A systematic literature review

of computational solutions applied to agriculture: visual disease diagnosis, postharvest classification, and intelligent IoT-based monitoring

Luis Chamba-Eras[1]
Sergio Jumbo-Gonza[2]
Walter Yaguana-Alejandro[3]
Geovanny Romero-Suquilanda[4]
Gerardo Herrera-Campoverde[5]

[1]Universidad Nacional de Loja, Facultad de la Energía, las Industrias y los Recursos Naturales No Renovables, Grupo de Investigación en Tecnologías de la Información y Computación, Carrera de Computación, Loja, Ecuador,
https://orcid.org/0000-0003-3069-9628
lachamba@unl.edu.ec

[2]Universidad Nacional de Loja, Facultad de la Energía, las Industrias y los Recursos Naturales No Renovables, Grupo de Investigación en Tecnologías de la Información y Computación, Carrera de Computación, Loja, Ecuador,
https://orcid.org/0000-0003-3069-9628
sergio.jumbo@unl.edu.ec

[3]Universidad Nacional de Loja, Facultad de la Energía, las Industrias y los Recursos Naturales No Renovables, Grupo de Investigación en Tecnologías de la Información y Computación, Carrera de Computación, Loja, Ecuador,
https://orcid.org/0009-0006-2614-8734
walter.yaguana@unl.edu.ec

[4]Universidad Nacional de Loja, Facultad de la Energía, las Industrias y los Recursos Naturales No Renovables, Grupo de Investigación en Tecnologías de la Información y Computación, Carrera de Computación, Loja, Ecuador,
https://orcid.org/0009-0004-7121-1574
geovanny.romero@unl.edu.ec

[5]Universidad Nacional de Loja, Facultad de la Energía, las Industrias y los Recursos Naturales No Renovables, Grupo de Investigación en Tecnologías de la Información y Computación, Carrera de Computación, Loja, Ecuador,
https://orcid.org/0009-0006-7925-2594
gerardo.herrera@unl.edu.ec

DOI: 10.61854/rccedia.v1n1.007
Sent: February 1, 2026
Approved: April 6, 2026
Published: Tuesday, July 14, 2026

Abstract

Agriculture faces increasing challenges related to timely disease diagnosis, postharvest quality assessment, and efficient farm management. In this context, computational solutions based on artificial intelligence, computer vision, and IoT technologies have emerged as key tools to support agricultural decision-making. This paper presents a systematic literature review (SLR) aimed at analyzing and synthesizing recent scientific evidence on computational approaches applied to (1) automatic disease detection in agricultural leaves and fruits, (2) surface classification and postharvest quality assessment of grains, and (3) farm monitoring and management through web platforms and IoT. The SLR followed the Kitchenham guidelines and the PRISMA protocol, applying explicit criteria for study identification, selection, quality assessment, and synthesis. A total of 263 initial records were identified; after screening and quality assessment, 16 primary studies were included: 6 on visual disease diagnosis (Axis A), 3 on postharvest grain classification (Axis B), and 7 on IoT-based farm monitoring (Axis C). The results indicate that deep learning (convolutional neural networks, attention mechanisms, and pretrained models) dominates visual diagnosis and postharvest classification. In addition, IoT-based infrastructures are widely adopted for agricultural monitoring, although usability evaluation and systematic application of standards such as ISO 9241 remain limited. Unlike previous reviews focused on generic technologies or specific crops, this SLR structures evidence by concrete agricultural problems, integrating visual diagnosis, postharvest classification, and IoT monitoring into a unified comparative analysis. Finally, the review identifies research gaps related to model generalization, dataset standardization, and validation under real-field conditions, which define future research directions in computational agriculture.

Keywords:
applied deep learning, postharvest grain classification, IoT agricultural monitoring.

Introduction

Contemporary agriculture faces growing challenges related to food security, environmental sustainability, and productive efficiency, intensified by climate change, pressure on natural resources, and the need to optimize production systems. In this context, computational solutions play a strategic role in supporting agricultural diagnosis, monitoring, and control processes, especially in scenarios where traditional methods are costly, subjective, or difficult to scale. Technologies such as artificial intelligence (AI), computer vision, deep learning, and web platforms integrated with IoT sensors are essential tools within the field of computing applied to agriculture (Ouamane et al., 2025; Yue et al., 2025).

One of the problems most frequently addressed in the recent literature is the early detection of diseases in agricultural crops because of their direct impact on productivity and profitability. In crops such as coffee, foliar diseases such as coffee leaf rust (Hemileia vastatrix) represent a significant threat to production. Several studies show that traditional visual diagnostic methods based on manual inspection have limitations in terms of accuracy, timeliness, and dependence on expert knowledge. In response to these limitations, approaches based on computer vision and deep neural networks have shown promising results for the automatic detection of foliar diseases, achieving high levels of accuracy compared with manual methods (Faisal et al., 2023; Muthusamy & Ramu, 2025).

Similarly, disease detection in agricultural fruits, such as cocoa pods, has been addressed through deep learning models aimed at recognizing visual patterns associated with defects and diseases. The literature reports the widespread use of convolutional neural networks (CNNs) trained with fruit images, as well as the use of preprocessing and data augmentation techniques to mitigate the scarcity of labeled datasets. However, challenges related to model generalization and the variability of image acquisition conditions in the field have also been identified, limiting their direct transfer to real-world production scenarios (Borromeo et al., 2025; Yue et al., 2025).

Another relevant aspect of the agricultural production cycle is the surface classification and quality assessment of grains during postharvest stages, particularly in crops such as maize. In this domain, pretrained deep learning models such as MobileNetV2 and EfficientNet have been widely studied because of their ability to balance accuracy and computational efficiency, facilitating their implementation in resource-constrained environments (Garg & Gill, 2024). The available evidence suggests that these approaches enable the automation of visual grain inspection with competitive results. Nevertheless, questions remain regarding the influence of the quality, quantity, and diversity of datasets on the models' generalization capability (Prakasa et al., 2021; Pasaribu et al., 2025; Kant, 2024).

In agricultural applications based on computer vision, one of the recurring challenges is the limited availability and imbalance of datasets, which affect the generalization capability of deep learning models. Several strategies have been proposed to mitigate this problem, such as data augmentation through synthetic data generation techniques. In particular, the use of generative adversarial networks (GANs) has been explored as an approach to enrich agricultural image datasets, although with mixed results regarding their impact on model performance (Rahman et al., 2024; Pasaribu et al., 2025).

Beyond visual diagnosis and postharvest classification, farm management constitutes an essential component of modern agriculture. In this area, IoT sensor-based web platforms enable the continuous monitoring of agricultural and environmental variables, supporting informed decision-making. However, the literature highlights that the effectiveness of these solutions depends not only on their technical functionality but also on aspects related to usability, software quality, and human-computer interaction. In this regard, standards such as ISO 9241 have been used to evaluate agricultural web platforms developed in simulated or prototype environments (Yeong et al., 2024; Nugroho et al., 2017).

Several previous systematic reviews have partially addressed these topics. Yue et al. (2025) conducted a review of deep learning applied to disease and pest recognition, concluding that CNNs and vision transformers are the dominant approaches. However, they also identified a lack of evaluations under real field conditions. Similarly, Rahman et al. (2024) reviewed the use of GANs for data augmentation in agriculture, identifying a significant gap in the evaluation of the quality of the generated synthetic images. Nevertheless, none of these reviews integrates visual disease diagnosis, postharvest classification, and IoT-based agricultural monitoring into a unified comparative analysis structured around specific agricultural problems. This thematic and methodological fragmentation hinders an integrated understanding of technological trends, predominant evaluation metrics, and research gaps from the perspective of applied computing.

In this context, the objective of this article is to systematically analyze and synthesize the recent scientific literature on the use of computational solutions applied to specific agricultural problems. In particular, the systematic literature review (SLR) focuses on: (1) the automatic detection of diseases in the leaves and fruits of agricultural crops, with an emphasis on coffee and cocoa; (2) the surface classification and quality assessment of agricultural grains in postharvest contexts, especially maize; and (3) the monitoring and management of agricultural farms through IoT sensor-based web platforms, considering usability and software quality criteria in simulated or prototype environments.

The time interval from 2017 to 2025 was deliberately selected to capture the period of greatest proliferation and maturity of modern deep learning architectures, such as vision transformers, attention models, and lightweight pretrained architectures, as well as the accelerated expansion of IoT infrastructures in agricultural environments. This period also coincides with the increasing publication of validation studies under semi-controlled conditions and the emergence of novel approaches such as agricultural digital twins. Nevertheless, it is acknowledged that this decision introduces a potential selection bias toward recent technologies, excluding earlier foundational studies that are also relevant for understanding the evolution of the field.

The main contributions of this article are: (1) to present an SLR structured around specific agricultural problems with complete PRISMA traceability; (2) to synthesize the approaches, architectures, and evaluation metrics used in computer vision- and deep learning-based solutions; (3) to analyze the balance between accuracy and computational efficiency of pretrained models applied to the surface classification of grains; and (4) to systematize software engineering practices and usability criteria reported in IoT-based agricultural web platforms, with an emphasis on the ISO 9241 standard.

This work is structured around the following research questions: RQ1, what computer vision and deep learning approaches have been reported for the automatic detection of diseases in the leaves and fruits of agricultural crops, particularly coffee and cocoa, and what metrics are used to evaluate them?; RQ2, which datasets, preprocessing techniques, and neural network architectures are most frequently used in agricultural visual diagnosis systems, and which achieve the best results in terms of accuracy and robustness?; RQ3, which pretrained deep learning models have demonstrated the best balance between accuracy and computational efficiency in the surface classification and quality assessment of agricultural grains, with an emphasis on maize, under computational resource constraints?; and RQ4, how are IoT sensor-based web platforms for agricultural farm monitoring and management designed and evaluated, and which usability, software quality, and regulatory criteria, such as ISO 9241, are considered in simulated or prototype environments?

Methods

Study design

A systematic literature review (SLR) was conducted to identify, analyze, and synthesize scientific evidence on the use of computational solutions applied to specific agricultural problems. The study adopts an explicit, reproducible, and structured approach, following the methodological process defined by Brereton et al. (2007) in the context of computing and related disciplines, and aligned with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol to ensure the traceability and transparency of the selection process.

This SLR is exploratory-descriptive in nature, as its objective is not to quantitatively compare homogeneous experimental results, but rather to identify trends, technological approaches, evaluation metrics, and research gaps within the analyzed domain. The study does not include primary experimentation or the implementation of new models, relying exclusively on the critical analysis of peer-reviewed scientific literature.

Literature search strategy

The literature search strategy was systematically designed to cover the domain of computing applied to agriculture, with an emphasis on visual disease diagnosis, postharvest quality classification, and intelligent agricultural farm monitoring. The scientific databases IEEE Xplore, Scopus, SpringerLink, ScienceDirect, and ACM Digital Library were searched, as they were selected for their broad coverage of publications in computer science and engineering.

The search strings were constructed using the PICOC (population, intervention, comparison, outcome, context) approach. Terms associated with the agricultural population and context (plant disease, coffee, cacao, maize, farm monitoring) were combined with terms related to computational interventions (computer vision, deep learning, pretrained models, IoT, web platform, usability, ISO 9241), using the Boolean operators AND and OR.

The search process was carried out iteratively, progressively refining the search strings based on the preliminary analysis of the retrieved results, with the aim of maximizing relevance and minimizing the retrieval of non-relevant studies. Searches were restricted to titles, abstracts, and keywords.

Inclusion and exclusion criteria

The retrieved studies were evaluated using predefined inclusion and exclusion criteria:

Inclusion criteria:

- Peer-reviewed scientific articles published in journals or conference proceedings.
- Studies published between 2017 and 2025, a period that captures the greatest maturity of modern deep learning architectures.
- Studies applying computational solutions (AI, computer vision, deep learning, web platforms, or IoT) to specific agricultural problems.
- Studies reporting quantitative evaluation metrics or formal system evaluation criteria.
- Publications available in English.

Revista Científica CEDIA - ISSN: 3151-8435

Exclusion criteria:

- Non-peer-reviewed articles (editorials, extended abstracts, or technical reports).
- Purely conceptual studies without experimental, simulated, or evaluative validation.
- Duplicate publications across databases.
- Studies whose focus is not clearly related to agricultural applications or applied computing.
- Studies with insufficient information to answer the research questions (quality score < 3; see the section "Quality Assessment of the Studies").

Study selection process and PRISMA flow

The study selection process was carried out independently by two researchers, applying three sequential analyses. Discrepancies between reviewers were resolved through consensus. The interobserver agreement level was calculated using Cohen's kappa coefficient ($\kappa = 0.82$), indicating substantial agreement according to the thresholds established by Landis & Koch (1977).

In the first analysis (title and abstract), 263 records retrieved from the five databases were evaluated. After removing 47 duplicates (N = 216 unique records), studies whose domain, focus, or context did not align with the research questions were excluded, resulting in 52 preselected studies. The second analysis consisted of a focused review of the introduction, methodology, and conclusions of these 52 studies, allowing the evaluation of methodological rigor and the consistency between objectives, methods, and results; 24 studies passed this stage. In the third analysis, a complete critical reading with systematic quality control was conducted, resulting in the final inclusion of 16 primary studies. Table 1 and Figure 1 summarize the selection flow according to the PRISMA 2020 protocol.

Table 1
Study Selection Flow According to the PRISMA Protocol

PRISMA Phase	Description	No. of records
Identification	Records retrieved from five databases (IEEE Xplore, Scopus, SpringerLink, ScienceDirect, ACM)	263
Duplicate removal	Records removed due to duplication across databases	47
Screening (analysis 1)	Evaluation by title and abstract; excluded due to non-relevant domain	216 screened, 164 excluded
Eligibility (analysis 2)	Review of the introduction, methodology, and conclusions; excluded due to methodological insufficiency	52 evaluated, 28 excluded
Included (analysis 3)	Full critical reading and quality assessment (QC ≥ 3); studies comprising the final synthesis	16 included

Note: Prepared by the authors

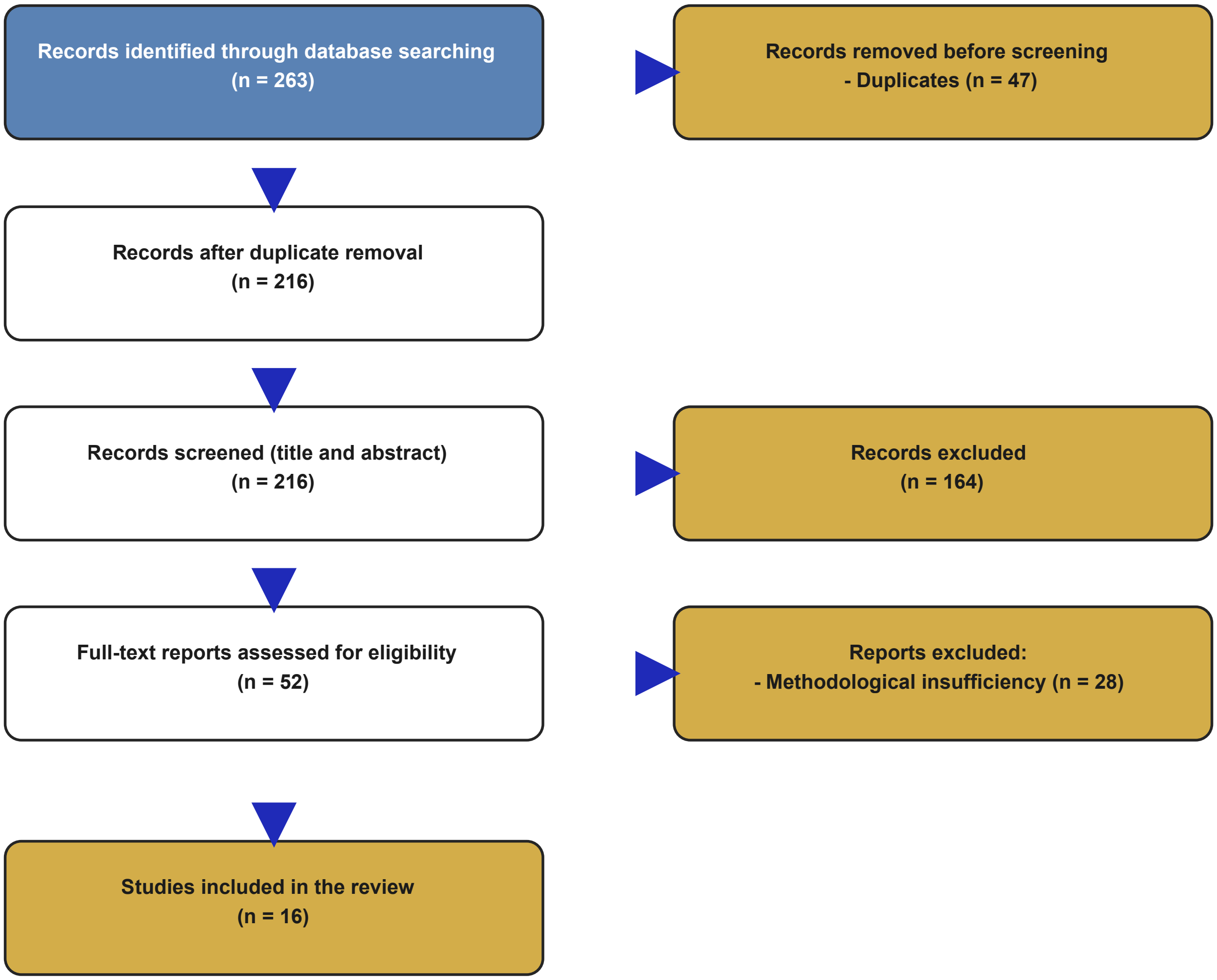

Figure 1

Study Selection Flow Diagram According to the PRISMA Protocol

Revista Científica CEDIA - ISSN: 3151-8435

Quality assessment of the studies

The quality assessment of the included studies was conducted using an evaluation instrument based on five quality criteria (QC), following the recommendations of Kitchenham et al. (2010). Each criterion was scored in a binary manner (0 = does not meet; 1 = meets), with a minimum acceptance threshold of 3 out of 5 points for inclusion in the final synthesis:

- QC1: Is the research objective or problem clearly defined and delimited?
- QC2: Is the methodology adequately described (approach, data, model, or system)?
- QC3: Is the dataset specified, including its size, source, or relevant characteristics?
- QC4: Are quantitative evaluation metrics (accuracy, F1-score, mAP, IoU, etc.) or formal system evaluation criteria reported?
- QC5: Are the conclusions supported by the reported experimental or evaluative results?

Each study was independently evaluated by the two reviewers, and the final score corresponded to the average of both evaluations. The eight studies excluded in the third analysis obtained scores below 3, mainly because they reported results without describing the dataset (QC3) or without quantitative metrics (QC4).

Information synthesis strategy

The information was synthesized using a qualitative-comparative approach structured according to the three predefined thematic axes. For each axis, a homogeneous extraction matrix was constructed including: (1) study identification (authors, year, source database); (2) crop or agricultural context; (3) problem addressed; (4) computational approach and architecture; and (5) evaluation metrics and reported values.

The thematic synthesis was conducted through inductive axial coding of the extracted data, identifying emerging patterns regarding dominant technological approaches, methodological trends, and recurring research gaps. No quantitative meta-analysis was performed because of the heterogeneity of approaches, metrics, and experimental contexts reported in the analyzed studies, which made the statistical aggregation of results unfeasible. The identified patterns were validated through triangulation between the two reviewer researchers.

Threats to validity

As with any SLR, this study presents potential threats to validity, which are addressed according to the classification proposed by Brereton et al. (2007). Threats to internal validity are related to potential biases in the selection and evaluation of studies; these were mitigated through explicit criteria, multiple independent analyses, systematic quality control, and the calculation of interobserver agreement ($\kappa = 0.82$).

Threats to external validity are associated with the generalizability of the findings, given that the analyzed studies correspond to diverse contexts, crops, and experimental scenarios. Threats to construct validity are related to the interpretation of the reported concepts and metrics and were mitigated through a clear definition of the analysis axes and a structured comparative synthesis.

Revista Científica CEDIA - ISSN: 3151-8435

Results

Table 2 presents the primary studies included in this SLR, organized by thematic axis. These studies constitute the final corpus on which the results and analysis presented in the following subsections are based.

Table 2

Primary studies included in the SLR

Axis A. Visual Diagnosis of Diseases in Leaves and Fruits: 6 Studies				
Study	**Crop**	**Problem addressed**	**Computational approach**	**Reported metrics**
Faisal et al. (2023)	Coffee	Foliar disease	CNN + feature fusion (MobileNetV3 + Swin Transformer)	*Accuracy, Precision, Recall, F1-score*
Muthusamy & Ramu (2025)	Coffee, Banana	Micronutrient deficiency + disease + severity	CNN + attention (D-PAN)	*Accuracy, Precision, Recall, F1-score, AUC*

Abinaya et al. (2023)	Multiple crops	Segmentation and classification	Autoencoder + U-Net (CAAR-UNet)	Pixel *Accuracy,* wmIoU, Dice, mBF1
Quamane et al. (2025)	Multiple crops	Disease detection	*Vision transformer*	*Accuracy, Precision, Recall, F1-score* (no AUC)
Noon et al. (2022)	Cotton	Detection + localization	YOLOX	*mAP, Precision*
Borromeo et al. (2025)	Cocoa	Fruit maturity and variety	VGG16 + SSD MobileNet	*Accuracy*

Axis B: Surface Classification and Postharvest Quality Assessment: 3 Studies				
Study	**Crop**	**Problem Addressed**	**Computational Approach**	**Reported Metrics**
Prakasa et al. (2021)	Maize	Seed quality	CNN + *features* (ResNet-50)	*F1-score* (primary), *Accuracy*
Pasaribu et al. (2025)	Maize, cassava, rice, sugarcane	Plant classification (impact of GAN -based data augmentation)	MobileNetV2, EfficientNet-B0, ResNet-18	*Accuracy*
Kant (2024)	Rice (transferable)	Grain classification	MobileNetV2	*Accuracy* (0.99), *Precision, Recall, F1-score* (0.99)

Axis C: Agricultural Farm Monitoring and Management (IoT + Web Platforms): 7 Studies				
Study	**Context**	**Problem Addressed**	**Technological Approach**	**Evaluation Criteria**
Yeong et al. (2024)	Farm, machinery	Agricultural monitoring	Web platform + IoT	Functionality
Rao et al. (2019)	General agriculture	Environmental monitoring	IoT	Functional validation
Aliparo et al. (2022)	Agricultural farm	Soil nutrients monitoring	IoT + sensors	Measurement accuracy
Yadav et al. (2025)	General agriculture	Sensor networks	IoT	Performance
Bhadra et al. (2025)	Greenhouse	Real-time monitoring	IoT + offline capability	Functionality
Hewage et al. (2017)	Agricultural management	Farm management	Software platform	Architecture
Archambault (2024)	Digital agriculture	Agricultural simulation	Digital twins	Modeling

Note: Prepared by the authors based on the included primary studies.

Results for RQ1: Computer vision and deep learning approaches for the detection of diseases in agricultural leaves and fruits

The six studies included in Axis A consistently report the use of deep learning, primarily through CNNs and derived architectures, as the dominant approach for the automatic detection of diseases in agricultural leaves and fruits. Four of the six studies employ CNNs as the core architecture, while the remaining two incorporate vision transformers and YOLOX, respectively, demonstrating an emerging methodological diversification.

For coffee foliar diseases, Faisal et al. (2023) report a model based on hybrid feature fusion through the combination of MobileNetV3 and Swin Transformer, together with a variational autoencoder (VAE), achieving an accuracy of 84.29%, with a precision of 84.67%, recall of 84.29%, and F1-score of 83.64% for disease classification using the real-world RoCoLe dataset. This publicly available dataset comprises 1,560 images of Robusta coffee leaves captured under field conditions. Muthusamy & Ramu (2025) propose the Dynamic Pyramid Attention Network (D-PAN) framework with dynamic attention mechanisms, designed not only for disease and micronutrient deficiency detection but also for severity level classification, reporting an accuracy of 98.80% for banana and 94.60% for coffee (evaluated on the Mendeley Co-Leaf DB dataset), with consistently strong performance in accuracy, precision, recall, and F1-score.

Other studies extend the approach to segmentation and classification tasks by incorporating hybrid architectures. Abinaya et al. (2023) employ the CAAR-UNet (Cascading Autoencoder with Attention Residual U-Net) model, evaluating performance using mean pixel accuracy (Am) of 95.26% and weighted mean Intersection over Union (wmIoU) of 0.7451, together with the Dice coefficient and mean Boundary F1-score (mBF1), demonstrating the use of metrics beyond traditional classification when precise localization of affected regions is required.

Regarding emerging architectures, Ouamane et al. (2025) report the use of optimized vision transformers on the public PlantVillage dataset, achieving an accuracy of 99.77%, precision of 99.61%, recall of 99.69%, and F1-score of 99.64%, also evaluated on the Taiwan Tomato and BananaLSD datasets, confirming their generalization capability. Likewise, object detection approaches are also represented in the literature: Noon et al. (2022) employ an improved YOLOX variant to handle multiple severity levels and coexisting diseases in cotton, using mAP as the primary evaluation metric, reporting 73.13% during training and 72.31% on the test set.

For agricultural fruits, specifically cocoa, Borromeo et al. (2025) report the use of VGG16 with transfer learning for the visual classification of pod maturity and variety, complemented by SSD MobileNet for object detection and implemented in a mobile application. The highest reported accuracy is 88.9% (combining variety and maturity classification with the largest available number of pods), decreasing to 76.7% when using the smallest dataset. Although the study focuses on maturity stage and variety, its inclusion demonstrates the application of computer vision directly to cocoa pod images as part of visual inspection systems.

Overall, the studies included in this axis predominantly employ classification-oriented evaluation metrics (accuracy, precision, recall, F1-score), with additional adoption of segmentation metrics (IoU, Dice) and detection metrics (mAP) when required by the task.

Results for RQ2: Datasets, architectures, and techniques used in agricultural visual diagnosis systems

The six studies in Axis A predominantly report the use of RGB image datasets captured under controlled or semi-controlled conditions. Most of the datasets used are domain-specific; however, some studies employ widely used public datasets. Faisal et al. (2023) use the public RoCoLe dataset (1,560 images of Robusta coffee leaves under real field conditions), whereas Ouamane et al. (2025) use the public PlantVillage dataset (more than 87,000 images). The remaining studies construct ad hoc datasets ranging from 102 to 1,112 original images (prior to augmentation). This heterogeneity limits reuse and direct comparison across studies and compromises the assessment of model generalization.

Deep CNN architectures, including variants such as ResNet, EfficientNet, and customized models, appear in four of the six studies. Attention mechanisms are incorporated in two studies (Muthusamy & Ramu, 2025; Abinaya et al., 2023), allowing the model to focus on task-relevant regions. Regarding preprocessing techniques, all studies report image normalization and data augmentation, including rotation, horizontal flipping, random cropping, and brightness adjustment. Pasaribu et al. (2025) demonstrate that synthetic augmentation improves accuracy by 3 to 8 percentage points in lightweight models.

In terms of performance, the reported accuracy values are heterogeneous, ranging from 72.31% (Noon et al., 2022, on cotton using a proprietary dataset) and 84.29% (Faisal et al., 2023, under real field conditions) to 99.77% (Ouamane et al., 2025, using the PlantVillage dataset). Accuracy values above 90% are concentrated in studies using controlled or large-scale public datasets (Muthusamy & Ramu, 2025: 94.60–98.80%; Abinaya et al., 2023: 95.26%; Borromeo et al., 2025: 88.9%). However, these values should be interpreted with caution, as they were obtained under conditions that may not be directly comparable and without cross-validation using independent field data, limiting the assessment of the models' actual robustness.

Results for RQ3: Pretrained models for surface classification and quality assessment of maize grains

The three studies in Axis B demonstrate the predominant use of deep learning models applied to digital images for grain classification and quality assessment. Prakasa et al. (2021) report a CNN-based system with feature classifiers to categorize maize seeds (BIMA-20 URI Hybrid variety) into two classes (good and poor quality), using a dataset of 232 laboratory-acquired images. The primary reported metric is the F1-score, with the best result of 0.983 obtained using ResNet-50, while most evaluated methods achieve F1-scores greater than 0.90. The study confirms that computer vision-based approaches enable the automation of postharvest inspection with consistent results.

Pasaribu et al. (2025) analyze the impact of synthetic data augmentation using GANs (DCGAN and WGAN-GP) on MobileNetV2, EfficientNet-B0, and ResNet-18 models applied to plant classification (cassava, maize, rice, and sugarcane). The main finding is that GAN-generated data negatively affected classification accuracy because of image artifacts and high Frechet Inception Distance (FID) values. The highest reported accuracy was 97.37%, obtained exclusively with the original dataset (245 images), whereas models trained with GAN-augmented data experienced reduced performance. WGAN-GP generated visually more realistic images than DCGAN but also failed to improve classification accuracy.

Although not focused exclusively on maize, Kant (2024) reports the application of MobileNetV2 with transfer learning for the classification of five rice grain varieties (Arborio, Basmati, Ipsala, Jasmine, and Karacadag), achieving an accuracy, precision, recall, and F1-score of 0.99. The inclusion of this study highlights the methodological transferability of lightweight pretrained architectures to surface classification problems involving agricultural grains under resource-constrained scenarios, which is relevant to the maize postharvest context.

Across the included studies, the primary evaluation metric is either the F1-score (Prakasa et al., 2021) or accuracy (Pasaribu et al., 2025; Kant, 2024). Low-complexity pretrained models (MobileNetV2 and EfficientNet-B0) demonstrate high accuracy using the original dataset (97.37%), but their advantage over more complex architectures (ResNet-18) is marginal when the training dataset is of adequate quality.

Results for RQ4: Web platforms and iot for agricultural farm monitoring and management

The seven studies in Axis C show that agricultural farm monitoring and management solutions are predominantly designed as IoT systems integrated with software platform components for data acquisition, transmission, storage, and visualization. Yeong et al. (2024) report a telematics monitoring platform for agricultural machinery, using a software architecture as the integration layer for IoT sensors. Rao et al. (2019) describe an IoT system integrated with UAVs and ZigBee sensor networks for real-time environmental monitoring, evaluated primarily in terms of functionality and scalability. Aliparo et al. (2022) implement an IoT system for monitoring soil nutrients (NPK), using measurement accuracy as the evaluation criterion. Yadav et al. (2025) report intelligent sensor networks with IoT connectivity, focusing on transmission performance and latency. Bhadra et al. (2025) implement an IoT system with offline monitoring capability for greenhouses, prioritizing continuous data availability.

Regarding web platforms and management components, Hewage et al. (2017) report a Farm Management Information System (FMIS) developed using agile software engineering practices to support operational planning. Archambault (2024) describes co-simulation models and digital twins for controlled agriculture, integrating computational representations of crops to support decision-making.

Critically, none of the seven studies in Axis C reports formal evaluations explicitly based on the ISO 9241 standard. The available evidence is concentrated on technical functionality, sensing performance, and connectivity criteria. This gap between technical functionality and usability evaluation is consistently observed across all studies in this axis.

Discussion

RQ1: Visual diagnosis approaches for leaves and fruits

The RQ1 results confirm that deep learning based on computer vision has become the dominant approach for the automated diagnosis of diseases in crops. The widespread adoption of CNNs and their variants (four of the six studies in Axis A) reflects their proven effectiveness in capturing visual patterns associated with phytosanitary symptoms. However, this dominance poses the risk of methodological blind spots: CNN models are known to be sensitive to domain variations that are not represented in the training dataset.

The emergence of vision transformers and object detectors such as YOLOX is not trivial; rather, it reflects the field's transition toward models with greater global attention capability and spatial localization, which is particularly relevant for agricultural scenarios where multiple diseases coexist or where the boundaries of affected regions are diffuse. Nevertheless, these models require greater computational capacity and larger training datasets, which may limit their deployment in field applications with constrained hardware.

From a technology transfer perspective, the main obstacle identified in the literature is the gap between evaluation environments and real field conditions. The reported accuracy values vary considerably across studies, ranging from 72.31% (Noon et al., 2022, using a proprietary cotton dataset) and 84.29% (Faisal et al., 2023, under real field conditions using the public RoCoLe dataset) to 99.77% (Ouamane et al., 2025, using the large-scale PlantVillage dataset). This heterogeneity reflects the critical influence of the dataset used (size, diversity, acquisition conditions) on the reported performance, rather than necessarily indicating the superiority of one architectural approach over another. The models are not systematically evaluated against variations in natural lighting, acquisition angles, occlusions caused by adjacent leaves, or images captured using different mobile devices. This external validity gap is the main barrier to the widespread adoption of these technologies by farmers.

RQ2: Data, Architectures, and robustness in visual diagnosis systems

The comparative analysis of the six studies in Axis A reveals a structural dependence on ad hoc datasets in four of the six studies, whose size and diversity are insufficient for training generalizable models. Two studies use public datasets: RoCoLe (Faisal et al., 2023) and PlantVillage (Ouamane et al., 2025), facilitating reproducibility and comparison with other studies. All studies mitigate data scarcity through data augmentation, which is a widely accepted practice but does not replace the diversity of real field conditions. The absence of standardized public repositories for diseases specific to crops in the Andean region (such as coffee leaf rust and cocoa moniliasis) is particularly critical.

The reported accuracy values should be interpreted with methodological caution. Faisal et al. (2023) report an accuracy of 84.29% under real field conditions, whereas Ouamane et al. (2025) report 99.77% using a large-scale controlled dataset, demonstrating that the evaluation context is a determining factor when interpreting reported performance. The adoption of complementary metrics such as macro F1-score (for imbalanced classes), the ROC curve, the precision-recall curve, and cross-validation protocols using independent datasets emerges as an urgent methodological need to strengthen the external validity of the reported results.

RQ3: Accuracy and efficiency in postharvest grain classification

The results of Axis B present nuanced findings regarding the impact of synthetic data augmentation. Prakasa et al. (2021) confirm that deep learning models achieve F1-scores above 0.90 (best: 0.983 using ResNet-50) for maize seed classification. Kant (2024) demonstrates that MobileNetV2 with transfer learning achieves an accuracy of 0.99 in rice grain classification, confirming its feasibility for postharvest applications under limited-resource conditions. The favorable balance between accuracy and computational efficiency offered by lightweight pretrained models makes them suitable for environments with hardware constraints and for small- and medium-scale applications.

However, the study by Pasaribu et al. (2025) provides a critical finding that contradicts the widespread assumption that synthetic data augmentation improves performance: the use of GAN-generated data (DCGAN and WGAN-GP) reduced classification accuracy compared with the original dataset, with the highest accuracy (97.37%) obtained using the original dataset of 245 images. This result highlights that the quality of synthetic data—measured by the absence of artifacts and by metrics such as Frechet Inception Distance (FID)—is more important than the quantity of generated images. Traditional data augmentation (rotations, flips, brightness adjustments) remains the most reliable strategy for small datasets.

From an applied computing perspective, these findings reinforce the potential of pretrained models to support automated postharvest inspection processes. However, they also highlight the need for more systematic evaluations under real operating conditions and for robustness analyses.

RQ4: Farm Monitoring and Management Through IoT and Web Platforms

The critical analysis of the seven studies in Axis C reveals a systematic disconnect between the technical advancement of IoT systems and their evaluation from the end-user perspective. The reported systems prioritize data acquisition, transmission, and availability, yet none explicitly evaluates ease of use, farmers' cognitive workload, or interface suitability for rural contexts with limited digital literacy. This omission is particularly critical considering that usability is a key determinant of technology adoption in the agricultural sector.

The absence of formal evaluations based on ISO 9241 in the analyzed corpus contrasts with the broad availability of tools such as the System Usability Scale (SUS), the NASA-TLX questionnaire, and heuristic inspection methods, which could easily be integrated into the evaluation protocols of agricultural IoT platforms. The emergence of approaches such as digital twins (Archambault, 2024) represents a promising trend toward simulation and decision support, although these approaches remain at an exploratory stage. Within the conceptual framework of Agriculture 4.0, the convergence of IoT, AI, cloud computing, and digital twins is reshaping agricultural management, and web platforms should be evaluated within this broader sociotechnical ecosystem.

Synthesis of implications and research gaps

Overall, the discussion reveals three critical cross-cutting gaps. First, the external validity gap: deep learning models and IoT systems are evaluated under controlled environments, and their transfer to real agricultural conditions requires more rigorous evaluation protocols and independent field datasets. Second, the standardization gap: the heterogeneity of datasets, metrics, and experimental protocols prevents direct comparison among studies and the identification of optimal approaches; the creation of public benchmarks for regionally relevant agricultural crops is an urgent need. Third, the usability gap: user-centered design and formal usability evaluation (ISO 9241, SUS, Nielsen's heuristics) remain absent from the analyzed IoT solutions, compromising the effective adoption of these technologies by agricultural stakeholders.

Limitations of the study

Despite the methodological rigor applied in this SLR, it is important to acknowledge several limitations that directly affect the interpretation of the results and the strength of the conclusions.

First, the SLR is based exclusively on studies published in English. This restriction introduces a language bias that may be particularly significant in the context of Latin American and Andean agricultural computing, where a growing body of research is published in Spanish and Portuguese in repositories such as Latindex, Redalyc, and CLACSO. The exclusion of this regional and gray literature may bias the findings toward technological approaches developed in agricultural contexts different from those of Latin America, thereby affecting the local relevance of the conclusions regarding model generalization and usability.

Second, the high heterogeneity among the 16 included studies, in terms of datasets, experimental protocols, and evaluation metrics, limited the possibility of performing direct quantitative comparisons or a statistical meta-analysis. Consequently, the identified trends are qualitative in nature and do not allow the establishment of the absolute superiority of one technological approach over another, thereby limiting the certainty of the comparative conclusions.

A third relevant limitation is the dependence on ad hoc datasets in the visual diagnosis and postharvest classification axes. The scarcity of standardized public datasets means that the reported accuracy results cannot be independently replicated or compared, weakening the internal validity of the conclusions regarding the performance of specific models.

Fourth, most of the analyzed studies evaluate their proposed solutions under controlled or semi-controlled environments, with limited validation under real field conditions. Therefore, the conclusions regarding the effectiveness of deep learning models and IoT systems are primarily valid within the experimental setting and cannot be directly extrapolated to real agricultural production scenarios without additional considerations regarding robustness and domain adaptation.

Finally, within the IoT monitoring axis, the limited evaluation of usability and the systematic absence of standards such as ISO 9241 prevent the formulation of robust conclusions regarding the user experience of agricultural web platforms. This constitutes an inherent methodological limitation of the analyzed corpus and does not necessarily reflect the state of the field as a whole, but rather the publication bias toward technical aspects over evaluative ones.

 Revista Científica CEDIA - ISSN: 3151-8435

Conclusions

This SLR analyzed and synthesized the scientific evidence from 16 primary studies published between 2017 and 2025, organized into three thematic axes on agricultural computing, leading to the following specific conclusions derived from the comparative analysis.

Regarding RQ1, the six studies in Axis A confirm that CNNs with attention mechanisms constitute the dominant architecture for the visual diagnosis of diseases in agricultural leaves and fruits. The reported accuracy values vary widely across studies, ranging from 72.31% (Noon et al., 2022) and 84.29% (Faisal et al., 2023) under field conditions or with small proprietary datasets, to 99.77% (Ouamane et al., 2025) using the large-scale public PlantVillage dataset. This variability reflects the influence of the dataset rather than the absolute superiority of any particular architecture. Vision transformers (Ouamane et al., 2025) and object detectors such as YOLOX (Noon et al., 2022) emerge as competitive alternatives for localization and classification tasks involving multiple coexisting diseases. However, none of the studies reports systematic robustness evaluations under real field conditions, representing the most critical gap within this axis.

Regarding RQ2, most of the studies in Axis A (four out of six) rely on ad hoc datasets, although two use well-established public datasets (RoCoLe and PlantVillage). The widespread use of data augmentation as a compensatory mechanism is a cross-cutting finding. The reported accuracy values (72% to 99.77%) should be interpreted in light of the evaluation context, as they do not necessarily reflect performance under real operating conditions. The standardization of datasets and the adoption of cross-validation protocols are concrete recommendations derived from this axis.

With respect to RQ3, the three studies in Axis B provide mixed evidence regarding synthetic data augmentation. Prakasa et al. (2021) confirm that deep learning models achieve F1-scores above 0.90 (best: 0.983) in maize classification. Kant (2024) demonstrates that MobileNetV2 achieves an accuracy of 0.99 in rice grain classification using transfer learning. Pasaribu et al. (2025) reveal that GAN-based data augmentation reduced accuracy compared with the original dataset (97.37%), calling into question the effectiveness of synthetic augmentation without image quality control. Nevertheless, the absence of evaluations of efficiency on edge devices and robustness under variations in image acquisition limits the practical validity of these findings.

Finally, in response to RQ4, the seven studies in Axis C demonstrate that agricultural IoT solutions prioritize technical functionality (data acquisition, transmission, and availability) over user-centered evaluation. The complete absence of evaluations based on ISO 9241 in the analyzed corpus represents a structural gap between technological development and effective adoption by farmers, particularly in contexts with low levels of digital literacy.

Future work

Based on the identified findings and limitations, four priority directions for future work are proposed. The first involves the development and adoption of public, standardized, and geographically diverse datasets for regionally relevant agricultural diseases (coffee leaf rust, cocoa moniliasis, and postharvest maize defects) to enable the evaluation of the generalization and reproducibility of visual diagnosis and classification models. The second involves validating models directly under real field conditions by incorporating environmental variability, visual noise, and operational constraints, using protocols that include robustness metrics, energy efficiency, and inference time on edge devices.

The third direction involves the systematic integration of usability criteria (ISO 9241, SUS, Nielsen's heuristics) from the early stages of designing agricultural web and IoT platforms, incorporating participatory design methodologies involving farmers. Finally, the fourth direction focuses on exploring emerging approaches such as foundation models with fine-tuning for agricultural domains, the integration of digital twins with predictive AI, and the application of federated learning techniques to preserve the privacy of agricultural data in distributed monitoring systems.

Acknowledgments

We extend our sincere appreciation to the Universidad Nacional de Loja (UNL), and particularly to the Computer Science program and the Information Technology and Computing Research Group (GITIC), for their invaluable support and the resources provided for conducting this research. This article is the result of an academic effort developed within the framework of the course "Technological Projects 1," taught during the academic period from September 2025 to February 2026. We are grateful for the guidance and intellectual encouragement received, which have been fundamental to the completion of this work and to our growth as researchers in the field of computer science.

Conflict of interest statement

The authors declare that they have no conflicts of interest.

Funding sources

This work is part of the research project "Development of an Artificial Intelligence Framework to Integrate and Strengthen the Core Functions of the Universidad Nacional de Loja and Its Scalability to the Public Universities of Ecuador" (Project Code: 32-DI-FEIRNR-2025), fully funded by the Universidad Nacional de Loja.

Declaration on the use of artificial intelligence

The authors declare the use of generative artificial intelligence during the research and writing process. According to the GAIDeT (2025) taxonomy, the following tasks were delegated to generative AI tools under full human supervision:

- Idea generation, definition of the research objective, formulation of research questions and hypotheses, writing of the literature review, research design, data curation and organization, data analysis, text generation, review and editing, text synthesis, formulation of conclusions, translation, and identification of limitations.

The generative AI tools used were ChatGPT 5.2 and NotebookLM.

Full responsibility for the final manuscript rests entirely with the authors.

The generative AI tools are not listed as authors and assume no responsibility for the final results.

Author contributions (CRediT):

Luis Chamba-Eras: Conceptualization; methodology; validation; formal analysis; investigation; resources; data curation; writing – original draft; writing – review and editing; supervision; project administration.

Sergio Jumbo-Gonza: Validation; formal analysis; investigation; resources; data curation; writing – review and editing; project administration.

Walter Yaguana-Alejandro: Validation; formal analysis; investigation; resources; data curation; writing – review and editing; project administration.

Geovanny Romero-Suquilanda: Validation; formal analysis; investigation; resources; data curation; writing – review and editing; project administration.

Gerardo Herrera-Campoverde: Validation; formal analysis; investigation; resources; data curation; writing – review and editing; project administration.

References

Abinaya, S., Kumar, K. U., & Alphonse, A. S. (2023). Cascading autoencoder with attention residual u-net for multi-class plant leaf disease segmentation and classification. *IEEE Access, 11*, 98153-98170. https://doi.org/10.1109/ACCESS.2023.3312718

Aliparo, J., Santos, C. D., Sese, C. M., Trinidad, J. ver, Beano, M. G., Capuno, M. E. A., Medina, O., Sigue, A., & Soriano, A. (2022). IoT-based assessment and monitoring of NPK content and fertility condition of soil. En *TENCON 2022 - 2022 IEEE Region 10 Conference (TENCON)* (pp. 1-6). https://doi.org/10.1109/TENCON55691.2022.9978040

Archambault, P. (2024). Co-simulation and crop representation for digital twins of controlled environment agriculture systems. En *Proceedings of the ACM/IEEE 27th International Conference on Model Driven Engineering Languages and Systems* (pp. 196-199). https://doi.org/10.1145/3652620.3676879

Bhadra, T., Mujakkir-Ul-Islam, & Alam, N. (2025). IoT-based real-time data monitoring of a green house farm with offline monitoring capability. En *2025 4th International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)* (pp. 111-116). https://doi.org/10.1109/ICREST63960.2025.10914444

Borromeo, K. O., Bonotan, M. R., Llagas, A. P., Juan, J. M. O., Sanchez, P. D. C., & Berro, N. K. A. (2025). Classification of cacao pods ripeness using convolutional neural networks. *Procedia Computer Science, 257*, 1197-1204. https://doi.org/10.1016/j.procs.2025.03.160

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software, 80*(4), 571-583. https://doi.org/10.1016/j.jss.2006.07.009

Faisal, M., Leu, J.-S., & Darmawan, J. T. (2023). Model Selection of Hybrid Feature Fusion for Coffee Leaf Disease Classification. *IEEE Access, 11*, 62281-62291. https://doi.org/10.1109/ACCESS.2023.3286935

GAIDeT. (2025). Generative Artificial Intelligence Disclosure Taxonomy (GAIDeT). https://panbibliotekar.github.io/gaidet-declaration/

Garg, D., & Gill, K. S. (2024). *Revolutionizing rice quality control: MobileNetV2 for grain classification. En 2024 2nd International Conference on Advances in Computation, Communication and Information Technology* (ICAICCIT) (pp. 809-812). https://doi.org/10.1109/ICAICCIT64383.2024.10912366

Hewage, P., Anderson, M., & Fang, H. (2017). An agile farm management information system framework for precision agriculture. En *Proceedings of the 9th International Conference on Information Management and Engineering* (pp. 75-80). https://doi.org/10.1145/3149572.3149583

Kant, V. (2024). Efficient rice grain classification using MobileNetV2: A deep learning approach. En *2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS)* (pp. 989-993). https://doi.org/10.1109/ICTACS62700.2024.10840597

Kitchenham, B., Pretorius, R., Budgen, D., Pearl Brereton, O., Turner, M., Niazi, M., & Linkman, S. (2010). Systematic literature reviews in software engineering – A tertiary study. *Information and Software Technology, 52*(8), 792-805. https://doi.org/10.1016/j.infsof.2010.03.006

Konakalla, S., Kunchala, L. V., & Yellampalli, S. S. (2021). Smart Agriculture using IoT & Filtering Technology. En *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-5). https://doi.org/10.1109/ICCCNT51525.2021.9579730

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Muthusamy, S., & Ramu, S. P. (2025). A dynamic attention-guided deep learning framework (d-pan) with reduced complexity for fine grained severity-aware classification of micronutrient deficiency and disease detection in banana and coffee leaves. *IEEE Access, 13*, 153474-153497. https://doi.org/10.1109/ACCESS.2025.3599359

Noon, S. K., Amjad, M., Qureshi, M. A., & Mannan, A. (2022). Handling Severity Levels of Multiple Co-Occurring Cotton Plant Diseases Using Improved YOLOX Model. *IEEE Access, 10*, 134811-134825. https://doi.org/10.1109/ACCESS.2022.3232751

Nugroho, L. E., Pratama, A. G. H., Mustika, I. W., & Ferdiana, R. (2017). Development of monitoring system for smart farming using Progressive Web App. En *2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE)* (pp. 1-5). https://doi.org/10.1109/ICITEED.2017.8250513

Ouamane, A., Chouchane, A., Himeur, Y., Miniaoui, S., Atalla, S., Mansoor, W., & Al-Ahmad, H. (2025). Optimized vision transformers for superior plant disease detection. *IEEE Access, 13*, 48552-48570. https://doi.org/10.1109/ACCESS.2025.3547416

Pasaribu, F. B., Dewi, L. J. E., Aryanto, K. Y. E., & Varnakovida, P. (2025). Effect of synthetic data augmentation on plant classification accuracy using MobileNetV2, EfficientNet-B0, and ResNet-18. En *2025 IEEE 5th International Conference on Software Engineering and Artificial Intelligence (SEAI)* (pp. 1-6). https://doi.org/10.1109/SEAI65851.2025.11108883

Prakasa, E., Prajitno, D. R., Nur, A., Sulistyo, K. A., & Rachmawati, E. (2021). Quality categorisation of corn (Zea mays) seed using feature-based classifier and deep learning on digital images. En *2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)* (pp. 498-505). https://doi.org/10.1109/3ICT53449.2021.9581464

Rahman, Z. U., Asaari, M. S. M., Ibrahim, H., Abidin, I. S. Z., & Ishak, M. K. (2024). Generative Adversarial Networks (GANs) for image augmentation in farming: A review. *IEEE Access, 12*, 179912-179943. https://doi.org/10.1109/ACCESS.2024.3505989

Rao, A., Shao, H., & Yang, X. (2019). The design and implementation of smart agricultural management platform based on UAV and wireless sensor network. En *2019 IEEE 2nd International Conference on Electronics Technology (ICET)* (pp. 248-252). https://doi.org/10.1109/ELTECH.2019.8839480

Yadav, S., T S, D., & Kokil, P. (2026). An IoT-enabled smart sensor system for real-time monitoring of soil and air conditions with integrated web-based farming guidance. *IEEE Sensors Journal, 26*(1), 1267-1275. https://doi.org/10.1109/JSEN.2025.3619663

Yeong, D. J., Panduru, K., & Walsh, J. (2024). Smart agriculture: Software platform for telematics monitoring in farm machinery. En *2024 35th Irish Signals and Systems Conference (ISSC)* (pp. 1-6). https://doi.org/10.1109/ISSC61953.2024.10602840

Yue, X., Qi, K., Na, X., Liu, Y., Yang, F., & Wang, W. (2025). Deep learning for recognition and detection of plant diseases and pests. *Neural Computing and Applications, 37*(17), 11265-11310. h https://doi.org/10.1007/s00521-025-11125-5