# Data-driven approach for AI-based crack detection: techniques, challenges, and future scope

Priti S. Chakurkar[1,2], Deepali Vora[3]*, Shruti Patil[4],
Sashikala Mishra[3] and Ketan Kotecha[3]

[1]Computer Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University) (SIU), Pune, Maharashtra, India, [2]School of Computer Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India, [3]Department of Computer Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University) (SIU), Pune, Maharashtra, India, [4]Department of Artificial Intelligence and Machine Learning (AIML), Symbiosis Institute of Technology, Symbiosis International (Deemed University) (SIU), Pune, Maharashtra, India

This article provides a systematic literature review on the application of artificial intelligence (AI) technology for detecting cracks in civil infrastructure, which is a critical issue affecting the performance and longevity of these structures. Traditional crack detection methods involve manual inspection, which is laborious and time-consuming, especially in urban areas. Therefore, automatic crack detection with AI technology has gained popularity due to its ability to identify degradation of roads in real-time, leading to increased safety and reliability. This review emphasizes two key approaches for crack detection: deep learning and traditional computer vision, with a focus on data-driven aspects that rely primarily on data from training datasets to detect and quantify the severity level of the crack. The article highlights the advantages and drawbacks of each approach and provides an overview of various crack detection models, feature extraction techniques, datasets, potential issues, and future directions. The research concludes that deep learning-based methods used for crack classification, localization and segmentation have shown better performance than traditional computer vision techniques, especially in terms of accuracy. However, deep learning methods require large amounts of training data and computational power, which can be a significant limitation. Additionally, the article identifies a lack of 3D datasets, unsupervised learning algorithms are rarely used to train crack detection model, and datasets having road images with variety of road textures such as asphalt and cement etc. as challenges for future research in this field. A need for 3D and combined texture datasets as challenges for future research in this field.

KEYWORDS

computer vision, crack detection, deep learning, image processing, segmentation

## 1. Introduction

Detecting cracks in vital infrastructure such as roads, bridges, and buildings cost millions of rupees annually. Catching flaws on highways and roads has also been a topic of attention for health safety to ensure driving safety. Natural calamities, such as foods and earthquakes, cause significant damage existing infrastructure. Because cracks are the most visible and widespread expression of civil work structure deterioration, developing operative methods

for monitoring and treating cracks is critical to maintaining the structural health of civil engineers (Li et al., 2020; Wang et al., 2021). One of the biggest problems in the world today is road safety. Given how frequently it is used, maintaining decent road pavement is essential to reducing accidents, and consequently, the number of fatalities. Overloading, seepage, inadequate and poor road surface drains, absence of appropriate road maintenance, a lack of proper design, and unsuitable climatic conditions, among other things, are significant reasons for rutting and degradation. Road distresses like cracking and disintegration obstruct and have a detrimental effect on traffic flow and safety, resulting in poor road performance. Early detection of road cracks is crucial to take adequate corrective actions before the issue gets out of hand and the pavement worsens. Maintenance processes typically involve a visible examination and estimate of the current state to maintain damaged infrastructure's structural and functional integrity. These damages could appear as minor or significant cracks that worsen over time, eventually leading to the structure's collapse or destruction. Cracks appear in many infrastructures (Cao M. T. et al., 2020; Munawar et al., 2021a; Wang et al., 2021) (tunnels, bridges, roads, pipelines, etc.) throughout their useful life that can reveal and enhance possible structural pathologies. Therefore, the detection of these cracks is vital in inspection work.

In most cases, the inspection is visual and performed by persons. This manual inspection is laborious, expensive, and longer duration. Due to the required knowledge, expertise, and experience, as well as human mistake brought on by fatigue and inattention, it also has a limited level of dependability, impartiality, and reproducibility. An accurate and efficient alternative to the human procedure is to use image recognition to monitor engineering structures and algorithms for machine learning to understand the photos and extract the crucial geometric details of the fracture. Automatic crack detection is essential to safeguard infrastructure's effectiveness and durability. Figure 1 depicts the real-time applications where computer-based crack detection is needed Chen Y. et al. (2021) suggested machine learning methods in surface defect detection is a key part in the quality inspection of industrial products. First, according to the use of surface features, the application of traditional machine vision surface defect detection methods in industrial product surface defect detection is summarized from three aspects: texture features, color features, and shape features. Secondly, the research status of industrial product surface defect detection based on deep learning technology from three aspects: supervised method, unsupervised method, and weak supervised method. Ali et al. (2021) studied surface cracks on the concrete structures as a key indicator of structural safety and degradation. To ensure the structural health and reliability of the buildings, frequent structure inspection and monitoring for surface cracks is important. Surface inspection conducted by humans is time-consuming and may produce inconsistent results due to the inspectors' varied empirical knowledge. The employment of deep learning algorithms using low-power computational devices for a hassle-free monitoring of civil structures.

Crack detection with digital image processing is the essential step toward automation in road health monitoring. Research and business have been discreetly moving toward developing and applying computerized road surface monitoring systems to reduce

expenses associated with manual inspection. An automated system for road crack detection must be built in four steps: acquisition of the image, pre-processing of an image, image segmentation, crack detection, and classification (Mohan and Poobal, 2018). Each of these steps has its importance in the system.

Artificial Intelligence (AI) based technology offers a more sophisticated approach for crack detection, which can execute various tasks (such as classification or regression) with exceptional performance. AI-based Crack features can be extracted using hand-crafted feature engineering with computer vision and automatic feature extraction with a deep learning approach—AI-based crack detection categories as computer vision-based crack detection and deep learning-based crack detection. Morphological operations, edge detection algorithms, support vector machine (SVM), and random structured forests (RSF) are the classical approaches used in the literature to extract hand-crafted features from asphalt pavement/road images. Wang et al. (2019) analyzed five classification algorithms, the support vector machines (SVM), neural networks (NN), random forests (RF), logistic regression (LR), and boosted tree (BT), to classify rail surface cracks. However, the practical application of this method is limited due to its slow convergence, over-fitting, and high computational cost, etc. Therefore, a fast and automatic deep learning-based feature extraction algorithm with CNN, transfer learning, and general adversarial network (GAN) (Zhong K. et al., 2020) are used to process the considerable monitoring data. Ghaderzadeh et al. (2022) proposed fully automatic deep learning based system for Acute Lymphoblastic Leukemia (ALL) diagnosis (ALL) diagnosis and subtype classification (early pre-B, pro-B, and pre-B ALL). The system's overall procedure would entail feeding the network pairs of segmented and original images, feature extraction using DenseNet-201, and then use the classification block to predict the ALL subtype based on the retrieved features. The model learns the intricate correlations and patterns between the input data and their corresponding subtypes through training on a sizable dataset of labeled images. Ghaderzadeh et al. (2021) suggested a deep learning-based model for COVID-19 identification using X-ray image. This model has the potential to enhance current testing strategies and aid in the pandemic response. Hosseini et al. (2023) study shows the proposed mobile application's potential to be an effective screening tool for hematologists and clinical professionals. The application can accurately detect B-ALL cases by utilizing preprocessing methods and deep learning algorithms, which can assist eliminate needless bone marrow biopsy cases and shorten the time required for B-ALL diagnosis. Gheisari et al. (2023) highlighted how deep learning (DL) techniques can be selected based on various issues and uses. It draws attention to the possibility for future work in developing DL frameworks and investigating novel applications, like forecasting natural disasters.

## 1.1. Significance

The importance of AI-based automatic crack detection systems is rising expeditiously hand in hand with sensor technology and the internet. Image sensors are economically helpful compared to

FIGURE 1
Crack detection application domains.
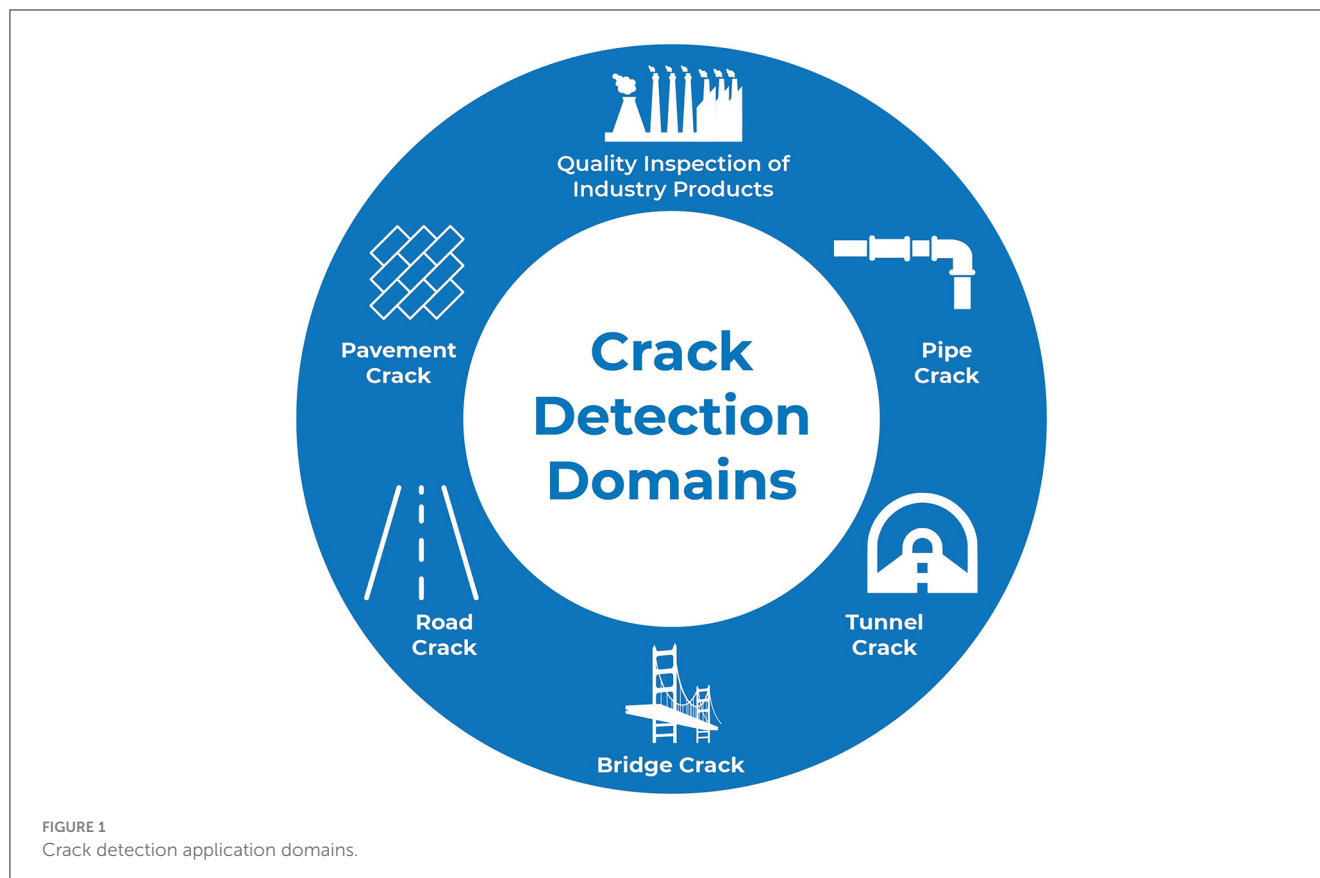
other sensors, which capture real-time images of civil structures. Crack features extracted from the images can help the timely and proactive management of structures. Research and business have been discreetly moving toward developing and applying AI-based crack detection systems to reduce expenses associated with manual inspection and ensure greater safety.

## 1.2. Motivation

There have been many reviews of the literature on image-based crack detection systems. Still, to our knowledge, challenges and probable solutions for Computer Vision-based and Deep Learning based crack detection system have yet to be the subject of a systematic literature review. Therefore, it is necessary to consider the systematic reviews of methodologies, datasets, and evaluation metrics used. This review article presents the challenges and future scope of the AI-based crack detection system, which guides researchers in constructing more effective and reliable crack detection systems in the future.

## 1.3. Prior research

Crack detection with image processing is a research area with many literature review articles. To our best knowledge, a systematic literature review has yet to be written on the data-driven approach of AI-based crack detection. Indeed, AI models' performance is

evaluated based on the quality of the input data. With a data-driven perspective, we studied how data is vital to improving the performance of AI-based crack detection. Hence, to achieve reliability and efficiency in crack detection, it becomes sensible to consider systematic reviews on the data-driven approach in this survey. Table 1 overviews survey papers studied in crack detection with image processing.

The study by Nguyen et al. (2022) evaluates the effectiveness of deep learning-based crack detection algorithms in locating cracks in asphalt pavement. This study recommends using pix2pix for crack segmentation, ResNet and DenseNet for crack classification, and Faster R-CNN for crack object detection. However, they suggest further research on unsupervised and semi-supervised learning techniques to improve fracture identification in asphalt pavement.

In a different article, Hamishebahar et al. (2022) provide a comprehensive literature review of deep learning-based crack detection research and evaluate studies that utilize the same publicly accessible datasets to determine effective crack detection strategies. Based on the trends and evaluated papers, the report suggests important avenues for future research in crack detection. Cao M. T. et al. (2020) provide a comprehensive overview of techniques for detecting road pavement cracks, including image processing, machine learning, and 3D imaging. The article compares and discusses deep learning neural networks for crack detection based on classification, object detection, and segmentation approaches, highlighting their significant improvement in detection performance. The study also evaluates

TABLE 1  Summary of existing surveys on crack detection.

| Ref. | Data set | Data collection | Preprocessing methods | AI based models | | | Learning | | | Performance metric | Challenges | Future direction |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Crack classification | Crack Object detection | Crack segmentation | Supervised | Semi/Weakly supervised | Unsupervised | | | |
| Nguyen et al. (2022) | X | ✓ | X | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | PathGANs metric is a new approach with high-grade results GPR images can be used to identify cracks inside the pavement in depth Supervised learning GANs should be used to solve crack segmentation. Semi supervised and unsupervised learning should be focus on |
| chch Hamishebahar et al. (2022) | ✓ | X | X | ✓ | ✓ | ✓ | X | X | X | X | ✓ | Deep edge detection crack segmentation task in unsupervised way, Complexity of the crack shapes in depth and the parameter correlations with the performances of the different proposed architectures and loss functions in the area. |
| Cao M. T. et al. (2020) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | X | ✓ | ✓ | ✓ | Comparison and analysis on deep learning methods and 3D image based methods Collecting 3D crack benchmark datasets will greatly benefit future study of the 3D crack detection. |
| König et al. (2022) | ✓ | X | X | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | Semi, weakly and unsupervised domain in crack detection will push the research forward and alleviate the issues that come with small datasets and difficult annotations. |

*(Continued)*

**TABLE 1 (Continued)**

| Ref. | Data set | Data collection | Preprocessing methods | AI based models | | | Learning | | | Performance metric | Challenges | Future direction |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Crack classification | Crack Object detection | Crack segmentation | Supervised | Semi/Weakly supervised | Unsupervised | | | |
| Li et al. (2022c) | ✓ | X | X | ✓ | ✓ | X | X | ✓ | ✓ | ✓ | ✓ | Analyzed crack detection through the pixel-level CIS methods to identify the improvement directions of the future work to better use DL method to enhance the performance and efficiency of CIS |
| Hsieh and Tsai (2020) | X | ✓ | X | ✓ | ✓ | ✓ | X | X | X | ✓ | ✓ | ML crack detection algorithms, reviewing 68 ML-based crack detection papers to identify the current development trend, pixel level crack segmentation. Performance comparisons among 8 DL crack segmentation models were then conducted using consistent evaluation metrics and real-world 3D pavement images under diverse conditions |
| Munawar et al. (2021b) | ✓ | ✓ | X | ✓ | X | X | X | X | X | X | X | Need to perform severity assessment of the detected cracks |

the performance of these approaches on widely-used benchmark datasets and covers performance evaluation measures.

Golding et al. (2022) propose using convolutional neural networks (CNN) as a deep learning-based technique for fracture detection in infrastructure and comparing grayscale and RGB models using various image processing methods. The results indicate that DL crack identification does not rely on color, as grayscale models perform similarly to RGB models, thresholding, and edge detection models perform worse than RGB models. König et al. (2022) discuss the importance of early surface crack detection and monitoring for structural health monitoring and provide a review of deep learning-based crack analysis algorithms. The study covers a range of tasks, including crack classification, detection, segmentation, and quantification, and offers thorough analyses of current fully, semi-, and unsupervised techniques. The review also includes measures used to assess algorithm performance and well-known datasets used for cracking.

Hu et al. (2021) discuss using deep learning models to identify asphalt pavement cracks, highlighting the issues with conventional artificial detection systems. They show that pavement fracture detection using the YOLOv5 series deep learning model has produced positive results, with the YOLOv5l model having the best detection accuracy (88.1%) and the YOLOv5s model having the quickest detection time (11.1 ms per image). In the study published by Li et al. (2022a) examine the importance of fracture detection in transportation infrastructure and the growth of deep learning-based techniques for crack image segmentation (CIS). They conduct a thorough analysis of over 40 papers on DL-based CIS methods released in the previous three years, categorizing them into ten themes based on backbone network design, including FCN, U-Net, multi-scale, attention mechanism, transformer, and weakly supervised learning, among others.

In this study, Hsieh and Tsai (2020) provide a thorough analysis of recent machine learning-based crack detection algorithms, with a special focus on pixel-level crack segmentation. They evaluate eight ML-based models using standardized evaluation criteria and 3D pavement photos with various conditions, showing that deeper backbone networks and skip connections improve performance in FCN models. The suggested algorithm tackles the false-positive issue as a necessary first step to enhance ML-based crack detection models.

## 1.4. Research goals

This systematic literature review (SLR) aims to understand recent developments in the field of computer vision-based road crack detection techniques and identify unresolved problems and obstacles within. The AI-based crack detection system survey's research goals are listed below.

- What are the different artificial intelligence-based approaches used for crack detection?
- What are the different datasets available for research purposes in AI-based crack detection?
- How to accurately measure the segmented crack parameters to assess the cracks' severity?

- What performance evaluation indicators are employed to assess the effectiveness of the AI-based crack-detecting system?
- How to optimize AI-based crack detection models be implemented with limited computational resources?

## 1.5. Work's contribution

The following are the critical findings of this in-depth literature review:

- To conduct a literature study integrating deep learning and computer vision-based crack detection systems.
- To provide a depth summary of the existing research on deep learning and computer vision-based crack detection systems with an emphasis on methods, data sets, applications, challenges, and future directions.
- We also provide an overview of publicly available datasets that can support research in the crack detection domain.
- To evaluate this method's potential applications for automatic road health monitoring, as well as its current advantages and disadvantages.

This article is a comprehensive and well-structured study on image-based road crack detection, covering various aspects such as a literature review on data-driven AI approaches, computer vision techniques, challenges, and future scope.

## 2. Research strategy

Published research is accessible based on core metrics used in the bibliometric analysis. As part of this process, we attempt to select the most well-known or active researchers and their affiliations, collaborative patterns, frequently used phrases, and numerous articles about them. We termed the initial step of our research process as "image-based crack detection." "The phrase crack-detection" is used in civil and structural sectors such as tunnels, bridges, and railway tracks for structural health monitoring.

## 2.1. Database selection and query formulation

This research aims at finding the use of artificial intelligence in the image-based crack detection domain; we started with bibliometric analysis for image-based crack detection as a primary keyword ANDing with "Computer Vision," "Deep Learning," "Machine Learning," Artificial Intelligence," as a secondary keyword in the research space. As a result, these primary and secondary keywords are used to frame the query, and relevant articles were initially gathered from the Scopus database. The primary and secondary keywords used in questions to select the data from Scopus databases using AND and OR Boolean operators are shown below.

TITLE-ABS-KEY (crack detection) and ((computer vision) or (deep learning) or (machine learning) or (artificial intelligence)) and(image).

After that, the filtration technique is used for the article collection to enhance the outcomes that satisfy our main objectives. This procedure's steps are removing duplication, applying exclusion and inclusion criteria, filtering based on titles and abstracts, and full-text screening. Inclusion criteria used to filter the relevant documents, such as

- Publication Articles must be published between the years 2012 to 2022.
- The subject area should be Computer science, decision sciences, material sciences, and multidisciplinary articles.
- The article should match a minimum of one of the search terms.
- The article type should be conference papers or review articles.

Exclusion criteria are foreign language research articles and articles unrelated to research questions.

## 2.2. Analysis of the information

In this review article, bibliometric analysis was done on filtered reports from the Scopus database using various parameters, including:

- Subject area
- Research trends in AI-based technologies
- Documents by country
- Keyword co-occurrences.

### 2.2.1. Subject area

To begin with, an analysis of the data revealed that the keyword "image-based crack detection" is most frequently utilized within the domain of Computer Science, followed by Engineering and Material Science. Material science is essential to the field of structural health monitoring (SHM) because of how much it affects a structure's performance, dependability, and durability. Since image-based crack detection is primarily a computer vision problem, it is unsurprising to observe the highest concentration of research activity within the Computer Science subject. This trend is visually represented in Figure 2.

### 2.2.2. Research trends in AI-based technologies

As shown in Figure 3, the graph demonstrates the trends that focus on the number of research articles retrieved from the Scopus database with primary keyword image-based crack detection with secondary keywords such as computer vision (CV), machine learning (ML), deep learning (DL), and artificial intelligence (AI). The number of research papers focusing on



FIGURE 2
Analysis of AI-based crack detection research in various subject areas.
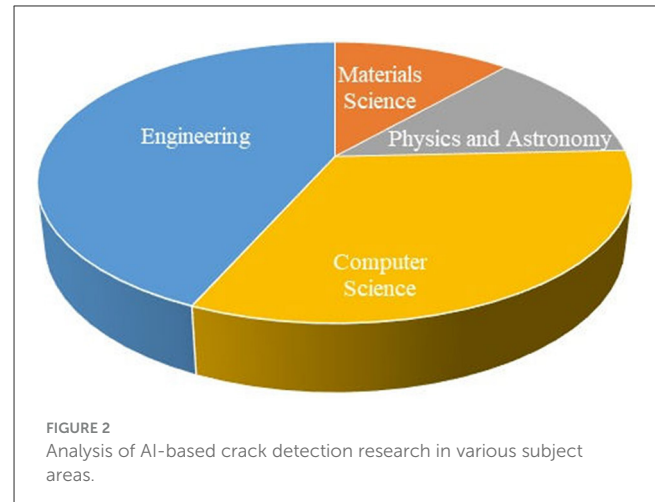
crack detection using CV keywords increased from below 50 in 2018 to over 150 in 2022, with more than 50 research papers already published in 2023. Similarly, AI keyword usage increased from around 20 research papers to approximately 50 in 2022, with about 30 published in 2023. ML keyword usage rose from 30 research papers in 2018 to over 100 in 2022, and around 60 were published in 2023. Notably, DL keyword usage increased from approximately 30 research papers in 2018 to 350 in 2022, and it has been used in over 160 research papers in 2023. These trends indicate a growing interest in and utilization of CV, AI, ML, and DL techniques in crack detection research.

### 2.2.3. Documents by country

The graph shown in Figure 4 represents a country-wise distribution of research papers focusing on AI-based crack detection. Among the countries mentioned, Hong Kong is associated with around 100 research papers, India with approximately 400 research papers, and notably, China with over 2600 research papers. The graph indicates a significant research output in AI-based crack detection from these countries. While the specific numbers for other countries are shown above, it is evident that multiple nations, including Italy, Germany, Japan, Australia, Canada, the United Kingdom, South Korea, and the United States, are actively engaged in this field. However, China has emerged as a dominant player with many research papers, showcasing its strong presence and advancing AI-based crack detection.

### 2.2.4. Keyword co-occurrences

Analysis of co-occurrence is performed on phrases using a VOS viewer. Keywords occurring with more than 50 are only considered for the analysis. 13,242 keywords out of 66 met the criteria. Keyword co-occurrence network visualization is shown in Figure 5 below and identifies terms that can be utilized more frequently in this research.
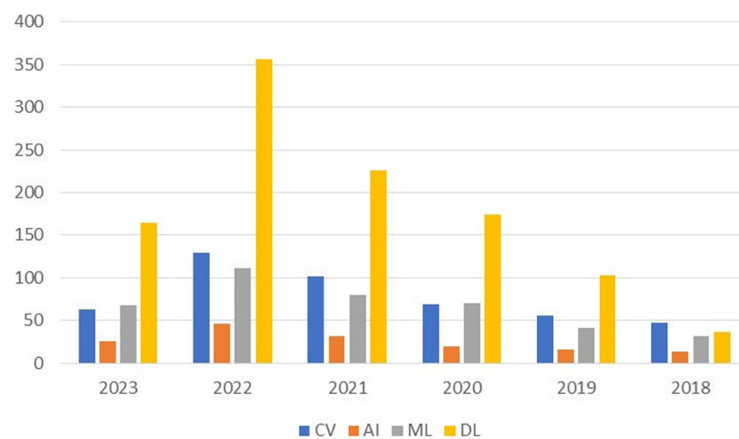
FIGURE 3
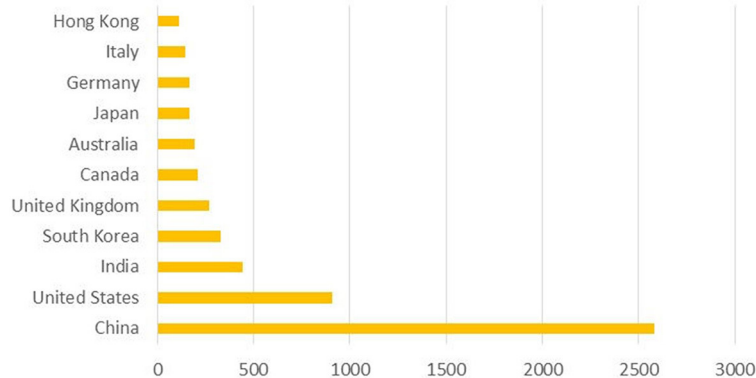Research trends in image-based crack detection with different AI-based technologies.



FIGURE 4
Country-wise distribution of research papers focusing on AI-based crack detection.

# 3. Data-driven AI-based imaging system for crack detection

Data is crucial in AI research, especially for image processing and crack detection tasks. Obtaining high-quality data is vital because the accuracy and effectiveness of AI models heavily rely on the data they are trained on Figure 6. Highlighted crack detection process from the data perspective, and the survey is presented in different data-oriented aspects used for crack detection.

Figure 6 shows an AI-based data-driven crack detection process. The process can be broken down into four stages, each of which is essential in detecting cracks. The first stage involves data collection, labeling, and dataset building. Data is collected using UAV-mounted cameras and handheld devices. The collected data is labeled at the pixel, object, and image levels, and private and public datasets are used. This stage is critical for ensuring that the data used for crack detection is accurate and reliable. Data augmentation, preprocessing, and learning algorithms are used in the second stage. Data augmentation is done using Generative Adversarial Networks (GANs), which help generate synthetic data

and improve the system's accuracy. Preprocessing is done using Histogram Equalization Filters, which normalize the images and improve their contrast. Learning algorithms such as supervised, weakly supervised, and unsupervised are used to train the system. The third stage involves crack classification, object detection, and segmentation. Transfer learning techniques such as VGG and MobiNet are used to develop the crack classification model. Crack object detection uses techniques such as YOLO, FASTER RCNN, and SSD.

Crack segmentation is performed using FCN, DEEPV LAN, and encoder-decoder techniques. These techniques are used to identify and locate cracks in the images accurately. The fourth and final stage uses evaluation metrics, attention techniques, and crack severity qualification. Evaluation metrics such as IOU, MAP, and AUC are used to assess the system's performance. Attention techniques such as SENET, CBAM, ECANET, and COORD ATT are used to improve the system's accuracy. Finally, crack severity qualification is done by measuring the detected cracks' length, width, and depth. This information can then be used to prioritize repairs and maintenance.

FIGURE 5
Network visualization diagram for keyword co-occurrence.
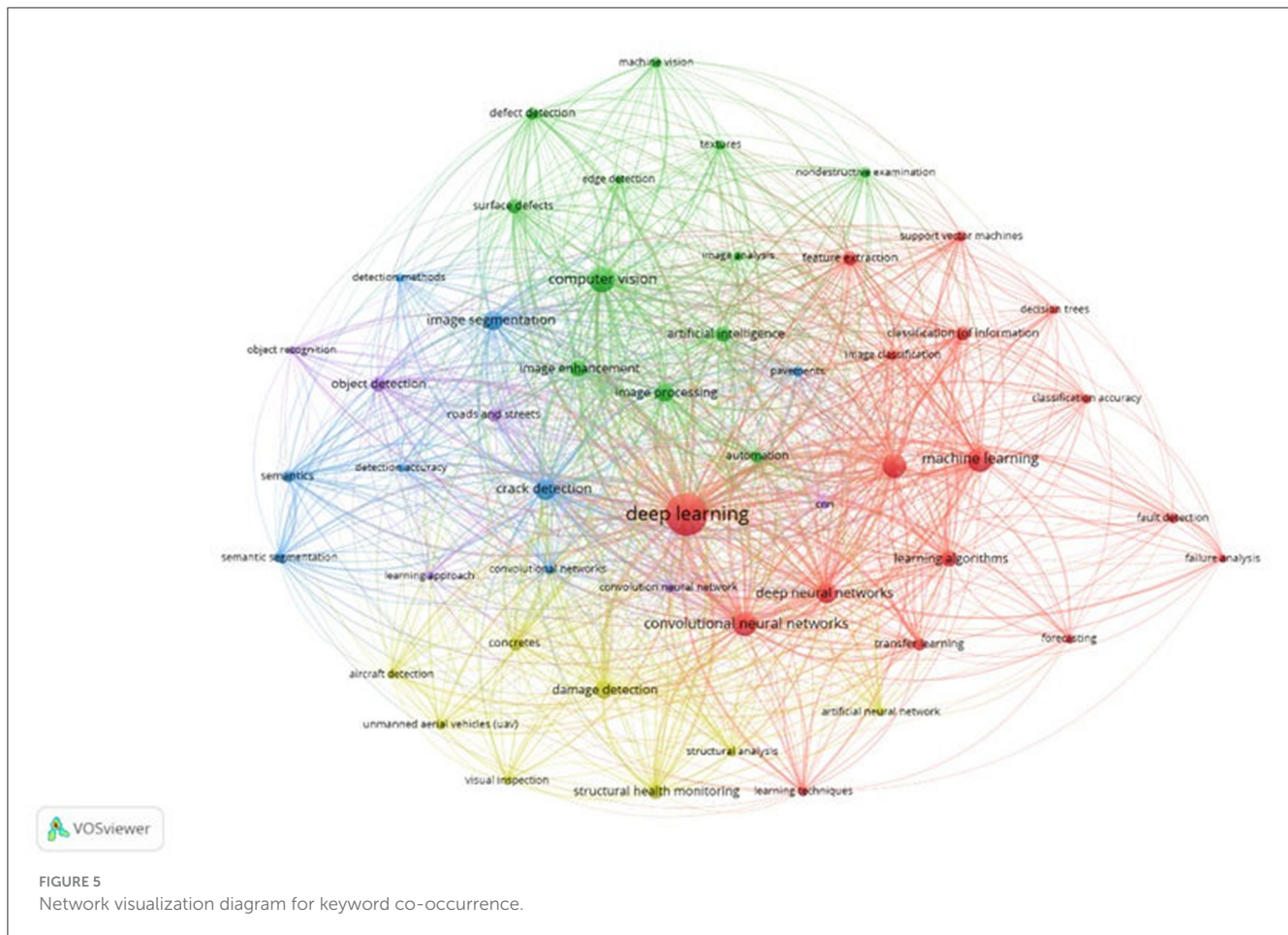
Overall, this approach leverages AI-based techniques for efficient and accurate crack detection, which can be applied in various industries and infrastructures, such as roads, bridges, and buildings, to ensure their safety and longevity.

## 4. Literature survey on data-driven AI-based imaging system for crack detection

The term "Image-based crack detection system" describes the full spectrum of activities, from taking images to classifying cracks according to their severity. This system provides the economic and engineering analysis tools required for making cost-effective maintenance, rehabilitation, and reconstruction decisions. A smart road monitoring system based on Industry 4.0 was developed. In addition to video, mobile data, weather data, and other sensor data, the monitoring system collects substantial amounts of data from the road environment infrastructure. Monitoring systems collect data on road environment defects that provide road environment safety data.

Consequently, road environment monitoring systems may include pavement and bridge crack detection. Internal, invisible faults and surface apparent defects include pavement and bridge defects. Surface obvious crack faults have been a long-standing issue that endangers public safety. The photographs of the road crack were taken in different lighting conditions (round the clock). A few of them included undesirable things, e.g., random particles, textures, heterogeneity, uneven lighting with variations in the road's surface, lines, boisterous environment, shadows, water, tire prints, oil slicks, etc. The result is a challenging problem in choosing a uniform threshold building. A successful preprocessing stage is crucial for getting good results because of the segmentation step. This action involves bringing edges, borders, or contrast to sharper focus analysis (Fu et al., 2020). There are numerous types of image enhancement like fuzzy edge removal, noise reduction, magnification, contrast stretching, filtering, artificial interpolation coloring, transform operations, histogram modeling, and fake color etc.

Automated image-based crack feature extraction methods can be divided into deep learning and traditional approach. The conventional systems mainly include the threshold approach, wavelet transform, the minimum path, etc. Traditional visual methods primarily rely on hand-craft features, with distinguishing abilities in the images, such as the grayscale, texture, and contour shape of defects, because of the road's intricacy, variety of topology, random forms, and sizes, as well as oil stains, weed stains, and other significant disruptions. Hand-craft features tend to fail, and the algorithm needs to be redesigned.
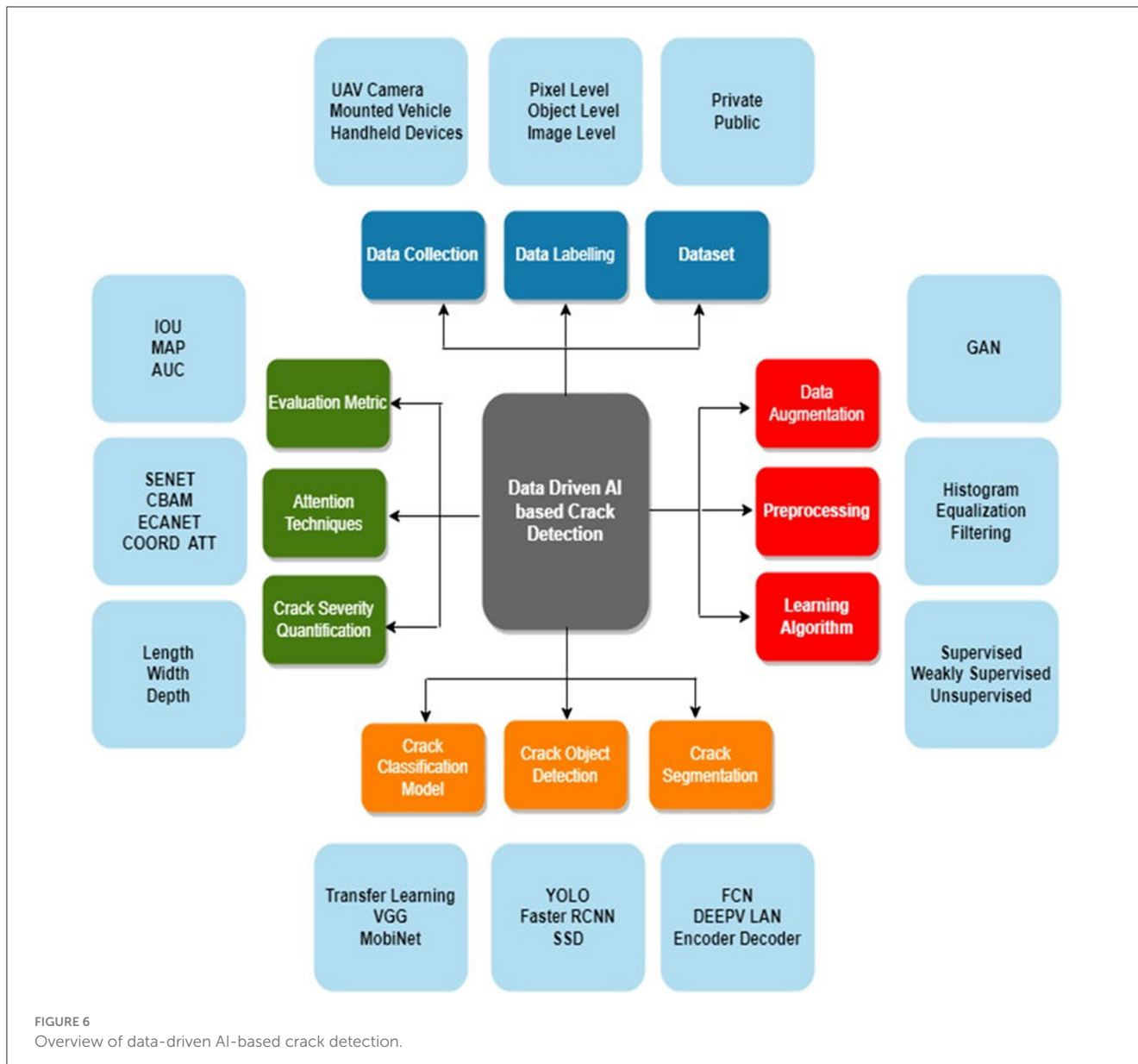
**FIGURE 6**
Overview of data-driven AI-based crack detection.

Therefore, traditional crack feature extraction methods have weak generalization ability and low efficiency for different road images in complex situations. Hence, research requirement is still active due to low robustness and fluctuating environment. Deep learning can resolve complex problems automatically with the help of AI. Various methods of detecting road cracks using deep learning have been classified according to their accuracy:

*Patch level crack detection – likely to be binary or multi-crack classification.*

*Object Level-Likely to be boundary box regression.*

*Pixel-level crack detection - likely to be a semantic segmentation task*

In recent studies, several methods for crack detection have been proposed using deep learning. The process may be based on the detection of objects or the segmentation of blocks of images. AI based crack detection process consists of four steps: data collection,

data pre-processing, dataset modeling and crack classification as shown in Figure 7.

## 4.1. Data collection

The data collection is the initial stage that feeds the input images to the entire crack detection cycle. Automating data acquisition has led to the development of complete systems carried out by vehicles (or, more recently, on devices such as smartphones or crewless aerial vehicles) for visual surface surveying. A computer, a Global Positioning System (GPS) sensor, and imaging sensors are all installed in the car. Several sensing technologies, sensor positions, and vehicles have been deployed to collect data to evaluate surface conditions. Table 2 summarizes the sensor systems for data collection.

FIGURE 7
AI-based crack detection tasks and algorithms.
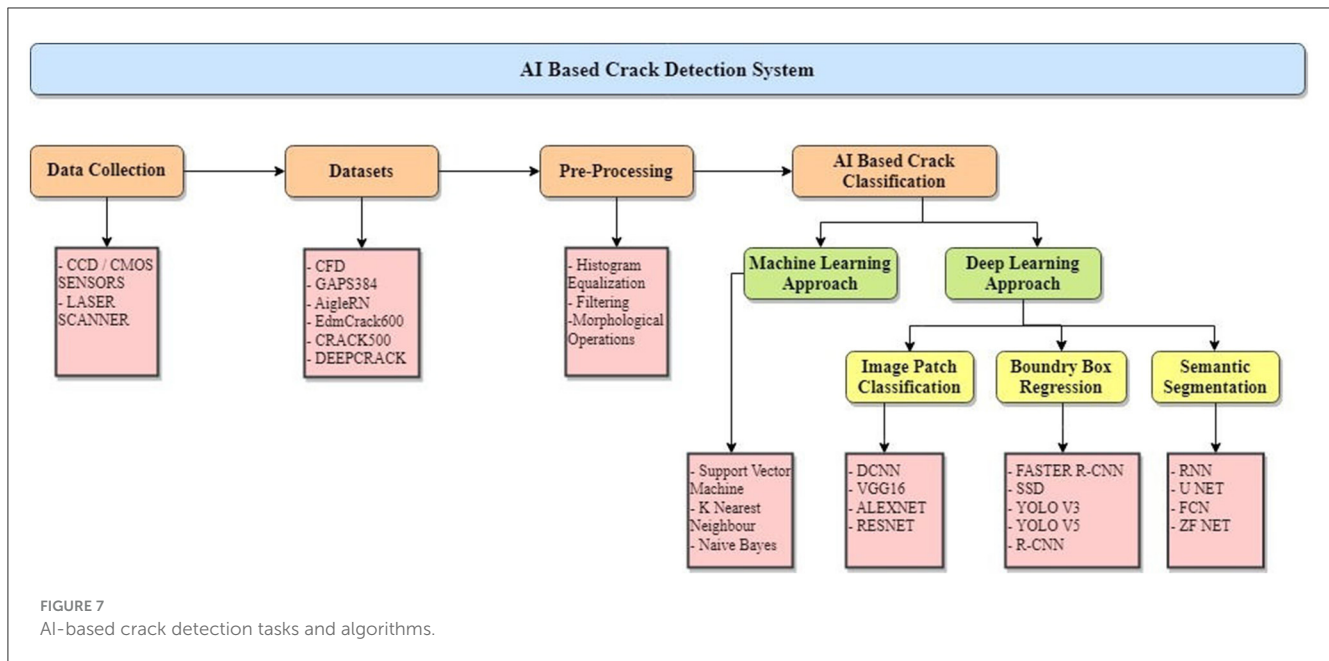
To ensure the safety, maintaining structural integrity, and lowering maintenance costs, cracks should be found and examined. Conventional techniques of fracture detection can be costly, time-consuming, and frequently necessitate stopping operations. However, the advent of AI-based data-driven crack detection systems, which can be quickly and precisely identify cracks, is a result of technological breakthroughs. Data collection, which uses sensor systems for fracture detection, is the basis of these systems. UAVs, camera-mounted vehicles, and handheld devices are just a few of the crack-detection sensor systems available for data collection. Each of these methods has its benefits and drawbacks, and choosing one depends on several considerations, including the level of accuracy that is sought, the difficulty of the terrain, and the available budget.

One of the most often used sensor systems for crack detection is crewless aerial vehicles (UAVs). They have cameras, sensors, and GPS and can take pictures at different angles and heights (Li et al., 2023). Depending on the type of camera and sensors being utilized, UAVs can capture both 2D and 3D images (Khaloo et al., 2018). The visual view can be in the top, front, or both directions at once. An ordinary UAV-based crack detection system typically consists of the UAV itself, a camera or sensor payload, a GPS, and a ground control station. A pilot controls the flight path and camera settings while controlling the UAV remotely. The application and required level of precision can influence the camera or sensor payload. High-resolution cameras and sensors are frequently utilized to take precise surface images for crack identification (Byrne et al., 2017). The UAV's location is tracked using GPS, ensuring reliable data collection. The ground control station tracks the UAV, gathers data, and performs data processing.

The capacity of UAVs to capture images from various angles and heights, which can give a more thorough perspective of the surface being investigated, is one benefit of employing them for data collection. UAVs can check substantial infrastructure projects like highways and bridges since they can swiftly and effectively cover large regions (Munawar et al., 2021a). However, the cost of purchasing and operating UAVs can be high, and their usage might be prohibited in some places due to laws or safety concerns.

Another alternative for gathering data on crack detection is a vehicle with surveillance cameras. Typically, they include a car with cameras or sensors that take pictures of the examined surface (Montero et al., 2015). Depending on the camera and sensors utilized, camera-mounted vehicles can record both 2D and 3D images. The top view, front view, or a combination of the two can be the image view. The car, a camera or sensor payload, and a computer system for data processing make up a conventional camera-mounted vehicle-based crack detection system. The driver, who controls the vehicle's direction and speed, is in charge of it. The application and required level of precision can influence the camera or sensor payload. High-resolution cameras and sensors are frequently utilized to take precise surface images for crack identification. The computer system does the data collection and processing. Data collecting with camera-mounted vehicles has the benefit of minimal energy usage and simple control. They are appropriate for checking both horizontal and vertical surfaces and can rapidly and effectively cover huge regions (Cao W. et al., 2020). The usage of camera-mounted vehicles, however, can be restricted in some situations, such as enclosed locations, and they might not be able to take pictures from specific angles or heights.

Another form of sensor system for crack detection uses handheld devices. An operator can use these portable tools to take pictures of surface cracks (Sony et al., 2019). Mobile devices with high-resolution cameras, such as smartphones and tablets, are the most often utilized handheld devices for crack detection. These tools can take pictures of the cracks in 2D and 3D (Chen et al., 2014). The portability and simplicity of use of mobile devices are advantages. They are a practical choice for small-scale crack detection applications because they are reasonably priced and easily accessible. The drawback of handheld equipment is that an operator is needed to take the images. This indicates that gathering data

TABLE 2 Sensor systems for data collection in road crack detection.

| Crack detection sensor system | Components | Image type | | Image view | | Advantages | Limitations |
|---|---|---|---|---|---|---|---|
| | | 2D | 3D | Top view | Front view | | |
| UAV (Ersoz et al., 2017; Tse et al., 2023) | Cameras, sensors, GPS, etc. RGB images, IMU | ✓ | ✓ | ✓ | ✓ | Easy to capture images in dangerous and/or unreachable places | Images collected by UAV can be more challenging due to lighting conditions, viewpoints, and scales |
| Camera mounted vehicle (Mei and Gui, 2020; Yang N. et al., 2022) | Camera, GPS | ✓ | ✓ | X | ✓ | Low energy consumption and easy control | Difficult to capture images in dangerous situations |
| Handheld devices (Kong et al., 2017; Jeong et al., 2020) | Smartphones | ✓ | X | X | ✓ | Simple and low cost | Need to compromise with complete view, and image quality, Not practical for large-scale projects |

can take a long time and may need to be more practical for large-scale projects. The accuracy of the crack identification algorithms may also need to be improved by the lower image quality of handheld devices compared to UAVs or vehicles with mounted cameras (Jordan et al., 2018). Terrestrial Laser Scanning (TLS) is used to assess road surface conditions. A laser line scanner, High-resolution, continuous 3D road surveys were produced by using a laser scanner or digital camera placed on a moving vehicle. Zhong M. et al. (2020) showed a multi-sensor, laser scanning approach for crack identification by merging to generate a 3D simulation of the highway, laser line scan data was combined with video that recognized fissures as small as a few centimeters in depth as well as their location.

Hassan et al. (2022) proposed three-dimensional reconstruction systems, the LCMS system (Laser Crack Measurement System), consisting of two three-dimensional laser profile-meters capable of measuring the cross-sectional profiles of a road with a resolution of 1 mm. This system collects intensity information, as well as geometric report of the road surface, which allows for characterizing and displaying images together with the shape (texture) of the road. The crack's depth, which is regarded as a crucial primitive, can be calculated with the use of the image's 3D representation. This feature is used to categorize crack types as small, moderate, or severe. With the advent of 3D road scanners, it is anticipated that, in the near future, a higher level of accuracy can be attained by utilizing the depth information from the collected 3D model of the road pavement. Since there are currently no benchmark datasets for 3D crack identification, collecting these datasets will be very helpful for future research.

Ultimately, data collection for AI-based crack detection systems relies on crack detection sensor systems. UAVs, vehicles with cameras placed on them, and handheld gadgets are the three most often employed sensor systems. Due to their swift ability to cover broad regions, UAVs are well-suited for large-scale projects, although they can be expensive to operate and require trained operators. Mid-scale projects can benefit from camera-mounted vehicles since they are convenient, inexpensive, and straightforward. Handheld devices offer a cost-effective alternative for small-scale projects, but they need an operator to take the pictures. The choice of which sensor system to utilize relies on the scope and complexity of the project. Each sensor system has advantages and limits. The project's size, complexity, and resource availability all affect the sensor system that will be used to collect data. AI-based crack detection technologies could revolutionize our ability to detect and maintain infrastructure, guaranteeing that our buildings, bridges, and roads are secure and safe for years to come.

## 4.2. Data labeling and dataset generation

Data labeling is the key to training the detection model, and the quality of labeling determines the accuracy. The three tracks of classification, detection, and segmentation are distinct and require different methods of data labeling. Image classification

annotation involves assigning a binary label to indicate the presence or absence of an object in the image at a high level. In pavement damage classification, a standard method is dividing a large image into smaller sub-images, called patch-level classification. Object detection annotation is done at the level of individual objects, requiring the identification of the object's category and its position in the image, typically represented by a bounding box. Segmentation annotation involves labeling each pixel in the image with its corresponding category or background and is done at a pixel-level granularity. Patch-level classification and object-level detection provide information about the localization of objects. Table 3 summarizes the datasets specified in the literature for crack detection with the level of annotations used. For data labeling, the labeling tools were studied in the literature to manually create image labels to produce a crack dataset in PASCAL VOC data format. The labeling tool will generate XML files based on the rectangular boxes, including crack category, size, and location. We should avoid other objects entering the rectangular box during the data labeling, leading to a high false detection rate. For many years, road crack detection systems have been a study issue. There are several public datasets available to assist us in conducting improved research. Shi et al. (2016) this study advances the field of road cracking detection through presenting a new annotated dataset (Crack Forest Dataset) and an effective detection technique that surpasses competing options in terms of noise reduction and processing time. The study's emphasis on two evaluation indicators, the "Continuity index (CI)" and "Crack detection accuracy" to assess accuracy, ensures an accurate assessment of the suggested method's efficacy. Liu et al. (2019) study makes a substantial contribution to crack detection and segmentation by presenting the DeepCrack benchmark dataset, a ground-breaking public benchmark aimed at promoting standardized evaluation and method comparison. The dataset includes several, thoroughly hand-annotated crack images in a variety of scales and settings. In contrast to earlier studies, which relied on limited, narrow evaluation datasets, the new benchmark database, which consists of 537 RGB color photos with manual annotations. Arya et al. (2021) proposed RDD2020 dataset to facilitate the development and improvement of deep learning algorithms in order to advance automatic road damage identification. It stimulates academics to look into cutting-edge methods for precisely identifying and locating road damage. The RDD2020 dataset is rigorously built to contain a varied collection of incidents of road damage, recorded across various road types, conditions, and environmental factors. Bounding boxes that exactly define sections of road damage are carefully inserted to each image in the dataset. For the sake of testing and refining deep learning algorithms, these annotations serve as the ground truth.

## 4.3. Data preprocessing in crack detection

Road image preprocessing is essential in automated pavement crack detection using image processing and deep learning algorithms. This process involves various techniques that aim to enhance image quality, reduce noise, and improve the visibility of cracks in the image. One of the primary steps is converting the colored image into grayscale, simplifying the image analysis process. Histogram equalization balances the brightness and contrast of the grayscale image, making cracks more visible to algorithms. Filters like Gaussian, Median, and Sobel are commonly applied to remove noise and highlight edges, making it easier for algorithms to detect cracks. Morphological transformation using techniques like erosion, dilation, opening, and closing helps to remove unwanted objects and enhance essential features. The ultimate goal of road image preprocessing is to improve the accuracy of automated pavement crack detection systems by highlighting the elements of interest, such as cracks, in the input image. The process can significantly improve the efficiency of assessing road conditions and ensuring safety.

Yang Z. et al. (2022) have developed a method for pavement crack localization and segmentation algorithms that aims to automate the detection of cracks on roads, improve safety, and reduce maintenance costs. The proposed technique involves a three-stage algorithm based on digital image processing and deep learning, which includes a two-dimensional discrete wavelet transform to obtain low and high-frequency coefficients of the wavelet, Retinex algorithm, guided filtering based on wavelet transform, and soft threshold filtering for denoising. The ultimate goal of this approach is to enhance the accuracy and efficiency of pavement crack detection and segmentation algorithms. On the other hand, Chen et al. (2022) proposed a crack detection method using image processing techniques to improve road safety and reduce maintenance costs. The method involves a potential crack region method that uses multiple thresholds for crack detection, in which histogram equalization is used to adjust the grayscale value distribution and enhance local contrast for better distinction between the crack and background. The technique also uses mean filtering to remove noise and improve image quality, which aims to improve the accuracy and efficiency of crack detection on road surfaces.

In the study, Li et al. (2022a) propose a new technique that combines convolutional neural networks and hybrid image processing to improve the accuracy of crack classification and segmentation on concrete bridge images. To achieve this, they use a bilateral filtering technique to sharpen the crack details and highlight their characteristics while minimizing the influence of factors such as rain or stains. The processed images are then converted to grayscale and subjected to contrast enhancement to enhance visual performance. This proposed technique aims to make the maintenance of concrete bridges more efficient and cost-effective. Similarly, Parrany and Mirzaei (2022) offer an image processing strategy that uses the Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm and non-linear diffusion filtering techniques to improve the accuracy and efficiency of surface crack identification in building structures, especially under non-uniform illumination conditions.

Automated crack segmentation aims to increase the precision and efficiency of detecting cracks in building structures. Nnolim (2020) proposes a method that uses saturation channel threshold, area classification, and a modified level set segmentation fusion with Canny edge detection. The studied article use global histogram equalization to normalize the distribution of intensities and the

**TABLE 3 Summary of datasets used for crack detection.**

| Dataset | Basic features | | | Advanced features | | | Dataset | Basic features | Advanced features |
|---|---|---|---|---|---|---|---|---|---|
| | Image/Sample no. | Resolution | Colored | Image/Sample no. | Resolution | Environmental or other interfering factors | | | |
| Crack forest dataset (CFD)1 | 118 | 480 × 320 | Crack Forest Dataset (CFD)1 | 118 | 480 × 320 | Lane lines, shadows, noise like oil stain | For Non-commercial research purposes only · | iPhone5 | China |
| AigleRN2 | 38 | 991 × 462 311 × 462 | AigleRN2 | 38 | 991 × 462 311 × 462 | Pre-processed to reduce the non-uniform illumination | With more complex texture than CFD, ·Similar small data sets like ESAR and LCMS | Professional Camera | French |
| Crack5003 | 500 | 2000 × 1500 | Crack5003 | 500 | 2000 × 1500 | Lane lines, oil/wet stain, tire brake marks, speckle noise, etc. | ·Covers almost all kinds of features except for shadow | LG-H345 Cell Phone | US |
| GAPs3 | 1969 | 1920 × 1080 | GAPs3 | 1969 | 1920 × 1080 | Little noise like oil stain, asphalt and rails, lane | – | JAI Pulnix TM2030 monochrome cameras, Kodak KAI2093 | Germany |
| Cracktree2003 | 206 | 800 × 600 | Cracktree2003 | 206 | 800 × 600 | only shadow & lane lines | – | Unknown | – |
| G454 | 122 | 2048 × 1536 | G454 | 122 | 2048 × 1536 | Brightness change, oil stain, tire brake marks, speckle noise, etc. | ·4 types: transverse, longitudinal, block, and alligator cracks | Unknown | China |
| EdmCrack6005 | 600 | 1920 × 1080 | EdmCrack6005 | 600 | 1920 × 1080 | Weather, illumination, shadow, texture difference | ·Only for academic research ·Perspective image taken from the rear camera of a car | Rear Camera of Car – GoPro 7 | Canada |
| JapanRoad6 | 9053 | 600 × 600 | JapanRoad6 | 9053 | 600 × 600 | Real street view perspective images captured from the front windshields of a car | ·License: CC BY-SA 4.0 ·For road defect detection ·PASCAL VOC annotation format | Front windshields of Car LG-5X | Japan |
| SDNET20187 | 230 (Cropped to 56092) | 4096x3840 (256 × 256) | SDNET20187 | 230 (Cropped to 56092) | 4096x3840 (256 × 256) | with a variety of obstructions, including shadows, surface roughness, scaling, edges, holes, and background debris | ·License: Creative Commons Attribution 4.0 ·54 bridge decks, 72 walls, 104 pavements ·Crack size: 0.06-25mm | 16-MP Nikon camera | US |
| Concrete crack images for classification 8 | 458(cropped to 40000) | 4032 × 3024 (227 × 227) | Concrete crack images for classification 8 | 458(cropped to 40000) | 4032 × 3024 (227 × 227) | Shadows, lighting spot, blurred, including thin & close-up cracks | ·Similar to SDNET2018, images collected from METU Campus Buildings | Professional Camera | Turkey |

wavelet-based enhancement algorithm to enhance crack detection by amplifying detail coefficients, suppressing approximation coefficients, and increasing contrast and sharpness. This method, known as the modified HSI-based crack detection, aims to improve crack detection accuracy and efficiency. On the other hand, Yu et al. (2022) propose a deep learning-assisted image processing technique. The approach involves using a mask filter to remove handwritten marks and a ratio filter to eliminate speckle linear noise from the input images. The use of deep learning algorithm is to classify images, identify regions of interest, then segment the images using threshold techniques, and perform crack quantification using image features. The proposed method aims to enhance the accuracy and speed of crack detection and quantification in concrete bridges.

The study proposes an advanced deep-learning fusion model (Feng et al., 2020) for more precise pavement crack detection and segmentation. The process starts with applying median and bilateral filters to eliminate noise, followed by contrast enhancement to improve the quality of the images by highlighting crack information. The fusion model combines convolutional neural networks and fully connected networks to achieve better results. Its main goal is to enhance the efficiency and accuracy of pavement crack detection and segmentation algorithms.

### 4.3.1. Data augmentation

Data augmentation techniques can be highly beneficial when you have limited training data for tasks like artificial crack detection. Data augmentation involves applying various transformations or modifications to existing data to create additional training examples, effectively expanding the dataset without the need for new annotations. This process helps improve the model's performance by providing more diverse and varied examples. Data augmentation is a machine learning technique that artificially generates additional synthetic training data through label-preserving transformations (Bayer et al., 2023). Though data augmentation techniques have achieved great success in computer vision applications, Data augmentation is performed by rotating the images with different angles, adding some noise, etc. The CrackNet, a deep-learning framework for crack detection and classification leverages data augmentation techniques to improve the model's performance and demonstrate the effectiveness of CrackNet on the CFD2014 dataset.

In order to improve the amount and quality of a training dataset and create a more effective deep learning model, data augmentation is an essential deep learning technique. In order to overcome the scarcity of training data and eliminate overfitting, Islam et al. (2022) used data augmentation and transfer learning. The first suggested data augmentation strategy is the Random-Resized-Crop Method, which randomly selects a section of the image to crop and resizes it to the required size. The second method, known as the Random-Rotation Method, rotates the image across a variety of random angles. Angles between 15 degrees and +15 degrees are chosen for our training dataset. The Color-Jitter Method, which randomly modifies the brightness, is the third data augmentation technique employed.

## 4.4. AI-based classification algorithms in crack detection

AI-based crack classification strategies employ various techniques to identify and classify image cracks accurately. One approach is image patch crack classification, where the image is divided into smaller patches, and each patch is analyzed independently to determine if it contains a crack. Another strategy involves boundary box regression, which aims to detect the presence of cracks by predicting the bounding boxes that encompass them. This allows for precise localization and identification of crack objects within the image. Additionally, semantic segmentation is employed to classify each pixel in the image as either crack or non-crack, enabling a fine-grained understanding of crack distribution. By combining these strategies, AI systems can detect, classify, and analyze cracks, facilitating efficient and reliable inspection and maintenance in various domains such as infrastructure, manufacturing, and construction.

### 4.4.1. Image patch classification

In recent years, there has been a growing interest in data-driven approaches for various applications. This approach utilizes machine learning techniques to analyze data and make predictions or decisions based on the insights generated from the data. In this summary, we will discuss several papers that focus on data-driven approaches and their applications. Table 4 represents the different parameters that affect the accuracy of the AI-based crack detection system.

The paper by Hammouch et al. (2022) proposes an automated methodology for detecting and classifying cracks in Moroccan flexible pavements using Convolutional Neural Networks (CNN) and transfer learning. The periodic inspections of Moroccan roads using high-resolution cameras and GPS/DGPS receivers have been conducted since 2011, but the manual processing of pavement surface image sequences is complex, time-consuming, and subjective. The proposed approach shows good crack detection and classification results using both the CNN and pre-trained VGG-19 models. The study demonstrates the potential of data-driven approaches, specifically CNN and transfer learning, in automating crack detection and classification to diagnose road networks effectively.

The paper by Amhaz et al. (2016) proposes a new algorithm for automatic crack detection from 2D pavement images using minimal path localization. The approach selects a set of minimal paths and introduces two post-processing steps to improve detection quality. The algorithm considers pavement images' photometric and geometric characteristics and is validated on synthetic and natural images from different acquisition systems. The proposed method is compared to five existing methods and is found to provide very robust and precise results in a wide range of situations in a fully unsupervised manner, surpassing the current state of the art. The study highlights the potential of data-driven approaches in improving crack detection in pavement images. The survey of Padsumbiya et al. (2022) proposes a method for automatic crack detection on concrete surfaces using a simple Convolutional Neural Network (CNN) and compares it with a

TABLE 4 AI based classification techniques for crack detection.

| References | Data size | Algorithm | Layer and epoch | Evaluation metric used | Classification result | Diversity in images |
|---|---|---|---|---|---|---|
| Silva and Lucena (2018) | 3,500 samples | CNN | 16 layers, 48 epoch, 3 parameters | Accuracy | Crack, No crack | Yes |
| Amhaz et al. (2016) | 269 real images | Dijkstra | NA-layers, NA-epoch,5 parameters | DICE similarity coefficient | transverse cracks, longitudinal cracks, and alligator cracks | No |
| Padsumbiya et al. (2022) | 40,000 | CNN | 4 layer, 5 epoch, 954,241 parameters | Precision, recall, and F1 score | Crack, No-crack | Yes |
| Ali et al. (2021) | 25,000 samples | CNN | 3 layers,20 epoch,2.7 million paramters | Accuracy, precision, recall, and F1-score | Crack, No-crack | Yes |
| Gopalakrishnan et al. (2017) | 760 samples | VGG-16 DCNN | 24 layers, 50 epochs, 144 million parameters | Accuracy, precision, recall, F1-score and Cohen's Kappa score | Crack, No-crack | No |
| Kim et al. (2021) | 40,000 images | CNN | 7 layers,20 epochs,5 parameters | Accuracy score, precision, recall, and F-Measure | Crack, No-crack | No |
| Maniat et al. (2021) | 48,000 images | CNN | 22 layers,100 epoch,20 million parameters | Precision, Recall, F1-Score | Transverse cracks, longitudinal cracks, and alligator cracks | No |
| Chen C. et al. (2021) | NA | SVM | NA | Accuracy | Transverse cracks, longitudinal cracks, and alligator cracks | Yes |
| Tabassum (2022) | 370 Images | CNN | 7 layers, 30 epoch, 813,475 parameters | Mean absolute error (MAE), mean square error (MSE), root mean squared error (RMSE), and r 2 value | Multiclass damage | No |
| Hammouch et al. (2022) | 3,287 images | CNN VGG19 | 19 layers,40 epoch, 2,00,996,51 parameters | Precision, Recall, F1-Score | Longitudinal crack, transverse crack, alligator crack | Yes |

Feed-Forward Fully Connected Neural Network. Using low-pixel density images, the proposed CNN is trained to detect and segregate cracked and non-cracked concrete surfaces. The model uses Max Pooling and optimization techniques and achieves a final accuracy of 97.8%. The study confirms the impact of Artificial Intelligence in Civil Engineering, where a simple neural network can carry out automatic crack detection, eliminating the need for high-cost digital image-capturing devices.

Ali et al. (2021) present a customized convolutional neural network (CNN) for crack detection in concrete structures and compare its performance with four existing deep learning methods based on training data size, data heterogeneity, network complexity, and number of epochs. The proposed CNN model and VGG-16 outperformed the other methods in terms of classification, localization, and computational time on a small amount of data. The evaluation considered various measures such as accuracy, precision, recall, and F1-score. The results indicate that increasing the training data size and reducing diversity reduced generalization performance and led to overfitting. The proposed CNN model and VGG-16 demonstrate superior crack detection and localization for concrete structures. The paper by Gopalakrishnan et al. (2017) suggests using a Deep Convolutional Neural Network (DCNN) trained on ImageNet to automatically detect cracks in Hot-Mix Asphalt (HMA) and Portland Cement Concrete (PCC) surfaced

pavement images, which also contain non-crack anomalies and defects. The study aims to train a classifier on combined HMA-surfaced and PCC-surfaced photos that have different surface characteristics. A single-layer neural network classifier (with "Adam" optimizer) trained on ImageNet pre-trained VGG-16 DCNN features yielded the best performance.

Kim et al. (2021) propose a shallow convolutional neural network (CNN) for surface concrete crack detection that can be employed using low-power computational devices. The proposed LeNet-5 architecture is optimized and trained using 40,000 images in the METU dataset, achieving a maximum accuracy of 99.8% with minimum computation. The model's hyperparameters are optimized for crack detection, and its performance is compared with various pre-trained deep-learning models. The study concludes that the proposed method can be incorporated with autonomous devices, such as crewless aerial vehicles, for real-time inspection of surface cracks. In this paper, Prasanna et al. (2014) present a new automated crack detection algorithm called the STRUM classifier, which is used to detect cracks on concrete bridges. The algorithm uses machine learning classification and multiple visual features that are spatially tuned to potential crack regions, and it employs robust curve fitting to spatially localize likely crack regions even in the presence of noise. The algorithm is demonstrated using a state-of-the-art robotic bridge scanning

system and actual bridge data from hundreds of crack regions over two bridges. The results show that the STRUM classifier outperforms a more typical image-based approach, with a peak performance of 95% accuracy. The crack density map for the bridge mosaic, provides a computational description and a global view of the spatial patterns of bridge deck cracking.

The study of Maniat et al. (2021) investigates using Google Street View (GSV) pavement images to evaluate pavement quality. A convolutional neural network (CNN) is developed to perform image classification on GSV pavement images by dividing them into smaller image patches and classifying them into different categories of pavement cracks. The study compares the results of the CNN with those of a commercial visual inspection company and shows that GSV images can be effectively used for pavement evaluation. The designed CNN is able to classify pavement images into different crack categories. The paper by Amhaz et al. (2016) presents a new algorithm for automatic crack detection from 2D pavement images, which relies on localizing minimal paths within each image. The proposed approach selects a set of minimal tracks and introduces two post-processing steps to improve the detection quality. Compared to five existing methods, the algorithm is validated on synthetic and natural images from five different acquisition systems. The results show that the proposed algorithm provides very robust and precise results in a wide range of situations in a fully unsupervised manner, which is beyond the current state of the art.

The article by Silva and Lucena (2018) describes the development of a machine learning-based model to detect cracks on concrete surfaces using a deep learning convolutional neural network (CNN) image classification algorithm. The model is intended to increase the level of automation on concrete infrastructure inspection when combined with crewless aerial vehicles (UAV). A relatively heterogeneous dataset with 3500 images of concrete surfaces, balanced between images with and without cracks, were used for this work. The model's accuracy were recorded for different experiments, and the best investigation yielded a model with an accuracy of 92.27%, showing the potential of using deep learning for concrete crack detection.

## 4.4.2. Boundary box regression in crack detection

Object detection is another crucial aspect of AI-based crack detection systems. It involves identifying and localizing objects of interest in an image or video. In the case of crack detection, the thing of interest is the crack itself. Object detection is usually performed using deep learning models such as YOLO (You Only Look Once) and Faster R-CNN (Faster Region-based Convolutional Neural Networks). These models use convolutional neural networks (CNNs) to extract features from the input image and then use these features to identify and localize the crack. Table 5 represents the review articles published on crack detection with an object detection approach.

Fang et al. (2020) present a hybrid approach to detecting cracks in raw images using deep learning models and Bayesian probabilistic analysis. The technique involves retraining an object detector to identify crack patches with a suitable signal-to-noise ratio, generating ground truths of crack patches using

a semi-automatic method, and using a Bayesian integration algorithm to suppress false detections. The algorithm uses a deep CNN to recognize the orientation of the crack segment in each detected patch, computes a Bayesian probability based on the accumulated evidence from seen adjacent patches within a neighborhood, and suppresses the patch lacking local supports as false detection. The proposed approach is evaluated on a comprehensive dataset of crack images and outperforms the state-of-the-art baseline approach on deep CNN classifier. Ablation experiments are conducted to demonstrate the effectiveness of the proposed techniques.

Hu et al. (2021) addressed the challenges in detecting cracks on asphalt pavement using traditional methods, which can be inefficient and miss detection. The study proposed using object detection based on the deep learning model YOLOv5 for pavement crack detection. The researchers collected 3,001 images of crack pavement with varying severity levels and used the YOLOv5 series models for training and testing. The results showed that the YOLOv5l model had the highest detection accuracy of 88.1%, and the YOLOv5s model had the shortest detection time of only 11.1 ms per image. The proposed approach could effectively detect cracks on asphalt pavement, which can improve road safety by identifying and repairing these cracks in time. Machine learning-based models can be difficult to generalize for various cracks, requiring the artificial design of pavement crack characteristics.

Li et al. (2019) propose a new method for crack inspection in aircraft structures, as existing methods are time-consuming and inaccurate. Their approach, called YOLOv3-Lite, utilizes depth-wise separable convolution, feature pyramids, and YOLOv3 to improve crack detection. By using depth-wise separable convolution, crack features are extracted, and parameters are reduced. The feature pyramid joins semantically strong features at high resolution for rich semantics, while YOLOv3 is used for bounding box regression. The results show that YOLOv3-Lite is 50% faster than YOLOv3 without any loss of detection accuracy, making it a state-of-the-art performance crack detection method for aircraft structures such as fuselage or engine blades.

Wan et al. (2021) introduce a solution to the issues of complex models and computational time-consuming problems in current deep learning-based methods for road damage detection. The proposed YOLO-LRDD model is a lightweight version of YOLOv5s that incorporates Shufe-ECANet with an ECA attention module as a new backbone network and utilizes BiFPN for reliable detection. The Focal-EIOU approach is applied in the training phase to obtain higher-quality anchor boxes, and the RDD2020 dataset is augmented with Chinese road scene samples for testing. Experimental results demonstrate that YOLO-LRDD outperforms several state-of-the-art object detection techniques in terms of accuracy and efficiency. Compared to YOLOv5s, YOLO-LRDD reduces model size by 28.8%, improves single image recognition speed by 22.3%, and is easier to implement in mobile devices due to its smaller and lighter model.

The paper by Hong et al. (2022) introduces the AugMoCrack network, a new method for detecting cracks at the level of bounding boxes. This approach focuses on identifying the location of crack objects through a morphological perspective, which sets it apart from other methods that require pixel-level

TABLE 5  Review of road crack object detection methods.

| Author | Types of cracks | Algorithm | Additional features | Advantage | Limitations |
|---|---|---|---|---|---|
| Hu et al. (2021) | Fatigue crack, longitudinal crack, and transverse crack: object detection | YOLOv5 series | Backbone modules include Focus and CSP structures (neck modules include FPNs and PANs). The prediction module uses GIOU_Loss | Detection accuracy of >85%. | Model can be enhanced by combining semantic segmentation technology. Anchor-free object detection is not used |
| Li et al. (2019) | Aircraft cracks | YOLOv3-Lite | Combines depth wise separable convolution, feature pyramids, and YOLOv3 | It can run on a mobile device due to its light-weight characteristics, YOLOv3-Lite is 50% faster than YOLOv3 | Detection accuracy is less compared to YOLOv3 |
| Wan et al. (2021) | Longitudinal cracks, lateral cracks, alligators, and potholes | YOLOv5 | Shuffle-ECANet by adding an ECA attention, BiFPN Focal-EIOU in order to get higher-quality anchor box. | Sample imbalance treatment enhances the detection ability of the model for Small objects to a certain extent | Optimizing the sample imbalance processing method, can improve the sensitivity to small objects |
| Hong et al. (2022) | Cracks:, Patch, segmentaion | YOLOv5 Frequency Augmentation Morphological Filtering Morphological attention loss | A weakly supervised crack detection method with data augmentation. LossCA, which considers the connectivity of the detected crack, and LossPA, which helps in the strict box detection of the proposed method | Tightly detecting the bounding box area while reducing false duplicate detections. Excellent crack-detection performance even with little crack-training data | Not efficient for handling the road images with different textures |
| Ren et al. (2022) | Longitudinal crack, transverse crack, alligator crack, and pothole: | YOLOV5-CoordAtt | The main novelty of the proposed method is that various attention modules were introduced into the original networks to improve the ability of extracting the features of pavement cracks. | Precision of this method is 4.36% higher than the original model | The confidence scores of pavement cracks are not so high. Increase in computational load |
| Fang et al. (2020) | Images are night view, wet pavement, distance view, marble base, | Faster-RCNN | A DCNN is introduced to recognize the crack orientation. A two-phase Bayesian integration algorithm is proposed to verify the connected true detections and suppress false detections | To retain high sensitivity to small cracks, the dense outputs of overlapping patch detections from Faster R-CNN are employed. | Bayesian integration increases the computational cost |
| Zhao et al. (2022) | Longitudinal, transverse, and bifurcation cracks | Faster-RCNN | Combined with sparse representation and compressed sensing, the dataset was preprocessed, and various preprocessing algorithms were compared | Crack detection effects in complex situations such as road marking interference, shallow cracks multiple cracks, and ambiguity were significantly improved. | Type and severity of damage (cracking) is not considered |

segmentation. The Poisson blending and high-frequency discrete cosine transform-based features to augment their training data. The network also employs morphological attention loss functions, considering neighbor connectivity and box area border to enhance detection accuracy. The researchers trained their network using two datasets and achieved a 4.5 and 2.5% increase in mean average precision (mAP) in the concrete crack and Crack500 datasets, respectively, compared to the baseline architecture. In a weakly supervised learning environment where training data

is limited, the AugMoCrack network outperforms state-of-the-art crack detection methods.

Ren et al. (2022) present a pavement crack detection method that employs YOLOv5 as the base model for real-time inspection. Attention modules were added to improve the accuracy of deep learning-based methods since these methods typically need help to achieve high accuracy when dealing with small-sized pavement cracks. The proposed method used self-built datasets from Linyi city, and the results demonstrated that YOLOv5-CoordAtt with

attention modules had a precision of 95.27%, which was higher than other conventional and deep learning methods. The proposed method proved accurate in detecting pavement cracks in various situations. In another study, Zhao et al. (2022) offer a deep learning-based approach to seeing road cracks that utilize image sparse representation and compressed sensing to preprocess the datasets. This method achieves high accuracy and efficiency in crack identification and is robust when dealing with various road crack images. The method was evaluated using different algorithms, and the results showed that it outperforms the original method by increasing the mean average precision (mAP) by up to 5%.

### 4.4.3. Semantic segmentation in crack detection

Segmentation is dividing an image into multiple segments or regions based on predefined criteria such as color, texture, or shape. In the case of crack detection, segmentation is used to identify and highlight the crack region in an image. Several segmentation techniques are available, such as thresholding, edge detection, and deep learning-based segmentation methods, such as U-Net and Fully Convolutional Networks (FCNs). Table 6 illustrates the segmentation methods used in the literature with a combination of convolutional and deconvolution layers to generate a pixel-wise segmentation map, which can be used to quantify the severity of the crack.

Fan et al. (2021) propose a pothole detection approach using single-modal semantic segmentation, employing a convolutional neural network to extract visual features, a channel attention module to enhance consistency, and an atrous spatial pyramid pooling module to integrate spatial context information. They also use a multi-scale feature fusion module to reduce the semantic gap between different feature channel layers. Qu et al. (2021) present a network model for automatically detecting cracks with uneven strength from a complex background, which uses hierarchical feature fusion and connected attention architecture to recover lost details and incomplete extracted cracks. The model uses an improved DCA-SE-ResNet-50 as the backbone network and combines depthwise separable convolution and dilated convolution for crack feature fusion. Moreover, they design an attention layer to integrate feature map2 with feature map4 and incorporate feature maps of the low and high convolutional layers in the side network to assist in obtaining the final prediction map. The proposed method achieves state-of-the-art performances with the best F-score over 0.86 and 12 FPS on CFD, Crack500, and DCD datasets, and experimental results demonstrate the effectiveness of both methods.

Ai et al. (2018) proposed a new automatic pavement crack detection method that leverages multi-scale neighborhood information and pixel intensity using a probabilistic generative (PGM) based practice. The PGM calculates the probability of a crack for each pixel, and a fusion algorithm merges the probability maps from PGM and SVM approaches into a fused map that detects cracks more accurately than any of the original probability maps. Additionally, a weighted dilation operation is proposed to improve crack continuity. The proposed method outperforms state-of-the-art pavement crack detection algorithms regarding precision, recall, f1-score, and receiver operating characteristics. In contrast,

Yang et al. (2019) address the challenge of automatic pavement crack detection by proposing a new network architecture called Feature Pyramid and Hierarchical Boosting Network (FPHBN) that integrates context information with low-level features and uses nested sample reweighting during training to balance the contributions of easy and hard samples to the loss. The paper also introduces an average intersection over union (AIU) measurement for crack detection evaluation.

Lin et al. (2023) proposed a deep learning based end-to-end segmentation approach focusing on the contextual semantic information and edge information on crack images. Three factors, including the extraction of multi-scale feature information, the spatiotemporal attention mechanism, and pyramidal pooling (PSA-Net), were used to construct this technique. Without expanding the number of network parameters, PSA-Net is a compact pavement crack detection model. By layered sample weighting to equalize the loss caused by simple and complex samples, it increases the accuracy of autonomous road crack detection.

## 4.5. Crack quantification

Crack quantification is the critical step after the crack skeleton is obtained from road image processing. Crack parameters such as length, width, and depth can be calculated to estimate the severity of the crack. Estimating crack severity will guide road maintenance officers to take the necessary actions to avoid mishaps. Figure 8 shows the crack severity categorization as low, medium, or high depending on the calculated values of the crack's length, width, and depth. These severity levels help in evaluating how the crack might affect the structural integrity of the civil infrastructure.

- Low severity: The absence or presence of relatively few interconnecting cracks with a width of 6 mm and a distance between cracks of <0.328 m. This type of crack can be so tight that determining its diameter might be difficult, if not impossible.
- Medium severity: The affected road section has a comprehensive, interconnected network of cracks. Crack widths range from 6 mm to 19 mm; they can also relate to any damage with an average width of <19 mm and near a pattern of mild severity cracking. The distance between cracks in this group is no more than 150 mm.
- High severity: The road area's crack pattern comprises moderately or severely developed linked cracks. Cracks bigger than 19 mm or any crack wider than 19 mm but near to medium to high-severity random cracking are associated with high severity.

Pavement cracks are a common problem on roads and highways, and their monitoring and quantification are essential for maintaining safe and efficient transportation networks. However, the complex texture of these cracks and the potential for noise and illumination to interfere with measurement accuracy have made traditional crack quantification methods challenging. The article by Sun et al. (2022) proposes a road crack monitoring and quantification method based on vehicle video to overcome

TABLE 6 Review of crack semantic segmentation[[Inline Image]] methods.

| Author | Types of cracks | Algorithm | Additional features | Advantage | Limitations |
|---|---|---|---|---|---|
| Fan et al. (2021) | Pothole: semantic segmentation | Encoder Decoder | Unet with ResNet as Backbone Channel attention module Multi scale feature fusion model | The quantitative comparisons method achieves the state-of-the-art (SoTA) performance on both RGB images and transformed disparity images, outperforming three SoTA single-modal semantic segmentation networks | Cost is more Crack parameters not calculated |
| Qu et al. (2021) | Cracks: semantic segmentation | DCA-SE-ResNet-50 | Feature fused (FF) module Attention layer Deeply-Supervised Nets (DSN) | Insensitive to noise crack marking, and can effectively distinguish stains and obstruction | Not very effective in detecting thin cracks |
| Ai et al. (2018) | Cracks-pixel level segmentation | A probabilistic generative model (PGM). | Support vector machine (SVM) multi-scale neighborhoods Information, Fusion Algorithm weighted Dilation to optimize the detected cracks | Maintains the continuity of crack | Produces fewer noise |
| Li et al. (2022b) | Transversal and longitudinal cracks, alligator Cracks: Segmentation | U-shaped network with the ResNet encoder | EDA: the high-efficiency dual attention module Multiscale attention hesitant fuzzy set (HFS) for crack severity classification | More suitable for the detection of crack images with unbalanced aspect ratios | Necessary to improve the attribute index and optimize the classification method of crack images, especially for classifying massive cracks and cracked cracks. |
| Yuan et al. (2022) | Cracks: Segmentation | Encoder Decoder | Residual detail attention (RDA) Attention gating strategy Cross-entropy loss function | Accuracy of crack extraction is improved with suppressed background artifacts | This method fails to achieve the desired results when dealing with fine cracks Integrity of fracture segmentation (e.g., blind inpainting), can be improved to extract more accurate and complete features |



FIGURE 8
Crack severity levels.

the limitations of traditional methods. The method includes automated vehicle-mounted equipment with GPS signals to capture crack images with location information, extracting morphological features of dynamic road cracks, and a calculation algorithm based on the United Kingdom scanning grid and projection method. The proposed method also improves the crack distress evaluation method through the analysis of different crack grades. The experimental results indicate strong reliability and adaptability with high-frequency and wide-range road detection, and the proposed method has the potential to improve the monitoring and quantification of pavement cracks and ensure the safety and longevity of transportation infrastructure.

On the other hand, Matarneh et al. (2023) developed an automated tool using the Hough transform algorithm for detecting and classifying pavement cracks to optimize road maintenance and prevent possible failures. The article reviews existing attempts to use the algorithm and proposes a simple, low-cost method that achieves high accuracy for detecting and classifying vertical, diagonal, and horizontal cracks. The article suggests that this low-cost image processing method has the

potential to automate pavement crack detection and guide long-term pavement maintenance decisions, which can reduce costs for highway agencies. The originality of the article lies in its successful testing of the Hough transform algorithm for automated cracks and distresses classification.

The article by Avendaño (2020) discusses the challenges of manual inspections for assessing damages in civil engineering structures. It highlights using image-based checks using cameras or crewless aerial vehicles (UAV) combined with image processing to overcome these challenges. The article presents an approach combining different aspects of the inspection, from data acquisition through crack detection to quantifying essential parameters. A convolutional neural network (CNN) is used to identify cracks, and different quantification methods are explored to determine the width and length of the damages. The results demonstrate a low to no false negative rate for crack identification using the CNN and the highest accuracy estimation for 0.2 mm cracks during quantification.

Similarly, Deng et al. (2023) propose an integrated framework for automatically detecting, segmenting and measuring road surface cracks. The approach involves using the YOLOv5 algorithm for crack detection and a modified Res-UNet algorithm for accurate segmentation at the pixel level—a novel crack surface feature quantification algorithm to determine the width and length of the cracks. The proposed method is validated using a road crack dataset containing complex environmental noise, and it shows higher accuracy for crack segmentation under complex backgrounds compared to other methods. The developed crack surface feature algorithm has an accuracy of 95% in identifying the crack length and a root mean square error of 2.1 pixels in identifying the crack width, with the accuracy being 3% higher in length measurement than that of the traditional method.

Ha et al. (2022) propose an integrated framework for the automated detection, classification, and severity assessment of road cracks to optimize pavement management systems. The proposed system expands the number of detected crack types to five (alligator, longitudinal, transverse, pothole, and patching). It includes the assessment of crack severity, which typically needs to be improved in related studies. The studied research article uses SqueezeNet, U-Net, and Mobilenet-SSD models to achieve an accuracy of 91.2% for both crack type and severity assessment. The proposed system uses U-Nets for linear and area cracking to improve object detection performance and automate the evaluation of crack severity. The suggested automated pavement management system better reflects each country's requirements for various crack types and severity standards.

Moreover, Carrasco et al. (2021) suggest a novel automated method for measuring the width of surface cracks in civil engineering infrastructure. The traditional visual inspection method and manual measurement with a crack-width comparator gauge are time-consuming and error-prone. Although algorithms for automatic crack detection have been developed, most still need to address the problem of crack width evaluation. The proposed method consists of three stages: anisotropic smoothing, segmentation, and stabilized central points by k-means adjustment. It allows the characterization of both crack width and curvature-related orientation and has been validated by assessing the surface

cracking of fiber-reinforced earthen construction materials. The preliminary results show that the proposal is robust, efficient, and highly accurate at estimating crack width in digital images, effectively detecting natural cracks as small as 0.15 mm width regardless of the lighting condition.

### 4.5.1. Method for calculating the road segment severity index

Boucetta et al. (2021) suggested three indices combined in the severity index computation unit: the alligator cracks index (ACI), the transverse cracks index (TCI), and the longitudinal cracks index (LCI). These indices are calculated based on the cracks observed in the gathered photos. The road network creation stage entails creating a graph from the collected road network data. The weighted network creation step generates a weighted graph based on the collected severity indices and road data. The severity indices are calculated using the edge weights of the road network

The road segment severity index (SI), is a composite of three indices: the ACI (Eq. 2), TCI (Eq. 3), and LCI (Eq. 4) of the same segment. Each of the three indices takes into account the number of cracks discovered in a road segment and their severity level, which is enabled by the suggested detection and classification methods discussed in the preceding divisions.

$$SI\,(i) = \frac{ACI(i) + TCI(i) + LCI(i)}{3} \quad (1)$$

$$ACI\,(i) = \frac{Cf_1 \times LAC + Cf_2 \times MAC + Cf_3 \times HAC}{\sum_{j=1}^{3} Cf_j} \quad (2)$$

LAC, MAC, and HAC denote the number of low, medium, and high severity alligator cracks in road segments, respectively.

$$TCI\,(i) = \frac{Cf_1 \times LTC + Cf_2 \times MTC + Cf_3 \times HTC}{\sum_{j=1}^{3} Cf_j} \quad (3)$$

LTC, MTC, and HTC denote the number of low, medium, and high severity transverse cracks in road segments, respectively.

$$LCI\,(i) = \frac{Cf_1 \times LLC + Cf_2 \times MLC + Cf_3 \times HLC}{\sum_{j=1}^{3} Cf_j} \quad (4)$$

LLC, MLC, and HLC are the counts of the low, medium, and high severity longitudinal cracks in road segments, respectively, and are coefficients associated with each cost parameter in each equation independently. Experts in the area are in charge of verifying the values of these coefficients to improve the optimization of the weight calculation algorithm.

## 5. Learning algorithms for crack detection

Learning algorithms are an essential part of AI-based crack detection systems. The primary goal of learning algorithms is to provide the system with the ability to learn from a training dataset to recognize and classify the crack images. There are three

types of learning algorithms: supervised, unsupervised, and semi-supervised. In supervised learning, the system is trained using labeled data, whereas, in unsupervised learning, the system is trained without labeled data. In semi-supervised learning, the system is trained with labeled and unlabeled data. The choice of learning algorithm depends on the availability of labeled data and the type of problem to be solved.

## 5.1. Supervised learning

Qu et al. (2022) introduce a new method for detecting cracks in concrete structures using deep learning and multiscale fusion techniques. The use of modified U-Net architecture achieves high accuracy rates of up to 98.67% on a diverse dataset of concrete images with various types of cracks. The study emphasizes the potential of combining deep learning and multiscale fusion techniques for accurate and efficient crack detection in concrete structures. Zou et al. (2018) proposed an end-to-end deep learning-based approach for crack segmentation in pavement images.VGG-16 and U-Net networks extract and refine hierarchical features for precise crack segmentation. The proposed method outperforms traditional methods and achieves state-of-the-art results on benchmark datasets. The study highlights the potential of deep learning-based approaches for accurate and efficient crack segmentation in pavement images. Zhang et al. (2016) present an automatic crack detection method for concrete images using a convolutional neural network (CNN). CNN architecture with two convolutional and two fully connected layers achieves a high accuracy rate of 94.4% on a dataset of concrete images with cracks. The study demonstrates the potential of deep learning-based approaches for automatic crack detection in concrete images. Chen Y. et al. (2021) and Sun et al. (2021) comprehensively review image processing-based techniques for crack detection in concrete pavements. The studied article examines crack detection techniques, including thresholding, edge detection, texture analysis, and machine learning-based methods. The study emphasizes the importance of developing efficient and accurate techniques for crack detection in concrete pavements to ensure their safety and durability. It provides insights for future research in the field.

## 5.2. Unsupervised learning

Unsupervised learning is a machine learning technique involving training models on data without explicit supervision or labeled data. Several studies have explored unsupervised learning techniques for various applications. Duan et al. (2020) proposed an unsupervised deep-learning framework for anomaly detection in industrial processes. The use of generative adversarial network (GAN) to learn the normal data distribution and identify anomalous samples. The proposed method outperformed traditional anomaly detection techniques and achieved high accuracy rates. The study demonstrates the potential of unsupervised learning techniques for anomaly detection in industrial processes.

Similarly, Li et al. (2021) proposed an unsupervised deep learning-based approach for anomaly detection in power systems. The variational autoencoder (VAE) learn the expected behavior of the power system and identify anomalies. The proposed method outperformed traditional anomaly detection techniques and achieved high accuracy rates. The study highlights the potential of unsupervised learning techniques for anomaly detection in power systems. Wu et al. (2021) proposed an unsupervised deep-learning framework for image classification. An autoencoder-based clustering approach group images with similar features and achieved high accuracy rates on benchmark datasets. The proposed method outperformed traditional clustering techniques and demonstrated the potential of unsupervised learning techniques for image classification.

Mubashshira et al. (2020) proposed an unsupervised learning-based approach for feature extraction in human action recognition. Use of a stacked denoising autoencoder (SDAE) to learn discriminative features from raw sensor data and achieved high accuracy rates on benchmark datasets. The proposed method outperformed traditional feature extraction techniques and demonstrated the potential of unsupervised learning techniques for human action recognition.

Methods for supervised crack detection today mainly rely on labeled data. Li et al. (2022b) presented an unsupervised reconstruction-based concrete crack detection approach based on nnU-Net. This approach works better when the normal sample distribution's variance is modest. In this instance, while the abnormal image cannot be entirely learned from the semantic information of the normal sample, the normal image can be improved rebuilt. In order to rebuild the original image, the input image first goes to the trained model, and the output image is then utilized as the input image of the final model. Following study, he suggested a reconstruction approach, and his selection of the loss function encourages the network to better rebuild the system.

Li et al. (2021) study's main goal is to address the weak generalization and low intelligence issues with crack detection. Considering efficiency and model simplicity, the studied article proposes a fused deep neural network model architecture with the K_means clustering algorithm. The K_means clustering technique is utilized to generate the pseudolabels that the AlexNet model uses to train its model, and this architecture is based on the original AlexNet model. This fused architecture's key benefit is that it avoids the disadvantages of supervised learning techniques by not requiring manually labeled ground truth images for model training. This model performs satisfactorily after being trained on cracked images acquired from cellphones and automated cars in a variety of environmental scenarios with different image quality.

## 5.3. Semi-Supervised learning

Semi-supervised learning is a machine learning approach that uses labeled and unlabeled data to improve model accuracy. In pavement evaluation and maintenance planning, detecting cracks is a critical task. The paper by Tang et al. (2022) discusses using Artificial Neural Networks (ANNs) for crack detection in pavement evaluation and maintenance planning. Most existing models use

the Fully Supervised Learning (FSL) approach, which relies on high-quality annotation for reasonable accuracy. However, this approach is costly and time-consuming, especially for complex networks. This paper proposes a Weakly Supervised Learning U-Net (WSL U-Net) for pavement crack segmentation. This approach uses weakly labeled images to train the network, significantly reducing the labor cost and human involvement in image annotation. The experimental results show that WSL U-Net outperforms some Semi-Supervised Learning (Semi-SL) and WSL methods and achieves comparable performance with its FSL version. The dataset cross-validation also demonstrates that WSL U-Net is more robust with fewer overfitting concerns and better generalization capability.

Detecting cracks at the pixel level is vital for building and road inspections, but it requires time-consuming pixel-level annotations. Previous work proposed a weakly-supervised approach but struggled with lighter-colored cracks and non-crack targets. Inoue and Nagayoshi (2023) proposed a data-driven annotation refinement approach, which is effective regardless of a dataset's pixel brightness profile. The method speeds up the annotation process by factors of 10 to 30 while maintaining detection accuracy on three crack segmentation datasets and one blood vessel segmentation dataset.

Zhu and Song (2020) present a weakly supervised approach for detecting and segmenting cracks in asphalt concrete bridge decks, which are challenging to detect using conventional methods due to their dark color and complex nature. The proposed method uses an autoencoder to differentiate data and highlight unlabeled data features for weakly supervised learning. K-means clustering is used to classify features, and semantic segmentation is performed under weak supervision to identify cracks. The proposed method is evaluated on a dataset of six types of defects on asphalt concrete bridge decks and outperforms existing methods reported in the references.

## 6. Challenges and future scope

In this article, we reviewed various road imaging methods focusing on the data that transforms through different image processing phases and outputs the segmented crack with its severity level. With a data-driven approach, we studied pre- and post-processing techniques used to improve image quality. We also looked at learning techniques used for based crack classifications. The following problem areas related to crack detection are majorly discussed in Table 7 with their future scope.

Furthermore, future dataset collection efforts will prioritize accuracy in computer vision tasks. Attention modules will enable models to focus on relevant features, while multiscale feature fusion will capture information at different levels of detail. Dataset creation will also facilitate the integration of object detection with segmentation. By emphasizing these aspects, dataset collection efforts can drive advancements and enhance the accuracy of computer vision applications in crack detection.

Moreover, addressing computational costs in object detection algorithms will be crucial in future dataset collection. This may involve exploring anchor-free techniques as alternatives to reduce computational complexity. Using unsupervised deep learning models can also help mitigate annotation overhead and computational requirements. By focusing on these objectives, dataset collection efforts can lead to developing more efficient object detection models without compromising accuracy and performance in crack detection applications.

Integrating distributed lightweight deep learning models will play a significant role in the future of real-time systems for crack detection. These models, designed for resource-constrained devices, will enhance the efficiency and responsiveness of real-time applications, particularly in road pavement. By leveraging distributed computing and lightweight architectures, real-time systems can achieve improved performance, low-latency processing, and benefit domains such as edge computing and IoT devices.

Lastly, advancements in image analysis techniques, automated detection systems, and the integration of non-destructive evaluation methods will shape the future of crack quantification. The aim is to enhance accuracy, automate the quantification process, and enable real-time monitoring for proactive maintenance and improved infrastructure safety in road pavement.

TABLE 7 Challenges and future scope in crack detection.

| Problem area | Challenges in crack detection | Future direction |
|---|---|---|
| Dataset | Imbalance image data samples | Transfer learning and semi-supervised GAN |
| | Heterogeneity in data samples | Extreme Learning Machine |
| | 3D dataset | 3D-Point-Cloud-Based Deep Neural Network |
| | Lacking labeled/annotated datasets | Deep active learning |
| Accuracy | Accuracy is compromised in varied road conditions like shadow, stains, oil spots etc. | Attention Modules Multi scale feature Fusion Combine object detection with Segmentation |
| Computational cost | Labeling cost Models are very large and heavy. It takes a lot of time and computing power to run Number of possible combinations for bounding boxes is huge, and these networks tend to be computationally demanding. | Anchor Free Object Detection Unsupervised deep learning models can be used |
| Real time system | Improve the performance in real time crack localization Need huge amounts of training data and powerful computational infrastructure | Distributed Lightweight Deep Learning Models |
| Crack quantification | Calculating the crack width, length, and depth Crack categorization based on severity | |

## 6.1. Future scope

In the context of crack detection in road pavement, the future scope of dataset collection is expected to drive advancements in research and application. Specialized datasets will support emerging techniques such as transfer learning, semi-supervised GANs, extreme learning machines, 3D-point-cloud-based deep neural networks, and deep active learning. These datasets will be diverse and domain-specific, covering many complexities. The aim is to leverage pre-trained models, explore advanced learning algorithms, and advance the field of 3D deep learning while reducing the effort required for labeling. This focus on dataset creation holds the potential to foster innovation and progress in machine learning and artificial intelligence, specifically for crack detection in road pavement.

### 6.1.1. Deep active learning

Deep active learning for crack detection combines deep learning models with active learning strategies to improve the efficiency and effectiveness of crack detection. It involves iteratively selecting the most informative samples from an unlabeled pool, annotating the samples by experts, and incorporating the labeled data into the training set to update the model. This process reduces annotation effort, improves model performance, and lowers labeling costs. Deep active learning provides flexibility, adaptability, and opportunities for exploring different active learning strategies to enhance crack detection. Overall, it enables efficient and accurate crack detection by leveraging deep and active learning strengths. Lv et al. (2020) developed an active learning architecture to minimize the labeling work required for flaw identification. An iteration pattern is used to train the detection model in the proposed system. An uncertainty sampling approach chose input photos based on their uncertainty levels for annotation of image data. Also, they developed an average margin approach to determine the sample ratios among defect categories to verify the sampling number for annotations.

### 6.1.2. Generative adversarial networks

GANs can be used for crack detection by generating realistic synthetic crack images and training a discriminator to differentiate between genuine and artificial cracks. The adversarial training process improves the accuracy of crack detection models. GANs can augment training data, enable unsupervised learning, assist in domain adaptation, and support semi-supervised learning. By leveraging the generative and discriminative capabilities of GANs, crack detection models can be enhanced for more effective infrastructure analysis and maintenance. Shim et al. (2020) proposed an automatic method for detecting cracks in concrete structures that combines transfer learning and data augmentation techniques. Their approach achieved high accuracy rates of 96.3 and 94.3%, respectively and is more efficient than traditional methods. The study highlights the significance of data augmentation and transfer learning in enhancing the accuracy and robustness of their proposed approach. The suggested technique can be valuable for engineers and practitioners working in crack detection in concrete structures.

### 6.1.3. Meta-learning

Meta-learning, known as "learning to learn," can enhance crack detection systems by enabling models to adapt to new crack types rapidly, learn from limited labeled data, and transfer knowledge across different structures or materials. It facilitates the efficient adaptation of crack detection models to changing crack patterns, optimizing hyperparameters, and guiding active learning and sample selection. With meta-learning, models can quickly generalize from previous crack detection tasks, allowing for effective crack identification and classification with minimal training data. The ability to learn empowers crack detection systems to leverage prior knowledge and experiences, accelerating their performance in novel scenarios. While the application of meta-learning to crack detection is still in its early stages, these principles hold promise for improving crack detection systems' efficiency, adaptability, and accuracy, contributing to safer and more reliable infrastructure management. Mundt et al. (2019) research contributes to the field of concrete defect recognition by introducing new meta-learning approaches and highlighting their effectiveness in finding optimized CNN architectures. These architectures demonstrate improved accuracy and parameter efficiency performance, addressing the challenges posed by the complex and varied nature of concrete defects in real-world scenarios. Further research and exploration in this area can unlock the full potential of meta-learning in crack detection.

### 6.1.4. Attention module

In recent studies, attention mechanisms have been utilized to improve the accuracy of automatic pavement crack detection using deep learning models. Xiang et al. (2020) proposed an end-to-end trainable deep convolution neural network that incorporates attention mechanisms to detect pavement cracks accurately. The network architecture includes a pyramid and spatial-channel combinational attention modules to refine crack features. At the same time, dilated convolution is used to avoid losing crack details during pooling operations. In addition, the Lovász hinge loss function is used to train the model on the CRACK500 dataset and evaluate it on three pavement crack datasets. The results show that the proposed method outperforms other methods regarding experimental precision.

Jing et al. (2022) proposed the AR-UNet network model to improve further the accuracy of crack detection, which introduces a convolutional block attention module (CBAM) in the encoder and decoder of the U-Net. The CBAM allows for effectively extracting global and local detail information, while the basic block prevents network degradation and layer growth. The method is tested on multiple datasets and achieves higher crack detection accuracy than existing methods.

Furthermore, Yu et al. (2022) proposed a U-shaped encoder-decoder network, RUC-Net, for automatic pavement crack detection. The scSE attention module and focal loss function were incorporated into the network to enhance detection accuracy. The proposed method was evaluated on three public datasets, demonstrating superior performance to other methods such as FCN, Unet, and SegNet. Additionally, studies were conducted on the CFD dataset to compare the effectiveness of different scSE

modules and their combinations in improving the performance of crack detection.

Wan et al. (2021) proposed CrackResAttentionNet, an encoder-decoder network-based architecture with position and channel attention modules to accurately detect pavement cracks with complex textures and different lighting conditions. The architecture outperformed popular models such as ENet, ExFuse, FCN, LinkNet, SegNet, and UNet regarding precision, mean IoU, recall, and F1 for both public and self-developed datasets. Ren et al. (2022) proposed an automatic pavement crack detection method using YOLOV5 as the base model and employing attention modules to improve detection accuracy for small cracks. The proposed CoordAtt module selectively attends to relevant features for better crack detection. The proposed method was evaluated on self-built datasets and outperformed conventional and deep learning methods, with a precision of 95.27%. Adding attention modules can effectively enhance the ability of crack detection under various situations.

# 7. Discussion

Detecting cracks in civil structures is a critical task to ensure the safety and longevity of our infrastructure. The development of advanced image processing technologies has made it possible to detect and analyze cracks non-invasively. In this study, we reviewed various image processing technologies and investigated the crack detection method based on image processing.

One of the main challenges in crack detection is data collection. Sensor systems such as UAVs, camera-mounted vehicles, and smartphones can be used to collect high-quality data. However, the 3D dataset is not widely available, so we are limited to using 2D data for our research. Furthermore, road image data can have a lot of noise, including uneven illumination, road lanes, stains, etc. Therefore, it requires preprocessing techniques such as grayscale conversion, histogram equalization, filtering, morphological operations, etc. Although this process can be time-consuming, obtaining high-quality data for effective crack detection is necessary. Labeling the data is another crucial aspect of crack detection. It can be done at image, block, or pixel levels, and it can be a tedious task. Moreover, combining textures such as asphalt and cement in datasets is often tricky, and there is a lack of small datasets available for Indian roads. Therefore, we plan to generate a combined textures dataset that includes Indian roads. Kanaeva and Ivanova (2021) diversity of road image data comprises images of various road surfaces, such as asphalt, cement, etc. Models can more easily generalize to situations in the real world when synthetic datasets are quickly modified to include a diverse variety of road picture data. In order to generate a robust and diversified dataset containing damaged and undamaged instances, synthetic dataset generation is a technique that is frequently used in machine learning and computer vision. Researchers can produce synthetic images that reflect the traits of the damage they seek to identify by employing a variety of techniques, including data augmentation, generative models, or simulations.

Learning algorithms play a significant role in crack detection and supervised learning algorithms such as CNN VGG16, Google net, transfer learning, and semi-supervised learning algorithms

that require partial labeling can be used. Unsupervised learning algorithms were rarely used in this field. Our research uses a semi-supervised learning approach to train our models effectively. Crack data augmentation is necessary to deal with the limited availability of small datasets. Generative adversarial networks (GANs) can be used to generate synthetic data for training, improving the models' accuracy.

Different crack classification approaches can be used at different levels, such as object detection with YOLO, Faster RCNN, segmentation thresholding, edge detection, and Unet FCN DEEPVLAB. However, real-time crack detection is still challenging, and accuracy is a significant problem due to the computational cost involved. To overcome this challenge, we plan to use object detection followed by segmentation, which has proven more effective.

Crack quantification is another essential aspect of crack detection. Crack parameters such as length, width, and depth need to be evaluated to determine the severity of the crack. However, depth calculation is complex, and there is a need for further research in this area. Finally, performance evaluation metrics such as F1 score, mean average precision (MAP), recall, precision, IOU, AUC, etc., are used to assess the accuracy of the models. We plan to use MAP and AUC as our evaluation metrics.

Hence, image processing-based crack detection is a promising technique for detecting and analyzing cracks in civil structures. Our research will address the challenges related to data collection, labeling, learning algorithms, crack classification, crack quantification, and performance evaluation metrics. Our findings will contribute to developing more accurate and efficient crack detection methods in the future.

# 8. Conclusions

This article comprehensively reviews AI-based image-processing technologies for crack detection in civil structures. The study explores and investigates the challenges in data-driven aspects of crack detection methods. The researchers carefully selected relevant research articles on crack detection systems and analyzed them in this review. The review first examines the element of data collection and dataset analysis. Researchers were observed to use camera-type images for research predominantly, and authentic datasets were predominantly utilized. This ensured efficiency and ease of implementation. The accuracy and error levels of the analyses were thoroughly assessed. The review also highlights methods relevant for future research in image processing-based crack detection systems. Machine learning (ML) and deep learning (DL) techniques have become mainstream technologies in developing more advanced pavement crack detection algorithms. DL-based approaches such as image patch classification, crack semantic segmentation, and boundary box regression models were compared using uniform evaluation metrics. Crack segmentation is identified as a significant research area in engineering, particularly in the field of image recognition technology. It plays a crucial role in extending the service life of civil structures and reducing safety hazards. While convolutional neural networks (CNNs) have achieved remarkable results in the segmentation and detection of road cracks, further improvements

are still necessary. The article extracts challenges in crack detection and suggests future directions for research. One area of focus is exploring methods to compress the network scale without compromising segmentation accuracy. Additionally, the physical quantification of structural cracks is an important aspect for further investigation.

## Author contributions

DV: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Supervision, Validation, Visualization, Writing—review and editing. PC: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing—original draft, Writing—review and editing. SP: Data curation, Formal analysis, Investigation, Resources, Supervision, Validation, Visualization, Writing—review and editing. SM: Conceptualization, Data curation, Formal analysis, Methodology, Supervision, Validation, Writing—review and editing. KK: Formal analysis, Funding acquisition, Investigation, Resources, Supervision, Validation, Visualization, Writing—review and editing.

## Funding

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

Ai, D., Jiang, G., Kei, L. S., and Li, C. (2018). Automatic pixel-level pavement crack detection using information of multi-scale neighborhoods. *IEEE Access* 6, 24452–24463. doi: 10.1109/ACCESS.2018.2829347

Ali, L., Alnajjar, F., Jassmi, H. A., Gocho, M., Khan, W., Serhani, M. A., et al. (2021). Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures. *Sensors* 21, 1688. doi: 10.3390/s21051688

Amhaz, R., Chambon, S., Idier, J., and Baltazart, V. (2016). Automatic crack detection on two-dimensional pavement images: An algorithm based on minimal path selection. *IEEE Trans. Int. Trans. Syst.* 17, 2718–2729. doi: 10.1109/TITS.2015.2477675

Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., and Sekimoto, Y. (2021). RDD2020: an annotated image dataset for automatic road damage detection using deep learning. *Data Brief* 36, 107133. doi: 10.1016/j.dib.2021.107133

Avendaño, J. (2020). *Identification and Quantification of Concrete Cracks Using Image Analysis and Machine Learning.* KTH VETENSKAP OCH KONST.

Bayer, M., Kaufhold, M. A., and Reuter, C. (2023). A survey on data augmentation for text classification. *ACM Comput. Surv.* 55, 1–39. doi: 10.1145/3544558

Boucetta, Z., El Fazziki, A., and El Adnani, M. (2021). A deep-learning-based road deterioration notification and road condition monitoring framework. *Int. J. Int. Eng. Syst.* 14, 503–515. doi: 10.22266/ijies2021.0630.42

Byrne, J., O'Keeffe, E., Lennon, D., and Laefer, D. F. (2017). 3D reconstructions using unstabilized video footage from an unmanned aerial vehicle. *J. Imag.* 3, 15. doi: 10.3390/jimaging3020015

Cao, M. T., Tran, Q. V., Nguyen, N. M., and Chang, K. T. (2020). Survey on performance of deep learning models for detecting road damages using multiple dashcam image resources. *Adv. Eng. Inf.* 46, 101182. doi: 10.1016/j.aei.2020.101182

Cao, W., Liu, Q., and He, Z. (2020). Review of pavement defect detection methods. *IEEE Access* 8, 14531–14544. doi: 10.1109/ACCESS.2020.2966881

Carrasco, M., Araya-Letelier, G., Velázquez, R., and Visconti, P. (2021). Image-based automated width measurement of surface cracking. *Sensors* 21, 7534. doi: 10.3390/s21227534

Chen, C., Seo, H., Jun, C., and Zhao, Y. (2022). A potential crack region method to detect crack using image processing of multiple thresholding. *Signal Image Vid. Proc.* 16, 1673–1681. doi: 10.1007/s11760-021-02123-w

Chen, C., Seo, H., Jun, C. H., and Zhao, Y. (2021). Pavement crack detection and classification based on fusion feature of LBP and PCA with SVM. *Int. J. Pave. Eng.* 23, 3274–3283. doi: 10.1080/10298436.2021.1888092

Chen, Y., Ding, Y., Zhao, F., Zhang, E., Wu, Z., Shao, L., et al. (2021). Surface defect detection methods for industrial products: a review. *Appl. Sci.* 11, 7657. doi: 10.3390/app11167657

Chen, Y., Naud, C. M., Rangwala, I., Landry, C. C., and Miller, J. R. (2014). Comparison of the sensitivity of surface downward longwave radiation to changes in water vapor at two high elevation sites. *Environ. Res. Lett.* 9, 114015. doi: 10.1088/1748-9326/9/11/114015

Deng, L., Zhang, A., Guo, J., and Liu, Y. (2023). An integrated method for road crack segmentation and surface feature quantification under complex backgrounds. *Remote Sens.* 15, 1530. doi: 10.3390/rs15061530

Duan, L., Geng, H., Pang, J., and Zeng, J. (2020). "Unsupervised pixel-level crack detection based on generative adversarial network," in *Proceedings of the 2020 5th International Conference on Multimedia Systems and Signal Processing (ICMSSP '20)* (New York, NY: Association for Computing Machinery), 6–10.

Ersoz, A. B., Pekcan, O., and Teke, T. (2017). Crack identification for rigid pavements using unmanned aerial vehicles. *IOP Conf. Ser.: Mater. Sci. Eng.* 236, 012101. doi: 10.1088/1757-899X/236/1/012101

Fan, J., Bocus, M. J., Hosking, B., Wu, R., Liu, Y., Vityazev, S., et al. (2021). "Multi-scale feature fusion: Learning better semantic segmentation for road pothole detection," in *2021 IEEE International Conference on Autonomous Systems (ICAS).* Piscataway, NJ: IEEE, 1–5.

Fang, F., Li, L., Gu, Y., Zhu, H., and Lim, J. H. (2020). A novel hybrid approach for crack detection. *Patt. Recognit.* 107, 107474. doi: 10.1016/j.patcog.2020.107474

Feng, X., Xiao, L., Li, W., Pei, L., Sun, Z., Ma, Z., et al. (2020). Pavement crack detection and segmentation method based on improved deep learning fusion model. *Mathematic. Prob. Eng.* 2020, 1–22. doi: 10.1155/2020/8515213

Fu, R., Xu, H., Wang, Z., Shen, L., Cao, M., Liu, T., et al. (2020). Enhanced intelligent identification of concrete cracks using multi-layered image preprocessing-aided convolutional neural networks. *Sensors* 20, 2021. doi: 10.3390/s20072021

Ghaderzadeh, M., Aria, M., and Asadi, F. (2021). X-ray equipped with artificial intelligence: changing the COVID-19 diagnostic paradigm during the pandemic. *BioMed Res. International*, 2021 doi: 10.1155/2021/9942873

Ghaderzadeh, M., Aria, M., Hosseini, A., Asadi, F., Bashash, D., Abolghasemi, H., et al. (2022). A fast and efficient CNN model for B-ALL diagnosis and its subtypes classification using peripheral blood smear images. *Int. J. Int. Syst.* 37, 5113–5133. doi: 10.1002/int.22753

Gheisari, M., Ebrahimzadeh, F., Rahimi, M., Moazzamigodarzi, M., Liu, Y., Dutta Pramanik, P. K., et al. (2023). Deep learning: applications, architectures, models, tools, and frameworks: a comprehensive survey. *CAAI Trans. Int. Technol.* 8, 1–26. doi: 10.1049/cit2.12180

Golding, V. P., Gharineiat, Z., Munawar, H. S., and Ullah, F. (2022). Crack detection in concrete structures using deep learning. *Sustainability* 14, 8117. doi: 10.3390/su14138117

Gopalakrishnan, K., Khaitan, S. K., Choudhary, A., and Agrawal, A. (2017). Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construct. Build. Mat.* 157, 322–330. doi: 10.1016/j.conbuildmat.2017.09.110

Ha, J., Kim, D., and Kim, M. (2022). Assessing severity of road cracks using deep learning-based segmentation and detection. *J. Supercomput.* 78, 17721–17735. doi: 10.1007/s11227-022-04560-x

Hamishebahar, Y., Guan, H., So, S., and Jo, J. (2022). A comprehensive review of deep learning-based crack detection approaches. *Appl. Sci.* 12, 1374. doi: 10.3390/app12031374

Hammouch, W., Chouiekh, C., Khaissidi, G., and Mrabti, M. (2022). Crack detection and classification in moroccan pavement using convolutional neural network. *Infrastructures* 7, 152. doi: 10.3390/infrastructures7110152

Hassan, S., O'Sullivan, D., Mckeever, S., Power, D., Mcgowan, R., and Feighan, K. (2022). "Detecting patches on road pavement images acquired with 3D laser sensors using object detection and deep learning," in *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2022) – Volume 5: VISAPP*, 413–420.

Hong, Y., Lee, S., and Yoo, S. B. (2022). AugMoCrack: augmented morphological attention network for weakly supervised crack detection. *Electr. Lett.* 58, 651–653. doi: 10.1049/ell2.12562

Hosseini, A., Eshraghi, M. A., Taami, T., Sadeghsalehi, H., Hoseinzadeh, Z., Ghaderzadeh, M., et al. (2023). A mobile application based on efficient lightweight CNN model for classification of B-ALL cancer from non-cancerous cells: a design and implementation study. *Inf. Med. Unlocked* 39, 101244. doi: 10.1016/j.imu.2023.101244

Hsieh, Y. A., and Tsai, Y. J. (2020). Machine learning for crack detection: review and model performance comparison. *J. Comput. Civil Eng.* 34, 04020038. doi: 10.1061/(ASCE)CP.1943-5487.0000918

Hu, G. X., Hu, B. L., Yang, Z., Huang, L., and Li, P. (2021). Pavement crack detection method based on deep learning models. *Wireless Commun. Mob. Comput.* 2021, 1–13. doi: 10.1155/2021/5573590

Inoue, Y., and Nagayoshi, H. (2023). Weakly-supervised crack detection. *IEEE Trans. Int. Trans. Syst.* 24, 1–13. doi: 10.1109/TITS.2023.3286934

Islam, M. M., Hossain, M. B., Akhtar, M. N., Moni, M. A., and Hasan, K. F. (2022). CNN based on transfer learning models using data augmentation and transformation for detection of concrete crack. *Algorithms* 15, 287. doi: 10.3390/a15080287

Jeong, J., Jo, H., and Ditzler, G. (2020). Convolutional neural networks for pavement roughness assessment using calibration-free vehicle dynamics. *Comput. Aided Civil Infrastr. Eng.* 35, 1209–1229. doi: 10.1111/mice.12546

Jing, P., Yu, H., Hua, Z., Xie, S., and Song, C. (2022). Road crack detection using deep neural network based on attention mechanism and residual structure. *IEEE Access* 11, 919–929. doi: 10.1109/ACCESS.2022.3233072

Jordan, S., Moore, J., Hovet, S., Box, J., Perry, J., Kirsche, K., et al. (2018). State-of-the-art technologies for UAV inspections. *IET Radar. Sonar Navig.* 12, 151–164. doi: 10.1049/iet-rsn.2017.0251

Kanaeva, I. A., and Ivanova, J. A. (2021). Road pavement crack detection using deep learning with synthetic data. *IOP Conf. Ser. Mat. Sci. Eng.* 1019, 320–325. doi: 10.1088/1757-899X/1019/1/012036

Khaloo, A., Lattanzi, D., and Cunningham, K., Dell'Andrea, R., and Riley, M. (2018). Unmanned aerial vehicle inspection of the Placer River Trail Bridge through image-based 3D modelling. *Struct. Infr. Eng.* 14, 124–136. doi: 10.1080/15732479.2017.1330891

Kim, B., Yuvaraj, N., Sri Preethaa, K. R., and Arun Pandian, R. (2021). Surface crack detection using deep learning with shallow CNN architecture for enhanced computation. *Neural. Comput. Appl.* 33, 9289–9305. doi: 10.1007/s00521-021-05690-8

Kong, Y., Yu, Z., Chen, H., Wang, Z., Chen, C., and Guo, B. (2017). "Detecting type and size of road crack with the smartphone" in *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)* (Guangzhou), 572–579.

König, J., Jenkins, M., Mannion, M., Barrie, P., and Morison, G. (2022). What's cracking? A review and analysis of deep learning methods for structural crack segmentation, detection and quantification. *arXiv preprint arXiv*: 2202.03714.

Li, B., Guo, H., Wang, Z., and Li, M. (2022a). Automatic crack classification and segmentation on concrete bridge images using convolutional neural networks and hybrid image processing. *Int. Trans. Infrastruct.* 1, liac016. doi: 10.1093/iti/liac016

Li, G., Ma, B., He, S., Ren, X., and Liu, Q. (2020). Automatic tunnel crack detection based on u-net and a convolutional neural network with alternately updated clique. *Sensors* 20, 717. doi: 10.3390/s20030717

Li, H., Wang, W., Wang, M., Li, L., and Vimlund, V. (2022b). A review of deep learning methods for pixel-level crack detection. *J. Traff. Trans. Eng.* 9, 945–968. doi: 10.1016/j.jtte.2022.11.003

Li, R., Yu, J., Li, F., Yang, R., Wang, Y., Peng, Z., et al. (2023). Automatic bridge crack detection using unmanned aerial vehicle and faster R-CNN. *Constr. Build. Mat.* 362, 129659. doi: 10.1016/j.conbuildmat.2022.129659

Li, W., Huyan, J., Gao, R., Hao, X., Hu, Y., Zhang, Y., et al. (2021). Unsupervised deep learning for road crack classification by fusing convolutional neural network and k_means clustering. *J. Trans. Eng. Pavem.* 147, 04021066. doi: 10.1061/JPEODX.0000322

Li, X., Yang, S., and Liu, H. (2022c). An unsupervised concrete crack detection method based on nnU-Net. *Lectur. Notes Electr. Eng.* 813, 609–623. doi: 10.1007/978-981-16-6963-7_55

Li, Y., Han, Z., Xu, H., Liu, L., Li, X., Zhang, K., et al. (2019). YOLOv3-lite: a lightweight crack detection network for aircraft structure based on depthwise separable convolutions. *Appl. Sci.* 9, 3781. doi: 10.3390/app9183781

Lin, X., Zhang, J., Wu, D., Meng, E., Liu, M., Li, M., et al. (2023). Intelligent identification of pavement cracks based on PSA-Net. *Front. Environ. Sci.* 11, 17. doi: 10.3389/fenvs.2023.1105467

Liu, Y., Yao, J., Lu, X., Xie, R., and Li, L. (2019). DeepCrack: a deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing* 338, 139–153. doi: 10.1016/j.neucom.2019.01.036

Lv, X., Duan, F., Jiang, J. J., Fu, X., and Gan, L. (2020). Deep active learning for surface defect detection. *Sensors* 20, 1650. doi: 10.3390/s20061650

Maniat, M., Camp, C. V., and Kashani, A. R. (2021). Deep learning-based visual crack detection using google street view images. *Neural Comput. Appl.* 33, 14565–14582. doi: 10.1007/s00521-021-06098-0

Matarneh, S., Elghaish, F., Al-Ghraibah, A., Abdellatef, E., and Edwards, D. J. (2023). An automatic image processing based on Hough transform algorithm for pavement crack detection and classification. *Smart Sust. Built Environ.* doi: 10.1108/SASBE-01-2023-0004

Mei, Q., and Gui, M. (2020). A cost effective solution for pavement crack inspection using cameras and deep neural networks. *Constr. Build. Mater.* 256, 119397. doi: 10.1016/j.conbuildmat.2020.119397

Mohan, A., and Poobal, S. (2018). Crack detection using image processing: a critical review and analysis. *Alexandria Eng. J.* 57, 787–798. doi: 10.1016/j.aej.2017.01.020

Montero, R., Victores, J. G., Martinez, S., Jardón, A., and Balaguer, C. (2015). Past, present and future of robotic tunnel inspection. *Autom. Constr.* 59, 99–112. doi: 10.1016/j.autcon.2015.02.003

Mubashshira, S., Azam, M. M., and Masudul Ahsan, S. M. (2020). "An unsupervised approach for road surface crack detection," in *2020 IEEE Region 10 Symposium (TENSYMP)*. Piscataway, NJ: IEEE, 1596–1599.

Munawar, H. S., Hammad, A. W., Haddad, A., Soares, C. A. P., and Waller, S. T. (2021a). Image-based crack detection methods: a review. *Infrastructures* 6, 115. doi: 10.3390/infrastructures6080115

Munawar, H. S., Ullah, F., Qayyum, S., and Heravi, A. (2021b). Application of deep learning on uav-based aerial images for flood detection. *Smart Cities* 4, 1220–1242. doi: 10.3390/smartcities4030065

Mundt, M., Majumder, S., Murali, S., Panetsos, P., and Ramesh, V. (2019). "Meta-learning convolutional neural architectures for multi-target concrete defect classification with the concrete defect bridge image dataset," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Piscataway, NJ: IEEE, 11188–11197.

Nguyen, S. D., Tran, T. S., Tran, V. P., Lee, H. J., Piran, M. J., Le, V. P., et al. (2022). Deep learning-based crack detection: a survey. *Int. J. Pavem. Res. Technol.* 16, 943–967. doi: 10.1007/s42947-022-00172-z

Nnolim, U. A. (2020). Automated crack segmentation via saturation channel thresholding, area classification and fusion of modified level set segmentation with Canny edge detection. *Heliyon* 6, e05748. doi: 10.1016/j.heliyon.2020.e05748

Padsumbiya, M., Brahmbhatt, V., and Thakkar, S. P. (2022). Automatic crack detection using convolutional neural network. *J. Soft Comput. Civil Eng.* 6, 1–17. doi: 10.22115/SCCE.2022.325596.1397

Parrany, A. M., and Mirzaei, M. (2022). A new image processing strategy for surface crack identification in building structures under non-uniform illumination. *IET Image Proc.* 16, 407–415. doi: 10.1049/ipr2.12357

Prasanna, P., Dana, K. J., Gucunski, N., Basily, B. B., La, H. M., Lim, R. S., et al. (2014). Automated crack detection on concrete bridges. *IEEE Trans. Autom. Sci. Eng.* 13, 591–599. doi: 10.1109/TASE.2014.2354314

Qu, Z., Cao, C., Liu, L., and Zhou, D. Y. (2021). A deeply supervised convolutional neural network for pavement crack detection with multiscale feature fusion. *IEEE Trans. Neural Netw. Learn. Syst.* 33, 4890–4899. doi: 10.1109/TNNLS.2021.3062070

Qu, Z., Wang, C. Y., Wang, S. Y., and Ju, F. R. (2022). A method of hierarchical feature fusion and connected attention architecture for pavement crack detection. *IEEE Trans. Int. Trans. Syst.* 23, 16038–16047. doi: 10.1109/TITS.2022.3147669

Ren, J., Zhao, G., Ma, Y., Zhao, D., Liu, T., Yan, J., et al. (2022). Automatic pavement crack detection fusing attention mechanism. *Electronics* 11, 3622. doi: 10.3390/electronics11213622

Shi, Y., Cui, L., Qi, Z., Meng, F., and Chen, Z. (2016). Automatic road crack detection using random structured forests. *IEEE Trans. Int. Transp. Syst.* 17, 3434–3445. doi: 10.1109/TITS.2016.2552248

Shim, S., Kim, J., Cho, G.-C., and Lee, S.-W. (2020). Multi-scale and adversarial learning-based semi-supervised semantic segmentation approach for crack detection in concrete structures. *IEEE Access* 8, 170939–170950. doi: 10.1109/ACCESS.2020.3022786

Silva, W. R. L. D., and Lucena, D. S. D. (2018). Concrete cracks detection based on deep learning image classification. *Proc. MDPI* 2, 489. doi: 10.3390/ICEM18-05387

Sony, S., and Laventure, S., and Sadhu, A. (2019). A literature review of next-generation smart sensing technology in structural health monitoring. *Struct. Control Health Monit.* 26, e2321. doi: 10.1002/stc.2321

Sun, M., and Zhao, H., and Li, J. (2022). Road crack detection network under noise based on feature pyramid structure with feature enhancement (road crack detection under noise). *IET Image Proc.* 16, 809–822. doi: 10.1049/ipr2.12388

Sun, Z., Pei, L., Yuan, B., Du, Y., Li, W., Han, Y., et al. (2021). "Pavement crack detection and quantification based on scanning grid and projection method," in *International Conference on Civil Engineering,* Singapore: Springer Singapore, 273–281.

Tabassum, N. (2022). *Image-Based Crack Detection by Extracting Depth of the Crack Using Machine Learning* (Doctoral dissertation). University of South Carolina.

Tang, Y., Qian, Y., and Yang, E. (2022). Weakly supervised convolutional neural network for pavement crack segmentation. *Int. Trans. Infr.* 1, liac013. doi: 10.1093/iti/liac013

Tse, K. W., Pi, R., Sun, Y., Wen, C. Y., and Feng, Y. (2023). A novel real-time autonomous crack inspection system based on unmanned aerial vehicles. *Sensors* 23, 3418. doi: 10.3390/s23073418

Wan, H., Gao, L., Su, M., Sun, Q., and Huang, L. (2021). Attention-based convolutional neural network for pavement crack detection. *Adv. Mat. Sci. Eng.* 2021, 1–13. doi: 10.1155/2021/5520515

Wang, J., He, X., Faming, S., Lu, G., Cong, H., Jiang, Q., et al. (2021). A real-time bridge crack detection method based on an improved inception-resnet-v2 structure. *IEEE Access* 9, 93209–93223. doi: 10.1109/ACCESS.2021.3093210

Wang, L., Zhuang, L., and Zhang, Z. (2019). Automatic detection of rail surface cracks with a superpixel-based data-driven framework. *J. Comput. Civil Eng.* 33, 1–9. doi: 10.1061/(ASCE)CP.1943-5487.0000799

Wu, T., Zhang, H., Liu, J., Chen, Y., Yi, J., Zhang, Z., et al. (2021). "Memory-augment convolutional Autoencoder for unsupervised pavement crack classification," in *2021 China Automation Congress (CAC),* Piscataway, NJ: IEEE, 2952–2956. doi: 10.1109/CAC53003.2021.9727812

Xiang, X., Zhang, Y., and El Saddik, A. (2020). Pavement crack detection network based on pyramid structure and attention mechanism. *IET Image Proc.* 14, 1580–1586. doi: 10.1049/iet-ipr.2019.0973

Yang, F., Zhang, L., Yu, S., Prokhorov, D., Mei, X., Ling, H., et al. (2019). Feature pyramid and hierarchical boosting network for pavement crack detection. *IEEE Trans. Int. Transp. Syst.* 21, 1525–1535. doi: 10.1109/TITS.2019.2910595

Yang, N., Li, Y., and Ma, R. (2022). An efficient method for detecting asphalt pavement cracks and sealed cracks based on a deep data-driven model. *Appl. Sci.* 12, 10089. doi: 10.3390/app121910089

Yang, Z., Ni, C., Li, L., Luo, W., and Qin, Y. (2022). Three-stage pavement crack localization and segmentation algorithm based on digital image processing and deep learning techniques. *Sensors* 22, 8459. doi: 10.3390/s22218459

Yu, G., Dong, J., Wang, Y., and Zhou, X. (2022). RUC-Net: a residual-Unet-based convolutional neural network for pixel-level pavement crack segmentation. *Sensors* 23, 53. doi: 10.3390/s23010053

Yuan, G., Li, J., Meng, X., and Li, Y. (2022). CurSeg: a pavement crack detector based on a deep hierarchical feature learning segmentation framework. *IET Intellig.Transp. Syst.* 16, 782–799. doi: 10.1049/itr2.12173

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

Zhao, M., Shi, P., Xu, X., Xu, X., Liu, W., Yang, H., et al. (2022). Improving the accuracy of an R-CNN-based crack identification system using different preprocessing algorithms. *Sensors* 22, 7089. doi: 10.3390/s22187089

Zhong, K., Teng, S., Liu, G., Chen, G., and Cui, F. (2020). Structural damage features extracted by convolutional neural networks from mode shapes. *Appl. Sci.* 10, 4247. doi: 10.3390/app10124247

Zhong, M., Sui, L., Wang, Z., and Hu, D. (2020). Pavement crack detection from mobile laser scanning point clouds using a time grid. *Sensors* 20, 4198. doi: 10.3390/s20154198

Zhu, J., and Song, J. (2020). Weakly supervised network based intelligent identification of cracks in asphalt concrete bridge deck. *Alexandria Eng. J.* 59, 1307–1317. doi: 10.1016/j.aej.2020.02.027

Zou, Q., Zhang, Z., Li, Q., Qi, X., Wang, Q., Wang, S., et al. (2018). Deepcrack: learning hierarchical convolutional features for crack detection. *IEEE Trans. Image Proc.* 28, 1498–1512. doi: 10.1109/TIP.2018.2878966