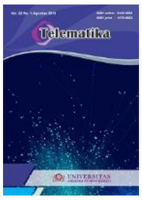
Telematika – Vol. 18, No. 1, February (2025) pp. 45-61 e-ISSN 2442-4528 | p-ISSN 1979-925X



Available online at :
http://ejournal.amikompurwokerto.ac.id/index.php/telematika/

Telematika

Accreditated SINTA "2" Kemenristek/BRIN, No. 85/M/KPT/2020

UNIVERSITAS
AMIKOM PURWOKERTO

Guava Disease Detection and Classification: A Systematic Literature Review

Muhammad Bayu Kurniawan[1*], Ema Utami[2]

[1,2] Department Master of Informatics,
Universitas Amikom Yogyakarta, Yogyakarta, Indonesia
E-mail: bayu.x@students.amikom.ac.id[1], ema.u@amikom.ac.id[2]

ARTICLE INFO

History of the article:
Received April 25, 2024
Revised February 28, 2025
Accepted February 28, 2025

Keywords:
Machine learning
Deep learning
Detection
Classification
Guava diseases

Correspondece:
E-mail:
bayu.x@students.amikom.ac.id

ABSTRACT

Guavas (Psidium guajava) are nutrient-rich fruits that provide significant health benefits. However, guava cultivation faces persistent threats from various diseases affecting both leaves and fruits, leading to substantial yield and quality losses. The early and accurate detection of these diseases is crucial but remains challenging due to economic constraints and limited infrastructure. While plant pathologists employ various diagnostic methods, these approaches are often time-consuming, costly, and sometimes inconsistent. Recent advancements in deep learning (DL) and machine learning (ML) have introduced innovative techniques for guava disease identification. This study conducts a Systematic Literature Review (SLR) to evaluate the existing research on guava leaf and fruit disease detection, focusing on dataset sources, identified disease categories, preprocessing and augmentation techniques, applied algorithms, and reported evaluation metrics. A comprehensive search was conducted across multiple databases, covering publications from 2017 to 2023, leading to the identification of 47 relevant studies. After applying exclusion criteria, 16 studies were selected for in-depth analysis. The findings highlight the most commonly used datasets, the predominant classification techniques, and the effectiveness of various deep learning models based on multiple performance metrics, providing insights into current research trends, existing limitations, and potential directions for future studies. This review serves as a valuable reference for researchers aiming to enhance the accuracy and efficiency of guava leaf and fruit disease diagnosis through data-driven approaches.

INTRODUCTION

Guavas (Psidium guajava) are rich in essential nutrients and widely cultivated due to their adaptability to diverse climates and soils (Angulo-López et al. 2021; Shihab et al. 2025). However, guava production faces significant challenges due to its susceptibility to various diseases, which reduce yield, lower fruit quality, and cause economic losses (Vitti et al. 2020). While pesticides are commonly used for disease control, their excessive and indiscriminate application raises environmental and health concerns. As a result, the need for efficient, sustainable, and early disease detection techniques has become increasingly important in agricultural research. Manual identification methods, although traditionally employed, are often time-consuming, subjective, and prone to human error, particularly in large-scale farming operations (Orchi et al. 2022). This has led researchers to explore more advanced computational approaches, particularly Machine Learning (ML) and Deep Learning (DL), to enhance the precision and efficiency of disease detection in guava plants. Therefore, automated approaches leveraging advanced computational techniques have gained attention as a promising alternative.

https://doi.org/ 10.35671/ telematika.v18i1.2901

Machine Learning (ML) has emerged as a powerful tool for analyzing agricultural data, including disease detection in crops (MacNish et al. 2025). ML models utilize algorithms that learn from data patterns, enabling the classification of diseases based on input features extracted from images of infected guava leaves or fruits. Unlike traditional image processing techniques, ML incorporates loss functions that optimize predictions and enhance classification accuracy over time (Akhtar et al. 2024). Various ML methods, such as decision trees, support vector machines (SVM), and k-nearest neighbors (KNN), have been explored for detecting guava diseases. Researchers have investigated the effectiveness of these techniques in improving early diagnosis and disease classification, with studies conducted by (Almadhor et al. 2021; Almutiry et al. 2021; Gaikwad, Rumma, and Hangarge 2021a; Mumtaz et al. 2023; Perumal et al. 2021), showcasing different ML applications in guava disease detection.

Deep Learning (DL), a subset of ML, provides a more advanced approach to disease detection by leveraging neural networks with multiple layers of processing (Upadhyay et al. 2025). Unlike conventional ML methods that require separate feature extraction and classification steps, DL integrates these processes through hierarchical feature learning. This allows DL models to handle complex, high-dimensional data with greater accuracy. One of the most commonly used DL techniques in agriculture is Convolutional Neural Networks (CNNs), which have been successfully applied to plant disease classification (Ashurov et al. 2024). DL's capability for automatic feature generation, advanced self-learning, and parallel processing makes it particularly suitable for analyzing unstructured image data in guava disease detection. Several studies, such as those by (Gaikwad, Rumma, and Hangarge 2021b; Mostafa et al. 2022; Mustak Un Nobi et al. 2023; Nandi et al. 2023; Rashid et al. 2023), have explored DL approaches for guava disease detection.

Despite the promising advancements in ML and DL for guava disease detection, several challenges remain. One of the key issues is the availability of high-quality datasets, as the accuracy of ML/DL models heavily depends on the diversity and representativeness of training data (Chiou et al. 2025). Additionally, different studies employ various preprocessing and data augmentation techniques, which may lead to inconsistencies in model performance evaluation. Another challenge is the selection of appropriate algorithms, as different ML/DL models exhibit varying degrees of effectiveness depending on factors such as dataset size, image quality, and computational resources. Furthermore, there is a lack of standardized evaluation metrics, making it difficult to compare the performance of different approaches. Overcoming these challenges is key to developing robust and scalable guava disease detection models.

To establish a comprehensive understanding of guava disease detection methodologies, this study conducts a Systematic Literature Review (SLR) focusing on ML and DL applications in this domain. This review aims to address key research questions (RQs) related to dataset sources, disease classification labels, preprocessing techniques, data augmentation strategies, detection and classification algorithms, and evaluation metrics. Specifically, it investigates the different methodologies employed in recent research, analyzes their effectiveness, and identifies challenges and trends in ML and DL applications for guava disease detection. By synthesizing existing studies, this review provides insights into the strengths and limitations of current approaches while offering guidance for future research in precision agriculture.

The present paper is organized into five distinct sections. The introduction is given in Section 1. The sections that follow are arranged as follows: Section 2 provides an overview of the research methodology employed to identify and analyze existing research, encompassing the criteria and research questions utilized. Section 3 presents a review of previous studies literature. Observations and comparisons are provided in Section 4. Ultimately, the paper's conclusion is presented in Section 5.

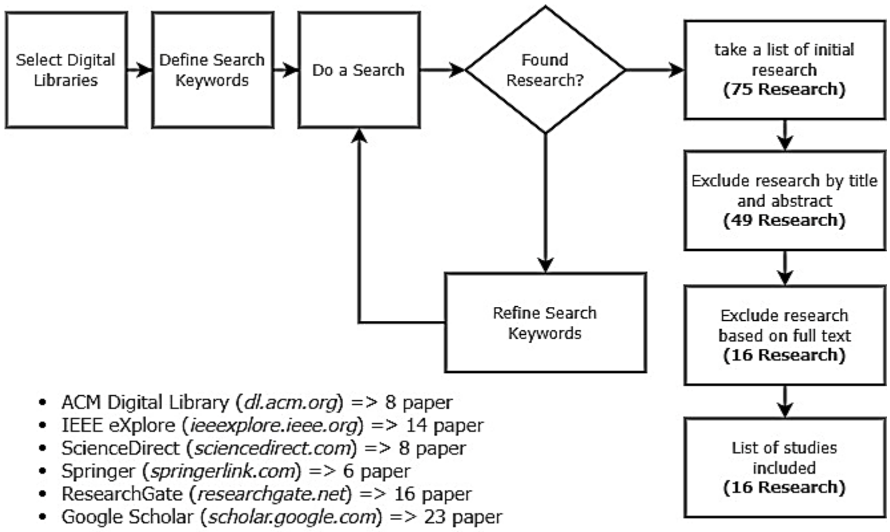


Figure 1. Selection process

RESEARCH METHODS

SLR is a structured methodology used to collect, analyze, and synthesize existing research to provide comprehensive insights into a specific field (Abbasi et al. 2022; Muhammad et al. 2024). This SLR systematically examines and synthesizes recent research on ML and DL for guava disease detection and classification. It explores their applications, identifies trends in algorithm selection, preprocessing techniques, and evaluation metrics, and highlights key challenges such as dataset limitations and performance inconsistencies. By providing a structured analysis of existing methodologies, this review offers insights into current research gaps and future advancements.

1. Research Question

The primary objective of this SLR is to examine the methodological approaches used in past studies on guava disease identification and classification. This review specifically explores data sources, disease classification labels, preprocessing techniques, data augmentation methods, detection and classification algorithms, and evaluation metrics. To maintain a clear focus on these objectives, the research questions presented in Table I were formulated.

Table 1. Research questions

ID	*Research Question*	*Motivation*
RQ1	What are the main sources for collecting datasets on guava diseases?	To identify the data sources used by researchers for guava disease image acquisition.
RQ2	What classes or labels of guava diseases have been identified in previous studies?	To determine the disease categories used in guava disease detection and classification.
RQ3	What preprocessing techniques are applied in studies?	To analyze the preprocessing methods used for enhancing image quality and feature extraction.
RQ4	What data augmentation techniques are employed in guava disease detection?	To examine the strategies used to expand datasets and improve model generalization.
RQ5	What algorithms are utilized for guava disease detection and classification?	To identify the ML and DL techniques applied in studies.
RQ6	What is the accuracy of existing guava disease detection and classification approaches?	To identify the accuracy of current methods in the detection and classification of guava diseases.

2. Search Strategy

We conducted a comprehensive search across various academic platforms, like the ACM Digital Library, IEEE eXplore, ScienceDirect, Springer, ResearchGate, and Google Scholar, using English language. The search process began with keyword-based queries in scientific repositories for relevant journal and conference papers. Specifically, we strategically combined terms such as “guava disease identification" or "guava disease detection" or "guava disease classification" or "guava disease" with keywords like "artificial intelligence" or "image processing" or "machine learning" or "deep learning" or "CNN". Our goal was to systematically gather academic literature related to guava disease detection and classification. This approach encompassed various research efforts and methodologies, providing comprehensive insights into the processes involved in addressing guava diseases. Figure 1 illustrates the search methodology and the number of studies identified at each stage.

3. Study Selection

During the paper screening process, it is crucial to establish clear screening criteria. As described by Shaffril et al. (2021), two types of criteria are commonly employed: inclusion criteria and exclusion criteria. Inclusion criteria outline the conditions that warrant a research paper's inclusion in the review, whereas exclusion criteria specify conditions that render a research paper ineligible for inclusion. The selection criteria outlined in Table 2 provide a structured framework for evaluating studies, offering a comprehensive overview of the factors considered during the planning phase of this SLR.

Table 2. Inclusion and exclusion criteria for selecting studies

ID	*Criteria*		*Description*
	Inclusion	*Exclusion*	
C1	X		The collected research must have been published between 2017 and December 2023.
C2	X		Papers must be written in English.
C3	X		The research topic must focus on detecting and classifying guava diseases.
C4		X	Research that doesn't meet the inclusion criteria.
C5		X	Studies without the full text available.
C6		X	That doesn't clearly outline the methodology used.
C7		X	That fails to address the research objectives.
C8		X	Duplicate publications from multiple sources.

4. Quality Assessment

A rigorous quality assessment of the reviewed papers is essential for ensuring research credibility (Maldonado-Canca et al. 2024; Reda et al. 2023). This study prioritizes high-quality papers published in internationally recognized journals and conferences. The assessment criteria include journal impact factors, citation rates, and indexing in reputable databases such as Scopus and Web of Science. Additionally, the relevance and methodological rigor of each study were considered to ensure reliability in synthesized findings. However, the limited availability of literature on guava disease detection and classification posed challenges in conducting a comprehensive quality assessment.

5. Data Extraction

During this phase, relevant information was extracted from the reviewed papers, including publication year, data sources, disease classification labels, preprocessing techniques, data augmentation techniques, detection and classification algorithms, and performance evaluation metrics. The extracted data was systematically compiled into a spreadsheet for further analysis.

6. Data Synthesis

This critical phase involved a rigorous evaluation of selected studies to ensure they effectively addressed the research questions while meeting the predetermined inclusion and exclusion criteria. Starting with an initial collection of 47 papers, a multi-step selection process was conducted, incorporating targeted research questions, specific inclusion criteria, and strategic filtering. Through this thorough assessment, a final set of 16 papers was identified, meeting stringent methodological standards. These selected studies underwent an in-depth analysis to verify their relevance, methodological rigor, and contribution to the overarching research objectives.

RESULTS AND DISCUSSION

The analysis of the 16 selected papers is categorized by year to illustrate trends in guava disease detection and classification. Figure 2 presents the distribution of research over different years.

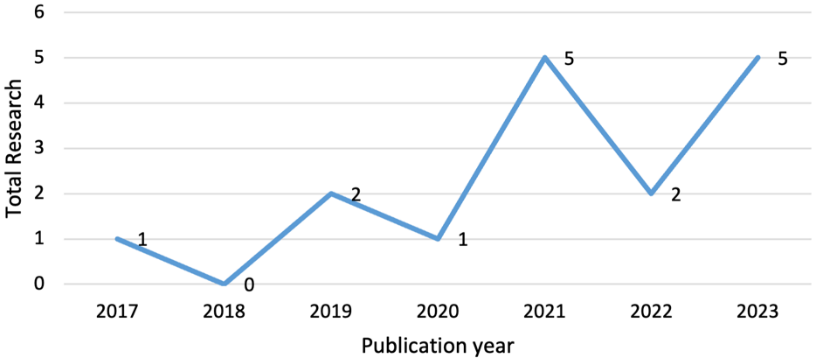


Figure 2. Research grouping by year of publication

1. RQ1: What Are the Main Sources for Collecting Datasets on Guava Diseases?

A dataset consists of structured information used for training and evaluating ML/DL models (Joseph et al. 2024; Moupojou et al. 2023). In guava disease detection, datasets typically consist of labeled images of healthy and diseased guava leaves and fruit, obtained either through manual collection using camera devices or from publicly available repositories.

The SLR identifies four primary sources used for collecting datasets on guava diseases: internet-based repositories, direct field observations, publicly available research databases, and data-sharing platforms. Several studies, such as those by (Farhan Al Haque et al. 2019; Perumal et al. 2021; S.Abirami 2017), utilize datasets obtained from the internet. This method provides a vast range of images, offering accessibility to a variety of guava disease samples. However, the quality and consistency of internet-sourced datasets can be inconsistent due to issues related to labeling accuracy, variations in image quality, and lack of standardization. In contrast, a significant portion of studies, including those by (Almutiry et al. 2021; Al Haque et al. 2019; Gaikwad et al. 2021a; Howlader et al. 2019; Mostafa et al. 2022; Rashid et al. 2023; Shakil et al. 2023; and Srinivas et al. 2021), relies on direct field observations. This method ensures high-quality and context-specific data collection, as images are captured under controlled conditions, reflecting real-world disease occurrences. However, field data collection is time-intensive and often limited by geographical constraints, requiring significant resources and expertise.

Additionally, some studies, such as those by (Almadhor et al. 2021; Mumtaz et al. 2023; Nobi et al. 2023; and Nandi et al. 2023), utilize datasets from Mendeley Data, a widely used research data-sharing platform. Mendeley provides structured and curated datasets contributed by researchers, which enhances data reproducibility and facilitates comparative analyses. However, reliance on existing datasets may introduce biases, as the data may not fully represent the diversity of guava diseases across different regions and climatic conditions. Lastly, a limited number of studies, such as Doutoum et al. (2023), employ datasets from Kaggle, a popular data science platform hosting a variety of publicly available datasets. Kaggle datasets often come preprocessed and well-structured, offering advantages in terms of usability and computational efficiency. However, the availability of guava disease-specific datasets on Kaggle is relatively limited, potentially restricting the scope of model training and validation. Table 3 provides a concise overview of the various sources used for gathering data.

Table 3. Overview of the data collection source

No	*Reference*	*Source*
1	(Haque et al. 2019; Perumal et al. 2021; Abirami 2017).	Internet
2	(Almutiry et al. 2021; Haque et al. 2019; Gaikwad et al. 2021b, 2021a; Howlader et al. 2019; Mostafa et al. 2022; Rashid et al. 2023; Shakil et al. 2023; Srinivas et al. 2021).	Observation
3	(Almadhor et al. 2021; Mumtaz et al. 2023; Nobi et al. 2023; Nandi et al. 2023).	Mendeley Data
4	(Doutoum et al. 2023).	Kaggle

This observation is derived solely from discussions related to RQ1: 23% of the reviewed studies sourced image data from Mendeley Data, and 6% obtained image data from a publicly accessible Kaggle. Meanwhile, 53% of studies utilized smartphones, drones, or similar devices to capture images from real-world environments, and 18% acquired their images from the internet. Figure 3 shows all of the data acquisition sources.

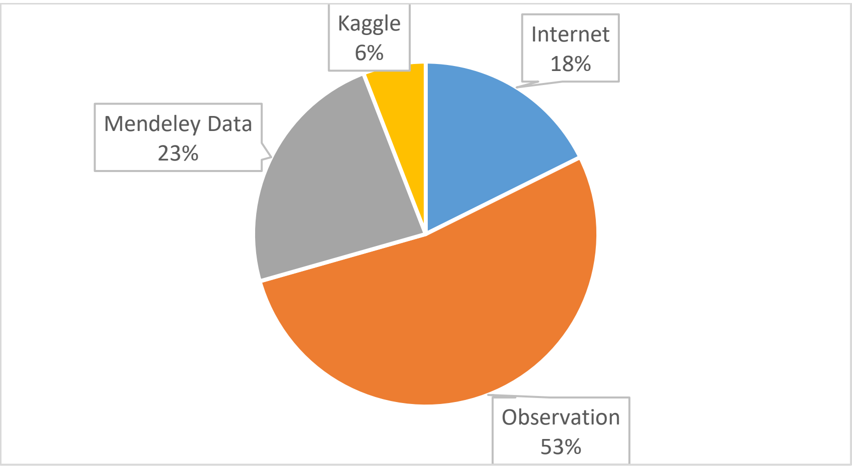


Figure 3. Utilization of different sources for acquiring data

2. RQ2: What Classes or Labels of Guava Diseases Have Been Identified in Previous Studies?

Dataset labeling is crucial for ML and DL models, enabling accurate pattern recognition and decision-making (El-Hasnony et al. 2022; Meng et al. 2020). The classification of guava diseases is essential for accurate detection and diagnosis. Research has categorized these diseases based on symptoms observed on leaves and fruits, typically distinguishing between healthy samples and those affected by fungal, bacterial, or physiological disorders. The SLR reveals varied disease labels, with some datasets spanning multiple categories and others targeting specific conditions.

Research conducted by Abirami (2017) applied KNN and SVM for diagnosing five disease classes in guava leaves: viburnum chindo, rust, algal leaf spot, powdery mildew, and curl. Similarly, a study by Howlader et al. (2019) proposed a DCNN model to classify healthy leaves, algal leaf spot, rust, and whitefly. Another study, conducted by Haque et al. (2019), employed a CNN model to identify four different disease classes: benign guava, fruit rot, fruit canker, and anthracnose. Meanwhile, Srinivas et al. (2021) utilized a DL method to identify wilt, rust, anthracnose, algal leaf spot, and black rot in guava leaves. Further research by Almadhor et al. (2021) explored machine learning methods for recognizing five categories of guava leaf conditions: normal, rust, canker, mummification, and dot. A study by Gaikwad et al. (2021a, 2021b) introduced a CNN model for identifying healthy leaves, insect-damaged leaves, pseudocercospora leaf spot, and rust. Similarly, Almutiry et al. (2021) suggested a multi-class classification approach for four types of guava diseases: fruit fly, anthracnose, algal spot, and styler end rot. Research by Perumal et al. (2021) applied SVM to classify healthy leaves, anthracnose, and bacterial blight.

Several other studies focused on DL methods for guava disease detection and classification. For instance, research by Mostafa et al. (2022), Doutoum et al. (2023), and Nobi et al. (2023) proposed models to classify guava conditions such as healthy, canker, mummification, dot, and rust. A study by Nandi et al. (2023) developed a methodology for identifying scab, red rust, phytophthora, disease-free fruit, disease-free leaf, and styler end rot in guava leaves and fruits. Similarly, Mumtaz et al. (2023) presented methods for detecting healthy leaves and leaf blight in guava. Research by Rashid et al. (2023) identified five distinct guava conditions: healthy, anthracnose, insect attack, nutrient deficiency, and wilt.Additionally, a study by Nobi et al. (2023) proposed a classification method for guava leaf diseases, including phytophthora, red rust, disease-free leaves, styler end rot, and scab. Furthermore, Shakil et al. (2023) identified five disease categories: phytophthora, fresh leaves, styler end rot, red rust, and scab, using guava leaf images.

Several studies have used slightly different terminologies to describe similar disease conditions. For instance, "Disease-Free Leaf" and "Fresh Leaf" both indicate healthy guava leaves, while "Insect Eaten" and "Insect Attack" refer to damage caused by insect feeding. Similarly, "Leaf Spot" is a general term, while "Algal Leaf Spot" and "Pseudocercospora Leaf Spot" specify infections caused by algae and fungi, respectively. These variations arise due to differences in dataset labeling approaches across studies. Table 4 provides a condensed summary of the guava disease categories identified and categorized in the reviewed studies.

Table 4. List of classes in the reviewed studies

No	*Reference*	*Classes*
1	(Abirami 2017)	Curl, powdery mildew, rust, viburnum chindo, algal leaf spot
2	(Howlader et al. 2019)	Healthy, algal leaf spot, rust, whitefly
3	(Haque et al. 2019)	Fruit canker, fruit rot, anthracnose, benign guava
4	(Srinivas et al. 2021)	Algal leaf spot, black rot, anthracnose, wilt, rust
5	(Gaikwad et al. 2021b)	Healthy, insect eaten, pseudocercospora leaf spot, rust
6	(Almadhor et al. 2021)	Normal/healthy, rust, canker, mummification, dot
7	(Almutiry et al. 2021)	Anthracnose, algal spot, styler end rot, fruit fly
8	(Perumal et al. 2021)	Healthy, anthracnose, bacterial blight
9	(Gaikwad et al. 2021a)	Healthy, insect eaten, leaf spot, rust
10	(Mostafa et al. 2022)	Healthy, rust, canker, mummification, dot
11	(Nandi et al. 2023)	Red rust, disease-free leaf, phytophthora, scab, disease-free fruit, styler end rot
12	(Mumtaz et al. 2023)	Healthy, leaf blight
13	(Doutoum et al. 2023)	Healthy, rust, canker, mummification, dot
14	(Rashid et al. 2023)	Healthy, anthracnose, insect attack, nutrient deficiency, wilt
15	(Nobi et al. 2023)	D1: healthy, canker, mummification, dot, rust D2: phytophthora, styler end rot, scab, disease-free, red rust
16	(Shakil et al. 2023)	Phytophthora, scab, fresh leaf, styler end rot, red rust

This observation is solely based on the analysis conducted regarding RQ2. As shown in Figure 4, the most frequently detected or classified categories are Healthy (13 evaluated studies) and Rust (9 evaluated studies), followed by Anthracnose (5 evaluated studies). Several diseases appear with the same frequency, including Canker, Mummification, Dot, and Styler End Rot (each appearing in 4 studies). Next are Algal Leaf Spot, Red Rust, Scab, Phytophthora, and Insect Eaten, each reported in 3 studies, followed by Wilt and Leaf Spot, which appear in 2 studies. Additionally, several conditions were detected only once across the reviewed studies. These include Curl, Powdery Mildew, Viburnum Chindo, Whitefly, Fruit Rot, Fruit Canker, Black Rot, Algal Spot, Fruit Fly, Bacterial Blight, Leaf

Blight, and Nutrient Deficiency. Some studies also identified Benign Guava (2 studies) and Disease-Free Leaves/Fruit, though inconsistencies in terminology were observed.

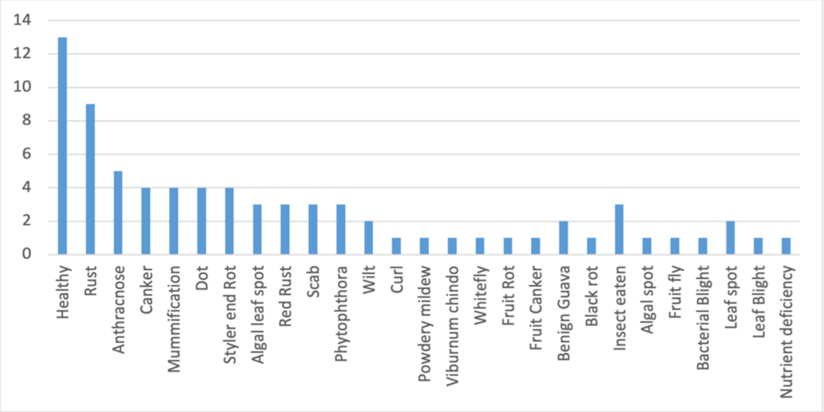


Figure 4. The number of various classes detected

3. RQ3: What Preprocessing Techniques Are Applied in Studies?

Preprocessing techniques are essential in image-based guava disease detection, as they enhance image quality, remove noise, and standardize inputs for machine learning models (Ojo and Zahid 2023; Ramadhan et al. 2024; Sarki et al. 2021). The SLR reveals that resizing is the most commonly applied preprocessing technique, appearing in seven out of nine reviewed studies. This suggests a strong emphasis on ensuring uniform image dimensions, which is crucial for deep learning models that require consistent input sizes.

Beyond resizing, contrast enhancement and histogram equalization improve image clarity, making disease symptoms more distinguishable. Studies such as Abirami (2017) and Perumal et al. (2021) apply contrast enhancement, while Mostafa et al. (2022) uses histogram equalization and unsharp masking to refine image sharpness. These techniques help highlight disease features that may otherwise be less visible in raw images. Noise reduction, implemented in Haque et al. (2019) and Doutoum et al. (2023), removes unwanted artifacts, enhancing classification accuracy. Similarly, Almadhor et al. (2021) addresses distortion reduction, an often-overlooked step that can enhance the quality of guava leaf and fruit images.

Some studies use color-space transformations for feature extraction. Mumtaz et al. (2023) applies RGB-to-YCbCr conversion to separate chrominance and luminance, aiding segmentation. Meanwhile, Almutiry et al. (2021) and Nobi et al. (2023) integrate normalization and thresholding, standardizing pixel intensity values across datasets to improve model generalization. Interestingly, binarization is only reported in Almutiry et al. (2021), suggesting that this technique is less favored in guava disease classification compared to other agricultural image-processing domains. The limited adoption of binarization may indicate that more complex preprocessing methods, such as differential shading improvement (Doutoum et al. (2023) and unsharp masking (Mostafa et al. (2022), are preferred for retaining critical disease features.

Table 5. Summary of different preprocessing techniques

No	*Reference*	*Operation Performed*
1	(Abirami 2017) (Perumal et al. 2021)	Resizing, contrast enhancement
2	(Gaikwad et al. 2021b; Howlader et al. 2019; Nandi et al. 2023; Rashid et al. 2023; Shakil et al. 2023)	Resizing
3	(Farhan Al Haque et al. 2019)	Cropping, noise removal
4	(Almadhor et al. 2021)	Resizing, image enhancement, reducing distortions
5	(Almutiry et al. 2021)	Resizing, thresholding, normalization, binarization

No	*Reference*	*Operation Performed*
6	(Mostafa et al. 2022)	Resizing, histogram equalization, unsharp masking
7	(Mumtaz et al. 2023)	RGB-to-YCbCr conversion
8	(Doutoum et al. 2023)	Resizing, noise reduction, differential shading improvement
9	(Nobi et al. 2023)	Resizing, normalization

This observation is based only on the talks in RQ3. Figure 5 illustrates the application of different preprocessing techniques. Resizing was the most frequently used technique, appearing in 7 studies. Normalization and contrast enhancement were applied in 2 studies each. Other techniques, including cropping, noise removal, image enhancement, distortion reduction, thresholding, binarization, histogram equalization, unsharp masking, RGB-to-YCbCr conversion, noise reduction, and differential shading improvement, were each used in 1 study.

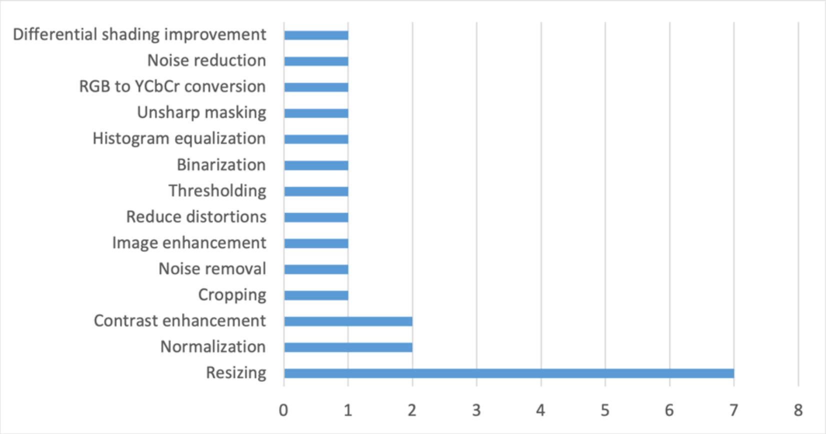


Figure 5. Graphical representation of various preprocessing techniques

4. RQ4: What Data Augmentation Techniques Are Employed in Guava Disease Detection?

Data augmentation is widely used in guava disease detection to enhance dataset diversity and improve model generalization (Liang et al. 2023; Hao et al. 2023). The SLR reveals that rotation and flipping are the most frequently applied techniques, appearing in five of the reviewed studies. These methods mitigate class imbalance and enhance robustness by introducing image orientation variations.

Several studies implement complex transformations beyond basic geometric modifications. Haque et al. (2019) employs shifting, shearing, re-scaling, and zooming, which artificially expand the dataset while preserving structural integrity. Similarly, Rashid et al. (2023) integrates random brightness adjustments and shear transformations, enhancing the model's ability to recognize diseases under different lighting conditions. Interestingly, affine transformation is only reported in Mostafa et al. (2022), suggesting limited adoption despite its effectiveness in maintaining geometric consistency.

Other methods, such as nearest fill interpolation (Haque et al. (2019) and random shear transformations (Rashid et al. (2023)), indicate a trend toward more adaptive augmentation strategies. Notably, Nandi et al. (2023) and Doutoum et al. (2023) do not specify any augmentation techniques, highlighting potential gaps in preprocessing approaches within some studies. Table 6 provides an overview of the different techniques of data augmentation utilized in various studies under evaluation.

Table 6. Overview of various data augmentation techniques

No	*Reference*	*Augmentation Operation*
1	(Farhan Al Haque et al. 2019)	Shifting, flipping, nearest fill, rotation, shearing, re-scaling, zooming
2	(Gaikwad et al. 2021b)	Rotation, tilted, and zooming
3	(Mostafa et al. 2022)	Affine transformation
4	(Mumtaz et al. 2023)	Flipping

No	*Reference*	*Augmentation Operation*
5	(Rashid et al. 2023)	Flipping, rotation (90-degree rotation, random rotation degree), random shear degree, and random brightness
6	(Nobi et al. 2023)	Shifting, rotation, flipping
7	(Shakil et al. 2023)	Rotation
8	(Nandi et al. 2023) (Doutoum et al. 2023)	-

This observation is based only on the talks in RQ4. Among the 16 reviewed studies, 9 implemented data augmentation techniques. Figure 6 illustrates the distribution of these techniques. The results indicate that rotation was the most frequently applied technique, appearing in 5 studies, followed by flipping in 4 studies. Shifting and zooming were each used in 2 studies, while shearing, random shear degree, affine transformation, random brightness, nearest fill, tilted, and re-scaling were each reported in 1 study.

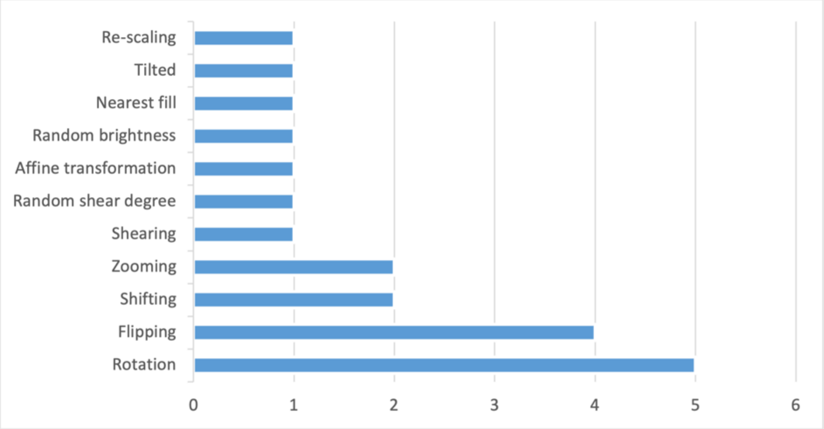


Figure 6. Utilization of various augmentation techniques

5. RQ5: What Algorithms Are Utilized for Guava Disease Detection and Classification?

The SLR reveals a diverse range of algorithms used for guava disease detection and classification. SVM and CNN are among the most frequently employed methods, highlighting their effectiveness in image-based classification tasks (Valero-Carreras et al. 2023). Among machine learning methods, SVM is the most frequently used, appearing in four studies (Gaikwad et al. 2021a; Mumtaz et al. 2023; Perumal et al. 2021; S.Abirami 2017), reaffirming its effectiveness in handling small to medium-sized datasets. KNN also remains a popular choice, applied in three studies, while ensemble techniques like Bagged Tree and Cubic SVM (Almadhor et al. 2021; Almutiry et al. 2021) indicate attempts to enhance classification accuracy through multiple decision boundaries.

Deep learning methods dominate, with CNN-based architectures leading the field. ResNet50 was utilized in three studies (Doutoum et al. 2023; Mostafa et al. 2022; Shakil et al. 2023), followed by AlexNet and CNN, both appearing in three studies. More advanced architectures such as InceptionResNetV2, EfficientNet, and Xception were explored in select studies (Doutoum et al. 2023; Nobi et al. 2023; Shakil et al. 2023), highlighting a shift toward deeper networks for enhanced feature extraction. Beyond mainstream architectures, a few studies introduced specialized deep learning frameworks tailored for guava disease classification. Rashid et al. (2023) proposed GIP-MU-Net, GLSM, and GMLDD, while Nobi et al. (2023) introduced GLD-Det, indicating a move toward custom-designed models fine-tuned for agricultural disease detection. Table 7 provides a concise summary of the varied detection or classification methods utilized in guava disease research.

Table 7. Overview of various algorithm methods

No	*Reference*	*Algorithm*
1	(Mumtaz et al. 2023; Abirami 2017)	SVM, KNN
2	(Howlader et al. 2019)	DCNN
3	(Al Haque et al. 2019; Srinivas et al. 2021)	CNN
4	(Gaikwad et al. 2021b)	CNN, AlexNet, SqueezeNet
5	(Almadhor et al. 2021)	Bagged Tree, Fine KNN, Boosted Tree, Cubic SVM, Complex Tree
6	(Almutiry et al. 2021)	Fine KNN, RUSBoosted Tree, Bagged Tree, Cubic SVM, W-KNN
7	(Perumal et al. 2021)	SVM
8	(Gaikwad et al. 2021a)	AlexNet, KNN, LDA, SVM
9	(Mostafa et al. 2022)	AlexNet, SqueezeNet, GoogleNet, ResNet50, ResNet101
10	(Nandi et al. 2023)	VGG16, GoogleNet, ResNet18, MobileNetV2, EfficientNetB2
11	(Doutoum et al. 2023)	VGG16, ResNet50, InceptionV3, EfficientNetB3
12	(Rashid et al. 2023)	GIP-MU-Net, GLSM, GMLDD
13	(Nobi et al. 2023)	GLD-Det, EfficientNetV2B2, EfficientNetB0, EfficientNetB2, EfficientNetB1, MobileNetV2
14	(Shakil et al. 2023)	InceptionResNetV2, ResNet50, Xception

This observation is solely based on the RQ5. Figure 7 illustrates the various approaches used for guava disease detection and classification across the reviewed studies. SVM was the most frequently utilized, appearing in four studies, followed by CNN, KNN, AlexNet, and ResNet50, each used in three studies. VGG16, Fine KNN, GoogleNet, SqueezeNet, MobileNetV2, EfficientNetB2, Bagged Tree, and Cubic SVM were each applied in two studies. Several other methods, including DCNN, Complex Tree, Boosted Tree, W-KNN, RUSBoosted Tree, LDA, ResNet18, ResNet101, GIP-MU-Net, InceptionResNetV2, InceptionV3, GMLDD, GLD-Det, EfficientNetV2B2, EfficientNetB0, GLSM, EfficientNetB3, Xception, and EfficientNetB1, were each employed in one study.

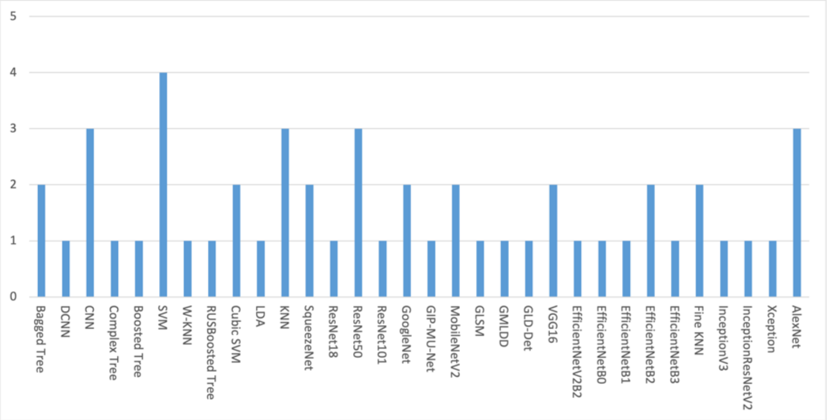


Figure 7. Various techniques utilized for detection and classification

6. RQ6: What Is the Accuracy of Existing Guava Disease Detection and Classification Approaches?

The reviewed studies report varying levels of detection and classification accuracy, with accuracy being the most commonly reported metric. The findings indicate that deep learning approaches generally outperform traditional machine learning models, with some achieving near-perfect accuracy. Reported accuracies range from 71% to 99.54%, demonstrating significant variability depending on the algorithm, dataset quality, and preprocessing techniques.

Among machine learning techniques, SVM consistently delivers high accuracy, making it one of the most reliable classifiers in guava disease detection. Abirami (2017) reported 97.2% accuracy, while Perumal et al. (2021) and Mumtaz et al. (2023) achieved 98.17% and 99.2%, respectively. This performance suggests that SVM remains a strong contender in cases where computational efficiency is a priority. Other machine learning models, such as Bagged Tree, also demonstrated strong results, with Almadhor et al. (2021) achieving 99% accuracy. Similarly, Cubic SVM performed well, reaching

97.52% (Almutiry et al. (2021). These findings indicate that traditional machine learning models can still achieve competitive accuracy when optimized appropriately.

Deep learning architectures dominate the top-performing models, with ResNet50 achieving the highest reported accuracy of 99.54% (Mostafa et al. (2022)). This suggests that ResNet-based architectures effectively capture complex features in guava disease classification. Other models demonstrating superior performance include EfficientNetB2 (99%) and Xception (98.88%), reinforcing the effectiveness of these architectures in handling large-scale image classification tasks. Certain studies also explored customized deep learning architectures, such as GIP-MU-Net (92.41%) (Rashid et al. (2023) and GLD-Det (98.06%) (Nobi et al. (2023), highlighting efforts to tailor models for agricultural disease detection. These specialized models may offer advantages in domain-specific optimizations.

Despite the dominance of deep learning, performance discrepancies exist among CNN architectures. While CNN-based models like AlexNet, ResNet, and EfficientNet consistently achieved high accuracy, some architectures, such as SqueezeNet (75.9%) and CNN (71%) (Srinivas et al. (2021)), exhibited significantly lower accuracy. These variations suggest that model choice, dataset quality, and preprocessing steps significantly impact performance. Table 8 showcases the highest accuracy attained across the evaluated studies.

Table 8. Detection and classification accuracy of the evaluated studies

No	*Reference*	*Algoritm*	*Best Accuracy*
1	(S.Abirami 2017)	SVM, KNN	SVM (97.2%)
2	(Howlader et al. 2019)	DCNN	98.74%
3	(Farhan Al Haque et al. 2019)	CNN	95.61%
4	(Srinivas et al. 2021)	CNN	71%
5	(Gaikwad et al. 2021b)	CNN, AlexNet, SqueezeNet	SqueezeNet (75.9%).
6	(Almadhor et al. 2021)	Bagged Tree, Fine KNN, Boosted Tree, Cubic SVM, Complex Tree	Bagged Tree (99%)
7	(Almutiry et al. 2021)	Fine KNN, RUSBoosted Tree, Bagged Tree, Cubic SVM, W-KNN	Cubic SVM (97.52%)
8	(Perumal et al. 2021)	SVM	98.17%
9	(Gaikwad et al. 2021a)	AlexNet, KNN, LDA, SVM	AlexNet (72.8%)
10	(Mostafa et al. 2022)	GoogleNet, SqueezeNet, AlexNet, ResNet50, ResNet101	ResNet-50 (99.54%)
11	(Nandi et al. 2023)	VGG16, GoogleNet, ResNet18, MobileNetV2, EfficientNetB2	EfficeintNet-b2 (99%)
12	(Mumtaz et al. 2023)	SVM, KNN	SVM (99.2%)
13	(Doutoum et al. 2023)	VGG16, ResNet50, Inception V3, EfficientNetB3	EfficientNet-b3 (94.93%)
14	(Rashid et al. 2023)	GIP-MU-Net, GLSM, GMLDD	GIP-MU-Net (92.41%)
15	(Nobi et al. 2023)	GLD-Det, EfficientNetV2B2, EfficientNetB0, EfficientNetB2, EfficientNetB1, MobileNetV2	D1: GLD-Det (98.06%) D2: GLD-Det (97.14%)
16	(Shakil et al. 2023)	InceptionResNetV2, ResNet50, Xception	Xception (98.88%)

This observation is solely based on the discussion in RQ6. Figure 8 compares the accuracy levels of existing models across three categories: ≤80%, 81–95%, and >95%, showcasing the majority with 70% of evaluated studies achieving guava disease diagnosis accuracy of over 95%. Meanwhile, 12% exhibited accuracy between 81 and 95%, with 18% achieving accuracy levels of 80% or less.

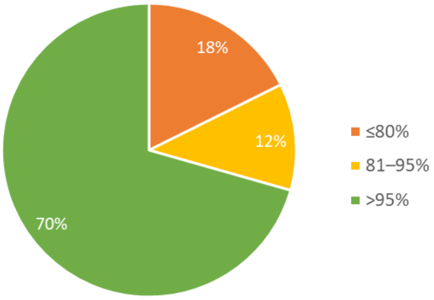


Figure 8. Classification and detection accuracies of the evaluated studies

OVERALL OBSERVATION AND COMPARISON

1. Overall Observation

The analysis conducted on the reviewed studies reveals several noteworthy observations regarding the methodologies employed in guava disease detection and classification. Observation 1 highlights that the image data utilized in these studies predominantly originated from real-world observations. Secondly, Observation 2 shows that the classes or labels most researched are healthy and rust. Thirdly, Observation 3 suggests that resizing was frequently employed for image pre-processing, while Observation 4 notes that rotation operations were commonly employed during data augmentation to increase the dataset size in the reviewed studies. Observation 5 reveals the prevalent utilization of SVM for classifying guava diseases in the reviewed publications. Despite SVM's popularity, CNN architectures based on transfer learning emerge as the most extensively employed approach across the studies. Observation 6 highlights that the majority of the research endeavors have yielded accuracy levels exceeding 95%.

2. Comparison

Table 9 presents a comparison of different research papers based on the data sources, disease classes, preprocessing techniques, data augmentation techniques, algorithms used, and their respective best accuracy outcomes.

Table 9. Comparison of various reviewed papers

No	*Reference*	*Data Source*	*Class or Label*	*Preprocessing*	*Augmentation*	*Algoritm*	*Best Accuracy*
1	(S.Abirami 2017)	Internet	Curl, powdery mildew, rust, viburnum chindo, algal leaf spot	Resizing, contrast enhancement	-	SVM, KNN	SVM (97.2%)
2	(Howlader et al. 2019)	Observation	Healthy, algal leaf spot, rust, whitefly	Resizing	-	DCNN	98.74%
3	(Farhan Al Haque et al. 2019)	Observation, Internet	Anthracnose, fruit rot, fruit canker, bening guava	Cropping, noise removal	Shifting, flipping, nearest fill, zoom rotation, shearing rescaling	CNN	95.61%
4	(Srinivas et al. 2021)	Observation	Anthracnose, algal leaf spot, wilt, rust,black rot	-	-	CNN	71%
5	(Gaikwad et al. 2021b)	Observation	Healthy, insect eaten, pseudocercospora leaf spot, rust	Resizing, image enhancement, reduce distortions	-	CNN	SqueezeNet (75.9%).
6	(Almadhor et al. 2021)	Mendeley Data	Normal, rust, canker, mummification, dot	Resizing	Rotation, tilted, zoom	Bagged Tree, Fine KNN, Boosted Tree, Cubic SVM, Complex Tree	Bagged Tree (99%)
7	(Almutiry et al. 2021)	Observation	Anthracnose, algal spot, styler end rot, fruit fly	Resizing, binarization, thresholding, normalization	-	Cubic SVM, Fine KNN, W-KNN, Bagged Tree, RUSBoosted Tree	Cubic SVM (97.52%)
8	(Perumal et al. 2021)	Internet	Healthy, anthracnose, bacterial blight	Resizing, contrast enhancement	-	SVM	98.17%
9	(Gaikwad et al. 2021a)	Observation	Healthy, insect eaten, leaf spot, rust	-	-	AlexNet, KNN, LDA, SVM	AlexNet (72.8%)
10	(Mostafa et al. 2022)	Observation	Healthy, rust, canker, mummification, dot	Resizing, histogram	Affine transformation	AlexNet, SqueezeNet, GoogleNet,	ResNet50 (99.54%)

No	*Reference*	*Data Source*	*Class or Label*	*Preprocessing*	*Augmentation*	*Algoritm*	*Best Accuracy*
				equalization, unsharp masking		ResNet50, ResNet101	
11	(Nandi et al. 2023)	Mendeley Data	Red rust, disease-free leaf, phytophthora, scab, disease-free fruit, styler end rot	Resizing	-	GoogleNet, ResNet-18, MobileNetV2, VGG-16, EfficientNetB2	EfficeintNetB2 (99%)
12	(Mumtaz et al. 2023)	Mendeley Data	Healthy, leaf blight	RGB-to-YCbCr conversion	Flipping	SVM and KNN	SVM (99.2%)
13	(Doutoum et al. 2023)	Kaggle	Healthy, rust, canker, mummification, dot	Resizing, noise reduction, differential shading improvement	-	VGG-16, ResNet50, Inception V3, EfficientNet-B3	EfficientNet-B3 (94.93%)
14	(Rashid et al. 2023)	Observation	Healthy, anthracnose, insect attack, nutrient deficiency, wilt	Resizing	Flipping, random brightness, rotation, random shear degree	GIP-MU-Net, GLSM, GMLDD	GIP-MU-Net (92.41%)
15	(Nobi et al. 2023)	Mendeley Data	D1: healthy, canker, mummification, dot, rust. D2: disease-free, phytophthora, red rust, scab, styler end rot	Resizing, normalization	Shifting, rotation, flipping	GLD-Det, MobileNetV2, EfficientNetV2B2, EfficientNetB0, EfficientNetB2, EfficientNetB1	D1: GLD-Det (98.06%), D2: GLD-Det (97.14%)
16	(Shakil et al. 2023)	Observation	Fresh leaf, styler end rot, phytophthora, red rust, scab	Resizing	Rotation	InceptionResNetV2, ResNet50, Xception	Xception (98.88%)

CONCLUSIONS

This study reviewed various machine learning and deep learning techniques for detecting and classifying guava diseases, analyzing 16 relevant studies. Our findings indicate that datasets collected directly from real-world environments were the most commonly used, ensuring practical applicability. Among the diseases studied, rust and anthracnose were the most frequently examined, highlighting their significance in guava disease management. In terms of preprocessing, resizing was the most widely applied technique, ensuring uniform input dimensions, while rotation was the most common data augmentation method, helping to enhance dataset diversity. Over the past six years, SVM has remained the most widely used algorithm, yet CNN architectures based on transfer learning have emerged as the dominant approach, demonstrating their effectiveness in feature extraction and classification. Most studies reported accuracy levels exceeding 95%, indicating the high performance of the methodologies employed in guava disease detection. However, a majority of studies relied on images of individual guava leaves, which may limit real-world applicability. Future research could explore multi-leaf image analysis to enhance disease identification accuracy. Additionally, advancements in guava disease detection could focus on staging disease severity, enabling more precise management strategies. Furthermore, integrating detection techniques with drones and mobile applications could facilitate early disease detection in large agricultural fields, improving efficiency and scalability.

REFERENCES

Abbasi, Rabiya, Pablo Martinez, and Rafiq Ahmad. 2022. "The Digitization of Agricultural Industry – a Systematic Literature Review on Agriculture 4.0." *Smart Agricultural Technology* 2(February):100042. doi: 10.1016/j.atech.2022.100042.

Akhtar, Mushir, M. Tanveer, and Mohd Arshad. 2024. “Advancing Supervised Learning with the Wave Loss Function: A Robust and Smooth Approach.” *Pattern Recognition* 155(November 2023). doi: 10.1016/j.patcog.2024.110637.

Almadhor, Ahmad, Hafiz Tayyab Rauf, Muhammad Ikram Ullah Lali, Robertas Damaševičius, Bader Alouffi, and Abdullah Alharbi. 2021. “Ai-Driven Framework for Recognition of Guava Plant Diseases through Machine Learning from Dslr Camera Sensor Based High Resolution Imagery.” *Sensors* 21(11):1–19. doi: 10.3390/s21113830.

Almutiry, Omar, Muhammad Ayaz, Tariq Sadad, Ikram Ullah Lali, Awais Mahmood, Najam Ul Hassan, and Habib Dhahri. 2021. “A Novel Framework for Multi-Classification of Guava Disease.” *Computers, Materials and Continua* 69(2):1915–26. doi: 10.32604/cmc.2021.017702.

Angulo-López, Jorge E., Adriana C. Flores-Gallegos, Cristian Torres-León, Karen N. Ramírez-Guzmán, Gloria A. Martínez, and Cristóbal N. Aguilar. 2021. “Guava (Psidium Guajava l.) Fruit and Valorization of Industrialization by-Products.” *Processes* 9(6):1–17. doi: 10.3390/pr9061075.

Ashurov, Asadulla Y., Mehdhar S. A. M. Al-Gaashani, Nagwan A. Samee, Reem Alkanhel, Ghada Atteia, Hanaa A. Abdallah, and Mohammed Saleh Ali Muthanna. 2024. “Enhancing Plant Disease Detection through Deep Learning: A Depthwise CNN with Squeeze and Excitation Integration and Residual Skip Connections.” *Frontiers in Plant Science* 15(January):1–16. doi: 10.3389/fpls.2024.1505857.

Chiou, Kuo-dung, Yen-xue Chen, Po-sung Chen, Ying-tzy Jou, and Shang-han Tsai. 2025. “Application of Deep Learning for Fruit Defect Recognition in Psidium Guajava L.” *Scientific Reports* 15. doi: 10.1038/s41598-025-88936-y.

Doutoum, Assad S., Recep Eryigit, and Bulent Tugrul. 2023. “Classification of Guava Leaf Disease Using Deep Learning.” *Wseas Transactions On Information Science And Applications* 20(October):356–63. doi: 10.37394/23209.2023.20.38.

El-Hasnony, Ibrahim M., Omar M. Elzeki, Ali Alshehri, and Hanaa Salem. 2022. “Multi-Label Active Learning-Based Machine Learning Model for Heart Disease Prediction.” *Sensors* 22(3). doi: 10.3390/s22031184.

Farhan Al Haque, A. S. M., Rubaiya Hafiz, Md Azizul Hakim, and G. M. Rasiqul Islam. 2019. “A Computer Vision System for Guava Disease Detection and Recommend Curative Solution Using Deep Learning Approach.” *2019 22nd International Conference on Computer and Information Technology, ICCIT 2019* (December). doi: 10.1109/ICCIT48885.2019.9038598.

Gaikwad, Sukanya S., Shivanand S. Rumma, and Mallikarjun Hangarge. 2021a. *Classification of Fungi Effected Psidium Guajava Leaves Using ML and DL Techniques. In Computer Vision and Machine Intelligence Paradigms for SDGs*. Vol. 967. edited by R. J. Kannan, S. M. Thampi, and S.-H. Wang.

Gaikwad, Sukanya S., Shivanand S. Rumma, and Mallikarjun Hangarge. 2021b. “Identification of Fungi Infected Leaf Diseases Using Deep Learning Techniques.” *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 6(12):5618–25.

Howlader, Md Rasel, Umme Habiba, Rahat Hossain Faisal, and Md Mostafijur Rahman. 2019. “Automatic Recognition of Guava Leaf Diseases Using Deep Convolution Neural Network.” *2nd International Conference on Electrical, Computer and Communication Engineering, ECCE 2019* 1–5. doi: 10.1109/ECACE.2019.8679421.

Joseph, Diana Susan, Pranav M. Pawar, and Kaustubh Chakradeo. 2024. “Real-Time Plant Disease Dataset Development and Detection of Plant Disease Using Deep Learning.” *IEEE Access* 12(January):16310–33. doi: 10.1109/ACCESS.2024.3358333.

Liang, Wen, Youzhi Liang, and Jianguo Jia. 2023. “MiAMix: Enhancing Image Classification through a Multi-Stage Augmented Mixed Sample Data Augmentation Method.” *Processes* 11(12):1–19. doi: 10.3390/pr11123284.

MacNish, Tessa R., Monica F. Danilevicz, Philipp E. Bayer, Mitchell S. Bestry, and David Edwards. 2025. “Application of Machine Learning and Genomics for Orphan Crop Improvement.” *Nature Communications* 16(1):982. doi: 10.1038/s41467-025-56330-x.

Maldonado-Canca, Luis Alfonso, Ana María Casado-Molina, Juan Pedro Cabrera-Sánchez, and Guillermo Bermúdez-González. 2024. “Beyond the Post: An SLR of Enterprise Artificial Intelligence in Social Media.” *Social Network Analysis and Mining* 14(1). doi: 10.1007/s13278-024-01382-y.

Meng, Yu, Yunyi Zhang, Jiaxin Huang, Chenyan Xiong, Heng Ji, Chao Zhang, and Jiawei Han. 2020. “Text Classification Using Label Names Only: A Language Model Self-Training Approach.” *EMNLP 2020 - 2020 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference* 9006–17. doi: 10.18653/v1/2020.emnlp-main.724.

Mohamed Shaffril, Hayrol Azril, Samsul Farid Samsuddin, and Asnarulkhadi Abu Samah. 2021. "The ABC of Systematic Literature Review: The Basic Methodological Guidance for Beginners." *Quality and Quantity* 55(4):1319–46. doi: 10.1007/s11135-020-01059-6.

Mostafa, Almetwally M., Swarn Avinash Kumar, Talha Meraj, Hafiz Tayyab Rauf, Abeer Ali Alnuaim, and Maram Abdullah Alkhayyal. 2022. "Guava Disease Detection Using Deep Convolutional Neural Networks: A Case Study of Guava Plants." *Applied Sciences (Switzerland)* 12(1). doi: 10.3390/app12010239.

Moupojou, Emmanuel, Appolinaire Tagne, Florent Retraint, Anicet Tadonkemwa, Dongmo Wilfried, Hyppolite Tapamo, and Marcellin Nkenlifack. 2023. "FieldPlant: A Dataset of Field Plant Images for Plant Disease Detection and Classification With Deep Learning." *IEEE Access* 11(March):35398–410. doi: 10.1109/ACCESS.2023.3263042.

Muhammad, Abdulhamid, Supavadee Aramvith, Khanita Duangchaemkarn, and Ming Ting Sun. 2024. "Brain MRI Image Super-Resolution Reconstruction: A Systematic Review." *IEEE Access* 12(November):156347–62. doi: 10.1109/ACCESS.2024.3478829.

Mumtaz, Sidrah, Mudassar Raza, Ofonime Dominic Okon, Saeed Ur Rehman, Adham E. Ragab, and Hafiz Tayyab Rauf. 2023. "A Hybrid Framework for Detection and Analysis of Leaf Blight Using Guava Leaves Imaging." *Agriculture (Switzerland)* 13(3). doi: 10.3390/agriculture13030667.

Mustak Un Nobi, Md., Md. Rifat, M. F. Mridha, Sultan Alfarhood, Mejdl Safran, and Dunren Che. 2023. "GLD-Det: Guava Leaf Disease Detection in Real-Time Using Lightweight Deep Learning Approach Based on MobileNet." *Agronomy* 13(9):2240. doi: 10.3390/agronomy13092240.

Nandi, Rabindra Nath, Aminul Haque Palash, Nazmul Siddique, and Mohammed Golam Zilani. 2023. "Device-Friendly Guava Fruit and Leaf Disease Detection Using Deep Learning." *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST* 490 LNICST:49–59. doi: 10.1007/978-3-031-34619-4_5.

Ojo, Mike O., and Azlan Zahid. 2023. "Improving Deep Learning Classifiers Performance via Preprocessing and Class Imbalance Approaches in a Plant Disease Detection Pipeline." *Agronomy* 13(3). doi: 10.3390/agronomy13030887.

Orchi, Houda, Mohamed Sadik, and Mohammed Khaldoun. 2022. "On Using Artificial Intelligence and the Internet of Things for Crop Disease Detection: A Contemporary Survey." *Agriculture (Switzerland)* 12(1). doi: 10.3390/agriculture12010009.

Perumal, P., Kandasamy Sellamuthu, K. Vanitha, and V. K. Manavalasundaram D A Professor. 2021. "Guava Leaf Disease Classification Using Support Vector Machine." *Turkish Journal of Computer and Mathematics Education* 12(7):1177–83.

Ramadhan, Nur Ghaniaviyanto, Adiwijaya, Warih Maharani, and Alfian Akbar Gozali. 2024. "Chronic Diseases Prediction Using Machine Learning With Data Preprocessing Handling: A Critical Review." *IEEE Access* 12(March):80698–730. doi: 10.1109/ACCESS.2024.3406748.

Rashid, Javed, Imran Khan, Ghulam Ali, Shafiq ur Rehman, Fahad Alturise, and Tamim Alkhalifah. 2023. "Real-Time Multiple Guava Leaf Disease Detection from a Single Leaf Using Hybrid Deep Learning Technique." *Computers, Materials and Continua* 74(1):1235–57. doi: 10.32604/cmc.2023.032005.

Reda, Oumaima, Naoual Chaouni Benabdellah, and Ahmed Zellou. 2023. "A Systematic Literature Review on Data Quality Assessment." *Bulletin of Electrical Engineering and Informatics* 12(6):3736–57. doi: 10.11591/eei.v12i6.5667.

S.Abirami, M. Thilagavathi. 2017. "Application of Image Processing in Diagnosing Guava Leaf Diseases." *International Journal of Scientific Research and Management (IJSRM)* 5(7):5927–33. doi: 10.18535/ijsrm/v5i7.19.

Sarki, Rubina, Khandakar Ahmed, Hua Wang, Yanchun Zhang, Jiangang Ma, and Kate Wang. 2021. "Image Preprocessing in Classification and Identification of Diabetic Eye Diseases." *Data Science and Engineering* 6(4):455–71. doi: 10.1007/s41019-021-00167-z.

Shakil, Rashiduzzaman, Bonna Akter, Aditya Rajbongshi, Umme Sara, Mala Rani Barman, and Aditi Dhali. 2023. "A Transfer Learning Approach to the Development of an Automation System for Recognizing Guava Disease Using CNN Models for Feasible Fruit Production." *Lecture Notes in Networks and Systems* 647 LNNS:127–41. doi: 10.1007/978-3-031-27409-1_12.

Shihab, Montasir Rahman, Nafiu Islam Saim, Mayen Uddin Mojumdar, Dewan Mamun Raza, Shah Md Tanvir Siddiquee, Sheak Rashed Haider Noori, and Narayan Ranjan Chakraborty. 2025. "Image Dataset for Classification of Diseases in Guava Fruits and Leaves." *Data in Brief* 59:111378. doi: 10.1016/j.dib.2025.111378.

Srinivas, B., P. Satheesh, P. Rama Santosh Naidu, and U. Neelima. 2021. *Prediction of Guava Plant Diseases Using Deep Learning*. Vol. 698. Springer Singapore.

Upadhyay, Abhishek, Narendra Singh Chandel, Krishna Pratap Singh, Subir Kumar Chakraborty, Balaji M. Nandede, Mohit Kumar, A. Subeesh, Konga Upendar, Ali Salem, and Ahmed Elbeltagi. 2025. "Deep Learning and Computer Vision in Plant Disease Detection: A Comprehensive Review of Techniques, Models, and Trends in Precision Agriculture." *Artificial Intelligence Review* 58(3). doi: 10.1007/s10462-024-11100-x.

Valero-Carreras, Daniel, Javier Alcaraz, and Mercedes Landete. 2023. "Comparing Two SVM Models through Different Metrics Based on the Confusion Matrix." *Computers and Operations Research* 152(April 2022):106131. doi: 10.1016/j.cor.2022.106131.

Vitti, Karine Alessandra, Lilian Maluf de Lima, and João Gomes Martines Filho. 2020. "Agricultural and Economic Characterization of Guava Production in Brazil." *Revista Brasileira de Fruticultura* 42(1):1–11. doi: 10.1590/0100-29452020447.

Xuejie Hao, Lu Liu, Rongjin Yang, Lizeyan Yin, Le Zhang and Xiuhong Li. 2023. "A Review of Data Augmentation Methods of Remote Sensing Image Target Recognition." *Remote Sensing*. doi: https://doi.org/10.3390/rs15030827.