



Review

Artificial intelligence applications in pavement infrastructure damage detection with automated three-dimensional imaging – A systematic review



Saleh Abu Dabous ^{a,b,*}, Mohamed Ait Gacem ^b, Waleed Zeiada ^{a,b}, Khaled Hamad ^{a,b}, Rami Al-Ruzouq ^a

^a Department of Civil and Environmental Engineering, University of Sharjah, Sharjah, United Arab Emirates

^b Sustainable Civil Infrastructure Systems Research Group, Research Institute of Sciences and Engineering, University of Sharjah, Sharjah, United Arab Emirates

ARTICLE INFO

Keywords:
Pavement
Road
Damage
Artificial intelligence
3D imaging

ABSTRACT

Pavement damage detection is vital in ensuring road-users safety and reducing economic losses related to maintenance and road incidents, thus making it a highly discussed research topic. Non-destructive pavement damage inspection methods have been rigorously explored, compared and integrated in the road health and safety practices. In particular, 3D imaging stands out due to its ability to extract the depth and geometric characteristics of the detected defects, thus facilitating performing accurate maintenance and corrective actions. During the past few years, with the rise of the highly accurate and robust artificial intelligence-driven analysis, researchers have increasingly adopted machine learning and deep learning models in pavement damage assessment. This paper systematically and exhaustively reviews the use of artificial intelligence-based 3D automated damage detection of pavement performed using laser scanners, stereo cameras/structure from motion and infrared sensors. It compiled 85 contributions published between 2011 to mid-2024. From which, the adopted artificial intelligence models, utilized 3D data collection hardware, utilized datasets, 3D data preprocessing techniques and application domains were extracted, compared and critically analyzed. The survey ultimately highlights the gaps and potential future research directions. Key findings highlight the need to explore more cost-effective and advanced 3D data collection methods, such as drones, in addition to enhancing 3D data preprocessing techniques to boost AI performance. Furthermore, the lack of comprehensive open 3D datasets is a significant gap that, if addressed, could help standardize research. Moreover, the research highlights the need for expanding the use of AI models to cover a wider range of pavement deformities.

1. Introduction

Roadway structures are prone to deterioration with time due to extensive use and exposure to environmental elements. Such deterioration can serve as a catalyst for roadway incidents, thus endangering the safety of its users [1], [2]. It has been particularly noted by [3], [4], [5], [6] that inadequate road surface conditions contribute to more road accidents with higher severity. For instance, it is estimated that in India alone, over 3597 road fatalities are linked to the presence of pavement potholes on yearly basis [7]. Moreover, road defects result in significant economic losses if not addressed, where it was estimated by the American Society of Civil Engineers (ASCE) that road defects cost the US economy \$ 67 billion per year [8].

The structural integrity of pavements can be influenced by several variables including, temperature, humidity, weather-induced

deterioration, and applied loads. The acquisition and quantification of geometric data pertaining to pavement damage is essential in facilitating the formulation of appropriate strategies for asphalt maintenance interventions. Moreover, the early detection of pavement deformities becomes vital during the early stages, since road surface deteriorates with a rate of 40 % during the first 3 quarters of its life. Such rate is accelerated when deformities are not detected early on, due to the penetration of water, which serves as an accelerant for further deterioration [9]. Traditional road defect detection methodologies rely on manual data collection by technicians, which requires substantial working hours to achieve basic approximation of road damage. Furthermore, the visual identification of road defects incurs notable costs and it is inherently prone to human error. Other traditional inspection methods include accelerometric and gyroscopic sensing [10], [11], [12], [13], which estimate the road surface condition by tracking the vibration of the

* Corresponding author at: Department of Civil and Environmental Engineering, University of Sharjah, Sharjah, United Arab Emirates.
E-mail address: sabudabous@sharjah.ac.ae (S. Abu Dabous).

inspection vehicle, as it goes through the road surface under inspection. While generating a fast assessment, such methods often provide shallow information regarding the detected defects. In addition, certain vibration patterns associated with road debris, inspection vehicle condition and passing over bridge joints can be mistakenly interpreted as a road defect. Furthermore, the geometric shape of certain road defects doesn't cause any noticeable vibration pattern, when interacting with the inspection vehicle, which results in leaving them unnoticeable until reaching an advanced stage. In addition, the inspection vehicle often requires high maintenance due to the constant physical interaction with certain road defects like potholes.

In recent years, there has been a rising need for more efficient and accurate approaches to perform road defect detection and assessment. Among the prevalent approaches, 3D imaging based automated surveys stand out, since they offer the capability to generate highly accurate quantitative and qualitative analysis of pavement conditions. 3D imaging is a form of non-destructive inspection, which generates images with depth estimation or point clouds that represent the geometric shape and elevation of an object. Pavement condition monitoring using 3D imaging has been first introduced in the literature in 1997 [14], [15], and ever since several 3D imaging techniques were developed like laser scanning, photogrammetry and stereovision [16]. Pavement damages are inherently three dimensional in nature. Hence, depth information is vital for estimating the strength of delamination and the possibility of future deterioration [17], thus making the accuracy of 3D imaging systems superior to traditional 2D imaging. Moreover, automated deformity detection using traditional image processing techniques faces several challenges, since such techniques often consist of several adjustable parameters, thus making the optimal parametric selection ununiform under different conditions [18].

Using artificial intelligence in 3D pavement damage assessment has gained a significant interest in the research community during recent years. This is due to artificial intelligence ability to assimilate large data, then perform accurate predictions with minimum bias. In addition, learning-based algorithms are highly adaptive, since their performance can improve with time as new data are present. Such techniques are particularly useful for analyzing 3D images since the latter often suffer from high sparsity and data variability originating from non-uniform sampling, relative pose and sensors range [19]. Hence, automatic feature extraction and pattern recognition becomes essential for overcoming such limitations. Traditional artificial intelligence algorithms like machine learning models (random forest, SVM, k-means, etc.) require a feature extraction stage prior to performing the analysis, while more advanced deep learning algorithms (ANN, CNN, etc.) perform autonomous feature extraction, thus making them more accurate and less complicated for adoption [20].

In order to further explore the potential of using artificial intelligence-based analysis in evaluating pavement damages using 3D images, a comprehensive and rigorous discussion regarding current research trends and obtained results is critical for promoting further research in this field. Researchers have previously reviewed pavement damage assessment performed using mobile laser scanning (MLS) [21], besides two-dimensional (2D) imaging techniques [20] [22] [23] [24] [25] [26] [27] [28] [29]. Such contributions offered limited discussion regarding the utilization of 3D imaging techniques. Other reviews discussed the hardware aspects of 3D imaging-based pavement damage assessment, without comprehensively discussing the utilization of artificial intelligence-based analysis [14], [16] [30]. Therefore, this paper presents a comprehensive and exhaustive systematic review that discusses artificial intelligence models used in the evaluation of several types of pavement damages. Such models are critically compared based on computational cost, training efficiency, noise robustness and performance considerations. The review paper also discusses the utilized 3D data acquisition systems, their respective resolutions and the type of datasets used for A.I. models training. In addition, the paper presents an overview of the utilized 3D data pre-processing techniques and their

respective pros and cons. Furthermore, the obtained performances of different artificial intelligence models, related to five distinctive pavement damage assessment application domains were reported.

The presented paper is organized as follows: chapter 2 presents an overview of the planning, execution and discussion activities involved in structuring the presented review paper. Chapter 3 presents an overview of the obtained survey results. This includes general survey statistics, which highlight the yearly and geographical distribution of the contributions, in addition to research trends based on utilized artificial intelligence models, and type of pavement deformity. Chapter 3 also critically discusses and presents the merits and limitations of the utilized artificial intelligence architectures, 3D data collection hardware and 3D data pre-processing techniques. Eventually, the chapter presents the application domains of pavement damage assessment and the performances achieved using different artificial intelligence algorithms. The chapter concludes by presenting an overview of the observed research gaps and possible new research directions, and finally chapter 4 presents the conclusion.

2. Review methodology

2.1. Systematic review protocol

Systematic reviews are widely conducted in scientific research to highlight, interpret and critically discuss contributions in a specific research area, based on a well-defined review strategy that ensures comprehensive assessment and unbiased evaluation. The presented review paper systematically and exhaustively analyzes the use of machine learning and deep learning in conjunction with 3D imaging in the field of pavement damage detection in the period between 2011 and 2022. The review methodology consisted of three stages namely, planning, execution and discussion as shown in Fig. 1. During the planning stage, the review scope and objectives were specified, the search databases and search keywords were selected, the inclusion/exclusion criteria were defined and a quality matrix was developed to categorize the papers based on their relevance. Following that, during the execution phase, the actual search process was performed, which included: Searching the previously defined databases using the pre-selected search terms, exporting all the search results, then merging all the resulting papers and deleting repeated entries, excluding the irrelevant studies based on the exclusion criteria, applying the quality matrix on the papers and extracting the relevant information from the papers. Throughout the execution stage, new search terms were added to make the review process more comprehensive. In addition, forward/reverse snowballing techniques were used to discover additional relevant contributions and new search terms, by investigating the references and citation of each candidate paper. Finally, during the discussion stage, the merits of each contribution were highlighted as well as their limitations and possible future research directions.

2.2. Research questions

The main goal of this study is to identify and critically analyze contributions in the field of machine learning and deep learning-based pavement damage assessment using 3D imaging techniques. To achieve that, the following research questions are set: (1) What are the different machine learning and deep learning algorithms used for 3D imaging-based pavement damage assessment? (2) What are the practical applications of machine learning and deep learning in the 3D imaging of pavement damage? (3) What are the datasets, data collection hardware and data preprocessing techniques used in the artificial intelligence-based 3D imaging of pavement damage? (4) What are the merits and limitations of each artificial intelligence method in relation to the observed pavement distresses?

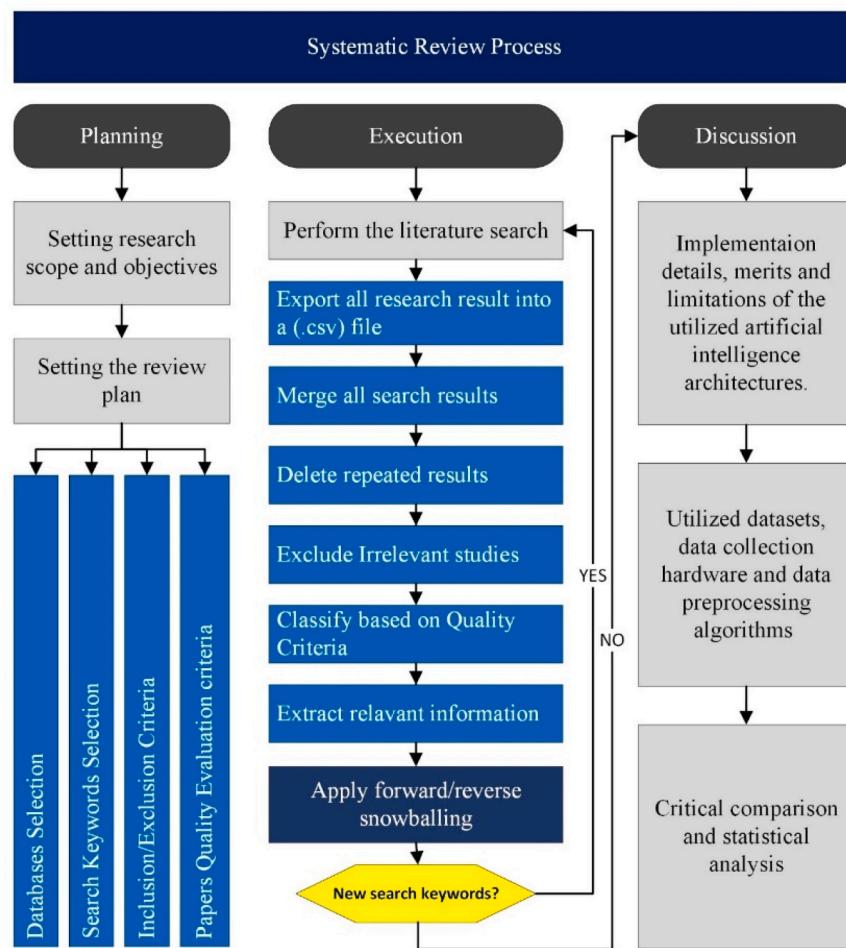


Fig. 1. Systematic review activities flow chart.

2.3. Search terms and databases

The search keywords listed in Fig. 2, were extracted from the previously set research questions, and new terms were added during the execution stage of the systematic review process as described in Fig. 1. In addition, different synonyms and pronunciations were considered to ensure a more inclusive search process. The selected search terms were: (“asphalt” OR “tarmac” OR “pavement”) AND (“crack” OR

“delamination” OR “damage” OR “pothole” OR “rut*” OR “friction” OR “rough*” OR “Manhole*” OR “ravel*” OR “IRI” OR “patch”) AND (“inspect” OR “inspection” OR “monitor” OR “monitoring” OR “evaluate” OR “evaluation”) AND (“artificial intelligence” OR “AI” OR “A.I.” OR “neural” OR “ANN” OR “CNN” OR “LSTM” OR “machine” OR “ML” OR “deep” OR “learning” OR “IoT” OR “internet of things” OR “Internet-of-things” OR “image-processing” OR “image processing” OR “computer vision” OR “computer-vision”) AND

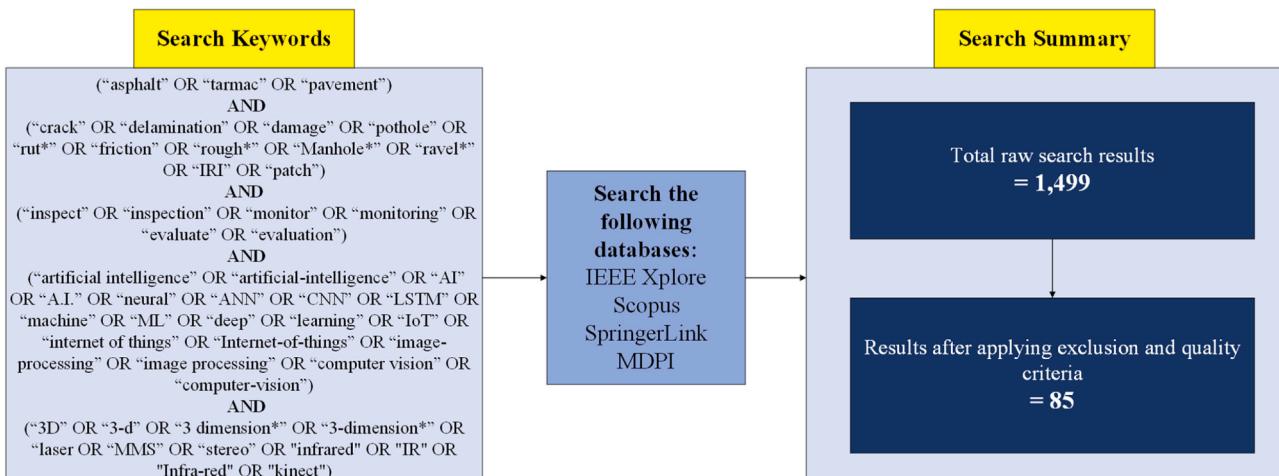


Fig. 2. Review process search keywords, databases and obtained contributions.

("3D" OR "3-d" OR "3 dimension*" OR "3-dimension*" OR "laser" OR "MMS" OR "stereo" OR "infrared" OR "IR" OR "Infrared" OR "kinect"). The search terms were then used to perform the review search in four widely used databases namely, IEEE Xplore, Scopus, MDPI and SpringerLink.

2.4. Inclusion/exclusion criteria

In the process of systematically evaluating the use of machine learning and deep learning in the 3D imaging of pavement damage, the following inclusion criteria were set: (1) The study evaluates pavement damage (2) The study uses a form of artificial intelligence (3) The study uses 3D imaging (or depth extraction) obtained through laser scanning, stereo imaging or infrared scanning. The exclusion criteria included: (1) Papers that do not specifically evaluate a form of pavement damage (e.g., general pavement extraction without damage assessment). (2) Papers that do not use a form of artificial intelligence (e.g., image processing and filtering techniques). (3) Papers that perform 3D imaging in methods other than laser scanning, stereo imaging or infrared scanning.

2.5. Papers quality evaluation

The papers obtained from the literature search were filtered based on their relevance to the review scope, connection to the stated inclusion criteria and overall quality and significance of work. The adopted scoring system involved granting a 100 score for papers strongly related to the scope, 50 for papers having minor relevance and 0 to non-relevant papers. Moreover, the relevant papers were further classified into high quality (score 100), medium quality (score 50), below average (score 25) and low quality (score 10), where quality refers to the clarity of presenting the methodology and the obtained results. The overall score of each paper was determined by calculating the summation of all scores, where papers having a score below 60 are excluded from the review, which corresponds to papers having no relevance to the review scope. Among the 1450 papers obtained from the raw search process described in the execution process stated in Fig. 1, only 85 papers qualified to be included in this review based on the inclusion/exclusion criteria and quality check evaluation.

2.6. Data extraction and synthesis

The data extraction process involved inspecting the type of artificial intelligence algorithm adopted by the paper, type of 3D imaging technique, type of investigated pavement damage, obtained results, type of dataset, resolution of the data collection hardware and data pre-processing techniques. The performance evaluation indices of the adopted artificial intelligence algorithms varied among different contributions. Hence, the equations and description related to each performance evaluation index stated in this paper are summarized in Table 1.

3. Overview of survey results

3.1. General review statistics

In the conducted review, a total of 85 research papers were studied, in the period between 2011 to mid-2024. It was observed that there has been a significantly increasing research focus on the topic of artificial intelligence-based pavement damage assessment using 3D imaging during the study period, where the number of contributions doubled on a biennial basis in the period between 2018 and 2022, shown in Fig. 3. Although there is a slight decline in 2023, the publication numbers remain strong through mid-2024, reflecting sustained attention and ongoing contributions to the field. Therefore, given the citation trend, it is expected to become a highly discussed research area in the upcoming years.

Table 1

Performance evaluation matrices of the artificial intelligence models stated in this review.

Models Evaluation Parameters	Equation	Description
Accuracy	$Accuracy = \frac{Tc}{Ttotal}$	Accuracy refers to the ratio between the total number of correct predictions (Tc) and total number of testing samples (Ttotal).
Recall	$Recall = \frac{TP}{TP + FN}$	Recall refers to the ratio between the total number of true positive predictions (TP) and the sum of total number of true positive (TP) and false negative (FN) predictions.
Precision	$Precision = \frac{TP}{TP + FP}$	Precision refers to the ratio between the total number of true positive predictions (TP) and the sum of total number of true positive (TP) and false positive (FP) predictions.
F1-score	$F1_score = \frac{2 * Precision * Recall}{Precision + Recall}$	F-score encompasses both precision and recall in its calculations through their harmonic mean. It is useful for datasets with class imbalance.
Coefficient of Determination (R^2)	$R^2 = 1 - \frac{RSS}{TSS}$	Coefficient of determination reflects how far the predictions are from the ground truth. It ranges between 0 and 1, where 0 indicates no similarity and 1 reflects a perfect match. It is calculated by estimating one minus the ratio between sum of squared residuals (RSS) and total sum of squares (TSS).
Root Mean Square Error (RMSE)	$RMSE = \sqrt{\frac{\sum (R - P)^2}{n}}$	Calculated by calculating the square root of average squared differences between real result (R) and predicted results (P). Where n is the total number of observations.
Euclidean error (Mean Square Loss)	$Euclidean\ Loss = \frac{1}{n} \sum (R - P)^2$	It measures the distance between the real result (R) and predicted results (P). Where n is the total number of observations.

Since 2020, most contributions focused on evaluating pavement cracks, followed by texture/friction/raveling related pavement deformities, pavement potholes/manholes and the detection of multiple pavement distresses using the same artificial intelligence model. Moreover, the evaluation of pavement rutting during 2022 was the most significant throughout the review period as shown in Fig. 4.

Convolutional Neural Network (CNN) based configurations were the most commonly adopted artificial intelligence architectures on an annual basis since 2018. Artificial Neural Network (ANN) based configurations started to gain some interest during the past three years. Other machine-learning architectures like Random Forest (RF), Support Vector Machine (SVM), Adaboost and Restricted Boltzmann Machines (RBM) were mainly adopted in the years prior to 2020. The use of Recurrent Neural Network (RNN) was suggested in 2018, but it was not adopted again until 2022 in the form of a Long Short-Term Memory (LSTM) architecture. Genetic Algorithm based architectures started to be adopted in the past two years, where they were used in conjunction with neural networks and DenseNet architectures. Researchers started to explore using Generative Adversarial Networks (GAN) in 2024 as shown in Fig. 5.

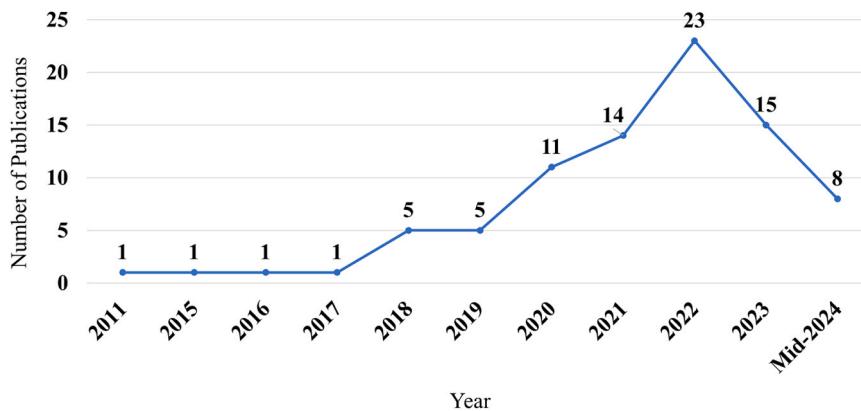


Fig. 3. Annual distribution of the reviewed publications.

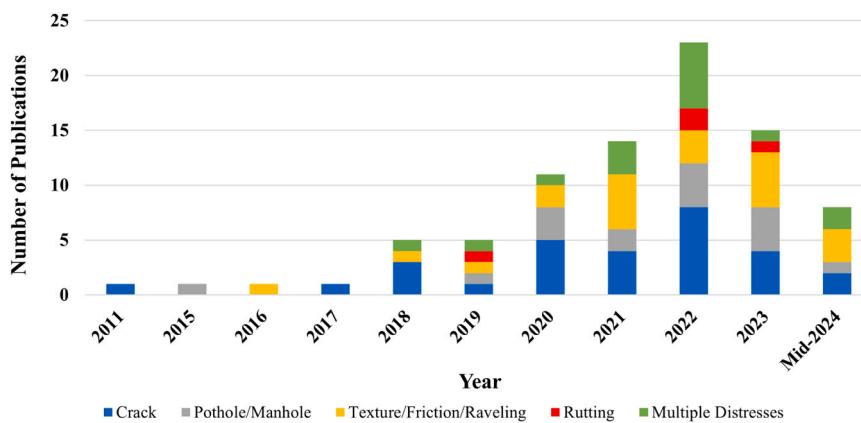


Fig. 4. Annual distribution based on the type of evaluated pavement damage.

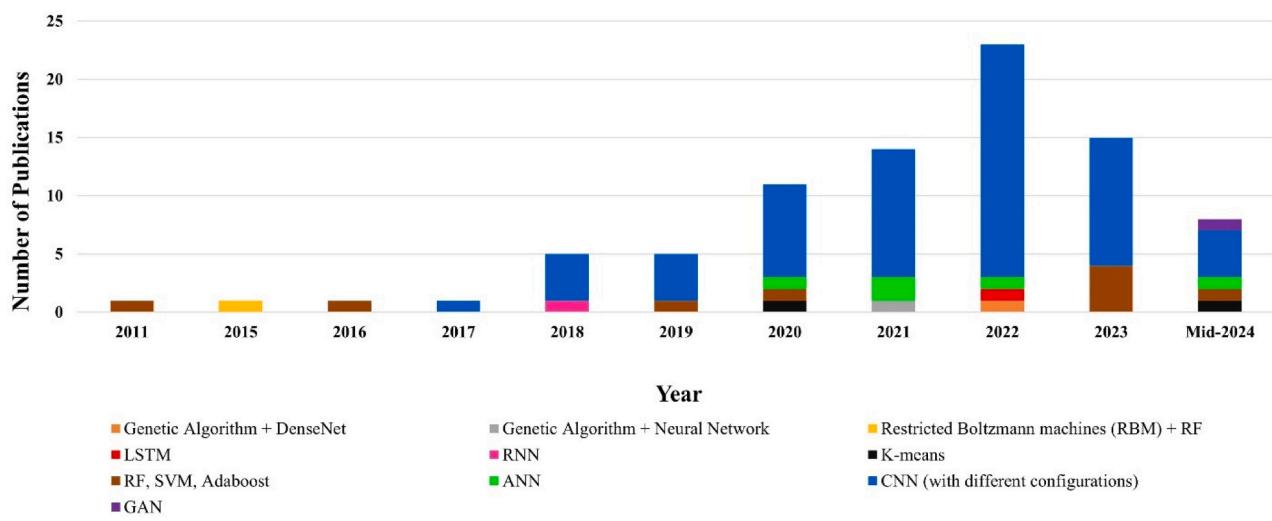


Fig. 5. Annual distribution based on the type of utilized A.I. algorithm.

Most of the contributions stated in the review originated from China (46 %) followed by the U.S.A. (27 %). The remaining contributions originated from: South Korea (4.7 %), New Zealand (2.4 %), Spain (2.4 %), India (2.4 %), Japan (2.4 %), KSA (1.18 %), Egypt (1.18 %), Greece (1.18 %), Netherlands (1.18 %), France (1.18 %), Germany (1.18 %), Thailand (1.18 %), Italy (1.18 %), Iran (1.18 %), China-U.S.A. (1.18 %) and Canada (1.18 %), as shown in Fig. 6. Contributions originating from China and the U.S.A., which jointly account for 74 % of the

contributions, mostly focused on evaluating pavement cracks, followed by the evaluation of texture/friction/raveling related pavement deformities. A less significant interest was shown in using artificial intelligence models in the detection of pavement potholes/manholes, rutting and the evaluation of multiple pavement distresses.

The studied contributions were mainly published in English (91.8 %), followed by Chinese (7.1 %) and German (1.18 %) as shown in Fig. 7(a). In addition, the most common publication types were journal

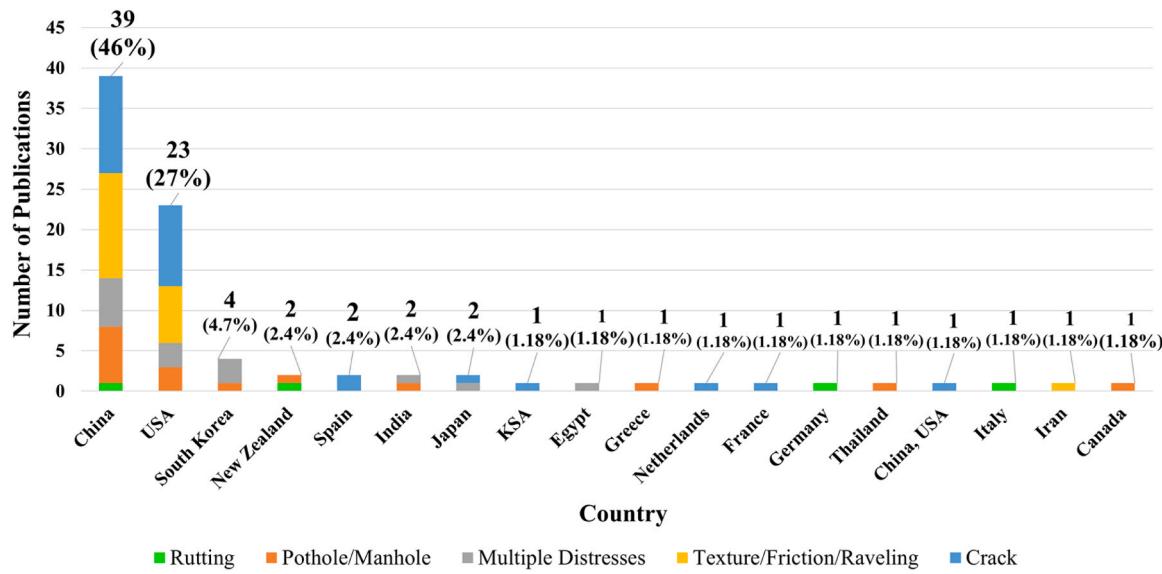


Fig. 6. Distribution based on country and evaluated damage type.

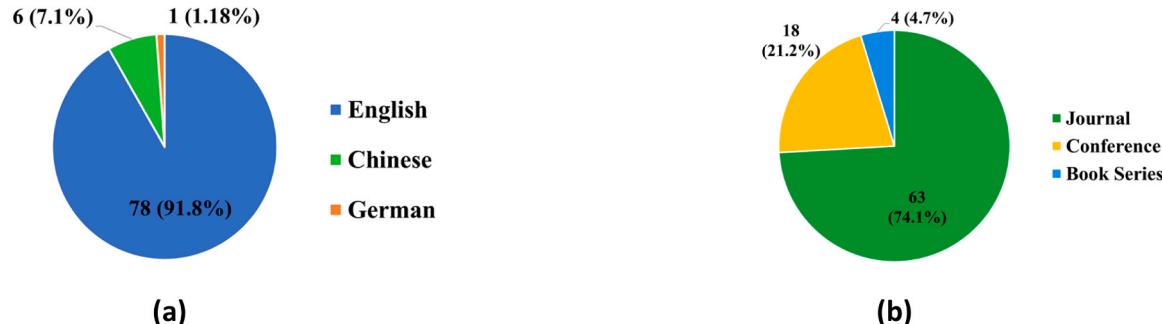


Fig. 7. (a) Contributions distribution based on publication language. (b) Contributions distribution based on the publication type.

articles (74.1 %), conference papers (21.2 %) and book series (4.7 %) as shown in Fig. 7(b). The most common publishers of the studied contributions are stated in Fig. 8.

3.2. Pavement 3D data acquisition and pre-processing

3.2.1. 3D data acquisition hardware

The process of 3D reconstruction of an object or a scene involves the acquisition of its depth, geometric characteristics and calibrated

distances. 3D imaging techniques of pavement damages include laser scanning, photogrammetry and stereovision [16]. Such systems are often mounted on dynamic inspection mediums like road vehicles, along with inertial and positioning systems like accelerometers, odometers and Global Positioning Systems (GPS) to keep a location-based record of the detected defects. In the conducted survey, the most common 3D data acquisition technique was based on a laser scanner (53 %). The remaining techniques included: stereo camera (15.3 %), single digital camera (9.4 %), laser + digital camera (8.2 %), digital camera + depth

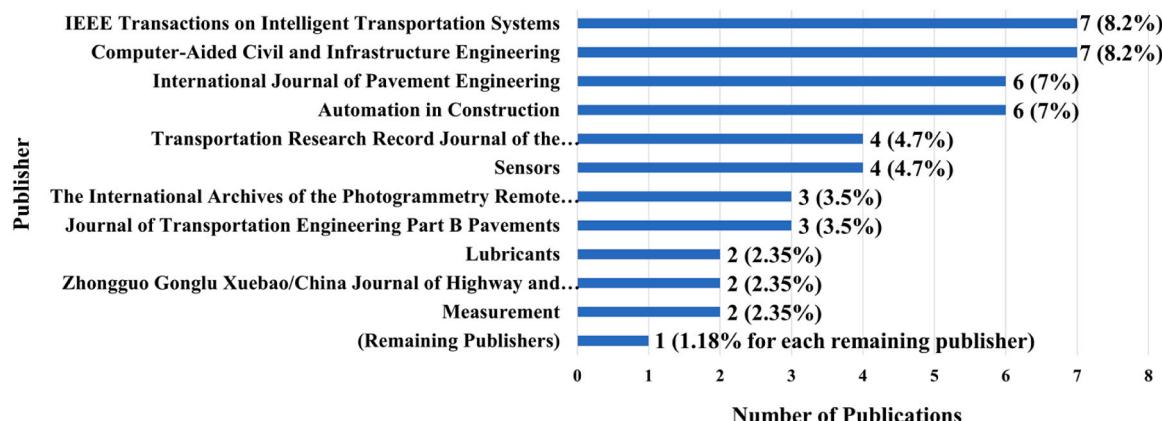


Fig. 8. Distribution of the contributions based on the publisher.

camera (5.88 %), digital camera + infrared camera (5.88 %) and laser, RGB-depth, stereo (2.35 %) as illustrated in Fig. 9. The contributions related to each 3D data acquisition techniques as well as their merits and limitations are stated in Table 2. Overall, the analyzed data in 90 % of the studied contributions were self-collected by the authors specifically for their own research work as illustrated in Fig. 10. External datasets were mostly related to 2D images, and they were used in transfer learning processes prior to evaluating the self-collected 3D pavement data. In addition, such datasets were mostly provided by government agencies and do not exist in an open format. The only open dataset cited by the contributions was RDD2020 pavement damages dataset [31], which includes various types of road distresses, such as pavement cracks and potholes, collected from three different countries: Japan, India, and the Czech Republic. Nevertheless, such dataset relied on vehicle mounted 2D smartphone camera and lacked depth information. Therefore, several researchers commonly pointed out a drawback in the datasets related to 3D pavement damage assessment, resembled in the lack of a comprehensive uniform 3D road damage open dataset in the literature [32]. Thus, making it challenging to benchmark and compare the performance of different assessment methodologies.

3D imaging can be performed by capturing 2D images then converting them into 3D spaces, like in stereo imaging and photometric stereo imaging. In stereo imaging, a pair of cameras determine the absolute 3D position of a specific point, through capturing then combining two-dimensional images. The calibration of such system requires the knowledge of lens and image characteristics. Nevertheless, the main advantage of stereo imaging is not requiring expensive equipment like in laser scanning applications, and being invariant to inspection vehicle vibration [33]. During the stereo imaging process, numerous 3D points coordinates are created using several stereo correspondence methods as described by [34]. Alternatively, stereo imaging is performed by structure from motion (SFM) technique, which creates the 3D scene by utilizing a single camera moving around the desired scene. As for photometric stereo imaging, the observed surface is illuminated from multiple light sources at a time. A camera captures 2D images from each illumination angle, then the collective 2D images are combined to form a 3D space. However, such method requires having all components stationary, thus preventing its use in dynamic long distance pavement assessment. Alternatively, flash LIDAR is used for dynamic imaging, where a short light pulse is sent and received by the camera, then by calculating the time of flight, the distance between the camera and observed surface can be evaluated in a single lighting exposure. In general, the transformation of 2D images into 3D images inherently faces several limitations resembled in being affected by occlusion and lighting conditions. Another limitation that faces photogrammetric based depth extraction is that it requires significant computational

resources to perform the points matching procedure from different viewpoints. In addition, the use of multiple cameras to obtain view variation, or dynamically move one camera to get different view perspectives, results in added costs and makes the system oversized, thus restricting its use for some applications [35].

Laser based 3D imaging is performed by transmitting a laser beam into the observed surface, then observing the reflected beam using a sensor placed at a known distance from the transmission source. Then, point depth and localization measurements can be made by utilizing the trigonometric triangulation [15] and assessing the laser beam's reflection angle. However, the use of such method can be limited by the high cost of equipment and the long-term maintenance costs of the inspection vehicle [36]. A common 3D laser imaging system is Mobile Mapping System (MMS), which primarily consists of a LiDAR laser scanning besides other inertial and distance measurement units. MMS is often mounted on an inspection vehicle, then used to perform dynamic profiling of pavement sections. The collected data are recorded in the form of point clouds having a dedicated coordinate in the 3D reconstructed mapping scene. The main limitation of laser-based 3D scanning systems is that the obtained data lack the spatial information, hence necessitating the use of complementary positioning systems [35].

Another 3D imaging technique is using an Infrared (IR) camera to estimate the pixel depth, by utilizing a similar methodology to the one previously described for the flash LIDAR. Therefore, the collective responses of pixels depths can be translated into a 3D space. Such process can be performed using low-cost Kinect sensor modules [37], which demonstrated their ability to detect different pavement cracking patterns including transverse, longitudinal and alligator cracks [38]. Nevertheless, such method suffers from IR saturation in the presence of direct sunlight [39]. Hence, it can be less reliable in the inspection of long pavement sections during daytime.

3.2.2. 3D data pre-processing techniques

The analysis of pavement damage 3D images using A.I. often requires the enhancement of the deformed regions and the correction of background pixels. In addition, the processing of 3D data can be computationally expensive. Hence, researchers have adopted data processing techniques to reduce the data size, thus perform faster evaluation. Such techniques often result in more accurate predictions, due to mitigating distortion. In the conducted survey, 68.85 % of the contributions stated using at least one form of 3D data pre-processing techniques. The most frequently used techniques are: data downsizing, data augmentation, patching technique, road surface extraction, 3D points height/intensity enhancement, noise reduction, 3D point cloud to image conversion, pavement slope image rectification, depth extraction, filling missing point-cloud points and clustering points to facilitate annotation as

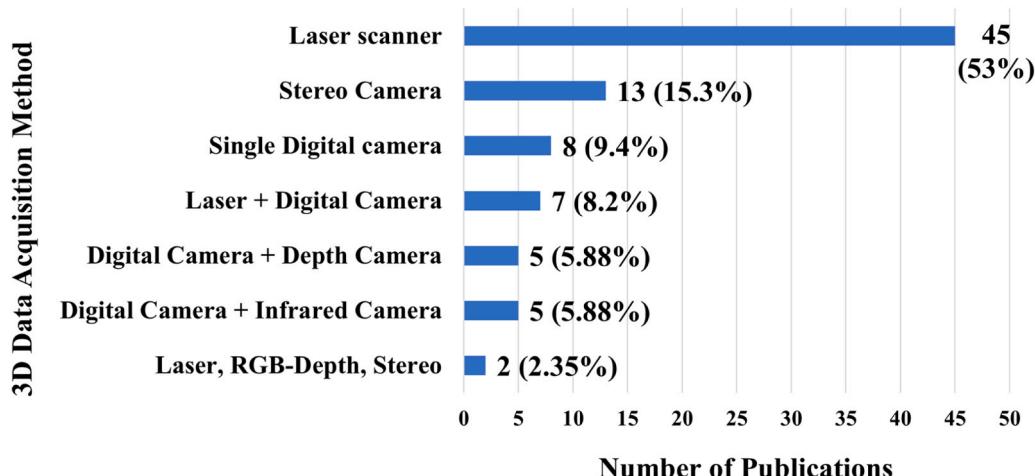


Fig. 9. Distribution based on the 3D data acquisition hardware.

Table 2

3D data acquisition techniques, their merits and limitations.

3D Data Acquisition Hardware	# Of Contributions (Total = 85)	Merits	Limitations	System Resolution	References
Laser Scanner	45 (53 %)	<ul style="list-style-type: none"> - Accurate and fast scanning. - Not affected by temperature saturations. 	<ul style="list-style-type: none"> - Expensive equipment. - Inspection vehicle requires constant maintenance. - Affected by the presence of water and dust on the pavement surface. 	<ul style="list-style-type: none"> < = 1 mm 1 mm transverse, 5 mm longitudinal 0.05 mm transverse, 0.01 mm longitudinal 0.05 mm horizontal, 0.01 mm vertical 5 mm 0.05 mm 1000 × 1600 3 mm 	[40] [41] [42] [43] [44] [18] [45] [46] [47] [48] [49] [50] [51] [52] [53] [54] [9] [55] [56] [57] [58] [59] [60] [61] [62] [63] [64] [65] [66] [67] [68] [69] [70] [71] [72] [73] [74] [75] [76] [77] [78] [79] [80] [81] [82]
Stereo Camera	13 (15.3 %)	<ul style="list-style-type: none"> - Less expensive than laser scanners. - Generates accurate depth estimation from RGB images. 	<ul style="list-style-type: none"> - Affected by lighting conditions and occlusion. - Requires calibration to match between the cameras. 	<ul style="list-style-type: none"> Not mentioned < = 1.2 mm 1920 × 1200 800 × 1312 4000 × 3000 1280 × 960 1920 × 1080 3840 × 2160 1730 × 1028 256 × 256 Multiple resolutions 	[83] [84] [35] [85] [86] [87] [88] [89] [90] [91] [92] [7]
Single Digital camera	8 (9.4 %)	<ul style="list-style-type: none"> - Low cost and compact system. 	<ul style="list-style-type: none"> Prone to mechanical failure since some configurations require moving the camera. - Affected by lighting conditions and occlusion. 	<ul style="list-style-type: none"> Not mentioned 0.25 mm 302 × 302 5472 × 3648 720 × 720 640 × 480 512 × 512 < = 1 mm Min. (2484 × 1739) Max. (6048 × 4231) 250 × 350 1000 × 855 	[93] [94] [95] [96] [97] [98] [99] [100] [101] [102,62,71] [103,50]
Laser + Digital Camera	7 (8.2 %)	<ul style="list-style-type: none"> - Reduced depth estimation errors. 	<ul style="list-style-type: none"> - High cost and bulky size. 	<ul style="list-style-type: none"> 250 × 350 1000 × 855 	[104] [105,43] [106], [107]
Digital Camera + Depth Camera	5 (5.88 %)	<ul style="list-style-type: none"> - Reduced depth estimation errors. 	<ul style="list-style-type: none"> - The system tends to be bulky due to the use of multiple cameras. 	<ul style="list-style-type: none"> Not mentioned 3700 × 10,000 3024 × 3024 2048 × 2048 	[108] [19] [109] [110] [111]
Digital Camera + Infrared Camera/ Infrared Camera	5 (5.88 %)	<ul style="list-style-type: none"> - Reduced depth estimation errors. 	<ul style="list-style-type: none"> - Prone to infrared saturation in the presence of direct sunlight. 	<ul style="list-style-type: none"> 640 × 480 255 × 360 	[112] [113] [114] [115] [116]
Laser, RGB-Depth, Stereo	2 (2.35 %)	<ul style="list-style-type: none"> - Reduced depth estimation errors. 	<ul style="list-style-type: none"> - High cost and bulky size. 	<ul style="list-style-type: none"> Not mentioned 640 × 480 	[117] [118]

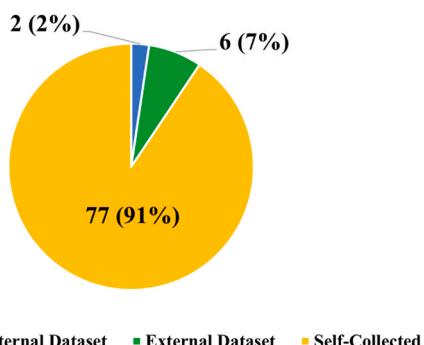


Fig. 10. Contributions distribution based on the dataset type.

illustrated in Fig. 11. The contributions related to each 3D data pre-processing technique as well as their merits and limitations are listed in Table 3.

The processing of large 3D images can be computationally expensive

and result in extended training time. Therefore, researchers often avoid using the raw 3D data in the training process, and instead use data downsizing operations to have a lighter dataset. One approach to achieve that is by using min-pooling technique, which reduces the data size by keeping only the minimum value in a certain data cluster. It was observed that such method makes the fine pavement cracks more visible and detectable in the downsized 3D images [55]. Another downsizing technique used for pavement crack detection enhancement involves assuming that pixels related to cracks would be darker than those surrounding the cracks. Hence, by observing the histogram of a specific area of interest within the 3D image, regions pertaining cracks would demonstrate a peak. Then, that area would be replaced by a single pixel having a value lower than non-crack regions. This results in a 3D image with lower resolution yet with smoother noise and lower overall size [9]. Researchers have also utilized 3D points down sampling methods like the Farthest Point Sampling (FPS), which relies on randomly sampling points, then calculating the Euclidean distance between that random point and all the other points. Eventually, the point having the largest distance is selected to be the starting point of the next sampling [119].

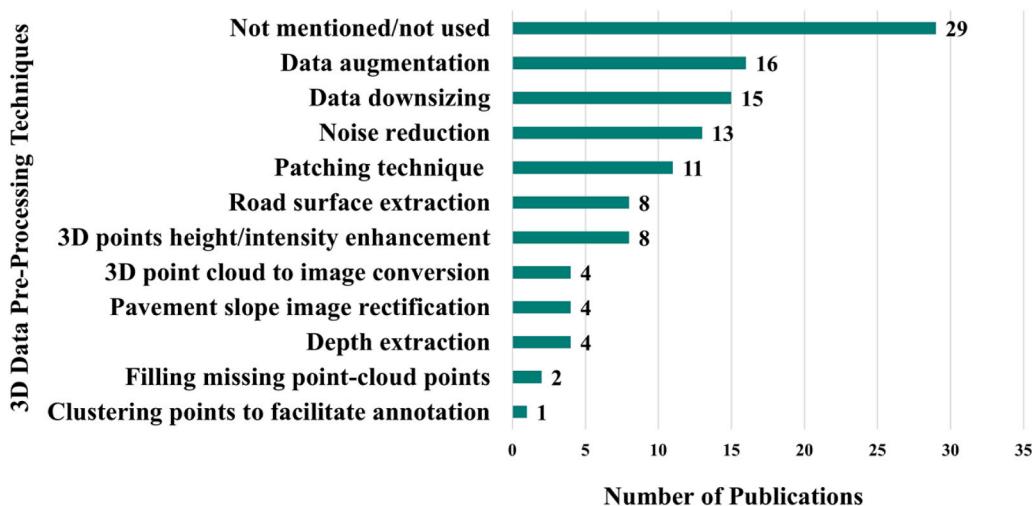


Fig. 11. Distribution of the utilized 3D data processing techniques.

Researchers in [78] used this method to reduce the density of 3D pavement cracks point clouds to only 1 % of the original set.

The performance of the artificial intelligence training models can be enhanced by applying augmentation algorithms to the dataset to increase the number of training samples and enhance the variability of the 3D images. This would ultimately result in a better generalization ability, when used to evaluate new data. The 3D images can be processed by applying rotation, mirroring and cropping into patches [33], [42], [50], [52], [87] [91]. Such techniques are also used to compensate for the class imbalance resulting from having varying numbers of 3D images per classification class, which results in a biased training process. Some processing techniques were aimed to divide the 3D images into patches to facilitate damage classification and patch-level segmentation. Slicing the data into smaller patches and then stitching them together was also used to manage large, high-resolution pavement images [90]. In the studied contributions, the patch sizes were arbitrary chosen by the authors according to the level of accuracy, and computational requirement set by their application [49] [102].

Another form of 3D data processing is separating the point clouds related to the road surface and the off-road surface. This is particularly used for datasets that involve stereo imaging, since they often have a wide field of view that covers both the pavement surface and the side environment [19]. Researchers have achieved road surface segmentation by applying Random Sample Consensus (RANSAC) plane segmentation, which has demonstrated its ability to tolerate the presence of substantial outliers in real-time evaluation scenarios [120]. Another technique for separating the pavement from side walk scenes is by performing pavement curb segmentation. In such method, the point clouds are aligned in vertical fixed width blocks, then by calculating the elevation gradient of the profile, the curb points can be estimated then highlighted as the boundary of the pavement section [77]. The extraction of road surface is also used on the 3D data collected using laser scanner. This is mainly adopted to remove the cloud points related to the pavement curb and off-road scenes, by utilizing a curb-based road surface extraction algorithm, that identifies curb related points by tracking elevation and slope variations [78], [121].

Another 3D data processing approach is 3D points height/intensity enhancement. This can be achieved by using deviation from the mean to extract the feature points from the 3D cloud points, where the elevation of each point from the mean is calculated. The resulting 3D space provides a better representation of the surface roughness. In addition, the intensity value of the points can be further enhanced by calculating the difference between a central point and its neighboring points [82] [69]. Other techniques include evaluating the representative elevation of the 3D surface by utilizing a median filter, then adjusting the elevation of

each pixel based on the local elevation. The elevation ranges can then be standardized by performing Z-score normalization [40]. Researchers have also utilized histogram equalization (HE) to enhance the pavement surface profile contrast. Such method spreads the frequency intensity values in an image to improve the contrast of low-contrast areas. A more advanced form is Contrast Limited Adaptive Histogram Equalization (CLAHE) [122], which performs the same frequency intensity spreading operations, but using multiple histograms of several regions within the same image rather than calculating one histogram for the whole image. In addition, it limits the contrast variation for specific regions, which helps in avoiding the amplification of noise contrast [49]. Gaussian High-Pass filters are also used to rectify the change in the 3D depth values, which occurs in the pixels related to pavement rutting [71]. In addition, depth maps initially reconstructed using Structure-from-Motion (SfM) and Multi-View Stereo (MVS) are in relative depth values, necessitating calibration to achieve real-world accuracy. This is done using a calibration board with known dimensions, placed in the scene; reference points marked on the board help calculate a scale factor that transforms these relative depths into absolute values. This calibration enhances the depth pixels, producing a detailed and accurate depth map of the pavement for evaluating texture depth and skid resistance effectively [110].

The conversion of the 3D data points into 2D depth images facilitates the learning process since most artificial intelligence algorithms are not designed to process raw 3D data points. In addition, it reduces the computational cost and the number of studied parameters. This can be achieved using inverse distance weighted (IDW) interpolation technique, in which the data points get vertically divided into georeferenced grids with known spacing. Then, the data points in each grid get interpolated by weight to form pixels [77]. Other noise reduction pre-processing techniques involve removing 3D point/pixels having invalid depth values, which result from hardware related errors [73] [89]. In addition, filtering techniques like median filtering and double-phase standard deviation filtering have demonstrated their effectiveness in suppressing noise in 3D data related to pavement surface texture [85]. Another pixel noise reduction algorithm is the connected domain component labeling algorithm, which is used to remove irrelevant noise areas from crack images [113]. In image segmentation, cracks are the primary focus, but these images often include noisy regions that need to be eliminated. The connected domain component labeling algorithm, specifically utilizing the eight-adjacency marking method, scans the clustered image to identify and label connected regions. For each cluster, the algorithm calculates the center coordinates and the average distance from the center to the edge of the cluster. It also determines the distance standard deviation to measure clustering

Table 3

3D data pre-processing techniques, their merits and limitations.

3D Data Pre-Processing Techniques	References	Merits	Limitations
Data downsizing	[101] [89] [9] [78] [71] [102] [40] [45] [88] [50] [55] [41] [111] [79] [114]	- Results in faster computation and lower hardware processing requirements.	- Could result in the loss of some critical information.
Data augmentation	[94] [88] [50] [42] [7] [84] [106] [43] [87] [33] [117] [103] [52] [91] [96] [81]	- Compensates for the class imbalance in a dataset. - Improves the generalization ability of the trained models by increasing the variability of the dataset.	- Some augmentation operations like cropping, result in the loss of critical information, which leads to incorrect evaluation.
Patching technique	[61] [62] [71] [44] [102] [49] [42] [109] [18] [46] [90]	- Enables faster processing of multiple patches instead of the full 3D image.	- The accuracy of such method depends on the patch size defined by the system designer.
Road surface extraction	[72] [73] [77] [78] [61] [62] [19] [60]	- Removes unnecessary information from the dataset and makes the training process focus only on pavement sections. - Reduces the computational cost.	- The processing stage can be time consuming. - Some processing techniques require subjective hand-crafted parameters, which require expertise.
3D points/ pixels height/ intensity enhancement	[102] [40] [45] [49] [82] [65] [110] [69]	- Improved deformity detection and localization.	- Some background pixels/points can be mistakenly interpreted as a deformity and get enhanced, thus leading to incorrect predictions.
Noise reduction	[85] [72] [73] [89] [9] [35] [113] [68] [75] [99] [76] [96] [80]	- Reduces the effect of outlier pixels/ points resulting in better deformity detection and localization.	- Some filtering techniques require handcrafted fine tuning.
3D point cloud to image conversion	[101] [77] [71] [44]	- The generated data is more suitable for A.I. models. - Lower computational cost.	- The interpolation process of points into pixels could result in the loss of important information.
Pavement slope image rectification	[72] [73] [89] [110]	- Performance enhancement.	- Some slope and curvature configurations are complex, thus making the techniques less accurate.
Depth extraction	[94] [86] [99] [118]	- Enables depth estimation from 2D images obtained from low-cost cameras rather than expensive laser scanners. - Creates a homogeneous dataset and an improved training process.	- Color based depth estimation is prone to errors due to illumination variations. - Interpolation techniques rely on the surrounding pixels/points to fill the missing data. However, the missing data point
Filling missing point-cloud points	[85] [101]		

Table 3 (continued)

3D Data Pre-Processing Techniques	References	Merits	Limitations
Clustering points to facilitate annotation	[49]	- Faster and easier data labeling.	could be significantly different from their surroundings, thus resulting in less accurate characterization. - Not efficient for labeling fine deformities like cracking.

consistency. Noise removal is achieved by deleting clusters with a standard deviation above a certain threshold, as these are likely to be uncorrelated noise rather than actual crack data. Another laser point cloud denoising technique used for pavement crack 3D evaluation is a combination of median filtering to remove outliers by addressing extreme values, followed by Gaussian filtering to smooth the data and reduce noise further. The median filter is effective in preserving edges while eliminating noise, particularly useful for the crack detection context, where maintaining sharpness is critical. Gaussian filtering then smooths the point cloud, enhancing the data's continuity and making it more suitable for subsequent feature extraction [68]. Another 3D data noise reduction technique is the combination of threshold filtering, moving window outlier detection, and discrete wavelet transform denoising [75]. Initially, threshold filtering is applied to remove obvious outliers by limiting the upper range of relative elevation data, with extreme points being replaced through interpolation. The moving window method further refines this process by detecting and replacing outliers within a sliding window matrix. Outliers are identified if they exceed three times the standard deviation from the mean, and these points are replaced with the mean value. Following outlier removal, discrete wavelet transform is employed for noise reduction, which stabilizes and optimizes the data by mitigating the influence of noise that blurs signal contours. This method effectively reconstructs the 3D pavement surface, closely aligning the processed data with the original pavement texture features.

The presence of pavement curvature and slopes in the captured 3D images often results in false classifications. Hence, researchers have applied preprocessing techniques to mitigate the effect of curvatures by creating a blurred version of the 3D image using a normalized box filter, then by subtracting the blurred version from the original 3D image, the edges in the 3D image get enhanced, thus resulting in a more accurate identification [72]. High-pass filters were also used to remove the curvature from 3D pavement images, which helped in having a better pavement raveling estimation [73]. Some processing techniques were used to extract depth information from 2D images. One approach is using a color filter and a convolving filter to extract color variations from the image, then translate them into depth information [94].

Missing 3D pixels/points can reduce the generalization ability of the trained artificial intelligence models. Therefore, some processing techniques focused on filling the missing data by deploying interpolation techniques, which replace the missing values with an interpolated value obtained from the mean of the surrounding pixels/points [101] [85]. Other processing techniques were used to facilitate the annotation and labeling of the 3D data. This was performed by clustering neighboring pixels into super-pixels by using unsupervised mean shift clustering [49].

3.3. Artificial intelligence models

The conducted review indicates that the most commonly used artificial intelligence models in the field of pavement damage assessment

using 3D imaging are: CNN (with different configurations) (74.1 %), ANN (4.7 %), RF, SVM, Adaboost (10.6 %), K-means, K-nearest neighbor (2.35 %), RNN (1.18 %), LSTM (1.18 %), Restricted Boltzmann machines (RBM) + RF (1.18 %), Genetic Algorithm + Neural Network (1.18 %), Genetic Algorithm + DenseNet (1.18 %), Deep Neural Network (DNN) (1.18 %) and Generative Adversarial Networks (GANs) (1.18 %). The contributions related to each artificial intelligence model, as well as their merits and limitations are listed in Table 4.

3.3.1. CNN based models

The significant advantage of CNN based pavement damage assessment over traditional image processing is that in CNN, the distinctive features in the data are automatically extracted and learnt, without the need for subjective human judgment, thus making it more accurate and

less prone to bias. In the conducted review, the CNN based contributions were divided into 8 categories according to the output generated by the models as shown in Table 5. The most common CNN output form was pixel-level/point-wise classification and segmentation (49 %). Other CNN output forms were: bounding boxes with classifications (13 %), 3D image (full) classification (13 %), regression (6.3 %), pixel-level/point-wise classification and segmentation in addition to bounding boxes (6.3 %), 3D image (patches) classification then segmentation (6.3 %), 3D image (full) classification + segmentation (4 %) and 3D image (full) classification + regression (2 %).

3.3.1.1. Pixel-level/point-wise classification and segmentation. In the pixel-level/point-wise classification and segmentation-based CNN contributions, the classification of pavement damage gets performed on

Table 4

Contributions distribution based on the adopted artificial intelligence technique and their corresponding merits and limitations.

Artificial Intelligence Technique	# Of Contributions (Total = 85)	References	Merits	Limitations
CNN (with different configurations)	63 (74.1 %)	[40] [41] [42] [43] [44] [48] [61] [117] [111] [45] [78] [79] [46] [97] [63] [102] [65] [66] [112] [108] [95] [47] [71] [48] [103] [109] [56] [58] [19] [49] [33] [50] [101] [93] [51] [62] [7] [57] [84] [106] [60] [87] [88] [94] [52] [86] [35] [67] [116] [107] [114] [98] [104] [115] [90] [100] [105] [91] [76] [118] [70] [92] [81]	- Ability to learn local features within 3D images. - Translation invariant pattern recognition. - Automatic feature extraction. - Ability to use transfer learning in the presence of limited data.	- Requires a large number of labeled data. - Prone to overfitting.
ANN	4 (4.7 %)	[53] [54] [64] [74]	- Simple architecture with low computation cost. - Ability to use transfer learning in the presence of limited data. - Automatic feature extraction from non-image data. - Faster to train with lower computational cost. - Ability to work with small datasets. - Less prone to overfitting.	- Can't capture spatial information from data. - Inefficient in tasks involving grid data like images. - Often requires manual feature extraction. - Can't capture spatial information from data. - Inefficient in tasks involving grid data like images. - Often requires manual feature extraction. - Difficulty in handling non-linearity in data. - Less efficient transfer learning capabilities. - Less capable of handling large datasets.
RF, SVM, Adaboost, generalized regression neural networks (GRNN) K-means, K-nearest Neighbor	9 (10.6 %)	[82] [9] [72] [73] [75] [69] [99] [96] [80]		
RNN LSTM	2 (2.35 %)	[83] [68]		
	1 (1.18 %)	[55]	- Designed for tasks involving sequential time series data. - Handles variable length sequences. - Maintains memory across time steps.	- Prone to vanishing and exploding gradients. - Takes more time to train due to its sequential nature. - Inefficient in capturing spatial information from grid data like images. - Does not have spatial invariance.
Restricted Boltzmann machines (RBM) + RF	1 (1.18 %)	[77]	- Performs unsupervised feature extraction. - Reduces the dimensionality of the data. - Captures complex dependencies within the data. - Ability to use transfer learning in the presence of limited data.	- Inefficient in capturing spatial information from grid data like images. - Requires manual feature extraction.
Genetic Algorithm + Neural Network Genetic Algorithm + DenseNet	1 (1.18 %)	[85]	- Can be used to optimize the parameters of other learning-based algorithms. - Can overcome noise by sampling several solutions.	- Takes a long time to train. - Does not perform well when there is a large search space. - Tuning the genetic algorithm parameters can be challenging.
DNN	1 (1.18 %)	[89]	- Rapid and detailed texture analysis.	- Need for further calibration to ensure consistency and avoid discrepancy in estimated texture depth.
GAN	1 (1.18 %)	[110]		
	1 (1.18 %)	[113]	- Generates highly realistic synthetic data, which can be used for data augmentation, reducing the reliance on large labeled datasets.	- The training process can be computationally expensive and often requires careful tuning of hyperparameters and loss functions.

Table 5

CNN based contributions, their limits and limitations.

CNN Output Form	# Of contributions (Total = 63 CNN Based contributions)	CNN architecture	Count	References	Merits	Limitations
Pixel-level/Point-wise Classification and Segmentation	29 (46 %)	CrackNet variations (CNN without pooling layers)	4	[40] [41] [47] [45]	- Accurate fine crack segmentation with low number of parameters.	- Slow processing and reduced generalization ability for some configurations.
		U-Net	4	[33] [97] [63] [104]	- Accurate detection of fine pavement deformities.	- More prone to overfitting due to the large number of parameters.
		Encoder-Decoder	3	[18] [103] [51]	- Robustness to quantity imbalances between cracked and non-cracked pixels. - Directly classifies the raw 3D data without pre-processing.	- Downsizing operations could result in some data loss and it is less accurate for shallow cracks detection.
		Graph convolution network variations	2	[79] [78]	- Enhanced receptive field compared to traditional CNNs and can be used to mitigate the information loss caused by the pooling operations. It also works with limited data.	- Less robust to noise.
		Simple CNN (Classification)	1	[35]	- Low number of parameters and fast pavement surface reconstruction.	- Has limited feature extraction capabilities compared to other CNN configurations.
		Densely Connected Convolutional Networks (DenseNets)	1	[86]	- More efficient feature extraction than conventional CNN.	- Computationally expensive with a large number of parameters.
		DeepLabv3 +	2	[106] [70]	- Robustness to patterns resembling pavement deformities.	- Less accurate segmentation at the edges.
		DeepLabv3 + (with graph attention layer GAL)	1	[87]	- Improved edge detection and segmentation.	- Computationally expensive.
		Encoder-Decoder (with residual connections)	1	[42]	- Robustness to patterns resembling pavement deformities.	- Increased network complexity.
		Encoder-Decoder (with pyramid pooling network)	1	[101]	- Lower parameters and faster computation.	- Some information can be lost during the pooling and down sampling operations. It is also less efficient in detecting thin cracks. - High computational complexity and prone to overfitting.
		Encoder-Decoder (CE-Net)	1	[90]	- Captures contextual semantic information, leading to high-level feature maps	- Prone to overfitting.
		Deep CNN	2	[65], [107]	- Enhanced pavement fine crack detection performance.	- Computationally expensive due to the lack of pooling layers.
		CapsNet (CNN with capsule layer)	1	[61]	- Overcomes spatial invariance limitations.	- More prone to overfitting due to the large number of parameters.
		U-Net (with graph network)	1	[111]	- Faster convergence speed compared to other U-Net configurations and better feature extraction.	- More prone to overfitting due to the large number of parameters.
		U-Net (with dense attention)	1	[66]	- Improved pavement crack localization and lower noise.	- Requires high computational resources to handle the fine-grained texture data and sophisticated wavelet decomposition
Bounding Boxes with Classifications	10 (15.87 %)	Deep fusion network	1	[67]	- Integrates multi-scale texture features from 3D point cloud data, thus significantly enhancing the accuracy of pavement friction estimation.	- Requires high computational resources due to the combined processing steps and sophisticated attention mechanisms used in segmentation.
		YOLOX	1	[81]	- Combines detection and segmentation for enhanced accuracy; robust to noise and varying conditions	-
		Multiple CNN architectures	1	[112]	-	-
		YOLO	6	[84] [19] [60] [88] [94] [115]	- Performs real-time evaluation.	- Less accurate than region-based CNN configurations.
		Faster R-CNN	1	[52]	- More accurate detection than YOLO.	- Slower processing.
		EfficientDet	1	[116]	- Designed to operate within the computational constraints of less powerful devices.	- Lower speed in high-performance scenarios compared to YOLO.
		Deep CNN (RetinaNet)	1	[91]	- Strong feature extraction capabilities, leading to high accuracy in complex image tasks.	- Requires significant computational resources, which may limit real-time application.

(continued on next page)

Table 5 (continued)

CNN Output Form	# Of contributions (Total = 63 CNN Based contributions)	CNN architecture	Count	References	Merits	Limitations
3D Image (Full) Classification	7 (11.1 %)	Deep CNN (Resnet 50)	1	[98]	- Strong feature extraction capabilities, leading to high accuracy in complex image tasks.	- Requires significant computational resources, which may limit real-time application.
		Simple CNN (Classification)	5	[109] [71] [46] [93] [62]	- Low parameters and simple architecture.	- Shallow feature extraction compared to other deeper CNN configurations.
Regression	6 (9.5 %)	Deep CNN	2	[44] [114]	- Better feature extraction and generalization ability.	- Higher computational cost with more parameters.
		Simple CNN (Regression)	2	[56] [95]	- Low parameters and simple architecture.	- Shallow feature extraction compared to other deeper CNN configurations.
Pixel-level/Point-wise Classification and Segmentation, Bounding Boxes	3 (4.76 %)	Deep CNN	4	[48] [105] [76] [92]	- Better feature extraction and generalization ability.	- Higher computational cost with more parameters.
		Faster R-CNN, YOLO Mask R-CNN, YOLO	2	[117] [50]	-	-
3D Image (Patches) Classification then Segmentation	3 (4.76 %)	Simple CNN (Classification)	3	[49] [43] [102]	- Computationally efficient.	- Less accurate than pixel-level segmentation
		PointNet+ +	2	[57] [58]	- Directly applied to the collected 3D data.	- Computationally expensive.
3D Image (Full) Classification + Segmentation	4 (6.3 %)	Convolutional block attention module (CBAM)	1	[100]	- Enhances feature representation by focusing on important spatial and channel-wise information.	- Increases model complexity, leading to potential slowdowns in processing speed.
		Encoder-Decoder CNN	1	[118]	- Robustness to quantify imbalances between cracked and non-cracked pixels. - Directly classifies the raw 3D data without pre-processing.	- Downsizing operations could result in some data loss and it is less accurate for shallow cracks detection.
3D Image (Full) Classification + Regression	1 (1.58 %)	Simple CNN (Classification, Regression)	1	[108]	- Low parameters and simple architecture.	- Shallow feature extraction compared to other deeper CNN configurations.

each pixel within a 3D image, or on each point in a 3D point cloud. A binary classifier is often used to categorize the pixels into damaged/undamaged categories. The collective classifications of all pixels/points create a segmentation between pavement areas pertaining a form of deformity and sound pavement sections. Traditional CNN architectures have limited pixel-level segmentation abilities due to the loss of information in the pooling stage, and the class imbalance between deformed and background regions [18]. Hence, researchers have introduced modified CNN architectures to improve the performance of deformity segmentation in 3D pavement images. Some researchers [41], [47] suggested performing pixel-based learning by deactivating the pooling layers from the CNN models used to extract the crack information from 3D pavement images. This is due to the downsizing nature of the pooling operation that merges or replaces pixels, which might result in losing pixels that could pertain critical information regarding the nature of certain deformities like cracking. The main limitation with such modification, is slow processing due to the large number of parameters. In addition, it relies on fixed operations, which require pre-designed filters, thus generating handcrafted features. This limits the learning process and generalization ability of the model. Such limitations can be addressed by utilizing a deeper network with less parameters and without using pooling layers, as demonstrated by CrackNet-V [40]. This can be achieved by reducing the filter sizes at the early layers, while still keeping stacked convolutional layers for deeper feature abstraction. A faster processing performance and more accurate pavement fine cracks detection can be achieved by having more hidden layers with less parameters and by replacing the handcrafted feature extraction process by learnable parameters throughout the layers, as demonstrated by CrackNet II [45].

Another CNN modification that doesn't involve using pooling layers, is replacing the scalar neurons by vector capsules in the architecture layers to get more detailed representation. Capsule networks (CapsNet)

have demonstrated their ability to overcome spatial invariance limitations experienced by traditional CNN architectures, when used to classify cracks in 3D pavement images [61]. Being spatial invariant enables CapsNet to classify objects from different viewpoints without explicitly including the viewpoints variations during the training process. Traditional CNNs are excellent for translational invariance, meaning that they can detect the presence of an object at different parts on an image. However, if the desired object has a new pose, which wasn't included during the training process, the CNN would experience a deteriorated performance. Such limitation is addressed by the spatially invariant CapsNets, since the activation vectors (capsules) automatically learn parameters like pose, orientation and rotation. In addition, capsule networks make use of dynamic routing, which enables the parallel communication between multiple capsules, thus enabling it to recognize multiple objects within the image even if they overlap. Such feature is particularly useful for pavement damage assessment, since pavement deformities tend to overlap in practice. The information loss caused by the use pooling operation, can be mitigated by utilizing dilated graph convolutions, which performs receptive field amplification by expanding the convolution kernel with low parameters. The authors in [78] utilized a similar methodology to perform 3D pavement crack segmentation, where they partitioned the processed input point clouds into cylindrical subsets. The features were then extracted from the cylindrical subsets using the dilated convolution operation. The detection performance of Graph Neural Networks can be further improved by deploying graph-widen module, which is a semi-supervised learning approach that first amplifies the intensity differences between crack points and non-crack points, then identifies the neighborhood surrounding the crack-points using nearest neighbors search (NNS) method, and finally it performs features extraction from the selected data including spatial density, elevation, intensity and local distribution. Such configuration contributes to improving the receptive field of the

graph neural network, and it is particularly useful when there is a scarcity in annotated data [79].

Encoder-decoder CNN configurations have also demonstrated their robustness to the quantity imbalances between cracked and non-cracked pixels in 3D pavement images. Such imbalance is commonly observed since cracks have a fine nature compared to the background. Encoder-decoder networks like EDNet [18] address such issue by fitting the pavement image with a reduced size version of the ground truth image. The scaling down operation reduces the pixel imbalance, hence improving the overall pixel-level segmentation. Moreover, encoder-decoder CNN architecture like DeepLabv3+ have demonstrated their ability to overcome the effect of non-crack patterns like pavement grooves and the disturbance of surface variations [106]. DeepLabV3+ has also demonstrated its effectiveness in 3D pavement rutting assessment by addressing the limitations of existing rutting depth measurement methods, which often fail to account for the influence of potholes, raveling, cracks, and bridge joints [70]. In addition, a highly accurate pavement crack segmentation performance was achieved using the raw 3D imaging data, and without utilizing any data preprocessing algorithms, which are often subjective in terms of parameters selection [42]. The encoder part of the network performs feature extraction from the input images, while the decoder performs expansion operations to match the output probability map with the input image size. Nevertheless, such configuration was observed to have a deteriorated performance when used to evaluate shallow cracks [18], [42], despite performing significant hyperparameter optimization on a wide range of model parameters [43]. This limitation can be addressed by using U-Net [33] [104], which borrows the same principles from an encoder-decoder configuration. Such architecture incorporates skip connections between the layers, which allows it to have a better performance in detecting fine details, since the information can be directly propagated from the early layers to the deeper ones, thus overcoming the problem of vanishing gradients. The performance of U-net based 3D pavement crack segmentation can be further improved by incorporating a Graph Network (GN) feature extraction branch. Such layer further improves the feature extraction capabilities of the model and boosts the convergence speed [111]. Graph Convolutional Networks have an enhanced receptive field compared to traditional CNNs, where they demonstrated a more efficient point-wise crack estimation from 3D data [78]. The crack localization ability of the U-Net decoder can be further improved by incorporating a channel attention block (CAB), which works on reducing noise and enhancing the salient features at the encoder stage prior to being passed to the decoder by modeling the source-target attention among the feature maps of the encoder-decoder, and by applying global attention pooling and additive feature fusion. Furthermore, the CAB block and the skip connections between the encoder-decoder compensate for the loss of information caused by the pooling layers, thus improving the crack segmentation and localization [66]. The semantic segmentation capability of DeepLabv3+ CNN architecture can be further improved by adding a Graph Attention Layer (GAL), which consists of two components. The first extracts the vertex and edge features, while the other updates the features representations. GAL layers were found to enhance the performance of CNNs mainly due to their ability to perform weight modulation operations, which create augmented activation values in the target areas, and reduced activation values in the background areas [87]. Another method to have a better pavement fine crack detection using a deep CNN configuration, is by dropping the fully connected network prior to the output, and relying instead on the output obtained from the last convolutional layer [65], [107]. Other CNN encoder-decoder configurations include performing pavement deformity feature extraction from 3D data through two parallel operations. In the first, the global information is extracted using a pyramid pooling module, while the second operation extracts the local information through encoder-decoder structure. Such configuration ensures the fusion of both local and global features, which results in a better deformity segmentation performance [101]. Another

encoder-decoder-based configuration is CE-Net, which was employed for 3D pavement cracks and potholes damage assessment using UAV remote-sensing images [90]. CE-Net, enhanced with DAC (Dense Atrous Convolution) and RMP (Residual Multi-kernel Pooling) modules, effectively captures contextual semantic information from high-resolution panoramic images.

CNN based pixel-level classification and segmentation can also be used for pavement surface height reconstruction, which could be used for estimating the presence of rutting. This was achieved in the literature by first extracting the depth information from 2D images by utilizing different light sources to get the matching points from different poses, then the obtained depth information was labeled using the data collected from a laser scanner, which served as the ground truth. Then, a simple CNN binary classifier was used to determine if the pixels were black, which reflects a certain depth. Such configuration resulted in a highly similar surface reconstruction to the one obtained using the laser scanning device [35]. Another CNN-based pixel-level classification and segmentation application include the evaluation of pavement skid resistance. This can be achieved through the use of 3D point cloud data and deep fusion networks. In [67], the developed fusion model processed fine-grained texture data from over 800 pavement sections to analyze the impact of multi-scale textures on pavement friction using two-dimensional wavelet decomposition. A multi-input fusion network with deep aggregation modules then integrated features from sub-images produced by this decomposition. Another point-wise classification and segmentation algorithm is YOLOX, which is used in combination with a Transformer-based model to perform pothole analysis from 3D pavement data obtained by a laser scanner [81]. YOLOX is first applied to detect pothole regions in the pavement images, utilizing its advanced object detection capabilities without relying on anchor boxes, which enhances detection accuracy and efficiency. Following the detection, the Transformer model further segments the identified potholes, allowing for precise extraction of their boundaries.

3.3.1.2. Bounding boxes with classifications. Some CNN based architectures perform pavement deformity localization and classification using an object detection configuration, where the algorithm delaminates the detected deformities in certain image patches by utilizing bounding boxes, rather than performing pixel-level classification. While such method generates less accurate characterization of the defects, it serves as a faster and less computationally expensive alternative for certain applications. The most common CNN based object detection algorithm is You Only Look Once (YOLO) [123], which performs object classification and localization in one-shot, where the whole image gets processed and evaluated in a single pass, thus enabling it to generate real-time predictions. Other CNN based object detection algorithm like Region Based CNN (R-CNN) [124] perform two-shot object detection, in which a set of predictions are first generated for each object classification class and localization, then a second processing stage is used to refine the predictions and select the most probable. In the original R-CNN, the algorithm creates around 2000 regions of proposal from each image prior to feeding them into the CNN classifier. Nevertheless, the network takes a long time to go through the classification of all regions. Hence, several region-based CNN configurations were developed throughout the years to address the processing speed limitation like Fast R-CNN [125] and Faster R-CNN [126]. In Fast R-CNN, the input image is directly plugged into the CNN classifier to generate feature maps, then the selective search process of the regions of proposal gets performed on the obtained feature maps. Therefore, the network goes through a lower number of classification instances. As for the Faster R-CNN, the slow process of creating the regions of proposals is performed by a separate network, then the obtained regions get reshaped and plugged into the CNN classifier, which results in a faster evaluation. Two-shot object detection configurations are generally more accurate than one-shot configurations. Nevertheless, they are less suited for real-time applications since

the network goes through an extra processing stage to generate predictions.

In terms of the on-shot object detection configurations, researchers found the YOLO algorithm fit for evaluating the presence of potholes in 3D stereo images [60], [84], [94]. One form of YOLO analysis, is utilizing two YOLO models, one evaluating the 2D RGB image and the other evaluating the depth information obtained through the disparity depth map, which is created by performing corrections to the collected stereo images. Then, the outputs of both networks are plugged into a fusion algorithm, like non-maximum suppression algorithm, to further discriminate between the overlapping classes, thus improving the overall performance [84]. Other researchers worked on estimating the volume of pavement potholes detected using a YOLO algorithm, by multiplying the horizontal and vertical laser depth dots confined in the estimated bounding box. This resulted in a 97.5 % volume estimation accuracy [88]. YOLO analysis was also used to identify various pavement distress types, including longitudinal cracking, transverse cracking, fatigue cracking, edge cracking, and potholes by applying it to a dataset comprising visible, infrared, and 3D fusion images with varying infrared ratios [115]. As for the two-shot object detection algorithms, researchers mainly used Faster-RCNN to estimate pavement manhole covers. An improved performance was achieved by replacing the VGG16 and ZF Net feature extractors used in the original Faster R-CNN, by a deeper ResNet50 feature extractor, which was found to be better in capturing the latent representations in the data [52]. Researchers in [116] used EffieicntDet to detect sub-pavement voids using infrared images captured by unmanned aerial vehicles. Such algorithm employs a uniform scaling methodology across network width, depth, and resolution to optimize performance, achieving a mean average precision of 0.85. The algorithm uses anchor boxes at various scales and aspect ratios to predict offsets and classify the content within each bounding box, distinguishing between different object classes. During training, it minimizes a loss function that combines classification accuracy and bounding box regression precision. Post-processing involves Non-Maximum Suppression (NMS) to refine the bounding boxes, eliminating redundant overlaps and retaining only the most confident detections. Other researchers have performed 3D pavement pothole damage assessment by utilizing a RetinaNet architecture [98]. This approach integrates structure from motion to generate a 3D point cloud structure from pavement video frames, allowing for the assessment of pothole severity based on depth. The trained model achieved an F1-score of 0.98 on a benchmark dataset, and a pothole depth estimation with a mean error of less than 5 %. Another deep CNN architecture used in 3D pavement pothole assessment is the modified RetinaNet, which involves a single-stage CNN that processes images of pavement surfaces to identify potholes and estimate their depth from UAVs or vehicle-installed cameras based on Structure from Motion (SfM) and 3D point cloud geometry [91].

3.3.1.3. 3D image (full) classification. Some CNN based contributions performed pavement damage classification using the full 3D image as input. While such approach lacks the localization of the detected deformities, it helps in providing fast and general assessment of the pavement condition. The classification results obtained from such approach can help maintenance teams in determining pavement sections requiring further visual inspection, or evaluation using more advanced processing techniques. Researchers successfully detected the presence of cracks, potholes, manholes, patches and raveling from 3D images using simple CNN classifiers [109] [71] [46] [93] [62]. Moreover, other researchers utilized deeper CNN configuration with more hidden layers to classify the severity level of pavement cracking by analyzing labeled full 3D images, without the use of patching or segmentation [44] [114].

3.3.1.4. Other output forms. Some research contributions utilized

multiple CNN configurations to evaluate pavement damage. For instance, both R-CNN and YOLO were used to evaluate the presence of pavement damage from 3D images, where a similar precision was obtained from both methods with varying output forms [50] [7]. Additionally, pavement cracks were extracted from 2D pavement scenes obtained through dash cameras using YOLO algorithm, then 3-dimensional crack depth estimation was performed using AdaBins, which is an encoder-decoder CNN modification that adaptively segments the depth ranges in a 3D scene into bins, thus giving a more global assessment of the data [117], [127]. Such process is found to have a better performance when transfer learning is applied to the AdaBins algorithm. Nevertheless, the conducted research work in this area suffers from scale-ambiguity, which stems from relying solely on the transfer learning-based depth estimation. Hence, it is recommended for future studies to use depth information associated with the input images utilized during the evaluation process, to ensure consistency and have uniform scaling [117].

Some CNN based contributions utilized simple CNN classifiers to evaluate the presence of damage in patches extracted from pavement 3D images. This resulted in a segmented 3D image with highlighted damaged areas. However, unlike the previously discussed segmentation techniques, this approach is not performed on the pixel level. Instead, it segments the damages on the patch level, whose size is determined by the system designer. Thus, making it more computationally efficient, but less accurate in detecting pavement deformities [49] [43] [102]. In such configurations, the performance of simple CNN classifiers was further improved by incorporating seed layers and region growing layers, which worked on amplifying the crack information by creating pixel-level crack contours [102]. In addition, Attention mechanisms like Dual Attention Convolutional Neural Network (DACNN) have demonstrated a substantially better pavement deformity classification and segmentation compared to the conventional VGG16, VGG19, ResNet50, and DenseNet121 architectures [49]. Such mechanisms give higher weights to relevant features, while the weights associated with irrelevant features are minimized. This is particularly useful, since it directs the training process to distinctive parts, thus allowing it to process large data with high performance. Simple CNN models were also used to solve regression problems related to pavement damage assessment from 3D images such as estimating the height of laser points in pavement surface profiling [56], pavement surface mean texture depth [95] and pavement skid resistance value [48] [105]. Other contributions classified pavement images into different roughness categories using a simple CNN classifier, then the roughness value was estimated using another regression CNN model [108]. Deeper CNN models like InceptionV4 module, were used in the 3D pavement skid resistance and texture assessment to predict pavement friction levels from surface texture data [76]. The network achieved a classification accuracy of 97.89 % on the test set, outperforming a conventional CNN model by 11.94 %.

The classification of 3D point clouds directly through CNN can be complex due to the points' disordered arrangement and the shallow depth of certain pavement damages. Hence, the analysis is often simplified by converting 3D point clouds into pixels or voxels [57]. This is achieved using CNN based tools like PointNet [128] and PointNet++ [119]. Such configurations were used to perform depth calculation on pavement damage images and contribute into classification [58] and segmentation [57] tasks. PointNet consists mainly of a transformation network and a max-pooling operation. The transformation network performs input alignment and uses spatial transformation network to achieve spatial invariance. The max-pooling operation extracts the global features from entire point clouds. However, the local features within the point cloud data remain unextracted. This limitation is addressed by PointNet++, which performs hierarchical feature learning to locally divide the point clouds. This is achieved through a sampling layer, which randomly selects a series of point clouds using farthest point sampling to define the local area. Then, the original PointNet algorithm is used to perform local feature extraction. Another

technique for full image segmentation and classification of 3D is by using Convolutional Block Attention Module (CBAM), which enhances feature extraction, selectively focusing on critical spatial and channel-wise details to distinguish real potholes from false artifacts [100]. Another full image classification and segmentation algorithm used in 3D pavement pothole assessment is the CNN encoder-decoder architecture [118]. This approach utilizes a consumer-grade RGB-D sensor paired with an edge computing device to collect 2D color images and corresponding 3D depth data of pavement surfaces. The encoder-decoder deep convolutional neural networks (DCNNs) are then trained on this heterogeneous RGB-D data to perform pothole segmentation.

3.3.2. ANN and DNN based models

ANN is a highly common learning-based artificial intelligence architecture. It consists mainly of an input layer, few hidden layers and output layer. The learning process is performed through the back-propagation process, in which the weights of the network get adjusted according to the error rate obtained by the output layer. Such process iterates until a minimum error rate is achieved. In the context of pavement damage assessment using 3D imaging, ANN models were mainly used in regression tasks related to the evaluation of pavement texture/friction/raveling related deformities. Particularly, ANN models were used to estimate the regression values of pavement surface texture [53] [54], pavement skid resistance and surface friction [64] and pavement surface roughness [74]. Generally, ANN models are well suited for applications having restricted computational requirements and small datasets since their architecture consists of shallow feature extraction with fewer parameters than other CNN based architectures. Moreover, the lack of spatial feature extraction layers makes ANN models less efficient in evaluating image-based data. In such data, ANN models convert 2D images into one-dimensional arrays, which results in exponentially increasing the number of training parameters. DNNs were also utilized to analyze texture depth and skid resistance of asphalt pavements by leveraging a Multiview stereo reconstruction method based on deep learning [110]. Cameras captured Multiview RGB-D and RGB images of various pavement types, which were then processed by the DNN. The network analyzed these images to reconstruct depth maps, which were compared against ground truth data to measure model precision. Key attributes such as image resolution, number of views, and pavement materials were studied to optimize model performance.

3.3.3. RF, SVM, adaboost, k-means, GRNN, RBM based models

Random Forest (RF) is a learning-based artificial intelligence algorithms that utilizes multiple randomly generated decision trees to perform classification and regression tasks. RF models are particularly useful in analyzing both linear and non-linear data. A decision tree is comprised of a set of nodes, which serve as a testing decision related to a feature attribute. Each node branches from a prior node, forming a tree like structure with a root and multiple branches. Such decision structure results in splitting the data into subsets related to the selected feature attributes. In the RF algorithm, multiple decision trees with random feature subsets are created, then the final results are obtained by voting on the results obtained from all decision trees or getting their average. Another decision tree-based algorithm is Adaboost, which similarly to RF, randomly creates multiple decision trees then performs a vote process. Except that in Adaboost the decision tree size is fixed to only one decision node and two result leaves referred to as stumps. Unlike RF models, Adaboost learns from previous errors by giving more weight to misclassified data points, so that the upcoming created stump would focus on learning those data points. Therefore, during the final decision-making process, larger stumps are given more vote weight, while in RF all decision trees are treated equally. As for Support Vector machine (SVM) learning-based algorithm, the main objective is to find a hyperplane that classifies data points into distinctive classes. In regression tasks, the same notion is applied in a continuous space, where the goal is

to find a hyperplane that has the minimum sum of squared errors with the continuous data points.

In the context of pavement damage assessment using 3D imaging, the previously mentioned algorithms were used to evaluate texture/friction/raveling related pavement deformities [72] [73]. While a similar performance was achieved using RF, SVM and Adaboost algorithms, random forest generally achieved a slightly higher performance. RF models often require feature extraction through manual or algorithmic means to achieve satisfactory results. For instance, RF models were used jointly with k-means clustering to segment pavement deformities from 3D laser profiles, where a 79 % classification accuracy was achieved. Linear regression and random forest were also used to assess the relationship between pavement texture characteristics and skid resistance in 3D pavement evaluation [75]. Linear regression offered a basic approach to modeling the connection between texture features and friction coefficients but was limited in capturing complex, non-linear patterns. Random forest improved upon this by employing an ensemble of decision trees, allowing for a better understanding of non-linear relationships and variable interactions. The k-means algorithm was used to pre-process the data, by clustering the 3D images segments into "damage" and "no damage". Then, RF models were used to perform further damage assessment [82]. Similarly, Restricted Boltzmann machines (RBM) were used to extract the high-level features from pavement manhole 3D images. Then, the obtained features were used to train a random forest, which generated a weighted vote for each image patch resembling the probability of having a manhole in a particular region within the image [77]. k-means clustering was also used along with region growing algorithms to extract and segment pavement pothole regions [83]. Another crack detection algorithm is K-Nearest Neighbor (KNN), which has been used to enhance pavement crack detection when combined with three-dimensional laser point cloud data. In this approach, M-estimator sample consensus (MSAC) first fits the point cloud data to a plane, separating crack points from the pavement surface. KNN then calculates the nearest neighbor distances for each point, using these distances to filter out non-crack areas and accurately identify crack regions. This method is particularly effective in determining crack depth and refining the detection of cracks by applying distance-based constraints [68]. As for SVM, the main utilization was in the field of pavement crack classification, where it was used to distinguish between 10 types of pavement surfaces [9]. SVM and Random Forest were used in the 3D pavement crack assessment to classify points in point clouds acquired from Mobile Laser Scanners [69]. First, local geometric features of the points were calculated using Principal Component Analysis (PCA), which provided eigenvalues and eigenvectors to assess point dispersion and spatial distribution. These features were then used to train the SVM and Random Forest models, with crack points labeled based on field annotations. Additionally, SVM was used in the 3D classification of pavement cracks by employing an RBF-type kernel to distinguish between different types of cracks based on their shape features [99]. The process involved applying grayscale conversion and binarization using the FCM thresholding method. Cracks were then classified using the K-means method, and relevant characteristics such as length, width, and orientation were extracted. The Hough transform was applied to identify rectilinear structures, which represent the cracks. Finally, SVM was utilized to classify the cracks according to their types, leveraging the extracted shape features. Another regression-based algorithm is Generalized Regression Neural Network (GRNN), which was used to predict the skid resistance and texture depth of asphalt pavements affected by eolian sand [96]. The process involved image processing to extract 3D texture features of the pavement surface. These features were then used as input parameters for the GRNN, which accurately predicted the texture depth with a high correlation coefficient ($R^2 = 0.98$). Random forest was also used in the 3D pavement skid resistance assessment to evaluate the anti-skid performance of roads based on pavement texture features, achieving a prediction accuracy of 0.78 for the dynamic friction coefficient at 70 km/h, indicating its

effectiveness in evaluating skid resistance under rapid tire sliding conditions [80].

3.3.4. RNN based models

It was observed that CNN based analysis is less effective in detecting hairline cracks. This is due to the nearly indistinguishable similarity between crack and noise patterns. Moreover, the pooling layers in the CNN perform downsizing of the data, thus resulting in the loss of critical pixel-based information. While this limitation is addressed by increasing the filters sizes, it results in higher computational requirements [55]. The issue of local receptive fields of a CNN is addressed by the use of Recurrent Neural Networks (RNN). Wherein, researchers consider the 3D images as sequences, and each sequence represent connected pixels. Then by deploying sequence generation and modeling, the probabilities of input sequences being related to cracks are obtained [55]. RNN differs from other learning-based configurations by having a memory unit that allows it to create dependencies among the input time-stamps. This makes it ideal for sequential data analysis, where context plays a major role in defining meaningful data. The advantage of having sequence-based pavement crack analysis is that a single sequence could be representative of all the needed crack information, and background pixels can be excluded from the sequence to reduce the computational cost. In addition, RNN has demonstrated its ability to process sequences with variable lengths, thus making it suitable for the evaluation of pavement cracks, whose lengths are inherently variable.

Another commonly used RNN based configuration is Long-Short Term Memory (LSTM) algorithm, which shares a similar structure to RNN with the addition of a forget gate. Such gate enables the network to erase irrelevant information, thus making it better in capturing long-term dependencies, and more immune to the vanishing gradients limitation suffered by RNN. LSTM models have demonstrated an ability to identify pavement rutting during pavement construction with an 84.91 % accuracy. Nevertheless, current LSTM based contributions exhibit high validation losses due to the low size of data used for training. In addition, the existing models rely only on the depth information obtained from the laser sensors. Hence, the integration of RGB based layers could improve the performance [59].

3.3.5. Genetic algorithm-based models

Genetic Algorithm (GA) is a search-based technique commonly used in solving optimization problems. Initially, a GA creates a random population set, which represent possible solutions to the addressed problem. Each individual possible solution is referred to as chromosome, and it consists of parameters referred to as genes. The search process is performed by iterating through the randomly generated population, while performing selection, crossover and mutation operations. In the selection process, a fitness function is used to quantify the chromosome's performance in solving the addressed problem. Then, the fittest solutions are selected for further processing by the subsequent stages. Then, in the crossover stage new solutions are created by exchanging the parameters of multiple solutions. As for the mutation stage, it serves in expanding the search area and maintaining parametric diversity by introducing random changes to the parameters. Hence, by performing several iterations, the obtained population sets get fitter than the previous iterations, until an optimum solution is obtained.

Genetic algorithms can be used to optimize the performance and the hyperparameters of neural network-based configurations. The performance of the latter is often affected by the random initialization of weights and thresholds in the network, which creates inconsistent results. Therefore, researchers have used Genetic algorithms to find the optimum neural network weights. Such configuration resulted in an enhanced pavement friction coefficient estimation performance using neural networks [85]. In addition, it has been used to optimize the hyperparameters of a DenseNet model used to evaluate the presence of different crack and pothole deformities in pavement 3D images [89].

3.3.6. Generative adversarial networks (GANs) models

Generative Adversarial Networks (GANs) are a class of machine learning frameworks where two neural networks, the generator and the discriminator, compete against each other. The generator creates synthetic data, while the discriminator attempts to distinguish between real and generated data. This adversarial process leads to the generation of high-quality data that closely resembles the real data. In the context of urban pavement crack detection, researchers have developed a method using a Conditional GAN (C-GAN) integrated with an image acquisition system that includes both visible light and infrared sensors [113]. The method involves fusing multisource images using a C-GAN with skip connections and dilated convolutional blocks. The loss function for the generator includes the Structural Similarity Index (SSIM) and the Sum of the Correlations of Differences (SCD) functions, ensuring that the fusion images retain more crack details and edge information. Additionally, the pixel-level crack detection model, performs semantic segmentation to identify cracks at the pixel level, achieving 85.1 % precision and 88.4 % recall in various types of crack detection.

3.4. Application domains

Artificial intelligence-based analysis and 3D imaging techniques were used to perform pavement damage assessment related to five distinctive applications namely, pavement crack evaluation (34.1 %), pavement texture/friction/raveling estimation (25.88 %), pavement Pothole/Manhole detection (18.8 %), multiple pavement distress detection (16.47 %) and pavement rutting evaluation (4.7 %) as illustrated in Table 6.

3.4.1. Pavement cracking evaluation

Pavement cracking involves the delamination of the pavement surface and the formation of line shaped cavities. Such deformity comes in several forms including, longitudinal crack which develop parallel to the pavement section, transverse cracks which develop perpendicular to the pavement section and craquelure, in which the pavement section breaks into several sharp angled pieces [82]. In the context of 3D imaging-based pavement cracking evaluation, the conducted survey shows that CNN based artificial intelligence models were the most commonly used algorithms. This can be related to the automatic feature extraction capabilities of the CNN based models, which make use of filters and convolutional layers that extract high level features from the data. Such operations learn features like edges, color variations and pixel/point patterns. Thus, making them exceptionally good in detecting fine cracking patterns in images. Besides CNN based configurations, researchers evaluated pavement cracking using SVM [9] and RNN [55] based models. The performances of all artificial intelligence-based 3D

Table 6

Pavement damage detection using artificial intelligence and 3D imaging applications.

Pavement Damage Type	# Of Contributions (Total = 85)	References
Crack	29 (34.1 %)	[55] [82] [9] [40] [41] [42] [43] [44] [18] [61] [117] [111] [45] [78] [79] [46] [97] [63] [102] [65] [66] [112] [113] [114] [68] [104] [69] [99] [116]
Texture/ Friction/ Raveling	22 (25.88 %)	[53] [54] [108] [95] [47] [71] [72] [73] [85] [48] [64] [103] [74] [56] [67] [110] [75] [105] [76] [96] [92] [80]
Pothole/ Manhole	16 (18.8 %)	[7] [83] [57] [84] [77] [106] [60] [87] [88] [94] [52] [98] [100] [118] [81] [91]
Multiple Distresses	14 (16.47 %)	[109] [58] [19] [49] [33] [50] [101] [93] [89] [51] [62] [107] [115] [90]
Rutting	4 (4.7 %)	[86] [59] [35] [70]

pavement crack evaluation methods are listed in Table 7. While a direct comparison between the obtained performances is not possible due to the lack of unified datasets and learning objectives, it can be generally inferred from the results that CNN classifiers, and particularly those which omit the pooling layers, have demonstrated the best pavement cracking detection and localization.

3.4.2. Pavement texture/friction/raveling evaluation

Pavement friction contributes greatly to the overall safety of the road network. It has been observed that a higher number of accidents with more severity are associated with lower friction numbers [129], [130], [131]. Pavement friction variation can occur due to several factors including raveling, in which the surface particles get dislodged, thus resulting in increased roughness and degraded ride experience. Raveling can result in significant risk to the passing vehicles, since the loose particles can bounce with high velocity and cause damage to the windshield glass [71].

In the context of 3D imaging-based pavement texture/friction/raveling evaluation, the conducted survey shows that CNN based artificial intelligence models were the most commonly used algorithms [71] [35] [48] [103] [47] [108] [95]. Other utilized artificial intelligence algorithms include ANN [53] [54] [64] [74], RF, SVM, Adaboost [72] [73] and Genetic Algorithm + Neural Network [85] as illustrated in Table 8. Given the absence of standardized datasets and uniform learning objectives, it remains difficult to directly select the best performing pavement texture/friction/raveling estimation algorithm. Nevertheless, RF, SVM, Adaboost algorithms standout since they achieved high precision and recall performances, despite using a low number of parameters. Thus, making them ideal for applications with limited computational resources.

3.4.3. Pavement multiple distresses evaluation

Some contributions investigated the presence of multiple pavement deformities rather than focusing on one type. Such approach is useful for general pavement condition assessment, where it serves as an initial evaluation stage. Nevertheless, models trained to detect multiple pavement deformities are often followed by visual inspection or other automated artificial intelligence-based techniques specifically trained on a certain type of deformities, for a more thorough evaluation.

In the context of 3D imaging-based pavement multiple distresses evaluation, the conducted survey shows that CNN based artificial intelligence models were the most commonly used algorithms. In particular, simple CNN classifiers have demonstrated highly accurate general assessment of the deformity type, while more advanced configurations like YOLO, R-CNN and U-Net were successful in performing both deformity classification and localization as illustrated in Table 9.

3.4.4. Pavement pothole/manhole evaluation

Potholes are a common form of road deformities which result in significant danger to the road user. Potholes often exist in the form of semi-spherical cavities on the pavement surface. Such deformity can result in an accelerated deterioration of the pavement section, since the depth and volume of potholes increase due to water accumulation and freezing melting cycles. Another common form of pavement cavities are manholes, which under normal conditions serve as channels for draining rainwater, passing power cable and communication lines. However, if the well cover is displaced for any reason, it would create a highly risky area for pedestrians and passing vehicles, especially under poor lighting conditions.

In the context of 3D imaging-based pavement pothole/manhole evaluation, the conducted survey shows that CNN based artificial intelligence models were the most commonly used algorithms. In

Table 7
Performances of pavement cracking detection models.

Type	Architecture	Ref.	Accuracy	Recall	Precision	F-score	RMSE
CNN (different configurations)	CrackNet variations (CNN without pooling layers)	[40]	-	90.12 %	84.31 %	87.12 %	-
		[41]	-	87.63 %	90.13 %	88.86 %	-
		[45]	-	89.06 %	90.20 %	89.62 %	-
	Simple CNN (Classification)	[43]	-	99 %	99.80 %	99.40 %	-
		[46]	94 %	-	-	-	-
		[102]	-	90.20 %	88.50 %	89.30 %	-
	Graph convolution network variations	[78]	-	77.10 %	79.50 %	78.30 %	-
		[79]	93.10 %	73.90 %	70 %	71.90 %	-
	Deep CNN	[44]	-	-	85.42 %	-	-
		[65]	88 %	-	-	-	-
		[114]	95.66 %	-	-	-	-
	U-Net	[97]	97.80 %	-	-	-	-
		[63]	-	-	-	89.92 %	-
		[104]	-	53.5 %	69.8 %	60.5 %	-
	PSPNet	[112]	66.8 %	-	-	36.2 %	-
	FPN		88 %	-	-	64.2 %	-
	FCN		63.8 %	-	-	34.7 %	-
	DEEPLabv3		89.3 %	-	-	65.8 %	-
	UNet		65 %	-	-	37.5 %	-
	UNet-ResNet101		74 %	-	-	40 %	-
	UNet-VGG19		70.4 %	-	-	41.5 %	-
	U-Net (with dense attention)	[66]	-	96.70 %	80.02 %	87.57 %	-
	U-Net (with graph network)	[111]	-	49 %	63 %	53 %	-
	Faster R-CNN, YOLO	[117]	-	-	57.68 %	63 %	-
	CapsNet (CNN with capsule layer)	[61]	-	95.30 %	81.10 %	88.20 %	-
	Encoder-Decoder	[18]	-	-	-	97.80 %	-
	Encoder-Decoder (with residual connections)	[42]	-	-	-	87.10 %	-
	EfficientDet	[116]	-	-	85 %	-	-
Other models	SVM	[9]	-	-	-	-	0.019
		[99]	95.54 %	-	-	-	-
	RNN	[55]	-	95 %	88.89 %	91.84 %	-
	RF	[82]	79 %	-	-	-	-
	SVM	[69]	-	51.70 %	49.43 %	50.54 %	-
	RF		-	74.44 %	30.79 %	43.74	-
	GAN (with skip connections and dilated convolutional blocks)	[113]	-	88.4 %	85.1 %	-	-
	K Nearest Neighbor (KNN)	[68]	-	-	-	-	1.32

Table 8

Performances of pavement texture/friction/raveling detection models.

Type	Architecture	Ref.	Accuracy	Recall	Precision	F-score	R2	Euclidean error	IoU
CNN (different configurations)	Simple CNN (Classification)	[71]	90.80 %	-	-	-	-	-	-
	Deep CNN	[48]	77.12 %	-	-	-	-	-	-
		[105]	-	81.07 %	83.29 %	82.16 %	-	-	-
		[76]	97.89 %	-	-	-	-	-	-
		[92]	-	-	-	-	-	-	0.858
	Encoder-Decoder	[103]	-	-	-	-	0.9216	-	-
	CrackNet variations (CNN without pooling layers)	[47]	-	90 %	90 %	-	-	-	-
	Simple CNN (Classification, Regression)	[108]	89.60 %	-	-	-	-	-	-
	Simple CNN (Regression)	[95]	-	-	-	-	-	0.0024	-
	Deep Fusion Network	[67, 95]	-	-	-	-	-	0.0935	-
Other models	ANN	[53]	-	-	-	-	0.95	-	-
		[54]	-	-	-	-	0.85	-	-
		[64]	-	-	-	-	-	-	-
		[74]	-	-	-	-	0.708	-	-
	RF	[72]	-	96.1 %	97.6 %	-	-	-	-
	SVM	-	-	95.6 %	97.6 %	-	-	-	-
	Adaboost	-	-	93.3 %	96.1 %	-	-	-	-
	RF	[73]	-	96.1 %	97.6 %	-	-	-	-
	SVM	-	-	95.6 %	97.6 %	-	-	-	-
	Adaboost	-	-	93.3 %	96.1 %	-	-	-	-
	Linear Regression	[75]	-	-	-	-	68.9 %	-	-
	RF	-	-	-	-	-	76.1 %	-	-
	RF	[80]	78 %	-	-	-	-	-	-
	Generalized regression neural networks (GRNN)	[96]	-	-	-	-	98 %	-	-
DNN	Genetic Algorithm + Neural Network	[85]	-	-	-	95.70 %	-	-	-
	DNN	[110]	-	-	77 %	-	-	-	-

Table 9

Performances of pavement multiple distresses detection models.

Type	Architecture	Ref.	Accuracy	Recall	Precision	F-score
CNN (different configurations)	Simple CNN (Classification)	[109]	-	-	-	-
		[49]	-	94.40 %	95.90 %	95.10 %
		[93]	-	-	-	-
		[62]	98.33 %	98.80 %	99.80 %	98.60 %
	Faster R-CNN	[50]	-	87.4 %	92.4 %	89.8 %
	YOLOv3	-	-	87.3 %	92.7 %	89.9 %
	Encoder-Decoder (with pyramid pooling network)	[101]	-	-	-	85.90 %
	Encoder-Decoder	[51]	-	-	-	92.54 %
	Encoder-Decoder (CE-Net)	[90]	99.61 %	-	-	-
	PointNet+ + YOLO	[58, 19]	90 %	-	-	-
Other models	U-net (Crack segmentation)	[115]	-	-	59 %	-
	U-net (Pothole segmentation)	[33]	-	84.47 %	85.91 %	85.19 %
	Simple CNN (Regression)	-	-	95.52 %	96.32 %	95.92 %
	Deep CNN	[107]	-	-	-	-
	Genetic Algorithm + DenseNet	[89]	88.20 %	-	-	-

particular, YOLO and R-CNN generated the most accurate pothole/manhole evaluation as illustrated in Table 10.

3.4.5. Pavement rutting evaluation

Rutting refers to the depression of pavement surface. Such deformity can occur during the road construction if different pavement layers are applied inconsistently. This results in having elevated and depressed regions, thus affecting the ride experience [59]. In the conducted survey, pavement height variations and surface profiling were evaluated using simple CNN classifiers, which performed pixel-level binary classification on 3D images obtained through stereo cameras achieving a 1.64×10^{-4} RMSE performance [35]. Pixel-matching between pavement stereo images was also performed using deeper CNN based configurations like DenseNet [86]. Besides stereo imaging, pavement rutting was also evaluated using laser scanning sensors, where the collected data were treated as time series instances, then used to train an LSTM algorithm

achieving an accuracy of 84.91 % [59]. Laser scanning based pavement rutting evaluation was also performed using an improved DeepLabV3 + network for detecting and correcting abnormalities, achieving a comprehensive detection accuracy of 81.63 % across five different pavement characteristics [70].

3.5. Potential future research directions

The findings obtained from the conducted survey indicated several gaps in the literature, areas of improvement and potentially new research directions, which could be summarized as follows:

- Artificial intelligence models utilization in new application domains: it was observed that only a limited number of artificial intelligence architectures were used in the evaluation of certain pavement deformities. Table 11 highlights the unexplored artificial intelligence

Table 10

Performances of pavement pothole/manhole detection models.

Type	Architecture	Ref.	Accuracy	Recall	Precision	F-score	IoU
CNN (different configurations)	YOLO	[84]	-	-	95.91 %	-	-
		[60]	-	-	95.20 %	97.50 %	-
		[88]	-	-	91.10 %	-	-
		[94]	-	92 %	84 %	88 %	-
		[81]	-	-	-	91.74 %	0.841
	Faster R-CNN	[52]	-	-	-	98.15 %	-
	DeepLabv3 + (with graph attention layer GAL)	[87]	98.67 %	-	-	85.64 %	-
	DeepLabv3 +	[106]	93.10 %	-	-	-	-
	PointNet++	[57]	97.50 %	-	-	-	-
	Mask R-CNN, YOLO	[7]	-	92.80 %	89.80 %	-	-
Other models	Deep CNN (RetinaNet)	[98]	-	96 %	100 %	98 %	-
		[91]	96.21 %	-	-	98 %	-
	Convolutional block attention module (CBAM)	[100]	96.62 %	94.67 %	93.12 %	93.89 %	-
Other models	Encoder-decoder CNN	[118]	-	-	-	-	0.82
	K-means	[83]	-	-	-	-	-
	Restricted Boltzmann machines (RBM) + RF	[77]	-	-	-	95.70 %	-

Table 11

Unexplored artificial intelligence architectures pertaining to different pavement damages.

Pavement damage type	Unexplored artificial intelligence architectures
Rutting	CNN without pooling layers, U-Net, Encoder-Decoder, Graph convolution network variations, DeepLabv3 +, DeepLabv3 + (with graph attention layer GAL), Encoder-Decoder (with residual connections), CapsNet (CNN with capsule layer), U-Net (with graph network), U-Net (with dense attention), R-CNN, Fast R-CNN, Faster R-CNN, Simple CNN (Regression), ANN, RF, SVM, Adaboost, K-means, RNN, Restricted Boltzmann machines (RBM), Genetic Algorithm configurations.
Pothole/Manhole	CNN without pooling layers, Graph convolution network variations, Densely Connected Convolutional Networks (DenseNets), Encoder-Decoder (with residual connections), CapsNet (CNN with capsule layer), U-Net (with graph network), U-Net (with dense attention), Simple CNN (Regression), ANN, RF, SVM, Adaboost, RNN, LSTM.
Multiple Distresses	CNN without pooling layers, U-Net, Graph convolution network variations, Densely Connected Convolutional Networks (DenseNets), DeepLabv3 +, DeepLabv3 + (with graph attention layer GAL), Encoder-Decoder (with residual connections), CapsNet (CNN with capsule layer), U-Net (with graph network), U-Net (with dense attention), ANN, Adaboost, K-means, RNN, LSTM, Restricted Boltzmann machines (RBM).
Texture/Friction/ Raveling	U-Net, Graph convolution network variations, Densely Connected Convolutional Networks (DenseNets), DeepLabv3 +, DeepLabv3 + (with graph attention layer GAL), Encoder-Decoder (with residual connections), Encoder-Decoder (with pyramid pooling network), CapsNet (CNN with capsule layer), U-Net (with graph network), U-Net (with dense attention), YOLO, R-CNN, Fast R-CNN, Faster R-CNN, PointNet++ , RF, SVM, Adaboost, K-means, RNN, LSTM, Restricted Boltzmann machines (RBM).
Cracks	Densely Connected Convolutional Networks (DenseNets), DeepLabv3 +, DeepLabv3 + (with graph attention layer GAL), Simple CNN (Regression), ANN, RF, SVM, Adaboost, K-means, LSTM, Restricted Boltzmann machines (RBM).

architectures pertaining to different types of pavement damages evaluation. The stated architectures hold a promising potential since they have demonstrated their ability to successfully analyze 3D data related to other types of pavement deformities as reported in the previous sections. For instance, it was observed that the use of sequence-based learning algorithms like RNN and LSTM was not thoroughly exploited, since such algorithms were only used to evaluate pavement cracking and rutting. Researchers can explore the feasibility of using such algorithms in analyzing other forms of pavement deformities such as pavement potholes/manholes,

texture/friction/raveling related deformities and the detection of multiple deformities.

- 3D data collection hardware: according to Fig. 9, most contributions utilized laser scanners to collect 3D pavement damage data. Nevertheless, the equipment related to such technique tends to be expensive and require constant maintenance. Hence, it is recommended for researchers to further explore alternative 3D data acquisition techniques like stereo imaging and structure from motion, which demonstrated their reliability in depth estimation using low-cost cameras. In addition, further research is needed to improve the data interpretation under different illumination conditions, where researchers can deploy new filtering and data processing techniques. Moreover, researchers could explore the feasibility of installing the 3D data hardware configurations mentioned in the previous sections, on new inspection mediums like Unmanned Aerial Vehicles (UAV). Researchers could further investigate 3D pavement damage assessment using microwave imaging techniques, known for their robustness in varying environmental conditions [132], [133]. These techniques are also highly suitable for integration with AI analysis in material evaluations, offering a promising potential for more accurate and reliable assessments.
- 3D data pre-processing: 3D data pre-processing plays a major role in enhancing the evaluation of pavement damages. Several 3D data processing techniques related to the extraction of road information have demonstrated a great performance in similar applications other than A.I. based pavement damage detection. This includes using Standard Hough Transform and logical constrains [134], parametric active contour or snake energy model [135], angular distance to the ground normal calculation [136], Otsu's thresholding, neighbor-counting filtering, region growing [137], voxel-based upward growing [138] and super-voxel with Hough Forest framework [139]. There exists a potential for researchers to explore the use of the aforementioned methods in improving the performances of the artificial intelligence models used for 3D pavement damage assessment. In addition, researchers could explore the use of Generative Adversarial Networks (GANs) in diversifying the collected dataset, thus enhancing the generalization ability of the trained models.
- Comprehensive open 3D pavement damages dataset: most of the studied contributions relied on self-collected 3D pavement data to perform their analysis, due to the lack of a comprehensive open 3D pavement damage dataset. Therefore, the creation of such dataset would significantly contribute towards enhancing the interpretation of the results obtained using different data analysis methodologies, and enable a more constructive direction of the research efforts.

4. Conclusion and future recommendations

The conducted systematic review aimed at analyzing the contribution published from 2011 to mid-2024 in the field of pavement damage assessment using 3D imaging and artificial intelligence methods. A total of 85 contributions were found relevant to the topic. From which, the adopted artificial intelligence models, utilized 3D data collection hardware, utilized datasets, 3D data pre-processing techniques and application domains were extracted, compared and critically analyzed. The conducted survey discussed the merits and limitations of different artificial intelligence techniques based on computational cost, training efficiency, noise robustness and performance considerations. In addition, it highlighted the implementation details of different 3D data acquisition systems, the obtained 3D data resolution and discussed their respective pros and cons. Moreover, it presented a general assessment of the utilized 3D data pre-processing techniques. Ultimately, five distinctive application domains were identified based on the type of detected pavement deformity. Then, the best performing artificial intelligence algorithms were highlighted in each. Generally, as pavement deformities detection problems are inherently complex due to the geometric nature of the inspected distresses and the impact of the surrounding data collection environment, it is evident that traditional inspection techniques are inadequate in terms of accuracy, speed and cost. Therefore, there has been a significantly growing research interest in recent years, in adopting artificial intelligence-based 3D imaging of pavement distresses, as an alternative approach which aligns with the global shift towards data-driven automation of inspection processes. Hence, the gaps and limitations highlighted by this survey are vital for developing robust future artificial intelligence-driven 3D pavement inspection applications. The findings obtained from the conducted survey indicated several gaps in the literature, areas of improvement and potentially new research directions, which could be summarized as follows:

- 3D data collection hardware: according to Fig. 9, most contributions utilized laser scanners to collect 3D pavement damage data. Nevertheless, the equipment related to such technique tends to be expensive and require constant maintenance. Hence, it is recommended for researchers to further explore alternative 3D data acquisition techniques like stereo imaging and structure from motion, which demonstrated their reliability in depth estimation using low-cost cameras. In addition, further research is needed to improve the data interpretation under different illumination conditions, where researchers can deploy new filtering and data processing techniques. Moreover, researchers could explore the feasibility of installing the 3D data hardware configurations mentioned in the previous sections, on new inspection mediums like Unmanned Aerial Vehicles (UAV). Researchers could further investigate 3D pavement damage assessment using microwave imaging techniques, known for their robustness in varying environmental conditions [132], [133]. These techniques are also highly suitable for integration with AI analysis in material evaluations, offering a promising potential for more accurate and reliable assessments.
- 3D data pre-processing: 3D data pre-processing plays a major role in enhancing the evaluation of pavement damages. Several 3D data processing techniques related to the extraction of road information have demonstrated a great performance in similar applications other than A.I. based pavement damage detection. This includes using Standard Hough Transform and logical constraints [134], parametric active contour or snake energy model [135], angular distance to the ground normal calculation [136], Otsu's thresholding, neighbor-counting filtering, region growing [137], voxel-based upward growing [138] and super-voxel with Hough Forest framework [139]. There exists a potential for researchers to explore the use of the aforementioned methods in improving the performances of the artificial intelligence models used for 3D pavement damage

assessment. In addition, researchers could explore the use of Generative Adversarial Networks (GANs) in diversifying the collected dataset, thus enhancing the generalization ability of the trained models.

- Comprehensive open 3D pavement damages dataset: most of the studied contributions relied on self-collected 3D pavement data to perform their analysis, due to the lack of a comprehensive open 3D pavement damage dataset. Therefore, the creation of such dataset would significantly contribute towards enhancing the interpretation of the results obtained using different data analysis methodologies, and enable a more constructive direction of the research efforts.
- Artificial intelligence models utilization in new application domains: it was observed that only a limited number of artificial intelligence architectures were used in the evaluation of certain pavement deformities. Some unexplored artificial intelligence architectures hold a promising potential in evaluating new types of pavement distresses, since they have demonstrated their ability to successfully analyze 3D data related to other types of pavement deformities as reported in the previous sections. The following list highlights the unexplored artificial intelligence architectures pertaining to different types of pavement damages evaluation:
 - **Rutting Evaluation:** CNN without pooling layers, U-Net, Encoder-Decoder, Graph convolution network variations, DeepLabv3 +, DeepLabv3 + (with graph attention layer GAL), Encoder-Decoder (with residual connections), CapsNet (CNN with capsule layer), U-Net (with graph network), U-Net (with dense attention), R-CNN, Fast R-CNN, Faster R-CNN, Simple CNN (Regression), ANN, RF, SVM, Adaboost, K-means, RNN, Restricted Boltzmann machines (RBM), Genetic Algorithm configurations.
 - **Pothole/Manhole Evaluation:** CNN without pooling layers, Graph convolution network variations, Densely Connected Convolutional Networks (DenseNets), Encoder-Decoder (with residual connections), CapsNet (CNN with capsule layer), U-Net (with graph network), U-Net (with dense attention), Simple CNN (Regression), ANN, RF, SVM, Adaboost, RNN, LSTM.
 - **Multiple Distresses Evaluation:** CNN without pooling layers, U-Net, Graph convolution network variations, Densely Connected Convolutional Networks (DenseNets), DeepLabv3 +, DeepLabv3 + (with graph attention layer GAL), Encoder-Decoder (with residual connections), CapsNet (CNN with capsule layer), U-Net (with graph network), U-Net (with dense attention), ANN, Adaboost, K-means, RNN, LSTM, Restricted Boltzmann machines (RBM).
 - **Texture/Friction/Raveling Evaluation:** U-Net, Graph convolution network variations, Densely Connected Convolutional Networks (DenseNets), DeepLabv3 +, DeepLabv3 + (with graph attention layer GAL), Encoder-Decoder (with residual connections), Encoder-Decoder (with pyramid pooling network), CapsNet (CNN with capsule layer), U-Net (with graph network), U-Net (with dense attention), YOLO, R-CNN, Fast R-CNN, Faster R-CNN, PointNet +, RF, SVM, Adaboost, K-means, RNN, LSTM, Restricted Boltzmann machines (RBM).
 - **Cracks Evaluation:** Densely Connected Convolutional Networks (DenseNets), DeepLabv3 +, DeepLabv3 + (with graph attention layer GAL), Simple CNN (Regression), ANN, RF, SVM, Adaboost, K-means, LSTM, Restricted Boltzmann machines (RBM).

CRediT authorship contribution statement

Mohamed Ait Gacem: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Saleh Abu Dabous:** Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Khaled Hamad:** Writing – review & editing, Methodology, Formal analysis, Data curation, Conceptualization. **Waleed Zeiada:** Writing – review & editing,

Methodology, Formal analysis, Data curation, Conceptualization. **Rami Al-Ruzouq**: Writing – review & editing, Methodology, Formal analysis, Data curation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors would like to thank Office of Vice Chancellor for Research and Graduate Studies and Research Institute of Science and Engineering at the University of Sharjah for funding this research study.

References

- [1] L. Studer, et al., Analysis of the relationship between road accidents and psychophysical state of drivers through wearable devices, *Appl. Sci.* 8 (8) (2018), <https://doi.org/10.3390/app8081230>.
- [2] H. Pérez-Acebo, R. Ziolkowski, H. Gonzalo-Orden, Evaluation of the radar speed cameras and panels indicating the vehicles' speed as traffic calming measures (TCM) in short length urban areas located along rural roads, *Energy* 14 (23) (2021), <https://doi.org/10.3390/en14238146>.
- [3] A. Alhasan, I. Nlenanya, O. Smadi, C.A. MacKenzie, Impact of pavement surface condition on roadway departure crash risk in Iowa, *Infrastructure* 3 (2) (2018), <https://doi.org/10.3390/infrastructures3020014>.
- [4] J. Lee, B. Nam, M. Abdel-Aty, Effects of pavement surface conditions on traffic crash severity, *J. Transp. Eng.* 141 (10) (2015) 4015020, [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000785](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000785).
- [5] Y. Li, C. Liu, L. Ding, Impact of pavement conditions on crash severity, *Accid. Anal. Prev.* 59 (2013) 399–406, <https://doi.org/10.1016/j.aap.2013.06.028>.
- [6] C. Chan, B. Huang, S. Richards, Investigating effects of asphalt pavement conditions on traffic accidents in Tennessee based on the pavement management system (PMS), *J. Adv. Transp.* 44 (2010) 150–161, <https://doi.org/10.1002/atr.129>.
- [7] A. Dhiman, R. Klette, Pothole detection using computer vision and learning, *IEEE Trans. Intell. Transp. Syst.* 21 (8) (2020) 3536–3550, <https://doi.org/10.1109/TITS.2019.2931297>.
- [8] G. B. J. B. C. R. C. S. C. G. D. L. F. H. H. A. H. others, R. Victor, 2013 Report Card for America's Infrastructure. American Society of Civil Engineers, 2013. doi: 10.1061/9780784478837..
- [9] M. Gavilan, et al., Adaptive road crack detection system by pavement classification, *Sensors* 11 (2011) 9628–9657, <https://doi.org/10.3390/s11009628>.
- [10] H.-W. Wang, C.-H. Chen, D.-Y. Cheng, C.-H. Lin, C.-C. Lo, A real-time pothole detection approach for intelligent transportation system, *Math. Probl. Eng.* 2015 (2015) 869627, <https://doi.org/10.1155/2015/869627>.
- [11] F. Gunawan, Detecting road damages by using gyroscope sensor, *ICIC Express Lett.* 12 (2018) 1089–1098, <https://doi.org/10.24507/icicel.12.11.1089>.
- [12] T. Lei, A.A. Mohamed, C. Claudel, An IMU-based traffic and road condition monitoring system, *HardwareX* 4 (2018) e00045, <https://doi.org/10.1016/j.hwx.2018.e00045>.
- [13] H.A. Jamakhandi, K.G. Srinivasa, Internet of Things based real time mapping of road irregularities, *Int. Conf. Circuits, Commun., Control Comput.* (2014) 448–451, <https://doi.org/10.1109/CIMCA.2014.7057842>.
- [14] S. Mathavan, K. Kamal, M. Rahaman, A review of three-dimensional imaging technologies for pavement distress detection and measurements, *IEEE Trans. Intell. Transp. Syst.* 16 (5) (2015) 2353–2362, <https://doi.org/10.1109/TITS.2015.2428655>.
- [15] J. Laurent, M. Talbot, M. Doucet, Road surface inspection using laser scanners adapted for the high precision 3D measurements of large flat surfaces, *Proc. Int. Conf. Recent Adv. 3-D. Digit. Imaging Model. (Cat. No. 97TB100134)* (1997) 303–310, <https://doi.org/10.1109/IM.1997.603880>.
- [16] S.N.A. Mukti, K.N. Tahar, Detection of potholes on road surfaces using photogrammetry and remote sensing methods (review), *Sci. Tech. J. Inf. Technol., Mech. Opt.* 22 (3) (2022) 459–471, <https://doi.org/10.17586/2226-1494-2022-22-3-459-471>.
- [17] J.-Y. Jung, H.-J. Yoon, H.-W. Cho, A study on crack depth measurement in steel structures using image-based intensity differences, *Adv. Civ. Eng.* 2018 (2018) 7530943, <https://doi.org/10.1155/2018/7530943>.
- [18] Y. Tang, A. Zhang, L. Luo, G. Wang, E. Yang, Pixel-level pavement crack segmentation with encoder-decoder network, *Measurement* 184 (2021) 109914, <https://doi.org/10.1016/j.measurement.2021.109914>.
- [19] M. Hedeaya, et al., A low-cost multi-sensor deep learning system for pavement distress detection and severity classification, *8th Int. Conf. Adv. Mach. Learn. Technol. Appl. (AMLTA2022)* (2022) 21–33, https://doi.org/10.1007/978-3-031-03918-8_3.
- [20] Y.-A. Hsieh, Y. Tsai, Machine learning for crack detection: review and model performance comparison, *J. Comput. Civ. Eng.* 34 (2020) 04020038, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000918](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000918).
- [21] L. Ma, Y. Li, J. Li, C. Wang, R. Wang, M. Chapman, Mobile laser scanned point-clouds for road object detection and extraction: a review, *Remote Sens.* 10 (2018) 1531, <https://doi.org/10.3390/rs10101531>.
- [22] C. Wenming, Q. Liu, Z. He, Review of pavement defect detection methods, *IEEE Access* (2020), <https://doi.org/10.1109/ACCESS.2020.2966881>.
- [23] S. Cano-Ortiz, P. Pascual-Muñoz, D. Castro-Fresno, Machine learning algorithms for monitoring pavement performance, *Autom. Constr.* 139 (2022), <https://doi.org/10.1016/j.autcon.2022.104309>.
- [24] Y. JiangRemote sensing and neural network-driven pavement evaluation: a review2022.
- [25] N. Ma, et al., Computer vision for road imaging and pothole detection: a state-of-the-art review of systems and algorithms, *Transp. Saf. Environ.* (2022) 1–16, <https://doi.org/10.1093/tse/tdac026>.
- [26] E. Ranyal, A. Sadhu, K. Jain, Road condition monitoring using smart sensing and artificial intelligence: a review, *Sensors* 22 (8) (2022), <https://doi.org/10.3390/s22083044>.
- [27] K. Gopalakrishnan, Deep learning in data-driven pavement image analysis and automated distress detection: a review, *Data* 3 (2018) 28, <https://doi.org/10.3390/data3030028>.
- [28] Y.-M. Kim, Y.-G. Kim, S.-Y. Son, S.-Y. Lim, B.-Y. Choi, D.-H. Choi, Review of recent automated pothole-detection methods, *Appl. Sci.* 12 (2022) 5320, <https://doi.org/10.3390/app12115320>.
- [29] S. Chambon, J.-M. Moliard, Automatic road pavement assessment with image processing: review and comparison, *Int. J. Geophys.* 2011 (2011) 989354, <https://doi.org/10.1155/2011/989354>.
- [30] T.B.J. Coenen, A. Golroo, A review on automated pavement distress detection methods, *Cogent Eng.* 4 (1) (2017) 1374822, <https://doi.org/10.1080/23311916.2017.1374822>.
- [31] Y. Fei, et al., Pixel-level cracking detection on 3D asphalt pavement images through deep-learning-based crackNet-V, *IEEE Trans. Intell. Transp. Syst.* 21 (1) (2020) 273–284, <https://doi.org/10.1109/TITS.2019.2891167>.
- [32] A. Zhang, et al., Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network, *Comput. -Aided Civ. Infrastruct. Eng.* 32 (10) (2017) 805–819, <https://doi.org/10.1111/mice.12297>.
- [33] S. Zhou, W. Song, Concrete roadway crack segmentation using encoder-decoder networks with range images, *Autom. Constr.* 120 (2020), <https://doi.org/10.1016/j.autcon.2020.103403>.
- [34] S. Zhou, W. Song, Deep learning-based roadway crack classification using laser-scanned range images: a comparative study on hyperparameter selection, *Autom. Constr.* 114 (2020) 103171, <https://doi.org/10.1016/j.autcon.2020.103171>.
- [35] Y.-A. Hsieh, Z. Yang, Y. Tsai, Convolutional neural network for automated classification of jointed plain concrete pavement conditions, *Comput. -Aided Civ. Infrastruct. Eng.* 36 (2020), <https://doi.org/10.1111/mice.12640>.
- [36] W. Zhu, W. Tan, L. Ma, D. Zhang, J. Li, M. Chapman, A capsnet approach to pavement crack detection using mobile laser scanning point clouds, *Int. Arch. Photogramm., Remote Sens. Spat. Inf. Sci.* XLIII-B1-2021 (2021) 39–44, <https://doi.org/10.5194/isprs-archives-XXXIII-B1-2021-39-2021>.
- [37] A. Badiley, R. Dalby, Z. Jiang, D. Sacharny, T.C. Henderson, Monocular road damage size estimation using publicly available datasets and dashcam imagery, *2022 IEEE Int. Conf. Multisens. Fusion Integr. Intell. Syst. (MFI)* (2022) 1–8, <https://doi.org/10.1109/MFI5806.2022.9913878>.
- [38] J. Chen, Y. Yuan, H. Lang, S. Ding, J. Lu, The improvement of automated crack segmentation on concrete pavement with graph network, *J. Adv. Transp.* 2022 (2022), <https://doi.org/10.1155/2022/2238095>.
- [39] A. Zhang, et al., Deep learning-based fully automated pavement crack detection on 3D asphalt surfaces with an improved cracknet, *J. Comput. Civ. Eng.* 32 (2018), [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000775](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000775).
- [40] L. Ma, J. Li, SD-GCN: Saliency-based dilated graph convolution network for pavement crack extraction from 3D point clouds, *Int. J. Appl. Earth Obs. Geoinf.* 111 (2022) 102836, <https://doi.org/10.1016/j.jag.2022.102836>.
- [41] H. Feng, et al., GCN-based pavement crack detection using mobile LiDAR point clouds, *IEEE Trans. Intell. Transp. Syst. PP* (2021) 1–10, <https://doi.org/10.1109/TITS.2021.3099023>.
- [42] B. Li, K. Wang, A. Zhang, E. Yang, G. Wang, Automatic classification of pavement cracks using deep convolutional neural network, *Int. J. Pavement Eng.* 21 (Jun. 2018) 1–7, <https://doi.org/10.1080/10298436.2018.1485917>.
- [43] S. Jiang, S. Gu, Z. Yan, Pavement crack measurement based on aerial 3D reconstruction and learning-based segmentation method, *Meas. Sci. Technol.* 34 (1) (Jan. 2022) 15801, <https://doi.org/10.1088/1361-6501/ac8e22>.
- [44] 陈泽斌, 罗文婷, and 李林, Automatic Identification of Pavement Crack Using Improved U-net Model 基于改进U-net 模型的路面裂缝智能识别, *J. Data Acquis. amp; Process. / Shu Ju Cai Ji Yu Chu Li* 35 (2) (2020) 260–269, <https://doi.org/10.16337/j.1004-9037.2020.02.007>.
- [45] T. Wen, H. Lang, S. Ding, J.J. Lu, Y. Xing, PCDNet: seed operation-based deep learning model for pavement crack detection on 3D asphalt surface, *J. Transp. Eng., Part B: Pavements* 148 (2) (2022) 4022023, <https://doi.org/10.1061/JPEODX.0000367>.
- [46] S. Luo, J. Yao, J. Hu, Y. Wang, S. Chen, Using deep learning-based defect detection and 3D quantitative assessment for steel deck pavement maintenance, *IEEE Trans. Intell. Transp. Syst.* (2022) 1–10, <https://doi.org/10.1109/TITS.2022.3169164>.

- [47] Y.-A. Hsieh, Y.-C.J. Tsai, DAU-net: dense attention U-net for pavement crack segmentation, 2021 IEEE Int. Intell. Transp. Syst. Conf. (ITSC) (2021) 2251–2256, <https://doi.org/10.1109/ITSC48978.2021.9564806>.
- [48] F. Liu, J. Liu, L. Wang, Asphalt pavement crack detection based on convolutional neural network and infrared thermography, *IEEE Trans. Intell. Transp. Syst.* 23 (11) (2022) 22145–22155, <https://doi.org/10.1109/TITS.2022.3142393>.
- [49] H. Abohamer, M. Elseifi, N. Dhakal, Z. Zhang, C.N. Fillastre, Development of a deep convolutional neural network for the prediction of pavement roughness from 3D images, *J. Transp. Eng., Part B: Pavements* 147 (4) (2021) 4021048, <https://doi.org/10.1061/JPEODX.0000310>.
- [50] Z. Tong, G. Jie, A. Sha, L. Hu, S. Li, Convolutional neural network for asphalt pavement surface texture analysis: convolutional neural network, *Comput.-Aided Civil Infrastruct. Eng.* 33 (2018), <https://doi.org/10.1111/mice.12406>.
- [51] G. Yang, K.C.P. Wang, A. Zhang, Y. Sun, B. Li, K.J. O'Connell, Field performance of deep-learning based fully automated cracking analysis and its potential for PCI surveys, *Highw. Pavements* (2019) 1–9, <https://doi.org/10.1061/9780784482476.001>.
- [52] Y.-A. Hsieh, Y. (James) Tsai, Automated asphalt pavement raveling detection and classification using convolutional neural network and macrotexture analysis, *Transp. Res. Rec.* 2675 (9) (2021) 984–994, <https://doi.org/10.1177/036119812111005450>.
- [53] G. Yang, A. Zhang, K. Wang, J. Li, W. Liu, Y. Liu, Deep-learning based non-contact method for assessing pavement skid resistance using 3D laser imaging technology, *Int. J. Pavement Eng.* (2022) 1–10, <https://doi.org/10.1080/10298436.2022.2147520>.
- [54] S. Dong, S. Han, C. Wu, O. Xu, H. Kong, Asphalt pavement macrotexture reconstruction from monocular image based on deep convolutional neural network, *Comput. -Aided Civ. Infrastruct. Eng.* 37 (13) (2022) 1754–1768, <https://doi.org/10.1111/mice.12878>.
- [55] S. Mokhtari, H.-B. Yun, A deep learning approach for pavement evaluation using 2D and 3D imaging systems, 18th Int. Conf. Road. Saf. Five Cont. (RSSC 2018) (2018).
- [56] S. Choi, S. Kim, H. Hong, Y. Kim, Road surface profiling based on artificial-neural networks, *RACS '20: Int. Conf. Res. Adapt. Converg. Syst.* (2020) 87–89, <https://doi.org/10.1145/3400286.3418282>.
- [57] Y. Li, S. Xue, Y. Mao, C. Zhao, X. Qin, Y. Liu, Classification of pavement disease 3D point cloud images based on deep learning network, 2021 6th Int. Conf. Intell. Comput. Signal Process. (ICSP) (2021) 812–815, <https://doi.org/10.1109/ICSP51882.2021.9408874>.
- [58] E. Eslami, H.-B. Yun, Improvement of multiclass classification of pavement objects using intensity and range images, *J. Adv. Transp.* 2022 (2022) 4684669, <https://doi.org/10.1155/2022/4684669>.
- [59] J. Guan, X. Yang, L. Ding, X. Cheng, V. Lee, C. Jin, Automated pixel-level pavement distress detection based on stereo vision and deep learning, *Autom. Constr.* 129 (2021) 103788, <https://doi.org/10.1016/j.autcon.2021.103788>.
- [60] R. Ghosh, O. Smadi, Automated detection and classification of pavement distresses using 3D pavement surface images and deep learning, *Transp. Res. Rec.: J. Transp. Res. Board* 2675 (2021), <https://doi.org/10.1177/036119812111007481>.
- [61] T. Wen, et al., Automated pavement distress segmentation on asphalt surfaces using a deep learning network, *Int. J. Pavement Eng.* (2022) 1–14, <https://doi.org/10.1080/10298436.2022.2027414>.
- [62] Y. Hu, T. Furukawa, A self-supervised learning technique for road defects detection based on monocular three-dimensional reconstruction, ASME 2019 Int. Des. Eng. Tech. Conf. Comput. Inf. Eng. Conf. (2019), <https://doi.org/10.1115/DETC2019-98135>.
- [63] A. Zhang, et al., Intelligent pixel-level detection of multiple distresses and surface design features on asphalt pavements, *Comput. -Aided Civ. Infrastruct. Eng.* 37 (2022), <https://doi.org/10.1111/mice.12909>.
- [64] D.Q. Thuyet, M. Jomoto, K. Hirakawa, Yin Lei LEI SWE, Development of an autonomous road surface damage inspection program using deep convolutional neural network, *J. JSCE* 10 (1) (2022) 235–246, <https://doi.org/10.2208/journalofjsce.10.1.235>.
- [65] S. Chaithavee, T. Chayakul, Classification of 3D point cloud data from mobile mapping system for detecting road surfaces and potholes using convolution neural networks, *Int. J. Geoinform.* 18 (6) (2022) 11–23, <https://doi.org/10.52939/jgg.v18i6.2455>.
- [66] J.C. Sungan Yoon, Convergence of stereo vision-based multimodal YOLOs for faster detection of potholes, *Comput., Mater. Contin.* 73 (2) (2022) 2821–2834, <https://doi.org/10.32604/cmc.2022.027840>.
- [67] Z. Xu, Road pothole extraction and safety evaluation by integration of point cloud and images derived from mobile mapping sensors, *Adv. Eng. Inform.* 42 (2019), <https://doi.org/10.1016/j.aei.2019.100936>.
- [68] L. Qing, K. Yang, W. Tan, J. Li, Automated detection of manhole covers in MLS point clouds using a deep learning approach. In: Proceedings of the IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, Sep. 2020, pp. 1580–1583. doi: 10.1109/IGARSS39084.2020.9324137..
- [69] R. Fan, H. Wang, Y. Wang, M. Liu, I. Pitas, Graph attention layer evolves semantic segmentation for road pothole detection: a benchmark and algorithms, *IEEE Trans. Image Process.* 30 (2021) 8144–8154, <https://doi.org/10.1109/TIP.2021.3112316>.
- [70] R. Li, C. Liu, Road damage evaluation via stereo camera and deep learning neural network, 2021 IEEE Aerosp. Conf. (2021) 1–7, <https://doi.org/10.1109/AERO50100.2021.9438528>.
- [71] S. Alshammari, S. Song, 3Pod: federated Learning-based 3 dimensional pothole detection for smart transportation, 2022 IEEE Int. Smart Cities Conf. (ISC2) (2022) 1–7, <https://doi.org/10.1109/ISC255366.2022.9922195>.
- [72] H. Zhang, et al., Efficient approach to automated pavement manhole detection with modified faster R-CNN, *Intell. Transp. Infrastruct.* 1 (2022), <https://doi.org/10.1093/iti/liac006>.
- [73] H. Brunken, C. Gühmann, Pavement distress detection by stereo vision, *Tech. Mess.* 86 (2019), <https://doi.org/10.1515/teme-2019-0046>.
- [74] G. de León, J. Cesbron, P. Klein, P. Leandri, M. Losa, Novel methodology to recover road surface height maps from illuminated scene through convolutional neural networks, *Sensors* 22 (17) (2022), <https://doi.org/10.3390/s22176603>.
- [75] C. Liu, N. Xu, Z. Weng, Y. Li, Y. Du, J. Cao, Effective pavement skid resistance measurement using multi-scale textures and deep fusion network, *Comput. -Aided Civ. Infrastruct. Eng.* 38 (8) (2023) 1041–1058, <https://doi.org/10.1111/mice.12931>.
- [76] N.N. Kulkarni, K. Raisi, N.A. Valente, J. Benoit, T. Yu, A. Sabato, Deep learning augmented infrared thermography for unmanned aerial vehicles structural health monitoring of roadways, *Autom. Constr.* 148 (2023) 104784, <https://doi.org/10.1016/j.autcon.2023.104784>.
- [77] S. Ramachandra, P. Kumar, S. Pasupunuri, J. Shinganmakkki, Evaluation of pavement surface distress using image processing and artificial neural network, *J. Test. Eval.* 51 (2023) 20220259, <https://doi.org/10.1520/JTE20220259>.
- [78] F. Liu, J. Liu, L. Wang, Deep learning and infrared thermography for asphalt pavement crack severity classification, *Autom. Constr.* 140 (2022) 104383, <https://doi.org/10.1016/j.autcon.2022.104383>.
- [79] E. Ranyal, A. Sadhu, K. Jain, AI assisted pothole detection and depth estimation, 2023 Int. Conf. Mach. Intell. GeoAnalytics Remote Sens. (MIGARS) (2023) 1–4, <https://doi.org/10.1109/MIGARS5735.2023.10064547>.
- [80] Y. Fujita, T. Tanaka, T. Hori, Y. Hamamoto, Classification model based on U-net for crack detection from asphalt pavement images, *J. Image Graph.* 11 (2023) 121–126, <https://doi.org/10.18178/jig.11.2.121-126>.
- [81] F. Liu, J. Liu, L. Wang, I.L. Al-Qadi, Multiple-type distress detection in asphalt concrete pavement using infrared thermography and deep learning, *Autom. Constr.* 161 (2024) 105355, <https://doi.org/10.1016/j.autcon.2024.105355>.
- [82] S. Feng, M. Gao, X. Jin, T. Zhao, F. Yang, Fine-grained damage detection of cement concrete pavement based on UAV remote sensing image segmentation and stitching, *Measurement* 226 (2024) 113844, <https://doi.org/10.1016/j.measurement.2023.113844>.
- [83] J. Dong, et al., CBAM-optimized automatic segmentation and reconstruction system for monocular images with asphalt pavement potholes, *IEEE Trans. Intell. Transp. Syst.* 25 (8) (2024) 10313–10330, <https://doi.org/10.1109/TITS.2024.3353257>.
- [84] G. Yang, A.A. Zhang, K.C.P. Wang, J.Q. Li, W. Liu, Y. Liu, Deep-learning based non-contact method for assessing pavement skid resistance using 3D laser imaging technology, *Int. J. Pavement Eng.* 24 (2) (2023) 2147520, <https://doi.org/10.1080/10298436.2022.2147520>.
- [85] E. Ranyal, A. Sadhu, K. Jain, Automated pothole condition assessment in pavement using photogrammetry-assisted convolutional neural network, *Int. J. Pavement Eng.* 24 (1) (2023) 2183401, <https://doi.org/10.1080/10298436.2023.2183401>.
- [86] G. Xu, X. Lin, S. Wang, Y. Zhan, J. Liu, H. Huang, Incep-frictionnet-based pavement texture friction level classification prediction method, *Lubricants* 12 (1) (2024), <https://doi.org/10.3390/lubricants12010008>.
- [87] Y.-T. Huang, M.R. Jahanshahi, F. Shen, T.G. Mondal, Deep learning-based autonomous road condition assessment leveraging inexpensive RGB and depth sensors and heterogeneous data fusion: pothole detection and quantification, *J. Transp. Eng. Part B: Pavements* 149 (2) (2023), <https://doi.org/10.1061/JPEODX-PVNG-1194>.
- [88] A. Wang, H. Lang, S. Ding, J. Lu, X. Hong, T. Wen, Rutting abnormality analysis method for 3D asphalt pavement surfaces based on semantic segmentation model, *Huanan Ligong Daxue Xuebao/J. South China Univ. Technol. (Nat. Sci.)* 51 (1) (2023) 134–144, <https://doi.org/10.12141/j.issn.1000-565X.210778>.
- [89] T. Ma, Z. Tong, Y.-M. Zhang, W.-G. Zhang, A three-dimensional reconstruction method of pavement macro-texture using a multi-view deep neural network, *Zhongguo Gonglou Xuebao/China J. Highw. Transp.* 36 (3) (2023) 70–80, <https://doi.org/10.19721/j.cnki.1001-7372.2023.03.005>.
- [90] A.-D. Wang, Y.-C. Peng, H. Lang, Y.-Y. Xing, Z. Chen, J. Lu, Pavement pothole extraction based on YOLOX-transformer two-step model, *Zhongguo Gonglou Xuebao/China J. Highw. Transp.* 36 (12) (2023) 304–317, <https://doi.org/10.19721/j.cnki.1001-7372.2023.12.023>.
- [91] G. Yang, K.C.P. Wang, J.Q. Li, G. Wang, A novel 0.1 mm 3D laser imaging technology for pavement safety measurement, *Sensors* 22 (20) (2022), <https://doi.org/10.3390/s22208038>.
- [92] Z. Yiwen, G. Yang, M. Cao, Neural network-based prediction of sideway force coefficient for asphalt pavement using high-resolution 3D texture data, *Int. J. Pavement Eng.* 23 (2021) 1–10, <https://doi.org/10.1080/10298436.2021.1884862>.
- [93] 彭毅, 李强, 战友, 杨广伟, and 王郴平, Pavement skid resistance evaluation based on 3D areal texture characterization 基于区域三维纹理特征的路面抗滑性能评估, *J. Southeast Univ. / Dongnan Daxue Xuebao* 50 (4) (2020) 667–676, <https://doi.org/10.3969/j.issn.1001-0505.2020.04.010>.
- [94] M. Fakhri, S.M. Karimi, J. Barzegaran, Predicting international roughness index based on surface distresses in various climate and traffic conditions using laser crack measurement system, *Transp. Res. Rec.* 2675 (11) (2021) 397–412, <https://doi.org/10.1177/036119812111017906>.

- [95] B. Horst, R. Lindenbergh, S. Puister, Mobile laser scan data for road surface damage detection, *Int. Arch. Photogramm., Remote Sens. Spat. Inf. Sci.* XLII-2/W13 (2019) 1141–1148, <https://doi.org/10.5194/isprs-archives-XLII-2-W13-1141-2019>.
- [96] Z. Wang, Development of an asphalt pavement raveling detection algorithm using emerging 3D laser technology and macrotexture analysis final report for NCHRP IDEA Project 163, 2016. [Online]. Available: <https://api.semanticscholar.org/CorpusID:36163270>.
- [97] Y.-C. Tsai, Y. Zhao, B. Pop-Stefanov, A. Chatterjee, Automatically detect and classify asphalt pavement raveling severity using 3D technology and machine learning, *Int. J. Pavement Res. Technol.* 14 (2020), <https://doi.org/10.1007/s42947-020-0138-5>.
- [98] G. Dai, Z. Luo, M. Chen, Y. Zhan, C. Ai, Reconstruction and intelligent evaluation of three-dimensional texture of stone matrix asphalt-13 pavement for skid resistance, *Lubricants* 11 (12) (2023), <https://doi.org/10.3390/lubricants11120535>.
- [99] P. del Rio, J. Grandio, B. Riveiro, P. Arias, Identification of relevant point cloud geometric features for the detection of pavement cracks using mls data, *Int. Arch. Photogramm., Remote Sens. Spat. Inf. Sci.* (2023) 107–112, <https://doi.org/10.5194/isprs-archives-XLVIII-1-W1-2023-107-2023>.
- [100] S. Sghaier, M. Krichen, I. Ben Dhaou, H. Elmannai, R. Alkanhel, Identification, 3D-reconstruction, and classification of dangerous road cracks, *Sensors* 23 (7) (2023), <https://doi.org/10.3390/s23073578>.
- [101] Y. Xue, P. Li, S. Jiang, N. Thom, X. Yang, Evaluation of skid resistance of asphalt pavement covered with aeolian sand in desert areas, *Wear* (2023) 204620, <https://doi.org/10.1016/j.wear.2023.204620>.
- [102] Y. Peng, Z. Zhang, Q. Li, G. Yang, A random forest evaluation model for pavement skid resistance based on comprehensive fractal, *Appl. Math. Mech.* 45 (4) (2024) 443–457, <https://doi.org/10.21656/1000-0887.440244>.
- [103] Y. Du, et al., A pothole detection method based on 3D point cloud segmentation, *Int. Conf. Digit. Image Process.* (2020) [Online]. Available: <https://api.semanticscholar.org/CorpusID:219842229>.
- [104] R. Li, X. Chen, Q. Dong, S. Wang, Z. Chu, X. guPoint cloud-based pavement crack extraction using MSAC and KNN Algorithm. 2024. doi: 10.1061/9780784485255.068..
- [105] A. Zhang, et al., Automated pixel-level pavement crack detection on 3d asphalt surfaces with a recurrent neural network: automated pixel-level pavement crack detection on 3D asphalt surfaces using CrackNet-R, *Comput.-Aid. Civil Infrastruct. Eng.* 34 (2018), <https://doi.org/10.1111/mice.12409>.
- [106] T. Patel, H.W. Brian, Y.Z. Guo, J.D. Van Der Walt, Y. Li, Sensor-based pavement layer change detection using long-short term memory (LSTM), *IOP Conf. Ser. Earth Environ. Sci.* 1101 (8) (2022) 82005, <https://doi.org/10.1088/1755-1315/1101/8/082005>.
- [107] Y. Yu, H. Guan, W. Mi, Automated detection of urban road manhole covers using mobile laser scanning data, *IEEE Trans. Intell. Transp. Syst.* 16 (2015) 1–12, <https://doi.org/10.1109/TITS.2015.2413812>.
- [108] Z. Sun, X. Hao, W. Li, J. Huyan, H. Sun, Asphalt pavement friction coefficient prediction method based on genetic-algorithm-improved neural network(GAI-NN) model, *Can. J. Civ. Eng.* 49 (2021), <https://doi.org/10.1139/cjce-2020-0051>.
- [109] J. Li, T. Liu, X. Wang, Advanced pavement distress recognition and 3D reconstruction by using GA-DenseNet and binocular stereo vision, *Measurement* 201 (2022) 111760, <https://doi.org/10.1016/j.measurement.2022.111760>.
- [110] H.-C. Dan, B. Lu, M. Li, Evaluation of asphalt pavement texture using multiview stereo reconstruction based on deep learning, *Constr. Build. Mater.* 412 (2024) 134837, <https://doi.org/10.1016/j.conbuildmat.2023.134837>.
- [111] H. Liang, et al., Automatic pavement crack detection in multisource fusion images using similarity and difference features, *IEEE Sens. J.* 24 (5) (2024) 5449–5465, <https://doi.org/10.1109/JSEN.2023.3267834>.
- [112] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: unified, real-time object detection. 2016. doi: 10.1109/CVPR.2016.91..
- [113] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2013, 10.1109/CVPR.2014.81.
- [114] R. Girshick, Fast R-CNN. In: *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1440–1448. doi: 10.1109/ICCV.2015.169.
- [115] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: towards real-time object detection with region proposal networks, *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (Jun. 2015), <https://doi.org/10.1109/TPAMI.2016.2577031>.
- [116] S. Bhat, I. Alhashim, P. Wonka, AdaBins: depth estimation using adaptive bins, *2021 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)* (2020) 4008–4017 [Online]. Available: <https://api.semanticscholar.org/CorpusID:227227779>.
- [117] R. Charles, H. Su, K. Mo, L. Guibas, PointNet: deep learning on point sets for 3D classification and segmentation. 2017. doi: 10.1109/CVPR.2017.16..
- [118] C. Qi, L. Yi, H. Su, L.J. Guibas, PointNet++: deep hierarchical feature learning on point sets in a metric space, *NIPS* (2017) [Online]. Available: <https://api.semanticscholar.org/CorpusID:1745976>.
- [119] D. Arya, et al., Deep learning-based road damage detection and classification for multiple countries, *Autom. Constr.* 132 (2021) 103935, <https://doi.org/10.1016/j.autcon.2021.103935>.
- [120] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, H. Omata, Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images: road damage detection and classification, *Comput.-Aided Civ. Infrastruct. Eng.* 33 (Jun. 2018), <https://doi.org/10.1111/mice.12387>.
- [121] D. Scharstein, R. Szeliski, R. Zabih, A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, *Proc. IEEE Workshop Stereo Multi-Baseline Vis.* (SMBV 2001) (2001) 131–140, <https://doi.org/10.1109/SMBV.2001.988771>.
- [122] Y.-C. Tsai, A. Chatterjee, Pothole detection and classification using 3D technology and watershed method, *J. Comput. Civ. Eng.* 32 (2018), [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000726](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000726).
- [123] I. Moazzam, K. Kamal, S. Mathavan, S. Ahmed, M. Rahman, 2013, Metrology and visualization of potholes using the microsoft kinect sensor. 2013. doi: 10.1109/ITSC.2013.6728408..
- [124] Y. Zhang, et al., A kinect-based approach for 3D pavement surface reconstruction and cracking recognition, *IEEE Trans. Intell. Transp. Syst.* (12) (May 2018) 1, <https://doi.org/10.1109/TITS.2018.2791476>.
- [125] S.M. Abbas, A. Muhammad, Outdoor RGB-D SLAM performance in slow mine detection. In: *Proceedings of the ROBOTIK 2012, 7th German Conference on Robotics*, 2012, pp. 1–6.
- [126] R. Raguram, J.-M. Frahm, M. Pollefeys, A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus 5303 (2008), https://doi.org/10.1007/978-3-540-88688-4_37.
- [127] L. Ma, Y. Li, J. Li, Z. Zhong, M.A. Chapman, Generation of horizontally curved driving lines in HD maps using mobile laser scanning point clouds, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 12 (5) (2019) 1572–1586, <https://doi.org/10.1109/JSTARS.2019.2904514>.
- [128] S.K. Shome, S.R.K. Vadali, Enhancement of diabetic retinopathy imagery using contrast limited adaptive histogram equalization, 2011. [Online]. Available: <https://api.semanticscholar.org/CorpusID:15587978>.
- [129] A. Alhasan, I. Nlenanya, O. Smadi, C.A. MacKenzie, Impact of pavement surface condition on roadway departure crash risk in Iowa, *Infrastructure* 3 (2) (2018), <https://doi.org/10.3390/infrastructures3020014>.
- [130] J.M. Pardillo Mayora, R. Jurado Piña, An assessment of the skid resistance effect on traffic safety under wet-pavement conditions, *Accid. Anal. Prev.* 41 (4) (2009) 881–886, <https://doi.org/10.1016/j.aap.2009.05.004>.
- [131] S. Najafi, G.W. Flintsch, A. Medina, Linking roadway crashes and tire–pavement friction: a case study, *Int. J. Pavement Eng.* 18 (2) (2017) 119–127, <https://doi.org/10.1080/10298436.2015.1039005>.
- [132] M. Ait Gacem, A. Zakaria, M. Ismail, U. Tariq, S. Yehia, Measurement of construction materials properties using Wi-Fi and convolutional neural networks, *IEEE Access* (2022), <https://doi.org/10.1109/ACCESS.2022.3226248>.
- [133] M.A. Gacem, A. Zakaria, M.H. Ismail, S. Yehia, U. Tariq, Concrete classification using Wi-Fi channel state information and convolutional neural networks, *Constr. Build. Mater.* 439 (2024) 137280, <https://doi.org/10.1016/j.conbuildmat.2024.137280>.
- [134] B. Riveiro, H. González-Jorge, J. Martínez-Sánchez, L. Díaz-Vilarino, P. Arias, Automatic detection of zebra crossings from mobile LiDAR data, *Opt. Laser Technol.* 70 (2015) 63–70, <https://doi.org/10.1016/j.optlastec.2015.01.011>.
- [135] P. Kumar, P. Lewis, C.P. McElhinney, P. Boguslawski, T. McCarthy, Snake energy analysis and result validation for a mobile laser scanning data-based automated road edge extraction algorithm, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 10 (2) (2017) 763–773, <https://doi.org/10.1109/JSTARS.2016.2564984>.
- [136] A. Hervieu, B. Soheilian, Semi-automatic road/pavement modeling using mobile laser scanning, *ISPRS Ann. Photogramm., Remote Sens. Spat. Inf. Sci.* II-3/W3 (Oct. 2013) 31–36, <https://doi.org/10.5194/isprsaannals-II-3-W3-31-2013>.
- [137] M. Cheng, H. Zhang, C. Wang, J. Li, Extraction and classification of road markings using mobile laser scanning point clouds, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 10 (3) (2017) 1182–1196, <https://doi.org/10.1109/JSTARS.2016.2606507>.
- [138] Y. Yu, J. Li, H. Guan, C. Wang, Automated extraction of urban road facilities using mobile laser scanning data, *IEEE Trans. Intell. Transp. Syst.* 16 (2015) 1–15, <https://doi.org/10.1109/TITS.2015.2399492>.
- [139] H. Wang, C. Wang, H. Luo, P. Li, Y. Chen, J. Li, 3-D point cloud object detection based on supervoxel neighborhood with hough forest framework, *Sel. Top. Appl. Earth Obs. Remote Sens.*, *IEEE J.* 8 (2015) 1570–1581, <https://doi.org/10.1109/JSTARS.2015.2394803>.