PhD Thesis

by Ruchi Tripathi

Submission date: 06-May-2026 03:55PM (UTC+0530)
Submission ID: 2953755573
File name: PhDthesis_final_ruchi.pdf (45.87M)
Word count: 31137
Character count: 179712

ROAD DAMAGE DETECTION FOR ADVANCED DRIVER ASSISTANCE SYSTEM

A Thesis Submitted
In Partial Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

by

Ruchi
(2K20/PHDEC/514)

Under the Supervision of
Prof. S. Indu
Professor, Department of ECE, DTU, Delhi
&
Dr. Rohit Kumar
Assistant Professor, Department of ECE, DTU, Delhi



Department of Electronics and Communication Engineering

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

March, 2026



DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I **Ruchi** hereby certify that the work which is being presented in the thesis entitled **"Road Damage Detection for Advanced Driver Assistance system "** in partial fulfillment of the requirements for the award of the Degree of Doctor in Philosophy, submitted in the Department of Electronics and Communication Engineering, Delhi Technological University is an authentic record of my own work carried out during the period from **January 2021 to March 2026** under the supervision of **Prof.S.Indu**, Professor in the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India and **Dr. Rohit Kumar**, Assistant Professor in the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other institute.

Candidate's Signature

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of our knowledge.

Signature of Supervisor **Signature of External Examiner**

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042

CERTIFICATE BY THE SUPERVISOR(s)

This is to certify that **Ruchi** (Enrollment No.: 2K20/PHDEC/514) has conducted the research work presented in this thesis titled **"Road Damage Detection for Advanced Driver Assistance system"** for the award of **Doctor of Philosophy** in the Department of Electronics and Communication Engineering, Delhi Technological University, in fulfillment of requirements of the award of the degree of Doctor of Philosophy is an authentic work carried under the supervision of **Prof. S.Indu** and **Dr. Rohit Kumar**. The matter presented in this synopsis has not been submitted elsewhere in part or fully to any other University or Institute for the award of any degree to the best of our knowledge and belief.

Prof. S. Indu
Department of ECE
Delhi Technological University
Delhi–110042, India

Dr. Rohit Kumar
Department of ECE
Delhi Technological University
Delhi–110042, India

Place: Delhi

Date:

ACKNOWLEDGEMENT

I am deeply grateful to the Almighty for giving me the strength, patience, and guidance to complete this research work and reach this important stage of my academic journey.

I extend my heartfelt gratitude to my supervisors, **Prof. S. Indu** , Professor Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, and **Dr. Rohit Kumar** Assistant Professor, Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, for their invaluable guidance, enthusiastic encouragement, and persistent support. Their valuable suggestions, timely feedback, and constant motivation played a major role in shaping this research work. I feel fortunate to have worked under their supervision.

I extend my sincere appreciation to the members of my Doctoral Research Committee for their valuable suggestions and constructive feedback, which helped in improving the quality of this work. I sincerely thank the external experts involved during my pre-PhD evaluation for their time, insightful comments, and constructive suggestions, which were helpful in strengthening the foundation of my research.

I am deeply thankful to **Prof. Neeta Pandey**, Head of the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, and other faculty members for their continuous support and cooperation during the course of this dissertation. I am thankful to all staff members of the Department of Electronics and Communication Engineering for their kind assistance and support during the entire duration of my research. I am grateful to the management of Delhi Technological University, Delhi, for providing the required resources, facilities, and supportive research environment for my research work.

My sincere appreciation also goes to my colleagues and friends in the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, for their constant support, encouragement, and timely assistance, especially Dr. Neha Sharma, Priyanka and Himani.

Above all, I owe my deepest appreciation to my husband, my beloved parents, family, and close friends for their unwavering encouragement, patience, and faith in me throughout my academic journey. Their love and belief have been my greatest

source of strength, inspiration, and motivation. My parents, Mr. Prem Narayan Tripathi and Mrs. Shyama Tripathi, along with my brothers, have shown undeterred faith in me since childhood and have been the strongest pillars of support, standing by me until the completion of this work. I am especially grateful to my husband, Mr. Saurabh Shukla, and my father-in-law, Mr. Umesh Chandra Shukla, for his unconditional and continuous motivation, support, and guidance, right from the decision to pursue doctoral research to its successful completion.

A special mention goes to my two daughters for their love, understanding, and support, and for patiently bearing with my absence during important phases of this journey. Last but not least, I sincerely thank everyone who, directly or indirectly, contributed to my academic and personal journey.

Ruchi

ABSTRACT

Road infrastructure plays an important role in ensuring road safety and efficiency, particularly with the growing integration of Advanced Driver Assistance Systems (ADAS) and intelligent vehicles. In order to reduce accidents and preserve driving comfort, accurate road surface condition detection and monitoring are important. The rapid deterioration of road surfaces due to cracks, potholes, and other structural damage poses significant risks to driving safety and incurs high maintenance costs. Traditional inspection techniques are labor-intensive and time-consuming, and also rely primarily on manual inspection. The aforementioned constraints underscore the growing demand for automated, accurate, and reliable road damage detection systems that operate efficiently in real-world conditions. The performance of vision-based road analysis has improved recently due to advances in deep learning. However, many current methods still struggle with generalization and robustness across diverse environmental conditions, and they lack reliable metrics such as uncertainty estimation. Moreover, the majority of existing systems only concentrate on detection accuracy, lacking risk-based decision-making processes that are necessary for realistic ADAS implementation.

Advanced Driver Assistance Systems (ADAS) rely heavily on road surface condition knowledge since different types of road damage have a direct impact on driving safety, passenger comfort, and vehicle stability. However, rather than offering a structured and reliable solution that can be applied in a range of difficult situations, current approaches for studying road damage usually focus on specific tasks like pothole recognition, crack segmentation, or damage categorization.

In this thesis, a complete framework for intelligent road damage segmentation and detection with improved decision support and reliability is presented. A modified

YOLOv4-based detection framework is implemented to detect potholes in various road condition images with accuracy and in real time. The model offers efficient and reliable pothole localization with high inference speed while maintaining strong detection accuracy. At the pixel level, an enhanced residual UNet-based architecture is proposed for accurate pothole segmentation, which can assess the fine-grained details of pothole regions that are crucial for detailed surface assessment and maintenance analysis. Additionally, the regions that affect detection judgments are highlighted by the Grad-CAM visualization, which is utilized to evaluate the model's predictions.

A further contribution of this work is the development of an uncertainty-aware deep learning model that tackles the problem of road crack segmentation. Road cracks are challenging to accurately segment because they are frequently thin, uneven, and visually similar to other surface patterns, including oil stains, textural differences, and shadows. To improve reliability, reliable decision-making in ADAS requires knowledge of model confidence in addition to localization performance, especially for complex damage patterns. Hence, to evaluate the reliability of model outputs, prediction uncertainty is explicitly quantified and connected with segmentation accuracy in the context of road crack segmentation, where uncertainty is estimated using stochastic dropout inference.

The final contribution of this thesis presents a framework for risk-aware multiclass pavement damage identification, aiming to bridge the gap between detection results and actionable decisions made by ADAS. The suggested system incorporates detection results with a risk evaluation process, in contrast to conventional methods that merely identify damage categories. An enhanced YOLOv11-Based architecture for unified multi-class road damage detection is proposed. This design can handle a variety of damage categories, including potholes, cracks, manhole covers, blurred or white road markings, and large-scale structural anomalies. Damage recognition is converted into actionable driving states by integrating a structured ADAS decision

logic into the framework. Depending on the type of damage discovered and the anticipated severity, the system classifies road conditions into operational levels, such as normal, warning, and emergency. The study extends the proposed methods to monitor dirt and gravel roads, which are commonly found in rural regions, in addition to paved road networks.

Extensive experiments on multiple datasets demonstrate that the proposed methods achieve state-of-the-art performance while being robust under a range of environmental conditions. Collectively, these contributions give a strong foundation for reliable, effective, and accurate road damage detection, providing useful information for preventative maintenance and raising the standard of safety in contemporary transportation networks.

TABLE OF CONTENTS

CANDIDATE'S DECLARATION i

CERTIFICATE BY THE SUPERVISOR ii

ACKNOWLEDGEMENT iii

ABSTRACT v

LIST OF FIGURES xii

LIST OF TABLES xv

LIST OF ABBREVIATIONS xvii

1 Introduction 1

1.1 Impact of Road Traffic Accidents on Public Safety 2

1.2 Road Damage Detection for Intelligent Transportation Systems . . . 4

1.2.1 ADAS-Oriented Perception Systems for Road Surface Condition Awareness 5

1.2.2 Limitations of Existing Road Damage Monitoring Approaches 7

1.2.3 Research Motivation 9

1.2.4 Problem Statement and Key Challenges 10

1.2.5 Research Objectives 10

1.2.6 Scope and Contributions of Thesis 11

1.2.7 Overall Research Framework 12

1.3 Thesis Organization 13

2 Literature Review 16

2.1 Pothole detection and segmentation Approaches 17

2.2 Road Crack Detection and Segmentation Methods 19
2.2.1 Traditional crack detection Methods 19
2.2.2 Deep Learning-Based Methods 20
2.3 Generic Road Damage Detection and Classification 22
2.3.1 Damage Detection in Road Infrastructure 22
2.3.2 Damage Detection in Disaster-Affected Scenarios 24
2.4 Unpaved Road Damage Detection 25
2.5 Vision-Based ADAS Systems Under Diverse Road and Driving Contexts 25
2.5.1 ADAS Performance Across Different Road Types 26
2.5.2 Influence of Weather Conditions on ADAS Reliability 26
2.5.3 Effect of Lighting Conditions on Vision-Based ADAS 27
2.5.4 ADAS Sensors for Road Damage Detection 27
2.6 Summary and Research Gaps 28

3 Pothole Detection and Segmentation for ADAS 33
3.1 Overview 33
3.2 Pothole Detection Using Improved YOLOv4 34
3.3 Pothole Segmentation using Enhanced Residual Convolutional Block based on U-Net for ADAS 39
3.3.1 Proposed Methodology 40
3.4 Experimental work and Results 44
3.4.1 Evaluation Metrics 45
3.4.2 Model Design Evaluation 46
3.4.2.1 Effect of Activation Function in the Residual Convolutional Block 46
3.4.2.2 Baseline Architecture Selection 46
3.4.2.3 Grad-CAM visualization 50
3.5 Summary 50

4 Road Crack Segmentation with Uncertainty Estimation 52
4.1 Overview 53
4.2 Proposed Methodology 55
4.2.1 Enhanced Residual Block 56

4.2.2 Dilated Depthwise Feature Extractor 58
4.2.3 Stochastic Dropout Inference 59
4.3 Experimental works and Results 61
4.3.1 Quantitative Performance Evaluation 63
4.3.2 Qualitative Visualisation on Road Crack Dataset 66
4.3.3 Uncertainty Estimation Analysis 68
4.3.4 Generalization Performance on Unseen Road Conditions . . 72

5 Risk-Aware Multiclass Pavement Damage Detection for ADAS 75
5.1 Overview 76
5.2 Proposed Distress-YOLO Framework 78
5.2.1 Overall Architecture 79
5.2.1.1 Proposed Hybrid Local–Global Attention Mechanism 80
5.2.1.2 Adaptive IoU Loss for Enhanced Localization Accuracy 83
5.3 Experimental Evaluation 85
5.3.1 Datasets 85
5.3.2 Evaluation Metrics 86
5.3.3 ADAS-Specific Evaluation 86
5.3.3.1 Risk–Coverage Analysis (RDD-2022) 87
5.3.3.2 Disaster Dataset ADAS Logic 87
5.4 Results and Discussion 88
5.4.1 Quantitative Performance on RDD2022 88
5.4.2 Risk-Aware Evaluation on RDD2022 91
5.4.3 Cross-Dataset Generalization on SODR 93
5.4.4 ADAS Severity Decision Analysis on Disaster Dataset . . . 94
5.4.5 Qualitative Analysis and Visualization 94
5.5 Summary 97

6 Unpaved Road Segmentation for Advanced Driver Assistance System 99
6.1 Overview 99
6.2 Methodology for Unpaved Road Damage Segmentation 101
6.3 Experimental work and Results 103

6.4 Summary 105

7 Conclusions, Future Scope and Social Impact **108**

7.1 Conclusions 108

7.2 Future Scope 110

7.3 Social Impact 110

Research Publications **114**

References **134**

LIST OF FIGURES

Figure 1.1	Distribution of Road Accident Types in India (2022) . .	3
Figure 1.2	Representative examples of common road surface damages	5
Figure 1.3	Major causes of road damage	6
Figure 1.4	Multi-sensor perception framework in Advanced Driver Assistance Systems (ADAS)	7
Figure 1.5	Traditional road inspection vs current technology	8
Figure 1.6	Conceptual transition from traditional road inspection to deep learning-based automated detection	8
Figure 1.7	Key challenges in vision-based road damage detection for ADAS	11
Figure 1.8	Overall research framework illustrating the end-to-end workflow adopted in this thesis.	13
Figure 2.1	Categorization of the literature review highlighting major research domains.	16
Figure 3.1	YOLOv4 network architecture	36
Figure 3.2	Representative sample pothole images from the dataset illustrating variations in road conditions	37
Figure 3.3	Quantitative performance comparison of trained models.	37
Figure 3.4	Qualitative results of different YOLO models for pothole detection (a) original images (b) Modified YOLOv4 detection result (c) YOLOv5 detection results	39
Figure 3.5	Proposed Convolution Block and Enhanced Residual Convolutional Block in ERCU-Net	41
Figure 3.6	Schematic of the proposed ERCU-Net architecture . . .	42
Figure 3.7	Proposed methodology of ERCU-Net	42

Figure 3.8	Prediction on CLAHE pre-processed images	43
Figure 3.9	Detailed configuration of encoder and decoder block of proposed ERCU-Net	44
Figure 3.10	Trade-off analysis between accuracy, model complexity, and inference speed for various segmentation architectures on the Pothole-600 dataset.	47
Figure 3.11	Pothole Segmentation Validation Loss Convergence . .	48
Figure 3.12	Sample qualitative demonstrations for the networks on the Pothole-600 test dataset. a) RGB Images b) Ground Truth c) Simple U-Net d) Attention U-Net e) Residual Attention U-Net f) MAFNet g) ERCU-Net	49
Figure 3.13	Grad-CAM visualization of proposed ERCU-Net Model a) RGB Images b) Ground truth images c) Corresponding heatmap images.	50
Figure 4.1	Proposed REDS-Net Model Architecture	56
Figure 4.2	Proposed methodology for road crack segmentation . . .	57
Figure 4.3	Enhanced Residual Block used in the proposed REDS-Net architecture.	58
Figure 4.4	Dilated Depthwise Feature Extractor module for crack segmentation.	59
Figure 4.5	Accuracy vs trainable parameter comparison	64
Figure 4.6	Qualitative comparison of crack segmentation results under diverse real-world scenarios. The columns represent different crack characteristics, including thin single cracks, complex interconnected cracks, low-contrast cracks, shadow-affected regions, and illumination variations. Rows correspond to (a) Original Image, (b) Ground Truth, (c) UNet (d) AttentionUnet, (e) Deeplabv3+ (f) Cracksegformer (g) REDS-Net(ours)	67
Figure 4.7	Image overlay visualization results of REDS-Net on crack segmentation test dataset	69

Figure 4.8	Uncertainty map visualisation of REDS-Net on crack segmentation test dataset	71
Figure 4.9	Uncertainty-Aware Performance Curve for REDS-Net	72
Figure 4.10	Visualization Results on External dataset	73
Figure 5.1	Proposed Methodology of Distress YOLO	79
Figure 5.2	Improved YOLOv11 Architecture	80
Figure 5.3	Internal architecture of the (a) C2PSA (Cross Stage Partial with Spatial Attention) block and (b) HLGA Module	81
Figure 5.4	Accuracy–speed trade-off of different YOLO variants for road damage detection	90
Figure 5.5	Precision-Recall (PR) and F1-score curves for the proposed model	90
Figure 5.6	Confusion matrix for the proposed Distress YOLO model	91
Figure 5.7	Risk–Coverage curve on the RDD-2022 dataset	92
Figure 5.8	Precision-Recall and F1-score curves for Distress-YOLO on SODR	94
Figure 5.9	The visualization results on the RDD2022: (a) Ground Truth, (b) YOLOv11 (baseline), (c) YOLOv11+AdaIoU (d) YOLOv11+HLGA (e) Distress-YOLO	95
Figure 5.10	Visualization Results on SODR dataset: a) original images b) Proposed Distress-YOLO	96
Figure 5.11	Qualitative visualization of ADAS severity logic on SODR dataset: a) original images b) ADAS decision overlays	97
Figure 6.1	Overall architecture of our proposed network	102
Figure 6.2	Proposed Methodology	103
Figure 6.3	IoU for different model architectures	104
Figure 6.4	F1 score for different model architectures	105
Figure 6.5	Sample qualitative demonstrations for the networks on the RTK test dataset. a) RGB Images b)Ground Truth images c) FPN d)LinkNet e)UNet f)UNet-Resnet18	106

LIST OF TABLES

Table 2.1 Comparison of Recent Deep Learning Methods for Road Damage Detection 30

Table 3.1 Comparative performance of the proposed Modified YOLOv4 with existing state-of-the-art object detection models for pothole detection. 38

Table 3.2 Effect of activation function in the residual convolutional block on mean IoU and average accuracy 46

Table 3.3 Comparative result analysis (%) on the Pothole-600 testing dataset 48

Table 4.1 Specifications of different crack datasets 62

Table 4.2 Comparison with baseline methods 64

Table 4.3 Influence of the DDFE module on crack segmentation performance 64

Table 4.4 Comparative results (%) of REDS-Net against state-of-the-art crack segmentation models 65

Table 4.5 Comparison of Model Complexity 66

Table 4.6 Quantitative evaluation of the model's uncertainty estimation 72

Table 4.7 Generalization result comparison on the NCD dataset. Metric values are reported in the remarks. 74

Table 5.1 Model performance comparison with baseline 89

Table 5.2 Per-class mAP@0.5 comparison 89

Table 5.3 Comparison with other object detectors 92

Table 5.4 Distress-YOLO performance on SODR dataset (fine-tuned for 50 epochs) 93

Table 5.5 ADAS severity decision distribution on disaster dataset . 94

Table 6.1 Comparison of segmentation models based on IoU and F1 score. Bold values indicate the best performance in the respective metrics. 105

LIST OF ABBREVIATIONS

ADAS	Advanced Driver Assistance System
AI	Artificial Intelligence
AIoU	Adaptive Intersection over Union
CNN	Convolutional Neural Network
DL	Deep Learning
ECE	Expected Calibration Error
FPS	Frames Per Second
FPN	Feature Pyramid Network
FN	False Negative
FP	False Positive
GAP	Global Average Pooling
Grad-CAM	Gradient-weighted Class Activation Mapping
GT	Ground Truth
HLGA	Hybrid Local Global Attention
IoU	Intersection over Union
SDI	Stochastic Dropout Inference
mAP	Mean Average Precision
NMS	Non-Maximum Suppression
PSA	Polarized Self-Attention
ReLU	Rectified Linear Unit
RDD	Road Damage Dataset
REDS-Net	Residual Enhanced Dilated Segmentation Network
DDFE	Dilated Depthwise Feature Extractor
ERCU-Net	Enhanced Residual Convolutional U-Net

ERCBs	Enhanced Residual Convolutional Blocks
RGB	Red, Green, Blue
ROI	Region of Interest
SODR	Social Media Image Dataset for Object Detection of Disaster-Induced Road Damage
SPP	Spatial Pyramid Pooling
SPPELAN	Spatial Pyramid Pooling with Efficient Layer Aggregation Network
PANet	Path Aggregation Network
RTK	Road Transverse Knowledge
NCD	Noise Crack Dataset
TP	True Positive
YOLO	You Only Look Once

Chapter 1

Introduction

Road infrastructure is a crucial element that supports a nation's social and economic growth [1]. Despite their significance, road networks are very vulnerable to deterioration brought on by aging, poor maintenance, and unintentional damage [2]. The probability of traffic accidents is greatly increased by hazardous road flaws as potholes, cracks, surface deformation, and structural breakdowns. In addition to endangering driving safety, these flaws significantly increase the risk of serious injuries, fatalities, and financial losses. For instance, an unexpected vehicle maneuver may result from the rapid appearance of a large pothole on a road, raising the risk of crashes and putting both drivers and passengers in danger [3]. Consequently, enhancing road safety and enabling intelligent transportation systems now depend heavily on the prompt and precise detection of road damage. Across the globe, traffic accidents remain one of the main causes of fatalities and serious injuries. The World Health Organization (WHO) reports that traffic accidents claim the lives of about 1.35 million people each year, and millions more sustain permanent injuries [4]. The situation is especially bad in developing countries, where deteriorating road conditions, a lack of infrastructure, and inadequate enforcement of traffic laws all contribute disproportionately to accident rates. In India, where over 90% of passenger travel occurs on roads, traffic and bad pavement are still important issues [5]. According to recent statistics [6], potholes alone cause more deaths in the

country than terrorist attacks. In order to improve driving safety, these numbers demonstrate the urgent necessity for proactive measures that monitor and alter road surface conditions.

Highways are a vital component of public infrastructure and are necessary for both economic growth and social mobility. Nevertheless, a variety of elements, such as temperature swings, heavy traffic, erosion from rainfall, and prolonged use, continuously degrade road surfaces. These factors eventually lead to the development of cracks and potholes in the structure. Minor issues that are not quickly identified and addressed can negatively affect driving comfort, road aesthetics, and, most importantly, traffic safety. Since these defects increase the likelihood of vehicle instability and collisions, they pose major hazards in their latter stages. Prompt detection and repair of road damage is crucial for lowering long-term maintenance costs, enhancing driving safety, and ensuring the reliable operation of modern transportation systems.

1.1 Impact of Road Traffic Accidents on Public Safety

About 150,000 people die in traffic accidents in India each year [7], which is equivalent to almost 400 deaths every day. This is a lot more than the 40,000 deaths that were recorded in well-known car markets like the US in 2016 [8]. The tremendous increase in the number of automobiles on Indian roads exacerbates this concerning trend. Road fatalities and injuries are expected to rise in the upcoming years if adequate preventive measures are not put in place, according to traffic research [9]. In order to predict hazards and improve road safety, intelligent and data-driven technologies are becoming more and more important in modern traffic planning and management. Accident frequency and severity can be significantly decreased by recognizing dangerous road conditions and developing evidence-based traffic policies.

To further understand the gravity of the situation, Figure 1.1 illustrates the distribution of road accidents in India for 2022 [10]. Only 5.8% of all accidents are non-injury accidents, with fatal accidents accounting for 33.8% of all instances, followed closely by serious injury accidents (31.1%) and minor injury accidents 29.3%. This shows the important need for effective road monitoring systems and smart detection tools, such as ADAS-enabled pothole detection solutions, to lower accident risk. It is clear that over 94% of traffic accidents lead to serious consequences.

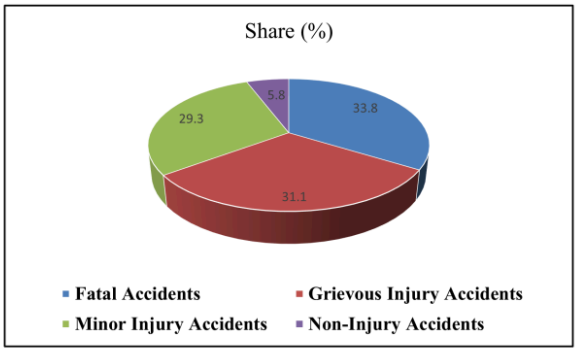


Figure 1.1: Distribution of Road Accident Types in India (2022)

Traffic accidents involve complicated interactions among human behavior, vehicle dynamics, and road conditions. These accidents can lead to serious injuries, fatalities, major property damage, or harm to the environment. They often result from sudden and unplanned events. Governments deal with heavy financial costs from these disasters, along with the tragic loss of life. Emergency response, infrastructure repair, and medical care add up to significant expenses. Traffic accidents are a major threat to public safety and sustainable transportation systems [11].

Therefore, to guarantee road safety, a comprehensive understanding of the interaction between the driver, vehicle, and road is required. In this sense, the primary objective of modern road safety efforts is to reduce accidents through improved traffic engi-

neering, raised driver awareness, and the judicious deployment of technology. In order to continuously evaluate road conditions, promote informed decision-making, and proactively reduce accident risks before serious consequences arise, automated and vision-based monitoring frameworks such as ADAS-enabled road damage and pothole detection systems offer a scalable and affordable option.

1.2 Road Damage Detection for Intelligent Transportation Systems

In modern intelligent transportation systems and ADAS environments, accurate road condition awareness has gained popularity. Overall traffic safety can be improved by using road damage information to support adaptive speed control, driver warning systems, and route planning modules [8].

Road surface degradation can take many different forms, such as rutting, alligator cracks, longitudinal and transverse cracks, potholes, and abnormalities in unpaved roads [12]. Potholes can cause tyre blowouts and destabilize vehicles; they are especially dangerous for two-wheelers. Cracks, on the other hand, damage pavement structural integrity and accelerate deterioration if left unrepaired. Representative examples of road damage, including potholes, cracks, and other pavement deteriorations, are shown in Figure 1.2. These illustrations show the variety of road surface distress patterns that any reliable automated detection system needs to be able to recognize in order to facilitate efficient maintenance scheduling and enhance traffic safety. Unpaved roads present an additional difficulty in rural and semi-urban locations since they are extremely susceptible to weathering and uneven wear, which makes driving dangerous and unpredictable [13].

In terms of ADAS, degraded road surfaces can impair modules for path planning, obstacle recognition, and lane detection. Water-filled potholes, uneven surfaces, and shadows all raise perceptual uncertainty, which, if left unchecked, could result in

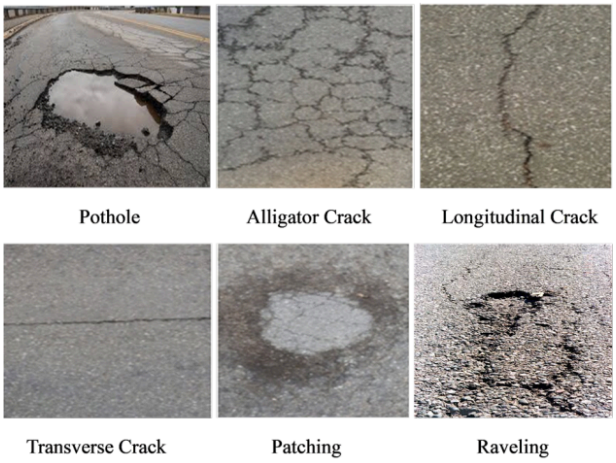


Figure 1.2: Representative examples of common road surface damages

inaccurate system reactions [14].

Numerous elements, such as high traffic volumes, poor drainage, deteriorating pavements, and environmental circumstances, might contribute to the incidence of this kind of damage. The main reasons for road damage are presented in Figure 1.3, which highlights the necessity of proactive and expandable road monitoring systems.

1.2.1 ADAS-Oriented Perception Systems for Road Surface Condition Awareness

Most existing research focuses on specific types of damage or is computationally inefficient for real-time applications such as ADAS. With the introduction of ADAS and the slow development of Autonomous Vehicles (AVs), the automotive industry has responded to these issues. These technologies heavily rely on perception systems that continuously monitor the driving environment, detect any hazards, and assist the driver in making judgments [15]. Road surface deterioration identification

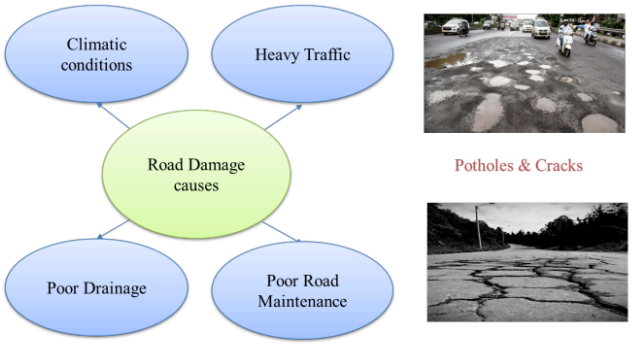


Figure 1.3: Major causes of road damage

has received relatively less attention, despite its clear link to driving comfort and safety, even as lane detection, pedestrian recognition, and accident avoidance have experienced notable gains. Conventional inspection techniques, such as manual surveys or specialized sensor vehicles, require a lot of resources and are not appropriate for widespread use in developing nations. This makes vision-based techniques that employ inexpensive cameras and ingenious algorithms essential for accurately identifying road damage. Deep learning has transformed computer vision in recent years, allowing for advances in visual comprehension, semantic segmentation [16], and object detection [17]. To increase driving comfort and safety, these techniques must be integrated with Advanced Driver Assistance Systems (ADAS). A variety of sensing modalities, such as cameras, LiDAR, RADAR, ultrasonic sensors, and infrared systems, are integrated by ADAS to monitor the environment around the vehicle continuously. As seen in Figure 1.4, these sensors provide features including adaptive cruise control, lane departure warning, emergency braking, park assist, traffic sign recognition, and collision avoidance [18]. They also provide thorough environmental coverage around the car. By incorporating reliable road damage detection into this multi-sensor perception framework, proactive driving modifications

can be enabled, thereby reducing accidents, and permitting early identification of hazardous road conditions.

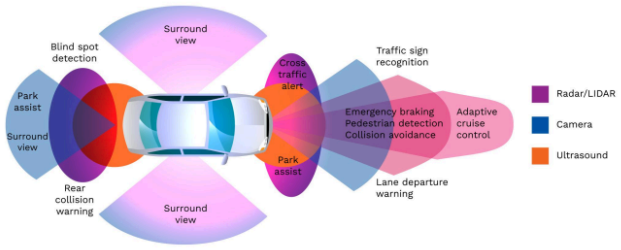


Figure 1.4: Multi-sensor perception framework in Advanced Driver Assistance Systems (ADAS)

1.2.2 Limitations of Existing Road Damage Monitoring Approaches

Despite government initiatives and regular maintenance programs to reduce pavement deterioration, traditional road inspection methods largely depend on manual surveys or specialized vehicles equipped with sensors, both of which are costly and time-consuming. Moreover, manual inspection cannot ensure the inspectors' personal safety at busy crossroads or on highways. For ongoing monitoring, these methods are not scalable, particularly in emerging nations with extensive and complex road networks. Traditional methods, as shown in Figure 1.5, are infeasible for nationwide continuous monitoring due to their high labor and equipment costs.

The shortcomings of traditional techniques of road inspection emphasize the necessity of scalable and automated solutions. Figure 1.6 shows the general transition and motivation towards intelligent vision-based technologies. The importance of deep learning frameworks for evaluating road conditions in real time is emphasized by this development.

To automate the identification of road conditions, researchers have explored several



Figure 1.5: Traditional road inspection vs current technology

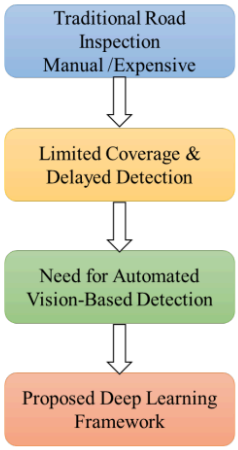


Figure 1.6: Conceptual transition from traditional road inspection to deep learning-based automated detection

methods, including grayscale threshold segmentation, wavelet transform, and edge and morphology detection [19]. It is still challenging to achieve reliable performance in real-world situations. The accuracy of the detection models often suffers due to changes in road surface textures, camera angles, lighting, and weather.

Recent progress in deep learning and computer vision offers a possible alternative. It allows for automated, vision-based evaluation of road conditions using images taken by smartphones or cameras attached to cars. By detecting cracks, potholes, and other surface flaws in real-time, these technologies greatly reduce the risk of accidents and simplify preventive maintenance. However, even with these improvements, there are still major research gaps. Existing models often do not consider the full extent of damages. They are not always reliable in various lighting and geographic conditions. Additionally, they seldom include uncertainty estimation, which is an important part of safety-critical systems like ADAS. Many studies overlook the need for multi-class detection systems that can handle various types of damage in real-world situations. They typically focus only on cracks or potholes. While unpaved roads are essential for rural mobility, the issue of monitoring them is still not well understood.

1.2.3 Research Motivation

Motivated by the limits of existing road inspection and vision-based detection methods, this study aims to offer a scalable and dependable framework for automatically identifying road damage in various real-world conditions. The growing use of autonomous driving technologies and Advanced Driver Assistance Systems (ADAS) requires reliable perception systems. These systems must accurately identify and locate various types of road-surface issues under different operating and environmental conditions. Improving driving comfort and safety requires not only finding the location and severity of damage but also estimating confidence reliably. This supports informed decision-making in applications where safety is critical.

1.2.4 Problem Statement and Key Challenges

Despite the significant progress in deep learning for identifying road damage, current methods are still not sufficient for trustworthy use in real-world Advanced Driver Assistance Systems (ADAS). Most current methods concentrate on particular types of damage. They work under limited environmental conditions and depend on predictions with low certainty. This reduces their usefulness in safety-critical driving situations.

Road damage detection systems need to work reliably in real-world ADAS applications, regardless of lighting, weather, or road surface conditions. However, to deliver real-time performance, they must strike a careful balance between efficiency and accuracy in detection. To evaluate prediction confidence and improve overall system reliability, we also need to measure uncertainty. Closing the gap between reliable ADAS-ready road monitoring solutions and current research prototypes requires tackling these issues. Vision-based road damage detection systems in ADAS face several major challenges, as seen in Figure 1.7. These restrictions include occlusions, tough environmental conditions, variations in light, and surface issues like water puddles, oil stains, and dirt. These problems highlight the need for dependable, uncertainty-aware, and real-time detection systems that work well in ADAS environments.

1.2.5 Research Objectives

The primary objectives of this research are as follows:

1. To design an efficient and accurate deep-learning model for pothole detection under diverse conditions.
2. To develop an efficient road crack detection model.
3. To build a generic model with improved accuracy, which can automatically

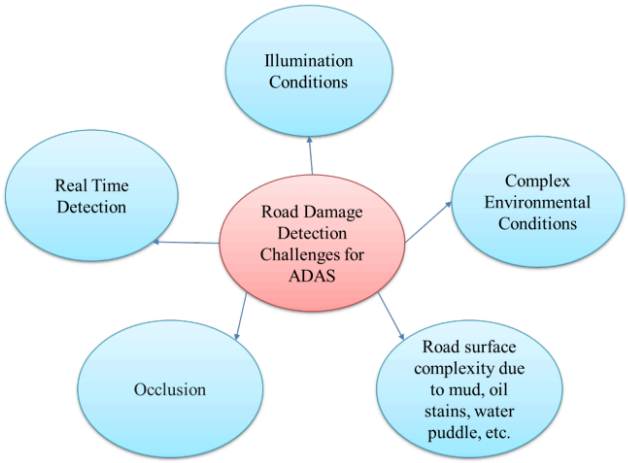


Figure 1.7: Key challenges in vision-based road damage detection for ADAS

detect and classify road damage.

4. Development of an effective and high-accuracy model for detecting damage on unpaved roads.

These goals focus on addressing the research limitations in robustness, generalization, and real-time use. They aim to gradually expand road damage detection from organized urban roads to unstructured rural areas.

1.2.6 Scope and Contributions of Thesis

To address the acknowledged issues, this thesis examines road damage identification from different viewpoints. First, it uses deep learning object detection frameworks to create effective pothole detection models. This includes segmentation-based methods for better boundary localization of potholes and road cracks. The research looks into uncertainty-aware crack segmentation to improve prediction accuracy, considering the limitations of deterministic models.

The scope is then expanded to include multi-class damage detection using improved YOLO variations. This allows for the detection of multiple damage categories within a single framework. The study concludes by addressing the challenge of segmenting dirt roads and extending intelligent road monitoring systems to better support rural road networks. Together, these contributions offer a comprehensive framework for identifying and evaluating the extent of road damage in support of ADAS and autonomous driving.

1.2.7 Overall Research Framework

The method used in this thesis is outlined in Figure 1.8. The process begins by collecting data from benchmark datasets that feature various road scenarios, such as unpaved roads, potholes, cracks, and damage from disasters. To improve resilience and generalization across different environmental conditions, the collected data is preprocessed with techniques like contrast enhancement, scaling, and data augmentation. Then, deep learning models for specific tasks are created and trained for identifying multiple types of damage, segmenting unpaved roads, recognizing potholes, and segmenting cracks while estimating uncertainty. Standard performance criteria are used to evaluate model performance, specifically looking at segmentation quality and detection accuracy. To ensure reliability and trust in safety-critical ADAS systems, the evaluation includes explainability analysis and uncertainty estimation. The research contributions in the thesis chapters are organized within this single framework.

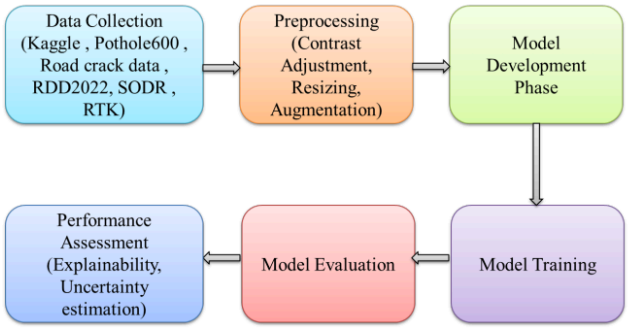


Figure 5. Overall Methodology Flow

Figure 1.8: Overall research framework illustrating the end-to-end workflow adopted in this thesis.

1.3 Thesis Organization

The thesis entitled, **"Road Damage Detection for Advanced Driver Assistance System"** comprises seven chapters. The organization of the research work is presented in a chapter-wise format as follows:

Chapter 1: Introduction

This chapter explains why road damage detection is important for ADAS and overall road safety. It discusses the effects of road accidents and the problems with traditional inspection methods. It also highlights the increasing need for effective and scalable road monitoring systems. The chapter also discusses the limitations of current vision-based methods. These include their failure to estimate uncertainty, reduced effectiveness in different real-world situations, and a lack of focus on detecting damage across multiple classes. It identifies research gaps and clearly outlines the study's objectives. This is followed by a brief overview of the thesis structure.

Chapter 2: Literature Review

This chapter reviews research on road damage detection. It includes topics such

as pothole identification, crack detection and segmentation, and multi-class damage classification. It highlights the pros and cons of each method as it shifts from traditional image processing techniques to modern deep learning methods. This chapter reviews earlier research and identifies key gaps. These gaps guide the current effort. They include the absence of uncertainty estimation, limited generalization in complex real-world situations, and a lack of focus on managing different damage categories.

Chapter 3: Pothole Detection and Segmentation for ADAS

This chapter describes the development of a strong framework for pothole segmentation and detection intended for ADAS applications. Firstly, a YOLO-based model is used to detect potholes accurately and in real time. Pothole regions are further segmented at the pixel level using an enhanced residual UNet-based segmentation model in order to accomplish accurate boundary extraction. For reliability and practical applicability in real-world driving scenarios, the framework is tested in a variety of lighting conditions, weather variations, and complex backgrounds.

Chapter 4: Road Crack Segmentation with Uncertainty Estimation

The suggested uncertainty-aware road crack segmentation model is presented in this chapter. Aiming for precise crack extraction, the developed system additionally estimates prediction uncertainty to improve reliability. The model is tested under difficult circumstances, such as intricate surface textures, shadows, and changes in lighting. The suggested method is evaluated over a variety of datasets, including unexplored datasets, to show its resilience as well as its capability for generalisation.

Chapter 5: Risk-Aware Multiclass Pavement Damage Detection for ADAS

In this chapter, a risk-aware framework for identifying and categorising various forms of road damage is presented. The suggested method uses an improved YOLO-based architecture to detect various pavement degradation categories in real time. The system includes a risk-based decision process in addition to detection, which helps

ADAS by classifying road conditions into normal, warning, and emergency scenarios, especially in environments affected by disasters. For practical deployment, the entire system is built to strike a compromise between computing efficiency and detection precision.

Chapter 6: Unpaved Road Segmentation for Advanced Driver Assistance System

This chapter concentrates on identifying damage on unpaved roads, an area that has received little attention in previous studies. To extend intelligent road monitoring systems to rural and semi-urban areas, a strong deep learning model is suggested for segmenting unpaved road anomalies.

Chapter 7: Conclusions, Future Scope and Social Impact

This chapter outlines the main conclusions of the study, emphasizes the contributions, and talks about possible avenues for further research. The social and practical effects of smart road damage detection systems on road safety and their role in better infrastructure maintenance are also discussed in this chapter.

Chapter 2

Literature Review

Road surface damage significantly impacts vehicle safety, driving comfort, and maintenance costs. Manual surveys and specialized sensor vehicles are examples of conventional inspection methods that are costly, time-consuming, and impractical for widespread use in underdeveloped nations. As a result, computer vision and deep learning-based automatic damage detection have emerged as a key area of study. This chapter highlights the restrictions that drive this research project by reviewing an extensive amount of literature in four important areas: pothole recognition, road crack detection, generic road damage detection, and damage detection on unpaved roads. Figure 2.1 presents the structured categorization of the literature surveyed in this thesis.

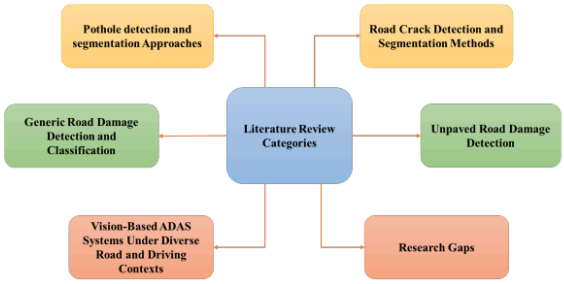


Figure 2.1: Categorization of the literature review highlighting major research domains.

2.1 Pothole detection and segmentation Approaches

Object detection is one of the most important tasks in autonomous driving since it enables the vehicle to avoid collisions when it recognizes obstacles. Potholes in the road may also result in collisions and damage to automobiles. Early studies on pothole identification mostly used traditional image-processing techniques, such as Hough transformations, edge detection, thresholding, and morphological procedures. These techniques were quite sensitive to surface noise, shadows, and changes in lighting, although they offered basic pothole localization. Pothole detection techniques are categorized in three approaches [20]: vibration-based approach [21], 3D reconstruction-based methods [22] and vision-based approach [23]. Pothole detection was traditionally done by a team of workers using recorded videos. This approach required a lot of time. The vibration method is unable to detect potholes in advance. When it gets over the pothole, it is able to make predictions. An affordable and simple pothole detection system using Raspberry Pi, camera and Micro SD card has been proposed [24].The authors employed OpenCV for real-time pothole detection, with the system deployed in a stationary configuration. In the case of 3D reconstruction method [25], laser scanning is used for pothole detection. This can detect the potholes in real time, but the method is costly.

Real-time and reliable solutions were made possible by the development of deep learning. The accuracy of detection was greatly increased using CNN-based classifiers and object detection frameworks like YOLO [26], SSD[27], and Faster R-CNN [28]. Deep Learning-based approaches for end-to-end road pothole detection train CNNs on annotated image data [29]. The first end-to-end single-modal CNN created for semantic segmentation was the fully convolutional network (FCN)[30]. Convolutions are created from all fully connected (FC) layers. A downsampling path and an upsampling path often make up an FCN. Convolutions, rectified linear units (ReLUs), and max pooling layers make up the converging path's usual CNN

design [31]. Convolutions and concatenations are used in a series of convolutions and concatenations along the expanded path to integrate the desired visual attributes and spatial information [32]. In order to extract context information for improved segmentation efficiency, PSPNet [33], DeepLabv3+ [34], and DenseASPP [35] use a pyramid pooling module. Moreover, GSCNN uses a two-branch structure made up of a shape branch and a regular branch, which can significantly enhance the semantic predictions on the boundaries [36].

As an alternative to the single-modal networks stated above, numerous data-fusion networks have also been proposed to increase the accuracy of semantic segmentation by extracting and fusing the features from multiple visual information modalities [37], [38]. To fuse the feature maps obtained from the RGB and depth branches, for example, FuseNet [39] and depth-aware CNN [40] leverage the well-known encoder-decoder architecture but apply distinct procedures. Additionally, by combining the features acquired from RGB and thermal images, RTFNet [41] was created to enhance semantic segmentation performance. Moreover, an element-wise addition fusion approach and an encoder-decoder architecture are used. Pan et al.[42] combined multispectral images taken by an unmanned aerial vehicle to detect potholes and cracks in asphalt roads. Zhen et al. have proposed a multimodal Attention Fusion network MAFNet [43] for the segmentation of potholes. However, compared to the ground truth, in the case of MAFNet, several edges still have some incorrectly detected pixels. Additionally, other research has utilized infrared imaging [44], depth maps [45], and stereo images [46] to enhance resilience in challenging situations, such as rain, snow, or low-light levels.

In spite of these developments, there are still important research gaps. The accuracy of pothole identification is still far from ideal in a variety of real-world scenarios, such as inconsistent lighting, bad weather, obstructions, and shifting camera angles. The performance of many current methods deteriorates in such challenging

circumstances, producing predictions that are not trustworthy. All of these drawbacks point to the critical need for a reliable and strong pothole detection model that can reliably produce high accuracy in a variety of situations, guaranteeing dependability and security in real-world Advanced Driver Assistance Systems (ADAS).

2.2 Road Crack Detection and Segmentation Methods

Road cracks are thin, elongated, and extremely uneven in shape, making them more difficult to detect than potholes. In contrast to potholes, cracks can differ greatly in breadth, continuity, and orientation, and frequently show no contrast with the surrounding pavement. This section provides a quick overview of common approaches and models for crack segmentation, such as traditional image processing, machine learning, and DNN.

2.2.1 Traditional crack detection Methods

Prior research on crack segmentation used standard image processing techniques like thresholding, filtering, and morphological approaches. These techniques mostly relied on variations in the background and the crack's color contrast. Images are typically thresholded at various values, noise-filtered, and crack-dilated to increase clarity, as seen in [47], and [48]. Gradient operators, thresholding [49], and Canny edge detection[50] were employed in traditional methods, but they were susceptible to background clutter and noise. Another common image processing technique utilized for crack segmentation in [51] was the Gabor filter. However, in real-world situations, different lighting conditions could affect these techniques and greatly reduce their effectiveness. More advanced studies have utilized pattern recognition techniques, such as 3D reconstruction and predefined filters, which were subsequently combined with machine learning methods like Support Vector Machines

(SVM), Random Forest, and AdaBoost, as demonstrated in [52], [53], and [54]. These techniques have proven useful in some situations, but in complicated situations, their reliance on feature selection and susceptibility to noise make them less trustworthy.

2.2.2 Deep Learning-Based Methods

Recent developments in deep learning methods have culminated in a fruitful period marked by noteworthy breakthroughs in several domains, such as segmentation, classification, and detection tasks. Crack detection was transformed by deep learning techniques, as models based on U-Net [55], FCN [56], SegNet[57], and GAN achieved more accurate pixel-level segmentation. Several DNN models [55] ,[58] used the encoder-decoder structure and U-Net architecture for crack segmentation. Hierarchical Feature Learning, which was extensively used for crack segmentation, was adopted by other DNN models [59] ,[60]. Zang et al. [61] developed a convolutional neural network model to detect fractures and non-cracks in road imagery using 500 3264 × 2448 images. Zhang et al. [62] constructed a CrackNet architecture without a pooling layer to classify cracks and non-cracks in asphalt pavement with a precision of 90.1%. The model was then enhanced to differentiate between cracks and sealed cracks [63]. Qu, Mei, Liu and Zhou [64] suggested a crack detection model of concrete pavement by employing finetuned LeNet-5 network to categorize the source images and VGG16 model [65] to extract features of crack for the detection process. In terms of F1 score, the model outperformed the VGG16 model [65] and U-net [66] by 25.2% and 2.8%, respectively. To identify crack, road marking, and intact area, Park, Bang and Kim [67] integrated segmentation and classification modules. The road surface is first extracted using a segmentation procedure to eliminate unwanted background, and then a unit patch is analyzed to

locate cracks in the derived road surface. Furthermore, to segment semantic cracks, a number of research have used encoder-decoder networks and fully convolutional networks (FCN) [68], and [69].A background removal module is introduced by Liu *et al.* [70] with the goal of reducing background noise interference. By improving the accuracy of crack feature extraction, this module strengthens the model's ability to absorb pertinent crack information. Pang *et al.* [71] integrates small-step shallow detail branches to achieve an outstanding balance between segmentation speed and accuracy. This incorporation successfully enhances the system's complex crack details, boosting its overall effectiveness. The Atrous space pyramid fusion attention module and the depth separable residual block are combined in ARD-Unet [72], allowing for accurate pixel-by-pixel identification of road cracks. The Transformer encoder module is integrated into a SegNet-like encoder-decoder architecture in CrackFormer II [73]. The accuracy of pavement crack segmentation is greatly increased by this integration, which makes it possible to retrieve contextual information from the feature channels efficiently. Though many studies have either used or improved upon pre-existing deep learning models, these methods might not necessarily produce effective or efficient outcomes in practical situations.

Xinwen Gao et al. proposed a new UNet based on multi-residual attention (MRA-UNet), which is used in road crack segmentation[74]. The accuracy of pavement crack segmentation is greatly increased by this integration, which makes it possible to retrieve contextual information from the feature channels efficiently. In order to address scale variation, such as micro-cracks versus vast cracks, more recent publications investigated transformer-based models, attention mechanisms, and multi-scale feature extraction. Feng Guo et al.[75] proposed a novel Crack Transformer (CT) model that models the long-range dependency and noisy pavement crack pixels at linear computational complexity and with high accuracy and efficiency. The model makes use of the Swin-Transformer and SegFormer with the hierarchical architecture

as the Encoder and a simplified Decoder structure in SegFormer. For reliable and accurate pixel-level pavement crack identification, Feng Guo et al. [76] suggested a Transformer-based semantic segmentation network that combines the UperNet with the attention module as the Decoder and the Swin Transformer as the Encoder. Qiao Dong et al. [77] has presented an enhanced sliding window algorithm that combines the YOLOv5 processing method for processing the two-dimensional images generated from point cloud data for automatic pavement crack detection with the random sample consensus (RANSAC) technique for directly extracting the crack information from point clouds.

However, there are still major challenges to overcome. Many studies generate false positives because they do not consider complicated backgrounds such as lane lines, oil stains, shadows, or rough textures. The majority of the works are also unsuitable for safety-critical applications since they do not consider uncertainty estimation. Furthermore, the evaluation of generalization across unexplored datasets is lacking. This study fills these gaps by creating a generalized crack segmentation model, adding uncertainty estimation, and confirming performance across several separate datasets to guarantee robustness and reliability.

2.3 Generic Road Damage Detection and Classification

This section provides a thorough analysis of the literature on the identification of road damage and the evaluation of infrastructure damage caused by disasters.

2.3.1 Damage Detection in Road Infrastructure

In recent years, many deep learning-based methods have been proposed , which can generally be categorized based on their architectural frameworks and the type of data used. The study of identifying and detecting road damage has made extensive use of deep learning techniques within the last ten years. Two-stage detectors, such

as Faster R-CNN [78], Cascade R-CNN [79], and Sparse R-CNN [80], achieve high accuracy at the expense of detection speed by first generating region proposals and then classifying them into certain object categories. One-stage detectors, such as the SSD [27], RetinaNet [81], EfficientDet [82], and YOLO series [83], [84], use a unified neural network to immediately translate image pixels to predictions to increase performance and decrease model parameters. Among them, YOLO offers the best balance of speed and accuracy. Its low memory usage and real-time performance make it ideal for transportation applications like road damage detection. For the classification of various types of road damage, Tang et al.[85] suggested a deep learning model called the iteratively optimized patch label inference network. A reliable CNN-based technique for detecting road damage on highways was published by Wang et al. [86], whereas Xie et al. [87], who focused on model optimization for mobile applications, provided an EfficientNet-based method for pavement crack segmentation. A generic framework for domain-adaptive road damage detection, known as DA-RDD, was introduced by Lin et al.[88] by combining unsupervised domain adaptation with road damage detection. By incorporating the squeeze-and-excitation (SE) and CA attention modules into YOLOv5, Wang et al.[89] increased the model's detection accuracy through ensemble learning. Wang et al. [90]improved the road area segmentation by fusing the image's contour data into a neural network. A cross-scale road object detector based on YOLOv3 was proposed by Shen et al. [91]. In order to identify and categorise road damage, Deepa et al. [92] integrated deep learning techniques with conventional image processing techniques. In order to address the issue of low accuracy in cross-scene road damage identification, Li et al. [93] suggested a transfer learning architecture. Road damage was identified in front-view images by Liu et al. [94] using the enhanced YOLOv5 model.

2.3.2 Damage Detection in Disaster-Affected Scenarios

The field of disaster-induced road damage is still notably under-represented, even though there are large datasets available for regular pavement distress identification. Images from social media platforms may provide important information for disaster response during calamities. Social media image datasets with various classification schemes are frequently discussed for particular response requirements. However, despite the availability of such resources, there remains a lack of specialized detection frameworks tailored to disaster-specific damage scenarios. Alam et al.[95] used social media photos to create a dataset for identifying the type of disaster, classifying informativeness, and determining the extent of damage. With an emphasis on collapse and blockage types, Wang et al.[96] presented the disaster-induced SODR dataset. An ensemble-based detection technique with attention-guided fusion was developed using this dataset to efficiently detect and categorize road damage in emergencies. Chen et al.[97] suggested a novel framework that combines Siamese damage identification and road extraction to evaluate road damage caused by disasters.

While these models had acceptable mean Average Precision (mAP) scores, they had drawbacks such as low geographic variety, uneven annotation quality, and class imbalance. Many methods fail to generalise across regions with varying road textures because they overfit to datasets from a single country. While prior studies have addressed road damage detection, none have systematically examined the generalizability of models across disaster-induced scenarios. This encourages us to focus on creating a general, high-accuracy model that can more robustly detect and categorise various types of damage, guaranteeing usefulness in large-scale implementations.

2.4 Unpaved Road Damage Detection

The field of unpaved road damage detection research is still in its early stages. Only paved asphalt or concrete roads are the subject of the majority of current research. Unpaved roads pose special difficulties because of things like uneven surfaces, loose dirt, dust, and lane markings. Furthermore, the construction of supervised learning models is limited by the lack of labelled datasets for unpaved roads. There are very few studies that try to visually assess unpaved roads, and even fewer that use deep learning methods. Ma *et al.* [98] reviewed the difficulty of detecting damages on unpaved roads in previous research, they emphasized the state-of-the-art techniques for road imaging and damage identification. Their effort focuses primarily on existing unpaved damage-detection techniques. The RoadFormer model, developed by Li *et al.* [99], uses a transformer-based architecture for RGB-normal semantic segmentation in road scene parsing and performs well under a variety of road conditions. Despite these developments, most approaches are applied to developed areas with well-maintained roadways, ignoring the difficulties presented by unpaved roads. Because of this, the issue of detecting damage in rural and unpaved road networks is still mostly unexplored. In order to fill this gap and provide reliable performance even with difficult textures and irregular road geometry, we are working to create a specialised model for unpaved road damage identification.

2.5 Vision-Based ADAS Systems Under Diverse Road and Driving Contexts

In order to enable reliable and secure driving decisions, Advanced Driver Assistance Systems (ADAS) significantly depend on an accurate perception of the road environment. The ability of emerging technologies to perceive the surroundings and anticipate risks more accurately and quickly than the human eye, as well as the

reduction of human error, are the primary reasons for the possible safety benefits [100]. Advanced driver assistance systems (ADAS) are a set of such technologies that introduce an augmented layer of safety through cooperation between the driver and the vehicle [101]. A thorough examination of the safety advantages requires an understanding of how the driving situation affects ADAS performance. Each ADAS's efficacy changes depending on the situation, such as whether the road surface is dry or wet. According to Scallon et al. [102] and Seacrist et al. [103], for example, drivers fitted with AEB who drive in dry conditions are more likely to benefit from AEB safety enhancements than drivers who drive in wet conditions.

2.5.1 ADAS Performance Across Different Road Types

Numerous studies have demonstrated that the type of road has a substantial impact on ADAS performance. Due to standardised lane markings, smoother geometry, and predictable traffic flow, motorways typically provide the best conditions for vision-based ADAS. This improves the performance of ADAS systems, Li et al. [104] and Mahdinia et al.[14]. However, abrupt bends, roundabouts, uneven lane widths, and deteriorating pavement conditions make rural roads extremely difficult to navigate. These factors increase perception mistakes and cause system responses to be delayed [105] and [106]. ADAS systems are adversely affected by the heavy traffic, frequent stops, complicated intersections, and varied road layouts found on urban roads [107]. These results emphasize the importance of accurately recognizing road damage in non-motorway contexts, demonstrating that vision-based ADAS systems remain highly sensitive to variations in road geometry and surface conditions.

2.5.2 Influence of Weather Conditions on ADAS Reliability

The reliability of Advanced Driver Assistance Systems (ADAS) perception systems is greatly affected by weather conditions. According to studies by Cafiso and Di

Graziano [108] and Bargman et al. [109], most ADAS features perform optimally in clear weather, where visibility, road friction, and the quality of sensor signals are at their highest. Rain, fog, snow, and storms can have a significant negative impact on system performance. They decrease visibility, obscure road objects, and reduce sensor accuracy [110] and [105]. Additionally, weather-related noise can lead to an increase in false detections and missed threats in camera-based Advanced Driver Assistance Systems (ADAS) [111]. Braking-related devices are often less effective in wet or icy conditions due to reduced friction coefficients and delayed response times, as noted by Seacrist et al.[103]. These limitations highlight the importance of developing weather-resistant visual perception models, especially for detecting road damage in real-world driving scenarios.

2.5.3 Effect of Lighting Conditions on Vision-Based ADAS

Lighting conditions also have a big impact on ADAS performance. Daytime is when vision-based systems typically achieve their highest accuracy because of sufficient illumination and strong visual contrast [112]. However, low light and nighttime conditions pose serious challenges, including reduced contrast, shadow artifacts, glare, and partial visibility of surface imperfections and lane markings [113]. Furthermore, at night, faded or obscured lane markers further reduce the accuracy of vision-based lane and road border recognition [114]. These findings highlight the need for illumination-invariant and uncertainty-aware models for road damage analysis since illumination variability remains a significant problem in camera-centric ADAS perception.

2.5.4 ADAS Sensors for Road Damage Detection

Advanced Driver Assistance Systems (ADAS) employ a range of sensory modalities, including cameras, LiDAR, radar, and inertial sensors, to perceive road environ-

ments. Among these, vision-based camera sensors have emerged as the most widely used and effective technique for identifying surface-level defects such as raveling, potholes, and cracks in roads. Because camera-based systems offer rich texture information and excellent spatial resolution, they can detect fine-grained road surface anomalies, such as thin fractures and shallow potholes. According to early research, monocular cameras could be used for pavement distress analysis due to their high frame rates and low cost of capturing detailed visual cues [115]. Camera-based ADAS systems have greatly increased detection accuracy and robustness in challenging conditions with the development of deep learning [116].

LiDAR sensors, on the other hand, have limitations in road damage analysis because of their sparse point density on flat road surfaces and decreased sensitivity to fine surface cracks, even though they are very effective for 3D scene reconstruction and obstacle detection [117]. LiDAR-based techniques are more suited for detecting large-scale deformations such severe potholes or subsidence, but they are less successful at segmenting thin cracks and are limited by high costs and computational overhead, which prevents them from being used in large-scale ADAS applications. Due to their low spatial resolution and incapacity to record surface texture information, radar sensors—which are frequently employed in ADAS for object detection and velocity estimation—are mostly inappropriate for detecting road damage [115]. For road damage identification, camera-based ADAS systems consistently offer the best trade-off between accuracy, cost, scalability, and real-time performance, according to the majority of available literature.

2.6 Summary and Research Gaps

This chapter surveyed the literature on road damage analysis, including ADAS-oriented vision systems, pothole detection and segmentation, road crack segmentation, road damage categorisation, and unpaved road damage detection. Conventional

image processing techniques mostly rely on manually created features and threshold-based processes, which have trouble with complicated lighting, shadows, occlusions, and changes in surface texture. Although machine learning techniques increase robustness somewhat, they still rely on manually created characteristics, which restricts their scalability and ability to generalize to new road conditions. Despite notable progress, several key research gaps remain. The following issues were discovered while reviewing the comprehensive literature review on road damage detection:

- An accurate and effective model should be developed because pothole detection accuracy is still far from satisfactory under diverse real-world conditions such as varying illumination, weather changes, occlusions, and camera viewpoints.
- Uncertainty estimation and generalization analysis are rarely employed in studies for accurate road crack identification, and most studies overlooked complex backgrounds, such as shadows or oil stains.
- A lot of research merely looks at whether damage is present or not. However, in order to act effectively in real-world situations, it is necessary to comprehend the type of damage and its extent.
- There is still much to learn about detecting road degradation on unpaved roads.
- Real-time development of road damage detectors is still open to research.

Table 2.1 summarizes recent deep learning based approaches proposed for road damage detection and segmentation. Although recent studies introduced advanced architectures such as fusion models, lightweight models, and transformer-based models to enhance damage feature extraction, some problems are still unsolved. Many approaches suffer from high computational cost, limited robustness under challenging road conditions, and difficulty in detecting fine-grained crack patterns.

Furthermore, existing approaches focus primarily on improving detection accuracy while overlooking the reliability of predictions. Current research has paid very little attention to uncertainty estimation, which is essential for safety-critical applications like Advanced Driver Assistance Systems (ADAS). For safer, more reliable deployment in intelligent transportation systems, this emphasizes the necessity of segmentation frameworks that are risk- and uncertainty-aware and capable of both accurately detecting road damage and producing confidence-aware predictions.

Table 2.1: Comparison of Recent Deep Learning Methods for Road Damage Detection

Study	**Method / Architecture**	**Dataset**	**Key Contribution**	**Limitations**
Kuang et al., 2024 [118]	Universal Crack Detection Framework	Multiple datasets	Generalized crack detection framework for intelligent road vehicles	High computational complexity
Liu et al., 2023 [73]	CrackFormer Transformer	Crack500	Captures long-range contextual crack features	Limited real-time performance
Feng et al., 2023 [119]	InconSeg Fusion Network	RGB-Depth datasets	Multi-modal obstacle segmentation	Modality inconsistency issues
Al-Maqtari et al., 2024 [120]	Lightweight Modular Network	Crack500, CFD	Lightweight architecture for efficient crack segmentation	Difficulty detecting thin cracks

Study	Method / Architecture	Dataset	Key Contribution	Limitations
Zheng et al., 2024 [121]	TV-Net	Crack datasets	Introduces tensor voting-based feature fusion and structural consistency loss to enhance crack structure representation and suppress noise interference	Limited evaluation on diverse road conditions and lacks uncertainty estimation
Wang et al., 2024 [96]	Ensemble YOLOv5 with Attention Mechanism	SODR, Road damage datasets	Introduces SODR dataset for disaster-induced road damage and proposes an attention-based ensemble YOLOv5 framework for improved detection accuracy	Limited to object detection and does not address pixel-level damage segmentation or uncertainty estimation
Lin et al., 2022 [88]	DA-RDD Framework	RDD2020, RDD2021	Introduces domain adaptation for road damage detection using image-level and instance-level feature alignment with adversarial learning	Requires complex training strategy and does not incorporate uncertainty-aware prediction for reliability assessment

Study	Method / Architecture	Dataset	Key Contribution	Limitations
Kong et al., 2025 [122]	Improved YOLOv8	Public pothole datasets	Introduces multi-scale feature fusion combining CNN and Transformer to improve pothole detection in complex environments	Model shows limited robustness for small-scale potholes and diverse road damage types
Li et al., 2024 [123]	Extended Mask R-CNN for Pothole Detection	Pothole dataset	Introduces a crowdsourced pothole dataset with diverse road conditions and proposes an extended Mask R-CNN model for accurate pothole detection and geometric feature extraction	Instance segmentation framework has high computational cost and may struggle with real-time deployment and small or partially occluded potholes
Feng et al., 2024 [124]	PotCrackSeg	RGB-D Road Obstacle Dataset	Proposes an RGB-D semantic segmentation network with dual semantic-feature complementary fusion to distinguish potholes and cracks for negative obstacle detection	Performance is sensitive to depth noise and relies on RGB-D sensors, which may limit robustness and practical deployment in diverse road environments

Chapter 3

Pothole Detection and Segmentation for ADAS

Potholes are one of the most common forms of road damage, and they can have a negative impact on vehicle conditions, comfort while driving, and road safety. Road potholes may increase the number of accidents, so detecting potholes and informing the concerned department can save the roads from worsening. For the Advanced Driver Assistance System (ADAS) to ensure safe driving, accurate segmentation of road potholes is a crucial feature. This chapter addresses the gap by enabling accurate pothole detection and fine-grained segmentation, which is essential for reliable severity assessment and real-time ADAS applications.

3.1 Overview

Roads are an indispensable means of transportation for any nation, furnishing communication facilities nationwide. Almost 90 percent of accidents happen due to poor road conditions all around the world [125]. In India, over 4000 people die yearly due to road potholes. In addition to being inconvenient, potholes on the road pose a safety risk because they can negatively impact driving comfort, the condition of vehicles, and traffic safety [126]. Vehicle traffic can cause the upper surface material of the road to move out and it generates small weak spots on the road. As time passes these spots worsen due to the continuous movement of vehicles over that

damaged surface and cause to emergence of new potholes [127]. Currently various computer vision algorithms are used efficiently for object detection and recognition task. The deep neural network performance has been rapidly enhanced for the task of object classification and detection. Learning-based strategies have increased the accuracy of image classification, object identification, and image segmentation tasks compared to traditional computer vision algorithms [128].

3.2 Pothole Detection Using Improved YOLOv4

This subsection concentrates on identifying potholes in road images, building on the broader goal of complete pothole research. Pothole identification was carried out utilising one of the distinctive real-time, fast and accurate representative YOLO on the pothole detection dataset. A modified YOLOv4 based object detection framework is adopted for road pothole detection in complex environments. YOLOv4 uses a Path Aggregation Network (PANet) and Spatial Pyramid Pooling (SPP) to enhance multi-scale feature representation, with CSPDarknet53 serving as the foundation for effective feature extraction. However, the original YOLOv4 framework with its large number of filters and multiclass loss formulation, involves a significant computational overhead, despite its great performance in multiclass object detection applications. Pothole detection and other single-class problems don't need this level of complexity. The network was specifically tailored for single-class learning by carefully modifying the YOLOv4 configuration file in order to get around these limitations. In particular, the following modifications were implemented:

- **Class Parameter Adjustment:**

  The classes parameter in each YOLO detection layer was set to **1**, since the model was designed to detect only potholes.

- **Filter Configuration Update:**

The number of filters in the convolutional layer preceding each YOLO layer was recalculated using the formula:

$$\text{filters} = (\text{classes} + 5) \times 3$$

For a single class, this becomes:

$$\text{filters} = (1 + 5) \times 3 = 18$$

As a result, 18 filters were used in each of the last three convolutional layers. This change successfully decreased pointless calculations without sacrificing detection accuracy.

- **Loss Function Simplification:**
  A binary cross-entropy loss function was used in place of the multiclass loss, allowing the network to focus on single-class detection. This modification enhanced gradient stability and facilitated faster and more precise convergence during training.

Figure 3.1 illustrates the overall architecture of the YOLOv4 framework used for pothole detection in this chapter.

The experiments were conducted on the publicly available Kaggle dataset [129]. The dataset includes 665 images that use shadow effects, varying lighting conditions, images of dry roads to replicate actual pothole situations, and potholes filled with water and moving vehicles. The model is trained on 80% of the dataset's images and then tested on the remaining 20%. Training datasets representing various road conditions and scenarios from various spatial locations were produced. Figure3.2 displays a few representative images from the collection.

The YOLO and its versions are trained on Google Colab, which offers access to a Tesla K80 GPU and 12 GB of RAM. The files needed for training are obj. names, which sets the names for classes, obj.data, which indicates how many classes there

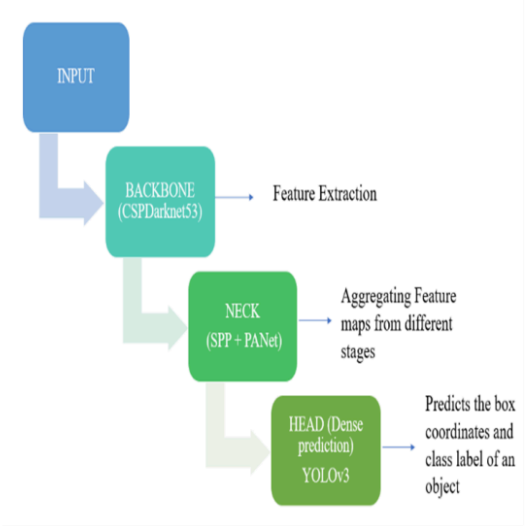


Figure 3.1: YOLOv4 network architecture

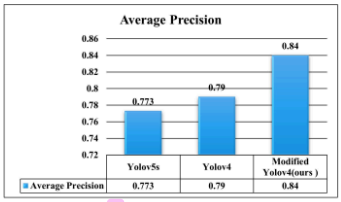
are, a path to train, test, and a backup folder. After every 100 iterations, the YOLO weights are saved in the backup folder. The configuration file is crucial for the training since it allows us to make adjustments in accordance with the demands of the model. Each YOLO model in our experiments is trained with a batch size of 32 and a learning rate of 0.001 for a maximum of 2,000 iterations. To test detection accuracy and computational efficiency under various real-world scenarios, the updated YOLOv4 model was trained and assessed on the selected pothole dataset.

Quantitatively, the suggested model outperformed the baseline YOLOv4 by more than 11% in mAP and 12.65% in recall, with an average precision of 84 %, average recall of 79 %, and mAP@0.5 of 82.39%. These enhancements show how well the model can locate and categorize abnormalities in the road surface. A reduction in false positives, which is necessary for the practical implementation of Advanced Driver Assistance Systems (ADAS), is also shown by the increase in precision. Fig-

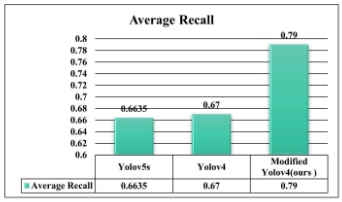


Figure 3.2: Representative sample pothole images from the dataset illustrating variations in road conditions

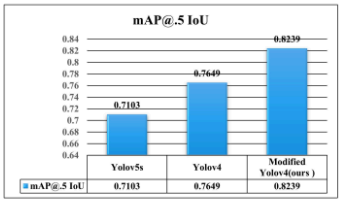
ure 3.3 shows the overall quantitative performance of the trained pothole detection models in terms of average precision, average recall, and mAP@0.5 IoU. These metrics provide a comprehensive evaluation of detection accuracy, resilience, and generalization.



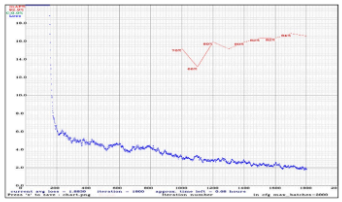
(a) Average Precision



(b) Average Recall



(c) mAP@0.5 IoU



(d) Overall performance comparison

Figure 3.3: Quantitative performance comparison of trained models.

Our results, when compared to other cutting-edge methods, demonstrate that our modified YOLOv4 model outperformed the original and YOLOv5 in terms of detection with the greatest average precision, Recall F1-score, and mAP@0.5 IoU. The comparison between our findings and the state-of-the-art methodology is publicized in Table 3.1. Our trained original YOLOv4, YOLOv5, and modified YOLOv4 have been compared to other cutting-edge efforts. The results unambiguously demonstrate that the proposed single-class YOLOv4 achieves a better balance between precision and recall.

Table 3.1: Comparative performance of the proposed Modified YOLOv4 with existing state-of-the-art object detection models for pothole detection.

Method	**Average Precision (%)**	**Average Recall (%)**	**mAP@0.5 (%)**
Gajjar et al. [130]	34.00	32.00	34.70
Shaghouri et al. [131]	88.00	71.00	–
Park et al. [26]	–	–	74.80
Our trained YOLOv4	77.30	66.35	71.03
Our trained YOLOv5	79.00	67.00	76.49
Proposed Modified YOLOv4	**84.00**	**79.00**	**82.39**

These quantitative advancements are further supported by qualitative analysis. The suggested modified YOLOv4 model exhibits strong pothole detection ability in a range of difficult real-world situations, as illustrated in Figure 3.4. The model can detect potholes with accuracy on wet road surfaces, in low light, and in nighttime illumination. Additionally, it efficiently manages several potholes in a single frame, reducing false detections and preserving constant localization accuracy.

Although detection is good at locating potholes, it is not precise enough at the pixel level to estimate their size and severity. To support severity estimation and informed driving decisions, ADAS systems require precise spatial information, including the area, shape, and boundaries of potholes. To achieve this, the research further extends to pothole segmentation, aiming to accurately detect, characterize, and evaluate pavement damage. This enables the facilitation of automated, precise, and economical road maintenance, safety, and repair planning.

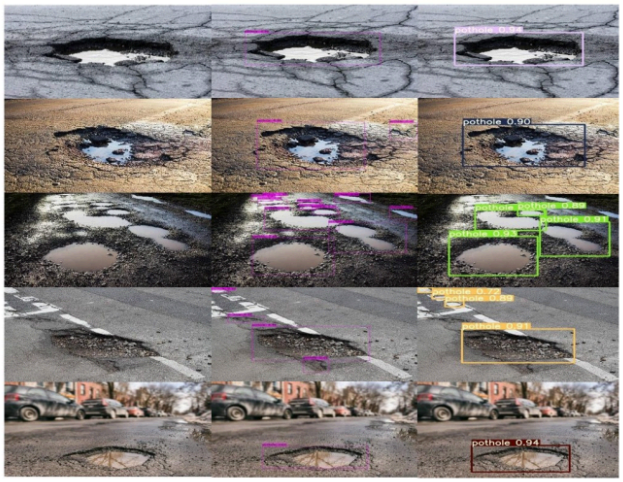


Figure 3.4: Qualitative results of different YOLO models for pothole detection (a) original images (b) Modified YOLOv4 detection result (c) YOLOv5 detection results

3.3 Pothole Segmentation using Enhanced Residual Convolutional Block based on U-Net for ADAS

The U-shaped topology of U-Net can well solve image segmentation tasks. However, due to the limitations of traditional convolution operations, U-Net cannot realise global semantic information[132]. Inspired by [133], a semantic segmentation neural network, which incorporates the benefits of residual learning and U-Net, Enhanced Residual Convolutional U-Net (ERCU-Net) is proposed to address the problem of correct detection of pixels at the edges of pothole areas. Enhanced Residual Convolutional Blocks (ERCBs) are incorporated into the encoder route of ERCU-Net, which builds upon the conventional U-Net segmentation framework. By incorporating 1×1 convolutions and residual connections into the shortcut path, the suggested ERCU-Net model expands on the conventional U-Net architecture to

enhance gradient flow and feature reuse. The proposed model employs skip connections to link the networks for the encoder and decoder. The residual convolution block uses a residual connection containing 1x1 convolution in its shortcut path, which allows the network to learn residual features and helps to improve the pothole segmentation performance by preserving more information from the input image. Mathematically, a residual connection can be defined as follows:

$$H(x) = F(x) + x \tag{3.1}$$

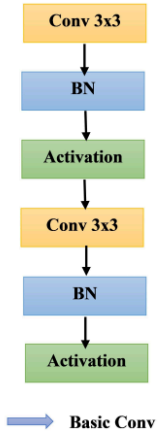
where x is the input to the layer, F(x) is the function of the layer, and H(x) is the output of the layer.

The network's capacity to capture potholes' fine-grained border features and global semantics is improved by this architecture. The structure of our basic convolutional block and residual convolutional block used in our proposed network is shown in Figure 3.5. The complete proposed network architecture is shown in Figure 3.6.

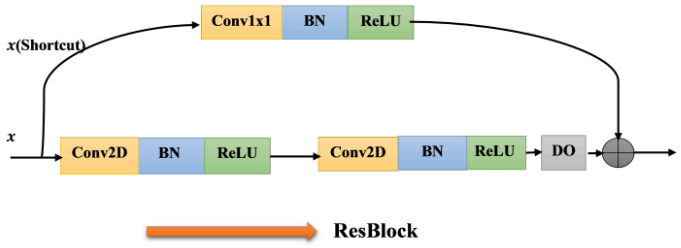
3.3.1 Proposed Methodology

In the proposed ERCU-Net model, residual connections are used with a shortcut path containing a 1x1 convolution with batch normalization and the activation function. The preceding feature map's depth is increased by applying the 1x1 convolution to the shortcut connectivity in the residual unit. The symmetric architecture of ERCU-Net includes expanding and contracting paths. The expanding path reconstructs spatial information, and the contracting path acquires context and learns high-level features. This symmetry makes it easier to encode and decode visual features, such as the pothole's edge, creating clearer boundaries. The overall proposed methodology is shown in Figure 3.7.

The input RGB images of the pothole with ground truth are fed to the preprocessing block and preprocessed using the CLAHE algorithm. Certain issues, like contrast



(a) Proposed Convolution Block



(b) Proposed Enhanced Residual Convolutional Block

Figure 3.5: Proposed Convolution Block and Enhanced Residual Convolutional Block in ERCU-Net

variability, feature extraction, and computing efficiency, are meant to be addressed using this algorithm. We have included pre-processed images using the CLAHE algorithm for more clarity in Figure 3.8, which shows the prediction of CLAHE pre-

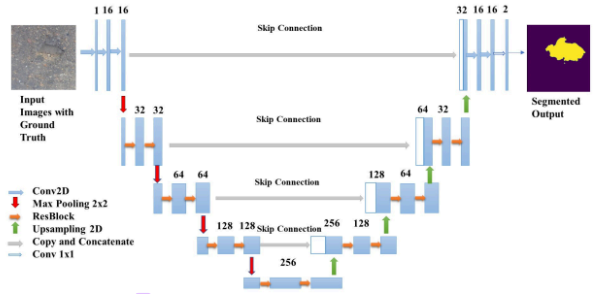


Figure 3.6: Schematic of the proposed ERCU-Net architecture

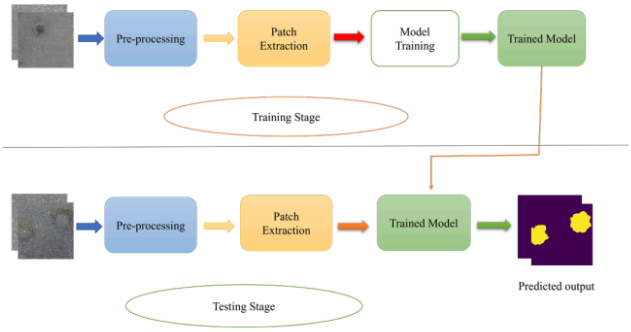


Figure 3.7: Proposed methodology of ERCU-Net

processed images for our proposed model. After this preprocessing, patch extraction takes place, which splits the images into patches. After training, the same steps are followed during the testing stage.

We have utilized a modified U-Net encoder block that consists of two 3x3 convolutions applied repeatedly, each followed by a rectified linear unit (ReLU), to create a single 3x3 convolution, following this convolutional block, there is a 2x2 max pooling operation with a stride of 2 for input stage downsampling. The proposed residual convolutional block is used in the remaining stages of the downsampling process. The 1x1 convolution layer in the shortcut path architecture provides dimensionality

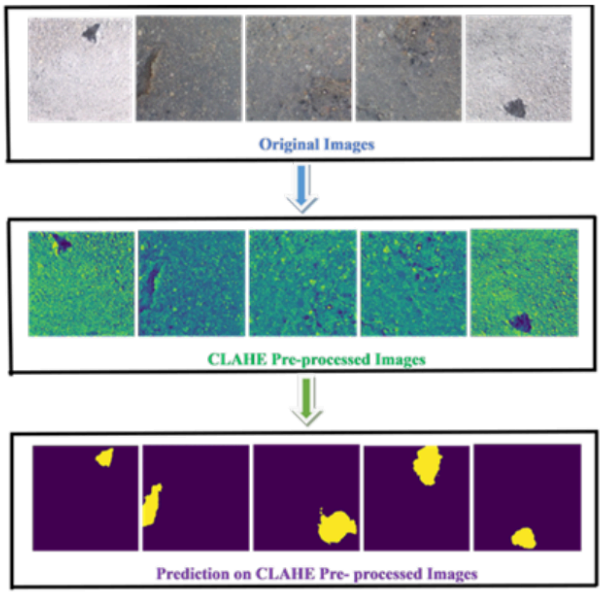


Figure 3.8: Prediction on CLAHE pre-processed images

reduction by lowering the number of filters while keeping crucial, feature-related information.

The decoder has five steps. The feature map resolution doubles for every step. In terms of resolution, the output matches the input exactly. The final step adds a sigmoid layer to determine the probability of each pixel falling into the background or pothole class. The detailed configuration of the encoder and decoder blocks used in the proposed model is shown in Figure 3.9.

Algorithm 3.1 depicts the training procedure for the proposed ERCU-Net model. The algorithm iterates the training epochs, modifies the model's parameters, and evaluates the model's performance on the validation set.

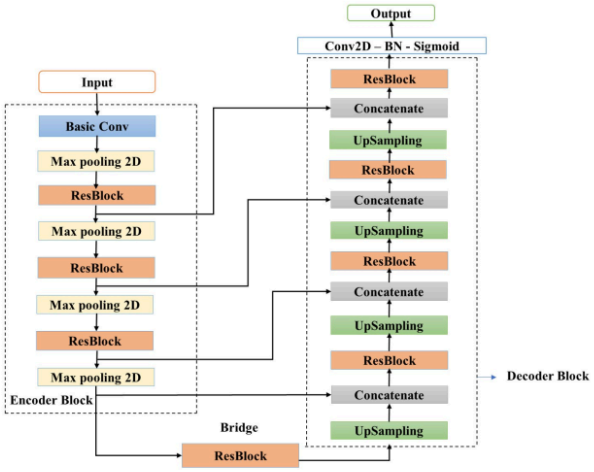


Figure 3.9: Detailed configuration of encoder and decoder block of proposed ERCU-Net

3.4 Experimental work and Results

For experimentation, the publicly available Pothole-600 dataset [36] was used, consisting of 600 RGB images with a resolution of 400×400 pixels. Following the data split strategy in [36], 240 images were used for training, and 180 each for testing and validation. To enhance model generalization, data augmentation techniques such as flipping, rotation, and cropping were applied, thereby expanding the training set to a total of 1,200 images.

The proposed ERCU-Net architecture was implemented using the Keras framework and trained on Google Colab equipped with a Tesla K80 GPU, 13 GB RAM, and 12 GB GDDR5 VRAM. The model, trained from scratch, incorporated residual convolutional blocks and employed a dropout rate of 20(%) in both the downsampling and upsampling paths to mitigate overfitting. Training was performed for 150 epochs with an initial learning rate of 0.001, using the Adam optimiser for efficient

Algorithm 3.1: Training Algorithm of ERCU-Net Model

Input: Pre-processed training dataset: $(x_i, y_i)_{i=1}^{N}$, where x_i - input image and y_i - corresponding ground truth mask for potholes

Output: Trained ERCU-Net model

```
Initialize ERCU-Net model
Define loss function
Choose adam optimizer with learning rate η = 1e − 3
Split dataset ← Train-val(70:30)
for each training epoch do
    for each mini-batch in the training set do
        Sample a mini-batch of input images (x_j) and corresponding masks (y_j)
        Perform forward pass through the ERCU-Net model to obtain predicted masks ŷ_j
        Compute the loss ℒ_j
        Compute the gradients ∂ℒ_j / ∂model parameters
        Update model parameters:
        model parameters ← model parameters − η · ∂ℒ_j / ∂model parameters
    end
    Evaluate model performance
end
```

convergence.

3.4.1 Evaluation Metrics

To assess the efficacy of semantic segmentation, we use two quantitative metrics: intersection over union (IoU) and accuracy. The following is how they are determined:

$$\text{Accuracy} = \frac{\text{True Positives}}{\text{True Positives + True Negatives}} \tag{3.2}$$

$$\text{IoU} = \frac{\text{True Positive}}{\text{True Positives + False Negatives + False Positives}} \tag{3.3}$$

In the ground-truth images of the dataset, there are two classes: pothole and unlabeled background. The average values, which are designated as Average Accuracy and mIoU, are the mean values across the two classes [43].

3.4.2 Model Design Evaluation

3.4.2.1 Effect of Activation Function in the Residual Convolutional Block

Multiple activation functions are tested under the same training settings to determine the best activation function for the residual convolutional block. To investigate the impact of activation functions on segmentation performance, the residual convolutional block is tested with various activation layers, including Softplus, Softsign, ReLU, and GeLU. The experimental results show that using ReLU within the residual block leads in higher mean IoU and average accuracy than alternative activation functions. Table 3.2 shows a quantitative comparison of various activation functions.

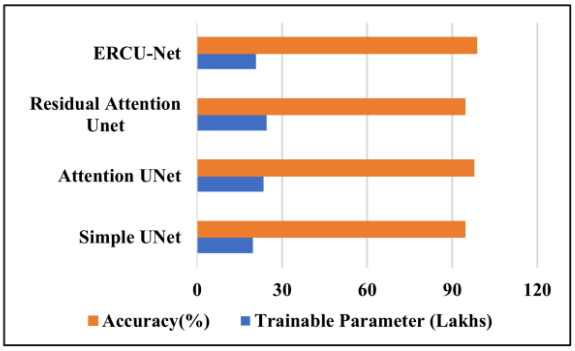
Table 3.2: Effect of activation function in the residual convolutional block on mean IoU and average accuracy

Activation Function	mIoU	Average Accuracy
Softplus	0.71	0.87
Softsign	0.86	0.97
GeLU	0.85	0.96
ReLU	**0.88**	**0.99**

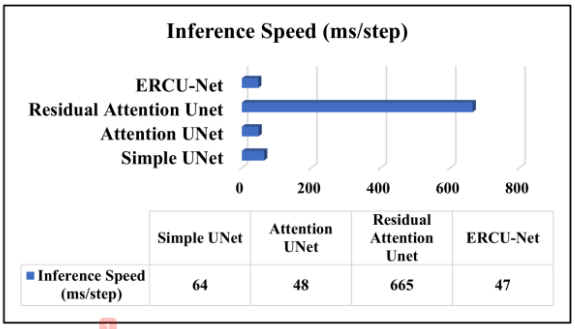
3.4.2.2 Baseline Architecture Selection

The residual-based UNet model is selected as the baseline model for the pothole segmentation problem due to its ability to effectively capture spatial features while maintaining computational efficiency. The performance of various models i.e. Simple UNet, Attention UNet, Residual Attention UNet, and ERCU-Net, is examined in terms of accuracy and trainable parameters and inference speed in Figure 3.10.

The quantitative comparison between ERCU-Net and other state-of-the-art segmentation models is summarized in Table 3.3. The results show that ERCU-Net outperforms all existing approaches on the Pothole-600 test dataset. It achieves a mean IoU of 88.45% and an average accuracy of 98.55% , which are superior to those obtained by SegNet, DeepLabv3+, ESPNet, GSCNN, and MAFNet.



(a) Accuracy vs trainable parameter



	Simple UNet	Attention UNet	Residual Attention Unet	ERCU-Net
Inference Speed (ms/step)	64	48	665	47

(b) Inference speed of trained models on pothole-600 dataset

Figure 3.10: Trade-off analysis between accuracy, model complexity, and inference speed for various segmentation architectures on the Pothole-600 dataset.

Figure 3.11 presents validation loss for different models on the Pothole-600 augmented dataset.

The enhanced segmentation performance of ERCU-Net is visibly demonstrated by the qualitative analysis, which is shown in Figure 3.12. The suggested model provides segmentation masks that are clearer and more accurate than those produced by existing architectures like U-Net, Attention U-Net, Residual Attention U-Net, and MAFNet.

Table 3.3: Comparative result analysis (%) on the Pothole-600 testing dataset

Model Name	Mean IoU	Average Accuracy
SegNet [57]	30.71	89.35
DeepLabv3+ [134]	55.31	97.39
ESPNet [135]	49.33	96.13
GSCNN [136]	52.71	97.71
MAFNet [137]	87.29	97.62
Simple U-Net	86.72	94.49
Attention U-Net	86.76	97.62
Residual Attention U-Net	82.63	94.54
ERCU-Net	**88.45**	**98.55**

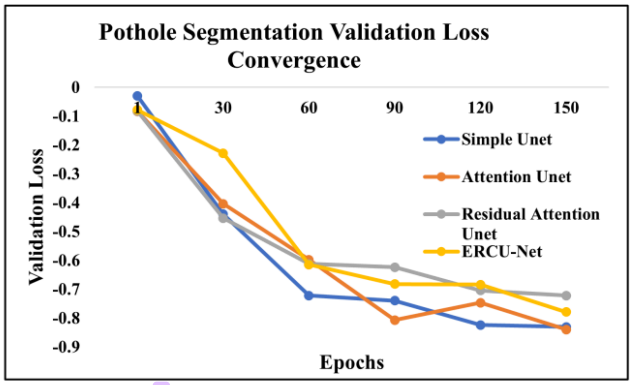


Figure 3.11: Pothole Segmentation Validation Loss Convergence

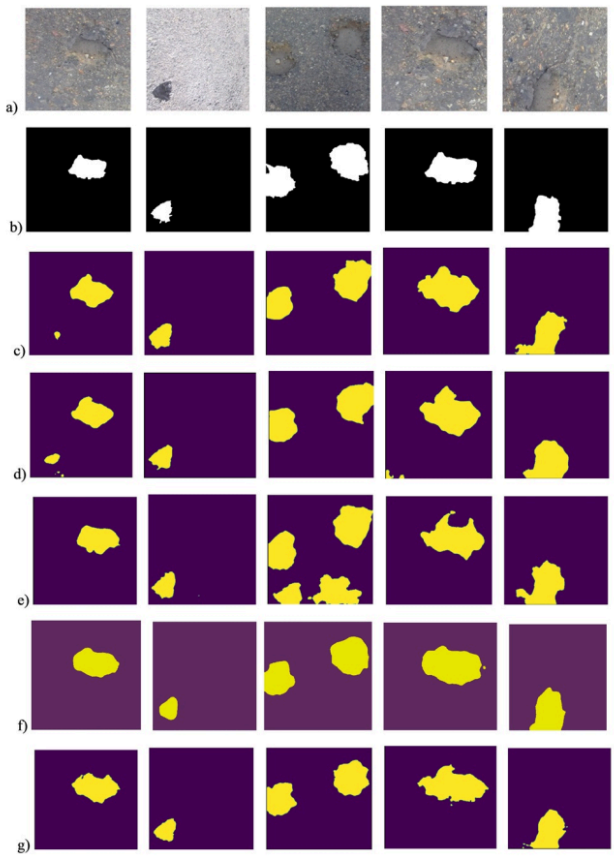


Figure 3.12: Sample qualitative demonstrations for the networks on the Pothole-600 test dataset. a) RGB Images b) Ground Truth c) Simple U-Net d) Attention U-Net e) Residual Attention U-Net f) MAFNet g) ERCU-Net

3.4.2.3 Grad-CAM visualization

Utilizing the Gradient-weighted Class Activation Mapping (Grad-CAM) technique, the heatmap is generated. The pothole regions are displayed as a heatmap by Grad-CAM, which uses the target's gradients to create the coarse localization map. The heatmap is created by utilizing the last convolution layer. Figure 3.13 demonstrates Grad-CAM visualization of the proposed ERCU-Net model.

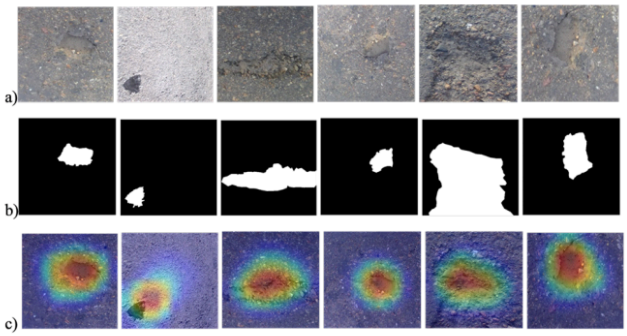


Figure 3.13: Grad-CAM visualization of proposed ERCU-Net Model a) RGB Images b) Ground truth images c) Corresponding heatmap images.

3.5 Summary

This chapter proposes the use of the ERCU-Net architecture for pothole segmentation and demonstrates its effectiveness on the Pothole-600 dataset of road images. Findings demonstrate that the ERCU-Net beats other cutting-edge techniques in terms of average accuracy and mIoU, making it a promising method for tasks requiring pothole segmentation. The analysis shows that ERCU-Net maintains a high level of computational efficiency while achieving good segmentation accuracy. Moreover, ERCU-Net outperforms other models in its exceptional capacity to precisely detect and define pothole borders, guaranteeing more accurate segmentation results and

distinct edge identification. More research in this field is still possible. Examining the model's real-time inference speed across various datasets is one possible avenue. In order to increase segmentation accuracy and robustness to environmental variability in pothole segmentation tasks, multimodal fusion approaches should be investigated.

Chapter 4

Road Crack Segmentation with Uncertainty Estimation

Road crack segmentation plays a crucial role in ensuring road safety and efficiency in the movement of Advanced Driver Assistance Systems (ADAS) on the road. Traditional segmentation methods often suffer from challenges such as poor generalization to unseen data and a lack of uncertainty estimation, which are critical for high-stakes applications like ADAS. This chapter proposes a robust road crack segmentation framework, REDS-Net (Residual Enhanced Dilated Segmentation Network), that addresses these issues by integrating the stochastic dropout inference (SDI) with a custom U-Net architecture. Traditional crack segmentation methods often suffer from limited robustness when subjected to real-world situations such as uneven sunlight, shadows, damp road surfaces and extremely thin or low-contrast cracks. The proposed method ensures reliable results for Advanced Driver Assistance Systems (ADAS) and road maintenance decision-making by concentrating on precise pixel-by-pixel segmentation of cracks and measuring the model's confidence in its predictions.

4.1 Overview

Public safety is seriously threatened by cracks on roads, pavements, and buildings, which can cause accidents and damage vehicles on roads [138]. Road crack identification and maintenance are essential to prevent further degradation and reduce maintenance cost [139],[140]. Automatic road crack identification methods have been made possible by the quick growth of computer technology, which includes large-capacity memory, high-performance CPUs, and digital image processing.

Deep learning has been an essential part of machine learning in recent years and has effectively tackled a number of computer vision problems [141],[142]. For segmentation tasks, encoder-decoder design has been the standard in recent years across a wide range of models, including Segnet [57] and Autoencoders [143]. Furthermore, attention methods were used in image segmentation models, which employed various strategies to give the image's most significant pixels more weight and its less significant pixels less weight. Numerous recent works [144] and [145] have proposed DNN techniques and models for crack segmentation. Despite this, cracks have unique features, such as edges, thickness, depth, shadows, and colour contrast, that set them apart from their surroundings in terms of shape and illumination. However, in low light, these features may resemble the crack, as in [146], quite a bit. The availability of many crack datasets made it possible for DNNs to learn the traits and patterns of cracks as a semantic segmentation work, or crack segmentation. Kheradmandi *et al.* [147] explain how digital image processing and depth learning extract road cracks. Pavement crack detection techniques fall into two primary categories: threshold-based and edge-based. Threshold-based techniques are fast and computationally effective [148], [49]. However, they are susceptible to noise, need careful threshold selection, and ignore the image's spatial information. The disadvantages of edge-based techniques [50],[149] are their computational complexity, sensitivity to noise, and difficulties in identifying cracks with high curvature or poor

continuity. In order to find concrete cracks, Cha *et al.* [150] employed a convolutional neural network (CNN) with a deep architecture and evaluated how well it performed in comparison to the Canny and Sobel edge detection techniques. Although road crack segmentation has advanced significantly, current techniques often fall short in generalizing over a variety of datasets and crack variations. Additionally, these methods do not include uncertainty estimation procedures and lack mechanisms to assess the confidence of their predictions, which are essential for detecting unclear predictions and enhancing decision-making in safety-critical applications like ADAS.

To address these gaps, we provide a new segmentation framework that incorporates uncertainty estimation into its design: REDS-Net (Residual Enhanced Dilated Segmentation Network) equipped with SDI to capture uncertainties. To efficiently capture complex crack properties while preventing gradient disappearance, the network makes use of Dynamic Feature Stabilizers (DFSs) and Enhanced Residual Block with skip connections. Moreover, the Dilated Depthwise Feature Extractor (DDFE) collects fine-grained characteristics that are essential for segmentation and improves receptive fields. Improved segmentation measures, such as accuracy, precision, recall, and IoU, are the outcome of combining these cutting-edge modules. The key contributions of this work are summarized as follows:

1. We proposed a novel residual-dilated architecture integrated with depthwise separable convolution to balance complexity and performance.
2. We incorporated pixel-wise uncertainty estimation into the segmentation pipeline, enhancing robustness and reliability in noisy conditions.
3. Extensive experiments are conducted across multiple datasets, including out of domain testing on the NCD dataset, demonstrating strong generalization.
4. We incorporated a comprehensive ablation study to validate the importance

of residual connections and dilated convolutions.

4.2 Proposed Methodology

The proposed model architecture is a robust U-Net-based design, enhanced with residual learning, a Dilated Depthwise Feature Extractor, and stochastic dropout inference to effectively segment cracks in input images. The first step is an input layer that receives RGB images with dimensions of 256 x 256 x 3. Each of the residual convolutional blocks in the downsampling path uses two convolutional layers with batch normalization and ReLU activation. A shortcut connection is then used to learn residual mappings. A max pooling layer is paired with each residual block to gradually reduce spatial dimensions while preserving significant characteristics.

In order to increase computational efficiency, the Dilated Depthwise Feature Extractor processes the feature maps at the bottleneck using depthwise separable convolutions. Dilated convolutions are then used to expand the receptive field and collect global context. Output feature maps of this module are of size 8 x 8 x 1024, which are then passed to the upsampling path.

Similar to the downsampling structure, the upsampling path gradually increases spatial resolution by employing upsampling layers. The upsampled features are concatenated with the corresponding encoder features at each stage in order to preserve spatial information that is lost during downsampling. For each concatenated feature map, a residual convolutional block is used to enhance it. The resulting feature maps have restored spatial dimensions.

After the last upsampling layer, stochastic dropout inference is used to improve generalization and provide uncertainty estimation. The output layer consists of a 1 x 1 convolution, which reduces the number of channels to match the number of classes, followed by batch normalization and a sigmoid activation function for pixel-wise classification. Combining feature refinement through residual learning,

detailed feature extraction through the dilated depthwise feature extractor, and uncertainty estimation through stochastic dropout inference, this architecture excels in road crack segmentation. The proposed model architecture is shown in Figure 4.1. The overall proposed methodology is shown in Figure 4.2. The Dataset, which consists of RGB images and corresponding ground truth masks, is first normalized, and then the dataset is divided into a ratio of 80:20 as training and testing sets for model training. After training, the performance of the model is evaluated based on accuracy, precision, recall, F1 score, and mIoU, and uncertainty is estimated using stochastic dropout inference.

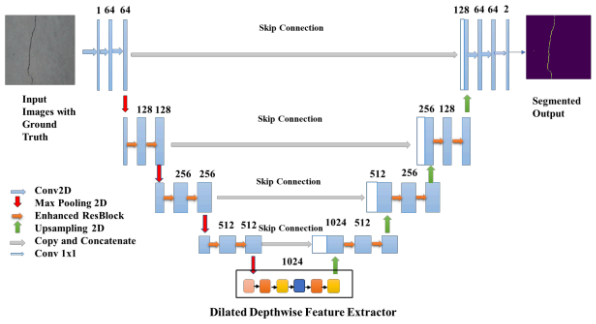


Figure 4.1: Proposed REDS-Net Model Architecture

The proposed REDS-Net architecture integrates three key modules designed to enhance feature representation, contextual understanding, and uncertainty estimation.

4.2.1 Enhanced Residual Block

The purpose of the enhanced residual convolutional block is to extract and refine features during backpropagation while preserving efficient gradient flow. The block receives a feature map with predetermined spatial dimensions as input. The block starts with a 2D convolutional layer that applies filters to extract features, followed by batch normalization, which stabilizes and accelerates training by normalizing

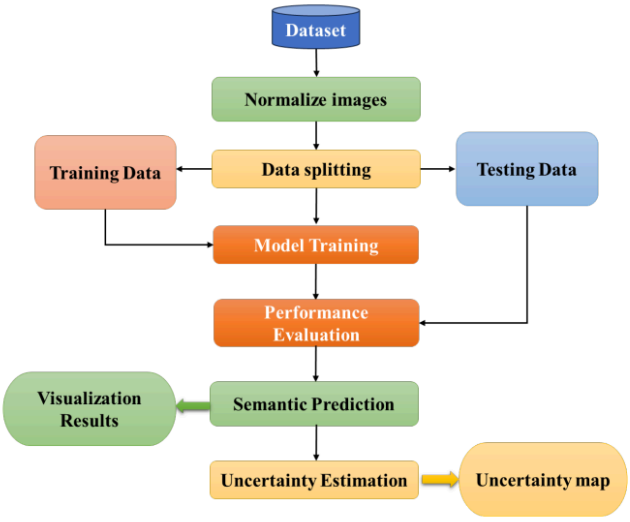


Figure 4.2: Proposed methodology for road crack segmentation

the activations. The second 2D convolution refines the features extracted by the first, followed by batch normalization and a dropout layer to prevent overfitting. The input dimensions are aligned with the main path by a shortcut connection that is made using a 1x1 convolution and batch normalization. To effectively enhance features, the output of the shortcut is appended to the output of the main path, creating a residual connection. Lastly, a ReLU activation is used to improve the feature representation while maintaining the dimensions of the input. Figure 4.3 depicts the structure of the enhanced residual convolution block.

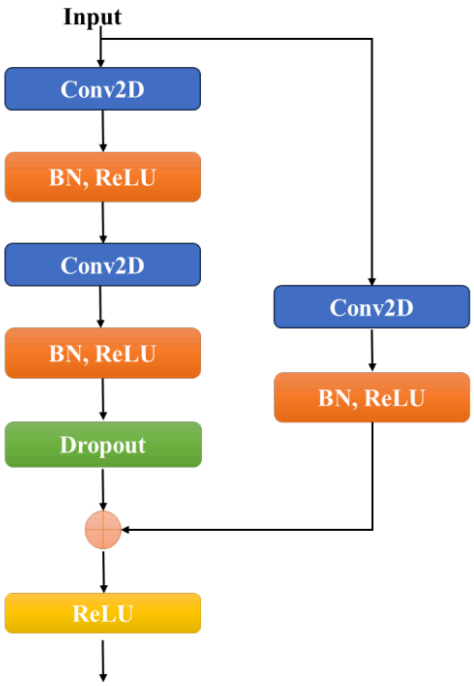


Figure 4.3: Enhanced Residual Block used in the proposed REDS-Net architecture.

4.2.2 Dilated Depthwise Feature Extractor

This module combines dilated convolutions with depthwise separable convolutions to improve feature extraction while maintaining spatial resolution. It starts with a Depthwise Separable Convolution and separates the spatial and depthwise convolutions to effectively learn spatial and channel-wise information. Next, batch normalization is applied for stable training, followed by a ReLU activation to introduce non-linearity. Then, without decreasing the spatial dimensions, a dilated convolution with a larger receptive field provides additional context. ReLU activa-

tion and batch normalization are again used to guarantee non-linearity and normalize the features. This module can be used for segmentation problems involving the extraction of dense features because the output maintains the same dimensions as the input. Figure 4.4 depicts the structure of the dilated depthwise feature extractor block used in the bottleneck part.

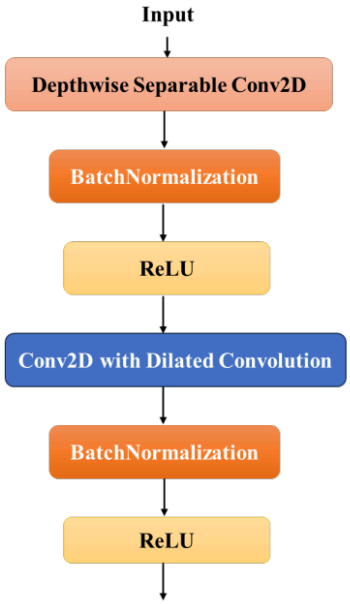


Figure 4.4: Dilated Depthwise Feature Extractor module for crack segmentation.

4.2.3 Stochastic Dropout Inference

In order to estimate uncertainty in the proposed model, stochastic dropout inference is applied during inference. This technique makes use of many stochastic forward passes while keeping dropout active. Given an input x and network parameters W,

the stochastic output for a single forward pass t is defined as follows:

$$\hat{y}_t = f(x; W \odot M_t) \tag{4.1}$$

Where $M_t \sim \text{Bernoulli}(1-p)$ represents a random binary mask applied to the weights with a probability (1- p), and $\odot$ denotes element-wise multiplication. To estimate the posterior predictive distribution $p(y \mid x, \mathcal{D})$, the model performs T stochastic forward passes, generating T predictions $\{\hat{y}_t\}_{t=1}^{T}$. The predictive mean, which is the final output of the model, is calculated as:

$$\mathbb{E}[y \mid x] \approx \frac{1}{T}\sum_{t=1}^{T}\hat{y}^t \tag{4.2}$$

The uncertainty in predictions is quantified using the variance of these T predictions:

$$\text{Var}[y \mid x] \approx \frac{1}{T}\sum_{t=1}^{T}\left(\hat{y}^t - \mathbb{E}[y \mid x]\right)^2 \tag{4.3}$$

In the segmentation model with SDI, for pixel i in the segmentation output, the predictive mean is given by:

$$\mathbb{E}[y_i \mid x] \approx \frac{1}{T}\sum_{t=1}^{T}\hat{y}_{i,t} \tag{4.4}$$

Where $\hat{y}_{i,t}$ represents the predicted value for pixel i during the t^{th} forward pass.

Through the use of SDI, the segmentation model can estimate the prediction as well as the associated uncertainty. The Dilated Depthwise Feature Extractor (DDFE) and residual blocks are merged with SDI in this model to provide precise feature representation and uncertainty quantification. The complete model training procedure is shown in Algorithm 4.1.

The most basic yet essential component of the training model is the loss function, which establishes how much the network's weights should change in order to reduce the error. We have used Binary Cross Entropy loss (BCE) in our experiments. The

Algorithm 4.1: Training Algorithm of REDS-Net for Road Crack Segmentation

Input: Pre-processed training dataset: $(x_i, y_i)_{i=1}^{N}$, where x_i - input image and y_i - corresponding ground truth mask for road cracks

Output: Trained REDS-Net model with SDI and DDFE

Initialize REDS-Net with DDFE
Loss function BCE
Choose Adam optimizer with learning rate $\eta = 1e-3$
Split dataset into train-val (80:20)
for *each training epoch* **do**
    for *each mini-batch in the training set* **do**
        Sample mini-batch (x_j, y_j)
        Perform forward pass through REDS-Net to get predicted masks $\hat{y}_j$
        Compute loss $\mathcal{L}_j$
        Compute gradients $\frac{\partial \mathcal{L}_j}{\partial \text{model parameters}}$
        Update model parameters:
        $$\text{model parameters} \rightarrow \text{model parameters} - \eta \cdot \frac{\partial \mathcal{L}_j}{\partial \text{model parameters}}$$
    end
    Evaluate model performance Accuracy, Precision, Recall, IoU, F1-Score
    Save the best model based on evaluation metrics F1-Score and IoU
    Apply Stochastic Dropout Inference (SDI) during evaluation
end

BCE loss is given by:

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)] \tag{4.5}$$

Where N is the total number of pixels ,y_i ($\in \{0, 1\}$) is the ground truth, and $\hat{y}_i$ is the predicted probability for the i^{th} pixel. This loss measures the discrepancy between true labels and predicted probabilities.

4.3 Experimental works and Results

The proposed architecture is developed using the Tensorflow-Keras framework. Training and testing of our network was conducted on Google Colab, utilizing a T4GPU, 12 GB of RAM, and 12 GB of VRAM. In our method, the input images are

Table 4.1: Specifications of different crack datasets

Dataset	Resolution	Training Set	Test Set
CFD	480×320	100	20
CRACK500	2000×1500	500	100
CrackTree200	Various sizes	175	25
DeepCrack	544×388	325	65
GAPs384	640×540	200	40
Forest	448×448	100	20

scaled to a resolution of 256 x 256 during network training. The model is trained using a batch size of 4 for 100 epochs. We utilize the binary cross-entropy as a loss function. The Adam optimizer with an initial learning rate of 0.001 is used for the training and testing of the model. At every newer highest F1score, the training model is saved.

In this study, six publicly available crack datasets— CFD [151], Crack500[146], CrackTree200 [152], DeepCrack [153], GAPs384 [154], and Forest [151] were combined to create a comprehensive and diverse dataset. The improved comprehensive dataset incorporates a broad range of surfaces to improve the resilience and generalizability of trained models, in contrast to the majority of crack segmentation datasets that gather data based on a single type of surface. This combined dataset has different types of crack surfaces. Additionally, each existing image contains a small percentage of crack pixels due to the features of specific cracks, which may lead to a class imbalance. In order to overcome this bias, we increased the amount of crack pixels in our sample using data augmentation techniques such as flipping, rotating, and random cropping to increase data diversity and quantity. Table 4.1 displays the specifications of each dataset in detail.

The aforementioned datasets were used for the experiments. The Accuracy, Recall, Precision, F-score, and mIoU matrices were calculated for every mini-batch. The F-score and mIoU are metrics that effectively represent the overall performance of the model. These values are calculated using the counts of True Positives (TP),

True Negatives (TN), False Positives (FP), and False Negatives (FN). They are computed in the manner described below.

$$\text{Fsc} = \frac{2 \text{ x Rec x Pre}}{\text{Rec + Pre}} \tag{4.6}$$

$$\text{Rec} = \frac{\text{True Positives}}{\text{True Positives + False Negatives}} \tag{4.7}$$

$$\text{Pre} = \frac{\text{True Positives}}{\text{True Positives + False Positives}} \tag{4.8}$$

$$\text{IoU} = \frac{\text{True Positive}}{\text{True Positives + False Negatives + False Positives}} \tag{4.9}$$

However, during evaluation, the segmentation results are significantly impacted by the threshold value. Hence, we have additionally calculated the uncertainty of our model using the stochastic dropout inference approach because the aforementioned matrices are insufficient for a reliable estimation of the performance.

4.3.1 Quantitative Performance Evaluation

For quantitative evaluation, we employed Accuracy, Precision, Recall, F1-score, and mIoU as the primary metrics. The initial comparison of REDS-Net was made with a number of baseline designs, such as U-Net, Attention U-Net, DeepLabV3+, and CrackSegFormer. As indicated in Table 4.2, the findings showed that REDS-Net continuously produced the greatest overall performance, with an F1-score of 80.85% and a mIoU of 69.37%. These findings suggest that feature discrimination and crack boundary delineation are much enhanced by the combination of the Dilated Depthwise Feature Extractor and the Enhanced Residual Block.

In addition, the model achieves higher accuracy with fewer trainable parameters compared to other architectures, emphasizing its computational efficiency, as depicted in Figure 4.5.

Table 4.2: Comparison with baseline methods

Baseline	Accuracy	Precision	Recall	F1-Score	mIoU
U-Net	96.50	79.64	73.55	74.60	65.10
Attention U-Net	96.80	79.60	70.50	72.90	64.60
DeepLabv3+	96.33	77.75	67.20	71.04	65.92
CrackSegFormer	96.80	81.94	70.18	74.67	67.90
REDS-Net (Ours)	**97.70**	**88.60**	**74.35**	**80.85**	**69.37**

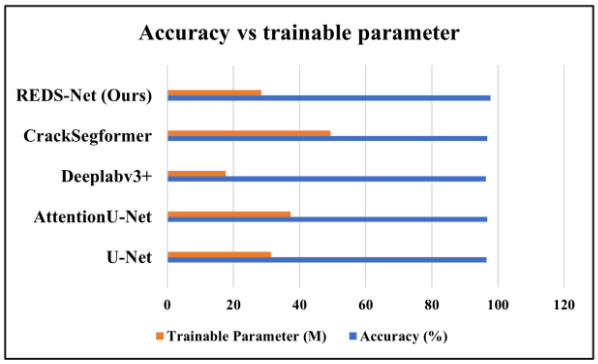


Figure 4.5: Accuracy vs trainable parameter comparison

The contribution of each module was examined by ablation experiments. A discernible decline in segmentation accuracy and IoU was noted upon removal of the DDFE module, indicating its significance in obtaining fine-scale contextual information. Table 4.3 presents the influence of the DDFE module on the crack segmentation dataset.

Table 4.3: Influence of the DDFE module on crack segmentation performance

Baseline	Accuracy	Precision	Recall	F1-Score	mIoU
w/o DDFE	96.97	84.25	70.94	75.37	66.31
REDS-Net (Ours)	**97.70**	**88.60**	**74.35**	**80.85**	**69.37**

For further validation, REDS-Net was benchmarked against state-of-the-art segmentation networks such as HED [155], SegNet [57], FFEDN [156], U2CrackNet

[157], DeepCrack [153], CracksegNet [158], SegFormer [159], HrSegNet [160], Hybrid-Segmentor [161], and TV-Net [121], As summarized in Table 4.4, REDS-Net surpassed most of these models in all major metrics. While TV-Net achieved competitive results, its reliance on manually tuned tensor fields limited adaptability across datasets. In contrast, REDS-Net maintained consistent performance across six different crack datasets, demonstrating strong generalization capability and adaptability to complex environments.

Table 4.4: Comparative results (%) of REDS-Net against state-of-the-art crack segmentation models

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**mIoU**
HED [155]	–	55.70	69.30	61.70	44.70
SegNet [57]	–	70.60	74.60	72.50	54.50
FFEDN [156]	–	66.50	73.90	70.21	54.20
CrackSegNet [158]	–	72.50	74.70	73.60	44.70
PSPNet [162]	–	65.93	67.68	66.79	59.53
U2CrackNet [157]	95.96	–	–	76.24	58.48
DeepCrack [153]	–	73.46	44.20	55.19	62.54
SegFormer [159]	96.50	75.00	71.90	73.10	58.30
HrSegNet [160]	96.90	80.00	72.40	75.70	61.20
Hybrid-Segmentor [161]	97.10	80.40	74.40	77.00	63.00
TV-Net [121]	**99.03**	64.65	**76.45**	68.48	–
REDS-Net (Ours)	97.70	**88.60**	74.35	**80.85**	**69.37**

To evaluate the computational efficiency of the REDS-Net model, we compared it with several crack segmentation models in terms of parameter counts, model size, and inference time. REDS-Net demonstrated the smallest model size (28.4 MB) and the fastest inference speed (0.105 s per image), as depicted in Table 4.5, validating its appropriateness for embedded ADAS systems and real-time traffic inspection. The model maintained excellent accuracy and segmentation quality in spite of its compact size. In contrast, models like FCN, SegNet, DeepCrack, CrackSegformer, Attention U-Net, and DeepLabv3+ require significantly more resources.

Table 4.5: Comparison of Model Complexity

Models	Params(M)	Model Size(MB)	Inference Time/Image(s)
FCN [163]	32.95	–	0.114
SegNet [57]	–	–	0.116
DeepCrack [164]	–	–	0.145
Z. Qu *et al.* [165]	58.72	–	–
U-Net	31.40	359.6	0.115
Attention U-Net	37.33	427.6	0.163
DeepLabv3+	40.35	385.4	0.136
CrackSegFormer	49.43	566.0	0.235
REDS-Net (Ours)	**28.40**	**325.7**	**0.105**

4.3.2 Qualitative Visualisation on Road Crack Dataset

Qualitative visualization, as shown in Figure 4.6 further demonstrates the model's accuracy in segmenting small cracks, even in difficult situations like low contrast surfaces, shadows, and varying lighting.

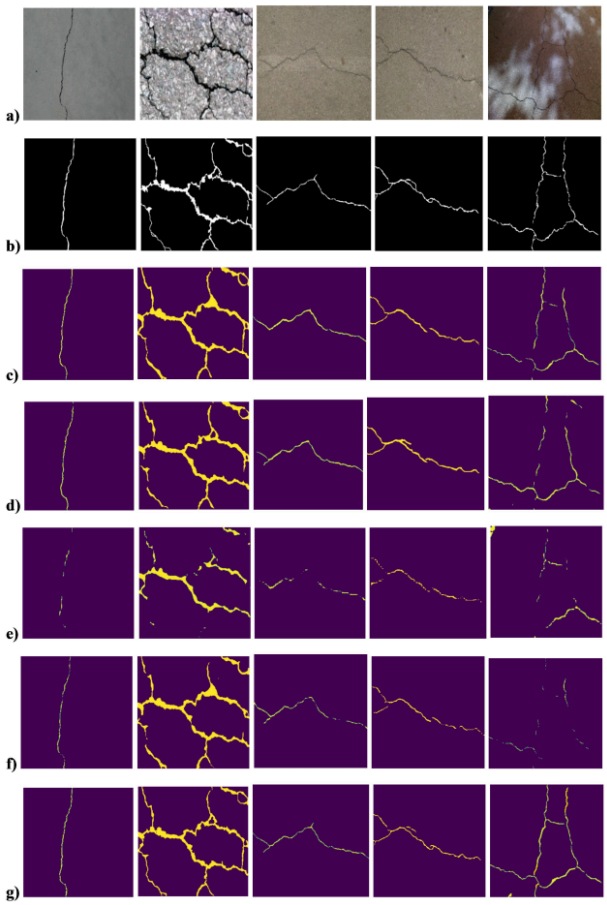


Figure 4.6: Qualitative comparison of crack segmentation results under diverse real-world scenarios. The columns represent different crack characteristics, including thin single cracks, complex interconnected cracks, low-contrast cracks, shadow-affected regions, and illumination variations. Rows correspond to (a) Original Image, (b) Ground Truth, (c) UNet (d) AttentionUnet, (e) Deeplabv3+ (f) Cracksegformer (g) REDS-Net(ours)

The predicted segmentation mask is superimposed on the original input image to improve interpretability. A clear visual depiction of the segmented output of our proposed model is provided by this overlay in Figure 4.7, which emphasises the identified road cracks immediately on the input. To maintain the visibility of the underlying road features for improved context, the segmentation mask is shown in green.

4.3.3 Uncertainty Estimation Analysis

During testing, REDS-Net uses Stochastic Dropout Inference (SDI) to model predicted uncertainty in road crack segmentation. Unlike deterministic inference, dropout layers are preserved throughout inference, permitting the network to provide several random predictions for the same input.

Given an input image x, the trained REDS-Net model f_θ performs T stochastic forward passes:

$$\hat{y}_t = f_\theta(x; \text{dropout}), \quad t = 1, 2, \ldots, T \tag{4.10}$$

where $\hat{y}_t \in [0, 1]^{H \times W}$ denotes the probabilistic crack segmentation output of the t-th forward pass.

The final predictive segmentation is obtained using the Monte Carlo mean:

$$\mu(x) = \frac{1}{T} \sum_{t=1}^{T} \hat{y}_t \tag{4.11}$$

Pixel-wise epistemic uncertainty is quantified using predictive variance:

$$\sigma^2(x) = \frac{1}{T} \sum_{t=1}^{T} (\hat{y}_t - \mu(x))^2 \tag{4.12}$$

The confidence behaviour of REDS-Net in challenging real-world conditions is analysed and visualized using the calculated prediction variance. Uncertainty maps

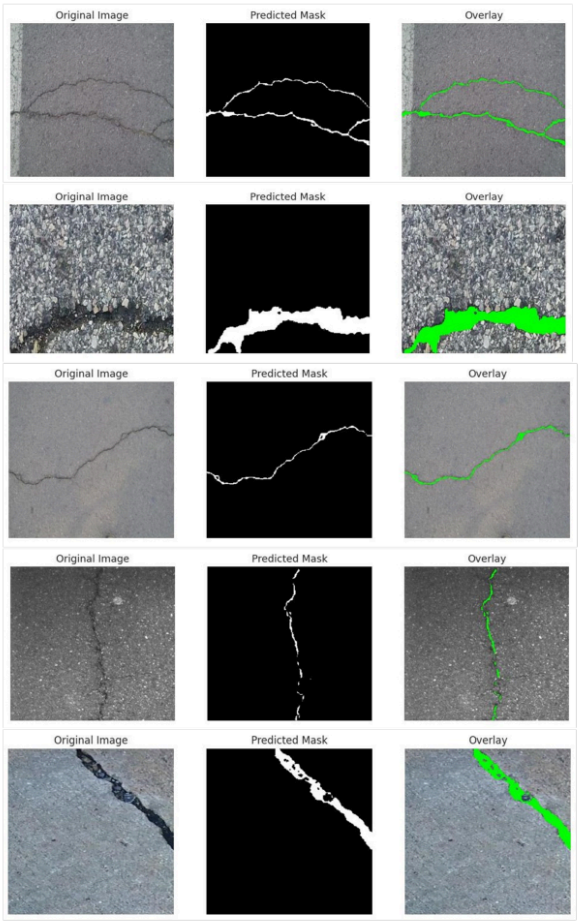


Figure 4.7: Image overlay visualization results of REDS-Net on crack segmentation test dataset

are generated by normalizing the variance values:

$$U(x) = \frac{\sigma^2(x) - \min(\sigma^2)}{\max(\sigma^2) - \min(\sigma^2)} \tag{4.13}$$

To evaluate the trustworthiness of the model, the uncertainty map is normalized to [0, 1] and displayed as a colourmap. Figure 4.8 shows the uncertainty map visualization of our proposed model, highlighting the uncertainty score on the top of the image. In these uncertainty map images, dark regions highlight the area of high predictive uncertainty. Each image also presents the uncertainty score, representing the average predictive variance across all pixels. Lower scores indicate confident predictions, while higher scores suggest ambiguity. This interpretability helps assess the model's reliability and offers actionable cues for ADAS systems to flag uncertain regions for human oversight or conservative decision-making.

To assess the accuracy of uncertainty predictions, we conducted a calibration analysis utilizing IoU-based correctness labels and image-level uncertainty scores at multiple thresholds (0.3 to 0.7). To determine how well the uncertainty estimates represent the model's prediction confidence, we calculated three common metrics: Expected Calibration Error (ECE), Negative Log-Likelihood (NLL), and Brier Score. Lower values of these metrics indicate better-calibrated and more reliable uncertainty estimates. Table 4.6 demonstrates that all metrics steadily improve as IoU thresholds rise, suggesting that REDS-Net's uncertainty estimations closely match segmentation quality in high-confidence areas. With an ECE of 0.1471, NLL of 1.1991, and Brier Score of 0.2721 at an IoU threshold of 0.7, REDS-Net demonstrates accurate and interpretable uncertainty estimation in challenging road crack scenarios.

We have also included an uncertainty-aware performance curve in Figure 4.9, to validate the effectiveness of REDS-Net's uncertainty estimation. As coverage increases, meaning more uncertain predictions are included, the average IoU gradually decreases. This confirms that REDS-Net reliably identifies confident predictions, making its uncertainty estimates useful for selective and robust crack detection in

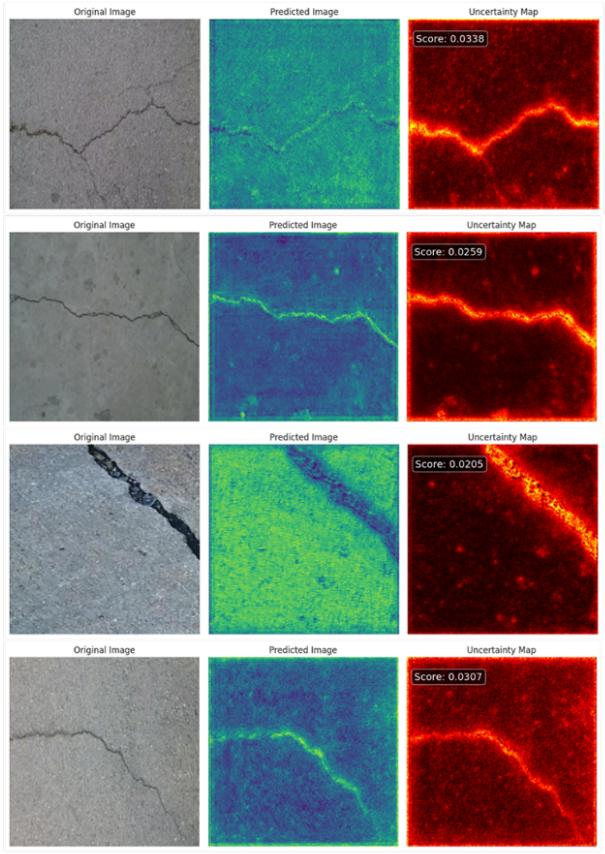


Figure 4.8: Uncertainty map visualisation of REDS-Net on crack segmentation test dataset

challenging ADAS scenarios.

Table 4.6: Quantitative evaluation of the model's uncertainty estimation

IoU Threshold	ECE	NLL	Brier Score
0.3	0.6850	1.6660	0.5162
0.4	0.6850	1.6660	0.5162
0.5	0.5635	1.4523	0.4428
0.6	0.4492	1.3182	0.3898
0.7	**0.1471**	**1.1991**	**0.2721**

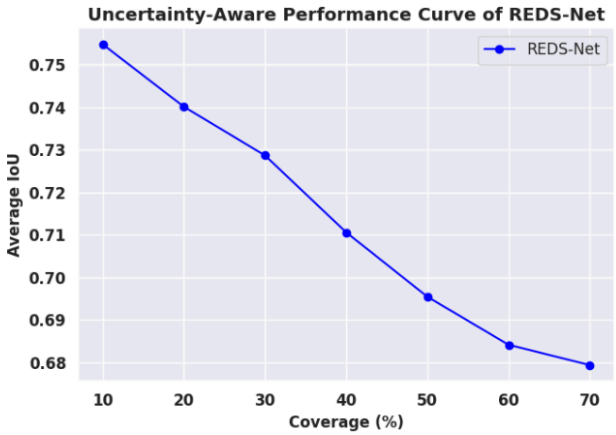


Figure 4.9: Uncertainty-Aware Performance Curve for REDS-Net

4.3.4 Generalization Performance on Unseen Road Conditions

The generalizability of the proposed model in various scenarios is evaluated in greater detail by performing a generalization test on the unseen dataset. Without training, we tested our model's generalizability on the NCD [166] dataset, which consists of 987 images for training and 385 for testing. The backgrounds of each picture are complicated and include things like zebra crossings, paints, stains, fences, waterstains, ice, snow, cars, shady spots, pavements, grass, low contrast, etc. Every image has at least one of the noises mentioned above. To see how well our model

performed on an external dataset, we tested it on the NCD test dataset. The precise visual depiction of this generalization experiment is shown in Figure 4.10.

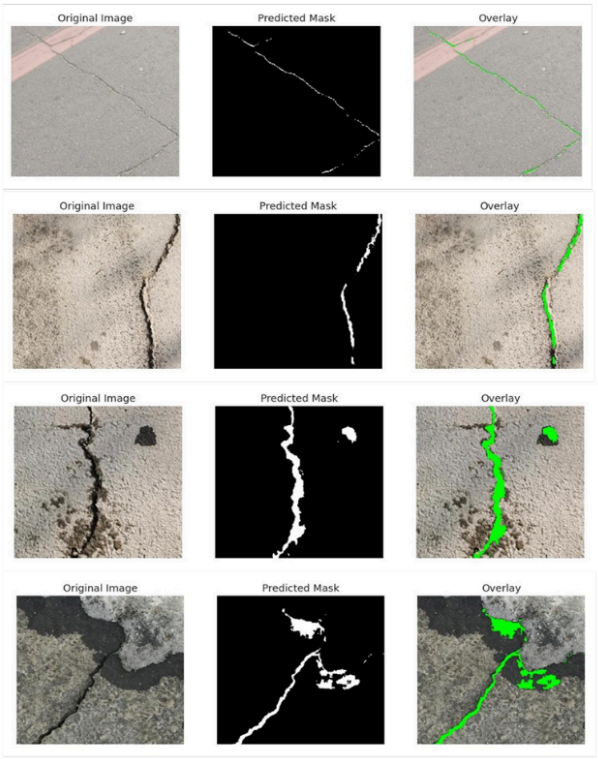


Figure 4.10: Visualization Results on External dataset

We conducted a detailed performance analysis on the unseen NCD dataset, through which the model demonstrates a precision of 61.43, a recall of 69.14, an F1-score of 61.13, an IoU of 45.10, and an overall accuracy of 98.21. The F1-score of 61.13 reflects consistent performance in both precision and recall. The IoU of 45.10 shows a reasonable overlap between the predicted crack regions and the

ground truth, which is an important metric for segmentation quality. Although the accuracy is very high, 98.21, this value is due to the large proportion of background pixels in the images. Table 4.7 shows the generalization results of our model on the unseen NCD dataset. Prior works report AIU values, by both training and testing on NCD. In contrast, REDS-Net reports IoU and Accuracy without any training on NCD, showcasing its superior generalization ability.

Table 4.7: Generalization result comparison on the NCD dataset. Metric values are reported in the remarks.

Method	**Dataset Used**	**Remarks**
ResNeSt [167]	NCD	AIU: 13.5; trained and tested on NCD
Zhao et al., 2023 [168]	NCD	AIU: 14.0; trained and tested on NCD
REDS-Net (Ours)	NCD	IoU: 45.10, Accuracy: 98.21%; evaluated without training on NCD

In summary, the proposed REDS-Net achieves accurate and reliable road crack segmentation with efficient uncertainty estimation. It outperforms existing models in both qualitative and quantitative analyses while maintaining computational efficiency. The model generalises well to unseen datasets, effectively detecting fine crack structures and providing reliable confidence estimates, making it suitable for real-world ADAS and automated inspection applications.

Chapter 5

Risk-Aware Multiclass Pavement Damage Detection for ADAS

Accurate pavement damage detection is crucial for intelligent transportation systems. However, most existing work solely focuses on object detection and classification, offering limited support for decision analysis in safety-critical scenarios. This chapter presents Distress-YOLO, a risk-aware pavement damage detection framework based on enhanced YOLOv11, explicitly designed toward ADAS-oriented severity and decision analysis. The proposed model demonstrates strong detection capability across both subtle road anomalies and severe hazards, enabling unified risk-aware interpretation under normal and disaster conditions. The proposed approach incorporates a Hybrid Local Global Attention (HLGA) module that is designed to enhance the performance of the object detection model by simultaneously capturing channel and spatial information, as well as integrating both local and global contextual information. Additionally, an adaptive loss formulation named Adaptive IoU (AIoU) is employed to dynamically adjust gradient focus based on anchor box quality, enhancing localization accuracy and robustness in complex scenes. Beyond conventional detection accuracy, the proposed framework is further assessed from a risk-aware ADAS perspective, enabling severity-driven interpretation of detected hazards. The RDD2022 dataset is used to evaluate prediction reliability under di-

verse pavement conditions using risk-coverage analysis. Furthermore, cross-dataset evaluation is performed on the Social Media Image Dataset for Object Detection of Disaster-Induced Road Damage (SODR) to examine model generalization under extreme conditions. On this disaster dataset, detected hazards are categorized into NORMAL, WARNING, and EMERGENCY operational states using a severity-driven ADAS decision logic, transforming detection outputs into actionable risk intelligence for safety-critical applications.

5.1 Overview

Road infrastructure is essential to maintaining effective and safe transportation systems [169]. However, over time, they can deteriorate due to continuous vehicular pressure, weather-induced erosion from rain and snow, and in extreme cases, natural or human-induced disasters. Pavement distresses such as cracks, potholes, blockages, and surface collapses not only degrade driving comfort but also pose serious safety risks, particularly for Advanced Driver Assistance Systems (ADAS) that rely on accurate environmental perception for decision-making.

In earlier studies, most of the object detection techniques were unable to distinguish various damage types using a unified framework reliably. Frequently, they are neither resilient nor adaptable to other road and disaster-induced conditions. Moreover, many models tend to overlook subtle and small-scale damage patterns, which are crucial for early maintenance. To address these limitations, attention mechanisms integrated into convolutional blocks have recently gained significant interest, as they enhance the model's ability to focus on informative features while suppressing irrelevant background details. These techniques enable the network to concentrate on critical damage-related features, which can improve both localization and classification accuracy.

Despite these advances, existing attention-based approaches continue to face chal-

lenges in balancing detection accuracy with real-time efficiency, especially in complex road environments. Furthermore, most studies remain centered on detection accuracy alone, without adequately addressing how detected damage information can be translated into risk-aware interpretations suitable for Advanced Driver Assistance Systems. To address these challenges, Distress YOLO, a framework based on the YOLOv11 architecture for risk-aware road damage identification, is proposed. The proposed model incorporates a Hybrid Local Global Attention (HLGA) module that efficiently models both local and global contexts. The HLGA module improves the network's ability to represent features while simultaneously extracting both channel-level and spatial information. Furthermore, for localization, we adopt the Adaptive IoU (AIoU) loss, which applies a dynamic weighting strategy that adjusts according to the quality of the anchor boxes. This enables better gradient allocation by emphasizing medium-quality samples, thereby improving bounding box regression and overall localization accuracy. Building upon these architectural enhancements, the proposed framework extends beyond conventional detection pipelines by integrating a severity-driven ADAS decision analysis. The key contributions of this study are summarized as follows:

1. We proposed a lightweight hybrid local-global attention module that can simultaneously incorporate local and global information to improve the performance of the network.

2. We incorporated the AIoU loss that adaptively focuses on informative samples based on anchor quality, addressing the limitations of static IoU thresholds in regression and improving bounding box accuracy under diverse conditions.

3. A risk-aware evaluation strategy is employed on the RDD2022 dataset using risk–coverage analysis to assess prediction reliability beyond conventional detection metrics.

4. Cross-dataset generalization is investigated by evaluating the proposed model on the SODR disaster-induced road damage dataset, demonstrating robustness under extreme and unseen road conditions.

5. A severity-driven ADAS decision analysis framework is developed to map detected damage into NORMAL, WARNING, and EMERGENCY operational states, enabling interpretable visualization of safety-critical decision behavior.

5.2 Proposed Distress-YOLO Framework

In the proposed work, an enhanced YOLOv11 framework is introduced to detect road damage under both standard and disaster-induced conditions. An adaptive IoU-based loss function is incorporated into the model to optimize bounding box regression by dynamically adjusting gradient emphasis under various damage scenarios, and the Hybrid Local-Global Attention (HLGA) module is implemented to enhance the detection of small and dispersed objects. Additionally, to ensure resilience across different domains, the suggested technique is assessed under cross-dataset generalization using a disaster-induced image dataset. Beyond conventional detection performance, the framework further supports risk-aware analysis for ADAS-oriented applications. Figure 5.1 illustrates the overall flow of the proposed methodology, which integrates YOLOv11 with the HLGA module and the adaptive loss mechanism. The model is initially trained on the RDD2022 dataset for both detection and reliability analysis to learn general road damage patterns and is subsequently fine-tuned on the disaster-induced SODR dataset for cross-dataset generalization and ADAS decision visualization. The final output demonstrates the model's ability to detect diverse damage types and to exhibit consistent risk-aware decision behaviour across varying road and disaster scenarios.

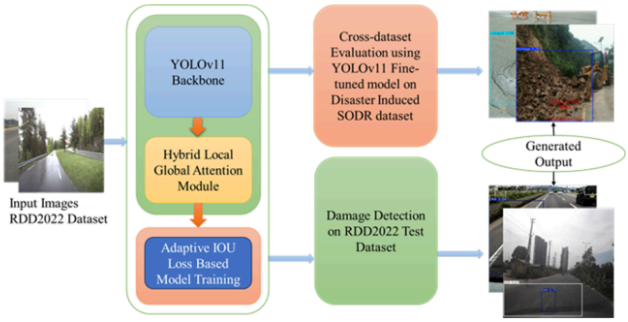


Figure 5.1: Proposed Methodology of Distress YOLO

5.2.1 Overall Architecture

The proposed Distress-YOLO framework is built upon YOLOv11 and introduces targeted architectural enhancements to support risk-aware pavement damage analysis. Figure 5.2 depicts the architecture of the enhanced YOLOv11 model. The backbone extracts multi-scale features from input images, and the proposed HLGA module refines these features by jointly modelling local spatial details and global contextual information. This enables the model to capture subtle damage patterns while maintaining scene-level awareness, which is important for severity estimation in ADAS applications. The neck aggregates refined multi-scale features, and the detection head performs unified damage classification and localization.

The HLGA module produces a more discriminative feature representation by combining local structural details with global contextual information to further improve the features. Figure 5.3 (a) depicts the C2PSA block, which enhances spatial-channel attention for effective feature fusion, while Figure 5.3(b) displays the HLGA module, which combines local and global attention to improve the accuracy of tiny damage detection. The HLGA module contains two parallel attention branches:

- **Global attention branch:** The input feature map first undergoes Global

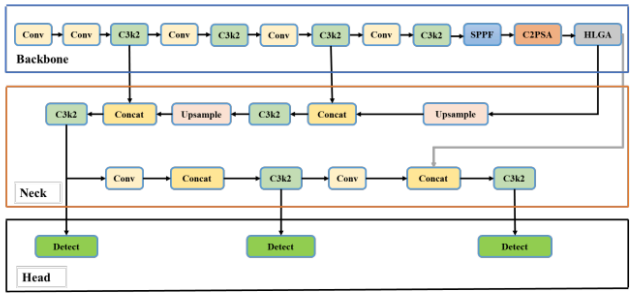


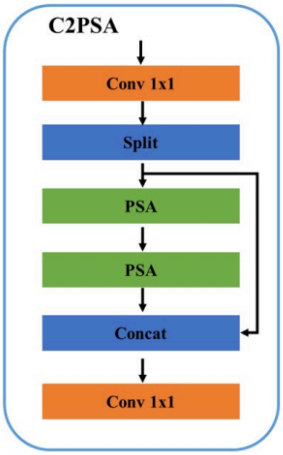
Figure 5.2: Improved YOLOv11 Architecture

Average Pooling to aggregate global contextual information across the spatial domain. The pooled features are passed through UNAP (Upsampling with Nearest Neighbor interpolation followed by Average Pooling) to match the spatial scale required for fusion. A 1D convolution is then applied along the channel dimension to model inter-channel dependencies.

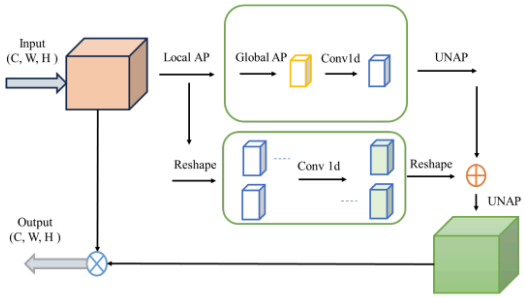
- **Local attention branch:** The input feature map is locally pooled to preserve fine-grained spatial structures. The resulting tensor is reshaped and passed through UNAP to adjust spatial resolution, followed by a 1D convolution to learn localized channel importance weights.

5.2.1.1 Proposed Hybrid Local–Global Attention Mechanism

The proposed Hybrid Local Global Attention (HLGA) module is designed to refine feature representations by combining local context information and global semantic dependencies. Unlike traditional channel attention mechanisms that rely solely on global pooling, HLGA incorporates both *local average pooling* for fine-grained spatial cues and *global average pooling* for capturing holistic image context. By dynamically combining these two sources of information, the module allows the network to focus on the most informative features across different receptive fields.



(a) C2PSA Block



(b) HLGA Module

Figure 5.3: Internal architecture of the (a) C2PSA (Cross Stage Partial with Spatial Attention) block and (b) HLGA Module

Given an input feature map $X \in \mathbb{R}^{B\times C\times H\times W}$, the HLGA module operates in the following steps.

5.2.1.1.1 Local and Global Feature Extraction: First, a local average pooling operation reduces the spatial resolution to a fixed size $L \times L$:

$$X_{\text{loc}} = \text{AAP}_{L\times L}(X) \in \mathbb{R}^{B\times C\times L\times L}. \tag{5.1}$$

Then, a global average pooling is applied to the locally pooled map:

$$X_{\text{glob}} = \text{AAP}_{1\times 1}(X_{\text{loc}}) \in \mathbb{R}^{B\times C\times 1\times 1}. \tag{5.2}$$

5.2.1.1.2 Adaptive Kernel Size Computation: Following the Efficient Channel Attention (ECA) principle, the 1D convolution kernel size k is computed adaptively from the channel dimension C:

$$t = \frac{\left|\log_2 C + b\right|}{\gamma}, \quad k = \begin{cases} t, & t \text{ odd}, \\ t+1, & \text{otherwise}, \end{cases} \tag{5.3}$$

where γ and b are constants.

5.2.1.1.3 Channel Attention Computation: Both local and global features are vectorized and passed through 1D convolutions:

$$Y_{\text{loc}} = \text{Conv1D}_k\left(\text{vec}(X_{\text{loc}})\right), \tag{5.4}$$

$$Y_{\text{glob}} = \text{Conv1D}_k\left(\text{vec}(X_{\text{glob}})\right). \tag{5.5}$$

The results are reshaped and normalized via the sigmoid activation:

$$A_{\text{loc}} = \sigma\left(\text{reshape}(Y_{\text{loc}})\right), \quad A_{\text{glob}} = \sigma\left(\text{reshape}(Y_{\text{glob}})\right). \tag{5.6}$$

5.2.1.1.4 Local–Global Fusion: The global attention map is upsampled to match the $L \times L$ local map and fused using a weighted sum:

$$A_{\text{fuse}} = (1 - \alpha) \cdot \text{Upsample}(A_{\text{glob}}, L \times L) + \alpha \cdot A_{\text{loc}}, \tag{5.7}$$

where $\alpha \in [0, 1]$ balances the contribution of each attention branch. The fused attention is adaptively pooled to (H, W), producing A_{final}.

5.2.1.1.5 Feature Refinement: Finally, the attention map is applied to the input feature map via element-wise multiplication:

$$Y = X \odot A_{\text{final}}. \tag{5.8}$$

By integrating local structural cues with global contextual information, the HLGA module provides a more discriminative feature representation, leading to improved object detection accuracy in the YOLO based architecture.

5.2.1.2 Adaptive IoU Loss for Enhanced Localization Accuracy

Furthermore, a dynamic IoU-based loss function, AIoU, is integrated into the detection head, which modifies the bounding box regression objective based on the spatial relationship between the ground-truth and predicted boxes. The proposed Adaptive IoU (AIoU) loss integrates an IoU-based regression term with an adaptive weighting strategy and a Distribution Focal Loss (DFL) component. The total bounding box regression loss can be expressed as:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{AIoU}} + \mathcal{L}_{\text{DFL}}, \tag{5.9}$$

where $\mathcal{L}_{\text{AIoU}}$ is the adaptive IoU loss term and $\mathcal{L}_{\text{DFL}}$ is the localization refinement term based on distance distributions.

5.2.1.2.1 Adaptive IoU Term: Given the predicted bounding boxes B_p and target bounding boxes B_t for foreground samples indicated by the mask M_{fg}, the IoU computation with the AIoU metric is:

$$\text{AIoU}(B_p, B_t) = R_{\text{IoU}} \cdot L_{\text{IoU}}, \tag{5.10}$$

where R_{IoU} is a relative weight term derived from geometric factors (e.g., center distance), and L_{IoU} is the base localization IoU value.

Let w_i be the classification confidence weight for the i-th positive sample:

$$w_i = \sum_c s_{ic}, \quad i \in M_{\text{fg}}, \tag{5.11}$$

where s_{ic} denotes the classification score for class c.

The AIoU loss term is then:

$$\mathcal{L}_{\text{AIoU}} = \frac{1}{S_{\text{pos}}} \sum_{i \in M_{\text{fg}}} \left[(1 - \text{AIoU}_i) \cdot w_i \right], \tag{5.12}$$

with $S_{\text{pos}} = \sum_{i \in M_{\text{fg}}} w_i$ being the total positive sample score sum.

5.2.1.2.2 Distribution Focal Loss Term: To improve bounding box regression accuracy, the predicted distance distribution D_p is compared with the target distance distribution D_t computed from anchor points and target boxes:

$$D_t = \text{bbox2dist}(A, B_t, r_{\max} - 1), \tag{5.13}$$

where A denotes the anchor points and $r_{\max}$ is the maximum discrete distance range.

The DFL term is:

$$\mathcal{L}_{\text{DFL}} = \frac{1}{S_{\text{pos}}} \sum_{i \in M_{\text{fg}}} \text{DFL}(D_{p,i}, D_{t,i}) \cdot w_i. \tag{5.14}$$

5.2.1.2.3 Final Form: Combining both terms, the final adaptive bounding box regression loss is:

$$\mathcal{L}_{\text{total}} = \frac{1}{S_{\text{pos}}} \sum_{i \in M_{\text{fg}}} \left[(1 - \text{AIoU}_i) \cdot w_i + \text{DFL}(D_{p,i}, D_{t,i}) \cdot w_i \right] \tag{5.15}$$

This formulation dynamically adjusts the penalty for localization errors based on box quality and class confidence, while also refining predictions using a distribution-aware regression strategy.

5.3 Experimental Evaluation

This section presents the experimental evaluation process to validate the effectiveness of the proposed Distress - YOLO framework under both normal pavement conditions and disaster-induced road scenarios.

5.3.1 Datasets

RDD 2022: We have used the RDD2022 dataset from kaggle [170], which comprises 47,420 road images gathered from six nations: China, Japan, India, the Czech Republic, Norway, and the United States. The dataset includes instances of road damage across 7 classes: D00 (Longitudinal Crack), D10 (Transverse Crack), D20 (Alligator Crack), D40 (Pothole), D43 (White Line Blur), D44 (Cross Walk Blur), and D50 (Manhole Cover). These classes cover different types of cracks and potholes, making the dataset suitable for training and evaluating road damage detection models [170].

SODR : SODR dataset was first presented by Wang et al. [96]. Images from social media and Google image searches, using terms such as collapse, rockfall, rockslide, and mudflow, along with roads, were used to build this dataset. A consistent resolution of 640 x 640 pixels was applied to the collected images, and 1552 images were annotated in two categories, blockage and collapse, using an object detection

format.

5.3.2 Evaluation Metrics

Several evaluation metrics, including Recall (R), Precision (P), Average Precision (AP), and Mean Average Precision (mAP), are employed in this research to assess the model's performance. They are computed in the manner described below.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5.16}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5.17}$$

where TP, FP, and FN indicate the number of true positives, false positives, and false negatives, respectively.

$$\text{AP}_i = \int_0^1 p_i(r)\, dr \tag{5.18}$$

where $p_i(r)$ is the precision as a function of recall for class i.

mAP is the mean of the AP values across all K classes:

$$\text{mAP} = \frac{1}{K} \sum_{i=1}^{K} \text{AP}_i \tag{5.19}$$

We report mAP@0.5 (IoU threshold 0.5) and mAP@0.5 : 0.95 (mean over IoU thresholds from 0.5 to 0.95 with a step of 0.05).

5.3.3 ADAS-Specific Evaluation

ADAS-oriented evaluation is conducted by assessing the prediction reliability on RDD2022 dataset and severity driven analysis on disaster-induced SODR dataset. Unlike standard detection metrics, this evaluation emphasizes severity estimation

and emergency decision-making.

5.3.3.1 Risk–Coverage Analysis (RDD-2022)

Let an image have size $H \times W$ with ground truth boxes $\mathrm{GT}_i = [x_1, y_1, x_2, y_2]_i$. Predicted boxes B_p are correct if

$$c_p = \begin{cases} 1 & \max_i \mathrm{IoU}(B_p, \mathrm{GT}_i) \geq \tau_{\mathrm{IoU}}, \\ 0 & \text{otherwise.} \end{cases} \tag{5.20}$$

The normalized area is $A_{\mathrm{norm}}(B_p) = \frac{(x_2 - x_1)(y_2 - y_1)}{H \cdot W}$. Sorting predictions by confidence s_p, the Risk–Coverage curve is:

$$\mathrm{Coverage}(k) = \frac{k}{N}, \quad \mathrm{Risk}(k) = 1 - \frac{1}{k} \sum_{j=1}^{k} c_j. \tag{5.21}$$

This analysis evaluates the reliability of high-confidence detections and is particularly relevant for safety-critical ADAS systems.

5.3.3.2 Disaster Dataset ADAS Logic

For each disaster box with confidence c, normalized area A_{norm}, and class weight w_{cls}:

$$\mathrm{Severity} = \min(1.0, c \cdot A_{\mathrm{norm}} \cdot w_{\mathrm{cls}}), \tag{5.22}$$

with $w_{\mathrm{cls}} = 1.4$ for blockage, 1.8 for collapse . The ADAS decision is:

$$\mathrm{Decision} = \begin{cases} \mathrm{EMERGENCY} & \text{if Severity} \geq 0.35, \\ \mathrm{WARNING} & 0.20 \leq \text{Severity} < 0.35, \\ \mathrm{NORMAL} & \text{if Severity} < 0.20. \end{cases} \tag{5.23}$$

Boxes are visualized with color-coded labels: green for *Normal*, orange for *Warning*, and red for *Emergency*, including class and severity score. This evaluation demonstrates that the model not only detects damage but also estimates severity, enabling

actionable ADAS alerts.

In order to develop a varied representation of pavement distress patterns, the model was first trained on the RDD2022 dataset [170]. Its generalization ability under different road and weather circumstances was then improved by fine-tuning it on the SODR dataset [96]. Every experiment was carried out in Google Colab on an NVIDIA Tesla T4 GPU with 16GB of RAM, and the PyTorch framework for 150 epochs. The training phase employed stochastic gradient descent (SGD) with a momentum of 0.937, a weight decay of 5×10^{-4}, and an initial learning rate of 0.01. Input images were resized to 512×512 pixels, and to increase robustness, common data augmentation methods were used, such as mosaic augmentation, random scaling, and horizontal flipping. A post-detection ADAS logic was used for disaster images in order to calculate the extent of damage and produce emergency-level notifications based on the type of damage and detection confidence.

5.4 Results and Discussion

This section evaluates the proposed Distress-YOLO framework from both detection accuracy and risk-aware ADAS decision-making perspectives. The analysis is conducted in three stages: (i) quantitative detection performance on the RDD2022 dataset, (ii) cross-dataset generalization on the disaster-induced SODR dataset, and (iii) qualitative and decision-level analysis using a severity-driven ADAS logic.

5.4.1 Quantitative Performance on RDD2022

In order to evaluate the efficacy of the proposed changes, the RDD2022 dataset was used to train and evaluate four different configurations of the YOLOv11 framework: the baseline YOLOv11 model, YOLOv11 with AIoU, YOLOv11 with HLGA Module, and the proposed Distress-YOLO, which combines both AIoU and HLGA module. Table 5.1 provides an overview of the performance comparison. The mAP@0.5

Table 5.1: Model performance comparison with baseline

Model Variant	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Baseline YOLOv11	0.541	0.480	0.481	0.245
YOLOv11 + AdaIoU	0.547	0.501	0.503	0.259
YOLOv11 + HLGA	0.552	0.499	0.515	0.262
YOLOv11 + HLGA + AdaIoU	**0.652**	**0.513**	**0.553**	**0.292**

Table 5.2: Per-class mAP@0.5 comparison

Model Variant	D00	D10	D20	D40	D44	D50	D43
Baseline YOLOv11	0.410	0.355	0.527	0.272	0.556	0.588	0.660
YOLOv11 + AdaIoU	0.434	0.398	0.548	0.272	0.578	0.621	0.672
YOLOv11 + HLGA	0.471	0.406	0.551	0.288	0.579	0.652	0.659
YOLOv11 + HLGA + AdaIoU	**0.484**	**0.408**	**0.583**	**0.367**	**0.604**	**0.717**	**0.708**

for the baseline YOLOv11 model is 0.481. Both mAP@0.5 and mAP@0.5:0.95 are improved by incorporating HLGA and AIoU, demonstrating their effectiveness in improving feature representation. The suggested changes perform better than the baseline, demonstrating the resilience of our design.

Class-wise detection performance is shown in Table 5.2. The findings demonstrate that our model consistently improves across the majority of damage categories compared to the baseline.The suggested model demonstrates significant gains for D40 (potholes) and D50 (manhole covers), demonstrating its ability to identify high-risk objects as well as irregular damages that are essential for road safety.

The accuracy–speed trade-off is further demonstrated in Figure 5.4, where different colours and circle sizes are used to represent each model. It is clear that while the suggested approach has a marginally lower frame rate than lightweight alternatives, it delivers noticeably better accuracy, which makes it more appropriate for use in ADAS and actual road monitoring situations.

The Precision–Recall (PR) curve shows how recall and precision vary across different confidence thresholds. Alongside it, the legend provides a per-class evaluation as well as the overall mean Average Precision (mAP). The point with the maximum F1-

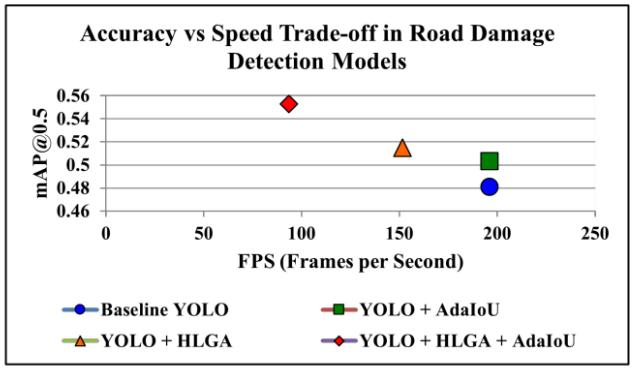


Figure 5.4: Accuracy–speed trade-off of different YOLO variants for road damage detection

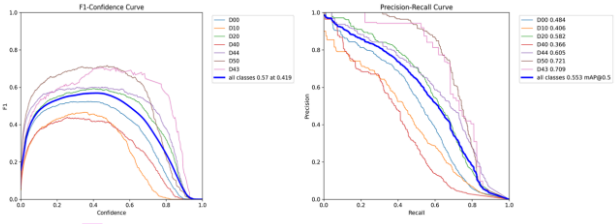


Figure 5.5: Precision-Recall (PR) and F1-score curves for the proposed model

score highlights the threshold that best balances precision and recall. Both the PR curve and the corresponding F1-score curve, which reflects this harmonic balance, are jointly illustrated in Figure 5.5.

A breakdown of detection results by class is provided by the confusion matrix, which displays true positives, false positives, and false negatives for each class. A clear picture of the reliability of per-class categorization is provided by the confusion matrix in Figure 5.6.

To demonstrate how well the proposed Distress-YOLO model detects damage, it is evaluated against the most advanced detectors. A quantitative comparison demon-

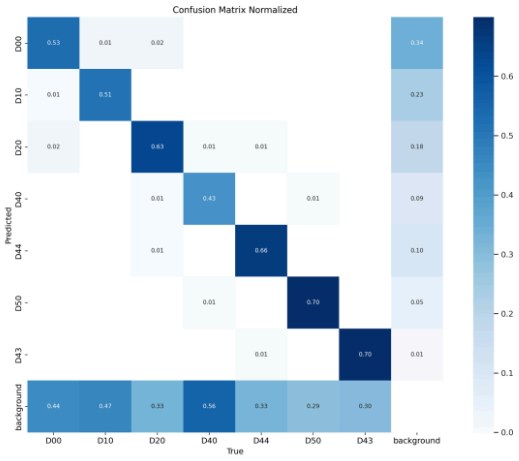


Figure 5.6: Confusion matrix for the proposed Distress YOLO model

strates that our approach outperforms current models in terms of resilience and detection accuracy, achieving competitive precision, recall, and mAP values. According to the comparison results in Table 5.3, the suggested Distress-YOLO performs significantly better in terms of F1-score and mAP@0.5 than traditional detectors like Faster R-CNN, RetinaNet, and CenterNet. Our model outperforms YOLOv5, YOLOv7, and YOLOv8 in terms of F1-score (57.41%) and total mAP@0.5 (55.3%), demonstrating its improved capacity to strike a balance between accurate localization and reliable detection across a range of road damages.

5.4.2 Risk-Aware Evaluation on RDD2022

In addition to traditional detection accuracy, we use a Risk–Coverage (RC) curve to assess model confidence reliability. Risk is defined as the percentage of inaccurate detections among retained predictions, and predictions are sorted according to confidence. The risk-coverage curve in Figure 5.7 evaluates the reliability of confidence

Table 5.3: Comparison with other object detectors

Model	mAP@0.5 (%)	F1-Score (%)
Faster R-CNN [171]	38.60	48.70
RetinaNet [172]	43.39	44.26
CenterNet [172]	43.83	48.33
YOLOv5 [173]	53.60	56.60
YOLOv7 [84]	53.74	–
YOLOv8 [174]	53.89	–
Distress-YOLO	**55.30**	**57.41**

estimates produced by our model on the RDD2022 dataset. The risk is negligible at low coverage, demonstrating the great dependability of high-confidence predictions. As coverage increases, lower confidence and more ambiguous detections are gradually included, leading to a natural increase in risk. The smooth and monotonic trend confirms that the confidence scores are well aligned with prediction difficulty, which is desirable for confidence-aware ADAS applications.

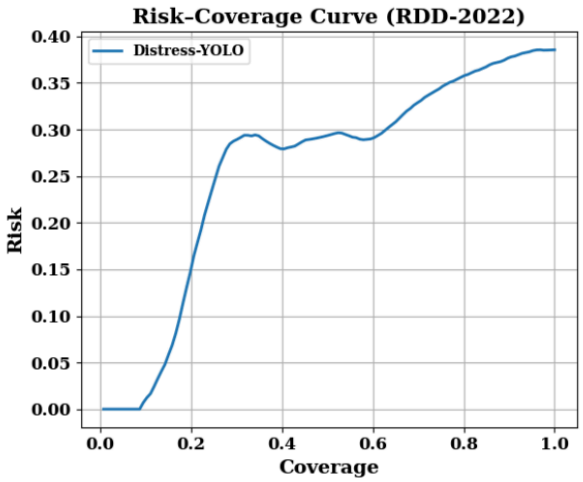


Figure 5.7: Risk–Coverage curve on the RDD-2022 dataset

Table 5.4: Distress-YOLO performance on SODR dataset (fine-tuned for 50 epochs)

Model	Class	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Distress-YOLO	**All**	**0.773**	**0.714**	**0.778**	**0.480**
	Blockage	0.699	0.606	0.687	0.376
	Collapse	0.848	0.822	0.869	0.585

5.4.3 Cross-Dataset Generalization on SODR

The proposed Distress-YOLO model was fine-tuned on the SODR dataset for 50 epochs to assess its generalization ability under various road conditions, basically on disaster-induced images. Unlike RDD2022, not all the YOLOv11 variants are trained in this case; only the final proposed model is fine-tuned and evaluated for generalization on the SODR dataset. Using the same hyperparameter settings as the primary experiment, the fine-tuning approach involved applying standard data augmentation and resizing the input images to 512×512 pixels. The fine-tuned Distress-YOLO model exhibits good detection performance, with an overall mAP@0.5 of 0.778 and mAP@0.5:0.95 of 0.48, after 50 epochs of training on the SODR dataset. The model strikes a balance between precise predictions and thorough coverage of object instances, achieving an overall precision of 0.773 and a recall of 0.714.

Both groups exhibit strong performance, according to a class-wise analysis. The model achieves a precision of 0.699, recall of 0.606, and mAP@0.5 of 0.687 for the Blockage class. The model performs even better for the Collapse class, with a precision of 0.848, recall of 0.822, and mAP@0.5 of 0.869. The overall performance of the Distress YOLO model on the SODR test set is compiled in Table 5.4.

The trade-off between precision and recall over confidence thresholds for every damage class is depicted by the Precision–Recall (PR) curve in Figure 5.8. Stronger performance is shown by a steeper curve towards the top-right, where greater recall is attained with little loss of precision. The F1 score curve provides a balanced measure

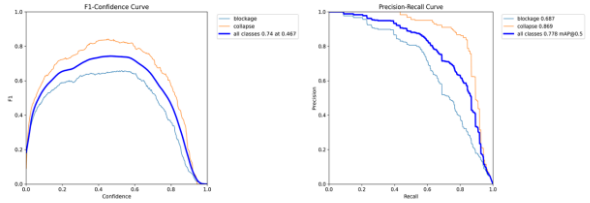


Figure 5.8: Precision-Recall and F1-score curves for Distress-YOLO on SODR

of accuracy by combining precision and recall into a single harmonic mean.

5.4.4 ADAS Severity Decision Analysis on Disaster Dataset

The ADAS logic described in Section Evaluation Metrics was applied to disaster-induced road damage images. Each detected instance was classified into NORMAL, WARNING, or EMERGENCY based on its computed severity. The resulting distribution of ADAS decisions is summarized in Table 5.5.

Table 5.5: ADAS severity decision distribution on disaster dataset

ADAS Decision	**Count**	**Percentage (%)**
NORMAL	192	47.5
WARNING	75	18.5
EMERGENCY	137	34.0
Total	**404**	**100**

5.4.5 Qualitative Analysis and Visualization

The qualitative visualization results are compared across YOLOv11 variants: baseline YOLOv11, YOLOv11 with AdaIoU, YOLOv11 with HLGA, and the final Distress-YOLO, which combines both modules. The final model produced clearer, more comprehensive detections and identified subtle cases not present in the ground truth, highlighting its resilience compared to the baseline, which missed multiple damages. The visualization results are shown in Figure 5.9.

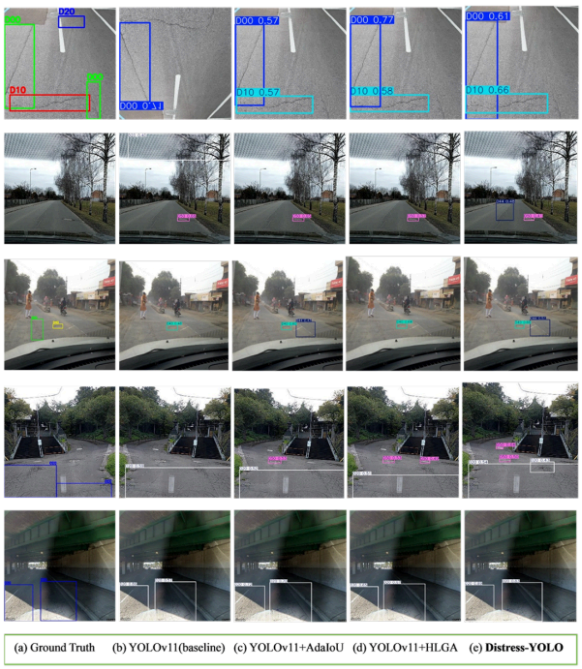


Figure 5.9: The visualization results on the RDD2022: (a) Ground Truth, (b) YOLOv11 (baseline), (c) YOLOv11+AdaIoU (d) YOLOv11+HLGA (e) Distress-YOLO

To further show the efficacy of the suggested approach, we offer qualitative visualizations on the SODR dataset, which includes the two primary damage categories of Collapse and Blockage. Comparisons between the original images and the predicted images produced by our improved model are displayed in Figure 5.10. Importantly, although the model was originally developed and trained on the RDD2022 dataset, its strong performance on SODR images highlights its cross-dataset generalization ability.



Figure 5.10: Visualization Results on SODR dataset: a) original images b) Proposed Distress-YOLO

The qualitative behaviour of the proposed ADAS severity logic on disaster-induced road images (SODR dataset) is shown in Figure 5.11. The original image and the corresponding ADAS overlay, which classifies detected risks into NORMAL, WARNING, and EMERGENCY categories, are displayed for each sample. The severity-aware prioritizing of dangers is highlighted by the colored bounding boxes, and the inset lists the number of ADAS decisions made for each image. These visual results show how the suggested ADAS logic efficiently converts detection outputs into actionable severity levels, allowing for quick risk assessment under disaster conditions.



Figure 5.11: Qualitative visualization of ADAS severity logic on SODR dataset: a) original images b) ADAS decision overlays

5.5 Summary

This work presented Distress-YOLO, a risk-aware pavement damage detection framework designed to bridge the gap between conventional object detection and ADAS-oriented decision-making. By enhancing the YOLOv11 architecture with a Hybrid Local–Global Attention (HLGA) module and an Adaptive IoU (AIoU) loss function, the proposed approach achieves improved feature representation and localization accuracy under diverse road conditions. Extensive experiments on the RDD2022 dataset demonstrate the effectiveness of the proposed architectural enhancements for accurate and reliable pavement damage detection. Furthermore, cross-dataset evaluation on the disaster-induced SODR dataset confirms the robustness and generalization capability of Distress-YOLO under extreme and previously unseen conditions.

Beyond detection performance, this study introduces a severity-driven ADAS decision analysis that maps detected hazards into interpretable operational states, namely NORMAL, WARNING, and EMERGENCY. Qualitative visualizations on

disaster scenarios illustrate how the proposed framework provides intuitive and actionable safety cues, enabling meaningful interpretation of road hazards in safety-critical environments.

Chapter 6

Unpaved Road Segmentation for Advanced Driver Assistance System

The visual assessment of damage on unpaved roads, which lack the structured features of paved roads, is one of the primary issues facing the Advanced Driver Assistance System (ADAS). Most existing solutions focus on well-maintained roads with high-quality data and neglect the vehicles operating on unpaved or unstructured roads. This chapter addresses the research gap related to road damage detection on unpaved roads for the advanced driver assistance system. Unpaved roads are characterized by irregular texture, low contrast between damaged and undamaged area, inconsistent lighting conditions and absence of well-defined regions of damage pattern, making the segmentation task more challenging. Most prior works focus on detecting specific damage types on structured and paved roads, leaving unpaved roads' conditions unexplored. This chapter aims to bridge the gap by proposing a deep learning based segmentation framework designed for unpaved road damage detection.

6.1 Overview

In the context of Advanced Driver Assistance Systems (ADAS), segmenting damages on unpaved roads is crucial since it has a direct effect on driving efficiency and vehicle

safety. It is more difficult to detect and segment damage on unpaved roads due to their uneven surfaces, varied textures, and fluctuating environmental conditions. Accurate segmentation of potholes, road cracks, and erosions is necessary for adaptive driving reactions, real-time notifications, and navigational adjustments on unpaved roads. Traditional manual inspection methods are labor-intensive and prone to human error [175]. In recent years, automatic damage detection on unpaved roads has received far too much attention. With the introduction of autonomous driving and ADAS, it is now essential to promptly assess road conditions in order to guarantee the safety of both vehicles and other road users. By combining powerful machine learning and deep learning models designed specifically for uneven road conditions, ADAS can enhance overall driving comfort, save vehicle wear and tear, and promote road safety in rural and underdeveloped areas. Automated damage detection systems, leveraging advanced technologies such as deep learning and computer vision, offer a practical alternative to manual methods. These technologies are particularly pertinent to Advanced Driver Assistance technologies (ADAS), which rely on precise and up-to-date road condition information to make crucial decisions like modifying vehicle behavior or alerting drivers to possible dangers [176, 177]. Even though road damage identification has advanced significantly, the majority of research has concentrated on paved roads because of their predictable surface conditions and the availability of high-quality data [178, 179]. Therefore, there is still a large research gap in the creation of reliable models for unpaved roads, which frequently have imperfections, different surface textures, and intricate environmental conditions. [180].

Recent developments in deep learning have shown promising results in addressing these issues, especially in semantic segmentation. Although models like U-Net and its variations have demonstrated cutting-edge performance in identifying road imperfections, such as cracks and potholes, their use on unpaved roads continues to be challenging because of the intricacy and unpredictability of these surfaces

[181, 99]. Furthermore, many current approaches ignore the larger context of unpaved road conditions in favor of concentrating on specific damage types [98].

To overcome these constraints, this study suggests a strong deep learning-based method for identifying surface deterioration on unpaved roads. The proposed approach uses a U-Net architecture with a ResNet-18 backbone to recognize a variety of surface defects, such as erosion, cracks, and ruts, by utilizing semantic segmentation capabilities. The RTK dataset is used to refine this architecture, which yields highly annotated images that are specific to unpaved road conditions [177]. The model provides a workable approach for incorporation into ADAS applications by taking into consideration the particular difficulties presented by unpaved road surfaces, resulting in high accuracy and generalization.

6.2 Methodology for Unpaved Road Damage Segmentation

The reliable detection of road damages, particularly on unpaved roads, is a critical task for enhancing Advanced Driver Assistance Systems (ADAS). Unpaved roads pose unique challenges such as varying textures, irregular lighting, and a wide range of surface damage types, including cracks and potholes. To address these challenges, we evaluated three state-of-the-art semantic segmentation architectures: U-Net, LinkNet, and Feature Pyramid Network (FPN), each with ResNet-18, InceptionV3, and VGG16 backbones. Based on the experimental results, U-Net with ResNet-18 was identified as the most effective model for the task. The U-Net architecture was selected as the backbone of our proposed model due to its ability to effectively capture both global context and fine spatial details through its encoder-decoder structure with skip connections. By integrating ResNet-18 as the encoder, the network benefits from residual connections, which alleviate the vanishing gradient problem and enable deeper feature extraction [66, 182]. The decoder complements the encoder by progressively upsampling the feature maps while concatenating

features from corresponding encoder layers. This combination allows the network to maintain spatial resolution while learning complex patterns necessary for accurate segmentation of road damages, such as cracks and potholes. Figure 6.1 illustrates the architecture of the suggested model, which combines ResNet-18 as the feature extractor in the encoder path with U-Net's encoder–decoder structure. ResNet-18's residual connections reduce vanishing gradients and improve gradient propagation, allowing for deeper feature learning.

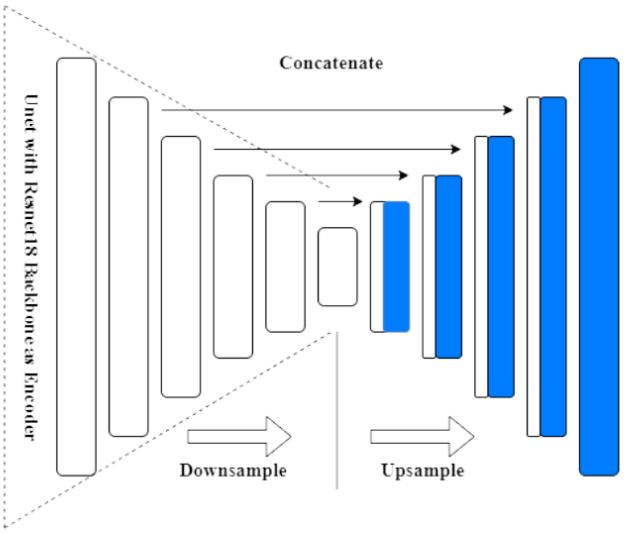


Figure 6.1: Overall architecture of our proposed network

Figure 6.2 provides an overall overview of the technique of the suggested approach, demonstrating the workflow from backbone selection, model training, and image preprocessing to evaluation. The input RGB images with corresponding ground truth images are first preprocessed according to the backbone. After preprocessing, extensive experiments have been done to finalise the best-suited backbone for the unpaved road segmentation, and then the model is evaluated based on segmentation

metrics: IoU and F1 score.

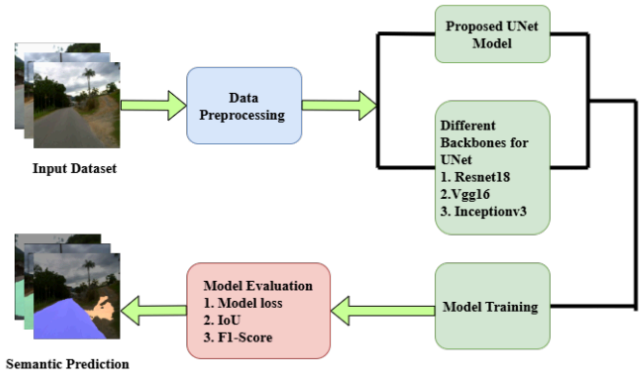


Figure 6.2: Proposed Methodology

6.3 Experimental work and Results

The experiments are conducted using the Road Transverse Knowledge (RTK) dataset [183], which comprises 701 images with ground truth images. 561 images are used for training, while 140 images are used for validation. The dataset includes images of unpaved roads taken in a range of environmental circumstances and capturing different kinds of degradation. We scaled the original images to 256 x 256 resolution for training purposes, even though they had a resolution of 352 x 288. The models are trained using the Adam optimizer with a learning rate of 1×10^{-3}. The proposed architecture is implemented using the Keras framework. Training and evaluation are conducted on Google Colab, utilizing a T4 GPU accelerator with 12 GB of RAM and 12 GB of GDDR5 VRAM.

The performance of the proposed models is assessed through quantitative analysis using standard evaluation metrics such as Intersection over Union (IoU) and F1 Score. The experiments are performed using three architectures—U-Net, LinkNet,

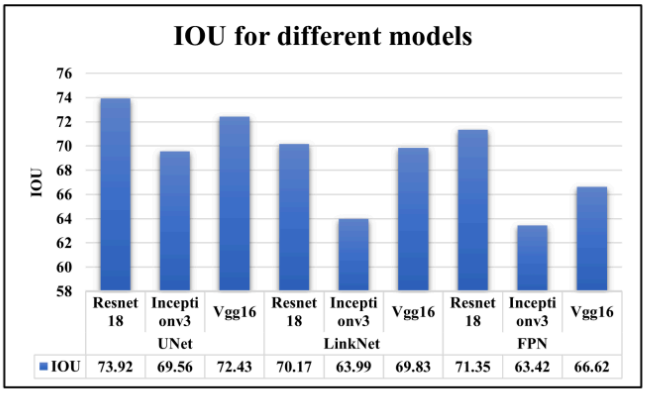


Figure 6.3: IoU for different model architectures

and Feature Pyramid Network (FPN)—each tested with ResNet-18, InceptionV3, and VGG16 as backbones. Figure 6.3 presents the IOU of different architectures, and Figure 6.4 presents the F1 scores for different network architectures on the RTK test dataset. The proposed approach leverages the strengths of U-Net with ResNet-18 to address the specific challenges of unpaved road damage detection, providing a reliable solution for ADAS.

Our experimental results demonstrated that the U-Net with ResNet-18 outperformed the other architectures in terms of IOU and F1 Score. The lightweight nature of ResNet-18 ensures that the model is computationally efficient, making it well-suited for real-time deployment in ADAS. This enables robust damage detection on unpaved roads, contributing to safer navigation in challenging environments. Table 6.1 represents the performance metrics comparison of different segmentation models based on IOU and F1 Score. Bold values indicate the best performance in respective metrics.

These findings are further supported by qualitative comparisons, as seen in Figure 6.5, which demonstrate that U-Net ResNet-18 generates better segmentation masks

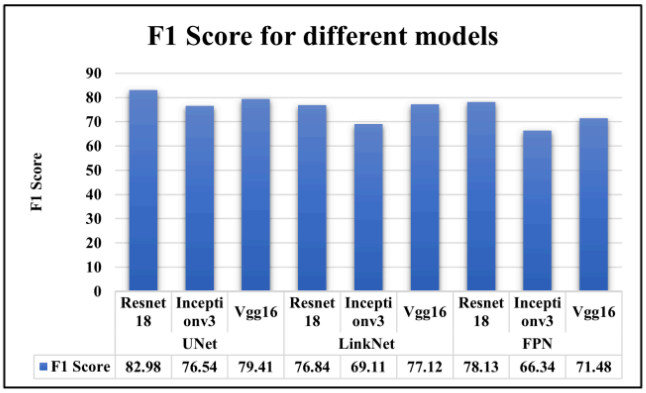


Figure 6.4: F1 score for different model architectures

Table 6.1: Comparison of segmentation models based on IoU and F1 score. Bold values indicate the best performance in the respective metrics.

Model	IoU (%)	F1 Score (%)
PAM-UNet [184]	65.81	79.41
DCWin-Attention [185]	67.72	–
Improved BiSeNet [186]	72.63	**83.21**
FPN with ResNet18	71.35	78.13
LinkNet with ResNet18	70.17	76.84
UNet with ResNet18	**73.92**	82.98

that successfully define the edges of potholes and cracks, even on dusty and uneven road surfaces where other architectures were unable to maintain consistent contours.

6.4 Summary

This chapter investigated the problem of unpaved road damage segmentation using semantic segmentation techniques. The U-Net with a ResNet-18 backbone model achieved superior performance in detecting damage on unpaved roads and outperforms LinkNet and FPN. The results highlight that the model is efficient enough to segment unpaved roads containing dirt or mud, which is helpful for ADAS driving

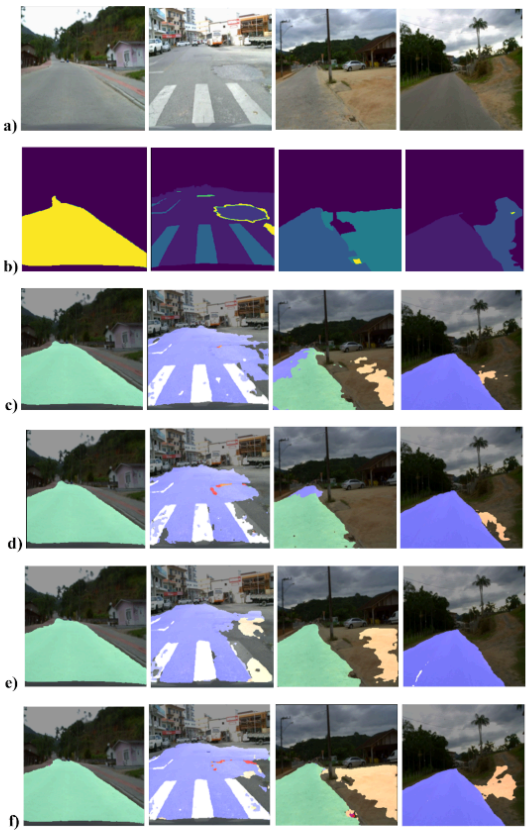


Figure 6.5: Sample qualitative demonstrations for the networks on the RTK test dataset. a) RGB Images b)Ground Truth images c) FPN d)LinkNet e)UNet f)UNet-Resnet18

on underdeveloped roads.

Chapter 7

Conclusions, Future Scope and Social Impact

This chapter provides an overview of the main research conclusions drawn from the theoretical and experimental contributions of the study. It also outlines potential directions for future research and discusses the societal and technical implications of the proposed work.

7.1 Conclusions

This thesis presents several deep learning-based approaches to address the problems of segmenting and detecting road damage in the context of autonomous driving and ADAS. The work started with the creation of a reliable framework for pothole detection that could handle a variety of real-world situations, such as complex road textures, occlusions, and shifting lighting. A better segmentation model, ERCU-Net, was introduced to achieve fine-grained localization. It performed better than normal U-Net variations in terms of segmentation accuracy and boundary preservation. Grad-CAM visualization, which highlighted the areas that affected the model's predictions, was added to the detection system to increase its transparency. Extensive tests under various lighting circumstances, shadows, background complexity, and surface textures showed that the suggested framework remains resilient in a variety of driving scenarios. The ADAS-ready systems that can operate reliably in

real-world situations are supported by this contribution.

Building on this, a generalized framework for road crack segmentation was suggested that incorporates uncertainty estimation. The generalization performance of the proposed model was assessed on several datasets, including unseen datasets. The findings showed that the framework performs well even in complex scenarios and with varying datasets, in addition to achieving accurate segmentation accuracy. Predictive uncertainty estimation improves system credibility and gives more information for lower-level decision-making modules. A risk-aware multiclass damage detection system was implemented to bridge the gap between perception and decision-making. This study did not limit the system to classification outputs; rather, detection results were transformed into operational driving states. With this integration, application-oriented system design is replacing purely analytical research. The validation on disaster-induced scenarios further highlighted the adaptability of the framework under extreme and uncommon road conditions.The extension of road damage assessments to unpaved and dirt road situations is another noteworthy result of this study. There is a lack of research in rural and unstructured contexts due to the majority of current studies concentrating on structured urban highways. By tackling these issues, this thesis expands the field of intelligent road monitoring and shows how adaptable the suggested approaches are to various types of infrastructure.

The study concluded by examining the relatively unexplored area of unpaved road segmentation and showed promising results in rural and off-road locations, improving precision, resilience, and generalizability. These contributions to this thesis collectively cover not only the improvement of detection models but also the creation of a comprehensive and practical framework that integrates robustness, uncertainty awareness, interpretability, and risk-based decision logic.

7.2 Future Scope

Even though the proposed framework performs well in a range of road conditions, there are still several interesting directions for future study:

1. Future studies can optimize the suggested models for real-time deployment in Advanced Driver Assistance Systems (ADAS). Model compression and lightweight architectures can boost computing efficiency for embedded and edge devices.

2. Adding more sensing modalities like LiDAR, RADAR, or depth information can increase the detection's resilience in challenging environmental conditions like low light, rain, fog, and dust.

3. Including damage severity grading and predictive maintenance modeling in the framework may aid in proactive road management strategies and infrastructure priorities.

4. The development of large datasets that cover geographically disparate locations, unfavorable climates, and scenarios of damage from disasters would enhance model generalization and real-world applicability.

7.3 Social Impact

The suggested study has important societal implications for transportation safety, infrastructure resilience, and sustainable mobility:

1. Automated pothole detection, crack segmentation, and multiclass damage identification enable early identification of hazardous road conditions, reducing the likelihood of accidents and improving overall driving safety.

2. By enhancing confidence estimates in safety-critical perception systems, uncertainty-aware detection algorithms enable the safe deployment of ADAS and autonomous driving technologies.

3. In disaster-related scenarios, multiclass damage detection can assist authorities and emergency response teams in prioritizing repairs and successfully restoring transportation networks.

4. In addition to organized urban roads, unpaved road segmentation and damage assessment enhance the sense of terrain in off-road and rural settings, allowing Advanced Driver Assistance Systems (ADAS) to operate with greater safety. Increased surface condition knowledge supports improved trajectory planning, driver assistance, and vehicle stability in scenarios without conventional lane markers.

In summary, this thesis enhances intelligent road surface monitoring by moving beyond simple detection models to more reliable and efficient systems that can be applied in the real world. It supports the development of safer and more responsive transportation solutions by offering a strong foundation for future studies in risk-aware and uncertainty-informed road analysis.

Research Publications

SCI/SCIE Journals (Accepted and published):

1. R. Tripathi, S. Indu, and R. Kumar, "ERCU-Net: Segmentation of road potholes using enhanced residual convolutional block based on U-Net for ADAS," *Signal, Image and Video Processing* , vol. 18, suppl. 1, pp. 385–394, 2024. (Publisher: Springer, Impact Factor: 2.1).

2. R. Tripathi, S. Indu, and R. Kumar, "REDS-Net: Uncertainty aware road crack segmentation based on residual enhanced dilated segmentation network for ADAS," *Nondestructive Testing and Evaluation*, pp. 1–23, 2025.(Publisher: Taylor & Francis, Impact Factor 4.2)

IEEE International Conferences

1. R. Tripathi, S. Indu, and R. Kumar, "Detection of road potholes using deep learning based improved YOLOv4 network," *IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)*, Delhi, India, 2022, pp. 1–6.

2. R. Tripathi, S. Indu, and R. Kumar,"Unpaved Road Damage Detection via Semantic Segmentation: Enabling Smarter ADAS," *IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Gwalior, India, 2025, pp. 1-5.

Extended Work: SCI/SCIE Journals (Communicated/under review)

- “Distress-YOLO: Risk-Aware Pavement Damage Detection and ADAS-Oriented Severity Analysis for Heterogeneous Road Hazards”, Road Materials and Pavement Design (2026). (Under Review) (Publisher: Taylor & Francis, SCIE Indexing, Impact Factor 3.0)

Bibliography

[1] L. R. Abdulrazzaq, M. N. Abdulkareem, M. R. M. Yazid, M. N. Borhan, and M. S. Mahdi, "Traffic congestion: Shift from private car to public transportation," *Civil Engineering Journal*, vol. 6, no. 8, pp. 1547–1554, 2020.

[2] O. Alfarraj, "Internet of things with bio-inspired co-evolutionary deep-convolution neural-network approach for detecting road cracks in smart transportation," *Neural Computing and Applications*, vol. 37, no. 35, pp. 29 061–29 076, 2025.

[3] R. Ali, J. H. Chuah, M. S. A. Talip, N. Mokhtar, and M. A. Shoaib, "Structural crack detection using deep convolutional neural networks," *Automation in Construction*, vol. 133, p. 103989, 2022.

[4] W. H. Organization, *Global status report on road safety 2018*. World Health Organization, 2019.

[5] R. Ohlan, A. Ohlan, R. Singh, and S. Kaur, "Trends in road traffic injuries mortality in india: An analysis from the global burden of disease study 1990–2021," *Journal of Prevention*, vol. 46, no. 1, pp. 59–82, 2025.

[6] R. Pal, A. Ghosh, R. Kumar, S. Galwankar, S. K. Paul, S. Pal, D. Sinha, A. Jaiswal, L. R. Moscote-Salazar, and A. Agrawal, "Public health crisis of road traffic accidents in india: Risk factor assessment and recommendations on prevention on the behalf of the academy of family physicians of india," *Journal of family medicine and primary care*, vol. 8, no. 3, pp. 775–783, 2019.

[7] 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, p. 107133, 2021.

[8] D. Arya, H. Maeda, S. K. Ghosh, D. Toshniwal, H. Omata, T. Kashiyama, and Y. Sekimoto, "Global road damage detection: State-of-the-art solutions," in *2020 IEEE International Conference on Big Data (Big Data)*. IEEE, 2020, pp. 5533–5539.

[9] O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, and D. J. Inman, "A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications," *Mechanical systems and signal processing*, vol. 147, p. 107077, 2021.

[10] Ministry of Road Transport and Highways, Government of India, "Road Accidents in India – 2022," https://morth.nic.in/sites/default/files/RA_2022_30_Oct.pdf, Oct. 2023, Transport Research Wing.

[11] K. Bansal, K. Mittal, G. Ahuja, A. Singh, and S. S. Gill, "Deepbus: Machine learning based real time pothole detection system for smart transportation using iot," *Internet Technology Letters*, vol. 3, no. 3, p. e156, 2020.

[12] S. Bouhsissin, H. Assemlali, and N. Sael, "Enhancing road safety: A convolutional neural network based approach for road damage detection," *Machine Learning with Applications*, vol. 20, p. 100668, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666827025000519

[13] M. Ansarinejad, K. Ansarinejad, P. Lu, Y. Huang, and D. Tolliver, "Autonomous vehicles in rural areas: A review of challenges, opportunities, and solutions," *Applied Sciences*, vol. 15, no. 8, 2025. [Online]. Available: https://www.mdpi.com/2076-3417/15/8/4195

[14] I. Mahdinia, R. Arvin, A. J. Khattak, and A. Ghiasi, "Safety, energy, and emissions impacts of adaptive cruise control and cooperative adaptive cruise control," *Transportation Research Record*, vol. 2674, no. 6, pp. 253–267, 2020.

[15] S. D. Pendleton, H. Andersen, X. Du, X. Shen, M. Meghjani, Y. H. Eng, D. Rus, and M. H. Ang, "Perception, planning, control, and coordination for autonomous vehicles," *Machines*, vol. 5, no. 1, 2017. [Online]. Available:

https://www.mdpi.com/2075-1702/5/1/6

[16] Y. Wang, M. Niu, K. Song, P. Jiang, and Y. Yan, "Normal-knowledge-based pavement defect segmentation using relevance-aware and cross-reasoning mechanisms," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 4, pp. 4413–4427, 2023.

[17] B. Su, H. Zhang, Z. Wu, and Z. Zhou, "Fsrdd: An efficient few-shot detector for rare city road damage detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 24 379–24 388, 2022.

[18] A. Afshari, A. Azadnia, M. Ghasemzadeh, S. Yazdani, A. Razzaghi, S. Daneshi, and K. Hushmandi, "Driving safer: a look at how men and women use advanced driver assistance systems," *Russian Open Medical Journal*, vol. 13, no. 4, p. 404, 2024.

[19] D. Arya, H. Maeda, S. K. Ghosh, D. Toshniwal, A. Mraz, T. Kashiyama, and Y. Sekimoto, "Deep learning-based road damage detection and classification for multiple countries," *Automation in Construction*, vol. 132, p. 103935, 2021.

[20] T. Kim and S.-K. Ryu, "A guideline for pothole classification," *International Journal of Engineering and Technology*, vol. 4, no. 10, pp. 618–622, 2014.

[21] K. Chang, J. Chang, and J. Liu, "Detection of pavement distresses using 3d laser scanning technology," in *Computing in civil engineering (2005)*, 2005, pp. 1–11.

[22] C. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," *Advanced engineering informatics*, vol. 25, no. 3, pp. 507–515, 2011.

[23] O. Mendoza, P. Melin, and G. Licea, "A new method for edge detection in image processing using interval type-2 fuzzy logic," in *2007 IEEE International Conference on Granular Computing (GRC 2007)*. IEEE, 2007, pp. 151–151.

[24] A. Bhat, P. Narkar, D. Shetty, and D. Vyas, "Detection of potholes using image processing techniques," *IOSR J. Eng*, vol. 2, pp. 52–56, 2018.

[25] R. Fan, X. Ai, and N. Dahnoun, "Road surface 3d reconstruction based on dense subpixel disparity map estimation," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 3025–3035, 2018.

[26] S.-S. Park, V.-T. Tran, and D.-E. Lee, "Application of various yolo models for computer vision-based real-time pothole detection," *Applied Sciences*, vol. 11, no. 23, p. 11229, 2021.

[27] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*. Springer, 2016, pp. 21–37.

[28] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.

[29] R. Fan and M. Liu, "Road damage detection based on unsupervised disparity map segmentation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 11, pp. 4906–4911, 2019.

[30] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[31] R. Fan, H. Wang, Y. Wang, M. Liu, and I. Pitas, "Graph attention layer evolves semantic segmentation for road pothole detection: A benchmark and algorithms," *IEEE transactions on image processing*, vol. 30, pp. 8144–8154, 2021.

[32] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.

[33] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890.

[34] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 801–818.

[35] M. Yang, K. Yu, C. Zhang, Z. Li, and K. Yang, "Denseaspp for semantic segmentation in street scenes," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3684–3692.

[36] R. Fan, H. Wang, M. J. Bocus, and M. Liu, "We learn better road pothole detection: from attention aggregation to adversarial domain adaptation," in *European conference on computer vision*. Springer, 2020, pp. 285–300.

[37] H. Wang, R. Fan, Y. Sun, and M. Liu, "Applying surface normal information in drivable area and road anomaly detection for ground mobile robots," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 2706–2711.

[38] R. Fan, H. Wang, P. Cai, and M. Liu, "Sne-roadseg: Incorporating surface normal information into semantic segmentation for accurate freespace detection," in *European Conference on Computer Vision*. Springer, 2020, pp. 340–356.

[39] C. Hazirbas, L. Ma, C. Domokos, and D. Cremers, "Fusenet: Incorporating depth into semantic segmentation via fusion-based cnn architecture," in *Computer Vision–ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part I 13*. Springer, 2017, pp. 213–228.

[40] W. Wang and U. Neumann, "Depth-aware cnn for rgb-d segmentation," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 135–150.

[41] Y. Sun, W. Zuo, and M. Liu, "Rtfnet: Rgb-thermal fusion network for semantic segmentation of urban scenes," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2576–2583, 2019.

[42] Y. Pan, X. Zhang, G. Cervone, and L. Yang, "Detection of asphalt pavement potholes and cracks based on the unmanned aerial vehicle multispectral imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 10, pp. 3701–3712, 2018.

[43] Z. Feng, Y. Guo, Q. Liang, M. U. M. Bhutta, H. Wang, M. Liu, and Y. Sun, "Mafnet: Segmentation of road potholes with multimodal attention fusion network

for autonomous vehicles," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–12, 2022.

[44] S. Gupta, P. Sharma, D. Sharma, V. Gupta, and N. Sambyal, "Detection and localization of potholes in thermal images using deep neural networks," *Multimedia tools and applications*, vol. 79, no. 35, pp. 26 265–26 284, 2020.

[45] Y. L. Chen, M. R. Jahanshahi, P. Manjunatha, W. Gan, M. Abdelbarr, S. F. Masri, B. Becerik-Gerber, and J. P. Caffrey, "Inexpensive multimodal sensor fusion system for autonomous data acquisition of road surface conditions," *IEEE Sensors Journal*, vol. 16, no. 21, pp. 7731–7743, 2016.

[46] R. Fan and M. Liu, "Road damage detection based on unsupervised disparity map segmentation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 11, pp. 4906–4911, 2019.

[47] M. Kamaliardakani, L. Sun, and M. K. Ardakani, "Sealed-crack detection algorithm using heuristic thresholding approach," *Journal of Computing in Civil Engineering*, vol. 30, no. 1, p. 04014110, 2016.

[48] L. Peng, W. Chao, L. Shuangmiao, and F. Baocai, "Research on crack detection method of airport runway based on twice-threshold segmentation," in *2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*. IEEE, 2015, pp. 1716–1720.

[49] N. Safaei, O. Smadi, B. Safaei, and A. Masoud, "Efficient road crack detection based on an adaptive pixel-level segmentation algorithm," *Transportation Research Record*, vol. 2675, no. 9, pp. 370–381, 2021.

[50] N.-D. Hoang, Q.-L. Nguyen, and D. Tien Bui, "Image processing–based classification of asphalt pavement cracks using support vector machine optimized by artificial bee colony," *Journal of Computing in Civil Engineering*, vol. 32, no. 5, p. 04018037, 2018.

[51] M. Salman, S. Mathavan, K. Kamal, and M. Rahman, "Pavement crack detection using the gabor filter," in *16th international IEEE conference on intelligent trans-*

portation systems (ITSC 2013). IEEE, 2013, pp. 2039–2044.

[52] M. R. Jahanshahi and S. F. Masri, "Adaptive vision-based crack detection using 3d scene reconstruction for condition assessment of structures," *Automation in Construction*, vol. 22, pp. 567–576, 2012.

[53] P. Prasanna, K. J. Dana, N. Gucunski, B. B. Basily, H. M. La, R. S. Lim, and H. Parvardeh, "Automated crack detection on concrete bridges," *IEEE Transactions on automation science and engineering*, vol. 13, no. 2, pp. 591–599, 2014.

[54] H. Oliveira and P. L. Correia, "Crackit—an image processing toolbox for crack detection and characterization," in *2014 IEEE international conference on image processing (ICIP)*. IEEE, 2014, pp. 798–802.

[55] C. Han, T. Ma, J. Huyan, X. Huang, and Y. Zhang, "Crackw-net: A novel pavement crack image segmentation convolutional neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 22 135–22 144, 2021.

[56] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[57] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[58] J.-M. Guo, H. Markoni, and J.-D. Lee, "Barnet: Boundary aware refinement network for crack detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 7343–7358, 2021.

[59] Q. Zhou, Z. Qu, and F.-r. Ju, "A lightweight network for crack detection with split exchange convolution and multi-scale features fusion," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 3, pp. 2296–2306, 2022.

[60] Z. Qu, C.-Y. Wang, S.-Y. Wang, and F.-R. Ju, "A method of hierarchical feature fusion and connected attention architecture for pavement crack detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 16 038–16 047,

2022.

[61] L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu, "Road crack detection using deep convolutional neural network," in *2016 IEEE international conference on image processing (ICIP)*. IEEE, 2016, pp. 3708–3712.

[62] A. Zhang, K. C. Wang, B. Li, E. Yang, X. Dai, Y. Peng, Y. Fei, Y. Liu, J. Q. Li, and C. Chen, "Automated pixel-level pavement crack detection on 3d asphalt surfaces using a deep-learning network," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 10, pp. 805–819, 2017.

[63] K. Zhang, H.-D. Cheng, and B. Zhang, "Unified approach to pavement crack and sealed crack detection using preclassification based on transfer learning," *Journal of Computing in Civil Engineering*, vol. 32, no. 2, p. 04018001, 2018.

[64] Z. Qu, J. Mei, L. Liu, and D.-Y. Zhou, "Crack detection of concrete pavement with cross-entropy loss function and improved vgg16 network model," *Ieee Access*, vol. 8, pp. 54 564–54 573, 2020.

[65] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[66] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *arXiv preprint arXiv:1505.04597*, 2015.

[67] S. Park, S. Bang, H. Kim, and H. Kim, "Patch-based crack detection in black box images using convolutional neural networks," *Journal of Computing in Civil Engineering*, vol. 33, no. 3, p. 04019017, 2019.

[68] H.-w. Huang, Q.-t. Li, and D.-m. Zhang, "Deep learning based image recognition for crack and leakage defects of metro shield tunnel," *Tunnelling and underground space technology*, vol. 77, pp. 166–176, 2018.

[69] S. Bang, S. Park, H. Kim, and H. Kim, "Encoder–decoder network for pixel-level road crack detection in black-box images," *Computer-Aided Civil and Infrastructure Engineering*, vol. 34, no. 8, pp. 713–727, 2019.

[70] T. Liu, L. Zhang, G. Zhou, W. Cai, C. Cai, and L. Li, "Bc-dunet-based segmentation

of fine cracks in bridges under a complex background," *Plos one*, vol. 17, no. 3, p. e0265258, 2022.

[71] J. Pang, H. Zhang, H. Zhao, and L. Li, "Dcsnet: a real-time deep network for crack segmentation," *Signal, Image and Video Processing*, pp. 1–9, 2022.

[72] Y. Gao, H. Cao, W. Cai, and G. Zhou, "Pixel-level road crack detection in uav remote sensing images based on ard-unet," *Measurement*, vol. 219, p. 113252, 2023.

[73] H. Liu, J. Yang, X. Miao, C. Mertz, and H. Kong, "Crackformer network for pavement crack segmentation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 9, pp. 9240–9252, 2023.

[74] X. Gao and B. Tong, "Mra-unet: balancing speed and accuracy in road crack segmentation network," *Signal, Image and Video Processing*, vol. 17, no. 5, pp. 2093–2100, 2023.

[75] F. Guo, Y. Qian, J. Liu, and H. Yu, "Pavement crack detection based on transformer network," *Automation in Construction*, vol. 145, p. 104646, 2023.

[76] F. Guo, J. Liu, C. Lv, and H. Yu, "A novel transformer-based network with attention mechanism for automatic pavement crack detection," *Construction and Building Materials*, vol. 391, p. 131852, 2023.

[77] Q. Dong, S. Wang, X. Chen, W. Jiang, R. Li, and X. Gu, "Pavement crack detection based on point cloud data and data fusion," *Philosophical Transactions of the Royal Society A*, vol. 381, no. 2254, p. 20220165, 2023.

[78] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 6, pp. 1137–1149, 2016.

[79] Z. Cai and N. Vasconcelos, "Cascade r-cnn: Delving into high quality object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 6154–6162.

[80] P. Sun, R. Zhang, Y. Jiang, T. Kong, C. Xu, W. Zhan, M. Tomizuka, L. Li, Z. Yuan, C. Wang *et al.*, "Sparse r-cnn: End-to-end object detection with learnable propos-

als," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 14 454–14 463.

[81] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.

[82] M. Tan, R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 10 781–10 790.

[83] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.

[84] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 7464–7475.

[85] W. Tang, S. Huang, Q. Zhao, R. Li, and L. Huangfu, "An iteratively optimized patch label inference network for automatic pavement distress detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 8652–8661, 2021.

[86] L. Wang and Z. Chen, "Road damage segmentation using convolutional neural networks on urban highways," *Transportation Research Part C: Emerging Technologies*, vol. 134, p. 103485, 2022.

[87] Q. Xie and Y. Zhou, "Efficientnet-based pavement crack detection for real-time applications," *Automation in Construction*, vol. 113, p. 103125, 2020.

[88] C. Lin, D. Tian, X. Duan, J. Zhou, D. Zhao, and D. Cao, "Da-rdd: Toward domain adaptive road damage detection across different countries," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 3, pp. 3091–3103, 2022.

[89] S. Wang, Y. Tang, X. Liao, J. He, H. Feng, H. Jiao, X. Su, and Q. Yuan, "An ensemble learning approach with multi-depth attention mechanism for road damage detection," in *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022, pp.

6439–6444.

[90] Q. Wang, J. Gao, and Y. Yuan, "Embedding structured contour and location prior in siamesed fully convolutional networks for road detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 1, pp. 230–241, 2017.

[91] L. Shen, H. Tao, Y. Ni, Y. Wang, and V. Stojanovic, "Improved yolov3 model with feature map cropping for multi-scale road object detection," *Measurement Science and Technology*, vol. 34, no. 4, p. 045406, 2023.

[92] D. Deepa and A. Sivasangari, "An effective detection and classification of road damages using hybrid deep learning framework," *Multimedia Tools and Applications*, vol. 82, no. 12, pp. 18 151–18 184, 2023.

[93] Y. Li, P. Che, C. Liu, D. Wu, and Y. Du, "Cross-scene pavement distress detection by a novel transfer learning framework," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 11, pp. 1398–1415, 2021.

[94] Y. Liu, F. Liu, W. Liu, and Y. Huang, "Pavement distress detection using street view images captured via action camera," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 1, pp. 738–747, 2023.

[95] F. Alam, F. Ofli, M. Imran, T. Alam, and U. Qazi, "Deep learning benchmarks and datasets for social media image classification for disaster response," in *2020 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*. IEEE, 2020, pp. 151–158.

[96] S. Wang, H. Jiao, X. Su, and Q. Yuan, "An ensemble learning approach with attention mechanism for detecting pavement distress and disaster-induced road damage," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 10, pp. 13 667–13 681, 2024.

[97] L.-A. Chen and S.-Y. Lin, "Enhancing road damage identification in satellite images through synthetic data," *International Journal of Disaster Risk Reduction*, vol. 116, p. 105091, 2025.

[98] N. Ma, J. Fan, W. Wang, J. Wu, Y. Jiang, L. Xie, and R. Fan, "Computer vision

for road imaging and pothole detection: a state-of-the-art review of systems and algorithms," *Transportation Safety and Environment*, vol. 4, no. 4, p. tdac026, 2022.

[99] J. Li, Y. Zhan, P. Yun, G. Zhou, Q. Chen, and R. Fan, "Roadformer: Duplex transformer for rgb-normal semantic road scene parsing," *IEEE Transactions on Intelligent Vehicles*, 2024.

[100] D. J. Fagnant and K. Kockelman, "Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations," *Transportation Research Part A: Policy and Practice*, vol. 77, pp. 167–181, 2015.

[101] J. Piao and M. McDonald, "Advanced driver assistance systems from autonomous to cooperative approach," *Transport reviews*, vol. 28, no. 5, pp. 659–684, 2008.

[102] J. M. Scanlon, R. Sherony, and H. C. Gabler, "Injury mitigation estimates for an intersection driver assistance system in straight crossing path crashes in the united states," *Traffic injury prevention*, vol. 18, no. sup1, pp. S9–S17, 2017.

[103] T. Seacrist, R. Sahani, G. Chingas, E. C. Douglas, V. Graci, and H. Loeb, "Efficacy of automatic emergency braking among risky drivers using counterfactual simulations from the shrp 2 naturalistic driving study," *Safety science*, vol. 128, p. 104746, 2020.

[104] Y. Li, Z. Li, H. Wang, W. Wang, and L. Xing, "Evaluating the safety impact of adaptive cruise control in traffic oscillations on freeways," *Accident Analysis & Prevention*, vol. 104, pp. 137–145, 2017.

[105] N. Strand, J. Nilsson, I. Karlsson, and L. Nilsson, "Exploring end-user experiences: self-perceived notions on use of adaptive cruise control systems," *IET intelligent transport systems*, vol. 5, no. 2, pp. 134–140, 2011.

[106] L. Yue, M. A. Abdel-Aty, Y. Wu, and A. Farid, "The practical effectiveness of advanced driver assistance systems at different roadway facilities: System limitation, adoption, and usage," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3859–3870, 2019.

[107] J. B. Cicchino, "Effectiveness of forward collision warning and autonomous emergency braking systems in reducing front-to-rear crash rates," *Accident Analysis &*

Prevention, vol. 99, pp. 142–152, 2017.

[108] S. Cafiso and A. Di Graziano, "Evaluation of the effectiveness of adas in reducing multi-vehicle collisions," *International journal of heavy vehicle systems*, vol. 19, no. 2, pp. 188–206, 2012.

[109] J. Bärgman, C.-N. Boda, and M. Dozza, "Counterfactual simulations applied to shrp2 crashes: The effect of driver behavior models on safety benefit estimations of intelligent safety systems," *Accident Analysis & Prevention*, vol. 102, pp. 165–180, 2017.

[110] B. D. Seppelt and J. D. Lee, "Making adaptive cruise control (acc) limits visible," *International journal of human-computer studies*, vol. 65, no. 3, pp. 192–205, 2007.

[111] S. H. Haus, R. Sherony, and H. C. Gabler, "Estimated benefit of automated emergency braking systems for vehicle–pedestrian crashes in the united states," *Traffic injury prevention*, vol. 20, no. sup1, pp. S171–S176, 2019.

[112] M. Ra, H. G. Jung, J. K. Suhr, and W.-Y. Kim, "Part-based vehicle detection in side-rectilinear images for blind-spot detection," *Expert Systems with Applications*, vol. 101, pp. 116–128, 2018.

[113] R. Anderson, S. Doecke, J. Mackenzie, and G. Ponte, "Potential benefits of autonomous emergency braking based on in-depth crash reconstruction and simulation," in *Proceedings of the 23rd International Conference on Enhanced Safety of Vehicles, US National Highway Traffic Safety Administration, Washington DC*, 2013.

[114] J. M. Scanlon, K. D. Kusano, and H. C. Gabler, "Lane departure warning and prevention systems in the us vehicle fleet: Influence of roadway characteristics on potential safety benefits," *Transportation Research Record*, vol. 2559, no. 1, pp. 17–23, 2016.

[115] Y. Safyari, M. Mahdianpari, and H. Shiri, "A review of vision-based pothole detection methods using computer vision and machine learning," *Sensors*, vol. 24, no. 17, p. 5652, 2024.

[116] J. Guan, X. Yang, L. Ding, X. Cheng, V. C. Lee, and C. Jin, "Automated pixel-level

pavement distress detection based on stereo vision and deep learning," *Automation in Construction*, vol. 129, p. 103788, 2021.

[117] I. Almasri, D. Manasreh, and M. D. Nazzal, "Ai meets adas: Intelligent pothole detection for safer av navigation," *Vehicles*, vol. 7, no. 4, p. 109, 2025.

[118] S. Kuang, Y. Liu, X. Wang, X. Qu, and Y. Wei, "An universal crack detection framework for intelligent road-perceptive vehicles," *IEEE Transactions on Intelligent Vehicles*, 2024.

[119] Z. Feng, Y. Guo, D. Navarro-Alarcon, Y. Lyu, and Y. Sun, "Inconseg: Residual-guided fusion with inconsistent multi-modal data for negative and positive road obstacles segmentation," *IEEE Robotics and Automation Letters*, vol. 8, no. 8, pp. 4871–4878, 2023.

[120] O. Al-maqtari, B. Peng, Z. Al-Huda, A. Al-Malahi, and N. Maqtary, "Lightweight yet effective: A modular approach to crack segmentation," *IEEE Transactions on Intelligent Vehicles*, 2024.

[121] W. Zheng, X. Jiang, Z. Fang, and Y. Gao, "Tv-net: A structure-level feature fusion network based on tensor voting for road crack segmentation," *IEEE Transactions on Intelligent Transportation Systems*, 2024.

[122] S. Kong, Q. Meng, X. Li, Z. Wang, X. Liu, and B. Li, "Road perception for autonomous driving: Pothole detection in complex environments based on improved yolov8," *IEEE Access*, 2025.

[123] L. Li, J. Liu, J. Xing, Z. Liu, K. Lin, and B. Du, "Road pothole detection based on crowdsourced data and extended mask r-cnn," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 9, pp. 12 504–12 516, 2024.

[124] Z. Feng, Y. Guo, and Y. Sun, "Segmentation of road negative obstacles based on dual semantic-feature complementary fusion for autonomous driving," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 4, pp. 4687–4697, 2024.

[125] J. Harvey, I. L. Al-Qadi, H. Ozer, and G. Flintsch, *Pavement, Roadway, and Bridge Life Cycle Assessment 2020: Proceedings of the International Symposium on Pave-*

ment. Roadway, and Bridge Life Cycle Assessment 2020 (LCA 2020, Sacramento, CA, 3-6 June 2020). CRC Press, 2020.

[126] R. Fan, U. Ozgunalp, B. Hosking, M. Liu, and I. Pitas, "Pothole detection based on disparity transformation and road surface modeling," *IEEE Transactions on Image Processing*, vol. 29, pp. 897–908, 2019.

[127] C. Koch, K. Georgieva, V. Kasireddy, B. Akinci, and P. Fieguth, "A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure," *Advanced Engineering Informatics*, vol. 29, no. 2, pp. 196–210, 2015.

[128] J. Kim, S. Kim, J. Jeong, H. Kim, J.-S. Park, and T. Kim, "Cbdn: cloud-based drone navigation for efficient battery charging in drone networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 11, pp. 4174–4191, 2018.

[129] A. Rahman and S. Patel, "Annotated potholes image dataset," *Kaggle*, 2020.

[130] K. Gajjar, T. van Niekerk, T. Wilm, and P. Mercorelli, "Vision-based deep learning algorithm for detecting potholes," in *Journal of Physics: Conference Series*, vol. 2162, no. 1. IOP Publishing, 2022, p. 012019.

[131] A. A. Shaghouri, R. Alkhatib, and S. Berjaoui, "Real-time pothole detection using deep learning," *arXiv preprint arXiv:2107.06356*, 2021.

[132] B. Li, S. Liu, F. Wu, G. Li, M. Zhong, and X. Guan, "Rt-unet: An advanced network based on residual network and transformer for medical image segmentation," *International Journal of Intelligent Systems*, vol. 37, no. 11, pp. 8565–8582, 2022.

[133] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual u-net," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 749–753, 2018.

[134] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 801–818.

[135] S. Mehta, M. Rastegari, A. Caspi, L. Shapiro, and H. Hajishirzi, "Espnet: Efficient spatial pyramid of dilated convolutions for semantic segmentation," in *Proceedings*

of the european conference on computer vision (ECCV), 2018, pp. 552–568.

[136] T. Takikawa, D. Acuna, V. Jampani, and S. Fidler, "Gated-scnn: Gated shape cnns for semantic segmentation," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 5229–5238.

[137] Z. Feng, Y. Guo, Q. Liang, M. U. M. Bhutta, H. Wang, M. Liu, and Y. Sun, "Mafnet: Segmentation of road potholes with multimodal attention fusion network for autonomous vehicles," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–12, 2022.

[138] H. S. Munawar, A. W. Hammad, A. Haddad, C. A. P. Soares, and S. T. Waller, "Image-based crack detection methods: A review," *Infrastructures*, vol. 6, no. 8, p. 115, 2021.

[139] D. Ai, G. Jiang, S.-K. Lam, P. He, and C. Li, "Computer vision framework for crack detection of civil infrastructure—a review," *Engineering Applications of Artificial Intelligence*, vol. 117, p. 105478, 2023.

[140] Y.-A. Hsieh and Y. J. Tsai, "Machine learning for crack detection: Review and model performance comparison," *Journal of Computing in Civil Engineering*, vol. 34, no. 5, p. 04020038, 2020.

[141] J. Gao and H. Tembine, "Distributed mean-field-type filters for traffic networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 507–521, 2018.

[142] Z. Wang, Y. Lu, Q. Li, X. Tao, Y. Guo, M. Gong, and T. Liu, "Cris: Clip-driven referring image segmentation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 11 686–11 695.

[143] J. Zhai, S. Zhang, J. Chen, and Q. He, "Autoencoder and its various variants," in *2018 IEEE international conference on systems, man, and cybernetics (SMC)*. IEEE, 2018, pp. 415–419.

[144] S. Xiao, K. Shang, K. Lin, Q. Wu, H. Gu, and Z. Zhang, "Pavement crack detection with hybrid-window attentive vision transformers," *International Journal of Applied*

Earth Observation and Geoinformation, vol. 116, p. 103172, 2023.

[145] W. Wang and C. Su, "Automatic concrete crack segmentation model based on transformer," *Automation in Construction*, vol. 139, p. 104275, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580522001480

[146] F. Yang, L. Zhang, S. Yu, D. Prokhorov, X. Mei, and H. Ling, "Feature pyramid and hierarchical boosting network for pavement crack detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 4, pp. 1525–1535, 2019.

[147] N. Kheradmandi and V. Mehranfar, "A critical review and comparative study on image segmentation-based techniques for pavement crack detection," *Construction and Building Materials*, vol. 321, p. 126162, 2022.

[148] Y. Hu, J. Wang, X. Ai, and X. Zhuang, "An improved multithreshold segmentation algorithm based on graph cuts applicable for irregular image," *Mathematical Problems in Engineering*, vol. 2019, no. 1, p. 3514258, 2019.

[149] N.-D. Hoang, "Classification of asphalt pavement cracks using laplacian pyramid-based image processing and a hybrid computational approach," *Computational Intelligence and Neuroscience*, vol. 2018, no. 1, p. 1312787, 2018.

[150] Y.-J. Cha, W. Choi, and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 5, pp. 361–378, 2017.

[151] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, "Automatic road crack detection using random structured forests," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 12, pp. 3434–3445, 2016.

[152] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, "Cracktree: Automatic crack detection from pavement images," *Pattern Recognition Letters*, vol. 33, no. 3, pp. 227–238, 2012.

[153] Y. Liu, J. Yao, X. Lu, R. Xie, and L. Li, "Deepcrack: A deep hierarchical feature learning architecture for crack segmentation," *Neurocomputing*, vol. 338, pp. 139–153, 2019.

[154] R. Stricker, M. Eisenbach, M. Sesselmann, K. Debes, and H.-M. Gross, "Improving visual road condition assessment by extensive experiments on the extended gaps dataset," in *2019 international joint conference on neural networks (IJCNN)*. IEEE, 2019, pp. 1–8.

[155] S. Xie and Z. Tu, "Holistically-nested edge detection," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1395–1403.

[156] C. Liu, C. Zhu, X. Xia, J. Zhao, and H. Long, "Ffedn: Feature fusion encoder decoder network for crack detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 15 546–15 557, 2022.

[157] P. Shi, F. Zhu, Y. Xin, and S. Shao, "U2cracknet: a deeper architecture with two-level nested u-structure for pavement crack detection," *Structural Health Monitoring*, vol. 22, no. 4, pp. 2910–2921, 2023.

[158] Y. Ren, J. Huang, Z. Hong, W. Lu, J. Yin, L. Zou, and X. Shen, "Image-based concrete crack detection in tunnels using deep fully convolutional networks," *Construction and Building Materials*, vol. 234, p. 117367, 2020.

[159] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, "Segformer: Simple and efficient design for semantic segmentation with transformers," *Advances in neural information processing systems*, vol. 34, pp. 12 077–12 090, 2021.

[160] Y. Li, R. Ma, H. Liu, and G. Cheng, "Real-time high-resolution neural network with semantic guidance for crack segmentation," *Automation in Construction*, vol. 156, p. 105112, 2023.

[161] J. M. Goo, X. Milidonis, A. Artusi, J. Boehm, and C. Ciliberto, "Hybrid-segmentor: A hybrid approach to automated fine-grained crack segmentation in civil infrastructure," *arXiv preprint arXiv:2409.02866*, 2024.

[162] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890.

[163] M. He and T. L. Lau, "Crackham: A novel automatic crack detection network based

on u-net for asphalt pavement," *IEEE Access*, vol. 12, pp. 12 655–12 666, 2024.

[164] Q. Zou, Z. Zhang, Q. Li, X. Qi, Q. Wang, and S. Wang, "Deepcrack: Learning hierarchical convolutional features for crack detection," *IEEE transactions on image processing*, vol. 28, no. 3, pp. 1498–1512, 2018.

[165] Z. Qu, C. Cao, L. Liu, and D.-Y. Zhou, "A deeply supervised convolutional neural network for pavement crack detection with multiscale feature fusion," *IEEE transactions on neural networks and learning systems*, vol. 33, no. 9, pp. 4890–4899, 2021.

[166] M. Sun, H. Zhao, and J. Li, "Road crack detection network under noise based on feature pyramid structure with feature enhancement (road crack detection under noise)," *IET Image Processing*, vol. 16, no. 3, pp. 809–822, 2022.

[167] H. Zhang, C. Wu, Z. Zhang, Y. Zhu, H. Zhang, Y. Lin, Y. Sun, T. He, J. Mueller, R. Manmatha, M. Li, and A. Smola, "Resnest: Split-attention networks," *arXiv preprint arXiv:2004.08955*, 2020.

[168] J. Zhao, J. Wang, J. Ma, and S. Zhao, "Road crack detection based on residual learning and data augmentation," *IET Image Processing*, vol. 17, no. 6, pp. 616–628, 2023.

[169] H. Ayman and M. W. Fakhr, "Recent computer vision applications for pavement distress and condition assessment," *Automation in Construction*, vol. 146, p. 104664, 2023.

[170] D. Arya, H. Maeda, S. K. Ghosh, D. Toshniwal, and Y. Sekimoto, "Rdd2022: A multi-national image dataset for automatic road damage detection," *Geoscience Data Journal*, vol. 11, no. 4, pp. 846–862, 2024.

[171] F. Kortmann, K. Talits, P. Fassmeyer, A. Warnecke, N. Meier, J. Heger, P. Drews, and B. Funk, "Detecting various road damage types in global countries utilizing faster r-cnn," in *2020 IEEE International Conference on Big Data (Big Data)*. IEEE, 2020, pp. 5563–5571.

[172] J. Chen, X. Yu, Q. Li, W. Wang, and B.-G. He, "Lag-yolo: Efficient road damage

detector via lightweight attention ghost module," *Journal of Intelligent Construction*, vol. 2, no. 1, pp. 1–10, 2024.

[173] G. Guo and Z. Zhang, "Road damage detection algorithm for improved yolov5," *Scientific reports*, vol. 12, no. 1, p. 15523, 2022.

[174] J. Li, Z. Qu, S.-Y. Wang, and S.-F. Xia, "Yolox-rdd: a method of anchor-free road damage detection for front-view images," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 10, pp. 14 725–14 739, 2024.

[175] Y. Huan, X. Li, and R. Wang, "Challenges in unpaved road detection and the role of automated systems," *Journal of Transportation Engineering*, vol. 17, no. 2, pp. 53–68, 2023.

[176] L. Wang and Y. Zhang, "Road damage detection using convolutional neural networks: A survey," *Computer Vision and Image Understanding*, vol. 200, p. 101945, 2022.

[177] G. Bosurgi, M. Modica, and O. Pellegrino, "An automatic pothole detection algorithm using pavement 3d data," *International Journal of Pavement Engineering*, vol. 24, no. 2, p. 2057978, 2023.

[178] R. Fan, X. Ai, and N. Dahnoun, "Road surface 3d reconstruction based on dense subpixel disparity map estimation," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 3025–3035, 2018.

[179] A. Dhiman and R. Klette, "Pothole detection using computer vision and learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 8, pp. 3536–3550, 2019.

[180] Y. Huang, M. Jahanshahi, F. Shen, and T. Mondal, "Deep learning–based autonomous road condition assessment leveraging inexpensive rgb and depth sensors and heterogeneous data fusion: Pothole detection and quantification," *Journal of Transportation Engineering, Part B: Pavements*, vol. 149, no. 2, p. 04023010, 2023.

[181] R. Fan, H. Wang, Y. Wang, M. Liu, and I. Pitas, "Graph attention layer evolves semantic segmentation for road pothole detection: A benchmark and algorithms," *IEEE transactions on image processing*, vol. 30, pp. 8144–8154, 2021.

[182] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[183] T. Rateke and A. Von Wangenheim, “Road surface detection and differentiation considering surface damages,” *Autonomous Robots*, vol. 45, no. 2, pp. 299–312, 2021.

[184] Y. Nie, K. An, X. Chen, L. Zhao, W. Liu, X. Wang, Y. Yu, W. Luo, K. Li, and Z. Zhang, “An improved u-net network for sandy road extraction from remote sensing imagery,” *Remote Sensing*, vol. 15, no. 20, p. 4899, 2023.

[185] W. Li, J. Zhang, J. Li, and L. Zhuo, “Unpaved road segmentation of uav imagery via a global vision transformer with dilated cross window self-attention for dynamic map,” *The Visual Computer*, pp. 1–19, 2024.

[186] C. Bai, L. Zhang, L. Gao, L. Peng, P. Li, and L. Yang, “Real-time segmentation algorithm of unstructured road scenes based on improved bisenet,” *Journal of Real-Time Image Processing*, vol. 21, no. 3, p. 91, 2024.

PhD Thesis

ORIGINALITY REPORT

19%	10%	18%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	link.springer.com Internet Source	6%
2	Ruchi, S. Indu, Rohit Kumar. "Detection of Road Potholes Using Deep Learning Based Improved YOLOv4 Network", 2022 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), 2022 Publication	2%
3	Ruchi Tripathi, S. Indu, Rohit Kumar. "ERCU-Net: segmentation of road potholes using enhanced residual convolutional block based on U-Net for ADAS", Signal, Image and Video Processing, 2024 Publication	2%
4	Submitted to Delhi Technological University Student Paper	1%
5	Paolo Ferro, Harinadh Vemanaboina, Chander Prakash. "Computational Techniques and Smart Manufacturing", CRC Press, 2026 Publication	<1%
6	Minh-Tu Cao, Quoc-Viet Tran, Ngoc-Mai Nguyen, Kuan-Tsung Chang. "Survey on performance of deep learning models for detecting road damages using multiple	<1%

dashcam image resources", Advanced Engineering Informatics, 2020
Publication

7	arxiv.org Internet Source	<1 %
8	R. N. V. Jagan Mohan, B. H. V. S. Rama Krishnam Raju, V. Chandra Sekhar, T. V. K. P. Prasad. "Algorithms in Advanced Artificial Intelligence - Proceedings of International Conference on Algorithms in Advanced Artificial Intelligence (ICAAAI-2024)", CRC Press, 2025 Publication	<1 %
9	Thangaprakash Sengodan, Sanjay Misra, M Murugappan. "Advances in Electrical and Computer Technologies", CRC Press, 2025 Publication	<1 %
10	Submitted to Wayne State University Student Paper	<1 %
11	d197for5662m48.cloudfront.net Internet Source	<1 %
12	Manoj Kumar, Tanweer Ali, Jaume Anguera, Suman Lata Tripathi. "Emerging Technologies in AI, Computation, Communication, and Cybersecurity - Proceedings of the First International Conference on Artificial Intelligence, Computation, Communication and Network Security (AICCoNS 2025)", CRC Press, 2026 Publication	<1 %
13	"ICT Analysis and Applications", Springer Science and Business Media LLC, 2026 Publication	<1 %

14	opus.lib.uts.edu.au Internet Source	<1%
15	Wenwen Zheng, Xiaoyan Jiang, Zhijun Fang, Yongbin Gao. "TV-Net: A Structure-Level Feature Fusion Network Based on Tensor Voting for Road Crack Segmentation", IEEE Transactions on Intelligent Transportation Systems, 2024 Publication	<1%
16	Kiran Dabral. "Integration of Self-Healing Functionalities in Structural Design for Serviceability of Concrete Structures under Marine Exposure", Universitat Politecnica de Valencia, 2025 Publication	<1%
17	"Advanced Intelligent Computing Technology and Applications", Springer Science and Business Media LLC, 2025 Publication	<1%
18	"Computer Vision and Robotics", Springer Science and Business Media LLC, 2026 Publication	<1%
19	Submitted to Monash University Student Paper	<1%
20	Abraham, Sophia. "Homotopy-Based Multi-Objective Optimization for Enhanced Machine Learning in Black-Box Environments", University of Notre Dame, 2025 Publication	<1%
21	escholarship.org Internet Source	<1%

22	Suneeta Satpathy, Álvaro Rocha, Sachi Nandan Mohanty, Tanupriya Choudhury. "Intelligent Data-Driven Systems with Innovations in Artificial Intelligence", CRC Press, 2025 Publication	<1 %
23	"Computer Vision and Image Processing", Springer Science and Business Media LLC, 2024 Publication	<1 %
24	"Intelligent IT Solutions to Promote Indigenous Innovations", Springer Science and Business Media LLC, 2025 Publication	<1 %
25	dspace.univ-ghardaia.dz Internet Source	<1 %
26	Hatem Taha, Hossam El-Habrouk, Wael Bekheet, Sayed El-Naghi, Marwan Torki. "Pixel-level pavement crack segmentation using UAV remote sensing images based on the ConvNeXt-UPerNet", Alexandria Engineering Journal, 2025 Publication	<1 %
27	"Computer Vision Systems", Springer Science and Business Media LLC, 2019 Publication	<1 %
28	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2018", Springer Science and Business Media LLC, 2018 Publication	<1 %
29	"Computer Vision – ECCV 2020 Workshops", Springer Science and Business Media LLC,	<1 %

2020
Publication

30	stax.strath.ac.uk Internet Source	<1 %
31	repository.rice.edu Internet Source	<1 %
32	Cenk Temizel, Salih Tutun, Celal Hakan Canbaz, Ekrem Alagoz et al. "Artificial Intelligence in the Energy Industry - Theory, Case Studies, and Applications", CRC Press, 2026 Publication	<1 %
33	Khumukcham Robindro Singh, Nazrul Hoque, Arnab Kumar Maji, Sabyasachi Mondal et al. "Emerging Trends and Future Directions in Artificial Intelligence, Machine Learning, and Internet of Things Innovations - A proceeding of NEIAIS — 2025", CRC Press, 2025 Publication	<1 %
34	Harvinder Singh, Priyanka Kaushal, Sarabjeet Kaur. "Advanced Computing and AI for Sustainable Environment, Energy, and Smart Manufacturing - AI-Driven Sustainability Across Environment, Energy, and Industry", CRC Press, 2026 Publication	<1 %
35	"Advanced Intelligent Computing Technology and Applications", Springer Science and Business Media LLC, 2024 Publication	<1 %
36	Submitted to University of Lancaster Student Paper	<1 %

37	Frank Y. Shih. "AI Deep Learning in Image Processing", CRC Press, 2025 Publication	<1 %
38	Xiaofeng Yang, Tonghe Wang. "AI-driven Medical Image Analysis in Precision Radiation Therapy", CRC Press, 2026 Publication	<1 %
39	"Smart Computing Paradigms: Intelligent Solutions for Sustainable Wellbeing", Springer Science and Business Media LLC, 2026 Publication	<1 %
40	Asali, Ehsan. "Spatiotemporal Multimodal Representation Learning for Activity Understanding and Behavior Monitoring.", University of Georgia Publication	<1 %
41	Erabati, Gopi Krishna. "Learning to Perceive: Scene Understanding for Autonomous Driving.", Universidade de Coimbra (Portugal) Publication	<1 %
42	"Computer Vision – ECCV 2020", Springer Science and Business Media LLC, 2020 Publication	<1 %
43	Maluleke, Musa. "Automated Pothole Detection for Road Maintenance: Selection Considerations.", University of Johannesburg (South Africa) Publication	<1 %
44	Nimpa Fondje, Cedric A.. "Bridging Domain Gaps for Cross-Spectrum and Long-range Face Recognition Using Domain Adaptive Machine Learning", The University of Nebraska - Lincoln, 2023	<1 %

Publication

45	Novakazi, Fjolle | Orlovksa, Julia. "Design for Perception - A Human-Centric Approach to the Design of Driving Automation Systems Based on the Driver's Perception", Chalmers Tekniska Hogskola (Sweden), 2024 Publication	<1 %
46	Shi Yi, Junjie Li, Xi Liu, Xuesong Yuan. "CCAFFMNet: Dual-spectral semantic segmentation network with channel-coordinate attention feature fusion module", Neurocomputing, 2021 Publication	<1 %
47	Yuxi Gao, Hongbin Cao, Weiwei Cai, Guoxiong Zhou. "Pixel-level road crack detection in UAV remote sensing images based on ARD-Unet", Measurement, 2023 Publication	<1 %
48	theses.lib.polyu.edu.hk Internet Source	<1 %
49	www.hindawi.com Internet Source	<1 %
50	"Advances in Visual Computing", Springer Science and Business Media LLC, 2025 Publication	<1 %
51	"Computer Vision and Image Processing", Springer Science and Business Media LLC, 2022 Publication	<1 %
52	Baoshan Sun, Jinze Liu, Liqing Gao, Kaiyu Bi, yong zhu, Jiabao Wen. "RDD-DETR: Lightweight Road Damage Detection with	<1 %

Hierarchical Feature Interaction", Engineering Research Express, 2025
Publication

53	Jaiteg Singh, S B Goyal, Rajesh Kumar Kaushal, Naveen Kumar, Sukhjit Singh Sehra. "Applied Data Science and Smart Systems - Proceedings of 2nd International Conference on Applied Data Science and Smart Systems 2023 (ADSSS 2023) 15-16 Dec, 2023, Rajpura, India", CRC Press, 2024 Publication	<1 %
54	Submitted to Liverpool John Moores University Student Paper	<1 %
55	Moeini, Mohammad. "Computational Analysis and Data-Driven Modeling of Hurricane Effects on Low-Rise Residential Buildings", The Pennsylvania State University, 2025 Publication	<1 %
56	"Computer Vision – ECCV 2022", Springer Science and Business Media LLC, 2022 Publication	<1 %
57	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2018", Springer Nature America, Inc, 2018 Publication	<1 %
58	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2020", Springer Science and Business Media LLC, 2020 Publication	<1 %
59	Ngeni, Frank Charles. "Multiple Object Tracking (MOT) to Solve Traffic Data Occlusion	<1 %

Problem Using DeepSORT and Quantum Computing", South Carolina State University, 2023
Publication

60	Chu, Ziping. "Robust Automated Tumour Segmentation Using Deep Learning.", University of New South Wales (Australia) Publication	<1 %
61	Corinne Mulley, John D. Nelson, Stephen Ison. "The Routledge Handbook of Public Transport", Routledge, 2021 Publication	<1 %
62	Jian Kang, Haiyan Guan, Dedong Zhang, Lingfei Ma, Lanying Wang, Yongtao Yu, Linlin Xu, Jonathan Li. "RPDNet: Street-level road pavement damage detection with a real-time anchor-free network", International Journal of Applied Earth Observation and Geoinformation, 2026 Publication	<1 %
63	Liang, Xiaolong. "Advanced Deep Learning Approaches for Real-Time Scene Classification, Polyp Detection, and Intestine Localization in Endoscopy Videos", University of Massachusetts Lowell, 2025 Publication	<1 %
64	Submitted to University College London Student Paper	<1 %
65	Zhang, Chenyu. "Multitask, Weakly-Supervised, and Gaze-Guided Deep Learning Methods for Automating Structural Inspection.", State University of New York at Stony Brook Publication	<1 %

66	Chen, Yujing. "Sensor Analytics using Multi-Task Learning and Federated Learning", George Mason University Publication	<1 %
67	Yu, Yunrui. "Reliable and Efficient Adversarial Attack Evaluation Strategies for Robust Deep Learning.", University of Macau Publication	<1 %
68	Ammar M. Okran, Hatem A. Rashwan, Adel Saleh, Domenec Puig. "Efficient crack segmentation with multi-decoder networks and enhanced feature fusion", Engineering Applications of Artificial Intelligence, 2025 Publication	<1 %
69	Fei Hu, Iftikhar Rasheed. "Deep Learning and Its Applications for Vehicle Networks", CRC Press, 2023 Publication	<1 %
70	Submitted to Khalifa University of Science Technology and Research Student Paper	<1 %
71	Rahul Kumar Chaurasiya, Varun Bajaj, Vishakha Chourasia. "Artificial Intelligence-based Signal Processing for Brain Activity Analysis", CRC Press, 2025 Publication	<1 %
72	www.researchgate.net Internet Source	<1 %
73	Lei Yang, Suli Bai, Yanhong Liu, Hongnian Yu. "Multi-scale triple-attention network for pixelwise crack segmentation", Automation in Construction, 2023 Publication	<1 %

74	"Computer Vision – ECCV 2018", Springer Science and Business Media LLC, 2018 Publication	<1 %
75	"Merged Papers", 2020 7th International Conference on Smart Structures and Systems (ICSSS), 2020 Publication	<1 %
76	"Pattern Recognition and Computer Vision", Springer Science and Business Media LLC, 2020 Publication	<1 %
77	Almansor, Ebtesam Hussain. "Intelligent and Proactive Approach for the Optimal Handling of Low Chatbot Quality of Services (CQoS)", University of Technology Sydney (Australia), 2023 Publication	<1 %
78	Deeksha Arya, Hiroya Maeda, Sanjay Kumar Ghosh, Durga Toshniwal, Alexander Mraz, Takehiro Kashiyama, Yoshihide Sekimoto. "Deep learning-based road damage detection and classification for multiple countries", Automation in Construction, 2021 Publication	<1 %
79	Feng Guo, Jian Liu, Chengshun Lv, Huayang Yu. "A novel transformer-based network with attention mechanism for automatic pavement crack detection", Construction and Building Materials, 2023 Publication	<1 %
80	Jun Peng, Weidong Wang, Wenbo Hu, Chengbo Ai, Xinyue Xu, Youyin Shi, Jin Wang, Zhifa Ran, Shi Qiu. "Automated detection and	<1 %

quantification of pavement cracking around manhole", Engineering Applications of Artificial Intelligence, 2024
Publication

81	Kanchan Naithani, Shrikant Tiwari, Shabnam Kumari. "AI for Product Development", CRC Press, 2025 Publication	<1 %
82	Peng, Han. "Action Recognition in Intelligent Systems", Northern Arizona University, 2023 Publication	<1 %
83	Sirisha Potluri, Suneeta Satpathy, Santi Swarup Basa, Antonio Zuorro. "AI in Agriculture for Sustainable and Economic Management", CRC Press, 2024 Publication	<1 %
84	Zhang, Zhang. "Enhancing Endoscopic Lesion Detection with Lightweight Transformers, Multi-Scale Attention, and Vision-Language Models", University of Massachusetts Lowell, 2025 Publication	<1 %

Exclude quotes	On	Exclude matches	< 10 words
Exclude bibliography	On		

PhD Thesis

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18

PAGE 19

PAGE 20

PAGE 21

PAGE 22

PAGE 23

PAGE 24

PAGE 25

PAGE 26

PAGE 27

PAGE 28

PAGE 29

PAGE 30

PAGE 31

PAGE 32

PAGE 33

PAGE 34

PAGE 35

PAGE 36

PAGE 37

PAGE 38

PAGE 39

PAGE 40

PAGE 41

PAGE 42

PAGE 43

PAGE 44

PAGE 45

PAGE 46

PAGE 47

PAGE 48

PAGE 49

PAGE 50

PAGE 51

PAGE 52

PAGE 53

PAGE 54

PAGE 55

PAGE 56

PAGE 57

PAGE 58

PAGE 59

PAGE 60

PAGE 61

PAGE 62

PAGE 63

PAGE 64

PAGE 65

PAGE 66

PAGE 67

PAGE 68

PAGE 69

PAGE 70

PAGE 71

PAGE 72

PAGE 73

PAGE 74

PAGE 75

PAGE 76

PAGE 77

PAGE 78

PAGE 79

PAGE 80

PAGE 81

PAGE 82

PAGE 83

PAGE 84

PAGE 85

PAGE 86

PAGE 87

PAGE 88

PAGE 89

PAGE 90

PAGE 91

PAGE 92

PAGE 93

PAGE 94

PAGE 95

PAGE 96

PAGE 97

PAGE 98

PAGE 99

PAGE 100

PAGE 101

PAGE 102

PAGE 103

PAGE 104

PAGE 105

PAGE 106

PAGE 107

PAGE 108

PAGE 109

PAGE 110

PAGE 111

PAGE 112

PAGE 113

PAGE 114

PAGE 115

PAGE 116

PAGE 117

PAGE 118

PAGE 119

PAGE 120

PAGE 121

PAGE 122

PAGE 123

PAGE 124

PAGE 125

PAGE 126

PAGE 127

PAGE 128

PAGE 129

PAGE 130

PAGE 131

PAGE 132

PAGE 133

PAGE 134

PAGE 135

PAGE 136

PAGE 137

PAGE 138

PAGE 139

PAGE 140

PAGE 141

PAGE 142

PAGE 143

PAGE 144

PAGE 145

PAGE 146

PAGE 147

PAGE 148

PAGE 149

PAGE 150

PAGE 151