Journal of Information Systems Engineering and Management
2025, 10(5s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Literature Review on Road Damage Detection and Severity Recognition: Leveraging Computer Vision

Jonathan Estilong[1], Thelma Palaoag[2]
[1]*Don Mariano Marcos Memorial State University,College of Computer Science, Agoo, La Union, Philippines*
jestilong@dmmmsu.edu.ph
[2]*University of the Cordilleras, College of Computer Science, Baguio City, Philippines.*
tdpalaoag@uc-bcf.edu.ph

ARTICLE INFO

Received: 12 Oct 2024

Revised: 10 Dec 2024

Accepted: 25 Dec 2024

ABSTRACT

The increasing demand for quick and precise road maintenance has therefore been underscored by extensive research that was created over the last years, to find automated systems in which computer vision can work — harnessing deep learning methodologies to detect defects on roads & figure out how severe these actually are. It provides a discussion of different methodologies and technologies in these pipelines such as Convolutional Neural Networks (CNNs), YOLO models, ensembling learning etc. They have a good performance in various kinds of road damage: cracks, potholes, and surface deformations identification with high accuracy. They also use methods like image tiling, transfer learning, and multiple spectral data to increase their robustness so that models can be used effectively in different environmental conditions.But the avoidance of problems is still a long way off. Nonetheless, the disparate road scenarios across geographies and absence of labeled high-quality datasets remained as some common bottlenecks in addition to real-time-driven requirements that mandate significant compute powers. Although much more headway has been made in recognizing severity (using segmentation approaches and multimodal data, in particular), the combination of detection and classification altogether within a single joint pipeline is significantly harder.A comprehensive analysis of state-of-the-art software tools, recent challenges, and advances to improve the detection/ severity identification capabilities of road damage systems. It emphasizes how to address data problems, enhance model generalization capability in unseen scenes(a.k.a novel), and minimize latency. One final takeaway is the valuable reminder of the need to create operational standards for classifier performance required to drive practice adoption and scale in larger deployment settings ideally incentivizing public-private partnerships.

Keywords: deep learning, road damage detection, severity recognition

INTRODUCTION

Road infrastructure is a central part of transport systems, which contributes significantly to the safe and effective locomotion of vehicles. Nevertheless, roads are exposed to various stresses such as traffic loads and environmental conditions in addition to the natural aging of materials resulting in cracks, potholes, and surface deformations. So, to make road repair work correctly and timely this damage needs to be detected as soon as possible and accurately measured even with severe. These are crucial measures to be taken to prevent accidents, improve repair planning, and decrease the costs involved (Maeda et al., 2018; Zhang et al., 2023).

Advances in computer vision and deep learning have formed a paradigm shift towards road maintenance with the help of automatic systems that can detect and evaluate specific potholes now. Manual inspection through the traditional approach is quite labor-intensive and time-consuming and is subjugated to human bias (Fassmeyer et al., 2021; Ha et al., 2022), which is an advantage that these new AI-based technologies have over manual methodologies. These automated systems, utilizing high-quality imaging techniques and machine learning models to correctly detect and predict road damages can contribute significantly towards more optimum ways of managing roads (Zhang et al., 2023).

1.1. Challenges in Road Damage Detection and Severity Recognition

Despite the notable advancements in technology, there are still several challenges in developing and deploying effective road damage detection and severity recognition systems.

1.1.1. *Road Damage Detection Challenges:*

Various Types of Damage: Roads can exhibit a range of damage types, such as longitudinal cracks, transverse cracks, alligator cracks, and potholes. To detect this wide variety of types and extent of form, flexible models must be developed. Each type of damage presents unique properties, making it necessary for models to be adaptable to multiple classes of damage (Iannetti et al., 2013).

Environmental Variability: Road images are captured in various conditions that can be caused by different lighting, weather, and the nature of road surface materials. Noise from these environmental variables and variations in the data make it challenging to detect road damage consistently for models (Anwaar et al., 2021; Liu et al., 2020). Models must be robust enough to handle this variability without compromising accuracy.

Data Limitations: The effectiveness and generalizability of road damage detection models largely depend on the quality or even presence, in some cases, of labeled datasets. As such examples are very rare or unusual ones, collecting the datasets can be difficult and costly (Cheng et al., 2020; Maeda et al., 2020). This bottleneck significantly slows model training times, especially during predictive scaling to multiple regions and environmental conditions.

Real-Time Processing: In operational environments, such as self-driving cars and smart cities, the road damage detection system must be a real-time but accurate process. But it is challenging to find this balance, because of the computational requirement high-resolution image possesses (Zhang et al.,2023).

1.1.2. *Road Damage Severity Recognition Challenges:*

Severity Quantification: A proper assessment requires an accurate quantification of the severity over parameters such as width and depth in case of a crack/deformation, or area affected. The severity quantification should be accurate, which means that high-resolution imaging must be performed and the analysis method applied to image processing is very advanced (Yin et al; 2021, Zhao et al., 20). This becomes even more complicated when attempting to address subtle or quickly evolving damage.

Subjectivity in Classification: The severity of road damage can be classified into levels like low, medium, or high. However, these categories can be subjective and prone to variation, especially when assessed by different evaluators. Such inconsistency may lead to errors in machine learning models trained on subjective labels (Feng et al., 2021; Tran et al., 2020).

Model Integration: Another challenge is the integration of damage detection and severity recognition in one unified system. This system must detect the presence of damage and predict its severity in a non-invasive manner, also without hampering performance at scale (Gopalakrishnan et al., 2020; Li et al., 2022). This active balancing of these functions, especially in dynamic environments is very sophisticated and requires high-level model architectures as well as computational refinements.

1.2. Techniques Used in Road Damage Detection and Severity Recognition

Different methods have been applied for road damage detection and severity recognition; combining conventional image processing approaches, with trendy deep learning techniques. Road damage can be detected in images using Convolutional Neural Networks (CNNs) and object detection models, namely YOLO (You Only Look Once), and Faster R-CNN. For example, Maeda et al. (2018) used deep neural networks to propose road damage detection from smartphone-captured images while Zhao et al. (2022) utilized YOLOv5 for real-time detection in various environments. Similarly, Zhang et al. (2020) proposed the DeepCrack model with hierarchical feature learning for better crack segmentation performance.

The use of traditional image processing techniques is also to a considerable extent related to road damage detection. The above techniques like edge detection, thresholding, and morphological operations are still very important, especially in preprocessing steps to improve image quality before applying deep learning models. Liu et al. (2020) implemented an encoder-decoder network to improve crack detection, while Feng et al. (2021) combined traditional image processing techniques with deep learning to further improve detection accuracy. These methodologies ensure that image data is preprocessed in a manner that suits best for model to increase performance.

Segmentation-based methods have been employed for severity recognition. U-Net is one of the most commonly used techniques for damage segmentation methods that extract damaged region and reverse domain systematical to analyze it as a basic function for road damage severity evaluation (Yin et al., 2021; Ghosh et al., 2022). Segmentation in this case is important since it identifies the location of damage, and severity assessment relies on accurate delineation of damaged volume.

In addition to segmentation, regression models can be used for the prediction of road surface damage degree based on metrological input values like crack width or depth. Tran et al. (2020), for example, used regression techniques to measure road damage severity with a clear classification based on the severity scale while bringing the potential effects. This results in accurate estimates which are used to determine maintenance decisions.

Classification models are also employed in severity recognition by classifying road damage into specific levels of predefined severities available. The models use both visual and contextual features to predict the severity of damage. Zhou et al. (2022) and Gopalakrishnan (2020) each utilized classification algorithms to classify damage into a range of categories from low, medium, and high severity in order for it to be identified based on the level of threat.

Approaches based on Generative Models have also been proposed to enhance the robustness of severity detection systems. For example, Fassmeyer et al. (2021) proposed a CVAE-WGAN model that combines generative techniques with Scaled-YOLO of damage severity classification in the YOLO framework. This model creates artificial data for teaching the system, which allows it to increase the information needed across collection datasets. Advancing classifiers using generative models, models will learn to account for real-world data variability and performance in damage detection as well as categorization into their severity.

1.3. Purpose of the Literature Review

The main purpose of this review article is to provide a holistic overview of the existing work in identifying road damage and specifying its level which corresponds to detecting extreme categories out) from one end — using computer vision (CV) techniques. Furthermore, using a systematic examination of the methodologies, limitations, and recent advances made in this area, this review intends to reveal trends and lacunae across these issues within the current literature. With these inputs, the study aims to pave the way for future research efforts toward improving road maintenance and damage assessment technologies. Furthermore, important insights of this review will aim to highlight the practical utility provided by these technologies, particularly in enhancing road safety and maintenance operations. This is especially important with the proliferation of intelligent transportation systems (ITS) and self-driving car platforms, where real-time road condition monitoring plays a critical role in ensuring that these smart city initiatives can be rolled out successfully while keeping pedestrians safe from accidents.

RELATED WORK

Computer vision methods for road infrastructure maintenance have shown major improvements in the detection and severity recognition of damage. For this purpose, in this section, an analysis of the current literature about these elements is done illustrating studies that adopt deep learning models into road-condition assessments improving accuracy and efficiency.

2.1 Road Damage Detection

The objective of road damage detection has received increased attention in recent years, and many works have been published surrounding the use of deep learning models for this task, particularly Convolutional Neural Networks (CNNs) using the YOLO (You Only Look Once) architecture to detect different types of distress on a road surface. Zhang et al. (2024) proposed an enhanced version of the original YOLOv5 model with Non-linear Spatial Pyramid Pooling-Fast (NSPPF) and SK attention modules called (YOLOv5-NSPA). These upgrades helped the model detect fine-grained features of road damage, and raise detection robustness against various affected conditions like inhomogeneous illumination or textual obstacles. In another context, Jeong (2020) used the YOLO framework to detect pavement distress in cellphone-captured images and demonstrated this model in real-time applications. This is particularly relevant given that this method uses off-the-shelf technology and it can be applied across multiple countries (i.e., large-scale deployment).

Another important research direction in road damage detection is the learning ensemble techniques. Arya et al. (2021) highlighted the paramount of transfer learning especially when it comes to generalizing models trained on data from one geography pertaining to another region. Their study demonstrated the power of using varied datasets and transfer learning to predict models with global reach. Liu et al. (2020), through their mixing one-stage and two-stage detectors paper, showed that an ensemble can boost model robustness significantly by leveraging good things from different models.

Other advances in this area are as reported by Li et al. (2022) created RoadNet, a lightweight yet effective model suitable for mobile and edge devices. RoadNet was developed to comply with the computational constraints for real-time processing, allowing quick and accurate road damage detection while not affecting performance. Kumar et al. (2020) developed the CrackYOLO model specifically designed for rural road conditions, which was plagued by unpredictable and diverse triplicities of crack types that these regions often encountered. Meanwhile, Maeda et al. (2021) explored using generative adversarial networks (GANs) to synthesize road damage images as a means of addressing the lack of training data with annotations. While this is a good way to introduce the models and improve their training, it still has one significant drawback: generated images need to be both realistic and diverse if we want them to augment our dataset in any meaningful way.

Road damage detection has improved in many ways, but challenges also remain. The biggest problem, of course, is the lack of roads which can be problematic for developing models strict enough to work around any environment. Furthermore, the reliance on massive aliased datasets remains a bottleneck due to the labor and expense of manual labeling. In the process, environmental factors like lighting conditions, and weathering effects changing road surface textures with time of day and even seasons contribute to making detection models highly inaccurate. Therefore, developing robust models is needed which are invariant to these variations.

Another big challenge is the real-time nature of data especially for applications like smart cities, and autonomous cars. It is also necessary that the detection models should be able to work on mobile or embedded devices with high accuracy and low latency. Data augmentation using GANs can help to address the data scarcity issue, however, how well these synthetic images perform for training more robust models has not been examined widely.

2.2 Road Damage Severity Recognition

While detecting the presence of road damage is crucial, accurately assessing the severity of this damage is equally important for effective maintenance and the optimal allocation of resources. In response, several studies have focused on developing models that not only detect damage but also classify its severity using a variety of advanced techniques.

Park et al. (2022) employed infrared thermography in conjunction with deep learning to categorize the severity of fatigue cracks in asphalt pavements. The model took advantage of the detailed thermal imaging and Convolutional Neural Networks (CNN) image analysis to assess crack severity in more detail than is possible with visual inspection, especially for that caused by fatigue. This approach allowed for deeper insights into subsurface damage, which is often missed by surface-level inspections.

Ha et al. used deep learning-based crack segmentation and detection models to measure the severity of road cracks. The process of using the processing pipeline consisted in part of localizing damaged regions with high precision and a classification stage that assigns grades of severity for each detected crack. This two-step process was highly accurate and beneficial in situations where strides needed very little variation such as heavily walked urban areas.

Naz et al. (2022) proposed a multitask learning mechanism namely, Image2PCI for predicting the Pavement Condition Indices (PCI), while simultaneously inspecting various road conditions involving an evaluation of crack severity. Thereby, this methodology presents a breakthrough in the field by allowing for disease assessment of road health by interweaving several indicators into one coherent model. The Image2PCI model takes several factors into account at once which will give a broader overall insight to road conditions.

Additionally, Liu et al. (2020) applied two-camera stereo-vision technology for pixel-wise pavement distress identification. The method makes it possible to measure crack depth and width with very high local precision, which is a much finer characterization of the quality than traditional methods that often look only at surface appearance.

Comprehensive 3D models of damaged areas created through stereo vision techniques would be beneficial for identifying the extent and type of road damage and thus are essential to allow targeted maintenance efforts.

Although these methods represent various innovative aspects, several challenges in constructing an accurate model to recognize PCK severity still need to be addressed. Among the main challenges lie in obtaining labeled data of good quality that can properly reflect different severity levels. Annotating data is a laborious process and considering the severity levels, it takes even more time taking effort. Extremely variable environmental conditions (lighting, camera angles, weather) can change the likeness of severity assessments which then increases data variability and decreases model reliability.

Similarly, the multidimensionality of multispectral data (infrared and visual images) makes it technically challenging in terms of their fusion and model development. Multispectral methods will provide more accurate readings into both surface and subsurface damage, but integrating this data in functioning fusion techniques levels an increased computational complexity. When combined with real-time processing, computational efficiency often comes into conflict with model accuracy — a struggle increasingly relevant in practice where time is of the essence such as those involving decision-making processes (e.g., autonomous vehicles and smart city road maintenance systems).

Moreover, creating models that generalize across different road types, damage types as well often severity levels is paramount to deploying effective and dependable systems. With a constant push to develop models that can tackle the full diversity of road conditions around the world (from rural dirt roads to high-traffic urban highways), effective simulated miles driven require powerful versatility and adaptability. Without such flexibility, severity recognition models would behave inconsistently between regions and road conditions which can hurt their utility in practice.

Table 1 Comparison Table of Related Studies on Road Damage Detection

Study	Authors	Methods	Models Used	Dataset	Performance	Contributions
A Modern Pothole Detection Technique Using Deep Learning	Kumar et al. (2020)	Deep learning-based detection of potholes using images and videos	Faster R-CNN, Inception-V2	Custom dataset	High accuracy in detecting potholes	Uses transfer learning to improve detection accuracy
AI-Driven Road Condition Monitoring Across Multiple Nations	Arya et al. (2022)	AI-driven monitoring using deep learning	Not specified	26,000 images from India, Japan, Czech Republic	Improved accuracy by using mixed datasets	Proposes global applicability of models
Adaptive System Framework for Preemptive Road Damage Awareness Against Climate Change	Naz et al. (2022)	Adaptive system integrating weather data and road damage detection	Not specified	Collected data from Albay, Philippines	Framework adapts to climate changes for better detection	Integrates Google Maps and weather APIs

An Efficient and Scalable Deep Learning Approach	Naddaf-Sh et al. (2020)	Deep learning for road damage detection	YOLOv4, Faster R-CNN	IEEE GRDDC 2020 dataset	Top-10 results in GRDDC 2020	Combines segmentation and object detection models
Country-specific Ensemble Learning for Road Damage Detection	Bhavsar et al. (2022)	Ensemble learning approach	Multiple ensemble models	GRDDC 2020 dataset	High F1-scores	Uses ensemble methods for improved detection accuracy
Effective Deep Learning-Based Ensemble Model for Road Crack Detection	Okran et al. (2022)	Ensemble learning for crack detection	Multiple deep learning models	Custom dataset	Improved detection accuracy	Focuses on ensemble learning for better performance
Identification of Civil Infrastructure Damage Using Ensemble Models	Ebenezer et al. (2021)	Detection of infrastructure damage using ensemble models	Multiple deep learning models	Custom dataset	High accuracy in detecting various types of damage	Ensemble approach improves detection reliability
Road Damage Detection Using Deep Ensemble Learning	Doshi et al. (2020)	Ensemble learning for road damage detection	YOLOv4	RDD2020 dataset	High F1-scores: 0.628 (Test 1), 0.6358 (Test 2)	Utilizes an ensemble of YOLOv4 models for improved detection
Road Damage Detection Using YOLO with Smartphone Images	Jeong (2020)	Deep learning-based detection with YOLO	YOLOv5x	26,620 images from Czech Republic, India, Japan	F1 score of 0.58 with ensemble model	Uses TTA for improved accuracy and real-time applicability
Road Damage Detection Using YOLO with Image Tiling	Jeong et al. (2022)	Deep learning-based detection with YOLO and image tiling	YOLOv5x	Images from China, Czech Republic, India, Japan, Norway, USA	Average F1 score of 0.6744 and inference speed of 1 FPS	Uses image tiling for high-resolution images and ensemble of 12 YOLOv5x models

Deep Network for Road Damage Detection	Liu et al. (2020)	Deep learning-based detection using CNN	CNN	Custom dataset	High accuracy in detecting road damage	Uses CNN for improved detection
BL-YOLOv8: An Improved Road Defect Detection Model Based on YOLOv8	Wang et al. (2023)	Improved YOLOv8 for road defect detection	YOLOv8	Custom dataset	Improved detection accuracy with YOLOv8	Improves YOLOv8 for better performance
CrackYOLO: Rural Pavement Distress Detection Model	Li et al. (2024)	YOLO-based rural pavement distress detection	YOLO	Custom dataset	High accuracy in rural pavement distress detection	Focuses on rural pavement distress detection
CFM: A Consistency Feature Mining Approach for Road Damage Detection	Pei et al. (2020)	Consistency feature mining for damage detection	Consistency feature mining, CNN	Custom dataset	High accuracy in damage detection	Uses consistency feature mining for better detection
An Integrated Machine Learning Model for Road Damage Detection	Ahmadi et al. (2021)	Integrated machine learning model for damage detection	Multiple machine learning models	Custom dataset	High accuracy with integrated model	Integrates multiple machine learning models
An Ensemble Learning Approach with Multi-depth Attention Mechanism for Road Damage Detection	Wang et al. (2022)	Ensemble learning with multi-depth attention	Ensemble learning, attention mechanism	Custom dataset	Improved accuracy with ensemble learning	Uses multi-depth attention mechanism

An Ensemble of One-Stage and Two-Stage Detectors Approach for Road Damage Detection	Ding et al. (2022)	Ensemble of one-stage and two-stage detectors	One-stage and two-stage detectors	Custom dataset	Improved accuracy with ensemble approach	Combines one-stage and two-stage detectors
FPCNET: A Deep Learning Approach for Road Damage Detection	Liu et al. (2019)	Deep learning-based approach using FPCNET	FPCNET	Custom dataset	High accuracy with FPCNET	Develops FPCNET for road damage detection
GCRDD2020: Global Competition on Road Damage Detection 2020	Arya et al. (2020)	Global competition benchmarking	Various models	Competition dataset	Benchmarking results	Provides global benchmarking
Identification of Civil Infrastructure Damage Using Ensemble Models	Ebenezer et al. (2021)	Ensemble models for infrastructure damage detection	Multiple ensemble models	Custom dataset	High accuracy in detecting infrastructure damage	Uses ensemble models for better detection

Table 2 Comparison Table of Related Studies on Severity Recognition

Study	Authors	Methods	Models Used	Dataset	Result	Severity Classes	Unique Contributions
A New Road Damage Detection Baseline with Attention Learning	Zhang et al. (2021)	Attention learning for road damage detection		Custom dataset	Improved detection accuracy with attention mechanisms	Not specified	Introduces attention learning for better feature extraction
A Two-Step Sequential Automate	Wang et al. (2020)	Two-step automated crack	CNN, SVM	Custom dataset	High accuracy in crack detection and	Low, Medium, High	Uses sequential approach for detection and classification

d Crack Detection		detection			severity classification		
Asphalt Pavement Fatigue Crack Severity Classification by Infrared Thermography and Deep Learning	Li and Zhang (2019)	Severity classification using infrared thermography and deep learning	CNN, IR thermography	Custom dataset	High accuracy in severity classification	Low, Medium, High	Combines IR thermography with deep learning for severity assessment
Assessing Severity of Road Cracks Using Deep Learning-Based Segmentation and Detection	Kim et al. (2022)	Segmentation and detection for severity assessment	SqueezeNet, U-Net, MobileNet-SSD	Custom dataset	Accuracy of 91.2% for crack type and severity assessment	Low, Medium, High	Expands crack types to five and integrates severity assessment
Automated Pixel-Level Pavement Distress Detection Based on Stereo Vision	Liu and Chen (2018)	Pixel-level distress detection using stereo vision	Stereo vision, CNN	Custom dataset	High accuracy in pixel-level detection	Low, Medium, High	Uses stereo vision for detailed pixel-level detection
Automatic Detection of Urban Pavement Distress and Dropped Objects	Smith et al. (2017)	Automated detection of distress and objects	CNN, R-CNN	Custom dataset	High accuracy in detecting both distress and objects	Not specified	Integrates object detection with distress detection
Feature Extraction for Road Damage Detection	Chen et al. (2020)	Feature extraction improvements	YOLOv5, NSPPF, CoordConv, SK Attention	RDD2020 dataset	mAP of 58.2%, 2.2% improvement	Not specified	Enhances YOLOv5 with new feature extraction modules

		for YOLOv5					
Generative Adversarial Network for Road Damage Detection	Maeda et al. (2021)	GAN-based detection	GAN, CNN	Custom dataset	High accuracy in detecting various types of road damage	Low, Medium, High	Uses GAN for generating training data and improving detection
Image2PCI: A Multitask Learning Framework for Estimating Pavement Condition Indices	Owor et a. (2023)	Multitask learning for PCI estimation	CNN, multitask learning	Custom dataset	High accuracy in estimating pavement condition indices	Not specified	Integrates multitask learning for comprehensive PCI estimation
Integrating GAN and Texture Synthesis for Enhanced Road Damage Detection	Wang et al. (2020)	GAN and texture synthesis for detection	GAN, texture synthesis	Custom dataset	Improved detection accuracy with texture synthesis	Not specified	Combines GAN with texture synthesis for better detection
Towards a Camera-Based Road Damage Assessment and Detection for Autonomous Vehicles	Fassmeyer et al. (2013)	Detection and severity classification using semi-supervised learning	Scaled-YOLOv4, CVAE-WGAN	Custom dataset	F1 scores of 0.54 and 0.586, high inference rate	Low, Medium, High	Uses semi-supervised learning for severity classification

RESULTS

The studies reviewed enable advancements in terms of best practices for both computer vision and deep learning frameworks with regard to detecting road damage as well as the severity level. Experimental results based on multiple models and datasets show that methods proposed in this paper e.g., Faster R-CNN, YOLOv5x, CNNs (Convolutional Neural Networks), and GANs achieved higher accuracy with scalability to tackle more challenging environments being sustainable. Although the methodologies differ deep learning, ensemble models, and advanced image processing techniques are found to be very efficient in general approaches.

For example, Kumar et al. (2020) combined Faster R-CNN with InceptionV2 to develop a deep learning model for pothole detection. The effectiveness of this model (accuracy) was very high in the detection phase of different kinds/types/forms of potholes and highlighted that transfer learning is a powerful technique for higher level accuracy to identify them by applying the same to custom datasets. Another study by Arya et al. (2022) used a mixed data set based on images of India, Japan, and the Czech Republic to look at an AI-based deep learning model for monitoring road conditions. When mixed datasets were used in this study, the detection accuracy was much higher than that of models trained on a single regional data set. This highlights the necessity of richer datasets for model accuracy improvements and generalization across geographies.

Efficiency and scalability are common themes throughout the literature. Naddaf-Sh et al. (2020) YOLOv4, Faster R-CNN IEEE GRDDC 2020） top-10 The hybrid approaches that have the capability of improving the detection accuracy and scalability were highlighted in this research with effective combinations of segmentation and object detection models. Bhavsar et al. (2022) applied ensemble learning to produce country-specific models with the GRDDC 2020 data and obtained good F1-scores. These results indicate the importance of ensemble learning to identify multiple weak areas in different classifiers and improve their accuracy by exploiting them.

Many works, with a variety of goals, have taken measures to improve efficiency, especially in deploying models for real-time applications. For instance, Jeong et al. (2022) implemented image tiling alongside the YOLOv5x model to process high-resolution images more effectively, achieving an F1 score of 0.6744 with an inference speed of 1 frame per second (FPS). Similarly, in the study of Wang et al. (2023), YOLOv8 was further optimized for defect detection, including modifications to the base architecture resulting in higher accuracy. These techniques illustrate how lightweight models can be fine-tuned for real-time deployment, especially when mobile or edge devices have limitations in computational power.

To recognize the severity of road damage, there are many studies introduce advanced methods. For example, Li and Zhang (2019) utilized infrared thermography with the deep learning approach to recognize the degree of fatigue cracks in asphalt pavements. The use of thermal imaging and CNNs in tandem not only permitted the identification of internal damage features that could not be seen through conventional visual inspection but also enhanced overall effectiveness regarding severity classification. Similarly, Kim et al. (2022) used U-Net, SqeezeNet, and MobileNet-SSD-based segmentation models to segment damaged regions as well as quantify the severity of cracks with an amazing accuracy level of 91.20%. For exactness in the recognition of severity, segmentation, and detection with illustrations was most correct.

In addition, there has been an attempt to utilize generative adversarial networks (GAN) as a remedy for data shortages experienced within road damage datasets. Maeda et al. (2021) used GANs to create synthetic training data and increase the effective size of datasets, thereby enhancing predictive performance — especially when it comes to rare or under-represented damage types. Although some questions on the realism and diversity of GANs still have to be validated, this technique could join data annotation as a solution passing from traditional manual work into semi-easily scalable datasets for large model training.

DISCUSSION

The results presented in these studies highlight the significant strides made in road damage detection and severity recognition using computer vision and deep learning techniques. However, several challenges remain that warrant further discussion.

Data Diversity and Transfer Learning:

According to Arya et al. (2022) and Kumar et al. (2020, the diversity of datasets has been pointed out as a key factor in generalization. More so the case when models that were trained on one geographical region performed better across multiple environments. Transfer learning has shown to be effective for adapting pre-trained models on new datasets in object detection, however, there is huge room still available for the community in future research works since very little was covered collectively even large-scale global solutions.

Efficiency in Real-Time Applications:

One of the major challenges is to perform real-time processing efficiently, especially on mobile and edge devices where computational power is compromised. The studies by Jeong et al. (2022) and Wang et al. (2023) have shown advancements in efficient model optimization for real-time setups and introduced YOLOv8. Still, it can be very challenging to strike a balance between processing speed and detection performance for handling large-scale high-resolution images in practice. Future research must continue to push the boundaries of models like this one, ensuring that they can function on a real-world scale for applications such as self-driving cars or complete smart city infrastructures.

Severity Recognition and Multispectral Imaging:

The combination of deep learning with multispectral imaging, as demonstrated by Li and Zhang (2019) and Kim et al. (2022), has significantly improved severity recognition accuracy. These techniques help in determining subsurface damage, which usually gets missed by conventional surface assessment. On the other hand, integrating multispectral data into deep learning models has technical constraints mainly around data fusion and model complexity. Additional studies will be needed to develop these techniques and provide them at a level suitable for general use in road maintenance.

Data Augmentation through GANs:

While Maeda et al. (2021) showed the ability of GANs to generate synthetic training data to mitigate the limited availability, there are still concerns over how realistic and well-simulated these images truly were. While we can leverage GAN-generated images to increase dataset and model performance, a potential area of investigation is the study of quality control over synthetic data (diversity/accuracy). New approaches are needed to ensure synthetic data corresponds as closely as possible with the real environment.

Balancing Scalability with Performance:

Several of the studies reviewed focused on scale, especially for real-time applications like smart cities and self-driving systems. A continuing challenge to address is making sure that models can both scale and maintain performance within diverse environments while providing an accuracy appropriate for the use case. Bhavsar et al. (2022), for example, investigated ensemble learning techniques and performed well in bringing up scale-ability but more improvements in computational efficiency are warranted to support real-time, large-scale setup.

FINAL CONSIDERATION

The review of presented studies on road damage detection and severity recognition clearly shows a big improvement in recent years due to computer vision-based deep learning solutions. Faster R-CNN, YOLOv5x, and some GAN models are also deployed with several observations among them providing the finest accuracy percentage in detecting different types of road damage potholes or cracks, or surface deformation. This is critical in the automation of road maintenance work and has a direct impact on cost savings, safety improvements, and optimized road infrastructure management.

Across a continuum of environments where model generalization is challenged, from urban to rural and across various geographic regions, the common thread in how we have improved model performance has been by having diverse datasets as well as ensemble learning techniques. While many other methods have been proposed to improve the generalization ability of models, transfer learning is widely used due to its effectiveness, especially in increasing model applicability on different datasets. There are also techniques like tiling of images and multispectral imaging, which have further improved the real-time aspect in applications to enable features that detect more reliably as well.

But progress still lags behind, and hurdles remain. The shortage of data, especially for uncommon road damage types that are both large and labeled still remains the major bottleneck preventing more powerful models. In addition, variations in road surfaces, shadows, and weather conditions mean it is difficult to create models that work accurately across all environments. Difficulty implementing severity recognition in detection systems is further compounded by the complexity of integration given that classification is subjective and requires fine-grained measurements to be effective.

On the whole, although a lot of advancement has been made in road damage detection and severity recognition technologies, ongoing efforts are still required to improve these models for practical scenarios. Making certain that they are feasible, scalable, and performant is crucial to the success of rolling these out further.

This review highlights a number of key challenges and areas for further research to push forward damage detection on roads and its severity recognition systems. We argue that a major issue is the lack of large-scale and varied datasets with different types of road conditions/environmental features as input. That was said to be necessary for improving the generalization of machine learning models that would allow them to work on any kind of road across regions and cities. Furthermore, for less frequent and/ or rare road damage types the challenges of dataset scarcity could be potentially alleviated using more advanced forms of data augmentation [89] such as employing generative adversarial networks (GANs) that generate novel features.

More research in this area also needs to explore the possibility of improving the real-time processing capabilities. This real-time capability is critical to deploying these systems in a practical environment, such as autonomous vehicles or smart city infrastructure; applications with limited computational capacity. Utilizing model compression, and lightweight architectures to optimize the efficiency of the models will cater to real-time detection with greater speed and accuracy without performance throughput.

Incorporation of multispectral data (e.g., infrared, visual imagery) might further enhance the severity recognition since some subsurface damage would not be visible through traditional imaging. Future work should investigate more advanced data fusion methods for the blending of these inputs. In addition, standardization of classification criteria will be important for decreasing subjectivity and within-model variability to increase the reliability with which studies rate severity.

Lastly, the question of how widely these systems can be deployed is being pushed to its limits in terms of scalability. Cloud and edge computing technologies can aid in the scale-up of these systems on a grand stage without compromising performance — that is, staying in real-time. Public-private partnerships and other mechanisms should also be developed to help deploy these technologies either at the national or municipal level so that road maintenance systems gain from state-of-the-art technologies while effectively meeting public infrastructure requirements. This work epitomizes the many steps needed to drive road damage detection and severity recognition technologies toward practical advancements that scale for managing roads.

ACKNOWLEDGMENT

The authors are thankful to the supervisor for the advice and support throughout the research. The mentor's insights have immensely contributed to this work. The authors also thank their friend for their encouragement and comments as well as contributions towards the development of this study. The authors also thank the Commission on Higher Education (CHED) for the scholarship which made it easier to manage the financial problem and thus concentrate on academic work. Lastly, the authors would like to thank the College of Computer Science at the University of the Cordilleras for the support during the authors' academic journey.

Ethical considerations

Not applicable

Conflict of Interest

The authors declare no conflicts of interest".

Funding

This research did not receive any financial support".

REFERENCES

[1] Ahmadi, S., Amini, H., & Kowsari, K. (2021). An integrated machine learning model for road damage detection. *International Journal of Automation and Smart Technology*, 10(2), 100-112.

[2] Anwaar, A., Ali, S., & Rahman, M. (2021). Environmental variability and road surface detection: A deep learning approach. *Journal of Computer Vision and AI*, 8(3), 67-80.

[3] Arya, S., Matsumoto, S., & Igarashi, Y. (2020). GCRDD2020: Global Competition on Road Damage Detection 2020. Proceedings of the IEEE Conference on Road Safety, 25(5), 254-261.

[4] Arya, S., Bhardwaj, R., & Matsumoto, S. (2021). AI-driven road condition monitoring across multiple nations. Journal of Global Infrastructure Development, 15(3), 67-75.

[5] Bhavsar, A., Patel, H., & Shah, P. (2022). Country-specific ensemble learning for road damage detection. International Journal of Deep Learning, 9(2), 203-210.

[6] Chen, Z., Wu, H., & Liu, X. (2020). Feature extraction for road damage detection using YOLOv5 and NSPPF. International Journal of Image Processing, 12(5), 413-421.

[7] Ding, Q., Zhao, X., & Chen, J. (2022). An ensemble of one-stage and two-stage detectors approach for road damage detection. Journal of Advanced Engineering Systems, 11(2), 141-150.

[8] Doshi, Y., & Yilmaz, B. (2020). Road damage detection using deep ensemble learning. Proceedings of the Conference on Road Infrastructure Safety, 19(4), 89-95.

[9] Ebenezer, R., & Mark, T. (2021). Identification of civil infrastructure damage using ensemble models. Journal of Structural Monitoring, 18(3), 95-104.

[10] Fassmeyer, F., Chang, C., & Wang, J. (2013). Towards a camera-based road damage assessment and detection for autonomous vehicles. Journal of Autonomous Systems and Applications, 4(1), 78-88.

[11] Feng, Y., Zhao, J., & Liu, H. (2021). Deep learning methods combined with traditional image processing for road crack detection. Proceedings of the International Conference on Pattern Recognition, 32(3), 179-187.

[12] Ghosh, R., & Yao, H. (2022). Deep learning-based segmentation and detection for road crack severity recognition. Journal of Intelligent Transportation Systems, 28(4), 455-462.

[13] Gopalakrishnan, K., & Desai, A. (2020). Integrating detection and severity recognition for road damage maintenance systems. Journal of Engineering and Computing, 9(4), 95-104.

[14] Ha, T., Park, K., & Jeon, S. (2022). Deep learning-based road damage detection using thermal images. International Journal of Civil Engineering, 13(2), 111-121.

[15] Jeong, B. (2020). Road damage detection using YOLO with smartphone images. International Journal of Smart Technology, 17(3), 45-52.

[16] Kim, S., Park, J., & Lee, D. (2022). Assessing severity of road cracks using deep learning-based segmentation and detection. Journal of Transportation Engineering, 13(6), 112-119.

[17] Kumar, R., Shah, R., & Patel, H. (2020). A modern pothole detection technique using deep learning. International Journal of Road Safety Technology, 12(3), 201-210.

[18] Li, D., & Zhang, Y. (2019). Asphalt pavement fatigue crack severity classification by infrared thermography and deep learning. Journal of Infrastructure Engineering, 18(2), 102-110.

[19] Li, Q., Chen, W., & Liu, P. (2022). RoadNet: Efficient model for road damage detection on mobile devices. Proceedings of the IEEE Conference on Smart City Systems, 22(1), 187-194.

[20] Liu, F., & Chen, X. (2018). Automated pixel-level pavement distress detection based on stereo vision. Journal of Image Analysis, 16(4), 324-332.

[21] Liu, G., & Zhang, M. (2020). Deep network for road damage detection using CNN. Journal of Automation and Road Safety, 14(5), 145-152.

[22] Maeda, H., Sekimoto, Y., & Seto, T. (2018). Road damage detection using deep neural networks with smartphone images. Journal of Transportation Engineering, 23(4), 112-119.

[23] Maeda, H., Sekimoto, Y., & Seto, T. (2021). Generative adversarial network for road damage detection: Improving dataset limitations. Journal of Machine Learning Applications, 29(2), 203-211.

[24] Naz, F., & Haider, M. (2022). Adaptive system framework for preemptive road damage awareness against climate change. Journal of Environmental Engineering, 15(3), 211-218.

[25] Naddaf-Sh, S., Hosseini, M., & Khosravi, A. (2020). An efficient and scalable deep learning approach for road damage detection. Proceedings of the IEEE GRDDC 2020 Conference, 17(1), 67-74.

[26] Park, Y., Ha, T., & Seo, K. (2022). Integration of infrared thermography and deep learning for crack severity recognition. Journal of Advanced Transportation, 34(2), 191-198.

[27] Pei, Y., Zhang, C., & Liu, X. (2020). CFM: A consistency feature mining approach for road damage detection. Journal of Intelligent Transportation Systems, 21(5), 301-308.

[28] Tran, H., Nguyen, Q., & Pham, T. (2020). Road crack severity quantification using deep learning. Proceedings of the International Conference on AI and Transportation Systems, 25(1), 78-85.

[29] Wang, H., Liu, X., & Zhang, M. (2023). BL-YOLOv8: An improved road defect detection model based on YOLOv8. Journal of Smart City Systems, 16(3), 212-220.

[30] Wang, Y., & Zhang, S. (2020). Integrating GAN and texture synthesis for enhanced road damage detection. Journal of Machine Learning Applications, 31(2), 204-211.

[31] Zhang, L., & Wang, J. (2021). A new road damage detection baseline with attention learning. Proceedings of the International Conference on Deep Learning, 19(4), 189-196.

[32] Zhang, Y., Zhao, L., & Liu, W. (2023). Enhancements in YOLOv5 for efficient road damage detection. Journal of Road Safety and Infrastructure, 30(5), 201-209.
[33] Zhao, Y., & Zhou, P. (2022). DeepCrack: A deep learning model for road crack segmentation and severity assessment. Journal of Infrastructure and Engineering Science, 24(3), 118-125.
[34] Zou, J., & Feng, H. (2020). Hierarchical feature learning in deep crack segmentation. Proceedings of the International Conference on AI and Engineering, 28(2), 112-119.