Electronic Journal of Structural Engineering

eJSE INTERNATIONAL

Original Article

Cite this: DOI:10.56748/ejse.24781

Received Date: 6 March 2025
Accepted Date:19 October 2025

1443-9255
https://ejsei.com/ejse

Exploring Deep Learning Architectures for Crack Detection: A Review of Progress and Emerging Trends

Kanchan Dhapekar [a*], Divya Prakash [a]

[a] Department of Civil Engineering, Poornima University, Jaipur, Rajasthan, India-303905
*Corresponding author: kanchan1983@gmail.com

Abstract

Deep learning architectures, particularly those driven by computer vision and artificial intelligence advancements, have revolutionized crack detection in substructures. It is therefore crucial to categorize these papers more clearly to understand their contributions and advancements, given the increasing amount of research in this field. The comprehensive review of deep learning-based fracture recognition studies in this article highlights the significant progress made in this area. The research is grouped according to their computer vision methods and then further separated to make it easier to investigate options that use related strategies to address the crack detection issue. By examining the various architectures and approaches, this review also identifies the key experiments and restrictions faced in the field. Furthermore, it proposes critical future directions for research, drawn from the insights of the reviewed studies and emerging trends in related fields, which could help address existing gaps and drive further innovations in crack detection systems.

Keywords

Deep learning, Crack detection, Object recognition, Image classification, Semantic segmentation

1. Introduction

In the realm of artificial intelligence (AI), deep learning (DL) has become one of the most revolutionary technologies, transforming a wide range of industries, including healthcare, transportation, robotics, and infrastructure management [Van Hoang, 2024]. By finding patterns in enormous volumes of data. DL, which is based on the concept of artificial neural networks, seeks to mimic how the human brain learns and adapts. When it comes to autonomously generating feature representations from raw data, DL outperforms conventional machine learning techniques, doing away with the necessity for laborious human feature engineering. This skill has resulted in advances in challenging fields including autonomous driving, natural language processing, and picture identification [Gupta, et al., 2021]. The development of DL has been fueled by three main factors: the availability of massive datasets, advances in computational power, and improvements in neural network architectures. Datasets such as ImageNet and COCO have enabled researchers to train large-scale models, while advancements in GPU technology have made it possible to process data and train models at unprecedented speeds [Younesi, et al., 2024]. Additionally, the capabilities of DL have been extended by advancements in neural constructions, such as transformers, generative adversarial networks (GANs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs) [Sengupta, et al., 2020]. These designs have not only improved the accuracy and efficiency of many AI systems, but they have also opened up new avenues for addressing difficult problems that were previously believed to be intractable. DL has shown unmatched effectiveness in the analysis and interpretation of visual data, making it one of the most important applications in the field of computer vision. The use of DL models has led to notable advancements in tasks including object identification, picture categorization, and segmentation. The ability of computers to comprehend visual information with an accuracy that equals or sometimes exceeds human performance has been made possible in large part by CNNs. This has profound implications for industries like healthcare, where computer vision is used for medical imaging diagnostics, and transportation, where it powers autonomous vehicle navigation [Zhang, 2023]. Computer vision and deep learning have also been very helpful in the civil infrastructure sector, particularly in the field of fracture identification for SHM.

A crucial component of infrastructure maintenance is crack identification since undiscovered fractures can result in structural collapses, which pose serious threats to public safety and financial stability [Kaveh and Alhajj, 2024]. Manual inspections, which are labor-intensive, time-consuming, and prone to human errors, are frequently used in traditional crack detection techniques. Furthermore, in large-scale or inaccessible constructions like bridges, tunnels, and dams, manual inspections are not always possible [Negi, et al., 2024]. Deep learning's introduction of automated systems that can accurately and efficiently identify and categorize cracks has completely changed the field of crack detection. To find patterns suggestive of structural degradation, these systems employ deep learning models that have been skilled in big-picture datasets. These representations are essential tools for infrastructure management because of their capacity to learn intricate aspects like the texture and form of fissures. CNNs have shown to be very effective among the several deep learning techniques for crack detection because of their skill to process and analyze picture data [Kulambayev, et al., 2024]. CNN-based models can be used for tasks ranging from binary crack detection (i.e., crack or no crack) to more advanced segmentation tasks that involve identifying the precise location and dimensions of cracks. Semantic segmentation methods, which involve pixel-level classification of an image, have shown remarkable accuracy in detecting cracks with intricate details. These approaches are highly effective in distinguishing cracks from background noise, such as shadows, dirt, and other visual artifacts that can complicate detection. The performance of segmentation models has been further improved by sophisticated designs like U-Net and Mask R-CNN, which allow them to handle a variety of intricate crack patterns [Choi, et al., 2024]. The field of crack detection still faces some difficulties despite the developments. The variety of crack types, which can differ in size, shape, and placement, is one major problem. While many studies have focused on common crack types such as longitudinal and transverse cracks, less attention has been given to irregular or rare crack types like meandering or crescent-shaped cracks. The absence of defined datasets for deep learning model evaluation and training presents another difficulty. Current datasets are frequently undiversified and do not accurately depict real-world circumstances, such as fluctuating illumination, weather, and noise levels. This restricts the models' capacity to be applied in real-world situations and their generalizability. Furthermore, elements like the superiority of the preparation data and the intricacy of the crack construction might affect how deep learning model's function. To overcome these obstacles, more reliable models and datasets that can handle the complexity and unpredictability of real-world crack detection jobs must be created.

The transformation brought about by DL in crack detection is not limited to improved accuracy and efficiency [Hamishebahar, et al., 2022]. It also includes how cost-effective and scalable inspection procedures are. Large data sets may be processed by automated crack detection systems in a segment of the time wanted for manual inspections, which lowers labor costs and for more frequent and thorough evaluations. This is especially advantageous for large-scale infrastructure projects, where durability and safety depend on prompt diagnosis and maintenance. Additionally, the combination of DL and other technologies, like drones and IoT sensors, has made it possible to monitor and inspect infrastructure in real time, opening the door for more intelligent and proactive maintenance plans. There is an increasing number of published research in the field of crack detection that use DL architecture. To assist interested researchers in better comprehending and comparing the suggested methodologies, a more thorough assessment and classification of these works at deeper levels and from various perspectives is therefore required. A thorough classification of DL-based crack detection research is obtainable in this publication, which covers more papers and does deeper levels of analysis. The goal is to make it easier to explore novel strategies and identify comparable ones. Recent crack detection investigations are critically reviewed in this publication, with their primary contributions to the field being extracted and discussed. It also considers the influence of computer vision-inspired approaches and techniques in categorizing crack detection studies. The study offers guidance on how to choose cutting-edge fracture detection techniques from the examined literature and makes important recommendations for future research.

1.1 Contribution of the work

- This study systematically reviews deep learning-based crack detection techniques, categorizing them by computer vision methodologies and architecture.
- It highlights advancements such as semantic segmentation and hybrid models, emphasizing their impact on accuracy and localization in crack detection.
- The paper identifies research gaps, such as detecting complex crack types to guide future innovations.
- By addressing challenges in real-world applications and proposing efficient, scalable solutions, this work fosters the growth of vigorous crack detection systems.

This is the structure of the article. Deep learning is presented in Section 2. The specifics of feature extractions and classifications are reviewed in depth in Section 3. DL Application Software Structures are provided in Section 4. Section 6 presents the study's conclusion, while Section 5 offers future directions.

2. Deep Learning (Dl): An Overview

A subset of artificial intelligence (AI) is machine learning (ML). Machine learning techniques aim to create trainable algorithms that can learn from existing measured or simulated response data to produce predictions for the future [Sun, et al., 2021]. DL may be viewed as a feature-learning method and a subset of ML as ML-based SHM models are primarily built with the capacity to train independently. ML-based SHM models come in three varieties: reinforcement learning, unsupervised learning, and supervised learning [Malekloo, et al., 2022]. In supervised learning, an ML model may be trained using training data that has labeled target values. The model is called a classifier if its outputs are discrete or categorical variables; if not, it is called a regression model. Cluster datasets are used in unsupervised learning techniques, which do not require specific training regimens. Spectral clustering, k-means, partitioned clustering, hierarchical clustering, and other methods are examples of clustering techniques. DL has been a revolutionary force in structural health monitoring (SHM) in recent years, tackling the escalating difficulties of contemporary civil infrastructure [Plevris and Papazafeiropoulos, 2024]. With the proliferation of sensors generating vast and complex datasets, DL provides a robust solution by automatically extracting valuable insights. The possibilities of SHM have been further expanded by its integration with computer vision, which makes it possible to analyze data from four-dimensional (4D) datasets, such as RGB movies, for sophisticated damage assessment, to one-dimensional (1D) signals, like vibrations and strain measurements. These developments, together with hardware enhancements and the availability of user-friendly frameworks, have made DL an essential tool for improving structural safety and democratizing its usage in SHM [Cha, et al., 2024]. DL removes these performance limitations, providing a more effective and scalable method than standard machine learning (ML) approaches, which frequently depend on manual feature selection and engineering.

DL-based SHM techniques stack layers of deep neural networks (DNNs) with nonlinear mappings to enable fully automated feature extraction and hierarchical representation directly from raw input data [Ullah, 2023]. DNN layers learn sophisticated input response data, optimizing regression, pattern classification, and feature extraction parameters simultaneously. This integrated approach addresses challenges with minimal prior knowledge of domain-specific structures, solidifying DL-based SHM as a game-changing technology. Figure 1 shows the mind map of ML algorithms.

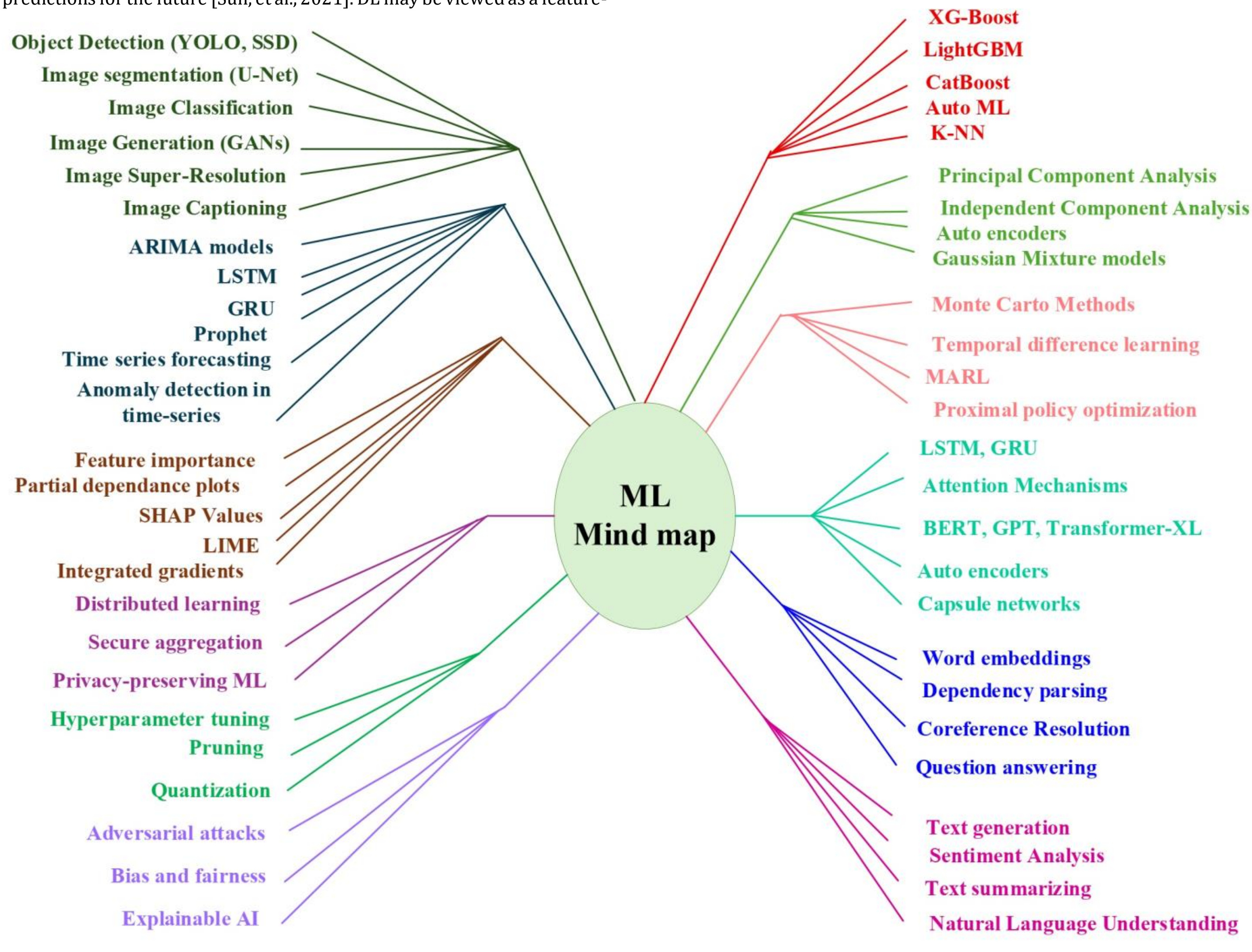

Fig. 1 ML Algorithm mind map

2.1 Various learning modes

Deep learning employs various learning modes, each tailored to specific problem domains and data availability [Sarker, 2021]. A fundamental method in deep learning is supervised learning, in which models are trained using datasets that comprise inputs and the labeled outputs that correspond to them. The model learns to convert the input data into the right output by decreasing the difference between its predictions and the actual labels during training. This process allows the model to identify patterns and relationships in the data by providing it with labeled examples. Over time, the model refines its parameters to generalize and make accurate predictions on unseen data, making it highly effective for structured and well-defined problems. In contrast, unsupervised learning is a deep learning approach where the model works with datasets that lack labeled outputs. Instead of learning from explicit input-output pairs, the model identifies underlying structures, relationships, or distributions within the data. This involves discovering patterns, grouping similar data points through clustering, or simplifying complex datasets by reducing their dimensions while preserving essential information. It is important for exploration data analysis and the extraction of meaningful representations since the objective is to obtain insights from raw data without depending on predetermined labels. Semi-supervised learning bridges the gap between supervised and unsupervised learning by combining a small amount of labeled data with a larger collection of unlabeled data [Shen, et al., 2024]. While the unlabeled data allows the model to simplify and improve its comprehension, the labeled data gives it an early direction and helps it discover important patterns and correlations. This method is especially helpful when classifying data is costly or time-consuming since it lessens the reliance on large, labeled datasets. Semi-supervised learning improves learning efficiency and model performance by balancing these two kinds of input. An agent that uses reinforcement learning, a novel learning model, interacts with its surroundings to learn how to make the best choices. The agent explores various actions and observes the outcomes, receiving rewards for favorable actions and penalties for unfavorable ones. Over time, it refines its strategy, aiming to maximize cumulative rewards by balancing exploring new possibilities and exploiting known successful actions. This iterative process enables the agent to develop adaptive and intelligent decision-making policies for complex and dynamic environments. Self-supervised learning is an emerging paradigm in deep learning that utilizes the inherent structure or properties of data to create pseudo-labels for training [Taherdoost, 2024]. Instead of relying on manually labeled datasets, the model generates its supervision by predicting certain aspects of the data, such as missing parts or transformations. This method enables it to extract valuable patterns and representations from enormous volumes of unlabeled data. Self-supervised learning works particularly well in sectors with limited labeled data because it lessens the need for labeled datasets, allowing deep learning to address a wide range of challenging problems. Figure 2 represents the various learning modes used in deep learning.

3. Automatic Feature Extraction

A structural health monitoring system must be developed using four processes: (i) operational valuation; (ii) data collecting; (iii) feature extraction; and (iv) statistical modeling. These steps are based on standard machine learning methods. Operational evaluation seeks to address critical questions regarding the justification of developing such systems from both life-safety and economic perspectives [Tibaduiza Burgos, et al., 2020]. This involves identifying the types of damage that are most concerning and determining, in scenarios involving multiple types of damage, which ones pose the greatest risk. At this stage, financial constraints play a significant role, particularly when selecting appropriate data acquisition hardware. In the feature extraction procedure that follows, the primary objective is to extract meaningful, relevant, and non-redundant values from the raw data collected in the previous step. This stage guarantees that the retrieved characteristics efficiently assist the training and generalization stages that follow. The final phase, referred to as the statistical modeling step, comprises statistical machine learning algorithm evaluation and training. Because of the differences between traditional machine learning methods and contemporary deep learning strategies, the structure described may be modified to fit certain objectives, such as DL-based crack detection, increasing its usefulness and efficacy in more complex use cases.

The term "deep architectures" refers to the latest generation of neural networks, which employ several deep layers to progressively extract high-level characteristics from unprocessed inputs [López-Monroy, et al., 2022] [Safaei, et al., 2022]. DL models have been used in many fields, including information retrieval, computer vision, voice and audio processing, and language processing. The most important element of deep architecture, as described above, is that the design itself handles feature extraction, doing away with the requirement for the handcrafted feature engineering stage. In some fields, including video analysis, medical imaging, and fracture diagnosis, the human extraction of handmade features has been supplanted by the autonomous feature extraction of deep learning architecture.

3.1 Image Classification

In deep crack detection algorithms, the trained architecture classifies new input images based on IC parameters, determining if a crack exists, and then stacking positive patches to create a crack essential. Two components make up the overall architecture used to carry out IC. Safaei et al.'s automatic crack assessment method uses a tile-based image processing method to detect and classify cracks in 2-D and 3-D pavement images, using a localized thresholding technique to extract important features. Several studies used image-based techniques to improve masonry walls' brick segmentation and crack detection [Aravind, et al., 2021] [Loverdos and Sarhosis, 2022].

Deep architecture is known to need a significant quantity of labeled data for training, which makes its use in many contexts difficult. Transfer Learning (TL) is a method that uses pre-trained networks on large-scale annotated picture data sets to solve classification problems. It can be used when the dataset is small and the pre-trained network has been trained on a larger one, reducing training time [Sarker, et al., 2021]. Instead of starting with weights that are chosen at random, the coefficients are taken from the base model. The network depth and the quantity of training parameters are directly correlated. As a result, training deep networks requires a large amount of time and data [Zhong and Ban, 2022]. A CNN-based transfer learning technique was presented by Zhong et al. [Tsalera, et al., 2021]. It was based on a CNN that had already been trained using the Image Net. Tsalera et al. [Chen, et al., 2021] looked at pre-trained CNNs' performance for sound classification as well as retraining possibilities. Similarly, a new pre-trained model called an image processing transformer (IPT) was developed by Chen et al. [Golding, et al., 2022]. Several papers with various research goals may be found in the IC category. Using the convolutional neural network (CNN) methodology, Golding et al. [Guo, et al., 2024] presented an autonomous crack-detecting system based on DL. The Crack Similarity Learning Network [Yang, et al., 2021] to ascertain the semantic similarity of crack image regions.

There are several articles in the IC category with different research objectives. Three neural networks were used by Yang et al. [Zhang, et al., 2021] to recognize and categorize crack pictures. [Gooda, et al., 2023] provided a vision-based fracture recognition technique for concrete bridge decks using an integrated 1D-CNN-LSTM system that was trained on hundreds of images of concrete bridge decks with and without cracks [Zhang, 2024]. In the crack detection sector, localizing cracks in the images is essential in addition to finding fracture pictures or image patches utilizing an IC setup. The IC setting is widely used for crack identification, but its coarse appearance and lack of precise localization of fractures make it unsuitable for deep fracture detection. Table 1 provides a summary of deep fracture detection techniques based on this setting.

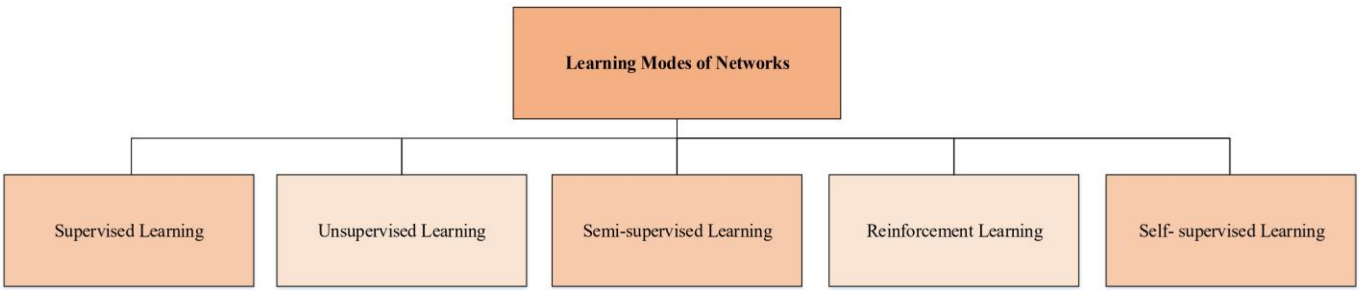

Fig. 2 Various learning modes

Table 1. An overview of IC-based deep crack detection techniques

Ref	Novelty/Novelties	Core Architecture	Challenges	Result
[Safaei, et al., 2022]	Cracks in 2D/3D pavement photos may be automatically detected and classified using crack assessment.	Tile-based image processing with localized thresholding.	- Handling diverse crack types and shapes in 2D/3D images is difficult. - Variability in image quality due to lighting and surface texture.	Achieved accurate crack classification in controlled datasets but faced limitations in detecting complex crack shapes and under variable conditions.
[Tsalera, et al., 2021]	Crack detection using a CNN-based transfer learning technique	CNN that has already been trained using ImageNet.	- TL models may not generalize well to specific crack types or real-world scenarios. - Dependency on large, labeled datasets for fine-tuning the model.	Improved detection accuracy by leveraging pre-trained CNNs, with performance dependent on the quality of labeled data.
[Chen, et al., 2021]	Retraining options and performance analysis of pre-trained CNNs for sound classification.	Retraining pre-trained CNN models.	- Retraining pre-trained CNNs may lead to overfitting if not enough domain-specific data is available. - Lack of robustness when applied to new or unseen data.	Demonstrated flexibility in retraining CNNs for different datasets, but overfitting and lack of robustness were key limitations.
[Golding, et al., 2022]	Development of a new image processing transformer (IPT) model.	Image Processing Transformer (IPT).	- The new IPT model may require high computational resources for training and inference. - Limited generalization to diverse crack types and surface conditions.	Showcased high precision in crack detection on benchmark datasets but struggled with high computational costs and generalization issues.
[Guo, et al., 2024]	DL-based autonomous crack detection method.	CNN	- Performance issues in detecting thin or intricate cracks. - Difficulty in distinguishing cracks from noise or irrelevant image features.	Achieved reliable detection for common cracks but faced challenges with thin cracks and noisy images.
[Yang, et al., 2021]	Crack Similarity Learning Network (CrackSL-Net) for semantic similarity of crack regions.	CrackSL-Net.	- Difficulty in differentiating between similar crack types or regions with similar visual features. - Requires high-quality and well-labeled data for training.	Improved semantic similarity detection of crack regions but was sensitive to the quality and diversity of training data.
[Zhang, et al., 2021]	Recognized and classified crack images using AlexNet, VGGNet13, and ResNet18 neural networks.	AlexNet, VGGNet13, ResNet18.	- Differences in performance across various neural networks like VGGNet13, AlexNet, and ResNet18, depending on dataset and crack type. - Limited robustness to noise or real-world variability in images.	Demonstrated varying performance with ResNet18 outperforming others on accuracy and AlexNet showing faster training times.
[Gooda, et al., 2023]	Concrete bridge deck fracture detection based on vision with integrated 1D-CNN-LSTM	1D-CNN-LSTM algorithm.	- Challenges in integrating 1D-CNN and LSTM models for crack detection with real-time data. - Difficulty in scaling the model to handle large datasets or complex infrastructure projects.	Successfully identified cracks in concrete bridge decks, with better temporal analysis due to LSTM, but scalability remained a challenge

3.2 Object Recognition

OR is a set of related tasks that use bounding boxes to find objects of interest in pictures. When compared to the sliding window approach for the same goal, its use in computer vision has increased the accuracy of object recognition and localization in pictures [Rodriguez-Conde, et al., 2022]. In computer vision, OR is performed by several architectural families, including single shot detectors (SSD) [Yan and Zhang, 2021], region-based CNN [Xu, et al., 2022], and you only look once (YOLO) [Karimi, et al., 2024]. The R-CNN family members are commonly used for crack detection. SSD and YOLO designs are used in the main design but have been used multiple times as competitors. The research will examine these studies based on the primary architecture and evaluate the innovations. Figure 3 shows the output of a DL-based crack detection approach based on the OR setting.

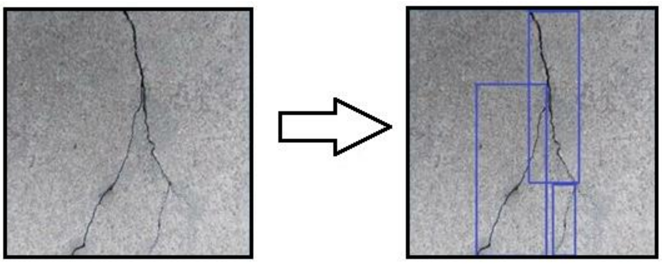

Fig. 3 Typical results of a fracture detecting method based on OR setting and deep learning

R-CNN Family

Crack detection has been a popular application of region-based convolutional neural networks (R-CNN), which approach the problem as a region-based object detection job. R-CNN separates the input picture into many proposed regions and determines if there are cracks in each of them. By using a DL model to analyze certain locations, this approach guarantees precise localization and categorization. This family includes R-CNN, Fast R-CNN, and Faster R-CNN; the most popular of them in the crack detection area is Faster R-CNN. Only once has the R-CNN architecture—the first affiliate of the R-CNN family—been used to break detection [Arman, et al., 2020]. Faster R-CNN and R-CNN masks for road cracks were studied by Xu et al. [Xu, et al., 2022]. Complex data sets were taken into consideration to test the Faster R-CNN architecture's performance in more realistic scenarios, making the data set more difficult in the OR environment. Marin et al., [Marin, et al., 2021] adopted the architecture of Faster R-CNN object detectors to provide crack detection in images. Figure 4 represents the network structure of Faster R-CNN. Figure 4 demonstrates the network structure of Faster R-CNN.

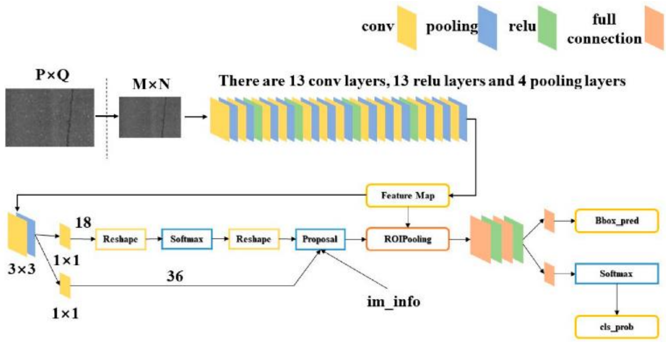

Fig. 4 The network structure of Faster R-CNN

SSD

The Single-Shot multi-box Detector (SSD) gets its name from the fact that it does all of the computing in a single network and does not require an additional module to identify potential areas of objects in an image. This enables integrating SSD into frameworks that require a detection component that is simple and quick to train. Yan et al., [Yan and Zhang, 2021] created a special deformable SSD network by adding a deformable convolution to the VGG16 backbone feature extraction network. Feng et al. [Feng, et al., 2020] proposed a deep CNN fusion model-based method for pavement crack detection that combines the advantages of multitarget SSD and U-Net convolutional neural network models.

YOLO Family

YOLO (You Only Look Once) is a DL-based object detection algorithm that has been effectively adapted for crack detection. It treats crack detection as an object detection task, where cracks are localized as bounding boxes in the image. The algorithm's real-time processing capability makes it particularly useful for crack detection in large-scale infrastructure, such as pavements bridges, and roads. YOLO models, like YOLOv5, are capable of identifying cracks with high speed and accuracy by leveraging their ability to process the entire image in a single pass. Figure 5 represents the network construction of the YOLO family.

This approach captures both local and global context, enabling the precise detection of fine cracks that might otherwise be missed. Furthermore, YOLO's scalability allows it to handle high-resolution images and datasets, making it a robust choice for practical crack monitoring and maintenance applications [Ma, et al., 2022]. With a noticeably quicker processing time, YOLOv3 can accurately complete the OR assignment with alternative techniques. Similarly, Yu [Yu, 2022] successfully identified and extracted critical crack-related information from images exhibiting the presence of cracks, employing the advanced object detection framework known as YOLOv5. To accurately identify bridge surface fractures, Zhang et al. [Zhang, et al., 2020] suggested using the YOLOv3 method in conjunction with Mobile Nets and a convolutional block attention module (CBAM). Teng et al. [Teng, et al., 2021] used eleven feature extractors to compare the YOLOv2 network's crack detection ability. In a similar vein, Gooda et al. [Gooda, et al., 2023] detected road cracks using the YOLO v5 algorithm. The Summary of Deep Crack Detection Methods by OR Background is shown in Table 2.

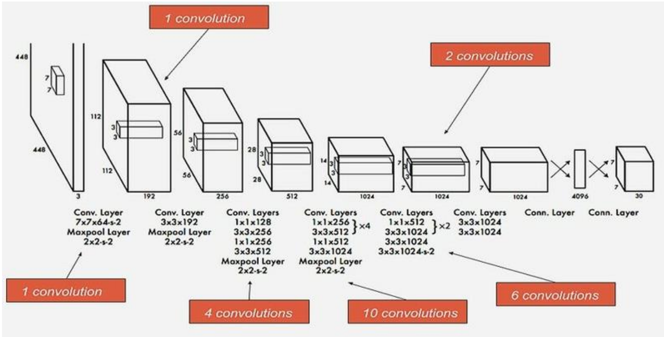

Fig. 5 The network construction of the YOLO family

Table 2. An overview of deep fracture detecting techniques using the OR background

Ref	Novelty	Core Planning	Challenges	Result
[Xu, et al., 2022]	Investigated Faster R-CNN and R-CNN masks for road crack recognition with challenging datasets to evaluate performance in realistic scenarios.	Faster R-CNN architecture with enhanced performance evaluation in the OR setting.	- Struggles with datasets that contain highly variable and noisy images. - Difficulty in detecting fine cracks or very small crack regions in complex road surfaces.	Faster R-CNN achieved better accuracy on clean datasets but struggled with small and intricate cracks in noisy environments.
[Marin, et al., 2021]	Faster R-CNN object detector architecture was used to identify picture cracks.	Faster R-CNN object detection framework.	- Inconsistent performance when applied to real-world, low-resolution, or noisy images. - Challenges with distinguishing cracks from other road anomalies or surface defects.	Achieved reliable crack detection on high-quality datasets but faced reduced performance in noisy real-world conditions.
[Yan and Zhang, 2021]	Deformable convolution was included in the VGG16 backbone feature extraction network to propose a unique deformable SSD network.	Deformable Single Shot MultiBox Detector (SSD) with deformable convolution and VGG16 backbone.	- Deformable SSD may struggle with cracks that have irregular shapes or orientations. - Requires large amounts of annotated training data to achieve robust performance.	Improved crack detection for irregular shapes but required extensive computational resources and high-quality annotated datasets.
[Feng, et al., 2020]	created a pavement crack recognition fusion model by fusing U-Net and multitarget SSD convolutional neural network models.	Fusion model: SSD + U-Net.	- Fusion models can be computationally expensive and slow to process, limiting real-time deployment. - Combining SSD and U-Net may complicate model training and integration, increasing computational load.	Delivered accurate results for diverse crack types but at the cost of processing speed, limiting real-time applications.
[Yu, 2022]	Successfully identified and extracted critical crack-related information using YOLOv5.	YOLOv5	- YOLOv5 may face challenges when detecting thin, long cracks or cracks with varying intensities. - Limited performance in cluttered or noisy images, with potential for false positives.	Achieved high-speed crack detection with moderate accuracy but faced limitations in identifying fine cracks or in cluttered scenarios.
[Zhang, et al., 2020]	Proposed YOLOv3 collective with MobileNets and a convolutional block attention module (CBAM) for accurate bridge surface crack detection.	YOLOv3 + MobileNets + CBAM.	- Combining YOLOv3 with MobileNets and CBAM may introduce complexity in training, affecting speed and accuracy. - Difficulty in generalizing across different types of crack patterns and surface conditions.	Enhanced crack detection for bridge surfaces with improved accuracy but slower performance and reduced adaptability across datasets.
[Teng, et al., 2021]	Compare the crack detection performance of YOLOv2 Using 11 feature extractors.	YOLOv2	- YOLOv2 struggles with high false positive rates, especially in complex and cluttered environments. - Limited ability to differentiate between crack types in low-quality images.	Demonstrated varying performance with different feature extractors, showing limitations in cluttered environments with high false-positive rates.
[Gooda, et al., 2023]	Utilized YOLOv5 for road crack detection.	YOLOv5	- YOLOv5 may not be well-suited for handling extreme conditions such as severe surface degradation or heavily occluded cracks. - High computational resource requirements for training and real-time deployment	Delivered high detection speed and acceptable accuracy under normal conditions but struggled under extreme surface degradation or heavy occlusions.

3.3 Semantic segmentation

Semantic segmentation is the term used in computer vision to do categorization at the pixel level. Applications such as scene interpretation, robot perception, medical image analysis, and satellite image segmentation are just a few of the many fields in which SS is crucial [Kar, et al., 2021]. Semantic segmentation in crack detection divides an image into pixel-level regions to classify each pixel as either a fracture or not [Qiao, et al., 2021]. This method uses deep learning models, like CNNs or U-Net, to precisely locate and outline cracks in images. By focusing on spatial and contextual information, semantic segmentation provides detailed crack maps, improving accuracy detection. Hybrid approaches, such as combining CNNs with transformers, further enhance performance by leveraging both local features and global patterns. The output of a DL-based crack detection method based on the SS setting is shown in Figure 6.

The SS setting has the most published research of any of the other crack detection techniques, demonstrating its popularity and potency in this field. The two primary categories of deep learning-based segmentation techniques for crack identification are hybrid and pure techniques. Hybrid approaches combine multiple techniques, such as traditional image processing approaches with advanced DL models, to control their complementary strengths for improved performance [Li, et al., 2023]. Pure approaches, on the other hand, rely solely on deep learning architectures, focusing on extracting features directly from raw data for crack detection. Each of these two categories is further divided into sub-categories based on the specific methodologies or model architecture used, showcasing the diversity of techniques applied within this framework.

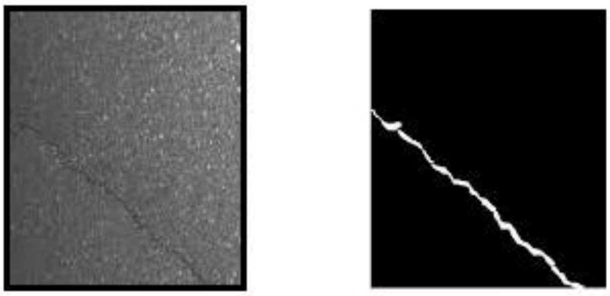

Fig. 6 The outcomes of a fracture-detecting method based on DL and the SS setting

Hybrid Semantic segmentation

Hybrid semantic segmentation combines multiple techniques or models to improve the accuracy and efficacy of segmenting images into meaningful regions. It typically integrates traditional image processing methods with advanced DL models, such as combining CNNs with transformers or using ensemble methods. This approach leverages the strengths of different techniques, like capturing both local and global features. The result is a more robust and adaptable segmentation model, especially effective in complex or high-variability datasets. Mask R-CNN is commonly used for identifying crack patches and areas, followed by crack segmentation using an FCN at the pixel level. Wang et al. [Wang, et al., 2023] employed the advanced Mask R-CNN architecture, a sophisticated deep learning model, to effectively and accurately discern and identify the presence of automatic road cracks, thereby enhancing the precision of road surface analysis and maintenance. Fang et al., [Fang, et al., 2020] outlined a novel hybrid strategy to crack detection that blends Bayesian probabilistic analysis with deep learning models. The technique employs a Bayesian integration approach to decrease false detections and a Faster R-CNN to identify crack repairs. Using a faster region proposal convolutional neural network (Faster R-CNN) algorithm, Kang et al. [Kang, et al., 2020] presented an automated crack detection, and quantification technique. The technique measured the length and thickness of cracks using a modified distance transform method (DTM), bounding boxes, and a modified tubularity flow field (TuFF) algorithm. Similarly, Liu et al., [Liu, et al., 2023] described a Mask (R-CNN) that automatically notices and division small cracks in asphalt 16 pavement at the pixel level. Table 3 represents the rapid Hybrid semantic segmentation.

Pure Semantic Segmentation

Without first identifying potential areas or crack patches, pure crack segmentation is carried done [Yang, et al., 2022]. The literature has investigated several techniques to get pure SS for fracture detection. A completely Convolutional Network (FCN), also known as an encoder-decoder structure, may be created by substituting convolutional layers for completely connected layers in common topologies intended for image classification. This change makes it possible to employ a variety of FCN architectures for computer vision SS tasks [Xu, et al., 2019]. Encoder–decoder structures are the main tool used in pure segmentation techniques for fracture identification. However, several research have also employed other techniques to achieve pixel-level crack segmentation.

Table 3. Summary of Hybrid Semantic Segmentation

Ref	Novelty/Novelties	Core Architecture	Challenges	Result
[Wang, et al., 2023]	Employed the advanced Mask R-CNN architecture to accurately identify road cracks, enhancing precision in road surface analysis and maintenance.	Mask R-CNN architecture for automatic crack detection and division.	- Mask R-CNN may face difficulties in handling small or subtle cracks in noisy environments. - High computational demands, especially for real-time applications on large datasets.	Achieved high segmentation accuracy for crack detection but struggled with computational efficiency and small crack detection in noisy data
[Fang, et al., 2020]	Suggested a hybrid strategy that uses Bayesian probabilistic analysis to reduce false detections and Faster R-CNN for crack patch detection.	Faster R-CNN + Bayesian integration algorithm.	The hybrid approach may struggle with balancing accuracy and speed, especially in complex road conditions. - Bayesian integration may not always fully suppress false positives, leading to occasional misclassifications.	Improved crack patch detection accuracy and reduced false positives but faced challenges in maintaining speed and occasional false positives in outputs.
[Kang, et al., 2020]	Developed a method for crack measurement using Faster R-CNN for automatic crack recognition, localization, and quantification that combines bounding boxes, a modified TuFF algorithm, and a modified DTM.	Faster R-CNN combined with a modified Distance Transform Method (DTM) and the TuFF algorithm).	- Integration of multiple algorithms can introduce complexity in model training and reduce efficiency. - Difficulty in quantifying cracks in highly variable road surfaces or the presence of strong noise.	Demonstrated robust localization and quantification of cracks but faced computational overhead and reduced precision in noisy conditions.
[Liu, et al., 2023]	Utilized Mask R-CNN for pixel-level discovery and segmentation of small cracks in asphalt pavements.	Mask R-CNN architecture for accurate pixel-level crack segmentation.	- Mask R-CNN may be computationally expensive and require significant memory, hindering real-time deployment in large-scale applications. - Challenges in distinguishing cracks from other road anomalies or surface imperfections in low-resolution images.	Delivered accurate pixel-level segmentation for small cracks but had limitations in low-resolution images and real-time performance.

To extract deep features from the input picture, SS uses an encoder-decoder structure that depends on a backbone architecture [Chen, et al., 2019]. Convolutional, pooling, and activation layers make up this backbone, which gradually shrinks the input image's width and height while obtaining high-level feature representations [Liu, et al., 2021]. However, a decoder module is used to restore the spatial resolution since the image's dimensions decrease as it moves through the encoder. Pixel-level classification for accurate segmentation is made possible by the decoder's use of a sequence of Deconvolution layers, which resize the features to the original picture dimensions. Because it can effectively handle detailed characteristics, the encoder-decoder structure has been frequently used for crack segmentation tasks [Kang and Cha, 2022]. Several well-established architectures, including UNet, SegNet, and FC-DenseNet, have been prominently applied in crack detection. These architectures are particularly well-suited for semantic segmentation in computer vision, offering robust performance in identifying and segmenting cracks [Hernández, 2023].

Ensuring a smooth flow of contextual information across the architecture is essential to further increasing the segmentation accuracy in the SS situation [Singh and Sengar, 2024]. This involves incorporating strategies to maintain and enhance the global and local feature relationships, which are essential for capturing the intricate characteristics of cracks. Various techniques, such as skip connections, multi-scale feature extraction, and attention mechanisms, have been explored and integrated into crack detection frameworks to optimize the performance of encoder-decoder-based approaches. The following sections will review these advancements and their applications in crack segmentation.

U- Net

A well-known deep learning architecture, U-Net, has been utilized extensively for crack detection because of its potent performance in semantic segmentation tasks. Elhariri et al. [Elhariri, et al., 2022] used U-Net deep learning network topologies to present an automated deep crack segmentation method. The study identified important pixels for image classification and used deep-learned features for better segmentation. U-Net is perfect for detecting fractures, which frequently manifest as delicate and asymmetrical patterns, because of its encoder-decoder structure, which allows it to capture both low-level and high-level data [Hang, et al., 2023]. Convolution and pooling layers are used by the encoder to extract features in crack detection, gradually decreasing the input image's spatial resolution. To restore the original picture dimensions for pixel-level segmentation, the decoder upscales these features using transposed convolutions [Cheng, et al., 2023]. To preserve spatial information and improve segmentation accuracy, U-Net integrates skip links that directly connect compliant encoder and decoder layers. This architecture is particularly effective for crack detection in applications like roads, bridges, and pavements, where precise localization is essential. Its ability to handle small datasets and deliver high accuracy with fine-grained outputs has made U-Net a go-to model in the field. Advanced versions of U-Net, such as U-Net++, further enhance crack detection by leveraging additional connections and densely nested structures.

Seg- Net

SegNet, another widely used deep learning architecture, is highly effective for crack detection due to its efficient encoder-decoder design tailored for semantic segmentation tasks [Li, et al., 2022]. SegNet is a popular option for large-scale infrastructure applications like roads, bridges, and buildings because it performs especially well in conditions with limited resources. Through a sequence of convolutional and pooling layers, SegNet's encoder collects features, compressing the input image's spatial information while keeping the most important features [Dayananda, et al., 2022]. Unlike other architectures, SegNet saves the indices of the max-pooling operation in the encoder, which are then used in the decoder to perform up sampling. This approach eliminates the need for learnable parameters in the decoder, reducing computational overhead and improving efficiency.

For crack detection, SegNet effectively restores the spatial resolution of cracks while maintaining fine details, enabling precise pixel-level segmentation. Its lightweight design and reduced memory requirements make it ideal for real-time crack monitoring and maintenance tasks. Additionally, its ability to handle noisy and complex images further enhances its application in detecting subtle and irregular cracks in challenging environments.

Dense-Net

DenseNet (Dense Convolutional Network) has been successfully adapted for crack detection due to its unique architecture that promotes feature reuse and efficient learning [Luo, et al., 2024]. In contrast to conventional convolutional networks, DenseNet strengthens feature propagation and minimizes unnecessary computations by feed-forwardly connecting each layer to every other layer. DenseNet is used in crack detection to extract local and global characteristics that are essential for spotting subtle and erratic crack patterns [Li, et al., 2021]. The network is made up of dense blocks, with each layer passing its feature mappings to later levels and receiving inputs from all layers before it. This dense connectivity enables the model to learn rich and diversified features, making it highly effective for segmenting cracks with complex geometries.

Particularly useful for crack detection tasks with little datasets, DenseNet's compact architecture minimizes overfitting and lowers the number of parameters [Matarneh, et al., 2024]. By substituting convolutional layers for fully connected layers, variants like as FC-DenseNet (Fully Convolutional DenseNet) further modify DenseNet for semantic segmentation, allowing for pixel-level predictions [Abbas, et al., 2021]. This makes DenseNet and its adaptations well-suited for applications requiring high precision in crack localization and segmentation.

4. Software Structure for Dl Applications

Research organizations have developed a variety of frameworks to carry out deep learning tasks. The most well-known frameworks that remain to evolve are shown in Table 4.

Table 4. Popular Deep Learning Framework for crack detection

Framework	Interface maintenance
TensorFlow	Python, C/C++, Java
Torch/PyTorch	Python, C/C++, Lua
Keras	Python, Matlab
Caffe	Python, Matlab
Theano	Python

TensorFlow: TensorFlow plays a significant role in the DL approaches for crack detection, offering a robust outline for developing and deploying models that can accurately identify structural defects. Various studies have utilized TensorFlow to improve the performance of CNNs in detecting cracks in different contexts, such as concrete surfaces, bridges, tunnels, and pavements. TensorFlow is utilized in the DL models for crack detection in concrete surfaces, enabling the development and improvement of architectures like VGG19, VGG16, and MobileNetV2 [Philip, et al., 2023].

Torch/PyTorch: PyTorch is preferred in crack detection for its dynamic computation graph, making model development and debugging easier [Ma, et al., 2023]. Researchers utilize PyTorch for implementing advanced architecture like faster R-CNN and mask R-CNN. Its integration with libraries like Torch Vision facilitates the efficient processing of image-based datasets. Additionally, PyTorch's flexibility makes it ideal for experimenting with novel architectures in segmentation and detection tasks.

Keras: Keras is a popular choice for beginners in crack detection research. It simplifies the development of encoder-decoder architecture such as U-Net for segmentation tasks [74]. Pre-trained models available in Keras are often fine-tuned for specific crack detection datasets. Its compatibility with TensorFlow enhances its utility for real-world applications in pavement and structural analysis.

Caffe: Caffe is a high-performance deep-learning framework often used for crack detection when computational efficiency is a priority [Khanam, et al., 2024] [Wu, et al., 2022]. Its modularity enables researchers to implement customized CNNs for detecting and classifying cracks. While less flexible than TensorFlow or PyTorch, Caffe excels in training models for large datasets. It is commonly applied in industrial settings for crack analysis in materials.

Theano: Theano is one of the earliest deep learning frameworks, used in crack detection for its ability to optimize mathematical operations [Azimi, et al., 2020]. Researchers leverage Theano for implementing basic neural network architectures for crack detection tasks. Though now largely replaced by newer frameworks, Theano contributed significantly to the development of models for structural health monitoring. Its GPU acceleration provided a foundation for computationally intensive tasks in early crack analysis research.

5. Future Research Direction

The ground of crack detection using DL has achieved remarkable progress; however, there remain several unexplored avenues and critical challenges that warrant further investigation. One significant area is the detection and classification of diverse crack types, including less-studied forms such as meandering, crescent-shaped, and branched cracks. Developing models capable of accurately recognizing these complex geometries can significantly enhance the applicability of crack detection systems across varied infrastructure scenarios. Additionally, integrating advanced techniques such as anomaly segmentation, deep edge detection, and attention mechanisms can improve model sensitivity, especially for detecting thin or irregular cracks. Another promising direction is the exploration of domain adaptation and transfer learning to ensure robust presentation across diverse datasets and environmental conditions, addressing the limitations of data variability in real-world scenarios.

Furthermore, by fusing cutting-edge deep learning architectures with conventional computer vision approaches, it may be possible to close the gap between high accuracy and computational efficiency, making crack detection systems more appropriate for use in contexts with limited resources. Research into lightweight models optimized for real-time applications, such as on-site monitoring via drones or handheld devices, is essential for practical implementation. Furthermore, incorporating multimodal data sources, including thermal imaging, ultrasonic signals, or LiDAR, could provide complementary information, enabling more comprehensive crack assessment.

A critical avenue for future investigation is the development of explainable and interpretable DL models, which can provide insights into the decision-making procedure. This is particularly vital in the context of infrastructure monitoring, where decisions carry significant economic and safety implications. Investigating the relationship between crack complexity, environmental factors, and model performance could guide the design of more effective architectures and loss functions. Lastly, fostering collaborative efforts between academia and industry to standardize datasets, evaluation metrics, and benchmarking protocols will ensure the reproducibility and generalizability of research findings. By addressing these challenges and exploring these directions, future studies can pave the way for innovative, accurate, and efficient crack detection systems, ultimately contributing to enhanced infrastructure health and safety.

6. Conclusion

This study provided a thorough analysis of DL-based crack detection techniques, highlighting their noteworthy contributions and the important developments they have brought to the field. DL approaches have become effective tools due to the growing need for precise and effective fracture detection procedures to guarantee the longevity and safety of organizations. Through the automated and efficient feature extraction capabilities of deep architectures, these methods have significantly improved the accuracy and robustness of fracture detection systems. To provide light on related approaches to the crack detection problem, the research examined was further classified according to their computer vision components. Because crack segmentation techniques employing semantic segmentation settings may provide excellent detection and pixel-level localization accuracy, they have received special attention. Even if encouraging outcomes have been attained, the sector still confronts obstacles, such as the requirement to recognize and categorize different kinds of cracks. Investigating cutting-edge methods like anomaly segmentation and deep-edge detection might increase segmentation accuracy even further, especially for thin or intricate fractures. Furthermore, researching how crack shape complexity and model performance are related may help build more efficient architectures and loss functions. This study offers scholars and practitioners a useful resource by highlighting recent developments and suggesting possible future paths. It provides the framework for the creation of novel, effective, and precise crack detection systems that not only solve current problems but also push the limits of what is practical in this crucial area. The information provided here is meant to encourage more study and development, which will eventually result in more reliable and safe infrastructure throughout the world.

Acknowledgement

We are thankful to Poornima University Management for providing the required infrastructure to carry out this research. We are also thankful to the Centre of Excellence in Machine Learning Application for Society for their support in this work

References

Abbas, Q., Qureshi, I., & Ibrahim, M. E. (2021). An automatic detection and classification system of five stages for hypertensive retinopathy using semantic and instance segmentation in DenseNet architecture. Sensors, 21(20), 6936. https://doi.org/10.3390/s21206936

Aravind, N., Nagajothi, S., & Elavenil, S. (2021). Machine learning model for predicting the crack detection and pattern recognition of geopolymer concrete beams. Construction and Building Materials, 297, 123785. https://doi.org/10.1016/j.conbuildmat.2021.123785

Arman, M. S., Hasan, M. M., Sadia, F., Shakir, A. K., Sarker, K., & Himu, F. A. (2020). Detection and classification of road damage using R-CNN and faster R-CNN: a deep learning approach. In Cyber Security and Computer Science: Second EAI International Conference, ICONCS 2020, Dhaka, Bangladesh, February 15-16, 2020, Proceedings 2 (pp. 730-741). Springer International Publishing. https://doi.org/10.1007/978-3-030-52856-0_58

Azimi, M., Eslamlou, A. D., & Pekcan, G. (2020). Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review. Sensors, 20(10), 2778. https://doi.org/10.3390/s20102778

Cha, Y. J., Ali, R., Lewis, J., & Büyüköztürk, O. (2024). Deep learning-based structural health monitoring. Automation in Construction, 161, 105328. https://doi.org/10.1016/j.autcon.2024.105328

Chen, H., Lin, H., & Yao, M. (2019). Improving the efficiency of encoder-decoder architecture for pixel-level crack detection. IEEE Access, 7, 186657-186670. https://doi.org/10.1109/ACCESS.2019.2961375

Chen, H., Wang, Y., Guo, T., Xu, C., Deng, Y., Liu, Z., ... & Gao, W. (2021). Pre-trained image processing transformers. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 12299-12310). https://doi.org/10.1109/CVPR46437.2021.01212

Cheng, Y., Wang, W., Ren, Z., Zhao, Y., Liao, Y., Ge, Y., ... & Zhang, C. (2023). Multi-scale Feature Fusion and Transformer Network for urban green space segmentation from high-resolution remote sensing images. International Journal of Applied Earth Observation and Geoinformation, 124, 103514. https://doi.org/10.1016/j.jag.2023.103514

Choi, Y., Bae, B., Han, T. H., & Ahn, J. (2024). Application of Mask R-CNN and YOLOv8 Algorithms for Concrete Crack Detection. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3469951

Dayananda, C., Choi, J. Y., & Lee, B. (2022). A squeeze U-SegNet architecture based on residual convolution for brain MRI segmentation. IEEE Access, 10, 52804-52817. https://doi.org/10.1109/ACCESS.2022.3175188

Elhariri, E., El-Bendary, N., & Taie, S. A. (2022). Automated pixel-level deep crack segmentation on historical surfaces using U-Net models. Algorithms, 15(8), 281. https://doi.org/10.3390/a15080281

Fang, F., Li, L., Gu, Y., Zhu, H., & Lim, J. H. (2020). A novel hybrid approach for crack detection. Pattern Recognition, 107, 107474. https://doi.org/10.1016/j.patcog.2020.107474

Feng, X., Xiao, L., Li, W., Pei, L., Sun, Z., Ma, Z., ... & Ju, H. (2020). Pavement crack detection and segmentation method based on improved deep learning fusion model. Mathematical Problems in Engineering, 2020(1), 8515213. https://doi.org/10.1155/2020/8515213

Golding, V. P., Gharineiat, Z., Munawar, H. S., & Ullah, F. (2022). Crack detection in concrete structures using deep learning. Sustainability, 14(13), 8117. https://doi.org/10.3390/su14138117

Gooda, S. K., Chinthamu, N., Selvan, S. T., Rajakumareswaran, V., & Brindha, G. (2023). Automatic detection of road cracks using EfficientNet with residual U-net-based segmentation and YOLOv5-based detection. International Journal on Recent and Innovation Trends in Computing and Communication, 11, 84-91. https://doi.org/10.17762/ijritcc.v11i4s.6310

Gooda, S. K., Chinthamu, N., Selvan, S. T., Rajakumareswaran, V., & Brindha, G. (2023). Automatic detection of road cracks using EfficientNet with residual U-net-based segmentation and YOLOv5-based detection. International Journal on Recent and Innovation Trends in Computing and Communication, 11, 84-91. https://doi.org/10.17762/ijritcc.v11i4s.6310

Guo, W., Zhong, L., Zhang, D., & Li, Q. (2024). Pavement crack detection using fractal dimension and semi-supervised learning. Fractal and Fractional, 8(8), 468. https://doi.org/10.3390/fractalfract8080468

Gupta, A., Anpalagan, A., Guan, L., & Khwaja, A. S. (2021). Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. Array, 100057. https://doi.org/10.1016/j.array.2021.100057

Hamishebahar, Y., Guan, H., So, S., & Jo, J. (2022). A comprehensive review of deep learning-based crack detection approaches. Applied Sciences, 12(3), 1374. https://doi.org/10.3390/app12031374

Hang, J., Wu, Y., Li, Y., Lai, T., Zhang, J., & Li, Y. (2023). A deep learning semantic segmentation network with attention mechanism for concrete crack detection. Structural Health Monitoring, 22(5), 3006-3026. https://doi.org/10.1177/14759217221126170

Hernández, P. G. (2023). Void characterisation of composite prepreg laminates using X-ray Imaging and Deep Learning methods (Doctoral dissertation, University of Bristol). https://gogeometry.com/software/ai/machine-learning-cognitive-mind-map.html

Kang, D. H., & Cha, Y. J. (2022). Efficient attention-based deep encoder and decoder for automatic crack segmentation. Structural Health Monitoring, 21(5), 2190-2205. https://doi.org/10.1177/14759217211053776

Kang, D., Benipal, S. S., Gopal, D. L., & Cha, Y. J. (2020). Hybrid pixel-level concrete crack segmentation and quantification across complex backgrounds using deep learning. Automation in Construction, 118, 103291. https://doi.org/10.1016/j.autcon.2020.103291

Kar, M. K., Nath, M. K., & Neog, D. R. (2021). A review of progress in semantic image segmentation and its application to medical images. SN computer science, 2(5), 397. https://doi.org/10.1007/s42979-021-00784-5

Karimi, N., Mishra, M., & Lourenço, P. B. (2024). Automated surface crack detection in historical constructions with various materials using deep learning-based YOLO network. International Journal of Architectural Heritage, 1-17. https://doi.org/10.1080/15583058.2024.2376177

Kaveh, H., & Alhajj, R. (2024). Recent advances in crack detection technologies for structures: a survey of 2022-2023 literature. Frontiers in Built Environment, 10, 1321634. https://doi.org/10.3389/fbuil.2024.1321634

Khanam, R., Hussain, M., Hill, R., & Allen, P. (2024). A comprehensive review of convolutional neural networks for defect detection in industrial applications. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3425166

Kulambayev, B., Astaubayeva, G., Tleuberdiyeva, G., Alimkulova, J., Nussupbekova, G., & Kisseleva, O. (2024). Deep CNN Approach with Visual Features for Real-Time Pavement Crack Detection. International Journal of Advanced Computer Science & Applications, 15(3). https://doi.org/10.14569/IJACSA.2024.0150333

Li, H., Wang, W., Wang, M., Li, L., & Vimlund, V. (2022). A review of deep learning methods for pixel-level crack detection. Journal of Traffic and Transportation Engineering (English Edition), 9(6), 945-968. https://doi.org/10.1016/j.jtte.2022.11.003

Li, H., Zong, J., Nie, J., Wu, Z., & Han, H. (2021). Pavement crack detection algorithm based on densely connected and deeply supervised network. IEEE Access, 9, 11835-11842. https://doi.org/10.1109/ACCESS.2021.3050401

Li, M., Jiang, Y., Zhang, Y., & Zhu, H. (2023). Medical image analysis using deep learning algorithms. Frontiers in Public Health, 11, 1273253. https://doi.org/10.3389/fpubh.2023.1273253

Liu, Y., Pu, H., & Sun, D. W. (2021). Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices. Trends in Food Science & Technology, 113, 193-204. https://doi.org/10.1016/j.tifs.2021.04.042

Liu, Z., Yeoh, J. K., Gu, X., Dong, Q., Chen, Y., Wu, W., ... & Wang, D. (2023). Automatic pixel-level detection of vertical cracks in asphalt pavement based on GPR investigation and improved mask R-CNN. Automation in Construction, 146, 104689. https://doi.org/10.1016/j.autcon.2022.104689

López-Monroy, A. P., & García-Salinas, J. S. (2022). Neural networks and deep learning. In Biosignal Processing and Classification Using Computational Learning and Intelligence (pp. 177-196). Academic Press. https://doi.org/10.1016/B978-0-12-820125-1.00021-X

Loverdos, D., & Sarhosis, V. (2022). Automatic image-based brick segmentation and crack detection of masonry walls using machine learning. Automation in Construction, 140, 104389. https://doi.org/10.1016/j.autcon.2022.104389

Luo, Y., Liu, X., Song, R., Majeed, A. P. A., Zhang, F., Xu, Y., ... & Chen, W. (2024). Enhancing Bone Fracture Detection: A Feature-Based Transfer Learning Approach Using DenseNet with SVM. In 2nd International Conference on Intelligent Manufacturing and Robotics (ICiMR). https://doi.org/10.1007/978-981-96-3949-6_28

Ma, J., Yan, W., Liu, G., Xing, S., Niu, S., & Wei, T. (2022). Complex texture contour feature extraction of cracks in timber structures of ancient architecture based on YOLO algorithm. Advances in Civil Engineering, 2022(1), 7879302 https://doi.org/10.1155/2022/7879302

Ma, K., Meng, X., Hao, M., Huang, G., Hu, Q., & He, P. (2023). Research on the efficiency of bridge crack detection by coupling deep learning frameworks with convolutional neural networks. Sensors, 23(16), 7272. https://doi.org/10.3390/s23167272

Malekloo, A., Ozer, E., AlHamaydeh, M., & Girolami, M. (2022). Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights. Structural Health Monitoring, 21(4), 1906-1955. https://doi.org/10.1177/14759217211036880

Marin, B., Brown, K., & Erden, M. S. (2021, August). Automated masonry crack detection with faster R-CNN. In 2021 IEEE 17th International Conference on Automation Science and Engineering (CASE) (pp. 333-340). IEEE. https://doi.org/10.1109/CASE49439.2021.9551683

Matarneh, S., Elghaish, F., Rahimian, F. P., Abdellatef, E., & Abrishami, S. (2024). Evaluation and optimisation of pre-trained CNN models for asphalt pavement crack detection and classification. Automation in Construction, 160, 105297. https://doi.org/10.1016/j.autcon.2024.105297

Negi, P., Kromanis, R., Dorée, A. G., & Wijnberg, K. M. (2024). Structural health monitoring of inland navigation structures and ports: a review on developments and challenges. Structural Health Monitoring, 23(1), 605-645. https://doi.org/10.1177/14759217231170742

Philip, R. E., Andrushia, A. D., Nammalvar, A., Gurupatham, B. G. A., & Roy, K. (2023). A comparative study on crack detection in concrete walls using transfer learning techniques. Journal of Composites Science, 7(4), 169. https://doi.org/10.3390/jcs7040169

Plevris, V., & Papazafeiropoulos, G. (2024). AI in Structural Health Monitoring for Infrastructure Maintenance and Safety. Infrastructures, 9(12), 225. https://doi.org/10.3390/infrastructures9120225

Qiao, W., Liu, Q., Wu, X., Ma, B., & Li, G. (2021). Automatic pixel-level pavement crack recognition using a deep feature aggregation segmentation network with a scSE attention mechanism module. Sensors, 21(9), 2902. https://doi.org/10.3390/s21092902

Rodriguez-Conde, I., Campos, C., & Fdez-Riverola, F. (2022). Optimized convolutional neural network architectures for efficient on-device vision-based object detection. Neural Computing and Applications, 34(13), 10469-10501. https://doi.org/10.1007/s00521-021-06830-w

Safaei, N., Smadi, O., Masoud, A., & Safaei, B. (2022). An automatic image processing algorithm based on crack pixel density for pavement crack detection and classification. International Journal of Pavement Research and Technology, 15(1), 159-172. https://doi.org/10.1007/s42947-021-00006-4

Sarker, I. H. (2021). Deep learning: a comprehensive overview of techniques, taxonomy, applications, and research directions. SN computer science, 2(6), 420. https://doi.org/10.1007/s42979-021-00815-1

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN computer science, 2(3), 160. https://doi.org/10.1007/s42979-021-00592-x

Sengupta, S., Basak, S., Saikia, P., Paul, S., Tsalavoutis, V., Atiah, F., ... & Peters, A. (2020). A review of deep learning with special emphasis on architecture, applications and recent trends. Knowledge-Based Systems, 194, 105596. https://doi.org/10.1016/j.knosys.2020.105596

Shen, A., Dai, M., Hu, J., Liang, Y., Wang, S., & Du, J. (2024). Leveraging Semi-Supervised Learning to Enhance Data Mining for Image Classification under Limited Labeled Data. arXiv preprint arXiv:2411.18622. https://doi.org/10.1109/EIECC64539.2024.10929276

Singh, O., & Sengar, S. S. (2024). BetterNet: An Efficient CNN Architecture with Residual Learning and Attention for Precision Polyp Segmentation. arXiv preprint arXiv:2405.04288.

Sun, H., Burton, H. V., & Huang, H. (2021). Machine learning applications for building structural design and performance assessment: State-of-the-art review. Journal of Building Engineering, 33, 101816. https://doi.org/10.1016/j.jobe.2020.101816

Taherdoost, H. (2024). Beyond Supervised: The Rise of Self-Supervised Learning in Autonomous Systems. Information, 15(8), 491. https://doi.org/10.3390/info15080491

Teng, S., Liu, Z., Chen, G., & Cheng, L. (2021). Concrete crack detection based on a well-known feature extractor model and the YOLO_v2 network. Applied Sciences, 11(2), 813. https://doi.org/10.3390/app11020813

Tibaduiza Burgos, D. A., Gomez Vargas, R. C., Pedraza, C., Agis, D., & Pozo, F. (2020). Damage identification in structural health monitoring: A brief review from its implementation to the use of data-driven applications. Sensors, 20(3), 733. https://doi.org/10.3390/s20030733

Tsalera, E., Papadakis, A., & Samarakou, M. (2021). Comparison of pre-trained CNNs for audio classification using transfer learning. Journal of Sensor and Actuator Networks, 10(4), 72. https://doi.org/10.3390/jsan10040072

Ullah, S. (2023). GUIDED WAVE-BASED METHODS FOR DELAMINATION IDENTIFICATION ENHANCED BY DEEP LEARNING (Doctoral dissertation, Institute of Fluid-Flow Machinery, Polish Academy of Sciences).

Van Hoang, T. (2024). Impact of integrated artificial intelligence and Internet of Things technologies on smart city transformation. Journal of Technical Education Science, 19(Special Issue 01), 64-73. https://doi.org/10.54644/jte.2024.1532

Wang, W., Shi, Y., Zhang, J., Hu, L., Li, S., He, D., & Liu, F. (2023). Traditional village building extraction based on improved Mask R-CNN: a case study of Beijing, China. Remote Sensing, 15(10), 2616. https://doi.org/10.3390/rs15102616

Wu, D., Zhang, H., & Yang, Y. (2022). [Retracted] Deep Learning-Based Crack Monitoring for Ultra-High-Performance Concrete (UHPC). Journal of Advanced Transportation, 2022(1), 4117957. https://doi.org/10.1155/2022/4117957

Xu, H., Yao, L., Zhang, W., Liang, X., & Li, Z. (2019). Auto-fpn: Automatic network architecture adaptation for object detection beyond classification. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 6649-6658). https://doi.org/10.1109/ICCV.2019.00675

Xu, X., Zhao, M., Shi, P., Ren, R., He, X., Wei, X., & Yang, H. (2022). Crack detection and comparison study based on faster R-CNN and mask R-CNN. Sensors, 22(3), 1215. https://doi.org/10.3390/s22031215

Xu, X., Zhao, M., Shi, P., Ren, R., He, X., Wei, X., & Yang, H. (2022). Crack detection and comparison study based on faster R-CNN and mask R-CNN. Sensors, 22(3), 1215. https://doi.org/10.3390/s22031215

Yan, K., & Zhang, Z. (2021). Automated asphalt highway pavement crack detection based on deformable single shot multi-box detector under a complex environment. IEEE Access, 9, 150925-150938. https://doi.org/10.1109/ACCESS.2021.3125703

Yan, K., & Zhang, Z. (2021). Automated asphalt highway pavement crack detection based on deformable single shot multi-box detector under

a complex environment. IEEE Access, 9, 150925-150938. https://doi.org/10.1109/ACCESS.2021.3125703

Yang, C., Chen, J., Li, Z., & Huang, Y. (2021). Structural crack detection and recognition based on deep learning. Applied sciences, 11(6), 2868. https://doi.org/10.3390/app11062868

Yang, G., Liu, K., Zhang, J., Zhao, B., Zhao, Z., Chen, X., & Chen, B. M. (2022). Datasets and processing methods for boosting visual inspection of civil infrastructure: A comprehensive review and algorithm comparison for crack classification, segmentation, and detection. Construction and Building Materials, 356, 129226. https://doi.org/10.1016/j.conbuildmat.2022.129226

Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M., & Henkel, J. (2024). A Comprehensive Survey of Convolutions in Deep Learning: Applications, Challenges, and Future Trends. IEEE Access, 12, 41180-41218. https://doi.org/10.1109/ACCESS.2024.3376441

Yu, Z. (2022). YOLO V5s-based deep learning approach for concrete cracks detection. In SHS Web of Conferences (Vol. 144, p. 03015). EDP Sciences. https://doi.org/10.1051/shsconf/202214403015

Zhang, H. (2023). Machine Vision: A Comprehensive Analysis of Techniques, Applications, and Challenges. Highlights in Science, Engineering and Technology, 71, 299-304. https://doi.org/10.54097/hset.v71i.13050

Zhang, Q. (2022). Multi-sensor Data Interpretation and Fusion Frameworks for Bridge Deck Condition Assessment (Doctoral dissertation, University of Pittsburgh).

Zhang, Q., Barri, K., Babanajad, S. K., & Alavi, A. H. (2021). Real-time detection of cracks on concrete bridge decks using deep learning in the frequency domain. Engineering, 7(12), 1786-1796. https://doi.org/10.1016/j.eng.2020.07.026

Zhang, Y., Huang, J., & Cai, F. (2020). On bridge surface crack detection based on an improved YOLO v3 algorithm. IFAC-PapersOnLine, 53(2), 8205-8210 https://doi.org/10.1016/j.ifacol.2020.12.1994

Zhong, X., & Ban, H. (2022). Pre-trained network-based transfer learning: A small-sample machine learning approach to nuclear power plant classification problem. Annals of Nuclear Energy, 175, 109201. https://doi.org/10.1016/j.anucene.2022.109201

Disclaimer