Automated quality control by AI-driven visual inspection techniques in the food processing industry

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 00 (2026) 000–000

www.elsevier.com/locate/procedia

20th CIRP Conference on Intelligent Computation in Manufacturing Engineering

Automated quality control by AI-driven visual inspection techniques in the food processing industry

Alberto Piovano[a], Elisa Verna[a], Gianfranco Genta[a]*, Maurizio Galetto[a]

[a]*Politecnico di Torino, Department of Management and Production Engineering, Corso Duca degli Abruzzi 24, 10129 Torino, Italy*

* Corresponding author. Tel.: +39-011-090-7257. *E-mail address*: gianfranco.genta@polito.it

Abstract

AI-driven visual inspection techniques in industrial food processing are critical for quality assurance and safety; therefore, the selection of the most suitable algorithm should be based on both operational and performance criteria. This study tests the convolutional neural network YOLOv8 on an industrial bakery dataset under typical shop-floor stress conditions. The algorithm performance is evaluated across high-volume and resource-constrained scenarios, by means of decision-oriented performance metrics: Risk-Weighted Total Cost of Quality, Throughput Efficiency, and Waste Reduction. Results show the robustness of YOLOv8 in detecting relevant defects, ease of use and economic sustainability, thus enabling its implementation and operational effort.

Peer-review under responsibility of the scientific committee of the 19th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 8-10 July, Gulf of Naples, Italy

Keywords: Quality control, Visual inspection, Artificial Intelligence

1. Introduction

The rhetoric of Industry 4.0 depicts food factories as connected, data-driven and increasingly autonomous. Yet, on many high-speed lines, quality control remains a persistent bottleneck: process data are digitised, while product conformity at the outlet is still “measured” through delayed and often manual checks. The enabling technologies are not missing, Artificial Intelligence (AI)/Machine Learning (ML), Internet of Things (IoT), smart sensing, cyber–physical systems and digital integration have long been identified as central to the Industry 4.0 agenda, including explicit links to food quality and safety [1]. The central question is how to establish a framework for making continuous and reliable quality decisions in real-time, where human inspection cannot guarantee consistency at industrial speeds and critical variables are frequently inferred only after processing, with slow or discontinuous procedures [2].

The bakery domain makes this tension particularly sharp because “appearance” is not a decorative attribute but a thermal footprint of the process. Browning results from non-enzymatic reactions and its relevance goes beyond sensory preference: it is also associated with nutritional implications and the formation of undesirable compounds such as acrylamide, turning what looks like an aesthetic judgement into a potential safety-related concern [3]. Moreover, browning is inseparable from transport phenomena: it depends on temperature and water activity within coupled heat and mass transfer dynamics that generate crust formation, moisture loss and structural changes [3].

The scientific literature pertaining to the quality assessment of baked goods has long acknowledged the limitations of conventional methods, which are commonly described as destructive, labour-intensive, time-consuming, error-prone and costly [4]. This has motivated a transition toward non-destructive imaging and computer-vision approaches.

Peer-review under responsibility of the scientific committee of the 20th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 8-10 July, Gulf of Naples, Italy

However, the step from laboratory prototypes to factory deployment exposes "ordinary" constraints that become decisive at scale. Standardising image acquisition is difficult; and in oven-adjacent environments, cameras are vulnerable to window fouling, condensation and drift, undermining repeatability and measurement stability [4], [5].

Prior to the widespread adoption of end-to-end deep learning pipelines, industrial visual inspection methods predominantly relied on engineered features and manually tuned decision rules. While these methods were often effective, they were susceptible to lack of robustness and costly maintenance. Chetima and Payeur in [6] highlight that manual inspection systems are product-specific, expert-dependent, and require trial-and-error tuning. Automating this process through feature selection and machine learning achieved comparable accuracy, demonstrating that tacit expertise can be technically replaced. From a production-technology perspective, the lesson is clear: accuracy is necessary, but industrialisation also depends on configurability, stability, and maintainability under inevitable process variability.

The shift to deep learning, particularly YOLO-type single-stage detectors, has reshaped real-time inspection by combining competitive detection with low latency suitable for in-line deployment. In a Quality 4.0 framing, such systems can support vertical and horizontal integration of quality information across shop-floor and enterprise layers, yet they also impose new requirements on datasets, competencies and infrastructure. A particularly informative "bridge" between classical vision and Quality 4.0 logic is offered by Görgülü [7], who demonstrates a hybrid, in-line monitoring system at the exit of a tunnel oven, combining cameras and sensors to read geometry, surface colour and moisture at high speed, explicitly criticising widely spaced sampling as non-representative and too slow for corrective action. Importantly, the contribution is not an algorithmic novelty claim but a production argument.

Applied evidence indicates that YOLOv8 is technically mature for defect-related tasks in food manufacturing. Kılcı and Koklu in [8], for example, reported high performance using YOLOv8 for biscuit quality classification, but rely on controlled conditions that exclude factory-realistic disturbances.

This observation directly addresses the research gap identified in the present contribution. Much of the current literature evaluates models through technical metrics alone, whereas manufacturing decisions depend on throughput, cost asymmetries, and waste, where false positives mean immediate scrap and false negatives risk safety and brand damage. Accordingly, this paper selects YOLOv8 not for novelty but for production suitability (low latency, edge feasibility, anchor-free robustness) and proposes a production-oriented evaluation framework that combines technical performance with manufacturing-relevant decision criteria. The original contribution lies in translating model assessment into an operational perspective, showing how AI-based visual inspection can be judged in terms of its impact on industrial decision-making rather than benchmark accuracy alone.

2. Methodology

2.1. Dataset and experimental setup

The study is grounded in an industrial bakery inspection setting, where products flow on a high-speed conveyor and visual checks must simultaneously support quality consistency and food-safety assurance. The dataset under consideration consists of images of crackers annotated into three mutually exclusive classes that reflect typical shop-floor decision needs (representative examples of the three classes are shown in Fig. 1.):

- Normal: compliant product.
- Cracked: surface/structural defect mainly affecting aesthetic/functional acceptability.
- Over-baked: excessive browning and texture alteration, treated as a proxy for critical thermal deviation. In bakery technology, this class is not merely an aesthetic issue: abnormal browning is associated with undesirable sensory profiles and can correlate with risk-relevant process deviations (e.g., acrylamide mitigation logic) [3].

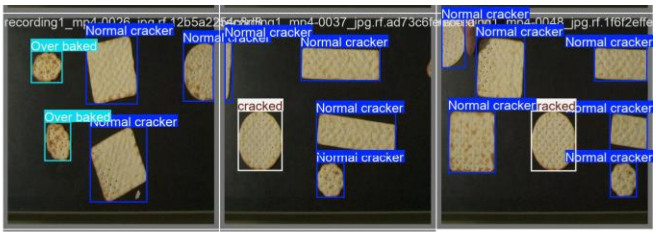

Fig. 1. Representative examples of the three product classes considered in this study: Normal, Cracked, and Over-baked

To ensure repeatability, all images were auto-oriented and resized to a uniform input scale of 640×640 px, matching the standard Ultralytics pipeline [9]. The experimental design uses two datasets: a Baseline set and a Stress set (shop-floor disturbance simulation). Dataset characteristics are summarised in Table 1, with fixed validation and test partitions to support a balanced robustness comparison.

Table 1. Dataset composition and augmentation scheme for Baseline vs. Stress conditions.

Data Property	Baseline Set	Stress Set	Rationale
Total images	658	1045	Higher variability for robustness testing.
Split (train/val/test)	537/81/40	924/81/40	Comparable validation/test sets for balanced evaluation.
Pre-processing	Auto-orient; resize 640×640 px	Auto-orient; resize 640×640 px	Uniform image scale.
Data augmentation	Horizontal flip	Horizontal flip; shear ±10° H/V; blur 4.3 px	Replicates realistic distortions and blur.

To evaluate robustness against engineering-plausible disturbances in SME production environments, the operational stress simulation applies two controlled perturbations, shear transformation (±10° horizontal/vertical) emulating belt vibrations and camera misalignment, and Gaussian blur (≈4.3 px) mimicking motion blur and lens fouling [10]. The Stress dataset was generated by augmenting primarily the training split while keeping validation/test sets comparable, deliberately isolating robustness evidence from augmentation-driven performance gains and testing whether YOLOv8 maintains stable behaviour under the imaging degradations repeatedly identified in food-vision literature as critical to industrial deployment.

2.2. From Technical KPIs to Managerial dimensions

The evaluation framework is deliberately hierarchical in nature. The behaviour of the model is validated by technical indicators and then translated into operational and economic implications by managerial dimensions.

2.2.1. Technical performance metrics

Model quality is assessed using standard detection metrics [11]:

- Precision (P) and Recall (R),
- Mean average precision at IoU threshold 0.50 (mAP@50) and averaged over IoU thresholds from 0.50 to 0.95 (mAP@50–95),
- Total latency (expressed in ms) and derived frames per second (FPS).

In addition, diagnostic analysis uses column-normalized confusion matrices to characterise class-dependent failure modes, as in food inspection the type of error (e.g., missed safety defect vs. false reject) dominates managerial relevance.

2.2.2. Managerial dimensions

Three dimensions are adopted as the primary decision layer.

(i) Risk-weighted Total Cost of Quality (TCQ).
By adapting traditional Cost of Quality models [12], a Risk-Weighted Total Cost of Quality (TCQ) is proposed. TCQ translates false decisions and speed constraints into costs, distinguishing between:

- Internal scrap cost driven by false positives (*FP*): compliant products rejected as defective.
- Aesthetic risk cost driven by false negatives on cracked products, i.e. FN_Q.
- Safety risk cost driven by false negatives on over-baked products (safety-weighted proxy), i.e. FN_S.
- Downtime/latency cost when inspection speed is slower than the target line speed.

A compact formulation is:

$$TCQ = FP \cdot c_{scrap} + FN_Q \cdot c_Q + FN_S \cdot c_s + T_{loss} \cdot c_D \quad (1)$$

where c_{scrap}, c_Q, c_S represent unit costs (expressed in €/unit) of production cost (materials, energy) lost due to scrap, non-conformance for aesthetic defects and non-conformance for safety defects, respectively. Furthermore, c_D is the downtime opportunity cost (expressed in €/h) and T_{loss} is the production time lost if the model's inference speed is slower than the line target. Notably, the "Cost of Latency" term explicitly monetises production losses arising when algorithm inference exceeds the line target, a variable typically overlooked in computer vision literature, yet critical for real-world deployment decisions [13].

(ii) Throughput Efficiency (TE)
Adapted from the Overall Equipment Effectiveness (OEE) framework [14], TE measures the system's ability to maintain the target line speed without buffering:

$$TE = \min\left(1, \frac{FPS_{model}}{FPS_{target}}\right) \quad (2)$$

Values TE < 1 indicate that the inspection system is structurally slower than the line.

(iii) Waste Reduction Index (WRI)
WRI links false rejects to circular-economy objectives by quantifying avoidable scrap relative to a baseline:

$$WRI = \frac{FP_{Baseline} - FP_{Stress}}{FP_{Baseline}} \quad (3)$$

WRI captures the waste penalty (or benefit) of operating under disturbances.

2.3. Scenario-based economic parametrization

To avoid treating “value” as universal, TCQ and TE are computed under three stylised scenarios representing distinct competitive logics reported in Table 2.

Table 2. Scenario parameterization across three competitive logics.

Parameter	Meas. Unit	Scenario A: Commodity	Scenario B: Brand Protection	Scenario C: SME
Strategic Priority	-	Speed	Zero-Defect	Cost and Waste Reduction
Batch Size	Units	100000	50000	20000
Target Line Speed	FPS = items/s	50	30	40
c_{scrap}	€/unit	0.01	0.05	0.02
c_Q	€/unit	0.05	0.50	0.15
c_S	€/unit	1.00	5.00	2.00
c_D	€/h	500	200	100
Operational Rationale	-	High-saturation lines where latency causes direct revenue loss via T_{loss}	Premium segments where safety failures FN_S trigger severe penalties.	Smaller batches, hardware limits and waste disposal is a critical KPI.

Scenario parameters (batch size, unit scrap cost, aesthetic/safety penalty levels, downtime opportunity cost) are set to reflect order-of-magnitude differences typical of Cost-of-Quality reasoning (Prevention–Appraisal–Failure (PAF) logic and the common "1–10–100" heuristic), with safety defects weighted substantially higher than aesthetic defects and internal scrap [14], [15].

3. Results

3.1. Technical detection performance and real-time capability

The experimental campaign was designed to answer a practical question: is YOLOv8 "good enough" not only in accuracy, but also in operability? In this sense, the results should be read less as an emphasis on high mAP and more as an assessment of industrial readiness under a realistic disturbance envelope. Table 3 summarises the global detection metrics.

Table 3. YOLOv8 detection performance and real-time capability under Baseline vs. Stress imaging conditions.

Condition	Precision (P)	Recall (R)	mAP@50	mAP@50-95	Total Latency [ms]
Baseline	0.881	0.942	0.974	0.953	24.2
Stress	0.928	0.935	0.979	0.962	23.8

Two points deserve emphasis. First, mAP@50 remains greater than 0.97 in both conditions. This places the model in a regime of algorithmic maturity: the detector has essentially learned the relevant visual manifold for this task, and further gains are likely to be incremental and dataset-specific rather than structurally transformative. Second, stress does not cause a meaningful collapse. The observed variations (e.g., recall from 0.942 to 0.935) are minor and compatible with a robust deployment, at least from a purely statistical viewpoint.

The central issue, however, is not accuracy, it is time. YOLOv8 processes one frame in 24.2 ms (Baseline) and 23.8 ms (Stress), corresponding to approximately 42.0 FPS. Indeed, it falls below the 50 Hz takt-time target in Scenario A, demonstrating that technical excellence does not automatically ensure operational compatibility.

3.2. Diagnostic analysis by class

At the aggregate level, stress does not induce performance collapse. However, robustness is not uniform across defect types, an aspect that matters in food manufacturing because different defects carry asymmetric risk. Fig. 2 (column-normalized confusion matrices) exposes the structure of YOLOv8's errors.

YOLOv8 maintains high stability on compliant products (Normal class), with 0.98 correct under Baseline conditions and 0.96 under Stress conditions, with shop-floor noise introducing only minor diffusion toward defect classes rather than an explosion of false alarms.

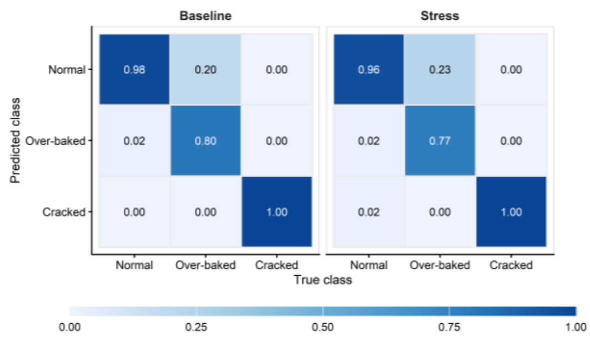

Fig. 2. True-class-normalized confusion matrices for YOLOv8 defect classification

Cracked detection is near-perfect (1.00 in both conditions), likely due to salient geometric cues robust to perturbation. The critical behaviour concerns Over-baked, which is the most consequential class from a food-safety and compliance perspective. In Fig. 1, the diagonal element is 0.80 in Baseline and drops to 0.77 in Stress. A substantial fraction of Over-baked products is absorbed into the Normal class (0.20 in Baseline; 0.23 in Stress). This is not a generic "accuracy reduction"; it is a specific operational risk, missed thermal deviation. So, the FP/FN "attitude" of the model is worth stating clearly. The matrices suggest a cautious but not indiscriminate profile, false rejections of Normal-class products are low (2–4%), yet the model still exhibits non-negligible false acceptances for Over-baked (20–23%). Hence, if one wants to call the system "conservative", that conservatism is class-dependent: it holds for protecting the Normal stream from excessive scrap, but it is weaker exactly where the risk is highest. This asymmetry is precisely why a purely metric-based reporting of aggregate metrics would be incomplete.

3.3. Performance becomes valuable only through context

The scenario analysis was introduced to prevent a common methodological shortcut: treating detection accuracy as a direct proxy for manufacturing value. Fig. 3 reports the risk-weighted TCQ per batch for the three scenarios, mainly driven by scrap waste (FP), safety-related false negatives on the Over-baked class (FN_S), and latency costs, while FN_Q costs are negligible.

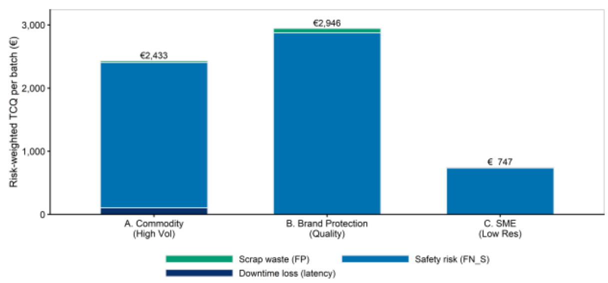

Fig. 3. Risk-weighted TCQ per batch across three scenarios, decomposed into FP, FN_S (safety), and latency costs; FN_Q costs are negligible.

3.3.1. Scenario A (Commodity)

Here, the target rate is 50 Hz, while YOLOv8 runs at about 42 FPS, i.e., a TE of roughly 0.83-0.84. The resulting TCQ is 2433 € per batch. This gap is not a "technical imperfection"; it is a system-level mismatch that generates an explicit downtime penalty in the TCQ decomposition (visible as the latency-related segment in Fig. 3). In high-volume regimes, even modest latency becomes economically amplified because time is monetised at scale.

3.3.2. Scenario B (Brand protection)

With a reduced target rate (30 Hz), latency is no longer binding (TE $\approx$ 1.4). The resulting TCQ rises to 2946 € per batch. It becomes dominated by risk-related costs, consistent with a premium logic where external failures and perceived inconsistency are disproportionately penalised. Importantly, this does not mean the system is "worse" than in Scenario A; it means the cost function is different. Scenario B is a reminder that TCQ is not an absolute property of the model, it is a property of the model embedded in a strategic environment.

3.3.3. Scenario C (SME)

At 40 Hz, YOLOv8 is approximately aligned with operational constraints (TE $\approx$ 1.03-1.05). The TCQ decreases to 747 € per batch, reflecting both the smaller batch size and the less punitive cost structure. Notably, the scrap-waste component remains limited in Fig. 3, indicating that YOLOv8 does not "buy" risk reduction by aggressively rejecting borderline products. For SMEs, this matters: waste is not only a cost item, but also a sustainability and handling burden.

From a sustainability angle, the WRI makes explicit the hidden cost of robustness: how many good products are unnecessarily discarded when operating conditions deteriorate. Under shop-floor stress, YOLOv8 generates $FP_{Stress} = 8$, compared with $FP_{Baseline} = 4$, which yields WRI = −100%. This negative value is indicative of a very specific phenomenon: the amount of avoidable scrap is doubled in comparison with the baseline reference (i.e., a 100% increase in false rejections). In summary, YOLOv8 demonstrates a heightened propensity for errors under stress, leading to the rejection of a significant proportion of compliant products, despite maintaining a headline accuracy that remains "mature" in terms of mAP.

When considered collectively, Fig. 2, 3 and Table 1 provide support for a concise yet substantial message, namely that YOLOv8 has achieved technical maturity and robustness; however, its industrial desirability remains contingent upon the specific scenario under consideration. In scenarios where takt time is stringent (Scenario A), the bottleneck shifts from accuracy to latency. In scenarios where risk penalties dominate (Scenario B), the residual weakness in the Over-baked class becomes the key managerial concern. In scenarios where resources are constrained (Scenario C), the model's balanced waste profile and near-alignment with line speed make it a plausible "good-enough" solution, provided that Over-baked misses are handled through process controls, threshold tuning, or complementary sensing.

4. Discussion and Conclusion

This discussion is organised around a simple claim: in industrial food inspection, technical robustness is necessary, but it is not a sufficient criterion for adoption. The evidence from the stress campaign shows that YOLOv8 can be "statistically impressive" and still be operationally or economically fragile, depending on the production logic in which it is embedded.

It is evident that YOLOv8 operates within a mature accuracy regime, also under shop-floor stress conditions. This is relevant because much of the food-vision literature still relies on controlled acquisition settings, while real deployment is constrained by illumination, noise, humidity, and calibration variability [16], [4]. However, robustness should not be interpreted as "problem solved". The stress diagnostics in Fig.1 show that the most critical class, Over-baked, remains the fragile point. In other words, robustness in mAP does not automatically mean robustness where it matters most: safety-relevant false acceptances are sparse events, but they dominate the overall risk. Consistently with the broader baked-goods literature, colour/browning is also a particularly sensitive attribute because it depends on reaction-driven surface development and is strongly influenced by process conditions [7].

The empirical evidence of this paper uses latency as a visible constraint (Scenario A), but the broader issue is more general: operability depends on where the algorithm is placed in the manufacturing system and what actions its output initiates. A detector is not a standalone artefact; it is a decision node connected to acquisition, computation, rejection mechanics, and process feedback. First, consider placement and architecture. Recent food-manufacturing deployments increasingly favour edge-based inspection architectures, precisely because inspection needs to be close to the physical process: it must acquire, infer, actuate (reject/accept), and report in real time [17]. This is not a stylistic implementation choice; it changes what "performance" means. A model that is excellent on a workstation but poorly integrated on the line may underperform a slightly weaker model that is properly embedded and synchronized.

Second, the "where" question is inseparable from the "what then?" question. A robust inspection system should not only classify; it should enable timely intervention. Real-time baked-goods monitoring studies explicitly discuss architectures where measured variables are displayed continuously and can generate control signals for process adjustment, reducing response time and keeping the process within narrower bounds [7]. In this view, vision is valuable not merely because it flags defects, but because it shifts quality control upstream, thereby reducing rework and stabilising the process, which is also the rationale behind in-line inspection along conveyors and ovens.

Third, model operability in production lines entails two hidden cost channels: a temporal one, where latency becomes economically critical not in absolute terms but relative to takt-time, forcing the system to absorb micro-stoppages and throughput reductions; and a material one, where false positives generate avoidable food scrap, as evidenced by

YOLOv8's WRI collapse under stress conditions, representing a tangible cost for SMEs and sustainability objectives that standard benchmarks fail to capture.

Once operability is framed as a joint problem of time alignment and material consequences, mAP alone cannot support deployment decisions. TE, WRI, and risk-weighted TCQ provide a compact translation from model behaviour to manufacturing outcomes, in line with the Quality 4.0 perspective that digital inspection is valuable only when technical outputs become actionable quality decisions across the system [18]. Together, these KPIs show that economic optimality is context-dependent and that hidden costs emerge from the model, context interaction rather than from the neural network alone.

Two limitations of the proposed approach should be acknowledged. First, the scenario cost parameters are approximations: realistic in magnitude, but not a substitute for plant-level accounting. Second, the study relies on a single dataset and a defined stress protocol; it cannot cover the full spectrum of drift mechanisms occurring on real lines, which are known challenges in food imaging [4]. Future work should prioritise physical deployment with explicit architectural variants and validation across multiple food categories where defect characteristics and risk asymmetries differ substantially. A further step is to extend from “inspection” to “intervention”, integrating detection outputs into process-control actions to shorten reaction times and stabilise critical quality variables.

Acknowledgements

This paper is part of the project NODES which has received funding from the MUR-M4C2 1.5 of PNRR funded by the European Union – NextGenerationEU (Grant agreement no. ECS00000036).

References

[1] A. Hassoun *et al.*, “Food traceability 4.0 as part of the fourth industrial revolution: key enabling technologies,” *Crit. Rev. Food Sci. Nutr.*, vol. 64, no. 3, pp. 873–889, 2024, doi: 10.1080/10408398.2022.2110033.

[2] A. Piovano, E. Verna, G. Genta, and M. Galetto, “Industry 4.0 enablers in food industry: a framework based on stakeholder challenges in quality and sustainability,” *International Journal of Quality and Reliability Management*, vol. 43, no. 2, pp. 349–372, Jan. 2026, doi: 10.1108/IJQRM-04-2025-0120/1300549.

[3] E. Purlis, “Browning development in bakery products – A review,” *J. Food Eng.*, vol. 99, no. 3, pp. 239–249, Aug. 2010, doi: 10.1016/j.jfoodeng.2010.03.008.

[4] S. J. Olakanmi, D. S. Jayas, and J. Paliwal, “Applications of imaging systems for the assessment of quality characteristics of bread and other baked goods: A review,” *Compr. Rev. Food Sci. Food Saf.*, vol. 22, no. 3, pp. 1817–1838, May 2023, doi: 10.1111/1541-4337.13131;SUBPAGE:STRING:FULL.

[5] E. Verna, A. Piovano, and M. Galetto, “Machine vision techniques for quality control in the wine industry,” *Discover Food 2025 5:1*, vol. 5, no. 1, pp. 425-457, Dec. 2025, doi: 10.1007/s44187-025-00706-x.

[6] M. M. Chetima and P. Payeur, “Automated tuning of a vision-based inspection system for industrial food manufacturing,” *2012 IEEE I2MTC - International Instrumentation and Measurement Technology Conference, Proceedings*, pp. 210–215, 2012, doi: 10.1109/I2MTC.2012.6229334.

[7] A. Görgülü, “Real-time quality analysis of baked goods using advanced technologies,” *J. Food Eng.*, vol. 388, p. 112359, Mar. 2025, doi: 10.1016/j.jfoodeng.2024.112359.

[8] O. Kılcı and M. Koklu, “Automated Classification of Biscuit Quality Using YOLOv8 Models in Food Industry,” *Food Analytical Methods 2025 18:5*, vol. 18, no. 5, pp. 815–829, Jan. 2025, doi: 10.1007/S12161-025-02755-5.

[9] “Home - Documentazione di Ultralytics YOLO.” Accessed: Nov. 12, 2025. [Online]. Available: https://docs.ultralytics.com/it/

[10] A. Mumuni and F. Mumuni, “Data augmentation: A comprehensive survey of modern approaches,” *Array*, vol. 16, p. 100258, Dec. 2022, doi: 10.1016/J.ARRAY.2022.100258.

[11] S. Dikici and R. J. Robinson, “Automated Defect Detection Using Image Recognition in Manufacturing,” *Journal of Data Science and Intelligent Systems*, Oct. 2024, doi: 10.47852/BONVIEWJDSIS42023833.

[12] L. J. Porter and P. Rayner, “Quality costing for total quality management,” *Int. J. Prod. Econ.*, vol. 27, no. 1, pp. 69–81, Apr. 1992, doi: 10.1016/0925-5273(92)90127-S.

[13] J. Hanhirova, T. Kämäräinen, S. Seppälä, M. Siekkinen, V. Hirvisalo, and A. Ylä-Jääski, “Latency and Throughput Characterization of Convolutional Neural Networks for Mobile Computer Vision,” *Proceedings of the 9th ACM Multimedia Systems Conference, MMSys 2018*, vol. 18, pp. 204–215, Mar. 2018, doi: 10.1145/3204949.3204975.

[14] A. Muffato Reis, E. Verna, L. Costa, S. D. Sousa, and M. Galetto, “A scenario-based framework for strategic inspection decision-making in high-volume production environments,” *International Journal of Quality & Reliability Management*, vol. 42, no. 6, pp. 1822–1850, May 2025, doi: 10.1108/IJQRM-03-2024-0100.

[15] N. Khadim *et al.*, “Quantifying the cost of quality in construction projects: an insight into the base of the iceberg,” *Quality & Quantity 2022 57:6*, vol. 57, no. 6, pp. 5403–5429, Jan. 2023, doi: 10.1007/S11135-022-01574-8.

[16] M. W. Alam *et al.*, “Emerging Trends in Food Process Engineering: Integrating Sensing Technologies for Health, Sustainability, and Consumer Preferences,” *J. Food Process Eng.*, vol. 48, no. 1, Jan. 2025, doi: 10.1111/jfpe.70035.

[17] S. W. Chen, C. J. Tsai, C. H. Liu, W. C. C. Chu, and C. T. Tsai, “Development of an Intelligent Defect Detection System for Gummy Candy under Edge Computing,” *Journal of Internet Technology*, vol. 23, no. 5, pp. 981–988, 2022, doi: 10.53106/160792642022092305006.

[18] H. Talaie, M. Ziaeian, and P. Malekinejad, “Towards quality 4.0 in home appliances: definitions, deployment scenarios, and future perspectives,” *Journal of Manufacturing Technology Management*, vol. 35, no. 7, pp. 1416–1438, Jul. 2024, doi: 10.1108/JMTM-02-2023-0044.