APPLICATIONS OF
MODELLING AND SIMULATION
http://arqiipubl.com/ams eISSN 2600-8084 VOL 8, 2024, 1-25

Identification of Concrete Cracks Using Deep Learning Models: A Systematic Review

Mazleenda Mazni[1,2*], Abdul Rashid Husain[1], Mohd Ibrahim Shapiai[3], Izni Syahrizal Ibrahim[4], Riyadh Zulkifli[1] and Devi Willieam Anggara[5]

[1]Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Johor, Malaysia
[2]Faculty of Mechanical Engineering, Universiti Teknologi MARA Cawangan Johor, Kampus Pasir Gudang, Masai, Malaysia
[3]Centre for Artificial Intelligence and Robotics, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
[4]Forensic Engineering Centre, Institute for Smart Infrastructure and Innovative Construction, Faculty of Civil Engineering, Universiti Teknologi Malaysia, Johor, Malaysia
[5]Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia

*Corresponding author: mazleenda@graduate.utm.my

Submitted 21 November 2023, Revised 21 December 2023, Accepted 30 December 2023, Available online 14 January 2024.

Abstract: Deep learning (DL) has grown in popularity in civil inspection, notably for crack diagnosis, as a means of guaranteeing the long-term stability and security of concrete structures. It is critical to identify cracks to conduct inspections and assessments while preserving the existing concrete frameworks. This article reviews and analyses existing literature on identification of cracks on concrete structures using DL, to enhance the clarity and understanding of the ongoing research efforts in this domain. A systematic review found 97 linked research papers from 2018 to the beginning of 2023, using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) Statement review process as a guide. The articles are categorised into several methods in identifying cracks, which include classification, segmentation, detection, and hybrid methods. Various issues in implementing DL in all the methods are discussed and several limitations, challenges and proposed solutions are presented. Finally, possible research directions are also discussed.

Keywords: Convolution neural network; Crack classification; Crack detection; Crack segmentation; Deep learning; Review; Structural health monitoring.

1. INTRODUCTION

Concrete is the most often utilised construction material on a global scale [1]. According to [2], the most widely used substance in the entire planet right now is concrete, second only to water, with three tonnes consumed per person per year. Concrete is used in construction at a rate of twice that of all other building materials combined and nowadays concrete contains cement, water, coarse and fine particles, and other mineral and chemical admixtures [3]. When concrete ages, cracks enable moisture, oxygen, carbon dioxide, chloride, and other hostile chemicals and gases to enter the crack gaps, causing considerable destruction of the structure, corrosion of steel, and degradation to the concrete, that all contribute to structural collapse of the member. On the other hands, cracks are form of basic imperfections that are detectable by the naked eye. These cracks might develop into something much larger and more dangerous over time. The complicated pattern of textures in the background makes it difficult to see some of the cracks in the surface. Based on [4], numerous sizes and shapes of cracks exist. For example, hairline cracks have a width of 0.1 mm and are visible on a background that is free of dust; however, they are difficult to observe when the lighting changes. Cracks considered to be fine can range in width from 0 to 1 mm (millimeters) to the point where in some instances, up to 5 mm wide cracks are considered safe and can be cured. In certain circumstances, significant harm can result from cracks wider than 5 mm and demand repair, rebuild or replacement [5]. It could be a warning sign that the structure is vulnerable to further harm in the future. However in [6], it was obvious that researchers devote a lot of time and resources to examining fractures wider than 3 millimeters, but they devote significantly less time and resources to studying micro-cracks.

Several primary causes of building crack [7], [8] are because of temperature, humidity, foundation, and quality of material. The temperature differential is significant, and this huge temperature difference contributes to the phenomenon of thermal expansion and cold contraction, which worsens the incidence of cracks in buildings. Variations in humidity levels can cause changes in structural tension within a building, which can lead to cracking when the tension surpasses a particular threshold.

The foundation of the structure ensures that the force acting on it is balanced; cracks will emerge if the layout is illogical, and the force is not uniform. The quality of material like concrete and steel bar materials indirectly affect wall structure, and crack size and shape vary between buildings according to pressure and usage conditions.

Concrete crack diagnosis is essential for the functionality and maintenance of concrete structures. Visual inspection is the foundation of the conventional crack detection method. A professional building inspector evaluates the quality of the structure's health dependent upon the crack's size and position. The alternative, manual detection, is a lengthy process combined with the outcomes of manual detection are subject to interpretation because inspectors make judgments about a building or structure condition based on their individual experiences and knowledge. Recently, Deep Learning (DL) applications for crack structure detection have grown in popularity. This is prompted by the many advantages of DL pertaining to data recognition, processing, and decision-making in numerous disciplines [9]. Recent advancements in information technology have made it possible to get massive volumes of data, all of which must still be manually processed and parsed, a process that is inefficient, time-consuming, and susceptible to mistakes. Automating these processes would allow for earlier detection of anomalies and prompt intervention to minimize any negative consequences.

Each year, the acquisition of numerous assets and instruments costs millions of dollars with the purpose of identifying material deterioration or changes in the structure of infrastructure systems including bridges, roads, waterways, and buildings, which are constantly subjected to severe strains from natural disasters, catastrophic events, and daily use [10]. These occurrences can result in the structure collapsing completely or causing significant damage [11]. For instance, in Zumrut, Turkey, a nine-story reinforced concrete structure [12], has been severely damaged by earthquakes, and four columns in the basement have begun cracking and fracturing unexpectedly. Due to the impact of pounding mixed with structural strains, earthquakes are capable of causing catastrophic damage to these masonry walls [11], including collapse. A case study of a humid basement in Konya, Turkey [13], discovered that several columns had exceeded their load-bearing limits due to axial compression loads, creep, and effects of corrosion, resulting in significant damage.

DL has demonstrated its ability to handle a wide range of problems linked with object identification and surveillance, which will improve the process of crack identification. DL is appealing because it can construct predictive models without pre-defining associations [14]. Through the use of DL methods, the capabilities and robustness of conventional methods have been substantially expanded [15]. With the rapid development of computer vision and machine learning (ML) technologies in the past few years, a large number of automatic approaches have been proposed as powerful tools to address the crack detection challenges in practice [16]. Also, superior performance in tackling the crack identification problem has been demonstrated by DL models, which differ from conventional machine learning techniques in that they can learn the representations of the data without the introduction of hand-crafted rules or prior knowledge [17], [18].

Since the debut of the technology, various contexts have been presented in consideration for the goal of using DL to identify structural cracks, including buildings [19]–[27]; bridges [14], [28]–[34]; roads [35]–[43]; railway systems [44]–[46]; tunnels [47]–[49]; dam [50], [51]; monument [52], [53], and nuclear power plant component [46]. Recent studies have shown that DL outperforms alternative methods for identifying cracks in concrete buildings; nevertheless, there is a critical lack of detailed analyses in this sector. This study addresses a knowledge gap in the field of research by conducting rigorous research into the use of DL to solve the difficulty of crack identification in concrete building constructions. In addition, the DL algorithms and architectures were investigated in this systematic review that have been used in recent years to address the specific difficulties associated with crack identification on concrete structures. For this purpose, it surveys the most recent research in the field by reviewing papers from the past five years that focus on crack identification and DL. The research has mostly aimed to address a narrow set of questions regarding network algorithms and learning techniques, including the datasets used by crack identification systems, as opposed to more general questions regarding this technology. This research sheds light on the problems inherent in deploying DL systems and suggests potential avenues for future investigation. It serves as a reference for researchers and academics working on improving existing systems and creating new crack identification frameworks for use in concrete construction.

The rest of the article is organized as follows: The systematic review approach is discussed in length in Materials and Methods, which contains detailed information about the methodology section and the approach of the PRISMA Statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). The section includes the research questions, the search terms, the data sources, and the techniques used for selection and data collection. In Results, we present a statistical analysis of the data and a summary of relevant prior studies. The following section provides discussion and organizes the literature to address each topic. It also conducts a critical analysis of the findings and highlights the most important points, including gaps, limitations, challenges, and future research opportunities. Finally, the Conclusion section provides a summary that concludes the paper. It also offers recommendations for future investigations.

2. MATERIAL AND METHODS

These article refers to a method that is widely utilised for the purpose of locating, analyzing, and interpreting significant pieces of research pertaining to a particular topic, region, or issue [54]. Kitchenham's [55] approach was employed as a reference for this systematic review implementation. This review is a methodical and deliberate procedure that comprises finding, selecting, and evaluating relevant research, as well as gathering and synthesizing data from the included studies for the research. This systematic study followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) Statement as a reference.

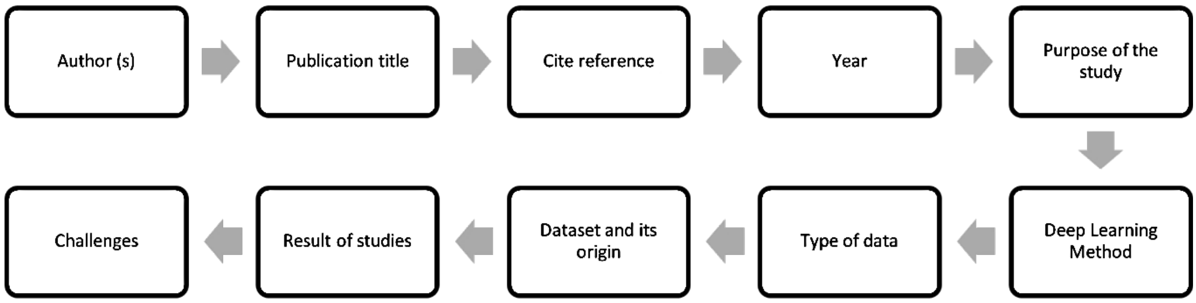

Figure 1. Criteria for inclusion items.

Comparing the PRISMA Statement to other systematic review techniques, there are three primary benefits [56]; (a) The technique establishes precise research questions that enable the execution of systematic research, (b) The technique specifies inclusion and exclusion standards for confirming eligibility, (c) With this strategy, a sizable body of scientific material is meant to be reviewed in a set amount of time. With the use of the PRISMA Statement, a systematic search for terms relating to concrete construction cracks can be undertaken with precision.

2.1 Study Protocol

This systematic review employed the search engines Scopus, Web of Science (WOS), and Google Scholar. Scopus has been recognized as the most efficient search engines for systematic literature reviews with a bigger field coverage and more search precision than other search engines [57]. With more than 25,000 journals from over 7,000 publishers located on every continent, Scopus is the world's most comprehensive collection of abstracts and citations from scientific literature. WOS, on the other hand, is an authoritative, publisher-independent citation database that has a comprehensive archive of the best quality research dating back over a century [58]. Within the Science Citation Index Expanded (SCIE), it is comprised of 9200 influential publications that cover 178 distinct scientific fields. According to Gusenbauer [59] in the Scientometrics journal, Google Scholar, the greatest comprehensive academic search engine in the world, has an estimated 389 million entries. The Scopus, WOS, and Google Scholar material are extensively evaluated and chosen by an independent review board of specialists in each discipline to maintain the highest quality standard. The results will next be filtered using two established high-indexed databases, Scopus and WOS, for the goal of quality standard conservation.

2.2 Inclusion Criteria

The scope of this investigation is narrowed to applied DL for locating cracks in concrete structures. Methods for detecting, classifying, and segmenting objects, as well as other crack identification tasks, such as hybrid detection, that made use of a DL system and were reported in an original English-language paper, were included. Subjects could only be limited in ways that made sense considering the three research questions. Since DL is a significant trend, only recent publications that match the search terms are displayed, with the first publication appearing in 2018. The articles chosen cover the period from 2018 to January 2023, which is a five-year period. For this study, only papers from journals, conference proceedings, and book sections were evaluated as sources of information. Throughout each inquiry, we collect data on a number of several items which are shown in Figure 1.

2.3 Eligibility and Exclusion Criteria

Five exclusion criteria are defined for the eligibility validation of the chosen document. These are the exclusion criteria:

E1. Works not included in either the Scopus or WOS databases.
E2. Works that are unrelated to the cracks on the concrete building; concrete pavement are not included.
E3. Only works that include an image or video of the structure's crack are permitted.
E4. Works that just make a claim without presenting any experiment or comparative results.
E5. Works presented or written in a language other than English.

The following sub-section displays the collected data based on the publishing distribution in five-year increments starts from January 2018 until January 2023, which is a period where the use of DL is grown and evolved as the main strategy for identifying a concrete buildings and structures. Since 2018, the study on DL used to identify the cracks has been grown popular [60].

2.4 Review Process

Based on the PRISMA statement there were four stages involved in the systematic review process. These are the identification stage, screening stage, eligibility checking stage, and selecting the included studies stage.

2.4.1 Identification Stage

The subject of the investigation is a combination of three primary search terms: ‘deep learning’, ‘crack’ and ‘concrete structures’. Each term can be searched using a variety of alternative terms. The 'OR' operator was used to select and combine the most pertinent and frequently used applicable terms. For instance, there were three keywords found to describe ‘deep learning’ are "deep learning", "deep network", and "deep architecture". There is also another way of expressing 'crack' and 'concrete structure' were its sole primary term. Additionally, the term 'crack' was used to refer to crack detection, crack classification, and crack segmentation and any other crack detection task. The 'AND' operator was used to concatenate individual search strings into a query engine. To include all possible verb forms of the keywords, the asterisk '*' has been introduced. To ensure that the most pertinent material is included, a full text search was conducted. The keywords used were: *(ALL("deep learning" OR "deep network*" OR "deep architecture*" AND ALL ("crack") AND ALL ("concrete structures"))* and it applied to the following three electronic databases: Google Scholar, Scopus and WOS.

2.4.2 Screening Stage

The first evaluation of the articles was conducted by analyzing the titles. Of the total number of publications obtained (n = 4,163), 2700 were omitted because they appeared in several databases, leaving 1463 studies. Due to their absence from the Scopus and WOS databases, 839 articles were eliminated during the screening stage (exclusion criteria E1). Several titles were found to be disconnected from the topic of the study. Therefore, the total number of authorized titles was 624.

2.4.3 Eligibility Stage

Each of the 624 documents was now acknowledged and based on the exclusion criteria 2 through 5 (E2 until E5), 527 manuscripts were rejected, and therefore 97 papers were accepted and critically reviewed.

2.4.4 Included Stage

For this systematic review, a total of 97 publications were selected and these items consist of journal publications and conference proceedings. The bulk of the articles (87 of them) were listed in the WOS database, while only 10 of the papers were indexed in Scopus. Figure 2 depicts the flowchart of the investigation.

2.5 Extraction and Analysis of Data

This systematic review study used by the Kitchenham approach in addition to the PRISMA Statement method for data extraction and analysis [54]. The procedure necessary to carry out a specified question for the identification, selection, critical evaluation, and analysis of pertinent data from studies that are included in the systematic review. The data (97 chosen publications) were picked systematically in accordance with the PRISMA declaration, which are described in full in Sections 2.1 through 2.4. The primary goal of this analysis was to evaluate the present status of research into the implementation of DL to identify cracks in concrete structures. Moreover, this paper aims to examine uses for deep neural networks and how they have been utilised to identify cracks in the concrete frameworks of structures. Thus, by providing knowledge of current practices, we can build on them to further enhance the existing conditions. This prompts the following three research questions (RQs):

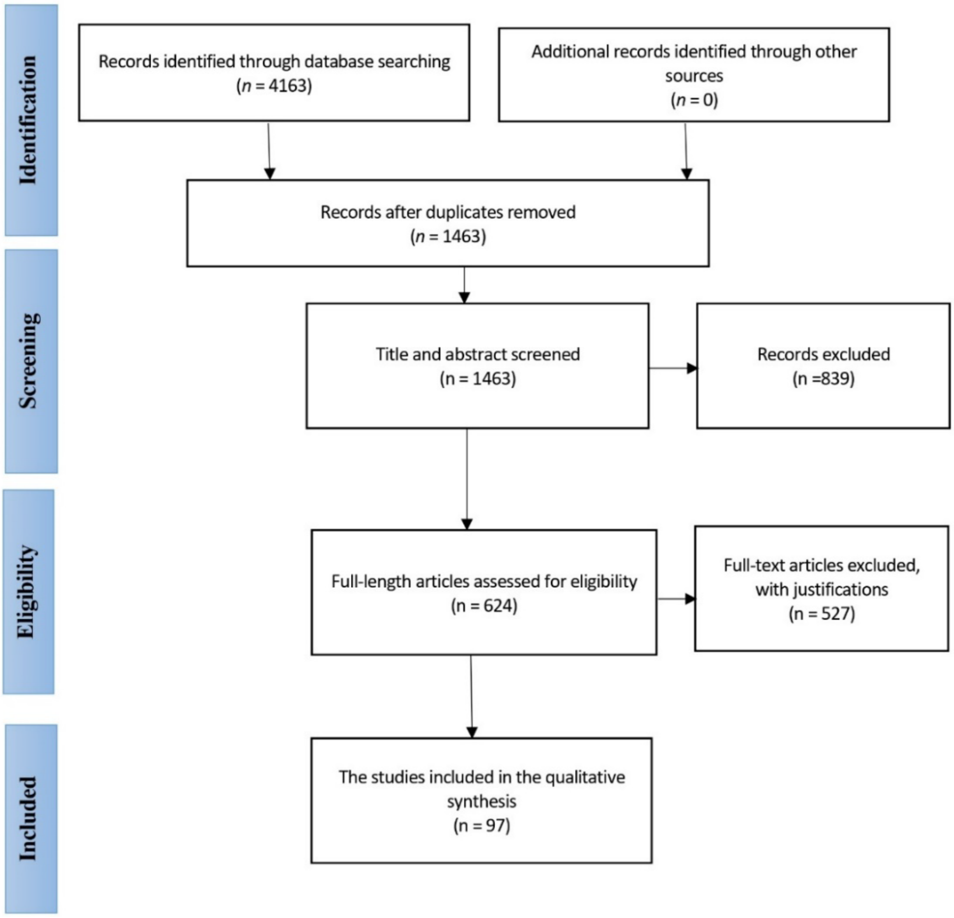

Figure 2. Diagrammatic depiction of Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) [61].

RQ1. Which type of DL technique has been applied to identify cracks in concrete structures?

RQ2. What was the type of dataset has been employed? There are several sub-questions that have been discussed in RQ2 such as: Does the information come from a self-collected source or an open source? Where does the data that were obtained by the individual themselves originate from?

RQ3. What are the limitations associated with crack identification in concrete structures?

Prior studies on DL for identifying cracks in concrete structures were gathered to help with the answers to these study topics. The data extraction and analysis (DA) process has been developed as a guideline to address all research questions as follows:

DA1. All DL techniques: (a) crack classification technique, (b) crack segmentation technique, (c) crack detection technique, (d) hybrid technique applied was collated, examined, and summarized in tabular form to put them in the spotlights of the various DL methods.

DA2. All datasets were collected, analyzed, and tabulated to demonstrate the comparison of datasets that used actual data from structures or open-source sources.

DA3. In-depth examination of the difficulties related with DL techniques were investigated, in addition to the difficulties reported in other research and uses for crack detection on concrete structures.

3. RESULTS

This section shows the collected data based on the publishing distribution in five-year increments, together with the citation analysis.

3.1 Publication Distribution

By using the extraction criteria specified in Section 2, Figure 3 depicts the number of publications published between 2018 and January 2023 in five-year increments. This steep dispersion of the publication trend suggests that interest in DL approaches has increased and that crack identification in concrete structures has garnered more attention and this trend was also found in [62]. The findings of the analysis indicated that there has been a significant rise in the sum of all journal and conference papers dedicated since 2018. The finding shows that 87 of the 97 articles were published in journals, while only 10 articles were published in conference proceedings.

Figure 4 presents an illustration of the total amount of publications in journals and conferences produced annually and Figure 5 illustrates the highest distribution of papers in journals. The majority of the papers focused on applying Artificial Intelligence (AI) to various aspects of the engineering and construction industries. The top five journals were *Automation in Construction*, *Applied Sciences, Computer-Aided Civil and Infrastructure Engineering*, *Sensors* and *IEEE Access*. All the articles were published by reputable publishers. In this case, the top three publishers that published many of the related articles were MDPI, Elsevier and Hindawi.

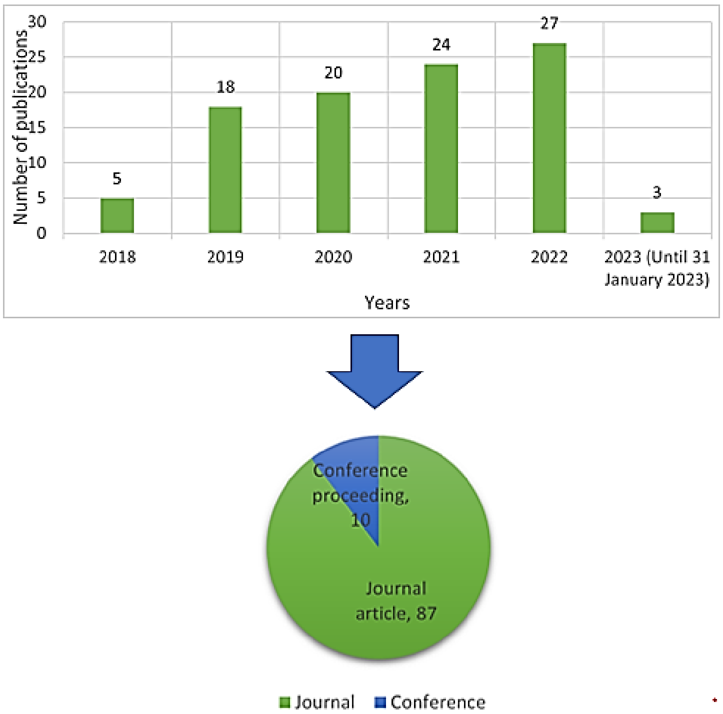

Figure 3. Publications distributed each year and the total number of papers distributed per kind of article (journal or conference paper).

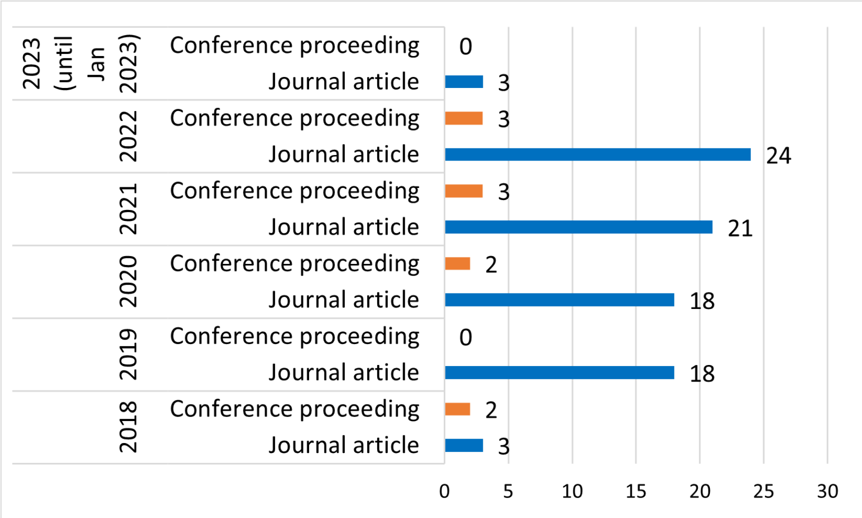

Figure 4. Journal and conference publication each year.

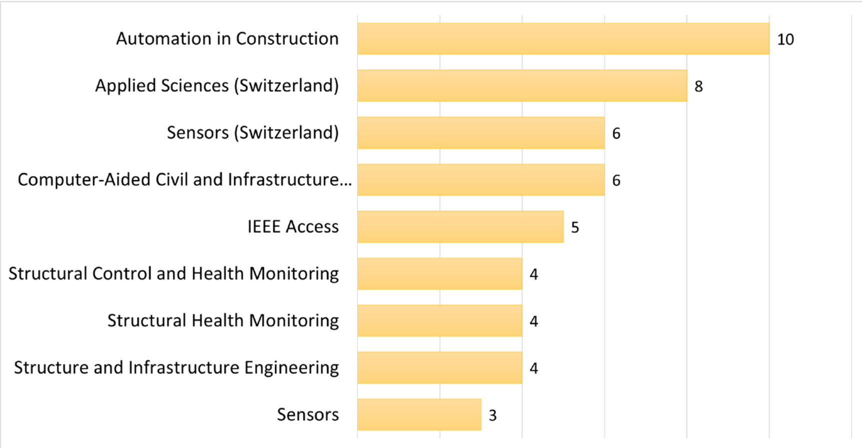

Figure 5. Number of articles published in journals.

3.2 Citation Analysis

Table 1 displays the leading articles with citation quantity for all the selected papers in descending order together with the title of article, year of publication, and total of citation for each publication. All of which were produced as a result of this systematic review procedure and were published between 2018 and January 2023. However, the top-cited papers were mostly from 2018 to 2021. To date, the three articles with the highest number of citations were "Autonomous Concrete Crack Detection Using Deep Fully Convolutional Neural Network" (414), "Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network" (250) and "Deep learning based image recognition for crack and leakage defects of metro shield tunnel" (187). It was found that 16 of the 97 chosen publications have not yet been cited, and the majority are from 2022 or early 2023.

4. DISCUSSION

DL has the capability to classify, i.e., labeling the image patch as crack or non-crack; segment, i.e., segmenting the image pixels into crack and non-crack pixels, detect, i.e., detecting the image of crack or another object. All the selected articles recommended a specific technique for identifying cracks in concrete. In these studies, four DL approaches are used to identify cracks in concrete structures: classification, detection, segmentation, and hybrid as shown in Figure 6. The term "hybrid" refers to a combination of classification and detection, classification and segmentation, or detection and segmentation.

We have categorized the body of literature based on their contributions. The articles to which each part refers are discussed in the subsections that correspond to those sections. Publications that offer a framework based on multiple architectures are categorized exhaustively under each of them. The majority of research (37%) was solely devoted to crack segmentation, 32% to crack classification, 18% to crack detection, and 13% to hybrid methods as depicted in Figure 7. Based on this record, crack segmentations and classifications are key concerns that have been the subject of a great deal of research, providing them an advantage over research that has concentrated just on crack detection in concrete structures. To respond to the RQs, a thorough examination of each publication has been carried out, with the necessary information being extracted. The problem being addressed, the primary strategy, the learning algorithm, the data, the network topology, and the way in which the data were prepared to train and test the network were both taken into consideration in each publication.

Table 1. Overview of the most-cited research paper.

Ref.	Publication title	Year	Cited by
[63]	Autonomous concrete crack detection using deep fully convolutional neural network	2019	414
[64]	Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network	2019	205
[65]	Deep learning based image recognition for crack and leakage defects of metro shield tunnel	2018	187
[66]	Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization	2019	181
[67]	Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques	2018	145
[68]	Image-based concrete crack detection using convolutional neural network and exhaustive search technique	2019	123
[69]	Image-based concrete crack assessment using mask and region-based convolutional neural network	2019	119
[20]	Automated vision-based detection of cracks on concrete surfaces using a deep learning technique	2018	109
[70]	Pixel-level crack delineation in images with convolutional feature fusion	2019	106
[71]	Concrete crack detection using context-aware deep semantic segmentation network	2019	82
[72]	Concrete bridge surface damage detection using a single-stage detector	2020	82
[73]	Deep learning–based autonomous concrete crack evaluation through hybrid image scanning	2019	78
[74]	Anomaly detection of defects on concrete structures with the convolutional autoencoder	2020	73
[34]	Robust pixel-level crack detection using deep fully convolutional neural networks	2019	71
[75]	Automatic crack detection for tunnel inspection using deep learning and heuristic image post-processing	2019	69
[37]	Crack detection and segmentation using deep learning with 3D reality mesh model for quantitative assessment and integrated visualization	2020	64
[76]	Vision-based autonomous crack detection of concrete structures using a fully convolutional encoder–decoder network	2019	55
[77]	A spatial-channel hierarchical deep learning network for pixel-level automated crack detection	2020	51
[78]	Concrete cracks detection based on FCN with dilated convolution	2019	43
[33]	Structural crack detection using deep learning–based fully convolutional networks	2019	38
[79]	Automatic pixel-level crack detection on dam surface using deep convolutional network	2020	38
[16]	Crack detection of concrete structures using deep convolutional neural networks optimized by enhanced chicken swarm algorithm	2022	37
[80]	Real-time tunnel crack analysis system via deep learning	2019	35
[17]	Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures	2021	35

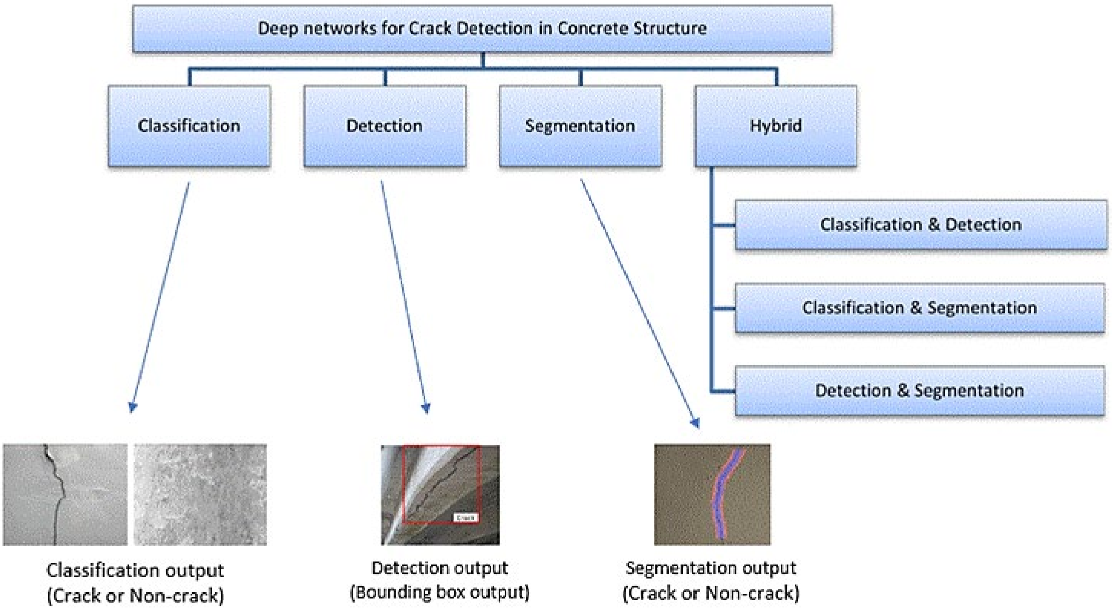

Figure 6. Deep learning techniques to identify cracks in concrete structures.

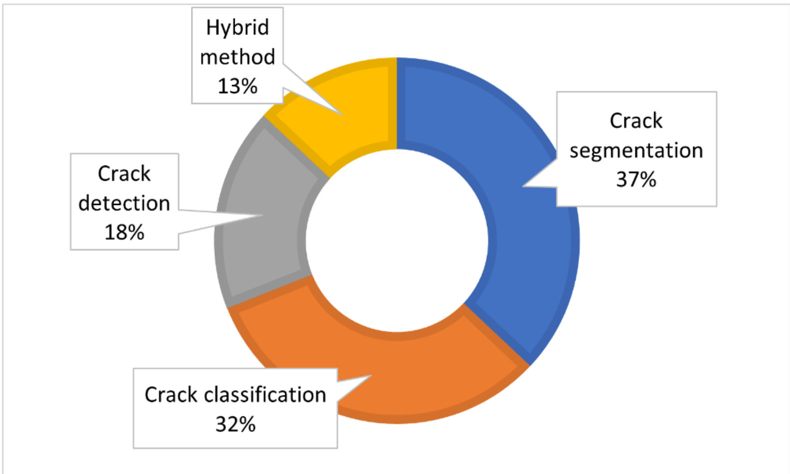

Figure 7. Analysis of crack identification techniques.

4.1 Methods of Deep Learning for Identifying Concrete Structural Cracks

During the 1980s, the CNN, was developed and has since become the most well-known, forward-thinking, and widely used algorithm for DL [81]. It failed to garner interest from researchers because of the era's limited access to computing resources, such as powerful computers and large data storage devices. On the other hand, the idea gained momentum as computing power and database retrieval and storage capabilities of machines improved [82]. Later on, CNNs were used effectively in the solution of classification issues and performed exceptionally well in computer vision applications [83]. Consequently, the detection, classification, segmentation, and hybrid method of a crack through crack recognition on concrete structures has generated a lot of interest.

According to the data, the majority of the chosen papers contribute to crack segmentation (37%) followed by crack classification (32%) and crack detection (18%). Only a small number of researchers employed hybrid strategies to identify cracks (13%). New networks have been researched and addressed for Classification, Segmentation, and Detection approaches, and hybrid methods involved have also been covered and tabulated in Table 2. The numerous networks used in various papers have also been categorized as "Various networks/ Others". For the hybrid approach, a new network has also been identified, but the category "Method to combine" has also been presented, indicating that there were two or three methods involved in identifying cracks, such as categorizing the network in the first phase and segmenting the crack in the second.

Table 2. Studies focus on the techniques for identifying cracks.

Technique	Category	Reference
Classification	New network/ Network configuration	[47], [76], [84]–[91]
	Various network/ others	[17], [44], [73], [92], [93], [100], [101], [102], [103], [104]
Segmentation	New network/ Network configuration	[31], [71], [98]–[101], [77]–[80], [94]–[97]
	Various network/ others	[102]–[107]
Detection	New network/ Network configuration	[108]–[113]
	Various network/ others	[114], [115]
Hybrid	New network/ Network configuration	[37], [116], [117]
	Method to combine	[34], [70], [118]–[120]
	Others	[121]–[123]

4.1.1 Crack Classification Technique

Reis and Khoshelham in [84] proposed the new network called ReCRNet (Residual CRack Detection Network) containing a classifier block, two ResBlocks, a Conv Layers block, and a Stem block. Personal computers and small datasets can benefit greatly from the advantages of this network's architecture. In this study, 71 raw images with cracks and 100 images without cracks compensate the dataset were acquired in Melbourne's historic Royal Exhibition Building. Moreover, in [85], a binary classification of crack patterns into isolated patterns and map patterns was suggested by contrasting the characteristics of structural and non-structural cracks. The ability to recognize crack patterns is accomplished via comparisons of similarity utilizing Deep Image Structure and Texture Similarity (DISTS) index. In the first step of the process, both the reference image and the damaged image are passed through VGG16, a 16-layer CNN. As a result of these tests, the proposed method has been shown to operate and provided accurate recognition of over 96% for crack patterns.

In Chow *et al.* [86], ResNet was utilised as the classifier, and utilize the training dataset to execute the model's training. Both processes were carried out from scratch. Additionally, new data has been obtained as a result of this research. According to findings, the planned inspection pipeline has a superior capacity to classify defects despite being exposed to a range of operating and environmental circumstances. These conditions include angle of capture, distance of camera, and lighting condition. The proposed inspection pipeline has a 95.6% testing accuracy on average. Another researcher, Li *et al.* [87] proposed a framework for high-classification based on ResNeXt with postprocessing also known as “ResNeXt+PP” that was offered in order to successfully discover cracks in concrete. According to the findings, the trained ResNeXt+PP has the ability as a means of automatically distinguishing non-cracks and cracks in the raw picture. The results show that the approach outperforms a variety of others. Additionally, an autonomous concrete structure's potential applications driver for bridge detection robot was addressed. Fu *et al.* in [88] proposed a combination use of conventional CNNs with a multi-layered image preprocessing strategy (MLP), and this uses the Otsu thresholding method and homomorphic filtering as its main components. The adaptability of crack recognition may be improved using the MLP-CNN architecture, in addition to clearly improving detection accuracy and noise immunity, implying that many categories and intensities of a uniform framework may be used to address background noise.

To execute end-to-end learning to classify each pixel, the suggested FCN-based segmentation approach was built atop encoder and decoder frameworks, as highlighted by Manjurul Islam and Kim in [76]. These novel networks have changed the header to VGGNet and set VGGNet as the benchmark. Therefore, the unknown test data was used to verify the optimized FCN model which resulted in scores of approximately 92% for F1 and accuracy, respectively. Following that, Protopapadakis *et al.* in [47] proposed deep CNNs and domain-specific heuristic post-processing strategies to deal with a variety of challenges, such as high accuracy demands, constrained hardware options, short period of time needed to complete a task, low-textured surface linings, bad and inconsistent lighting, an excess of noise , and inadequate training data. In addition, the framework that has been proposed obtains the good performance rates that it achieves in a substantially shorter amount of time compared to the execution time required by other methods. Other than that, it should be noted that Bayesian optimization was utilised to improve the classifiers' hyperparameters for the experimental comparisons. As a foundational element of an autonomous robotic inspection, this mechanism was designed and constructed and it was tested and certified at the Egnatia Motorway tunnels in Metsovo, Greece.

Wang *et al.* in [89] presented an incremental extreme learning machine (ELM) with a multilayer sparse feature representation, that is capable of favorable feature learning and classification. The given ELM-based crack model may be trained effectively without having to laboriously adjust all the network's parameters, in contrast to frequently used DL-based techniques. Research demonstrates that the suggested concrete crack classification model provides superior training efficiency and preferable crack identification accuracy compared to other recently established crack detectors. Finally, in 2018, Sharma *et al.* in [90] suggested a stand-alone CNN that uses a classifier known as the Support Vector Machine (SVM) to categorize data and CNN to extract features. If a given picture patch is part of a certain crack class, the integrated model will automatically extract features and determine this. It was determined based on the findings, that the suggested method was capable of identifying cracks on the surface of concrete with an accuracy of 90.76 %, utilizing a dataset of 550 photos that were collected from diverse places via a digital camera. Kim *et al.* proposed Conv2D ResNet exponential model [91] and evaluated the performance of the other CNN model with several activation layers and categorised the kind of wall defect by transfer learning. According to implementation findings, the suggested Conv2D ResNet model with an exponential activation layer performed better than others Conv2D models like Xception, VGG19, and DenseNet with an F-Score of 0.997826. The suggested Conv2D ResNet exponential model's possible results include selecting the activation layer function that would forecast the kind of wall defect with the greatest degree of accuracy.

A learning algorithm's optimal hyperparameter values are chosen, and these values are then applied to any data set. This process is known as "hyperparameter tuning". Additionally, it can be deduced that by fine-tuning the hyperparameters of the models for the crack classification task, it is possible to obtain promising levels of accuracy by employing fewer epochs and the ideal number of convolutional layers [17]. Next, Cheng and Wang in [67] demonstrated hyper-parameters such as kernel dimension and stride size which can affect the model performance in sewer pipe defects classification namely faster region-based convolutional neural network (faster R-CNN).

On eight datasets of various sizes, derived from two public datasets, Ali *et al*. in [17] utilized and compared four networks, including ResNet-50, VGG-16, VGG-19, and Inception V3 models. In terms of classification and computational time on a little quantity of data, research's customized CNN and VGG-16 models was better than the alternatives. The findings showed that these two models exhibit improved crack classification for concrete structures. Nonetheless, "detection" has been utilized throughout this article when "classification" is more appropriate.

Monirul Islam *et al*. in [124] utilized four transfer learning models for the experimental setup containing VGG16, ResNet18, DenseNet161, and AlexNet. The lack of training data and overfitting have been addressed using data augmentation techniques as transfer learning, random-resized-crop, random-rotation, color-jitter, and random-horizontal-flip. AlexNet exceeds all other models when it comes to performance indicators, obtaining accuracy rates of 99.90%, P rates of 99.92%, R rates of 99.80%, and an F1-score of 99.86%. Additionally, it displayed the time taken for each model's training period. In this instance, AlexNet provided the lowest time. Following this, the researcher has also utilised a variety of networks in order to categorise the cracks in the concrete structure. For example, Mohammed *et al.* in [125] has selected three well-known open-source CNN models (Model1, Model2, and Model3) and Google Colab was used to train the models. A dataset of 40,000 images of 227 by 227 pixels that contains concrete cracks has been used to train the three selected models. Results indicate Model2 with batch normalization had the most outstanding classification accuracy and lowest loss among the three models considered for concrete fracture detection, according to a thorough comparison.

O'Brien [92] applied VGG16 model with weights transferred from ImageNet and to discover the best model for the novel dataset, the transfer model was trained in a variety of scenarios. The different models are tested using 30 images, each 3072 x 4096 pixels in size, all of which were taken from the underground infrastructure at the European Centre for Nuclear Research (CERN). The different models were trained using over 10000 images, validated on 2500 images, all of which are 256 x 256 pixels in size. The effectiveness of the suggested methods shows that a CNN crack classifier can thrive in the challenging tunnel environment.

On the other hand, Zoubir applied VGG16 network based on optimizing a loss function such as Binary and Multi-Class Cross-Entropy loss [93] that calculates the difference between the expected outputs and the ground truth using back-propagation. On the suggested dataset, a VGG16 network was trained using three Transfer Learning algorithms with varied layers. The performance of the model in each learning configuration was evaluated based on classification metrics, computational time, and generalizability. Experiments revealed that retraining the classification layers and the last two convolutional layers of the VGG16 network significantly improved classification performance. The gradient descent optimizer, such as stochastic gradient descent, and adaptive optimizers are typically used in the optimization method to update the learning parameters of the network.

Lee *et al*. suggested method consists of two sub-networks called the crack-component-aware (CCA) network and the crack-region-aware (CRA) network in [126]. The CCA network was tasked with learning the crack gradient component, whereas the CRA network was tasked with learning a region-of-interest by differentiating between important cracks and noise like scratches. Results from experiments confirm that the proposed method is superior to current best practices in crack recognition. Kim *et al*. in [127] has provided a framework for the use of stereo vision using one telephoto lens and one wide-angle lens, allowing accurate crack as well as effective 3D reconstruction. Additionally, a reliable depth estimate method for flat surfaces is suggested. Such surfaces are common in concrete bridges. The efficacy of the suggested technique has been proven in situ using a concrete bridge that is currently in use. A series of crack candidate regions (CCRs) was used to train the classification model using ResNet-18. The bounding box known as the CCR, may or may not have a crack in it.

Padsumbiya *et al*. in [128] proposed the usage of Max Pooling and appropriate optimization techniques. Through input images, the model was trained to distinguish between cracked and non-cracked concrete surfaces. Using pixel-level data, the proposed model predicts and labels images with cracks on concrete surfaces and images without cracks. The final accuracy attained by the proposed CNN model is 97.8%. Due to its affordability and processing effectiveness, the proposed model is a novel method for identifying cracks on low-pixel density images of concrete surfaces, which eliminates the need for expensive digital image capture equipment.

Jang *et al*. in [73] implemented Google Neural Network while keeping the benefits of hybrid pictures. To enhance fracture detection while reducing false alarms, a hybrid picture combines visual and infrared thermography data. Spatial scanning with the infrared camera, vision camera, and continuous-wave line laser installed on an autonomous vehicle allows for efficient inspection of large-scale concrete-made facilities like dams and bridges. Despite their widespread usage, crack identification methods that rely on the judgement of experts are generally inefficient, time-consuming, and inaccurate.

In this part, we have conducted a literature study utilizing CNNs to classify concrete cracks. CNN has been used in a variety of ways as part of a larger framework. or as complete classifier. A certain proportion of them is predicated on the new network or the new design of the network. The characteristics of the data informed the development of the model in every way. Real and synthetic images were used in CNN's training sessions. In addition, different methods of pre-processing techniques have been used by the researchers for the management of unbalanced data.

4.1.2 Crack Segmentation Technique

Zheng *et al*. in [95] proposed the new network in segmentation called "Lightweight SegNet"; the bottleneck depth-separable convolution with residuals and high-precision lightweight bridge concrete crack technique was both used. The root mean square prop (RMSProp) method was utilised to optimize the training process and the cross-entropy loss function is chosen as the evaluation function. The MIoU index has improved to 77.76%, and the accuracy has increased to 97.95% when researchers have compared their results to the most recent techniques DeepCrack and CrackU-net. Next, optimization techniques and hyper parameters have been used to optimise the model's parameters. Xavier's initialization is used to initialise the parameters of the FC layer and convolution layer. The "U-net + Atrous Spatial Pyramid Pooling" architecture has been suggested by [96] and consists of a mix of U-net (an upgraded inception module) and an Atrous Spatial Pyramid Pooling (ASPP) module. The upgraded U-net model yields binary images of concrete cracks, which the network then uses to determine the crack widths. In 2019, Song *et al*. [80] also used the ASPP module as the decoder, which will serve as the benchmark in the next operations regarding datasets; and adopt ResNet18 as the encoder. The encoder's priority has been on feature extraction and inference time in order to identify tunnel cracks, while the decoders have to be on receptive field for semantic segmentation in order to achieve superior segmentation performance.

With the help of a ring-shaped climbing robot, Jang *et al.* [31] demonstrated DL-based semantic segmentation and Euclidean distance transform (EDT)-based crack quantification techniques. Additionally, an automated digital crack map of the region of interest (ROI) on the designated bridge pier has been created. The proposed strategy was tested by experiments to ensure its efficacy and in-situ test results from the Jang-Duck bridge in South Korea were used as proof. New image-based crack identification model of tunnel lining utilising Residual U-Net (ResU-Net) network was suggested by Hou *et al*. in [97]. To address this issue of model deterioration, the residual learning units were included into the U-Net network's encoding route. A lining crack dataset was constructed using information from a highway tunneling project in Western China. The collection contains photos from a variety of western China tunnel constructions, including a highway tunnel 8100 mm in diameter. Three assessment criteria, pixel accuracy (PA), intersection over union (IoU), and Dice coefficient (Dice), provided results that are superior than those of conventional U-Net: 98.67%, 56.45%, and 68.09%, respectively.

In [77], Pan *et al*. introduced a novel network, dubbed a spatial-channel hierarchical network (SCHNet), that was practical to allow the automatic and reliable pixel-level concrete crack segmentation. The self-attention mechanism is specifically present in SCHNet with a base net Visual Geometry Group 19 (VGG19), and it is implemented by three parallel modules: the spatial attention module, the channel attention module, and the feature pyramid attention module. In addition to taking into account semantic interdependencies of channel dimensions and in spatial, it may also adaptively incorporate local aspects into their global dependencies. While comparing state-of-the-art models such as DeepLab-v2, U-net, Ding, Dilated FCN and PSPNet, it was found that SCHNet outperformed the others as compared to a higher degree of generalisation and resistance to disturbances under different conditions. Rough surfaces, holes, and darkness all fell under this category.

Feng in [79] introduced a technique of crack identification on dam surface (CDDS) utilizing deep convolution network. Pixel-level crack detection is a possible means of supplying more comprehensible and reliable detection findings for dam condition evaluation. They started by sending an unmanned aerial vehicle (UAV) along a specified path to take pictures of the dam's surface. To continue, they trimmed the original raw photos. The photos were cropped, and crack locations were manually labelled in order to generate the crack dataset, after which the CDDS network architecture was developed. Finally, the crack dataset was used to train, verify, and test the CDDS network. CDDS network performance was verified by comparing projected results with those of ResNet152-based, SegNet, UNet, and FCN. The test dataset findings showed that the CDDS network was more effective in identifying surface cracks in dams.

In [71], Zhang *et al*. introduced a context-aware overlapping-patch fusion (CAOPF) approach which aggregates results from pixel-level prediction from numerous local overlapping image patches in order to segment cracks in structural infrastructure under different scenarios with overlapping patterns. As its output, a deep convolutional semantic segmentation network provides a prediction for each of the image's local patches. The predictions from several picture patches are then integrated using a CAOPF technique. Methods of data preparation, such as data augmenting data splitting have also been used in this study. Another new network configuration for crack segmentation, in [78], Zhang *et al*. reported an FCN-based method that created dilated convolutions with a multi-branch fusion technique and varying dilation rates. Using a multi-branch feature fusion method with a dilated convolution architecture with varying dilation rates was shown to help reduce the issue of lost detail in experimental settings. However, adding more layers to the network did not enhance the performance of the segmentation. Incorrect segmentation occurs when dealing with photos with complicated backgrounds since lacks semantic label examples with such backgrounds and the dataset is rather basic.

Çelik *et al*. [98] proposed sigmoid optimized CrackNet or SO-CrackNet. Squeeze-and-excitation blocks that have a sigmoid activation function that has been modified are added to this network to improve it. They added a stretch coefficient to the sigmoid function and declared it a trainable parameter, allowing for more differentiated calibration of the feature map throughout network training. Encoder and decoder subnetworks in this network are not only connected sequentially from the bottom to the top layers, but also at the same depth level through intermediate layers. The second employment solution was to adopt kernel initialization through transfer learning (TL). By copying, geo-metrically editing, and pasting crack masks onto new concrete background images, the Copy-Edit-Paste Transfer Learning (CEP TL) method generates thousands of semisynthetic images used to pretrain the network. This CEP TL method significantly improves the performance of models.

Wang *et al*. in [99] proposed SegCrack that uses a top-down pathway with lateral connections to output pixel-level crack identification and adopts a hierarchically structured transformer encoder to output multiscale features. The images from the training set showed a sample imbalance phenomenon in the background and crack pixels. To conduct targeted training on hard examples, the online hard example mining (OHEM) [129] was adopted in the study as the dataset on concrete damage contained a large number of easy examples and a small number of difficult examples. The automatic selection of these difficult examples can result in more efficient training. Each mini-batch contains hundreds of thousands of candidate pixels, and the OHEM was used to subsample the candidate pixels using a lossy distribution. The robustness test results show that SegCrack's performance was stable against a wide range of real-world image corruptions, but the severity of structural damage was not assessed using the segmentation results.

Falaschetti *et al.* in [100] suggested a vision system based on two accurate and lightweight CNN models implemented on a low-cost, low-power hardware, specifically the OpenMV Cam H7 Plus. The proposed architectures (Model 1 and Model 2) were built on the Python-programmable OpenMV Cam H7 Plus platform to suit the specific crack detection task's criteria for a low-power, low-cost, real-time image sensor. As a result, the suggested CNN architectures function well on this platform, requiring just 6 to 26% of the memory required by LeNet, and consistently delivering improved accuracy in all tested instances and on both datasets, with only a minimal increase in inference time. Moreover, Siriborvornratanakul *et al.* in [101] suggested DeepLabV3-ResNet101 was employed as a basis model, and several training methods and loss functions were then tested. The experiment showed that it was feasible to develop downstream models for low-level crack detection on top of an existing DL architecture for semantic segmentation. Although the unique loss function is essential to deal with the drastically unbalanced nature of low-level crack detection, major design changes to the underlying model are not required.

Model training, model prediction, and dataset development were all optimized for the Mask R-CNN model in 2021 by following the guidelines specified for evaluating the effectiveness of these processes [130]. Owing to the original dataset included thermal pictures, which are a mix of blurred and other excellent photos with non-uniform intensities of images due to unequal lighting and illumination, median filtering was used [131]. Mask R-CNN that included a morphological closure operation was introduced by Hongwei *et al.* in [132]. Mask R-CNN models' training hyper-parameters were established by trial and error during the model construction phase, and the morphological closure operation was then added to the M-R-101-FPN model to create a combined model for the testing phase. Although an image database is provided, the number of images is minimal to draw any firm conclusions about the universal nature of the situations being studied.

Researchers have been applying various networks in segmentation of crack in concrete structure. In [102], two approaches were presented by Droguett *et al.* to segment crackss in pictures taken of bridge structures. The first technique substituted a lightweight CNN, such as LeNet 5, for the DeepLabV3+ classifier. Using the second technique, a DenseNet model's layer number is to be decreased. To combine higher-level characteristics with lower spatial resolution with their lower-level, higher-resolution counterparts, skip connections were used in its construction. By advancing low-level details and enabling attention on abstract invariant aspects, these links enhanced deep representational properties. With a lean up-sampling approach, the resultant architecture was able to carry out very accurate semantic segmentation. This drastically lowers the amount of memory needed and makes it possible to run apps that were designed to work with smaller crack images. Yu *et al.* in [106] has contrasted two networks, the U-Net model and the Mask Region with CNN (Mark RCNN). Through up-sampling and skip connection, the U-Net model increases the segmentation accuracy of the model, increasing its Recall, Kappa, and Dice by 6.88%, 1.94%, and 7.72%, respectively. When it comes to very thin and undetectable cracks, Mark RCNN outperforms U-Net in the placement requirements scene, preventing the problem of missing cracks.

In their subsequent research, Jiang *et al.* [103] introduced the Deeplabv3+ model, a cutting-edge DL model for semantic image segmentation, to segment the cracks in visual images. They used the ResNet-101 infrastructure. The network may easily converge to a state in which all pixels are deemed background because the number of pixels in the background is substantially more than that of the cracks. To address this problem, researchers utilized the full-size original photos as input instead of compressing them, and then created a weighted cross-entropy loss where crack pixels are given a greater weight than background pixels by a factor of 10.

Yoon *et al.* suggested a high-resolution image-based crack width measurement approach utilizing the VDSR algorithm to address the challenge of measuring crack width caused by low resolution in photos taken at safe intervals for crack width inspection with unmanned aircraft [133]. The goal was to analyse crack widths by increasing the image resolution matching the crack region at a safe distance interval. A high-resolution algorithm increases the resolution of photographs acquired at 3 m intervals on walls with cracks to measure crack widths. The crack width measurement was compared to a general photograph and a high-resolution conversion image. Thus, the high-resolution picture's crack width measurement matched the actual value. These discoveries may help to overcome camera resolution and distance limits during plant safety checks to make UAVs more viable [1].

Wan *et al.* in [104] employed single shot multibox detector (SSD) technique and sliding window technology for bridge crack identification. It has been demonstrated that four types of cracks achieve good detection results with a precision of more than 95% and a recall of more than 75% using model training on the training set and parameter tweaking on the validation set. This approach can implement a direct mark on the crack image, and the crack hits were all well marked with the sliding window, which typical neural network algorithms cannot perform. Moreover, Zhou *et al.* [106] applied multi-scale dilated convolutional method to improve the accuracy of segmentation. The encoder-decoder structure of U-Net serves as the foundation for the proposed network. To increase the receptive field without using additional parameters, cascade multi-scale dilated convolutions have been used in the network's centre. The multi-scale and multi-level side network features were integrated for the final prediction in the decoder stage using the feature fusion module. It was demonstrated that the suggested strategy can robustly identify a variety of crack types and obtain superior outcomes on these datasets. Additionally, it demonstrated the value of multi-scale data in crack detection.

Hong *et al.* introduced an AugMoCrack network, which is a bounding box-level crack detection method for weakly supervised crack detection achieved by augmenting the training data with Poisson blending and high-frequency discrete cosine transform-based features. [105]. The proposed AugMoCrack detects the box position of the crack object from a morphological standpoint, including neighbour connection inside the crack pixels and crack-area fitting. In addition, new morphological attention loss functions for considering neighbour connectivity and the box area border have been proposed. The proposed network also outperforms previous state-of-the-art crack-detection methods in an environment with insufficient training data and weak supervision. Ye *et al*. [107] proved the advantage of DL-based methods in crack segmentation by contrasting the performance of the FCN called Ci-Net with that of conventional edge detection algorithms. They designed a DL-based crack detection tool for smartphones in 2022 [134] and created a pruned crack recognition network by reducing DNN size using the pruning method.

4.1.3 Crack Detection Technique

To identify concrete cracks from images and subsequently quantify their orientations and maximum widths, Li and Zhao suggested a DenseNet-121-based CedNet in [108]. The weight and bias parameters of the CedNet were initialised using a transfer learning technique. The CedNet works well in tests, and this method was determined to be effective for correcting distorted pictures and measuring crack sizes and orientations. Testing was done on the derived algorithms for crack identification and measuring and compared CedNet to FCN and Mask R-CNN. As a consequence, CedNet outperformed the competition in research comparing how well it could find thin cracks in concrete.

Li *et al*. in [109] employed upgraded You Only Look Once version 3 (YOLOv3) and the crack recognition network with two modules, the crack encoder-decoder network (CEDNet) and the crack residual refinement network (CRRNet). After the sleeper on the ballast bed is extracted using the grey projection approach, the upgraded YOLOv3 network was used to discover cracks on the sleepers and segment them. The sleeper is entered into CEDNet for the extraction of crack features in order to forecast the coarse crack saliency map. The prediction graph is fed into CRRNet in order to optimise its local region and edge information. A combined loss function comprising binary cross-entropy (BCE), structural similarity index measure (SSIM), and intersection over union is used to increase the crack identification model's accuracy (IOU). Results demonstrate that this approach is capable of correctly detecting the sleeper crack image.

Kumar and Ghosh in [112] demonstrated dual-channel CNN called Dual-Channel Deep Convolutional Network (DuCCNet), and close variants of it was extremely promising for drone-based civil structure surveillance, as it has been optimized to counter the presence of random rotations, zooming, and intensity scaling. They utilised a camera to gather pertinent images, all of which were shot on the campus of the National Institute of Technology in Durgapur. Due to this, a dual-channel deep CNN was developed that demonstrates great accuracy (92.25%) and robustness in detecting concrete cracks in realistic settings. Furthermore, to improve computer vision-based automated crack identification, Bae *et al*. proposed SrcNet in [113]. SrcNet is an end-to-end deep super-resolution crack network. When cracks in large-scale civil infrastructure are photographed by autonomous robots, the quality of the images may be compromised by factors such as poor pixel and resolution motion blur. Thus, SrcNet has been used to substantially boost the detectability of cracks by enhancing the digital image's raw pixel resolution. Experimental verification of SrcNet was accomplished via the use of digital images captured by an unmanned aerial vehicle and a climbing robot of South Korean concrete bridges. Compared to detection results obtained using raw digital pictures, the suggested SrcNet shows a 24% improvement in detectability for cracks.

Moreover, the EfficientNetB0 was enhanced by Su and Wang in [110] to use transfer learning for surface crack detection in concrete. The weights for ImageNet were pre-trained, and supervised learning was then utilised to fine-tune the network parameters using the Adam optimizer. The testing phase also made use of crack images captured in other locations to assess the model's scalability. The detection findings were compared to the InceptionV3, DenseNet201, and MobileNetV2 models and it demonstrates that this model significantly brings down the number of variables in the meantime maintaining an accuracy at a very high degree (0.9911) and a strong capacity for generalization. Bayesian optimization was used by Liang in VGG-16 in [66] to identify a minimal dataset collection. The VGG-16 achieves the highest testing accuracy when overfitting tendency is present, below the limited sample size and it is effectively handled by choosing the values of the hyperparameters. This allows for the VGG-16 to achieve its goal. Based on the findings, it can be concluded that Bayesian optimization produces models with good robustness curves. In [114], Li *et al*. introduced the concept of a variable network, which involves first using a faster region CNN (Faster R-CNN) to create coarse crack region localization and classification, and then using edge extraction to create fine crack edge detection, achieving both great efficiency and accuracy. Observably, the associated fine edge detection output image makes quality of tunnel evaluation with size estimation and crack length uncomplicated.

Kumar and Batchu [115] applied the YOLO-v3/TYOLO-v3 DL model in a multi-drone real-time damage detection system (DDS) for tall buildings. Input images and videos can be processed in real-time, allowing for a speedy determination to be made. Models of the YOLO family only need to be trained once, and they can be run on any PC with a GPU. Furthermore, the data captured in the few prior systems, including videos and photos, will be transferred via the Wi-Fi channel, making it more dependable than the existing systems. In the event of a communication breakdown, the information will be destroyed and must be resent. Instead, the DSS just sends out the damage data that it already has. Cracks larger than 0.2 mm were consistently recognized from a distance of around 0.4 m, as shown by the YOLO-v3 DL model. From this experiment, they can infer that drone-based DDS can be used to examine the spalling of high-rise concrete buildings and the penetration of various sized and shaped cracks in concrete.

Other research by Deng *et al*. in [111] presented three distinct regular detectors: the region-based fully convolutional networks (R-FCN), the Faster R-CNN, and the Faster R-CNN based on a feature pyramid network (FPN). Implications of the proposed method were examined by comparing the findings of the proposed detector with those of conventional detectors. The findings demonstrate that the mean average precisions (mAPs) attained by the Faster R-CNN, R-FCN, and FPN-based Faster R-CNN for crack identification are improved by the inclusion of deformable modules. Furthermore, these detectors use deformable modules that allow them to detect out-of-plane cracks, otherwise invisible to standard detectors.

4.1.4 Hybrid Detection Technique

In the hybrid category, a new network has been established. Hybrid detection approaches consisting of crack detection and segmentation were presented by Kalfarisi *et al*. in [37]. One approach combines the faster region-based convolutional neural network (FRCNN) with structured random forest edge detection, where was quicker than either method alone (SRFED). The FRCNN was used to detect the bounding boxes containing the cracks, and the SRFED was used to segment the cracks within those boxes. The next method straightforwardly employs Mask R-CNN for crack segmentation and classification. Using 3D reality mesh-modeling technology, both methods have been combined to provide a quantitative evaluation and unified depiction of the investigated structure. Real-world tunnels, bridges, and constructions are used to test and illustrate the effectiveness of the suggested methods.

Kee *et al*. proposed Mask Region-based Denoised Deformable Convolutional Network (R-DDCN) [116] to identify cracks for accurate instance segmentation and image classification. The use of denoised deformable convolution improves the modelling capability of the convolution layer. It utilizes the current deformable convolution with non-local means as a mechanism for denoising to maximize the spatial sample locations' enhancement with filtered offsets. As a result, it can enhance the capability of CNNs to represent unknown geometric transformations and enhance the model's precision with a larger computing demand. Yadav *et al*. [117] proposed 3SCNet (3ScaleNetwork), an unique deep convolutional neural network, for crack detection. The SLIC (Simple Linear Iterative Clustering) segmentation approach generates the cluster of similar pixels, while the LBP (Local Binary Pattern) locates the texture pattern in the crack image. SLIC, LBP, and grey pictures are given to 3SCNet to generate a feature vector pool. On a public historical building crack dataset, this multi-scale feature fusion (3SCNet+LBP+SLIC) technique achieved the maximum sensitivity, specificity, and accuracy of 99.47 %, 99.75 %, and 99.69 %, respectively.

Crack delineation network (CDN) has been developed by Ni *et al*. in [70]. The feature pyramid network (FPN) was used to accomplish feature map fusion to create the CDN model after the output of a generic pretrained CNN model, GoogLeNet CNN. The next step was to utilize the fused feature maps to detect and segment cracks at the pixel level using consecutive convolutional layers. The concrete crack image database was used for both CDN training and evaluation. For fully autonomous structural crack identification, the results show that the proposed framework can accurately and rapidly highlight cracks in images. However, the researchers note that quickly and accurately autonomously identifying cracks at the pixel level is a challenging task. By converting well-known image classification architectures' fully connected layers into convolutional filters, Alipour *et al*. in [34] introduced CrackPix (Fully Convolutional Network for Pixel-Level Crack Detection), which utilizes well-known image classification architectures for dense predictions. FCN was used for the purpose of segmenting cracks at the pixel level, according to the fundamental principles of fully convolutional networks. The feature extraction has also been transferred using the VGG-16 architecture, apart from the upsampling layer, which was trained using bilinear upsampling initially. CrackPix was capable to accurately recognizing over 92% of crack pixels and 99.9% of non-crack pixels in the validation set, according to a sensitivity study. Fully convolutional neural networks (FCNs) with an encoder and decoder architecture have been presented by Kim *et al.* [118], which conduct pixel-wise classification to reliably detect cracks. The researcher then utilised the described approach to create a neural network for crack segmentation, which they named HiRes3DNet (see also [135]). However, images captured in the real world do not include physical length data, despite the fact that the suggested measuring technique was tested using ideal data comprising the physical length of a pixel.

In [119], Yang and Ji used the refined VGG-16 model as a classifier to classify the crack images and save computer resources for non-crack photos in further processing. The following step is the crack semantic segmentation, which was accomplished using the U-Net++. In order to properly classify the gathered images, the network parameters, knowledge of the general model framework, and data collected can be transferred to the crack detection job using transfer learning. Results from the experiments showed that the suggested approach is able to identify the cracks at the pixel level. Furthermore, Guo *et al*. [120] used Mask-RCNN and U-Net model to classify and segmentize the crack on the special structures called High-performance fiber-reinforced cementitious composites (HPFRCCs). A trained version of Mask-RCNN was used to identify and locate cracks; ultimately, the detected cracks were examined by the trained U-net in order to quantify the cracks based on pixel value. In order to enhance the accuracy of crack detection and segmentation, three methods of data augmentation were used: horizontal, vertical, and horizontal plus vertical mirroring. This study demonstrates that the given method for crack identification reaches an accuracy of 99.2%.

A two-stage transfer learning pre-processing strategy was proposed by Li *et al* [121] that included both cross-domain and within-domain transfer learning (from ImageNet to structural ImageNet). In order to boost the model interpretability without depending on thorough manual data annotations, researchers employed gradient-weighted class activation mapping (Grad-CAM) to classify and weakly-supervised identify the underlying surface cracks on an arch dam. Despite extensive environmental background inference, to demonstrate, by experiment, that the suggested system can recognize structural faults in concrete dams with good accuracy and recall without needing refined human annotations.

One of the most common ways to achieve data oversampling is via data augmentation techniques. In addition, GAN has been used to synthesize data. In [122], Chen *et al*. used a Generative Adversarial Network (GAN) to simulate a large number of crack images that matched the characteristics of genuine samples. In order to segment fractures when pictures were identified as cracks in the classification test and the background, DeepLab v3+, a state-of-the-art pixel-level segmentation approach, was applied. The amount of training and validation images has been increased using five conventional data augmentation techniques: random rotation, random scaling, random X/Y inversion, random Y-direction translation, and random X-direction translation. Three neural networks, VGGNet13, ResNet18, and AlexNet, were used by Yang *et al*. in [123] to classify crack photos. The experiments demonstrate the ResNet18 model yields the finest outcomes. In addition, the trained YOLOv3 model was able to detect the crack area with a high degree of precision. To further enhance the model's generalizability, the study additionally makes use of data augmentation techniques such as adjusting the data's brightness, saturation, horizontal flip, vertical flip, random cut, rotation 180 degrees, and random zoom aspect ratio. The results of this research suggest that DL, a new methodology, might one day replace conventional methods of crack detection and identification.

4.2 Pre-Processing Techniques

Pre-processing is vital in DL to optimize input data for better model performance. Techniques like data normalization ensure balanced scaling across features, preventing dominance. Data augmentation enhances training set diversity through transformations, reducing overfitting. Noise reduction eliminates unwanted artifacts or noise. Dimensionality reduction reduces computational load by reducing input features. Handling missing data ensures complete and accurate training data. Appropriate pre-processing strategies should be chosen based on the specific problem, dataset, and DL model, with experimentation and analysis to find the best approach. The work in [136] emphasizes the relevance of data augmentation in overcoming the dearth of labelled picture data, providing a comprehensive review and taxonomy of various image data augmentation methods across computer vision tasks, including empirical comparisons, challenges, and future perspectives. Table 3 provides an overview of the literature cited in this section, which the mentioned pre-processing techniques include leveraging pre-trained networks [6], data augmentation involving various transformations like flipping, rotating, scaling, blurring, and color adjustment [20], [67], [137], [138], [139], [140], image enhancement [141], feature extraction [142], median filtering [131] [143], sliding-window assembly, resizing, interpolation, and image cropping [65], random rotation-based augmentation [92], [144], leveraging pre-trained weights [72], geometric operations such as rotation and flipping [74] , confetti noise method [137], and initializing weights and biases in pre-existing models [108]. These techniques are employed to enhance data quality, increase diversity, extract relevant features, and improve model performance in various tasks, including crack segmentation and detection.

Table 3. Pre-processing techniques involved.

Crack classification	
Reference	***Pre-processing technique***
[6]	Leveraging the knowledge and learned representations from pre-trained networks
[20]	Data augmentation (blurring, color adjustment, and rotation)
[67]	Data augmentation (scaling, rotating, flipping horizontally and vertically, & tuning colors)
[68]	Mean reduction and image cropping
[91]	Leveraging pre-trained models with different activation layers
[141]	Image enhancement
[92]	Rotation-based data augmentation in crack
[137]	Image augmentation and random 50% confetti noise method
[94]	Image resolution changes
[145]	Data imbalance handling
[146]	Object-level data augmentation
[147]	Rotation-based data augmentation
[148]	Contrast adjustment and noise addition
Crack segmentation	
[48]	Horizontal and vertical flips for the purpose
[63]	Loading a pre-trained model and feature extraction
[65]	Sliding-window assembly, resizing and interpolation, and image enhancement using morphological operations.
[69]	Data augmentation (right-left flipping or random up-down)
[131]	Median filtering
[143]	Data augmentation
[149]	Image acquisition, manual annotation, image conversion, random cropping
[144]	Random rotation-based data augmentation
Crack detection	
[64]	Data augmentation (horizontal flipping)
[66]	Data augmentation
[72]	Leveraging pretrained weights from a similar dataset
[74]	Geometric operations (rotation and flipping to sets of images)
[108]	Initializing the weight and bias parameters on pre-existing model
[142]	Feature extraction
[138]	Data augmentation (flipping, rotating, and scaling)
[139]	Data augmentation method (combining multiple images for small objects)
[140]	Data augmentation (applying filters to generate additional training examples)

4.3 Dataset-Based Analysis

There are substantial data pertaining to the dataset that each researcher has used, and these data are based on 97 selected papers. 61% of the researchers utilised self-collected images or videos as their main data source, 31% of the researchers used open sources data that could be retrieved via the internet or any other platform, and just 8% used both platforms to obtain the data. Table 4 shows that most of the authors have used a self-collected crack image dataset through camera, CCTV (image or video), smartphone, and tactile sensor to collect crack images. It is not recommended that smart phones be used as embedded devices mainly because of their slow speed performance and a lack of adaptability to challenging circumstances. However, smartphones have the capability to identify photos [150]. Furthermore, some researchers also prefer to install the camera on Unmanned Aerial Vehicles (UAVs), robots or drones. Volume and quality of the dataset are important issues which are necessary for the learning component of neural networks [151]. The scale of the data set has a considerable impact on the model's accuracy. If only a small dataset is available for analysis, it is unreasonable to predict that a model trained on a huge dataset will perform more effectively than one that was trained on a small dataset [150], [152]. Poor data quality, a lack of training data, and location-specific characteristics like the position and angle of capture, which impede the localization of equipment, all contribute to the less-than-ideal accuracy of these prediction models [153]. Creating a larger dataset comprised of high-quality images and videos gathered from a variety of perspectives and places might increase the model's accuracy and stability. However, in this review paper, none of the studies make any mention of the specific number of datasets that are necessary for identification of cracks using DL. Furthermore, the system's ambiguity will be increased in circumstances in which the detection of certain pieces of equipment requires knowledge of supplementary tools on which they rely [154] and these issues will have to be accounted for in the model. In addition, the review analysis reveals in Table 5 the articles were published by researchers from 22 countries. The top three countries were China (48 articles, 49.8%) South Korea (17 articles, 17.5%) and India (6 articles, 6.2%).

Table 4. Crack image database.

Self-collected images	
Type	**References**
Camera	[37], [44], [48], [69], [70], [72], [74], [80], [85], [86], [89], [90] [148], [141], [93], [127], [95], [96], [130]–[132], [111], [115], [138]–[140], [142] [123], [155]
CCTV image/ video	[67], [102]
Smartphones	[68], [108], [134], [137], [139], [147]
Tactile Sensor	[103]
Unmanned Aerial Vehicles (UAVs)	[73], [79], [146], [124], [156], [133], [113]
Robot	[31], [157]
Drone	[84]
Open source / internet images	
Type	**References**
Middle East Technical University (METU) dataset	[76], [78], [158], [118], [99], [116], [17], [159]
SDNET2018 Dataset	[77], [100], [91], [6], [17], [119]
CrackForest Dataset (CFD)	[126], [71], [106], [160]
Collected from the Internet	[48], [161], [125], [50],
Crack500 dataset	[105], [106], [143]
Fire Crack Dataset (FCD)	[126]
AigleRN Dataset	[126]
Tomorrows Road Infrastructure Monitoring, Management Dataset (TRIMMD)	[71]
Customized Field Test Dataset (CFTD)	[71]
COCO dataset	[162]
CrackTree260	[101]
European Centre for Nuclear Research (CERN) dataset	[92]
DeepCrack	[143]
Bochum Crack Dataset (BCD)	[98]

Table 5. The number of papers and citations from the nations contributing to the study subject.

Country of origin	Number of documents	Total of citation	Country of origin	Number of documents	Total of citation
China	48	1021	Canada	1	15
South Korea	17	330	Chile	1	2
India	6	422	German	1	0
Australia	3	496	Greece	1	0
Singapore	3	81	Ireland	1	0
Thailand	3	84	Malaysia	1	0
United State	3	106	Mexico	1	145
United Arab Emirates	2	208	Morocco	1	0
Bangladesh	2	20	United Kingdom	1	69
Japan	2	32	Vietnam	1	18
Italy	2	26	Virginia	1	3

5. LIMITATION AND CHALLENGES

This section includes a study of the problem to be solved, a discussion of the limitations of present work, including a discussion of how DL may be used with high-tech civic tools. Firstly, the suggested models are restricted by the size of the crack picture. Developing a new database system necessitates a limited access to databases including bridge images in [147], as is the case with Thailand's Department of Highways. Meanwhile, a new cloud-based data gathering system is being built to take photos and temporarily store data for an extension of this study. It should be emphasized that the provided picture database is still rather limited in comparison to the sample size needed for generalization [132]. The suggested method's balanced accuracy, F1 score, IoU, and robustness can be improved in the future by expanding the database to include other types of cracks. Furthermore, Huang *et al.* in [65], following action in creating a more diversified group of defect images is to acquire additional lining photographs from various metro tunnel lines, which would alleviate the restriction that all picture data originates from a single set of tunnel images.

The second limitation is the computation time that has been highlighted in [119] and [120]. Further work is needed to speed up the computing process. During research that was carried out, a viable strategy is to substitute more effective techniques for the Mask-RCNN model, as the Mask-RCNN model contains a mask generation branch that significantly increases computation time. Additionally, with a small dataset, the Mask-RCNN is incapable of generating masks on crack. It is hypothesized that increased computation efficiency can be achieved by utilizing fast object detection models. In another case, the amount of learnable parameters has a substantial impact on the model's computing time. The ideal amount of training samples for the customized CNN and VGG-16 models varied significantly with illumination shadowing, concrete surface type, and crack patterns, as shown by Ali *et al.* in [17]. Nonetheless, when comparable image attributes are examined at each epoch, expanding the sample size of training data without enough variation, both models performed less well. Additionally, accuracy with great potential can be achieved via optimizing the number of convolutional layers in the models and reducing the number of epochs, adjusting the models' hyperparameters for the fracture classification job, as well.

Additionally, the appearance of foreign objects and shadows in the crack images will significantly complicate the DCNN detection results as highlighted by [44]. As a result, future research will likely enlarge length or breadth within the range of detection, the database for crack, and incorporate new detrimental environmental circumstances. Similarly, as demonstrated in [47], [63], [125] and [85], designing DL based techniques to identify concrete crack physical properties and varying crack size estimate remains a problem, even more so when the test image contains a significant number of noisy crack-like features. But, in [44], the results illustrate the robustness and environmental adaptability of the Inception-ResNet-v2 network such as bright or dimmer lighting, interference from background noise; however, when human variables cause visual blurring, accuracy and F1 score alter considerably with standard deviation.

In [6], extensions to the current algorithm include analyse microcracks which is smaller than 3 mm and cracks in a range of additional building materials, including corrugated rods, bricks, masonry, and mortar blocks. The importance of this research rests in the lengthy and intricate method employed to optimise a model's performance given a batch of photos. Meanwhile, Chen *et al.* in [122] exhibits the suggested the Transfer Learning on three-stage detection method's excellent detection performance includes CNN classifier, GAN, and DeepLab v3+. However, only photos with a white backdrop have been generated thus far. Typically, the context of different crack images is complicated in real engineering. A viable option is to combine the quantity of cracked photographs with real-world background, resulting in a sufficient several cracked images with the complicated background.

The majority of research on identifying cracks in concrete structures has been conducted on static images [67], [77]. The solution on these problem has been highlighted by Cheng and Wang in [67], to consider the accumulation of a big quantity of inspection videos, DL-based techniques for interpreting time series frames or videos. It has also been agreed in [77], however, there is still a major obstacle to properly recognizing structural danger from dynamic photos with a high degree of noise. Therefore, instead of running a model on an actual engineering project, they utilised a publicly available data set. They could attempt to generate their own dataset and conduct crack detection to further expand the model's usefulness. UAVs are an excellent instrument for doing physical inspections of civil engineering because they can rapidly record many images or videos of cracks and their characteristics under a variety of structural systems [37], [73], [77], [127], [156].

The inability to identify cracks in real time is another drawback of this area of research [71], [108]. Song *et al.* in [80] has developed the tunnel crack dataset, annotated with semantic segmentation terms. To develop a comprehensive tunnel crack recognition and analysis system, the researcher also presented the rapid tunnel crack identification approach using semantic segmentation in computer vision. Yet another major barrier to advancement in image semantic segmentation is the need for substantial human effort to perform extra annotations for segmentation. During a research that was carried out in [121], there should be some thought given to real-time detection in order to facilitate the installation of various image capture devices that can automatically gather, analyse, and identify structural damage to a dam's structure. In the study of [17], [31] and [103], integration with IoT devices was proposed, along with the usage of a prototype for a real-time robotic video inspection system. As a result, these real-time robots will be expanded to detect a variety of different sorts of damage, including rust, exposure, spalling, and efflorescence [31]. Furthermore, Alzubaidi *et al.* [163] provided an overview of the Internet of Drones (IoD), encompassing its applications, deployments, and integration in domains such as unmanned aerial vehicles, wireless sensor networks, and mobile computing, which will be explored in the present study.

Additionally, crack severity characterization is often overlooked, but it is crucial. This point was also made in [66], [141], [110] and [121]. Defects detection using CNNs and transfer learning would therefore significantly increase the level of intelligent management for cracks in concrete structures, which will benefit all parties involved. A few corrupted bounding box images might be retained because of the algorithm's rapid execution rate. This problem has brought to light by Kumar *et al.* [115] and need to be solved in the future by adding a buffer. In 2022, a new bounding-box object augmentation (BoxAug) method was proposed by Lee *et al.* [146] to enhance the effectiveness of DL models in identifying flaws on the facades of residential buildings. This may lead to situations where the bounding boxes of different classes are inappropriately copied into new locations. In the original image's trees and roads, for instance, the bounding box of a crack might be found. The BoxAug dataset was modified, however it was still unable to prevent overfitting problem due to the entirely heterogeneous background. Additionally, the dataset used in this research did not completely capture all patterns, including those related to weather, time of day, distance, and shadows. As a result, BoxAug's effectiveness can be increased by incorporating more original picture datasets reflecting various factors.

Despite the efficacy of the suggested strategy in determining the number and width of cracks, the lack of depth information in the photos prevented computer vision technology from measuring the crack depth [125], [127], [111], [120]. Based on [148], [149], and [111], 3D photos may be used instead of 2D images to extract internal information about cracks or to increase the accuracy of severity level assessment since 2D images lack depth information about cracks. To prepare for a potent network for automatic crack identification, a large picture dataset comprising a range of various sorts of cracking patterns is needed.

To improve the performance of the network, research on the effects of various hyperparameters used during the training phase has to be performed [118], [131]. In order to train the YOLOV2 network end-to-end to recognize concrete fractures in real-world images, Deng *et al.* in [138] employed stochastic gradient descent with momentum (SGDM). The hyperparameter issue also occurred in [95], where the number of hyperparameters was reduced and consideration was given to adjusting as little as possible during the implementation of this method, but some hyperparameters still needed to be set, and the training and verification sets, respectively, served as the sources for these hyperparameters. Additionally, by maximizing the number of convolutional layers and minimizing the number of epochs in the models for the crack classification problem, as well as by improving model performance on the crack classification problem by adjusting model hyperparameters, promising levels of accuracy may be reached [17], however, the model's computation time is significantly influenced by the number of learnable parameters.

Table 6 highlights the challenges of the studies reported in this review, as well as the solutions that have been proposed to address those problems. Most of the model is relevant to static photos, but the functionality may be expanded to include videos with time series integration. Nevertheless, the researchers face major hurdles in directly perceiving threat to the structure from noisy moving pictures. One frequent problem of nearly all vision-based techniques is their inability to discern the depth of damage, which is caused by the flattened form of photographic images. In the presence of shadows and foreign objects, the outcomes of the CNN detection algorithm will be considerably affected while working with defect data.

Table 6. Challenges and possible solutions.

Challenge	**Solution/ Possible solution**	**Reference**
Limited datasets.	Develop a cloud-based data collection system to facilitate the acquisition of more data. Focus on obtaining additional lining images to expand the dataset.	[147], [132]
High computation time.	Use fast object detection models.	[120]
Presence of shadows and foreign objects on the crack images (noisy crack-like features).	Extend the crack database to include a wider variety of crack images. Define detectable length and width ranges for crack classification to account for variations in crack sizes. Add additional unfavourable environmental conditions to the training data to make the model more resilient to changes in illumination and environment.	[44], [125], [63], [47], [85],

Limited focus on static images in concrete crack detection studies.	Utilize extensive collection of inspection videos for crack analysis. Employ publicly available datasets instead of real-world engineering projects. Incorporate Unmanned Aerial Vehicles (UAVs) for civil engineering assessments.	[67], [77], [115], [30], [86], [64], [164]
Lack of real-time detection.	Utilizing hardware accelerators or specialized processing units to speed up the detection process. Binocular vision and target matching algorithms will be combined to enable this real-time monitoring. Employing DL models optimized for real-time performance, such as lightweight architectures or model compression techniques.	[121], [17], [103], [31]
Classifying the severity of cracks.	Utilize CNNs with transfer learning for intelligent crack detection in concrete structures. Implement CNNs models with a multi-class classification approach to classify the severity of cracks. Apply advanced feature extraction methods, such as texture analysis or edge detection, to extract relevant crack features and feed them into a classification model.	[110], [141], [66], [121], [6], [122]
Rapid execution rate of the algorithm leads to the retention of corrupted bounding box image.	Adding a buffer to prevent the retention of corrupted bounding box images in the future. Introduce the bounding-box object augmentation (BoxAug) method.	[146], [115]
Lack of depth information in 2D photos hinders computer vision technology from accurately measuring the crack depth and assessing degree of crack severity.	Utilize 3D photos instead of 2D images to extract internal information about cracks and improve the accuracy of severity level assessment. Build a large dataset comprising diverse types of cracking patterns to prepare a robust network for automatic crack identification.	[120], [111], [127], [125], [149], [148]
Network performance needs improvement	Reduce the number of hyperparameters and adjust them as minimally as possible during implementation. Set remaining hyperparameters based on the training and verification sets. Maxime the number of convolutional layers and minimising the number of epochs to improve model performance for crack classification. Adjust model hyperparameters to achieve promising levels of accuracy.	[118], [131], [138], [17]

6. CONCLUSIONS

This article has reviewed articles related to identification of cracks in concrete structures by providing methodical answers to the research questions that were presented in the literature review planning protocol. Depending on the scope of the study and the amount of data available, the articles included in this systematic review have used a variety of DL methods and architectures. Four areas related to the research field have been covered including classification, segmentation, detection, and hybrid methods. Various issues have been addressed, and limitations, challenges and possible solutions have been proposed. In light of the results of this systematic review, attention should be focused on the following areas with the purpose of making further contributions to the field:

a) By offering real data for the research purpose, the vast majority of works on crack identification made use of camera or CCTV photos or videos to gather data about flaws in preparation for analysis. Indeed, several of this research used robots, UAVs, or drones equipped with cameras. Several of them captured the images using sensor data and smartphone technology. For deep neural networks to be trained to ac-accomplish exact feature learning, it is required to employ defect data that has been precisely labelled.

b) Most of the research on locating cracks in concrete has relied on static images. In order to facilitate real-time detection, it is suggested that the accuracy of the defect learning networks be improved in operational scenarios. The formation of a brand-new network and many factors were adjusted to analyze what variables affect a measure of the model's efficiency in terms of both accuracy and computing expense. More imbalance control strategies, such as the algorithmic level method, should be investigated. Greater emphasis should be placed on evaluating network parameters; number of layers, different network initialization methods, activation and loss functions, network hyper-parameters, data augmentation techniques, kernel and stride size, the training method being used for the model, and regularization methods.

c) A decrease in the amount of work required for the computation. Another constraint is the reliance on GPU, which is necessary for some approaches as a result of the massive size of the dataset's picture archive prevents them from working as effectively without it. It would be beneficial to build effective learning algorithms to cut down on the amount of time spent on training as well as the amount of memory and processing resources required.

It is envisaged that this paper can assist researchers in understanding the crack identification techniques, which include classification, segmentation, detection, and hybrid approaches. It can also benefit the research community in addressing and realizing the full potential of computer vision and DL-based crack detection, research challenges and to identify future research opportunities.

ACKNOWLEDGEMENT AND FUNDING

The authors would like to express our sincere gratitude for the financial support received for this research work. The project benefited from the UTM HIR (Q.J13000.245108G87) and Ministry of Higher Education (MOHE) FRGS (R.J130000.78085F400) research grants. Additionally, we acknowledge the support provided by the Universiti Teknologi MARA Cawangan Johor (UiTM Johor), under Grant (600-TNCPI 5/3/DDN (01) (001/2021). Their support has played a crucial role in the successful completion of this research.

DECLARATION OF CONFLICTING INTERESTS

The authors declare no potential conflicts of interest with respect to the research and publication of this article.

REFERENCES

[1] K. Hrabová, B. Teplý and P. Hájek, Concrete, sustainability and limit states, *IOP Conference Series: Earth and Environmental Science*, 290(1), 2019, 012049.

[2] C. Gagg, Cement and concrete as an engineering material: an historic appraisal and case study analysis, *Engineering Failure Analysis*, 40, 2014, 114-140.

[3] N. Subramanian, *Design of Reinforced Concrete Structures*. New Delhi, India: Oxford University Press, 2013.

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

[5] M. H. Zeng, Z. M. Wu and Y. J. Wang, A stochastic model considering heterogeneity and crack propagation in concrete, *Construction and Building Materials*, 254, 2020, 119289.

[6] J. A. Guzmán-Torres, M. Z. Naser and F. J. Domínguez-Mota, Effective medium crack classification on laboratory concrete specimens via competitive machine learning, *Structures*, 37, 2022, 858-870.

[7] M. Zheng, Z. Lei and K. Zhang, Intelligent detection of building cracks based on deep learning, *Image and Vision Computing*, 103, 2020, 103987.

[8] V. P. Golding, Z. Gharineiat, H. S. Munawar and F. Ullah, Crack detection in concrete structures using deep learning, *Sustainability*, 14(13), 2022, 8117.

[9] J. Wang, Y. Ma, L. Zhang, R. X. Gao and D. Wu, Deep learning for smart manufacturing: Methods and applications, *Journal of Manufacturing Systems*, 48, 2018, 144-156.

[10] H. S. Munawar, A. W. A. Hammad, A. Haddad, C. Alberto, P. Soares and S. T. Waller, Image-based crack detection methods: A review, 6(8), 2021, 115.

[11] M. Murat Maras, Experimental behavior of injected geopolymer grout using styrene-butadiene latex for the repair and strengthening of masonry walls, *Advances in Structural Engineering*, 24(11), 2021, 2484-2499.

[12] H. H. Korkmaz, A. Yakut and A. Bayraktar, Analysis of a multi-story reinforced concrete residential building damaged under its self-weight, *Engineering Failure Analysis*, 98, 2019, 38-48.

[13] A. Nakipoglu, A. Unal and M. Kamanli, Damage analysis of Golyazi Apartments building failed under axial loads: A case study in Konya, Turkey, *Structures*, 30, 2021, 409-417.

[14] D. Liang, X. F. Zhou, S. Wang and C. J. Liu, Research on concrete cracks recognition based on dual convolutional neural network, *KSCE Journal of Civil Engineering*, 23(7), 2019, 3066-3074.

[15] I. H. Sarker, Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions, *SN Computer Science*, 2(6), 2021, 1-20.

[16] Y. Yu, M. Rashidi, B. Samali, M. Mohammadi, T. N. Nguyen and X. Zhou, Crack detection of concrete structures using deep convolutional neural networks optimized by enhanced chicken swarm algorithm, *Structural Health Monitoring*, 21(5), 2022, 2244-2263.

[17] L. Ali, F. Alnajjar, H. Al Jassmi, M. Gochoo, W. Khan and M. A. Serhani, Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures, *Sensors*, 21(5), 2021, 1-22.

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

[19] D. Kang, S. S. Benipal, D. L. Gopal and Y. J. Cha, Hybrid pixel-level concrete crack segmentation and quantification across complex backgrounds using deep learning, *Automation in Construction*, 118, 2020, 103291.

[20] B. Kim and S. Cho, Automated vision-based detection of cracks on concrete surfaces using a deep learning technique, *Sensors*, 18(10), 2018, 3452.

[21] W. R. L. da Silva and D. S. de Lucena, Concrete cracks detection based on deep learning image classification, *Proceedings*, 2(8), 2018, 489.

[22] S. Li and X. Zhao, Image-based concrete crack detection using convolutional, *Advances in Civil Engineering*, 2019, 6520620.

[23] F. Fang, L. Li, Y. Gu, H. Zhu and J. H. Lim, A novel hybrid approach for crack detection, *Pattern Recognition*, 107, 2020, 107474.

[24] S. Dorafshan, R. J. Thomas and M. Maguire, Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete, *Construction and Building Materials*, 186, 2018, 1031-1045.

[25] Z. Wang, G. Xu, Y. Ding, B. Wu and G. Lu, A vision-based active learning convolutional neural network model for concrete surface crack detection, *Advances in Structural Engineering*, 23(13), 2020, 2952-2964.

[26] Y. Liu, J. K. W. Yeoh and D. K. H. Chua, Deep learning–based enhancement of motion blurred UAV concrete crack images, *Journal of Computing in Civil Engineering*, 34(5), 2020.

[27] X. Wu and X. Liu, Building crack identification and total quality management method based on deep learning, *Pattern Recognition Letters*, 145, 2021, 225-231.

[28] H. Xu, X. Su, H. Xu and H. Li, Autonomous Bridge crack detection using deep convolutional neural networks, *Advances in Computer Science Research: Proceedings of the 3rd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA)*, 2019, 274-284.

[29] S. Li and L. Sun, Detectability of bridge-structural damage based on fiber-optic sensing through deep-convolutional neural networks, *Journal of Bridge Engineering*, 25(4), 2020, 04020012.

[30] E. McLaughlin, N. Charron and S. Narasimhan, Automated defect quantification in concrete bridges using robotics and deep learning, *Journal of Computing in Civil Engineering*, 34(5), 2020, 04020029.

[31] K. Jang, Y. K. An, B. Kim and S. Cho, Automated crack evaluation of a high-rise bridge pier using a ring-type climbing robot, *Computer-Aided Civil and Infrastructure Engineering*, 36(1), 2021, 14-29.

[32] S. E. Park, S. H. Eem and H. Jeon, Concrete crack detection and quantification using deep learning and structured light, *Construction and Building Materials*, 252, 2020, 119096.

[33] X. W. Ye, T. Jin and P. Y. Chen, Structural crack detection using deep learning–based fully convolutional networks, *Advances in Structural Engineering*, 22(16), 2019, 3412-3419.

[34] M. Alipour, D. K. Harris and G. R. Miller, Robust pixel-level crack detection using deep fully convolutional neural networks, *Journal of Computing in Civil Engineering*, 33(6), 2019, 04019040.

[35] T. S. Tran, V. P. Tran, H. J. Lee, J. M. Flores and V. P. Le, A two-step sequential automated crack detection and severity classification process for asphalt pavements, *International Journal of Pavement Engineering*, 23(6), 2022, 2019-2033.

[36] T. U. Ahmed, M. Shahadat Hossain, M. J. Alam, and K. Andersson, An integrated CNN-RNN framework to assess road crack, *22nd International Conference on Computer and Information Technology (ICCIT 2019)*, Dhaka, Bangladesh, 2019.

[37] R. Kalfarisi, Z. Y. Wu and K. Soh, Crack detection and segmentation using deep learning with 3D reality mesh model for quantitative assessment and integrated visualization, *Journal of Computing in Civil Engineering*, 34(3), 2020, 04020010.

[38] Q. Mei and M. Gül, A cost effective solution for pavement crack inspection using cameras and deep neural networks, *Construction and Building Materials*, 256, 2020, 119397.

[39] A. Zhang *et al.*, Deep learning–based fully automated pavement crack detection on 3D asphalt surfaces with an improved cracknet, *Journal of Computing in Civil Engineering*, 32(5), 2018, 04018041.

[40] K. Zhang, H. D. Cheng and B. Zhang, Unified Approach to pavement crack and sealed crack detection using preclassification based on transfer learning, *Journal of Computing in Civil Engineering*, 32(2), 2018, 04018001.

[41] K. Zhang, Y. Zhang and H. D. Cheng, Self-supervised structure learning for crack detection based on cycle-consistent generative adversarial networks, *Journal of Computing in Civil Engineering*, 34(3), 2020, 04020004.

[42] Q. Yang, W. Shi, J. Chen and W. Lin, Deep convolution neural network-based transfer learning method for civil infrastructure crack detection, *Automation in Construction*, 116, 2020, 103199.

[43] V. Mandal, L. Uong, and Y. Adu-Gyamfi, Automated road crack detection using deep convolutional neural networks, *Proceedings of 2018 IEEE International Conference on Big Data,* Seattle, USA, 2019, 5212–5215.

[44] W. Wang *et al.*, Automated crack severity level detection and classification for ballastless track slab using deep convolutional neural network, *Automation in Construction*, 124, 2021, 103484.

[45] B. Xia, J. Cao, X. Zhang and Y. Peng, Automatic concrete sleeper crack detection using a one-stage detector, *International Journal of Intelligent Robotics and Applications*, 4(3), 2020, 319-327.

[46] F. C. Chen and M. R. Jahanshahi, NB-CNN: Deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion, *IEEE Transactions on Industrial Electronics*, 65(5), 2018, 4392-4400.

[47] E. Protopapadakis, A. Voulodimos, A. Doulamis, N. Doulamis and T. Stathaki, Automatic crack detection for tunnel inspection using deep learning and heuristic image post-processing, *Applied Intelligence*, 49(7), 2019, 2793-2806.

[48] J. S. Lee, S. H. Hwang, I. Y. Choi and Y. Choi, Estimation of crack width based on shape-sensitive kernels and semantic segmentation, *Structural Control and Health Monitoring*, 27(4), 2020, 1-21.

[49] Y. Ren *et al.*, Image-based concrete crack detection in tunnels using deep fully convolutional networks, *Construction and Building Materials*, 234, 2020, 117367.

[50] Y. Li *et al.*, A deep residual neural network framework with transfer learning for concrete dams patch-level crack classification and weakly-supervised localization, *Measurement*, 188, 2021, 110641.

[51] M. Flah, A. R. Suleiman and M. L. Nehdi, Classification and quantification of cracks in concrete structures using deep learning image-based techniques, *Cement and Concrete Composites*, 114, 2020, 103781.

[52] N. Wang, X. Zhao, L. Wang and Z. Zou, Novel system for rapid investigation and damage detection in cultural heritage

conservation based on deep learning, *Journal of Infrastructure Systems*, 25(3), 2019, 04019020.
[53] M. E. Hatır, İ. İnce and M. Korkanç, Intelligent detection of deterioration in cultural stone heritage, *Journal of Building Engineering*, 44, , 2021, 102690. doi:.
[54] B. Kitchenham, Procedures for performing systematic reviews, *Empirical Software Engineering*, 33, 2004, 1-26.
[55] B. Kitchenham *et al.*, Systematic literature reviews in software engineering-A tertiary study, *Information and Software Technology*, 52(8), 2010, 792-805.
[56] P. C. Sierra-Correa and J. R. Cantera Kintz, Ecosystem-based adaptation for improving coastal planning for sea-level rise: A systematic review for mangrove coasts, *Marine Policy*, 51, 2015, 385-393.
[57] M. Tober, PubMed, ScienceDirect, Scopus or Google Scholar – Which is the best search engine for an effective literature research in laser medicine?, *Medical Laser Application*, 26, 2011, 139-144.
[58] M. G. M. Hamami and Z. H. Ismail, A systematic review on particle swarm optimization towards target search in the swarm robotics domain, *Archives of Computational Methods in Engineering*, 2020, 0123456789.
[59] M. Gusenbauer, Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases, *Scientometrics*, 118, 2019, 177-214.
[60] L. Ali, F. Alnajjar, W. Khan, M. A. Serhani and H. Al Jassmi, Bibliometric analysis and review of deep learning-based crack detection literature published between 2010 and 2022, *Buildings*, 12(4), 2022, 432.
[61] D. Moher, A. Liberati, J. Tetzlaff and D. G. Altman, Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement, *BMJ*, 339(7716), 2009, 332-336.
[62] Y. Hamishebahar, H. Guan, S. So and J. Jo, A comprehensive review of deep learning-based crack detection approaches, *Applied Sciences*, 12(3), 2022, 1374.
[63] C. V. Dung and L. D. Anh, Autonomous concrete crack detection using deep fully convolutional neural network, *Automation in Construction*, 99, 2019, 52-58.
[64] S. Li, X. Zhao and G. Zhou, Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network, *Computer-Aided Civil and Infrastructure Engineering*, 34(7), 2019, 616-634.
[65] H. Wei Huang, Q. Tong Li and D. Ming Zhang, Deep learning based image recognition for crack and leakage defects of metro shield tunnel, *Tunnelling and Underground Space Technology*, 77, 2018, 166-176.
[66] X. Liang, Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization, *Computer-Aided Civil and Infrastructure Engineering*, 34(5), 2019, 415-430.
[67] J. C. P. Cheng and M. Wang, Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques, *Automation in Construction*, 95, 2018, 155-171.
[68] S. Li and X. Zhao, Image-based concrete crack detection using convolutional neural network and exhaustive search technique, *Advances in Civil Engineering*, 2019, 6520620.
[69] B. Kim and S. Cho, Image-based concrete crack assessment using mask and region-based convolutional neural network, *Structural Control and Health Monitoring*, 26(8), 2019, 1-15.
[70] F. T. Ni, J. Zhang and Z. Q. Chen, Pixel-level crack delineation in images with convolutional feature fusion, *Structural Control and Health Monitoring*, 26(1), 2019, 1-18.
[71] X. Zhang, D. Rajan and B. Story, Concrete crack detection using context-aware deep semantic segmentation network, *Computer-Aided Civil and Infrastructure Engineering*, 34(11), 2019, 951-971.
[72] C. Zhang, C. Chen Chang and M. Jamshidi, Concrete bridge surface damage detection using a single-stage detector, *Computer-Aided Civil and Infrastructure Engineering*, 35(4), 2020, 389-409.
[73] K. Jang, N. Kim and Y. K. An, Deep learning–based autonomous concrete crack evaluation through hybrid image scanning, *Structural Health Monitoring*, 18(5–6), 2019, 1722-1737.
[74] J. K. Chow, Z. Su, J. Wu, P. S. Tan, X. Mao and Y. H. Wang, Anomaly detection of defects on concrete structures with the convolutional autoencoder, *Advanced Engineering Informatics*, 45, 2020, 101105.
[75] E. Protopapadakis, A. Voulodimos, A. Doulamis, N. Doulamis and T. Stathaki, Automatic crack detection for tunnel inspection using deep learning and heuristic image post-processing, *Applied Intelligence*, 49(7), 2019, 2793-2806.
[76] M. M. Manjurul Islam and J. M. Kim, Vision-based autonomous crack detection of concrete structures using a fully convolutional encoder–decoder network, *Sensors*, 19(19), 2019, 4251.
[77] Y. Pan, G. Zhang and L. Zhang, A spatial-channel hierarchical deep learning network for pixel-level automated crack detection, *Automation in Construction*, 119, 2020, 103357.
[78] J. Zhang, C. Lu, J. Wang, L. Wang and X. G. Yue, Concrete cracks detection based on FCN with dilated convolution, *Applied Sciences*, 9(13), 2019, 2686.
[79] C. Feng, Automatic pixel-level crack detection on dam surface using deep convolutional network, *Sensors*, 20(7), 2020, 2069.
[80] Q. Song *et al.*, Real-time tunnel crack analysis system via deep learning, *IEEE Access*, 7, 2019, 64186-64197.
[81] K. Fukushima, Neocognitron: A hierarchical neural network capable of visual pattern recognition, *Neural Networks*, 1(2), 1988, 119-130.
[82] M. D. Zeiler and R. Fergus, Visualizing and understanding convolutional networks, *Lecture Notes in Computer Science* (*European Conference on Computer Vision),* 8689, 2014, 818-833.
[83] Y. Le Cun et al., Handwritten digit recognition with a back-propagation network, *Advances in Neural Information Processing Systems*, 2, 1989, 396-404.
[84] H. C. Reis and K. Khoshelham, ReCRNet: a deep residual network for crack detection in historical buildings, *Arabian Journal of Geosciences*, 14(20), 2021, 2112.
[85] Y. Liu and J. K. W. Yeoh, Automated crack pattern recognition from images for condition assessment of concrete

structures, *Automation in Construction*, 128, 2021, 103765.
[86] J. K. Chow *et al.*, Artificial intelligence-empowered pipeline for image-based inspection of concrete structures, *Automation in Construction*, 120, 2020, 103372.
[87] G. Li, X. Ren, W. Qiao, B. Ma and Y. Li, Automatic bridge crack identification from concrete surface using ResNeXt with postprocessing, *Structural Control and Health Monitoring*, 27(11), 2020, e2620.
[88] R. Fu *et al.*, Enhanced intelligent identification of concrete cracks using multi-layered image preprocessing-aided convolutional neural networks, *Sensors*, 20(7), 2020, 2021.
[89] B. Wang, Y. Li, W. Zhao, Z. Zhang, Y. Zhang and Z. Wang, Effective crack damage detection using multilayer sparse feature representation and incremental extreme learning machine, *Applied Sciences*, 9(3), 2019, 614.
[90] M. Sharma, W. Anotaipaiboon and K. Chaiyasarn, Concrete crack detection using the integration of convolutional neural network and support vector machine, *Science and Technology Asia*, 23(2), 2018, 19-28.
[91] B. Kim *et al.*, Deep learning activation layer-based wall quality recognition using Conv2D ResNet exponential transfer learning model, *Mathematics*, 10(23), 2022, 4602.
[92] D. O. Brien, J. Andrew Osborne, E. Perez-Duenas, R. Cunningham and Z. Li, Automated crack classification for the CERN underground tunnel infrastructure using deep learning, *Tunnelling and Underground Space Technology*, 131, 2023, 104668.
[93] H. Zoubir, M. Rguig, M. El Aroussi, A. Chehri, R. Saadane and G. Jeon, Concrete bridge defects identification and localization based on classification deep convolutional neural networks and transfer learning, *Remote Sensing*, 14(19), 2022, 4882.
[94] T. T. Le, V. H. Nguyen and M. V. Le, Development of deep learning model for the recognition of cracks on concrete surfaces, *Applied Computational Intelligence and Soft Computing*, 2021, 8858545.
[95] X. Zheng, S. Zhang, X. Li, G. Li and X. Li, Lightweight bridge crack detection method based on segnet and bottleneck depth-separable convolution with residuals, *IEEE Access*, 9, 2021, 161649-161668.
[96] W. Qiao, H. Zhang, F. Zhu and Q. Wu, A crack identification method for concrete structures using improved U-Net convolutional neural networks, *Mathematical Problems in Engineering*, 2021, 6654996.
[97] S. K. Hou, Z. G. Ou, P. X. Qin, Y. L. Wang and Y. R. Liu, Image-based crack recognition of tunnel lining using residual U-Net convolutional neural network, *IOP Conference Series: Earth and Environmental Science*, 861(7), 2021, 072001.
[98] F. Çelik and M. König, A sigmoid-optimized encoder–decoder network for crack segmentation with copy-edit-paste transfer learning, *Computer-Aided Civil and Infrastructure Engineering*, 37(14), 2022, 1875-1890.
[99] W. Wang and C. Su, Automatic concrete crack segmentation model based on transformer, *Automation in Construction*, 139, 2022, 104275.
[100] L. Falaschetti, M. Beccerica, G. Biagetti, P. Crippa, M. Alessandrini and C. Turchetti, A lightweight CNN-based vision system for concrete crack detection on a low-power embedded microcontroller platform, *Procedia Computer Science*, 207, 2022, 3948-3956.
[101] T. Siriborvornratanakul, Downstream semantic segmentation model for low-level surface crack detection, *Advances in Multimedia*, 2022, 3712289.
[102] E. Lopez Droguett, J. Tapia, C. Yanez and R. Boroschek, Semantic segmentation model for crack images from concrete bridges for mobile devices, *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 236(4), 2020, 570-583.
[103] J. Jiang, G. Cao, D. F. Gomes and S. Luo, Vision-guided active tactile perception for crack detection and reconstruction, *29th Mediterranean Conference on Control and Automation (MED2021)*, Puglia, Italy, 2021, 930-936.
[104] C. Wan *et al.*, Crack detection for concrete bridges with imaged based deep learning, *Science Progress*, 105(4), 2022, 1-21.
[105] Y. Hong, S. J. Lee and S. B. Yoo, AugMoCrack: Augmented morphological attention network for weakly supervised crack detection, *Electronics Letters*, 58(17), 2022, 651-653.
[106] Q. Zhou, Z. Qu, and F. Rong Ju, A multi-scale learning method with dilated convolutional network for concrete surface cracks detection, *IET Image Processing*, 16(5), 2022, 1389-1402.
[107] X. -W. Ye, J. Tao and P. -Y. Chen, Structural crack detection using deep learning–based fully convolutional networks, *Advances in Structural Engineering*, 22, 2019, 3412-3419.
[108] S. Li and X. Zhao, Automatic crack detection and measurement of concrete structure using convolutional encoder-decoder network, *IEEE Access*, 8, 2020, 134602-134618.
[109] L. Li *et al.*, Crack detection method of sleeper based on cascade convolutional neural network, *Journal of Advanced Transportation*, 2022, 7851562.
[110] C. Su and W. Wang, Concrete cracks detection using convolutional neural network based on transfer learning, *Mathematical Problems in Engineering*, 2020, 7240129.
[111] L. Deng, H. H. Chu, P. Shi, W. Wang and X. Kong, Region-based CNN method with deformable modules for visually classifying concrete cracks, *Applied Sciences*, 10(7), 2020, 2528.
[112] B. Kumar and S. Ghosh, Detection of concrete cracks using dual-channel deep convolutional network, *11th International Conference on Computing, Communication and Networking Technologies (ICCCNT2020)*, Kharagpur, India, 2020.
[113] H. Bae, K. Jang and Y. K. An, Deep super resolution crack network (SrcNet) for improving computer vision–based automated crack detectability in in situ bridges, *Structural Health Monitoring*, 20(4), 2021, 1428-1442.
[114] C. Li, P. Xu, L. Niu, Y. Chen, L. Sheng and M. Liu, Tunnel crack detection using coarse-to-fine region localization and edge detection, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(5), 2019, e1308.

[115] P. Kumar, S. Batchu, N. Swamy S. and S. R. Kota, Real-time concrete damage detection using deep learning for high rise structures, *IEEE Access*, 9, 2021, 112312-112331.

[116] K. W. Kee, K. H. Lim, C. H. Lim, W. L. Lim and H. E. Yap, Cracks identification using mask region-based denoised deformable convolutional network, *Multimedia Tools and Applications*, 82(3), 2023, 4387-4404.

[117] D. P. Yadav *et al.*, A novel multi-scale feature fusion-based 3SCNet for building crack detection, *Sustainability*, 14(23), 2022, 16179.

[118] J. J. Kim, A. R. Kim and S. W. Lee, Artificial neural network-based automated crack detection and analysis for the inspection of concrete structures, *Applied Sciences*, 10(22), 2020, 8105.

[119] Q. Yang and X. Ji, Automatic pixel-level crack detection for civil infrastructure using Unet++ and deep transfer learning, *IEEE Sensors Journal*, 21(17), 2021, 19165-19175.

[120] P. Guo, W. Meng and Y. Bao, Automatic identification and quantification of dense microcracks in high-performance fiber-reinforced cementitious composites through deep learning-based computer vision, *Cement and Concrete Research*, 148, 2021, 106532.

[121] Y. Li *et al.*, A deep residual neural network framework with transfer learning for concrete dams patch-level crack classification and weakly-supervised localization, *Measurement*, 188, 2022, 110641.

[122] G. Chen, S. Teng, M. Lin, X. Yang and X. Sun, Crack detection based on generative adversarial networks and deep learning, *KSCE Journal of Civil Engineering*, 26(4), 2022, 1803-1816.

[123] C. Yang, J. Chen, Z. Li and Y. Huang, Structural crack detection and recognition based on deep learning, *Applied Sciences*, 11(6), 2021, 2868.

[124] H. J. Woo, D. M. Seo, M. S. Kim, M. S. Park, W. H. Hong and S. C. Baek, Localization of cracks in concrete structures using an unmanned aerial vehicle, *Sensors*, 22(17), 2022, 6711.

[125] M. A. Mohammed, Z. Han and Y. Li, Exploring the detection accuracy of concrete cracks using various CNN models, *Advances in Materials Science and Engineering*, 2021, 9923704.

[126] J. Lee, H. S. Kim, N. Kim, E. M. Ryu and J. W. Kang, Learning to detect cracks on damaged concrete surfaces using two-branched convolutional neural network, *Sensors*, 19(21), 2019, 4796.

[127] H. Kim, S. H. Sim and B. F. Spencer, Automated concrete crack evaluation using stereo vision with two different focal lengths, *Automation in Construction*, 135, 2022, 104136.

[128] M. Padsumbiya, V. Brahmbhatt and S. P. Thakkar, Automatic crack detection using convolutional neural network, *Journal of Soft Computing in Civil Engineering*, 6(3), 2022, 1-17.

[129] A. Shrivastava, A. Gupta and R. Girshick, Training region-based object detectors with online hard example mining, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Las Vegas, USA, 2016, 761-769.

[130] J. Guo, Q. Wang and Y. Li, Evaluation-oriented façade defects detection using rule-based deep learning method, *Automation in Construction*, 131, 2021, 103910.

[131] D. Andrushia A, N. Anand, E. Lubloy and G. Prince Arulraj, Deep learning based thermal crack detection on structural concrete exposed to elevated temperature, *Advances in Structural Engineering*, 24(9), 2021, 1896-1909.

[132] H. Huang, S. Zhao, D. Zhang and J. Chen, Deep learning-based instance segmentation of cracks from shield tunnel lining images, *Structure and Infrastructure Engineering*, 18(2), 2022, 183-196.

[133] J. Yoon, H. Shin, M. Song, H. Gil and S. Lee, A crack width measurement method of UAV Images using high-resolution algorithms, *Sustainability*, 15(1), 2023, 0478.

[134] X. W. Ye, Z. X. Li and T. Jin, Smartphone-based structural crack detection using pruned fully convolutional networks and edge computing, *Smart Structures and Systems*, 29(1), 2022, 141-151.

[135] W. Li, G. Wang, L. Fidon, S. Ourselin, M. J. Cardoso and T. Vercauteren, On the compactness, efficiency, and representation of 3D convolutional networks: Brain parcellation as a pretext task, *Lecture Notes in Computer Science*, 10265, 2017, 348-360.

[136] S. Yang, W. Xiao, M. Zhang, S. Guo, J. Zhao and F. Shen, Image data augmentation for deep learning: A survey, *ArXiv*, 2204(08610), 2022.

[137] X. Gao, C. Huang, S. Teng and G. Chen, A deep-convolutional-neural-network-based semi-supervised learning method for anomaly crack detection, *Applied Sciences*, 12(18), 2022, 9244.

[138] J. Deng, Y. Lu and V. C. S. Lee, Imaging-based crack detection on concrete surfaces using You Only Look Once network, *Structural Health Monitoring*, 20(2), 2021, 484-499.

[139] C. Li, W. Pan, R. K. L. Su and P. C. Yuen, Multiple structural defect detection for reinforced concrete buildings using YOLOv5s, *Transactions Hong Kong Institution of Engineers*, 29(2), 2022, 141-150.

[140] L. Yu, S. He, X. Liu, S. Jiang and S. Xiang, Intelligent Crack detection and quantification in the concrete bridge: A deep learning-assisted image processing approach, *Advances in Civil Engineering*, 2022, 1813821.

[141] J. Zhu, C. Zhang, H. Qi and Z. Lu, Vision-based defects detection for bridges using transfer learning and convolutional neural networks, *Structure and Infrastructure Engineering*, 16(7), 2020, 1037-1049.

[142] S. Teng, Z. Liu, G. Chen and L. Cheng, Concrete crack detection based on well-known feature extractor model and the YOLO_v2 network, *Applied Sciences*, 11(2), 2021, 0813.

[143] M. A. Mohammed, Z. Han, Y. Li, Z. Al-Huda, C. Li and W. Wang, End-to-end semi-supervised deep learning model for surface crack detection of infrastructures, *Frontiers in Materials*, 9, 2022, 1058407.

[144] P. Shokri, M. Shahbazi, D. Lichti and J. Nielsen, Vision-based approaches for quantifying cracks in concrete structures, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 43(B2), 2020, 1167-1174.

[145] E. A. Felix and S. P. Lee, Systematic literature review of preprocessing techniques for imbalanced data, *IET Software*, 13(6), 2019, 479-496.

[146] K. Lee, S. Lee and H. Y. Kim, Bounding-box object augmentation with random transformations for automated defect detection in residential building façades, *Automation in Construction*, 135, 2022, 104138.

[147] P. Kruachottikul, N. Cooharojananone, G. Phanomchoeng, T. Chavarnakul, K. Kovitanggoon and D. Trakulwaranont, Deep learning-based visual defect-inspection system for reinforced concrete bridge substructure: a case of Thailand's department of highways, *Journal of Civil Structural Health Monitoring*, 11(4), 2021, 949-965.

[148] X. Xu and H. Yang, Vision measurement of tunnel structures with robust modelling and deep learning algorithms, *Sensors*, 20(17), 2020, 4945.

[149] J. Ji, L. Wu, Z. Chen, J. Yu, P. Lin and S. Cheng, Automated pixel-level surface crack detection using U-net, *Lecture Notes in Computer Science*, 11248, 2018, 69-78.

[150] S. Arabi, A. Haghighat and A. Sharma, A deep-learning-based computer vision solution for construction vehicle detection, *Computer-Aided Civil and Construction Engineering*, 35(7), 2020, 753-767.

[151] R. Miotto, F. Wang, S. Wang, X. Jiang and J. T. Dudley, Deep learning for healthcare: review, opportunities and challenges, *Briefings in Bioinformatics*, 19(6), 2018, 1236-1246.

[152] Y. Xue and Y. Li, A fast detection method via region-based fully convolutional neural networks for shield tunnel lining defects, *Computer-Aided Civil and Construction Engineering*, 33(8), 2018, 638-654.

[153] S. Wei, P. W. Tien, J. K. Calautit, Y. Wu and R. Boukhanouf, Vision-based detection and prediction of equipment heat gains in commercial office buildings using a deep learning method, *Applied Energy*, 277, 2020, 115506.

[154] D. Roberts and M. Golparvar-fard, Automation in Construction End-to-end vision-based detection , tracking and activity analysis of earthmoving equipment fi lmed at ground level, *Automation in Construction*, 105, 2019, 102811.

[155] P. Kumar, A. Sharma and S. R. Kota, Automatic multiclass instance segmentation of concrete damage using deep learning model, *IEEE Access*, 9, 2021, 90330-90345.

[156] M. S. Saeed, Unmanned aerial vehicle for automatic detection of concrete crack using deep learning, *International Conference on Robotics, Electrical and Signal Processing Techniques*, Dhaka, Bangladesh, 2021, 624-628.

[157] R. Zulkifli, A. R. Husain, I. S. Ibrahim, M. Mazni and N. H. A. M. Fauzan, Analysis of the hybrid adhesion mechanism of the wall climbing robot, *Lecture Notes in Electrical Engineering*, 921, 2022, 155-169.

[158] C. V. Dung and L. D. Anh, Autonomous concrete crack detection using deep fully convolutional neural network, *Automation in Construction*, 99, 2019, 52-58.

[159] M. M. Islam, M. B. Hossain, M. N. Akhtar, M. A. Moni and K. F. Hasan, CNN based on transfer learning models using data augmentation and transformation for detection of concrete crack, *Algorithms*, 15(8), 2022, 0287.

[160] H. Yu, L. Zhu, D. Li, Q. Wang, X. Liu and C. Shen, Comparative study on concrete crack detection of tunnel based on different deep learning algorithms, *Frontiers in Earth Science*, 9, 2022, 817785.

[161] T. Yamane and P. J. Chun, Crack detection from a concrete surface image based on semantic segmentation using deep learning, *Journal of Advanced Concrete Technology*, 18(9), 2020, 493-504.

[162] B. Kim and S. Cho, Automated multiple concrete damage detection using instance segmentation deep learning model, *Applied Sciences*, 10(22), 2020, 8008.

[163] L. Alzubaidi et al., Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, *Journal of Big Data*, 8, 2021, 53.

[164] S. Jiang and J. Zhang, Real-time crack assessment using deep neural networks with wall-climbing unmanned aerial system, *Computer-Aided Civil and Infrastructure Engineering*, 35(6), 2020, 549-564.