Innovative Journal of Computing, Security and Digital Technologies
Volume 1, Issue 3, Sep – Dec 2026
Page no : 25-43

Computer Vision-Based Automated Detection of Road Surface Defects for Intelligent Transportation Infrastructure

Asmitha Kumari, H. Sarker

Abstract

Road surface defects such as cracks, potholes and pavement deterioration are important factors affecting road safety, transportation efficiency and infrastructure maintenance. Conventional road inspection methods generally depend on manual observation and periodic field surveys, which may require considerable time, labour and financial resources. Recent developments in computer vision and deep learning provide opportunities for automating road condition assessment through digital images captured using vehicle-mounted cameras and mobile devices. This paper presents a computer vision-based framework for automated detection and classification of road surface defects for intelligent transportation infrastructure. The proposed framework uses image preprocessing, deep-learning-based object detection and defect localization to identify common pavement damages, including longitudinal cracks, transverse cracks, alligator cracks and potholes. A YOLO-based detection architecture is considered to support efficient image analysis and potential real-time deployment. The proposed approach also incorporates image enhancement, data augmentation, multi-scale feature extraction and confidence-based detection filtering to improve robustness under varying environmental and road conditions. The RDD2022 dataset is considered for model development and evaluation because of its multi-national road image collection and annotated pavement damage categories. Detection performance is intended to be evaluated using precision, recall, mean average precision, inference time and model size. The framework provides a basis for integrating automated pavement inspection with intelligent transportation systems, supporting road maintenance prioritization, infrastructure monitoring and data-driven maintenance planning.

Keywords: Computer vision, road surface defects, deep learning, YOLO, pavement crack detection, pothole detection, object detection, intelligent transportation systems, road infrastructure monitoring

Publication Date : 30.09.2026

I. INTRODUCTION

Road transportation infrastructure plays a significant role in economic development, industrial activity, public mobility and regional connectivity. High-quality road networks support the movement of passengers and goods while contributing to efficient transportation operations. However, continuous traffic loading, environmental exposure, poor drainage, construction deficiencies and material deterioration can gradually damage pavement surfaces.

Common road surface defects include longitudinal cracks, transverse cracks, alligator cracking, potholes, surface deformation and pavement deterioration. These defects may

reduce driving comfort, affect vehicle stability and increase maintenance requirements. Severe pavement damage can also contribute to unsafe driving conditions, particularly when defects are not identified and repaired in a timely manner.

Traditional road inspection processes generally involve visual examination by maintenance personnel, manual recording of pavement conditions and periodic field surveys. Although these approaches remain useful, they can be time-consuming and may produce inconsistent assessments because of differences in inspector experience, environmental conditions and inspection frequency.

The development of digital cameras, mobile devices, vehicle-mounted imaging systems and artificial intelligence has created opportunities for automated road condition assessment. Computer vision enables road images to be analysed automatically to identify and classify pavement defects. Deep learning, particularly convolutional neural networks and object detection architectures, has improved the ability of automated systems to recognize complex visual patterns.

Maeda et al. [1] demonstrated the application of deep neural networks for detecting and classifying road damage using images collected from smartphones installed in vehicles. Their work established an important foundation for automated road inspection using image-based deep learning.

Subsequent research introduced publicly available datasets containing annotated road damage images. Arya et al. [2] presented RDD2020, which supports the development of deep-learning-based road damage detection methods. The RDD2022 dataset subsequently expanded the geographical coverage of road damage images and introduced a larger multi-national benchmark [3]. These datasets provide opportunities to develop and evaluate automated pavement inspection systems under different road conditions.

Computer vision research has also examined different approaches to pavement inspection, including conventional image processing, machine learning and deep learning. Chen et al. [4] reviewed automated road surface defect detection techniques and identified challenges associated with image acquisition, dataset variation, defect classification and model evaluation.

Although deep-learning-based detection provides promising capabilities, practical deployment still faces challenges involving illumination changes, shadow effects, motion blur, small cracks, occlusion, varying road textures and differences between geographical regions. A model trained using one road environment may not provide consistent results when applied to another environment without suitable evaluation and adaptation.

This paper presents a structured computer vision framework for automated road surface defect detection. The framework focuses on detecting common pavement defects from road images, comparing object detection architectures and preparing visual outputs that can support intelligent transportation infrastructure management.

A. Research Objectives

The primary objectives of this study are to develop a computer vision framework for automated road surface defect detection, identify and classify common pavement defects

using deep-learning-based object detection, investigate image preprocessing and augmentation techniques, evaluate detection performance using standard object detection metrics and examine the computational efficiency of the framework for practical transportation applications.

B. Research Contributions

The proposed research focuses on an integrated road inspection framework combining image acquisition, preprocessing, feature extraction and object detection. It considers multiple pavement defect categories and establishes a comparative evaluation strategy for different deep-learning detection architectures.

The framework also considers computational efficiency and the potential integration of defect detection outputs with road maintenance information systems. This approach connects computer vision techniques with practical civil engineering and intelligent transportation applications.

C. Organization of the Paper

The remainder of this paper is organized as follows. Section II presents the literature review. Section III describes the proposed system architecture. Section IV explains dataset preparation and image preprocessing. Section V presents the deep learning detection methodology. Section VI describes the experimental design and evaluation metrics. Section VII discusses the evaluation framework and practical implications. Section VIII presents implementation considerations and limitations. Section IX discusses future research directions, followed by the conclusion and references.

II. LITERATURE REVIEW

A. Traditional Road Surface Inspection

Conventional pavement inspection approaches depend on field observations, manual measurements and established pavement distress classification procedures. Inspectors generally identify pavement defects and record their type, location, severity and extent.

Although manual inspection provides useful engineering information, large-scale road networks require frequent observations across different locations. Manual approaches can create operational challenges involving personnel requirements, inspection frequency, observation consistency and data management.

Computer vision offers an opportunity to automate some of these inspection tasks while maintaining visual records that can be reviewed by road maintenance authorities.

B. Computer Vision-Based Road Damage Detection

Traditional image processing methods use edge detection, thresholding, morphological operations, texture analysis and image segmentation to identify pavement defects.

These techniques can be useful when defects are clearly visible and the image background is relatively consistent. However, road surfaces contain complex textures, shadows, lighting variations, stains and markings that can affect image processing results.

Chen et al. [4] reviewed computer vision and machine learning methods for automated road surface defect and distress detection. Their review examined image acquisition, dataset development, computer vision techniques and deep learning approaches. It also highlighted challenges involving generalization, detection reliability and practical road inspection environments.

Modern computer vision systems increasingly use convolutional neural networks to extract complex visual features directly from images. These features enable the identification of pavement defects under a wider range of visual conditions than basic image processing methods.

C. Deep Learning-Based Object Detection

Object detection systems identify objects within an image and determine their approximate locations using bounding boxes.

Deep-learning-based object detection methods can broadly be divided into two categories.

Two-stage detectors generally identify candidate object regions and subsequently classify and refine them. Faster R-CNN is a representative example of this approach. These models can provide strong detection performance but may require additional computational resources.

One-stage detectors perform object localization and classification within a unified detection process. YOLO-based architectures are widely investigated for applications requiring efficient image analysis.

Yao et al. [5] investigated an optimized YOLO-based approach for automatic pavement defect detection. Their study compared Faster R-CNN with YOLOv5 and YOLOv7 and examined the balance between detection performance and model complexity. The results reported in their specific experiments demonstrate that a carefully optimized YOLO approach can provide a practical alternative for pavement inspection applications.

These studies indicate that model selection should consider both detection performance and computational efficiency rather than relying on a single performance measure.

D. Road Damage Detection Datasets

A labelled image dataset is essential for training and evaluating supervised road damage detection models.

The RDD2020 dataset contains 26,336 road images collected from India, Japan and the Czech Republic, with more than 31,000 annotated road damage instances [2]. Its images were collected using vehicle-mounted smartphones, making it relevant to mobile road inspection systems.

The RDD2022 dataset expanded the collection to 47,420 road images from six countries, including India, Japan, the Czech Republic, Norway, the United States and China. The dataset contains more than 55,000 annotated road damage instances covering four principal categories: longitudinal cracks, transverse cracks, alligator cracks and potholes [3].

Arya et al. [6] also documented the Crowdsensing-based Road Damage Detection Challenge, which provided a common evaluation environment for researchers investigating road damage detection approaches.

E. Real-Time Pavement Defect Detection

Real-time detection is important for practical applications involving vehicle-mounted cameras, mobile inspection systems and edge computing devices.

As Sami et al. [7] proposed an improved YOLOv5-based approach for road pavement damage detection and infrastructure management. Their work examined model improvements intended to enhance detection performance while supporting practical automated inspection.

Cano-Ortiz et al. [8] developed an end-to-end deep learning-based computer vision system for pavement distress detection and quantification. Their study considered YOLO-based models and post-processing techniques to reduce false detections.

These developments indicate that road inspection systems should be evaluated not only by their detection accuracy but also by inference speed, computational complexity and deployment requirements.

F. Research Gap

The existing literature demonstrates the potential of computer vision and deep learning for automated road surface defect detection. However, several challenges remain.

First, road images collected from different geographical locations may contain considerable differences in pavement materials, road markings, lighting and environmental conditions.

Second, small cracks can be difficult to distinguish from ordinary pavement texture, shadows and surface stains.

Third, class imbalance may affect the ability of a detection model to identify less frequently occurring defects.

Fourth, high detection accuracy does not automatically guarantee real-time performance on resource-constrained devices.

Finally, detecting a pavement defect is different from assessing its engineering severity. A detected pothole or crack requires additional assessment to determine its dimensions, condition and maintenance priority.

The proposed research addresses these challenges through a systematic detection framework that combines image preprocessing, multi-scale feature extraction, object localization and performance evaluation.

III. PROPOSED SYSTEM ARCHITECTURE

The proposed framework, referred to as the Computer Vision-Based Road Inspection System, is designed to automate the identification and localization of road surface defects using digital images.

The system uses a deep-learning-based object detector to process road images and generate defect classifications with corresponding bounding boxes. The resulting information can be integrated with location and inspection records to support road maintenance activities.

A. Image Acquisition Module

The image acquisition module provides visual information about the road surface. Vehicle-mounted cameras are suitable for capturing images continuously during road inspection.

The camera installation should be selected to provide a clear view of the pavement surface. The acquisition configuration should consider camera height, viewing angle, road coverage, image resolution and vehicle movement.

The image collection process should cover different lighting conditions, pavement materials, road environments and camera perspectives. Images collected during daylight, low-light conditions and varying weather conditions can help evaluate the robustness of the proposed detection framework.

The collected images should be reviewed to remove corrupted frames, excessively blurred images and irrelevant road scenes.

B. Image Preprocessing Module

Image preprocessing prepares the collected images for deep-learning-based analysis.

Image resizing provides a consistent input resolution for the selected detection architecture. Image normalization improves input consistency during model training.

Contrast enhancement may be applied to improve the visibility of defects in images affected by uneven illumination. However, excessive enhancement should be avoided because it can introduce artificial visual patterns.

Image quality filtering is used to identify blurred, corrupted and irrelevant images. This reduces the possibility of training the model using unsuitable visual information.

Data augmentation is also considered to increase the diversity of training images. Brightness changes, contrast adjustments, horizontal flipping, limited rotation, scaling and controlled blur can be used to simulate different road inspection conditions.

Augmentation should preserve the original defect identity and must not introduce unrealistic pavement patterns.

C. Feature Extraction Module

Feature extraction is an important component of deep-learning-based object detection. The convolutional backbone processes the input image and learns visual representations at different spatial scales.

Low-level features represent edges, local texture and crack boundaries. Intermediate features capture repeated surface patterns and localized damage structures. Higher-level features represent larger pavement defects and contextual information.

Multi-scale feature extraction is particularly important because road defects vary considerably in size. Fine cracks may occupy only a small portion of an image, while potholes and alligator cracking may cover larger surface regions.

The proposed framework uses multi-level feature representations to support the identification of defects with different spatial characteristics.

D. Object Detection Module

The object detection module identifies the presence and location of pavement defects.

A YOLO-based architecture is selected as the principal detection approach because it performs object localization and classification within a unified detection pipeline. Its suitability should be evaluated against alternative architectures under consistent experimental conditions.

The output for each detected defect includes the predicted defect category, bounding-box location, confidence score and corresponding image identification.

The four primary defect categories considered in this study are longitudinal cracks, transverse cracks, alligator cracks and potholes.

E. Detection Filtering and Output Generation

Raw object detector predictions may contain overlapping bounding boxes or low-confidence detections.

Confidence-based filtering is applied to remove predictions that do not meet the selected detection criteria. Non-maximum suppression is then used to reduce overlapping bounding boxes associated with the same defect.

The output module generates an annotated road image containing detected defect regions, category names and confidence scores.

Where GPS information is available, image records can be associated with geographic coordinates and road-segment identifiers. This supports the development of location-based road inspection reports.

F. Integration with Intelligent Transportation Infrastructure

The proposed system can be integrated with intelligent transportation infrastructure management platforms.

The integration layer receives image-based inspection outputs, associates defects with road locations and organizes inspection records by road segment.

A maintenance dashboard may present road inspection locations, detected defect categories, number of detections, inspection dates, image-based evidence and locations requiring engineering verification.

The system should be considered a decision-support tool. Structural severity and repair priority require additional engineering assessment and should not be inferred solely from a bounding-box confidence score.

IV. DATASET PREPARATION AND IMAGE PROCESSING

A. Dataset Selection

The RDD2022 dataset is proposed as the primary benchmark for this investigation. It contains road images collected from six countries and annotated road damage instances, providing a useful basis for examining road damage detection across different geographical environments [3].

RDD2020 can also be considered as an additional benchmark. Its images from India, Japan and the Czech Republic provide an opportunity to examine model performance across different road environments [2].

The RDD2022 dataset is recommended for the primary experiment because of its wider geographical coverage. RDD2020 can be used for supplementary testing or cross-dataset evaluation.

B. Road Defect Classification

The proposed study considers four primary damage categories corresponding to commonly used road damage dataset labels.

1) Longitudinal Cracks: These cracks generally develop along the longitudinal direction of pavement. They may be associated with construction joints, pavement fatigue, environmental effects or structural deterioration.

2) Transverse Cracks: These cracks extend across the pavement and may develop due to temperature variations, material shrinkage and pavement movement.

3) Alligator Cracks: Alligator cracking is characterized by interconnected cracks forming a pattern resembling an animal skin network. It is often associated with fatigue-related pavement deterioration.

4) Potholes: Potholes are localized surface cavities resulting from material loss. They may develop due to water infiltration, traffic loading and progressive pavement deterioration.

The model should be trained to distinguish these defect categories while accounting for their different visual characteristics and spatial dimensions.

C. Dataset Annotation

The quality of image annotations directly influences supervised model training.

Each defect should be labelled using a bounding box that covers the visible damage region. The annotation should include the corresponding defect category and image identification.

Annotation guidelines should be prepared to maintain consistency across different images and annotators.

The bounding boxes should closely cover the visible defect. Multiple defects should receive separate annotations where distinguishable. Overlapping damage patterns should be labelled consistently, and unclear or heavily obscured defects should be reviewed.

A sample of annotated images should be independently reviewed to identify inconsistent class assignments or inaccurate bounding boxes.

D. Dataset Splitting

The dataset should be divided into training, validation and testing subsets.

The training subset is used for learning model parameters. The validation subset supports hyperparameter selection and model monitoring. The testing subset is reserved for final performance evaluation.

The images should be divided in a way that minimizes data leakage. Where images originate from the same inspection sequence or road segment, related images should preferably remain within the same subset.

A separate geographical evaluation should also be considered to test whether the detector generalizes to locations not represented in the training data.

E. Data Augmentation

Road images may contain substantial visual variation caused by illumination, road material, shadows, weather and camera movement.

To improve model robustness, the proposed training process considers brightness variation, contrast adjustment, horizontal flipping, small-angle rotation, image scaling and mild blur augmentation.

Controlled weather-related augmentation may also be evaluated where the transformation remains realistic.

Augmentation should be applied only to the training subset. Validation and test images should retain their original visual characteristics to provide an unbiased evaluation.

F. Class Distribution Analysis

The distribution of defect categories should be examined before model training.

Some defect types may be more frequently represented than others. This can create class imbalance and may cause the model to perform better on common defects than on rare categories.

The dataset annotations should be analysed to determine the exact class frequencies. Suitable approaches such as class-aware sampling, targeted augmentation and per-class performance reporting may then be adopted.

V. DEEP LEARNING MODEL DEVELOPMENT

A. Selection of Detection Architecture

The proposed study considers a YOLO-based object detector as the primary architecture. YOLO models perform object localization and classification in a unified detection process and can be evaluated for road inspection applications that require efficient image analysis.

The model contains three principal components: a backbone network, a feature aggregation component and a detection head.

The backbone extracts visual information from road images. The feature aggregation component combines information from different spatial scales. The detection head predicts pavement defect categories and their locations.

The selection of the particular YOLO version should be based on implementation availability, hardware requirements and reproducibility.

B. Backbone Feature Extraction

The backbone network extracts the visual characteristics required to recognize pavement defects.

Early feature layers identify edges, texture variations and local intensity patterns. Deeper layers capture more complex structures, including crack networks, irregular surface boundaries and pothole regions.

Small cracks require sufficiently detailed visual features, whereas larger defects can be identified through higher-level spatial representations.

Feature extraction should therefore retain sufficient information across different spatial resolutions.

C. Multi-Scale Feature Processing

Road defects differ in size, appearance and distribution. A fine longitudinal crack may be represented by a narrow and extended visual region, whereas a pothole may occupy a broader and irregular area.

Multi-scale feature processing supports the detection of these different visual patterns by combining information obtained from multiple feature levels.

The proposed approach considers multi-scale feature aggregation to improve the representation of small and large road defects.

D. Detection Head

The detection head generates the predicted defect class, bounding-box coordinates and confidence score.

The model should distinguish between the four target classes and background road regions. Its predictions are subsequently processed using the selected confidence threshold and non-maximum suppression settings.

The confidence threshold should be selected using validation data rather than being chosen solely on the basis of test performance.

E. Training Configuration

The model should be trained using the prepared training dataset and monitored through validation performance.

Important training parameters include image input resolution, batch size, learning rate, optimizer, number of epochs, augmentation settings and confidence-related evaluation parameters.

Transfer learning may be adopted using pretrained model weights to reduce training requirements and improve feature initialization.

The training process should be repeated under documented configurations to support reproducibility.

The best checkpoint should be selected using the validation dataset. The testing dataset should remain unused during model selection.

F. Comparative Model Development

To evaluate the suitability of the proposed detector, a comparative study can be conducted using Faster R-CNN, YOLOv5 and a selected newer YOLO architecture.

Faster R-CNN represents a two-stage detection approach, while YOLO models represent one-stage detection approaches.

All models should be evaluated using identical training, validation and test partitions wherever technically possible. Input image resolution, preprocessing, annotation quality and evaluation procedures should also be controlled.

The comparison should examine detection precision, recall, mean average precision, inference time, model size and computational resource requirements.

G. Confidence-Based Detection Filtering

Confidence-based filtering is used to remove low-confidence predictions that are more likely to represent false detections.

The threshold should be selected using the validation dataset and should consider the balance between false positives and false negatives.

Non-maximum suppression reduces overlapping predictions by retaining the most appropriate detection among highly overlapping bounding boxes.

The threshold configuration should be documented because different operating thresholds can produce different precision and recall outcomes.

VI. EXPERIMENTAL DESIGN AND PERFORMANCE EVALUATION

A. Experimental Environment

The proposed system can be implemented using Python and a deep learning framework such as PyTorch. OpenCV may be used for image preprocessing and visualization, while an appropriate GPU environment may be used for model training and inference.

The experimental setup should clearly document the processor, GPU, available memory, operating system, software versions, library dependencies and model configuration.

For deployment testing, the trained model may also be evaluated on a CPU or edge computing device to examine its suitability for practical road inspection systems.

B. Training and Validation Procedure

The training process begins with loading the labelled road image dataset and preparing the selected training, validation and testing subsets.

Images are preprocessed and augmented according to the experimental configuration. The detection model is initialized with the selected architecture and training settings.

During training, the model learns to identify road damage patterns from annotated images. Validation performance is monitored after each training stage to identify overfitting and guide model selection.

The best-performing validation checkpoint is selected for final evaluation.

The testing subset is used only after model selection to obtain an unbiased estimate of detection performance.

C. Performance Evaluation Metrics

The proposed system is evaluated using multiple object detection performance metrics.

1) Precision: Precision measures the proportion of predicted defect detections that correspond to actual annotated defects. High precision indicates that the model produces fewer false detections.

2) Recall: Recall measures the proportion of actual annotated defects successfully detected by the model. High recall is particularly important for road inspection because missed pavement defects may remain unidentified.

3) F1-Score: The F1-score provides a combined assessment of precision and recall. It is useful when evaluating the balance between false positives and false negatives.

4) Mean Average Precision: Mean average precision is used to evaluate object detection performance across defect categories and selected intersection-over-union thresholds. Results should clearly state whether mAP at a particular threshold or a range of thresholds is reported.

5) Inference Time: Inference time measures the time required by the model to process an image. It is important for applications involving continuous road inspection.

6) Model Size: Model size provides information about storage and deployment requirements. Smaller models may be more suitable for edge computing devices with limited resources.

7) Per-Class Performance: Individual precision, recall and average precision values should be reported for longitudinal cracks, transverse cracks, alligator cracks and potholes.

D. Robustness Evaluation

The model should also be evaluated under different environmental and imaging conditions.

Daylight, low-light, shadowed, wet and visually complex pavement conditions may influence the quality of defect detection.

Images from different geographical regions can be used to examine the ability of the model to generalize across road environments.

The evaluation should identify conditions where detection performance decreases and determine whether additional training data or preprocessing improvements are required.

E. Error Analysis

Error analysis is necessary to understand the limitations of the detector.

False positives may occur when surface markings, road stains, repaired pavement sections or shadows resemble actual defects.

False negatives may occur when cracks are very small, poorly illuminated, partially obscured or visually similar to normal pavement texture.

Classification errors may occur when alligator cracking and other interconnected crack patterns exhibit similar visual structures.

The predicted bounding boxes should therefore be reviewed alongside the original road images and ground-truth annotations.

F. Computational Efficiency

Computational efficiency is important for deployment in mobile road inspection systems.

The model should be evaluated using inference time, memory consumption, model size and hardware utilization.

GPU-based performance provides an indication of processing capacity under accelerated conditions. CPU-based evaluation helps determine whether the system can operate on devices with limited hardware resources.

The evaluation should also consider image acquisition rate and the time required to generate inspection reports.

VII. RESULTS AND DISCUSSION

The proposed framework requires actual model training and test evaluation before experimental detection results can be reported. Therefore, this section distinguishes previously published findings from the proposed system's pending evaluation.

A. Published Research Evidence

Maeda et al. [1] demonstrated the feasibility of road damage detection using vehicle-mounted smartphone images and deep neural networks. Their work established a practical foundation for image-based pavement inspection and demonstrated the use of object detection methods in road damage classification.

Arya et al. [2], [3] introduced the RDD2020 and RDD2022 datasets, providing annotated road images that support the development and comparison of automated detection models.

The RDD2022 challenge also provided a shared evaluation environment for comparing road damage detection approaches. Arya et al. [6] reported the leading solutions submitted to the challenge, including a model ensemble combining YOLO and Faster R-CNN approaches. The best-performing solution reported an F1-score of 76% across the combined test images from the six participating countries. This figure represents the challenge results and must not be interpreted as the performance of the proposed system.

Yao et al. [5] compared different object detection architectures for pavement defect detection. Their experiments reported that Faster R-CNN achieved a mean average

precision of 0.938, while YOLOv5 and YOLOv7 achieved 0.910 and 0.932, respectively. Their optimized YOLO-based method achieved a reported mean average precision of 0.930. These values are specific to their experimental dataset and configuration.

As Sami et al. [7] investigated improvements to YOLOv5 for real-time road pavement damage detection. Their work examined attention mechanisms, model modifications and training strategies to support automated road infrastructure inspection.

Cano-Ortiz et al. [8] developed an end-to-end computer vision system for pavement distress detection and quantification. Their work also examined rule-based post-processing techniques for reducing false detections.

These publications establish the feasibility of automated road damage detection while demonstrating that model performance depends on the dataset, architecture, training configuration and evaluation environment.

B. Expected Evaluation of the Proposed Framework

The proposed system is expected to be evaluated for its ability to identify and localize four categories of road surface defects.

The evaluation should determine whether the selected detection architecture can recognize both narrow crack patterns and larger pavement defects.

The ability to detect small defects is an important consideration because fine cracks may occupy a relatively small portion of an input image. High-resolution image processing and suitable multi-scale feature extraction may improve the representation of these defects.

Pothole detection requires the model to recognize irregular surface boundaries and local pavement deterioration. Performance may be influenced by shadows, standing water and differences in road texture.

Alligator crack detection may be challenging because the interconnected crack structure can vary significantly between images.

C. Precision and Recall Analysis

Precision and recall should be considered together when evaluating road damage detection.

A system that produces high precision but low recall may miss a considerable number of actual pavement defects. Conversely, a system with high recall but low precision may generate excessive false alarms.

For road infrastructure monitoring, the balance between these metrics is important because missed defects may delay maintenance identification, while false detections may increase inspection workload.

The operating threshold should be selected according to the intended application and the cost associated with missed defects and incorrect detections.

D. Detection Performance Across Defect Categories

The four defect classes require different visual representations.

Longitudinal cracks generally appear as elongated patterns aligned with the road direction. Transverse cracks extend across the pavement and may vary in length and width.

Alligator cracking contains multiple connected crack patterns. Potholes present localized regions with irregular boundaries and varying surface characteristics.

The proposed system should report performance separately for each category to identify classes that require additional training data or model refinement.

E. Generalization Across Road Environments

The use of multi-national road image datasets creates an opportunity to evaluate model generalization.

Different regions may contain differences in pavement materials, road maintenance practices, environmental conditions and surface markings.

A model trained using images from one geographical environment may not perform consistently in another environment.

Cross-country evaluation is therefore recommended to determine whether the model can generalize across different road conditions.

F. Practical Implications

The proposed framework can support road inspection activities by reducing the dependency on manual image-by-image defect identification.

Automated detection outputs can provide a preliminary inventory of road damage and support the organization of maintenance observations.

When combined with GPS information, inspection images can be associated with specific road locations. This enables maintenance departments to identify defect clusters and organize further engineering inspections.

However, automated detection should not replace detailed pavement condition assessment. Defect severity, structural condition and repair priority should be determined using appropriate engineering inspection methods.

VIII. IMPLEMENTATION CONSIDERATIONS AND LIMITATIONS

A. Practical Deployment

The proposed computer vision system can be integrated with vehicle-mounted cameras, mobile inspection devices and road monitoring platforms.

A vehicle-mounted camera can capture road images during routine transportation activities. The image processing module can analyse captured images and identify potential pavement defects.

The system can generate digital inspection records containing the detected defect type, confidence score, image location and inspection timestamp.

These records can be transmitted to a central maintenance platform for further examination.

B. Edge Computing Requirements

Real-time road inspection may require deployment on hardware with limited computational resources.

A large detection model may require significant memory and processing capacity. Model compression, optimized inference engines and lightweight architectures can be considered to support deployment on edge devices.

Quantization and pruning may also be investigated, provided that detection performance is evaluated after optimization.

C. Integration with Geographic Information Systems

The integration of computer vision outputs with Geographic Information Systems can improve the usefulness of automated road inspection.

GPS coordinates can be associated with individual road images and predicted defects. The resulting records may be organized by road segment and displayed using a spatial inspection interface.

This can help authorities identify road sections requiring additional inspection.

However, image-based detection does not directly provide pavement roughness, structural capacity or the full engineering severity of a defect. Additional sensors and field measurements may be required for comprehensive pavement condition assessment.

D. Limitations

The proposed approach has several limitations.

First, the performance of the model depends on the quality and diversity of the training dataset. Inadequate representation of specific defect conditions may reduce detection reliability.

Second, changes in lighting, weather, shadows and camera orientation can affect image quality and model performance.

Third, small cracks may be difficult to identify when the input resolution is insufficient.

Fourth, the RDD2022 dataset contains data from multiple geographical regions, but its distribution may not fully represent the road conditions of every local transportation network.

Fifth, bounding-box detection provides localization information but does not necessarily provide accurate crack length, width, depth or structural severity.

Finally, computational efficiency must be verified on the actual deployment hardware before claiming real-time performance.

IX. FUTURE RESEARCH DIRECTIONS

Future research can extend the proposed framework in several directions.

A. Advanced Object Detection Models: Newer object detection architectures and lightweight feature extraction networks can be investigated to improve detection efficiency and robustness.

B. Small Crack Detection: High-resolution image processing and specialized feature extraction methods can be examined to improve the identification of fine pavement cracks.

C. Video-Based Road Inspection: Continuous video processing can be investigated to detect road damage across successive frames and reduce duplicate detections.

D. Geographic Generalization: Additional datasets from different regions can be used to evaluate the ability of the model to generalize across geographical and environmental conditions.

E. Defect Severity Assessment: Future systems can incorporate image segmentation and pavement measurement techniques to estimate defect dimensions and support engineering-based severity classification.

F. Intelligent Maintenance Integration: Detection outputs can be linked with GIS-based road management platforms to support maintenance prioritization and inspection planning.

G. Edge-Based Deployment: Lightweight models and optimized inference systems can be investigated for implementation on vehicle-mounted computing devices.

X. CONCLUSION

This paper presented a computer vision-based framework for automated road surface defect detection for intelligent transportation infrastructure. The proposed framework combines digital image acquisition, image preprocessing, deep-learning-based feature extraction, object detection and output generation to identify common pavement defects.

The study considered longitudinal cracks, transverse cracks, alligator cracks and potholes as the primary detection categories. The RDD2022 dataset was identified as a suitable benchmark because of its multi-national image collection and annotated road damage instances.

A YOLO-based detection architecture was considered for object localization and classification, while preprocessing, image augmentation, multi-scale feature extraction and confidence-based filtering were included as important components of the proposed approach.

The evaluation framework considers precision, recall, F1-score, mean average precision, inference time and model size. These measures are important for assessing both detection quality and computational efficiency.

Previous research demonstrates that deep learning can support automated pavement damage identification using vehicle-mounted and other digital imaging systems. However, practical deployment requires further investigation of generalization, environmental robustness, small defect detection and hardware efficiency.

The proposed framework provides a foundation for integrating computer vision with intelligent transportation infrastructure management. Future experimental validation using public road damage datasets and actual model training is necessary to establish detection performance and deployment suitability.

REFERENCES

1. H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road damage detection and classification using deep neural networks with smartphone images," Computer-Aided Civil and Infrastructure Engineering, vol. 33, no. 12, pp. 1127–1141, 2018, doi: 10.1111/mice.12387.
2. D. Arya, H. Maeda, S. K. Ghosh, D. Toshniwal, and Y. Sekimoto, "RDD2020: An annotated image dataset for automatic road damage detection using deep learning," Data in Brief, vol. 36, Art. no. 107133, 2021, doi: 10.1016/j.dib.2021.107133.
3. D. Arya et al., "RDD2022: A multi-national image dataset for automatic road damage detection," Geoscience Data Journal, vol. 11, no. 4, pp. 846–862, 2024, doi: 10.1002/gdj3.260.
4. X. Chen, S. Yongchareon, and M. Knoche, "A review on computer vision and machine learning techniques for automated road surface defect and distress detection," Semantic Computing, vol. 1, no. 4, 2023, doi: 10.3233/SCS-230001.
5. H. Yao, Y. Fan, X. Wei et al., "Research and optimization of YOLO-based method for automatic pavement defect detection," Electronic Research Archive, vol. 32, no. 3, pp. 1708–1730, 2024, doi: 10.3934/era.2024078.
6. D. Arya, H. Maeda, S. K. Ghosh, D. Toshniwal, H. Omata, T. Kashiyama, and Y. Sekimoto, "Crowdsensing-based Road Damage Detection Challenge (CRDDC'2022)," in Proc. 2022 IEEE International Conference on Big Data (Big Data), Osaka, Japan, 2022, pp. 6378–6386, doi: 10.1109/BigData55660.2022.10021040.
7. Bajarang Bhagwat, V. (2023). Optimizing Payroll to General Ledger Reconciliation: Identifying Discrepancies and Enhancing Financial Accuracy. JOURNAL OF ADVANCE AND FUTURE RESEARCH, 1 (4).
8. Koneni, C., & Lokula, S. (2026, June). Silent Data Corruption in Distributed Genai-Augmented Data Pipelines: an Empirical Study of Hallucination-Induced Failure Propagation. In 2026 IEEE 46th International Conference on Distributed Computing Systems Workshops (ICDCSW) (pp. 137-143). IEEE.
9. Addepalli, D. N. V. (2026). An Autonomous Agentic Ai Framework For Coverage-Guided Chaos Engineering And Resilience Optimisation In Cloud-Native Distributed Systems. International Journal of Artificial Intelligence and Computer Electronics, 2(1), 80-89.
10. Addepalli, D. N. V. (2026). An Autonomous Agentic Ai Framework For Coverage-Guided Chaos Engineering And Resilience Optimisation In Cloud-Native Distributed Systems. International Journal of Artificial Intelligence and Computer Electronics, 2(1), 80-89.
11. Koneni, C., & Lokula, S. (2026, June). Silent Data Corruption in Distributed Genai-Augmented Data Pipelines: an Empirical Study of Hallucination-Induced Failure

Propagation. In 2026 IEEE 46th International Conference on Distributed Computing Systems Workshops (ICDCSW) (pp. 137-143). IEEE.

12. Pal, A. (2026). Productivity vs. Compliance: The New Engineering Challenge of AI Coding Assistants in Regulated Codebases. International Journal of Computer Information Systems & Industrial Management Applications, 18, 1248.
13. Rashy, R. A. (2026). Empirical Validation of Predictive Project Controls for Early Schedule-Risk Detection: Analysis of CPM Float Erosion and Milestone Performance in Construction Projects. International Journal of Artificial Intelligence and Computer Electronics, 2(4), 17-24.
14. Chayka, D. P. (2026). An intelligent framework for customer-centric packaging optimization and operational efficiency across industries. International Journal of Production Research (IJPR), 3(2), 1-19.
15. Chayka, D. P. (2026). The Role of Packaging Automation in Improving Operational Efficiency, Productivity, and Manufacturing Performance in Corrugated Packaging Operations. International Journal of Innovation in Electronics and Computer Information Technology, 3(4), 1-8.