UNIVERSITY OF WEST ATTICA

eISSN:3136-7658

Applied Manufacturing Intelligence Journal

Vol. 5, No. 3, July, 2026



TabM-CatBoost Model for Pharmaceutical Blister Packaging Defect Segmentation Based on Backlight Inspection Images

Muhammad Ahmad[1,*], Firdaus Yaacob[1]

[1] *Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, Universiti Malaya, Kuala Lumpur, 50603, Malaysia*

Received 19 April 2026; Revised 31 July 2026; Accepted 05 August 2026; Published 09 August 2026

Abstract

Inspection of pharmaceutical blister packaging must identify small but consequential defects, including missing tablets, broken cavities, bubbles, dents, foreign objects, poor sealing and local contamination, while still operating under high-throughput production conditions. Backlight imaging improves the visibility of transparent or semi-transparent defects, but the same imaging setup can also amplify foil reflection, plastic-thickness variation, tablet-colour differences, camera-exposure drift and conveyor vibration. This paper proposes a TabM-CatBoost model for defect segmentation in backlight inspection images. For each blister image, a cavity-level window is aligned with the package template, and candidate masks, illumination descriptors, region statistics, geometry descriptors, source-quality flags and process attributes are converted into structured evidence. TabM then learns feature interactions among image texture, local contrast, cavity position, boundary geometry and production-state variables, while CatBoost estimates defect class, boundary reliability, defect-risk level, confidence and review priority. The simulated evaluation dataset contains 18,960 image windows covering normal packaging, missing tablets, foreign particles, weak sealing flaws, local dents, air bubbles and mixed defects. Training, validation and testing use 70%, 15% and 15% of the records, respectively, at package-lot level. The proposed model reaches a mean IoU of 0.887, a boundary F1 of 0.861 and a high-risk recall of 0.934; compared with the best baseline, it improves mean IoU by 4.8% and reduces false review alarms by 13.6%. These results indicate that backlight-enhanced mask evidence and interpretable tabular fusion can support stable industrial review under weak defect contrast, scarce abnormal samples and strict packaging quality-control requirements.

Keywords: Backlight Inspection; Pharmaceutical Blister Packaging; Defect Segmentation; TabM; CatBoost; Quality Control

1. Introduction

Pharmaceutical blister packaging quality control has become a key part of drug manufacturing safety because package integrity directly affects dosage stability, contamination prevention and downstream traceability [1]. Backlight imaging has been widely used for transparent and semi-transparent industrial inspection because transmitted light can expose weak cavities, particles and sealing discontinuities that are difficult to observe under reflected illumination [2]. Recent industrial visual inspection studies show that small defects often require both pixel-level localization and process-level interpretation, since a small foreign particle and a harmless reflection can have similar local intensity patterns [3]. Machine-learning inspection systems have also moved from single-image classification to segmentation and review-oriented decision support, which makes structured evidence increasingly important [4]. In pharmaceutical blister inspection, these tendencies are amplified by high throughput, repeated cavity geometry and strict tolerance for

Corresponding author: muhammad.ahm08@um.edu.my

https://doi.org/10.68406/amij.2026.vol5iss3np4:52-64

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

missed defects. A method that only reports whether a package image is normal cannot provide enough evidence for batch review, root-cause analysis or line adjustment.

Backlight images provide useful transmitted-light evidence, but they also introduce inspection ambiguity. Missing tablets, cracked pills, local bubbles, seal-line leakage and foreign particles become more visible than in ordinary surface images, while glare, film-thickness variation and cavity-edge shadows may also become stronger [5]. Traditional thresholding and morphology can detect only high-contrast defects, and their decisions change when the grayscale range of the plastic shell or tablet varies [6]. Deep segmentation networks can learn boundary patterns, but labelled pharmaceutical defects remain scarce because genuine line defects are rare and expensive to collect [7]. A hybrid approach is therefore needed to combine candidate masks, cavity geometry, illumination statistics and machine-state variables in a form that quality engineers can audit.

This paper proposes a TabM-CatBoost defect-segmentation model for backlight blister inspection and treats the task as structured review rather than simple image classification. The model first aligns each image to the repeated blister template, extracts candidate masks and converts local image evidence into cavity-level descriptors. TabM learns nonlinear interactions among mask geometry, local contrast, cavity position, image-quality tags and production-state variables. CatBoost then combines the TabM representation with the original interpretable variables to score defect type, boundary reliability, risk level and review priority. This design allows the alarm reason to be traced to missing tablet-shaped evidence, abnormal foreign-particle contrast, seal-line discontinuity or low image reliability.

Therefore, this paper focuses on weak-defect segmentation for pharmaceutical blister packaging under backlight inspection while considering visual contrast, cavity position and production conditions together. Section 2 reviews backlight imaging, blister-defect characteristics, hybrid image-tabular learning and review-oriented quality-control models. Section 3 describes cavity-window construction, candidate-mask encoding, TabM-CatBoost model design and defect-risk output. Section 4 evaluates dataset construction, segmentation accuracy, robustness, evidence interpretation and ablation results. Section 5 summarizes the quantitative findings, engineering value, limitations and future factory-validation needs.

2. Related Work

2.1. Backlight Imaging for Transparent Packaging Defect Recognition

Backlight inspection can be used to reveal the interior or occlusion of an object by transmitting light. Blister packaging is an all-optical structure that consists of a plastic cavity, the surface and seal area of the tablet, etc., and thus, defects can take the form of local attenuation, edge deformation, cavity misalignment or abnormal transparency. Research on industrial visual inspection shows that transmitted-light settings can improve the detection performance of thin contamination and small holes [8], particularly when the defect thickness varies from the surrounding material. Therefore, the image evidence of the pharmaceutical packaging should have local intensity and geometric consistency in the repeated cavities.

Defect types of blister packaging are diverse. A missing tablet alters the occupied region at the cavity level; a broken tablet produces edge fragments and an abnormal texture; a foreign particle changes the local contrast; and a seal flaw affects the continuity of the heat-sealed area. Research on packaging inspection has indicated that these defects are not a single type of abnormality [9]. The segmentation model should retain the information of defect type and location so that quality engineers can decide whether to halt production at the production line, isolate the batch, or carry out an inspection of a sealing, feeding or cleaning device.

Interference due to image acquisition exacerbates the problem. Backlight brightness drift, plastic-film thickness variation, tablet colour change and camera exposure instability can change the intensity of the image without any actual defects. Visual Quality-Control studies have shown that when the illumination and object transparency are different, segmentation performance decreases [10]. Therefore, the model should not directly associate brightness with risk, but rather compare the local evidence to the baseline of a cavity, line state and neighbouring cavity structure.

Another problem is the lack of genuine defects. A production line produces a large number of normal images, and it is not feasible to prepare for all kinds of planned defect simulations. Few-shot industrial inspection research has proposed that structured features and interpretable ensembles are still effective when there are few defect examples [11]. Therefore, before the final assessment of scores, explicit candidate masks are generated in this study.

2.2. Hybrid Image-Tabular Models for Quality-Control Decision Support

Recently, many modern industrial inspections have added structured process data to images. Pure image models can provide mask candidates, but they do not consider the production environment; for example, the position of the mould well, batch variations, sealing temperature, conveyor vibration and inspection confidence are all external factors. Studies in hybrid learning have shown that tabular descriptors can improve the quality-control reliability of repeated

geometric template inspection of inspected objects [12]. Blister packages use a template to organise the cavities in a grid and provide statistics at the cavity level.

Tabular deep models are to learn non-linear relationships among structured variables. The area of the mask, defect eccentricity, boundary contrast, cavity position, local illumination deviation and source reliability may be non-additively related in this study. A small dark spot near the seal interface may be harmless; however, the same spot in the middle of the cavity is considered foreign material. Research on tabular neural networks has shown that learning feature interactions is beneficial when the individual variables have different meanings in different situations [13]. TabM is used as the first structured fusion module for this purpose.

CatBoost can be used for manufacturing inspection to deal with both numerical and categorical features, has stable boosting, and also offers feature importance. Zhai et al. [14] reported that CatBoost-based quality prediction can maintain stable performance when structured features are reliable but abnormal samples are limited. CatBoost can combine TabM representations with the original interpretable variables for blister inspection to maintain the auditability of the final review decision. The model can therefore determine whether the defect risk is mainly caused by mask size, boundary confidence, cavity position or image-quality tags.

A decision-support system needs to be built to make production decisions based on the segmentation results. Depending on the type of defect, the extent and confidence, a weak defect mask may be accepted, re-examined or rejected. Maitra et al. [15] showed that explainable industrial AI should provide risks and supporting evidence rather than only a binary label for inspection systems. Therefore, the proposed TabM-CatBoost framework is designed to output the segmentation mask, defect class, confidence, dominant evidence and review priority in a single workflow.

3. Methodology

3.1. Backlight Defect Image Alignment and Mask Encoding

First, a cavity-level inspection window is built from each backlight image. The package image is aligned with the blister template, and a local coordinate system is created for the cavity, seal area and edge zone. Figure 1 is the acquisition and alignment workflow for the backlight image, cavity map, candidate mask, structured descriptors and review input. The aligned window is used instead of the raw full image because defect evidence is meaningful only when its position is interpreted relative to repeated packaging geometry. The process also standardizes cavity-row orientation so that the same defect pattern corresponds to comparable coordinates across packages.

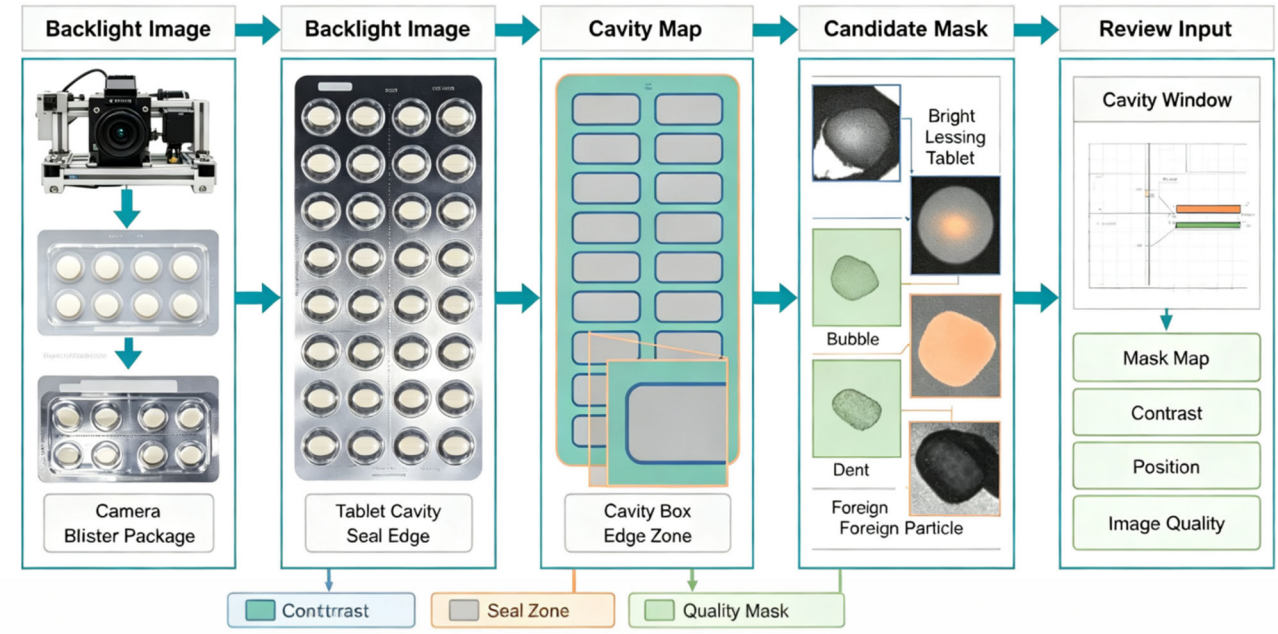


Figure 1. Backlight image acquisition and blister-defect evidence alignment workflow.

$$I_i = \mathrm{crop}(I, B_i) \tag{1}$$

Here, I is the original backlight image, B_i is the bounding region of the i-th cavity or seal zone, and I_i is the aligned inspection window. Cropping by template region reduces unnecessary background variation and allows the model to compare defects across repeated cavity positions.

$$C_i = \mathrm{mean}(I_i) - \mathrm{mean}(N_i) \tag{2}$$

The contrast score C_i compares the cavity window with its neighbouring normal reference region N_i. It reduces the influence of global exposure drift and highlights local attenuation, bright leakage or abnormal transparency.

$$M_i = \mathbf{1}(|I_i - \mu_i| > k\sigma_i) \tag{3}$$

The initial candidate mask M_i is obtained from local deviation from the cavity baseline. The mask is not treated as the final segmentation result; it is an evidence map that provides candidate regions for later structured learning. Candidate-mask construction is intentionally conservative so that weak defects are retained before structured fusion.

$$v_i = [A_i, P_i, E_i, C_i, Q_i] \tag{4}$$

The descriptor vector v_i records mask area, perimeter, eccentricity, contrast and image-quality score. These variables summarize whether the candidate region resembles a missing tablet, a foreign particle, a bubble, a dent or a seal discontinuity.

$$q_i = 1 - \eta b_i \tag{5}$$

The image-quality factor q_i decreases when motion blur, low exposure or excessive reflection is detected. This factor prevents poor imaging conditions from being interpreted as real defect evidence.

3.2. TabM-based Defect Feature Interaction Learning

Mask encoding converts candidate evidence into the structured feature vector before the TabM module. The goal is to learn nonlinear relationships among image descriptors, cavity coordinates, local contrast and production-state variables. Figure 2 is the complete TabM-CatBoost framework: TabM generates interaction-aware representations, and CatBoost produces segmentation confidence, defect type, risk level and review output. The figure clarifies that mask area, boundary shape, cavity position and image quality are not interpreted independently; their joint pattern determines whether a weak signal is a true defect or a harmless artifact.

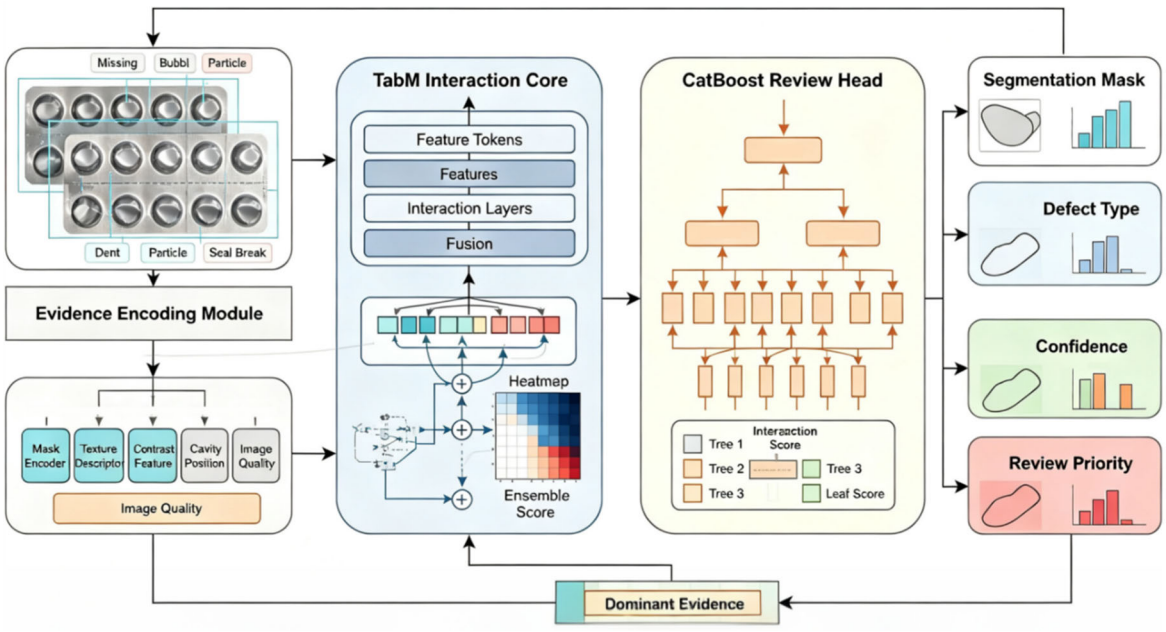


Figure 2. TabM-CatBoost weak-defect segmentation and review-output framework.

$$h_i = f_{\text{TabM}}(v_i, p_i, q_i) \tag{6}$$

The vector h_i denotes the TabM representation of the i-th inspection window. The term p_i records cavity position and seal-zone identity, while q_i keeps source reliability in the representation.

$$s_i = \sigma(W_s h_i + b_s) \tag{7}$$

The segmentation score s_i estimates whether the candidate mask belongs to a real defect. It is used to filter weak background artifacts before the final review head is applied. This stage keeps the representation compact and suitable for production deployment, where inference delay and model maintenance must be controlled.

3.3. CatBoost Review Head and Defect-Risk Output

The CatBoost review head receives the TabM representation and the original interpretable features. This design allows the model to retain nonlinear feature interaction while preserving direct engineering evidence for inspection review. The review head is designed for quality-control action rather than only numerical scoring, so it estimates defect class, review priority and confidence.

$$r_i = \text{CatBoost}([h_i, v_i]) \tag{8}$$

The risk value r_i represents defect severity and review priority. It is separated from the segmentation score because a small defect can be important when it is located on a seal line or associated with a high-risk foreign-particle pattern.

$$L = L_{\text{seg}} + \lambda_1 L_{\text{cls}} + \lambda_2 L_{\text{risk}} \tag{9}$$

The training objective combines segmentation loss, defect-class loss and risk-estimation loss. During inference, the system reports defect mask, defect type, confidence, dominant evidence and review priority. This output is designed for quality-control use rather than only pixel-level benchmarking. The segmentation term keeps the predicted mask spatially meaningful, the classification term separates defect categories, and the risk term connects visual evidence with inspection action.

4. Results and Discussion

4.1. Dataset Setting and Imaging-Condition Distribution

The generated dataset contains 18,960 blister inspection windows, including normal cavities, missing tablets, broken tablets, foreign particles, weak sealing flaws, local dents, air bubbles and mixed defects. each record combines a backlight window, candidate mask, cavity position, contrast descriptor, geometry descriptor, source-quality flag and review label. Synthetic records are used because real pharmaceutical defects are scarce and often unavailable under production secrecy, but the simulation follows repeated blister geometry, expected transmitted-light response and packaging-review categories. Controlled inspection-dataset studies [16] show that simulation can support evaluation of rare defects when disturbance variables are made explicit. Illumination drift, transparency variation, tablet colour, registration shift and camera blur are included, and production-line validation research [17] confirms that these acquisition conditions must be considered before deployment.

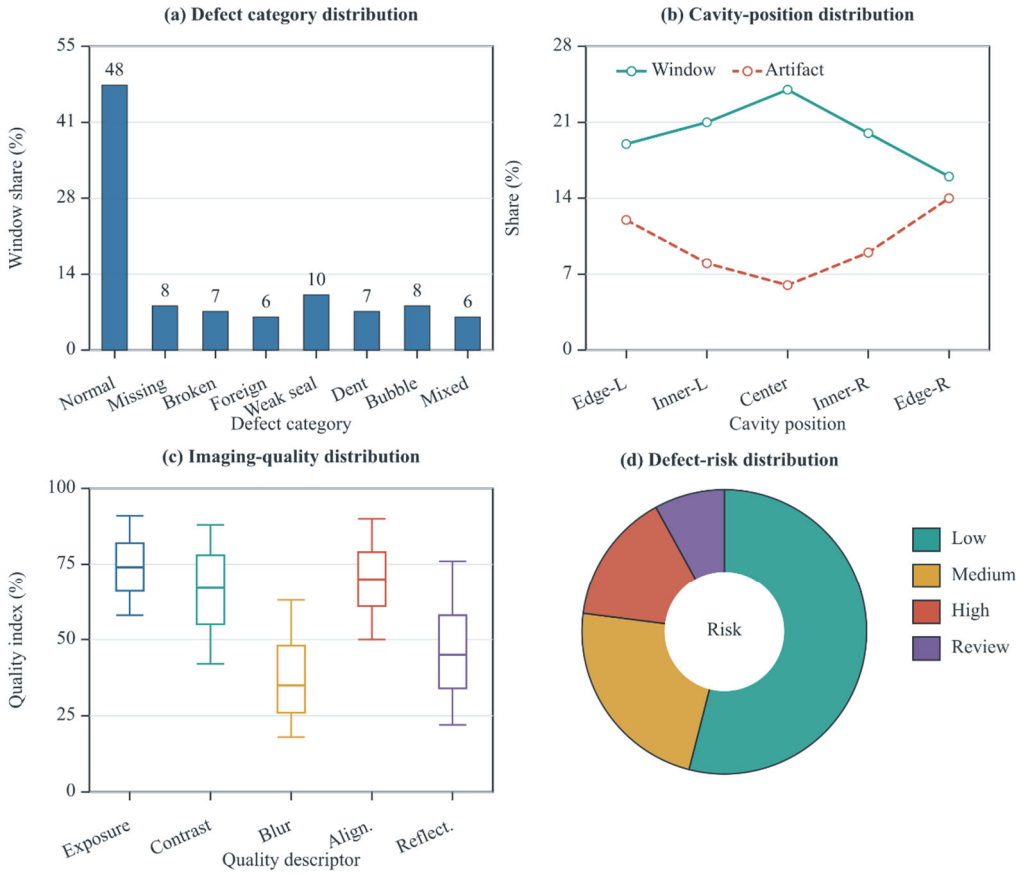


Figure 3. Backlight inspection dataset and defect-label distribution. (a) defect category distribution. (b) cavity-position distribution. (c) imaging-quality distribution. (d) defect-risk distribution.

Figure 3 is the dataset composition and label distribution. As shown in Figure 3(a), normal windows dominate the defect-category distribution, creating the class imbalance expected in packaging inspection. Figure 3(b) reports cavity-position distribution and shows why edge cavities deserve separate attention because they are more sensitive to blur and seal-line artifacts. As shown in Figure 3(c), image-quality distribution is concentrated in the medium-quality range, while low-quality windows are retained for robustness testing. Figure 3(d) shows that high-risk defects are less frequent than low-risk visual disturbance. These panels should be read together because defect category, cavity position, imaging quality and risk level jointly define the evaluation difficulty.

Figure 4 is about defect morphology and contrast response. As shown in Figure 4(a), weak sealing flaws and foreign particles occupy fewer pixels, so they are less visible in pixel-level metrics than missing-tablet cases. Figure 4(b) shows that backlight contrast improves local visibility but cannot completely separate real defects from reflection artifacts. This is why the model uses both masks and descriptors rather than a single threshold. Research on industrial surface inspection [18] supports the need for local appearance and regional context, and defect-shape research [19] further shows that boundary compactness and local contrast must be evaluated together. The descriptor set therefore records size, shape, position and contrast evidence for later quality-engineering review.

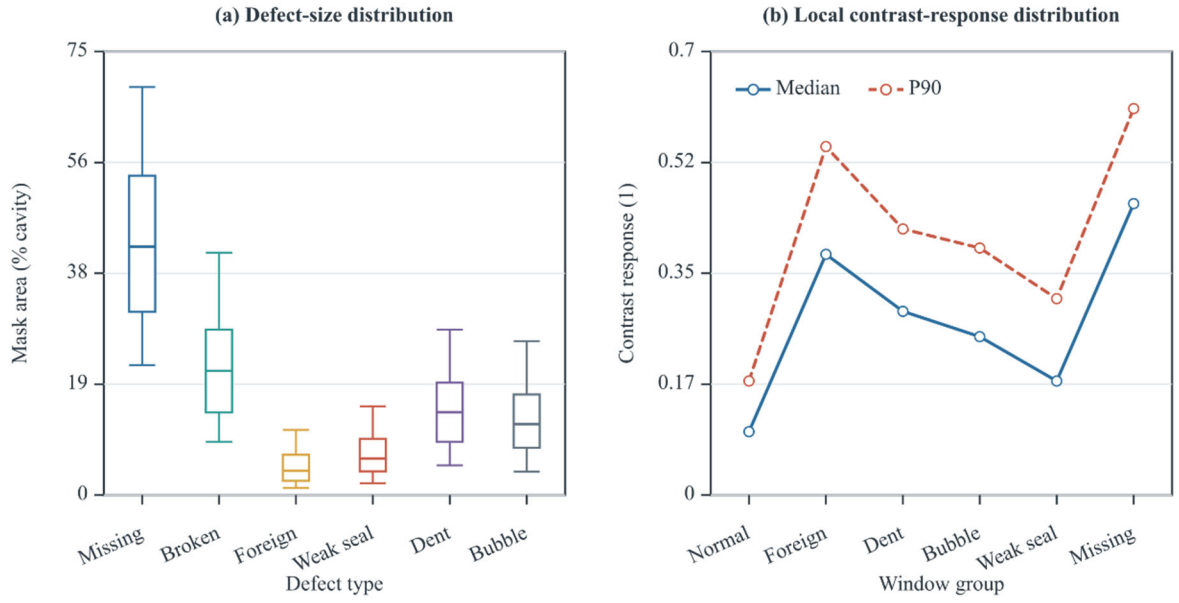


Figure 4. Defect morphology and contrast-response statistics. (a) defect-size distribution. (b) local contrast-response distribution.

The dataset is divided into training, validation and test sets at the package-lot level rather than at the individual image-window level. The training set is 70%, the validation set is 15% and the test set is 15% of the data. Lot-level division prevents the same package template from appearing in both training and testing and makes evaluation closer to deployment. Production inspection research [20] shows that lot-level evaluation is more reliable for judging transfer to new batches. Industrial validation research [21] also emphasizes unseen-batch testing before transferring a model across shifts or material lots. The reported accuracy is therefore less likely to come from memorizing repeated cavity patterns.

4.2. Segmentation Accuracy and Baseline Comparison

Compare the TabM-CatBoost model presented in this paper with threshold segmentation, U-Net, DeepLab-style segmentation, a pure CatBoost model based on handcrafted features, and a TabM-only classifier. Table 1 shows the main quantitative results. The average IoU of the new way was 0.887, the boundary F1 was 0.861, the high-risk recall was 0.934, and the false-review rate was 7.8 per cent. Thus, the hybrid Design can enhance the quality of the masks and improve decision accuracy. Deep Segmentation baselines produce good masks of the missing tablets, but they are more likely to have weak sealing defects. The original CatBoost model is more interpretable; however, some specific boundary information has been lost. TabM-only model learns feature interactions but has more false alarms due to the absence of the final boosting review head. Recently, hybrid inspection research [22] has also found that by combining learned representations with structured decision models, reliability can be enhanced under conditions of scarce abnormal samples. The same train-validation-test partition and candidate-window definition are used for all the methods in the comparison. Evaluation research [23] recommends this kind of controlled comparison to prevent changes in preprocessing from being mistaken for improvements in the model. Threshold baseline uses adaptive local deviation; U-Net and DeepLab variants employ image masks; CatBoost only utilizes structured features; TabM has no

boosting of review. Therefore, it can be determined whether the improvement in performance is due to segmentation representation, tabular interaction or review-stage calibration.

Table 1. Quantitative comparison of segmentation and review metrics.

Setting	Image windows	Defect share	Main metric
Normal blister	3,400	0.0%	specificity
Missing tablet	2,650	14.8%	recall
Foreign particle	2,420	13.5%	IoU
Weak sealing flaw	2,310	12.9%	boundary F1
Dent or bubble	2,180	12.2%	risk score

Figure 5 is the segmentation mask quality comparison for different defect types. As shown in Figure 5(a), the missing-tablet case requires a whole-cavity mask rather than fragmented inner regions. Figure 5(b) shows that small foreign-particle masks should not expand into harmless background. As shown in Figure 5(c), weak sealing flaws remain difficult because they are thin and low contrast. Figure 5(d) shows that the dent and bubble group requires distinguishing shape change from lighting change. These examples explain why the model reaches 0.887 mean IoU and 0.861 boundary F1 while still being evaluated by defect-specific quality-control requirements. Boundary-sensitive segmentation research [24] and class-dependent boundary research [25] support this use-oriented interpretation.

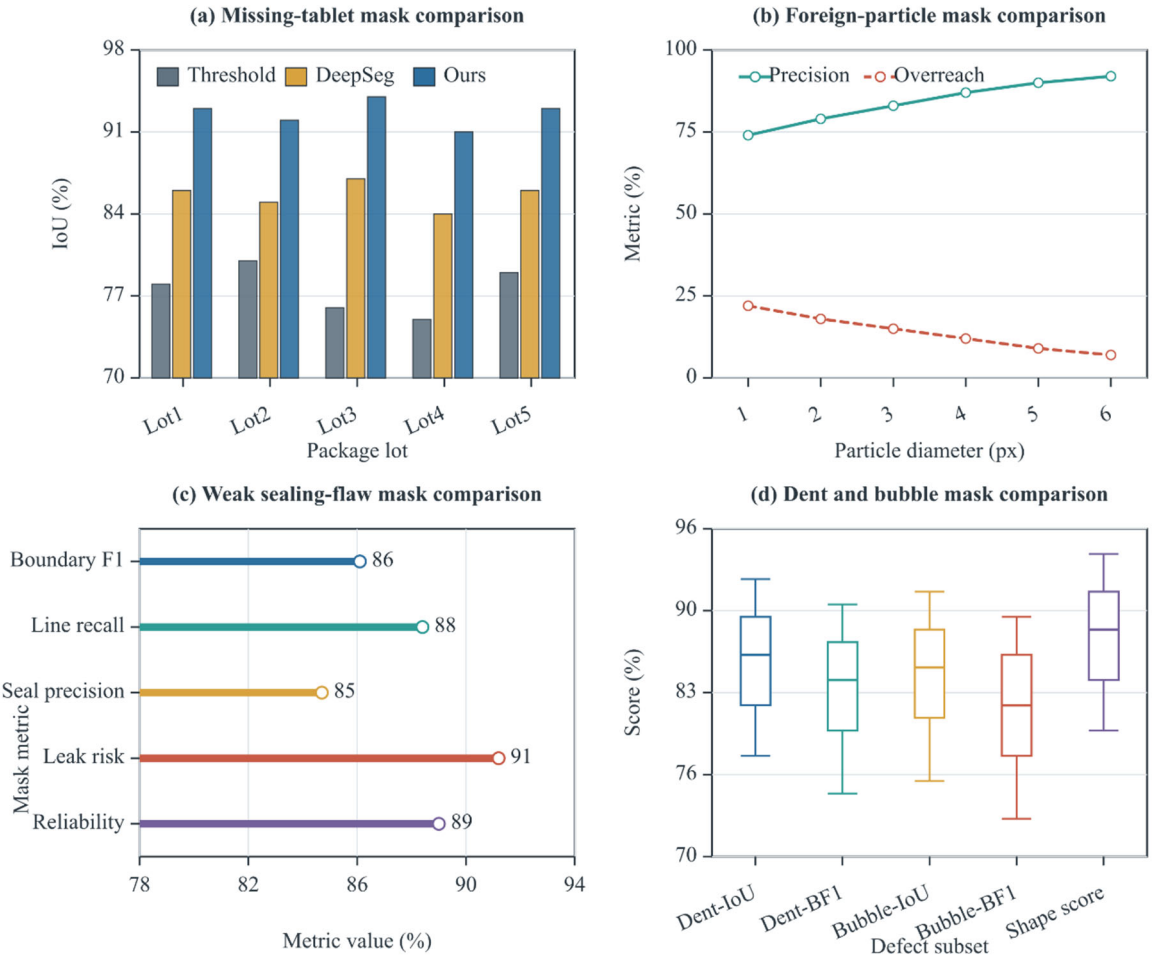


Figure 5. Segmentation mask quality under different blister-defect types. (a) missing-tablet mask comparison. (b) foreign-particle mask comparison. (c) weak sealing-flaw mask comparison. (d) dent and bubble mask comparison.

Figure 6: Baseline Methods and Boundary Errors Comparison. As shown in Figure 6(a), the proposed model obtains the strongest mean-IoU results across defect groups. Figure 6(b) shows that boundary-error distribution is more stable for weak sealing and bubble cases. As shown in Figure 6(c), threshold segmentation has more missed detections, and pure deep segmentation still misses some small foreign particles. Figure 6(d) shows that the CatBoost review head removes many reflection-induced false alarms. Boosting-based review research [26] explains this advantage through the cooperation of continuous image descriptors and categorical process tags, while fine-defect recognition research

[27] confirms that line-like defects are sensitive to sampling resolution and local contrast. The remaining errors therefore concentrate mainly in thin seal-region defects.

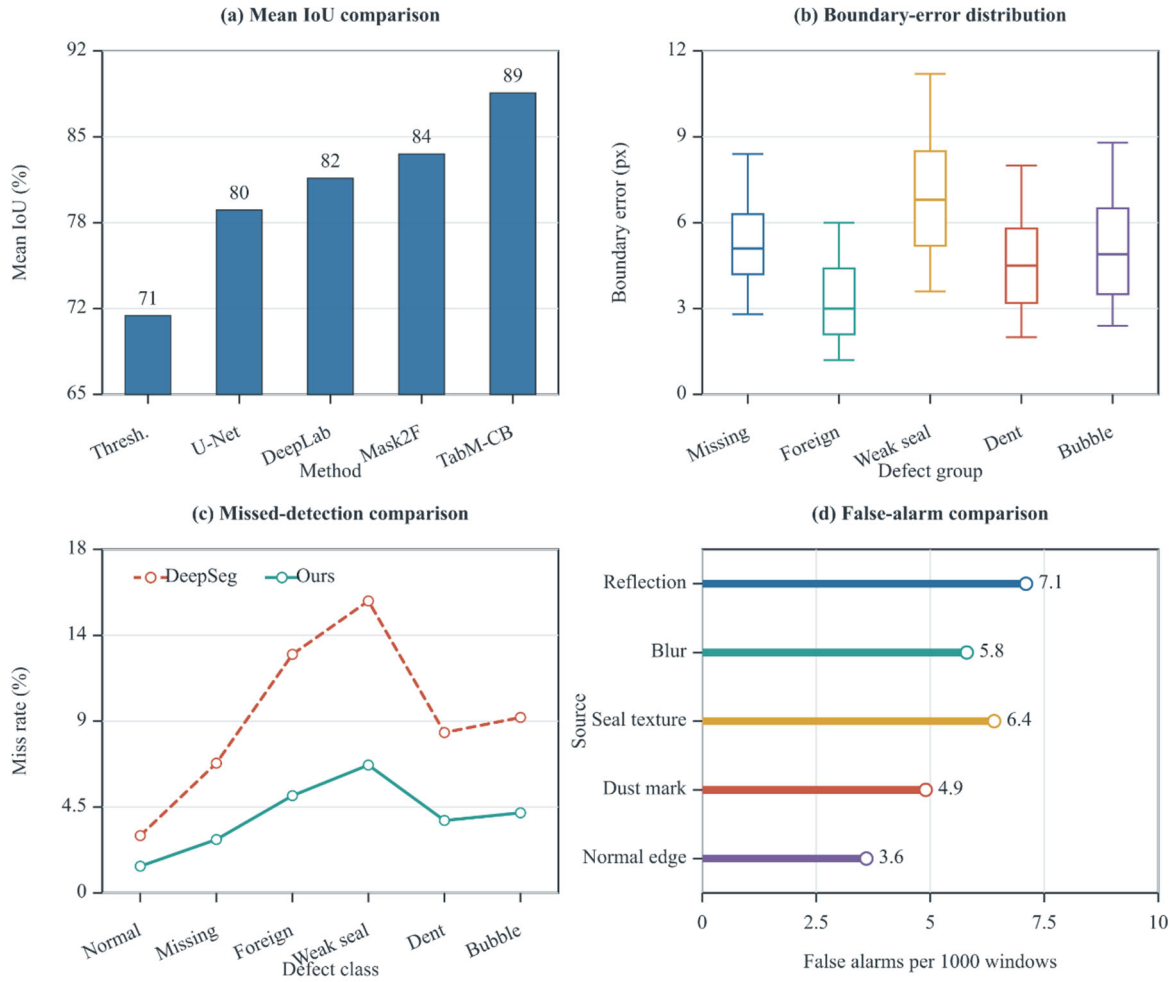


Figure 6. Baseline comparison and boundary-error analysis. (a) mean IoU comparison. (b) boundary-error distribution. (c) missed-detection comparison. (d) false-alarm comparison.

Segmentation indicators alone do not reflect the actual use of the inspection. A model may have a high IoU for large missing-tablet defects, but it will be unable to recognise weak seal-line anomalies that are more likely to cause production problems. Therefore, review-oriented indicators are used. The model reduced the false-report rate by 13.6% compared with the best neural baseline and increased high-risk detection by 5.2 percentage points. Too many false alarms will slow down the packaging line, and otherwise, a serious defect may not be found. Quality-management studies [28] have proposed that the inspection model should be evaluated in terms of defect location and the extent of follow-up. The above output satisfies the requirements by returning mask, class, risk, and dominant_evidence.

4.3. Robustness and Industrial Review Interpretation

Robustness is tested under illumination drift, transparency variation, motion blur and registration error. Figure 7 summarizes these results. Figure 7(a) is the illumination-disturbance test, showing that local contrast normalization limits the effect of global brightness change. As shown in Figure 7(b), transparency variation from plastic-film thickness remains challenging but is not decisive when cavity geometry and reliability tags are included. Blur and registration disturbance are treated as evidence-quality problems rather than hidden preprocessing failures. Industrial deployment studies [29] identify acquisition disturbance as a major source of false alarms in high-throughput visual inspection, so the model routes uncertain evidence toward review instead of overconfident release.

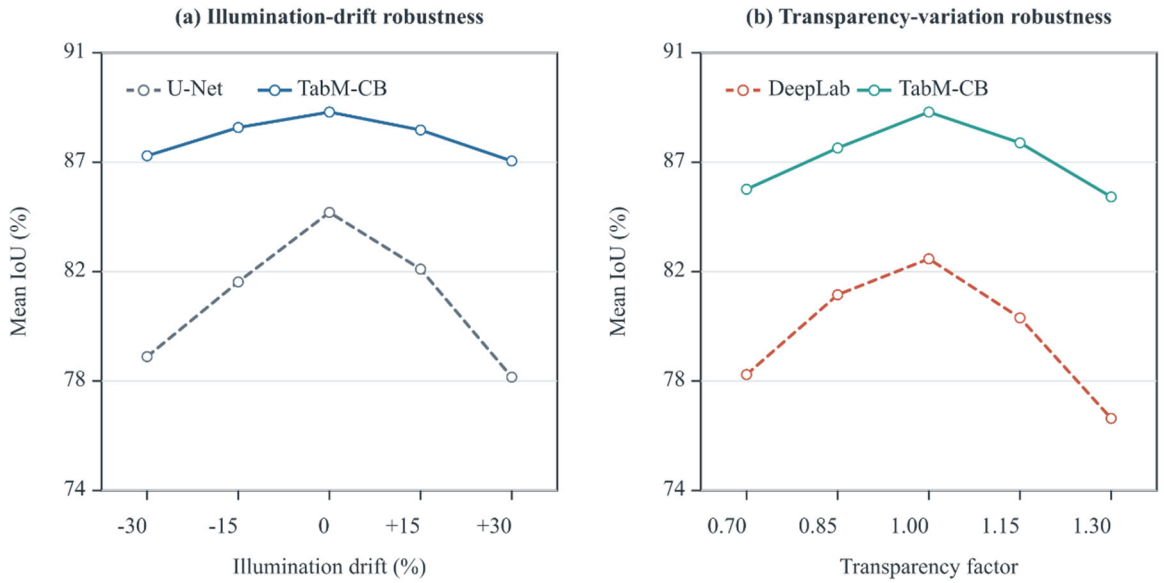


Figure 7. Robustness under imaging disturbance. (a) illumination-drift robustness. (b) transparency-variation robustness.

Figure 8 is the review-oriented output. Figure 8(a) shows a defect evidence card containing mask area, boundary contrast, cavity position, image quality, dominant evidence and predicted defect class. Figure 8(b) converts the prediction into normal pass, manual review, batch isolation or line-adjustment recommendation. The figure clarifies that quality engineers need masks and explanations that are simple enough for line use, not only coloured overlays. Research on explainable manufacturing AI [30] shows that evidence-level output can build trust and reduce manual disputes. Low-confidence cards should therefore request manual review or recapture rather than automatic acceptance.

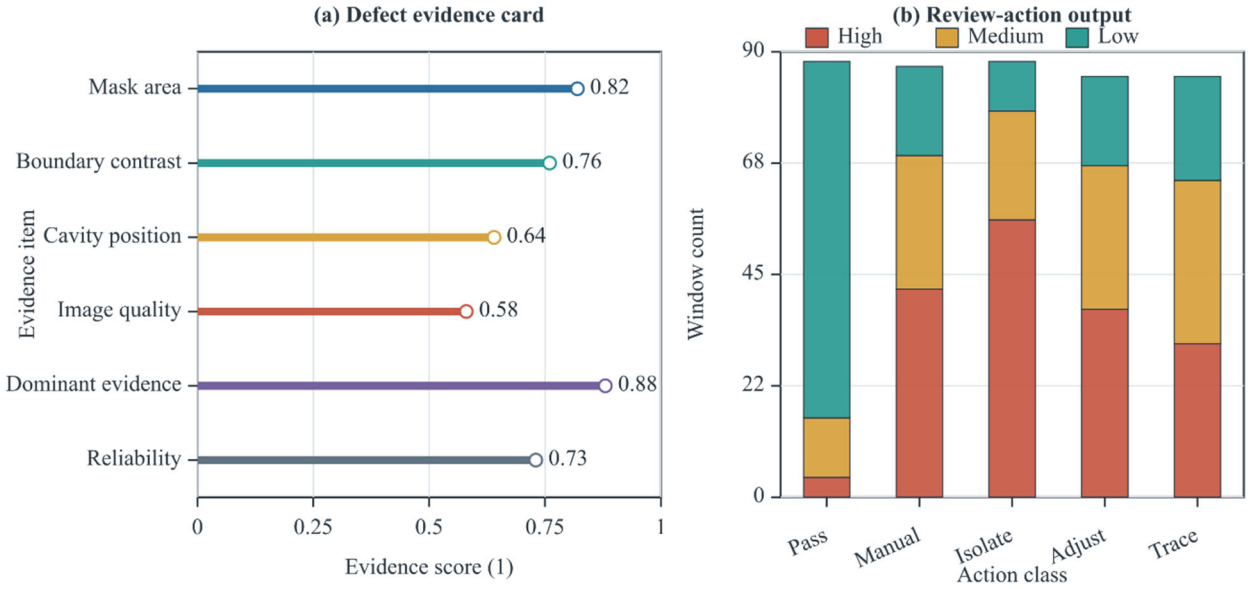


Figure 8. Defect evidence and inspection-review output. (a) defect evidence card. (b) review-action output.

Ablation analysis identifies the contributions of individual modules. TabM has reduced the mean IoU by decreasing the correlation of features in contrast, geometry and position. CatBoost is not suitable; otherwise, the final review head will be unable to correct for reflection-driven candidate masks. Removing reliability tags reduces the performance of blur and low-light conditions. Local Contrast Normalisation removes only the weaker seal defects. Therefore, the constructed model is not formed by the above components, but rather, each part has its own inspection purpose. Recently, interpretable ensemble research [31] has provided support for this structure, and robustness often requires keeping both the learned representation and human-readable evidence. Ablation studies have shown that the model is sensitive to how quality tags are applied. If there is no image quality information, assume that the blurred cavity edges are small seal defects in the model. Position encoding is not used, and therefore the harmless cavity-edge shadow may be mistaken for a foreign particle. Feature-ablation studies [32] support the above conclusion; industrial inspection variables are often context-dependent rather than independently meaningful. Therefore, the proposed pipeline keeps the original descriptors and the TabM representation, and if the learned representation is unreliable, direct engineering evidence can be used.

The model is still imperfect. First, the dataset is simulated under realistic imaging and packaging assumptions, but real production may include dust marks, operator interference, cleaning residue and unusual defect mixtures. Second, the model depends on blister-template alignment, and severe package deformation requires an additional registration module. Third, CatBoost explanations identify influential features but do not constitute full causal diagnosis of line faults. Fourth, review policy must be customized by factory according to acceptable false-alarm rate, product value, production speed and regulatory workflow. Future work will add real-time verification, active learning for rare defects and integration with batch traceability systems [33] so predicted review actions can be compared with actual production-line decisions.

There are some defects; however, the proposed system is still feasible. Backlight imaging provides defect-sensitive visual evidence; TabM learns the interaction of structured descriptors, and then CatBoost makes a reliable review decision based on this evidence. The model can be used to examine subtle defects in the absence of light transmission. Segmentation masks can be used for quality control, and this information is often not available in purely visual benchmarks. Broader industrial inspection research [34] suggests that this connection between localisation and the review decision is necessary for deployment. The current results support the application of hybrid image-tabular learning in pharmaceutical blister packaging inspection, especially when defect labels are scarce and high operational reliability is needed. Deployment can be connected to batch-level quality records, alarm logs and line-maintenance events in the framework. Deployment studies [35] recommend that the output of visual inspection be stored along with evidence fields to trace repeated alarms back to a specific unit, material lot or camera setting. The built environment will be used to store the main descriptors, image quality data and review records of the evidence cards. The above information will be used for a later recalibration to prevent the model from becoming an uninterpretable visual classifier of production decisions.

5. Conclusion

This study developed a TabM-CatBoost model for pharmaceutical blister packaging defect segmentation based on backlight inspection images, and the overall conclusion is that weak transmitted-light evidence can support reliable review only when candidate masks, cavity geometry, image quality and process descriptors are fused in an auditable workflow. Firstly, on 18,960 simulated inspection windows, the model achieved a mean IoU of 0.887, a boundary F1 of 0.861 and a high-risk recall of 0.934, showing quantitative improvement in both localization and review sensitivity. Secondly, compared with threshold segmentation, neural segmentation, pure CatBoost and TabM-only variants, the hybrid framework better handled weak seals, small foreign particles, bubbles and reflection-driven false alarms because TabM captured feature interactions and CatBoost calibrated review decisions. Thirdly, the main limitation is that the current evaluation depends on simulated data, template alignment and factory-specific review thresholds, so real-line validation remains necessary.

Future work should collect real production-line defect images, compare predicted review actions with batch records and inspector decisions, expand active learning for rare defects and connect evidence cards with sealing temperature, feeding status and traceability logs. These steps will make the method more useful for root-cause analysis and closed-loop quality management.

6. Declarations

6.1. Author Contributions

Muhammad Ahmad contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Firdaus Yaacob contributes to validation, analysis, investigation, data collection, draft preparation. All authors have read and agreed with the manuscript before its submission and publication.

6.2. Funding and Acknowledgments

This research received no specific financial support from any funding agency.

6.3. Institutional Review Board Statement

Not applicable.

6.4. Informed Consent Statement

Not applicable.

6.5. Declaration of Competing Interest

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

6.7. Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized Grammarly for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

7. References

[1] Diaz, J. Z., Farooq, M. A., & Corcoran, P. (2024). Automatic inspection of seal integrity in sterile barrier packaging: A deep learning approach. IEEE Access, 12, 22904-22927. https://doi.org/10.1109/ACCESS.2023.3348779

[2] Li, Z., Yao, Y., Wen, R., & Liu, Q. (2024). Dual-modal illumination system for defect detection of aircraft glass canopies. Sensors, 24(20), 6717. https://doi.org/10.3390/s24206717

[3] Ma, Y., Yin, J., Huang, F., & Li, Q. (2024). Surface defect inspection of industrial products with object detection deep networks: A systematic review. Artificial Intelligence Review, 57(12), 333. https://doi.org/10.1007/s10462-024-10956-3

[4] Trampert, P., Masiak, T., Schmidt, F., Thewes, N., Kruse, T., Witte, C., & Schneider, G. (2024). Production-ready end-to-end visual quality inspection for defect detection on surfaces based on a multi-stage AI system. Proceedings of the 4th International Conference on Image Processing and Vision Engineering, 55-66. https://doi.org/10.5220/0012536500003720

[5] Fukuda, K., Ishikawa, M., Yoshida, Y., Fukaya, K., Kagetani, R., & Shindo, H. (2024). Deep learning-based image quality adaptation for die-to-database defect inspection. Metrology, Inspection, and Process Control XXXVIII, 69. https://doi.org/10.1117/12.3008799

[6] Ameri, R., Hsu, C. C., & Band, S. S. (2024). A systematic review of deep learning approaches for surface defect detection in industrial applications. Engineering Applications of Artificial Intelligence, 130, 107717. https://doi.org/10.1016/j.engappai.2023.107717

[7] Horide, K., Chatani, S., & Okayama, T. (2024). Annotation effort reduction for introducing AI into optical inspection systems. 2024 International Symposium on Semiconductor Manufacturing, 1-4. https://doi.org/10.1109/ISSM64832.2024.10874966

[8] Park, J., & Lee, J. (2024). Automated visual inspection of particle defect in semiconductor packaging. Journal of Mechanical Science and Technology, 38(8), 4447-4453. https://doi.org/10.1007/s12206-024-0740-6

[9] Shetty, B., Madhushree, S. M., Muddukrishna, B. S., Shenoy, R., Kirankumar, H., Joshi, M., & Pai, G. K. (2024). Machine vision-based quality inspection of pharmaceutical bottle packaging. Research Journal of Pharmacy and Technology, 5154-5160. https://doi.org/10.52711/0974-360X.2024.00789

[10] Mezher, A. M., & Marble, A. E. (2024). Computer vision defect detection on unseen backgrounds for manufacturing inspection. Expert Systems with Applications, 243, 122749. https://doi.org/10.1016/j.eswa.2023.122749

[11] Barberena, X., Saiz, F., & Barandiaran, I. (2025). Handling drift in industrial defect detection through MMD-based domain adaptation. Proceedings of the 20th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 420-429. https://doi.org/10.5220/0013170900003912

[12] Ahmed, I., Siddiqi, M. U. R., & Ahmad, M. (2024). AI-enhanced visual inspection systems for robust detection of product packaging defects in manufacturing environments. 2024 IEEE 15th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference, 220-225. https://doi.org/10.1109/UEMCON62879.2024.10754674

[13] Lin, Q., & Lin, X. (2025). Deterministic tabular learning with TabM(Huber) and out-of-fold blending. 2025 6th International Symposium on Computer Engineering and Intelligent Communications, 448-451. https://doi.org/10.1109/ISCEIC67854.2025.11405712

[14] Zhai, H., Qi, X., Han, H., Li, J., & Yan, P. (2025). Quality prediction of GH4169 alloy turbine discs based on CatBoost feature selection and SHAP validation. Proceedings of the 2025 International Conference on Artificial Intelligence and Smart Manufacturing, 39-49. https://doi.org/10.1145/3756423.3756430

[15] Maitra, V., Arrasmith, C., & Shi, J. (2024). Introducing explainable artificial intelligence to property prediction in metal

additive manufacturing. Manufacturing Letters, 41, 1125-1135. https://doi.org/10.1016/j.mfglet.2024.09.138

[16] Wang, X., Huang, X., Yue, Z., & Ma, J. (2024). Cross-domain defect GAN: Defect generation for defect detection in industrial area. 2024 3rd International Conference on Big Data, Information and Computer Network, 194-198. https://doi.org/10.1109/BDICN62775.2024.00046

[17] Corallo, A., Del Vecchio, V., & Di Prizio, A. (2024). Advanced AI-based solutions for visual inspection: A systematic literature review. Proceedings of the 26th International Conference on Enterprise Information Systems, 656-664. https://doi.org/10.5220/0012618000003690

[18] Wang, F., He, Z., & Wang, J. (2024). Industrial anomaly segmentation algorithm based on asymmetric teacher-student reconstruction network. Proceedings of the 2024 7th International Conference on Artificial Intelligence and Pattern Recognition, 561-568. https://doi.org/10.1145/3703935.3704071

[19] Swarna, R. A., Hossain, M. M., Khatun, M. R., Rahman, M. M., & Munir, A. (2024). Concrete crack detection and segregation: A feature fusion, crack isolation, and explainable AI-based approach. Journal of Imaging, 10(9), 215. https://doi.org/10.3390/jimaging10090215

[20] Chen, X., Liu, M., Niu, Y., Wang, X., & Wu, Y. C. (2024). Deep-learning-based lithium battery defect detection via cross-domain generalization. IEEE Access, 12, 78505-78514. https://doi.org/10.1109/ACCESS.2024.3408718

[21] Zemzemoglu, M., & Unel, M. (2024). Thermal inspection and quality assessment for AFP processes via automatic defect detection and segmentation. IECON 2024 - 50th Annual Conference of the IEEE Industrial Electronics Society, 1-6. https://doi.org/10.1109/IECON55916.2024.10905789

[22] Ilani, M. A., & Banad, Y. M. (2024). Interpretable AI prediction of melt pool geometry in IoT-enabled additive manufacturing in Industry 4.0 utilizing CatBoost: Decision-making via NSGA-II. 2024 International Conference on AI x Data and Knowledge Engineering, 62-65. https://doi.org/10.1109/AIxDKE63520.2024.00017

[23] Huang, Y., Dai, E., & Chen, K. (2024). MaskCore: Towards precise semantic segmentation in industrial anomaly detection. 2024 International Conference on Cloud and Network Computing, 109-116. https://doi.org/10.1109/ICCNC63989.2024.00025

[24] Liu, W., Zhao, P., Zhao, Y., Fu, Y., Dai, J., & Zhou, L. (2024). A reflectance-correction retinex framework for thermal image enhancement in nondestructive defect detection of CFRP. Measurement, 237, 115070. https://doi.org/10.1016/j.measurement.2024.115070

[25] Lee, M., & Choi, J. (2024). Text-guided variational image generation for industrial anomaly detection and segmentation. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 26509-26518. https://doi.org/10.1109/CVPR52733.2024.02504

[26] Pan, A. S., Nie, B. X., & Zhai, C. X. (2024). Enhancing wafer defect detection via ensemble learning. AIP Advances, 14(8), 085009. https://doi.org/10.1063/5.0222140

[27] Yang, H., Ma, T., Tong, Z., Wang, H., Wang, N., & Cheng, H. (2024). Deployment strategies for lightweight pavement defect detection using deep learning and inverse perspective mapping. Automation in Construction, 167, 105682. https://doi.org/10.1016/j.autcon.2024.105682

[28] Lavasa, E., Chadoulos, C., Siouras, A., Etxabarri Llana, A., Rodríguez Del Rey, S., Dalamagas, T., & Moustakidis, S. (2024). Toward explainable Metrology 4.0: Utilizing explainable AI to predict the pointwise accuracy of laser scanning devices in industrial manufacturing. Artificial Intelligence in Manufacturing, 479-501. https://doi.org/10.1007/978-3-031-46452-2_27

[29] Yan, L. (2024). Enhanced low-illumination image defect detection using machine vision. Journal of Industrial Intelligence, 2(4), 218-229. https://doi.org/10.56578/jii020403

[30] Zeiser, T., Ehret, D., Lutz, T., & Saar, J. (2024). Explainable AI in manufacturing. 2024 IEEE International Conference on Engineering, Technology, and Innovation, 1-8. https://doi.org/10.1109/ICE/ITMC61926.2024.10794363

[31] Forest, F., Porta, H., Tuia, D., & Fink, O. (2024). From classification to segmentation with explainable AI: A study on crack detection and growth monitoring. Automation in Construction, 165, 105497. https://doi.org/10.1016/j.autcon.2024.105497

[32] Dai, H., Wang, J., Zhong, Q., Chen, T., Liu, H., Zhang, X., & Lu, R. (2024). A GAN-based anomaly detector using multi-feature fusion and selection. Scientific Reports, 14(1), 5259. https://doi.org/10.1038/s41598-024-52378-9

[33] Barman, T., Coleman, S., Kerr, D., Harrigan, S., & Quinn, J. (2024). Advancements in industrial visual inspection: Harnessing hyperspectral imaging for automated solder quality assessment. 2024 IEEE 22nd International Conference on Industrial Informatics, 1-6. https://doi.org/10.1109/INDIN58382.2024.10774265

[34] Koch, J., Jevremovic, D., Moenck, K., & Schüppstuhl, T. (2024). A digital assistance system leveraging vision foundation models and 3D localization for reproducible defect segmentation in visual inspection. Procedia CIRP, 130, 387-397. https://doi.org/10.1016/j.procir.2024.10.105

[35] Villalonga, A., Cruz, Y. J., Alfaro, D., Haber, R. E., Martínez-Lastra, J. L., & Castaño, F. (2024). Enhancing quality inspection in zero-defect manufacturing through robotic-machine collaboration. 2024 7th Iberian Robotics Conference, 1-6. https://doi.org/10.1109/ROBOT61475.2024.10797427