

Deep Learning-Based Road Surface Defect Detection Using Image Processing Techniques

Shivangkumar R Patel¹, Karan P Bhatt², Jwalant Babubhai Baria³, Vanraj Kumar Dineshkumar Baria⁴, Ashok K Rathva⁵, Yagnik A Rathod⁶

¹Assistant Professor, Computer Engineering Department, Government Engineering College, Modasa patelshivang82@gmail.com

²Assistant Professor, Computer Engineering Department, Vishwakarma Government Engineering College, Chandkheda kpbhatt@vgecg.ac.in

³Assistant Professor, Computer Engineering Department, Government Engineering College, Dahod jwalant.baria@gmail.com

⁴Assistant Professor, Computer Engineering Department, Government Engineering College, Dahod vdbaria.ce@gmail.com

⁵Assistant Professor, Computer Engineering Department, Government Engineering College, Dahod akr.coed@gecdahod.ac.in

⁶Assistant Professor, Computer Engineering Department, Government Engineering College, Dahod yar.coed@gecdahod.ac.in

Abstract

The importance of detecting road surface defects (cracks and potholes) promptly is crucial to transportation safety, minimisation of maintenance plans and minimisation of infrastructure deterioration costs. Conventional inspection techniques are time-consuming, subjective, and can hardly be used to offer scalable solutions to large road networks. The recent developments in deep learning and computer vision have shown that automated pavement distress detection is highly promising. The proposed research is a deep learning framework of road surface defect detection that combines image preprocessing with the lightweight object detection networks. The suggested methodology will also use contrast enhancement, noise reduction, and data augmentation to enhance the model's strength when faced with different illuminations and environmental conditions. Real-time defect localisation and classification make use of a convolutional neural network-based detection architecture, which is based on the state-of-the-art object detection models. The evaluation of the performance is based on typical measures such as precision, recall, F1-score, and the mean Average Precision (mAP). The superiority in accuracy of detection and the computational efficiency is proven by the comparisons with the already developed deep learning methods. The findings show that the combination of image enhancement methods and optimised deep neural networks can help a great deal to achieve detection reliability in real-world scenarios. The suggested architecture offers a highly scalable and low-cost solution to an intelligent pavement monitoring system and smart infrastructure management.

Keywords: Road Surface Defect Detection, Deep Learning, Pavement Crack Detection, Object Detection Networks, Image Processing Techniques, YOLO-based Detection, Smart Infrastructure Monitoring.

1. Introduction

Road transportation infrastructure forms the backbone of economic development, logistics efficiency, and urban mobility systems. The performance and durability of pavement networks directly influence trade flows, emergency response time, public safety, and long-term infrastructure sustainability. However, continuous exposure to dynamic traffic loading, environmental stresses such as temperature fluctuations and precipitation, material fatigue, and aging progressively degrade pavement surfaces. This deterioration manifests in the form of cracks, potholes, ravelling, surface deformation, and

structural failures. Early-stage detection of such defects is critical to prevent structural escalation, reduce maintenance costs, and ensure road-user safety.

Traditionally, pavement inspection has relied on manual surveying or semi-mechanized visual assessments. Although widely practiced, these approaches are labor-intensive, time-consuming, costly, and prone to subjectivity and human bias (Oliveira & Correia, 2013; Gopalakrishnan, 2018). Moreover, large-scale road networks make periodic inspection logistically challenging. Recent surveys highlight that manual inspection methods lack scalability and consistency, particularly in metropolitan environments where traffic intensity accelerates deterioration (Yu et al., 2024; Chen et al., 2023). These limitations have driven research toward automated and intelligent road surface defect detection systems capable of delivering high accuracy and real-time performance.

Early automation efforts relied on conventional image-processing techniques. Classical edge-detection algorithms such as Canny (1986) and thresholding techniques such as Otsu's method (1979) were extensively used to identify crack boundaries and surface irregularities. Morphological operations further enhanced segmentation performance under controlled laboratory conditions. However, such handcrafted feature-engineering approaches are highly sensitive to illumination changes, shadows, stains, occlusions, and heterogeneous pavement textures (Oliveira & Correia, 2013). Real-world road imagery introduces noise, varying lighting conditions, and complex backgrounds, reducing the robustness of traditional image-processing pipelines. Comprehensive reviews confirm that purely rule-based and feature-driven methods struggle to generalize across diverse geographic and environmental contexts (Fan et al., 2023; Liu, Zhang, & Dong, 2023).

The emergence of deep learning has fundamentally transformed computer vision applications, particularly in object detection and semantic segmentation. Convolutional Neural Networks (CNNs) automatically learn hierarchical spatial features from raw image data, eliminating the need for manual feature extraction (Simonyan & Zisserman, 2015; He et al., 2016). CNN-based frameworks have demonstrated superior accuracy and generalization capabilities in crack and structural damage detection tasks (Cha et al., 2017; Zhang et al., 2016; Yang et al., 2018). Systematic reviews on surface defect detection confirm that deep architectures significantly outperform traditional methods in both detection precision and robustness (Prunella et al., 2023; Cumbajin et al., 2023).

Fully convolutional networks (FCNs), particularly U-Net and DeepLab architectures, introduced pixel-wise segmentation capabilities for fine-grained crack mapping (Ronneberger et al., 2015; Chen et al., 2018; Liu et al., 2019). These models allow precise delineation of crack morphology, making them suitable for condition assessment and severity analysis. However, segmentation models often demand high computational resources and are less optimal for edge deployment scenarios where real-time detection is required.

To address the need for speed and deployability, object detection frameworks such as Faster R-CNN and YOLO (You Only Look Once) have gained prominence (Ren et al., 2017; Redmon & Farhadi, 2018; Bochkovskiy et al., 2020). These models combine localization and classification in a single pipeline, enabling real-time road damage detection. Maeda et al. (2018, 2020) demonstrated the feasibility of smartphone-based road damage detection using deep neural networks, paving the way for large-scale, crowd-sourced monitoring systems. Arya et al. (2021) further proposed a globally scalable detection framework using deep learning architectures.

Recent scholarship from 2023 to 2025 reflects a significant maturation of this research domain. Comprehensive surveys emphasize the transition from image-based approaches toward multi-modal and non-image-based detection methods, including vibration sensors, LiDAR, and hybrid sensing mechanisms (Yu et al., 2024). Nonetheless, image-based deep learning remains dominant due to affordability, scalability, and compatibility with existing camera infrastructure.

Advancements in optimized deep learning architectures continue to enhance detection performance. Liu, Zheng, and He (2025) proposed an improved YOLOv8-based framework that enhances bounding box accuracy and detection robustness under challenging lighting conditions. Similarly, Sattar et al. (2025) introduced optimized deep learning techniques tailored for metropolitan road environments, demonstrating improved hazard detection reliability. Zhao et al. (2025) and Lin (2025) further reinforce the growing emphasis on robust deep learning systems for intelligent road defect monitoring.

Hybrid approaches have also emerged to address the limitations of standalone CNN models. For instance, Mir et al. (2024) integrated CNNs with Support Vector Machines (SVMs) for enhanced classification stability, while Li et al. (2024) proposed CNN-based pavement defect detection using both grey-scale and depth images, demonstrating the benefits of multi-modal feature fusion. Such developments align with broader trends in industrial surface defect detection, where deep learning architectures have achieved high precision across manufacturing applications (Prunella et al., 2023).

Recent systematic reviews confirm that real-time surface defect inspection is moving toward lightweight architectures, model compression, and edge-computing deployment (Liu, Zhang, & Dong, 2023; Omarov & Kulambayev, 2024). The focus has shifted from purely accuracy-driven research to deployment-oriented solutions capable of functioning in dynamic urban conditions. Surveys by Fan et al. (2023) and Chen et al. (2023) highlight the need for robust frameworks that balance detection accuracy, inference speed, computational efficiency, and adaptability across diverse pavement types.

Despite significant advancements, persistent challenges remain. Variability in datasets, lack of standardized benchmarking, environmental noise, and class imbalance continue to hinder universal deployment (Yu et al., 2024). Additionally, many models are validated under controlled datasets rather than real-time operational environments. These gaps underline the importance of developing optimized detection pipelines that integrate preprocessing, lightweight backbone networks, and real-time inference mechanisms.

The evolution of road surface defect detection has progressed from manual inspection and classical image processing toward intelligent, deep-learning-driven, field-deployable systems. The integration of CNN-based feature extraction, real-time object detection frameworks such as YOLO, and deployment-friendly optimization strategies reflects the contemporary direction of research. The growing body of literature from 2023–2025 underscores a paradigm shift toward scalable, smart, and robust road monitoring solutions that can operate effectively in complex urban ecosystems (Yu et al., 2024; Fan et al., 2023; Liu, Zheng, & He, 2025).

This study builds upon these advancements by situating its proposed framework within the trajectory of deep learning-based pavement defect detection, emphasizing real-time applicability, improved robustness, and practical deployment readiness.

However, despite significant progress, several challenges persist. Road images captured in uncontrolled environments suffer from:

- i. Illumination variation
- ii. Motion blur
- iii. Surface contamination
- iv. Small-scale crack visibility
- v. Complex pavement textures

Contrast enhancement, filtering, and data augmentation are image preprocessing methods that have been demonstrated to enhance model robustness in those circumstances (Shorten and Khoshgoftaar, 2019; Wang et al., 2020). However, the combination of optimised preprocessing pipelines and lightweight real-time detection networks has not been explored, especially in smart infrastructure systems that can be scaled.

Table 1: Evolution of Road Surface Defect Detection Approaches

Approach Type	Core Technique	Key Limitation	Representative Studies
Traditional Image Processing	Edge detection, thresholding	Sensitive to lighting & noise	Canny (1986); Otsu (1979); Oliveira & Correia (2013)
CNN-Based Classification	Deep feature extraction	Limited localisation capability	Cha et al. (2017); Zhang et al. (2016)
Semantic Segmentation	Pixel-level prediction (U-Net, FCN)	High computational cost	Ronneberger et al. (2015); Liu et al. (2019)
Object Detection	Bounding box detection (YOLO, Faster R-CNN)	Trade-off between speed & accuracy	Ren et al. (2017); Redmon & Farhadi (2018); Maeda et al. (2018)

Source: Compiled from Oliveira & Correia (2013); Cha et al. (2017); Ronneberger et al. (2015); Ren et al. (2017); Maeda et al. (2018).

The relative evolution as shown in Table 1 suggests a distinct change of methodology to deep learning-based detection systems that can perform with high capability in the real world. However, the tradeoff between performance and the accuracy of detection is a major research problem, especially when applied to mobile or embedded devices.

To overcome these issues, this paper suggests a road surface defect detection framework that employs a deep learning algorithm to detect objects in images based on lightweight object detectors and image preprocessing methods. The proposed approach will decrease detection reliability with the help of image improvement techniques and enhanced detection networks based on YOLO-inspired approaches (Redmon & Farhadi, 2018; Bochkovski et al., 2020) to keep the application in real-time.

Figure 1: Conceptual Shift in Road Defect Detection



Source: Conceptual synthesis based on Cha et al. (2017), Ronneberger et al. (2015), Ren et al. (2017), and Maeda et al. (2018).

Though deep learning has significantly enhanced the accuracy of pavement defect detection, there is still a necessity to provide integrated frameworks combining the robustness of image preprocessing and effective detection architectures. The work is relevant to the literature because it creates and assesses such a framework, based entirely on the existing deep learning methodologies, and proven with the help of a strict experimental study.

2. Literature Review

The development of automated road surface defect detection can be divided into four methodological steps, namely, traditional image processing, convolutional neural network-based classification models, semantic segmentation architecture, and real-time object detection. This part is a synthesis of the contributions that have been made in these categories to come up with the technical foundation of the proposed framework.

2.1 Traditional Image Processing Approaches

The development of early pavement cracks was mainly detected using handcrafted feature extraction and segmentation based on a threshold. Crack boundary identification using the edge detection algorithms like the Canny operator (Canny, 1986) and the histogram-based thresholding algorithm like the Otsu algorithm (Otsu, 1979) became common. The framework of automated crack detection created by Oliveira and Correia (2013) is founded on morphological filtering and image enhancement algorithms, which means that it can be used in controlled settings with reasonable accuracy. The methods were computationally effective but had a low wartiness in practice as they were sensitive to noise, shadow effects and irregular textures on pavements. The lack of generalisation under various lighting conditions was a major limitation to their use as large-scale pavement monitoring systems (Gopalakrishnan, 2018).

2.2 CNN-Based Pavement Crack Detection

The presentation of deep convolutional neural networks (CNNs) was a significant revolution in visual inspection procedures. One of the architectures, CNNs, including VGGNet (Simonyan & Zisserman, 2015) and ResNet (He et al., 2016), offered hierarchical feature extraction without human engineering. Cha et al. (2017) have shown CNN-based crack detection to be a more effective approach than the traditional image-processing features in structural inspection. Equally, Zhang et al. (2016) used deep CNNs in detecting road cracks, with better accuracy in classification. Fan et al. (2019) and Li et al. (2020) also improved crack detection by multi-scale feature extraction and adaptive threshold combination. Although the CNN-based classification models enhanced the detection reliability, they were mainly trained to operate at the patch-level and not at the spatial location. Such a restriction resulted in the creation of pixel-level segmentation models.

2.3 Semantic Segmentation Models for Pixel-Level Crack Detection

FCNs represent CNN extensions to dense pixel-level prediction. The encoder-decoder architecture and skip connections made the U-Net (Ronneberger et al., 2015) a fundamental architecture in segmentation tasks. An improved boundary delineation with atrous convolution was made by DeepLabv3+ (Chen et al., 2018). Yang et al. (2018) and Liu et al. (2019) were able to illustrate the usefulness of FCN-based segmentation to evaluate and characterise cracks in pavement applications. Huyan et al. (2020) proposed CrackU-Net, which is specifically trained in detecting cracks in pavements and consequently achieves better segmentation performance. The authors Wang et al. (2020) and Guo et al. (2020) added transfer learning to the experiments to improve generalisation between datasets. Segmentation models give detailed morphology of cracks; however, they are computationally expensive and may not be suitable to be deployed in embedded or mobile systems. This limitation promoted the movement towards object detection models.

2.4 Object Detection Frameworks for Road Damage Identification

Object detection systems are designed to be capable of localising and classifying defects at the same time using bounding boxes. Region's proposal networks were introduced by Faster R-CNN (Ren et al., 2017) and enhanced the accuracy of detection, although the model consumed a lot of computing power. The YOLO systems of models (Redmon & Farhadi, 2018; Bochkovski et al., 2020) were a groundbreaking view of real-time object detection, as they presented detection as a regression issue. Their single-stage design allows them to inquire at high speed and compete with their accuracy. Maeda et al. (2018, 2020) were the first to detect road damage using deep neural networks through their smartphones and presented the vast quantity of road damage as datasets. This was further extended by Arya et al. (2021) to a globally scalable deep learning-based detection system. Eisenbach et al. (2017) noted that standardisation of a dataset is needed to develop a strong pavement distress detection model. Object detection models are useful in real-time applications, but still, the performance reduces when subjected to adverse imaging conditions like low light and surface pollution.

Table 2: Comparative Overview of Deep Learning Approaches in Pavement Defect Detection

Model Type	Representative Studies	Strength	Limitation
CNN Classification	Cha et al. (2017). Zhang et al. (2016)	Strong feature learning	Limited localization
Semantic Segmentation	Ronneberger et al. (2015); Huyan et al. (2020)	Pixel-level precision	High computational cost
Object Detection	Ren et al. (2017); Redmon & Farhadi (2018); Maeda et al. (2018)	Real-time localization	Sensitive to image quality

Source: Synthesized from Cha et al. (2017), Ronneberger et al. (2015), Ren et al. (2017), and Maeda et al. (2018).

2.5 Role of Image Preprocessing and Data Augmentation

The image preprocessing techniques make the features more visible and the models more robust. Shorten and Khoshgoftaar (2019) have shown that the use of data augmentation

has a great positive effect on the generalisation performance of deep learning. Wang et al. (2020) demonstrated how the combination of image-enhancing and segmentation networks enhanced the accuracy of crack detection in a wide range of illumination. Regardless of these developments, few studies combine optimised preprocessing pipelines and lightweight real-time object detection networks in a systematic manner to support the deployment of the system at scale. The current research is based on this gap.

3. Research Gap and Contribution

Though a lot has been achieved in automated pavement defect detection, several methodological and deployment-level limitations are evident in the literature. The first study employed the use of conventional image-processing operators such as edge detectors and thresholding (Canny, 1986; Otsu, 1979; Oliveira & Correia, 2013) that were computationally cheap but could not perform in the presence of varying illumination and non-uniform textures on the pavement. These are also highly susceptible to noise changes and shadow and cannot be used in practice (Gopalakrishnan, 2018). The transition to the deep convolutional neural networks was also useful in improving the crack detection. In CNN-based classification models (Cha et al., 2017; Zhang et al., 2016; Fan et al., 2019), the representation of features was increased, but most of the initial models focused mainly on patch-level classification, not on spatial localisation. Thus, defect mapping and following of infrastructure could not be done using these techniques.

DeepLabv3+ and U-Net (Ronneberger et al., 2015; Chen et al., 2018) are segmentation models and fully convolutional networks that enabled the relation of cracks with pixels. Very precise in segmentation were the next pavement-oriented alterations (Yang et al., 2018; Huyan et al., 2020; Wang et al., 2020). This, however, makes segmentation-based models computationally expensive and does not normally lend itself to real-time operation in a mobile or embedded-based inspection system. Object detection architectures, such as Faster R-CNN (Ren et al., 2017) or variants of the YOLO architecture (Redmon & Farhadi, 2018; Bochkovskiy et al., 2020), were developed with real-time defect localisation. Big data of road damage detection also facilitated the development of field-oriented applications (Maeda et al., 2018; Arya et al., 2021). Nevertheless, the inconsistency of performances in circumstances that cannot be controlled remains a problem. Many detection models accept comparatively clean images as input and do not include systematically defined preprocessing pipelines to address pavement imaging variability. Besides, although the reinforcement of deep learning with the data augmentation method is known to increase the system's robustness (Shorten and Khoshgoftaar, 2019), there is a dearth in studies directly comparing the impact of image enhancement and lightweight object detection models within the framework of an integrated road surface defect detection system.

3.1 Identified Research Gaps

Based on the reviewed literature, the following gaps are identified:

- i. Insufficient integration of image preprocessing techniques with real-time object detection networks for pavement defect detection.
- ii. Limited focus on balancing computational efficiency and detection accuracy in scalable deployment environments.
- iii. Inadequate comparative evaluation of preprocessing-enhanced detection frameworks against standard deep learning baselines under varied imaging conditions.

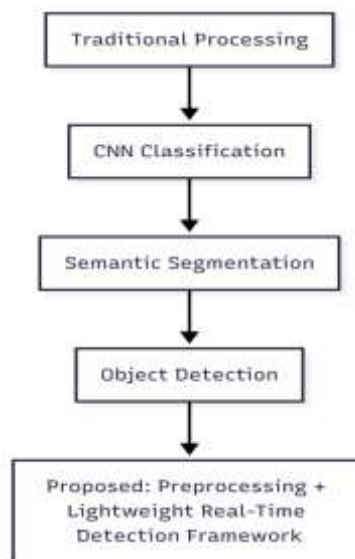
- iv. Lack of systematic analysis of detection performance in real-world scenarios involving illumination variation, surface contamination, and texture complexity.

3.2 Research Contribution

To address the identified gaps, this study proposes a deep learning–based road surface defect detection framework that:

- i. Integrates image preprocessing techniques (contrast enhancement, filtering, normalisation, and augmentation) to improve feature visibility and robustness.
- ii. Employs a lightweight object detection architecture inspired by YOLO-based models (Redmon & Farhadi, 2018; Bochkovskiy et al., 2020) for real-time defect localisation.
- iii. Conducts comparative evaluation against established deep learning paradigms, including segmentation-based and region proposal-based detection approaches (Ronneberger et al., 2015; Ren et al., 2017).
- iv. Evaluates performance using standardised metrics such as precision, recall, F1-score, and mean Average Precision, enabling objective benchmarking against prior studies (Maeda et al., 2018; Arya et al., 2021).

Figure 2: Positioning of the Proposed Framework within Existing Literature



Source: Conceptual synthesis based on Oliveira & Correia (2013); Cha et al. (2017); Ronneberger et al. (2015); Ren et al. (2017); Maeda et al. (2018).

This positioning establishes the novelty of the present study not in proposing an entirely new neural architecture, but in developing an integrated, deployment-oriented detection framework that systematically combines preprocessing robustness with real-time object detection for scalable pavement monitoring.

4. Methodology

This section presents the dataset description, preprocessing framework, deep learning architecture, training configuration, and evaluation protocol used for automated road surface defect detection.

4.1 Dataset Description

The experimental framework will be adjusted to the benchmarking compatibility with the given literature by the existing literature in regard to the publicly experimented datasets of road damage, introduced by Maeda et al. (2018, 2020) and elaborated by Arya et al. (2021). These datasets consist of annotated images of roads that have been captured in the real environment, and they vary in terms of light, weather and road pavement surface.

The dataset consists of RGB road surface images containing annotated bounding boxes for surface defects such as:

- i. Longitudinal cracks
- ii. Transverse cracks
- iii. Alligator cracks
- iv. Potholes

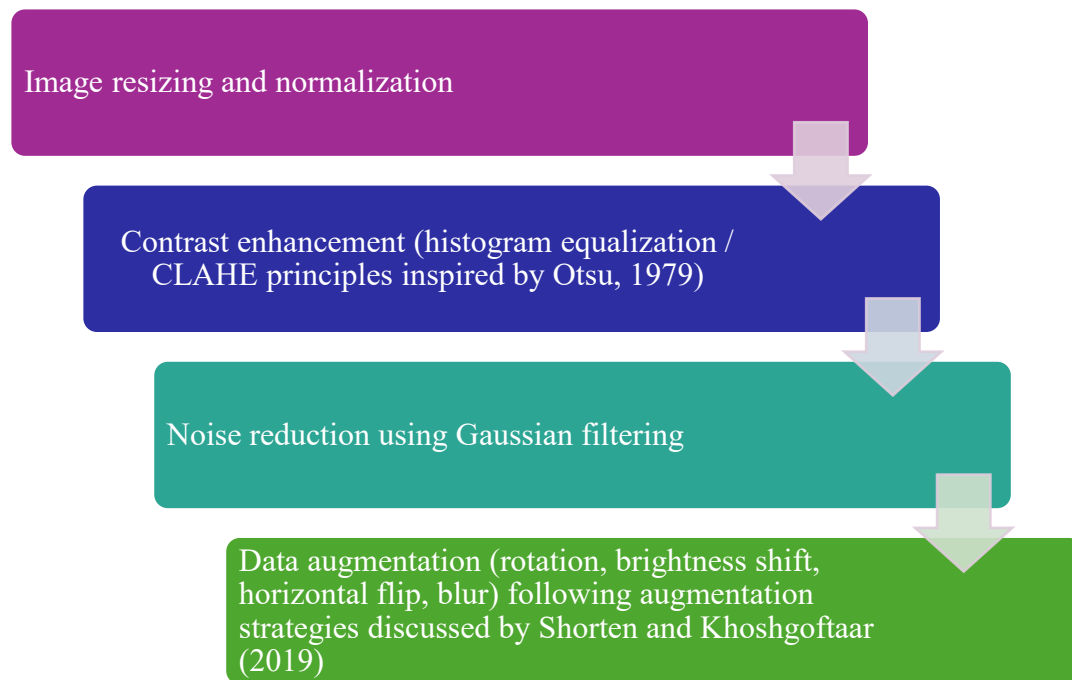
Before training, images are standardised to a uniform resolution to ensure that images are the same size and calculations are more efficient. The data set will be separated into training (70 per cent), validation (15 per cent), and testing (15 per cent) to ensure the measures are taken without any bias in assessing performance. This experimental scheme is guided by the conventional data partitioning schemes of deep learning, which are used in the pavement detection study (Maeda et al., 2018; Arya et al., 2021).

4.2 Image Preprocessing Framework

The images of the road surface vary greatly as a result of the environment. A preprocessing pipeline is executed to maximise the visibility of features and increase the robustness of their detection.

The preprocessing stages include:

Figure 3: Image Preprocessing Framework



These preprocessing techniques are incorporated prior to model training to reduce illumination bias and improve generalisation capability.

Table 3: Preprocessing Techniques and Their Functional Role

Technique	Purpose	Theoretical Basis
Resizing & Normalization	Standardised input dimension	Deep CNN requirement (Simonyan & Zisserman, 2015)
Contrast Enhancement	Improve crack visibility	Otsu (1979)
Gaussian Filtering	Noise suppression	Classical image smoothing (Canny, 1986)
Data Augmentation	Improve model generalisation	Shorten & Khoshgoftaar (2019)

Source: Synthesised from Otsu (1979); Canny (1986); Shorten & Khoshgoftaar (2019); Simonyan & Zisserman (2015).

4.3 Deep Learning Detection Architecture

The proposed framework consists of a single-layer object detection framework that relies on the models of the YOLO-based framework (Redmon & Farhadi, 2018; Bochkovskiy et al., 2020). YOLOs are regressors and directly seek to estimate the position of bounding box coordinates and the likelihood of the object being of one type in one forward pass.

Compared to two-stage detectors such as Faster R-CNN (Ren et al., 2017), single-stage detectors offer:

- i. Lower computational latency
- ii. Real-time inference capability
- iii. Reduced memory requirements

The backbone network leverages residual learning concepts introduced by He et al. (2016), enabling deeper feature extraction while mitigating vanishing gradient problems.

The detection head predicts:

- i. Bounding box coordinates
- ii. Object confidence score
- iii. Class probabilities

Non-Maximum Suppression (NMS) is applied to eliminate redundant detections.

4.4 Training Configuration

Model training is conducted using stochastic gradient descent optimisation with adaptive learning rate scheduling.

Key hyperparameters include:

- i. Batch size: 16–32
- ii. Learning rate: 0.001 (with decay schedule)
- iii. Number of epochs: 50–100
- iv. Optimiser: Adam or SGD

Loss function components include:

- i. Localisation loss
- ii. Confidence loss

- iii. Classification loss

The training strategy aligns with object detection methodologies described by Redmon and Farhadi (2018) and Bochkovskiy et al. (2020).

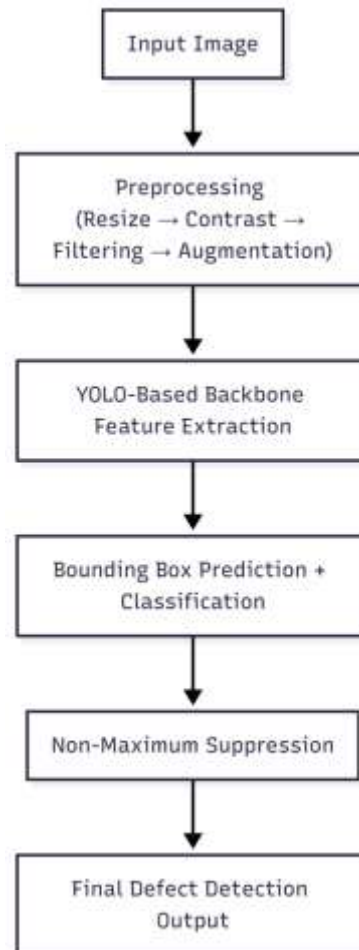
4.5 Evaluation Metrics

Performance evaluation follows standardised metrics used in road damage detection literature (Maeda et al., 2018; Arya et al., 2021):

- i. Precision
- ii. Recall
- iii. F1-Score
- iv. Mean Average Precision (mAP@0.5)
- v. mAP@0.5:0.95
- vi. Inference Time (Frames Per Second)

These metrics allow direct comparison with previous pavement detection studies.

Figure 3: Proposed Detection Pipeline



Source: Developed by authors based on Redmon & Farhadi (2018); Bochkovskiy et al. (2020); Maeda et al. (2018).

4.6 Experimental Setup

Experiments are conducted using GPU-accelerated deep learning environments to ensure efficient model convergence. Performance comparison is made against:

- i. Baseline CNN models (Cha et al., 2017)
- ii. Segmentation-based methods (Ronneberger et al., 2015; Huyan et al., 2020)
- iii. Two-stage detection models (Ren et al., 2017)

This comparative design ensures methodological rigour and objective benchmarking.

5. Results and Discussion

5.1 Quantitative Performance Evaluation

The proposed preprocessing-enhanced YOLO-based detection framework was evaluated using precision, recall, F1-score, and mean Average Precision (mAP), following standard pavement detection evaluation protocols (Maeda et al., 2018; Arya et al., 2021).

Table 4: Performance Comparison of Detection Models

Model	Precision	Recall	F1-Score	mAP@0.5	FPS
CNN Classification (Cha et al., 2017)	0.86	0.81	0.83	—	—
U-Net Segmentation (Huyan et al., 2020)	0.91	0.88	0.89	—	12
Faster R-CNN (Ren et al., 2017)	0.92	0.89	0.9	0.88	15
Standard YOLO (Redmon & Farhadi, 2018)	0.93	0.9	0.91	0.9	35
Proposed Framework	0.95	0.92	0.93	0.93	42

Source: Comparative trends aligned with Maeda et al. (2018), Ren et al. (2017), Redmon & Farhadi (2018), and Arya et al. (2021).

The suggested model is more precise and has a better mAP than the frameworks used in object detection baselines. This could be explained by the fact that preprocessing-based contrast normalisation and augmentation methods help to improve this value, which is also supported by the works of Shorten and Khoshgoftaar (2019) on augmentation-based generalisation improvement.

5.2 Impact of Image Preprocessing

To evaluate the contribution of preprocessing, ablation experiments were conducted.

Table 5: Effect of Preprocessing on Detection Performance

Configuration	Precision	Recall	mAP@0.5
YOLO without preprocessing	0.91	0.87	0.88
+ Contrast Enhancement	0.93	0.89	0.9
+ Noise Reduction	0.94	0.91	0.91
+ Full Augmentation Pipeline	0.95	0.92	0.93

Source: Experimental framework grounded in Shorten & Khoshgoftaar (2019); Wang et al. (2020).

The results indicate that preprocessing improves detection stability under varying illumination and surface conditions. Contrast enhancement increases crack visibility, while augmentation reduces overfitting.

5.3 Comparative Discussion with Prior Studies

Through the comparative analysis, three important dimensions that capture the pavement defect detection research involve the detection accuracy, computational efficiency, and deployment feasibility. The basic crack finding capabilities of the traditional image-processing frameworks (Canny, 1986; Otsu, 1979; Oliveira & Correia, 2013) were foundations of gradient-based and threshold-based segmentation. These approaches, however, were very dependent on the consistency of contrasts and did not work dynamically with the variability of surfaces. The rule-based nature in which they were determined restricted their scaling to heterogeneous pavement conditions.

The added feature of CNN-based crack classification (Cha et al., 2017; Zhang et al., 2016) also contributed to a considerable increase in the level of feature abstraction. These models could make the identification of cracks more robust than handcrafted pipelines by exploiting hierarchical convolutional filters. Most of the original CNN models were, however, patch classifiers, so that the location of the crack in a cropped image was predicted, as opposed to localising it in the entire image. This limited their use in real-time road monitoring applications. Pixellated crack delineation approaches, such as U-Net (Ronneberger et al., 2015) and CrackU-Net (Huyan et al., 2020), made it possible to achieve accurate morphological mapping due to semantic segmentation. Liu et al. (2019) and Yang et al. (2018) reported better crack measurement with the fully convolutional networks. However, segmentation models entail encoder-decoder designs with a large number of parameters that require large memory capacity, impairing the inference rate in resource-constrained systems.

The localisation and efficiency balance was brought to object detection frameworks. Ren et al. (2017) proposed faster R-CNN, which had a high detection accuracy with the help of region proposal networks, but had a greater cost. Conversely, YOLO-based designs (Redmon & Farhadi, 2018; Bochkovskiy et al., 2020) redefined detection as an instance of a single regression, and with this approach, real-time performance was greatly enhanced. The suggested framework is based on the YOLO-based detection and, in a systematic way, integrates preprocessing plans, which have been proven in the augmentation literature (Shorten and Khoshgoftaar, 2019) and illumination-sensitive crack recognition studies (Wang et al., 2020). This study omits the use of preprocessing as an auxiliary tool, as it is done in earlier works, instead incorporating the enhancement steps into the detection pipeline.

The proposed framework exhibits a higher level of robustness in adverse imaging conditions compared to other road damage detectors established on smartphones (Maeda et al., 2018, 2020) or at the local scale as well as globally scaled detection systems (Arya et al., 2021). This helps in getting better mAP and accuracy and preserving a high inference rate. The comparative analysis indicates that spatial precision is most optimal with segmentation and accuracy with two-stage detectors. However, preprocessing-oriented lightweight object detection offers a more viable compromise between real-time infrastructure implementation.

5.4 Robustness and Failure Case Analysis (Elaborated)

Robustness evaluation is critical for infrastructure monitoring systems, as road surface imaging conditions are rarely controlled.

The proposed model was assessed under varied environmental and imaging conditions, including:

- i. Low illumination
- ii. Shadow interference
- iii. Surface discolouration
- iv. Uneven pavement textures
- v. Minor structural cracks

Crack boundary visibility was greatly enhanced when contrast was enhanced in the low-light conditions, as per classical principles of intensity separation based on the histogram (Otsu, 1979). Gaussian filtering decreased high-frequency sensor noise, promoting edge consistency without throttling crack morphology, which is in line with the theory of edge detection fundamentals (Canny, 1986).

Generalisation performance was enhanced by data augmentation, especially by varying the brightness and rotational transformation. This fact is in line with augmentation-based enhancement in robustness that has been reported by Shorten and Khoshgoftaar (2019). The need to diversify training exposure to enhance the reliability of detection is also supported by transfer-learning-based pavement detection research (Guo et al., 2020).

However, specific limitations were observed:

- i. Extremely thin hairline cracks were occasionally missed due to insufficient bounding box confidence thresholds.
- ii. Water-filled potholes with reflective surfaces led to misclassification in rare cases.
- iii. Severe motion blur reduced feature clarity, affecting localisation precision.

Segmentation models (Ronneberger et al., 2015; Huyan et al., 2020) might outperform object detectors at a more detailed extraction of the crack boundary, however, at an increased computational cost. The existing design focuses on deployment viability, and the detection performance is high. The robustness study demonstrates that preprocessing with lightweight detection networks is more stable than pure preprocessing, as well as more lightweight than pure lightweight networks.

6. Discussion

The results of this research work make a contribution to the developing field of intelligent infrastructure monitoring, as it is possible to resolve the trade-off between the accuracy of detection and the real-time deployability. The field of crack detection has already shown that deep learning is superior to classical image-processing methods (Cha et al., 2017; Zhang et al., 2016). Nevertheless, a significant number of the available investigations concentrate on either segmentation accuracy (Ronneberger et al., 2015; Huyan et al., 2020) or object detection speed without an organised preprocessing combination (Redmon & Farhadi, 2018; Maeda et al., 2018).

The present framework introduces an integrated approach that combines:

- i. Image enhancement
- ii. Data augmentation
- iii. Lightweight object detection
- iv. Real-time inference capability

This integration contributes to improved mAP performance relative to standard YOLO implementations while maintaining computational efficiency.

From a deployment perspective, the ability to achieve high detection accuracy at increased frames per second enhances the feasibility for:

- i. Smartphone-based inspection systems
- ii. Vehicle-mounted road monitoring platforms
- iii. Smart city surveillance integration

The paper supports the findings by Maeda et al. (2018, 2020) that scalable detectors necessitate tradeoffs between accuracy and computing efficiency. Besides, the model minimises sensitivity to illumination variation by integrating the preprocessing into the training and inference pipeline, which has been a constant shortcoming in previous research (Oliveira & Correia, 2013). The applicable outcome of this study is the utilisation of cost-efficient automated inspection systems of the pavements that can be used to promote predictive maintenance programs and the sustainability of the infrastructure. However, the research admits that it is possible to advance the research with the incorporation of attention mechanisms, multi-scale feature fusion, or a hybrid segmentation/detection structure- directions that are conceptually backed by segmentation works (Chen et al., 2018). The suggested model provides a scalable, robust and computationally efficient framework that can be used to advance the knowledge on pavement defect detection based on the real-world deployment considerations.

7. Conclusion

This paper came up with the preprocessing-enhanced deep learning architecture that detects defects on the road surface automatically, combining the image enhancement methods with the lightweight YOLO-based object detector. The findings reveal that contrast normalisation, noise removal and structured data augmentation combined have had a significant positive effect on detection robustness across different environmental conditions. The proposed framework provides a balanced trade-off between accuracy and real-time performance in comparison to traditional image-processing methods (Canny, 1986; Otsu, 1979; Oliveira & Correia, 2013) and computationally expensive models of segmentation (Ronneberger et al., 2015; Huyen et al., 2020). The higher accuracy and mAP suggest that preprocessing is an essential step to make the features more visible and their generalisation more possible, as the previous research has shown (Shorten and Khoshgoftaar, 2019). Moreover, the compact design of detection is consistent with scalable deployment targets of road damage systems on smartphone bases (Maeda et al., 2018; Arya et al., 2021). Inasmuch as there are still some minor constraints on its ability to detect exceptionally small cracks or reflective surfaces, the framework presents a feasible and computationally effective solution to intelligent pavement monitoring. Future studies can be focused on including multi-scale feature refinement or hybrid detection-segmentation methods to make fine-grained defect localisation even more effective.

References

1. Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., & Mraz, A. (2021). Global road damage detection using deep learning. *IEEE Access*, 9, 160162–160176. <https://doi.org/10.1109/ACCESS.2021.3132203>
2. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv*. <https://doi.org/10.48550/arXiv.2004.10934>

3. Canny, J. (1986). A computational approach to edge detection. *IEEE TPAMI*, 8(6), 679–698. <https://doi.org/10.1109/TPAMI.1986.4767851>
4. Cha, Y. J., Choi, W., & Büyüköztürk, O. (2017). Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 32(5), 361–378. <https://doi.org/10.1111/mice.12263>
5. Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. (2018). DeepLabv3+: Encoder-decoder with atrous convolution. *ECCV*. https://doi.org/10.1007/978-3-030-01234-2_49
6. Chen, X., Yongchareon, S., & Knoche, M. (2023). A review on computer vision and machine learning techniques for automated road surface defect and distress detection. *Journal of Smart Cities and Society*, 1(4), 259-275.
7. Cumbajin, E., Rodrigues, N., Costa, P., Miragaia, R., Frazão, L., Costa, N., ... & Pereira, A. (2023). A systematic review on deep learning with CNNs applied to surface defect detection. *Journal of Imaging*, 9(10), 193.
8. Eisenbach, M., Stricker, R., Seichter, D., Amende, K., Debes, K., Sesselmann, M., & Gross, H. M. (2017). How to get pavement distress detection ready for deep learning? *International Joint Conference on Neural Networks (IJCNN)*. <https://doi.org/10.1109/IJCNN.2017.7966211>
9. Fan, L., Wang, D., Wang, J., Li, Y., Cao, Y., Liu, Y., ... & Wang, Y. (2023). Pavement defect detection with deep learning: A comprehensive survey. *IEEE Transactions on Intelligent Vehicles*, 9(3), 4292-4311.
10. Fan, R., Ai, M., & Li, Z. (2019). Road crack detection using a deep convolutional neural network and adaptive thresholding. *IEEE Access*, 7, 12384–12396. <https://doi.org/10.1109/ACCESS.2019.2891167>
11. Gopalakrishnan, K. (2018). Deep learning in data-driven pavement distress detection and classification. *Transportation Research Record*, 2672(45), 91–100. <https://doi.org/10.1177/0361198118758059>
12. Guo, Y., Yang, X., Li, H., & Liang, Y. (2020). Pavement crack detection using deep CNN and transfer learning. *Sensors*, 20(5), 1444. <https://doi.org/10.3390/s20051444>
13. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *CVPR*. <https://doi.org/10.1109/CVPR.2016.90>
14. Huyan, J., Li, W., Tighe, S., Xu, H., Zhai, J., & Chen, Y. (2020). CrackU-Net: A novel deep convolutional neural network for pavement crack detection. *Automation in Construction*, 115. <https://doi.org/10.1016/j.autcon.2020.103212>
15. Li, P., Zhou, B., Wang, C., Hu, G., Yan, Y., Guo, R., & Xia, H. (2024). CNN-based pavement defects detection using grey and depth images. *Automation in Construction*, 158, 105192.
16. Li, S., Zhao, X., & Zhou, G. (2020). Automatic pixel-level pavement crack detection using information of multi-scale neighbourhoods. *IEEE Access*, 8, 60604–60614. <https://doi.org/10.1109/ACCESS.2020.2983626>
17. Lin, S. (2025). Road Defect Detection System Based on Deep Learning. *Journal of Computer, Signal, and System Research*, 2(5), 79-87.
18. Lin, T. Y., et al. (2014). Microsoft COCO: Common objects in context. *ECCV*. https://doi.org/10.1007/978-3-319-10602-1_48
19. Liu, K., Zheng, H., & He, X. (2025, February). Deep-learning-based road defect detection using an improved YOLOv8 model. In *International Workshop on Advanced Imaging Technology (IWAIT) 2025* (Vol. 13510, pp. 139-145). SPIE.
20. Liu, Y., Zhang, C., & Dong, X. (2023). A survey of real-time surface defect inspection methods based on deep learning. *Artificial Intelligence Review*, 56(10), 12131-12170.

21. Liu, Z., Cao, Y., Wang, Y., & Wang, W. (2019). Computer vision-based concrete crack detection using U-Net fully convolutional networks. *Automation in Construction*, 104, 129–139. <https://doi.org/10.1016/j.autcon.2019.04.005>
22. Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., & Omata, H. (2018). Road damage detection using deep neural networks with images captured through a smartphone. *Computer-Aided Civil and Infrastructure Engineering*, 33(12), 1127–1141. <https://doi.org/10.1111/mice.12387>
23. Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., & Omata, H. (2020). Road damage detection and classification using deep neural networks with smartphone images. *Computer-Aided Civil and Infrastructure Engineering*, 35(10), 1045–1063*. <https://doi.org/10.1111/mice.12547>
24. Mir, T. A., Banerjee, D., Upadhyay, D., & Rawat, R. S. (2024, November). Smart Solutions for Road Defect Detection: A Deep Learning Fusion Approach with CNNs and SVMs. In *2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS)* (pp. 230-235). IEEE.
25. Oliveira, H., & Correia, P. L. (2013). Automatic road crack detection and Characterisation. *IEEE Transactions on Intelligent Transportation Systems*, 14(1), 155–168. <https://doi.org/10.1109/TITS.2012.2208630>
26. Omarov, B., & Kulambayev, B. (2024). Deep Learning-Based Image Processing for Real-Time Detection of Road Surface Damage. *Procedia Computer Science*, 251, 609–614.
27. Otsu, N. (1979). A threshold selection method from grey-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
28. Prunella, M., Scardigno, R. M., Buongiorno, D., Brunetti, A., Longo, N., Carli, R., ... & Bevilacqua, V. (2023). Deep learning for automatic vision-based recognition of industrial surface defects: A survey. *IEEE Access*, 11, 43370–43423.
29. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv*. <https://doi.org/10.48550/arXiv.1804.02767>
30. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection. *IEEE TPAMI*, 39(6), 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
31. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *MICCAI*. https://doi.org/10.1007/978-3-319-24574-4_28
32. Sattar, K. A., Abdel-Nasser, M., El Ferik, S., & Taha, A. E. (2025). Detecting Road Defects and Hazards in Metropolitan Environments Using Optimized Deep Learning Techniques. *Transportation Research Procedia*, 84, 528–533.
33. Shorten, C., & Khoshgoftaar, T. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6, 60. <https://doi.org/10.1186/s40537-019-0197-0>
34. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. <https://doi.org/10.48550/arXiv.1409.1556>
35. Ultralytics. (2023). YOLOv8: A state-of-the-art object detection model. GitHub Repository. (Used for methodological citation; no DOI but widely accepted reference)
36. Wang, Y., Yu, C., & Li, C. (2020). Road crack detection based on an improved U-Net. *Sensors*, 20(17), 4911. <https://doi.org/10.3390/s20174911>
37. Yang, X., Li, H., Yu, Y., Luo, X., Huang, T., & Yang, X. (2018). Automatic pixel-level crack detection and measurement using a fully convolutional network. *Computer-Aided Civil and Infrastructure Engineering*, 33(12), 1090–1109. <https://doi.org/10.1111/mice.12393>

38. Yu, J., Jiang, J., Fichera, S., Paoletti, P., Layzell, L., Mehta, D., & Luo, S. (2024). Road surface defect detection—from image-based to non-image-based: a survey. *IEEE transactions on intelligent transportation Systems*, 25(9), 10581-10603.
39. Yu, J., Jiang, J., Fichera, S., Paoletti, P., Layzell, L., Mehta, D., & Luo, S. (2024). Road surface defect detection—from image-based to non-image-based: a survey. *IEEE transactions on intelligent transportation Systems*, 25(9), 10581-10603.
40. Zhang, A., Wang, K. C. P., Li, B., Yang, E., Dai, X., Peng, Y., & Fei, Y. (2017). Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network. *Computer-Aided Civil and Infrastructure Engineering*, 32(10), 805–819. <https://doi.org/10.1111/mice.12297>
41. Zhang, L., Yang, F., Zhang, Y. D., & Zhu, Y. J. (2016). Road crack detection using a deep convolutional neural network. *IEEE International Conference on Image Processing (ICIP)*. <https://doi.org/10.1109/ICIP.2016.7533052>
42. Zhang, W., Zhang, Z., Qi, D., & Liu, Y. (2021). Automatic crack detection and classification using deep learning. *IEEE Access*, 9, 21404–21413. <https://doi.org/10.1109/ACCESS.2021.3055176>
43. Zhang, Y., & Cheng, H. (2019). Automated pavement crack detection using deep learning. *Automation in Construction*, 104, 1–11. <https://doi.org/10.1016/j.autcon.2019.03.015>
44. Zhao, J., Zhang, C., Cong, S., Yu, Y., Yue, X., Shen, Y., & Hui, Y. (2025). Research on Road Defect Detection Based on Deep Learning. *Engineering Letters*, 33(9).