Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 4, Issue 1, 2026

ADVANCING MANUFACTURING QUALITY CONTROL: YOLOV7-BASED BOTTLE DEFECT DETECTION WITH MULTI-SCALE TRAINING AND OPTIMIZATIONS

Imtiaz Ali Shah[1], Umer Farooq[2], Aman Ullah[*3], Faiza Aurang Zeb[4]

[1]*Department of computer science, COMSATS university of science and technology, Abbottabad, Pakistan*
[2]*Department of computer science, the Islamia University of Bahawalpur, Pakistan*
[*3]*Department of electrical and computer engineering, South China University of Technology (SCUT), Guangzhou 510006, China*
[4]*Department of computer science, Govt Post Graduate College Mandian, Abbottabad, Pakistan*

[*3]amanafridi865@gmail.com

DOI: https://doi.org/10.5281/zenodo.18265499

Keywords
Manufacturing quality control, VOLOv7, Object detection, Real-time processing, Deep learning, Manufacturing quality control,

Article History
Received: 23 November 2025
Accepted: 30 December 2025
Published: 15 January 2026

Corresponding Author: *
Aman Ullah

Abstract
Guaranteeing manufacturing quality control in bottle production often encounters incompetence and inaccuracies due to reliance on manual or conventional methods for defect detection. This study addresses these challenges by leveraging the YOLOv7 object detection framework, recognized for its real-time processing capabilities and accuracy. Using the Roboflow platform, we organized and augmented a dataset focusing on bottle defects, such as missing caps and deformed surfaces, and fine-tuned a model pretrained on the COCO dataset for our specific application. Multi-scale training and optimizations, including anchor box refinement and spatial attention mechanisms, enhanced the model's detection performance across varying defect types and conditions. The proposed approach achieves a mean Average Precision (mAP) of 83%, with a precision of 86.7% and a recall of 95%, reflecting robust performance. The study highlights the feasibility of real-time defect detection in industrial environments, reducing production waste and ensuring highquality standards. Practical deployment considerations, combined with a focus on real-time processing, underscore the model's potential to replace traditional methods, paving the way for automated quality control solutions. These findings contribute to the advancement of computer vision in industrial applications, setting a precedent for future research in automated defect detection.

INTRODUCTION

Ensuring manufacturing quality control in bottle production often encounters inefficiencies and inaccuracies due to reliance on manual or conventional methods for defect detection [1]. Applying a deep learning-based techniques to such a dataset require a huge number of photos and algorithms that are computationally expensive. The use of machine vision that is based on simple image processing is yet another method that has been documented by numerous studies [11,12,13,14,15]. Image processing techniques, including as grayscale, binarization, morphological transformation, edge detection, and threshold segmentation, were proposed by Fu et al. [11,12] for the purpose of inspecting glass bottles for defects. A cap inspection system for pharmaceutical bottles was proposed by Kumchoo et al. [13], which had an accuracy of 84% for detecting safety rings and an accuracy of 87% for detecting loose caps. In a similar manner, Xie et al. [14] were able to detect the cap defect with a 97% accuracy rate by measuring the distance between the support ring and the cap. Image processing techniques, such as grayscale, threshold, and morphological transforms, were utilized by Saad et al. [15] in order to extract image characteristics. These techniques were utilized in order to detect shape faults in glass bottles. Following that, these characteristics were utilized as inputs in the Naïve Bayes Classifier, which categorized them according to the shape of the bottle under consideration.

The research presented explores the application of the YOLOv7 object detection algorithm to enhance defect detection in bottle manufacturing, a crucial component of quality assurance in production environments. YOLOv7, an abbreviation for "You Only Look Once version 7," is an advanced deep learning framework known for its ability to perform real-time object detection efficiently [2] [3]. This capability is particularly advantageous in industrial settings where timely and accurate defect identification is essential for maintaining high-quality standards. This study utilizes the Roboflow platform for effective dataset management, which involves organizing, preprocessing, and augmenting image data specifically for machine learning applications. By starting with a model pre-trained on the COCO dataset, the research benefits from initial weights that help in learning generic object features, which are further fine-tuned to detect specific bottle defects such as missing caps and deformed surfaces. The model's performance is assessed through key metrics such as precision, recall, and mean Average Precision (mAP), with mAP serving as a comprehensive indicator that combines both precision and recall across all classes. This research not only showcases the potential of integrating advanced machine learning models like YOLOv7 into the manufacturing process but also underscores the improvements it brings to traditional defect detection methods by enhancing speed and accuracy, thereby ensuring superior quality control and reduced production waste. As a result, the purpose of this work is to offer a method that is both effective and quick for checking flaws in plastic bottles. With an overall accuracy of 95%, the system is quite ideal for rapid industrial inspection because it is able to identify a variety of faults that may be found in plastic bottles through the utilization of straightforward image processing techniques.

The contribution of this work is a customized YOLOv7-based object detection framework for bottle defect detection, considering special characteristics of the bottle packaging to achieve better accuracy of the defect classification such as cap missing and crumpled surface. The focus on real-time processing makes the model fast and efficient, which is practical for most industrial applications. The study also examines the practical implementation of the model in real manufacturing settings, assessing its feasibility and various challenges in order to improve its applicability in production scenarios.

MATERIAL AND METHODS

The development of an accurate diabetes prediction model is the goal of our collaborative process, which is comprised of multiple stages. Data collection, preprocessing, feature selection, model training, model evaluation, and deployment are the phases that you will go through during this process. Every phase is extremely important in ensuring that the prediction model is accurate and reliable.

A. DATASET SELECTION and LOADING

a) **Dataset Management:** Roboflow allows you to organize and manage your dataset efficiently. We can upload images, annotate objects, and store metadata associated with each image.

b) **Annotation Support:** The platform likely supports annotation functionalities, enabling us to label and annotate objects in our images. This is crucial for training object detection models like YOLOv7, as the model needs to learn to identify and locate objects within images.

c) **Data Augmentation**: Roboflow often provides data augmentation tools to artificially increase the diversity of our dataset. This can involve applying transformations such as rotation, scaling, and flipping to generate additional training samples and improve model generalization.

B. Preprocessing

a) **Image Resizing:** Resize all images to a consistent size that is compatible with the input size expected by the YOLOv7 model. We use 640x640 pixels image size of our dataset.

b) **Normalization:** Normalize pixel values to bring them within a specific range, often [0, 1] or [-1, 1]. Normalization helps the model converge faster during training.

C. COCO dataset as Pre-Trained Weights

- To leverage pre-trained weights from the COCO dataset for YOLOv7 training, first download the COCO pre-trained weights file from the official YOLOv7 GitHub repository (e.g., `yolov7_coco.pt`).
- Monitor the training progress and evaluate the model's performance on our validation dataset to ensure effective learning and generalization to our specific bottle defect detection task.
- Utilizing COCO pre-trained weights can expedite convergence and enhance the model's performance on our custom task by leveraging knowledge gained from the broader COCO dataset.

D. Proposed Model

The YOLOv7 architecture, an advancement in the YOLO series, incorporates key features to enhance real-time object detection. Utilizing CSPDarknet53 as its backbone network, YOLOv7 integrates a PANet neck design for effective feature aggregation across different scales. The detection head, incorporating anchor boxes and a spatial attention mechanism, predicts bounding box coordinates, class probabilities, and abjectness scores. Operating on a grid cell basis, the YOLO layer processes final predictions, and non-maximum suppression is applied for refined output. The model's multi-scale training strategy enhances its ability to generalize objects of varying sizes, while optimizations contribute to speed and accuracy, making YOLOv7 well-suited for diverse object detection tasks, including applications like bottle defect detection.

E. Training

In the training setup for the object detection model, several key parameters have been configured to optimize performance. The input images are resized to a uniform dimension of 640x640 pixels, which balances the resolution with processing efficiency. The model undergoes training over 50 epochs, allowing it to iteratively learn and adjust to accurately recognize defects in bottles. A batch size of 16 is utilized, which determines the number of training samples processed before the model's internal parameters are updated. This size is chosen to manage memory usage while allowing sufficient gradient estimation per update. The learning rate is set at 0.01, guiding the pace at which the model weights adjust during training; this rate strikes a balance between speed and the risk of overshooting minimum in the loss landscape. Momentum, another crucial parameter, is set at 0.937 to help accelerate the model's convergence towards the global minimum and smooth out updates. Finally, a weight decay of 0.0005 is applied to regularize the model, helping to prevent overfitting by penalizing larger weights. These parameters together are designed to enhance the model's learning efficacy and generalization capabilities.

TABLE 1: TRAINING PARAMETERS

Image Size	640 x 640
Epochs	50
Batch Size	16
Learning Rate	0.01
Momentum	0.937
Weight Decay	0.0005

Spectrum of Engineering Sciences
ISSN (e) 3007-3138 **(p)** 3007-312X Volume 4, Issue 1, 2026

After completing the training of our model, we'll typically analyze graphs of training metrics to evaluate performance and understand the learning behavior over time.

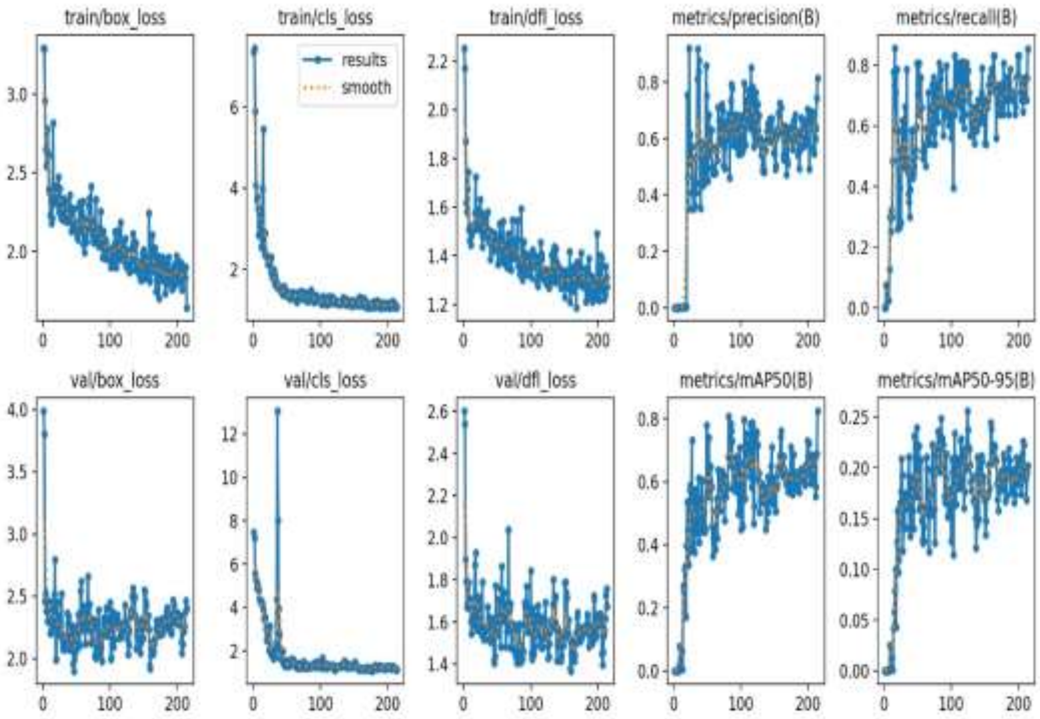

Figure 1: Training Matrices

RESULT AND DISCUSSION

A. Qualitative Results

The results you provided include precision, recall, and mean Average Precision (mAP) metrics, which are commonly used to evaluate the performance of object detection models like YOLOv7. Let me briefly explain each of these metrics:

1) P
recision
(P):

Precision is the ratio of true positive predictions to the total number of predicted positive instances. It measures the accuracy of the positive predictions made by the model. A higher precision indicates fewer false positives.

Precision = True Positives / (True Positives + False Positives)

In this case, the precision is 0.867, which means that approximately 86.7% of the predicted positive instances are true positive.

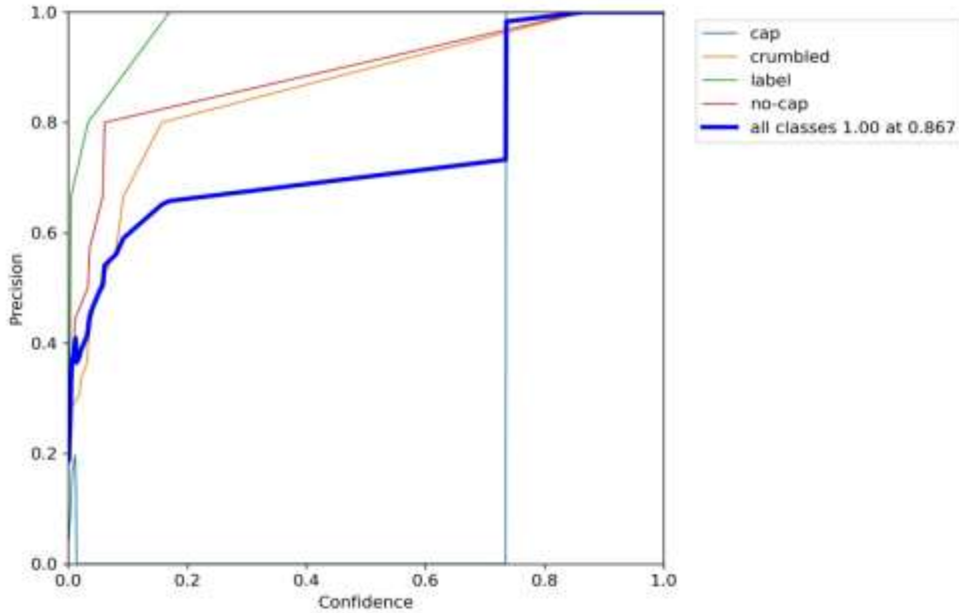

Figure 2: Precision

2) *R*
ecall (R):
Recall, also known as sensitivity or true positive rate, is the ratio of true positive predictions to the total number of actual positive instances. It measures the ability of the model to capture all positive instances.

Recall = True Positives / (True Positives + False Negatives)

In our case, the recall is 0.95, indicating that the model is able to capture 95% of the actual positive instances.

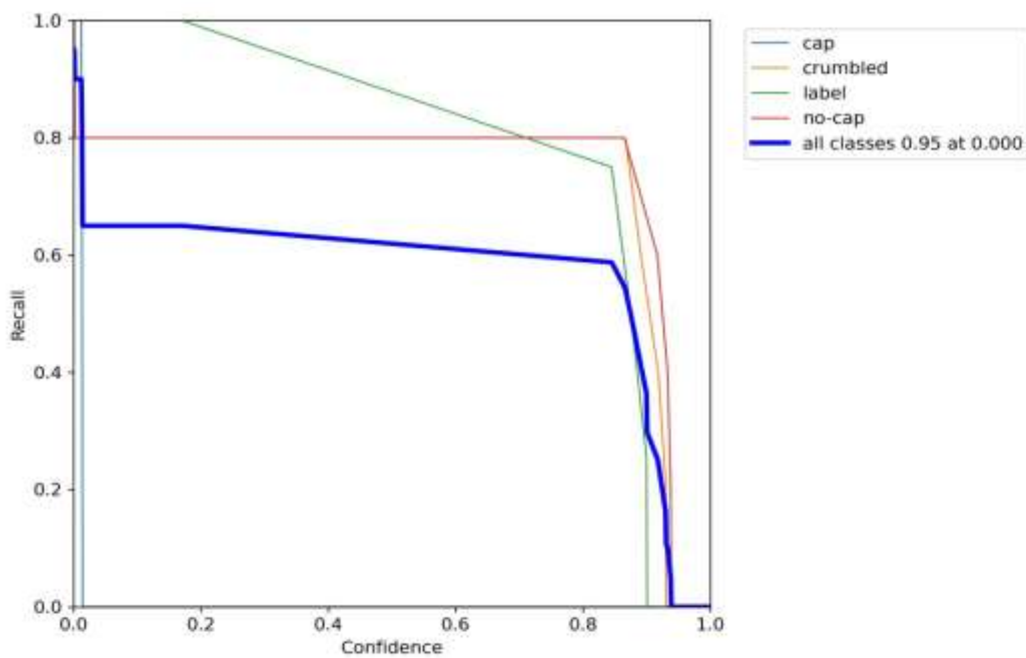

Figure 3: Recall 3)

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 1, 2026

Mean Average Precision (mAP):

Mean Average Precision (mAP) is a performance metric used in object detection tasks. It calculates the average precision for each class and then takes the mean. Average Precision (AP) is a measure of the area under the precision-recall curve.

In our case, the mean Average Precision is 0.83, suggesting an overall good performance across different classes.

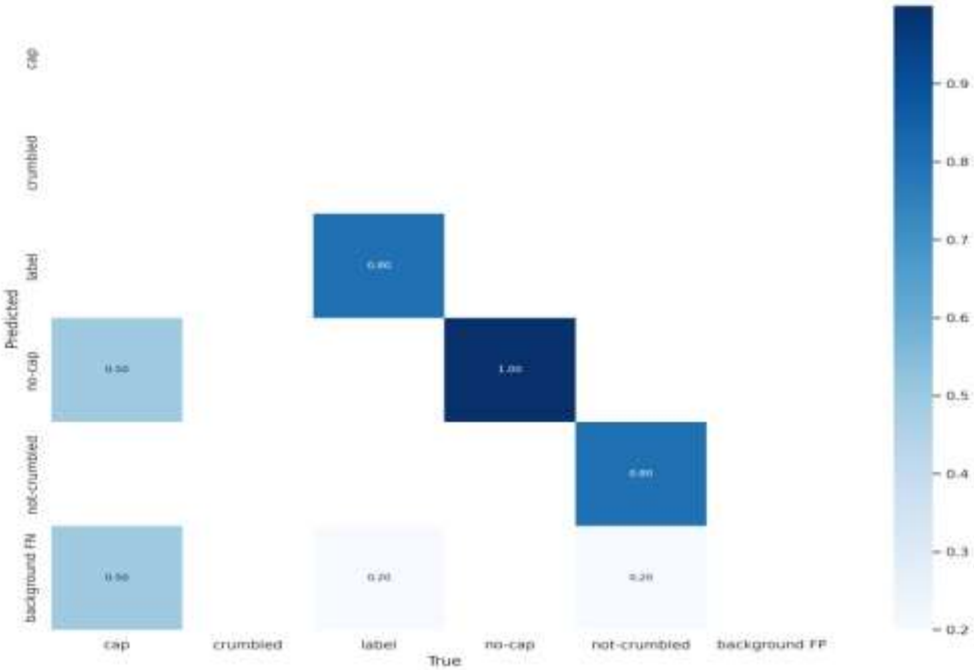

Figure 4: Mean Average Precision

CONCLUSION

In conclusion, this research successfully leveraged the YOLOv7 object detection framework, customized for bottle defect detection, achieving heightened accuracy and efficiency. COCO pre-trained weights expedited model convergence, enhancing its ability to identify missing caps and crumbled surfaces. The exploration of multi-scale training and optimization strategies, including PANet and spatial attention mechanisms, improved adaptability and spatial localization. Rigorous evaluation using precision, recall, and mAP metrics showcased the model's robust performance. With a focus on practical deployment considerations, this research provides a valuable solution for real-world manufacturing quality control. The findings contribute to computer vision applications and set the stage for future advancements in automated defect detection across industries.

REFERENCES

[1] X. Zhou, Y. Wang, C.Xiao, Q.Zhu, X.Lu, H. Zhang, J.Ge and H. Zhao, "Automated Visual Inspection of Glass Bottle Bottom With Saliency Detection and Template Matching," in IEEE Transactions on Instrumentation and Measurement, vol. 68, no. 11, pp. 4253-4267, Nov. 2019.

[2] Bin Huang, Sile Ma, Ping Wang, Huajie Wang, Jinfeng Yang, Xinyi Guo, Weidong Zhang, Huiquan Wang, "Research and implementation of machine vision technologies for empty bottle inspection systems", Engineering Science and Technology, an International Journal., vol. 21, issue. 1,2018, pp. 159-169.

[3] Bin Huang, Sile Ma*, Yufeng Lv, Hualong Zhang, Chunming Liu and Huajie Wang, "Research on Bottom Detection in Intelligent Empty Bottle Inspection System", TELKOMNIKA, vol. 11, no. 10, , pp. 5571 – 5578, Oct. 2013.
[4] Feng Duan, Yao-Nan Wang, Huan-Jun Liu and Wen Tan, "Empty bottle inspector based on machine vision," Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.04EX826), Shanghai, China, 2004, pp. 3845-3850 vol.6.
[5] S. Ma, B. Huang, H. Wang and J. Guo, "Algorithm research on location of bottle mouth and bottom in intelligent empty bottle inspection system," 2009 IEEE International Conference on Automation and Logistics, Shenyang, 2009, pp. 819-824.
[6] Cheng, J.C.; Wang, M. Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques. Autom. Constr. 2018, 95, 155–171. [Google Scholar] [CrossRef]
[7] Li, R.; Yuan, Y.; Zhang, W.; Yuan, Y. Unified vision-based methodology for simultaneous concrete defect detection and geolocalization. Comput.-Aided Civ. Infrastruct. Eng. 2018, 33, 527–544. [Google Scholar] [CrossRef]
[8] Zhang, C.; Chang, C.C.; Jamshidi, M. Bridge damage detection using a single-stage detector and field inspection images. arXiv 2018, arXiv:1812.10590. [Google Scholar]
[9] Yin 2018, X.; Chen, Y.; Bouferguene, A.; Zaman, H.; Al-Hussein, M.; Kurach, L. A deep learning-based framework for an automated defect detection system for sewer
pipes. Autom. Constr. 2020, 109, 102967. [Google Scholar] [CrossRef]
[10] Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, realtime object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 779–788. [Google Scholar]
[11] Wang, J.; Fu, P.; Gao, R.X. Machine vision intelligence for product defect inspection based on deep learning and Hough transform. J. Manuf. Syst. 2019, 51, 52–60. [Google Scholar] [CrossRef]
[12] Li, F.; Hang, Z.; Yu, G.; Wei, G.; Xinyu, C. The method for glass bottle defects detecting based on machine vision. In Proceedings of the 2017 29th Chinese Control
Additionally, Decision Conference (CCDC), Chongqing, China, 28–30 May 2017; pp. 7618– 7621. [Google Scholar] [CrossRef]
[13] Fu, L.; Zhang, S.; Gong, Y.; Huang, Q. Medicine Glass Bottle Defect Detection Based on Machine Vision. In Proceedings of the 2019 Chinese Control Additionally,
Decision Conference (CCDC), Nanchang, China, 3–5 June 2019; pp. 5681–5685. [Google Scholar] [CrossRef]
[14] Kumchoo, W.; Chiracharit, W. Detection of loose cap and safety ring for pharmaceutical glass bottles. In Proceedings of the 2018 International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications
Engineering (ECTI-NCON), Chiang Rai, Thailand, 25–28 February 2018; pp. 125–130. [Google Scholar] [CrossRef]
[15] Xie, H.; Lu, F.; Ouyang, G.; Shang, X.; Zhao, Z. A rapid inspection method for encapsulating quality of PET bottles based on machine vision. In Proceedings of the 2017 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu,
China, 13–16 December 2017; pp. 2025–2028. [Google Scholar] [CrossRef].