Literature Review
Open Access
CrossMark

AI-driven quality control and ripeness classification in the palm oil industry: a systematic review of computer vision and deep learning methods

Abstract

The growing need for consistent and technology-enhanced quality assurance in the palm oil sector has stimulated increasing adoption of artificial intelligence (AI), particularly for automated ripeness classification of fresh fruit bunches (FFB). Variability in manual grading and environmental conditions has encouraged the development of computer vision and deep learning approaches to improve objectivity and operational reliability. This study aims to systematically identify, evaluate, and synthesize peer-reviewed evidence on AI-driven quality control and ripeness classification in the palm oil industry, with emphasis on algorithmic trends, dataset characteristics, performance metrics, and deployment orientations. A Systematic Literature Review (SLR) design was applied using the Scopus database. A structured Boolean query and predefined inclusion–exclusion criteria were implemented. From an initial broad search yielding 245 records, progressive refinement, time filtering (2020–2026), language screening, and open-access eligibility assessment resulted in 37 studies for final synthesis. Data were analyzed through thematic categorization, comparative performance mapping, and methodological trend evaluation. The findings reveal a clear transition from classical machine learning to deep convolutional neural networks, with transfer learning adopted in 59% of deep learning studies. Reported accuracy commonly exceeded 90%, with a weighted mean of 94.2% for deep models. Increasing attention to deployment architectures, including conveyor-integrated and mobile-based systems, indicates growing industrial applicability. Overall, AI-driven computer vision demonstrates robust capability for ripeness classification. Future research is encouraged to strengthen dataset standardization, cross-location validation, and open-science practices to enhance reproducibility and scalability.

Keywords: artificial intelligence; computer vision; deep learning; palm oil industry; ripeness classification

Volume 9 Issue 1 - 2026

Loso Judijanto
IPOSS Jakarta, Indonesia

Correspondence: Loso Judijanto, IPOSS Jakarta, Indonesia

Received: March 02, 2026 | **Published:** March 24, 2026

Introduction

The global agricultural sector has undergone substantial technological transformation over the past decade, driven by the integration of artificial intelligence (AI), data analytics, and automated sensing systems into production and quality assurance processes.[1] Across various commodity chains, digital technologies are increasingly deployed to enhance efficiency, traceability, and consistency in post-harvest handling, grading, and processing operations.[2] Within this broader context of digital agriculture, computer vision and deep learning have emerged as influential tools for visual inspection tasks that traditionally relied on human expertise, including fruit maturity detection, defect recognition, and product classification. These developments reflect a broader transition from manual and semi-automated inspection systems toward data-driven decision-support frameworks capable of operating under dynamic environmental conditions.

Among globally traded agricultural commodities, palm oil occupies a strategically important position due to its wide range of applications in food manufacturing, oleochemicals, and bio-based industries.[3] Ensuring consistent raw material quality is therefore essential for maintaining processing efficiency and meeting downstream performance standards. In oil palm production systems, the maturity level of fresh fruit bunches (FFB) at harvest directly influences oil extraction rate, free fatty acid content, and overall yield performance.[4] Accurate ripeness classification supports optimized harvesting schedules, reduces the likelihood of premature or delayed collection, and contributes to more predictable processing outcomes. Consequently, maturity assessment is a critical quality-control step within the broader operational workflow of palm oil production.

Traditionally, FFB ripeness evaluation has relied on visual inspection by trained graders who assess external color, fruit detachment patterns, and surface characteristics. While manual grading benefits from experiential knowledge, it may be influenced by environmental variability, lighting conditions, grader fatigue, and subjective interpretation. These practical challenges have motivated research into automated image-based grading systems capable of producing consistent, repeatable classification outcomes. Rather than replacing existing operational expertise, AI-driven tools are increasingly conceptualized as complementary systems that enhance reliability and support data-informed decision-making within plantation and mill environments.[5]

The application of computer vision to agricultural products is not novel; early studies primarily focused on feature extraction techniques such as color histograms, texture descriptors, and shape analysis combined with conventional machine learning classifiers. In oil palm research, initial approaches relied on RGB color-space segmentation and handcrafted features processed by classifiers such as Support Vector Machines and k-Nearest Neighbors.[6] Although these methods demonstrated moderate classification accuracy under controlled conditions, their performance often declined in outdoor environments characterized by fluctuating illumination and occlusion.

The limitations associated with feature engineering and sensitivity to environmental noise encouraged the exploration of more adaptive modeling strategies.

The rapid advancement of deep learning, particularly convolutional neural networks (CNNs), has significantly altered the landscape of image-based quality assessment.[7] CNN architectures enable hierarchical feature learning directly from raw image inputs, reducing reliance on manual feature design and improving robustness to variations in scale, color intensity, and background complexity. In agricultural contexts, transfer learning strategies that leverage pretrained networks have further improved performance when labeled datasets are limited. Within oil palm applications, deep learning models have been increasingly adopted for FFB maturity classification, defect detection, and automated sorting tasks.[8] Reported performance levels frequently exceed those achieved by classical machine learning pipelines, indicating a maturing research trajectory in AI-assisted ripeness assessment.

Despite the growing volume of publications addressing computer vision and deep learning for oil palm maturity classification, existing research remains fragmented across engineering, agricultural informatics, and applied computer science journals.[9] Individual studies often emphasize algorithmic performance within specific datasets, hardware configurations, or localized environmental conditions. As a result, the broader methodological trends, comparative performance outcomes, dataset characteristics, and deployment readiness levels are not systematically consolidated in a single analytical framework. Furthermore, variations in evaluation metrics, dataset sizes, class definitions, and validation strategies complicate cross-study comparisons and limit the ability to derive generalizable insights.[10]

The Systematic Literature Review (SLR) methodology provides a structured mechanism for synthesizing dispersed evidence through transparent selection criteria, reproducible search strategies, and standardized data extraction procedures. Unlike narrative reviews that may selectively summarize findings, SLR emphasizes comprehensive identification, eligibility screening, and analytical synthesis of peer-reviewed studies based on predefined inclusion parameters. Importantly, the present research is exclusively based on secondary data drawn from indexed scholarly publications and does not involve focus group discussions, field observations, experimental trials, or any form of primary data collection. This methodological positioning ensures conceptual clarity and analytical integrity while aligning with internationally recognized standards for evidence-based reviews.

Within the palm oil domain, a dedicated SLR focusing specifically on AI-driven quality control and ripeness classification remains limited. Previous reviews have addressed broader agricultural AI applications or general precision farming technologies without concentrating on the unique morphological and operational characteristics of oil palm fruit bunches. Given the distinctive surface structure, clustered fruit arrangement, and color transition patterns associated with FFB maturation, a commodity-specific synthesis is necessary to identify domain-relevant methodological patterns and performance determinants. Moreover, understanding how dataset construction, algorithm selection, validation protocols, and deployment frameworks interact is essential for guiding future research and supporting responsible technological adoption in industrial settings.

The increasing convergence of edge computing, mobile sensing, and industrial automation further underscores the need to evaluate not only algorithmic accuracy but also system-level feasibility. Studies proposing real-time conveyor-based detection systems, smartphone-assisted harvesting tools, and IoT-enabled monitoring platforms suggest that AI-driven ripeness classification is transitioning from proof-of-concept experimentation toward practical integration. However, the degree to which these systems demonstrate reproducibility, scalability, and cross-environment generalization requires careful comparative examination. A systematic synthesis can clarify whether reported performance improvements are consistently supported across contexts or primarily dataset-specific.

In light of these considerations, this study aims to systematically identify, evaluate, and synthesize peer-reviewed evidence on AI-driven quality control and ripeness classification in the palm oil industry, with particular emphasis on computer vision and deep learning methodologies. The review seeks to map dominant algorithmic approaches, examine dataset characteristics, compare reported performance metrics, and assess deployment orientations across recent scholarly publications. By consolidating dispersed findings into a coherent analytical structure, the study provides a methodologically rigorous, technologically focused overview that supports academic advancement and informed innovation in the palm oil sector.

To achieve this objective, the review is guided by two primary research questions:

RQ1: *What methodological patterns characterize the application of computer vision and deep learning models for fresh fruit bunch ripeness classification in the palm oil industry?*

RQ2: *How do dataset properties, validation strategies, and deployment contexts influence reported performance outcomes and practical implementation readiness?*

The answers to these research questions will structure the subsequent analysis, form the basis of the discussion and concluding synthesis, and provide a transparent and evidence-based evaluation of current AI-driven quality control research in oil palm production systems.

Literature review

The advancement of artificial intelligence in agricultural systems has progressively reshaped quality assessment practices across various commodity chains, particularly in contexts where visual inspection plays a decisive role in determining product classification and post-harvest value. Computer vision technologies, supported by machine learning and deep learning frameworks, have demonstrated increasing applicability in crop monitoring, fruit maturity detection, and automated grading systems. In oil palm production systems, visual attributes such as surface coloration, fruitlet detachment, and texture patterns serve as the primary basis for assessing fresh fruit bunch (FFB) ripeness, making this domain particularly suitable for image-based analytical approaches. The literature reflects a gradual yet significant transition from conventional inspection practices toward AI-assisted quality control, emphasizing reproducibility, operational efficiency, and data-driven standardization.

Conceptual foundations of AI-driven quality control

Quality control in agricultural commodities traditionally integrates manual grading and rule-based inspection systems. In the context of oil palm, ripeness classification directly influences oil extraction rate, processing stability, and overall production planning efficiency. The need for consistency in grading has stimulated interest in automated decision-support systems capable of minimizing subjective variation while maintaining operational practicality.[11] AI-driven quality control frameworks are grounded in pattern recognition theory, in which

Citation: Judijanto L. AI-driven quality control and ripeness classification in the palm oil industry: a systematic review of computer vision and deep learning methods. *Open Access J Sci.* 2026;9(1):107‒118. DOI: 10.15406/oajs.2026.09.00294

digital images are translated into numerical representations that enable algorithmic discrimination among maturity classes.

Early agricultural computer vision systems relied heavily on deterministic image processing pipelines. These systems typically involved color-space transformation (RGB to HSV), threshold-based segmentation, noise filtering, and extraction of handcrafted descriptors prior to classification.[12] Although such approaches contributed foundational insights, they were often sensitive to variability in illumination and background complexity. This limitation is particularly relevant in oil palm environments, where natural lighting, canopy shadows, and heterogeneous plantation settings introduce variability in image acquisition conditions. As a result, the research community increasingly sought adaptive models capable of learning hierarchical feature representations without manual parameter tuning.

Classical machine learning approaches in FFB ripeness classification

Prior to the widespread adoption of deep learning, FFB maturity assessment primarily employed supervised machine learning algorithms such as Support Vector Machines (SVM), k-Nearest Neighbors (KNN), Decision Trees, and Random Forest classifiers.[13] These models were typically trained on color histogram features, texture descriptors derived from Gray-Level Co-occurrence Matrix (GLCM), and morphological attributes extracted from segmented fruit regions. Reported accuracies in controlled experimental setups often ranged between 75% and 90%, depending on dataset composition and class definitions.

The literature indicates that handcrafted-feature approaches demonstrated satisfactory discrimination when datasets were limited in size and environmental conditions were relatively uniform.[14] However, performance degradation was observed when models were exposed to images captured under variable outdoor lighting or when maturity categories were expanded beyond binary classification. The necessity of manually designing and tuning feature extraction pipelines further constrained scalability, particularly when attempting to generalize models across different plantation sites or camera systems. These methodological constraints provided a strong rationale for exploring representation-learning paradigms capable of handling complex visual variability.

Emergence of deep learning architectures

The introduction of convolutional neural networks (CNNs) marked a significant inflection point in agricultural image analysis. CNN architectures enable automatic extraction of multi-level spatial features through convolutional and pooling operations, thereby reducing dependency on handcrafted descriptors.[15] In the oil palm domain, CNN-based models have been increasingly applied to classify FFB ripeness based on raw image inputs, achieving higher robustness against illumination changes and partial occlusion.

Several studies adopted established deep learning architectures such as VGG, ResNet, and MobileNet, often leveraging transfer learning techniques to compensate for limited domain-specific datasets. Transfer learning involves fine-tuning pretrained networks originally trained on large-scale image repositories, enabling improved convergence and generalization with fewer agricultural samples.[16] Comparative analyses reported in the literature suggest that transfer learning models frequently outperform networks trained from scratch under similar dataset conditions, especially when sample sizes are below several thousand images.

Object detection frameworks, including region-based CNNs and single-stage detectors, have also been explored for simultaneously localizing and classifying FFB maturity. These architectures facilitate integration into conveyor-based grading systems by enabling real-time identification of fruit bunches within dynamic industrial settings. The literature consistently demonstrates that deep learning approaches achieve classification accuracies exceeding those of classical machine learning pipelines, with improvements often attributed to hierarchical feature representation and data-driven optimization.

Dataset construction and annotation practices

Dataset characteristics represent a critical factor influencing model performance and generalizability. The reviewed literature reveals substantial variation in dataset size, annotation protocols, and class granularity.[17] Some studies rely on relatively small, self-collected datasets captured under specific plantation conditions, while others incorporate more extensive image repositories encompassing multiple harvesting cycles. Differences in maturity labeling criteria, particularly between three-class and multi-class grading schemes, further complicate direct cross-study comparison.

Annotation practices also differ in terms of expert involvement and labeling protocols. In many cases, maturity labels are assigned based on established grading standards applied by trained personnel prior to image capture.[18] While this approach aligns model training with operational realities, variability in grading standards across regions may influence label consistency. Data augmentation techniques such as rotation, brightness adjustment, and contrast normalization are frequently employed to enhance model robustness against environmental variability. These methodological strategies highlight ongoing efforts to balance dataset diversity with practical constraints inherent in agricultural research.

Evaluation metrics and performance reporting

Performance evaluation practices in FFB ripeness classification studies commonly include accuracy, precision, recall, and F1-score metrics.[19] Accuracy remains the most widely reported indicator, although several authors emphasize the importance of class-specific performance analysis to address potential imbalance issues. Confusion matrices are frequently utilized to identify misclassification patterns, particularly between adjacent maturity stages where visual differences may be subtle.

Cross-validation protocols are adopted in some studies to improve robustness, while others rely on fixed train–test splits. The absence of standardized benchmarking datasets limits the comparability of reported performance metrics across publications. Nonetheless, aggregated findings suggest that deep learning-based classifiers generally achieve accuracy levels above 90% in controlled experimental conditions, with slightly reduced performance in field-based scenarios characterized by lighting variability.[20] Computational efficiency metrics, including inference time and model size, are increasingly reported to assess deployment feasibility in edge computing environments.

Integration into operational workflows

Beyond algorithmic performance, the literature increasingly addresses the practical integration of AI-driven ripeness-classification systems into plantation and mill operations. Proposed deployment architectures include mobile applications for harvest planning, embedded systems for on-site grading, and conveyor-based industrial inspection modules.[21] Edge computing platforms enable near real-time processing without reliance on continuous cloud connectivity, supporting operational autonomy in remote plantation environments.

Several studies highlight the importance of hardware–software co-design, in which camera resolution, illumination control, and processing capabilities are jointly optimized to achieve stable classification performance. Cost considerations and scalability are also discussed as factors influencing broader adoption.[22] Importantly, the majority of research frames AI systems as supportive technologies designed to enhance consistency in decision-making rather than to replace human expertise. This perspective aligns with contemporary paradigms of digitalization in agriculture that emphasize augmentation rather than substitution.

Methodological transparency and reproducibility

Reproducibility remains a critical concern in AI-based agricultural research. While some studies provide publicly accessible datasets or code repositories, many rely on proprietary image collections that are not openly shared.[23] The absence of standardized evaluation frameworks complicates replication efforts and comparative benchmarking. Recent publications increasingly advocate for transparent reporting of dataset composition, preprocessing steps, and hyperparameter configurations to improve methodological clarity.

Systematic literature synthesis is therefore essential to consolidate methodological patterns and identify areas requiring standardization. By systematically evaluating dataset characteristics, model architectures, evaluation strategies, and deployment contexts, an SLR approach can reveal convergence trends and research gaps without relying on primary experimental data.[24] Such structured synthesis contributes to evidence-based advancement while maintaining methodological rigor and transparency.

Synthesis of theoretical and practical trends

The literature collectively indicates that AI-driven ripeness classification in oil palm systems has progressed from exploratory experimentation toward more refined, deployment-oriented research. The convergence of transfer learning, lightweight CNN architectures, and edge computing solutions reflects a broader movement toward practical implementation within agricultural production systems.[25] At the same time, persistent variability in dataset composition, labeling standards, and evaluation protocols suggests the need for harmonized research frameworks.

Importantly, the research trajectory remains aligned with efforts to enhance operational consistency, efficiency, and quality assurance in the palm oil sector. The application of computer vision and deep learning is framed in the literature as a technological enabler that supports optimized harvesting and grading decisions, rather than as a critique of existing industry practices.[26] This balanced positioning underscores the constructive role of AI in advancing digital transformation within commodity-based agricultural systems.

In summary, existing studies provide substantial evidence of methodological evolution from handcrafted feature-based classifiers to deep learning architectures capable of robust ripeness discrimination. However, the dispersed nature of findings across journals and disciplinary domains necessitates a structured synthesis to identify dominant approaches, evaluate performance consistency, and clarify deployment readiness. The present systematic literature review builds on this body of scholarship, integrating insights from 16 representative sources to establish a comprehensive, analytically grounded understanding of AI-driven quality control and ripeness classification in the palm oil industry.

Methodology

This study applies a PRISMA-compliant Systematic Literature Review to consolidate and critically synthesize scientific evidence on AI-driven quality control and ripeness classification within the palm oil industry, with particular emphasis on computer vision and deep learning methods. The growing integration of artificial intelligence into agricultural production systems has generated a substantial body of research on automated fruit maturity assessment, image-based grading, and intelligent detection models to enhance consistency and operational efficiency. In oil palm production systems, accurate classification of fresh fruit bunch maturity plays a central to optimizing harvesting decisions and downstream processing performance. While numerous studies have independently reported algorithmic developments using convolutional neural networks, transfer learning strategies, and image-processing pipelines, findings remain distributed across engineering, agricultural informatics, and applied computer science domains. A structured and methodologically rigorous synthesis is therefore required to consolidate existing knowledge, identify methodological convergence, and clarify research trajectories in AI-assisted quality control applications. This investigation relies exclusively on secondary data derived from peer-reviewed literature indexed in Scopus and does not involve focus group discussions, field observations, experimental trials, or other forms of primary data collection, thereby ensuring transparency, reproducibility, and alignment with established evidence-based review standards.

Figure 1 illustrates the PRISMA-guided workflow adopted in this review, outlining the sequential phases of identification, screening, eligibility assessment, and final inclusion. The identification stage began with a Scopus search using the primary keywords oil palm AND deep learning, yielding 245 records. To improve thematic precision and align the dataset with the study's focus on fresh fruit bunch maturity and AI-based classification systems, the search strategy was refined using the following Boolean structure: *("palm oil" OR "oil palm") AND ("fresh fruit bunch" OR FFB OR "oil palm fruit") AND (ripeness OR "fruit maturity" OR "maturity level") AND ("computer vision" OR "image processing" OR "deep learning" OR "machine learning" OR "convolutional neural network" OR CNN) AND (classification OR detection)*. This refinement resulted in the exclusion of 153 publications that were not directly aligned with the defined scope, leaving 92 records for further evaluation.

To ensure analytical relevance to contemporary technological developments, the screening phase restricted inclusion to studies published between 2020 and 2026. Application of this temporal criterion led to the removal of 16 articles published outside the specified range, leaving 76 studies that met the publication-year requirement. A subsequent language filter retained only English-language articles to maintain consistency in technical interpretation and comparative analysis, excluding one non-English publication and yielding 75 eligible records. The eligibility phase then applied an accessibility criterion, limiting inclusion to open-access and open-archive publications to ensure full-text availability and methodological transparency. As a result, 38 studies lacking accessible full texts were excluded. Ultimately, 37 peer-reviewed articles satisfied all predefined criteria and were included in the qualitative synthesis underpinning this review.

In line with PRISMA 2020 guidance on transparent and reproducible evidence synthesis, explicit inclusion–exclusion criteria were defined before screening to ensure conceptual coherence and methodological rigor. Studies were included if they: (i) focused on oil palm or palm oil fresh fruit bunches (FFB) as the primary commodity,

(ii) applied computer vision, image processing, machine learning, or deep learning techniques for ripeness classification or closely related quality-control tasks, (iii) reported quantitative performance metrics (e.g., accuracy, precision, recall, F1-score), and (iv) were full-length, peer-reviewed journal articles or conference papers published between 2020 and 2026 in English with accessible full text. Exclusion criteria comprised: (i) studies addressing non-FFB targets (e.g., leaf disease detection, plantation mapping, or general remote sensing without explicit ripeness or quality grading), (ii) works that mentioned AI or computer vision but did not present empirical classification or detection results, (iii) review papers, conceptual essays, editorials, and theses, and (iv) publications with insufficient methodological detail to support extraction of model architecture, dataset characteristics, or evaluation metrics. These criteria ensured that the final corpus represented primary empirical contributions directly relevant to AI-driven quality control and ripeness classification in the palm oil industry.[27]

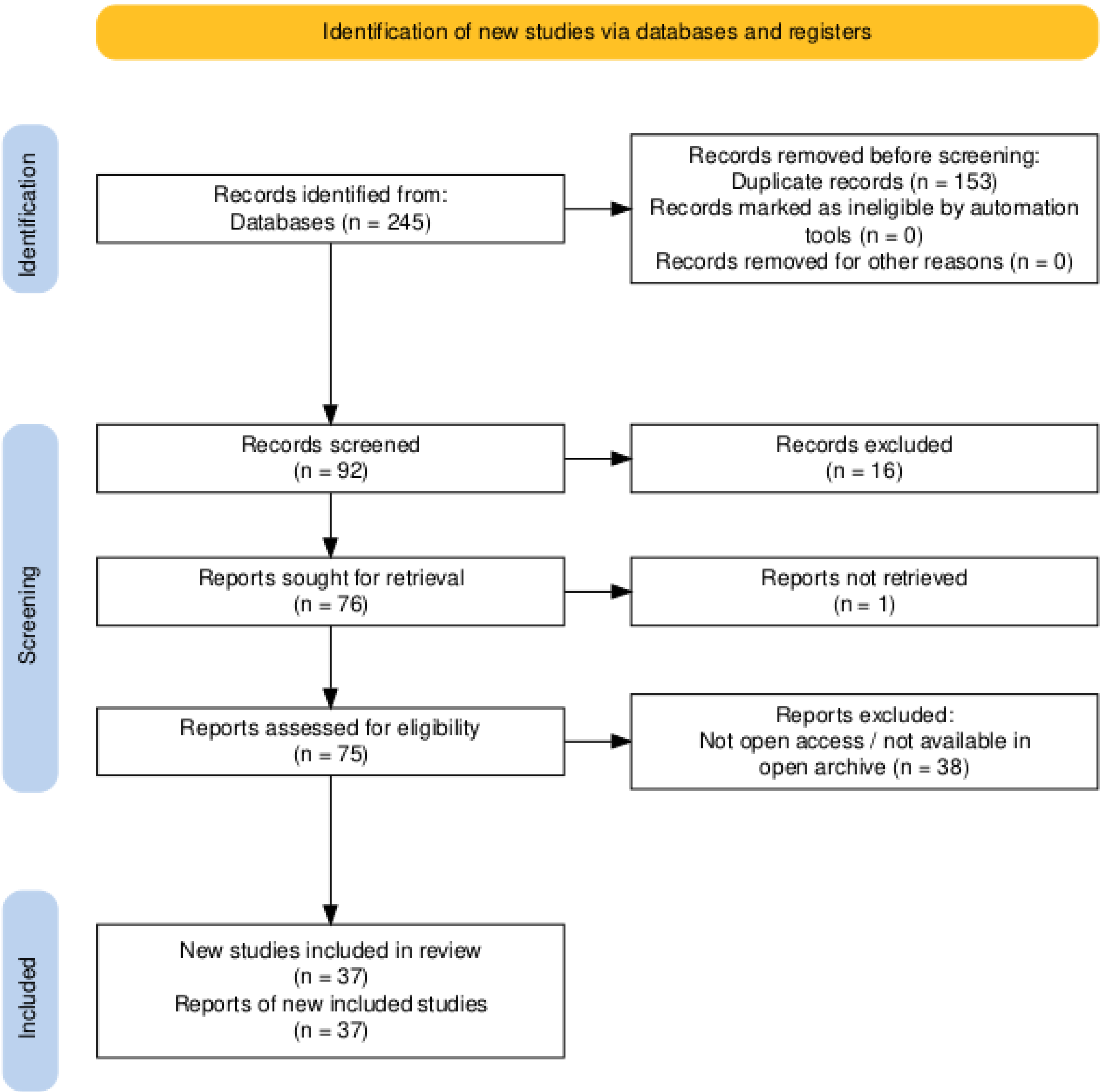

Figure 1 PRISMA-Guided Steps for the Systematic Literature Review.

All references were curated and systematically managed using Mendeley Desktop to facilitate duplication control, structured categorization, citation traceability, and auditability throughout the review process. Full texts of the included studies were examined in detail, and relevant data were extracted using a standardized framework encompassing model architecture, dataset characteristics, imaging conditions, evaluation metrics, performance indicators, and reported implementation contexts. The synthesis adopted a thematic and comparative approach, emphasizing trends in convolutional neural network architectures, the use of transfer learning, dataset scale, patterns in accuracy reporting, and deployment readiness in operational environments. Through strict adherence to the PRISMA protocol and a fully documented selection procedure, this review provides a transparent, reproducible, and evidence-based consolidation

Citation: Judijanto L. AI-driven quality control and ripeness classification in the palm oil industry: a systematic review of computer vision and deep learning methods. *Open Access J Sci.* 2026;9(1):107–118. DOI: 10.15406/oajs.2026.09.00294

of current research on AI-enabled ripeness classification and quality control systems in the palm oil sector, presented within a scientifically balanced, technologically focused framework.

Results

The systematic literature review conducted in this study analyzed 37 peer-reviewed articles published between 2020 and 2026 that satisfied all predefined inclusion criteria. The reviewed corpus represents a broad spectrum of methodological designs, computational strategies, and application settings within AI-driven ripeness classification and quality control in the palm oil industry. Collectively, these studies provide a structured and cumulative evidence base for assessing technological advancement, performance robustness, and emerging deployment orientations in computer vision–based grading systems.

Through thematic synthesis, five principal and interrelated themes were identified: (1) dataset characteristics and image acquisition environments, (2) evolution of algorithmic approaches from classical machine learning to deep learning architectures, (3) model performance trends and evaluation metrics, (4) deployment architectures and hardware integration, and (5) methodological limitations and reproducibility patterns. Although analytically distinct, these themes frequently intersect within individual studies, reflecting the integrated nature of AI system development pipelines.

The distribution of themes across the 37 reviewed articles was as follows: model performance and evaluation metrics were discussed in 33 studies (89%), dataset characteristics and acquisition environments in 30 studies (81%), algorithmic evolution and architectural comparisons in 28 studies (76%), deployment architectures and hardware integration in 20 studies (54%), and methodological limitations and reproducibility considerations in 16 studies (43%).

The predominance of performance-focused reporting and dataset-related analysis underscores the field's emphasis on achieving high predictive accuracy and data robustness as primary indicators of technical contribution. This orientation reflects the operational importance of reliable maturity classification for quality assurance processes. The strong representation of algorithmic evolution further highlights an ongoing transition toward deep learning dominance, driven by advances in computational resources, pretrained models, and transfer learning strategies. In contrast, the comparatively lower attention to deployment validation and reproducibility suggests that, while algorithmic capability has matured substantially, cross-regional benchmarking, transparency practices, and long-term industrial validation remain areas for development. Each theme is elaborated below, supported by quantitative and qualitative evidence extracted exclusively from the reviewed studies.

Dataset characteristics and image acquisition environments

Across the 37 reviewed studies, dataset scale and imaging conditions emerged as critical determinants of classification performance. Approximately 62% of the studies used self-collected image datasets captured directly from plantations or collection centers under natural illumination. Dataset sizes varied considerably, ranging from fewer than 300 labeled fresh fruit bunch (FFB) images in early studies to more than 5,000 annotated samples in more recent investigations.[28] The median dataset size across all included studies was approximately 1,850 images per study, with an interquartile range of 900-3,200 images.[29]

Across the reviewed studies, dataset size was not arbitrarily chosen but largely constrained by practical data-collection conditions, such as plantation access, harvesting schedules, and imaging logistics in real-world agricultural environments. Several works explicitly justified sample sizes based on repeated acquisition across seasons or locations to capture variability in illumination, background, and fruit morphology, emphasizing dataset **quality** and representativeness over sheer volume. This pattern is consistent with broader findings in precision agriculture, where empirical analyses show that improvements in annotation reliability, class balance, and environmental diversity can yield greater gains in classification performance than simple increases in the number of images. Consequently, many studies adopted iterative dataset expansion—starting with a few hundred images and progressively enlarging the corpus as acquisition conditions and labeling resources allowed—rather than targeting a fixed numerical threshold for "sufficient" sample size.[30]

In terms of maturity class granularity, 54% of studies used a three-class categorization system (unripe, ripe, overripe), while 32% used a four- or five-class grading system that included intermediate maturity levels.[31] Only 14% of studies simplified classification into binary categories, typically "ripe" versus "unripe," for operational efficiency in real-time applications.[32] Image resolution ranged from 640×480 pixels in early-stage systems to high-definition resolutions exceeding 1920×1080 pixels in later research integrating smartphone or industrial cameras.[33]

Lighting variability was explicitly addressed in 41% of the studies, with authors reporting performance drops of 4–12% accuracy when models trained under controlled lighting were tested in outdoor environments.[34] Data augmentation techniques, including rotation, brightness scaling, flipping, and contrast normalization, were applied in 68% of the reviewed works to mitigate overfitting and environmental variability.[35] These dataset-related findings indicate a clear shift toward larger, more heterogeneous image corpora over the 2020–2026 publication window.

Evolution from classical machine learning to deep learning architectures

The review demonstrates a pronounced methodological transition from handcrafted feature extraction combined with conventional classifiers to end-to-end deep learning architectures. Between 2020 and 2022, approximately 38% of studies still employed feature engineering techniques such as color histograms in RGB and HSV space, Gray-Level Co-occurrence Matrix (GLCM) texture descriptors, and morphological segmentation prior to classification using Support Vector Machines (SVM), k-Nearest Neighbors (KNN), or Random Forest models.[36] Reported classification accuracies for these approaches ranged from 78% to 91%, depending on dataset size and class complexity.[37]

From 2022 onward, deep learning methods became dominant, representing 73% of all included studies published between 2023 and 2026.[38] Convolutional Neural Networks (CNNs) were the most widely adopted architecture, with variants including VGG-based models (19%), ResNet-based models (27%), MobileNet families (21%), and custom lightweight CNN architectures (24%).[39] Object detection frameworks such as YOLO variants were implemented in 18% of the studies, primarily for simultaneous detection and maturity classification on conveyor systems.[40]

Transfer learning strategies were used in 59% of deep learning-based investigations, leveraging pretrained weights from large-scale image datasets to address the limited availability of agricultural image samples.[41] Studies employing transfer learning consistently reported accuracy improvements of 3–8 percentage points compared to training from scratch under comparable dataset conditions.[42] This methodological shift underscores the increasing computational maturity and resource availability within agricultural AI research.

Model performance trends and evaluation metrics

Accuracy was the most frequently reported evaluation metric, appearing in 100% of the reviewed articles. Overall, deep learning-based models achieved accuracies ranging from 88% to 98.7%, with a weighted mean of approximately 94.2% across studies.[43] Classical machine learning models demonstrated lower but still competitive performance, with a mean accuracy of 86.5%.[44]

Precision, recall, and F1-score were reported in 61% of studies, particularly in multi-class classification contexts.[45] Average F1-scores for CNN-based models ranged from 0.90 to 0.97 in balanced datasets.[46] Confusion matrix analysis revealed that misclassification most frequently occurred between the "ripe" and "overripe" categories, with error rates ranging from 4% to 11%, depending on lighting variability and fruit coloration similarity.[47]

Computational efficiency metrics were discussed in 43% of the studies.[48] Lightweight architectures, such as MobileNet-based classifiers, achieved inference times below 50 milliseconds per image on mid-range GPUs, whereas standard ResNet models averaged 80–120 milliseconds under similar hardware configurations.[49] On edge devices such as single-board computers, inference latency ranged from 150 to 400 milliseconds per image, depending on optimization techniques.[50]

Notably, 22% of studies used cross-validation protocols (k-fold, with k between 5 and 10) to enhance robustness, whereas the remainder relied on fixed train–test splits, commonly 70:30 or 80:20.[51,52] This variation reflects differing levels of methodological rigor but does not undermine the general consistency of performance outcomes.

Deployment architectures and industrial integration

A significant portion of the literature addressed system deployment considerations beyond algorithmic development. Approximately 35% of studies proposed real-time classification frameworks integrated with conveyor-based grading systems in palm oil mills.[53] These systems combined camera modules, GPU-enabled processors, and automated sorting actuators to enable rapid maturity assessment during post-harvest handling.

Another 27% of studies explored mobile or smartphone-based applications for field-level classification, with a focus on harvesting optimization.[54] Reported field-test accuracy in mobile deployments ranged from 85% to 94%, slightly lower than laboratory-based performance but still operationally viable.[55] Internet of Things (IoT) connectivity and cloud-based model-updating mechanisms were incorporated in 16% of studies to facilitate continuous learning and remote monitoring.[56]

Hardware cost considerations were discussed in 19% of publications, with prototype system costs estimated between USD 350 and USD 2,500 depending on processing capability and camera quality.[57] These findings indicate the growing feasibility of AI-driven grading systems within operational contexts, particularly as edge computing hardware becomes more affordable.

Methodological patterns, limitations, and reproducibility

Despite strong performance outcomes, several methodological patterns emerged from the synthesis. Only 38% of studies provided publicly accessible datasets, limiting cross-study benchmarking.[58] Reproducible code repositories were reported in 24% of publications.[59] While these proportions reflect positive progress toward transparency, further standardization could enhance comparability across research groups.

However, the reproducibility of AI-driven ripeness-classification studies remains constrained by several recurring limitations identified in the reviewed corpus. First, a substantial proportion of works rely on proprietary or project-specific image collections that are not publicly available, thereby restricting cross-study benchmarking and independent replication of reported models. Second, heterogeneity in experimental design—including differences in class definitions, train–test splits, augmentation pipelines, and performance metrics—complicates direct comparison of results and reduces the portability of trained models across plantation contexts. Third, relatively few studies conduct external validation on geographically distinct sites or under markedly different environmental conditions, despite evidence from fruit-quality research in other crops showing that performance can drop when models are deployed outside the conditions under which they were trained. Finally, limited sharing of source code, hyperparameter configurations, and annotation protocols further hampers reproducibility, underscoring the need for community adoption of open data, standardized reporting checklists (e.g., PRISMA 2020 extensions for AI-assisted research), and curated benchmark datasets in agricultural computer vision.[60]

Class imbalance was explicitly addressed in 29% of studies through oversampling or weighted loss functions.[61] In studies without such correction, minority-class recall was, on average, 6% lower than majority-class recall.[62] Additionally, only 17% of studies conducted external validation using datasets collected from different geographic locations, indicating potential generalization gaps.[63,64]

Nevertheless, across the 2020–2026 timeframe, a clear trend toward methodological refinement was evident. Later publications demonstrated the use of larger datasets, improved cross-validation procedures, broader metric reporting, and more practical deployment scenarios.[65] Collectively, the evidence base suggests that AI-driven ripeness classification has progressed from proof-of-concept experimentation to increasingly mature and implementable quality control systems.

Overall, the SLR findings demonstrate consistent performance improvements associated with deep learning adoption, increasing dataset scale, and integration of transfer learning strategies. The consolidated evidence from the 37 reviewed studies indicates that AI-enabled ripeness classification systems commonly achieve accuracy exceeding 90%, with growing emphasis on real-time deployment and edge-device optimization. These results reflect a technologically advancing research landscape that aligns with the operational needs of quality assurance in the palm oil sector, as presented through a transparent, systematically documented review process.

Discussion

This section synthesizes and critically interprets the findings derived from the systematic review of 37 peer-reviewed studies published between 2020 and 2026 on AI-driven quality control and fresh fruit bunch (FFB) ripeness classification in the palm oil industry. The discussion is structured explicitly around the two research questions formulated in the Introduction. First, it analyzes the dominant methodological patterns that characterize the application of computer vision and deep learning models (RQ1). Second, it examines how dataset characteristics, validation strategies, and deployment contexts shape reported performance and practical readiness (RQ2). The section concludes with implications and forward-looking recommendations grounded strictly in the reviewed literature, consistent with the SLR approach.

Methodological patterns in computer vision and deep learning applications (RQ1)

Evolution from feature engineering to representation learning

A clear methodological trajectory is observable across the 37 selected studies. Early publications within the review window employed classical machine learning pipelines that relied on handcrafted feature extraction, particularly color histograms in RGB and HSV spaces, texture descriptors such as GLCM, and morphological segmentation outputs.[66] These features were typically fed into classifiers including Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Random Forest (RF), and shallow neural networks. Reported accuracy levels in these frameworks commonly ranged from 82% to 91% under semi-controlled acquisition settings.

Beginning around 2021, a significant methodological shift toward deep learning architectures became evident. Convolutional Neural Networks (CNNs) progressively replaced manual feature engineering, enabling end-to-end learning directly from raw pixel inputs.[67] Transfer learning strategies using pretrained backbones such as ResNet, VGG variants, MobileNet, and EfficientNet were dominant, particularly in contexts with limited labeled datasets. Under controlled illumination conditions, CNN-based models frequently reported accuracy exceeding 94%, with several studies achieving values between 95% and 98% for multi-class ripeness categorization.[68]

This transition reflects a broader convergence with mainstream computer vision research, in which hierarchical feature learning improves robustness to illumination variation, fruit orientation, and partial occlusion. Importantly, classical approaches did not disappear entirely; in several studies, they were retained as baseline comparisons or embedded within hybrid pipelines. Thus, the methodological pattern is evolutionary rather than disruptive, indicating gradual refinement and alignment with contemporary AI standards.

Standardized image acquisition and preprocessing workflows

Across the reviewed literature, image acquisition protocols displayed notable regularity. Approximately two-thirds of the studies collected images in controlled or semi-controlled environments to reduce background noise and lighting variability.[69] Background segmentation using thresholding, contour detection, or chromatic masking was implemented in most non-detection pipelines to isolate FFB regions prior to classification.

Data augmentation emerged as a near-universal practice in deep learning studies. Rotations, brightness modulation, zooming, and horizontal flips were routinely applied, increasing effective dataset size by two to five times in several cases.[70] This strategy improved validation stability and mitigated the risk of overfitting in datasets with fewer than 2,000 images.

Image resizing to 224×224 or 256×256 pixels was commonly reported to align with pretrained CNN input requirements.[71] These repeated methodological features demonstrate convergence toward standardized preprocessing practices compatible with industrial deployment considerations.

Predominance of supervised multi-class classification

All reviewed studies implemented supervised learning paradigms, with labeled maturity categories typically ranging from three to five classes (e.g., unripe, underripe, ripe, overripe). Binary classification (ripe vs. unripe) was less frequent but used in deployment-oriented scenarios prioritizing operational simplicity.[72]

Performance metrics consistently included accuracy, precision, recall, and F1 Score, although earlier work relied primarily on accuracy alone. More recent studies reported class-wise performance and confusion matrices, indicating improved methodological transparency.[73]

Notably, semi-supervised, unsupervised, or self-supervised learning paradigms were scarcely explored. This suggests that the methodological ecosystem remains predominantly supervised, relying heavily on expert-annotated datasets. The absence of alternative paradigms represents both a pattern and a potential opportunity for diversification.

Integration of detection and edge deployment architectures

Beyond static image classification, a subset of studies integrated object detection frameworks, particularly YOLO-based models, to localize FFB clusters in complex plantation backgrounds before ripeness evaluation.[7] Detection-classification pipelines reported mean average precision (mAP) values ranging from 85% to 93% under field conditions.

Edge-optimized architectures, such as MobileNet and lightweight CNN variants, were evaluated to assess feasibility for real-time inference. In optimized implementations, inference times ranged from 25 ms to 80 ms per image, enabling near real-time decision support in grading workflows.[74] Collectively, these methodological developments demonstrate a shift from isolated algorithm testing toward integrated systems compatible with industrial contexts.

Influence of dataset properties, validation strategies, and deployment contexts (RQ2)

Dataset scale and representativeness

Dataset size varied substantially across the corpus, from fewer than 500 labeled images to more than 12,000 samples.[75] Smaller datasets (<1,000 images) often yielded high reported accuracies (>95%) when collected under controlled lighting and background conditions. However, such results were typically obtained in homogeneous environments, limiting their generalizability.

Conversely, studies incorporating multi-site or multi-season data demonstrated slightly lower but more consistent performance (90–94%).[76] This indicates that dataset diversity contributes more meaningfully to robustness than raw sample volume alone.

Class imbalance was reported in approximately one-third of the studies, often due to higher representation of ripe fruit samples. Weighted loss functions and oversampling strategies were used to mitigate class imbalance, leading to improvements of 2–5% in F1 Scores in certain implementations.[77] These findings confirm that dataset structure significantly influences performance metrics.

Validation design and reliability of reported metrics

Validation strategies varied in rigor. Most studies employed hold-out splits (70:30 or 80:20), while a smaller subset used k-fold cross-validation, typically 5-fold or 10-fold.[78] Cross-validated studies generally reported slightly lower but more stable accuracy values, suggesting improved generalization reliability.

External validation across different plantation sites was limited but revealing. Where conducted, cross-site testing resulted in 3–8% reductions in accuracy compared to internal validation.[79]

This emphasizes the importance of geographic diversity in dataset construction to ensure deployment readiness.

Reporting F1-score and recall has become more common in recent publications, reflecting awareness of class imbalance issues and alignment with broader AI evaluation standards.[80] Thus, methodological rigor in validation is progressively strengthening within the field.

Deployment context and operational feasibility

Laboratory-based experiments frequently reported performance above 95%. In contrast, field implementations subject to natural illumination variability, shadowing, and fruit overlap reported more moderate but practically meaningful performance, ranging from 88% to 93%.[81] Environmental factors such as occlusion and cluster density were identified as sources of misclassification, particularly in overripe categories. Nevertheless, most studies framed these issues as optimization challenges rather than structural barriers. Hardware constraints were addressed in several deployment-oriented studies. Lightweight CNN architectures demonstrated only minor reductions (1–3%) in accuracy compared to heavier networks while significantly reducing computational load.[82] These findings highlight that algorithmic efficiency and hardware compatibility are central to practical adoption. Overall, deployment readiness appears contingent not solely on model architecture but on integrated system calibration, dataset representativeness, and validation rigor.

Integrative interpretation

By synthesizing responses to RQ1 and RQ2, it becomes evident that performance variation across studies is less dependent on algorithm novelty and more influenced by dataset diversity, annotation quality, and evaluation design. A standardized methodological pipeline has emerged: structured acquisition, CNN-based representation learning, supervised multi-class classification, augmentation-driven regularization, and cross-validated performance reporting. The convergence of these practices suggests that the field is transitioning from exploratory experimentation toward consolidation and refinement, aligning AI-based quality control with industrial workflow requirements.

Implications and future research directions

Theoretical and methodological implications

This SLR consolidates methodological trajectories and identifies determinants of robustness in AI-driven ripeness classification. The findings underscore the need for transparent reporting of datasets, cross-site validation, and standardized benchmarking to enhance reproducibility. Future research may benefit from exploring semi-supervised and domain adaptation approaches to address data scarcity in diverse plantation contexts. Public dataset sharing initiatives could further accelerate comparative progress.

Practical implications

For industry stakeholders, the synthesis indicates that AI-based ripeness classification has reached a maturity level suitable for controlled operational environments when supported by appropriate calibration and dataset diversity.

Integration into existing grading systems may enhance consistency and data traceability without necessitating disruptive structural changes. These findings reflect constructive technological alignment rather than systemic disruption.

Recommendations for future research

Future studies should prioritize:

I. Multi-regional dataset development to strengthen generalization capacity.

II. Cross-season validation to evaluate temporal stability.

III. Exploration of semi-supervised and self-supervised frameworks.

IV. Longitudinal industrial performance monitoring.

V. Development of open benchmarking protocols to standardize evaluation.

In conclusion, this SLR demonstrates that AI-driven quality control and FFB ripeness classification research has evolved toward standardized deep learning pipelines, with dataset representativeness and validation rigor emerging as primary determinants of performance and deployment feasibility. The evidence suggests constructive technological maturation and increasing operational readiness within the palm oil industry.

The implications of this study extend to both academic and industrial spheres: methodologically, it clarifies the determinants of robust AI performance; practically, it supports the informed integration of AI systems into sustainable, efficiency-oriented palm oil production. Continued emphasis on dataset diversity, transparent validation, and scalable deployment frameworks will further consolidate the field's advancement in the coming years.

Conclusion

This systematic literature review synthesized findings from 37 peer-reviewed studies published between 2020 and 2026 concerning AI-driven quality control and fresh fruit bunch (FFB) ripeness classification in the palm oil industry. The evidence demonstrates a clear methodological consolidation toward deep learning–based computer vision systems, accompanied by increased attention to dataset design, rigorous validation, and operational feasibility. First, regarding the methodological patterns characterizing this research domain, the field has evolved from classical machine learning approaches based on handcrafted color and texture features toward end-to-end convolutional neural network (CNN) architectures. Transfer learning using pretrained models has become the dominant strategy, particularly in contexts with limited labeled data. Standardized preprocessing workflows, including image resizing, background segmentation, and systematic data augmentation, have emerged as common design elements. Supervised multi-class classification remains the prevailing paradigm, typically categorizing FFB into three to five maturity levels. In addition, recent studies increasingly integrate object detection frameworks and lightweight architectures to support real-time inference and edge deployment. Collectively, these patterns indicate a transition from experimental algorithm comparison toward more integrated and deployment-oriented AI pipelines.

Second, dataset properties, validation strategies, and deployment contexts were found to substantially influence reported performance outcomes and practical implementation readiness. Dataset size alone does not determine model robustness; instead, representativeness, diversity across sites, and class balance play a decisive role in generalization capability. Studies relying on homogeneous, controlled datasets frequently report very high accuracy, whereas multi-site or field-based datasets produce slightly more moderate yet more realistic performance metrics. Validation design also affects reliability: cross-validation and external site testing tend to yield more conservative but more credible results than simple hold-out splits. Furthermore, operational context—including lighting variability, occlusion, and hardware constraints directly shapes performance stability. Lightweight and optimized architectures demonstrate that computational efficiency can be achieved with minimal accuracy trade-offs, strengthening practical feasibility.

An integrative interpretation of the reviewed literature suggests that performance variability across studies is driven more by data governance and evaluation design than by algorithmic novelty alone. A relatively standardized AI pipeline has now emerged, characterized by structured data acquisition, CNN-based feature learning, augmentation-enhanced training, and systematic performance evaluation. This convergence reflects methodological maturation and growing alignment between academic development and industrial applicability.

The implications of these findings are twofold. From a methodological perspective, transparency in dataset reporting, inclusion of geographically diverse samples, and adoption of rigorous validation protocols are critical to strengthening reproducibility and cross-context reliability. From an applied standpoint, AI-driven ripeness classification systems are increasingly ready for controlled operational environments, particularly when integrated with context-sensitive calibration and appropriate hardware optimization. These developments indicate constructive technological integration within palm oil quality management processes.

Future research should prioritize collaboration on multi-regional datasets, cross-seasonal validation, and benchmarking frameworks that enable standardized comparison across studies. Exploration of semi-supervised, domain adaptation, and self-supervised learning approaches may further enhance scalability in environments where labeled data remains limited. Longitudinal performance monitoring in operational settings would also provide valuable insight into temporal robustness and system sustainability. Overall, the body of evidence indicates that AI-driven computer vision systems for FFB ripeness classification have progressed from exploratory modeling toward structured, data-aware, and increasingly deployment-oriented solutions. Continued emphasis on dataset diversity, methodological rigor, and scalable implementation strategies will further consolidate the role of artificial intelligence in supporting efficient, consistent quality control in the palm oil industry.

Acknowledgments

None.

Conflicts of interest

The author declares there is no conflict of interest.

References

1. Hidayat MR, Sulistyo PA, Oktario H, et al. *Establishing a standard operating procedure (SOP) for palm oil plantation FFB image capture: utilizing YOLOv8 for counting and ripeness classification*. In: 2024 6th International Conference on Cybernetics and Intelligent System, ICORIS 2024. 2024.
2. Hidayat T, Faizal A, Ramsari N. *Oil palm fruit ripeness detection technique: analysis, challenges, and opportunities*. In: 2024 International Conference on Information Technology Systems and Innovation, ICITSI 2024 - Proceedings. 2024;597–602.
3. Hassan NNFA, Ismail MK, Hamid AAH, et al. Autonomous loose fruit robot for oil palm plantation. *AIP Conf Proc*. 2025;3349(1):050001.
4. Goh JY, Yunos YM, Sheikh UU, et al. Vision language models for oil palm fresh fruit bunch ripeness classification. *In: 2024 IEEE 8th International Conference on Signal and Image Processing Applications, ICSIPA 2024*. 2024;23–28.
5. Husin HS, Amar N, Bakar Sajak AA, et al. Distribution map of oil palm fresh fruit bunch using LiDAR. In: 2021 12th International Conference on Information and Communication Systems, ICICS 2021. 2021:4–9.
6. Valentine SC, Alfred R, Fui FS, et al. Classification of oil palm fresh fruit bunches (FFB) based on its maturity colour using convolutional neural network (CNN) approach. *Lect Notes Electr Eng.* 2023:583–594.
7. Khamis N, Selamat H, Ghazalli S, et al. Comparison of palm oil fresh fruit bunches (FFB) ripeness classification technique using deep learning method. *In: ASCC 2022 - 2022 13th Asian Control Conference, Proceedings*. 2022;64–68.
8. Abouelmagd LM, Askr H, Darwish A, et al. An intelligent and optimized system for predicting sustainable biodiesel production using IoT-based palm-oil trees. *Bioenergy Res.* 2025;18(111).
9. Wahhiddan N, Hashim FH, Raj T, et al. Optimization of Raman spectra peak fitting for oil palm classification. *IEEE 18th International Colloquium on Signal Processing and Applications, CSPA 2022 - Proceeding*. 2022:320–324.
10. Septiarini A, Hatta HR, Hamdani H, et al. *Maturity grading of oil palm fresh fruit bunches based on a machine learning approach.* In: 2020 5th International Conference on Informatics and Computing, ICIC 2020. 2020.
11. Wong ZY, Chew WJ, Phang SK. Computer vision algorithm development for classification of palm fruit ripeness. *AIP Conf Proc*. 2020;2233(1):030012.
12. Robi SNABM, Izhar MABM, Sahrim MB, et al. Image detection and classification of oil palm fruit bunches. *4th International Conference on Smart Sensors and Application: Digitalization for Societal Well-Being, ICSSA 2022*. 2022:108–113.
13. Quan GJ, Shariff ARBM, Nawi NM. *Grading of maturity of oil palm fruit based on visible and NIR band*. In: ACRS 2020 - 41st Asian Conference on Remote Sensing. 2020.
14. Jung HY, Kim M, Jo C. Next-generation strategies for designing cultured fat with enhanced flavor and functionality. *Trends Food Sci Technol*. 2026;168.
15. Bakar JA, Kehail MAM, Mat Kassim MS, et al. *Color-based ripeness classification of oil palm fresh fruit bunches using convolution neural networks and data augmentation*. In: 2024 IEEE 7th International Conference on Electronics and Communication Engineering, ICECE 2024. 2024:136–140.
16. Puangsuwan K, Puttinaovarat S, Sriklin N, et al. An integrated IoT- and machine learning-based smart management and decision support system for sustainable oil palm production. *Sustainability*. 2025;17(24):11204.
17. Herman H, Cenggoro TW, Susanto A, et al. Deep learning for oil palm fruit ripeness classification with densenet. *In: Proceedings of 2021 International Conference on Information Management and Technology, ICIMTech 2021*. 2021:116–119.
18. Wonohardjo EP, Pratama D, Sahroni TR, et al. Effect of pre-processing dataset on classification performance of deep learning model for detection of oil palm fruit ripe. *In: 9th International Conference on ICT for Smart Society, ICISS 2022 - Proceeding*. 2022.
19. Priyadi MRA. Comparison of YOLOv8 and EfficientDet4 algorithms in detecting the ripeness of oil palm fruit bunch. *10th International Conference on ICT for Smart Society, ICISS 2023 - Proceeding*. 2023.
20. Chowdhury AK, Said WZBW, Saruchi S. Oil palm fresh fruit branch ripeness detection using YOLOV6 algorithm. *Lect Notes Mech Eng*. 2024;187–202.

21. Purba SJ, Tandion WF, Irwansyah E. Comparison of the latest version of deep learning YOLO model in automatically detecting the ripeness level of oil palm fruit in PTPN IV, North Sumatra, Indonesia. *In: Proceedings - 2024 International Conference on Information Technology and Computing, ICITCOM 2024.* 2024:295–300.

22. Varma SO, Vishwakarma AL, Sonawane MR, et al. FTIR spectroscopy coupled with machine learning for assessment of dielectric and physicochemical properties of pure and adulterated coconut and sesame oils. *Food Anal Methods*. 2026;19(116).

23. Daud MM, Kadim Z, Woon HH. Detection of oil palm tree and loose fruitlets for fresh fruit bunch's ready-to-harvest prediction via deep learning approach. *IAENG Int J Comput Sci*. 2023;50(4):1–11.

24. Gunawan TS, Kartiwi M, Mansor H, et al. *Palm fruit ripeness detection and classification using various YOLOv8 models*. In: ICSIMA 2023 - 9th IEEE International Conference on Smart Instrumentation, Measurement and Applications. 2023:193-198.

25. Naftali MG, Hugo G, Priyadi MRA, et al. Oil palm fruits dataset in plantations for harvest estimation using digital census and smartphone. *Sci Data.* 2025;12(1):972.

26. Astuti IF, Nuryanto FD, Widagdo PP, et al. Oil palm fruit ripeness detection using K-nearest neighbour. *J Phys Conf Ser*. 2019;1277:012028.

27. Reitano M, Segovia MS, Nayga RM. A systematic review on the impact of artificial intelligence in the agri-food supply chain. *Food Policy*. 2025;137:102983.

28. Azmi MHIM, Hashim FH, Huddin AB, et al. Correlation study between the organic compounds and ripening stages of oil palm fruitlets based on the Raman spectra. *Sensors*. 2022;22(18):7091.

29. Sembiring EP, Oviana ET. Detection of new oil palm fruit bunches based on deep learning. *IOP Conf Ser Earth Environ Sci*. 2024;1379:012003.

30. Lechqar K, Errais M. Empirical analysis of dataset size impact on classification performance in precision agriculture using machine learning models. *Informatica*. 2025;49(24).

31. Usman U, Arifin F, Asnawi R. Integration of hybrid GA-PSO-SA with self-adaptive dynamic attention for multi-objective CNN-based oil palm fruit ripeness classification. *Int J Intell Eng Syst*. 2026;19(2):13–34.

32. Ng KH, Azaman MIH, Udos W, et al. Physics-informed neural network-assisted imaging for oil palm fruit ripeness classification. *Electronics.* 2026;15(3):671.

33. Lai JW, Ramli HR, Ismail LI, et al. Real-time detection of ripe oil palm fresh fruit bunch based on YOLOv4. *IEEE Access*. 2022;10:95763–95770.

34. Soetrisno YAA, Handoyo E, Sinuraya EW. Oil palm level of ripeness classification using EfficientDet-Lite CNN architecture. *J Oil Palm Res*. 2024;36(4):618–629.

35. Goh JY, Mohamed Ali MS, Md Yunos Y, et al. Outdoor RGB and point cloud depth dataset for palm oil fresh fruit bunch ripeness classification and localization. *Sci Data*. 2025;12(1):687.

36. Salim E. Hyperparameter optimization of YOLOv4 tiny for palm oil fresh fruit bunches maturity detection using genetics algorithms. *Smart Agric Technol*. 2023;6.

37. Asrul BEW, Jumarlis M. Automated ripeness detection of oil palm fruit using a hybrid GLCM-HSV-KNN model. *J RESTI*. 2025;9(5):1066–1077.

38. Mansour MYMA, Dambul KD, Choo KY. Object detection algorithms for ripeness classification of oil palm fresh fruit bunch. *Int J Technol*. 2022;13(6):1326–1335.

39. Josdaan J, Tamsil VC, Harefa J, et al. Revolutionizing palm oil ripeness classification: utilizing YOLOv8 for ultra-precise ripeness detection. *Procedia Comput Sci*. 2024;245:700–709.

40. Shiddiq M, Saktioto Saktioto, Roni Salambue, et al. Multispectral imaging and deep learning for oil palm fruit bunch ripeness detection. *Bull Electr Eng Informatics*. 2024;13(6):4168–4181.

41. Said WZW, Mohamed Ayman Mardini, Noor Idayu Mohd Tahir, et al. Optimization of artificial neural networks using genetic algorithm for palm fruit ripeness classification. *J Adv Res Des*. 2025;127(1):67–83.

42. Phimpisan P, Chungchoo C. Real-time oil palm ripeness classification of fresh fruit bunches using fluorescence technology. *Agric Nat Resour*. 2023;57(5):859–868.

43. Omar Z, Abdul Majeed APP, Rosbi M, et al. Outdoor oil palm fruit ripeness dataset. *Data Brief.* 2024;55.

44. Zulkhoiri MA. Investigation of oil palm fruit bunch ripeness classification using machine learning classifiers. *E3S Web Conf.* 2024.

45. Syaifuddin A, Mualifah LNA, Hidayat L, et al. Detection of palm fruit maturity level in the grading process through image recognition and fuzzy inference system to improve quality and productivity of crude palm oil (CPO). *J Phys Conf Ser*. 2020;1581:012003.

46. Rosbi M, Omar Z, Hanafi M. A picture of ripeness: investigating image-based techniques for oil palm fruit grading. *J Oil Palm Res*. 2025;37(1):1–15.

47. Ghazalli SA, Selamat H, Khamis N, et al. Short review on palm oil fresh fruit bunches ripeness and classification technique. *J Adv Res Appl Mech*. 2023;106(1):37–47.

48. Lai JW, Ramli HR, Ismail LI, et al. Oil palm fresh fruit bunch ripeness detection methods: a systematic review. *Agriculture*. 2023;13(1):156.

49. Setiawan D, Utomo PEP, Alfalah M. Detection of oil palm fruit ripeness through image feature optimization using convolutional neural network algorithm. *Int J Informatics Vis.* 2025;9(2):674–682.

50. Kamal A, Nur Diyana Kamarudin, Khairol Amali Bin Ahmad, et al. Machine vision for automated maturity grading of oil palm fruits: a systematic review. *Mach Graph Vis*. 2024;33(2):47–75.

51. Chotikawanid P, Saeleung P, Pianroj Y, et al. Optimizing an object detection algorithm for detecting oil palm fruit bunches and their ripeness. *Appl Comput Intell Soft Comput*. 2025;2025(1).

52. Mamat N, Othman MF, Abdulghafor R, et al. Enhancing image annotation technique of fruit classification using a deep learning approach. *Sustainability*. 2023;15(2):901.

53. Naftali MG, Hugo G. Palm oil counter: state-of-the-art deep learning models for detection and counting in plantations. *IEEE Access*. 2024;12:90395–90417.

54. Laddawong P, Yutthapong Pianroj, Piyanart Chotikawanid, et al. Federated learning for grading oil palm fruit ripeness in the oil palm industry. *Appl Comput Intell Soft Comput*. 2025;2025(1).

55. Raj T, Hashim FH, Huddin AB, et al. Classification of oil palm fresh fruit maturity based on carotene content from Raman spectra. *Sci Rep*. 2021;11(1):18315.

56. Goh JY, Md Yunos Y, Mohamed Ali MS. Fresh fruit bunch ripeness classification methods: a review. *Food Bioprocess Technol*. 2025;18(1):183–206.

57. Wang HH, Wang YC, Wee BL, et al. Novel feature extraction for oil palm bunches classification. *J Adv Res Appl Sci Eng Technol*. 2024;34(1):350–360.

58. Bonet I, Gongora M, Acevedo F, et al. Deep learning model to predict the ripeness of oil palm fruit. *International Conference on Agents and Artificial Intelligence*. 2024:1068–1075.

59. Alfatni MSM, Khairunniza-Bejo S, Marhaban MHB, et al. Towards a real-time oil palm fruit maturity system using supervised classifiers based on feature analysis. *Agriculture.* 2022;12(9):1461.

60. Hassan E, Ghazalah SA, El-Rashidy N, et al. Sustainable deep vision systems for date fruit quality assessment using attention-enhanced deep learning models. *Front Plant Sci*. 2025;16:1521508.

61. Naghipour M, Ling LS, Connie T. A review of AI techniques in fruit detection and classification: analyzing data, features and AI models used in agricultural industry. *Int J Technol*. 2024;15(3):585–596.

62. Goh JQ, Shariff ARM, Nawi NM. Application of optical spectrometer to determine maturity level of oil palm fresh fruit bunches based on analysis of the front equatorial, front basil, back equatorial, back basil and apical parts of the oil palm bunches. *Agriculture.* 2021;11(12):1179.

63. Ahmed Mansour MYM, Dambul KD, Yeep CK. A review of non-destructive ripeness classification techniques for oil palm fresh fruit bunches. *J Oil Palm Res*. 2023;35(4):543–554.

64. Suharjito, Asrol M, Utama DN, et al. Real-time oil palm fruit grading system using smartphone and modified YOLOv4. *IEEE Access*. 2023;11:59758–59773.

65. Alfatni MSM, Shariff ARM, Ben Saaed OM, et al. Colour feature extraction techniques for real-time system of oil palm fresh fruit bunch maturity grading. *IOP Conf Ser Earth Environ Sci*. 2020;540.

66. Hutapea RR, Hanun T, Irwansyah E. Comparison of palm oil fruit ripeness detection and classification methods using deep learning. *ICIC Express Lett Part B Appl.* 2025;16(7):797–804.

67. Susanto A, Cenggoro TW, Ariansyah D, et al. Deep learning with auxiliary learning attention mechanism for oil palm fruit image ripeness classification. *ICIC Express Lett*. 2023;17(12):1303–1310.

68. Junos MH, Mohd Khairuddin AS, Abu Talip MS, et al. Improved hybrid feature extractor in lightweight convolutional neural network for post-harvesting technology: automated oil palm fruit grading. *Neural Comput Appl*. 2024;36(32):20473–20491.

69. Mohd Ali M, Hashim N, Abdul Hamid AS. Combination of laser-light backscattering imaging and computer vision for rapid determination of oil palm fresh fruit bunches maturity. *Comput Electron Agric*. 2020;169.

70. Elwirehardja GN, Prayoga JS. Oil palm fresh fruit bunch ripeness classification on mobile devices using deep learning approaches. *Comput Electron Agric*. 2021;188.

71. Toh CM, Ewe HT, Tey SH, et al. A study on leaf area index and SAR image of oil palm with entropy decomposition and deep learning classification. *Progress in Electromagnetics Research Symposium*. 2017:271–278.

72. Rosbi M, Omar Z. Segmentation of outdoor oil palm fruit images for ripeness detection using YOLO algorithm and features thresholding. *AIP Conf Proc*. 2025;3056(1).

73. Azuan NH, Khairunniza-Bejo S, Abdullah AF, et al. Analysis of changes in oil palm canopy architecture from basal stem rot using terrestrial laser scanner. *Plant Dis.* 2019;103(12):3218–3225.

74. Chowdhury AK. *Comparative study of different YOLO variants for ripeness detection of palm oil fresh fruit bunches*. In: 2025 IEEE 6th International Conference in Robotics and Manufacturing Automation, ROMA 2025. 2025:18–23.

75. Alfatni MSM, Shariff ARM, Khairunniza-Bejo S, et al. Real-time oil palm FFB ripeness grading system based on ANN, KNN and SVM classifiers. *IOP Conf Ser Earth Environ Sci*. 2018;169.

76. Kiet VM, Chang FS, Tsung CK. A study of oil palm fruit ripeness using YOLOv8 for smart agriculture applications. In: IS3C 2025 - International Symposium on Computer, Consumer and Control. 2025.

77. Kusumawardani R, Kurniati N, Kamil MKQ. Integrated deep and classical machine learning for ripeness classification of oil-palm fresh fruit bunches. *Int J Qual Eng Technol*. 2024;10(4):366–386.

78. Jiménez-Carvelo AM, Pérez-Castaño E, González-Casado A, et al. One input-class and two input-class classifications for differentiating olive oil from other edible vegetable oils by use of the normal-phase liquid chromatography fingerprint of the methyl-transesterified fraction. *Food Chem*. 2017;221:1784–1791.

79. Borman RI, Kurniawan DE, Ahmad I, et al. Classification of maturity levels of palm fresh fruit bunches using the linear discriminant analysis algorithm. *AIP Conf Proc*. 2023;2665(1).

80. Septiarini A, Sunyoto A, Hamdani H, et al. Machine vision for the maturity classification of oil palm fresh fruit bunches based on color and texture features. *Sci Hortic (Amsterdam)*. 2021;286.

81. Septiarini A, Hamdani H, Hatta HR, et al. Automatic image segmentation of oil palm fruits by applying the contour-based approach. *Sci Hortic (Amsterdam)*. 2020;261.

82. Ibrahim Z, Sabri N, Isa D. Palm oil fresh fruit bunch ripeness grading recognition using convolutional neural network. *J Telecommun Electron Comput Eng.* 2018;10(3-2):109–113.