ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

From Human Eyes to Machine Vision: A Structured Review and Implementation Framework for AI-Driven Real-Time Defect Detection and Sorting in Bakery Production

Austin Iriaremien Okosun

Department of Computer Science, Edo State University Iyamho, Edo State, Nigeria

DOI: https://doi.org/10.51583/IJLTEMAS.2026.151100021

Received: 11 November 2026; Accepted: 16 November 2026; Published: 23 November 2026

ABSTRACT

Visual quality control in many bakeries still depends on human inspectors whose judgements drift with fatigue and line speed. This review assesses the evidence on computer vision and artificial intelligence (AI) for detecting and sorting defective baked goods and translates it into a deployment framework. Eight predefined search strings were run on 16 November 2026, after scoping searches on 30 October 2026, and records were screened against explicit criteria following PRISMA 2020 reporting principles; citation searching added further studies. From 57 unique records and 6 additional reports, 19 primary studies published between 2000 and 2026 were included, covering bread (10), biscuits (6), muffins, pizza and tortillas. Reported performance was mostly between 87% and 99% under controlled imaging, yet a structured appraisal exposed weaknesses: only two studies tested systems on a production line, one reported processing speed, and two confirmed that training and test images came from different physical products. Accuracy fell sharply when conditions changed, from 96% to 79% for ungraded muffins and from 99% in training to 56.6% on whole-loaf test images in a staling study, and deep learning did not consistently outperform classical methods on small datasets. No study came from an African bakery, a gap this review identifies rather than fills. Building on these recurring weaknesses, a six-stage implementation framework is proposed, together with a two-step validation protocol based on expert review and a shadow-mode pilot. Recommendations are offered for researchers, bakery managers and vocational educators.

Keywords: Computer vision, Defect detection, Bakery quality control, Deep learning, Food automation

INTRODUCTION

For most buyers of bread, biscuits and pastries, the first test of quality is visual. Crust colour, shape and surface integrity signal freshness and care, while a scorched top or a cracked biscuit signals the opposite. Crust colour is the visible outcome of Maillard reactions and caramelisation, so small shifts in oven temperature, humidity or residence time can move a batch outside its acceptable range (Purlis, 2010). Visual defects are therefore a routine by-product of a thermal process that is hard to hold steady over long production runs, not an occasional accident.

Removing defective items has traditionally been the job of people standing beside the conveyor, a practice that scales poorly. An inspector can work reliably at around twenty parts per minute under favourable conditions (Joshi et al., 2020), whereas one documented biscuit line carries 3,600 pieces per minute across thirty lanes (MVTec Software, 2013). Judgement also varies: in an early muffin-grading study, decisions differed between inspectors by 20 to 30% (Abdullah et al., 2000). Because about 14% of food produced worldwide is lost between harvest and retail (Food and Agriculture Organization [FAO], 2019), late rejection and customer returns of baked goods carry real economic and environmental cost.

Computer vision offers non-destructive, objective inspection of every item (Brosnan & Sun, 2004; Sun, 2016), and deep learning has made it more flexible by learning features directly from images (LeCun et al., 2015; Zhou et al., 2019). Recent reviews caution, however, that strong accuracy on held-out images does not by itself

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

establish industrial readiness or sound reject decisions on a working line (Ghonimy & Barakat, 2026; Xu & Lu, 2026). This gap between laboratory accuracy and dependable sorting is the focus of the present review.

Statement of the Problem

Many bakeries, particularly small and medium enterprises (SMEs) in Africa, still rely on visual checks by staff, sometimes supported by checkweighers or metal detectors that reveal nothing about colour, shape or surface damage. Research on machine vision for baked goods spans more than two decades, but it is scattered across disciplines, uses inconsistent defect definitions and metrics, and says little about the hardware, data, integration and skills needed to move from prototype to routine production. Without a transparent, critical account of what the evidence does and does not show, bakery managers lack a sound basis for investment, and training institutions lack a basis for preparing the workforce such systems require.

Aim and Objectives

The aim of this study is to critically synthesise evidence on AI-driven computer vision for defect detection and sorting of baked goods and to translate that evidence into a testable implementation framework. The specific objectives are to:

i. identify, through a documented search, primary studies on image-based inspection of bakery products;

ii. compare their tasks, methods and reported performance, and appraise how far that performance can be trusted in production;

iii. identify the technical, data-related and contextual gaps that limit real-time deployment, including the absence of evidence from African settings; and

iv. propose a staged implementation framework, with a protocol for validating it, and the workforce competencies it requires.

The review contributes a transparent search and screening record, a side-by-side summary of reported performance across 19 bakery-specific studies, a structured appraisal of their evaluation practices, and a framework whose stages are linked explicitly to the weaknesses found.

LITERATURE REVIEW

Visual Quality and the Limits of Conventional Inspection

Quality assurance in baking has long relied on trained staff, with oven time and temperature settings as the main production guides (Khan et al., 2017). Attention declines over a shift, standards vary between individuals, and conveyor speed leaves little time for careful judgement. Sensors such as checkweighers, light curtains and laser height gauges brought objectivity to weight and dimensions, but they cannot see uneven browning, scorching, blistering, cracks or missing toppings. Conventional laboratory methods for these attributes are destructive, laborious and expensive, which has driven interest in imaging (Olakanmi et al., 2023).

From Hand-Crafted Features to Deep Learning

Classical machine vision combines controlled lighting, a camera, segmentation and features chosen by the designer, such as colour statistics, shape descriptors or texture measures, with a statistical classifier such as discriminant analysis or a support vector machine (SVM). These pipelines are interpretable, run on modest hardware and can be trained on small datasets, but features tuned to one product, camera and lighting set-up often need re-engineering when conditions change.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

Convolutional neural networks (CNNs) learn layered representations directly from labelled images (LeCun et al., 2015). Large networks trained on ImageNet (Krizhevsky et al., 2012) and residual architectures (He et al., 2016) made deep models practical, and transfer learning allows a network pre-trained on general images to be fine-tuned on a small product dataset (Pan & Yang, 2010). Augmentation (Shorten & Khoshgoftaar, 2019) and dropout (Srivastava et al., 2014) reduce overfitting. For production lines, single-stage detectors such as YOLO locate and classify many items in one pass (Redmon et al., 2016); unsupervised anomaly detection learns what a good product looks like when defects are rare (Bergmann et al., 2019); and lightweight networks run on embedded processors beside the line (Howard et al., 2017).

From Classification to Sorting

A model that labels images correctly is only one part of a sorting system. Each image must be captured at the right moment, usually triggered by a conveyor encoder; the decision must arrive within the time an item takes to travel from camera to reject point; and an air jet or pusher must act on the right item. Xu and Lu (2026) call this acquisition-to-action performance and argue that, together with independent validation and well-chosen thresholds, it decides whether an inspection system can be trusted. Thresholds matter because rejecting good products and passing defective ones carry different costs. Saliency methods such as Grad-CAM can show which image regions drove a decision (Selvaraju et al., 2017), and defect logs linked to lane and batch can point operators to drifting oven conditions.

Related Reviews and the Gap Addressed

General reviews of computer vision in food (Brosnan & Sun, 2004; Kakani et al., 2020; Zhou et al., 2019) cover bakery products only briefly. Olakanmi et al. (2023) reviewed imaging for bread quality and noted that illumination, humidity and noise limit imaging methods. Martínez-Lara et al. (2026) screened 142 records and evaluated 41 articles on computer vision for physical properties of baked products, finding only one industrial case study and identifying non-standardised imaging hardware, narrow datasets and cost for SMEs as the main barriers. Ghonimy and Barakat (2026) observed that bakery vision studies rarely document production-line trials or external validation. None of these reviews combines a structured appraisal of evaluation practice in bakery defect detection with a deployment framework for resource-constrained settings, which is the gap this paper addresses.

METHODOLOGY

Review Design

A structured narrative review was conducted and reported following the relevant items of PRISMA 2020 (Page et al., 2021): search strategy, eligibility criteria, selection counts and a flow diagram. Meta-analysis was not attempted because the included studies differ in product, task, imaging set-up and metric, so pooled estimates would be misleading.

Information Sources and Search Dates

Scoping searches on 30 October 2026 identified key studies and reviews. The main search was run on 16 November 2026 using an AI-assisted academic web search tool that retrieves records from publisher platforms and indexing services, including ScienceDirect, SpringerLink, Wiley Online Library, IEEE Xplore, MDPI, PubMed Central, the Directory of Open Access Journals and ResearchGate. Eight search strings (Table 1) were run, and the first page of ranked results for each string (nine or ten records) was exported for screening. Searches were supplemented by backward citation searching of included studies and reviews and by known-item searches to complete bibliographic details.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

Table 1: Search Strings and Records Retrieved (16 November 2026)

No.	Search string (as entered)	Records
S1	"machine vision" bakery products quality inspection journal	9
S2	computer vision bread defect detection deep learning	9
S3	biscuit cookie defect inspection image processing classification	9
S4	bread crust colour browning image analysis baking computer vision	9
S5	bread mould detection image classification neural network	9
S6	pizza tortilla cake computer vision quality classification	10
S7	YOLO object detection bakery bread biscuit production line	10
S8	bread crumb image analysis texture classification machine learning	9
	Total	**74**

Eligibility Criteria

Studies were included if they (i) were primary studies applying image-based analysis to the quality or defects of bakery products, including bread, biscuits, cookies, cakes, muffins, pizza bases and flatbreads; (ii) were published in a peer-reviewed journal or conference proceedings; (iii) appeared between January 2000 and November 2026; (iv) were written in English; and (v) reported at least one quantitative outcome. Records were excluded if they addressed non-food or non-bakery products, did not use imaging, were patents, code repositories, datasets, blogs or preprints, or did not provide enough detail to extract the task, method and outcome. Reviews were excluded from synthesis but used as background.

Selection Process

The 74 retrieved records contained 17 duplicates, leaving 57 for title and abstract screening. Of these, 34 were excluded: 11 were not about food or bakery products, 9 were patents, code repositories, blogs or dataset portals, 2 concerned non-bakery foods, 3 did not involve image-based inspection, and 9 were reviews or general methodological works. Of the 23 records assessed in full, 10 were excluded: 1 was published before 2000, 3 were preprints, and 6 lacked sufficient detail. Thirteen studies were included from the main search. Citation and known-item searching yielded 7 further reports, of which 1 could not be verified and 6 met the criteria, giving 19 included studies (Figure 1).

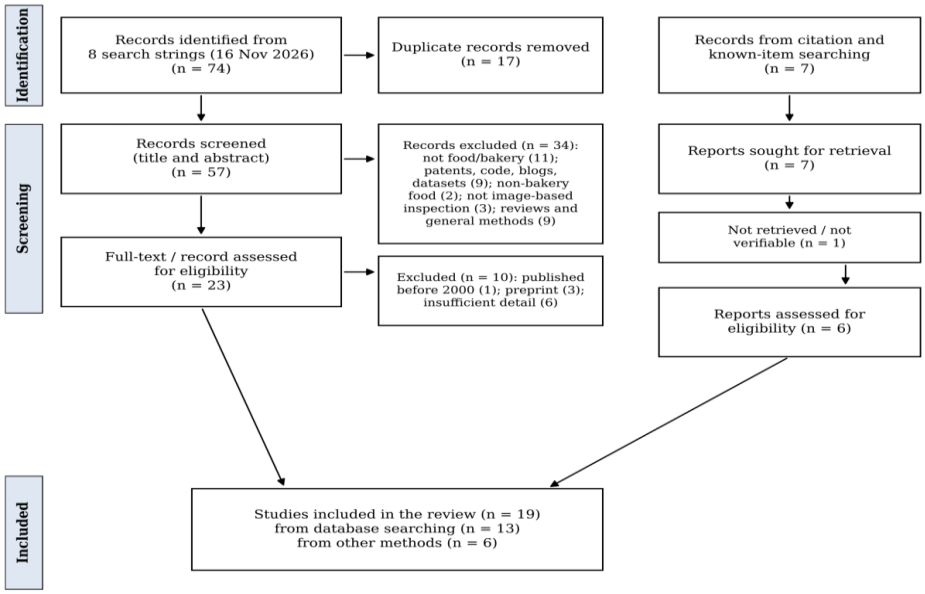


Figure 1: PRISMA 2020 Flow Diagram of Study Selection

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

Data Extraction and Appraisal

For each study, the product, task or defect, imaging and method, dataset size, evaluation design and principal reported result were extracted, using full texts where accessible and otherwise abstracts and publisher records. Because no validated risk-of-bias tool exists for food-inspection vision studies, five appraisal criteria were defined from the deployment concerns raised by Xu and Lu (2026) and Ghonimy and Barakat (2026): (C1) evaluation on a held-out test set or by cross-validation; (C2) training and test data separated by physical item or batch, preventing leakage; (C3) testing on a moving line or in production; (C4) processing speed or latency reported; and (C5) a dataset of at least 1,000 images. Each criterion was rated yes, no or unclear. Screening, extraction and appraisal were carried out by a single reviewer with AI assistance, and every included record was checked against its publisher page.

Framework Development

The weaknesses that recurred across studies were mapped to the stages of a production deployment, from image capture to rejection and monitoring, and checked against the barriers reported in related reviews. The resulting framework is presented as a proposition to be tested, and a validation protocol is specified in Section IV.F.

RESULTS

Characteristics of the Included Studies

The 19 studies examined bread (10), biscuits and cookies (6), muffins (1), pizza (1) and corn tortillas (1). Seven were published before 2015, three between 2015 and 2019, and nine from 2020 onward. Eleven relied on hand-crafted features with statistical or classical machine-learning classifiers, and eight used deep learning, all published from 2020. Tasks covered bake level and crust colour (5), physical defects such as cracks, breakage, incomplete pieces or topping distribution (5), contamination and mould (3), crumb structure, staling or sensory traits (4), counting and tracking (1) and overall quality grading (1). No study was conducted in, or used products from, an African bakery.

Reported Performance

Table 2 summarises the task, method, data and principal result of each study. Most reported accuracies between 87% and 99%, but the figures are not directly comparable because classes, metrics and imaging conditions differ widely.

Table 2: Summary of Included Studies and Reported Performance Metrics

Study	Product	Task	Imaging / method	Data	Principal reported result
Abdullah et al. (2000)	Muffins	Colour grading	Colour camera; discriminant analysis	200 samples	96% (pre-graded), 79% (ungraded); inspectors varied 20–30%
Sun & Brosnan (2003)	Pizza	Topping quality	Colour imaging; fuzzy logic	25 samples	24% error over 5 classes; 100% for 2 classes
Gonzales-Barron & Butler (2006)	Bread	Crumb structure	6 thresholding methods vs. k-means	135 slice images	Small threshold shifts caused large changes in crumb features
Nashat & Abdullah (2010)	Biscuits	Bake level (8 classes)	SVM with Wilks' λ selection	NR	87.25% and 86.75% (two SVM schemes)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

Mery et al. (2010)	Tortillas	Quality grade (5 classes)	2,300+ features; SVM	750 items; 1,500 images	About 95% (cross-validated)
Srivastava et al. (2014)	Biscuits, cookies	Quality parameters	Low-cost ARM processor and USB camera	NR	94% and 82% for two tasks
Nashat et al. (2014)	Biscuits	Cracks	Canny–Deriche filter, pyramid scheme, SVM	NR	>97% correct; sensitivity 96%, specificity 98%
Wang et al. (2015)	Sandwich biscuits	Breakage, cream overflow	Real-time feature-based system	On-line runs	Error rates of 0.33‰ and 0.32‰; <40 ms per biscuit; 6-h run
Khan et al. (2017)	Biscuits	Shape and bake level	HMAX shape + opponent colour; SVM	NR	Outperformed non-bio-inspired descriptors
Castro et al. (2017)	Bread	Crust colour acceptability	CIELab image analysis; response surface	Designed experiment	Colour linked to consumer acceptance; process optimised
Cotrim et al. (2020)	Bread	Crust browning (7 stages)	Short-CNN (Inception v3)	374 image fragments	Reliable stage separation; 91.6% less training memory than AlexNet
Yin et al. (2021)	Toast bread	Foreign contaminants	Modified U-Net; also NIR spectroscopy	NR	Vision: 95%, 93%, 92% (iron, plastic, hair)
Treepong & Theera-Ampornpunt (2023)	Bread	Early mould	22 CNNs; smartphone with 50× lens	≈1,490 images per bread type	Best F1 0.995 (white) and 0.997 (whole wheat)
Argun et al. (2023)	Bread	Staling (days old)	CNN vs. random forest	NR	CNN test: 92.3% slices, 56.6% whole loaves (training 99%)
Sisman et al. (2024)	Bread	Detection and counting	YOLOv8 + DeepSORT on Jetson Xavier NX	NR	96% detection (training phase)
Kılcı & Koklu (2025)	Finger biscuits	Defect (2 and 3 classes)	YOLOv8 n/s/m/l/x	4,990 images	96.78% (binary), 96.99% (3-class)
Shilaskar et al. (2025)	Bread	Fungus detection	Decision tree, random forest vs. U-Net	NR	U-Net 95.10% pixel-level accuracy
Serfa Juan et al. (2026)	Bread	Crumb quality	BQe-CNN vs. SVM, decision tree, random forest	3,673 images	CNN 96.0% vs. random forest 95.6% (test)
Azmat et al. (2026)	Bread	24 sensory traits	Colour-texture features; SVR	54 breads	Mean R = 0.64; R ≥ 0.75 for 7 traits

Note. NR = not reported in the accessible record; SVM = support vector machine; SVR = support vector regression; NIR = near-infrared. Results are as reported by the original authors and are not directly comparable.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

Appraisal of Evaluation Practice

Table 3 shows that evaluation practice was much weaker than headline accuracy suggests. Nine studies clearly reported a held-out or cross-validated evaluation, while six did not make their evaluation design clear in the accessible record. Only two studies confirmed that training and test images came from different physical products (Azmat et al., 2026; Treepong & Theera-Ampornpunt, 2023); in the rest, images of the same item could have appeared on both sides of the split, which inflates accuracy. Only two systems were tested on a moving line or in production (Sisman et al., 2024; Wang et al., 2015), only one reported processing speed (Wang et al., 2015), and only four used datasets of 1,000 images or more.

Table 3: Appraisal of Evaluation Practice in the Included Studies

Study	C1 Test set / CV	C2 Item-level split	C3 On-line test	C4 Speed reported	C5 ≥1,000 images
Abdullah et al. (2000)	U	U	N	N	N
Sun & Brosnan (2003)	N	U	N	N	N
Gonzales-Barron & Butler (2006)	–	–	N	N	N
Nashat & Abdullah (2010)	U	U	N	N	U
Mery et al. (2010)	Y	U	N	N	Y
Srivastava et al. (2014)	U	U	N	N	U
Nashat et al. (2014)	U	U	N	N	U
Wang et al. (2015)	Y	U	Y	Y	U
Khan et al. (2017)	U	U	N	N	U
Castro et al. (2017)	–	–	N	N	–
Cotrim et al. (2020)	Y	U	N	N	N
Yin et al. (2021)	Y	U	N	N	U
Treepong & Theera-Ampornpunt (2023)	Y	Y	N	N	Y
Argun et al. (2023)	Y	U	N	N	U
Sisman et al. (2024)	N	U	Y	N	U
Kılcı & Koklu (2025)	Y	U	N	N	Y
Shilaskar et al. (2025)	U	U	N	N	U
Serfa Juan et al. (2026)	Y	U	N	N	Y
Azmat et al. (2026)	Y	Y	N	N	N
Studies meeting criterion	**9 / 19**	**2 / 19**	**2 / 19**	**1 / 19**	**4 / 19**

Note. Y = yes; N = no; U = unclear from the accessible record; – = not applicable (measurement or process-optimisation study); CV = cross-validation.

Contradictory Findings and Methodological Weaknesses

Three patterns qualify the generally positive results. First, accuracy is strongly condition-dependent. The muffin system of Abdullah et al. (2000) fell from 96% on pre-graded samples to 79% on ungraded ones; the staling model of Argun et al. (2023) reached 99% in training but only 56.6% on whole-loaf test images, against

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

92.3% on slices; and the pizza system of Sun and Brosnan (2003) was perfect for two classes but erred 24% of the time across five. Performance therefore depends heavily on how finely defects are graded and how closely test conditions match training conditions, and Olakanmi et al. (2023) reach the same conclusion about illumination.

Second, deep learning did not consistently outperform classical methods. Serfa Juan et al. (2026) found their attention-based CNN only 0.4 percentage points more accurate than a random forest (96.0% vs. 95.6%), Azmat et al. (2026) obtained their best results with support vector regression on hand-crafted colour-texture features, and Khan et al. (2017) improved accuracy by changing descriptors rather than classifiers. By contrast, Argun et al. (2023) and Shilaskar et al. (2025) reported clear gains from deep models. Among the YOLOv8 variants compared by Kılcı and Koklu (2025), all results fell within about 1.2 percentage points, and the largest model never outperformed the medium-sized one. On the small, controlled datasets typical of this field, image acquisition and data quality appear to matter more than model architecture.

Third, several reports contain gaps or inconsistencies that limit how far their figures can be relied on. Serfa Juan et al. (2026) give accuracies of 92% and 88% in the abstract but 96% in the results; the confusion counts in Kılcı and Koklu (2025) imply an accuracy slightly different from the one reported; and Sisman et al. (2024) report detection accuracy only for the training phase. Studies that address these concerns stand out: Treepong and Theera-Ampornpunt (2023) split data by bread slice, repeated training five times and reported precision, recall and F1, while Wang et al. (2015) ran their system on-line for six hours and reported error rates and processing time.

Gaps in the Evidence

Five gaps emerge. (1) Datasets are small and private: only four studies used 1,000 or more images, and none reported an openly shared bakery defect benchmark. (2) Robustness to lighting, product variation and batch effects is rarely tested. (3) Line-level evidence is scarce: two of 19 studies tested systems in production, and none reported false-reject rates, escape rates or payback. (4) Internal defects such as under-baked centres are almost absent, although crumb-based studies (Gonzales-Barron & Butler, 2006; Serfa Juan et al., 2026) show that sliced-product imaging is feasible. (5) No included study came from Africa. These findings agree with the scarcity of industrial case studies reported by Martínez-Lara et al. (2026).

Proposed Implementation Framework and Validation Protocol

The six-stage framework in Table 4 maps each recurring weakness to a deployment stage, so that each stage produces evidence that can be checked before further investment.

Table 4: Six-Stage Framework for Vision-Based Defect Detection and Sorting

Stage	Key activities	Weakness addressed (evidence)	Outputs and indicators
1. Problem definition and acquisition design	Agree a defect catalogue with quality staff; enclosed, diffuse lighting; camera after oven and before packaging; encoder triggering	Illumination sensitivity and inconsistent defect definitions (Abdullah et al., 2000; Olakanmi et al., 2023)	Written defect definitions with reference images; repeatable image capture
2. Data collection and annotation	Capture across shifts, batches and seasons; label against definitions; double-label a subset; split by item and production day	Small datasets and possible leakage (C2, C5 in Table 3)	Versioned dataset; inter-annotator agreement
3. Model development	Benchmark a classical baseline against a fine-tuned CNN or detector; anomaly detection where defects are rare	Deep learning not consistently superior (Section IV.D)	Per-class precision, recall and F1 on unseen production days

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

4. Edge deployment	Lightweight model on an industrial PC or embedded GPU; measure end-to-end latency at line speed	Speed rarely reported (C4)	Latency within camera-to-reject travel time
5. Rejection control and integration	Link to PLC and actuator; set thresholds from cost of false rejects vs. escapes; shadow-mode trial first	Few on-line tests (C3); grading granularity	False-reject and escape rates; throughput unchanged
6. Monitoring and improvement	Defect dashboards by lane, shift and batch; drift checks; periodic relabelling and retraining	Accuracy loss when conditions change (Argun et al., 2023)	Trend reports used in oven adjustment; model version log

Note. PLC = programmable logic controller; GPU = graphics processing unit.

Because the framework is derived from a modest and heterogeneous evidence base, it should be validated before it is treated as guidance. A two-step protocol is proposed. In the first step, a panel of 10 to 15 experts, including bakery quality managers, automation engineers, food scientists and vocational instructors, rates the relevance and feasibility of each stage and activity on a five-point scale over two modified Delphi rounds, with consensus defined in advance as at least 70% of ratings at 4 or 5 (Hsu & Sandford, 2007). In the second step, the revised framework is piloted in one or more SME bakeries. The vision system runs in shadow mode for four to six weeks alongside existing inspection, and both are compared with an independent audit of sampled products. Pre-specified outcomes are false-reject rate, escape rate, end-to-end latency, effect on throughput, and installation and running costs relative to the value of reduced waste and returns.

DISCUSSION

What the Evidence Supports

Nineteen studies are a small base for deployment guidance, and the heterogeneity of tasks and metrics prevents any pooled estimate of performance. The framework does not depend on such an estimate, however. It rests on weaknesses that recur across most of the included studies and that independent reviews have reported for broader sets of bakery and food-inspection research (Ghonimy & Barakat, 2026; Martínez-Lara et al., 2026; Xu & Lu, 2026). The evidence therefore supports two claims: that vision systems can classify common bakery defects accurately under controlled conditions, and that the steps needed to make that accuracy dependable in production have rarely been tested. It does not yet show what accuracy, speed or return on investment a bakery should expect in routine operation, which is why the framework includes its own validation protocol.

The African Context: An Identified Gap

None of the included studies was conducted in Africa, so this review cannot report how vision systems perform in African bakeries; it can only identify the absence of such evidence. Constraints commonly faced by African SMEs, including capital cost, unreliable electricity and scarce technical expertise, are therefore offered as reasons to test specific design choices rather than as findings. The included studies suggest which choices are worth testing: low-cost processors and cameras (Srivastava et al., 2014), inexpensive smartphone optics (Treepong & Theera-Ampornpunt, 2023) and embedded detection hardware (Sisman et al., 2024). A plausible starting configuration is a single camera station aimed at the costliest defect, running offline on an edge device with an uninterruptible power supply. Where automatic rejection is not affordable, an alert that tells a worker which item to remove would still provide consistency and data. Building openly licensed image collections of locally produced breads and snacks and running pilots of the kind described in Section IV.F are the most direct ways to close this gap.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

Workforce Development

A vision system shifts work rather than removing it. Someone must define defects, label images, maintain cameras and lighting, interpret dashboards and decide when a model needs retraining. These tasks require food quality knowledge, basic data and machine-learning literacy and familiarity with industrial controls. Technical and vocational education and training (TVET) institutions, polytechnics and university programmes in food science, mechatronics and computing could build these competencies through short modules, practical projects and placements with local bakeries, and university–industry partnerships could also produce the local datasets the field lacks.

LIMITATIONS

Several limitations apply. Only the first page of ranked results was screened for each search string, so relevant studies may have been missed, and the search did not use subscription databases such as Scopus or Web of Science. Screening, extraction and appraisal were performed by one reviewer with AI assistance rather than by two independent reviewers. Some data were extracted from abstracts and publisher records because full texts were not accessible, which may explain several unclear ratings in Table 3. Only English-language studies were included, the appraisal criteria have not been formally validated, and the framework remains untested until the proposed validation is carried out.

CONCLUSION AND RECOMMENDATIONS

Across 19 studies spanning more than two decades, computer vision has repeatedly classified bakery defects with high accuracy under controlled conditions, but the evidence that this accuracy holds on working production lines is thin. Few studies separated training and test data by product, tested systems on-line or reported speed, and none came from Africa. Deep learning broadened what can be detected but did not consistently beat well-designed classical methods on small datasets. The proposed six-stage framework and its validation protocol are intended to close the gap between laboratory accuracy and dependable sorting. It is recommended that (a) researchers split data by product and production day, report line-level metrics and share datasets; (b) bakery managers start with a narrowly scoped, shadow-mode pilot before automating rejection; (c) African universities and bakeries jointly build local datasets and pilot evidence; and (d) training institutions integrate machine vision, data annotation and basic AI literacy into food technology and engineering programmes.

DECLARATIONS

Use of artificial intelligence. Generative AI tools were used to run the documented literature searches, support screening and data extraction and edit the language. The author checked every included record against its publisher page, verified the extracted data, reviewed and revised the manuscript, and takes full responsibility for its content, interpretation and conclusions. No AI tool is listed as an author.

Conflict of interest. The author declares no conflict of interest.

Funding. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Ethical approval. Not applicable. This study reviews published literature and did not involve human participants or animals.

Data availability. No new data were generated. The search strings, record counts and extracted study data are reported in Tables 1 to 3 and Figure 1, and all included studies are listed in the References.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

REFERENCES

1. Abdullah, M. Z., Aziz, S. A., & Dos Mohamed, A. M. (2000). Quality inspection of bakery products using a color-based machine vision system. *Journal of Food Quality, 23*(1), 39–50. https://doi.org/10.1111/j.1745-4557.2000.tb00194.x
2. Argun, M. Ş., Türk, F., & Kurt, A. (2023). Non-destructive prediction of bread staling using artificial intelligence methods. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi, 12*(4), 985–993. https://doi.org/10.17798/bitlisfen.1308493
3. Azmat, M., García-Gómez, M. B., Cernadas, E., Fernández-Delgado, M., Quiroga-García, M., & Romero-Rodríguez, M. A. (2026). Assessing traits in bread using image analysis and machine learning. *Food Research International, 235*, Article 119131. https://doi.org/10.1016/j.foodres.2026.119131
4. Bergmann, P., Fauser, M., Sattlegger, D., & Steger, C. (2019). MVTec AD: A comprehensive real-world dataset for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 9592–9600). IEEE.
5. Brosnan, T., & Sun, D.-W. (2004). Improving quality inspection of food products by computer vision: A review. *Journal of Food Engineering, 61*(1), 3–16. https://doi.org/10.1016/S0260-8774(03)00183-3
6. Castro, W., Oblitas, J., Chuquizuta, T., & Avila-George, H. (2017). Application of image analysis to optimization of the bread-making process based on the acceptability of the crust color. *Journal of Cereal Science, 74*, 194–199. https://doi.org/10.1016/j.jcs.2017.02.002
7. Cotrim, W. da S., Minim, V. P. R., Felix, L. B., & Minim, L. A. (2020). Short convolutional neural networks applied to the recognition of the browning stages of bread crust. *Journal of Food Engineering, 277*, Article 109916. https://doi.org/10.1016/j.jfoodeng.2020.109916
8. Food and Agriculture Organization of the United Nations. (2019). *The state of food and agriculture 2019: Moving forward on food loss and waste reduction*. FAO.
9. Ghonimy, M., & Barakat, H. (2026). Artificial intelligence and computer vision for automated quality inspection of processed foods: A comprehensive review. *Foods, 15*(22), Article 4064. https://doi.org/10.3390/foods15224064
10. Gonzales-Barron, U., & Butler, F. (2006). A comparison of seven thresholding techniques with the k-means clustering algorithm for measurement of bread-crumb features by digital image analysis. *Journal of Food Engineering, 74*(2), 268–278.
11. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). IEEE.
12. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). *MobileNets: Efficient convolutional neural networks for mobile vision applications* (arXiv:1704.04861). arXiv.
13. Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research, and Evaluation, 12*(10), 1–8.
14. Joshi, K. D., Chauhan, V., & Surgenor, B. (2020). A flexible machine vision system for small part inspection based on a hybrid SVM/ANN approach. *Journal of Intelligent Manufacturing, 31*, 103–125. https://doi.org/10.1007/s10845-018-1438-3
15. Kakani, V., Nguyen, V. H., Kumar, B. P., Kim, H., & Pasupuleti, V. R. (2020). A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research, 2*, Article 100033. https://doi.org/10.1016/j.jafr.2020.100033
16. Khan, B., Han, F., Wang, Z., & Iqbal, A. (2017). Automatic quality inspection of bakery products based on shape and color information. *Journal of Harbin Institute of Technology, 24*(5), 88–96. https://doi.org/10.11916/j.issn.1005-9113.16179
17. Kılcı, O., & Koklu, M. (2025). Automated classification of biscuit quality using YOLOv8 models in food industry. *Food Analytical Methods, 18*, 815–829. https://doi.org/10.1007/s12161-025-02755-5
18. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems, 25*, 1097–1105.
19. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

20. Martínez-Lara, N.-D., Filomena-Ambrosio, A., & Garzón-Castro, C. L. (2026). Use of computer vision systems in baked products: Potential tool for measuring physical properties. *Journal of Food Measurement and Characterization, 20*(5), 6925–6947. https://doi.org/10.1007/s11694-026-04154-8
21. Mery, D., Chanona-Pérez, J. J., Soto, A., Aguilera, J. M., Cipriano, A., Vélez-Rivera, N., Arzate-Vázquez, I., & Gutiérrez-López, G. F. (2010). Quality classification of corn tortillas using computer vision. *Journal of Food Engineering, 101*(4), 357–364. https://doi.org/10.1016/j.jfoodeng.2010.07.018
22. MVTec Software. (2013, September 3). *Machine vision inspects biscuits – Machine Vision Technology Ltd.* https://www.mvtec.com/knowledge-base/news/article/machine-vision-inspects-biscuits-machine-vision-technology-ltd
23. Nashat, S., & Abdullah, M. Z. (2010). Multi-class colour inspection of baked foods featuring support vector machine and Wilk's λ analysis. *Journal of Food Engineering, 101*(4), 370–380.
24. Nashat, S., Abdullah, A., & Abdullah, M. Z. (2014). Machine vision for crack inspection of biscuits featuring pyramid detection scheme. *Journal of Food Engineering, 120*, 233–247.
25. Olakanmi, S. J., Jayas, D. S., & Paliwal, J. (2023). Applications of imaging systems for the assessment of quality characteristics of bread and other baked goods: A review. *Comprehensive Reviews in Food Science and Food Safety, 22*, 1817–1838. https://doi.org/10.1111/1541-4337.13131
26. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71
27. Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering, 22*(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191
28. Purlis, E. (2010). Browning development in bakery products: A review. *Journal of Food Engineering, 99*(3), 239–249. https://doi.org/10.1016/j.jfoodeng.2010.03.008
29. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779–788). IEEE.
30. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 618–626). IEEE.
31. Serfa Juan, R. O., Kaufman, R. C., Bean, S. R., Chen, R. Y., Sutton, T. L., Tilley, M., Abacan, S. F., & Gerken, A. R. (2026). AI-based bread quality assessment using image processing techniques and the developed BQe-CNN. *International Journal of Food Science and Technology, 61*(1), vvag010. https://doi.org/10.1093/ijfood/vvag010
32. Shilaskar, S., Bhatlawande, S., Bhalerao, S., Chakole, S., & Bagmare, S. (2025). Fungus detection on baked breads: A comparative approach of machine learning and deep learning. In *Proceedings of the 1st International Conference on Lifespan Innovation (ICLI 2025)* (pp. 175–183). Atlantis Press. https://doi.org/10.2991/978-94-6463-831-8_22
33. Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data, 6*, Article 60. https://doi.org/10.1186/s40537-019-0197-0
34. Sisman, H. I., Guney, E., & Bayilmis, C. (2024). A real-time tracking system for bread production based on YOLOv8 and DeepSORT. In F. P. Garcia, A. Jamil, A. A. Hameed, A. Ortis, & I. S. Ramirez (Eds.), *Recent trends and advances in artificial intelligence* (Lecture Notes in Networks and Systems, Vol. 1138, pp. 469–479). Springer. https://doi.org/10.1007/978-3-031-70924-1_35
35. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research, 15*, 1929–1958.
36. Srivastava, S., Boyat, S., & Sadistap, S. (2014). A robust machine vision algorithm development for quality parameters extraction of circular biscuits and cookies digital images. *Journal of Food Processing, 2014*, Article 376360. https://doi.org/10.1155/2014/376360
37. Sun, D.-W. (Ed.). (2016). *Computer vision technology for food quality evaluation* (2nd ed.). Academic Press.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue XI November 2026

38. Sun, D.-W., & Brosnan, T. (2003). Pizza quality evaluation using computer vision—Part 2: Pizza topping analysis. *Journal of Food Engineering, 57*(1), 91–95. https://doi.org/10.1016/S0260-8774(02)00276-5
39. Treepong, P., & Theera-Ampornpunt, N. (2023). Early bread mold detection through microscopic images using convolutional neural network. *Current Research in Food Science, 7*, Article 100574. https://doi.org/10.1016/j.crfs.2023.100574
40. Wang, Y., Shi, C., Zhang, C., & Liao, Q. (2015). A real-time computer vision system for biscuit defect inspection. In *Proceedings of the 10th International Conference on Computer Vision Theory and Applications (VISAPP 2015)* (Vol. 1, pp. 531–536). SciTePress. https://doi.org/10.5220/0005306705310536
41. Xu, N., & Lu, Y. (2026). Artificial intelligence-enabled computer vision for food inspection: From validated measurements to process control. *Foods, 15*(20), Article 3591. https://doi.org/10.3390/foods15203591
42. Yin, J., Hameed, S., Xie, L., et al. (2021). Non-destructive detection of foreign contaminants in toast bread with near infrared spectroscopy and computer vision techniques. *Journal of Food Measurement and Characterization, 15*, 189–198. https://doi.org/10.1007/s11694-020-00627-6
43. Zhou, L., Zhang, C., Liu, F., Qiu, Z., & He, Y. (2019). Application of deep learning in food: A review. *Comprehensive Reviews in Food Science and Food Safety, 18*(6), 1793–1811. https://doi.org/10.1111/1541-4337.12492