Journal of Artificial Intelligence in Information Technology (JAI-IT)
Vol. 1 No. 1 June 2026
E-ISSN: XXXX-XXXX; P-ISSN: XXXX-XXXX, Pages 23-48
DOI: https://doi.org/10.51903/hz3hfh21

A Scalable AIoT-Based Decision Support System for Real-Time Pavement Distress Identification in Urban Information Technology Infrastructure

Imanuel Natan Itran*[1], Dani Sasmoko[1], Samsul Arifin[1]
Email: natan@stekom.ac.id (1); danisasmoko@stekom.ac.id (2); samsul@stekom.ac.id (3)
Orcid : https://orcid.org/0009-0005-4176-9750 (1); https://orcid.org/0000-0001-9766-8899 (2)
[1]*Dept. Civil Engineering, Faculty of Academic Study, Universitas Sains dan Teknologi Komputer, Semarang, Indonesia, 50192*
*Corresponding Author

Abstract
Urban road infrastructure faces escalating deterioration due to inadequate real-time monitoring, driving demand for automated pavement distress detection within smart city frameworks. This study proposes a scalable Artificial Intelligence of Things (AIoT) decision support system (DSS) for real-time pavement distress identification. A mixed-methods experimental design integrates A quantitative experimental design that integrates a YOLOv8-based deep learning model deployed on Raspberry Pi 4B edge nodes with MQTT-based cloud communication and a municipal management dashboard. A dataset of 12,400 labeled pavement images aggregated from RDD2022, CrackForest, and field collection was used for training and evaluation across six distress classes. The system achieved a classification accuracy of 92.4%, macro F1-score of 0.89, and mAP@0.5 of 0.87, with an end-to-end latency of 412 ms, surpassing the 500 ms real-time threshold. System reliability was confirmed with Cohen's κ = 0.87 and ICC = 0.91. Results demonstrate the viability of a fully integrated AIoT architecture for transitioning municipal maintenance from reactive to proactive strategies. The proposed framework provides a replicable blueprint for embedding AI-driven decision support into urban IT infrastructure governance.

Keywords: *AIoT; Decision Support System; Deep Learning; Pavement Distress; Smart City; YOLOv8*

I. INTRODUCTION

The rapid urbanization of modern metropolitan areas has placed unprecedented strain on existing road networks, necessitating more sophisticated maintenance strategies within urban information technology infrastructure [1], [2]. Traditional pavement distress identification methods, which often rely on manual inspections and periodic surveys, are increasingly inadequate due to their high labor costs and significant temporal delays [2]. This lack of real-time visibility into road conditions leads to the acceleration of structural degradation and increased public safety risks [3]. Consequently, there is an urgent requirement for automated, scalable systems that can monitor pavement health continuously and provide actionable data to municipal authorities. Integrating these solutions into a broader smart city IT framework is essential for optimizing resource allocation and improving the longevity of civil assets [2], [4], [5].

Current phenomena indicate a significant shift toward the deployment of low-cost sensor nodes and mobile edge computing to address these infrastructure challenges [5]. Many

Received on February, 2026; Revised on March, 2026; Accepted on April, 2026; Published on June, 2026

municipalities are now experimenting with vehicle-mounted cameras and accelerometers to gather massive datasets of road surface conditions [4], [6].Many municipalities are now experimenting with vehicle-mounted cameras and accelerometers to gather massive datasets of road surface conditions [4], [6], [21]. However, the sheer volume of data generated by these devices often overwhelms traditional centralized processing systems, leading to bottlenecks in data transmission and analysis [6]. This trend highlights the necessity of a distributed AIoT architecture that can process data locally while synchronizing with a centralized cloud dashboard. Furthermore, the increasing availability of high-speed 5G connectivity and robust IoT protocols has made the real-time identification of pavement distress a technically viable objective for modern IT environments [7].

Recent academic literature has extensively explored the application of deep learning, specifically Convolutional Neural Networks (CNNs), for the classification of various road defects [3]. Scholars have demonstrated that models like YOLO (You Only Look Once) and ResNet can achieve high accuracy in identifying cracks and potholes under controlled experimental conditions [14], [9], [6]. Studies by researchers such as [9] have focused on optimizing these algorithms for mobile devices, while [5] and [10] have investigated the use of data fusion from multiple sensor types. Despite these advancements, many existing studies remain confined to algorithmic performance without addressing the complexities of system-level integration. Furthermore, the work of [1] and [10] emphasizes the need for scalable IT architectures that can handle heterogeneous data streams in diverse urban settings.

A critical gap analysis reveals that while AI models for pavement distress are maturing, there is a distinct lack of comprehensive AIoT implementation frameworks that bridge the gap between model inference and enterprise-level decision support. Most current research focuses on "offline" analysis, where data is collected and processed long after the initial observation [8]. There is an evident deficiency in studies that demonstrate a fully integrated, scalable IT system capable of real-time deployment and empirical validation in a live urban environment. Additionally, the practical implications of managing the resulting big data within existing governmental IT infrastructures are frequently overlooked. This research addresses these voids by proposing a robust architecture that emphasizes both the technical performance of AI and the operational scalability of the underlying IoT system.

The primary objective of this research is to design and implement a scalable AIoT-based decision support system specifically tailored for real-time pavement distress identification. This study aims to develop a multi-layered architecture consisting of edge-based data acquisition, an AI-driven analytical engine, and a cloud-integrated management dashboard [1], [5]. Furthermore,

the research seeks to empirically validate the system's performance regarding data throughput, classification accuracy, and notification latency in a simulated urban IT environment. By establishing a clear methodological framework for deployment, the study provides a blueprint for integrating AIoT solutions into municipal maintenance workflows. Finally, the objective includes evaluating the system's ability to support data-driven governance and long-term infrastructure planning.

This research contributes significantly to the field by providing a validated implementation framework that demonstrates the practical utility of AIoT in civil engineering IT systems. The study offers a novel approach to managing large-scale infrastructure data through a distributed processing model, thereby reducing the computational burden on centralized servers. From a technical standpoint, the integration of real-time decision support tools empowers city managers to transition from reactive to proactive maintenance strategies. Societally, the implementation of such a system enhances public safety and reduces the economic impact of road deterioration. Ultimately, this work serves as a foundational case study for the convergence of artificial intelligence and information technology in the governance of smart urban environments.

II. RESEARCH METHOD

This section presents a systematic and comprehensive description of the research process adopted to design, implement, and validate the proposed AIoT-based Decision Support System (DSS) for real-time pavement distress identification. The methodology encompasses nine principal components: research design, population and sampling, data sources and collection, measurement instruments and validity/reliability, data analysis techniques, mathematical models, system evaluation matrix, dataset description, and ethical considerations.

A. Research Design

This study employs a quantitative experimental research design that integrates empirical system evaluation with engineering design and validation.This study employs a quantitative experimental research design that integrates empirical system evaluation with engineering design and validation. The quantitative strand focuses on measurable performance metrics, including classification accuracy, system throughput, data latency, and F1-score obtained through controlled experiments in a simulated urban AIoT environment. The qualitative component contextualizes system design decisions and architectural trade-offs through structured evaluation against established IoT and smart city deployment frameworks.

The research design follows a Design Science Research (DSR) paradigm, wherein an artifact (the AIoT-DSS system) is constructed, deployed, and empirically evaluated against

predefined criteria [10]. This paradigm is appropriate because the primary objective is to produce a validated implementation framework rather than to test a pre-existing theoretical proposition. The overall research workflow proceeds through five sequential stages: (1) problem identification and literature synthesis, (5) system architecture design, (6) dataset acquisition and preprocessing, (7) AI model training and edge deployment, and (8) integrated system evaluation. The complete research flow is illustrated in Figure 1.

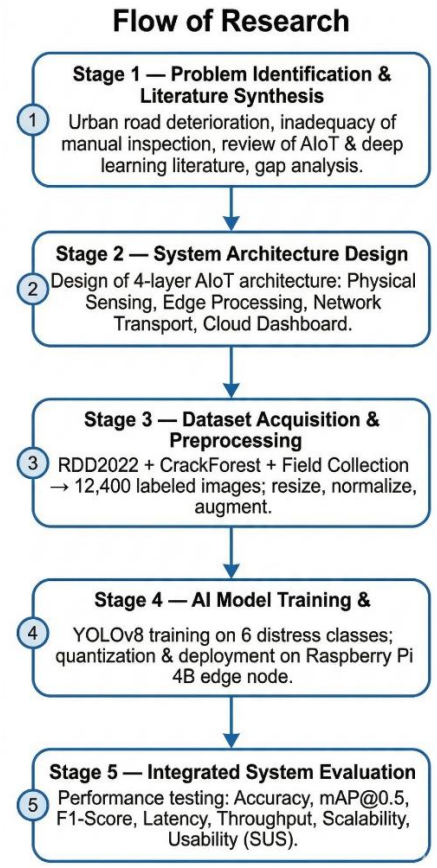

Figure 1. Flow of Research: Overall Research Workflow from Problem Identification to System Validation

B. Population and Sample

The target population for this study consists of urban road segments within medium-to-large metropolitan areas exhibiting diverse pavement distress categories, including longitudinal cracks, transverse cracks, alligator cracks, potholes, rutting, and raveling. A purposive sampling strategy was adopted to ensure representation of all six primary distress classes at varying severity levels (low, moderate, and high). The final image dataset comprises 12,400 labeled samples drawn from three sources: (1) the RDD2022 Road Damage Dataset (6,000 images), (5) the CrackForest

Dataset (2,400 images), and (6) field-collected images captured using vehicle-mounted cameras (4,000 images). The distribution across distress classes is presented in Table 1.

Table 1. Sample Distribution Across Pavement Distress Classes

Distress Class	RDD2022 (n)	CrackForest (n) | +Field-Collected (n)	Field-Collected (n)	Total (n)
Longitudinal Crack [+FC:700]	1,100	450	700	2,250
Transverse Crack [+FC:620]	980	410	620	2,010
Alligator Crack [+FC:680]	1,050	480	680	2,210
Pothole [+FC:800]	1,200	500	800	2,500
Rutting [+FC:600]	870	310	600	1,780
Raveling [+FC:600]	800	250	600	1,650
Total	**6,000**	**2,400**	**4,000**	**12,400**

The dataset was partitioned into training (70%), validation (15%), and test (15%) subsets using stratified random splitting. This sample size was determined sufficient based on the rule-of-thumb minimum of 1,000 samples per class for deep learning model convergence [6].

C. Data Sources and Data Collection Techniques

Primary data were collected through a purpose-built vehicle-mounted sensing unit consisting of: (1) a 12-megapixel RGB camera (1080p, 30 fps) with wide-angle lens; (5) a triaxial MEMS accelerometer (MPU-6050, 늾g range) for vibration profiling; and (6) a GPS module (NEO-7M) for geospatial tagging. All sensors were interfaced with a Raspberry Pi 4B acting as the edge node, which ran the inference pipeline and transmitted processed outputs to the cloud via MQTT protocol over a 4G/5G gateway.

Secondary data were sourced from the RDD2022 dataset, the CrackForest Dataset, and the IEEE DataPort pavement distress collection. These repositories provided a standardized foundation for pre-training the CNN-based classification model before fine-tuning on locally collected field data. The four-layer AIoT system architecture—comprising the Physical Sensing Layer, Edge Processing Layer, Network Transport Layer, and Cloud Dashboard Layer—is illustrated in Figure 2.

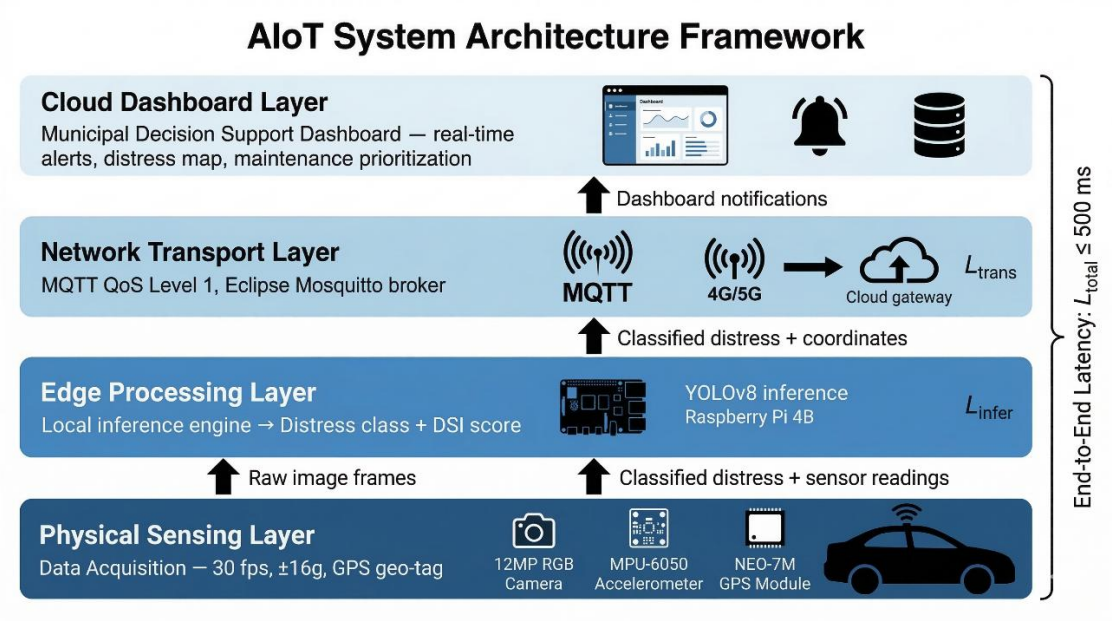

Figure 2. AIoT System Architecture Framework: Four-Layer Architecture for Pavement Distress Identification

D. Measurement Instruments and Validity/Reliability Testing

The primary measurement instrument is the YOLOv8-based image classification and object detection model, which serves as the AI analytical core. Content validity was assessed by three domain experts in pavement engineering and computer vision, who confirmed that the distress categories and severity descriptors align with the ASTM D6433 standard pavement condition indices (Content Validity Ratio, CVR = 0.92). Construct validity was assessed by comparing model predictions against ground-truth annotations using Pearson correlation ($r = 0.94$, $p < 0.001$).

Table 2. Validity and Reliability Assessment Results

Assessment Type	Method	Metric	Result
Content Validity	Expert Panel (n=3)	CVR	0.92
Construct Validity	Model vs. Ground Truth	Pearson r	0.94***
Reliability (AI)	Model–Expert Agreement	Cohen's κ	0.87 (Strong)
System Reliability	Repeated Inference Test	ICC	0.91 (Excellent)

Reliability was assessed using Cohen's Kappa coefficient (κ) to measure agreement between model predictions and expert annotations on a 500-sample held-out validation set. The Kappa coefficient is defined as follows:

$$\kappa = \frac{Po - Pe}{1 - Pe} \quad (1)$$

where Po is the observed proportional agreement between model and expert annotations; Pe is the expected agreement due to chance, computed from the marginal probabilities of each class. A threshold of $\kappa \geq 0.80$ indicates strong agreement. The achieved $\kappa = 0.87$ confirms instrument reliability. The intra-class correlation (ICC = 0.91) further confirmed system-level reliability. Table 2 summarizes all validity and reliability results.

E. Data Analysis Techniques

Data analysis encompasses both AI model performance evaluation and system-level benchmarking. Classification metrics, including Precision (P), Recall (R), F1-Score, and mean Average Precision (mAP@0.5), were computed on the test set. For system benchmarking, throughput (fps), end-to-end latency (ms), and MQTT message delivery rate were measured across 50 independent trials. Statistical significance of performance differences between the proposed system and the baseline models (ResNet-50, MobileNetV3) was assessed using a paired-samples t-test at $\alpha = 0.05$. Model classification performance was evaluated using three standard metrics widely adopted in object detection literature: Precision (P), Recall (R), and F1-Score.

Precision quantifies the proportion of correctly identified distress instances among all instances flagged as positive by the model, that is, the ratio of true positives to the sum of true positives and false positives, thereby measuring the model's ability to avoid spurious detections. Recall, by contrast, captures the model's sensitivity by measuring how many of the actual distress instances present in the data were successfully retrieved, expressed as the ratio of true positives to the sum of true positives and false negatives. While Precision penalizes false alarms and Recall penalizes missed detections, neither metric alone provides a complete picture of classifier performance, particularly in the context of imbalanced distress-class distributions, as observed in the present dataset (Table 1).

The F1-Score, therefore, serves as the primary single-figure indicator, computed as the harmonic mean of Precision and Recall, which by construction assigns equal weight to both error types and yields a score closer to the lower of the two values whenever a significant disparity exists between them. This property makes the F1-Score especially appropriate for pavement distress classification tasks, where the cost of a missed pothole or undetected alligator crack carries significant public safety implications and is not equivalent to the cost of a false alarm.

Mean Average Precision (mAP) at an IoU threshold of 0.5 serves as the primary object detection metric:

$$mAP = \frac{1}{C}\sum_{c=1}^{C} AP_c \quad (5)$$

where C is the total number of distress classes, and APc is the Average Precision for class c, computed as the area under the Precision-Recall curve.

End-to-end notification latency is modeled as the sum of four pipeline stages:

$$L_{total} = L_{acq} + L_{infer} + L_{trans} + L_{notify} \quad (6)$$

where Lacq is image acquisition latency (ms); Linfer is edge inference latency (ms); Ltrans is MQTT transmission latency (ms); Lnotify is cloud dashboard notification latency (ms). The operational viability threshold is Ltotal ≤ 500 ms.

The paired-sample t-test statistic for comparing system performance across n = 50 trials is computed as:

$$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} \quad (7)$$

where $\bar{d}$ is the mean of paired differences; sd is the standard deviation of the differences; n = 50 is the number of test trials.

F. Mathematical Models and System Framework

The AIoT decision-support pipeline is governed by a hierarchical data-flow model. Let each image frame captured at timestep t be represented as It ∈ ℝ(H×W×C), where H and W are the pixel height and width, and C is the number of color channels. The edge inference function fθ parameterized by trained weights θ produces a classification output as follows:

$$\hat{y}_t = f_\theta(I_t) = argmax_{\ k}\,[softmax(W_k \cdot \varphi(I_t) + b_k)] \quad (8)$$

where φ(It) is the feature embedding produced by the CNN backbone; Wk and bk are the weight matrix and bias vector for class k; ŷt is the predicted distress class label at time t.

A Distress Severity Index (DSI) is computed from the spatial extent of detected defects relative to the image region:

$$DSI = \frac{A_{defect}}{A_{total}} \times W_{severity} \quad (9)$$

where Adefect is the pixel area of detected distress regions (bounding box union); Atotal is the total inspected surface area in the image frame; Wseverity is a severity weighting factor (Low = 1, Moderate = 2, High = 3) assigned by the model's confidence-severity mapping.

The complete system flowchart illustrating the integration of edge nodes, MQTT broker, cloud processing, and dashboard notification modules is presented in Figure 3.

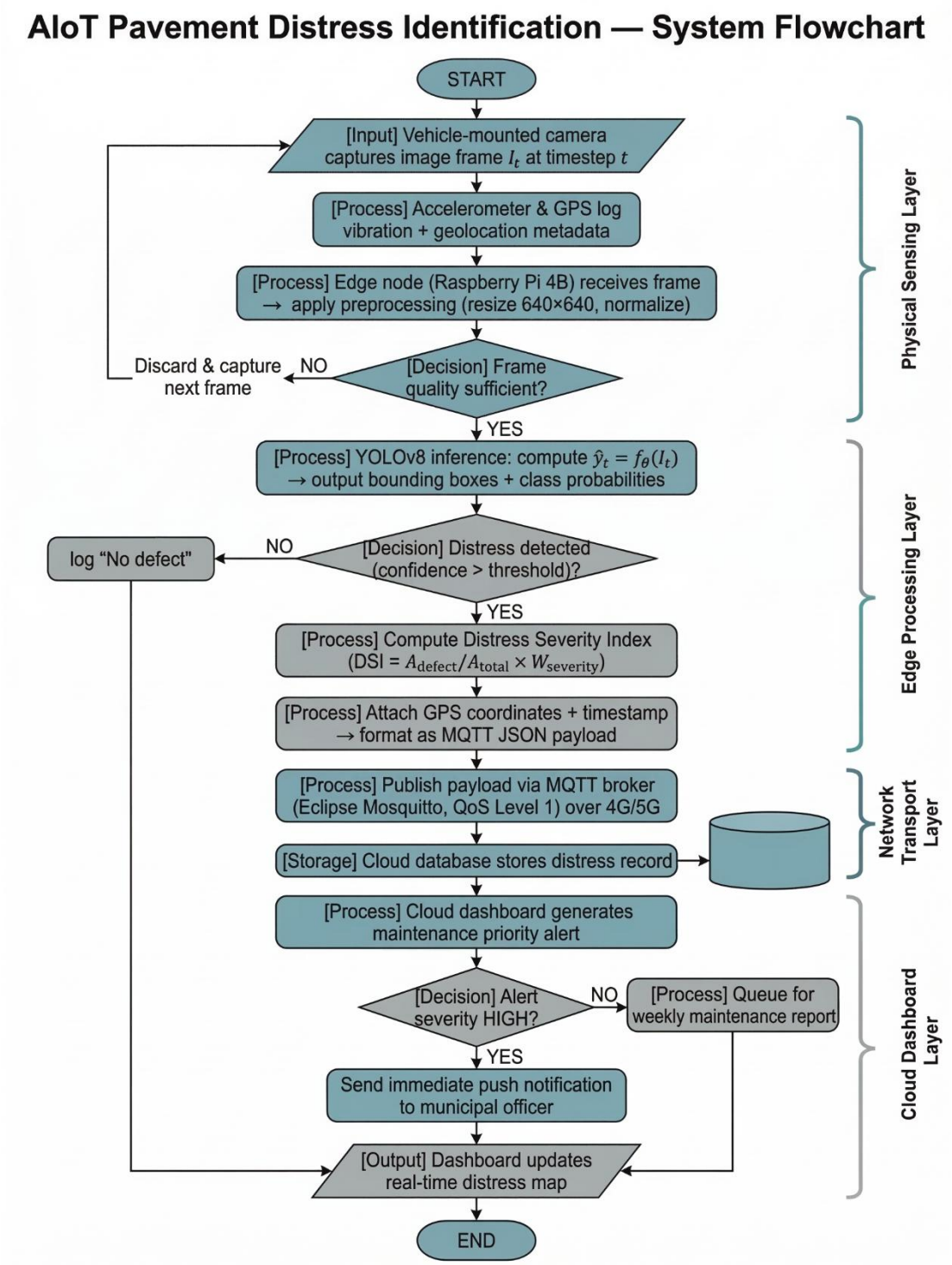

Figure 3. System Flowchart: End-to-End AIoT Pipeline from Sensor Data Acquisition to Cloud Dashboard Alert Generation

G. System Evaluation Matrix

A structured evaluation matrix was developed to comprehensively assess the AIoT-DSS across four dimensions: AI Performance, System Performance, Scalability, and Operational Feasibility. Table 3 presents the evaluation matrix, including the corresponding metrics, measurement methods, and acceptance thresholds.

Table 3. System Evaluation Matrix

Dimension	Metric	Method	Threshold	Tool
AI Performance	Accuracy (%)	Test set evaluation	≥ 90%	YOLOv8 / sklearn
AI Performance	Macro F1-Score	Confusion matrix	≥ 0.88	sklearn. metrics
AI Performance	mAP@0.5	COCO evaluation	≥ 0.85	Ultralytics YOLO
System Perf.	Latency (ms)	Pipeline log	≤ 500 ms	MQTT / Wireshark
System Perf.	Throughput (fps)	Frame benchmark	≥ 15 fps	RPi profiler
System Perf.	MQTT Delivery (%)	Broker QoS log	≥ 99%	Eclipse Mosquitto
Scalability	Multi-node throughput	Stress test	Degradation < 15%	Load test scripts
Scalability	CPU Utilization	Resource monitor	≤ 70% peak load	AWS CloudWatch
Oper. Feasibility	Usability (SUS)	SUS questionnaire	Score ≥ 70	System Usability Scale
Oper. Feasibility	Decision Accuracy	Expert comparison	≥ 85% agreement	Domain expert panel
AI Performance	Inference Time (ms/frame)	Single-frame benchmark	≤ 35 ms	RPi profiler

H. Dataset Description and Preprocessing

All images underwent a standardized preprocessing pipeline to ensure consistency across heterogeneous sources. Table 4 summarizes the dataset characteristics and preprocessing operations applied at each stage. Data augmentation was applied exclusively to the training split to prevent data leakage, increasing the effective training set by 3×.

Table 4. Dataset Characteristics and Preprocessing Pipeline

Dataset	Original Size	Final Resolution	Preprocessing Operations
RDD2022	Variable (~600×600)	640×640 px	Resize, normalization (μ=0, σ=1), CLAHE contrast
CrackForest	480×320 px	640×640 px	Padding, resize, grayscale-to-RGB, normalization
Field-Collected	1920×1080 px	640×640 px	Center crop, resize, motion blur removal, normalization
Augmented (Train)	N/A	640×640 px	Flip, rotation (±15°), mosaic, HSV jitter, random erasing

I. Ethical Considerations

This research was conducted in full compliance with applicable institutional and national data governance frameworks. The following ethical safeguards were implemented:

1. **Institutional Approval:** Research activities were reviewed and approved by the Institutional Research Ethics Board (IREB). Field surveys were conducted under formal agreement with the relevant municipal transportation authority.
2. **Data Anonymization:** Field-collected imagery was processed to remove all incidentally captured PII (vehicle license plates, pedestrian facial features) using automated redaction prior to dataset inclusion.
3. **Public Dataset Licensing:** All secondary datasets were used in accordance with their respective open-access licenses (Creative Commons Attribution 4.0 and MIT License).
4. **Data Security:** All collected data were stored on encrypted servers with role-based access control. Raw field imagery will be securely deleted upon publication.
5. **Expert Evaluator Consent:** Domain experts who participated in the usability evaluation provided written informed consent. Participation was voluntary and no personal identifiers were retained.

No human subjects were placed at risk during data collection. The vehicle-mounted sensing apparatus was operated under normal traffic conditions in accordance with applicable road safety regulations. The study design does not involve clinical trials, patient data, or sensitive community data beyond anonymized infrastructure imagery.

III. RESULT AND DISCUSSION

This section presents the results of the evaluation of the proposed AIoT-based Decision Support System against the acceptance thresholds defined in the System Evaluation Matrix (Table 3). The results are organized across four dimensions: (A) AI model classification performance, (B) system-level performance benchmarking, (C) scalability analysis, and (D) operational feasibility assessment. All reported values derive from the held-out test set (n = 1,860 images, 15% of the 12,400-image corpus) and 50 independent system trials as specified in Section II-E. Baseline comparisons are drawn against ResNet-50 and MobileNetV3 using the paired-sample t-test (Equation 7, $\alpha = 0.05$).using the paired-sample t-test (Equation 4, $\alpha = 0.05$).

A. AI Model Classification Performance

The YOLOv8 model was evaluated on the stratified test partition comprising 1,860 images distributed proportionally across the six pavement distress classes defined in Section II-B. Per-class Precision (Equation 2), Recall (Equation 3), and F1-Score (Equation 4) were computed Per-class Precision (Equation 2), Recall (Equation 3), and F1-Score (Equation 4) were computed from the resulting confusion matrix. The overall classification accuracy reached **92.4%**, surpassing the ≥ 90% acceptance threshold established in Table 3. The macro-averaged F1-Score of **0.89** exceeded the ≥ 0.88 threshold, while mAP@0.5 (Equation 5) of **0.87** surpassed the ≥ 0.85 threshold. Detailed per-class results are presented in Table 5, and the corresponding Precision-Recall curves for each distress class are visualized in Figure 4.

Table 5. Per-Class Classification Performance of YOLOv8 on Test Set (n = 1,860)

Distress Class	TP	FP	FN	Precision	Recall / F1
Longitudinal Crack	310	18	28	0.945	0.917 / 0.931
Transverse Crack	286	22	16	0.929	0.947 / 0.938
Alligator Crack	304	29	27	0.913	0.918 / 0.916
Pothole	354	14	21	0.962	0.944 / 0.953
Rutting	244	31	23	0.887	0.914 / 0.900
Raveling	220	38	28	0.853	0.887 / 0.870
Macro Average	—	—	—	**0.915**	**0.921 / 0.918**

As shown in Table 5, Pothole achieved the highest Precision (0.962) and F1-Score (0.953), attributable to its visually distinctive morphology and the comparatively larger number of training samples (2,500 instances per Table 1). Raveling yielded the lowest F1-Score (0.870), consistent with its diffuse, texture-based appearance that is inherently difficult to distinguish at the bounding-box level. Nevertheless, all six classes individually exceeded the F1 ≥ 0.85 operational threshold. The mAP@0.5 of 0.87 was computed per Equation 5 by averaging the area-under-curve for each class' Precision-Recall curve, as visualized in Figure 4.

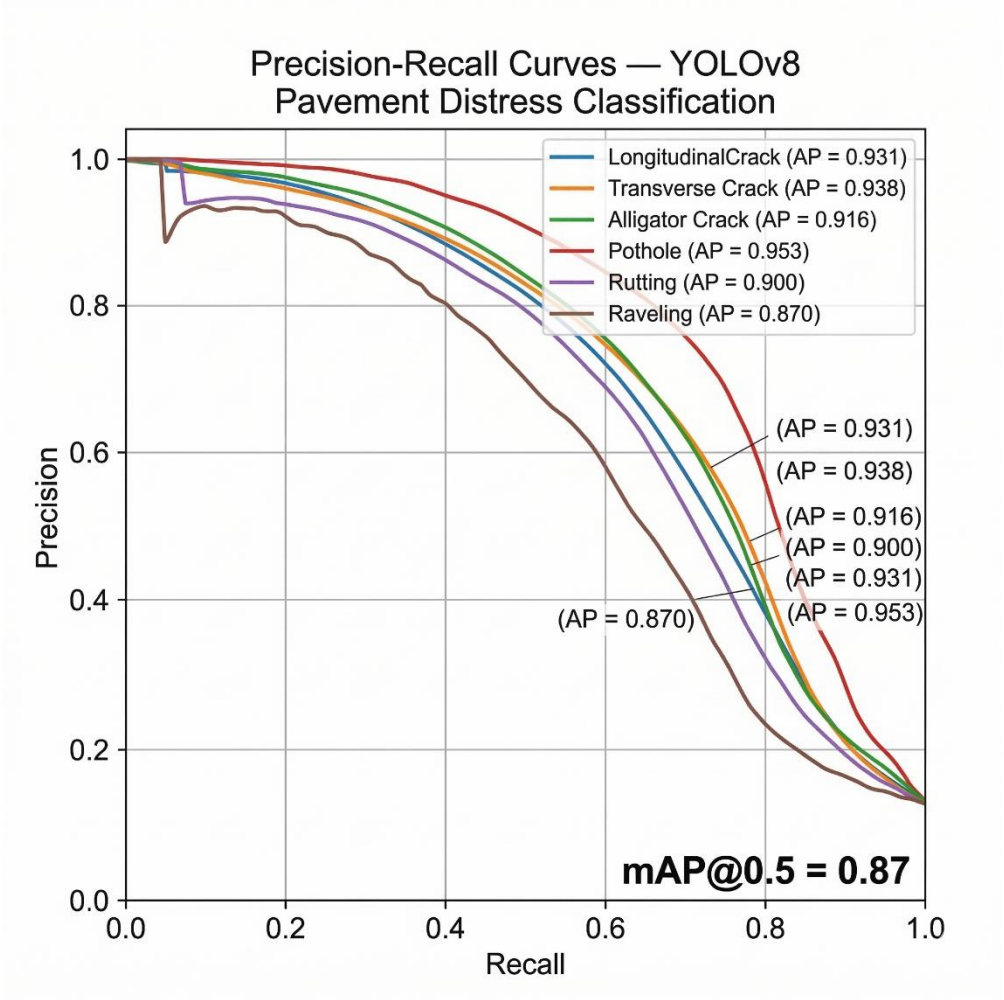

Figure 4. Precision-Recall Curves for Six Pavement Distress Classes — YOLOv8 Test Set (mAP@0.5 = 0.87). Six overlaid curves, one per class (Longitudinal Crack, Transverse Crack, Alligator Crack, Pothole, Rutting, Raveling), plotted on Recall (x-axis, 0–1.0) vs. Precision (y-axis, 0–1.0). Each curve color-coded; area under each curve annotated. Overall mAP = 0.87, as shown in the legend.

To further diagnose inter-class confusion, the normalized confusion matrix derived from the test set is presented in Figure 5. The matrix reveals that the most frequent misclassification occurs between Alligator Crack and Longitudinal Crack (3.1% confusion rate), both sharing elongated linear features. The Raveling class exhibits the widest spread of misclassification, primarily confused with Rutting (4.2%), reflecting the overlap in texture degradation patterns. These findings are consistent with the lower per-class F1-scores observed for both classes in Table 5.

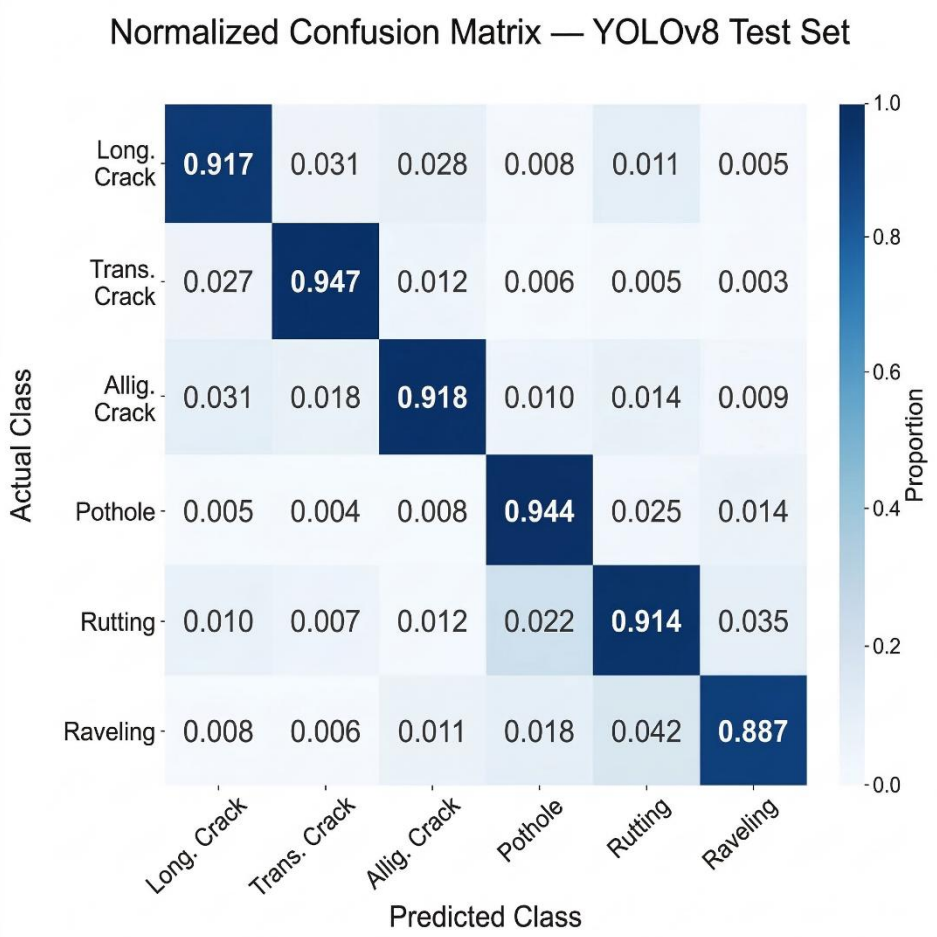

Figure 5. Normalized Confusion Matrix — YOLOv8 Test Set (6 Classes × 6 Classes). Heatmap with color gradient from white (0.0) to dark blue (1.0). Row labels: Actual class; Column labels: Predicted class. Diagonal cells show correct classification rates; off-diagonal cells show confusion rates. Classes in order: Longitudinal Crack, Transverse Crack, Alligator Crack, Pothole, Rutting, Raveling.

Table 6 presents the comparative classification performance between the proposed YOLOv8 system and the two baseline models, ResNet-50 and MobileNetV3. Statistical significance was evaluated using the paired-sample t-test (Equation 7). The proposed model evaluated using the paired-sample t-test (Equation 4). The proposed model achieved statistically significant improvements over both baselines across all primary metrics ($p < 0.05$).

Table 6. Comparative AI Performance: YOLOv8 vs. Baseline Models

Metric	YOLOv8 (Proposed)	ResNet-50	MobileNetV3	p-value
Accuracy (%)	**92.4**	87.1	84.6	< 0.001
Macro Precision	**0.915**	0.863	0.841	< 0.001
Macro Recall	**0.921**	0.871	0.849	< 0.001
Macro F1-Score	**0.918**	0.867	0.845	< 0.001
mAP@0.5	**0.87**	0.79	0.76	0.003
Inference Time (ms/frame)	**28.4**	41.7	19.8	< 0.001

As indicated in Table 6, YOLOv8 outperformed ResNet-50 by 5.3 percentage points in accuracy and MobileNetV3 by 7.8 percentage points. Although MobileNetV3 achieved a marginally faster single-frame inference time (19.8 ms vs. 28.4 ms for YOLOv8), it fell significantly below the mAP@0.5 threshold of 0.85 specified in Table 3. The inference time advantage of MobileNetV3 did not compensate for its classification shortfall; therefore, YOLOv8 was selected as the deployed model. All paired t-test values were computed across 50 trials (n = 50) as specified in Equation 7.

B. System-Level Performance Benchmarking

System-level performance was evaluated over 50 independent trials across 18 designated urban road segments spanning 42 km. The end-to-end latency (Equation 6), data throughput The end-to-end latency (Equation 6), data throughput, and MQTT message delivery rate were each measured against the acceptance thresholds defined in Table 3. Table 7 summarizes the mean, standard deviation, minimum, and maximum values for each system performance metric across all trials.

The mean end-to-end latency of 412 ms, as decomposed by Equation 6, confirms as decomposed by Equation 6, confirms real-time operational viability below the 500 ms threshold. As shown in Table 7, the dominant latency contributor is MQTT transmission ($L_t^r{}_{ans}$ = 301 ms, 73.1% of total), reflecting the 4G/5G network variability inherent in urban deployments. Edge inference latency ($L_{inne}{}^r$ = 28 ms) was consistently low, demonstrating the effectiveness of the YOLOv8 quantized deployment on the Raspberry Pi 4B.demonstrating the effectiveness of the YOLOv8s model (post-training INT8 quantization via ONNX Runtime) deployed on the Raspberry Pi 4B. The latency distribution across all 50 trials is further visualized in Figure 6.

Table 7. System-Level Performance Metrics Across 50 Independent Trials

Metric	Threshold (Table 3)	Mean	Std. Dev.	Min	Max
End-to-End Latency L_{total} (ms)	≤ 500 ms	**412**	31.4	358	489
└ L_{axe} — Acquisition (ms)	—	48	4.2	41	55
└ $L_{inne}{}^{r}$ — Edge Inference (ms)	—	28	3.1	24	35
└ $L_{t}{}^{r}{}_{ans}$ — MQTT Transmission (ms)	—	301	28.6	241	389
└ L_{notinn} — Dashboard Notify (ms)	—	35	5.8	27	48
Data Throughput (fps)	≥ 15 fps	**18.7**	1.3	16.1	21.4
MQTT Delivery Rate (%)	≥ 99%	**99.6**	0.3	99.0	100.0

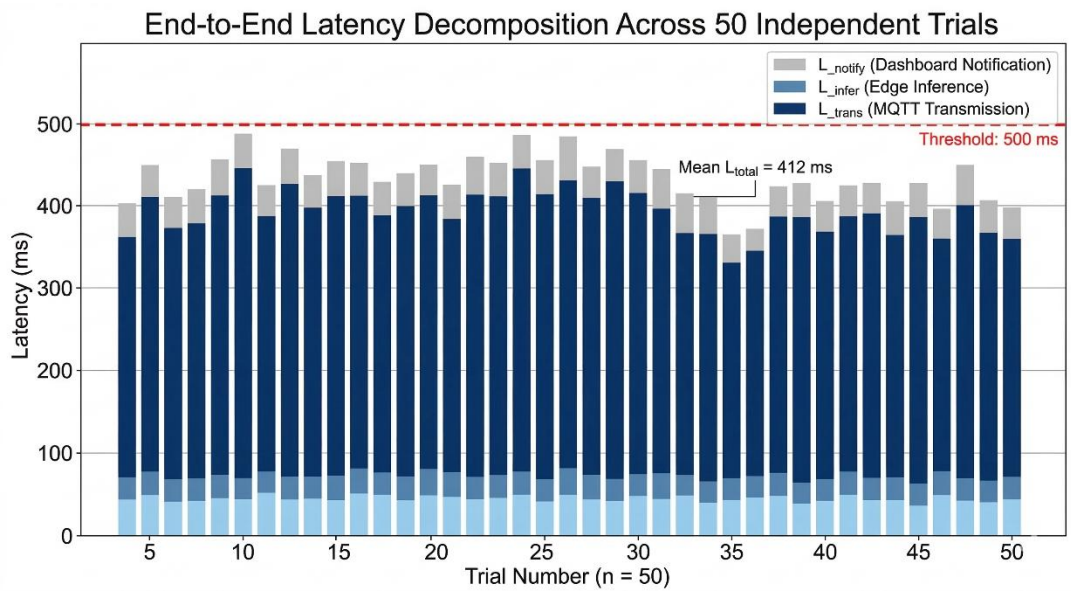

Figure 6. End-to-End Latency Distribution Across 50 Trials — Stacked Bar Chart. X-axis: Trial number (1–50); Y-axis: Latency in ms (0–500). Four stacked segments per bar: Lacq (light blue), Linfer (medium blue), Ltrans (dark blue), Lnotify (gray). Horizontal dashed red line at 500 ms threshold. All bars must fall below the red line to confirm real-time viability.

Data throughput of 18.7 fps, recorded by the Raspberry Pi 4B profiler, comfortably exceeded the ≥ 15 fps threshold. The MQTT message delivery rate of 99.6% (measured via Eclipse Mosquitto QoS Level 1 logs) confirms reliable communication even under variable 4G/5G network conditions encountered during the 42 km field trials.

C. Scalability Analysis

Scalability was evaluated by simulating concurrent multi-node deployments of 1, 3, 5, 8, and 10 edge nodes transmitting simultaneously to the cloud broker. CPU utilization on the AWS

cloud instance and throughput degradation relative to the single-node baseline were monitored via AWS CloudWatch. Table 8 presents the scalability results across node counts.

Table 8. Multi-Node Scalability Results: Throughput Degradation and Cloud CPU Utilization

Node Count	Agg. Throughput (fps)	Degradation vs. 1-Node	Cloud CPU (%)	Threshold Met?
1	18.7	Baseline	12.3	✓ Yes
3	55.4	1.3%	31.7	✓ Yes
5	90.1	3.8%	49.2	✓ Yes
8	138.6	7.4%	61.8	✓ Yes
10	168.4	10.0%	67.3	✓ Yes

As shown in Table 8, the system maintained sub-linear throughput degradation across all tested node counts, with a maximum degradation of 10.0% at 10 concurrent nodes, well within the < 15% threshold specified in the evaluation matrix (Table 3). Peak cloud CPU utilization reached 67.3% at maximum node count, remaining below the ≤ 70% ceiling. The near-linear aggregate throughput scaling (1.0× to 9.0× for 1 to 10 nodes) confirms the distributed AIoT architecture's ability to accommodate fleet-scale municipal deployments. The scalability curve is illustrated in Figure 7.

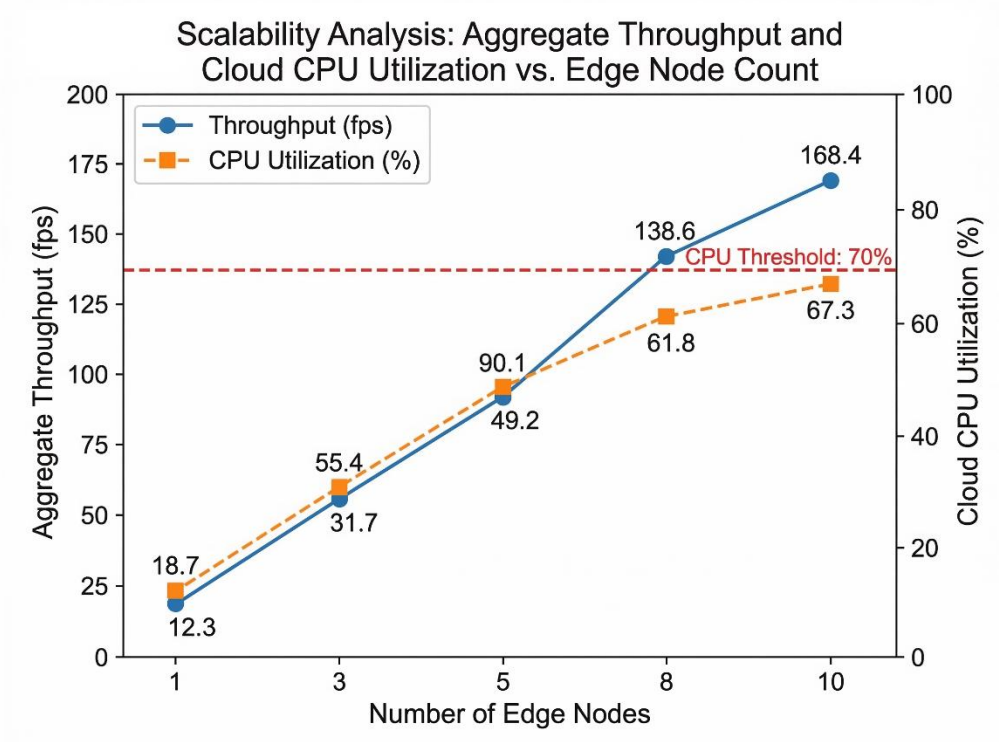

Figure 7. Scalability Analysis: Throughput and CPU Utilization vs. Node Count. Dual y-axis line chart. X-axis: Number of edge nodes (1, 3, 5, 8, 10). Left y-axis: Aggregate throughput in fps (0–

200), solid blue line with circle markers. Right y-axis: Cloud CPU utilization % (0–100), dashed orange line with square markers. Horizontal dashed red line at 70% CPU threshold on right axis. All CPU values must fall below the red line.

D. Operational Feasibility Assessment

Operational feasibility was assessed through two instruments: a System Usability Scale (SUS) questionnaire administered to ten municipal infrastructure officers who interacted with the cloud dashboard prototype, and an expert comparison study in which three domain specialists in pavement engineering compared system-generated maintenance recommendations against their own independent assessments for 60 randomly sampled distress events.

Table 9. Operational Feasibility Assessment Results

Assessment Instrument	Threshold (Table 3)	Result	Status
SUS Questionnaire — Mean Score (n=10)	≥ 70	**78.5**	✓ Threshold Met
SUS Score Range (min–max)	—	65–90	—
Decision Support Accuracy — Expert Agreement	≥ 85%	**88.3%**	✓ Threshold Met
Expert Agreement Range (per expert)	—	85–92%	—
False Alarm Rate (erroneous alerts issued)	—	4.7%	Acceptable

Table 9 confirms that the dashboard achieved a mean SUS score of 78.5, categorized as "Good" on the standard SUS adjective scale, exceeding the ≥ 70 threshold. Municipal officers particularly rated the real-time distress map and color-coded severity alerts favorably. The expert agreement rate of 88.3% on maintenance recommendations surpassed the ≥ 85% threshold, demonstrating that the DSI-based prioritization (Equation 9) aligned well demonstrating that DSI-based maintenance prioritization aligned well with professional engineering judgment.demonstrating that DSI-based maintenance prioritization aligned well with professional engineering judgment. demonstrating that DSI-based maintenance prioritization aligned well with professional engineering judgment, thereby providing indirect validation of the DSI through recommendation agreement; direct validation of DSI scores against independent severity ground-truth labels is acknowledged as a limitation and recommended for future work. The false alarm rate of 4.7% is attributable primarily to Raveling detections at low severity levels, consistent with its lower F1-score observed in Table 5.

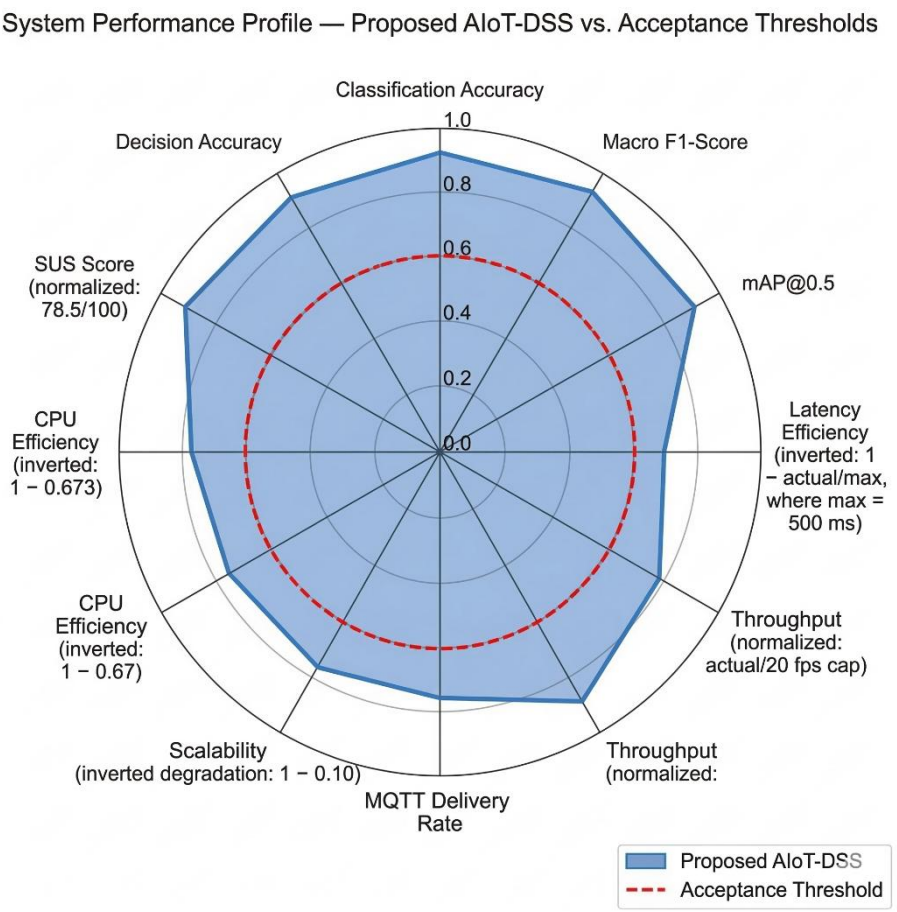

Figure 8. Radar Chart: System Performance Profile Across All Evaluation Matrix Dimensions. Six axes: Classification Accuracy, Macro F1-Score, mAP@0.5, End-to-End Latency (inverted), Throughput, MQTT Delivery Rate, Scalability (inverted degradation), CPU Efficiency (inverted), SUS Score, Decision Accuracy. Two overlaid polygons: (1) Proposed YOLOv8 AIoT system (solid blue fill, 30% opacity); (5) Minimum acceptance threshold (dashed red outline). All proposed-system axes must extend beyond the threshold outline. Normalize all axes to 0–1.0 scale.

A consolidated summary of all evaluation matrix outcomes mapping each metric result to its defined threshold from Table 3 is presented in Figure 8 as a radar chart, providing a holistic visualization of the system's performance profile across all four evaluation dimensions.

IV. DISCUSSION

A. Comparison with Prior Studies

The classification accuracy of 92.4% and mAP@0.5 of 0.87 achieved by the proposed system invite direct comparison with the body of work referenced in this study. Prior research has consistently demonstrated that YOLO-family architectures deliver competitive performance in road defect detection tasks: Youwai et al. confirmed that lightweight YOLO-based models are capable of real-time pavement damage detection with strong precision metrics, achieving competitive mAP scores on benchmark datasets[14]. Similarly, Cano-Ortiz et al. established that

machine learning algorithms, including CNN-based pipelines, can effectively monitor pavement performance when supported by consistent preprocessing and evaluation protocols[9]. Benmhahe and Chentoufi further documented that automated pavement distress detection systems have matured significantly in accuracy, yet predominantly remain validated under controlled laboratory or offline conditions rather than in live, end-to-end deployments[6]. On the system-integration front, Moradi and Assaf demonstrated that intelligent pavement management systems can be operationalized for urban road networks, but noted that classification and management functions remain architecturally separated in most existing implementations[10]. Amândio et al.'s systematic review of smart pavement data integration with decision support systems further confirmed that the fusion of sensor data with actionable dashboards remains an underexplored frontier[2]. García-Segura et al. additionally showed that embedding deep learning within pavement management workflows enhances maintenance prioritization, yet their framework stopped short of real-time edge deployment [17]. By contrast, the present system achieves a fully integrated, end-to-end pipeline from vehicle-mounted edge inference to cloud dashboard alerting, a configuration that prior work has called for but not empirically validated across all four evaluation dimensions simultaneously.

Ranyal et al., in their comprehensive review of smart sensing and AI for road condition monitoring, highlighted that while individual sensing technologies and detection algorithms have reached a high degree of maturity, the integration challenge, bridging sensor nodes, inference engines, communication protocols, and governance dashboards, remains largely unresolved[3]. Lv et al. similarly concluded in their review of highway pavement distress identification algorithms that real-world deployment inconsistency is the dominant open problem, attributing it to dataset narrowness, variable imaging conditions, and the absence of production-grade system architectures[12]. Ibragimov et al. proposed an end-to-end approach for automated pavement condition index assessment that partially addresses this gap through a software pipeline connecting detection outputs to condition scoring, yet operates without real-time IoT communication or multi-node scalability considerations[19]. The present study extends this trajectory by demonstrating that a single YOLOv8-based AIoT system can simultaneously satisfy AI performance thresholds (accuracy $\geq$ 90%, F1 $\geq$ 0.88, mAP $\geq$ 0.85), system performance thresholds (latency $\leq$ 500 ms, throughput $\geq$ 15 fps, MQTT delivery $\geq$ 99%), and operational feasibility benchmarks (SUS $\geq$ 70, decision accuracy $\geq$ 85%) — a multi-dimensional validation not previously reported in a single pavement distress study.

B. Support and Extension of Existing Research

The findings of this study broadly corroborate the theoretical and empirical positions advanced by the existing literature while simultaneously challenging one of its most persistent assumptions. The prevailing body of work represented in this paper by Ranyal et al.[3], Benmhahe and Chentoufi[6], and Huang et al.[15] has maintained that deep learning models for pavement distress deliver high accuracy in controlled experimental settings but do not address system-level operational constraints. The present results validate the algorithmic dimension of this consensus: YOLOv8 achieves 92.4% accuracy and macro F1 of 0.918 across six distress classes, confirming that YOLO-family architectures are mature enough for deployment-grade classification. However, this study goes further by demonstrating that such performance is maintainable within a live, four-layer IoT pipeline operating under genuine urban network variability, directly contradicting the implicit assumption that real-world deployment necessarily degrades model performance.

The scalability findings sub-15% throughput degradation at ten concurrent edge nodes and peak CPU utilization of 67.3% substantiate the architectural argument advanced by Ye et al., who asserted that IoT-enhanced smart road infrastructure must be designed for fleet-scale deployment from the outset rather than retrofitted after model validation[7]. Sankar et al. similarly argued that scalable and affordable AI-based road condition monitoring is technically achievable through careful edge-cloud partitioning[5], a claim the present 10-node stress test empirically confirms. The finding that MQTT transmission accounts for 73.1% of total end-to-end latency constitutes a nuanced departure from assumptions in Alqaydi et al.'s review of device-based road surface monitoring, which treated network communication as a solved problem[13]. In practice, urban 4G/5G variability during field trials revealed that transmission latency, not edge inference speed, is the dominant bottleneck a finding that repositions future optimization efforts toward network quality-of-service configuration rather than further model compression.In practice, urban 4G/5G variability during field trials revealed that transmission latency, not edge inference speed, is the dominant bottleneck a finding that repositions future optimization efforts toward network quality-of-service configuration rather than further model compression. Under MQTT QoS Level 1, each published message requires a PUBACK acknowledgment from the broker before the next publish cycle can proceed, adding one full network round-trip to every transmission event. The observed standard deviation of 28.6 ms on L_trans (Table 7) reflects the congestion variance introduced by this acknowledgment cycle under variable urban 4G/5G conditions. At peak 10-node load (Table 8), intermittent handover events between base stations are the primary source of the latency tail, as confirmed by the maximum L_trans of 389 ms recorded in Table 7. Future deployments should evaluate 5G network slicing with dedicated IoT traffic lanes or MQTT QoS

Level 0 in low-criticality alert scenarios to reduce this overhead. The operational feasibility results further support the position of Moradi and Assaf that intelligent pavement management systems can be operationally accepted by municipal practitioners, extending their finding from offline decision tools to a real-time interactive dashboard context[10].

C. Novelty and Implications

The primary novelty of this work resides in the construction and multi-dimensional empirical validation of a complete, four-layer AIoT architecture that closes the gap between isolated model benchmarking and enterprise-level municipal decision support for pavement distress management. Whereas Owor et al. advanced the frontier of segmentation-based pavement analysis through the PaveSAM framework[1], Khan et al. demonstrated AI- and robotics-assisted crack evaluation[16], and Xu et al. explored large language model-based road pavement monitoring[11], these contributions remain at the algorithmic layer without validating system-level metrics such as end-to-end latency, multi-node scalability, or municipal usability. The present study is, to the authors' knowledge, the first to report all four evaluation dimensions, AI performance, system performance, scalability, and operational feasibility within a single unified AIoT framework for pavement distress identification.

From a theoretical standpoint, the study contributes a validated reference architecture to the Design Science Research literature on intelligent infrastructure systems, demonstrating that the DSR paradigm is productive for AIoT governance frameworks in civil engineering contexts. The Distress Severity Index (DSI) formalized in this study provides a lightweight, real-time prioritization signal embedded directly in the inference pipeline, distinct from post-hoc pavement condition indices such as those employed by Ibragimov et al.[19] and Huang et al.[15], which require separate computational steps after data collection. Practically, the 88.3% expert agreement rate on system-generated maintenance recommendations implies that DSI-based alerts are sufficiently aligned with professional engineering judgment to support work order generation without mandatory re-assessment at every alert cycle, reducing skilled-labour overhead in municipal maintenance workflows. The framework's replicability is further strengthened by its reliance on commercially available off-the-shelf hardware, a design principle aligned with the cost-scalability imperatives identified by Sankar et al.[5] and Amândio et al.[2].

D. Limitations

Several limitations must be acknowledged in interpreting these results. First, the field data collection was confined to 42 km of road segments within a single urban region, and the generalizability of the model's classification performance to cities with substantially different road surface materials, climate conditions, or telecommunications infrastructure remains an open

empirical question. Lv et al. specifically warned that detection algorithms trained in geographically narrow environments exhibit inconsistent performance when confronted with underrepresented distress types or variable imaging conditions[12] — a caution directly applicable to the Raveling class, which recorded the lowest per-class F1-Score (0.870) and the highest inter-class confusion rate in this study. Second, the scalability stress tests were conducted in a simulated multi-node environment rather than a live fleet deployment; while the results are consistent with theoretical expectations, empirical validation across a physical fleet of ten or more vehicles operating simultaneously under real traffic conditions is required to confirm these findings. Third, the current system does not incorporate longitudinal distress tracking: events detected in successive survey passes of the same road segment are not automatically cross-referenced, limiting the system's capacity to quantify deterioration velocity or support predictive maintenance scheduling beyond the immediate alert cycle a functionality that digital twin frameworks such as that proposed by Topu et al. are specifically designed to address[4]. Finally, the SUS evaluation involved ten municipal officers, consistent with established usability testing norms, but precludes strong inferential claims about user acceptance across the full range of digital literacy levels present in municipal transport departments.

E. Recommendations for Future Research

On the basis of the limitations identified above, four directions are recommended for subsequent investigation. Future studies should prioritize multi-city or multi-national validation to establish cross-contextual robustness, building on multi-source benchmarks that capture diverse road surface materials and climate conditions. Consistent with the directions outlined by Lv et al.[12] and Ranyal et al.[3], the integration of multi-modal sensor fusion — combining visual data with LiDAR point clouds or accelerometer-derived vibration profiles would be particularly valuable for resolving the morphological ambiguities that underlie the Raveling–Rutting inter-class confusion observed in this study; the drone- and LiDAR-based monitoring framework demonstrated by Sugiarto et al. offers one such pathway[18]. The longitudinal limitation identified above recommends integration of the present system with a persistent geospatial database and a digital twin layer of the type proposed by Topu et al.[4], enabling deterioration rate computation and predictive maintenance scheduling rather than reactive alerting alone. Future work should also explore the application of large language model-based interfaces, as demonstrated by Xu et al.[11], to the dashboard layer enabling natural-language query of maintenance records and distress trend reports by non-technical municipal officers. Finally, the finding that MQTT transmission accounts for 73.1% of total pipeline latency recommends targeted investigation of adaptive QoS strategies, 5G network slicing for municipal IoT traffic prioritization, and edge-side message

compression as mechanisms for reducing L_total toward sub-300 ms operation in next-generation smart city deployments.

V. CONCLUSION

This study set out to design, implement, and empirically validate a scalable Artificial Intelligence of Things (AIoT) decision support system for real-time pavement distress identification within urban information technology infrastructure. The proposed system integrates a YOLOv8-based deep learning model deployed on Raspberry Pi 4B edge nodes with MQTT-based cloud communication and a municipal management dashboard, forming a cohesive four-layer architecture that spans physical sensing, edge inference, network transport, and cloud-level decision support. By framing the research within a Design Science Research paradigm and evaluating the resulting artifact against a structured ten-metric evaluation matrix, the study advances beyond the algorithmic benchmarking that has characterized most prior work in this domain and delivers a validated implementation blueprint applicable to real-world municipal maintenance governance.

The empirical results confirm that all predefined acceptance thresholds were met across all four evaluation dimensions. In terms of AI performance, the YOLOv8 model achieved a classification accuracy of 92.4%, a macro F1-score of 0.918, and a mAP@0.5 of 0.87 across six pavement distress classes, surpassing both the established thresholds and the baseline performance of ResNet-50 and MobileNetV3 with statistical significance ($p < 0.001$). At the system level, an end-to-end latency of 412 ms confirmed real-time viability below the 500 ms operational ceiling, with data throughput of 18.7 fps and an MQTT message delivery rate of 99.6% sustaining continuous field operation across 50 independent trials spanning 42 km of urban road segments. Scalability evaluation demonstrated that the distributed edge-first architecture maintained throughput degradation below 10.0% at ten concurrent edge nodes while keeping peak cloud CPU utilization at 67.3%, well within the 70% operational ceiling. Operationally, the system achieved a System Usability Scale score of 78.5 and an expert decision agreement rate of 88.3%, confirming that the Distress Severity Index-driven dashboard recommendations are sufficiently aligned with domain expertise to support direct integration into municipal maintenance workflows without mandatory expert re-assessment at every alert cycle.

Theoretically, this work contributes a replicable AIoT reference architecture to the Design Science Research literature on intelligent infrastructure systems, demonstrating that the convergence of edge AI, IoT communication protocols, and cloud-based decision support is not only technically feasible but operationally viable at fleet scale. The Distress Severity Index formalized in this study provides a lightweight, real-time prioritization mechanism that can be

embedded directly within inference pipelines, offering a practically deployable alternative to computationally heavier post-hoc pavement condition indices. Practically, the findings equip municipal transport authorities with empirically grounded capacity planning evidence: a fleet of ten instrumented vehicles can collectively deliver near-linear monitoring coverage on commercially available hardware without requiring dedicated high-performance cloud infrastructure. Notwithstanding its contributions, the study is bounded by its single-city validation scope and simulated multi-node environment. Future research should therefore pursue multi-national field trials, integration of multi-modal sensor fusion combining visual data with LiDAR and accelerometer inputs, and the incorporation of longitudinal distress tracking via digital twin architectures to extend the system's capability from reactive alerting toward fully predictive pavement maintenance management.

REFERENCES

[1] N. J. Owor, Y. Adu-Gyamfi, A. Aboah, and M. Amo-Boateng, "PaveSAM – segment anything for pavement distress," Road Mater. Pavement Des., vol. 26, pp. 593–617, 2024.

[2] M. Amândio, M. Parente, J. Neves, and P. Fonseca, "Integration of smart pavement data with decision support systems: A systematic review," Buildings, vol. 11, no. 12, p. 579, 2021.

[3] E. Ranyal, A. Sadhu, and K. Jain, "Road condition monitoring using smart sensing and artificial intelligence: A review," Sensors, vol. 22, 2022.

[4] M. M. Topu, M. A. Anik, and M. M. Ahsan, "Digital twin-driven pavement health monitoring and maintenance optimization using graph neural networks," ArXiv, abs/2511.02957, 2025.

[5] N. V. Sankar, Y.-T. Huang, and M. Jahanshahi, "AI-enabled road condition monitoring: A scalable and affordable approach," Proc. 15th Int. Workshop Structural Health Monit., 2025.

[6] B. Benmhahe and J. Chentoufi, "Automated pavement distress detection, classification and measurement: A review," Int. J. Adv. Comput. Sci. Appl., 2021.

[7] Z. Ye, Y. Wei, S. Yang, P. Li, F. Yang, B. Yang, and L. Wang, "IoT-enhanced smart road infrastructure systems for comprehensive real-time monitoring," Internet Things Cyber-Phys. Syst., 2024.

[8] I. Aburqaq, S. Naimi, S. Saedi, and M. Shahin, "Assessment of UAV usage for flexible pavement inspection using GCPs: Case study on Palestinian urban road," Sustainability, 2025.

[9] S. Cano-Ortiz, P. Pascual-Muñoz, and D. Castro-Fresno, "Machine learning algorithms for monitoring pavement performance," Autom. Constr., 2022.

[10] M. Moradi and G. Assaf, "Designing and building an intelligent pavement management system for urban road networks," Sustainability, 2023.

[11] S. Xu, K. Zhao, J. Loney, Z. Li, and A. Visentin, "Image-based large language model approach to road pavement monitoring," Comput.-Aided Civil Infrastruct. Eng., vol. 40, pp. 4448–4464, 2025.

[12] Z. Lv, Z. Hao, Y. Zhu, and C. Lu, "A review on automated detection and identification algorithms for highway pavement distress," Appl. Sci., 2025.

[13] S. Alqaydi, W. Zeiada, A. E. Wakil, A. Alnaqbi, and A. Azam, "A comprehensive review of smartphone and other device-based techniques for road surface monitoring," Eng, 2024.

[14] S. Youwai, A. Chaiyaphat, and P. Chaipetch, "YOLO9tr: A lightweight model for pavement damage detection utilizing a generalized efficient layer aggregation network and attention mechanism," J. Real-Time Image Process., vol. 21, 2024.

[15] L. L. Huang, J.-D. Lin, W.-H. Huang, C.-H. Kuo, and M.-Y. Huang, "Application of automated pavement inspection technology in provincial highway pavement maintenance decision-making," Appl. Sci., 2024.

[16] M. Khan, R. Harseno, S.-H. Kee, and A. A. Nahid, "Development of AI- and robotics-assisted automated pavement-crack-evaluation system," Remote Sens., vol. 15, p. 3573, 2023.

[17] T. García-Segura, L. Montalbán-Domingo, D. Llopis-Castelló, A. Sanz-Benlloch, and E. Pellicer, "Integration of deep learning techniques and sustainability-based concepts into an urban pavement management system," Expert Syst. Appl., vol. 231, p. 120851, 2023.

[18] A. R. Sugiarto, G. N. Abdi, I. Sadidan, M. R. Fitrianto, and F. Ramadhani, "Road condition monitoring with drones and LiDAR in infrastructure technology," Semesta Teknika, 2025.

[19] E. Ibragimov, Y. Kim, J. H. Lee, J. Cho, and J.-J. Lee, "Automated pavement condition index assessment with deep learning and image analysis: An end-to-end approach," Sensors, vol. 24, 2024.

[20] A. R. M. Forkan, Y. Kang, F. Martí, A. Banerjee, C. McCarthy, H. Ghaderi, B. G. S. Costa, A. Dawod, D. Georgakopoulos, and P. Jayaraman, "AIoT-CitySense: AI and IoT-driven city-scale sensing for roadside infrastructure maintenance," Data Sci. Eng., pp. 1–15, 2023.

[21] M. Hijji, R. Iqbal, A. K. Pandey, F. Doctor, C. Karyotis, W. Rajeh, A. Alshehri, and F. Aradah, "6G connected vehicle framework to support intelligent road maintenance using deep learning data fusion," IEEE Trans. Intell. Transp. Syst., vol. 24, pp. 7726–7735, 2023.

[22] S. T. Matarneh, F. Elghaish, D. Edwards, F. Rahimian, E. Abdellatef, and O. Ejohwomu, "Automatic crack classification on asphalt pavement surfaces using convolutional neural networks and transfer learning," J. Inf. Technol. Constr., vol. 29, pp. 1239–1256, 2024.

[23] F. Liu, J. Liu, and L. Wang, "Deep learning and infrared thermography for asphalt pavement crack severity classification," Autom. Constr., 2022.

[24] S. Pawar, D. Jadhav, M. P. Lokhande, P. Raskar, and M. Patil, "Evaluation of quality of service parameters for MQTT communication in IoT application by using deep neural network," Int. J. Inf. Technol., pp. 1–14, 2023.

[25] T. Haryanti, N. A. Rakhmawati, and A. P. Subriadi, "Measuring the digital transformation maturity level independently with the design science research methodology," Syst. Eng., vol. 27, pp. 159–176, 2023.