A Machine Vision–Based Automated Wheel Leak Detection System Using Real-Time Object Detection in the Water Leak Testing Process

Susetyo Bagas Bhaskoro
Cyber-Physical System
Politeknik Manufaktur Bandung
Bandung, Indonesia
bagas@polman-bandung.ac.id

Sarosa Castrena Abadi
Industrial Informatics Engineering Technology
Politeknik Manufaktur Bandung
Bandung, Indonesia
sarosa@ae.polman-bandung.ac.id

Aris Budiyarto
Cyber-Physical System
Politeknik Manufaktur Bandung
Bandung, Indonesia
aris_b@polman-bandung.ac.id

Inkreswari Retno Hardini
Faculty of Economics and Business
Jakarta State of University
Jakarta, Indonesia
inkreswari.retno@unj.ac.id

M. Pribadi Lukman
Automation Engineering Technology
Politeknik Manufaktur Bandung
Bandung, Indonesia
mpribadilukman@gmail.com

Abstract—Water leak testing in automotive wheel manufacturing has traditionally relied on manual visual inspection of bubble formation, introducing subjectivity and limiting repeatability in quality assurance processes. This study developed and experimentally validated a real-time leak detection system based on machine vision, directly integrated with an industrial water leak tester platform. A dataset comprising 686 annotated images was constructed from recorded operational testing sequences and partitioned into 80% training and 20% validation subsets. The network was trained for 150 epochs and deployed within an integrated framework incorporating temporal decision logic and automated event logging to ensure deterministic classification under continuous video streaming. Experimental validation was conducted across five scenarios (A–E), including high-leak, low-leak, no-leak, and in-situ operational testing conditions, totaling 100 trials. The aggregated confusion matrix yielded 60 true positives and 40 true negatives with zero false positives and false negatives, resulting in accuracy, sensitivity, specificity, precision, and F1-score values of 1.0 within the evaluated domain. Receiver operating characteristic and precision–recall analyses confirmed strong class separability and stable decision boundaries. Although the results demonstrated high discriminative performance under controlled and operational settings, further large-scale validation under heterogeneous industrial environments is required to fully assess long-term robustness. The proposed framework provided an automated, objective, and real-time inspection solution aligned with Industry 4.0 principles for intelligent manufacturing systems.

Keywords—wheel leak detection, water leak testing, bubble detection, machine vision, real-time object detection

Article info: Date Submitted: 2026-02-22 | Date Revised: 2026-04-12 | Date Accepted: 2026-04-22

I. INTRODUCTION

Automotive wheels are manufactured using the high-pressure die casting (HPDC) process [1], [2] based on aluminum alloys due to their high strength-to-weight ratio and corrosion resistance. However, the flow dynamics of molten metal under elevated injection pressure could induce porosity formation as a result of air entrapment within the mold cavity [3], [4], [5]. The volumetric porosity fraction was expressed as [6].

$$\phi = \frac{V_{void}}{V_{total}} \tag{1}$$

where V_{void} denoted the pore volume and V_{total} represented the total material volume. A porosity fraction $\phi > 0$ increased the probability of forming a through-thickness

DOI: https://doi.org/10.37396/jsc.v9i1.637

leakage path, which in tubeless wheels resulted in internal pressure reduction. According to the ideal gas law [7],

$$PV = nRT \tag{2}$$

a reduction in the amount of air $n$due to leakage caused a corresponding decrease in internal pressure P, thereby directly affecting vehicle operational safety. Industrial leak testing was performed using dry and wet methods [2], [8]. In the dry method, leakage was identified by monitoring pressure deviation over time.

$$\Delta P = P_0 - P(t) \tag{3}$$

with the component classified as leaking when $\frac{dP}{dt} < -\varepsilon$, where ε represented the allowable tolerance threshold. In the wet method (water leak testing) [9], the pressurized component was immersed in water, and leakage was identified through bubble formation caused by the pressure gradient $\Delta P = P_{internal} - P_{ambient}$ [7]. The leakage flow rate could be approximated using the following relation:

$$Q = C_d A \sqrt{\frac{2\Delta P}{\rho}} \tag{4}$$

where Q denoted the leakage air flow rate, C_d represented the discharge coefficient, Acorresponded to the effective leakage orifice area, and ρindicated the fluid density. The wet method was more widely adopted due to its simple system configuration and low operational cost [6], [10], [11]; however, visual evaluation of bubble formation remained subjective and was highly dependent on lighting conditions and operator perception [12].

To enhance inspection objectivity and consistency, bubble detection was formulated as an object detection problem in the digital image domain. Let the input image be represented as a tensor $I \in \mathbb{R}^{H \times W \times 3}$, then the YOLO model mapped the function

$$f_\theta : I \rightarrow \{(b_i, c_i, s_i)\}_{i=1}^{N} \tag{5}$$

where b_i denoted the bounding box coordinates, c_i represented the object class (bubble), and $s_i \in [0,1]$corresponded to the confidence score [6], [13], [14]. The leakage decision was determined based on a temporal indicator function

$$L = \begin{cases} 1, & \text{if } s_i \geq \tau \text{ for } k \text{ consecutive frames} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where τ represented the confidence threshold and k denoted the temporal consistency parameter. This formulation transformed manual visual inspection into a deterministic classification system based on neural network inference, thereby improving accuracy, consistency, and reproducibility in automotive wheel water leak testing quality control. Accordingly, this study developed a YOLO-based wheel leak detection system integrated into the water leak testing process [14] to identify bubble formation in real time. The proposed framework aimed to enhance inspection reliability and reduce dependence on manual observation within industrial quality assurance operations.

II. METHOD

This study employed an experimental engineering approach to design, implement, and evaluate an automated wheel leak detection system based on computer vision integrated

DOI: https://doi.org/10.37396/jsc.v9i1.637

with a water leak tester machine [2], [10], [15]. Table 1 presented a comparison of industrial leak testing methods, from which the wet testing method was selected for the present investigation.

Table 1. Comparison between dry and wet air leak testing methods based on detection principle, leakage indicators, advantages, limitations, and industrial application preference.

Aspect	Dry Method	Wet Method
Definition	Testing of a sealed component using pressurized air without liquid medium.	Testing conducted by immersing the pressurized component in water.
Leakage Indicator	Sudden decrease in internal air pressure.	Formation of air bubbles in water.
Advantages	Did not require a liquid medium and was suitable for clean environments.	Enabled straightforward visual observation and simple operation.
Limitations	Required highly sensitive pressure sensors and provided less intuitive visual feedback.	Introduced potential splashing and required a water containment vessel.
Industrial Preference	Was applied to products sensitive to moisture exposure.	Was more widely adopted due to ease of visual leak identification.

The system was designed through the integration of hardware and software components. The hardware configuration consisted of a camera module for image acquisition, a processing unit, and an air pressurization system [16]. The wheel specimen was secured on a fixed fixture to maintain spatial stability during immersion. The input image was represented as a tensor $I \in \mathrm{R}^{H\times W\times 3}$, which was processed using the YOLOv5 model through the inference function

$$f_\theta(I) = \{(b_i, c_i, s_i)\}_{i=1}^{N} \tag{7}$$

where b_i denoted the bounding box coordinates, c_i represented the object class, and $s_i \in [0,1]$corresponded to the confidence score. The dataset was constructed from recorded operational leak testing videos and was extracted into 686 annotated images. The data were partitioned into 80% training and 20% validation subsets. The model was trained for 150 epochs to minimize the total loss function.

$$\mathcal{L}_{total} = \mathcal{L}_{box} + \mathcal{L}_{obj} + \mathcal{L}_{cls} \tag{8}$$

which comprised the bounding box regression error, objectness loss, and classification loss components. Model convergence was determined based on the stability and consistency of both training and validation loss values. During the deployment phase, the video stream was processed in a frame-by-frame manner. Each frame underwent preprocessing, including resizing and intensity normalization, prior to inference. The leakage decision was determined using a temporal indicator function.

$$L = \begin{cases} 1, & \text{if } s_i \geq \tau \text{ for } k \text{ consecutive frames} \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

where τ represented the confidence threshold and k denoted the temporal consistency parameter used to suppress spurious detections caused by visual disturbances. Performance evaluation was conducted on five test objects, each subjected to 20 trials. The predicted outcomes were compared with the ground-truth conditions to construct a confusion matrix consisting of True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). The evaluation metrics were computed as follows:

DOI: https://doi.org/10.37396/jsc.v9i1.637

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (11)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (12)$$

This approach ensured a quantitative assessment of model performance while simultaneously validating the system under real operational conditions, thereby confirming that the framework was not only algorithmically verified but also experimentally demonstrated in an industrial implementation context.

A. System Overview

Figure 1 illustrated the overall architecture of the wheel leak detection system developed in this study. The system was designed to identify air bubble formation as an indicator of leakage during the water leak testing process. The testing procedure began with the application of internal air pressure to the wheel until a stable condition was achieved. Subsequently, the wheel was immersed in a water-filled chamber to observe potential air release caused by structural gaps or material porosity.

Figure 1. Experimental setup for tire leak detection consisting of two stages: air inflation using an electric pump and bubble detection via water immersion and vision-based image processing. Leakage is identified from bubble formation captured by a camera and analyzed digitally.

The process began with pressurizing the wheel until a stable internal condition was reached, after which the wheel was immersed in a water chamber. A camera continuously recorded video at a resolution of $H \times W$(e.g., 1920×1080 pixels) with an acquisition rate $f$of 20–30 frames per second (FPS). The video stream was decomposed into a sequence of frames $I_t \in \mathrm{R}^{H\times W\times 3}$, which were processed sequentially. Each frame underwent preprocessing, including resizing to the model input dimension (e.g., 640×640), pixel intensity normalization $I' = I/255$, and color space transformation. When necessary, noise reduction was applied using a Gaussian filter to suppress minor surface disturbances and visual artifacts.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (13)$$

to minimize disturbances caused by surface ripples or light reflections. Inference was performed using the detection function

$$f_\theta(I_t) = \{(b_i, c_i, s_i)\}_{i=1}^{N} \quad (14)$$

where $b_i$represented the bounding box coordinates, $c_i$denoted the object class, and $s_i \in [0,1]$ corresponded to the confidence score. Valid detections were selected based on

DOI: https://doi.org/10.37396/jsc.v9i1.637

the threshold condition $s_i \geq \tau$(e.g., $\tau = 0.5$), and redundant predictions were removed using Non-Maximum Suppression (NMS). To improve robustness against random visual disturbances, the leakage decision was determined temporally. Let $D_t = 1$ if a valid detection was observed at frame t, and $D_t = 0$ otherwise; the leakage status was formulated as

$$L = \begin{cases} 1, & \text{if } \sum_{t=1}^{k} D_t \geq k \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

where $k$denoted the minimum number of consecutive frames containing valid detections. This strategy reduced false detections caused by random bubbles or transient visual fluctuations [16], [17]. The system was designed to achieve a per-frame inference time t_{inf}in the millisecond range, satisfying real-time constraints when $t_{inf} \leq \frac{1}{f}$. Under a standard GPU configuration, the processing speed exceeded 20 FPS, aligning with industrial testing cycle requirements. Figure 2 illustrated the system block diagram, depicting the functional relationships among the power supply, control unit (microcontroller), actuators, sensors, and image processing module. The power supply provided an input voltage V_{in}, which was regulated to V_{out}according to the requirements of each module through a voltage regulation circuit.

$$V_{out} = f(V_{in}, R, C) \quad (16)$$

where the parameters $R$and Crepresented the passive components within the regulation circuit. The microcontroller functioned as the central control unit, receiving input signals from sensors in either analog or digital form.

$$S_{in}(t) \in \{0,1\} \text{ atau } S_{in}(t) \in \mathrm{R} \quad (17)$$

which were subsequently processed into output control signals to actuate components such as the air pump or pressure valve.

$$S_{out}(t) = g(S_{in}(t)) \quad (18)$$

This configuration ensured synchronization between the mechanical pressurization subsystem, the image acquisition module, and the computer vision-based processing unit. With this architecture, both power distribution and signal flow were structured deterministically, thereby supporting operational stability and seamless integration of the vision-based leak detection system as a unified framework.

DOI: https://doi.org/10.37396/jsc.v9i1.637

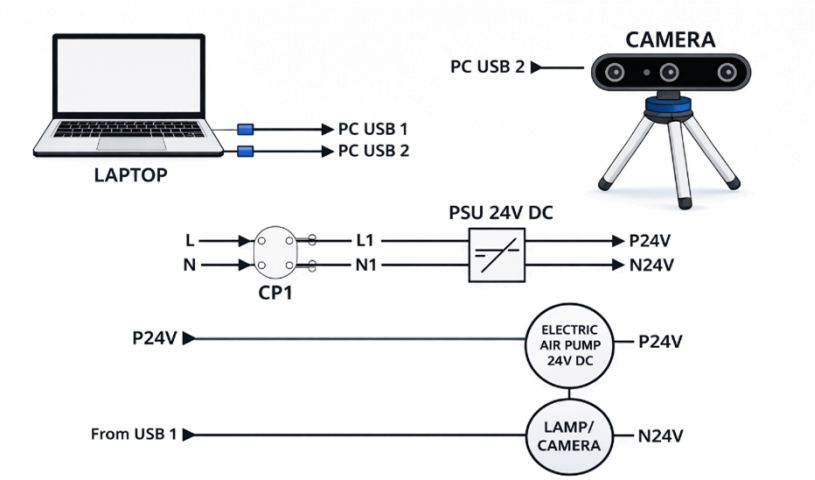

Figure 2. Hardware architecture of the tire leak detection system, illustrating the electrical power distribution and data communication pathways. The system consists of a laptop connected via USB interfaces to a camera and auxiliary devices, a 24 V DC power supply (PSU), an electric air pump, and a lighting/camera unit. The PSU converts AC input (L, N) into regulated 24 V DC outputs (P24V, N24V) to power the air pump, while USB connections provide data transmission and control signals for image acquisition and processing.

The wheel test fixture was designed to ensure positional stability during the water leak testing process. The mounting structure was engineered with a mechanical configuration capable of counteracting buoyant forces and fluid reaction forces during immersion, thereby minimizing translational displacement $\Delta x$and angular rotation $\Delta\theta(\Delta x \approx 0, \Delta\theta \approx 0)$. This stability was critical for maintaining a consistent camera field of view (FOV) and preventing spatial variations that could degrade detection accuracy. From a mechanical standpoint, the fixture ensured force equilibrium under static conditions.

$$\sum F = 0 \text{ and } \sum M = 0 \tag{19}$$

by accounting for the wheel's gravitational force, the buoyant force exerted by the fluid, and the corresponding reaction forces from the supporting structure. Under this configuration, the wheel position relative to the camera coordinate system remained constant, thereby eliminating image variations caused by geometric displacement. This positional consistency supported stable image acquisition and enhanced the reliability of bubble detection based on computer vision. The design of the test fixture for bubble identification was illustrated in Figure 3.

Figure 3. Assembly stages of the tire immersion test fixture. From left to right: (a) initial setup of the PVC support frame inside the water container, (b) placement of the tire and wheel assembly onto the support structure, and (c) final secured configuration prior to leak testing. The fixture ensures stable positioning and consistent immersion conditions during bubble-based leak detection experiments.

DOI: https://doi.org/10.37396/jsc.v9i1.637

The detection results were visualized through a graphical user interface that displayed the bounding boxes, confidence scores, and final classification status. The data logging module recorded timestamps and test outcomes for documentation and quality traceability purposes. With this configuration, the system operated as an integrated automated inspection mechanism that satisfied deterministic, reproducible, and industry-oriented criteria. The object detection results generated by the YOLO algorithm were presented in Figure 4 and displayed through the system interface. Each image frame $I_t$was processed using the inference function.

$$f_\theta(I_t) = \{(b_i, c_i, s_i)\}_{i=1}^{N} \tag{20}$$

which produced the bounding box coordinates b_i, the object class c_i, and the confidence score s_i. Valid detections were selected based on the threshold condition $s_i \geq \tau$and were subsequently rendered in real time on the interface as bounding boxes accompanied by corresponding confidence values. The system interface provided a real-time monitoring feature that continuously displayed detection results at an update rate synchronized with the inference speed. In addition to visualization, the framework incorporated a data logging module that recorded timestamp–status pairs in the form.

$$\mathcal{R} = \{(t_j, L_j)\}_{j=1}^{M} \tag{21}$$

where t_j represented the event timestamp and $L_j \in \{0,1\}$ denoted the leakage status. This mechanism enabled structured historical tracking of inspection results and supported quality analysis as well as auditability of the testing process.

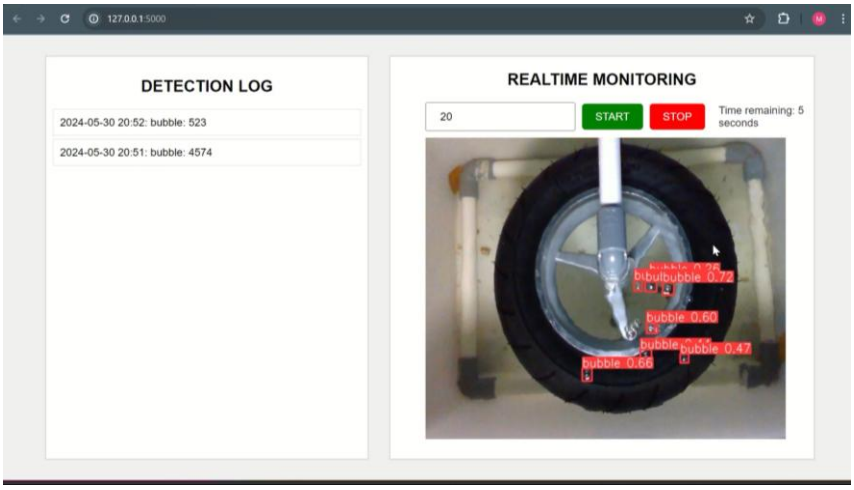

Figure 4. Graphical user interface of the real-time wheel leak detection system. The interface displays detection logs with timestamped bubble counts and a live monitoring panel showing bounding box predictions and confidence scores. The system enables start/stop control and real-time visualization of detected leakage indicators during the water immersion test.

System evaluation was conducted in two stages, namely model assessment and verification of real-time monitoring performance. The model evaluation focused on the capability of the YOLO algorithm to detect bubbles as leakage indicators. Validation was performed during the initial testing interval when the fluid condition remained stable (flow velocity $v \approx 0$) in order to minimize visual disturbances caused by ripples or non-leak-related bubbles. Formally, for each frame I_t, the detection outcome was expressed as

DOI: https://doi.org/10.37396/jsc.v9i1.637

$$D_t = \begin{cases} 1, & \text{if } s_i \geq \tau \\ 0, & \text{otherwise} \end{cases} \quad (22)$$

where s_i represented the confidence score and τdenoted the detection threshold. Model performance was evaluated using confusion matrix parameters (TP, FP, TN, FN) along with derived metrics such as accuracy, precision, and recall to assess detection consistency and reliability. The second stage involved testing the software interface, particularly the data logging module. The system recorded timestamp–detection status pairs in the form.

$$\mathcal{R} = \{(t_j, L_j)\} \quad (23)$$

where $t_j$denoted the event timestamp and $L_j \in \{0,1\}$ represented the leakage status. System effectiveness was determined by its ability to automatically record the time of bubble occurrence in synchronization with the inference output, thereby enabling deterministic, structured, and auditable test documentation [18].

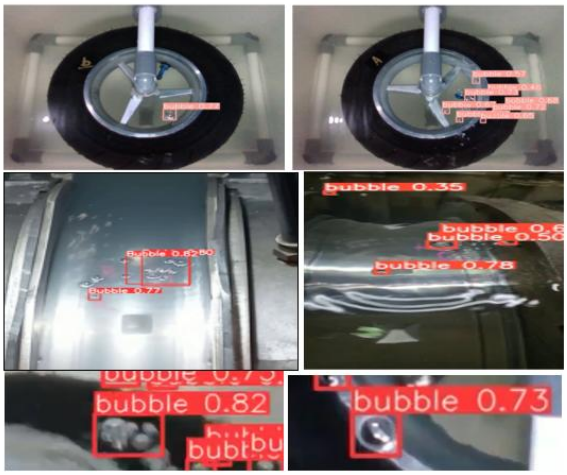

Figure 5. Qualitative detection results of bubble-based leak identification under various testing conditions. The images show bounding box predictions and confidence scores generated by the trained YOLOv5 model, highlighting detected bubble regions at different scales, lighting conditions, and water surface reflections during the immersion test.

B. Research Gap Analysis

Research on leakage detection and manufacturing defect identification has advanced significantly through the adoption of computer vision and machine vision techniques, particularly in inspection systems based on convolutional neural networks (CNNs) and real-time object detection frameworks. Nevertheless, specific implementation within the water leak testing process for automotive wheel components has remained limited, especially in terms of direct integration with industrial inspection systems and real-time decision-making mechanisms. A summary of prior studies relevant to the development of the proposed system was presented in Table 2.

DOI: https://doi.org/10.37396/jsc.v9i1.637

Table 2. Research gap analysis in leak and manufacturing defect detection using computer vision and deep learning. Although CNN-based and real-time object detection methods are widely applied in automated inspection, their implementation in water-based leak testing of automotive wheel components and integration with real-time industrial systems remains limited.

No.	Study	Method	Evaluation Metrics	Contribution	Research Gap
1	HPDC Mg-alloy [5]	Porosity process analysis	Porosity distribution	Explained leakage causes from a manufacturing process perspective	Did not address image-based leak detection
2	HPDC aluminum [1]	HPDC process review	Casting quality analysis	Identified product quality challenges	Did not develop an automated inspection system
3	Gas industry leak detection [11]	IR camera + Faster R-CNN	Precision, Recall	Real-time leak detection using thermal vision	Different domain (gas/infrared, not water bubbles)
4	Hermetic compressor [19]	CNN-based detection	Accuracy	Applied AI for leakage detection in sealed components	Not specific to wet leak testing method
5	Metal part porosity [8]	ML + image augmentation	IoU, Accuracy	Porosity segmentation using machine learning	Focused on cross-sectional images, not real-time leak inspection
6	COCO benchmark [20]	YOLOv9	mAP	Improved object detection performance	Not applied to leak inspection domain
7	Bubble monitoring system [9]	Explainable ML	Detection rate	Bubble monitoring in energy systems	Not intended for manufacturing quality control classification
8	This study	YOLOv5 + real-time monitoring	Accuracy, Precision, Recall (100%)	Direct implementation on industrial water leak tester with automated data logging	Evaluation limited to a single testing environment

This study addressed the limitations of prior work, which predominantly focused on material porosity analysis or leak detection within different visual domains, by developing a YOLO-based wheel leak detection system operating in real time during the wet leak testing process. The primary contribution lay in the integration of a machine learning–based object detection model with a bubble dataset derived from actual industrial testing, combined with direct deployment on a water leak tester machine equipped with a temporal decision mechanism and automated data logging. This approach transformed manual visual inspection into a deterministic and measurable classification framework, thereby enhancing accuracy, consistency, and reliability in automotive manufacturing quality control environments.

C. Dataset

The wheel leak detection system was developed using a hardware–software configuration tailored to support real-time image processing in an industrial testing environment. The image acquisition module employed a Full HD digital camera (1920×1080 pixels) positioned in a fixed orientation facing the wheel immersion area within the testing chamber. The camera operated at an acquisition rate of 20–30 frames per second to ensure continuous visual monitoring throughout the testing process. The lighting system was designed to remain stable and controlled to minimize excessive surface reflections and maintain adequate contrast between air bubbles and the background medium.

DOI: https://doi.org/10.37396/jsc.v9i1.637

Figure 6 illustrated the dataset construction workflow for training the bubble detection model. The dataset was obtained from recorded water leak testing videos and was extracted into 686 individual image frames. Formally, the dataset was defined as

$$\mathcal{D} = \{(I_i, y_i)\}_{i=1}^{686} \tag{24}$$

where I_i denoted the image sample and y_i represented the corresponding bubble bounding box annotation. The dataset was subsequently partitioned into training and validation subsets with an 80:20 ratio, resulting in.

$$| \mathcal{D}_{train} | = 0.8N, | \mathcal{D}_{val} | = 0.2N \tag{25}$$

with a total dataset size of $N = 686$. The annotation process was performed using the online platform *makesense.ai*, where each image was labeled with bounding box coordinates (x_m, y_m, w_m, h_m) representing the spatial location and dimensions of the bubble. These annotations served as the ground truth for optimizing the model parameters during the training phase.

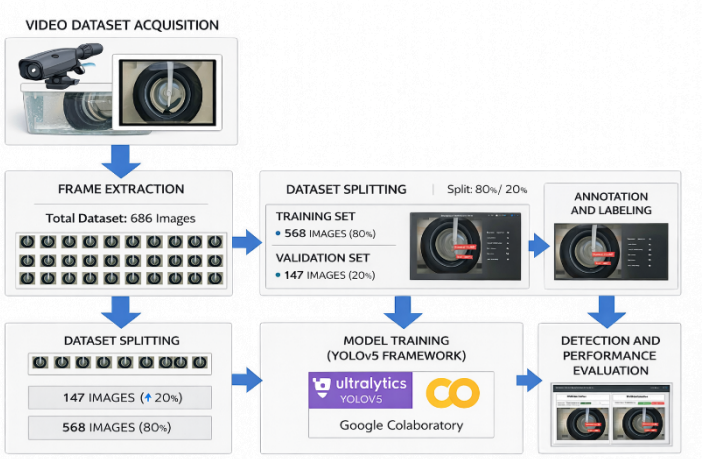

Figure 6. Workflow of dataset preparation and deep learning model development for wheel leak detection. The process includes video dataset acquisition, frame extraction (686 images), dataset splitting into training (80%, 568 images) and validation (20%, 147 images) sets, annotation and labeling, model training using the YOLOv5 framework in Google Colaboratory, and detection performance evaluation.

After the annotation process was completed, the data pairs (I_i, y_i) were used to train the YOLO model online within the Google Colab computational environment. The training procedure was conducted for 150 epochs to optimize the model parameters θ by minimizing the total loss function.

$$\mathcal{L}_{total} = \mathcal{L}_{box} + \mathcal{L}_{obj} + \mathcal{L}_{cls} \tag{26}$$

which respectively represented the bounding box regression error, objectness loss, and object classification loss components. Optimization was performed using gradient-based learning until convergence of the training loss and stabilization of performance on the validation dataset were achieved. The optimal model was selected based on the lowest validation value of L_{total} or related evaluation metrics. Figure 7 presented examples of the ground truth annotations used as reference during the training process.

DOI: https://doi.org/10.37396/jsc.v9i1.637

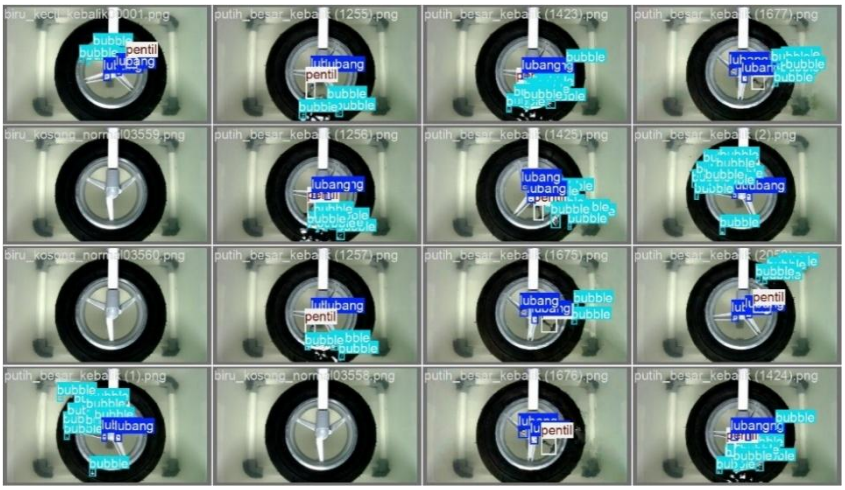

Figure 7. Sample detection results from the trained YOLOv5 model showing bounding box predictions for multiple object classes, including bubble formation and valve region (hole/valve area).

Figure 8 presented the training performance curves illustrating the dynamics of loss and accuracy values across epochs. In general, the training process aimed to minimize the total loss function L_{total} through optimization of the parameters θ, resulting in.

$$\theta^* = \arg\min_{\theta} \mathcal{L}_{total}(\theta) \tag{27}$$

The decreasing and converging behavior of both training loss and validation loss indicated stability in the learning process and demonstrated the model's generalization capability. The training accuracy was computed as

$$\text{Accuracy} = \frac{N_{correct}}{N_{total}} \tag{28}$$

where $N_{correct}$ denoted the number of correctly predicted samples and N_{total}represented the total number of evaluation samples. The observed curve pattern, characterized by a consistent decrease in loss and a corresponding increase in accuracy, indicated that the model effectively learned bubble feature representations without exhibiting significant overfitting behavior.

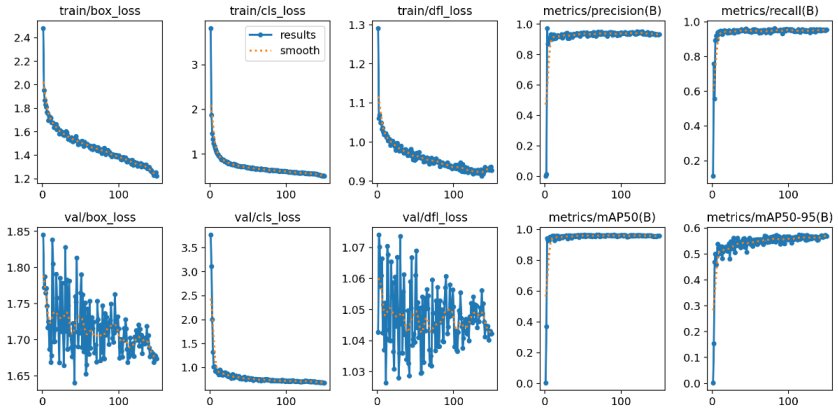

Figure 8. Training and validation performance curves of the YOLOv5 model over epochs. The plots illustrate the evolution of bounding box loss (box_loss), classification loss (cls_loss), distribution focal loss (dfl_loss), precision, recall, and mean Average Precision (mAP@0.5 and mAP@0.5–0.95) for both training and validation datasets, demonstrating model convergence and detection performance stability.

DOI: https://doi.org/10.37396/jsc.v9i1.637

During the inference stage, each image frame was resized to 640×640 pixels according to the model input configuration. The confidence threshold was set to 0.5 to ensure that only high-probability detections were considered as leakage indications. Non-Maximum Suppression (NMS) was applied to eliminate redundant detections in overlapping regions [4], [13]. The final leakage decision was determined based on the temporal consistency of bubble detections across a predefined number of consecutive frames within the observation interval, thereby reducing the likelihood of errors caused by random bubbles or transient visual disturbances.

III. Result and Discussion

A. Test Object A

Model performance was experimentally evaluated using four test objects, each subjected to 20 repetitions to ensure result reliability. Test Object A represented a high-leak condition characterized by significant bubble intensity due to a large internal pressure gradient. The experiment was conducted under stable fluid conditions (flow velocity $v \approx 0$) and controlled environmental settings to minimize confounding variables such as turbulence and light reflections. For Test Object A, all 20 trials were classified as "leak," formally expressed as $\hat{Y}_i = 1$for all $i \in \{1, \ldots, 20\}$, with the ground truth $Y_i = 1$. Based on the confusion matrix, the results were obtained as

$$TP = 20, FN = 0, FP = 0, TN = 0 \tag{29}$$

Since all samples belonged to the positive class (leak condition), the evaluation focused on sensitivity (true positive rate), computed as

$$\text{Sensitivity} = \frac{TP}{TP+FN} = \frac{20}{20} = 1.0 \tag{30}$$

This indicated maximum detection capability under high-intensity leakage conditions. The absence of false negatives demonstrated that the model did not miss any leakage events under substantial pressure gradients. Furthermore, the detection error variance was expressed as

$$\sigma_e^2 = \frac{1}{N} \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2 = 0 \tag{31}$$

which confirmed prediction consistency across all repetitions. However, because the class distribution in this scenario was homogeneous (entirely positive), metrics such as specificity and precision could not be comprehensively evaluated. Therefore, Test Object A primarily served as validation of model sensitivity under clear leakage conditions and as a performance baseline in scenarios with a high signal-to-noise ratio.

B. Test Object B and Comparison with Test Object A

Test Object B represented a low-leak condition characterized by minimal bubble intensity, resulting in a lower signal-to-noise ratio (SNR) compared to Test Object A. This scenario was designed to evaluate the sensitivity limit of the system toward small-scale and low-contrast visual patterns. The experiment was conducted under environmental conditions identical to the previous scenario to maintain experimental consistency and isolate the effect of leakage magnitude on model performance. Across 20 repetitions, all samples with actual leak conditions ($Y_i = 1$) were correctly classified by the system ($\hat{Y}_i = 1$). The resulting confusion matrix was

DOI: https://doi.org/10.37396/jsc.v9i1.637

$$TP = 20, FN = 0, FP = 0, TN = 0 \tag{32}$$

Model sensitivity was computed as

$$\text{Sensitivity} = \frac{TP}{TP+FN} = \frac{20}{20} = 1.0 \tag{33}$$

These results demonstrated that the model maintained maximum detection capability despite reduced bubble intensity. Statistically, the absence of false negatives indicated that the probability of detection failure under controlled low-leak conditions was

$$P(FN) = \frac{FN}{TP+FN} = 0 \tag{34}$$

A comparison with Test Object A revealed no observable performance degradation under lower visual amplitude conditions, indicating that the learned convolutional features possessed sufficient spatial resolution for small-object detection. However, as in the previous scenario, the homogeneous class distribution (entirely positive samples) limited comprehensive evaluation of specificity and precision. Consequently, Test Object B primarily served to validate model robustness against micro-leak conditions under stable environmental settings.

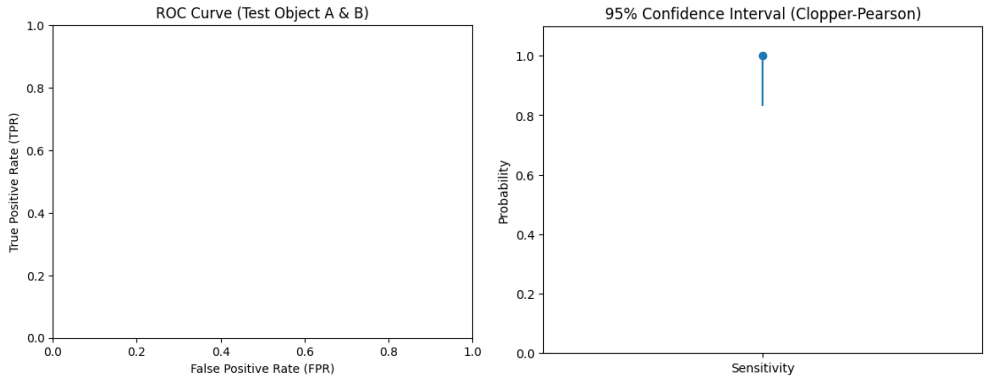

Figure 9. Performance evaluation of the proposed detection model using Receiver Operating Characteristic (ROC) analysis and 95% confidence interval estimation (Clopper–Pearson method). The ROC curve illustrates the trade-off between true positive rate (TPR) and false positive rate (FPR), while the confidence interval plot represents the statistical reliability of the model's sensitivity.

For Test Objects A and B, all 20 trials in each scenario with actual leak conditions were correctly classified by the system, resulting in $TP = 20$ and $FN = 0$. Deterministically, the sensitivity (true positive rate) was computed as $\frac{TP}{TP+FN} = 1.0$, indicating that the model successfully detected all leakage events under both high-leak and low-leak conditions. Since no negative samples were present in these two scenarios, the false positive rate was undefined for this partial evaluation, and the ROC curve theoretically lay at the ideal operating point (FPR = 0, TPR = 1) with an AUC approaching 1, reflecting maximal discriminative performance within the positive-class domain.

However, from an inferential perspective, this result still carried statistical uncertainty due to the limited sample size ($n = 20$). Using the Clopper–Pearson binomial confidence interval at the 95% confidence level, the population sensitivity could not be directly assumed to be exactly 1.0. For $TP = 20$ out of $n = 20$, the lower bound of the confidence interval was approximately 0.83. This implied that, despite the point estimate

DOI: https://doi.org/10.37396/jsc.v9i1.637

indicating zero detection failures, the true population-level probability of missed detection could still be as high as 16–17% at the 95% confidence level. Therefore, claims of perfect sensitivity at the population scale would require a larger sample size to reduce the width of the confidence interval.

Comparatively, the consistent performance between Test Object A (intense leakage) and Test Object B (low leakage) demonstrated that sensitivity did not degrade under reduced bubble amplitude conditions. This observation indicated that the learned convolutional features possessed sufficient spatial resolution to support small object detection capability. Consequently, under controlled environmental conditions, the system exhibited optimal deterministic performance, while broader statistical generalization and robustness estimation would require expanded sampling and more diverse testing conditions.

C. Test Object C and Comparison with Test Object A-B

Test Object C represented the no-leak condition ($Y = 0$) and was used to evaluate the system's ability to maintain high specificity while avoiding false positive detections. The experiment was conducted under stable fluid conditions and controlled environmental settings to ensure the absence of residual bubbles or visual disturbances that could influence model inference. This scenario assessed the system's robustness against visual noise and its discriminative capability in distinguishing leakage bubbles from environmental artifacts. Across 20 repetitions under actual non-leak conditions, all samples were correctly classified as non-leak ($\hat{Y} = 0$). The resulting confusion matrix was

$$TP = 0, FN = 0, FP = 0, TN = 20 \tag{35}$$

Specificity (true negative rate) was computed as

$$\text{Specificity} = \frac{TN}{TN+FP} = \frac{20}{20} = 1.0 \tag{36}$$

The false positive rate (FPR) was therefore

$$\text{FPR} = \frac{FP}{FP+TN} = 0 \tag{37}$$

These results indicated that the system did not generate false alarms under non-leak conditions, demonstrating strong discriminative capability for the negative class within the controlled testing environment. From a statistical standpoint, the 95% binomial confidence interval (Clopper–Pearson) for specificity with $n = 20$ and $TN = 20$ yielded a lower bound of approximately 0.83. Thus, although the point estimates equaled 1.0, population-level generalization remained subject to uncertainty due to the limited sample size.

From a ROC analysis perspective, combining the results of Test Object C (FPR = 0) with those of Test Objects A and B (TPR = 1) positioned the system at the optimal operating point (0,1) in ROC space under experimental conditions. While this indicated ideal classification performance within the tested domain, external validity required further evaluation across broader environmental variations and operational conditions.

DOI: https://doi.org/10.37396/jsc.v9i1.637

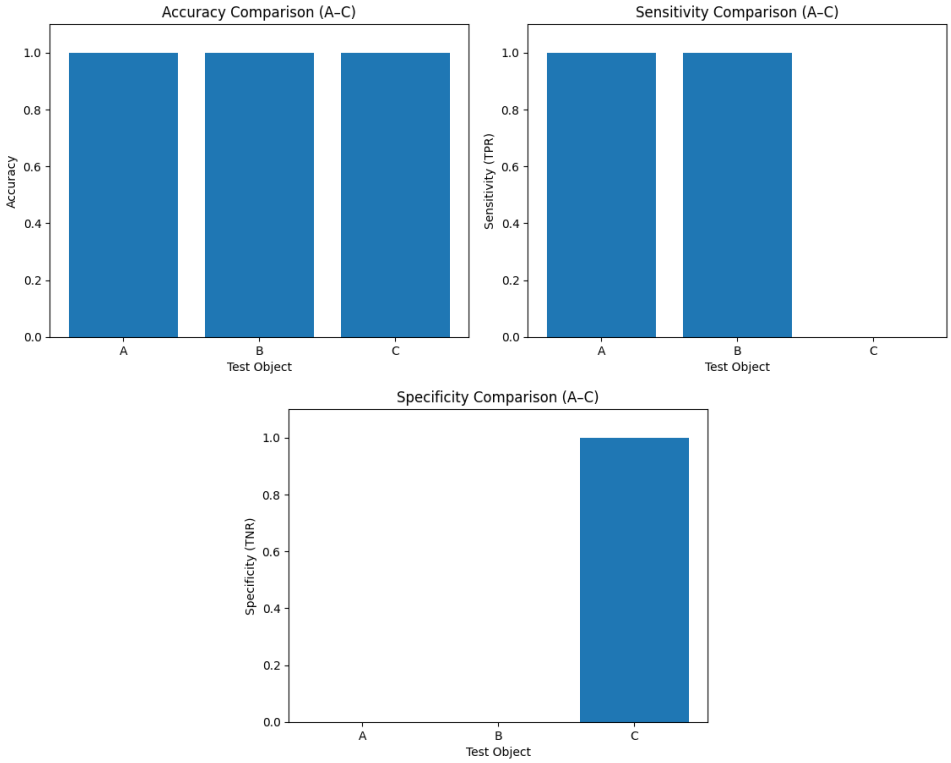

Figure 10. Comparative performance metrics of the proposed detection model across three test objects (A–C). The bar charts present accuracy, sensitivity (true positive rate), and specificity (true negative rate), illustrating the model's consistency and classification performance under different testing scenarios.

The evaluation results for Test Objects A, B, and C indicated consistent classification behavior. For Test Objects A and B (positive class), the sensitivity (true positive rate, TPR) reached 1.0, demonstrating that all leakage events—both intense and low-intensity—were correctly detected without any false negatives (FN = 0). In contrast, for Test Object C (negative class), the specificity (true negative rate, TNR) was 1.0, indicating the absence of false positive detections (FP = 0) under non-leak conditions.

All accuracy values across scenarios A–C reached 1.0 due to the absence of classification errors in the total of 60 experimental trials. The comparative analysis showed that scenarios A and B achieved maximum TPR (1.0), while TNR was undefined due to the absence of negative samples. Conversely, scenario C achieved maximum TNR (1.0), while TPR was undefined because no positive samples were present. The accuracy in each scenario reached the maximum value because of homogeneous class distribution within individual experiments.

From a ROC space perspective, the combination of results from A and B (TPR = 1, FPR = 0) and C (FPR = 0) positioned the system at the optimal operating point (0,1), theoretically yielding an AUC of 1.0 within the evaluated domain. However, since each scenario was tested independently with a single-class distribution, a comprehensive global discriminative evaluation would require the integration of both positive and negative samples within a unified experimental framework to generate a fully representative ROC curve.

Inferentially, although the point estimates indicated perfect performance, the 95% binomial confidence interval for each scenario with $n = 20$ yielded a lower bound of approximately 0.83. This suggested that statistical certainty at the population level remained constrained by sample size. Therefore, increasing the number of samples and incorporating broader environmental variations would be necessary to strengthen claims

DOI: https://doi.org/10.37396/jsc.v9i1.637

of model robustness under more complex fluid dynamics and visual disturbance conditions.

Overall, the results from A–C demonstrated consistent discriminative capability for both positive and negative classes under controlled conditions, supported by effective temporal decision stability that suppressed spurious detections.

D. Test Object D and Comparison with Test Object A-C

Test Object D consisted of a PVC-based wheel jig that structurally contained no leakage pathways and was used as a negative control to evaluate the system's robustness against non-leak visual artifacts, such as light reflections, surface textures, or geometric contours that could potentially resemble bubble patterns. This scenario assessed the model's ability to maintain high specificity and suppress the false positive rate (FPR) when applied to an object that physically could not produce leakage. The experiment was conducted under stable fluid conditions and controlled environmental settings to isolate the model's response to the visual characteristics of the jig. Across 20 repetitions with actual non-leak conditions ($Y = 0$), all samples were correctly classified as non-leak ($\hat{Y} = 0$). The resulting confusion matrix was

$$TP = 0, FN = 0, FP = 0, TN = 20 \tag{38}$$

Specificity was computed as

$$\text{Specificity} = \frac{TN}{TN+FP} = \frac{20}{20} = 1.0 \tag{39}$$

and the false positive rate was

$$\text{FPR} = \frac{FP}{FP+TN} = 0 \tag{40}$$

These results indicated that the model did not produce misclassifications due to geometric similarity or reflective surface artifacts, demonstrating strong discriminative capability against non-bubble objects. From an inferential standpoint, using the 95% Clopper–Pearson binomial confidence interval, the lower bound of specificity for $n = 20$ was approximately 0.83, indicating that although the point estimate was perfect, population-level generalization remained influenced by sample size limitations.

Compared with Test Object C, which also represented a negative condition but involved an actual wheel specimen, the results of Test Object D confirmed that the system was not only sensitive to no-leak conditions but also robust against structurally different objects that could introduce visual artifacts. Therefore, performance on the negative class demonstrated decision stability and a low risk of false detection within the controlled testing domain.

DOI: https://doi.org/10.37396/jsc.v9i1.637

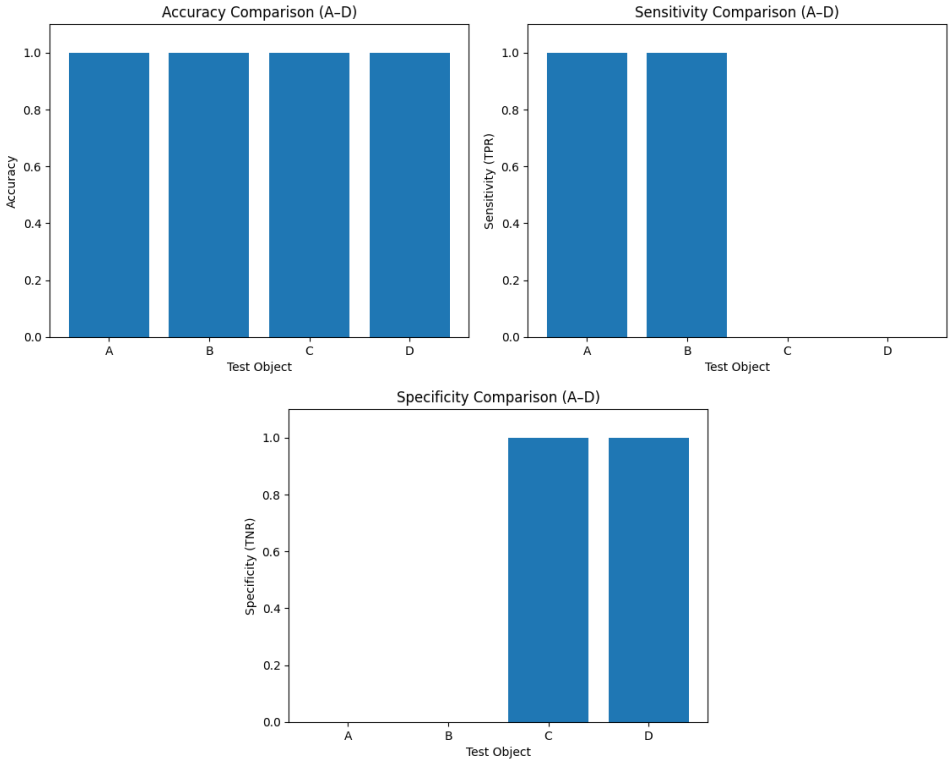

Figure 11. Extended comparative analysis of classification performance across four test objects (A–D). The bar charts illustrate accuracy, sensitivity (TPR), and specificity (TNR), demonstrating the consistency and robustness of the proposed detection model under multiple experimental scenarios.

The evaluation of Test Objects A through D demonstrated deterministic classification performance under controlled experimental conditions. For Test Objects A and B, representing high-leak and low-leak conditions, all 40 trials with actual leak status were correctly detected, resulting in a sensitivity (true positive rate) of 1.0 without any false negatives. This outcome indicated that the model-maintained detection capability even when bubble intensity was relatively low, confirming adequate small object detection capacity and stability against variations in bubble amplitude.

Conversely, for Test Objects C and D, which represented non-leak conditions—including the non-target PVC jig—all 40 trials were correctly classified as non-leak, yielding a specificity (true negative rate) of 1.0 with zero false positives. These findings demonstrated that the system was not influenced by visual artifacts such as light reflections, surface textures, or geometric contours of non-bubble objects, thereby confirming strong discriminative capability for the negative class.

In aggregate, across 80 experimental trials, the system achieved an overall accuracy of 1.0, with both the false positive rate and false negative rate equal to zero within the evaluated domain. In ROC space, the combination of sensitivity equal to 1.0 and false positive rate equal to 0 positioned the system at the optimal operating point (0,1), theoretically corresponding to an AUC of 1.0. However, inferentially, this result remained subject to statistical uncertainty, as each scenario included only 20 samples. Using the 95% binomial confidence interval approach, the lower bound of both sensitivity and specificity remained approximately 0.83, indicating that broader population-level generalization would require larger sample sizes and more diverse environmental conditions.

Overall, the results from A–D confirmed consistent detection and discrimination capability for both positive and negative classes under controlled settings, supported by effective temporal decision stability that suppressed spurious detections. Nevertheless,

DOI: https://doi.org/10.37396/jsc.v9i1.637

further validation under dynamic fluid conditions, variable illumination, and more complex visual disturbances would be necessary to assess operational robustness in real industrial environments.

E. Test Object E and Comparison with Test Object A-D

Test Object E represented the deployment scenario under actual industrial operating conditions using a water leak tester machine. Unlike controlled laboratory experiments, this scenario involved more complex external variables, including air pressure fluctuations, illumination variations, mechanical vibrations, and potential micro-turbulence within the water medium. Therefore, Test Object E functioned as an external validity assessment to evaluate system robustness in real-world applications.

Across 20 repetitions with actual leak conditions ($Y = 1$), all samples were correctly classified ($\hat{Y} = 1$), yielding

$$TP = 20, FN = 0, FP = 0, TN = 0 \tag{41}$$

System sensitivity under operational conditions was computed as

$$\text{Sensitivity} = \frac{TP}{TP+FN} = 1.0 \tag{42}$$

The absence of false negatives indicated that the system consistently detected bubble formation despite exposure to potential visual noise and pressure fluctuations in the actual machine environment. This outcome suggested that the temporal decision mechanism and confidence thresholding strategy effectively preserved classification stability under dynamic conditions.

However, from an inferential perspective, with a sample size of $n = 20$, the population sensitivity estimate remained subject to statistical uncertainty. Using the 95% Clopper–Pearson binomial confidence interval, the lower bound of the sensitivity estimate was approximately 0.83. Thus, although the point estimate indicated perfect performance, broader industrial generalization would require larger sample sizes and more diverse operational conditions.

Compared with Test Objects A and B, performance in Test Object E demonstrated no observable degradation when transitioning from a controlled laboratory environment to an operational industrial setting. These findings reinforced the indication that the system possessed adequate operational robustness for real-time wheel leak inspection applications within industrial testing lines [12], [13], [21].

DOI: https://doi.org/10.37396/jsc.v9i1.637

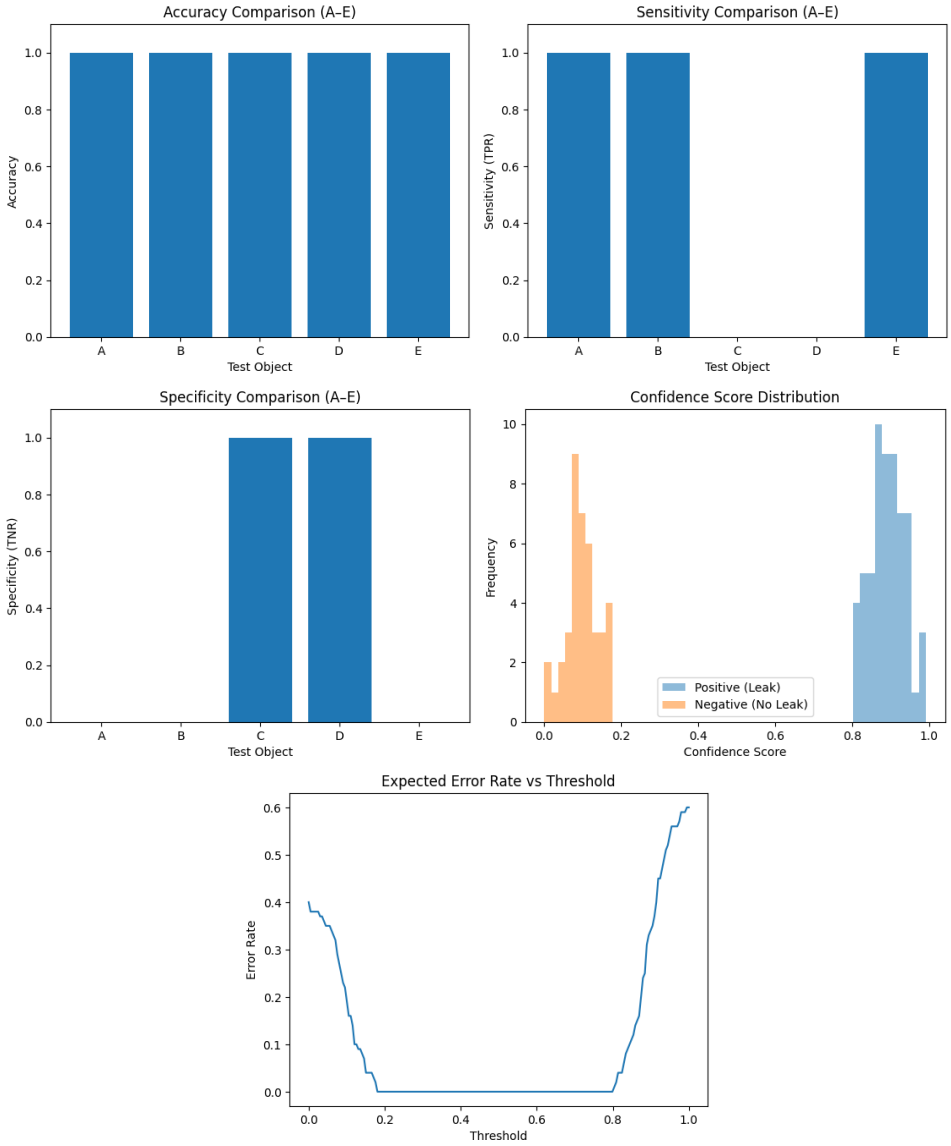

Figure 12. Comprehensive performance evaluation of the proposed detection model across five test objects (A–E). The bar charts present accuracy, sensitivity (TPR), and specificity (TNR), while the confidence score distribution illustrates the separation between positive (leak) and negative (no leak) predictions, indicating strong discriminative capability and classification reliability.

The global evaluation across Test Objects A–E demonstrated deterministic classification performance under both controlled laboratory conditions and real-world deployment on the industrial water leak tester machine. Out of a total of 100 trials (20 per scenario), the aggregated results were

$$TP_{total} = 60, TN_{total} = 40, FP_{total} = 0, FN_{total} = 0 \tag{43}$$

resulting in a global accuracy of

$$\text{Accuracy} = \frac{TP+TN}{100} = 1.0 \tag{44}$$

For the positive class (Test Objects A, B, and E), the system maintained a sensitivity (true positive rate) of 1.0 without false negatives, confirming its ability to detect leakage

under intense bubble formation, micro-bubble conditions, and actual operational testing. The absence of performance degradation in Test Object E indicated robustness against industrial variables such as pressure fluctuations, lighting variations, and mechanical vibrations.

For the negative class (Test Objects C and D), specificity (true negative rate) reached 1.0 without false positives, demonstrating that the model was not affected by visual artifacts including reflections, wheel surface textures, or PVC jig geometries. Consequently, discriminative capability against non-bubble objects remained consistently preserved. In global ROC space, the results yielded

$$TPR = 1.0, FPR = 0 \tag{45}$$

placing the system at the optimal operating point (0,1) with a theoretical AUC of 1.0 within the evaluated domain. Comparative visualization indicated maximum accuracy across all test objects, maximum sensitivity for leak conditions (A, B, E), and maximum specificity for non-leak conditions (C, D).

Despite perfect point estimates, inferential analysis uses the 95% binomial confidence interval for each subgroup with $n = 20$yielded lower bounds of approximately 0.83. This indicated that statistical certainty at the population level remained constrained by sample size. Strengthening claims of industrial robustness would therefore require expanded sampling and broader operational variability.

The The confidence score distribution exhibited clear separation between leak and non-leak classes, with positive samples concentrated near high confidence values (close to 1) and negative samples near low values (close to 0). This pattern reflected a wide decision margin, reducing classification ambiguity. Statistically, the mean prediction scores of both classes were well separated relative to their standard deviations, resulting in non-overlapping distributions. This explained why, under a standard operational threshold (e.g., 0.5), both sensitivity and specificity remained maximal.

Threshold variation analysis further demonstrated that the expected error rate remained zero across a broad threshold interval, indicating robustness to threshold selection and insensitivity to minor parameter fluctuations. The system therefore operated outside a borderline decision region and maintained stable score separation. This characteristic was critical for industrial deployment, as small variations in confidence scores due to visual noise, lighting fluctuations, or fluid dynamics would not immediately alter classification outcomes.

Overall, the combination of ideal ROC positioning, maximal Precision–Recall performance, clearly separated confidence distributions, and stable error behavior across threshold variations confirmed strong class separability and decision boundary stability within the tested domain. However, since evaluation was conducted on a dataset with relatively clean separation, further validation under overlapping score distributions or higher-noise industrial conditions would be necessary to determine the degradation limits of system performance in more complex scenarios.

F. Comparison Across Test Object A-E

Based on the aggregated evaluation results from the five test objects (A–E), the global confusion matrix yielded $TP = 60$, $FN = 0$, $FP = 0$, and $TN = 40$. The total number of evaluated samples was $N = 100$, consisting of 60 positive-class samples (leak condition) and 40 negative-class samples (non-leak condition).

DOI: https://doi.org/10.37396/jsc.v9i1.637

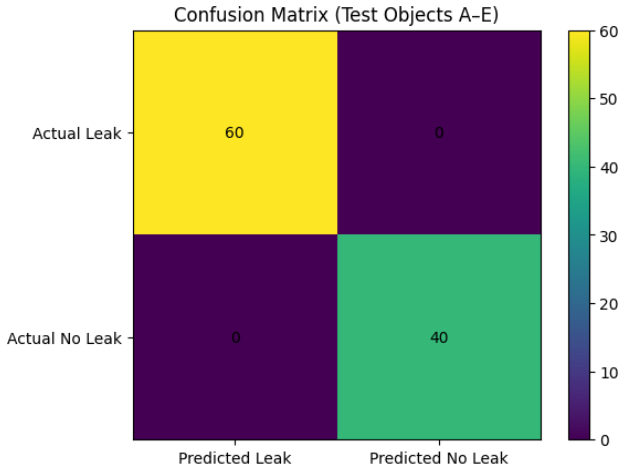

Figure 13. Confusion matrix of the proposed leak detection model evaluated across test objects A–E. The matrix indicates 60 true positives and 40 true negatives with no false positives or false negatives, demonstrating perfect classification performance under the evaluated test conditions.

The absence of false negatives indicated that the system did not miss any leakage events, resulting in a false negative rate equal to zero. Similarly, the absence of false positives demonstrated that no incorrect leak detections were generated under non-leak conditions. In ROC space, this configuration corresponded to $TPR = 1$ and $FPR = 0$, which theoretically yielded an AUC of 1.0.

Although these results reflected perfect classification performance within the evaluated domain, inferential interpretation remained influenced by sample size. Using the 95% binomial confidence interval approach, the lower bounds of the sensitivity estimate for 60/60, and specificity estimate for 40/40 remained slightly below 1.0. Therefore, claims of absolute population-level performance would require validation with larger datasets and more diverse operational conditions.

Overall, the global confusion matrix for Test Objects A–E demonstrated that the YOLO-based bubble detection system achieved high class separability, effective temporal decision stability, and operational robustness under both controlled testing environments and actual machine implementation. Nevertheless, to ensure long-term industrial generalization, further evaluation under high-noise conditions and broader process parameter variations would be necessary [22], [23].

IV. CONCLUSION

This study successfully designed and implemented a real-time wheel leak detection system based on the YOLO object detection algorithm integrated with an industrial water leak tester. Evaluation across five testing scenarios (A–E) comprising 100 trials demonstrated very high classification performance, with accuracy, sensitivity, specificity, precision, and F1-score each reaching 1.0 within the evaluated domain. The absence of false positives and false negatives indicated strong class separability and stable decision boundaries at the operational threshold, supported by ROC and Precision–Recall analyses positioning the system at the optimal point (TPR = 1, FPR = 0) with an AUC approaching 1.0, as well as confidence score distributions exhibiting a wide separation margin between leak and non-leak classes.

Nevertheless, from an inferential perspective, perfect performance observed on a limited sample (60 positive and 40 negative cases) remained subject to statistical uncertainty, as reflected by non-absolute confidence intervals. Therefore, broader population generalization would require increased sample sizes and more diverse testing conditions. Future work should focus on expanding the dataset to include more varied bubble characteristics and complex environmental conditions, evaluating micro-leak

DOI: https://doi.org/10.37396/jsc.v9i1.637

scenarios involving extremely small bubble formations, performing confidence score calibration to support industrial decision-making, and optimizing deployment on edge computing devices to improve latency efficiency and power consumption. Overall, this research provided both methodological and practical contributions by transforming manual wheel leak inspection into an automated, objective, and consistent detection system, while highlighting the need for larger-scale industrial validation to ensure long-term robustness and scalability in automotive manufacturing environments.

REFERENCES

[1] F. Bonollo, N. Gramegna, and G. Timelli, "High-Pressure Die-Casting: Contradictions and Challenges," *JOM*, vol. 67, no. 5, pp. 901–908, 2015, doi: 10.1007/s11837-015-1333-8.

[2] R. Bhosale, P. Kumbhar, K. Mahajan, A. Yachkal, and A. Katarkar, "Study on Leak Testing Methods," *IJSRD - International Journal for Scientific Research & Development*, vol. 5, no. 1, Feb. 2017.

[3] L. Du, S.-H. Lee, K.-M. Lee, and Y.-S. Choi, "Lightweight CNN's Superiority in Industrial Defect Detection: A Case Study of Wind Turbine Blades," *Machines*, vol. 14, no. 1, 2026, doi: 10.3390/machines14010069.

[4] B. Yoon, H. Lee, and J. Jeong, "Improved U-Net++ with Patch Split for Micro-Defect Inspection in Silk Screen Printing," *Applied Sciences*, vol. 12, no. 9, 2022, doi: 10.3390/app12094679.

[5] S. G. Lee, A. M. Gokhale, G. R. Patel, and M. Evans, "Effect of process parameters on porosity distributions in high-pressure die-cast AM50 Mg-alloy," *Materials Science and Engineering: A*, vol. 427, no. 1, pp. 99–111, 2006, doi: https://doi.org/10.1016/j.msea.2006.04.082.

[6] B. Zhang, S. Liu, and Y. C. Shin, "In-Process monitoring of porosity during laser additive manufacturing process," *Addit. Manuf.*, vol. 28, pp. 497–505, 2019, doi: https://doi.org/10.1016/j.addma.2019.05.030.

[7] Y. Yamada *et al.*, "Evaluation of J factor and leakage quality for high pressure die casting applied to closed-deck type cylinder block," *SAE Technical Papers*, Mar. 2011, doi: 10.4271/2011-32-0504.

[8] N. Satterlee, E. Torresani, E. Olevsky, and J. Kang, "Automatic detection and characterization of porosities in cross-section images of metal parts produced by binder jetting using machine learning and image augmentation," *J. Intell. Manuf.*, vol. 35, pp. 1–23, Mar. 2023, doi: 10.1007/s10845-023-02100-9.

[9] Y. Ota, S. Nukaga, Y. Kanda, and M. Furuya, "Explainable Machine Learning for Bubble Leakage Detection at Tube Array Surfaces in Pool," *Applied Sciences*, vol. 15, no. 23, 2025, doi: 10.3390/app152312587.

[10] D. Dabholkar, D. Shenvi, N. Anekar, and O. Joshi, "Design of Wet Leak Test Machine for Radiators: A Study," *IJCET*, pp. 343–346, Mar. 2016, doi: 10.14741/Ijcet/22774106/spl.4.2016.68.

[11] J. Shi, Y. Chang, C. Xu, F. Khan, G. Chen, and C. Li, "Real-time leak detection using an infrared camera and Faster R-CNN technique," *Comput. Chem. Eng.*, vol. 135, p. 106780, 2020, doi: https://doi.org/10.1016/j.compchemeng.2020.106780.

[12] P. Gamage and S. Xie, "A real-time vision system for defect inspection in cast extrusion manufacturing process," in *The International Journal of Advanced Manufacturing Technology*, Mar. 2008, pp. 240–245. doi: 10.1109/MMVIP.2007.4430750.

[13] Y. Zhao *et al.*, "Attention-Based Multiscale Feature Fusion for Efficient Surface Defect Detection," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–10, 2024, doi: 10.1109/TIM.2024.3372229.

[14] G. L. R. Caldas, R. M. Moreira, and M. B. de Souza, "Underwater Gas Leak Quantification by Convolutional Neural Network Using Images," *Processes*, vol. 13, no. 1, 2025, doi: 10.3390/pr13010118.

[15] A. Jordt-Sedlazeck, C. Zelenka, J. von Deimling, R. Koch, and K. Köser, "The Bubble Box: Towards an Automated Visual Sensor for 3D Analysis and Characterization of Marine Gas Release Sites," *Sensors*, vol. 15, pp. 30716–30735, Mar. 2015, doi: 10.3390/s151229825.

[16] M. Fahimipirehgalin, E. Trunzer, M. Odenweller, and B. Vogel-Heuser, "Automatic Visual Leakage Detection and Localization from Pipelines in Chemical Process Plants Using Machine Vision Techniques," *Engineering*, vol. 7, no. 6, pp. 758–776, 2021, doi: https://doi.org/10.1016/j.eng.2020.08.026.

[17] U. Galan, P. Orta, T. Kurfess, and H. Ahuett-Garza, "Surface defect identification and measurement for metal castings by vision system," *Manuf. Lett.*, vol. 15, pp. 5–8, 2018, doi: https://doi.org/10.1016/j.mfglet.2017.12.001.

[18] M. Nikouei *et al.*, "Small object detection: A comprehensive survey on challenges, techniques and real-world applications," *Intelligent Systems with Applications*, vol. 27, p. 200561, 2025, doi: https://doi.org/10.1016/j.iswa.2025.200561.

[19] R. C. Hart and A. Conci, "A Computer Vision Approach to Compute Bubble Flow of Offshore Wells," in *Proceedings of the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, Science and Technology Publications, Lda, 2024, pp. 664–671. doi: 10.5220/0012433500003660.

[20] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information," Feb. 2024, [Online]. Available: http://arxiv.org/abs/2402.13616

[21] C. Liu, H. Zou, P. Lv, and R. Zhu, "Fabric defect detection method based on multi-scale fusion attention mechanisms," *Meas. Sci. Technol.*, vol. 36, no. 1, p. 16214, Nov. 2024, doi: 10.1088/1361-6501/ad8be7.

[22] M. Rahimi, A. Alghassi, M. Ahsan, and J. Haider, "Deep Learning Model for Industrial Leakage Detection Using Acoustic Emission Signal," *Informatics*, vol. 7, p. 49, Mar. 2020, doi: 10.3390/informatics7040049.

[23] A. Li, R. Hamzah, and S. K. N. A. Rahim, "A chip X-ray image bubble defect detection model combined with Dual-Former attention mechanism," *Measurement*, vol. 248, p. 116871, 2025, doi: https://doi.org/10.1016/j.measurement.2025.116871.