



Systematic Review

A Systematic Review of Methods and Algorithms for the Intelligent Processing of Agricultural Data Applied to Sunflower Crops

Valentina Arustamyan ^{1,*}, Pavel Lyakhov ², Ulyana Lyakhova ², Ruslan Abdulkadirov ¹, Vyacheslav Rybin ³ and Denis Butusov ^{3,*}

¹ North-Caucasus Center for Mathematical Research, North-Caucasus Federal University, 355029 Stavropol, Russia; rabdulkadirov@ncfu.ru

² Department of Mathematical Modeling, North-Caucasus Federal University, 355029 Stavropol, Russia; ljahov@mail.ru (P.L.); uljahovs@mail.ru (U.L.)

³ Computer-Aided Design Department, Saint Petersburg State Electrotechnical University “LETI”, 197022 Saint Petersburg, Russia; vgrybin@etu.ru

* Correspondence: vaababoshina@ncfu.ru (V.A.); dnbutusov@etu.ru (D.B.)

Abstract

Food shortages are becoming increasingly urgent due to the growing global population. Enhancing oil crop yields, particularly sunflowers, is key to ensuring food security and the sustainable provision of vegetable fats essential for human nutrition and animal feed. However, sunflower yields are often reduced by diseases, pests, and other factors. Remote sensing technologies, such as unmanned aerial vehicle (UAV) scans and satellite monitoring, combined with machine learning algorithms, provide powerful tools for monitoring crop health, diagnosing diseases, mapping fields, and forecasting yields. These technologies enhance agricultural efficiency and reduce environmental impact, supporting sustainable development in agriculture. This systematic review aims to assess the accuracy of various machine learning technologies, including classification and segmentation algorithms, convolutional neural networks, random forests, and support vector machines. These methods are applied to monitor sunflower crop conditions, diagnose diseases, and forecast yields. It provides a comprehensive analysis of current methods and their potential for precision farming applications. The review also discusses future research directions, including the development of automated systems for crop monitoring and disease diagnostics.

Keywords: sunflower disease diagnosis; crop monitoring; sunflower disease classification; agricultural image analysis; intelligent monitoring systems



Academic Editor: Randy Goebel

Received: 3 September 2025

Revised: 3 October 2025

Accepted: 24 October 2025

Published: 27 October 2025

Citation: Arustamyan, V.; Lyakhov, P.; Lyakhova, U.; Abdulkadirov, R.; Rybin, V.; Butusov, D. A Systematic Review of Methods and Algorithms for the Intelligent Processing of Agricultural Data Applied to Sunflower Crops. *Mach. Learn. Knowl. Extr.* **2025**, *7*, 130. <https://doi.org/10.3390/make7040130>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The world population is projected to increase by 2 billion over the next 30 years [1] according to the United Nations (UN). Approximately 713–757 million people are projected to experience food insecurity in 2023 according to the latest report of the “State of Food Security and Nutrition in the World” (SOFI) in July 2024, which was published by the United Nations’ Food and Agriculture Organization (FAO) and the World Food Programme (WFP). This situation represents an increase of approximately 152 million compared to the year 2019 [2]. Furthermore, about 2.33 billion people currently experience moderate to severe food insecurity. Reports by FAO, the World Health Organization (WHO), and other UN organizations indicate that if current trends persist, approximately 582 million individuals will be affected by chronic undernourishment by the year 2030 [3].

Consequently, the necessity to enhance both the quantity and quality of food is more urgent than ever before, requiring the undertaking of appropriate research efforts. Oilseeds, including but not limited to soybean, oil palm, sunflower, and groundnut, have the potential to contribute to the alleviation of malnutrition. These crops are among the most significant sources of vegetable fats globally, ranking among the top ten crops in terms of total calorie production [4]. Sunflower (*Helianthus annuus*) is considered to be the optimal crop for cultivation in different climatic zones, ranging from temperate to subtropical regions. Sunflower meal contains up to 50% protein, as well as various amino acids, vitamins B and E, minerals, antioxidants, fiber, lipids, etc., making it an important component of human and farm animal diets [5].

The yield of any crop may be reduced by up to 40% as a consequence of various diseases, pests, and phytopathogens [6]. The common pesticides have been associated with a multitude of adverse consequences, including the diminution of biodiversity [7], the degradation of soil, water and air quality, and the emergence of health hazards that can progress to chronic diseases [8]. Phytosanitary regulations are defined as rules, requirements, and measures that are implemented to safeguard crops from diseases, pests, weeds, and other such harmful organisms. These regulations are followed through the responsible use of agrochemicals and biological protection methods with the overarching aim of causing minimal damage to the environment. The rapid implementation of phytosanitary measures necessitates continuous monitoring of agricultural fields to ensure that if a disease is detected, appropriate measures can be taken to prevent the infestation of the entire field. The monitoring of a single field is a complex and resource-intensive process, as disease symptoms and pests may be similar or combined, making accurate diagnosis difficult. Moreover, it is impractical to systematically check every plant during the growing season, especially if access to individual fields is difficult.

The utilization of computer vision and artificial intelligence technologies for the purpose of automated crop monitoring is a growing trend [9]. Despite the fact that a number of reviews have been conducted on the utilization of intelligent systems in sunflower cultivation, these have frequently been found to possess certain limitations that hinder the acquisition of a comprehensive overview of the situation. For instance, review [10] provides a comprehensive analysis of hyperspectral images; however, it is noted that this technology is currently inaccessible to farmers and offers only a limited comparison with more pragmatic alternatives. In a similar vein, [11] offers significant insights into the application of deep learning for disease diagnosis. However, the primary focus on Indian oil crops, as well as the omission of contemporary hybrid models, imposes limitations on its global relevance and applicability. These works highlight a clear gap in research: the lack of a systematic review that is globally relevant and covers the full range of modern data sources (such as UAVs and Sentinel-2) and intelligent processing methods specifically designed for sunflower cultivation.

The main goal of this systematic review is to investigate methods and algorithms for processing sunflower images and data in order to monitor plant health during the growing season and prevent disease emergence. Within the scope of this paper, the most relevant and significant algorithms are discussed, such as CNN [12,13] and multilayer perceptron (MLP) algorithms [14], as well as RF algorithms [15,16], and SVM [13]. In the field of machine learning, several algorithms have been employed for sunflower disease classification, including federated learning (FL) [6], K-Nearest Neighbour (KNN) algorithm, K-Means, and others. For segmentation tasks, models such as YOLO [17], UNet [18], Feature Pyramid Network (FPN) [19], and SVM [13] were chosen. This review is the first systematic synthesis to combine the analysis of satellite and UAV data and classification and segmentation methods for disease diagnosis and yield prediction in sunflowers. Unlike

existing reviews, this work systematically compares algorithm accuracy and synthesizes findings from 50 recent papers (2020–2025) gathered from various journals and conferences. The review focuses on the practical applicability of the algorithms to sunflower crops in the agricultural sector.

The rest of the paper is organized as follows. The Section 2 provides a detailed exposition of the search methods employed, the databases utilized, and the selection criteria for identifying suitable papers. The Section 3 presents a characterization of the most common sunflower diseases and a description of the datasets most frequently used in the reviewed papers, including Sentinel-2 satellite and UAV data. The Section 4 presents various metrics used to evaluate the accuracy of the methods in the reviewed works, as well as vegetation indices that are widely used for vegetation monitoring and plant health assessment. The Section 5 provides a comprehensive overview of the existing intelligent processing methods, including deep neural networks, random forests and other algorithms. These methods are utilized for various purposes, such as the classification and processing of sunflower data, disease diagnosis and yield prediction. The Section 6 provides a comprehensive overview of the YOLO neural network architecture, which is regarded as the most prevalent segmentation method. Additionally, it offers a comprehensive summary of studies that employ alternative methods for segmenting sunflower crops and lesions. The Section 7 presents the results of an analysis of advanced methods of sunflower data processing. It also highlights promising directions for further research, including the improvement of model accuracy and its adaptation to different agronomic conditions. The Section 8 considers the main problems of the reviewed works, as well as the issues of processing and integration of data from different sources. This section puts forward several suggestions for the enhancement of future research, including the improvement of the interpretability of the models, the use of time series data, and the integration of additional factors such as weather and soil characteristics. This is followed by the Section 9, which summarizes the results of the review, the conclusions that were drawn, and suggestions for future research in the area of sunflower crop monitoring using intelligent systems.

2. Strategy for Finding Studies for Systematic Review and Criteria for Selecting Papers

A meticulous search strategy and rigorous criteria were employed to undertake a systematic review of intelligent data processing methods and algorithms applied to sunflower crops by the PRISMA methodology [20] and the Kitchenham and Charters guidelines [21]. PRISMA 2020 checklist is available in Supplementary Materials. This review was retrospectively registered in the Open Science Framework (OSF), registration DOI: <https://doi.org/10.17605/OSF.IO/MEVSX>. The authors did not use any AI-assisted tools for the preparation of this manuscript. This comprehensive review systematically analyzes a wide range of studies on sunflower image analysis, selected through clear inclusion criteria based on relevance and publication date. The primary objective is to synthesize research that explores various machine learning techniques—including deep neural networks and other data processing methods—for applications such as mapping and disease detection. The study encompasses data captured by cameras, satellites, and UAVs.

The search for relevant papers was conducted through a range of major scientific databases including Google Scholar, IEEE Xplore, SpringerLink, Scopus and Wiley. This review incorporated studies that were published between 2020 and February, 2025. As illustrated in Table 1, the primary keywords were utilized to identify studies for a systematic review of methods and algorithms for the intelligent processing of agricultural data concerning sunflower crops. The search process also encompassed papers that addressed

the application of remote sensing techniques, such as Sentinel-1 and Sentinel-2, as well as UAV-derived data for the monitoring and diagnosis of sunflower diseases.

Table 1. Keywords for paper selection in the application of methods and algorithms for intelligent agricultural data processing to sunflower crops.

Target	Processing Method	Result
Sunflower	Machine learning	Recognition
Sunflower diseases	Artificial intelligent	Classification
UAV	Neural network	Segmentation
Satellite data	Deep learning	Detection
Vegetation indices		

The selection and assessment of papers for further inclusion in the review study were determined by the criteria outlined in Table 2. A key limitation identified in this review is the narrowly focused scope of the existing research. Due to the specificity of the field related to the application of machine learning techniques for sunflower crop monitoring and disease detection, there are a limited number of studies that meet all the strict selection criteria. Consequently, the systematic review encompasses papers that may not fully align with established criteria, particularly with regard to aspects such as completeness of method description, quality of data, or clarity of validation of results. This is attributable to the necessity of generating a representative sample of studies, considering the restricted number of available publications in the field. Selection of scientific publications was carried out according to a step-by-step scheme for processing studies devoted to the sunflower data processing methods, presented in Figure 1.

Table 2. Paper selection criteria for further inclusion in the review study.

Inclusion Criterion of the Study	Exclusion Criteria of the Study
Use of artificial intelligence algorithms for classification and segmentation of sunflower crop data	Lack of detailed description of the developed methods and algorithms necessary for verification of the obtained results
Papers describing machine learning methods for classification, segmentation and diagnosis of sunflower diseases	Papers that employ proprietary or poorly documented methods that prevent reproducibility and clear categorization
Works with open access for analysis and verification of results	Works without full text availability

The search strategy for studies relevant to sunflowers and the criteria for paper selection ensured the inclusion in the review of the most relevant and methodologically sound studies. The selection of works, accompanied by detailed descriptions of the algorithms and characteristics of the data used provided the foundation for further analysis. The majority of the papers utilize specific types of data required for the detection and monitoring of diseases in sunflower crops. The following analysis will address these data types in greater detail.

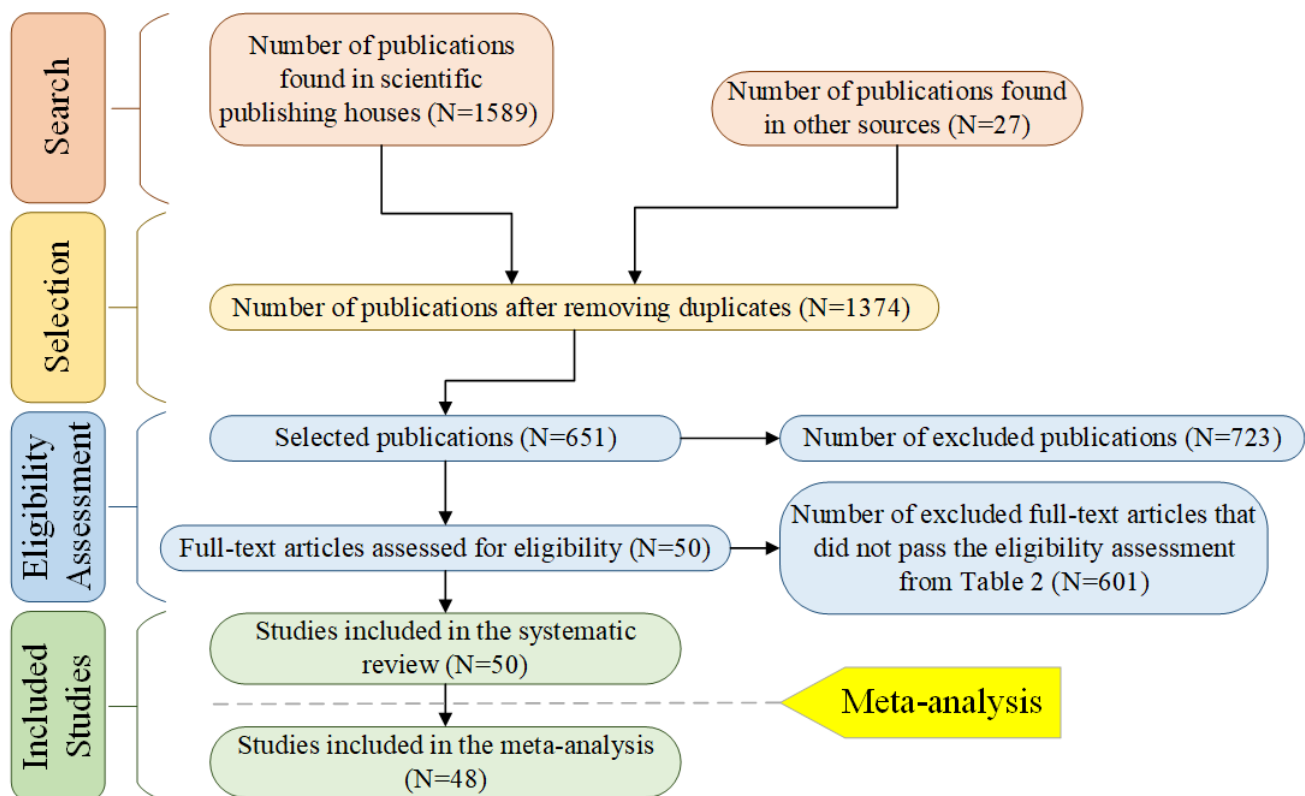


Figure 1. Selection scheme for studies devoted to the sunflower in a systematic review according to the PRISMA recommendation.

3. Data Used in Sunflower Crop Detection and Condition Studies

The studies selected in this systematic review utilize disparate sunflower data for crop analysis, mapping, or disease classification. In most cases, four to five different diseases affecting sunflowers are used in the works, each of which is caused by a different pathogen. The classification of these diseases primarily utilizes datasets collected with conventional cameras to obtain high-resolution images. The entire field is subject to monitoring through the utilization of UAV or satellite data. In addition, the utilization of multispectral imaging facilitates the acquisition of novel insights into the domain of crop analysis. This encompasses the assessment of green mass condition, water stress, and the correlation between plant condition and the prevailing growing season.

3.1. Common Diseases Affecting Sunflower Crops

Sunflower datasets under consideration in the papers are typically characterized by the presence of multiple disease categories, including, but not limited to, healthy plants, false powdery mildew, powdery mildew, rust, *Alternaria*, *Septoria*, *sclerotinia*, bacterial, *Verticillium*, *Fusarium* and others. Sunflower gray rot, a disease of considerable severity, is caused by the fungus *Botrytis cinerea* and affects various parts of the plant. Downy mildew (*Plasmopara halstedii*) affects young plants under high humidity, causing chlorotic spots and white patches on leaf undersides, leading to stunting and deformation [22,23]. Rust (*Puccinia helianthi*) is characterized by orange pustules on leaves, reducing photosynthetic rate [24]. The powdery mildew (*Golovinomyces cichoracearum*) manifests as white powdery plaque on leaves and stems, negatively impacting yield quality. Furthermore, it has been demonstrated that the plant's immune system is also compromised.

Alternaria (*Alternariaster helianthi*) causes brown necrotic lesions on leaves during early growth stages [25]. *Septorios* (*Septoria helianthi*) is characterized by the presence of gray or brown spots on leaves, which coalesce as the disease progresses, leading to leaf

withering. Sclerotiniosis (*Sclerotinia sclerotiorum*) is a fungal pathogen that affects stems, baskets and roots, causing them to rot. The presence of white mycelium with black sclerotia has been observed in the affected areas. The bacterium *Pseudomonas syringae*, commonly known as bacterial spotting, produces watery spots that darken and cause leaf desiccation. Verticillium wilt (*Verticillium dahliae*) and Fusarium wilt cause root infection and plant wilting [26,27]. The fungus *Phoma macdonaldii*, the causative agent of Phoma black stem, and Cercosporosis (*Cercospora helianthin*) produce stem lesions and leaf spots, respectively. The conditions and periods of occurrence of diseases affecting sunflowers are presented in Figure 2. Diseases that are contingent upon warm weather and high humidity are indicated by the color blue, while those diseases that require dry conditions are indicated by the color yellow. As illustrated in Figure 3, the following examples of plants have been observed to be affected by the diseases considered.

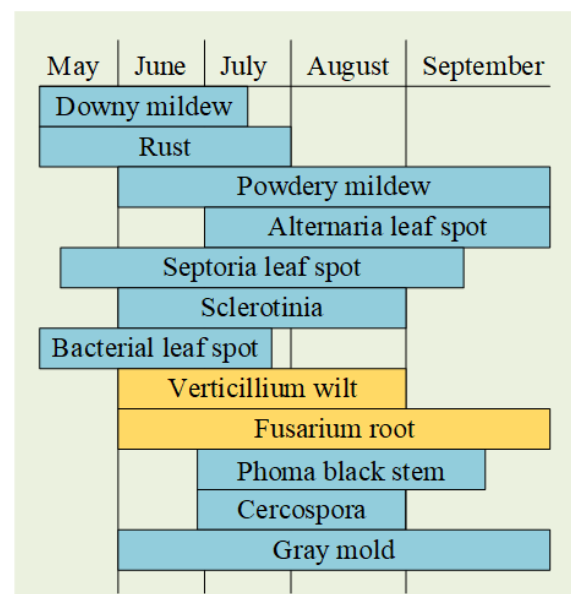


Figure 2. Periods, necessary conditions for the development of various sunflower diseases.

It should be noted that the review did not encompass all extant infections and fungi, given that the reviewed papers employed data derived from a circumscribed set of diseases. Multiple diseases can affect the field concurrently, which can complicate subsequent diagnosis. It is imperative to meticulously monitor crop growth and expeditiously implement remedial measures to eradicate insects and diseases, thereby averting further infections and ensuring optimal yield retention.

3.2. Sunflower Photo Datasets and Methods of Obtaining Them

There are currently two publicly available sunflower datasets: BARI-Sun and a set of synthetic images generated for segmentation tasks. The majority of the reviewed papers are authored by individuals who have created their datasets using conventional camera, UAV, or satellite data (Figure 4). The use of a multispectral camera in agricultural data collection facilitates the analysis of hidden plant characteristics, including growing season and water stress. Existing sunflower datasets and imaging techniques for various crop monitoring applications are presented further.

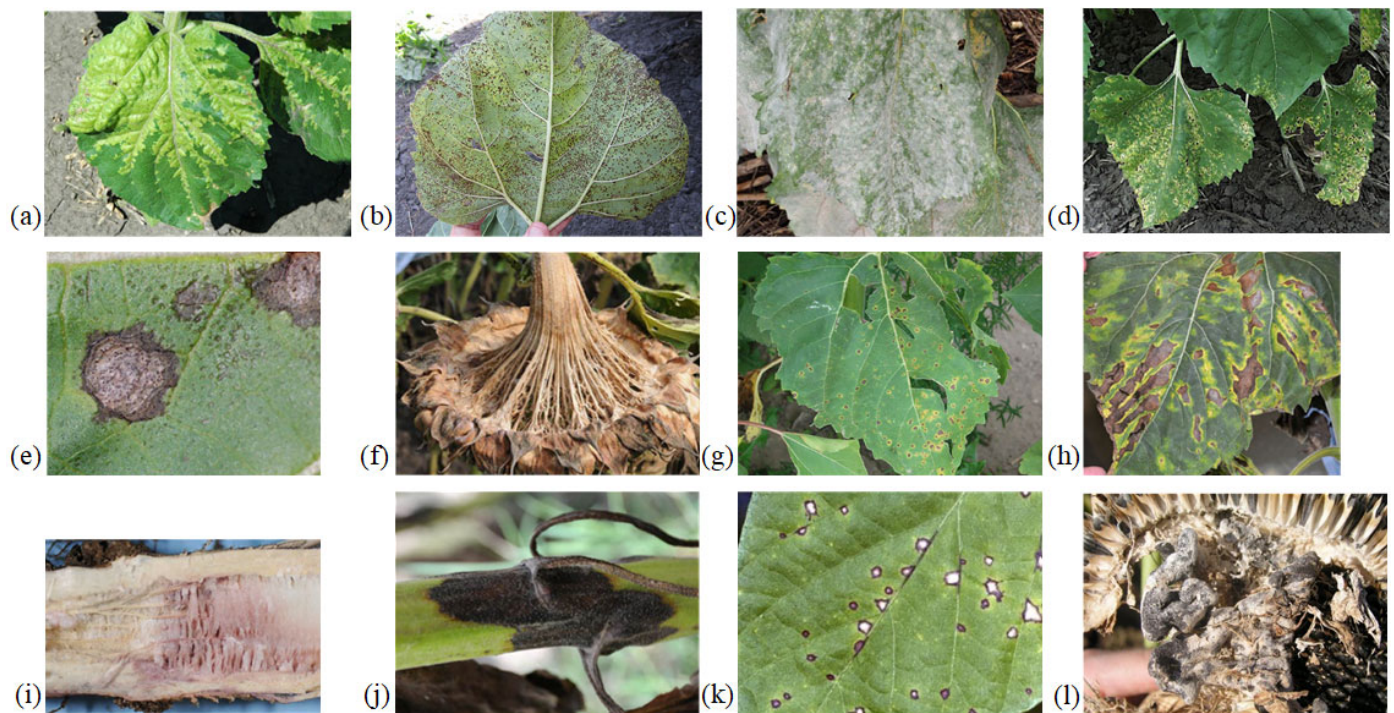


Figure 3. Examples of sunflower parts affected by different diseases: (a) downy mildew, (b) rust, (c) powdery mildew, (d) alternaria leaf spot, (e) septoria leaf spot, (f) sclerotinia, (g) bacterial leaf spot, (h) verticillium wilt, (i) fusarium root, (j) phoma black stem, (k) cercospora, (l) gray mold.

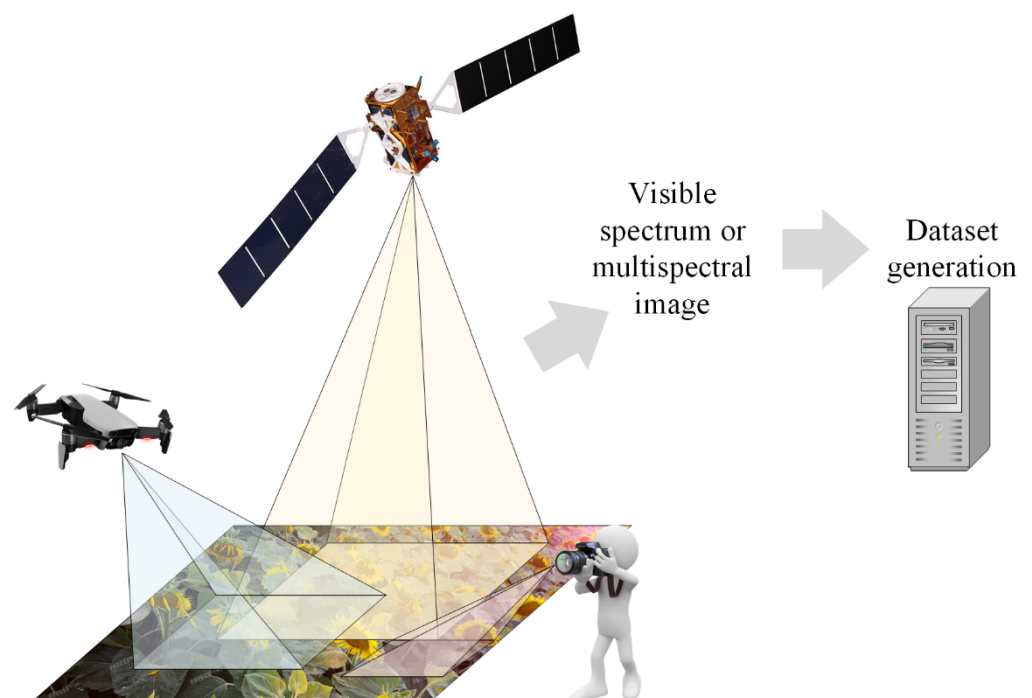


Figure 4. General scheme of the data collection process to generate the sunflower dataset.

3.2.1. Known Datasets

The PlantVillage is one of the most widely used datasets created for the automatic plant disease detection and classification using machine learning and computer vision methods [28]. This dataset contains over 54,000 images across 38 categories including both healthy leaf images and those of affected crops such as tomatoes, potatoes, corn, and others. BARI-Sunflower or BARI-Sun is the most commonly used open dataset in the

field of sunflower disease recognition [12,29–37]. In 2021, a large dataset of healthy and affected sunflowers was collected at the demonstration farm of the Bangladesh Agricultural Research Institute (BARI) in Gazipur, Bangladesh, consisting of 467 original images divided into four categories: fresh leaf, downy mildew, leaf scars, and gray mold [38]. The authors further augmented the dataset using various techniques, such as rotation, scaling, and shearing, increasing the dataset by 1668 images. The photographs are taken up close (Figure 5), allowing phytophysiolologists to compare the structures of different sunflower species. The heterogeneous weather and lighting conditions in the images complicate the task of disease detection by the naked eye.

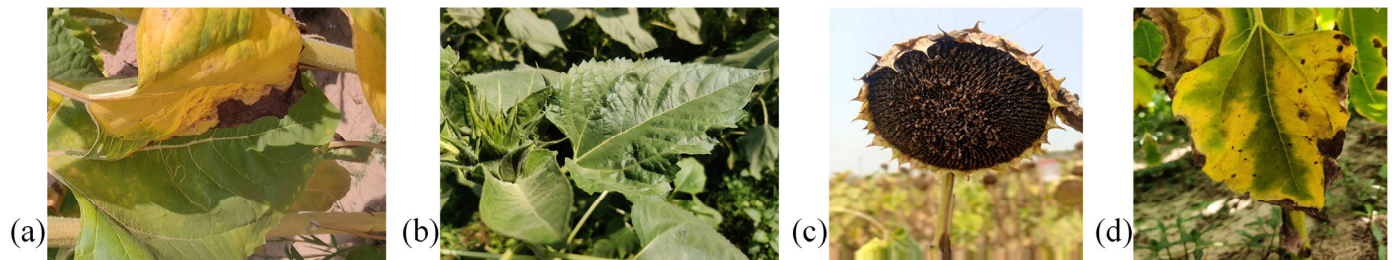


Figure 5. Example images from the dataset [38], where (a) Downy Mildew, (b) Fresh Leaf, (c) Gray Mold, (d) Leaf Scars.

The synthetic image dataset generated using cGAN is presented in the work [18]. The original images were collected using agricultural robots equipped with a four-channel camera that includes RGB channels and a near-infrared spectroscopy (NIR) channel. This dataset contains data for beets and sunflowers, categorized into crops, soil, and weeds. The sunflower data were collected in spring, over a month, starting from the seedling phase and continuing until the end of the period suitable for chemical treatments. Synthetic scenes in the multispectral images were created by replacing the objects (crops and weeds) with generated analogs, maintaining the realism of the scene. This approach increased the training dataset size and increased the training set diversity for segmentation [36].

The BARI-Sun dataset has gained widespread use in the classification of diseases due to its open access, extensive disease class representation, and high-quality images. Models trained on this dataset will be highly versatile due to the varying shooting conditions and diverse classes, making them easy to fine-tune. The generation of scenes or plants enables the qualitative and realistic augmentation of any dataset, and in combination with other augmentation techniques, it significantly expands the dataset.

3.2.2. Methods of Acquiring Sunflower Images from Agricultural Lands

In recent years, remote sensing methods have gained significant popularity in agriculture. Thanks to multispectral and hyperspectral imaging technologies, farmers and agronomists can obtain valuable information about crop health, growth dynamics, and potential stress responses [39]. Multi-spectral imaging techniques enable the analysis of various plant characteristics, opening new possibilities for precision agriculture [40]. Multispectral images are created by visualizing several electromagnetic spectrum ranges, typically from four to ten channels. The visible light, near-infrared, and red bands are most commonly used to diagnose plant sunlight absorption and assess the state of their photosynthetic apparatus. Hyperspectral images capture reflected light across hundreds of narrow spectral bands, allowing for a detailed analysis of plant chemical composition, pigment levels, moisture, and nutrients.

These data can be acquired in three main ways: through satellite platforms [40], UAVs [17], and ground-based spectrometric sensors. Satellite images provide global moni-

toring and allow for long-term crop analysis, but their spatial resolution is relatively low, making them suitable for monitoring large-scale fields. UAVs equipped with multispectral cameras offer a more expensive solution but provide high-resolution images of any part of the field from any altitude. Ground-based sensors are placed directly within the crop canopy, offering precise measurements of the spectral characteristics of individual plants and soil. Additionally, these sensors can be used to calibrate aerial and satellite data.

The spectral data obtained can be useful for calculating vegetation indices such as the normalized difference vegetation index (NDVI) [41], enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), normalized difference moisture index (NDMI), and others [42], as well as for detecting pathogens, pests, optimizing irrigation or fertilization, and predicting yield using time-series data. Remote sensing methods, including multispectral and hyperspectral imaging, have become an integral part of modern agriculture, providing highly accurate diagnostics of crop conditions. The variety of data sources enables the collection of information both for individual fields and entire regions. These data are actively used to calculate vegetation indices, making them key tools for precision farming and decision-making based on objective plant health parameters.

3.2.3. Sentinel-2 Data

Sentinel-2 data, provided by the European Space Agency (ESA), is widely used for monitoring various agricultural crops, including sunflowers [15,43–47]. These data offer high spatial (10–30 m per pixel) and temporal resolution. Consequently, time-series analysis allows for tracking phenological changes, assessing crop conditions, and forecasting yields. The multispectral images provided by Sentinel-2 satellites are created across 13 spectral bands, including visible, near-infrared, and shortwave infrared bands, and are updated every five days [40]. The analysis of satellite data enables the determination of vegetation status using indices like NDVI, EVI, and the normalized difference red edge index (NDRE). These indices are employed for phenological analysis, monitoring crop growth stages, detecting stress, and mapping crop fields. Additionally, the study of time-series satellite data enables crop development modeling, yield prediction, and assessment of climatic impacts for optimizing crop growth conditions, as well as tracking deviations from the norm, such as pests, diseases, or adverse weather conditions. Given the accessibility of Sentinel-2 data and their extensive coverage [15], there is an opportunity to develop powerful tools for efficient agricultural management, minimizing losses due to diseases and enhancing crop yields.

Due to its high spatial and temporal resolution, satellite data is a vital tool in crop monitoring. Its multispectral capabilities allow for calculating key vegetation indices and analyzing the phenological stages of plant growth. Regular data updates enable the creation of time-series necessary for identifying deviations, forecasting yields, and assessing the impact of climatic factors. The broad availability and coverage of satellite imagery create the foundation for the development of scalable agricultural management systems, capable of effectively minimizing crop losses and ensuring sustainable industry development.

3.2.4. Unmanned Aerial Vehicles Data

Drones equipped with multispectral and hyperspectral cameras provide highly accurate data on the condition of vegetation on agricultural lands, their stress levels, and the need for water and fertilizers [13,17,19,42,48,49]. The flexibility and timeliness of this data collection method are ensured by the ability to conduct necessary research at any time, regardless of weather conditions, while satellite data is updated every five days and may be unreadable due to cloud cover or precipitation [13]. Additionally, detailed UAV images can have resolution up to several centimeters per pixel, which is crucial for analyzing

soil heterogeneity, identifying areas with low fertility, detecting pathogenic lesions, and evaluating the effectiveness of agrotechnical measures and phytosanitary interventions. Using UAVs for crop monitoring reduces costs associated with fertilizers, irrigation, and plant protection, as well as enhancing yield by providing timely responses to emerging issues. Combined with satellite monitoring data, UAVs form the foundation for developing intelligent agricultural management systems, making the process more efficient and sustainable.

In the studies reviewed, UAV imagery, Sentinel-2 satellite data, and ground-based sensors are most commonly used to detect diseases at various plant growth stages. A significant portion of the research on sunflowers is dedicated to disease recognition, so this systematic review includes descriptions of common pathogens and their visual manifestations, which are crucial for subsequent automated image processing. The analysis of the presented datasets confirms the need for specialized methods of data processing that account for the specifics of agricultural objects. The next section presents statistical methods for evaluating intelligent systems and the application of vegetation indices for crop condition monitoring.

4. Methods for Statistical Evaluation of Intelligent Systems and Vegetation Indices

To process sunflower data, whether in the form of multispectral satellite images or simple disease photographs, different techniques such as machine learning algorithms are required. Several metrics are used to evaluate the accuracy of the developed models. This section also includes the characterization of some vegetation indices that improve the accuracy of diagnostics and forecasting.

4.1. Accuracy Metrics for Algorithms

Evaluating the accuracy of machine learning methods and neural network algorithms applied to sunflower image processing is crucial for ensuring the reliability and effectiveness of algorithms in real agricultural conditions. Accuracy metrics not only help assess model performance but also allow for comparisons of different approaches, enabling the selection of the most suitable method for a specific task [50]. One of the main metrics is accuracy, which measures the proportion of correctly classified objects among all objects in the test dataset:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)$$

where TP are the true positives, TN are the true negatives, FP are the false positives, FN are the false negatives. In multiclass tasks, such as diagnosing various sunflower diseases, accuracy may be insufficient. In these cases, more specific metrics such as precision and recall are used to provide more detailed information about the model's performance for each class. Precision measures the proportion of correctly classified objects of a certain class among all objects classified as that class:

$$Precision = \frac{TP}{TP + FP}. \quad (2)$$

For a model classifying diseases, precision indicates how many of the predicted disease cases are actually diseases [51]. Recall on the other hand measures the proportion of correctly classified objects of a given class among all objects that truly belong to that class:

$$Recall = \frac{TP}{TP + FN}. \quad (3)$$

This metric is crucial when minimizing the number of missed cases is important (e.g., undiagnosed diseases) [23]. To provide a more comprehensive evaluation of model performance, the F1-score is often used, which is the harmonic mean of precision and recall, helping to balance both metrics:

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (4)$$

The F1-score is especially useful in tasks with imbalanced classes, where some diseases or plant conditions occur less frequently, and the model might tend to ignore such cases [52]. The R^2 coefficient, or coefficient of determination, measures how well the model explains the variation of the dependent variable [44]. Values range from 0 to 1, with 1 indicating perfect fit:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (\hat{y}_i - \bar{y})^2}. \quad (5)$$

where y_i are actual values, $\hat{y}_i$ are predicted values, $\bar{y}$ is the mean of actual values. Time-series-based metrics, such as Root Mean Squared Error (RMSE), are commonly used for yield forecasting and assessing plant growth dynamics [45]. These metrics measure the average deviation of forecasts from actual values, which is essential for evaluating the accuracy of long-term forecasts, such as yield predictions. The lower the RMSE value, the higher the model's accuracy:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2}. \quad (6)$$

IoU (Intersection over Union) [48] is used in segmentation and object detection tasks and shows the intersection between the predicted region and the true region of the label:

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}, \quad (7)$$

where A is the predicted region and B is the true region. Mean Average Precision at IoU threshold 0.5 (mAP@0.5) represents the average precision across all classes at the IoU threshold of 0.5. This metric is used to evaluate the quality of object detection models [53]. First, Average Precision (AP) is calculated for each class, then averaged:

$$\text{mAP} = \frac{1}{N} \sum_{k=1}^{k=n} AP_k, \quad (8)$$

where N is the number of classes. Therefore, selecting and using the appropriate accuracy metrics is key to developing reliable and effective monitoring and diagnostic systems for agriculture, especially when working with large datasets from various sources, such as satellites, UAVs, and ground sensors.

4.2. Vegetation Indices

Vegetation indices are numerical indicators calculated based on the reflectance of vegetation in various spectral bands, such as blue, red, green, near-infrared, and red edge. These indices are used to assess the condition of vegetation, its productivity, moisture content, stress levels, and other agronomically significant parameters. Normalized Difference Vegetation Index (NDVI) is one of the most widely used indices for evaluating vegetation density and health [41]. It is based on the difference in reflectance between near-infrared

(NIR) and red (Red) bands. Chlorophyll in leaves strongly absorbs red light, while the leaf structure reflects near-infrared light. NDVI can be calculated as:

$$\text{NDVI} = \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}}, \text{NDVI} \in [-1; 1], \quad (9)$$

where *NIR* is the reflectance in the near-infrared band ($\approx 780\text{--}900$ nm) and *Red* is the reflectance in the red band ($\approx 630\text{--}680$ nm). NDVI values range from -1 to 0 , indicating water bodies, clouds, or snow, which interfere with visibility. Values between 0 and 0.2 indicate low vegetation or bare soil, values between 0.2 and 0.5 represent sparse vegetation, and values above 0.5 indicate dense, healthy vegetation. Despite its widespread use due to the simplicity of its calculation and interpretation, NDVI is sensitive to soil and atmospheric conditions, such as haze, clouds, and precipitation.

Enhanced Vegetation Index (EVI) was developed as an alternative to NDVI to correct for atmospheric effects and background soil noise [42]. This index accounts for aerosol correction coefficients and coefficients that normalize the sensitivity of spectral channels:

$$\text{EVI} = G \frac{\text{NIR} - \text{Red}}{\text{NIR} + C_1 \times \text{Red} - C_2 \times \text{Blue} + L}, \quad (10)$$

where *Blue* is the blue channel ($\approx 450\text{--}490$ nm), $G = 2.5$ is the contrast enhancement factor, $C_1 = 6$ and $C_2 = 7.5$ are aerosol correction coefficients, and $L = 0.5$ is the background correction factor. The contrast enhancement coefficient G enhances the difference between *NIR* and *Red* spectra, making the index more sensitive to changes in vegetation cover. The value of $G = 2.5$ was empirically determined from multiple studies to provide optimal contrast between vegetation and background objects such as soil. For dense forests, the coefficient may be increased to 3 , while for sparse vegetation or areas with high atmospheric noise, it may be decreased. The coefficients C_1 and C_2 adjust for the impact of atmospheric particles (aerosols) on measurements, especially for the red and blue channels. The background reflection coefficient L adjusts for the influence of soil, reducing errors associated with variations in reflectance on different soil types. L can vary from 0 to 1 , where 1 corresponds to sparse vegetation, deserts, and semi-deserts, and 0 corresponds to dense vegetation, such as forests. In agricultural landscapes, the vegetation density varies depending on the growth stage, and L would be set at 0.5 , representing grasslands and pastures, which can be further adjusted after analyzing vegetation cover for more accurate EVI calculation.

To correct for soil influence on spectral reflectance in sparse vegetation areas, Soil Adjusted Vegetation Index (SAVI) is used:

$$\text{SAVI} = \frac{(\text{NIR} - \text{Red}) \times (1 + L)}{\text{NIR} + \text{Red} + L}. \quad (11)$$

SAVI helps correct for soil background in dry areas but is less useful in regions with high vegetation density, unlike NDVI. Normalized Difference Moisture Index (NDMI) is used to assess moisture content in vegetation and soil [42]. It is calculated based on the reflectance in the near-infrared and shortwave infrared (SWIR) bands:

$$\text{NDMI} = \frac{\text{NIR} - \text{SWIR}}{\text{NIR} + \text{SWIR}}, \quad (12)$$

where *SWIR* is the shortwave infrared channel ($\approx 1550\text{--}1750$ nm). High NDMI values indicate good vegetation hydration, while low values indicate water stress or drought. NDMI is widely used for monitoring droughts, assessing crop conditions, and predicting fires and soil degradation zones. A limitation of NDMI is its requirement for *SWIR* data, which makes it less accessible. In general, the choice of index depends on the task at hand—

whether a general assessment is needed, vegetation density, or an analysis of water balance. Combining vegetation indices with machine learning methods and geographic information systems opens new possibilities for precision agriculture and ecosystem monitoring.

Vegetation indices have been identified as the most significant tools for evaluating vegetation condition and monitoring crop health. The utilization of these sensors facilitates a thorough examination of the crop's condition, enabling the discernment of stress conditions, the identification of diseases, and the estimation of yield. The integration of machine learning methodologies with vegetation indices serves as a foundational element in the development of intelligent monitoring systems, which, in turn, can effectively support agricultural decision-making processes. An integrated approach that incorporates both the accuracy metrics of machine learning architectures and the calculation of vegetation indices can more accurately assess current plant health, diagnose diseases, and make yield predictions. The comprehension and adept utilization of these methodologies and indices are imperative for the development of robust intelligent systems within the agricultural domain. Further research endeavors should prioritize the integration of disparate data sources and the refinement of processing algorithms to enhance the precision of monitoring and forecasting under real field conditions. The subsequent section will present the most common methods of sunflower data processing, including machine learning, clustering, and segmentation methods.

5. Methods of Machine Learning for Sunflower Image Processing

Artificial intelligence technologies enable efficient disease recognition, determining the extent of plant damage, and yield prediction. Machine learning helps automate tasks such as image classification, segmentation, anomaly detection, and other data processing tasks obtained from various remote sensing platforms (satellites, UAVs, ground sensors). In this context, algorithms such as CNN, SVM, RF, and others are actively used, which efficiently handle large data volumes and complex classification and prediction tasks. Figure 6 illustrates the schematic of machine learning methods used for sunflower image processing. This section considers the most effective machine learning methods employed for sunflower image processing and their applicability for agricultural monitoring.

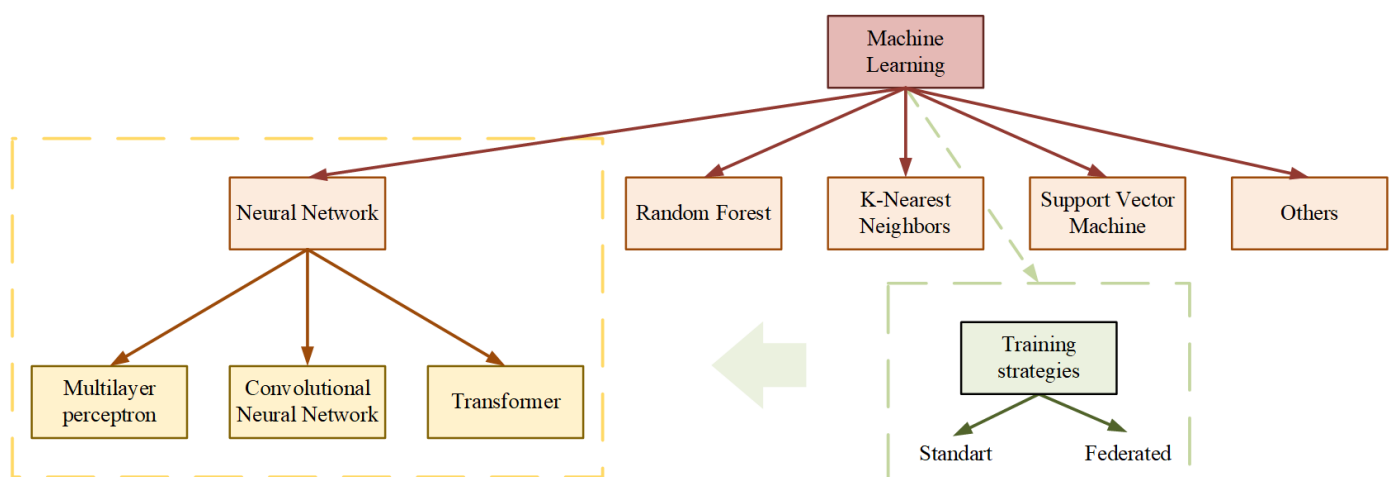


Figure 6. Schematic of machine learning methods used in the reviewed studies.

5.1. Convolutional Neural Networks

CNNs remain one of the most popular and widely used deep learning model for image processing of crop diseases. CNN architectures consist of an input layer that accepts a multidimensional array (image) (Figure 7). Convolutional layers take this image and

extract local features using a defined filter. After each convolution, an activation function, typically ReLU (Rectified Linear Unit), is applied, adding nonlinearity and helping the network learn more complex dependencies. After convolutional layers, a pooling layer (subsampling) is typically applied, which reduces the spatial dimensions of the feature map, lowering computational costs and preventing overfitting. These components of the architecture may be repeated several times, allowing the network to extract more complex, abstract features. Next, a flattening procedure is required, in which a multi-dimensional tensor of features is transformed into a one-dimensional vector. The fully connected layer then processes this vector and outputs the classification or prediction result. The output layer uses an activation function depending on the task; for multi-class classification, the Softmax function is used.

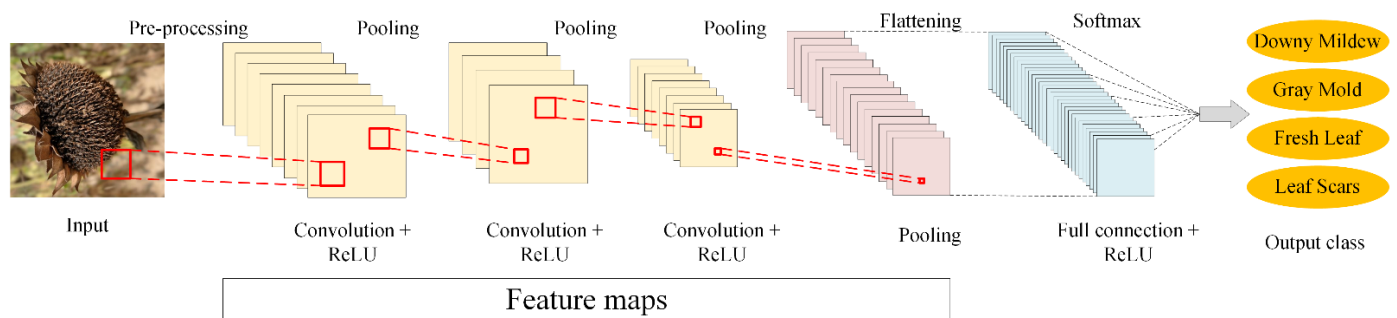


Figure 7. General CNN architecture applied to the BARI-Sun dataset [38].

In the study [12], the authors proposed a hybrid model based on VGG19 and a simple CNN, consisting of a convolutional layer, a pooling layer (max and average), and a fully connected layer, combined with transfer learning, which achieved 93.00% accuracy on the BARI dataset. Other models, such as Xception, Inception, MobileNet, ResNet, and DenseNet, were considered but showed lower results in combination with CNN. Authors of [14] assessed sunflower seed yield using MLP, radial basis function, artificial neural networks (ANN), CNN, and linear regression. The best result was shown by the convolutional neural network with an R^2 of 0.914, MAPE of 4.950, MAE of 0.163, and RMSE of 1.699. Authors of [54] focused on using ResNet152 to improve pathogen detection accuracy in sunflower fields. The dataset included rust, downy mildew, powdery mildew, and alternaria leaf blight with 405 photos per category. The test accuracy of ResNet152 was 98.02% on the original dataset. In [29] a comparative analysis of 38 neural network models trained with transfer learning on datasets of cauliflower and sunflower diseases is presented. Some models achieved 100% recognition accuracy on each of the datasets, raising questions about the reliability of the study's results. The EfficientNet_V2B2 model achieved an accuracy of 96.83%, which is a relatively high result. Similarly, ref. [30] demonstrated the advantage of EfficientNet over other deep models, with an accuracy of 97.90% using the well-known k-fold cross-validation algorithm on the BARI-Sunflower dataset.

Other works, such as [39], explored the use of deep fuzzy dimension reduction, CNN, and graph networks with an attention module, achieving optimal multispectral visualization, analysis, and classification of hyperspectral images. In the study by [36] a CNN architecture called LeafNet was proposed for recognizing various datasets, including [38]. This model, based on the popular VGG19 architecture, augmented with an attention module, achieved 96.00% accuracy in detecting sunflower diseases. Authors of [55] proposed a seven-layer CNN model for classifying sunflower diseases, such as alternaria leaf blight, downy mildew, phoma blight, and verticillium wilt. After feature extraction by the main model, SVM was used to classify leaves into two categories: healthy and diseased. The overall disease recognition accuracy for sunflowers was 98.00%. Other studies, such as

those by [23], focused on detecting downy mildew and the pests causing it, achieving a detection accuracy of 98.16% using CNN. Reference [56] introduced an ensemble of VGG-16 and MobileNet to detect sunflower diseases such as alternaria leaf blight, downy mildew, phoma blight, and verticillium wilt, which achieved 89.20% accuracy. This research was extended by the authors in [57]. Authors of [33,34] also used the classic CNN approach for training and classification with a slight dataset augmentation [38], achieving recognition accuracy of 97.88%. Various image preprocessing techniques, such as histogram equalization, contrast stretching, gamma correction, Gaussian noise, and Gaussian filters, improved the accuracy of disease recognition in sunflowers to 98.70% [34]. In [58] ResNet152 is used for the classification of four sunflower leaf diseases with similar symptoms, achieving an accuracy of 98.80%.

CNNs are one of the most efficient and widely used models for image processing in agriculture. These networks perform well in classification and segmentation tasks, extracting important features from images. In the context of sunflowers, CNNs are used to diagnose diseases such as powdery mildew and rust, and to predict yield. The advantages of using this algorithm include high accuracy in image classification, the ability to handle large amounts of data and generalize well, and the ability to extract complex abstract features. The disadvantages of CNNs can be highlighted as the need for large volumes of labeled data and high computational requirements when using deep architectures.

5.2. Random Forest Algorithm

The RF classifier is very popular for classification tasks in precision agriculture. It is an ensemble learning method based on the aggregation of multiple decision trees to improve prediction accuracy and reduce model variability. This algorithm was proposed in 2001 [59] and has since become one of the most popular tools in machine learning. The algorithm constructs multiple decision trees, each trained on a random subsample of the input training data (bootstrapping) (Figure 8). A random subset of features is selected at each tree node split, which helps to create diversity between the trees and reduce the correlation between them. This allows the algorithm to avoid overtraining and increases its ability to generalize. A voting method is used for classification tasks, and predictions are averaged for regression tasks. The algorithm can process a large amount of data with a large number of features without considering noise and missing information, but it is quite difficult to interpret and has a high computational cost when training on large datasets.

Sunflower crop mapping using the RF algorithm shows a high classification accuracy of 80% in Europe and the USA. This result depends on the creation of a consistent predictor dataset derived from monthly and seasonal composites using Sentinel-1 data. Automatic mapping can also be performed using a wavelet-inspired neural network, which outperformed traditional RF and SVM, achieving an overall accuracy result of 97.89% [16]. The authors also introduced factor analysis to reduce the hyperspectral image, and the wavelet transform allowed the extraction of important features which, together with the CNN spectral attention mechanism, achieved a high result in the mapping automation task. The combination of CNN and RF achieved an overall accuracy of 81.54% for seven sunflower disease classes, including downy mildew, rust, verticillium wilt, phoma black stem, alternaria leaf spot, sclerotinia stem rot, and phomopsis stem canker [16]. The proposed hybrid model used well-known back-propagation and gradient descent and consisted of two convolutional layers, max-pooling, and 14 decision trees. A sunflower disease dataset consisting of 5 classes (gray mold, downy mildew, leaf scar, and leaf rust) collected from different sources was seeded using the k-means algorithm [60], which showed an underperforming result for disease trait extraction. Using Sentinel-1 data, the work [43] focuses on sunflower area mapping and shows a 5% decrease in area planted. The RF

machine learning model based on multi-temporal metrics of satellite data takes into account phenological characteristics of sunflowers, such as the directional behavior of the plant during flowering. A paper [41] investigates the performance of different machine learning algorithms and datasets for mapping sunflower, soybean and maize using Sentinel-1 and Sentinel-2 satellite data. RF, MLP, and XGBoost classifiers were used to analyze the data, and the best result was demonstrated by the InceptionTime classifier, which was developed specifically for time series analysis and achieved a maximum accuracy of 91.60%.

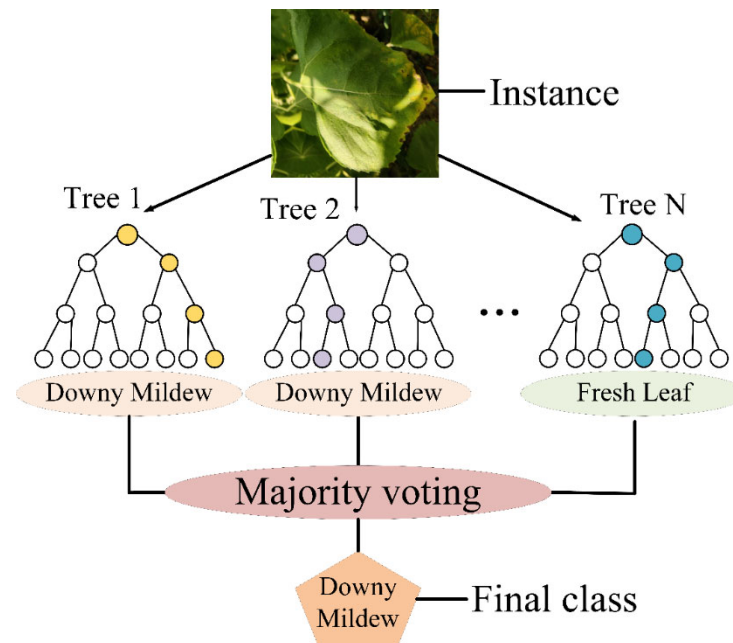


Figure 8. Main scheme of the Random Forest algorithm. Colored nodes represent individual decision trees within the ensemble, and arrows indicate the direction of data flow.

RF is an ensemble method that uses multiple decision trees for classification and prediction. It is a powerful tool due to its high accuracy, robustness to overfitting, and ability to handle noisy data. However, RF has lower accuracy than more sophisticated neural networks when dealing with large numbers of classes or high-dimensional data. In addition, training a random forest can require significant computational resources when dealing with large datasets.

5.3. Federated Learning

The FL method is distinguished by the fact that models are not trained on a single server but rather distributed across multiple devices or nodes (e.g., cell phones, drones, local servers). The training process occurs on a local device, with the model weights and gradients subsequently being transmitted to a central server for merging, as illustrated in Figure 9. The efficacy of this approach has been demonstrated in the context of sensitive data management and when dealing with substantial volumes of data. In the domain of precision agriculture, federated learning has emerged as a significant technological advancement. This innovative approach enables the local analysis of data from UAVs, field sensors, and satellites. It facilitates individual adaptation of models to diverse geographical regions, obviating the necessity for centralized data collection. The implementation of such a method has been demonstrated to enhance privacy, reduce network load, and enable customization of models to align with the unique characteristics of each farm. The subsequent works employ a unified dataset encompassing images of rust, downy mildew, Verticillium wilt, Septoria leaf spot, and powdery mildew. In the paper [61], the utilization

of the classical FL approach in conjunction with a CNN comprising three convolution layers, four max-pooling layers, and two fully connected layers is proposed for sunflower disease detection. This is achieved by the combination of local models that have been trained on data from five clients into a single global model using Federated Averaging. The model demonstrates high precision, recall, and F1-score, with accuracy metrics ranging from 85.31% to 91.86%. Addressing the imbalance in the data through various methods, such as reasoning the input data, has been demonstrated to enhance overall classification accuracy and reduce prediction error for rare classes.

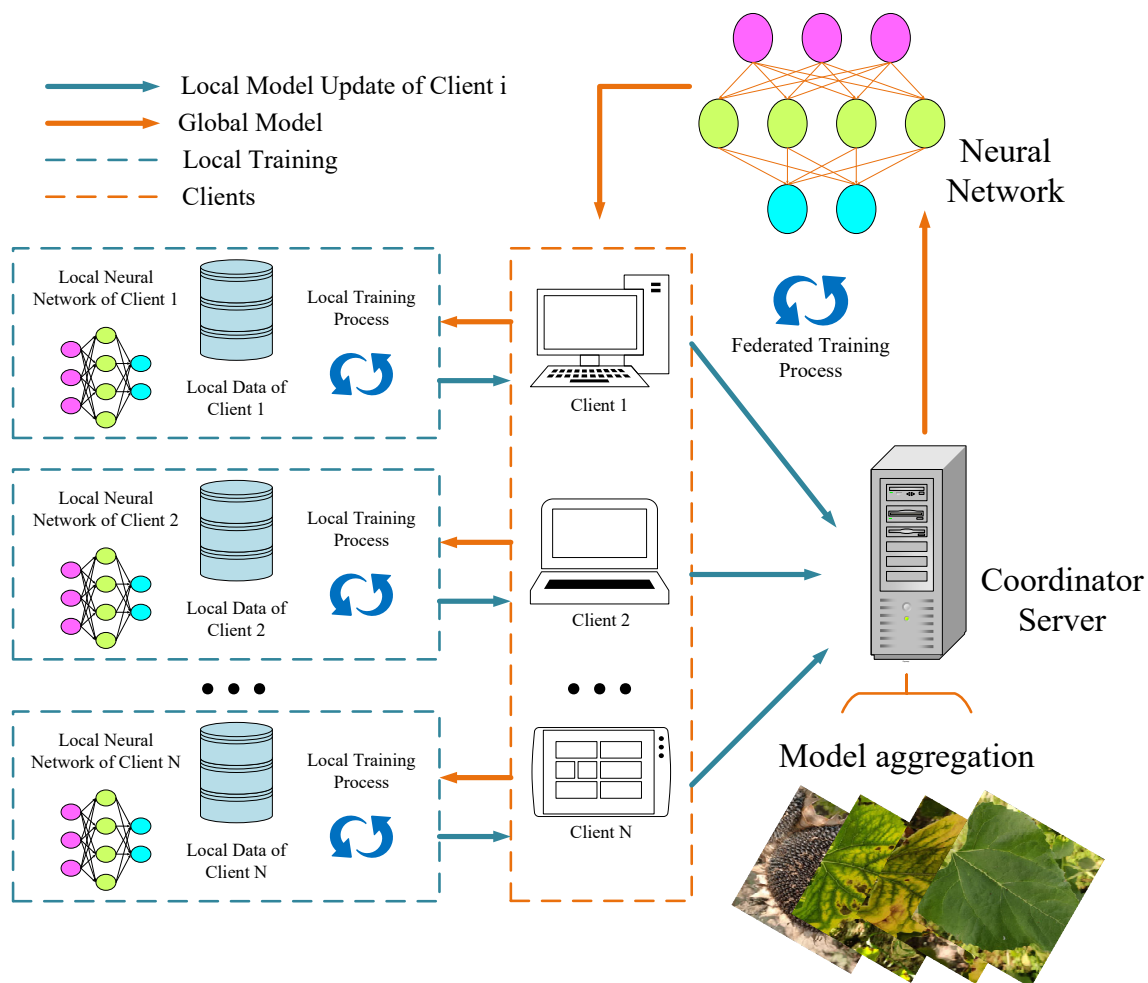


Figure 9. General flowchart of the Federated Learning method.

A study [50] focuses on the dynamic improvement of the model with an increasing number of clients and investigates the impact of federated averaging strategies. The experiment involves the use of FL in conjunction with the CNN (three convolutional layers with ReLU activation, three max pooling layers, and two fully connected layers with ReLU and Softmax activation) approach. The effectiveness of incremental federated learning under conditions of uneven data distribution is validated with a gradual increase in accuracy, starting from 82.95% in the first stage and reaching 94.87% by the final version of the model. The findings demonstrate that the employment of diverse data pooling and averaging mechanisms can exert a substantial influence on the ultimate classification outcome. A paper [51] examines the role of local and global learning in sunflower disease classification. The analysis of data from five clients with differing information representations demonstrates that FL assists in addressing imbalances in data, enhancing the quality and precision of classification from 88.76% to 95.72%. The primary distinction of the study [6] is that it

employs a four-tiered classification system to categorize disease severity, deviating from the conventional approach of employing class divisions. The efficacy of the method is enhanced by the integration of a hybrid CNN with an SVM, wherein the CNN comprises three convolutional layers with ReLU activation, three pooling layers, two fully connected layers with ReLU activation, and dropout and Softmax layers. This approach yielded macro-averaged accuracies ranging from 87.00% to 90.35%, with weighted accuracies ranging from 90.41% to 94.87%.

The FL approach has the potential to enhance the confidentiality of data, given that it does not involve the transmission of information outside the device. Moreover, federated learning serves to reduce the load on central servers and the network, thus conserving resources. A salient benefit of this approach is its capacity to adapt models to local data, thereby facilitating their utilization for tasks such as the monitoring of sunflower crops from drones or mobile devices. However, the process of synchronizing data from multiple devices can be challenging, and the devices must possess substantial computing capabilities. Furthermore, the quality of the model may be adversely affected by the limited data collected on a single device, rendering such systems less versatile.

5.4. Support Vector Machine

Support vector machines (SVMs) are a machine learning algorithm that is utilized for the purposes of classification and regression. The algorithm's objective is to identify a hyperplane that maximizes the separation of the data into distinct classes by providing the largest possible margin between them. The objects in closest proximity to the plane's boundary, also known as the partition, and which determine its position, are designated as support vectors (Figure 10). Moreover, it is possible to transform a space that is linearly indivisible by making it of higher dimensionality when employing the Kernel Trick method. In the domain of agriculture, this methodology is extensively employed for the analysis of images and information of various modalities. It demonstrates a high degree of training efficiency with small datasets, yielding optimal recognition results. However, when confronted with a substantial volume of processed data, this approach can exhibit significant resource demands. In the study [35], a hybrid approach of applying support vector machines (SVMs) in conjunction with convolutional neural networks (CNNs) for the multi-class classification of sunflower diseases is presented. This approach comprises three convolutional layers, three subsampling (max-pooling) layers and two fully connected layers. The second of these employs an SVM for final classification. The final model achieved an accuracy of 83.59% and an F1-score of 83.45%, with individual class accuracies ranging from 80.65% to 85.37%.

SVMs are frequently employed for classification with limited data and in tasks with clearly defined boundaries between classes, such as the identification of healthy and diseased plants. They are less computationally expensive and can be effective with small amounts of data. However, SVMs perform poorly on high-dimensional or complex data where the boundaries between classes are not obvious. Furthermore, to use SVMs efficiently, it is necessary to tune many hyperparameters.

5.5. K-Nearest Neighbor

The KNN algorithm is predicated on the identification of the k closest points in the feature space for a given point to be classified. The determination of the class of a new point is achieved through the analysis of the classes of its nearest neighbor points, with the selection of the most frequent class among them. The authors [62] utilized 280 artificially generated images and classified six types of flowering sunflowers using the k -neighbor algorithm, achieving an accuracy of 88.52%.

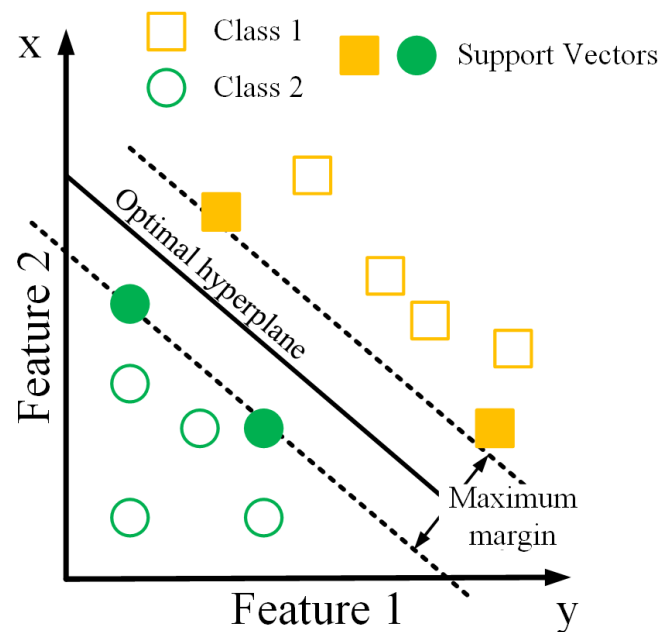


Figure 10. General scheme of the SVM method.

In the context of large and multi-class datasets, KNN is subject to a disadvantage in that the enhancement of the number of features results in a concomitant deterioration in the quality of classification. Additionally, the algorithm is prone to latency when processing voluminous datasets, due to the necessity of recalculating distances for all points within the training set for each new prediction. KNN is also susceptible to noise in the data, which can result in a reduction in accuracy. Notwithstanding the aforementioned limitations, the algorithm remains a viable option for classification tasks in environments characterized by limited resources and data, such as the early detection of sunflower diseases.

5.6. Other Methods for Processing Sunflower Data

K-Means is a popular data clustering algorithm that divides an input sample into k groups, or clusters, based on their similarity and without prior knowledge of class membership. The fundamental principle of the algorithm is as follows: a specified number of clusters is allocated, k initial points (centroids) in the feature space are arbitrarily selected, the belonging of clusters to the nearest centroids is determined (via the Euclidean distance), and a new centroid is calculated for each cluster as the average of all points that fall into it. The final two steps of this process are repeated iteratively until the centroids no longer change (i.e., until convergence) or until a predetermined number of iterations is reached. The efficacy of the algorithm in the domain of plant disease recognition has been demonstrated. The capabilities of the algorithm include lesion extraction (segmentation), classification of leaf texture and color to identify disease features, and optimization of data processing.

A study [40] was conducted to evaluate the suitability of NDVI maps derived from Sentinel-2 satellite data for identifying sunflower crops. The study utilized the k-means clustering algorithm to analyze the data, aiming to ascertain the effectiveness of the method. The method was employed to analyze NDVI data to identify homogeneous zones (subfields) within crop areas. The experimental results demonstrated the efficacy of the proposed method for the monitoring and management of sunflower crops in individual field plots. An R^2 regression coefficient of 0.73 was determined between yield and NDVI at approximately 63 days of crop maturity. The authors [63] utilized an enhanced YOLO_v4 model to detect verticillium wilt, powdery mildew, and rust. To enhance the efficacy of the algorithm, the

k-means clustering algorithm was employed, thereby generating a novel anchor frame size and rendering the network capable of detecting minute lesions. Furthermore, the backbone feature extraction network was substituted with MobileNet_V1, MobileNet_V2, and MobileNet_V3 to mitigate the computational intricacy of the proposed methodology. The metrics of precision, recall, and F1 achieved 0.84, 0.84, and 0.84, respectively, and the accuracy was 96.46%, which outperforms the original YOLO_v4. The authors of the dataset [38] developed a method [60] to segment affected sunflower patches using k-means clustering, then extract features and classify using Logistic Regression (LR), Naïve Bayes (NB), J48 [64] (a decision tree algorithm based on recursive attribute partitioning), K-star, and RF. The latter algorithm demonstrated superior performance in comparison to the other algorithms, attaining an average accuracy of 90.68%.

A study [65] investigated the potential for early detection of sunflower infestation by the parasitic weed *Orobanche cumana* using hyperspectral imaging. Logistic regression was utilized to differentiate between infested and uninfested plants, achieving an accuracy of up to 89.00%. Furthermore, analysis of a reduced spectral set demonstrated the potential of employing multispectral cameras to address the weed search problem. The objective of the present study [46] was to investigate the effect of climatic conditions on satellite signals using long time series of optical and Sentinel-1 C-band Synthetic Aperture Radar (SAR) data. The authors processed six years of data from 940 sunflower fields. The application of backscatter analysis, in conjunction with NDVI, GAI, and Fcover indices, in combination with radar signals, to monitor crop phenology, demonstrated the efficacy of SAR data in detecting anomalies in sunflower development, particularly when employing dual-orbit data and accounting for radar capture geometry. Recent years have seen a surge in popularity of ViT architectures in ensemble with deep neural networks, which have achieved a noteworthy 97.02% accuracy on the BARI-Sunflower dataset [37]. In the analysis of the BARI-Sunflower dataset, the BeiT architecture, which is based on a bi-navel encoder for image transformer, demonstrated the most optimal result, achieving an accuracy of 99.58% [66]. The Near-Real-Time Cumulative Sum (CuSum-NRT) algorithm [47] applied to Sentinel-1 SAR data utilizes the sensitivity of SAR data to soil, moisture, and crop condition to track changes throughout the growing season. A range of seasonal patterns of change were identified. Firstly, the VV polarization (the radar signal is emitted vertically and received vertically) was found to be more sensitive during periods of bare soil. Secondly, the VH polarization (the radar signal is emitted vertically and received horizontally) was found to be more sensitive during vegetation growth. In their study, [42] advanced a novel methodology for yield prediction, integrating time series data on vegetation indices and plant physiological parameters, including fresh weight, dry weight, and moisture factor. The utilization of physiological parameters yielded superior outcomes in comparison to the scenario where they were disregarded, resulting in an RMS error of 0.4075. The findings of such a study may be generalized to other crops in the future.

The presented alternative methods, in contrast to neural network approaches, offer distinct advantages in the processing of complex and heterogeneous data, as well as in tasks where classification or extraction of hidden patterns in the data is required. For instance, the K-Means algorithm can effectively segment images and identify anomalies in vegetation, which aids in the early detection of sunflower diseases. Nevertheless, such methods may be less effective at processing high-dimensional or complex data than state-of-the-art deep neural networks. Furthermore, the integration of machine learning and time series analysis techniques provides novel perspectives for the monitoring of changes in crop health throughout the season. These methods use information collected at different phases of plant growth to forecast yields or identify stress conditions in plants. However, it should be highlighted that each of these approaches has drawbacks, such as the difficulty in

interpreting the results and the need for large volumes of pre-processed data. A technique for segmenting sunflower patches impacted by the condition was created in [60] by the dataset's authors [38]. This method employed k-means clustering for the extraction of features and classification using Logistic Regression (LR), Naïve Bayes (NB), J48 [64] (a decision tree algorithm based on recursive attribute partitioning), K-star and RF.

Machine learning methodologies, encompassing CNN, RF, SVM, KNN, and FL, have been demonstrated to be efficacious in the domain of sunflower image processing, particularly in the context of disease diagnosis, yield prediction and crop mapping. Nevertheless, despite the high accuracy, there are several challenges, such as the necessity for large amounts of labeled data for training, the high computational complexity and the sensitivity of the models to noise in the data. Moreover, the adaptation of machine learning methods to differing agronomic conditions and discrete tasks is a challenge that persists at the present time. Prospects in this area are related to further improvement of models, including the integration of additional factors such as climate and soil data to improve the accuracy of predictions and monitoring. Furthermore, there is a necessity to develop methods that are less dependent on large amounts of data and more adaptable to real-world conditions. Segmentation is an integral step in the image preprocessing pipeline. This step has been shown to increase disease recognition accuracy and allow for precise lesion diagnosis in every section of an agricultural site.

6. Segmentation Methods Applied to Sunflower Crop Analysis

Image segmentation is currently one of the most significant data processing steps in machine learning and computer vision, particularly in the context of crop monitoring, such as sunflowers. This process facilitates the isolation and classification of specific regions within an image, a procedure that is imperative for precise disease diagnosis, plant health evaluation, and yield estimation. Segmentation is a process that can be used to divide an image into different zones corresponding to different objects or features. It can also be used to improve the accuracy and efficiency of processing data from different remote sensing technologies, such as satellite imagery, UAV data, or multispectral cameras. In order to accurately analyze the condition of sunflower crops, it is necessary to identify specific areas, such as disease-affected parts of the plant or healthy areas. Thus, the use of these methods in combination with more advanced approaches can significantly improve the accuracy and efficiency of sunflower data processing. Furthermore, the execution of segmentation in an optimal manner is instrumental in enhancing the precision of yield forecasts and optimizing agronomic measures.

Image segmentation necessitates the implementation of diverse approaches and techniques, encompassing both conventional algorithms and sophisticated neural networks. In order to facilitate a comprehensive understanding of the subject matter and to present the diversity of methods, the section has been divided into several subparts. This will facilitate a more detailed examination of the architectures and approaches employed, including the well-known You Only Look Once (YOLO) neural network architecture and other methodologies employed for vegetation segmentation and lesion detection.

6.1. Utilization of YOLO for the Purpose of Segmentation

YOLO is a popular model for the classification, detection, and segmentation of objects in images. The algorithm's working principle, based on a convolutional architecture, can be outlined as follows: The input image is segmented into a grid of cells. Each cell is responsible for analyzing the object within its own grid cell. If the object's center is found within a cell, that cell is responsible for detecting it. Once a cell has found the center of the object, a bounding rectangle is constructed. YOLO performs a regression of the bounding

rectangle, predicting its parameters. The intersection over union (IOU) metric measures the degree to which the predicted bounding rectangle matches the real one, thus helping to eliminate erroneous predictions.

The network divides the image into cells, with each cell predicting the bounding boxes and confidence level. The cells then determine the probability of an object belonging to a particular class. In instances where multiple model cells are employed to predict an object, the result is often a multitude of repetitive and redundant prediction frames. In order to address this issue, YOLO employs a Non-Maximum Suppression (NMS) algorithm. The fundamental principle of the method is that the box with the highest probability score is selected from a group of overlapping bounding boxes. All other boxes with a high degree of overlap with the selected box, but with lower scores, are suppressed and excluded from further processing. This process is repeated iteratively, resulting in the retention of only the final boundary frames that represent the most accurate predictions. This significantly improves the quality of object recognition. (Figure 11). The accuracy of YOLO models is evaluated by means of the mean average precision (mAP) metric.

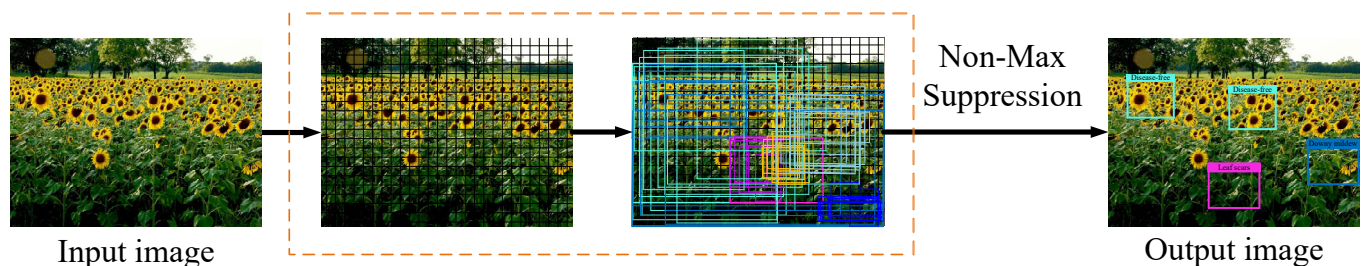


Figure 11. General scheme of how the YOLO architecture works.

Each new YOLO model introduces certain innovations that improve the accuracy and efficiency of the model. YOLO_v12 is the latest model released by Ultralytics on 18 February 2025. The YOLO_v9, YOLO_v10, and YOLO_v11 models were also released in 2024. The latest YOLO models are capable of detection, classification, segmentation, human pose estimation, object tracking, and the ability to orient bounding boxes to detect the presence of an object without its exact location [67]. Wide applications of the algorithm have also been found in precision agriculture. Different versions of YOLO are used to detect weeds or detect diseases, which, together with the use of UAVs, allows for large-scale crop monitoring. For example, the authors [17] used a drone to collect over one and a half thousand sunflower images to detect and classify sunflower capitula. The network divides the image into cells, each of which predicts the bounding boxes and confidence level, and then determines the probability that an object belongs to a particular class. The precision, recall, and mAP@0.5 achieved were 92.30%, 89.70%, and 93.00%, respectively, outperforming the original YOLO model. The work in [31] was based on the older YOLO_v5 model, extended with an attention pyramid module. The resulting system processes non-obvious feature information inferred from the first module of the model and highlights affected plant patches in the image in the second module of the model on three datasets: Katri-Twelve, BARI-Sunflower [38], and FGVC8, with an average F1 score of 95.27%. The same architecture was used in another study [32] where in combination with the Dwarf Mongoose Optimization Algorithm, 95.87% accuracy was achieved on the dataset [38]. The authors [53] collected their dataset of photos of weeds in fields using a UGV robot, segmented them using U-Net, and further labeled them. YOLO_v7 and YOLO_v8 were used for the detection task on sunflower crops, which showed a mAP@0.5 score of 0.971 in detecting sunflower crops, which is higher than Efficientnet, Retinanet and

Detection Transformer. The highest recognition score for the two classes was 0.770, leading to the conclusion that only the labeled dataset is novel.

The YOLO segmentation method has proven to be one of the most effective for image processing tasks, including sunflower crop analysis. YOLO allows the simultaneous detection and segmentation of objects in an image, making it particularly useful for the rapid and accurate detection of lesions and diseases on plants and enabling the real-time analysis of large amounts of data, which is essential for monitoring in agricultural environments. The main challenge is accuracy when processing small objects or complex backgrounds where objects can be difficult to distinguish. In addition, YOLO models require significant computational resources, which can limit their use on devices with limited capabilities. Despite these drawbacks, YOLO remains one of the most popular tools for agricultural segmentation.

6.2. Other Segmentation Methods

Different versions of YOLO are used for weed detection or disease detection, which, together with the use of UAVs, enables large-scale crop monitoring. For example, the authors [17] used a drone to collect more than one and a half thousand sunflower images to detect and classify sunflower capitula. YOLO_v7-tiny was taken as the main machine learning model, which was further refined by changing the activation function to SiLU and improving the feature map for the capitula detection task, as well as introducing deformable convolution and the SimAM attention mechanism. The achieved precision, recall and mAP@0.5 were 92.30%, 89.70%, and 93.00%, respectively, outperforming the original YOLO model.

The joint application of FPN with ResNet18 showed an accuracy higher than the U-Net architecture by an average of 1.05 percentage points. Despite their simplicity and efficiency, these methods may be limited in the case of complex images with high noise or a large number of classes. They require considerable effort to optimize the parameters and may not always handle subtle differences between objects. However, their use in combination with more sophisticated neural network architectures can provide good results for analyzing data with less computational effort.

Image segmentation is a key step in data processing for crop monitoring, including sunflowers. Segmentation methods have greatly improved the accuracy of disease diagnosis and crop health assessment. However, despite the successes, there are several challenges such as high computational requirements, the need for accurate parameter tuning and the sensitivity of the methods to data quality. The need for large amounts of labeled data for model training is also an important limitation. Prospects for segmentation lie in the further development of more adaptive and efficient methods, such as improving neural network models to deal with small objects or complex backgrounds. One of the most important directions is the automatic annotation of large image sets using large language models, which will significantly speed up the process of generating training data and increase the availability of high-quality labeled sets for research.

The analysis of different segmentation methods, their comparative evaluation and the identification of the best approaches serve as a basis for the selection of the most effective solutions in agricultural practice. Systematic comparison allows us to summarize the results of different studies, to highlight promising directions and to assess their practical applicability in real agricultural conditions.

7. Systematic Comparison and Synthesis of Results

Due to the significant heterogeneity of the reviewed studies, which address different tasks (e.g., classification, segmentation and yield prediction), use different datasets and

report incompatible metrics (e.g., accuracy, mAP and R^2), a formal quantitative meta-analysis with pooled effect sizes was not feasible. Instead, this section presents a structured qualitative synthesis and systematic comparison of the results. The results were stratified by main task and algorithm family to identify common trends, performance benchmarks and the most promising approaches described in the literature.

To ensure a structured and transparent summary of the results of the 47 selected studies, the following procedure was adopted. First, key performance indicators (e.g., accuracy and F1 score) and important study characteristics (e.g., algorithm used, data type and task, e.g., classification, segmentation or yield prediction) were systematically extracted from each article and compiled into a comparative table (Appendix A). Next, the studies were stratified by task. Due to the fundamental heterogeneity of the tasks and evaluation metrics, it was not possible to aggregate the results statistically. Instead, the studies were grouped by major task, such as disease classification and recognition, object detection, image segmentation, yield prediction and crop mapping. This allowed for meaningful comparisons within groups. Within each task group, the most relevant and frequently mentioned metrics were used to compare algorithms. For classification, the focus was on accuracy and the F1 score. For detection and segmentation, the focus was on the mAP@0.5 and the IoU. For yield prediction, the focus was on the R^2 and RMSE. This qualitative comparison across specific tasks revealed trends, the most effective models and common challenges. The synthesis focused on identifying high-level trends, such as the dominance of certain architectures, the influence of data sources, and how methods have evolved over the five-year review period.

The following sections present the results of this systematic comparison, stratified by algorithm and task. For instance, the RF method is widely used for classification in precision agriculture and has been applied to analyze data from the Sentinel-2 platform in combination with vegetation indices. In one study, the use of RF resulted in a sunflower crop classification accuracy of over 98%. This approach can effectively categorize the different growth stages of sunflowers and determine plant health, including stress conditions and illnesses. Vegetation indices, such as the NDVI, greatly enhance classification outcomes and in the detection of anomalies in the field.

To categorize the stages of sunflower development, some studies have employed SVM and LR algorithms trained on a variety of vegetable indices. When tested on separate plots, the model's accuracy reached 98%. The highest accuracy values were obtained with index time series taken from Sentinel-2. This illustrates how these methods can be used to evaluate crop phenological stages in actual field settings. The use of U-Net and FPN models for semantic image segmentation with the goal of differentiating sunflowers at various growth stages is also highlighted in the review.

Converged architectures continue to be one of the most effective tools for image processing in agriculture. CNNs have been used in studies to classify diseases in sunflowers with extremely high accuracy. On the BARI-Sun dataset, for instance, employing CNNs with pre-trained models (VGG19, Res-Net152) produced a detection accuracy of up to 98.70%. Classifying data from the BARI-Sunflower dataset using the EfficientNet_V2B2 architecture yielded an accuracy of 96.83%, one of the best results in the field. In contrast to other architectures like Inception or Xception, this model offered superior results at a lower computational cost. CNN and SVM hybrid models demonstrated a high 98.00% classification accuracy for sunflower diseases like downy mildew and Alternaria leaf blight.

YOLO models were used for the segmentation task of identifying damaged plant patches. For example, YOLO_v4 combined with k-means clustering showed an accuracy of 96.46% in detecting diseases such as powdery mildew and Verticillium wilt. In another study, using an older YOLO_v5 model with pyramidal attention, an average F1 score

of 95.87% was achieved on three datasets including BARI sunflower. The YOLO_v5 architecture combined with pyramidal compressed attention and multiple convolutional layers achieved a higher accuracy of 98.25% compared to the other works reviewed.

Figure 12 presents a graph illustrating the distribution of the different methods over time, showing the average accuracy (precision) of these methods from one year to the next. The methods using CNNs represent the most popular and accurate solutions, with more than 96%, and lead in terms of prevalence in the studies. Early detection techniques, such as the use of hyperspectral imagery and machine learning, help to detect diseases at an early stage, which significantly reduces disease damage. RF also performs well, with an accuracy of around 92–96%. They are consistently used for sunflower disease detection, although much less frequently than CNNs. Federated learning methods show significant growth in 2024, indicating a growing interest in these methods, despite their lower accuracy at the moment. YOLO is also used for detection tasks and, despite the relatively high accuracy of the method (over 96%), remains less popular than the other algorithms. The graph shows that in recent years there has been an increase in the popularity of deep learning methods such as CNN and FL, and some increase in interest in YOLO methods, reflecting changing preferences in machine learning for analyzing sunflower data. Comparative table of existing sunflower data processing methods presented in the systematic comparison is presented in Appendix A. Please note that the reported accuracy values are derived from diverse studies using different datasets, problem formulations, and are often achieved by hybrid or ensemble methods. Therefore, the values are intended to illustrate general performance trends rather than for direct numerical comparison between algorithms.

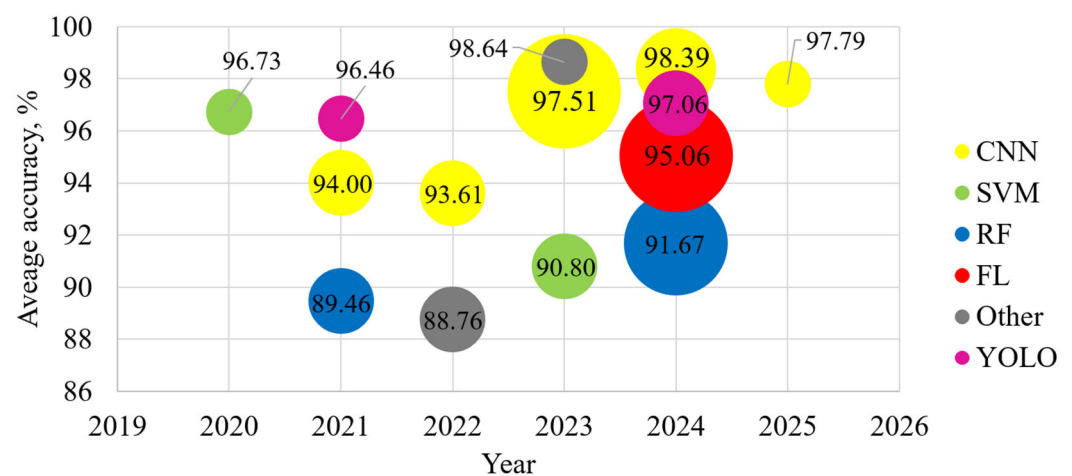


Figure 12. Average accuracy of different methods developed in the papers reviewed.

One of the key factors affecting prediction accuracy is the quality of the input data. Many studies show that high spatial and spectral resolution of satellite imagery improves classification results. Using data with a resolution of 10 m per pixel allows detailed analysis of crops, while data with a resolution of 30 m gives less accurate results. In addition, combining different data sources, such as satellite imagery and drone data, can reduce prediction errors. Transfer learning is actively used to improve accuracy with a limited dataset, significantly reducing the amount of training data required and speeding up the model training process.

This systematic comparison demonstrates that machine learning and satellite monitoring methods have high efficiency in agricultural tasks. In addition to increasing classification and prediction accuracy, the use of contemporary algorithms lowers the cost of agricultural field monitoring. More precise and dependable systems for yield forecasting

and plant disease detection will be developed in the future as a result of the advancement of machine learning techniques and their integration with other data sources.

8. Problems and Future Directions

Scaling and transferring developed models to actual agrotechnological processes is hindered by a number of issues and limitations in current research, despite notable advancements in the application of machine learning and satellite data processing methods to sunflower crop monitoring. The small and inconsistent datasets are one of the primary drawbacks of these studies, particularly when it comes to sunflowers. Most studies use local or proprietary datasets collected under specific geographical or climatic conditions, making it difficult to generalize the resulting models. In addition, information on plant phenological stages, agronomic practices, or soil types is often missing, reducing the applicability of models in other regions. Machine learning models, especially deep neural networks, are sensitive to changes in external conditions: illumination, viewing angle, and atmospheric conditions. This systematic comparison demonstrates that machine learning and satellite monitoring methods have high efficiency in agricultural tasks. The application of modern algorithms allows not only to improve the accuracy of classification and prediction, but also to reduce the cost of monitoring agricultural fields. In the future, the development of machine learning methods and their integration with other data sources will contribute to the creation of more accurate and reliable systems for yield forecasting and plant disease detection. This implies a limitation in the use of models trained in one agro-ecological zone and applied in another. Many models are trained on specific sunflower varieties, local conditions, and farming techniques. Depending on the region, downy mildew may be accompanied by a different disease or pest or have slightly different symptoms. Therefore, when the model is transferred to a different zone, the accuracy is drastically reduced. This highlights the need to develop more versatile, adaptive solutions that can be pre-trained or adapted without having to be completely re-trained from scratch.

Another problem with most research in this area is that researchers study disease symptoms that have already occurred or estimate yields after key phenological phases have occurred. Early forecasting tools, i.e., models that analyze weather data, soil types, vegetative field indices, and information from the previous season, will be important for agronomy. This will allow yield forecasts to be made before flowering and the possibility of disease or pest outbreaks to be known in advance. Such a tool for analyzing an agricultural object can be developed by combining data with different spatial and temporal resolutions, structures and presentation formats. To date, most models work with one type of data, either RGB images, vegetation indices, or radar images. However, integrating data from different platforms (Sentinel-1, Sentinel-2, hyperspectral, UAV, and weather stations) into multimodal models can significantly improve the accuracy and robustness of intelligent systems.

Insufficient accuracy of models used for sunflower crop monitoring can have several serious consequences, both at the level of scientific development and in applied agricultural practice. Low prediction accuracy can lead to false positive or false negative predictions, which in the first case means unnecessary expenditure on fungicide treatment of a healthy field area, and in the second case—lost time to control the pathogen or pest and consequently an increase in the area affected. An error in yield prediction can affect seed and fertilizer purchases and harvest logistics. Underforecasting will lead to underutilization of capacity, and overforecasting will lead to a shortage of resources at harvest. With the gradual adoption of technology in the agronomic sector, the loss of farmer confidence due to inaccurate forecasts may slow the progress of precision farming in general. If the problem of low model accuracy is overfitting, such a tool will not adapt well to other regions or

seasons. Lack of accuracy in the tasks of early disease detection or phenological phase estimation may lead to late application of agronomic and phytosanitary measures such as fertilization, irrigation and protection, thereby reducing potential yields. All of this results in high financial costs for fertilizers and fungicides that are applied to an unspecified, possibly healthy, area that is not in the correct vegetative phase, thereby reducing potential yield. For researchers worldwide, low precision and limited model validation will make it difficult to reproduce and verify results, thereby reducing the value of such publications to the community.

The prospects for further research include the development of more interpretable models that can explain their predictions. Integrating different types of data, such as weather station data, soil analysis and satellite imagery, as well as multispectral and hyperspectral data, is also an important direction. This will allow for more comprehensive models that take into account all aspects of crop growth and development. Such models will be able to capture any differences in spectral characteristics and will greatly improve the accuracy and sensitivity of multi-class disease classification models. Time series analysis will allow vegetation dynamics to be monitored and deviations from typical growth trajectories to be identified. The use of temporal patterns rather than single images increases the robustness of the models to noise and allows the detection of complex dependencies that are not visible when analyzing single images.

New approaches such as transformer networks and self-supervised learning allow us to work with large amounts of unlabeled data and train models with representations that are useful for subsequent tasks. This is particularly important given the limited number of labeled examples for sunflower. Such methods also allow the development of more versatile models that are robust to differences in sensor type, season, and growing conditions. For practical applications, it is important for agronomists, it is important not only to get a prediction, but also to understand the basis of the model's solution. Explainable artificial intelligence methods help to identify the most important attributes, track down errors, and increase confidence in decision support systems in the agricultural sector. It is also promising to develop a universal, generalized model that does not lose accuracy when applied to a different region under different conditions or that does not require extensive retraining.

Despite the high accuracy of modern machine learning models, several critical issues remain unresolved. One key challenge is the limited transferability of trained models across different geographical regions, especially ones with distinct soil types and climatic conditions. This raises the question of how well these models generalize beyond the datasets on which they were trained. Another important consideration is the trade-off between interpretability and performance. Can models be made more transparent without sacrificing accuracy? Additionally, there is still no consensus on the minimum volume and quality of input data required to ensure robust model performance under real-world field conditions.

Future research should focus on addressing these questions, particularly through developing interpretable model architectures and validating them under actual agroecological scenarios. In practical applications, especially in resource-constrained agricultural settings, models must demonstrate high accuracy in controlled environments and maintain stability in the presence of noise, incomplete input, or cloud-covered satellite data. This is essential for integrating them into decision-support systems used by agronomists and farmers, who often lack access to advanced AI infrastructure.

This review confirms that combining machine learning techniques with satellite and UAV monitoring creates a powerful toolkit for agricultural applications. These technologies enhance classification accuracy and forecasting precision while reducing operational

costs associated with crop monitoring. Moving forward, improvements in algorithmic efficiency and integration of heterogeneous data sources, such as spectral, temporal, and climatic data, will enable the development of robust systems for early disease detection and yield prediction. These advancements could significantly impact sunflower production, encouraging broader adoption among farmers and contributing to national agricultural productivity.

Furthermore, while this review concentrated on well-established classification, segmentation and yield prediction methods, several significant and emerging areas of research were beyond its immediate scope. Future research in the field of intelligent sunflower monitoring should prioritize studying these new areas to improve the reliability and applicability of models. One key area is addressing domain shifts and interregional issues to develop models that can be generalized for different geographical regions and growing conditions. Other promising areas include using multitemporal data to gain a dynamic understanding of phenology and to enable the early detection of stress, which goes beyond analyzing a single point in time. Another promising area is studying data fusion methods, particularly between SAR and optical/hyperspectral data, in order to overcome limitations caused by cloud cover and to enrich feature sets. Additionally, applying advanced learning paradigms such as transfer learning and semi-supervised and self-supervised learning will reduce dependence on large manually labeled datasets and enhance agri-data processing methods for more accurate predictions and diagnostics. Methods can be made more applicable by using modern architectures, such as spatio-temporal transformers and graph neural networks, to model complex relationships in the field more accurately. From a practical standpoint, integrating explainable artificial intelligence (XAI) methods will build trust and provide agronomists and farmers with practical recommendations. Solving these challenges is critical to creating the next generation of scalable, reliable and user-centric artificial intelligence solutions for precision agriculture.

9. Conclusions

This systematic review summarizes and critically evaluates recent advances in the application of machine learning and remote sensing techniques to sunflower crop monitoring, with a focus on yield prediction, disease diagnosis, and crop health assessment. The analysis covers a wide range of algorithms, including traditional classifiers (e.g., SVM, RF), deep learning methods (e.g., CNN, U-Net), and data preprocessing techniques such as vegetation index calculation and principal component-based feature selection. The reviewed studies used a variety of data sources, ranging from multispectral drone imagery to Sentinel-2 satellite data and ground-based observations, demonstrating the growing trend towards the integration of multiscale information in precision agriculture tasks. A review study showed that the models exhibit moderately high classification accuracy under supervised conditions: the accuracy value reaches 98.80% for multi-class disease diagnosis and 98.00% for crop mapping tasks under favorable conditions. However, when the models are transferred to new regions or seasons, a decrease in accuracy is observed, indicating the limited generalizability of existing approaches.

Despite the progress made, the review identified several methodological limitations: small sample sizes, limited validation on heterogeneous data, and lack of consistent protocols for assessing model quality. Only a fraction of the reviewed studies provides an extended set of metrics (such as F1 score or accuracy), making it difficult to compare results objectively. In addition, most papers do not address the issues of model interpretability and practical implementation scenarios, which are critical for agricultural users. Therefore, there is a need to establish and implement standardized requirements and criteria for the development and further evaluation of artificial intelligence algorithms for solving agricul-

tural problems. In addition, to address the identified shortcomings, future research should pay attention to generating large and diverse datasets, including spatial and temporal diversity, using rigorous validation schemes (e.g., cross-validation with data from other regions), and providing comprehensive reporting on model quality. The data obtained from the review can serve as the foundation for developing decision support systems for the agricultural sector. These systems would integrate satellite and UAV monitoring with machine learning algorithms. Given the challenges of global food security, creating scalable, adaptive, and interpretable AI systems suitable for practical application at all stages of the sunflower production cycle is an urgent task.

Promising directions for further development of artificial intelligence algorithms in sunflower crop management are integration of auxiliary information sources such as meteorological data or soil characteristics, use of temporal sequences to model crop phenological development, and introduction of interpretable, transformer, and tensor-based deep learning architectures. This review demonstrates the high potential of intelligent data processing for sustainable sunflower production and identifies key scientific and practical challenges that need to be addressed to achieve practical impact. Answers to the questions posed by the review study will lead to more robust, scalable, and interpretable solutions in the agricultural domain.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/make7040130/s1>. PRISMA 2020 checklist is supplemented.

Author Contributions: Conceptualization, V.A.; methodology, P.L. and U.L.; software, V.A.; validation, V.A., V.R. and P.L.; formal analysis, V.A. and D.B.; investigation, V.A. and R.A.; resources, R.A., V.R. and D.B.; data curation, P.L. and V.R.; writing—original draft preparation, V.A.; writing—review and editing, P.L., D.B. and U.L.; visualization, V.A.; supervision, P.L.; project administration, P.L.; funding acquisition, D.B. All authors have read and agreed to the published version of the manuscript.

Funding: The authors express their gratitude to North-Caucasus Federal University for the support of small scientific groups and individual scientists. Research was conducted with the support of the Russian Science Foundation (project No. 23-71-10013).

Data Availability Statement: No new data were created.

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

Abbreviations

The following abbreviations are used in this manuscript:

UAV	Unmanned Aerial Vehicle
UN	United Nations
FAO	Food and Agriculture Organization
SVM	Support Vector Machine
RF	Random Forest Algorithm
CNN	Convolutional Neural Network
ViT	Visual Transformer
MP	Multilayer Perceptron
FL	Federated Learning
KNN	K-Nearest Neighbour Algorithm
FPN	Feature Pyramid Network
NDVI	Normalized Difference Vegetation Index
EVI	Enhanced Vegetation Index
SAVI	Soil-Adjusted Vegetation Index
NDMI	Normalized Difference Moisture Index

NDRE	Normalized Difference Red Edge Index
IoU	Intersection over Union
mAP@0.5	Mean Average Precision at IoU threshold 0.5
NIR	Near-Infrared Band
R ²	Regression Coefficient
RMSE	Root Mean Square Error
SAR	Synthetic Aperture Radar

Appendix A

Table A1. Comparative table of existing sunflower data processing methods presented in the systematic comparison.

Work	Base Method	Year	Accuracy	F1	Precision	Recall	R ²	RMSE	IoU	mAP@0.5	Dataset	Aim
[12]	CNN	2023	93.00	0.93	0.93	0.93	-	-	-	-	BARI	recognition diseases
[13]	SVM, CNN	2020	96.73	-	-	-	-	-	-	-	Custom (UAV)	mapping
[65]	LR	2022	89.00	0.89	0.91	0.87	-	-	-	-	Custom (satellite)	recognition diseases
[17]	YOLO	2024	-	-	0.923	-	-	-	-	0.93	Custom (UAV)	detection
[14]	CNN, LR	2022	-	-	-	-	0.914	1699	-	-	Custom (sensor)	prediction of yield
[62]	KNN	2022	88.52	-	-	-	-	-	-	-	Custom	detection
[54]	CNN	2022	98.02	-	-	-	-	-	-	-	Custom	recognition diseases
[15]	RF	2024	98.00	0.94	-	-	-	-	-	-	Custom (satellite)	mapping
[43]	RF	2024	0.96	-	-	-	-	-	-	-	Custom (satellite)	mapping
[29]	CNN	2023	1.00	-	-	-	-	-	-	-	BARI	recognition diseases
[16]	CNN	2023	97.89	-	-	-	-	-	-	-	Custom (satellite)	crop classification
[40]	K-Means	2023	-	-	-	-	0.73	43.40	-	-	Custom (satellite)	mapping
[30]	CNN	2023	97.60	0.976	0.976	0.976	-	-	-	-	BARI	recognition diseases
[31]	YOLO	2024	98.25	0.9803	0.9859	0.9749	-	-	-	-	BARI	recognition diseases
[39]	CNN	2024	1.00	-	-	-	-	-	-	-	[18]	segmentation
[18]	cGAN	2021	-	-	-	-	-	-	-	-	[18]	segmentation
[36]	CNN	2025	96.00	0.9600	0.9600	0.9600	-	-	-	-	BARI	recognition diseases
[41]	RF, MLP	2024		0.9060	0.9505	0.8787	-	-	-	-	Custom (satellite)	mapping
[44]	RFR	2023	-	-	-	-	0.84	284.5	-	-	Custom (satellite)	prediction of yield
[45]	RFR, LR, SVM	2022	-	-	-	-	0.99	0.366	-	-	Custom (satellite)	prediction of yield
[48]	U-Net, RF	2021	88.23	0.9108	-	-	-	-	0.8362	-	Custom (UAV)	detection
[32]	YOLO	2024	95.87	94.64	-	-	-	-	-	-	BARI	recognition diseases
[37]	CNN, ViT	2024	97.02	-	-	-	-	-	-	-	BARI	recognition diseases

Table A1. Cont.

Work	Base Method	Year	Accuracy	F1	Precision	Recall	R ²	RMSE	IoU	mAP@0.5	Dataset	Aim
[63]	YOLO	2021	96.46	0.84	0.84	0.85	-	-	-	-	Custom	recognition diseases
[49]	YOLO	2022	-	-	0.6210	0.9760	-	-	-	-	Custom (UAV)	detection
[53]	YOLO	2023	-	-	-	-	-	-	-	0.971	Custom	detection
[46]	time-series analysis	2024	-	-	-	-	-	-	-	-	Custom (satellite)	yield monitoring
[19]	FPN, CNN	2023	98.64	-	0.9996	0.9852	-	-	0.9271	-	[18]	segmentation
[55]	CNN + SVM	2023	98.00	0.9302	-	0.9279	-	-	-	-	Custom	recognition diseases
[6]	FL, CNN	2024	94.87	0.9484	0.9480	0.9496	-	-	-	-	Custom	recognition diseases
[47]	Cumulative sum	2024	-	-	-	-	-	-	-	-	-	yield monitoring
[50]	FL, CNN	2024	98.00	0.9614	0.9598	0.9630	-	-	-	-	Custom	recognition diseases
[51]	FL, CNN	2024	95.72	0.9571	0.9571	0.9572	-	-	-	-	Custom	recognition diseases
[23]	CNN	2024	98.16	0.9851	0.9862	0.9847	-	-	-	-	Custom	recognition diseases
[52]	RF, CNN	2024	81.54	0.8903	0.9079	0.8947	-	-	-	-	Custom	recognition diseases
[60]	RF	2021	90.68	-	0.7559	-	-	-	-	-	Custom	recognition diseases
[56]	CNN	2021	89.20	-	-	-	-	-	-	-	Custom	recognition diseases
[33]	CNN	2023	97.88	0.95	0.91	0.91	-	-	-	-	BARI	recognition diseases
[34]	CNN	2023	98.70	0.99	0.99	0.99	-	-	-	-	BARI	recognition diseases
[61]	FL + CNN	2024	91.86	0.9183	0.9183	0.9188	-	-	-	-	Custom	recognition diseases
[58]	CNN	2021	98.80	0.8145	0.7941	0.7941	-	-	-	-	Custom	recognition diseases
[42]	time-series analysis	2024	-	-	-	-	-	0.4075	-	-	Custom (UAV)	recognition diseases
[35]	CNN, SVM	2023	83.59	0.8345	-	-	-	-	-	-	Custom	recognition diseases
[68]	CNN, RF	2024	92.21	0.9220	0.8411	0.9220	-	-	-	-	Custom	prediction of yield
[57]	CNN	2022	89.20	-	-	-	-	-	-	-	Custom	recognition diseases

References

1. Population | United Nations. Available online: <https://www.un.org/en/global-issues/population> (accessed on 18 January 2024).
2. Hunger Numbers Stubbornly High for Three Consecutive Years as Global Crises Deepen: UN Report. Available online: <https://www.who.int/news/item/24-07-2024-hunger-numbers-stubbornly-high-for-three-consecutive-years-as-global-crises-deepen-un-report> (accessed on 3 February 2025).
3. FAO. *The State of Food Security and Nutrition in the World*; Food and Agriculture Organization of the United Nations: Rome, Italy, 2023.
4. FAO. *The Plants That Feed the World: Baseline Data and Metrics to Inform Strategies for the Conservation and Use of Plant Genetic Resources for Food and Agriculture*; Food and Agriculture Organization of the United Nations: Rome, Italy, 2022.
5. Adeleke, B.S.; Babalola, O.O. Oilseed Crop Sunflower (*Helianthus annuus*) as a Source of Food: Nutritional and Health Benefits. *Food Sci. Nutr.* **2020**, *8*, 4666–4684. [\[CrossRef\]](#)
6. Singh, H.; Kukreja, V.; Mehta, S.; Gupta, S.; Aeri, M. AI in Agriculture: Federated Learning CNNs for Sunflower Leaf Disease Diagnosis. In Proceedings of the 2024 5th International Conference for Emerging Technology, INCET, Belgaum, India, 24–26 May 2024. [\[CrossRef\]](#)
7. Agrawal, M.; Agrawal, S. Rice Plant Diseases Detection Using Convolutional Neural Networks. *Int. J. Eng. Syst. Model. Simul.* **2023**, *14*, 30–42. [\[CrossRef\]](#)

8. Umapathi, R.; Ghoreishian, S.M.; Sonwal, S.; Rani, G.M.; Huh, Y.S. Portable Electrochemical Sensing Methodologies for On-Site Detection of Pesticide Residues in Fruits and Vegetables. *Coord. Chem. Rev.* **2022**, *453*, 214305. [[CrossRef](#)]
9. Ali, Z.A.; Zain, M.; Pathan, M.S.; Mooney, P. Contributions of Artificial Intelligence for Circular Economy Transition Leading toward Sustainability: An Explorative Study in Agriculture and Food Industries of Pakistan. *Environ. Dev. Sustain. A Multidiscip. Approach Theory Pract. Sustain. Dev.* **2024**, *26*, 19131–19175. [[CrossRef](#)]
10. Centorame, L.; Ilari, A.; Del Gatto, A.; Foppa Pedretti, E. A Systematic Review on Precision Agriculture Applied to Sunflowers, the Role of Hyperspectral Imaging. *Comput. Electron. Agric.* **2024**, *222*, 109097. [[CrossRef](#)]
11. Sehgal, S.; Roy, A. A Study of Deep Learning Techniques on Oilseed Crops. In Proceedings of the 3rd IEEE 2022 International Conference on Computing, Communication, and Intelligent Systems, ICCIS, Greater Noida, India, 4–5 November 2022; pp. 526–532. [[CrossRef](#)]
12. Ghosh, P.; Mondal, A.K.; Chatterjee, S.; Masud, M.; Meshref, H.; Bairagi, A.K. Recognition of Sunflower Diseases Using Hybrid Deep Learning and Its Explainability with AI. *Mathematics* **2023**, *11*, 2241. [[CrossRef](#)]
13. Song, Z.; Zhang, Z.; Yang, S.; Ding, D.; Ning, J. Identifying Sunflower Lodging Based on Image Fusion and Deep Semantic Segmentation with UAV Remote Sensing Imaging. *Comput. Electron. Agric.* **2020**, *179*, 105812. [[CrossRef](#)]
14. Khalifani, S.; Darvishzadeh, R.; Azad, N.; Seyed Rahmani, R. Prediction of Sunflower Grain Yield under Normal and Salinity Stress by RBF, MLP and, CNN Models. *Ind. Crop. Prod.* **2022**, *189*, 115762. [[CrossRef](#)]
15. Qadir, A.; Skakun, S.; Kussul, N.; Shelestov, A.; Becker-Reshef, I. A Generalized Model for Mapping Sunflower Areas Using Sentinel-1 SAR Data. *Remote. Sens. Environ.* **2024**, *306*, 114132. [[CrossRef](#)]
16. Farmonov, N.; Amankulova, K.; Szatmari, J.; Sharifi, A.; Abbasi-Moghadam, D.; Mirhoseini Nejad, S.M.; Mucsi, L. Crop Type Classification by DESIS Hyperspectral Imagery and Machine Learning Algorithms. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2023**, *16*, 1576–1588. [[CrossRef](#)]
17. Jing, R.; Niu, Q.; Tian, Y.; Zhang, H.; Zhao, Q.; Li, Z.; Zhou, X.; Li, D. Sunflower-YOLO: Detection of Sunflower Capitula in UAV Remote Sensing Images. *Eur. J. Agron.* **2024**, *160*, 127332. [[CrossRef](#)]
18. Fawakherji, M.; Potena, C.; Pretto, A.; Bloisi, D.D.; Nardi, D. Multi-Spectral Image Synthesis for Crop/Weed Segmentation in Precision Farming. *Robot. Auton. Syst.* **2021**, *146*, 103861. [[CrossRef](#)]
19. Poleshchenko, D.; Mikhailov, I.; Petrov, V. On the Segmentation of Sunflower Plants in UAV Photos. In Proceedings of the 2023 25th International Conference on Digital Signal Processing and its Applications, DSPA, Moscow, Russia, 29–31 March 2023. [[CrossRef](#)]
20. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* **2021**, *372*, n71. [[CrossRef](#)]
21. Kitchenham, B.; Charters, S.M. *Guidelines for Performing Systematic Literature Reviews in Software Engineering, Version 2.3*; Keele University: Staffs, UK, 2007.
22. Gascuel, Q.; Martinez, Y.; Boniface, M.C.; Vear, F.; Pichon, M.; Godiard, L. The Sunflower Downy Mildew Pathogen *Plasmopara halstedii*. *Mol. Plant Pathol.* **2015**, *16*, 109–122. [[CrossRef](#)] [[PubMed](#)]
23. Banerjee, D.; Sharma, N.; Upadhyay, D.; Singh, M.; Chythanya, K.R. Decoding Sunflower Downy Mildew: Leveraging Hybrid Deep Learning for Scale Severity Analysis. In Proceedings of the 2024 5th International Conference for Emerging Technology, INCET, Belgaum, India, 24–26 May 2024. [[CrossRef](#)]
24. Baka, Z.A.M. Light and Electron Microscopy of the Micromorphology and Development of Pycniospores and Aeciospores of the Sunflower Rust, *Puccinia helianthi*. *Micron* **2025**, *189*, 103733. [[CrossRef](#)] [[PubMed](#)]
25. Shree, D.; Manjula, C.P.; Prasanna Kumar, M.K.; Palanna, K.B.; Harish, J.; Kumar, P.V.D.; Khan, F. Population Frequency Distribution and Introgression of Alternaria Species Causing Leaf Blight of Sunflower, India. *Physiol. Mol. Plant Pathol.* **2025**, *136*, 102527. [[CrossRef](#)]
26. Montecchia, J.F.; Fass, M.I.; Cerrudo, I.; Quiroz, F.J.; Nicosia, S.; Maringolo, C.A.; Di Rienzo, J.; Troglia, C.; Hopp, H.E.; Escande, A.; et al. On-Field Phenotypic Evaluation of Sunflower Populations for Broad-Spectrum Resistance to Verticillium Leaf Mottle and Wilt. *Sci. Rep.* **2021**, *11*, 11644. [[CrossRef](#)]
27. Kandel, H.; Endres, G.; Buetow, R. *Sunflower Production Guide*; North Dakota Agricultural Experiment Station and North Dakota State University Extension: Fargo, ND, USA, 2020.
28. Hughes, D.; Salathe, M. An Open Access Repository of Images on Plant Health to Enable the Development of Mobile Disease Diagnostics. *arXiv* **2015**, arXiv:1511.08060.
29. Asha Rani, K.P.; Gowrishankar, S. Pathogen-Based Classification of Plant Diseases: A Deep Transfer Learning Approach for Intelligent Support Systems. *IEEE Access* **2023**, *11*, 64476–64493. [[CrossRef](#)]
30. Gulzar, Y.; Ünal, Z.; Aktaş, H.; Mir, M.S. Harnessing the Power of Transfer Learning in Sunflower Disease Detection: A Comparative Study. *Agriculture* **2023**, *13*, 1479. [[CrossRef](#)]

31. Dai, G.; Tian, Z.; Fan, J.; Sunil, C.K.; Dewi, C. DFN-PSAN: Multi-Level Deep Information Feature Fusion Extraction Network for Interpretable Plant Disease Classification. *Comput. Electron. Agric.* **2024**, *216*, 108481. [\[CrossRef\]](#)
32. Thilagavathi, T.; Arockiam, L. Segmentation of Sunflower Leaf Disease Using Improved YOLO Network with IDMO Model. *Int. J. Intell. Syst. Appl. Eng.* **2024**, *12*, 600–611.
33. Sathi, T.A.; Hasan, M.A.; Alam, M.J. SunNet: A Deep Learning Approach to Detect Sunflower Disease. In Proceedings of the 7th International Conference on Trends in Electronics and Informatics, ICOEI, Tirunelveli, India, 11–14 April 2023; pp. 1210–1216. [\[CrossRef\]](#)
34. Soheli, A.; Alam Sarker, M.M.; Das, U.C.; Das, P.K.; Siddiquee, S.M.T.; Haider Noori, S.R.; Thien Le, N. Sunflower Disease Identification Using Deep Learning: A Data-Driven Approach. In Proceedings of the 2023 26th International Conference on Computer and Information Technology, ICCIT, Cox's Bazar, Bangladesh, 13–15 December 2023. [\[CrossRef\]](#)
35. Banerjee, D.; Kukreja, V.; Vats, S.; Jain, V.; Goyal, B. AI-Driven Sunflower Disease Multiclassification: Merging Convolutional Neural Networks and Support Vector Machines. In Proceedings of the 2023 4th International Conference on Electronics and Sustainable Communication Systems, ICESC, Coimbatore, India, 6–8 July 2023; pp. 722–726. [\[CrossRef\]](#)
36. Parek, S.; Dilshad, N.; Lee, J.W. A Channel Attention-Driven Optimized CNN for Efficient Early Detection of Plant Diseases in Resource Constrained Environment. *Agriculture* **2025**, *15*, 127. [\[CrossRef\]](#)
37. Baboshina, V.A.; Lyakhov, P.A.; Kaplun, D.I. Ensemble of Visual Transformer and Deep Neural Networks for Recognizing Sunflower Diseases from Photographs. In *Lecture Notes in Networks and Systems, Proceedings of the NIELIT's International Conference on Communication, Electronics and Digital Technology, Guwahati, India, 6–17 February 2024*; Springer: Berlin/Heidelberg, Germany, 2024; Volume 1023, pp. 15–24. [\[CrossRef\]](#)
38. Sara, U.; Rajbongshi, A.; Shakil, R.; Akter, B.; Sazzad, S.; Uddin, M.S. An Extensive Sunflower Dataset Representation for Successful Identification and Classification of Sunflower Diseases. *Data Brief* **2022**, *42*, 108043. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Peng, M.; Liu, Y.; Qadri, I.A.; Bhatti, U.A.; Ahmed, B.; Sarhan, N.M.; Awwad, E.M. Advanced Image Segmentation for Precision Agriculture Using CNN-GAT Fusion and Fuzzy C-Means Clustering. *Comput. Electron. Agric.* **2024**, *226*, 109431. [\[CrossRef\]](#)
40. Marino, S. Understanding the Spatio-Temporal Behaviour of the Sunflower Crop for Subfield Areas Delineation Using Sentinel-2 NDVI Time-Series Images in an Organic Farming System. *Heliyon* **2023**, *9*, e19507. [\[CrossRef\]](#)
41. Maleki, S.; Baghdadi, N.; Bazzi, H.; Dantas, C.F.; Ienco, D.; Nasrallah, Y.; Najem, S. Machine Learning-Based Summer Crops Mapping Using Sentinel-1 and Sentinel-2 Images. *Remote Sens.* **2024**, *16*, 4548. [\[CrossRef\]](#)
42. Athar, U.; Ali, M.; Zafar, Z.; Khurshid, H.; Berns, K.; Fraz, M.M. Merging UAV-Derived Metrics with Crop Physiology to Estimate Sunflower Yield. In Proceedings of the 2024 International Conference on Frontiers of Information Technology, FIT, Islamabad, Pakistan, 9–10 December 2024. [\[CrossRef\]](#)
43. Qadir, A.; Skakun, S.; Becker-Reshef, I.; Kussul, N.; Shelestov, A. Estimation of Sunflower Planted Areas in Ukraine during Full-Scale Russian Invasion: Insights from Sentinel-1 SAR Data. *Sci. Remote Sens.* **2024**, *10*, 100139. [\[CrossRef\]](#)
44. Amankulova, K.; Farmonov, N.; Mucsi, L. Time-Series Analysis of Sentinel-2 Satellite Images for Sunflower Yield Estimation. *Smart Agric. Technol.* **2023**, *3*, 100098. [\[CrossRef\]](#)
45. Amankulova, K.; Farmonov, N.; Mukhtorov, U.; Mucsi, L. Sunflower Crop Yield Prediction by Advanced Statistical Modeling Using Satellite-Derived Vegetation Indices and Crop Phenology. *Geocarto Int.* **2023**, *38*, 2197509. [\[CrossRef\]](#)
46. Baup, F.; Fieuzal, R.; Ygorra, B.; Gorrab, A.; Riazanof, S.; Martin-Comte, A.; Gross, K.; Frappart, F. A 6-Year Analysis of Sentinel-1, Sentinel-2 and Landsat-8 Over Sunflower Crops and an Experimental Field in Southwestern France. In Proceedings of the IGARSS 2024—2024 IEEE International Geoscience and Remote Sensing Symposium, Athens, Greece, 7–12 July 2024; pp. 10121–10125. [\[CrossRef\]](#)
47. Ygorra, B.; Baup, F.; Fieuzal, R.; Martin-Comte, A.; Gross, K.; Riazanoff, S.; Frappart, F.; Wigneron, J.-P. CuSum-Nrt as a Crop Monitoring System: A Sentinel-1 Application to Sunflower and Sorghum in Southwestern France. In Proceedings of the IGARSS 2024—2024 IEEE International Geoscience and Remote Sensing Symposium, Athens, Greece, 7–12 July 2024; pp. 10117–10120. [\[CrossRef\]](#)
48. Li, G.; Han, W.; Huang, S.; Ma, W.; Ma, Q.; Cui, X. Extraction of Sunflower Lodging Information Based on UAV Multi-Spectral Remote Sensing and Deep Learning. *Remote Sens.* **2021**, *13*, 2721. [\[CrossRef\]](#)
49. Poleshchenko, D.; Petrov, V.; Mikhailov, I. Detection of Sunflower Plants in UAV Photos. In Proceedings of the 2022 4th International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency, SUMMA, Lipetsk, Russia, 9–11 November 2022; pp. 385–388. [\[CrossRef\]](#)
50. Sharma, H.; Kukreja, V.; Mehta, S.; Nisha Chandran, S.; Garg, A. Plant AI in Agriculture: Innovative Approaches to Sunflower Leaf Disease Detection with Federated Learning CNNs. In Proceedings of the 2024 5th International Conference for Emerging Technology, INCET, Belgaum, India, 24–26 May 2024. [\[CrossRef\]](#)
51. Joshi, J.; Aeri, M.; Kukreja, V.; Mehta, S. Revolutionizing in Agriculture: Federated CNN Models for Sunflower Leaf Diseases. In Proceedings of the 2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO, Noida, India, 14–15 March 2024. [\[CrossRef\]](#)

52. Yashu; Singh, A. Sunflower Sentry: Advanced Hybrid Model for Early Disease Detection. In Proceedings of the 2024 5th International Conference on Data Intelligence and Cognitive Informatics (ICDICI), Tirunelveli, India, 18–20 November 2024; pp. 552–557. [\[CrossRef\]](#)
53. Abuhani, D.A.; Hussain, M.H.; Khan, J.; Elmohandes, M.; Zualkernan, I. Crop and Weed Detection in Sunflower and Sugarbeet Fields Using Single Shot Detectors. In Proceedings of the 2023 IEEE International Conference on Omni-Layer Intelligent Systems, COINS, Berlin, Germany, 23–25 July 2023. [\[CrossRef\]](#)
54. Dawod, R.G.; Dobre, C. ResNet Interpretation Methods Applied to the Classification of Foliar Diseases in Sunflower. *J. Agric. Food Res.* **2022**, *9*, 100323. [\[CrossRef\]](#)
55. Rana, G.; Singh, R.; Pal, A.; Gupta, R. Enhancing Sunflower Disease Identification with CNN-SVM Integration. In Proceedings of the 2023 3rd International Conference on Smart Generation Computing, Communication and Networking, SMART GENCON, Bangalore, India, 29–31 December 2023. [\[CrossRef\]](#)
56. Sirohi, A.; Malik, A. A Hybrid Model for the Classification of Sunflower Diseases Using Deep Learning. In Proceedings of the 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), London, UK, 28–30 April 2021; pp. 58–62. [\[CrossRef\]](#)
57. Malik, A.; Vaidya, G.; Jagota, V.; Eswaran, S.; Sirohi, A.; Batra, I.; Rakhra, M.; Asenso, E. Design and Evaluation of a Hybrid Technique for Detecting Sunflower Leaf Disease Using Deep Learning Approach. *J. Food Qual.* **2022**, *2022*, 9211700. [\[CrossRef\]](#)
58. Dawod, R.G.; Dobre, C. Classification of Sunflower Foliar Diseases Using Convolutional Neural Network. In Proceedings of the 2021 23rd International Conference on Control Systems and Computer Science Technologies, CSCS, Bucharest, Romania, 26–28 May 2021; pp. 476–481. [\[CrossRef\]](#)
59. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
60. Rajbongshi, A.; Biswas, A.A.; Biswas, J.; Shakil, R.; Akter, B.; Barman, M.R. Sunflower Diseases Recognition Using Computer Vision-Based Approach. In Proceedings of the IEEE Region 10 Humanitarian Technology Conference, R10-HTC, Bangalore, India, 30 September–2 October 2021. [\[CrossRef\]](#)
61. Suryavanshi, A.; Mehta, S.; Malhotra, S.; Choudhary, A.; Jain, V. Growing Insights: Federated Learning CNN's in Combatting Sunflower Leaf Diseases. In Proceedings of the 2024 Asia Pacific Conference on Innovation in Technology (APCIT), Mysore, India, 26–27 July 2024; pp. 1–6. [\[CrossRef\]](#)
62. Kaur, R.; Jain, A.; Kumar, S. Optimization Classification of Sunflower Recognition through Machine Learning. *Mater. Today Proc.* **2022**, *51*, 207–211. [\[CrossRef\]](#)
63. Chen, S.; Lv, F.; Huo, P. Improved Detection of Yolov4 Sunflower Leaf Diseases. In Proceedings of the 2021 2nd International Symposium on Computer Engineering and Intelligent Communications (ISCEIC), Nanjing, China, 6–8 August 2021; pp. 56–59. [\[CrossRef\]](#)
64. Hermawan, D.R.; Fahrio Ghanial Fatihah, M.; Kurniawati, L.; Helen, A. Comparative Study of J48 Decision Tree Classification Algorithm, Random Tree, and Random Forest on In-Vehicle Coupon Recommendation Data. In Proceedings of the 2021 International Conference on Artificial Intelligence and Big Data Analytics, Bandung, Indonesia, 27–29 October 2021; pp. 1–6. [\[CrossRef\]](#)
65. Atsmon, G.; Nehurai, O.; Kizel, F.; Eizenberg, H.; Nisim Lati, R. Hyperspectral Imaging Facilitates Early Detection of Orobancha Cumana Below-Ground Parasitism on Sunflower under Field Conditions. *Comput. Electron. Agric.* **2022**, *196*, 106881. [\[CrossRef\]](#)
66. Baboshina, V.A.; Lyakhov, P.A.; Lyakhova, U.A.; Pismennyy, V.A. Bidirectional Encoder Representation from Image Transformers for Recognizing Sunflower Diseases from Photographs. *Comput. Opt.* **2025**, *49*, 435–442. [\[CrossRef\]](#)
67. Tomassini, S.; Ali Akber Dewan, M.; Liaqat Ali, M.; Zhang, Z. The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection. *Computers* **2024**, *13*, 336. [\[CrossRef\]](#)
68. Rani, L.; Gagandeep; Veeramanickam, M.R.M.; Pandey, B. Innovative Fusion for Sunflower Leaf Disease Identification: CNN and Random Forest Strategies. In Proceedings of the 2024 IEEE AITU: Digital Generation, Conference Proceedings—AITU, Astana, Kazakhstan, 3–4 April 2024; pp. 65–70. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.