




Article

On Precision Agriculture: Enhanced Automated Fruit Disease Identification and Classification Using a New Ensemble Classification Method

Abid Mehmood ¹, Muneer Ahmad ^{2,*} and Qazi Mudassar Ilyas ^{3,*}

¹ Department of Management Information Systems, College of Business Administration, King Faisal University, Al Ahsa 31982, Saudi Arabia

² Department of Human and Digital Interface, Woosong University, Daejeon 34606, Republic of Korea

³ Department of Information Systems, College of Computer Sciences and Information Technology, King Faisal University, Al Ahsa 31982, Saudi Arabia

* Correspondence: muneer.ahmad@wsu.ac.kr (M.A.); qilyas@kfu.edu.sa (Q.M.I.)

Abstract: Fruits are considered among the most nutrient-dense cash crops around the globe. Since fruits come in different types, sizes, shapes, colors, and textures, the manual classification and disease identification of a large quantity of fruit is time-consuming and sluggish, requiring massive human intervention. We propose a multilevel fusion method for fruit disease identification and fruit classification that includes intensive fruit image pre-processing, customized image kernels for feature extraction with state-of-the-art (SOTA) deep methods, Gini-index-based controlled feature selection, and a hybrid ensemble method for identification and classification. We noticed certain limitations in the existing literature of adopting a single data source, in terms of limited data sizes, variability in fruit types, variability in quality, and variability in disease type. Therefore, we extensively aggregated and pre-processed multi-fruit data to simulate our proposed ensemble model on comprehensive datasets to cover both fruit classification and disease identification aspects. The multi-fruit imagery data contained regular and augmented images of fruits including apple, apricot, avocado, banana, cherry, fig, grape, guava, kiwi, mango, orange, peach, pear, pineapple, and strawberry. Similarly, we considered normal and augmented images of rotten fruits including beans (two categories), strawberries (seven categories), and tomatoes (three categories). For consistency, we normalized the images and designed an auto-labeling mechanism based on the existing image clusters to label inconsistent data to appropriate classes. Finally, we verified the auto-labeled data with a complete inspection to correctly assign it to the relevant classes. The proposed ensemble classifier outperforms all other classification methods, achieving 100% and 99% accuracy for fruit classification and disease identification. Further, we performed the analysis of variance (ANOVA) test to validate the statistical significance of the classifiers' outcomes at $\alpha = 0.05$. We achieved F-values of 32.41 and 11.42 against F-critical values of 2.62 and 2.86, resulting in p -values of 0.00 (<0.05) for fruit classification and disease identification.

Keywords: precision agriculture; ensemble method; fruit classification; disease identification; imagery data



Citation: Mehmood, A.; Ahmad, M.; Ilyas, Q.M. On Precision Agriculture: Enhanced Automated Fruit Disease Identification and Classification Using a New Ensemble Classification Method. *Agriculture* **2023**, *13*, 500. <https://doi.org/10.3390/agriculture13020500>

Academic Editor: Xanthoula Eirini Pantazi

Received: 15 January 2023

Revised: 6 February 2023

Accepted: 16 February 2023

Published: 20 February 2023



Copyright: © 2023 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

Food is one of the most fundamental requirements for human life and survival. Agriculture plays a vital role in fulfilling food requirements globally. The rate of growth in agricultural production has declined recently, according to the Food and Agricultural Organization (FAO) [1]. Global food security is seriously threatened by this trend and other risks like climate change, population growth, rural-to-urban migration, and biofuel demand [2]. As a result, to increase plant yield for food production, a higher yielding and more sustainable environment is required [3,4].

Given that there will hardly be an increase in the size of land available for agricultural use in the future, the best way to deal with the problem of food insecurity is by increasing the productivity of existing farmland. For this purpose, innovative methods to increase agricultural yield will be continuously required. Furthermore, since 30% of the global agricultural workforce is predicted to move to work in other sectors of the economy between 2017 and 2030 [5], the adoption of agricultural technology becomes even more critical. Therefore, technology has been adopted, modified, and optimized increasingly over the past century to increase agricultural yields [6]. Precision Agriculture (PA) is gaining popularity in this setting of sustainably increasing yields. PA is a broad term that encompasses all approaches used to improve the accuracy and control of farming management [7]. These techniques include GPS navigation of tractors, robotics, remote sensing, data analytics, and unmanned aerial and terrestrial vehicles [8]. PA relies heavily on early and precise plant disease detection.

Fruit production and cultivation comprise a sizeable portion of all agricultural activity and are essential to the human diet. On the other hand, fruit diseases (e.g., gray mold, apple scab, canker, black knot, brown rot) are among the key factors that can negatively influence the quality of fruits and reduce their production. Such diseases essentially affect the quality and quantity of fruits by interfering with several processes, such as plant growth, the development of flowers and fruits, and absorbent capacity [9,10]. As a result, early detection and precise classification of fruit diseases is an intriguing problem related to increasing the potential economic value of fruits [11]. The manual inspection of plants and fruits to detect any signs of disease involves various issues. Firstly, it demands ongoing monitoring, disease identification expertise, and other related efforts and time [12]. Furthermore, this method might not detect all infections.

As far as automated techniques for the detection and classification of fruit diseases are concerned, we can categorize them as laboratory methods (e.g., serological methods and molecular techniques [13]) and image processing methods. As image processing methods are inexpensive and non-destructive while yielding superior performance, they have been widely used [14,15]. Fruit plants exhibit various symptoms from their early to late stages, which are, in most cases, distinctly observable to the naked eye on infected fruits and plants [16]. Thus, image processing algorithms can be created to diagnose these conditions quickly, accurately, and affordably using typical digital photographs [17,18].

The use of image processing techniques and machine learning algorithms in disease detection and recognition is a growing research field with great potential to address the challenges involved in the early detection of diseases [13,18,19]. These methods have been developed to achieve accuracies exceeding human-level perception.

As far as the workflow of disease identification and classification using image processing is concerned, it comprises almost identical phases in all approaches. Specifically, the captured image passes through four steps: image pre-processing, image segmentation, feature extraction, and classification. To this end, early works employed classical image processing procedures and ‘hand-crafted’ feature extraction from leaf and fruit images. The extracted features were then used to train shallow classifier algorithms such as support vector machines, principal component analysis, maximum likelihood classification, k-nearest neighbors, naïve Bayes, decision trees, random forest, and artificial neural networks (e.g., [20–23]). Another recent line of work has mainly focused on deep learning architectures comprising convolutional neural networks to automate feature extraction and image classification (e.g., [24–27]).

Despite the exciting results reported by the aforementioned methods, the reliance on a single classifier is deemed to affect the performance of disease detection and classification. In this way, the classification methods based on a single method underperform when applied to large heaps of fruit because of the diversity in texture, shape, color, size, disease, and maturity levels. Here, we propose a hybrid ensemble method for fruit disease identification and fruit classification that employs a multilevel fusion method, intensive fruit image pre-processing, customized image kernels for feature extraction using state-of-the-art

deep methods, and Gini-index-based controlled feature selection. The major contributions of this study are as follows:

1. The hybrid ensemble classification method is inspired by multilevel fusion methodology that includes intensive fruit image pre-processing, adoption of customized image kernels for feature extraction with state-of-the-art (SOTA) deep methods, and Gini-index-based controlled feature selection.
2. Considering different use cases of convolutional neural networks (CNNs) simulated on multiple imagery data, we evaluated the performance of Squeeze Net with defined/customized image kernels.
3. Although the performance of state-of-the-art (SOTA) methods is architecture-dependent, the performance of Squeeze Net was noted to be better with relatively faster training time. It ultimately helped to improve classification accuracy by empowering the weaker classifiers.
4. The ensemble model outperforms individual classifiers with the help of the best feature extractor, Squeeze Net, among other convolution networks, including Inception-V3, VGG-16, VGG-19, and DeepLoc, simulated on larger sets of imagery data.
5. The proposed ensemble method lays the foundation for efficient solutions to similar problems on embedded devices by adopting Squeeze Net (in general) and MobileNetV2 (in particular) for remote imagery data.

The rest of the manuscript is organized as follows. A brief description and the limitations of related work are presented in Section 2. Section 3 presents the proposed method with a description of the essential components of the architecture. Section 4 discusses the significant results, and Section 5 concludes the research work.

2. Related Work

There are two approaches for fruit disease classification. The first approach, applicable at an early stage of plant and fruit development, attempts to identify diseases in plants and leaves [24,28,29]. The second approach involves analyzing the skin and appearance of the fruit for disease identification and classification. The scope of this study is limited to the latter approach. In the following, we provide a brief overview of methods and techniques proposed by researchers for disease identification and classification in several fruits, such as citrus, apple, mango, papaya, and strawberry.

Citrus is one of the most popular fruit in the world. Citrus fruit may suffer from various diseases, such as canker, greening, anthracnose, scab, black spot, and melanose. Cubero et al. provide an excellent review of the application of machine learning techniques for various issues related to citrus fruit, including disease detection [30]. Sharif et al. used the support vector machine (SVM) algorithm to identify and classify these diseases in citrus fruit [31]. The proposed method identifies the lesion spots on the fruit's skin using a segmentation method followed by the classification of diseases. The proposed method is tested on three datasets and achieved 97%, 90.4%, and 89% accuracy for these datasets. Abdulridha et al. used unmanned aerial vehicles to capture hyperspectral images of a citrus orchard for real-time identification of the citrus canker disease [32]. The radial basis function achieved a classification accuracy of 94%, 96%, and 100% for asymptomatic, early, and late symptoms of citrus canker, respectively. The corresponding results for the k-nearest neighbors method were 94%, 95%, and 96%. Rauf et al. prepared a dataset comprising 759 images of healthy and unhealthy citrus leaves and fruit to help other researchers develop methods and techniques for citrus disease detection [33].

Apple is another fruit consumed throughout the world for the whole year. The most common diseases in apple fruit include black rot canker, apple scab, powdery mildew, core rot, white/brown/root rot, collar rot, and seedling blight [34]. Turkoglu et al. proposed an ensemble classification method for apple disease detection [35]. The study applied deep feature extraction on pre-trained models and used the SVM and long short-term memory (LSTM)-based convolutional neural network for disease classification. The authors used a real-time dataset comprising leaves and fruit images. The proposed method achieved

96.4% accuracy with SVM and 99.2% using LSTM. Tian et al. proposed a deep learning model for detecting anthracnose lesions in apple fruit [36]. The authors highlight the difficulty in collecting sufficient data for fruit diseases and use Cycle-Consistent Adversarial Network to augment the available data. The authors used the YOLO-V3 model for disease identification. Fan et al. also used a deep learning model for segregating defective apples from healthy fruit [37]. The study also employed conventional machine learning algorithms, but reported better results using deep learning models. Al-Shawwa and Abu-Naser developed a traditional expert system for detecting all apple fruit diseases mentioned above [34]. A limitation of the study is that the detailed design is not discussed. Moreover, the results are not reported in the study to compare with other solutions.

Strawberries provide great economic value to farmers. However, diseases and pests pose a significant threat to the strawberry fruit. Dong et al. used a deep learning model to identify nine diseases and pests in strawberry fruit, including root rot, slug, thrips, powdery mildew, sunburn, botrytis cinerea, vegetable green insect, virus disease, and aphid [38]. The authors used an enhanced AlexNet by transfer learning that reduced the learning time for the proposed model. The classification accuracy was improved using the inner product and vector max norm. Further details of the proposed model are provided by [39]. Siedliska et al. propose using hyperspectral imaging to detect fungal infections in strawberries [40]. The authors used visible, near-infrared, and shortwave-infrared (VNIR/SWIR) spectroscopy to capture images of two groups of strawberries; a control group without any treatment and a group inoculated with fungi. The study used four classification methods: backpropagation neural network, random forest, naïve Bayes, and support vector machine. The backpropagation neural network achieved the best results.

Habib et al. developed a near real-time computer vision model to detect and identify five diseases in papaya fruit: brown spot, black spot, phytophthora blight, anthracnose, and powdery mildew [41]. A user-captured image from the field is pre-processed by: (a) resizing the image to a fixed size, (b) enhancing contrast through histogram equalization, and (c) converting the image from color to $L^*a^*b^*$ color space. The k-means clustering algorithm is used for image segmentation into healthy and unhealthy fruit regions. The unhealthy regions are then subjected to a feature extraction phase comprising gray level co-occurrence matrix and statistical features. Finally, a support vector machine classifier is used to identify the disease. Andrushia and Patricia use ant colony optimization for feature identification and support vector machine to identify anthracnose and stem end rot diseases of the mango fruit [42].

Several review articles have also addressed the problem of automatic fruit disease identification and classification. Naranjo-Torres et al. have reviewed the application of the convolutional neural network in the problem domain [43]. Bhargava and Bansal have addressed the fruit disease identification problem while reviewing fruit and vegetable grading and quality inspection [15]. Wani et al. provide an excellent review of conventional and deep learning techniques for plant and fruit disease detection [44].

3. Methodology

We propose a multilevel fusion methodology for fruit disease identification and fruit classification that includes intensive fruit image pre-processing, customized image kernels for feature extraction with state-of-the-art (SOTA) deep methods, Gini-index-based controlled feature selection, and hybrid ensemble method for identification and classification.

Figure 1 presents the salient components of the proposed methodology in two segments. The first segment (represented with pink objects) shows a systemic flow of multi-fruit classification. The multi-fruit imagery data contains regular and augmented images of various fruits, including apple, apricot, avocado, banana, cherry, fig, grape, guava, kiwi, mango, orange, peach, pear, pineapple, and strawberry. We identified images of fruits with different shapes, sizes, quality, saturation, and labels that required intensive pre-processing to normalize the fruit for an enhanced classification. In addition, we designed an auto-labeling mechanism based on the existing image clusters to label inconsistent data to belong

to appropriate classes. Lastly, we verified the auto-labeled data with a complete inspection to correctly assign it to the relevant classes. Our proposed methodology can be generalized to any larger set of diverse fruit types to meet the agricultural industry requirements for bulk fruit classification.

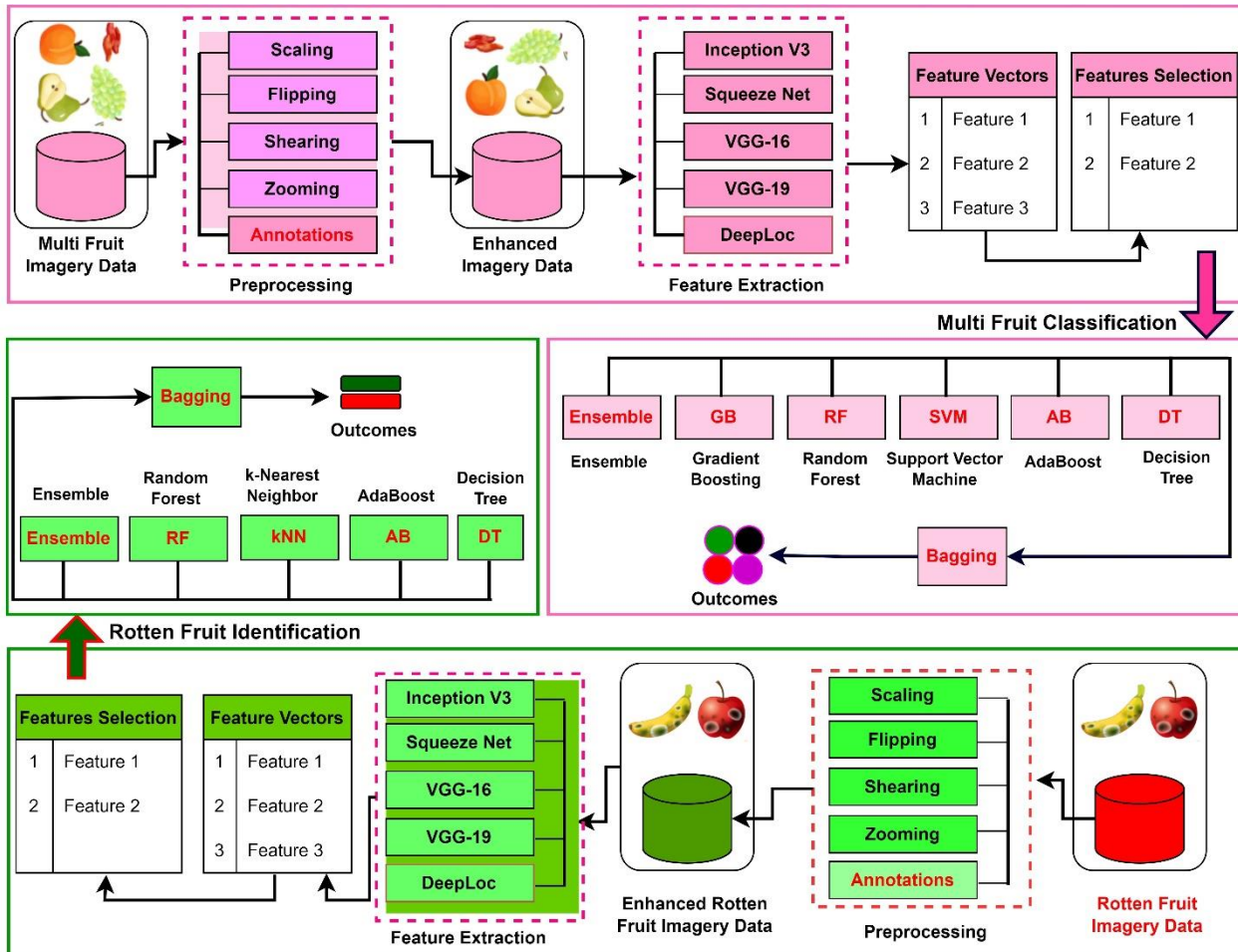


Figure 1. Block diagram of the proposed methodology.

The state-of-the-art (SOTA) convolutional neural networks generate enormous feature sets of image data that are computationally expensive for conventional classification methods. We employed the custom image kernels to extract the relevant and most significant features. Further, we pruned the extracted features using the Gini index, information gain, and entropy of feature values. Eventually, the pruned features were used as inputs for the ensemble classifier. We used decision trees, AdaBoost, support vector machine, random forest, and gradient boosting methods. We employed individual classifiers on diverse imagery data, although classification outcomes fluctuated. The classification accuracy of an individual classifier on certain imagery data was very different on another data set and vice versa. The variations in the outcomes of single classification methods require devising a new ensemble classification method based on the majority voting policy of several distinct classifiers.

The second segment (represented by green objects) depicts a systemic flow of multi-fruit disease classification. The existing multi-fruit disease imagery data also contains regular and augmented images of different fruits including beans (two categories), strawberries (seven categories), and tomatoes (three categories). Here we also noticed that fruit disease data comes in different shapes, sizes, quality, saturation, and labels. Therefore, as with the fruit imagery data, we performed intensive pre-processing to bring the disease imagery data to a consistent format acceptable for an enhanced classification. In the second

segment, we devised an auto-labeling method for correctly labeling unlabeled augmented data (ensuring complete inspection of annotations). At this stage, we also encountered challenges of achieving stable and consistent classification outcomes from single classification methods for various fruit diseases. Here, the proposed ensemble method equally bestowed the significant classification outcomes, tested on various disease data. We illustrate the process in the subsequent section.

Let us consider a set $I_{INPUT} = \{I_1, I_2, I_3, I_4 \dots I_K, \forall K \in N\}$ of input fruit images. Let $I_{OUT} = \{O_1, O_2, O_3, O_4 \dots O_M, \forall M \in N\}$ be a scaled set of I_{INPUT} images. Let $I_K, \forall K \in N$ carry a scale $w \times h$, where $w, h \in [N, N]$. If W, H are new dimensions of $I_K, W \times H$ is a new scale of $O_M, \forall M \in N$, mapped with $O_M(x, y) = O_M(W, H) = I_K(w/W, h/H)$, where $x \in [0, N]$ and $y \in [0, N]$. Further, we consider any fruit image $I(a, b) \in I_{INPUT}, 0 \leq a, b \leq N - 1$ of size $N \times N$. Let $I'(a', b') \in I_{OUT}$, where $0 \leq a', b' \leq M - 1$ of size $M \times M \forall M \in N$. We construct a mapping function $F: I_{INPUT} \rightarrow I_{OUT}$ defines a zoomed image $Z(I'(a', b')) = I(ka, kb)$, where $k \in N$.

Now, we consider a fruit image $J(p, q) \in I_{INPUT}, 0 \leq p, q \leq N - 1$ of size $N \times N$. Let $J'(p', q') \in I_{OUT}$, where $0 \leq p', q' \leq M - 1$ of size $M \times M \forall M \in N$. Here, we consider a new mapping function $S: I_{INPUT} \rightarrow I_{OUT}$ defines a sheared image Z in the x-direction by $J'(p', q') = J(p, kq)$, where $k \in R$. Similarly, we can also construct another mapping function $S': I_{INPUT} \rightarrow I_{OUT}$ that defines a sheared image Z' in the y-direction by $J'(p', q') = J(kp, q)$, where $k \in R$. Let us consider a fruit image $F(r, s) \in I_{INPUT}, 0 \leq r, s \leq N - 1$ of size $N \times N$. Let $F'(r', s') \in I_{OUT}$, where $0 \leq r', s' \leq M - 1$ of size $M \times M \forall M \in N$. Now, we construct a flip function $T: I_{INPUT} \rightarrow I_{OUT}$ that defines a flip image in the x-direction by $F'(r', s') = F(r, s-k)$, where $k \in R$. Further, let function $T': I_{INPUT} \rightarrow I_{OUT}$ for a flipped image in the y-direction by $F'(r', s') = F(r-k, s)$, where $k \in R$. Algorithms 1–4 describe the pre-processing steps.

Algorithm 1: Preprocessing (Scaling) of Fruit Images

Inputs: $I_{INPUT} = \{I_1, I_2, I_3, I_4 \dots I_K, \forall K \in N\}$ of input fruit images.

Outputs: $I_{OUT} = \{O_1, O_2, O_3, O_4 \dots O_M, \forall M \in N\}$

Start PROC_SCALE

Let $w, h = I_K \in I_{INPUT}, K \leq N$, where w, h are width and height of current image

Let $w', h' = 100, 100 \forall I_K \in I_{INPUT}, K \leq N$, where w' and h' are new dimensions of image

Let $s_x = w'/w, s_y = h'/h$, where s_x and s_y are revised scales of image $I_K, K \leq N$

For A in range (h'):

For B in range (w'):

scaled $I_K \rightarrow populate(x, y)$, where populate is a procedure that scales $I_K, K \leq N$

save $I_K \rightarrow O_M, \forall M \in N, O_M \in I_{OUT}$

End For

End For

End PROC_SCALE

Algorithm 2: Preprocessing (Zooming) of Fruit Images

Inputs: $I_{INPUT} = \{I_1, I_2, I_3, I_4 \dots I_K, \forall K \in N\}$ of input fruit images.

Outputs: $I_{OUT} = \{O_1, O_2, O_3, O_4 \dots O_M, \forall M \in N\}$

Start PROC_ZOOM

Let $a, b = I_K \in I_{INPUT}, K \leq N$, where a, b are old dimensions of current image

Let $a', b' = I_M \in I_{OUT}, M \leq N$, where a' and b' are new zoomed dimensions of image

ForX in range (a'):

ForY in range (b'):

Zoomed $I_K \rightarrow generate(a, b)$, where generate will map new dimensions to old dimensions $\forall I_K, K \leq N$

save $I_K \rightarrow O_M, \forall M \in N, O_M \in I_{OUT}$

End For

End For

End PROC_ZOOM

Algorithm 3: Preprocessing (Shearing) of Fruit Images**Inputs:** $I_{\text{INPUT}} = \{I_1, I_2, I_3, I_4 \dots I_K, \forall K \in \mathbb{N}\}$ of input fruit images.**Outputs:** $I_{\text{OUT}} = \{O_1, O_2, O_3, O_4 \dots O_M, \forall M \in \mathbb{N}\}$ *Start PROC_SHEAR**Let $p, q = J(p, q) \in I_{\text{INPUT}}, p, q \leq N$, where p, q are old coordinates of any image**Let $p', q' = J'(p', q') \in I_{\text{OUT}}, p', q' \leq N$, where p' and q' are new sheared dimensions**ForS in range (p'):**ForT in range (q'):**Define S: $I_{\text{INPUT}} \rightarrow I_{\text{OUT}}$ for sheared image Z in x-direction by $J'(p', q') = J(p, kp)$, where $k \in \mathbb{R}$* *Define S': $I_{\text{INPUT}} \rightarrow I_{\text{OUT}}$ for sheared image Z' in y-direction by $J'(p', q') = J(kq, q)$, where $k \in \mathbb{R}$* *Save $Z, Z' \rightarrow O_M, \forall M \in \mathbb{N}, O_M \in I_{\text{OUT}}$* *End For**End For**End PROC_SHEAR***Algorithm 4: Preprocessing (Flipping) of Fruit Images****Inputs:** $I_{\text{INPUT}} = \{I_1, I_2, I_3, I_4 \dots I_K, \forall K \in \mathbb{N}\}$ of input fruit images.**Outputs:** $I_{\text{OUT}} = \{O_1, O_2, O_3, O_4 \dots O_M, \forall M \in \mathbb{N}\}$ *Start PROC_FLIP**Let $r, s = F(r, s) \in I_{\text{INPUT}}, r, s \leq N$, where r, s are old coordinates of any image**Let $r', s' = F'(r', s') \in I_{\text{OUT}}, r', s' \leq N$, where r' and s' are new flipped dimensions**ForA in range (r'):**ForB in range (s'):**Define T: $I_{\text{INPUT}} \rightarrow I_{\text{OUT}}$ for Flipped image F in x-direction by $F'(r', s') = F(r, r-k)$, where $k \in \mathbb{R}$* *Define T': $I_{\text{INPUT}} \rightarrow I_{\text{OUT}}$ for Flipped image F in y-direction by $F'(r', s') = F(r-k, s)$, where $k \in \mathbb{R}$* *Save $F, F' \rightarrow O_M, \forall M \in \mathbb{N}, O_M \in I_{\text{OUT}}$* *End For**End For**End PROC_FLIP*

Let us now consider $I_{\text{OUT}} = \{O_1, O_2, O_3, O_4 \dots O_M, \forall M \in \mathbb{N}\}$ and $I'_{\text{OUT}} = \{O_1, O_2, O_3, O_4 \dots O_M, \forall M \in \mathbb{N}\}$ as pre-processed images of normal and rotten fruit, respectively. We extract feature vectors F , and F' of said datasets, and analyze with classification methods described in Algorithms 5 and 6 below.

Algorithm 5: Ensemble Classification of pre-processed Images*Start: PROC_Fruits-Classification***Inputs:** Feature Vector F of $I_{\text{OUT}} = \{O_1, O_2, O_3, O_4 \dots O_M, \forall M \in \mathbb{N}\}$ images of different fruits**Outputs:** Classification Outcomes vector of different fruits C achieved for P instances*Let us consider a collection $I_{\text{OUT}} = \{O_1, O_2, O_3, O_4 \dots O_M, \forall M \in \mathbb{N}\}$ of processed images**Let us extract the features by applying image embedding to extract $F = \{F_1, F_2, F_3, F_i\}$ features of images $\forall i \leq N$* *Analyze P_i instances with features F using AdaBoost classifier where each P_i in P* *Analyze P_i instances with features F using the Decision tree classifier, each P_i in P* *Analyze P_i instances having features F using Support Vector Machine (SVM) classifier, each P_i in P* *Analyze P_i instances having features F using Random Forest classifier, each P_i in P* *Analyze P_i instances with features F using GradientBoost where each P_i in P* *Analyze the individual performance of all classifiers on P_i attributes of P for $i \leq N$* *Output the classification as a result $C (Y \leq 5)$ of Y classifiers**End**End: PROC_Fruits-Classification*

Algorithm 6: Ensemble Fruit Disease Classification of pre-processed Images**Start: PROC_Fruits-Disease_Classification****Inputs: Feature Vector F' of $Y_{OUT} = \{O_1, O_2, O_3, O_4 \dots O_M, \forall M \in N\}$ images of rotten fruits****Outputs: Classification Outcomes of fruit disease vector C' achieved for P instances**Let us consider a collection $I'_{OUT} = \{O_1, O_2, O_3, O_4 \dots O_M, \forall M \in N\}$ of processed fruit disease imagesLet us extract the features by applying image embedding to extract $F = \{F'_1, F'_2, F'_3, F'_i\}$ features of fruit diseases images $\forall i \leq N$ Analyze P_i instances with features F' using AdaBoost classifier where each P_i in P Analyze P_i instances with features F using the Decision tree classifier, each P_i in P Analyze P_i instances having features F using k-Nearest Neighbor (kNN) classifier, each P_i in P Analyze P_i instances having features F using Random Forest classifier, each P_i in P Analyze the individual performance of all classifiers on P_i attributes of P for $i \leq N$ Output the classification as a result C' ($Y \leq 4$) of Y classifiers**End****End: PROC_Fruits-Disease_Classification**

Further, we implemented the bagging strategy to seek a classification model with the lowest variance compared to single classification methods. The proposed ensemble model achieved the best approximation of bootstrap samples. Let us consider M bootstrap samples of size L as described,

$$\{S_1^1, S_2^1, S_3^1, \dots, S_L^1\}, \{S_1^2, S_2^2, S_3^2, \dots, S_L^2\}, \{S_1^3, S_2^3, S_3^3, \dots, S_L^3\}, \dots, \{S_1^M, S_2^M, S_3^M, \dots, S_L^M\} \quad (*)$$

In the above expression given by (*), S_L^M represents L -th training examples of M -th bootstrap samples. Based on M samples, we can fit about M weak learners *Learner1* (\cdot), *Learner2* (\cdot), *Learner3* (\cdot), \dots , *LearnerM* (\cdot) to seek lower variance by averaging their outcomes. The ensemble model thus becomes:

$$S_M(\cdot) = \arg_T \max [card(m | Learner_m(\cdot) = T)] \quad (1)$$

Using equation 1, we apply soft voting policy, aggregate, and average the probabilities of models presented in Figure 1, and retain the class with the highest mean probability. In addition, the chosen models for voting policy adopted parallelism and provided the flexibility to include as many models as required.

4. Results and Discussion

In the following Section 4.1 we present the numerical results as the performance analysis of different classification methods, while Section 4.2 further illustrates the findings.

4.1. Results

We reviewed several publicly available fruit imagery datasets based on the reported literature on fruit classification and disease identification. We noticed certain limitations in adopting a single data source regarding limited data sizes, variability in fruit types, variability in quality, and variability in disease type. Eventually, we extensively aggregated and pre-processed multi-fruit data to simulate our proposed ensemble model on comprehensive datasets to cover fruit classification and disease identification. The multi-fruit imagery data contained regular and augmented images of fruit including apple, apricot, avocado, banana, cherry, fig, grape, guava, kiwi, mango, orange, peach, pear, pineapple, and strawberry. Similarly, the multi-fruit disease imagery data also contains regular and augmented images of different fruits including beans (two categories), strawberries (seven categories), and tomatoes (three categories). We also identified images of fruits with different shapes, sizes, quality, saturation, and labels that required intensive pre-processing to normalize the fruit for an enhanced classification. In addition, we designed an auto-labeling mechanism based on existing image clusters to label inconsistent data to the appropriate classes. Lastly, we verified the auto-labeled data with a complete inspection to assign the relevant classes correctly.

We employed the following evaluation measures to investigate the performance of the classification methods:

True Positive (TP): The outcome of the model when the model correctly predicts the positive class

False Positive (FP): The outcome of the model when the model incorrectly predicts the positive class

True Negative (TN): The outcome of the model when the model correctly predicts the negative class

False Negative (FN): The outcome of the model when the model incorrectly predicts the positive class

(a) **Accuracy:** The proportion of true results to the total number of cases examined,

$$\text{Classification Accuracy (CA)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \times 100\% \quad (2)$$

(b) **Precision:** The proportion of predicted positives to be truly positive,

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (3)$$

(c) **Recall:** The proportion of actual positives correctly classified,

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (4)$$

(d) **AUC:** How well the probabilities from the positive classes are separated from the negative classes.

(e) **F1-Score:** The overall performance of the model is measured

$$\text{F1 - Score} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \times 100\% \quad (5)$$

Figure 2 shows the performance of different fruit classification methods. The decision tree classifiers performed the worst, with an average of 0.85 for recall, F1 score, and classification accuracy. The SVM and random forest classifiers depict similar outcomes on almost all evaluation metrics, with evaluation measures of 0.93 to 0.95. The proposed ensemble classifier outperforms all other classification methods, achieving 100% for AUC, CA, F1 score, precision, and recall.

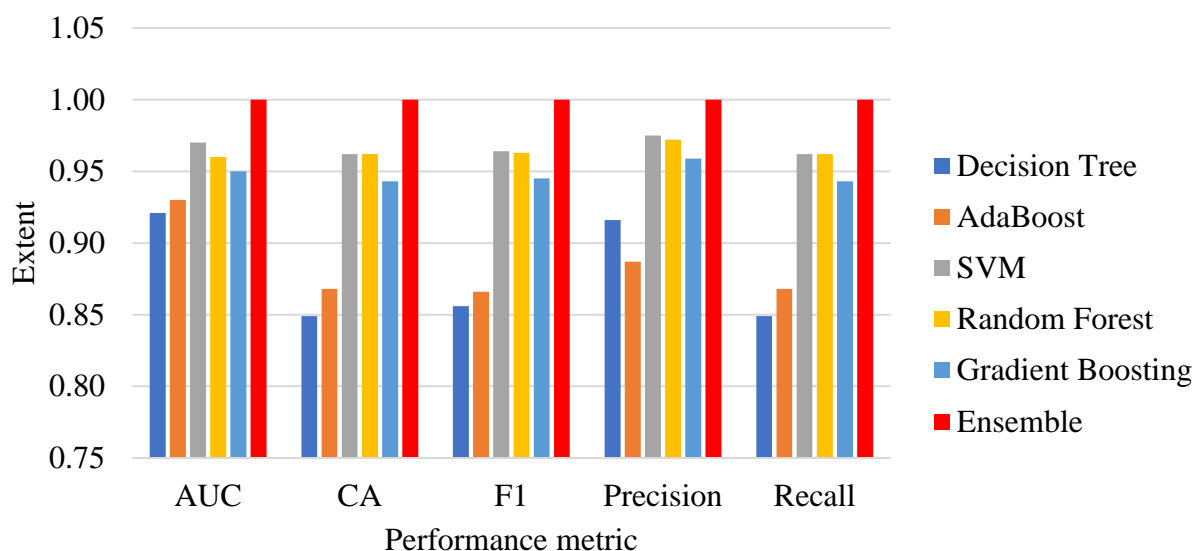


Figure 2. Performance analysis of fruit classification methods.

Table 1 presents the misclassification (partial confusion matrix) of different classification methods (only cells colored red refer to misclassifications). The decision tree classifier depicts a more significant misclassification for some fruits, i.e., fig, guava, kiwi, mango, orange, and pear. Gradient boosting and random forest misclassified guava, mango, and pear. The ensemble classification had no misclassification except for guava (5%) and outperformed all single classification methods. Almost all classification methods carried a more considerable misclassification for the guava fruit. Similarly, mango was also misclassified to some extent by decision trees, AdaBoost, SVM, gradient boost, and random forest algorithms.

Table 1. Misclassification of classification methods.

Fruit	Model					
	Decision Tree	AdaBoost	SVM	Random Forest	Gradient Boosting	Ensemble
Apple	100%	100%	100%	100%	100%	100%
Apricot	100%	20%	100%	100%	100%	100%
Avocado	100%	100%	100%	100%	100%	100%
Banana	100%	100%	100%	100%	100%	100%
Cherry	100%	100%	100%	100%	100%	100%
Fig	14%	17%	100%	100%	100%	100%
Grape	100%	100%	100%	100%	100%	100%
Guava	20%	50%	17%	13%	33%	5%
Kiwi	20%	25%	100%	100%	100%	100%
Mango	14%	17%	17%	13%	33%	100%
Orange	50%	100%	100%	100%	100%	100%
Peach	100%	100%	100%	100%	100%	100%
Pear	40%	17%	100%	13%	17%	100%
Pineapple	100%	100%	100%	100%	100%	100%
Strawberry	100%	100%	100%	100%	100%	100%

Figure 3 presents the ROC analysis of the classification methods. The AdaBoost and decision tree classifier methods performed the worst. Gradient boosting and random forest classification methods show similar true positive/false positive rates. In contrast, the ensemble classification method stands out with a significant curve leaning towards the upper left corner of the ROC graph.

Further, we calculated the performance metrics of different classification methods for fruit disease identification. We recorded the following observations.

Figure 4 demonstrates the performance analysis of different classification methods for fruit disease identification. Here, we can notice that the decision tree classifier underperformed with an average of 0.67 for recall, precision, F1 score, and CA, while it achieved 0.86 for the area under the curve. Random forest showed an intermediate performance compared to other classification methods, with an average score of 0.78 for recall, precision, F1, and CA, while it scored 0.96 for the AUC metric. Similarly, the kNN performed well for all evaluation metrics. We can notice the outperformance of the ensemble classification method that scored 0.99 for the AUC and 0.93 for other evaluation metrics.

Table 2 shows the misclassification of classification methods for fruit disease identification (only cells colored red refer to misclassifications). We can see that the decision tree classifier had a more significant misclassification of 51%, 26%, 9%, 3%, and 6%, for Beans-1, Beans-2, Strawberry-1, Strawberry-2, and Strawberry-4, respectively. Similarly, AdaBoost missed fruits Beans-1, Beans-2, and Strawberry-5 with a more significant 77%, 19%, and 82%, respectively. The kNN and random forest misclassified Beans-1 and Beans-2 by 82% and 76%, respectively. The ensemble classifier stood out with misclassifications of 3% and 7% for Beans-1 and Beans-2, respectively. Almost all classification methods misclassified Beans-1 and Beans-2 to varying extents.

Figure 5 shows the receiver operating characteristic (ROC) curve analysis of the classification methods. The AdaBoost and Decision Tree classifiers performed the worst,

and kNN and Random Forest classification methods show similar true positive/false positive rates. In contrast, the ensemble classification method stands out with a significant curve leaning towards the upper left corner of the ROC graph.

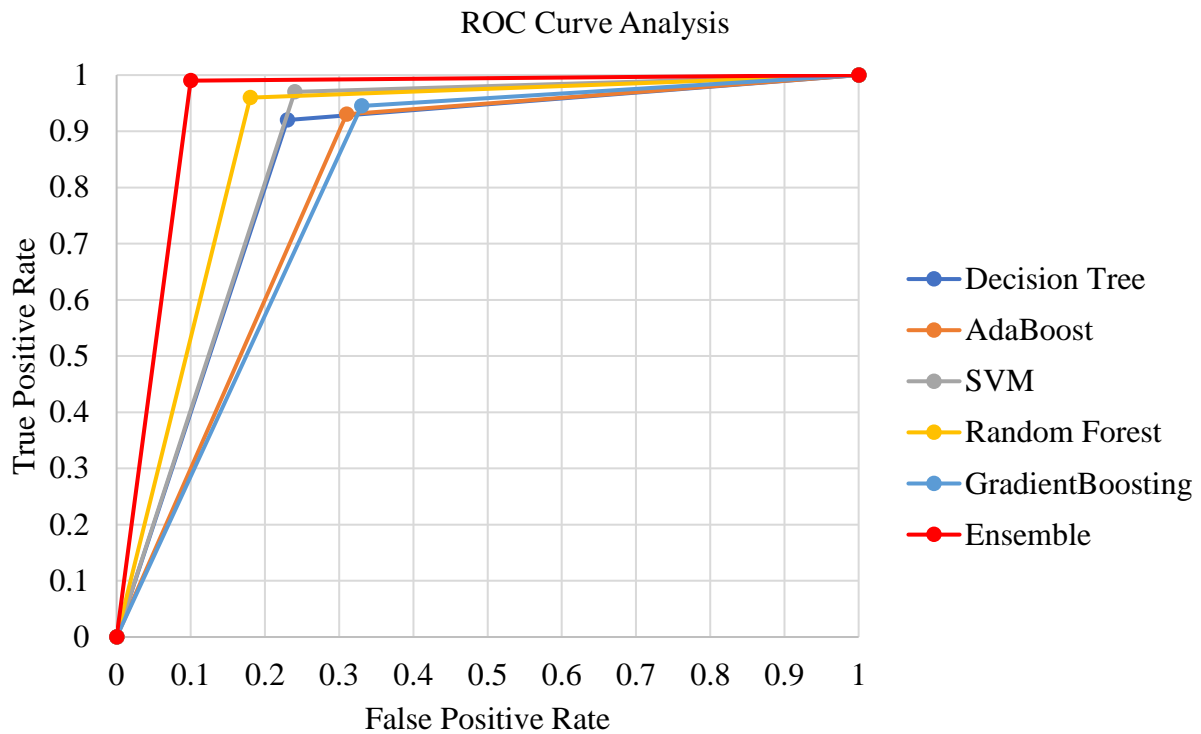


Figure 3. Receiver operating characteristic curve analysis.

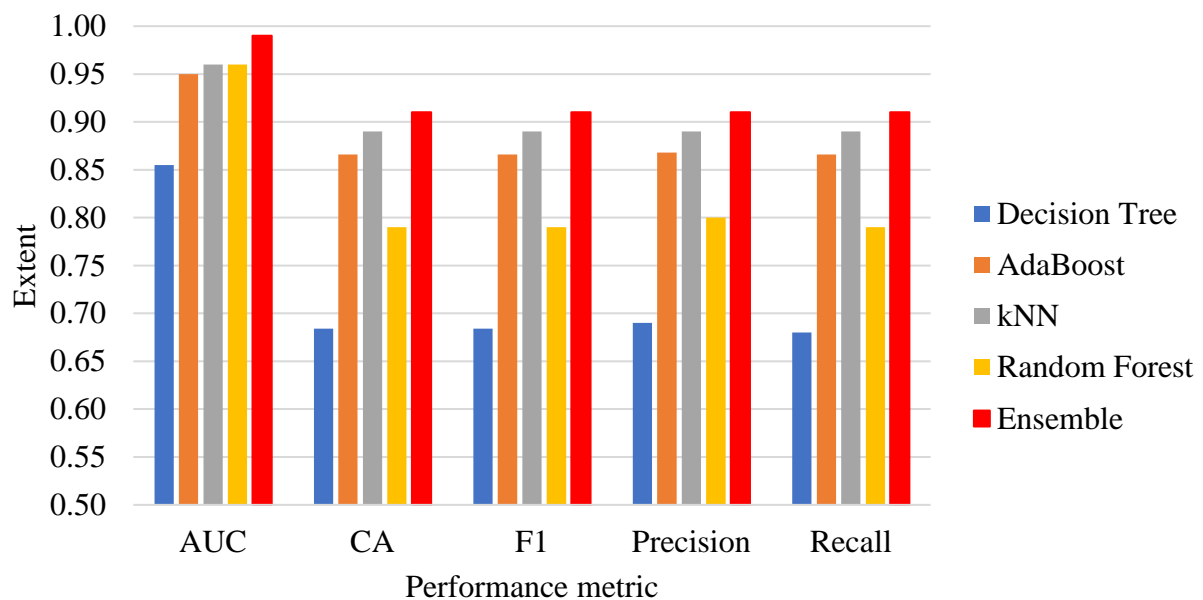


Figure 4. Performance metrics for fruit disease identification.

Table 2. Misclassification of classification methods for fruit disease identification.

Fruit	Model				
	Decision Tree	AdaBoost	kNN	Random Forest	Ensemble
Beans-1	51%	77%	82%	76%	3%
Beans-2	26%	19%	11%	24%	7%
Strawberry-1	9%	4%	7%	100%	100%
Strawberry-2	3%	100%	100%	100%	100%
Strawberry-3	100%	100%	100%	100%	100%
Strawberry-4	6%	100%	30%	50%	100%
Strawberry-5	100%	82%	70%	100%	100%
Strawberry-6	100%	100%	100%	100%	100%
Strawberry-7	6%	12%	100%	60%	100%
Tomato-1	100%	100%	100%	100%	100%
Tomato-2	100%	100%	100%	100%	100%
Tomato-3	100%	100%	100%	100%	100%

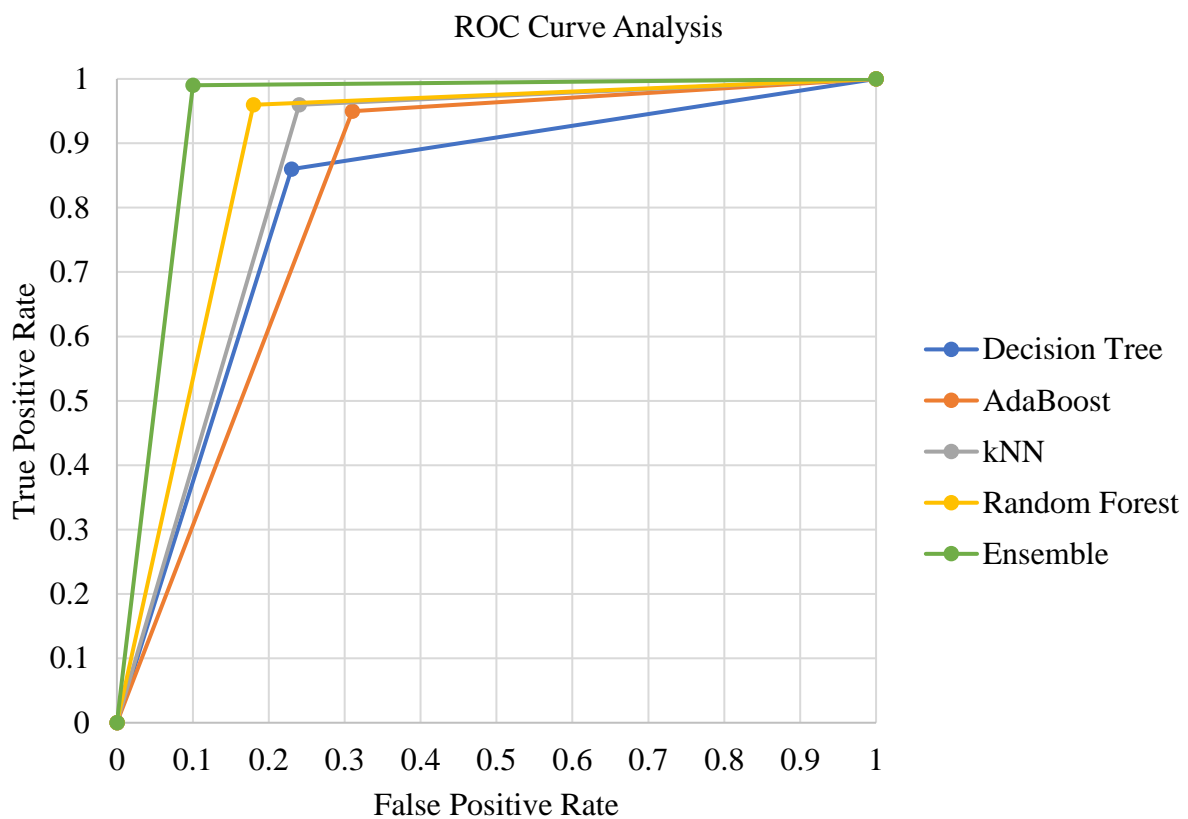


Figure 5. ROC analysis of classification methods for fruit disease identification.

4.2. Discussion

To analyze the performance of our proposed ensemble classification method, we compiled a larger dataset of over 20,000 images. A careful review of the literature reported the employment of various datasets. The reported datasets varied to a larger extent regarding the size, shape, quality, and variety of fruits for classification and disease identification.

We segregated our datasets into four major segments. We applied bootstrapping on each segment to identify an optimal method for feature extraction. The study considered Inception V3, Squeeze Net, VGG-16, VGG-19, and DeepLoc as the best candidate networks for feature extraction.

Table 3 highlights the comparative analysis of different feature extraction methods. We employed the convolution layers of these models. Based on the images' embedding, we noticed the following cumulative statistics:

Table 3. Comparative analysis of feature extraction methods.

Segment 1					
Model	AUC	CA	F1	Precision	Recall
Inception V3	0.930	0.868	0.866	0.887	0.868
Squeeze Net	0.980	0.962	0.963	0.972	0.962
VGG-16	0.950	0.906	0.907	0.928	0.906
VGG-19	0.930	0.868	0.840	0.832	0.868
DeepLoc	0.971	0.943	0.945	0.960	0.943
Segment 2					
Model	AUC	CA	F1	Precision	Recall
Inception V3	0.921	0.849	0.856	0.916	0.849
Squeeze Net	0.969	0.943	0.942	0.956	0.943
VGG-16	0.950	0.906	0.916	0.958	0.906
VGG-19	0.950	0.906	0.908	0.931	0.906
DeepLoc	0.938	0.868	0.869	0.900	0.868
Segment 3					
Model	AUC	CA	F1	Precision	Recall
Inception V3	0.998	0.962	0.964	0.975	0.962
Squeeze Net	0.998	0.962	0.962	0.970	0.962
VGG-16	0.992	0.962	0.964	0.975	0.962
VGG-19	0.992	0.962	0.964	0.975	0.962
DeepLoc	0.989	0.962	0.964	0.975	0.962
Segment 4					
Model	AUC	CA	F1	Precision	Recall
Inception V3	0.998	0.935	0.939	0.955	0.935
Squeeze Net	0.989	0.946	0.951	0.964	0.946
VGG-16	0.987	0.935	0.940	0.957	0.935
VGG-19	0.987	0.927	0.934	0.956	0.927
DeepLoc	0.984	0.904	0.908	0.927	0.904

Figure 6 describes the performance metrics of different feature extraction models. For our particular domain of fruit classification and disease identification, we recorded that Squeeze Net outperformed the other models with an extent of 0.99 for AUC, 0.95 for CA, F1 score, and recall. In addition, it showed a precision of 0.97 compared to the other models.

Figure 7 shows the selection of features based on information gain, gain ratio, and Gini index. The principal components are represented with letters "PC", for instance "PC3" represents the third principal component. The Gini index is considered an excellent split criterion for most machine learning algorithms. This is the reason that we restricted our findings to the Gini index. In addition, the other measures, i.e., information gain and gain ratio, depicted similar characteristics. We can see the features sorted in descending order with respect to the Gini index referring to the most robust features (candidates) for performance analysis of ensemble classification models and other classifiers.

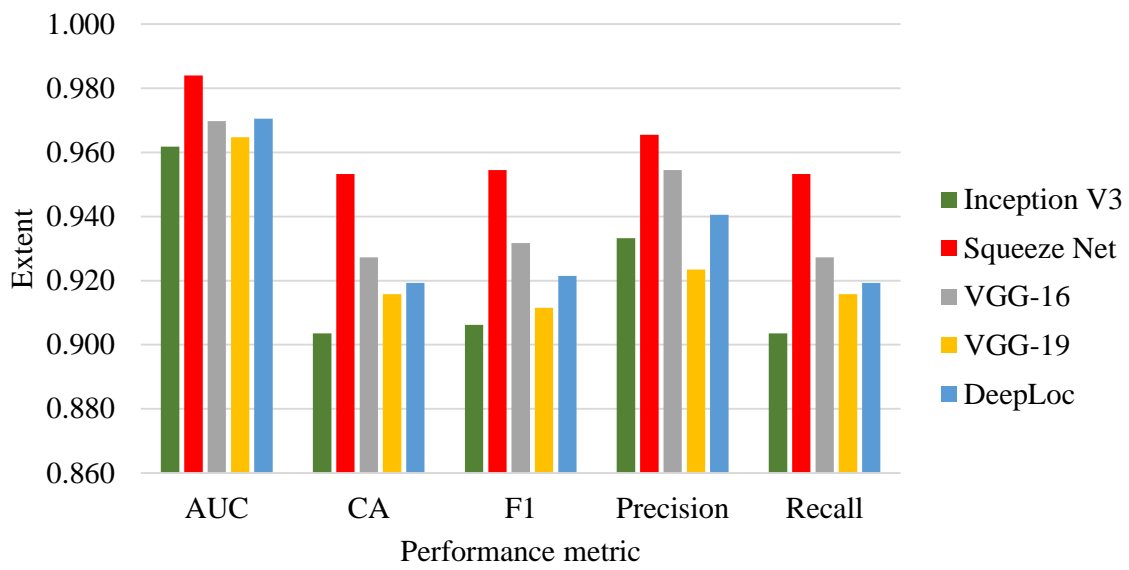


Figure 6. Cumulative performance analysis of models.

		Info. gain	Gain ratio	Gini
1	N PC3	1.147	0.573	0.105
2	N PC7	1.102	0.551	0.105
3	N PC8	1.088	0.544	0.107
4	N PC2	1.066	0.533	0.098
5	N PC6	1.031	0.515	0.095
6	N PC10	1.005	0.502	0.088
7	N PC4	0.996	0.498	0.090
8	N PC5	0.983	0.491	0.088
9	N PC9	0.972	0.486	0.090
10	N PC12	0.960	0.480	0.088
11	N PC1	0.869	0.434	0.078
12	N PC13	0.816	0.408	0.072
13	N PC15	0.707	0.353	0.064
14	N PC11	0.694	0.347	0.064
15	N PC14	0.639	0.319	0.056

Figure 7. Feature selection based on characteristics.

Furthermore, we validated the feature selection using central data tendency measures as shown in Figure 8. Figure 8 is a high-resolution figure prepared at 600 pixels per inch. It may be zoomed in to view the contents.

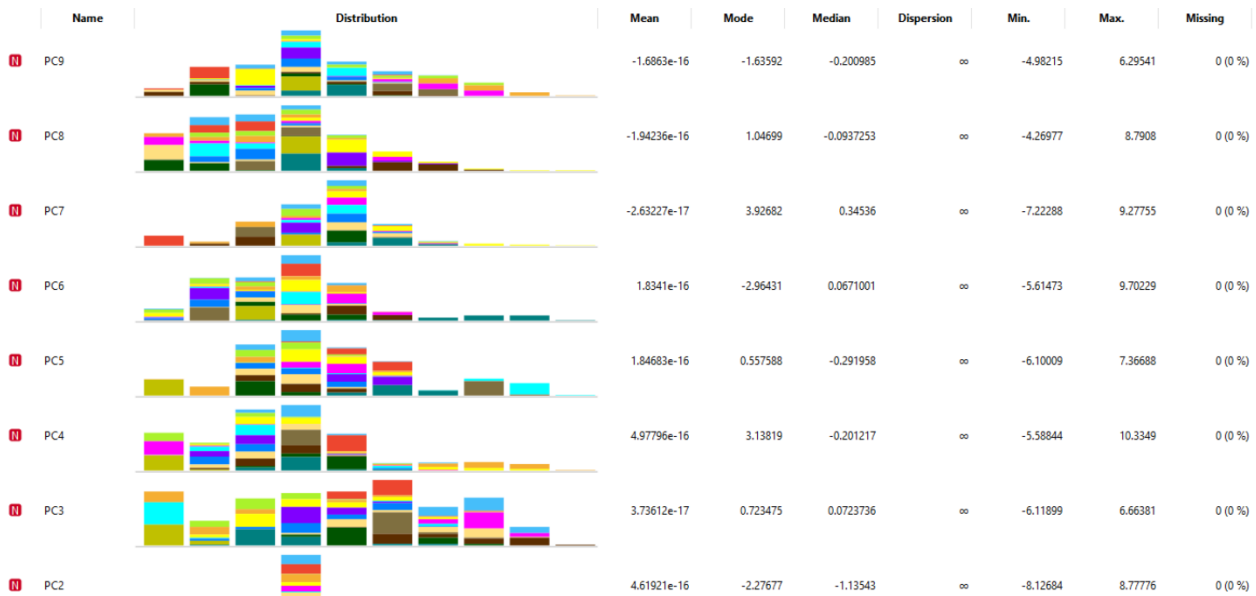


Figure 8. Basic consideration of measures for feature selection.

Figure 8 presents an exponential number format which can be read as follows, the subsequent signed number (either plus or minus) succeeding “e” is raised to the power of 10, for instance number 8×10^{-3} will be read as 8×10^{-3} . The central data tendency measure gives the data distribution with respect to mean, mode, and median. It also glimpses the data dispersion, minimum, maximum, and missing values (if any) in certain features. We can validate the chosen strongest candidates concerning variance in the data and their distance from the mean. This gives us a second opinion regarding the selection of better candidate features.

In addition, we also measured the relevance of features with each other as follows:

Figure 9 shows the impact between features. For instance, we can identify different fruit clusters among principal components (1) and (2). Apple and cherry fruits are well segregated, while mango, peach, pear, and strawberry overlap at certain data points of two features. This is another dimension that gives us the co-relevance among principal components to understand the dispersion of data into clusters shared by two or more principal components.

Figure 10 presents the classification accuracy and error of different classification methods. It should be noted that hyphen sign with vertical/horizontal scale represent minus sign, for instance -8 can be read as (−, 8) for clarity. The actual accuracy is represented with a green bar and the residual error with red bars. We can notice that the ensemble method has negligible (or no) errors in the correct classification of fruits. Further, the random forest and SVM classifiers show a similar residual error of 0.45. The decision tree classifier carries the highest residual error in the correct classification of fruits.

Figure 11 shows the classification accuracy and error of methods for fruit disease identification. The decision tree classifier carries the most significant residual error, while kNN and AdaBoost contain an error in an intermediate range of 0.10. The random forest algorithm shows a 0.21 error. The ensemble classifier produced the slightest error in fruit disease identification.

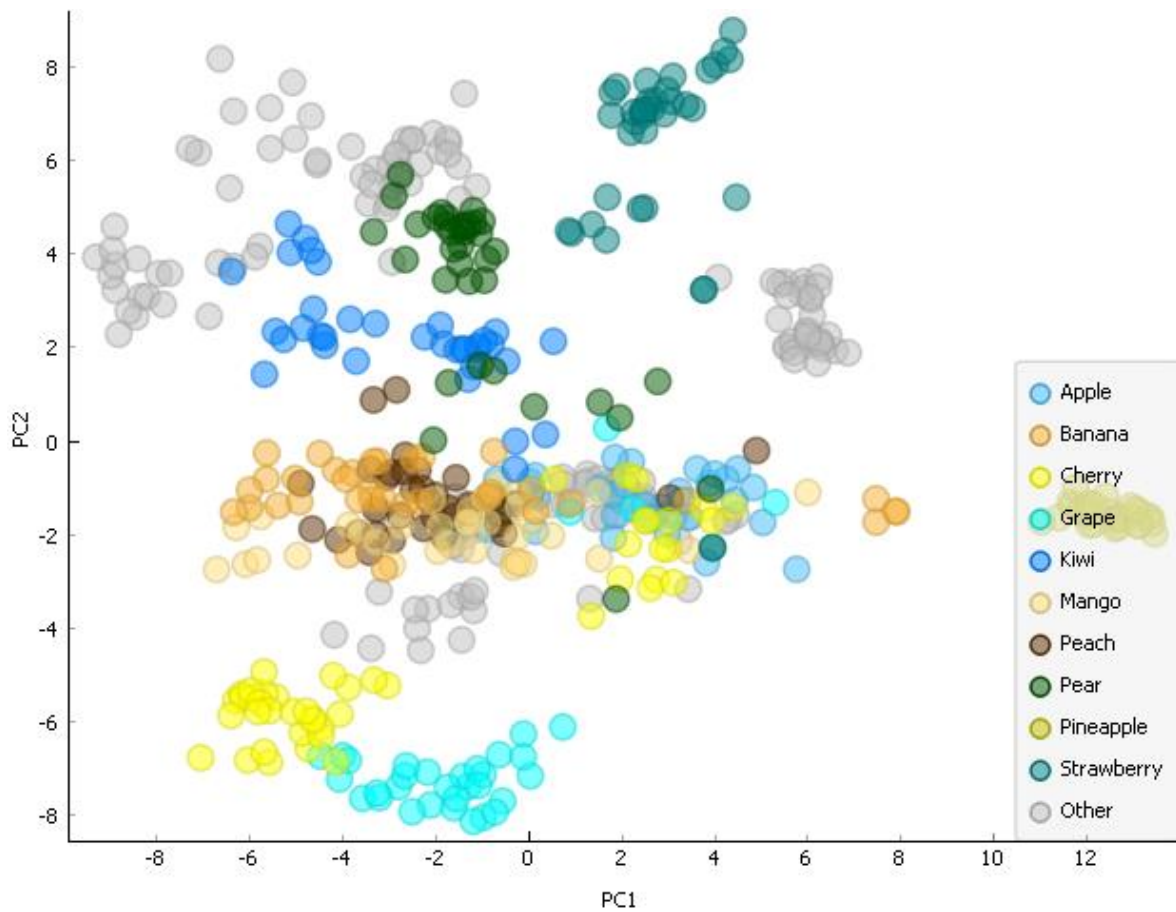


Figure 9. Scatter plot analysis of selected features.

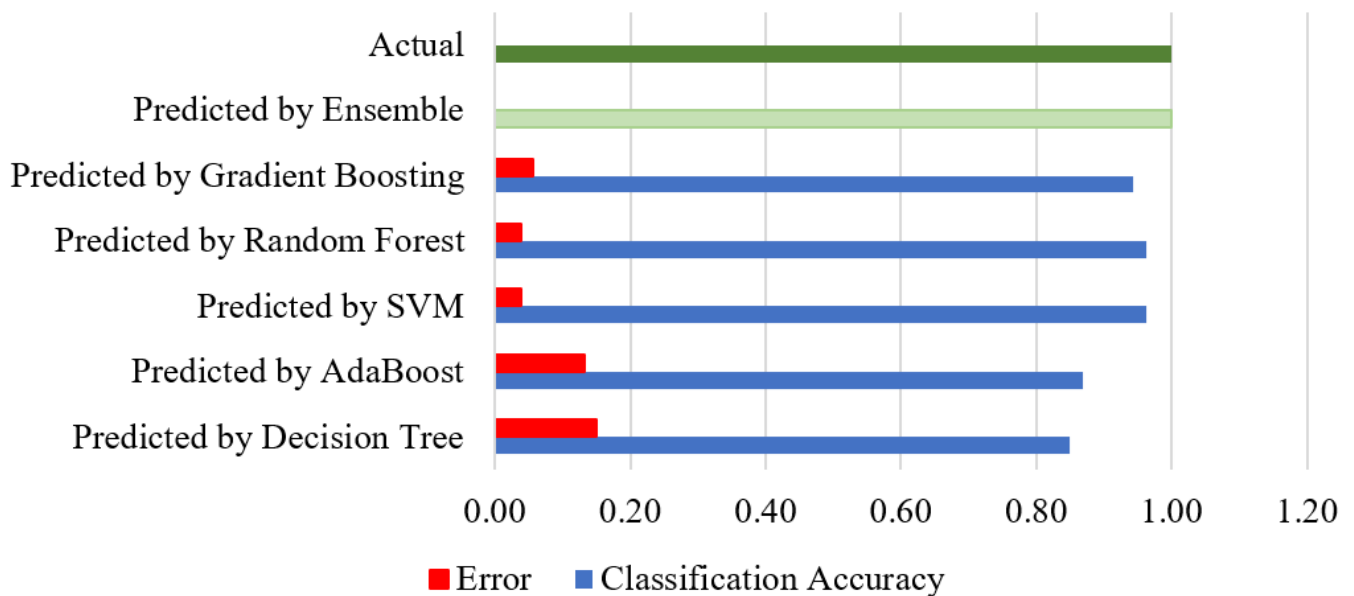


Figure 10. Classification accuracy and error in fruit classification.

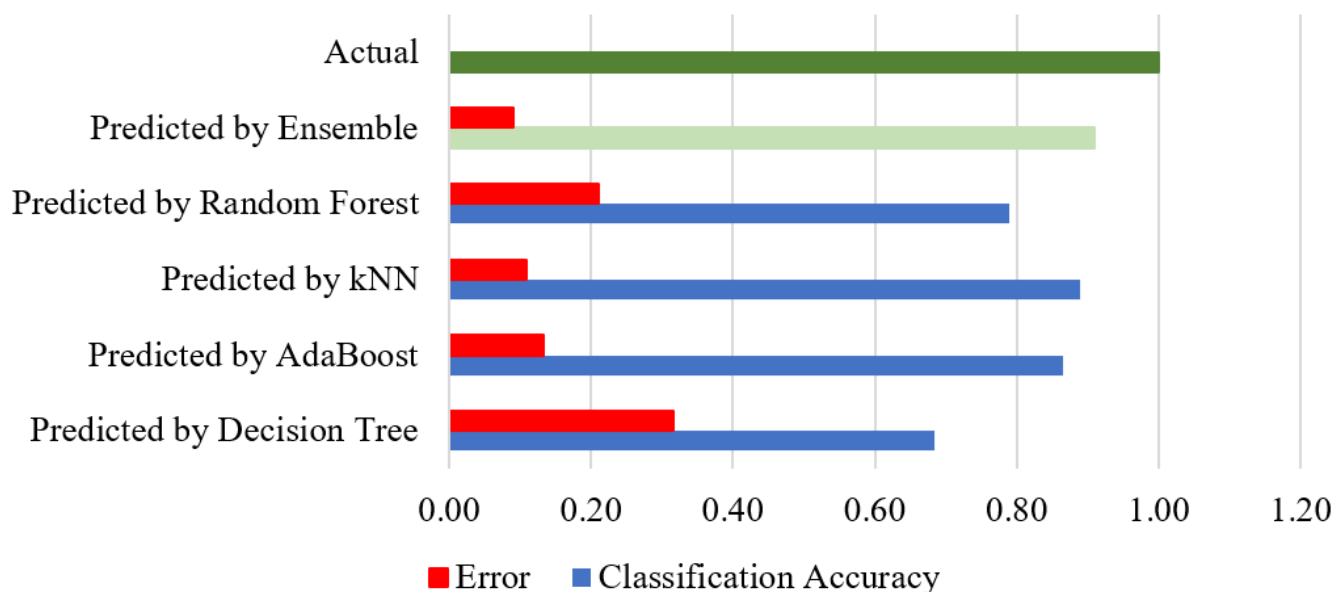


Figure 11. Classification accuracy and error in fruit disease identification.

Further, we performed the analysis of variance (ANOVA) test on our simulated results (including both fruit classification and disease identification). ANOVA employs the F-test to investigate the variability between group means within and between the groups. We achieved the following outcomes:

In Table 4, chosen on a 95% confidence interval ($\alpha = 0.05$), the F value is 32.41, which is significantly larger than the F-critical value of 2.62, achieving a p -value of 0.00, which is significantly smaller than 0.05.

Table 4. ANOVA: Statistical significance of outcomes of fruit classification methods ($\alpha = 0.05$).

SUMMARY						
Groups	Count	Sum	Average	Variance		
Decision Tree	5	4.39	0.878	0.0013647		
AdaBoost	5	4.41	0.883	0.0007402		
SVM	5	4.83	0.966	3.28×10^{-5}		
Random Forest	5	4.81	0.963	0.0000222		
Gradient Boosting	5	4.74	0.948	4.6×10^{-5}		
Ensemble	5	5	1	0		
Source of Variation	SS	df	MS	F	p -value	F crit
Between Groups	0.05957	5	0.01191	32.4103	0.00000	2.620
Within Groups	0.00882	24	0.00036			
Total	0.068401	29				

Similarly, we performed the ANOVA test on the outcomes of SOTA methods used for feature extraction.

In Table 5, chosen on a 95% confidence interval ($\alpha = 0.05$), the F value is 11.42, which is significantly larger than the F-critical value of 2.866, achieving a p -value of 0.00, significantly smaller than 0.05.

Table 5. ANOVA: Statistical significance of outcomes of disease identification methods ($\alpha = 0.05$).

SUMMARY						
Groups	Count	Sum	Average	Variance		
Decision Tree	5	3.59	0.718	0.0058		
AdaBoost	5	4.41	0.883	0.0013		
kNN	5	4.52	0.904	0.0008		
Random Forest	5	4.13	0.826	0.0053		
Ensemble	5	4.63	0.926	0.0012		
Source of Variation	SS	df	MS	F	p-value	F crit
Between Groups	0.138120	4	0.0345	11.424	0.000054	2.866
Within Groups	0.060448	20	0.0030			
Total	0.19856	24				

5. Conclusions

We proposed a hybrid ensemble classification method for fruit classification and disease identification, inspired by multilevel fusion methodology that includes intensive fruit image pre-processing, adoption of customized image kernels for feature extraction with state-of-the-art (SOTA) deep learning methods, and Gini-index-based controlled feature selection. Realizing the limitations of publicly available fruit imagery datasets regarding data sizes, variability in fruit types, quality, and diseases, we extensively aggregated and pre-processed multi-fruit data to achieve augmented images of various fruits, including apple, apricot, avocado, banana, cherry, fig, grape, guava, kiwi, mango, orange, peach, pear, pineapple, beans (two categories), strawberries (seven categories), and tomatoes (three categories). The ensemble classifier outperformed all classification methods, achieving 100% and 99% classification accuracies for fruit classification and disease identification. We validated the statistical significance of the classifiers' outcomes at $\alpha = 0.05$ and achieved F-values of 32.41 and 11.42 against F-critical values of 2.62 and 2.86, resulting in p -values of 0.00 (<0.05) for fruit classification and disease identification. As a future work direction, we consider implementing the hybrid ensemble algorithm to embedded devices for remote classification of image sensory data. Another limitation of the study regards multiple simultaneous diseases in a fruit. New data sets need to be developed to test the performance of our approach for such instances.

Author Contributions: Conceptualization, A.M., M.A. and Q.M.I.; methodology, Q.M.I., A.M. and M.A.; formal analysis, A.M., M.A. and Q.M.I.; investigation, A.M., M.A. and Q.M.I.; resources, A.M.; data curation, A.M., M.A. and Q.M.I.; writing—original draft preparation, A.M., M.A. and Q.M.I.; project administration, A.M.; funding acquisition, A.M. All authors have read and agreed to the published version of the manuscript.

Funding: The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number INST030.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Food and Agricultural Organization (FAO). Crop Production and Natural Resource Use. Available online: <http://www.fao.org/3/y4252e/y4252e06.htm> (accessed on 7 December 2022).
2. Calicioglu, O.; Flammini, A.; Bracco, S.; Bellù, L.; Sims, R. The Future Challenges of Food and Agriculture: An Integrated Analysis of Trends and Solutions. *Sustainability* **2019**, *11*, 222. [CrossRef]
3. Rathore, S.S.; Shekhawat, K.; Singh, R.K.; Babu, S.; Singh, V.K. Diversification for Restoration of Ecosystems and Sustainable Livelihood. *Sustain. Agric. Syst. Technol.* **2022**, *21*–36. [CrossRef]

4. Hunter, M.C.; Smith, R.G.; Schipanski, M.E.; Atwood, L.W.; Mortensen, D.A. Agriculture in 2050: Recalibrating Targets for Sustainable Intensification. *Bioscience* **2017**, *67*, 386–391. [[CrossRef](#)]
5. Research for AGRI Committee - The EU Farming Employment: Current Challenges and Future Prospects. Available online: [https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU\(2019\)629209](https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU(2019)629209) (accessed on 7 December 2022).
6. Horton, P.; Long, S.P.; Smith, P.; Banwart, S.A.; Beerling, D.J. Technologies to deliver food and climate security through agriculture. *Nat. Plants* **2021**, *7*, 250–255. [[CrossRef](#)]
7. Cisternas, I.; Velásquez, I.; Caro, A.; Rodríguez, A. Systematic literature review of implementations of precision agriculture. *Comput. Electron. Agric.* **2020**, *176*, 105626. [[CrossRef](#)]
8. Lassoued, R.; Macall, D.M.; Smyth, S.J.; Phillips, P.W.B.; Hessel, H. Expert Insights on the Impacts of, and Potential for, Agricultural Big Data. *Sustainability* **2021**, *13*, 2521. [[CrossRef](#)]
9. Ali, M.M.; Yousef, A.F.; Li, B.; Chen, F. Effect of Environmental Factors on Growth and Development of Fruits. *Trop. Plant Biol.* **2021**, *14*, 226–238. [[CrossRef](#)]
10. Paradiso, R.; Proietti, S. Light-Quality Manipulation to Control Plant Growth and Photomorphogenesis in Greenhouse Horticulture: The State of the Art and the Opportunities of Modern LED Systems. *J. Plant Growth Regul.* **2022**, *41*, 742–780. [[CrossRef](#)]
11. Wieme, J.; Mollazade, K.; Malounas, I.; Zude-Sasse, M.; Zhao, M.; Gowen, A.; Argyropoulos, D.; Fountas, S.; Van Beek, J. Application of hyperspectral imaging systems and artificial intelligence for quality assessment of fruit, vegetables and mushrooms: A review. *Biosyst. Eng.* **2022**, *222*, 156–176. [[CrossRef](#)]
12. Khan, M.A.; Akram, T.; Sharif, M.; Alhaisoni, M.; Saba, T.; Nawaz, N. A probabilistic segmentation and entropy-rank correlation-based feature selection approach for the recognition of fruit diseases. *Eurasip J. Image Video Process.* **2021**, *2021*, 1–28. [[CrossRef](#)]
13. Manavalan, R. Automatic identification of diseases in grains crops through computational approaches: A review. *Comput. Electron. Agric.* **2020**, *178*, 105802. [[CrossRef](#)]
14. Ouhami, M.; Hafiane, A.; Es-Saady, Y.; El Hajji, M.; Canals, R. Computer Vision, IoT and Data Fusion for Crop Disease Detection Using Machine Learning: A Survey and Ongoing Research. *Remote Sens.* **2021**, *13*, 2486. [[CrossRef](#)]
15. Bhargava, A.; Bansal, A. Fruits and vegetables quality evaluation using computer vision: A review. *J. King Saud Univ.-Comput. Inf. Sci.* **2021**, *33*, 243–257. [[CrossRef](#)]
16. Zhang, X.; Qiao, Y.; Meng, F.; Fan, C.; Zhang, M. Identification of Maize Leaf Diseases Using Improved Deep Convolutional Neural Networks. *IEEE Access* **2018**, *6*, 30370–30377. [[CrossRef](#)]
17. Thakur, P.S.; Khanna, P.; Sheorey, T.; Ojha, A. Trends in vision-based machine learning techniques for plant disease identification: A systematic review. *Expert Syst. Appl.* **2022**, *208*, 118117. [[CrossRef](#)]
18. Ngugi, L.C.; Abelwahab, M.; Abo-Zahhad, M. Recent advances in image processing techniques for automated leaf pest and disease recognition—A review. *Inf. Process. Agric.* **2020**, *8*, 27–51. [[CrossRef](#)]
19. Shashi; Singh, J. Automated Disease Detection and Classification of Plants Using Image Processing Approaches: A Review. *Lect. Notes Netw. Syst.* **2021**, *203*, 641–651. [[CrossRef](#)]
20. Johannes, A.; Picon, A.; Alvarez-Gila, A.; Echazarra, J.; Rodriguez-Vaamonde, S.; Navajas, A.D.; Ortiz-Barredo, A. Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. *Comput. Electron. Agric.* **2017**, *138*, 200–209. [[CrossRef](#)]
21. Krithika, N.; Selvarani, A.G. An individual grape leaf disease identification using leaf skeletons and KNN classification. In Proceedings of the 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 17–18 March 2017; Volume 138, pp. 1–5. [[CrossRef](#)]
22. Islam, M.; Anh, D.; Wahid, K.; Bhowmik, P. Detection of potato diseases using image segmentation and multiclass support vector machine. In Proceedings of the 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), Windsor, ON, Canada, 30 April–3 May 2017; pp. 1–4. [[CrossRef](#)]
23. Qin, F.; Liu, D.; Sun, B.; Ruan, L.; Ma, Z.; Wang, H. Identification of Alfalfa Leaf Diseases Using Image Recognition Technology. *PLoS ONE* **2016**, *11*, e0168274. [[CrossRef](#)]
24. Li, L.; Zhang, S.; Wang, B. Plant Disease Detection and Classification by Deep Learning—A Review. *IEEE Access* **2021**, *9*, 56683–56698. [[CrossRef](#)]
25. Bierman, A.; LaPlumm, T.; Cadle-Davidson, L.; Gadoury, D.; Martinez, D.; Sapkota, S.; Rea, M. A High-Throughput Phenotyping System Using Machine Vision to Quantify Severity of Grapevine Powdery Mildew. *Plant Phenomics* **2019**, *2019*, 1–13. [[CrossRef](#)] [[PubMed](#)]
26. JBarbedo, J.G. Factors influencing the use of deep learning for plant disease recognition. *Biosyst. Eng.* **2018**, *172*, 84–91. [[CrossRef](#)]
27. Fuentes, A.; Yoon, S.; Kim, C.S.; Park, S.D. A Robust Deep-Learning-Based Detector For Real-Time Tomato Plant Diseases and Pests Recognition. *Precis. Agric.* **2017**, *17*, 2022. [[CrossRef](#)] [[PubMed](#)]
28. Vishnoi, V.K.; Kumar, K.; Kumar, B. Plant disease detection using computational intelligence and image processing. *J. Plant Dis. Prot.* **2020**, *128*, 19–53. [[CrossRef](#)]
29. Buja, I.; Sabella, E.; Monteduro, A.; Chiriaco, M.; De Bellis, L.; Luvisi, A.; Maruccio, G. Advances in Plant Disease Detection and Monitoring: From Traditional Assays to In-Field Diagnostics. *Sensors* **2021**, *21*, 2129. [[CrossRef](#)]
30. Cubero, S.; Lee, W.S.; Aleixos, N.; Albert, F.; Blasco, J. Automated Systems Based on Machine Vision for Inspecting Citrus Fruits from the Field to Postharvest—A Review. *Food Bioprocess Technol.* **2016**, *9*, 1623–1639. [[CrossRef](#)]

31. Sharif, M.; Khan, M.A.; Iqbal, Z.; Azam, M.F.; Lali, M.I.U.; Javed, M.Y. Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. *Comput. Electron. Agric.* **2018**, *150*, 220–234. [[CrossRef](#)]
32. Abdulridha, J.; Batuman, O.; Ampatzidis, Y. UAV-Based Remote Sensing Technique to Detect Citrus Canker Disease Utilizing Hyperspectral Imaging and Machine Learning. *Remote Sens.* **2019**, *11*, 1373. [[CrossRef](#)]
33. Rauf, H.T.; Saleem, B.A.; Lali, M.I.U.; Khan, M.A.; Sharif, M.; Bukhari, S.A.C. A citrus fruits and leaves dataset for detection and classification of citrus diseases through machine learning. *Data Brief* **2019**, *26*, 104340. [[CrossRef](#)]
34. Al-Shawwa, M.; Abu-Naser, S.S. Knowledge Based System for Apple Problems Using CLIPS. *Int. J. Acad. Eng. Res.* **2019**, *3*, 1–11. Available online: www.ijeais.org/ijaer (accessed on 22 November 2022).
35. Turkoglu, M.; Hanbay, D.; Sengur, A. Multi-model LSTM-based convolutional neural networks for detection of apple diseases and pests. *J. Ambient Intell. Humaniz. Comput.* **2019**, *13*, 3335–3345. [[CrossRef](#)]
36. Tian, Y.; Yang, G.; Wang, Z.; Li, E.; Liang, Z. Detection of Apple Lesions in Orchards Based on Deep Learning Methods of CycleGAN and YOLOV3-Dense. *J. Sens.* **2019**, *2019*, 1–13. [[CrossRef](#)]
37. Fan, S.; Li, J.; Zhang, Y.; Tian, X.; Wang, Q.; He, X.; Zhang, C.; Huang, W. On line detection of defective apples using computer vision system combined with deep learning methods. *J. Food Eng.* **2020**, *286*, 110102. [[CrossRef](#)]
38. Dong, C.; Zhang, Z.; Yue, J.; Zhou, L. Classification of strawberry diseases and pests by improved AlexNet deep learning networks. In Proceedings of the 2021 13th International Conference on Advanced Computational Intelligence (ICACI), Wanzhou, China, 14–16 May 2021; pp. 359–364. [[CrossRef](#)]
39. Dong, C.; Zhang, Z.; Yue, J.; Zhou, L. Automatic recognition of strawberry diseases and pests using convolutional neural network. *Smart Agric. Technol.* **2021**, *1*, 100009. [[CrossRef](#)]
40. Siedliska, A.; Baranowski, P.; Zubik, M.; Mazurek, W.; Sosnowska, B. Detection of fungal infections in strawberry fruit by VNIR/SWIR hyperspectral imaging. *Postharvest Biol. Technol.* **2018**, *139*, 115–126. [[CrossRef](#)]
41. Habib, T.; Majumder, A.; Jakaria, A.; Akter, M.; Uddin, M.S.; Ahmed, F. Machine vision based papaya disease recognition. *J. King Saud Univ.-Comput. Inf. Sci.* **2020**, *32*, 300–309. [[CrossRef](#)]
42. Andrushia, A.D.; Patricia, A.T. Artificial bee colony based feature selection for automatic skin disease identification of mango fruit. In *Nature Inspired Optimization Techniques for Image Processing Applications*; Springer: Berlin/Heidelberg, Germany, 2019; Volume 150. [[CrossRef](#)]
43. Naranjo-Torres, J.; Mora, M.; Hernández-García, R.; Barrientos, R.J.; Fredes, C.; Valenzuela, A. A Review of Convolutional Neural Network Applied to Fruit Image Processing. *Appl. Sci.* **2020**, *10*, 3443. [[CrossRef](#)]
44. Wani, J.A.; Sharma, S.; Muzamil, M.; Ahmed, S.; Sharma, S.; Singh, S. Machine Learning and Deep Learning Based Computational Techniques in Automatic Agricultural Diseases Detection: Methodologies, Applications, and Challenges. *Arch. Comput. Methods Eng.* **2021**, *29*, 641–677. [[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.