classes are uneven (Breton and Eng 2019).

$$Accuracy = \frac{TN + TP}{FP + TP + FN + TN} \quad (14)$$

**Table 7** Disease detection and identification using image dataset

S. No.	Database used	Dataset size	Methods and Algorithms	Results/Findings	Reference
1.	Plant Village	54,305	CNN (DenseNet-121, VGG-16)	99.81% accuracy in leaf disease detection	Dey et al. (2024)
2.	New plant disease dataset	87,000 images	YOLOv4	98% accuracy across 38 disease classes	Aldakheel et al. (2024)
3.	LeafSnap	30,000 images	Deep Learning (ResNet50)	95% accuracy in plant species and disease detection	Demilie (2024)
4.	Custom dataset	7,623 training, 1,906 validation images	DenseNet169, Xception, InceptionV3, MobileNetV2, ResNet50V2	99.42% validation accuracy with MobileNetV2 and ResNetV2	Kau and Bansal (2024)
5.	Images collected from the field as well as Internet	7199	An integrated model was suggested for one- and two-stage target detection networks. Target frame size was clustered in candidate frame generation to enhance small target recognition in the Faster-RCNN-based two-stage network. The internal structure was optimized for the single-stage network, which uses YOLO.	This model attained 85.2% mAP, significantly higher than other comparative models, among the 37 pests and eight diseases found.	Li et al. (2023a, b, c)
6.	Images collected from the field & Internet	2626	A position attention block was developed to improve model feature extraction for the region of interest. The feature map's position information can be easily extracted, and an attention map can be generated.	Compared to other models of a similar nature, the accuracy of the ResNet model, which is based on positional attention blocks, was 96.4%.	Cheng et al. (2023)
7.	Plant Pathology 2021-FGVC8 dataset and field images	4182	An upgraded version of Faster region-based convolutional neural network technology, Faster R-CNN, was proposed. A complex Res2Net and feature pyramid network design extracted reliable, multi-dimensional features.	The upgraded Faster R-CNN structure outperforms previous object recognition algorithms in the annotated apple leaf disease dataset with 63.1% average precision.	Gong and Zhang (2023)
8.	Images collected from the field	-	A method of re-parameterization was presented to improve the boosting accuracy of convolutional neural networks. A dynamic pruning gate was used to regulate the network structure dynamically.	Subclasses with low detection accuracy were improved by data augmentation and verified by ablation tests for model correctness. Finally, the model was 0.94 accurate.	Liu et al. (2023a, b)
9.	Aerial images collected from cotton field	-	K-means clustering and dynamic time warping technique was used to compare Long term short memory (LSTM) models for predicting canopy features of cotton crop.	Results demonstrated an increase in accuracy with time used for fine tuning of LSTM models for forecasting of all canopy features.	Gadepally et al. (2023)



Table 7 (continued)

S. No.	Database used	Dataset size	Methods and Algorithms	Results/Findings	Reference
10.	Aerial images collected from cotton field	-	Ant colony optimization algorithm was used to generate an optimal path for edge computing platform for real time detection of cotton root rot using YOLOV5.	The results showed that the precision, mAP and recall values at an IoU of 50% reached 45%, 30% and 41%, respectively.	Qian et al. (2022)
11.	AgriPK dataset	5135	A Compact Convolutional Transformer (CCT)-based methodology was presented to categorize the image dataset and compare it to state-of-the-art methods.	The proposed model was 97.2% accurate, while Mobile Net, ResNet152v2, and VGG-16 were 95, 92, and 90% accurate.	Jajja et al. (2022)
12.	PlantVillage and Embrapa	-	A hybrid model of a CNN and a Vision Transformer, PlantViT, has been proposed to identify plant diseases from images of leaves on the PlantVillage and Embrapa datasets.	The proposed model achieved 98.61% and 87.87% accuracy on the PlantVillage and Embrapa datasets.	Thakur et al. (2022)
13.	PlantVillage	-	A new data augmentation method based on deep threshold multi-feature extraction convolution GAN with Mixed Attention and Markovian Discriminator (MMDGAN) was proposed for tomato disease leaf classification	Frechet Inception Distance (FID) scores of healthy tomato leaf image, Leaf Mold, Leaf Curl, and Spider Mite generated by MMDGAN were 159.31, 164.47, 230.38, and 254.98 respectively.	Zhang et al. (2022)
14.	PlantVillage	54,305	The proposed neural network consists of two fully connected layers, four convolutional layers, and a final softmax layer to offer a feature representation for different plant species.	Remarkable performance of the proposed model with an accuracy of 90.0%.	Reddy et al. (2021)
15.	PlantVillage	61,486	B4 and B5 models of EfficientNet architecture achieved the highest values compared to other deep learning models in original and augmented datasets.	B5 and B4 models achieved an accuracy of 99.91% and 99.97%, respectively, and a precision of 98.42% and 99.39%, respectively.	Atila et al. (2021)
16.	PlantVillage	17,641	A healthy class of tomato leaves and nine classes of diseased tomato leaves from the Plant Village dataset make up the suggested architecture derived from the baseline CNN design.	The accuracy of the model was 97.2%.	Sembiring et al. (2021)
17.	PlantVillage	7,733	Pre-trained deep models like VGG-s and AlexNet are used initially to extract the deep features, which are then improved using the transfer learning principle.	The model's accuracy, sensitivity, G-measure, and precision performance were 97.8%, 97.6%, 97.6%, and 97.6%, respectively.	Khan et al. (2020)



**Table 7** (continued)

S. No.	Database used	Dataset size	Methods and Algorithms	Results/Findings	Reference
18.	PlantVillage	61,486	A deep learning-based mathematical model with a region proposal network and Chan-Vese (CV) was proposed to identify and detect plant diseases that increase precision, generality, and training effectiveness.	The approach's accuracy, 83.57%, is higher than the conventional method's.	Guo et al. (2020)
19.	PlantVillage	76,000	LeNet, AlexNet, and Inception V3 were used to categorize 38 varieties of plant diseases.	Inception V3 performs better, with an accuracy of 98% compared to other models.	Adit et al. (2020)
20.	PlantVillage	61,486	The TensorFlow object detection platform was used to apply three DL meta-architectures, including the SSD, Faster RCNN, and RFCN.	The SSD model had the greatest mean average precision (mAP), 73.07%.	Saleem et al. (2020a, b)
21.	PlantVillage	54,306	Developing the deep siamese convolutional network addressed the issue of tiny picture datasets.	The identification of the Esca, Black rot, and Chlorosis diseases on the grape leaves was accurate to above 90%.	Goncharov et al. (2019)
22.	Hyperspectral image data set	111	The 3D DCNN model was used in this research with a class of hyperspectral imaging of injected and mock-inoculated stem images, and the algorithm disproportionately selects the spatial regions with obvious disease symptoms by using the idea of a saliency map to illustrate the most sensitive pixel positions.	The F1 score and the model's classification accuracy were 0.87 and 95.73%, respectively.	Nagasubramanian et al. (2019)
23.	Sugar beet data set	155	The suggested method included imaging-based expert systems and demonstrated how altering CNN parameters following the picture and areas to be identified might improve the performance of the Faster R-CNN architecture.	The approach's accuracy of classification rate was determined to be 95.48%.	Ozguven and Adem (2019)
24.	PlantVillage	9,000	Developing deep learning models using progressively bigger picture datasets shows a clear route for smartphone-assisted crop disease detection worldwide.	The accuracy of the model was 99.4%.	Ashqar and Abu-Naser (2019)
25.	Powdery mildew data set	20	A semantic segmentation method based on convolutional neural networks (CNN). The proposed method performs better than segmentation techniques currently in use, like Random Forest, K-means, and GBDT.	This model achieved intersection over a union of 72.11%, average pixel accuracy of 96.08%, and Dice accuracy of 83.45%.	Lin et al. (2019)
26.	Oryza sativa database	619	Support Vector Machine (SVM) was employed as a classifier, while a pre-trained deep CNN was used as a feature extractor. This model serves as a preventative strategy and an early warning system.	The accuracy of the model was 91.4%.	Shrivastava et al. (2019)

**Table 7** (continued)

S. No.	Database used	Dataset size	Methods and Algorithms	Results/Findings	Reference
27.	Aerial image data set	3,000	Unmanned aerial vehicle (UAV) photos of NLB disease lesions were used to train the Mask R-CNN model.	The mean intersect over union (IOU) was 0.73, with an average accuracy of 0.96 at an IOU of 0.61 over an IOU threshold range of 0.50 to 0.95.	Stewart et al. (2019)
28.	Cassava disease data set	720	A CNN object detection model was trained to recognize foliar symptoms of diseases in cassava ( <i>Manihot esculenta</i> Crantz). Thereafter, the model was deployed in a mobile app, and its performance was tested on images and videos of 720 diseased leaflets in an agricultural field.	A decreased performance for real-world images and video as measured with the F1-score. The F1-score dropped 32% for pronounced symptoms in real-world images.	Ramcharan et al. (2019)
29.	Turkey plant data set	1,965	The work considers using pre-trained deep models for feature extraction and further fine-tuning. SVM, ELM, and KNN approaches are then used to classify the collected features through deep feature extraction.	The average accuracy ratings of the AlexNet, VGG16, and VGG19 models' fc6 layers was 95.6%, higher than those of the other layers.	Turkoglu and Hanbay (2019)
30.	IPPN phenotyping data set	345	Five CNN models were developed and trained—namely Inception VGG16, ResNet v2, ResNet50, VGG19, and Xception—on the PlantVillage dataset of tomato leaf pictures.	ResNet50 emerged as the most effective prediction tool, with an accuracy of 90.4%.	Verma et al. (2019)
31.	Real-world plant disease data set	21,917	A thin Deep Neural Networks (DNN) method that can function on Internet of Things (IoT) devices with limited resources. The DNN model was further simplified and minimized the size of the input and computation cost of the model.	The proposed models can reliably identify plant disease while utilizing little computer resources with an accuracy of 98.0%.	Ale et al. (2019)
32.	Apple disease image data set	640	A densely connected neural network (DenseNet) was used to optimize the lower-resolution feature layers of the YOLO-V3 model based on picture data augmentation.	The YOLO-V3 model's detection performance was improved, and the feature usage in the neural network was significantly improved using DenseNet with an accuracy of 99.4%.	Tian et al. (2019)
33.	PlantVillage augmented	50,000	The technological barriers to autonomous plant disease classification have mostly been removed, using small picture datasets for training.	The accuracy of the model was 96.0%.	Barbedo (2018b)
34.	PlantVillage	54,000	The four classifier labels employed are Bacterial Spot, Leaf Curl Virus, Late Blight, and Healthy Leaf. Twenty repetitions are used to fit the retrieved characteristics into the neural network.	The most accurate artificial neural network design for identifying plant diseases performs best, with 98.59% accuracy.	Dhakar and Shakya (2018)

**Table 7** (continued)

S. No.	Database used	Dataset size	Methods and Algorithms	Results/Findings	Reference
35.	Saitama Research Center data set	60,000	A deep learning method for leaf localization using wide-angle photos taken on-site.	At 2.0 frames per second, the technique has a detection performance of 78.0% in the F1 measure.	Cap et al. (2018)
36.	Open plant database	87,848	Convolutional neural network models were created to identify plant diseases and diagnose them using straight-forward photos of healthy and sick leaves.	The best model architecture identified the appropriate [plant, disease] pair with a success percentage of 99.53%.	Ferentinos (2018)
37.	PlantVillage	18,160	The dataset that was formerly referred to as the IPPN dataset is now known by its authors as the PRL dataset. This dataset was chosen because it has annotations for each of the three tasks completed in the validation experiments.	The Plant/Ara2013-Canon, Plant/Tobacco, and Plant/Ara2012 leaf counting datasets show good performance with $R^2$ values of 0.90, 0.74, and 0.85 for the actual and projected leaf counts, respectively.	Ubbens and Stavness (2018)
38.	PlantVillage	13,262	Tomato crop disease classification was done using photos from the PlantVillage dataset and pre-trained deep learning architecture, specifically AlexNet and VGG16 net.	The classification accuracy for VGG16 net was 97.29%, while for AlexNet, it was 97.49%.	Rangarajan et al. (2018)
39.	Cornell University data set	18,222	The biggest publicly accessible image set annotated for specific plant disease, the three data sets collectively comprise 18,222 photos tagged with 105,705 NLB lesions.	The accuracy of the model was 85.0%.	Wiesner-Hanks et al. (2018)
40.	Soybean image data set	12,673	This study observes the viability of using CNN to identify plant diseases in leaf photos captured in their natural habitat. A model built on the LeNet architecture was developed to categorize plant diseases of soybeans.	The constructed approach achieved 99.32% classification accuracy.	Wallelign et al. (2018)
41.	PlantVillage	299	A brand-new approach to combine data from several degrees of abstraction was developed to enhance transfer learning convergence. In order to automatically identify veins and colours that lead to diseased leaves, the algorithm extracts low-level information from the raw data.	The detection of leaf scorch has accuracy and a true-positive rate of 98.60% and 1.47% respectively.	Cruz et al. (2017)
42.	PlantVillage	3700	It uses a hybrid convolution neural network (CNN) to detect banana disease and the classification to help farmers by enabling fertilizers to avoid the disease in its early stages.	The proposed technique demonstrates 99.7% accuracy when compared to related deep-learning techniques.	Amara et al. (2017)

**Table 7** (continued)

S. No.	Database used	Dataset size	Methods and Algorithms	Results/Findings	Reference
43.	Korean tomato disease data set	5,000	AlexNet and SqueezeNet, two distinct deep-learning network designs, were evaluated in turn. Training and validation for these networks were carried out on the Nvidia Jetson TX1.	The accuracy of the model was 97.2%.	Durmus et al. (2017)
44.	PlantVillage	14,828	CNN was used to train classifiers. Visualization was utilized to understand symptoms and locate unhealthy leaf sections to test the deep model.	The accuracy of the model was 99.18%.	Brahimi et al. (2017)
45.	Tomato diseased image data set	5,000	These meta-architectures are each combined with “deep feature extractors” like VGG net and Residual Network (ResNet). A technique was suggested for local and global class annotation and data augmentation.	The suggested system can identify nine distinct diseases and pests with an accuracy of 87.0%.	Fuentes et al. (2017)
46.	PlantVillage	2,265	This study presents a systematic evaluation of the performances of deep models fine-tuned by transfer learning and shallow networks created from scratch.	The deep VGG16 model with transfer learning training is the most accurate, with a total accuracy of 90.4% on the stand test set.	Wang et al. (2017)
47.	Cassava disease image data set	11,670	Transfer learning was utilized for training a deep convolutional neural network to detect two types of pest damage and three diseases.	For brown leaf spot, red mite damage, green mite damage, cassava brown streak disease, and cassava mosaic disease, the top trained models had 98%, 96%, 95%, and 96% accuracy rates (CMD).	Ramcharan et al. (2017)
48.	Images collected from the Internet	33,469	The deep CNN training was conducted using Caffe, a deep learning framework created by Berkley Vision and Learning Centre.	The constructed model produced accuracy between 91% and 98%, with an average of 96.3%.	Sladojevic et al. (2016)

## 10.6 ROU-AUC score

The ROC (receiver operating characteristic) curve depicts the connection between the rate of true positives and the false positive, which is used to compute this statistic (1- specificity) (Breton and Eng 2019). A binary classification statistic called Area Under ROC Curve, or ROC-AUC, demonstrates how effectively a model can discriminate between negative and positive target classes. The ROC-AUC score might be supportive in indicating the performance if the relevance of negative and positive classes is equal for a specific problem.

## 10.7 IoU

The Deep Learning community uses the Intersection over Union (IoU) (also known as Area Ratio Overlap) to measure the effectiveness of object detection models. Considering the

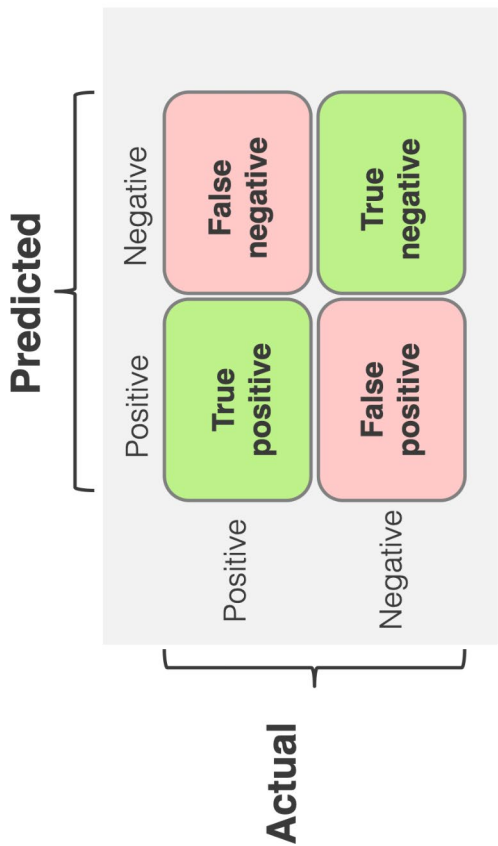


Fig. 11 Confusion matrix with actual and predicted classes

size of the bounding boxes, this measure goes one step further than the Centroid of Rectangles (CoR). It is possible to think of the IoU as a generalized instance of the CoR. The (x, y) coordinates of the predicted box of bounding from a detector won't almost certainly match the (x, y) coordinates of the corresponding ground truth bound box in practice. As a result, this assessment measure rewards overlapping bounding boxes for the ground truth and anticipated boundaries. IoU ranges from 0 (no overlap) to 1 (perfect overlap) (Breton and Eng 2019).

$$IOU (GT, PB) = \frac{TP}{TP + FP + FN} \quad (15)$$

PB is the Predicted bounding box, GT is the Ground truth bounding box, TP & FP are True Positive and False Positive, and FN is False Negative.

### 10.8 mAP

An indicator of object detection accuracy across all classes in a given database is the mean Average Precision (mAP) (Padilla et al. 2020).

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (16)$$

Where N is the total number of courses being assessed and  $AP_i$  is the average precision in the IT class.

## 11 Challenges and the way forward

The decline in agricultural production and productivity adversely affects human being and animals. Addressing this issue will require the application of modern technology. This investigation demonstrates that a wide range of parameters affect image segmentation-based technology. The dents and alterations causing diseases of the plants may be identified using this method in contrast to diseases that result in damages that cannot be seen from the photos of the plants (Loey et al. 2020). This investigation also reveals a deficient database that might be utilized to offer context for contrasting the captured photos (Barbedo 2018b). The segmentation of large-scale images under complex and real-world natural scenes will continue to be a focus and a challenging area of research because various natural environment factors, such as wind velocity, illumination, temperature, humidity, and background, influence the acquisition and obtaining images of disease. Another difficulty is that different diseases might have various symptoms and features that are quite similar (Barbedo 2018a).

The absence of appropriate tools for use in picture detection is the other problem. Most field specialists lack the necessary tools to interpret the photos they collect, making it challenging to gather precise information and recognize diseases (Ashqar and Abu-Naser 2019). Stringent data validity measures have led to low adoption rates in certain areas for agricultural technologies. For instance, during the fourth and sixth international conferences on soft computing and machine learning, several rules hindered the use of machine learning in

specific areas (Durmus et al. 2017). Because some of the outcomes from the ML functions do not comply with the necessary criteria, the rules forbid their usage in actual applications.

The above-mentioned difficulties demonstrate the wide range of applications for image-based detection but also make it less practical. The first option is to give sufficient information that can be utilized to reliably identify the diseases without confounding those closely related. Numerous diseases that have not been officially reported have been brought on by weather changes, global warming, and other effects. The answer is to expand the number of scientists involved and advocate for an improved method of information gathering (Sladojevic et al. 2016). Improved methods for recording information about the diseases are another potential option. Suppose data captioning is enhanced to include the fine features of the photos captured and the distinctions that define them. In that case, the problem of insufficient information about the disease can be resolved (Barbedo 2018a). The images should be carefully examined to determine whether one is affected or damaged.

The image-based detection method makes extracting and detecting diseases simpler due to high accuracy, minimal hassles, and less data duplication. For particular plants, such as tomatoes, a high accuracy rate is required to utilize the photos to identify the diseases that affect them and the amount of damage (Fuentes et al. 2017). Utilizing contemporary information storage techniques might provide a solution. For instance, using cloud computing might improve accessibility and storage accuracy. The alternative solution is to train those responsible factors for research and information analysis. The precision of the technique is increased by a trained DL algorithm (Rangarajan et al. 2018). The second option would be comprehending the phenotypes employed in disease detection (Ubbens and Stavness 2018). Typically, the phenotype used to identify diseases is a product of the weather and climate (Stewart et al. 2019; Rangarajan et al. 2018). Updating the systems would be the alternative way to guarantee that the data gathered is recent. The utilization of the technology is influenced by the considerable uncertainty surrounding disease detection. For instance, various uncertainties are connected to using Bayesian DL (Hernandez and López 2020). It implies that when used alone, this strategy is unreliable. In dealing with inaccurate and sluggish disease detection procedures, CNN techniques may also be useful (Singh et al. 2018). The techniques have various advantages and have been used to identify rice-affecting diseases (Li et al. 2019a, b).

The inefficiencies in procedures might be reduced by combining some approaches. For instance, the use of deep learning models with meta-architectures offers remedies for the problems encountered when utilizing different techniques for disease identification (Saleem et al. 2020a, b). The practice of deep convolutional generative adversarial networks would be the alternative method for identifying and analyzing the pictures (Li et al. 2019a, b). The adversarial networks' participation improves the detection method's accuracy. ViT can also be applied for disease identification as they have the potential to perform better for large datasets. Since many resources are required for image-based detection, the authorities should ensure they are available.

## 12 Discussion

Deep learning techniques have significantly improved plant disease identification by extracting intricate features from images and learning hierarchical representations. This review paper explores the employment of computer vision and deep learning and their collaborative application in plant disease identification. It highlights the need for efficient methods to monitor and diagnose plant health. It deliberates the role of various imaging technologies, including RGB and hyperspectral imaging, in capturing detailed plant visual information. The paper also evaluates imaging cameras and sensors for plant disease detection, highlighting their advantages and limitations. It discusses the historical perspective of traditional image-processing techniques and their transformative impact on deep learning. The review also explores advanced computer vision models, such as RNN and CNN, and their impact on accuracy and robustness in plant disease identification.

Academic research in object detection has yielded models that significantly benefit agricultural applications. Knowing their licensing is crucial for practical implementation. Paper; highlight some notable models, their uses in agriculture, and their licenses.

AlexNet and VGG-16 are widely used for plant disease identification. AlexNet showcased CNNs' potential, while VGG-16 improved accuracy with deeper architectures. Both models have permissive licenses—Apache 2.0 for AlexNet and BSD for VGG-16 making them free to use in production. ViTs divide images into patches and analyze relationships using self-attention mechanisms, capturing long-range dependencies which was developed by Google. ViTs are available under the Apache 2.0 license, allowing free use in production environments. Models like StyleGAN and BigGAN generate synthetic images of diseased plants, enhancing training datasets. StyleGAN is available under the NVIDIA Source Code License for non-commercial use, while BigGAN is under the Apache 2.0 license, allowing commercial use. SimCLR and BYOL use large amounts of unlabeled data to learn useful representations, which can be fine-tuned for plant disease detection. Both models are available under the Apache 2.0 license, suitable for commercial production use. CLIP and ALIGN integrate visual and textual information for zero-shot classification and image captioning. These models classify new diseases based on textual descriptions. CLIP is under the MIT license, while ALIGN's licensing can vary and should be checked for commercial use. EfficientNet and MobileNet balance accuracy and computational efficiency, ideal for mobile and edge computing in agriculture. Both models are available under the Apache 2.0 license, allowing for free use in commercial environments.

Understanding the licensing of these models helps practitioners make informed decisions about their implementation. Models like AlexNet, VGG-16, ViT, BigGAN, SimCLR, BYOL, CLIP, EfficientNet, and MobileNet are accessible for commercial use due to their permissive licenses. Utilizing these state-of-the-art models enhances plant disease detection systems, contributing to sustainable and productive farming practices.

However, challenges persist, such as the requirement for diverse and large datasets for model training, which limits the generalizability of models. Future research should focus on creating standardized datasets and fostering collaboration among researchers to address this issue. The interpretability of deep learning models is also crucial, as their inherent complexity poses challenges in understanding their decision-making processes. Addressing this interpretability gap is essential for gaining the trust of end-users, especially in agricultural settings where decisions based on disease identification models directly affect crop yield



and food security. Researchers and practitioners must explore model compression techniques, lightweight architectures, and edge computing solutions to make these technologies more accessible and feasible for real-world deployment.

## **12.1 Outcomes**

The thorough analysis of “Deep Learning–Computer Vision for Plant Disease Detection” summarizes the current knowledge of DL and CV, which offers insightful information to agricultural and plant science academicians, practitioners, and policymakers. The review’s outcomes include:

### **12.1.1 Knowledge synthesis**

The study provides a collective resource for researchers initiating or advancing in the field by integrating information on employing deep learning-driven computer vision in plant disease detection.

### **12.1.2 Guidance for practitioners**

A comprehensive evaluation of imaging methodologies, sensor configurations, camera selection, and model construction provides practitioners with the knowledge needed to facilitate well-informed assessments, allowing them to select the most appropriate strategies for their particular use cases.

### **12.1.3 Dataset evaluation**

Researchers acquire a more profound comprehension of the accessible plant disease datasets, enabling them to make knowledgeable judgments regarding dataset selection and emphasizing the significance of filling in the current data diversity and quality gaps.

### **12.1.4 Performance assessment**

In plant disease identification, discussing performance metrics facilitates a standardized method for evaluating the accuracy, precision, and recall of classification and object recognition models. This study helps researchers to examine their models more successfully.

### **12.1.5 Identification of challenges**

Future study initiatives are guided by the identification of constraints, such as a lack of annotated datasets, complex architectures of DL models, and problems with interpretability. In order to solve these issues, the paper promotes interdisciplinary cooperation and the investigation of novel solutions.

### 12.1.6 Roadmap for future research

The review study suggests a research roadmap, highlighting the need for improvements in explainable AI, more interaction between plant science, advanced deep learning models, and computer vision communities, and an emphasis on practical deployment factors.

## 13 Conclusion

Plant disease and pest detection approaches based on deep learning combined edge detection and feature extraction have wide advancing projections and high potential, in contrast to standard image processing techniques, which handle these jobs in various phases and linkages. Even though the technology for detecting pests and diseases in plants is advancing quickly and has been affecting agricultural research and its applications, some way is still there to go before it is fully developed for use in the actual natural environment, and there are still some issues that need to be fixed. This review made it possible to map the many deep-learning research studies on disease diagnosis with multiple data modalities. The following conclusion can be drawn from the study:

- Spectral imaging may be a crucial tool for determining the health of a crop because this is related to the extent of disease severity, the extent of spectral sensitivity to stress, and variations at different crop growth stages. Hyperspectral and multispectral pictures were highly helpful for disease identification and offered greater accuracy.
- The technical features (brightness, resolution, etc.), sample grounding settings (field or laboratory), and sample features can all have an impact on the spectrum reflectance (size, texture, humidity, etc.). It is necessary to do more research on reflectance based on crop vegetation indicators during all crop development and infection stages.
- Intelligent image segmentation and enormous data processing will be of utmost importance for identifying and treating agricultural diseases due to the quick growth of big data, IoT, and artificial intelligence technologies.
- Indeed, in the agricultural industry, neural networks and deep learning models showed an extensive potential to monitor crop health and its development and capture abnormalities, exceeding conventional machine learning methods. Therefore, combining many crucial components can result in an effective disease detection system.
- The choice of a learning framework and algorithm is required for multimodal deep learning applications. Multimodal fusion has recently demonstrated significant promise and is being employed more often in various fields, including object identification, sentiment analysis, human-robot interaction, and healthcare.

Agricultural practices can be transformed through enhanced precision, speed and scalability by integrating DL models into the identification of plant diseases. Therefore, eliminations of plant diseases would be easier since they can accurately diagnose in real-time translating into increased agricultural productivity and food security. Nonetheless, more research into diverse high quality datasets improvement on the efficiency of models as well as practical deployment strategies across different agricultural contexts needs to be done if these potential benefits are to be known in entirety. In order to overcome current challenges and realize

the full potential of DL models, researchers, practitioners and technology developers must work together so that we can have an agricultural future where advanced technology works hand in hand with the agricultural sector to ensure global food security issues are adequately addressed.

**Author contributions** Abhishek Upadhyay: Conceptualization, Methodology, Data curation, Software, Writing-Original draft preparation. Narendra Singh Chandel: Writing-Review and Editing, Visualization, Investigation. Krishna Pratap Singh and Subir Kumar Chakraborty: Visualization, Supervision, Resources, Writing- Review. Balaji M. Nandede: Resources, Visualization, Investigation. Mohit Kumar, A. Subeesh, and Konga Upendar: Resources, Writing-Review and Editing. Ali Salem and Ahmed Elbeltagi: Resources, Investigation, Funding acquisition, Writing-Review and Editing.

**Funding** Open access funding provided by University of Pécs. No funding was obtained for this study.

**Data availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abdulridha J, Ampatzidis Y, Kakarla SC, Roberts P (2019) Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques. *Precis Agric* 21:955–978
- Adao T, Hruška J, Pádua L, Bessa J, Peres E, Morais R, Sousa JJ (2017) Hyperspectral imaging: a review on UAV-Based sensors, Data Processing and Applications for Agriculture and Forestry. *Remote Sens* 9(11):1110. <https://doi.org/10.3390/rs9111110>
- Adit VV, Rubesh CV, Bharathi SS, Santhiya G, Anuradha R (2020) A comparison of Deep Learning algorithms for Plant Disease classification. *Advances in Cybernetics, Cognition, and Machine Learning for Communication Technologies. Lecture Notes in Electrical Engineering*, vol 643. Springer, Singapore, pp 153–161
- Agarwal M, Singh A, Arjaria S, Sinha A, Gupta S (2020) ToLeD: Tomato leaf disease detection using convolutional neural network. *Procedia Comput Sci* 167:293–301
- Aggarwal A, Mittal M, Battineni G (2021) Generative adversarial network: an overview of theory and applications. *Int J Inf Manag Data Insights* 1(1). <https://doi.org/10.1016/j.jjimei.2020.100004>
- Agustika DK, Mercuriani I, Purnomo CW, Hartono S, Triyana K, Iliescu DD, Leeson MS (2022) Fourier transform infrared spectrum pre-processing technique selection for detecting PYLCV-infected Chilli plants. *Spectrochim Acta Mol Biomol Spectrosc* 278:121339
- Ahmad JI, Hamid M, Yousaf S, Shah ST, Ahmad MO (2020) Optimizing pretrained convolutional neural networks for tomato leaf disease detection. *Complexity* 2020:8812019. <https://doi.org/10.1155/2020/8812019>
- Ait Nasser A, Akhloufi MA (2022) A review of recent advances in deep learning models for chest disease detection using radiography. *Diagnostics* 13(1):159. <https://doi.org/10.3390/diagnostics13010159>
- Al-bayati JSH, Üstündağ BB (2020) Evolutionary feature optimization for plant leaf disease detection by deep neural networks. *Int J Comput Intell Syst* 13(1):12–23

- Al-Saddik H, Simon JC, Brousse O, Cointault F (2017) Multispectral band selection for imaging sensor design for vineyard disease detection: case of Flavescence dorée. *Adv Anim Biosci* 8:150–155
- Albetis J, Jacquin A, Goulard M, Poilvé H, Rousseau J, Clenet H, Dedieu G, Duthoit S (2018) On the potentiality of UAV multispectral imagery to detect Flavescence dorée and grapevine trunk diseases. *Remote Sens* 11:23
- Aldakheel EA, Zakariah M, Alabdallal AH (2024) Detection and identification of plant leaf diseases using YOLOv4. *Front Plant Sci* 15:1355941. <https://doi.org/10.3389/fpls.2024.1355941>
- Ale L, Sheta A, Li L, Wang Y, Zhang N (2019) Deep learning-based plant disease detection for smart agriculture. In: *Proc IEEE Globecom Workshops*, Waikoloa, HI, USA, pp 1–6
- Alshammari H, Gasmi K, Ltaifa IB, Krichen M, Ammar LB, Mahmood A (2022) Olive disease classification based on Vision Transformer and CNN models. *Comput Intell Neurosci* 2022:3998193. <https://doi.org/10.1155/2022/3998193>
- Amara J, Bouaziz B, Algergawy A (2017) A deep learning-based approach for banana leaf diseases classification. In: *Datenbanksysteme für Business, Technologie und Web*, Stuttgart
- Anand R, Veni S, Aravindh J (2016) An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method. *Proc 2016 Int Conf Recent Trends Inf Technol ICRTIT*. <https://doi.org/10.1109/ICRTIT.2016.7569531>
- Anasta N, Setyawan F, Fitriawan H (2021) Disease detection in banana trees using an image processing-based thermal camera. *IOP conf ser: Earth Environ Sci*. IOP Publishing, Bristol, UK, p 012088
- Anonymous (2018b) India outranks US, China with world's highest net cropland area. Archived from the original on 18 November 2018. Retrieved January 17, 2024
- Anonymous (2018a) ILO modelled estimates database. ILOSTAT. International Labour Organization. Accessed February 07, 2024. <https://www.ilo.org/industries-and-sectors/agriculture-plantations-other-rural-sectors#events>
- Ashourloo D, Mobasheri MR, Huete A (2014) Developing two spectral disease indices for detection of wheat leaf rust (*Puccinia Triticina*). *Remote Sens* 6:4723–4740
- Ashqar B, Abu-Naser S (2019) Image-based tomato leaves disease detection using deep learning. *Int J Eng Res* 2(12):10–16
- Atanassova S, Nikolov P, Valchev N, Masheva S, Yorgov D (2019) Early detection of powdery mildew (*Podosphaera Xanthii*) on cucumber leaves based on visible and near-infrared spectroscopy. *AIP Conf Proc* 2075:160014
- Atila U, Uçar M, Akyol K, Uçar E (2021) Plant leaf disease classification using EfficientNet deep learning model. *Ecol Inf* 61:101182
- Atole RR, Park D (2018) A multiclass deep convolutional neural network classifier for detection of common rice plant anomalies. *Int J Adv Comput Sci Appl*
- Azadbakht M, Ashourloo D, Aghighi H, Radiom S, Alimohammadi A (2019) Wheat leaf rust detection at canopy scale under different LAI levels using machine learning techniques. *Comput Electron Agric*. <https://doi.org/10.1016/j.compag.2018.11.016>
- Baranowski P, Jedryczka M, Mazurek W, Babula-Skowronska D, Siedliska A, Kaczmarek J (2015) Hyperspectral and thermal imaging of oilseed rape (*Brassica napus*) response to fungal species of the genus *Alternaria*. *PLoS ONE* 10
- Barbedo JGA (2018a) Factors influencing the use of deep learning for plant disease recognition. *Biosyst Eng* 172:84–91
- Barbedo JGA (2018b) Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Comput Electron Agric* 153:46–53
- Barbedo JGA, Koenigkan LV, Santos TT (2016) Identifying multiple plant diseases using digital image processing. *Biosyst Eng* 147:104–116
- Bazi Y, Bashmal L, Al Rahhal MM, Dayil RA, Ajlan NA (2021) Vision transformers for remote sensing image classification. *Remote Sens* 13(3):1–20. <https://doi.org/10.3390/rs13030516>
- Bengio Y, Courville A, Vincent P (2013) Representation learning: a review and new perspectives. *IEEE Trans Pattern Anal Mach Intell* 35(8):1798–1828
- Berra EF, Gaulton R, Barr S (2017) Commercial off-the-shelf digital cameras on unmanned aerial vehicles for multitemporal monitoring of vegetation reflectance and NDVI. *IEEE Trans Geosci Remote Sens* 55:4878–4886
- Bhandari M, Shahi TB, Neupane A, Walsh KB (2023) BotanicX-AI: identification of tomato leaf diseases using an explanation-driven deep-learning model. *J Imaging* 9(2). <https://doi.org/10.3390/jimaging9020053>
- Bi S, Zhang Y, Dong M, Min H (2019) An embedded inference framework for convolutional neural network applications. *IEEE Access* PP:1–1. <https://doi.org/10.1109/ACCESS.2019.2956080>

- Bierman A, LaPlumm T, Cadle-Davidson L, Gadoury D, Martinez D, Sapkota S et al (2019) A high-throughput phenotyping system using machine vision to quantify severity of grapevine powdery mildew. *Plant Phenom* 2019:1–13. <https://doi.org/10.34133/2019/9209727>
- Bock CH, Barbedo JG, Del Ponte EM, Bohnenkamp D, Mahlein AK (2020) From visual estimates to fully automated sensor-based measurements of plant disease severity: Status and challenges for improving accuracy. *Phytopathol Res* 2:1–30
- Brahimi M, Boukhalfa K, Moussaoui A (2017) Deep learning for tomato diseases: classification and symptoms visualization. *Appl Artif Intell* 31(4):299–315
- Brahimi M, Arsenovic M, Laraba S, Sladojevic S, Boukhalfa K, Moussaoui A (2018) Deep learning for plant diseases: detection and saliency map visualization. In: Zhou J, Chen F (eds) *Human and Machine Learning. Human–Computer Interaction Series*. Springer, Cham. [https://doi.org/10.1007/978-3-319-90403-0\\_6](https://doi.org/10.1007/978-3-319-90403-0_6)
- Breton M, Eng P (2019) Overview of two performance metrics for object detection algorithms evaluation. *Defence Research and Development Canada Reference Document*
- Brown CD, Davis HT (2006) Receiver operating characteristics curves and related decision measures: a tutorial. *Chemometr Intell Lab Syst* 80(1):24–38
- Calderon R, Navas-Cortés JA, Lucena C, Zarco-Tejada PJ (2013) High-resolution airborne hyperspectral and thermal imagery for early detection of Verticillium wilt of olive using fluorescence, temperature, and narrow-band spectral indices. *Remote Sens Environ*. <https://doi.org/10.1016/j.rse.2013.07.031>
- Calderon R, Montes-Borrego M, Landa BB, Navas-Cortés JA, Zarco-Tejada PJ (2014) Detection of downy mildew of opium poppy using high-resolution multi-spectral and thermal imagery acquired with an unmanned aerial vehicle. *Precis Agric* 15:639–661
- Cap HQ, Suwa K, Fujita E, Kagiwada S, Uga H, Iyatomi H (2018) A deep learning approach for on-site plant leaf detection. In *Proc. IEEE 14th Int Colloq Signal Process & Its Appl (CSPA)*, Batu Feringghi, pp. 118–122
- Carranza-Rojas J, Goeau H, Bonnet P, Mata-Montero E, Joly A (2017) Going deeper in the automated identification of herbarium specimens. *BMC Evol Biol* 17:1–14
- Chakraborty SK, Chandel NS, Jat D, Tiwari MK, Rajwade YA, Subeesh A (2022) Deep learning approaches and interventions for futuristic engineering in agriculture. *Neural Comput Appl* 34(23):20539–20573
- Chandel NS, Rajwade YA, Dubey K, Chandel AK, Subeesh A, Tiwari MK (2022) Water stress identification of winter wheat crop with state-of-the-art AI techniques and high-resolution thermal-RGB imagery. *Plants* 11(23):3344
- Chandel NS, Chakraborty SK, Chandel AK, Dubey K, Subeesh A, Jat D, Rajwade YA (2024) State-of-the-art AI-enabled mobile device for real-time water stress detection of field crops. *Eng Appl Artif Intell* 131:107863
- Chen Y, Lin Z, Zhao X, Wang G, Gu Y (2014) Deep learning-based classification of hyperspectral data. *IEEE J Sel Top Appl Earth Obs Remote Sens* 7:2094–2107
- Chen Y, Jiang H, Li C, Jia X, Ghamisi P (2016) Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Trans Geosci Remote Sens* 54:6232–6251
- Chen T, Zhang J, Chen Y, Wan S, Zhang L (2019a) Detection of peanut leaf spots disease using canopy hyperspectral reflectance. *Comput Electron Agric* 156:677–683
- Chen Y, Li L, Yu L, Kholy AE, Ahmed F, Gan Z, Cheng Y, Liu J (2019b) UNITER: UNiversal Image-Text Representation Learning. *ArXiv. /abs/1909.11740*
- Cheng S, Cheng H, Yang R, Zhou J, Li Z, Shi B, Lee M, Ma Q (2023) A high performance wheat disease detection based on position information. *Plants* 12(5). <https://doi.org/10.3390/plants12051191>
- Cilia C, Panigada C, Rossini M, Meroni M, Busetto L, Amaducci S, Boschetti M, Picchi V, Colombo R (2014) Nitrogen status assessment for variable rate fertilization in maize through hyperspectral imagery. *Remote Sens* 6:6549–6565
- Costa JM, Grant OM, Chaves MM (2013) Thermography to explore plant-environment interactions. *J Exp Bot* 64:3937–3949
- Cruz A, Luvisi A, De Bellis L, Ampatzidis Y (2017) Vision-based plant disease detection system using transfer and deep learning. In *Proc ASABE Annu Int Meet*, Spokane, WA, USA, 2017
- Cui R, Li J, Wang Y, Yang S, Yu K, Zhao Y (2022) Hyperspectral imaging coupled with dual-channel convolutional neural network for early detection of apple valsa canker. *Comput Electron Agric* 2022:107411. <https://doi.org/10.1016/j.compag.2022.107411>
- Damos P (2015) Modular structure of web-based decision support systems for integrated pest management: a review. *Agron Sustain Dev* 35:1347–1372. <https://doi.org/10.1007/s13593-015-0319-9>
- Dandawate Y, Kokare R (2015) An automated approach for classification of plant diseases towards development of futuristic decision support system in Indian perspective. In *Proc IEEE Int Conf Adv Comput Commun Inf (ICACCI)*, Kochi, India
- Dash J, Pearse G, Watt M (2018) UAV multispectral imagery can complement satellite data for monitoring forest health. *Remote Sens* 10:1216

- DeChant C, Wiesner-Hanks T, Chen S, Stewart EL, Yosinski J, Gore MA et al (2017) Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning. *Phytopathology* 107:1426–1432. <https://doi.org/10.1094/PHYTO-11-16-0417-R>
- Deery D, Jimenez-Berni J, Jones H, Sirault X, Furbank R (2014) Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy* 4:349–379
- Dell'Aquila A (2009) Digital imaging information technology applied to seed germination testing: a review. *Agron Sustain Dev* 29:213–221. <https://doi.org/10.1051/agro:2008039>
- Demilie WB (2024) Plant disease detection and classification techniques: a comparative study of the performances. *J Big Data* 11(1):1–43. <https://doi.org/10.1186/s40537-023-00863-9>
- Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L (2009) ImageNet: A large-scale hierarchical image database. In *Proc IEEE Conf Comput Vis Pattern Recognit*, Miami, FL, USA, 20–25 June 2009; pp. 248–255
- Deng L, Mao Z, Li X, Hu Z, Duan F, Yan Y (2018) UAV-based multispectral remote sensing for precision agriculture: a comparison between different cameras. *J Photogramm Remote Sens* 146:124–136
- Dey AK, Sharma M, Meshram MR (2016) Image processing-based leaf rot disease detection of betel vine (*Piper betle* L.). *Procedia Comput Sci*. <https://doi.org/10.1016/j.procs.2016.05.262>
- Dey P, Mahmud T, Nahar SR, Hossain MS, Andersson K (2024) Plant disease detection in precision agriculture: Deep learning approaches. In *2nd Int Conf Intell Data Commun Technol Internet Things (IDCIoT)*, Bengaluru, India, 2024, pp. 661–667. <https://doi.org/10.1109/IDCIoT59759.2024.10467525>
- Dhakal A, Shakya S (2018) Image-based plant disease detection with deep learning. *Int J Comput Trends Technol* 61(1):26–29
- Dhingra G, Kumar V, Joshi HD (2018) Study of digital image processing techniques for leaf disease detection and classification. *Multimedia Tools Appl* 77(15):19951–20000. <https://doi.org/10.1007/s11042-017-5445-8>
- Di Gennaro SF, Battiston E, Di Marco S, Facini O, Matese A, Nocentini M, Palliotti A, Mugnai L (2016) UAV-based remote sensing to monitor grapevine leaf stripe disease within a vineyard affected by esca complex. *Phytopathol Mediterr* 55:262–275
- Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Hounsby N (2020) An image is worth 16×16 words: transformers for image recognition at scale. *ArXiv*. <http://arxiv.org/abs/2010.11929>
- Durmaz H, Günes EO, Kırıcı M (2017) Disease detection on the leaves of tomato plants using deep learning. In *Proc 6th Int Conf Agro-Geoinformatics*, Fairfax, VA, USA, 7–10 August 2017; pp. 1–5
- Esgario JGM, Krohling RA, Ventura JA (2020) Deep learning for classification and severity estimation of coffee leaf biotic stress. *Comput Electron Agric* 169:105162
- Everingham M, Van Gool L, Williams CK, Winn J, Zisserman A (2010) The Pascal Visual object classes (VOC) challenge. *Int J Comput Vis* 88:303–338
- Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognit Lett* 27(8):861–874
- Fazari A, Pellicer-Valero OJ, Gómez-Sanchis J, Bernardi B, Cubero S, Benalia S, Zimbalatti G, Blasco J (2021) Application of deep convolutional neural networks for the detection of anthracnose in olives using VIS/NIR hyperspectral images. *Comput Electron Agric* 187:106252. <https://doi.org/10.1016/j.compag.2021.106252>
- Ferentinos KP (2018) Deep learning models for plant disease detection and diagnosis. *Comput Electron Agric* 145:311–318
- Fergus R (2012) Deep learning methods for vision. *CVPR 2012 Tutorial*
- Fuentes A, Yoon S, Kim S, Park D (2017) A robust deep learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors* 17(9):2022
- Gadepally KC, Dhal SB, Bhandari M, Landivar J, Kalafatis S, Nowka K (2023) A Deep Transfer Learning-based approach for forecasting spatio-temporal features to maximize yield in cotton crops. In *2023 57th Annual Conference on Information Sciences and Systems (CISS)*, IEEE, pp 1–4. <https://doi.org/10.1109/CISS56502.2023.10089748>
- Gago J, Douthe C, Coopman R, Gallego P, Ribas-Carbo M, Flexas J, Escalona J, Medrano H (2015) UAVs challenge to assess water stress for sustainable agriculture. *Agric Water Manag* 153:9–19
- Gallo R, Ristorto G, Daglio G, Berta G, Lazzari M, Mazzetto F (2017) New solutions for the automatic early detection of diseases in vineyards through ground sensing approaches integrating LiDAR and optical sensors. *Chem Eng Trans* 58:673–678
- Geetharamani G, Pandian JA (2019) Identification of plant leaf diseases using a nine-layer deep convolutional neural network. *Comput Electr Eng* 76:323–338
- Geipel J, Link J, Wirwahn J, Claupein W (2016) A programmable aerial multispectral camera system for in-season crop biomass and nitrogen content estimation. *Agriculture* 6(4):4
- Gevaert CM, Suomalainen J, Tang J, Kooistra L (2015) Generation of spectral–temporal response surfaces by combining multispectral satellite and hyperspectral UAV imagery for precision agriculture applications. *IEEE J Sel Top Appl Earth Obs Remote Sens* 8:3140–3146

- Girshick R (2015) Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 11–18 December 2015
- Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 20–23 June 2014
- Gomez MS, Vergara A, Montenegro F, Alonso Ruiz H, Safari N, Raymaekers D, Ocimati W, Ntamwira J, Tits L, Omondi AB et al (2020) Detection of banana plants and their major diseases through aerial images and machine learning methods: a case study in DR Congo and Republic of Benin. *J Photogramm Remote Sens* 169:110–124
- Goncharov P, Ososkov G, Nechaevskiy A, Nestsiaenia I (2019) Disease detection on the plant leaves by deep learning. In Selected Papers from the XX International Conference on Neuro-informatics, Advances in Neural Computation, Machine Learning, and Cognitive Research II, pp 151–159, Moscow, Russia
- Gong X, Zhang S (2023) A high-precision detection method of apple leaf diseases using improved faster R-CNN. *Agric (Switzerland)* 13(2). <https://doi.org/10.3390/agriculture13020240>
- Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. In Advances in Neural Information Processing Systems, pp 2672–2680
- Granun E, Pérez-Bueno ML, Calderón CE, Ramos C, de Vicente A, Cazorla FM, Barón M (2015) Metabolic responses of avocado plants to stress induced by *Rosellinia Necatrix* analysed by fluorescence and thermal imaging. *Eur J Plant Pathol* 142:625–632
- Grinblat GL, Uzal LC, Lares MG, Granitto PM (2016) Deep learning for plant identification using vein morphological patterns. *Comput Electron Agric* 127:418–424
- Gruner E, Astor T, Wachendorf M (2019) Biomass prediction of heterogeneous temperate grasslands using an SfM approach based on UAV imaging. *Agronomy* 9:54
- Gudkov SV, Matveeva TA, Sarimov RM, Simakin AV, Stepanova EV, Moskovskiy MN, Dorokhov AS, Izmailov AY (2023) Optical methods for the detection of plant pathogens and diseases (review). *AgriEngineering* 5(4):1789–1812. <https://doi.org/10.3390/agriengineering5040110>
- Guo Y, Zhang J, Yin C et al (2020) Plant disease identification based on deep learning algorithm in smart farming. *Discrete Dyn Nat Soc* 2020:2479172
- Hernandez S, López JL (2020) Uncertainty quantification for plant disease detection using bayesian deep learning. *Appl Soft Comput* 96:106597
- Hinton GE, Salakhutdinov R (2006) Reducing the dimensionality of data with neural networks. *Science* 313(5786):504–507
- Hlaing CS, Zaw SMM (2018) Tomato plant diseases classification using statistical texture feature and colour feature. In Proc IEEE/ACIS 17th Int Conf Comput Inf Sci, Singapore
- Hu W, Huang Y, Wei L, Zhang F, Li H (2015) Deep convolutional neural networks for hyperspectral image classification. *J Sens* 2015:258619. <https://doi.org/10.1155/2015/258619>
- Hu F, Zhou M, Yan P, Li D, Lai W, Bian K, Dai R (2019) Identification of mine water inrush using laser-induced fluorescence spectroscopy combined with one-dimensional convolutional neural network. *RSC Adv* 9:7673–7679. <https://doi.org/10.1039/C9RA00805E>
- Huang Z, Zeng Z, Liu B, Fu D, Fu J (2020) Pixel-BERT: aligning image pixels with text by deep multi-modal transformers. *arXiv*. <https://arxiv.org/abs/2004.00849>
- Hughes DP, Salathe M (2015) An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowd sourcing. <https://arxiv.org/abs/1511.08060>
- Iqbal F, Lucieer A, Barry K (2018) Simplified radiometric calibration for UAS-mounted multispectral sensor. *Eur J Remote Sens* 51:301–313
- Jajja AI, Abbas A, Khattak HA, Niedbala G, Khalid A, Rauf HT, Kujawa S (2022) Compact Convolutional Transformer (CCT)-based approach for whitefly attack detection in cotton crops. *Agric (Switz)* 12(10). <https://doi.org/10.3390/agriculture12101529>
- 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>
- Jin X, Jie L, Wang S, Qi H, Li S (2018) Classifying wheat hyperspectral pixels of healthy heads and *Fusarium* head blight disease using a deep neural network in the wild field. *Remote Sens* 10:395. <https://doi.org/10.3390/rs10030395>
- Jin H, Li Y, Qi J, Feng J, Tian D, Mu W (2022) GrapeGAN: unsupervised image enhancement for improved grape leaf disease recognition. *Comput Electron Agric* 198:107055. <https://doi.org/10.1016/j.compag.2022.107055>



- Kalischuk M, Paret ML, Freeman JH, Raj D, Da Silva S, Eubanks S, Wiggins DJ, Lollar M, Marois JJ, Mellinger HC et al (2019) An improved crop scouting technique incorporating unmanned aerial vehicle-assisted multispectral crop imaging into conventional scouting practice for gummy stem blight in watermelon. *Plant Dis* 103:1642–1650. <https://doi.org/10.1094/PDIS-12-18-2267-RE>
- Karthik R, Joshua Alfred J, Joel Kennedy J (2023) Inception-based global context attention network for the classification of coffee leaf diseases. *Ecol Inf* 77:102213. <https://doi.org/10.1016/j.ecoinf.2023.102213>
- Kaur K, Bansal K (2024) Enhancing plant disease detection using advanced deep learning models. *Indian J Sci Technol* 17(17):1755–1766. <https://doi.org/10.17485/IJST/v17i17.536>
- Kerkech M, Hafiane A, Canals R (2018) Deep learning approach with colourimetric spaces and vegetation indices for vine disease detection in UAV images. *Comput Electron Agric* 155:237–243. <https://doi.org/10.1016/j.compag.2018.10.015>
- Kerkech M, Hafiane A, Canals R (2020) Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach. *Comput Electron Agric* 174:105446. <https://doi.org/10.1016/j.compag.2020.105446>
- Khakimov A, Salakhutdinov I, Omolikov A, Utaganov S (2022) Traditional and current-prospective methods of agricultural plant diseases detection: a review. *IOP Conf Ser Earth Environ Sci* 951(1):012002. <https://doi.org/10.1088/1755-1315/951/1/012002>
- Khan MA, Akram T, Sharif M, Saba T (2020) Fruit diseases classification: exploiting a hierarchical framework for deep features fusion and selection. *Multimed Tools Appl* 79(35–36):25763–25783. <https://doi.org/10.1007/s11042-020-08841-z>
- Khot LR, Sankaran S, Carter AH, Johnson DA, Cummings TF (2015) UAS imaging-based decision tools for arid winter wheat and irrigated potato production management. *Int J Remote Sens* 37:125–137. <https://doi.org/10.1080/01431161.2015.1112015>
- Krithika N, Selvarani AG (2017) An individual grape leaf disease identification using leaf skeletons and KNN classification. In: *Proceedings of the International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS 2017)*, Coimbatore, India, 2017-January, pp. 1–5. <https://doi.org/10.1109/ICIIECS.2017.8275951>
- Kumar S, Singh SK (2020) Occluded thermal face recognition using bag of CNN (BoCNN). *IEEE Signal Process Lett* 27:975–979. <https://doi.org/10.1109/LSP.2020.2984815>
- Kumar M, Chandel NS, Singh D, Rajput LS (2023) Soybean disease detection and segmentation based on Mask-RCNN algorithm. *J Exp Agric Int* 45(5):63–72. <https://doi.org/10.9734/jeai/2023/v45i52132>
- Kussul N, Lavreniuk M, Skakun S, Shelestov A (2017) Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geosci Remote Sens Lett* 14:778–782. <https://doi.org/10.1109/LGRS.2017.2681128>
- Lawaniya H (2020) Computer vision. *IET Computer Vision*. <https://github.com/himanshu6670/iot-object-detection->
- Lee SH, Chan CS, Wilkin P, Remagnino P (2015) Deep-Plant: Plant identification with convolutional neural networks. In: *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Quebec City, QC, Canada, 27–30 September 2015; pp. 452–456
- Lee SH, Chan CS, Mayo SJ, Remagnino P (2017) How deep learning extracts and learns leaf features for plant classification. *Pattern Recogn* 71:1–13
- Lei R, Jiang H, Hu F, Yan J, Zhu S (2017) Chlorophyll fluorescence lifetime imaging provides new insight into the chlorosis induced by plant virus infection. *Plant Cell Rep* 36:327–341
- Li J, Mi Y, Li G, Ju Z (2019a) CNN-based facial expression recognition from annotated RGB-D images for human–robot interaction. *Int J Humanoid Rob* 16(04):504–505
- Li LH, Yatskar M, Yin D, Hsieh CJ, Chang KW (2019b) VisualBERT: A simple and performant baseline for vision and language. <https://arxiv.org/pdf/1909.11740.pdf>
- Li M, Zhang Z, Lei L, Wang X, Guo X (2020a) Agricultural greenhouses detection in high-resolution satellite images based on convolutional neural networks: comparison of faster R-CNN, YOLO v3 and SSD. *Sens (Switzerland)* 20(17):1–14. <https://doi.org/10.3390/s20174938>
- Li X, Yin X, Li C, Zhang P, Hu X, Zhang L, Wang L, Hu H, Dong L, Wei F, Choi Y, Gao J (2020b) Oscar: object-semantics aligned pre-training for vision-language tasks. *ArXiv. /abs/2004.06165*
- Li M, Zhou G, Chen A, Yi J, Lu C, He M, Hu Y (2022) FWDGAN-based data augmentation for tomato leaf disease identification. *Comput Electron Agric* 194. <https://doi.org/10.1016/j.compag.2022.106779>
- Li J, Xu M, Xiang L, Chen D, Zhuang W, Yin X, Li Z (2023a) Large language models and foundation models in smart agriculture: Basics, opportunities, and challenges. *ArXiv. /abs/2308.06668*
- Li M, Cheng S, Cui J, Li C, Li Z, Zhou C, Lv C (2023b) High-performance plant pest and disease detection based on model ensemble with inception module and cluster algorithm. *Plants* 12(1). <https://doi.org/10.3390/plants12010200>



- Li X, Li X, Zhang S, Zhang G, Zhang M, Shang H (2023c) SLViT: shuffle-convolution-based lightweight vision transformer for effective diagnosis of sugarcane leaf diseases. *J King Saud Univ - Comput Inform Sci* 35(6):101401. <https://doi.org/10.1016/j.jksuci.2022.09.013>
- Liang P-S, Haff RP, Hua S-ST, Munyaneza JE, Mustafa T, Sarreal SBL (2018) Nondestructive detection of zebra chip disease in potatoes using near-infrared spectroscopy. *Biosyst Eng* 166:161–169
- Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Zitnick CL (2014) Microsoft COCO: Common objects in context. In: *Lecture Notes in Computer Science, Proceedings of the European Conference on Computer Vision—ECCV 2014, Zurich, Switzerland, 6–12 September*; Springer: Cham, Switzerland, 2014; Volume 8693, pp. 740–755
- Lin K, Gong L, Huang Y, Liu C, Pan J (2019) Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. *Front Plant Sci* 10:155
- Liu W, Angelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC (2016) SSD: Single shot multibox detector. In: *Proceedings of the European Conference on Computer Vision—ECCV 2016, Amsterdam, The Netherlands, 8–16 October 2016*
- Liu W, Wang Z, Liu X (2017) A survey of deep neural network architectures and their applications. *Neuro-computing* 234:11–26
- Liu B, Zhang Y, He D, Li Y (2018) Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry* 10:11. <https://doi.org/10.3390/sym10010011>
- Liu C, Cao Y, Wu E, Yang R, Xu H, Qiao Y (2023a) A discriminative model for early detection of anthracnose in strawberry plants based on hyperspectral imaging technology. *Remote Sens* 15(18):4640
- Liu Y, Liu J, Cheng W, Chen Z, Zhou J, Cheng H, Lv C (2023b) A high-precision plant disease detection method based on a dynamic pruning gate friendly to low-computing platforms. *Plants* 12(11). <https://doi.org/10.3390/plants12112073>
- Loey M, ElSawy A, Afify M (2020) Deep learning in plant diseases detection for agricultural crops. *Int J Service Sci Manage Eng Technol* 11(2):41–58
- Lopez-Lopez M, Calderon R, Gonzalez-Dugo V, Zarco-Tejada PJ, Fereres E (2016) Early detection and quantification of almond red leaf blotch using high-resolution hyperspectral and thermal imagery. *Remote Sens*. <https://doi.org/10.3390/rs8040276>
- Lowe A, Harrison N, French AP (2017) Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods* 13:80
- Ludovisi R, Tauro F, Salvati R, Khoury S, Mugnoz Scarascia G, Harfouche A (2017) UAV-based thermal imaging for high-throughput field phenotyping of black poplar response to drought. *Front Plant Sci* 8:1681
- MacPherson J, Voglhuber-Slavinsky A, Olbrisch M et al (2022) Future agricultural systems and the role of digitalization for achieving sustainability goals: a review. *Agron Sustain Dev* 42:70. <https://doi.org/10.1007/s13593-022-00792-6>
- Maes WH, Steppe K (2019) Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci* 24:152–164
- Mahajan U, Bundel BR (2016) Drones for Normalized Difference Vegetation Index (NDVI) to estimate crop health for precision agriculture: A cheaper alternative for spatial satellite sensors. In: *International Conference on Innovative Research in Agriculture, Food Science, Forestry, Horticulture, Aquaculture, Animal Sciences, Biodiversity, Ecological Sciences, and Climate Change*. Krishi Sanskriti Publications, New Delhi, India
- Mahlein AK (2016) Plant disease detection by imaging sensors: parallels and specific demands for precision agriculture and plant phenotyping. *Plant Dis* 100:241–251
- Mahlein AK, Oerke EC, Steiner U, Dehne HW (2012) Recent advances in sensing plant diseases for precision crop protection. *Eur J Plant Pathol* 133:197–209
- Mahlein AK, Alisaac E, Al Masri A, Behmann J, Dehne HW, Oerke EC (2019) Comparison and combination of thermal, fluorescence, and hyperspectral imaging for monitoring Fusarium head blight of wheat on spikelet scale. *Sensors* 19:2281
- Mahmud B, Hong G, Fong B (2023) A study of human–AI symbiosis for creative work: recent developments and future directions in deep learning. *ACM Trans Multimedia Comput Commun Appl* 20(2):47. <https://doi.org/10.1145/3542698>
- Mahmud BU, Al Mamun A, Hossen M, Hong GY, Jahan B (2024) Light-weight deep learning model for accelerating the classification of mango-leaf disease. *Emerg Sci J* 8:28–42. <https://doi.org/10.28991/ESJ-2024-08-01-03>
- Marcassa LG, Gasparoto M, Belasque J, Lins E, Dias Nunes F, Bagnato VS (2006) Fluorescence spectroscopy applied to orange trees. *Laser Phys* 16:884–888
- Martin T, Gasselini P, Hostiou N et al (2022) Robots and transformations of work in farms: a systematic review of the literature and a research agenda. *Agron Sustain Dev* 42:66. <https://doi.org/10.1007/s13593-022-00796-2>

- Martinelli F, Scalenghe R, Davino S et al (2015) Advanced methods of plant disease detection: a review. *Agron Sustain Dev* 35:1–25. <https://doi.org/10.1007/s13593-014-0246-1>
- Martinez-Martinez V, Gomez-Gil J, Machado ML, Pinto FAC (2018) Leaf and canopy reflectance spectrometry applied to the estimation of angular leaf spot disease severity of common bean crops. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0196072>
- Mathew A, Amudha P, Sivakumari S (2021) Deep learning techniques: an overview. *Adv Intell Syst Comput* 1141:599–608. [https://doi.org/10.1007/978-981-15-3383-9\\_54](https://doi.org/10.1007/978-981-15-3383-9_54)
- Mattupalli C, Moffet C, Shah K, Young C (2018) Supervised classification of RGB aerial imagery to evaluate the impact of a root rot disease. *Remote Sens* 10:917
- Matveyeva TA, Sarimov RM, Simakin AV et al (2022) Using fluorescence spectroscopy to detect rot in fruit and vegetable crops. *Appl Sci* 12:3391
- Maurício J, Domingues I, Bernardino J (2022) Comparing vision transformers and convolutional neural networks for image classification: a literature review. *Appl Sci* 13:5521. <https://doi.org/10.3390/app13095521>
- McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 5:115–133
- Melnyk P, You Z, Li K (2019) A high-performance CNN method for offline handwritten Chinese character recognition and visualization. *Soft Comput* 24:7977–7987
- Meunkaewjinda A, Kumsawat P, Attakitmongkol K, Srikaew A (2008) Grape leaf disease detection from color imagery using hybrid intelligent system. In: *Proc IEEE 5th Int Conf Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pp 513–516, Krabi
- Mohanty SP, Hughes DP, Salathé M (2016) Using deep learning for image-based plant disease detection. *Front Plant Sci* 7:1419. <https://doi.org/10.3389/fpls.2016.01419>
- Moshou D, Bravo C, Oberti R, West J, Bodria L, McCartney A, Ramon H (2005) Plant disease detection based on data fusion of hyperspectral and multi-spectral fluorescence imaging using Kohonen maps. *Real-Time Imaging* 11:75–83
- Nagasubramanian K, Jones S, Singh AK, Sarkar S, Singh A, Ganapathysubramanian B (2019) Plant disease identification using explainable 3D deep learning on hyperspectral images. *Plant Methods* 15:98
- Nebiker S, Lack N, Abächerli M, Läderach S (2016) Light-weight multispectral UAV sensors and their capabilities for predicting grain yield and detecting plant diseases. *ISPRS Int Arch Photogramm Remote Sens Spat Inf Sci* 41:963–970
- Ngugi LC, Abelwahab M, Abo-Zahhad M (2021) Recent advances in image processing techniques for automated leaf pest and disease recognition – A review. *Inform Process Agric* 8(1):27–51. <https://doi.org/10.1016/j.inpa.2020.04.004>
- Nguyen C, Sagan V, Maimaitiyiming M, Maimaitijiang M, Bhadra S, Kwasniewski MT (2020) Early detection of plant viral disease using hyperspectral imaging and deep learning. *Sensors* 21:742. <https://doi.org/10.3390/s21030742>
- Nhamo L, Ebrahim GY, Mabhaudhi T, Mpandeli S, Magombeyi M, Chitakira M, Magidi J, Sibanda M (2020) An assessment of groundwater use in irrigated agriculture using multi-spectral remote sensing. *Phys Chem Earth A/B/C* 115:102810
- Nijland W, De Jong R, De Jong SM, Wulder MA, Bater CW, Coops NC (2014) Monitoring plant condition and phenology using infrared-sensitive consumer-grade digital cameras. *Agric Meteorol* 184:98–106
- O'Shea K, Nash R (2015) An introduction to convolutional neural networks. <https://arxiv.org/abs/1511.08458>
- Oerke EC, Herzog K, Toepfer R (2016) Hyperspectral phenotyping of the reaction of grapevine genotypes to Plasmopara Viticola. *J Exp Bot*. <https://doi.org/10.1093/jxb/erw318>
- Oppenheim D, Shani G (2017) Potato disease classification using convolutional neural networks. *Adv Anim Biosci* 8:244–249
- Ozguven MM (2018) Determination of sugar beet leaf spot disease level (*Cercospora Beticola* Sacc.) With image processing technique by using drone. *Curr Investig Agric Curr Res* 5:621–631
- Ozguven MM, Adem K (2019) Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms. *Phys* 5:535:122537
- Padilla R, Netto SL, da Silva EAB (2020) A survey on performance metrics for object-detection algorithms. In: *International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 237–242. <https://doi.org/10.1109/IWSSIP48289.2020.9145130>
- Padol PB, Sawant SD (2016) Fusion classification technique used to detect downy and powdery mildew grape leaf diseases. In: *Proceedings of the International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)*, 2017:298–301
- Parmar N, Vaswani A, Uszkoreit J, Kaiser L, Shazeer N, Ku A, Tran D (2018) Image transformer. In: *Proceedings of the International Conference on Machine Learning (ICML)*

- Paulus S, Dupuis J, Mahlein AK, Kuhlmann H (2013) Surface feature-based classification of plant organs from 3D laser-scanned point clouds for plant phenotyping. *BMC Bioinformatics* 14:1–12. <https://doi.org/10.1186/1471-2105-14-55>
- Pavlovskaya NE, Gagarina IN (2018) In: Borodin DB, Gneusheva IA, Gorkova IV, Solokhina IYu, Kostromicheva EV, Lushnikov AV, Yakovleva IV, Ageyeva NY (eds) *Agro-biological substantiation of the technology of growing vegetable products with the use of biological means of protection*. Monograph. Publishing House of the Federal State Budgetary Educational Institution of Higher Education OSAU, Orel
- Pires RDL, Gonçalves DN, Oruê JPM et al (2016) Local descriptors for soybean disease recognition. *Comput Electron Agric* 125:48–55. <https://doi.org/10.1016/j.compag.2016.04.015>
- Polder G, Blok PM, de Villiers HAC, van der Wolf JM, Kamp J (2019) Potato virus Y detection in seed potatoes using deep learning on hyperspectral images. *Front Plant Sci* 10:1–12. <https://doi.org/10.3389/fpls.2019.00116>
- Powers DM (2011) Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *J Mach Learn Res* 2(1):37–63
- Prakash K, Saravanamoorthi P, Sathishkumar R, Parimala M (2017) A study of image processing in agriculture. *Int J Adv Netw Appl* 22:3311–3315
- Proctor C, He Y (2015) Workflow for building a hyperspectral UAV: challenges and opportunities. *ISPRS - Int Arch Photogramm Remote Sens Spat Inf Sci* 40:415–419
- Pydiapati R, Burks TF, Lee WS (2006) Identification of citrus disease using colour texture features and discriminant analysis. *Comput Electron Agric* 52(1–2):49–59. <https://doi.org/10.1016/j.compag.2006.02.001>
- Qian Q, Yu K, Yadav PK, Dhal S, Kalafatis S, Thomasson JA, Hardin IVRG (2022) Cotton crop disease detection on remotely collected aerial images with deep learning. In: *Autonomous Air and Ground Sensing systems for Agricultural optimization and phenotyping VII*. SPIE 12114:23–31. <https://doi.org/10.1117/12.2623039>
- Qin J, Wang B, Wu Y, Lu Q, Zhu H (2021) Identifying pine wood nematode disease using UAV images and deep learning algorithms. *Remote Sens* 13:162. <https://doi.org/10.3390/rs13010162>
- Qiu R, Yang C, Moghimi A, Zhang M, Steffenson BJ, Hirsch CD (2019) Detection of fusarium head blight in wheat using a deep neural network and colour imaging. *Remote Sens* 11(22):2658. <https://doi.org/10.3390/rs11222658>
- Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, Sastry G, Askell A, Mishkin P, Clark J, Krueger G, Sutskever I (2021) Learning transferable visual models from natural language supervision. arXiv preprint. <https://arxiv.org/pdf/2103.00020.pdf>
- Radočaj P, Radočaj D, Martinović G (2024) Image-based leaf disease recognition using transfer deep learning with a novel versatile optimization module. *Big Data Cogn Comput* 8(6):52. <https://doi.org/10.3390/bdcc8060052>
- Ramachandran P, Parmar N, Vaswani A, Bello I, Levskaya A, Shlens J (2019) Stand-alone self-attention in vision models. In: *Proceedings of NeurIPS*
- Ramcharan A, Baranowski K, McCloskey P, Ahmed B, Legg J, Hughes DP (2017) Deep learning for image-based cassava disease detection. *Front Plant Sci* 8:1852. <https://doi.org/10.3389/fpls.2017.01852>
- Ramcharan A, McCloskey P, Baranowski K (2019) A mobile-based deep learning model for cassava disease diagnosis. *Front Plant Sci* 10:272. <https://doi.org/10.3389/fpls.2019.00272>
- Rangarajan A, Purushothaman R, Ramesh A (2018) Tomato crop disease classification using pre-trained deep learning algorithm. *Procedia Comput Sci* 133:1040–1047. <https://doi.org/10.1016/j.procs.2018.07.196>
- Raza S-, Clarkson JP, Rajpoot NM (2015) Automatic detection of diseased tomato plants using thermal and stereo visible light images. *PLoS ONE* 10. <https://doi.org/10.1371/journal.pone.0123262>
- Reddy SRG, Varma GPS, Davuluri RL (2021) Optimized convolutional neural network model for plant species identification from leaf images using computer vision. *Int J Speech Technol*. <https://doi.org/10.1007/s10772-021-09882-6>
- Redmon J, Farhadi A (2017) YOLO9000: Better, faster, stronger. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017*
- Redmon J, Farhadi A (2018) YOLOv3: An incremental improvement. arXiv preprint. <https://arxiv.org/abs/1804.02767>
- Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016*
- Ren S, He K, Girshick R, Sun J (2017) Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell* 39:1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>

- Rivera-Romero CA, Palacios-Hernández ER, Trejo-Durán M, Rodríguez-Liñán MdC, Olivera-Reyna R, Morales-Saldaña JA (2020) Visible and near-infrared spectroscopy for detection of powdery mildew in Cucurbita pepo L. leaves. *J Appl Remote Sens* 14:044515. <https://doi.org/10.1117/1.JARS.14.044515>
- Roth L, Streit B (2017) Predicting cover crop biomass by lightweight UAS-based RGB and NIR photography: an applied photogrammetric approach. *Precis Agric* 19:93–114. <https://doi.org/10.1007/s11119-017-9448-6>
- Rothe PR, Kshirsagar RV (2014) Automated extraction of digital images features of three kinds of cotton leaf diseases. In: *Proceedings of the International Conference on Electronics, Communication and Computer Engineering (ICECCE)*, 2014:67–71. <https://doi.org/10.1109/ICECCE.2014.7086637>
- Roy AM, Bose R, Bhaduri J (2022) A fast accurate fine-grain object detection model based on YOLOv4 deep neural network. *Neural Comput Appl* 34(5):3895–3921. <https://doi.org/10.1007/s00521-021-06651-x>
- Saari H, Akujärvi A, Holmlund C, Ojanen H, Kaivosoja J, Nissinen A, Niemeläinen O (2017) Visible, very near IR and short-wave IR hyperspectral drone imaging system for agriculture and natural water applications. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 42:165–170
- Sahu P, Chug A, Singh AP, Singh D, Singh RP (2021) Challenges and issues in plant disease detection using deep learning. In: Dua M, Jain A (eds) *Handbook of Research on Machine Learning Techniques for Pattern Recognition and Information Security*. IGI Global, pp 56–74. <https://doi.org/10.4018/978-1-7998-3299-7.ch004>
- Saleem MH, Khanchi S, Potgieter J, Arif KM (2020a) Image-based plant disease identification by deep learning meta-architectures. *Plants* 9:1451
- Saleem M, Atta BM, Ali Z, Bilal M (2020b) Laser-induced fluorescence spectroscopy for early disease detection in grapefruit plants. *Photochemical Photobiological Sci* 19:713–721
- Sambasivam G, Opiyo GD (2021) A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks. *Egypt Inf J* 22:27–34
- Sandino J, Pegg G, Gonzalez F, Smith G (2018) Aerial mapping of forests affected by pathogens using UAVs, hyperspectral sensors, and artificial intelligence. *Sensors* 18:944
- Sankaran S, Maja JM, Buchanon S, Ehsani R (2013) Huanglongbing (citrus greening) detection using visible, near infrared and thermal imaging techniques. *Sensors* 13:2117–2130
- Sasaki Y (2007) The truth of the F-measure. *Teach Tutor Mater* 1:1–5
- Schikora M, Neupane B, Madhogaria S et al (2012) An image classification approach to analyze the suppression of plant immunity by the human pathogen *Salmonella Typhimurium*. *BMC Bioinformatics* 13:171
- Schoofs H, Delalieux S, Deckers T, Bylemans D (2020) Fire blight monitoring in pear orchards by unmanned airborne vehicles (UAV) systems carrying spectral sensors. *Agronomy* 10:615
- Sembiring A, Away Y, Arnia F, Muharar R (2021) Development of concise convolutional neural network for tomato plant disease classification based on leaf images. *Journal of Physics: Conference Series* 1845:012009
- Sethy PK, Barpanda NK, Rath AK, Behera SK (2020) Deep feature-based rice leaf disease identification using support vector machine. *Comput Electron Agric* 175:105527
- Shewale MV, Daruwalla RD (2023) High performance deep learning architecture for early detection and classification of plant leaf disease. *J Agric Food Res* 14. <https://doi.org/10.1016/j.jafr.2023.100675>
- Shin J, Chang YK, Heung B, Nguyen-Quang T, Price GW, Al-Mallahi A (2020) Effect of directional augmentation using supervised machine learning technologies: a case study of strawberry powdery mildew detection. *Biosyst Eng* 194:49–60
- Shoaib M, Shah B, Ali A, Ullah A, Alenezi F, Gechev T, Hussain T, Ali F (2023) An advanced deep learning models-based plant disease detection: a review of recent research. *Front Plant Sci* 14:1158933. <https://doi.org/10.3389/fpls.2023.1158933>
- Shrivastava V, Pradhan M, Minz S, Thakur M (2019) Rice plant disease classification using transfer learning of deep convolution neural network. *ISPRS - International archives of the photogrammetry, remote sensing and spatial Information sciences XLII*. –3/W6:631–635
- Shuaibu M, Lee WS, Schueller J, Gader P, Hong YK, Kim S (2018) Unsupervised hyperspectral band selection for apple Marssonina blotch detection. *Comput Electron Agric* 148:45–53
- Signoroni A, Savardi M, Baronio A, Benini S (2019) Deep learning meets hyperspectral image analysis: a multidisciplinary review. *J Imaging* 5:52
- Singh A, Ganapathysubramanian B, Sarkar S, Singh A (2018) Deep learning for plant stress phenotyping: Trends and future perspectives. *Trends Plant Sci* 23:883–898
- Singh V, Sharma N, Singh S (2019) A review of imaging techniques for plant disease detection. *Artif Intell Agric* 4:229–242. <https://doi.org/10.1016/j.aiia.2020.10.002>

- Sivakumar ANV, Li J, Scott S, Psota E, Jhala AJ, Luck JD, Shi Y (2020) Comparison of object detection and patch-based classification deep learning models on mid-to late-season weed detection in UAV imagery. *Remote Sens* 12(13). <https://doi.org/10.3390/rs12132136>
- 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* 2016:3289801
- Smigaj M, Gaulton R, Barr SL, Suárez JC (2015) UAV-borne thermal imaging for forest health monitoring: Detection of disease-induced canopy temperature increase. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 40:349–354
- Song X, Zhang G, Liu F, Li D, Zhao Y, Yang J (2016) Modeling spatio-temporal distribution of soil moisture by deep learning-based cellular automata model. *J Arid Land* 8:734–748
- Stewart EL, Wiesner-Hanks T, Kaczmar N et al (2019) Quantitative phenotyping of Northern Leaf Blight in UAV images using deep learning. *Remote Sens* 11:2209
- Su W, Zhu X, Cao Y, Li B, Lu L, Wei F, Dai J (2019) VL-BERT: pre-training of generic visual-linguistic representations. *ArXiv*. <https://arxiv.org/abs/1908.08530>
- Sugiura R, Tsuda S, Tamiya S, Itoh A, Nishiwaki K, Murakami N, Shibuya Y, Hirafuji M, Nuske S (2016) Field phenotyping system for the assessment of potato late blight resistance using RGB imagery from an unmanned aerial vehicle. *Biosyst Eng* 148:1–10. <https://doi.org/10.1016/j.biosystemseng.2016.04.010>
- Sundararajan SK, Sankaragomathi B, Priya DS (2019) Deep belief CNN feature representation-based content-based image retrieval for medical images. *J Med Syst* 43:1–9
- Suproteem K, Sarkara JD, Ehsanib R, Kumara V (2016) Towards autonomous phytopathology: Outcomes and challenges of citrus greening disease detection through close-range remote sensing. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm, Sweden, 16–20 May 2016
- Tabbakh A, Barpanda SS (2023) A deep features extraction model based on the transfer learning model and vision transformer TLMViT for plant disease classification. *IEEE Access* 11:45377–45392. <https://doi.org/10.1109/ACCESS.2023.3273317>
- Thai HT, Tran-Van NY, Le KH (2021) Artificial cognition for early leaf disease detection using vision transformers. In: *Proceedings of the International Conference on Advanced Technologies for Communications (ATC)*, pp. 33–38. <https://doi.org/10.1109/ATC52653.2021.9598303>
- Thakur PS, Khanna P, Sheorey T, Ojha A (2022) Vision transformer for plant disease detection: PlantViT. In: Raman B, Murala S, Chowdhury A, Dhall A, Goyal P (eds) *Computer Vision and Image Processing, CVIP 2021. Communications in Computer and Information Science*, vol 1567. Springer, Cham. [https://doi.org/10.1007/978-3-031-11346-8\\_43](https://doi.org/10.1007/978-3-031-11346-8_43)
- Thomas S, Kuska MT, Bohnenkamp D, Brügger A, Alisaac E, Wahabzada M, Behmann J, Mahlein AK (2018) Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective. *J Plant Dis Prot* 125:5–20
- Tian Y, Yang G, Wang Z, Li E, Liang Z (2019) Detection of apple lesions in orchards based on deep learning methods of cycle GAN and YOLOv3-dense. *J Sens* 2019:7630926
- Tiwari VM, Tarum G (2017) Plant leaf disease analysis using image processing technique with modified SVM-CS classifier. *Int J Eng Manage Technol* 5:11–17
- Tsaftaris SA, Minervini M, Scharr H (2016) Machine learning for plant phenotyping needs image processing. *Trends Plant Sci* 21:989–991
- Turkoglu M, Hanbay D (2019) Plant disease and pest detection using deep learning-based features. *Turkish J Electr Eng Comput Sci* 27:1636–1651
- Ubbens J, Stavness I (2018) Corrigendum: deep plant phenomics: a deep learning platform for complex plant phenotyping tasks. *Front Plant Sci* 8:2245
- Upadhyay A, Chandel NS, Chakraborty SK (2024) Disease control measures using vision-enabled agricultural robotics. In: Chouhan SS, Singh UP, Jain S (eds) *Applications of Computer Vision and Drone Technology in Agriculture 4.0*. Springer, Singapore. [https://doi.org/10.1007/978-981-99-8684-2\\_10](https://doi.org/10.1007/978-981-99-8684-2_10)
- Valasek J, Thomasson JA, Balota M, Oakes J (2016) Exploratory use of a UAV platform for variety selection in peanut. In: *Proceedings of the Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping*, Baltimore, Maryland, 18–19 April 2016. 98660F
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need. In: *NIPS*
- Verma S, Chug A, Singh AP, Sharma S, Rajvanshi P (2019) Deep learning-based mobile application for plant disease diagnosis: a proof of concept with a case study on tomato plant. *Applications of Image Processing and Soft Computing Systems in Agriculture*, pp 242–271
- Viljanen N, Honkavaara E, Näsi R, Hakala T, Niemeläinen O, Kaivosoja J (2018) A novel machine learning method for estimating biomass of grass swards using a photogrammetric canopy height model, images and vegetation indices captured by a drone. *Agriculture* 8:70



- Waheed A, Goyal M, Gupta D, Khanna A, Hassanien AE, Pandey HM (2020) An optimized dense convolutional neural network model for disease recognition and classification in corn leaf. *Comput Electron Agric* 175:105456
- Wallelign S, Polceanu M, Buche C (2018) Soybean plant disease identification using convolutional neural network. In: *Proc. Thirty-First International Florida Artificial Intelligence Research Society Conference (FLAIRS-31)*, pp. 146–151, Melbourne, FL, USA
- Wang M, Xiong Y, Ling N, Feng X, Zhong Z, Shen Q, Guo S (2013) Detection of the dynamic response of cucumber leaves to fusaric acid using thermal imaging. *Plant Physiol Biochem* 66:68–76
- Wang G, Sun Y, Wang J (2017) Automatic image-based plant disease severity estimation using deep learning. *Comput Intell Neurosci* 2017:2917536
- Wang D, Vinson R, Holmes M, Seibel G, Bechar A, Nof S, Tao Y (2019) Early detection of tomato spotted wilt virus by hyperspectral imaging and outlier removal auxiliary classifier generative adversarial nets (OR-AC-GAN). *Sci Rep* 9:4377
- Wang F, Rao Y, Luo Q, Jin X, Jiang Z, Zhang W, Li S (2022) Practical cucumber leaf disease recognition using improved Swin Transformer and small sample size. *Comput Electron Agric* 199. <https://doi.org/10.1016/j.compag.2022.107163>
- Wiesner-Hanks T, Stewart EL, Kaczmar N (2018) Image set for deep learning: field images of maize annotated with disease symptoms. *BMC Res Notes* 11:440
- Wu Q, Chen Y, Meng J (2020) DCGAN-based data augmentation for tomato leaf disease identification. *IEEE Access* 8:98716–98728. <https://doi.org/10.1109/ACCESS.2020.299700>
- Xia C, Wang L, Chung BK, Lee JM (2015) In situ 3D segmentation of individual plant leaves using a RGB-D camera for agricultural automation. *Sensors* 15:20463–20479
- Xie C, Yang C, He Y (2017) Hyperspectral imaging for classification of healthy and gray mold diseased tomato leaves with different infection severities. *Comput Electron Agric* 135:154–162
- Xing N, Yeung SH, Cai C, Ng TK, Wang W, Yang K, Yang N, Zhang M, Chen G, Ooi BC (2021) SINGA-Easy: An easy-to-use framework for multi-modal analysis. *arXiv*. <https://doi.org/10.1145/1122445.1122456>
- Yadav S, Sengar N, Singh A, Singh A, Dutta MK (2021) Identification of disease using deep learning and evaluation of bacteriosis in peach leaf. *Ecol Inform* 61:101247
- Yang X, Guo T (2017) Machine learning in plant disease research. *Eur J Biomed Res* 3:6–9
- Yang D, Li S, Peng Z, Wang P, Wang J, Yang H (2019a) MF-CNN: traffic flow prediction using convolutional neural network and multi-features fusion. *IEICE Trans Inf Syst* 102(8):1526–1536
- Yang N, Yuan M, Wang P, Zhang R, Sun J, Mao H (2019b) Tea diseases detection based on fast infrared thermal image processing technology. *J Sci Food Agric* 99:3459–3466
- Ye H, Huang W, Huang S, Cui B, Dong Y, Guo A, Ren Y, Jin Y (2020) Recognition of banana fusarium wilt based on UAV remote sensing. *Remote Sens* 12:938
- Yu K, Leufen G, Hunsche M, Noga G, Chen X, Bareth G (2013) Investigation of leaf diseases and estimation of chlorophyll concentration in seven barley varieties using fluorescence and hyperspectral indices. *Remote Sens*. <https://doi.org/10.3390/rs6010064>
- Yu K, Anderegg J, Mikaberidze A, Karisto P, Mascher F, McDonald BA et al (2018) Hyperspectral canopy sensing of wheat septoria tritici blotch disease. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2018.01195>
- Yu S, Xie L, Huang Q (2023) Inception convolutional vision transformers for plant disease identification. *Internet Things* 21:100650. <https://doi.org/10.1016/j.iot.2022.100650>
- Zaman-Allah M, Vergara O, Araus JL, Tarekne A, Magorokosho C, Zarco-Tejada PJ, Hornero A, Alba AH, Das B, Craufurd P et al (2015) Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods* 11:35
- Zhang D, Zhou X, Zhang J, Lan Y, Xu C, Liang D (2018a) Detection of rice sheath blight using an unmanned aerial system with high-resolution colour and multispectral imaging. *PLoS ONE* 13
- Zhang K, Wu Q, Liu A, Meng X (2018b) Can deep learning identify tomato leaf disease? *Advances in Multimedia* 2018:6710865
- Zhang X, Qiao Y, Meng F, Fan C, Zhang M (2018c) Identification of maize leaf diseases using improved deep convolutional neural networks. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2018.2844405>
- Zhang X, Han L, Dong Y, Shi Y, Huang W, Han L, González-Moo P, Ma H, Ye H, Sobehi T (2019) A deep learning-based approach for automated yellow rust disease detection from high-resolution hyperspectral UAV images. *Remote Sens* 11:1554
- Zhang Z, Yan Y, Dai X, Zhou D, Gai K (2021) Multi-modal pre-training for dense video captioning. <https://arxiv.org/pdf/2103.06561.pdf>
- Zhang L, Zhou G, Lu C, Chen A, Wang Y, Li L, Cai W (2022) MMDGAN: a fusion data augmentation method for tomato-leaf disease identification. *Applied Soft Computing* 123. Elsevier Ltd. <https://doi.org/10.1016/j.asoc.2022.108969>

- Zhao Y, Gu Y, Qin F, Li X, Ma Z, Zhao L, Li J, Cheng P, Pan Y, Wang H (2017) Application of near-infrared spectroscopy to quantitatively determine relative content of *Puccinia striiformis* f. sp. *tritici* DNA in wheat leaves in incubation period. *J. Spectrosc.* 2017:9740295
- Zhao H, Jia J, Koltun V (2020) Exploring self-attention for image recognition. *CVPR* 12:55–63
- Zhao Y, Sun C, Xu X, Chen J (2022) RIC-Net: a plant disease classification model based on the fusion of Inception and residual structure and embedded attention mechanism. *Comput Electron Agric* 193. <https://doi.org/10.1016/j.compag.2021.106644>
- Zhou J, Zhou J, Ye H, Ali ML, Nguyen HT, Chen P (2020) Classification of soybean leaf wilting due to drought stress using UAV-based imagery. *Comput Electron Agric* 175:105576
- Zhu N, Liu X, Liu Z, Hu K, Wang Y, Tan J, Huang M, Zhu Q, Ji X, Jiang Y (2018) Deep learning for smart agriculture: concepts, tools, applications, and opportunities. *Int J Agric Biol Eng* 11:32–44
- Zhu F, He M, Zheng Z (2020) Data augmentation using improved cDCGAN for plant vigor rating. *Comput Electron Agric* 175:105603. <https://doi.org/10.1016/j.compag.2020.105603>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

**Abhishek Upadhyay<sup>1</sup>**  · **Narendra Singh Chandel<sup>2</sup>** · **Krishna Pratap Singh<sup>3</sup>** · **Subir Kumar Chakraborty<sup>4</sup>** · **Balaji M. Nandede<sup>5</sup>** · **Mohit Kumar<sup>6</sup>** · **A. Subeesh<sup>2</sup>** · **Konga Upendar<sup>7</sup>** · **Ali Salem<sup>8,9</sup>** · **Ahmed Elbeltagi<sup>10</sup>**

✉ Abhishek Upadhyay  
abhiupadhyay7798@gmail.com

✉ Ali Salem  
salem.ali@mik.pte.hu

Narendra Singh Chandel  
narendracae@gmail.com

Krishna Pratap Singh  
kp.kp24@gmail.com

Subir Kumar Chakraborty  
subir8275@gmail.com

Balaji M. Nandede  
bm8ciae@gmail.com

Mohit Kumar  
mohit.agengg@sknau.ac.in

A. Subeesh  
subeesh.a@icar.gov.in

Konga Upendar  
k.upendar007@gmail.com

Ahmed Elbeltagi  
ahmedelbeltagy81@mans.edu.eg

<sup>1</sup> Department of Farm Machinery and Power Engineering, ICAR – Central Institute of Agricultural Engineering, Bhopal 462038, Madhya Pradesh, India

<sup>2</sup> Agricultural Mechanization Division, ICAR- Central Institute of Agricultural Engineering, Bhopal 462038, Madhya Pradesh, India

<sup>3</sup> Indian Council of Agricultural Research, New Delhi 110001, India

<sup>4</sup> Technology Transfer Division, ICAR- Central Institute of Agricultural Engineering, Bhopal 462038, Madhya Pradesh, India

- <sup>5</sup> Southern Region Farm Machinery Training & Testing Institute, Anantapur, Andhra pradesh, India
- <sup>6</sup> Sri Karan Narendra Agriculture University, Jobner, Jaipur 303329, Rajasthan, India
- <sup>7</sup> Infyz solutions, Hyderabad, Telangana, India
- <sup>8</sup> Civil Engineering Department, Faculty of Engineering, Minia University, Minya, Egypt
- <sup>9</sup> Structural Diagnostics and Analysis Research Group, Faculty of Engineering and Information Technology, University of Pécs, Pécs, Hungary
- <sup>10</sup> Agricultural Engineering Dept, Faculty of Agriculture, Mansoura University, Mansoura 35516, Egypt