Multi-Sensor Spatial Data Fusion for Road Condition Monitoring Digital Twins

Deepak Satheesan [1], Songnian Li [1], Michael Chapman [1]

[1] Department of Civil Engineering, Faculty of Engineering and Architecture Science, Toronto Metropolitan University, 350 Victoria Street, Toronto – (dsatheesan, snli, mchapman)@torontomu.ca

Keywords: Digital twin, Pavement condition monitoring, Multi-sensor data, Deep learning, Anomaly detection

Abstract

Pavement Management Systems (PMS) are essential for evaluating and maintaining transportation infrastructure; however, conventional monitoring methods are often labour-intensive, costly, and inaccurate. The growing need for reliable. timely pavement condition data has driven the development of automated, data-driven approaches. This study presents a low-cost and scalable framework for pavement condition monitoring that integrates multimodal sensing with a digital twin (DT) environment. Smartphones equipped with inertial measurement unit (IMU) sensors, GPS, and cameras are used to collect synchronized vibration and visual data during normal driving conditions. Vibration signals are analysed to detect anomalies associated with pavement surface irregularities, while video data are processed using a deep learning-based object detection model to identify surface distress. A late fusion approach combines the outputs from both modalities to improve detection reliability and provide comprehensive condition assessment. The system enables spatial mapping of detected distresses and supports real-time visualization through a web-based DT dashboard. Results demonstrate that multimodal sensing compensates for the limitations of individual sensors, enhancing both detection accuracy and robustness. The proposed framework offers a practical solution for efficient pavement monitoring. It supports data-driven decision-making for proactive infrastructure management, with potential for future expansion through crowdsourced data and additional sensing technologies.

1. Introduction

A Pavement Management System (PMS) is a systematic framework designed to evaluate, maintain, and manage pavement infrastructure efficiently. Pavements deteriorate over time due to traffic and weather, making timely inspection and accurate condition assessment essential for cost-effective maintenance and proactive asset management. Reliable data also enables the prediction of future performance, guiding maintenance planning. Conventional pavement monitoring relies on visual inspections or specialized vehicles, which are labour-intensive, costly, and prone to human error (Valipour et al., 2024). In Canada, pavements in unknown condition contribute to part of the $141.7 billion replacement value of transportation infrastructure (Statistics Canada, 2024). This lack of reliable data can lead to economic losses from poorly targeted funding, more roads in poor condition, and higher vehicle repair costs due to deteriorated road surfaces.

To address these challenges, studies have begun to focus on cost-effective, automated, and rapid data-driven solutions that provide actionable insights with minimal expertise and infrastructure requirements (Sattar et al., 2018). With the advent of IoT and sensor technologies, various sensor modalities, including vision, vibration, and laser-based sensors, have been integrated into pavement condition assessment systems (Al-Sabaeei et al., 2024; Chhabra et al., 2021; Sattar et al., 2021). However, most existing research still relies on a single data modality, which limits the comprehensiveness and accuracy of distress detection and classification. In contrast, multimodal fusion, which combines multiple sensor types, has been widely utilized in other fields, such as structural health monitoring, human activity recognition, and autonomous vehicles, where it has proven effective at compensating for sensor data gaps and enhancing model robustness (Nong et al., 2023). These successes highlight the potential of multi-modal fusion for pavement condition monitoring, which remains largely underexplored.

On the other hand, Digital Twin (DT) technology has emerged as a promising framework for real-time infrastructure management, offering enhanced decision-making capabilities through data integration and visualization. Beyond these capabilities, a DT enables continuous synchronization with the physical asset, supporting real-time updates, scenario simulation, predictive analytics, efficient data management, and interactive user inputs that can directly modify the model. It can also facilitate advanced functions such as automated maintenance workflows and repair planning. However, the application of DT for pavement condition monitoring is still in its early stages, and most existing studies focus primarily on structural models rather than on timely surface distress detection or decision support systems (Talaghat et al., 2024).

The current study aims to leverage the complementary strengths of multiple sensing modalities to enhance pavement condition assessment, while integrating these capabilities within a digital twin framework to support data-driven decision-making. By leveraging smartphones as an accessible, low-cost, and scalable platform for multimodal sensing, the study aims to enhance the accuracy, efficiency, and practicality of pavement monitoring. The following sections present the methodology, results and discussion, and conclusions.

2. Methodology

The current study leverages smartphones equipped with inertial measurement unit (IMU) sensors, global positioning system (GPS), and cameras to collect multi-sensor data for detecting and spatially mapping pavement surface distress. Smartphones mounted on vehicle windshields simultaneously record vibration signals and video of the road surface, enabling complementary analysis of pavement conditions through vibration-based and vision-based approaches.

A custom mobile application was developed using Android Studio Ladybug 2024.2.1, to enable concurrent acquisition of sensor data and video. The application records tri-axial measurements from the accelerometer, gyroscope, and magnetometer, along with corresponding GPS coordinates and timestamps, and stores them in a structured format. Video data are recorded simultaneously and indexed using the recording start timestamp to ensure synchronization with the sensor data.

Data collection was conducted using a Google Pixel 9 smartphone mounted on the windshield of a 2023 Kia Forte EX Premium vehicle. The device was secured using a suction cup mount to maintain a stable position and provide a clear forward-facing view of the road surface. This configuration enables consistent recording of road imagery and corresponding vibration responses as the vehicle traverses through pavement irregularities (Figure 1).

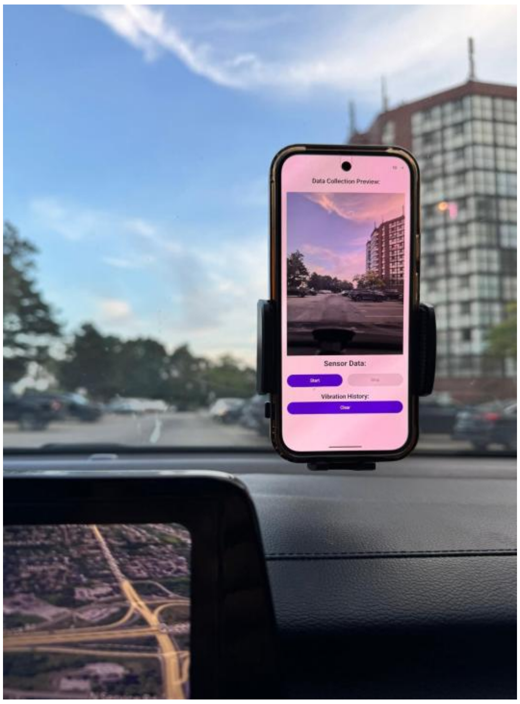

Figure 1. Smartphone placement

Vibration signals are processed to detect abnormal patterns associated with pavement anomalies, while video frames are analysed using a deep learning-based object detection model to identify visible surface distress. The outputs from these two modalities are integrated using a late fusion approach based on the Dempster–Shafer evidence theory to improve detection reliability. The overall workflow of the proposed methodology is illustrated in Figure 2.

2.1 Vibration data processing

Raw IMU data are processed through a multi-stage pipeline to extract road-induced vibration characteristics while minimizing the influence of sensor noise, device orientation, and vehicle dynamics. The preprocessing framework consists of sequential steps, including sensor orientation calibration, temporal resampling, coordinate transformation, condition indexing, and statistical deviation analysis (Singh et al., 2017; Wu et al., 2020).

Sensor orientation calibration is performed first to determine the device's orientation relative to the Earth reference frame. A 30-second stationary period at the start of each data collection run is used to estimate Euler angles (roll, pitch, yaw) from gravity-dominated accelerometer measurements combined with tilt-compensated magnetometer readings. Calibration quality is verified by ensuring that the accelerometer standard deviation remains below 0.5 m/s² and that the measured gravity magnitude approximates 9.81 m/s². This step is critical for accurately transforming device-frame accelerations into Earth-fixed coordinates and ensuring consistent extraction of vertical vibration components regardless of smartphone mounting orientation.

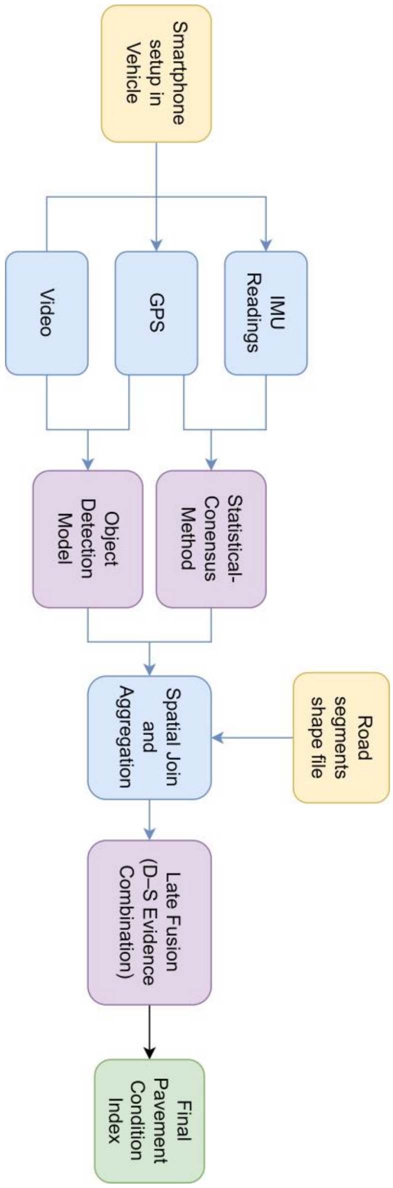

Figure 2. Overall methodology

Following calibration, temporal resampling is applied to convert irregularly sampled sensor data to a uniform 50 Hz frequency using linear interpolation. This standardization removes timing inconsistencies and enables the reliable application of frequency-based filtering and analysis techniques. The resampled linear acceleration signals are then transformed from the device coordinate system to the Earth-fixed reference frame using rotation matrices derived from the calibrated Euler angles. This transformation ensures that vertical acceleration components are consistently represented, comparisons across different data-collection setups.

To quantify road surface irregularities, the vertical component of linear acceleration (Vci), is computed by comparing the instantaneous three-dimensional acceleration vector with a rolling baseline representing normal driving conditions over a 30-second window (Sattar et al., 2021). The index is calculated as the normalized dot product between the current and baseline vibration vectors, providing a measure of deviation from typical vibration patterns. Higher deviations indicate the presence of potential pavement anomalies. In addition to Vci, statistical deviation measures are computed for each Earth-frame axis (East, North, Vertical) relative to their respective mean values, along with the overall acceleration magnitude. These features capture both directional and overall vibration characteristics and support robust anomaly detection.

Road surface anomalies are identified using an ensemble of complementary statistical methods. A rolling-window z-score approach detects local deviations by flagging points that exceed 3 standard deviations from the mean within a 50-sample window. Spike detection is performed by analysing first-order differences and identifying values exceeding the 99th percentile, capturing abrupt changes associated with surface defects. A median absolute deviation (MAD) method is applied to detect outliers robustly, with modified z-scores computed using a conservative threshold of 5.0. A Hampel filter further identifies localized anomalies based on deviations from the median within a seven-sample window. Finally, DBSCAN clustering is used to detect outliers as points in low-density regions without assuming any predefined data distribution. A consensus-based decision mechanism is employed to improve detection reliability, where a data point is classified as an anomaly if at least two of the five methods classify it as such. This approach balances sensitivity and false-positive reduction, resulting in robust identification of pavement surface irregularities. The final processed dataset includes georeferenced vibration features and anomaly indicators, which are subsequently integrated with visual detection results for multimodal analysis.

2.2 Vision Data processing

Vision data were collected using the smartphone camera, which captures continuous video of the road surface as the vehicle moves. Surface distresses, such as cracks and potholes, and repairs appear as visible anomalies in the frames. The objective of analysing this video data is to detect and classify these surface defects using image processing and deep learning, enabling spatial mapping and quantification of pavement condition. The smartphone application records video simultaneously with vibration and GPS data, with each video file saved using a timestamp corresponding to the start of recording. This ensures temporal alignment between vision and vibration data, allowing multimodal analysis and fusion of complementary information from both sensing modalities.

The model was trained on a merged dataset compiled from three publicly available pavement distress repositories: UAPD (Unified Asian Pavement Distress), RDD202 (Road Damage Detection 2022), and PID (Pavement Image Dataset) (Arya et al., 2024; Majidifard et al., 2020a, 2020b; Zhong et al., 2022; Zhu et al., 2022). All datasets were converted to YOLO format (normalized bounding box coordinates) and harmonized into six unified defect classes: alligator crack, longitudinal crack, transverse crack, oblique crack, pothole, and repair. The merged dataset was partitioned into training (70%), validation (15%), and test (15%) subsets using stratified random sampling to maintain consistent class distributions and ensure adequate representation of rare defect types. An additional class, maintenance hole (MH), was added via transfer learning, using a pre-trained checkpoint and freezing the first 10 backbone layers to preserve low-level feature extractors (edges, textures, gradients) while fine-tuning detection-specific layers on road defect patterns.

The YOLOv8n (nano) variant was selected for its computational efficiency and real-time suitability. An input resolution of 640 × 640 pixels was used to balance detection accuracy and processing speed. The general architecture of YOLOv8 is illustrated in Figure 3 (Yaseen, 2024). The YOLOv8n model used in this study follows the same architecture with reduced depth and width. Training was conducted on Google Colab with GPU acceleration (T4/V100) using the Adam optimizer with learning rate scheduling for stable convergence. A composite loss function combining Complete IoU (CIoU) for bounding box localization, binary cross-entropy for defect classification, and Distribution Focal Loss (DFL) for bounding box corner refinement (weighted 7.5:0.5:1.5), balanced localization, classification, and boundary accuracy. Extensive data augmentation—including mosaic augmentation, random horizontal flipping, translation (±10%) and scaling (0.5–1.5×), HSV colour jittering, RandAugment, random rotation (±10°), and random erasing (40%)—was applied to improve generalization under real-world variations in lighting, occlusion, and viewing angles.

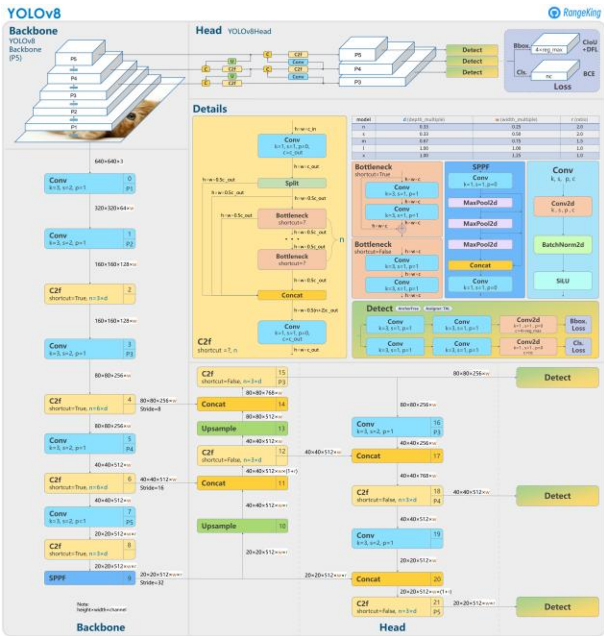

Figure 3. The architecture of YOLOv8

The model was trained with a batch size of 16 for up to 50 epochs with early stopping (patience = 20), an initial learning rate of 0.001 with cosine decay, and weight decay of 0.0005. Total training time was approximately 1 hour 52 minutes (6,700 seconds), averaging 2.2 minutes per epoch. During inference, Non-Maximum Suppression with an IoU threshold of 0.7 removed duplicate detections, a confidence threshold of 0.25 filtered low-confidence predictions, and a maximum of 300 detections per frame was allowed. Hyperparameters were selected using established, empirically validated practices to ensure stable training and reliable performance.

2.3 Late Fusion Approach

Smartphone-captured videos were processed frame-by-frame using the trained YOLOv8n model to detect road defects. Each detection was assigned GPS coordinates by matching frame timestamps (calculated as frame number/fps) to the synchronized sensor data. Defect density was computed as the ratio of the bounding box area to the effective road surface area (25% of each frame), and the Pavement Condition Index (PCI) for each frame was estimated using an automated pavement condition assessment method (Satheesan et al., 2024). Frame-level PCI values were then spatially aggregated to road segments using GeoPandas nearest-neighbour spatial joins, and segment-level statistics—including mean, standard deviation, minimum, and maximum—were calculated from all frames mapped to each segment.

Segment-level PCI and Vci measurements were first normalized via Rank Fusion, converting raw scores to percentile ranks (0–100), where higher ranks indicate worse pavement condition (PCI ranked in descending order, Vci in ascending order). These percentile ranks were then transformed into belief mass distributions over five condition categories (Excellent, Good, Fair, Poor, Critical) using overlapping triangular fuzzy membership functions. Each category spanned a 20-percentile range (0–20, 20–40, 40–60, 60–80, 80–100) with 50% overlap between adjacent categories. For a given rank, the membership function assigned partial belief to both the primary category and the adjacent transitional category, representing uncertainty at category boundaries and enabling smooth transitions between condition states. The Dempster–Shafer evidence theory was applied to combine the PCI and Vci belief distributions, calculating fused belief for each category while quantifying uncertainty arising from disagreement between modalities (Hassani et al., 2024). The final condition for each road segment was assigned to the category with the maximum combined belief, and a fused score on a 0–10 scale was computed as a belief-weighted average (Excellent = 10, Good = 7.5, Fair = 5, Poor = 2.5, Critical = 0).

2.4 Digital twin prototype framework

Multi-sensor observations are temporally aligned to enable cross-verification and sensor fusion. Detections from the late fusion approach are used to determine the condition rating of each road segment, with each detected event timestamped and geographically referenced to support spatial mapping of pavement conditions.

The digital twin prototype was developed using a file-based data architecture comprising CSV, JSON, and Shapefile formats for structured storage of tabular, detection, and geospatial data. The processing pipeline was implemented in Python, SciPy for signal processing, and GeoPandas for spatial analysis. Geospatial processing involved associating GPS-tagged observations with road segment geometries using nearest-neighbour spatial joins, followed by aggregation of condition indicators at the segment level. The user interface is implemented using a Dash web framework, with Plotly providing interactive visualizations, including synchronized time-series plots and geospatial maps.

A web-based interface provides interactive access to the digital twin, supporting both real-time-like inspection and analytical assessment. The framework comprises two integrated components: (i) an operational module enabling synchronized multimodal visualization for data validation and state inspection, and (ii) an analytical module that derives segment-level condition metrics through spatial aggregation and multimodal fusion. Together, these components enable both micro-level state interrogation and macro-level network assessment, supporting maintenance planning and infrastructure management decisions.

3. Results and Discussion

The proposed methodology was applied to a 1.5 km test section along Lakeshore East Boulevard (Figure 4), with manual field inspections conducted to validate the detected pavement distress and condition assessments (Figure 5).

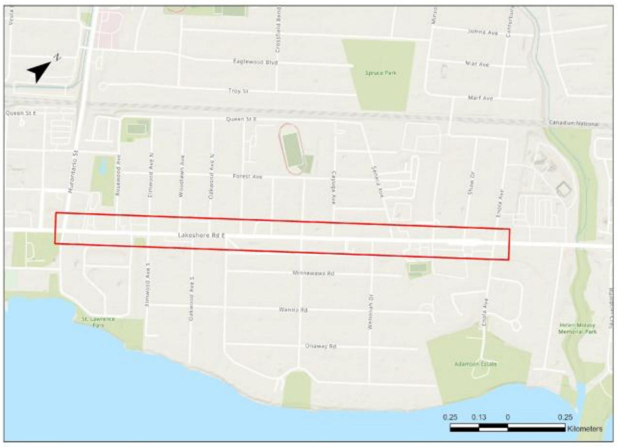

Figure 4. Study area

Figure 5. Field photographs

3.1 Vibration data

Raw smartphone IMU data underwent multi-stage preprocessing to isolate road-induced vibrations from confounding factors (Figure 6). Stage 1 resampled data to a uniform 50 Hz frequency. Stage 2 applied a high-pass filter (0.3 Hz cutoff) that removed the dominant gravity component (~9.8 m/s² on the Y-axis), reducing the mean acceleration to near zero across all axes. Stage 3 performed gravity removal via low-pass filtering and subtraction to eliminate residual static components, ensuring zero-mean linear acceleration. Stages 4 and 5 exhibit visually subtle changes because they operate on already-cleaned data: the band-pass filter (0.5-20 Hz) removes only small residual components outside the road defect frequency range (reducing noise by 3-4%), while coordinate transformation redistributes signal energy from device frame (X/Y/Z) to Earth frame (East/North/Vertical) without altering magnitude. This energy-preserving rotation ensures orientation-independent measurements critical for crowdsourced data collection. Stage 6 calculates the Vibration Condition Index (Vci) using a 30-second rolling baseline, and Stage 7 applies ensemble anomaly detection to identify 80 high-confidence road surface irregularities, demonstrating the pipeline's effectiveness in extracting meaningful road condition metrics from noisy smartphone sensor data.

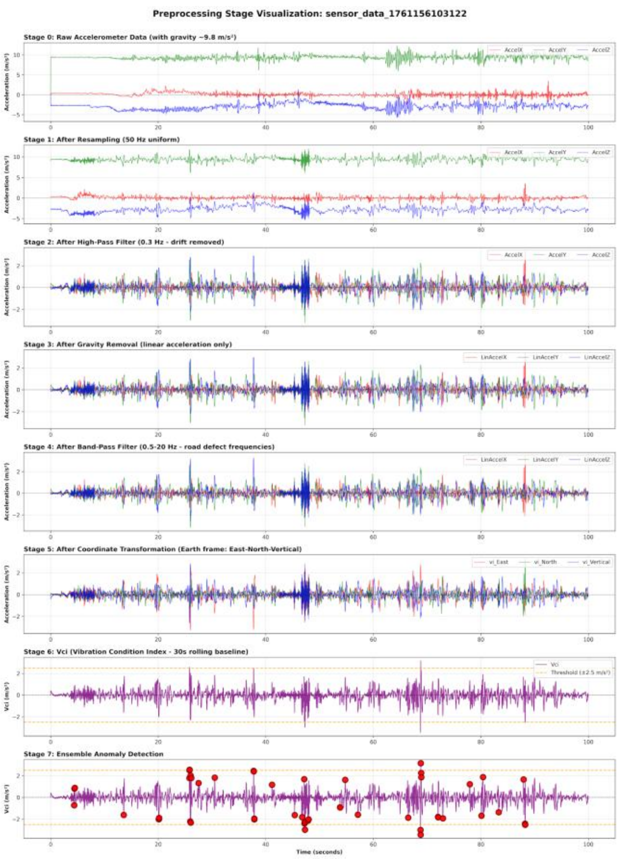

Figure 6. Vibration data results

3.2 Vision data results

The YOLOv8n model achieved strong overall detection performance on the validation set, with mAP@0.5 of 68.87%, mAP@0.5–0.95 of 43.03%, precision of 73.05%, recall of 65.58%, and F1-score of 69.1%, demonstrating competitive accuracy with state-of-the-art road defect detection systems while maintaining computational efficiency. Training showed rapid convergence, with mAP@0.5 improving by 47% over the first 20 epochs and with training and validation losses closely matching, indicating robust generalization. Figure 7 shows training losses. Class-specific analysis highlighted challenges due to dataset imbalance: longitudinal cracks dominated, and minority classes (oblique cracks, repairs) underperformed, while pothole detection was limited by visual ambiguity, lighting, and 2D context constraints. Figure 8 shows the confusion matrix from training results. The precision-favoured operating point prioritizes reliability, mitigating false positives critical for maintenance dispatches, and can be adjusted post-deployment to tune recall. The model offers a favourable trade-off between accuracy, real-time performance (10–15 FPS on smartphone CPUs), multi-dataset generalization, and integration with vibration-based monitoring. Figure 9 shows the batch results after training.

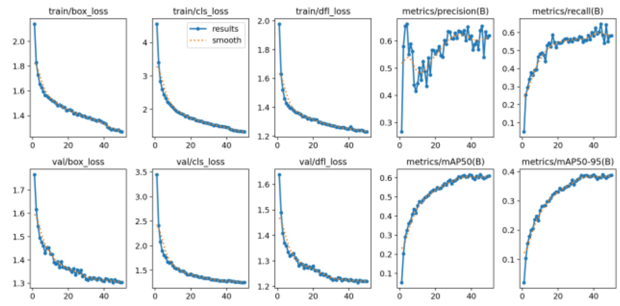

Figure 7. Training losses

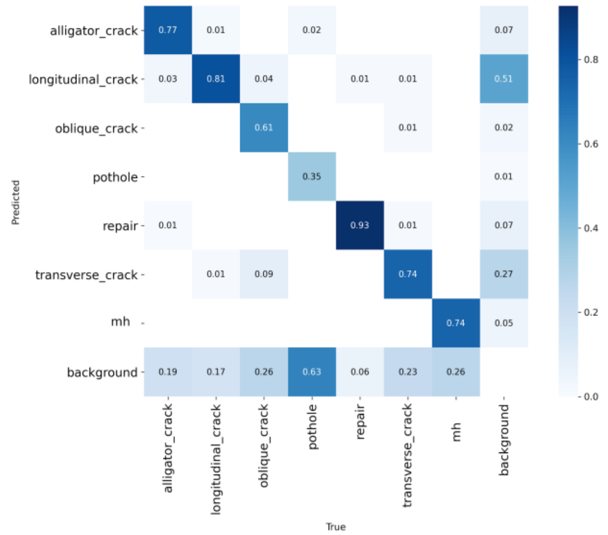

Figure 8. Confusion Matrix

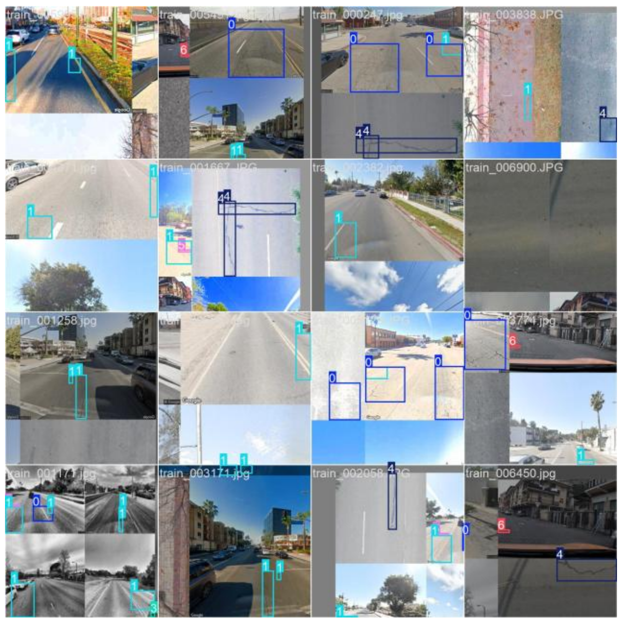

Figure 9. Training batch results

The YOLOv8n model detected road defects across 292 frames, 8.9% of 3,294 total frames. Sample frame from the survey is shown in Figure 10. Frame-level PCI analysis revealed predominantly excellent pavement conditions with a mean PCI of 7.5-8.0, where 62% of analysed frames scored as "Excellent" (PCI > 8.5) and only 4% as "Very Poor" (PCI < 2). Longitudinal cracks were the most frequently detected defect type, followed by transverse and oblique cracks, with defect density averaging 18-20% of the effective road surface area, reaching up to 82% in severely distressed sections. The majority of the surveyed route demonstrated well-maintained pavement suitable for continued service.

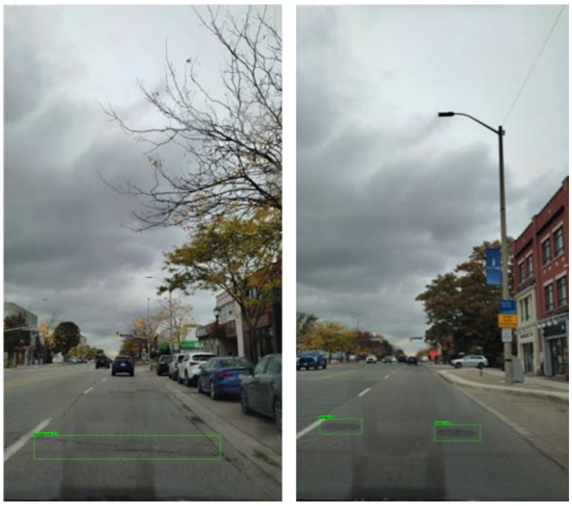

Figure 10. Distress detection results from vision data

3.3 Data Fusion result

The late fusion analysis combined visual PCI assessments with vibration-based Vci measurements using the Dempster–Shafer evidence theory across road segments along Lakeshore Road East (Figure 11). The results revealed notable discrepancies between the two modalities, supporting the need for multi-modal assessment. In several segments, vibration data indicated significant distress despite high PCI values from visual inspection, resulting in a downgrade of the fused condition rating from “Excellent” to lower categories. This suggests that vibration measurements capture subsurface defects and ride-quality degradation that are not detectable with vision-based methods alone, thereby reducing the risk of overestimating pavement condition.

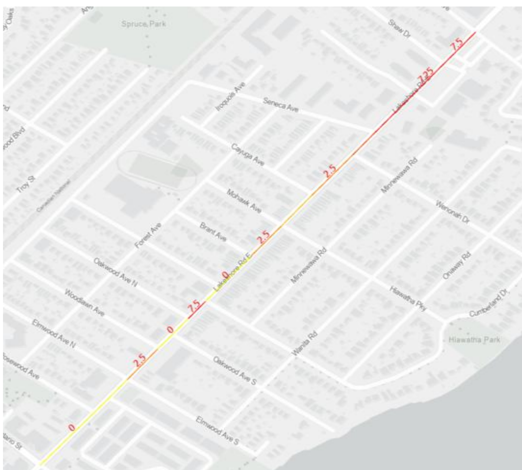

Figure 11. Data fusion result

Conversely, segments showed low uncertainty and high-confidence classification, highlighting their priority for maintenance. The fusion framework further quantified uncertainty, enabling identification of segments with conflicting evidence or insufficient data that may require additional inspection. Overall, the results demonstrate that integrating vision and vibration data provides a more comprehensive and reliable assessment of pavement condition than single-modality approaches.

3.4 Digital Twin

The developed digital twin framework enabled integrated visualization and assessment of pavement conditions through two complementary interfaces. The operational interface (Figure 12) facilitated synchronized inspection of video, sensor signals, and detected anomalies, allowing detailed examination of pavement responses at specific locations. This enabled the identification of localized events such as road depressions and transient vibration spikes that may not be evident from a single data source.

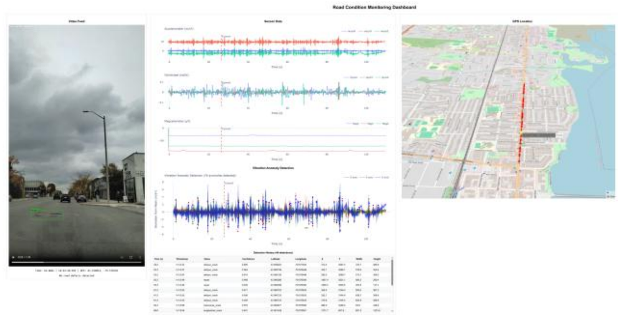

Figure 12. Operational interface of the digital twin

The analytical interface (Figure 13) provided segment-level condition assessment based on fused PCI and Vci indicators. Spatial visualization of condition ratings revealed variability in pavement performance along the study section, highlighting segments requiring maintenance attention. The integration of multi-modal data improved the robustness of condition assessment by combining surface-level defect detection with vibration-based indicators of ride quality.

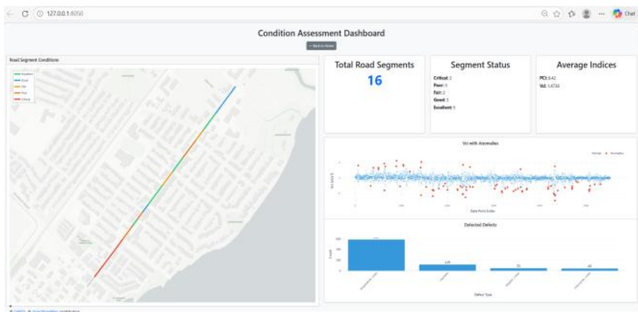

Figure 13. Analytical interface of the digital twin

The system implements four of five dimensions of Digital Twin framework (Tao et al., 2019), which are Physical Entity (road network with smartphone sensors), Virtual Entity (web dashboard), Services (YOLO detection, anomaly detection, data fusion), and Data (multimodal sensor integration), with partial connection implementation. The prototype operates at Level 1 (Descriptive monitoring) maturity with emerging Level 2 (Diagnostic analysis) capabilities through multi-method ensemble anomaly detection and uncertainty quantification but lacks Level 3 predictive modelling of deterioration progression (Rasheed et al., 2020).

4. Conclusion

This study presents a scalable and cost-effective pavement condition monitoring framework that integrates smartphone-based vision and vibration sensing to enable reliable multimodal distress detection. By combining complementary sensing modalities, the proposed approach mitigates the limitations of single-source methods and improves the robustness of pavement condition assessment. Integration within a GIS-enabled digital twin environment further allows spatially explicit visualization, continuous state representation, and enhanced decision support for maintenance planning. The vision-based component, implemented using a YOLOv8n model, demonstrated strong performance in detecting visible surface distress with high precision and satisfactory accuracy, while maintaining computational efficiency suitable for near real-time deployment. When combined with vibration-based indicators, the system captures both surface-level defects and underlying ride-quality variations, providing a more comprehensive representation of pavement condition.

Despite these advantages, several limitations remain. The fusion framework based on the Dempster–Shafer evidence theory requires careful design and calibration of belief functions, and its assumption of independence between sensing modalities may not always hold in practice. In addition, performance may vary under different vehicle dynamics, environmental conditions, and sensor configurations, which were not exhaustively explored in this study. The current implementation also operates as a Digital Shadow, with unidirectional data flow and limited predictive capability.

Future work will focus on improving model generalization and detection performance through advanced training strategies and expanded datasets, as well as validating results against survey-grade pavement evaluation methods. Enhancements to the fusion framework, including data-driven and learning-based approaches, will be explored to improve scalability and adaptability. Further developments will also aim to enable automated data transfer, real-time processing, and the incorporation of additional sensing modalities, thereby supporting the transition toward a fully autonomous, predictive digital twin for pavement management.

References

Al-Sabaeei, A.M., Souliman, M.I., Jagadeesh, A., 2024. Smartphone applications for pavement condition monitoring A review. Constr. Build. Mater. 410, 134207. https://doi.org/10.1016/j.conbuildmat.2023.134207

Arya, D., Maeda, H., Ghosh, S.K., Toshniwal, D., Sekimoto, Y., 2024. RDD2022: A multi-national image dataset for automatic road damage detection. Geosci. Data J. gdj3.260. https://doi.org/10.1002/gdj3.260

Chhabra, R., Rama Krishna, C., Verma, S., 2021. A survey on state-of-the-art road surface monitoring techniques for intelligent transportation systems. Int. J. Sensor Netw. 36(2), 123–145.

Hassani, S., Dackermann, U., Mousavi, M., Li, J., 2024. A systematic review of data fusion techniques for optimized structural health monitoring. Inf. Fusion 103, 102136. https://doi.org/10.1016/j.inffus.2023.102136

Majidifard, H., Adu-Gyamfi, Y., Buttlar, W.G., 2020a. Deep machine learning approach to develop a new asphalt pavement condition index. Constr. Build. Mater. 247, 118513. https://doi.org/10.1016/j.conbuildmat.2020.118513

Majidifard, H., Jin, P., Adu-Gyamfi, Y., Buttlar, W.G., 2020b. Pavement image datasets: A new benchmark dataset to classify and densify pavement distresses. Transp. Res. Rec. J. Transp. Res. Board 2674, 328–339. https://doi.org/10.1177/0361198120907283

Nong, X., Luo, X., Lin, S., Ruan, Y., Ye, X., 2023. Multimodal deep neural network-based sensor data anomaly diagnosis method for structural health monitoring. Buildings 13, 1976. https://doi.org/10.3390/buildings13081976

Rasheed, A., San, O., Kvamsdal, T., 2020. Digital twin: Values, challenges and enablers from a modeling perspective. IEEE Access 8, 21980–22012. https://doi.org/10.1109/ACCESS.2020.2970143

Satheesan, D., Talib, M., Li, S., Yuan, A.X., 2024. An automated method for pavement surface distress evaluation. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLVIII-M-4, 47–53. https://doi.org/10.5194/isprs-archives-XLVIII-M-4-2024-47-2024

Sattar, S., Li, S., Chapman, M., 2021. Developing a near real-time road surface anomaly detection approach for road surface monitoring. Meas. J. Int. Meas. Confed. 185, 109990. https://doi.org/10.1016/j.measurement.2021.109990

Sattar, S., Li, S., Chapman, M., 2018. Road surface monitoring using smartphone sensors: A review. Sensors (Switzerland) 18, 3845. https://doi.org/10.3390/s18113845

Singh, G., Bansal, D., Sofat, S., Aggarwal, N., 2017. Smart patrolling: An efficient road surface monitoring using smartphone sensors and crowdsourcing. Pervasive Mob. Comput. 40, 71–88. https://doi.org/10.1016/j.pmcj.2017.06.002

Statistics Canada, 2024. Canada’s core public infrastructure survey: Replacement values. https://www150.statcan.gc.ca/n1/daily-quotidien/241021/dq241021b-eng.htm

Talaghat, M.A., Golroo, A., Kharbouch, A., Rasti, M., Heikkilä, R., Jurva, R., 2024. Digital twin technology for road pavement. Autom. Constr. 168, 105826. https://doi.org/10.1016/j.autcon.2024.105826

Tao, F., Zhang, H., Liu, A., Nee, A.Y.C., 2019. Digital twin in industry: State-of-the-art. IEEE Trans. Ind. Inform. 15, 2405–2415. https://doi.org/10.1109/TII.2018.2873186

Valipour, P.S., Golroo, A., Kheirati, A., Fahmani, M., Amani, M.J., 2024. Automatic pavement distress severity detection using deep learning. Road Mater. Pavement Des. 25, 1830–1846. https://doi.org/10.1080/14680629.2023.2276422

Wu, C., Wang, Z., Hu, S., Lepine, J., Na, X., Ainalis, D., Stettler, M., 2020. An automated machine-learning approach for road pothole detection using smartphone sensor data. Sensors (Switzerland) 20, 5564. https://doi.org/10.3390/s20195564

Yaseen, M., 2024. What is YOLOv8: An in-depth exploration of the internal features of the next-generation object detector. arXiv preprint arXiv:2408.15857. https://doi.org/10.48550/arXiv.2408.15857

Zhong, J., Zhu, J., Huyan, J., Ma, T., Zhang, W., 2022. Multi-scale feature fusion network for pixel-level pavement distress detection. Autom. Constr. 141, 104436. https://doi.org/10.1016/j.autcon.2022.104436

Zhu, J., Zhong, J., Ma, T., Huang, X., Zhang, W., Zhou, Y., 2022. Pavement distress detection using convolutional neural networks with images captured via UAV. Autom. Constr. 133, 103991. https://doi.org/10.1016/j.autcon.2021.103991