Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

AI-Driven Pill Blister Pack Defect Detection using Enhanced YOLOv8 and Hybrid CNN Architecture

Dr. C. Sridevi
Associate Professor with the Department of Electronics Engineering, MIT Campus, Anna University, Chrompet, Chennai – 600044

Nithyakamini P G, Ifrath Fathima M, Kirthana K
Department of Electronics Engineering, MIT Campus, Anna University, Chrompet, Chennai – 600044, India

ABSTRACT Ensuring defect-free pharmaceutical blister packaging is critical for patient safety, regulatory compliance, and manufacturing efficiency. Conventional manual inspection techniques are labor-intensive, inconsistent, and unsuitable for high-speed production lines. Although recent advances in deep learning have enabled automated visual inspection, detecting subtle blister defects such as cracks, stains, missing tablets, and foreign particle contamination remains challenging under real-world conditions. This paper presents an AI-driven real-time defect detection framework for pharmaceutical blister packs using enhanced YOLOv8 architectures and a hybrid convolutional neural network (CNN) approach. The proposed system evaluates three models: baseline YOLOv8, CBS-YOLOv8 incorporating Coordinate Attention, BiFPN, and SimSPPF modules, and a hybrid YOLOv8 + DenseNet121 architecture for refined defect classification. Experimental results demonstrate that CBS-YOLOv8 achieves the highest detection accuracy with an mAP@0.5 of 96.75%, while the hybrid model improves classification precision for visually similar defect types. The system is further integrated with an OpenCV-based live video inspection pipeline, enabling real-time defect detection and visualization. The proposed framework provides a scalable, accurate, and industry-ready solution aligned with Pharma 4.0 quality assurance requirements.

INDEX TERMS - Pharmaceutical blister pack inspection, deep learning–based defect detection, YOLOv8 object detection, attention-enhanced feature fusion, hybrid CNN architectures, real-time visual inspection.

I. INTRODUCTION

Pharmaceutical blister packaging plays a vital role in ensuring dosage accuracy, hygiene, and protection of solid oral medications. Even minor defects such as missing tablets, cracks, discoloration, or foreign particle contamination can compromise drug safety and lead to regulatory violations and costly product recalls. In high-speed pharmaceutical manufacturing environments, manual visual inspection is still widely practiced; however, it is slow, subjective, and prone to human fatigue and inconsistency.

Traditional machine vision systems based on rule-based image processing and thresholding techniques offer limited robustness. These methods are highly sensitive to lighting variations, reflections from blister materials, and differences in tablet appearance. As pharmaceutical production lines continue to scale, there is a growing need for intelligent, automated inspection systems capable of real-time operation with high accuracy and reliability.

Recent advancements in deep learning and computer vision have significantly improved object detection and classification capabilities. Single-stage detectors such as You Only Look Once (YOLO) provide high-speed inference suitable for real-time industrial applications. However, baseline YOLO models often struggle with fine-grained defects, particularly small cracks, stains, and visually similar defect categories.

To address these challenges, this paper proposes a multi-model AI-driven blister pack defect detection framework. The system evaluates a baseline YOLOv8 detector, an enhanced CBS-YOLOv8 architecture incorporating attention and feature fusion mechanisms, and a hybrid YOLOv8 + DenseNet121 model for refined defect classification. Additionally, real-time inspection is achieved through OpenCV-based live video feed integration, simulating industrial conveyor belt environments.

The key contributions of this work are summarized as follows:

- Development of an automated deep learning-based defect detection system for pharmaceutical blister packs.
- Integration of Coordinate Attention, BiFPN, and SimSPPF modules to enhance YOLOv8 detection performance.
- Implementation of a hybrid two-stage YOLOv8 + DenseNet121 architecture for fine-grained defect classification.
- Comparative performance evaluation of YOLOv8, CBS-YOLOv8, and hybrid architectures.
- Real-time deployment using OpenCV for continuous industrial inspection.

The remainder of this paper is organized as follows. Section II reviews related work in blister pack inspection and industrial defect detection. Section III describes the proposed methodology. Section IV presents experimental results and analysis. Section V discusses observed challenges and limitations. Section VI concludes the paper and outlines the future work.

II. RELATED WORK

Automated inspection of pharmaceutical packaging has been an active research area due to its direct impact on product quality and patient safety. Early approaches relied on handcrafted image features such as edge detection, color histograms, and morphological operations combined with traditional classifiers for blister pack inspection [4], [6]. While effective under controlled conditions, these systems lacked robustness to lighting variations, surface reflections, and product variability commonly observed in industrial environments.

With the emergence of deep learning, convolutional neural networks (CNNs) have become the dominant approach for visual defect detection. Several studies have demonstrated the effectiveness of CNN-based classifiers for detecting broken or missing tablets in blister packs by learning discriminative features directly from data [5], [13]. However, many of these approaches operate on static images and lack real-time capability required for high-speed production lines.

Recent research has focused on object detection frameworks such as YOLO and Faster R-CNN for industrial inspection. YOLO-based models enable simultaneous localization and classification, making them suitable for real-time pharmaceutical inspection tasks [2], [6]. Improved variants incorporating attention mechanisms and feature fusion have shown enhanced performance for detecting small and subtle defects [8], [11].

Despite these advancements, most existing works either focus solely on detection accuracy or lack comparative evaluation of enhanced architectures. Furthermore, limited research integrates hybrid detection and classification pipelines with real-time deployment. The proposed work addresses these gaps by evaluating multiple YOLOv8-based architectures, incorporating attention-enhanced detection, and demonstrating real-time industrial applicability.

III. PROPOSED METHODOLOGY

The proposed system is designed as an end-to-end AI-driven framework for automated blister pack defect detection. The overall pipeline consists of image acquisition, preprocessing, defect detection, defect classification, and real-time visualization.

I. A. Overall System Architecture

Input images are obtained either from a pre-collected dataset or a live camera feed. After preprocessing, the images are passed through one of the detection architectures: YOLOv8, CBS-YOLOv8, or hybrid YOLOv8 + DenseNet121. Detected defects are classified and visualized using bounding boxes and confidence scores. For real-time operation, the pipeline is integrated with OpenCV for continuous frame processing. The overall architecture of the proposed blister pack defect detection system is illustrated in Fig. 1, highlighting the sequential flow from image acquisition to real-time defect visualization.

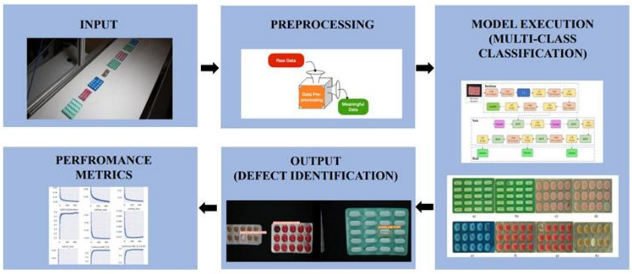

Fig. 1. Overall architecture of the proposed AI-driven pill blister pack defect detection system.

II. B. Defect Detection Using YOLOv8

The baseline YOLOv8 model serves as the reference detector. It employs an anchor-free detection head with efficient feature extraction and fast inference. YOLOv8 demonstrates reliable performance for detecting obvious defects such as missing and broken tablets but shows reduced sensitivity for subtle surface-level defects. The baseline YOLOv8 architecture employed for defect detection is shown in Fig. 2, consisting of a backbone, neck, and anchor-free detection head.

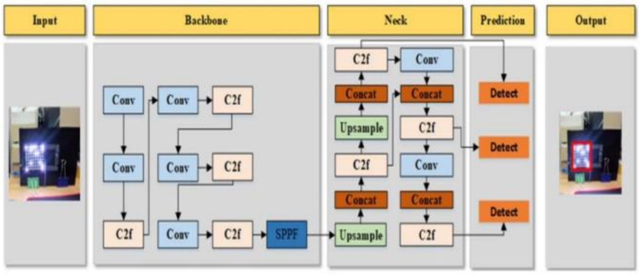

Fig. 2. Architecture of the baseline YOLOv8 object detection model.

III. C. CBS-YOLOv8 Architecture

To enhance detection accuracy, CBS-YOLOv8 integrates Coordinate Attention (CA), Bi-Directional Feature Pyramid Network (BiFPN), and Simplified Spatial Pyramid Pooling Fast (SimSPPF) modules. Coordinate Attention improves spatial focus on defect regions, BiFPN strengthens multi-scale feature fusion, and SimSPPF enhances contextual representation with minimal computational overhead. This architecture significantly improves detection of fine-grained defects. To enhance detection accuracy for subtle defects, the CBS-YOLOv8 architecture shown in Fig. 3 incorporates attention and multi-scale feature fusion mechanisms.

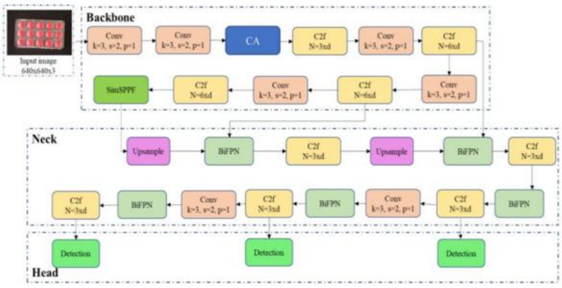

Fig. 3. CBS-YOLOv8 architecture integrating Coordinate Attention, BiFPN, and SimSPPF modules.

IV. D. Hybrid YOLOv8 + DenseNet121 Model

The hybrid model employs a two-stage pipeline where YOLOv8 performs defect localization and DenseNet121 classifies cropped regions of interest (ROI). DenseNet121's dense connectivity enables detailed texture and pattern learning, improving differentiation between visually similar defects such as stains and colour mismatch. As illustrated in Fig. 4, YOLOv8 performs defect localization, while DenseNet121 refines defect classification at the ROI level.

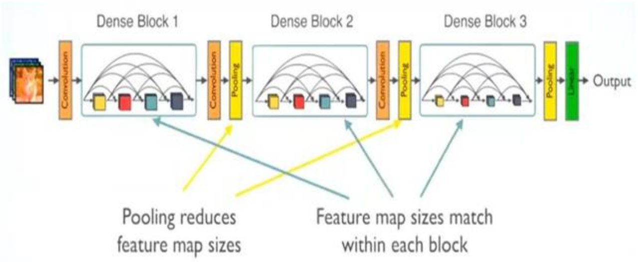

Fig. 4. Hybrid YOLOv8 + DenseNet121 architecture for two-stage defect detection and classification.

V. E. Real-Time Detection Using OpenCV

An OpenCV-based live video feed module captures frames from a camera and processes them in real time. Detected defects are displayed with bounding boxes and labels, simulating industrial conveyor belt inspection. To validate the practical feasibility of the proposed system, the trained models were integrated with an OpenCV-based live video pipeline. The system successfully detects and classifies blister pack defects in real time with bounding boxes and confidence scores, as illustrated in Fig. 5.

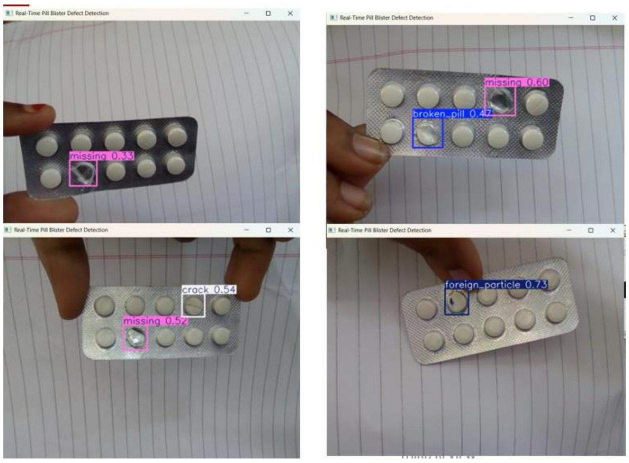

Fig. 5. Real-time pill blister pack defect detection using an OpenCV-based live video inspection pipeline.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

VI. A. Experimental Setup

The system was trained and evaluated using a manually annotated dataset consisting of approximately 5,500 blister pack images covering five defect categories: broken, cracked, missing, stained, and foreign particle. The dataset was split into training, validation, and testing sets.

VII. B. Evaluation Metrics

The performance of the proposed defect detection models was quantitatively evaluated using standard object detection metrics, including Precision, Recall, and mean Average Precision (mAP@0.5). Precision and Recall were computed for each defect class to assess classification reliability and detection completeness, respectively, while mAP@0.5 was used as an overall indicator of detection accuracy at an Intersection over Union (IoU) threshold of 0.5. In addition to accuracy metrics, real-time performance was evaluated using Frames Per Second (FPS), measured as the average inference speed during continuous frame processing. The computed metrics were subsequently used for comparative analysis across different model architectures, as presented in the following subsection.

VIII. C. Comparative Results

This section presents a comparative evaluation of the proposed defect detection architectures to assess their effectiveness for pharmaceutical blister pack inspection. The baseline YOLOv8, CBS-YOLOv8, and hybrid YOLOv8 + DenseNet121 models are compared in terms of detection accuracy, inference speed, and class-level discrimination. Detection performance in terms of Precision, Recall, and mean Average Precision (mAP@0.5) is summarized in Table I, while real-time inference capability is evaluated using Frames Per Second (FPS), as reported in Table II. Together, these metrics provide a comprehensive assessment of both detection reliability and real-time feasibility. CBS-YOLOv8 achieved the highest detection accuracy, while the hybrid model improved classification reliability for subtle defects.

TABLE I: PERFORMANCE COMPARISON OF DEFECT DETECTION MODELS

Model	Precision	Recall	mAP@0.5
YOLOv8	88.48%	88.37%	90.25%
CBS-YOLOv8	94.48%	92.13%	96.75%
YOLOv8+ DenseNet121	92.21%	90.35%	94.53%

TABLE II: INFERENCE SPEED COMPARISON OF DEFECT DETECTION MODELS

Model	Inference Speed (FPS)
YOLOv8	42.6
CBS-YOLOv8	36.8
YOLOv8 + DenseNet121	28.4

In addition to detection accuracy, inference speed is a critical factor for real-time industrial deployment. Table II presents the average inference speed of the evaluated models measured in frames per second (FPS). The baseline YOLOv8 model achieves the highest inference speed due to its single-stage architecture. CBS-YOLOv8 incurs a marginal reduction in FPS owing to the integration of attention and feature fusion modules, while still maintaining real-time performance. The hybrid YOLOv8 + DenseNet121 model exhibits the lowest FPS as a result of the additional second-stage CNN classification.

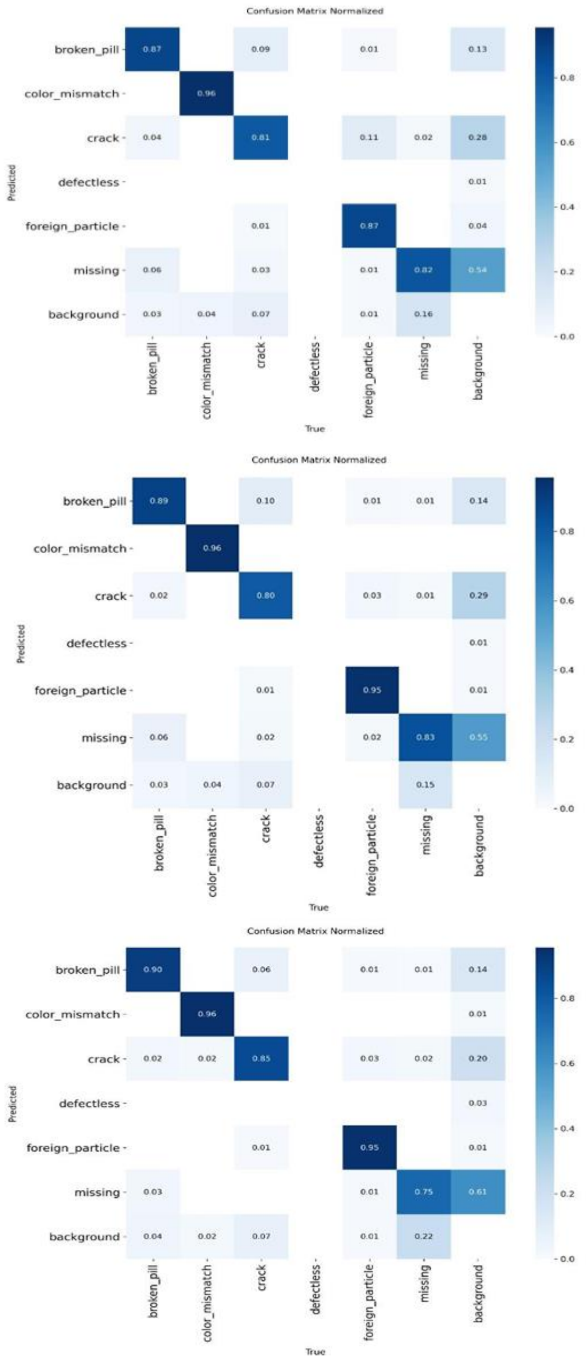

Fig. 6. Confusion matrices comparing YOLOv8, CBS-YOLOv8, and hybrid models.

The confusion matrices shown in Fig. 6 provide further insight into the class-level performance of the evaluated models. The baseline YOLOv8 model exhibits higher misclassification rates for visually similar defect categories, particularly stains and colour mismatch. In contrast, CBS-YOLOv8 demonstrates improved class separation, with a noticeable reduction in false positives and false negatives across all defect classes. The hybrid YOLOv8 + DenseNet121 model further refines classification consistency for subtle defects; however, this improvement comes at the cost of reduced inference speed, as reflected in the FPS comparison.

Overall, the comparative analysis indicates that CBS-YOLOv8 offers the most balanced trade-off between detection accuracy, class-level reliability, and real-time performance, making it well-suited for deployment in high-speed pharmaceutical blister pack inspection systems.

IX. D. Qualitative Results

In addition to quantitative evaluation, qualitative analysis was conducted to assess the robustness of the proposed defect detection models under challenging real-world conditions. This analysis focuses on model performance under variations in illumination and blister pack orientation, which are commonly encountered in industrial environments. Figure 7 illustrates detection results obtained under normal lighting conditions with tilted blister orientations. The model successfully localizes the blister pack and accurately identifies multiple defect types, demonstrating reliable performance despite perspective variations.

Fig. 7. Qualitative detection results under normal lighting conditions and tilted blister orientation.

Figure 8 presents detection outputs captured under low-light conditions, where reduced illumination and background noise are present. Even in these challenging scenarios, the model consistently detects stained and broken pills with high confidence scores.

Fig. 8. Qualitative detection results under low-light conditions demonstrating robust defect detection.

The consistent performance observed in both Fig. 7 and Fig. 8 highlights the robustness of the proposed architecture against illumination changes, surface reflections, and geometric distortions. The integration of attention-enhanced feature extraction and multi-scale learning enables the model to focus on defect-relevant regions even when visual cues are weak or partially degraded.

Overall, the qualitative results confirm that the proposed defect detection framework generalizes well beyond controlled dataset conditions and is suitable for real-time deployment in practical pharmaceutical manufacturing environments.

X. D. Qualitative Multi-Defect Detection Results

In addition to the robustness analysis under varying illumination and orientation, qualitative evaluation was performed to demonstrate the capability of the proposed model to detect multiple defect types within a single blister pack. Figure 9 illustrates a representative detection output where the model simultaneously identifies a defective blister pack along with multiple pill-level defects, including stained and broken pills.

The model accurately localizes the entire blister pack region and assigns a high confidence score to the defective pack classification. Furthermore, distinct bounding boxes are generated for individual pill defects, each labelled with its corresponding defect class and confidence value. This demonstrates the model's ability to perform multi-level detection, combining global blister-level assessment with fine-grained pill-level defect localization.

The results highlight the effectiveness of the attention-enhanced feature extraction and multi-scale learning mechanisms in handling complex real-world inspection scenarios, where multiple defects may coexist within a single package. Such capability is critical for practical pharmaceutical quality assurance systems, as it enables comprehensive inspection without requiring multiple passes or separate models.

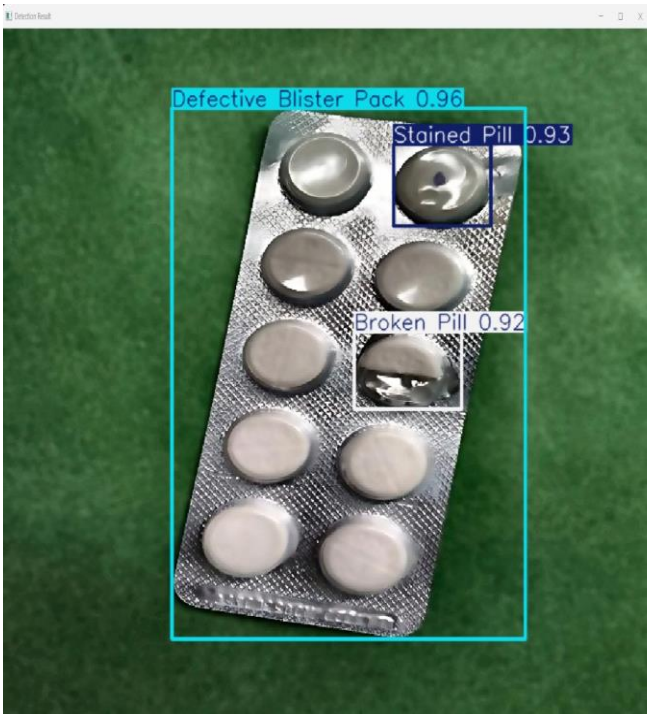

Fig. 9. Qualitative example of multi-defect detection within a single blister pack, showing simultaneous identification of defective blister pack, stained pill, and broken pill with corresponding confidence scores.

V. DISCUSSION AND COMMON CHALLENGES

The experimental results demonstrate that deep learning-based inspection significantly improves the accuracy and reliability of pharmaceutical blister pack defect detection compared to conventional methods. However, several practical challenges were observed during model training, evaluation, and real-time deployment, which are discussed in this section.

One of the primary challenges encountered was visual similarity among defect classes. Defects such as stains and colour mismatch often exhibit comparable texture and intensity patterns, especially under reflective blister surfaces. Although the CBS-YOLOv8 model improved spatial focus through Coordinate Attention and multi-scale feature fusion, minor misclassifications were still observed in borderline cases. The hybrid YOLOv8 + DenseNet121 architecture helped mitigate this issue by performing refined ROI-level classification, enabling better discrimination between visually similar defect types.

Lighting variation and surface reflections posed another significant challenge. Blister packs are typically made of transparent or semi-reflective materials, which can produce glare under industrial lighting conditions. These reflections occasionally caused false positives, particularly for crack-like patterns. While data augmentation and CBS feature fusion improved robustness, complete elimination of reflection-induced errors remains challenging without controlled illumination or polarized lighting setups.

Small and subtle defects, such as micro-cracks and tiny foreign particles, also affected detection performance. The baseline YOLOv8 model occasionally missed these defects due to limited spatial attention. In contrast, CBS-YOLOv8 demonstrated superior performance by enhancing feature representation at multiple scales. This confirms the importance of attention mechanisms and hierarchical feature fusion in industrial defect detection tasks.

From a deployment perspective, the trade-off between accuracy and inference speed was observed. The hybrid model introduced additional computational overhead due to the second-stage CNN classification, leading to a slight reduction in FPS compared to single-stage YOLO-based models. However, this trade-off is acceptable in pharmaceutical quality assurance applications, where detection accuracy and reliability are prioritized over marginal speed improvements.

Finally, dataset limitations influenced overall performance. Although more than 5,500 images were manually annotated, certain defect classes exhibited lower variability in terms of lighting, orientation, and packaging type. Expanding the dataset with more diverse real-world samples is expected to further improve generalization and robustness.

VI. CONCLUSION AND FUTURE SCOPE

This paper presented an AI-driven real-time defect detection system for pharmaceutical blister packs using enhanced YOLOv8 architectures and a hybrid deep learning approach. The proposed framework systematically evaluated three models - baseline YOLOv8, CBS-YOLOv8, and a hybrid YOLOv8 + DenseNet121 architecture—to address the challenges of detecting subtle, small-scale, and visually similar defects in high-speed pharmaceutical production environments.

Experimental results demonstrated that CBS-YOLOv8 achieved the highest overall detection accuracy, with an mAP@0.5 of 96.75%, owing to the integration of Coordinate Attention, BiFPN, and SimSPPF modules. The hybrid YOLOv8 + DenseNet121 model further enhanced defect classification precision, particularly for visually ambiguous defects such as stains and colour mismatch. Real-time testing using an OpenCV-based live video pipeline confirmed the feasibility of deploying the proposed system for continuous industrial inspection.

The study highlights the effectiveness of combining attention-enhanced object detection with CNN-based fine-grained classification for pharmaceutical quality assurance. By reducing reliance on manual inspection and minimizing subjectivity, the proposed system aligns with the principles of Pharma 4.0 and intelligent manufacturing.

Future work will focus on several key extensions. First, the system can be deployed on Edge AI platforms such as NVIDIA Jetson devices to enable low-latency, on-site inspection. Second, expanding the dataset to include a wider variety of blister materials, tablet shapes, and lighting conditions will improve model generalization. Third, integrating transformer-based attention mechanisms or next-generation YOLO variants may further enhance detection of extremely subtle defects. Finally, coupling the detection system with automated rejection mechanisms and production analytics dashboards can enable fully autonomous quality control on pharmaceutical production lines.

REFERENCES

[1] [1] Base paper A. Vijayakumar, V. Subramaniyaswamy, J. A. S. Koilraj, M. Rajappa, K. Kotecha, and A. Kulkarni,“Real-time visual intelligence for defect detection in pharmaceutical packaging,” Scientific Reports, vol. 14, no. 18811, 2024, doi: 10.1038/s41598-024-69701-z.

[2] [2] R. Patgiri, V. Ajantha, S. Bhuvaneswari, and V. Subramaniyaswamy, “Intelligent defect detection system in pharmaceutical blisters using YOLOv7,” in Proc. 2nd Int. Conf. Emerging Trends in Information Technology and Engineering (ICETITE), Thanjavur, India, 2024, pp. 1–7, doi: 10.1109/IC-ETITE58242.2024.10493735.

[3] [3] S. J. Kim and D. S. Cho, “Medical-pills detection using YOLOv11: A proof-of-concept study for pharmaceutical automation,” Clinical Case Reports and Studies, vol. 10, no. 2, 2025, doi: 10.59657/2837-2565.brs.25.252.

[4] [4] C. E. Pardo, L. F. Sosa, E. A. Gutierrez, and F. R. Jiménez, “Classification system for blister pack of pills,” in Proc. IEEE Andean Council Int. Conf. (ANDESCON), Tunja, Colombia, 2014, pp. 1–6, doi: 10.1109/ANDESCON.2014.49.

[5] [5] L. E. Sylvester, P. Prea, S. Gobee, and V. Durairajah, “Vision-based pill blister package inspection system using CNN,” in Proc. 13th Int. Conf. Biomedical Engineering and Technology (ICBET), Tokyo, Japan, 2023, pp. 1–6, doi: 10.1145/3620679.3620694.

[6] [6] T. T. H. Vu, D. L. Pham, and T. W. Chang, “A YOLO-based real-time packaging defect detection system,” Procedia Computer Science, vol. 217, pp. 886–894, 2023, doi: 10.1016/j.procs.2022.12.285.

[7] [7] E. Dehaerne, B. Dey, H. Esfandiar, L. Verstraete, H. S. Suh, S. Halder, and S. De Gendt, “YOLOv8 for defect inspection of hexagonal directed self-assembly patterns: A data-centric approach,” arXiv preprint, arXiv:2307.15516, 2023.

[8] [8] Z. Wang, K. Liao, Y. Xu, and X. Gao, “CBAM-YOLOv8: An improved YOLOv8 model for object detection using attention mechanisms,” IEEE Access, vol. 12, pp. 13240–13251, 2024.

[9] [9] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2017, pp. 4700–4708.

[10] [10] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal loss for dense object detection,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 42, no. 2, pp. 318–327, Feb. 2020.

[11] [11] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, “Path aggregation network for instance segmentation,” in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), 2018, pp. 8759–8768.

[12] [12] M. Tan, R. Pang, and Q. V. Le, “EfficientDet: Scalable and efficient object detection,” in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), 2020, pp. 10781–10790.

[13] [13] P. Voigtlaender, L. Leal-Taixé, and B. Leibe, “Deep learning for visual inspection in manufacturing: A comprehensive survey,” IEEE Trans. Industrial Informatics, vol. 18, no. 6, pp. 3605–3620, Jun. 2022.

[14] [14] J. Kang, Y. Cen, K. Wang, and Y. Liu, “CFIS-YOLO: A lightweight multi-scale fusion network for edge-deployable defect detection,” arXiv preprint, arXiv:2504.11305, 2025.