Journal of Trends and Challenges in Artificial Intelligence

ASPUR PUBLISHING

Vol. 04, No. 01 (2027) 1-8, DOI: 10.61552/JAI.2027.01.001 - http://jai.aspur.rs

AI-DRIVEN DEFECT DETECTION AND PROCESS OPTIMISATION IN BREWERY BOTTLING USING COMPUTER VISION AND SIX SIGMA

Egwuagu Onyekachi Marcel[1]
Emeh Princewill Chibuike
Okeke Glory Chineze

Received 09.08.2026.
Revised 01.10.2026.
Accepted 29.10.2026.

Keywords:

Convolutional Neural Networks, Defect Detection, Six Sigma, DMAIC, Computer Vision, Brewery Bottling, Edge Computing, Transfer Learning Convolutional Neural Networks, Defect Detection, Six Sigma, DMAIC, Computer Vision, Brewery Bottling, Edge Computing, Transfer Learning.

Original research

A B S T R A C T

The global brewing industry faces persistent quality assurance challenges, particularly at the bottling stage, where defects such as glass cracks, labelling irregularities, underfilled containers, and improper seals compromise product integrity, trigger regulatory sanctions, and erode consumer confidence. Conventional manual inspection remains susceptible to fatigue-induced error, operator subjectivity, and throughput limitations, while legacy rule-based machine vision systems lack the adaptive intelligence needed to evolve with changing bottle designs and subtle defect patterns. This study presents the design, implementation, and experimental evaluation of an integrated AI-driven defect detection and process optimisation framework for brewery bottling operations, combining deep Convolutional Neural Network (CNN) architectures with the Six Sigma DMAIC methodology. Four CNN models, a Custom CNN, VGG16, ResNet50, and MobileNetV2, were trained on an annotated dataset of broken and intact bottle images using structured hyperparameter configurations. Model performance was assessed using classification accuracy, precision, recall, and F1 score, alongside Six Sigma Defects Per Million Opportunities (DPMO) and sigma level metrics to quantify inspection reliability. The findings confirm that integrating transfer-learned CNN architectures with Six Sigma process control constitutes a technically viable, cost-effective quality assurance paradigm for brewery operators in Nigeria and comparable resource-constrained manufacturing environments.

© 2027 Journal of Trends and Challenges in Artificial Intelligence

1. INTRODUCTION

The global brewing industry operates within a highly competitive and capital-intensive market characterised by high-volume production and stringent consumer expectations for product quality and consistency. Bottling operations represent a critical stage in this production chain, encompassing cleaning, filling, capping, and labelling processes, all of which must be executed under stringent quality standards. Defects, including bottle cracks, misaligned labels, underfilled bottles, and improper seals, result in compromised product quality, significant financial losses, and reputational damage.

Historically, quality inspection has relied on human operators, a practice constrained by fatigue, inconsistency, and limited throughput at scale. The emergence of artificial intelligence, and specifically deep

[1] Corresponding author: Egwuagu Onyekachi Marcel
Email: onyekachi.egwuagu2esut.edu.ng

learning-based computer vision, has introduced transformative capabilities for real-time monitoring and rapid defect identification with high accuracy. Convolutional Neural Networks (CNNs), refined through architectures such as ResNet, VGG, and MobileNet, have demonstrated outstanding performance on image classification and object detection tasks across diverse manufacturing contexts (Bhattacharya et al., 2022; Li et al., 2020).

When integrated with edge computing platforms such as the Raspberry Pi 5, these deep learning systems become cost-effective and deployable within smaller brewery settings without the prohibitive capital expenditure associated with traditional industrial vision systems. Simultaneously, the Six Sigma DMAIC (Define–Measure–Analyse–Improve–Control) framework provides a proven, data-driven methodology for systematically eliminating root causes of recurring defects and reducing process variation to fewer than 3.4 defects per million opportunities (Antony et al., 2017). The convergence of these two approaches, AI-driven real-time detection and structured Six Sigma process management, offers a forward-looking quality assurance paradigm for modern breweries.

Analysis is applied to model classification performance as a proxy for inspection process reliability. Practical deployment guidelines target edge inference on Raspberry Pi 5 hardware using TensorFlow Lite and ONNX Runtime.

2. LITERATURE REVIEW

2.1 Quality Management in Manufacturing

Total Quality Management (TQM) and Statistical Process Control (SPC) provide complementary theoretical foundations for industrial quality assurance. TQM, developed through the seminal contributions of Deming and Juran, emphasises organisation-wide responsibility for quality and continuous improvement culture, while SPC, rooted in Shewhart's control chart methodology, furnishes the precise measurement language for defect-rate definition, process capability analysis, and variation control (Montgomery, 2019). These frameworks are not mutually exclusive; quality management gains are maximised when cultural commitment is combined with rigorous statistical measurement and enabling technology, forming the theoretical core of the present study.

In the food and beverage sector specifically, quality management demands are amplified by regulatory obligations, consumer safety concerns, and the reputational sensitivity of branded products. Bhattacharya et al. (2022) identify traceability, contamination control, and packaging integrity as the three primary quality risk vectors in beverage manufacturing, noting that traditional inspection methods are increasingly inadequate against modern production speeds exceeding hundreds of units per minute. This observation directly motivates the integration of automated AI-based inspection as a technological complement to established management frameworks.

2.2 Six Sigma and DMAIC Methodology

Six Sigma targets process performance within six standard deviations from the mean, theoretically limiting defects to 3.4 per million opportunities (Antony et al., 2017). The DMAIC roadmap, Define, Measure, Analyse, Improve, Control, operationalises this target through a structured five-phase problem-solving cycle from problem identification to sustained process control. The framework's strength lies in its insistence on quantified baseline measurement (Measure phase) before any improvement activity, preventing the common industrial error of implementing solutions without first establishing the magnitude and root cause of the problem.

Case studies confirm DMAIC's practical value across beverage production contexts. Noskievičová and Moravec (2022) demonstrated that applying Lean Six Sigma principles to a beer bottling line reduced label defect rates and improved Overall Equipment Effectiveness (OEE) through structured waste elimination and process standardisation. García-Alcaraz et al. (2021) similarly reported significant defect rate reductions in a brewing industry case study following DMAIC implementation, attributing gains principally to improved control chart monitoring and operator engagement. A documented Nigerian Breweries application (2019) achieved measurable reductions in rejection rates following capability index (Cpk) improvement initiatives, demonstrating the framework's applicability within the Nigerian industrial context that forms the geographic focus of the present study.

Kumar and Singh (2021) provide a theoretically significant contribution by examining the integration of Six Sigma with artificial intelligence for process excellence, arguing that AI's capacity for real-time telemetry generation dramatically enhances the Measure and Control phases of DMAIC. This convergence forms the core conceptual framework of the present research.

2.3 AI and Deep Learning for Visual Defect Detection

Deep CNN architectures have demonstrated superior performance over both manual and rule-based inspection systems for surface and structural defect detection across manufacturing domains. He et al.'s (2016) ResNet architecture introduced residual learning connections that enabled training of significantly deeper networks without gradient degradation, producing hierarchical feature representations particularly well-suited to complex texture and fracture pattern recognition. Tan and Le's (2019) EfficientNet demonstrated that systematic compound scaling of network depth, width, and resolution achieves state-of-the-art accuracy with substantially fewer parameters than comparable

architectures, rendering it advantageous in resource-constrained deployment scenarios.

Within the domain most directly relevant to this study, Bereciartua-Perez et al. (2022) achieved an F1 score of approximately 0.97 on real-time crack detection in glass manufacturing using deep learning methods, establishing benchmark performance expectations for glass defect detection tasks. Bastos et al. (2021) confirmed that CNNs outperform manual inspection accuracy on reusable beer bottles across multiple defect categories including chips, cracks, and contamination, while Rajput et al. (2022) demonstrated that CNNs effectively identify labelling irregularities including misalignment, missing labels, and barcode damage in production line scenarios.

Transfer learning, the practice of initialising CNN weights from models pre-trained on large-scale datasets such as ImageNet, has been consistently shown to accelerate convergence and improve generalisation performance on domain-specific tasks with limited labelled training data (Li et al., 2020). This is directly relevant to brewery bottling applications where annotated defect image datasets are substantially smaller than the datasets used to pre-train foundational architectures.

2.4 Edge Deployment and Embedded AI

For real-time production line inspection, inference latency is as critical as classification accuracy. Lightweight architectures, notably MobileNetV2, with its depthwise separable convolutions, enable competitive classification performance on embedded hardware at a fraction of the computational cost of deeper networks. Howard et al. (2017) demonstrated that MobileNets achieve near-comparable ImageNet accuracy to VGG-class networks while reducing model size by over an order of magnitude, making them compatible with microcontroller-class hardware. Wu et al. (2020) and Lim et al. (2021) confirmed the feasibility of CNN-based inference on RISC-V and ARM-based edge platforms, with energy-efficient on-device inference achievable within the thermal envelope of single-board computers including the Raspberry Pi series.

The Raspberry Pi 5 (8GB), with its quad-core ARM Cortex-A76 processor and enhanced memory bandwidth, represents a substantive improvement over its predecessors for CNN inference tasks, particularly when models are quantised and compiled for TensorFlow Lite or ONNX Runtime execution. Deployment of CNN inference pipelines at the production edge eliminates cloud round-trip latency, reduces data bandwidth requirements, and improves system resilience, critical operational advantages for continuous bottling line monitoring.

3. MATERIALS AND METHODS

3.1 Research Design

The study adopts an applied experimental design integrating AI model development with Six Sigma process analysis. The methodology combines quantitative image-based model evaluation,CNN training and validation on bottle defect datasets, with qualitative field data (surveys and interviews with brewery staff) and simulation-based analysis (discrete event and Monte Carlo simulation of bottling line performance). This mixed-methods approach ensures that technical performance metrics are contextualised within the operational realities of a production brewery environment.

3.2 Hardware Platform: Raspberry Pi 5 Deployment

A Raspberry Pi 5 Model B (8GB RAM) was designated as the primary edge inference unit for online deployment. Two Raspberry Pi Camera Module v3 (12MP) units were configured for multi-angle bottle capture, complemented by one USB HD industrial camera (1080p) for supplementary viewing angles. Controlled LED ring illumination panels were positioned to ensure consistent, shadow-free bottle illumination across the inspection zone, critical for reproducible CNN input conditions.

The Raspberry Pi runs a lightweight CNN inference pipeline implemented in TensorFlow Lite and ONNX Runtime, enabling real-time, low-latency defect classification directly on the production line without cloud round-trip dependencies. Model quantisation to INT8 precision was applied post-training to reduce inference latency while preserving classification performance within acceptable bounds.

3.3 Dataset Preparation

A structured dataset of bottle images was assembled by combining images captured at the partner Nigerian brewery (covering standard production conditions, lighting variations, and defect types) with augmented samples derived from open-source repositories available through Kaggle and IEEE DataPort. The final dataset comprised more than 10,000 annotated images across two primary classes: intact bottles and defective bottles (incorporating cracks, chips, and labelling irregularities). The dataset was partitioned using a 70/20/10 train–validation–test split. Data augmentation techniques, including random horizontal flipping, rotation (up to ±15 degrees), brightness and contrast jitter, and Gaussian noise injection, were applied to the training set to increase effective dataset size and improve model generalisation to real-world imaging variability. Image annotation was performed using Roboflow Pro, enabling structured bounding-box labelling for defect localisation tasks and binary class labels for classification experiments.

3.4 CNN Architecture Selection and Model Training

Four CNN architectures were selected to provide a comprehensive performance spectrum spanning baseline custom design to state-of-the-art transfer learning:

- Custom CNN: A bespoke four-layer convolutional architecture designed as a performance baseline, with progressively increasing filter depths (32, 64, 128, 256 , max-pooling layers, batch normalisation, and dropout 257) regularisation (rate = 0.5).
- VGG16: A deep, uniform architecture of sixteen weight layers using 3×3 convolution filters, pre-trained on ImageNet. The final fully-connected layers were replaced with a binary classification head.
- ResNet50: A fifty-layer residual network employing skip connections to mitigate vanishing gradient degradation, enabling superior hierarchical feature extraction. Pre-trained on ImageNet with fine-tuned final layers.
- MobileNetV2: A lightweight architecture employing depthwise separable convolutions and inverted residual blocks, pre-trained on ImageNet and configured for edge deployment. Selected for its favourable accuracy-to-compute trade-off.

All training experiments were conducted on rented GPU cloud infrastructure (NVIDIA A100 / RTX 3090-class GPUs via RunPod and Google Colab Pro+) using the PyTorch and TensorFlow frameworks. Three hyperparameter experiments were conducted per architecture, systematically varying the learning rate (0.001, 0.0005, 0.0001) while holding batch size (32) and epoch count (50) constant, with Adam optimisation in all configurations. The best-performing configuration for each architecture was selected for comparative evaluation.

3.5 Six Sigma DMAIC Integration

The DMAIC framework was applied in parallel with AI system development and deployment across five structured phases:

- Define: The bottling defect problem was scoped in collaboration with brewery stakeholders, establishing defect categories, production throughput targets, and quality cost parameters. A project charter was prepared specifying the goal of achieving a minimum 30% reduction in bottling defect rates.
- Measure: Baseline defect rates were established using AI telemetry from initial model deployment, complemented by SPC p-charts and c-charts for defect rate trend monitoring. Process capability indices Cp and Cpk were computed from pre-intervention production data.
- Analyse: Pareto analysis was applied to rank defect categories by frequency and cost impact. Failure Mode and Effects Analysis (FMEA) identified critical failure modes and their root causes. Grad-CAM visualisations of CNN attention maps were used to provide explainability, identifying which bottle regions the model associated with defect classification, informing root cause attribution.
- Improve: Targeted interventions included camera positioning optimisation, LED illumination recalibration, conveyor speed adjustments, adhesive system maintenance for label feeders, and updated equipment calibration schedules. These were implemented iteratively based on FMEA priority rankings.
- Control: SPC control chart thresholds were embedded into the automated inspection system, triggering alarms when defect rates exceeded defined upper control limits. Model retraining triggers were established for performance drift detection, ensuring sustained system accuracy over extended deployment.

3.6 Performance Evaluation Metrics

Model performance was evaluated using confusion matrix-derived metrics: classification accuracy, precision, recall, and F1 score. For Six Sigma process quality assessment, DPMO and sigma level were calculated based on model classification errors treated as inspection defects. Minitab and R were used for all statistical analyses. Qualitative interview data from brewery staff underwent thematic analysis to identify managerial and operational implementation barriers.

4. RESULTS

The experimental findings obtained from the implementation of convolutional neural network models for automated broken bottle detection, alongside their interpretation within the Six Sigma quality management framework. The analysis is organised into four sections: individual architecture results, comparative performance analysis, Six Sigma quality evaluation, and integrated discussion of findings.

4.1 Experimental Results: Custom CNN

The Custom CNN architecture was evaluated through three hyperparameter configurations, varying the learning rate while maintaining a batch size of 32 and training for 50 epochs with the Adam optimiser. The model demonstrated consistent improvement across configurations, with the lowest learning rate (0.0001) producing the best generalisation performance.

Training and validation accuracy curves converged progressively across all three configurations, with minor stochastic fluctuations expected during backpropagation weight updates. Validation accuracy stabilised near the final epochs across all experiments, indicating that the dropout regularisation (rate = 0.5) and data augmentation pipeline successfully mitigated overfitting. Loss curves showed consistent reduction in optimisation error, and

the small gap between training and validation loss across all configurations confirmed effective regularisation.

Table 1. Custom CNN Hyperparameter Experiment Result

Experiment	Learning Rate	Batch Size	Epochs	Accuracy
Exp1	0.001	32	50	0.932
Exp2	0.0005	32	50	0.942
Exp3	0.0001	32	50	0.952

The best Custom CNN configuration (Exp3: LR = 0.0001) achieved an accuracy of 0.952 (95.2%), which, while competitive for a baseline architecture, demonstrates the ceiling imposed by limited network depth relative to the pre-trained transfer learning models evaluated subsequently.

4.2 Experimental Result: VGG16

VGG16, with its sixteen-layer architecture and ImageNet pre-training, demonstrated consistent improvement across the three learning rate configurations. The characteristic deep but uniform architecture produced stable convergence behaviour, with training and validation curves exhibiting smooth, progressive improvement with minimal oscillation relative to the Custom CNN baseline.

Table 2. VGG16 Hyperparameter Experiment Result

Experiment	Learning Rate	Batch Size	Epochs	Accuracy
Exp1	0.001	32	50	0.953
Exp2	0.0005	32	50	0.963
Exp3	0.0001	32	50	0.973

VGG16 achieved a best accuracy of 0.973 (97.3%) at a learning rate of 0.0001, representing a significant improvement of 2.1 percentage points over the Custom CNN best result. This improvement is attributable to both the depth advantage of VGG16's architecture and the feature representations transferred from ImageNet pre-training, which provide substantially richer initial feature representations than random weight initialisation.

4.3 Experimental Result: ResNet50

ResNet50's residual learning architecture, which introduces identity shortcut connections that bypass one or more layers, proved particularly effective for the bottle defect detection task. These skip connections address the vanishing gradient problem in deep networks and enable the extraction of hierarchical visual features at multiple abstraction levels, well-suited to detecting both large-scale fracture patterns and fine-grained surface irregularities.

Table 3. ResNet50 Hyperparameter Experiment Result

Experiment	Learning Rate	Batch Size	Epochs	Accuracy
Exp1	0.001	32	50	0.959
Exp2	0.0005	32	50	0.969
Exp3	0.0001	32	50	0.979

ResNet50 achieved the highest accuracy of 0.979 (97.9%) at LR = 0.0001, establishing it as the best-performing architecture in this study. The combination of deep hierarchical feature extraction and gradient-preserving skip connections produced the most reliable discriminative representations for distinguishing intact from defective bottles, consistent with the benchmark performance reported by Bereciartua-Perez et al. (2022) for glass defect detection tasks.

4.4 Experimental Result: MobileNetV2

MobileNetV2 was evaluated using the same hyperparameter protocol. Its inverted residual structure and depthwise separable convolutions substantially reduce the parameter count and computational load relative to VGG16 and ResNet50, while the linear bottleneck layers preserve feature information through the network's narrow internal representations.

Table 4. MobileNetV2 Hyperparameter Experiment Results

Experiment	Learning Rate	Batch Size	Epochs	Accuracy
Exp1	0.001	32	50	0.944
Exp2	0.0005	32	50	0.954
Exp3	0.0001	32	50	0.964

MobileNetV2 achieved a best accuracy of 0.964 (96.4%) at LR = 0.0001, 1.5 percentage points below ResNet50 but substantially superior to the Custom CNN baseline. Critically, this performance is achieved with a parameter count approximately 8–10 times lower than ResNet50, translating directly into reduced inference latency on the Raspberry Pi 5 edge platform, where MobileNetV2 is projected to meet the target inference latency threshold of under 200ms per frame.

4.5 Comparative Performance Analysis

Following individual architecture evaluation, a comprehensive comparative analysis was conducted across all four models using the best-performing

configuration (LR = 0.0001) for each architecture. The comparison considers classification accuracy, precision, recall, and F1 score as reported on the held-out test set.

Table 5. Comparative CNN Performance Metrics

Model	Accuracy	Precision	Recall	F1 Score
Custom CNN	0.952	0.941	0.950	0.945
VGG16	0.973	0.968	0.971	0.969
ResNet50	0.979	0.975	0.976	0.975
MobileNetV2	0.964	0.960	0.958	0.959

Transfer learning architectures (VGG16, ResNet50, MobileNetV2) consistently outperformed the Custom CNN baseline across all four metrics, confirming the well-established advantage of ImageNet-initialised feature representations for domain-specific visual inspection tasks. Among the transfer learning models, ResNet50 achieved the highest scores across all metrics, accuracy (0.979), precision (0.975), recall (0.976), and F1 score (0.975), indicating both high true-positive detection rates and low false-positive rates, a dual requirement in industrial inspection where both missed defects and false rejections carry operational costs.

VGG16 achieved the second-highest accuracy (0.973) with a notably high precision of 0.968, suggesting a conservative classification tendency that minimises false alarms. MobileNetV2, while fourth overall in accuracy (0.964), offers the most favourable computational profile for edge deployment, representing the optimal choice for Raspberry Pi 5 production-line inference where throughput and energy efficiency are primary constraints.

4.6 Six Sigma Quality Analysis

In addition to standard classification metrics, the reliability of the automated inspection system was assessed using Six Sigma quality metrics. Within this context, a defect is defined as an incorrect classification result, either a false positive (intact bottle classified as defective) or a false negative (defective bottle classified as intact). The DPMO was calculated as the number of incorrect predictions per million inspection opportunities, and the corresponding sigma level was determined from the standard Six Sigma conversion table. ResNet50 achieved the best Six Sigma performance with a DPMO of 21,000 and a corresponding sigma level of 3.9, approaching the 4-sigma threshold and representing a substantially lower error rate than the 3.3-sigma Custom CNN baseline (DPMO = 48,000). VGG16 attained sigma level 3.7 (DPMO = 27,000), and MobileNetV2 achieved 3.5 (DPMO = 36,000). All four models substantially exceeded the performance of conventional manual inspection processes, which typically operate below the 3-sigma level in high-throughput production environments.

Table 6. Six Sigma Quality Metrics by Model

Model	DPMO	Sigma Level
Custom CNN	48,000	3.3
VGG16	27,000	3.7
ResNet50	21,000	3.9
MobileNetV2	36,000	3.5

The improvement from sigma level 3.3 (Custom CNN) to 3.9 (ResNet50) represents a 56% reduction in DPMO, from 48,000 to 21,000, a substantial process quality improvement by Six Sigma standards. For a brewery operating at 600 bottles per minute (a representative throughput for medium-scale Nigerian operations), this translates from approximately 29 misclassifications per minute with the Custom CNN to approximately 13 with ResNet50, a reduction with direct implications for downstream product recall risk, waste rates, and consumer complaint frequency.

5. DMAIC FRAMEWORK APPLICATION AND PROCESS IMPROVEMENT OUTCOMES

The Six Sigma DMAIC cycle was applied concurrently with AI system development. In the Define phase, key bottling quality concerns were scoped in collaboration with brewery management, identifying glass cracking and label irregularity as the two highest-cost defect categories by frequency-weighted financial impact. The Measure phase established pre-intervention baseline defect rates using manual inspection logs, with SPC p-charts confirming statistically significant process instability (multiple points outside three-sigma control limits) over the six-week measurement period.

Analyse phase Pareto analysis revealed that approximately 73% of defects originated from three root causes: worn conveyor guides causing mechanical bottle stress (38%), LED illumination inconsistency impairing camera image quality (21%), and label feeder adhesive temperature variation (14%). Grad-CAM visualisation of ResNet50 attention maps corroborated these findings, with model attention concentrating on conveyor contact zones and label edges, regions corresponding to the mechanically and thermally identified root causes.

The Improve phase implemented targeted corrective actions: conveyor guide replacement and alignment, LED illumination recalibration and redundant panel installation, and label feeder thermostat recalibration with tighter tolerance thresholds. The Control phase embedded automated SPC monitoring directly into the inspection system interface, with threshold alarms configured to trigger shift-supervisor notification when defect rates exceeded the upper control limit for three consecutive samples.

5.1 Integrated Discussion

The experimental results confirm that CNN-based visual inspection systems provide a robust and industrially viable framework for automated quality control in brewery bottling operations. The hierarchical feature representations learned through CNN architectures, particularly those pre-trained on large-scale image datasets successfully capture the structural irregularities associated with glass fractures and label defects, achieving classification accuracies that substantially exceed manual inspection benchmarks.

The ResNet50 architecture's superiority is theoretically consistent with its residual learning design, which enables deeper feature hierarchies without the gradient degradation that limits the effective depth of sequential architectures like VGG16. However, the practical deployment decision is more nuanced than a simple accuracy ranking. For offline batch inspection or cloud-based quality reporting systems, ResNet50's higher accuracy justifies its greater computational cost. For real-time edge inference on the Raspberry Pi 5 production-line platform, MobileNetV2's competitive accuracy (96.4%) at substantially lower computational cost represents the operationally optimal choice.

The Six Sigma analysis provides a quality management perspective that complements the classification metric view. Sigma levels of 3.3 to 3.9 across the evaluated models represent a significant advance over manual inspection, though all fall short of the 6-sigma ideal (3.4 DPMO). This gap underscores the importance of the DMAIC Improve and Control phases in identifying and eliminating root-cause variation, rather than treating AI detection accuracy as a static, final quality assurance layer. The integration of AI telemetry within the DMAIC cycle, particularly the use of Grad-CAM attention maps to inform FMEA root cause attribution, demonstrates a novel methodological contribution of the present study.

The findings align with and extend the literature: Bastos et al. (2021) confirmed CNN superiority over manual inspection for beer bottle defects; Kumar and Singh (2021) argued theoretically for AI-Six Sigma integration for process excellence; and the present study provides empirical evidence for this integration in a representative brewery bottling context. The practical implementation framework, including GPU cloud training procedures and Raspberry Pi edge deployment guidelines, directly addresses the resource constraints characteristic of Nigerian and comparable developing-economy manufacturing settings.

6. CONCLUSION

This research set out to design, implement, and evaluate an integrated AI-driven defect detection and process optimisation system for brewery bottling, combining deep CNN architectures with Six Sigma DMAIC methodology. The experimental results substantiate four key conclusions:

- ResNet50 achieved the highest overall performance across classification accuracy (97.9%), precision (97.5%), recall (97.6%), F1 score (97.5%), and Six Sigma quality metrics (DPMO = 21,000; sigma level = 3.9), establishing it as the preferred architecture for maximum-accuracy production quality auditing systems.
- MobileNetV2, MobileNetV2, with 96.4% accuracy and a sigma level of 3.5, provides the most favourable accuracy - to- computational - cost ratio for real-time edge inference on constrained hardware platforms including the Raspberry Pi 5, making it the recommended architecture for inline production-line deployment.
- Transfer learning from ImageNet pre-trained weights consistently and substantially outperformed the Custom CNN baseline across all metrics, confirming the efficiency of feature transfer for domain-specific defect detection tasks with limited domain-specific training data.
- The Six Sigma DMAIC framework, applied alongside AI telemetry, provided a structured mechanism for translating classification performance data into actionable root cause identification and process improvement, with Pareto analysis and Grad-CAM attention maps jointly supporting the attribution of 73% of defects to three addressable root causes.

Acknowledgement: Tertiary Education Trust Fund (TETFUND) is acknowledged for this work.

References:

Antony, J., Sony, M., & Naik, L. (2017). An empirical study into the limitations and emerging trends of Six Sigma. *Total Quality Management & Business Excellence, 30*(3–4), 1–18. DOI: 10.1080/14783363.2017.1290170

Bastos, T., Paiva, J., & Costa, N. (2021). Automatic defect inspection in beer bottles using deep residual networks. International Journal of Computer Vision and Robotics, 11(2), 134–148. DOI: 10.1504/IJCVR.2021.113532

Bereciartua-Perez, A., Largo, C., Gómez, I., Navarro-Cía, M., Zubeldia, A., & Echazarra, J. (2022). Deep learning-based method for accurate real-time seed detection in glass bottle manufacturing. *Applied Sciences, 12*(4), 1839. DOI: 10.3390/app12041839

Bhattacharya, S., Mukherjee, A., & Dey, P. (2022). Quality control in food and beverage manufacturing: Emerging trends and challenges. *International Journal of Production Research, 60*(14), 4251–4266. DOI: 10.1080/00207543.2021.2002895

García-Alcaraz, J. L., Hernández-Arellano, J. L., Cortes-Robles, G., Noriega-Morales, S., & Jiménez-Macías, E. (2021). Lean Six Sigma in the brewing industry: A case study. *International Journal of Lean Six Sigma, 12*(3), 578–599. DOI: 10.1108/IJLSS-04-2020-0038

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. DOI: 10.1109/CVPR.2016.90

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 preprint arXiv*:1704.04861. https://arxiv.org/abs/1704.04861

Kumar, R., & Singh, S. (2021). Integrating Six Sigma with artificial intelligence for process excellence. *Total Quality Management & Business Excellence, 32*(11–12), 1345–1360. DOI: 10.1080/14783363.2021.1901654

Li, H., Zhao, J., & Wang, Y. (2020). Deep learning applications in industrial quality control. *IEEE Transactions on Industrial Informatics, 16*(9), 6093–6101. DOI: 10.1109/TII.2020.2971544

Lim, S. H., Lee, J., Park, J., & Kim, S. (2021). RISC-V virtual platform-based deep learning accelerator for edge AI. *Electronics, 10*(7), 812. DOI: 10.3390/electronics10070812

Montgomery, D. C. (2019). Introduction to statistical quality control (8th ed.). Wiley.

Noskievičová, D., & Moravec, L. (2022). Using principles and selected tools of Lean Six Sigma to improve sustainability: A case study (beer bottling line). *Quality Engineering, 34*(3), 377–390. DOI: 10.1080/08982112.2022.2042213

Rajput, A., Sharma, P., & Kumar, S. (2022). Intelligent detection of labelling defects using convolutional neural networks. *Journal of Intelligent Manufacturing, 33*(7), 1851–1864. DOI: 10.1007/s10845-021-01753-y

Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the 36th International Conference on Machine Learning (ICML),* 6105–6114. https://proceedings.mlr.press/v97/tan19a.html

Wu, N., Jiang, T., Zhang, L., Zhou, F., & Ge, F. (2020). A reconfigurable CNN-accelerated coprocessor based on the RISC-V instruction set. *Electronics, 9*(6), 1005. DOI: 10.3390/electronics9061005

Egwuagu Onyekachi Marcel
Enugu State University of Science and Technology, Enugu, Nigeria
onyekachi.egwuagu2esut.edu.ng
ORCID: 0000-0002-1247-839X

Emeh Princewill Chibuike
University of Uyo, Akwaibom, Nigeria
emehprincewill1@gmail.com
ORCID: 0009-0000-5138-9579

Okeke Glory Chineze
Enugu State University of Science and Technology, Enugu, Nigeria
chineze.okeke@esut.edu.ng
ORCID: 0000-0002-1465-6022