



Pavement Distress Detection Using Deep Learning Based Methods: A Survey on Role, Challenges and Opportunities

Ankit Khatri^(✉), Ravi Khatri, Abhishek Kumar, and Kuldeep Kumar

Dr B R Ambedkar National Institute of Technology Jalandhar, Jalandhar, Punjab, India
{ankitk.cs.21,ravik.cs.21,abhishekk.cs.21,kumark}@nitj.ac.in

Abstract. Roadways have always been one of the most used modes of transportation, and their contribution to the nation's economy is also huge. To meet the demands of the growing global population and an increase in urbanization, there has been an exponential rise in the number of vehicles plying on the roads as well as the length of the roads. With this increase in traffic, coupled with other issues like heavy rainfall, the material used for the construction of the road, etc., the condition of the roads deteriorates with cracks and potholes developing on them, which may lead to serious accidents. For effective maintenance of roads and to reduce the associated risks, these defects must be detected. With the advent of Deep Learning (DL) in the recent past and its applications in various sectors, we have comprehensively explored various approaches, particularly using DL in this study, along with the associated challenges in adopting such techniques and future opportunities in this domain. Based on our analysis, using object detection-based models turned out to outperform other approaches.

Keywords: Pavement distress · Deep learning · CNN · Object recognition

1 Introduction

India, with 11% of the world's share of fatalities due to road accidents, sadly tops this list [1]. According to Ministry of Road Transport and Highways data, the majority of these deaths occur between the ages of 18 and 60 [1], which is the most economically productive age. Thus, it leads to a great loss to the country in the form of its most precious asset, i.e., human resources. The World Health Organization has stated that among the major causes of deaths in the world, road accidents stand at the eighth position. These deaths incur a huge burden in the form of treatment costs as well as hamper the productivity of the dead, impaired, and even their family members. Although there are a variety of causes that lead to such accidents, which include overspeeding, driver's negligence, bad condition of roads, the existence of cracks and potholes, lack of strong vehicle standards in the country, etc. [36]. Out of these, one of the major reasons for this massive number of fatalities is the poor condition of roads with multiple defects, which are caused by heavy rainfall, which results in inundation of water over the pavements,

the rising number of vehicles plying on them, the lack of quality construction material used and unstable soil. The significance of our study is as follows:

For the efficient maintenance of the roads and to ensure the safety of people on these roads, the magnitude of such road damage must be known. There are multiple ways to assess the extent of road damage, which may be broadly classified into manual, semi-automated, and fully automated.

In the case of manual assessment, the assessors ply the road using a vehicle moving slowly. They note down the extent of such defects by visual assessment. But this approach takes a lot of time and requires extensive human interference.

In the case of semi-automated assessment [2–4], the evaluation of road damage is done through the collection of images using automated systems. These images are collected using a swift-moving vehicle. These are then sent for manual defect detection, which again requires human intervention.

In fully automated road damage assessment, complex and modern sets of sensors [3–5] are installed on the moving vehicle. These sensors gather the images of the roads, which are then sent for automated defect detection. But this kind of fully automated system is very costly and many countries lack the funds to implement such systems. However, in today's world, mobile phone devices have very good cameras and impressive computation power, which presents an economical and efficient solution to the above problem. For example, Mertz et al. [6] have used mobile phone devices to collect images of roads that were positioned on various frequently plying vehicles. Mobile phone devices with sensors were used by Casas-Avellaneda and Lopez-Parra [7] to detect potholes on the roads. A mobile application was developed by Maeda et al. [8] for automatic road evaluation in Japan.

Motivation for this Study: Roads and infrastructure are crucial to national growth. Metaled or unmetalled roads drive economic growth. Natural disasters, environmental causes, human interference, inexpensive building materials, and normal wear and tear degrade roads, causing cracks, potholes, and other deformations that cause accidents, human and animal deaths, economic slowness, etc. It's important to repair these problems and identify them first. There have been manual and automated approaches to achieve it. After DL's excellent performance in image issues, various research has analyzed its ability to automatically recognise road damage. Our paper includes all significant and state-of-the-art methodologies implemented in this domain, together with their advantages, limitations, and performance over key evaluation measures. This study aims to help future scholars obtain a complete source of important studies in this field.

This study includes all recent and state-of-the-art road damage identification solutions presented by different authors. We've also included the benefits, cons, and challenges of these studies.

The remainder of the paper is organized as follows: First, we have discussed a few basic concepts and terminologies in Sect. 2. In Sect. 3, we have reviewed research articles comprehensively related to the problem domain. In Sect. 4, we have included the challenges in this domain and the contributions made by our study. Finally, we have concluded the paper in Sect. 5 along with future scope.

2 Basic Concepts and Terminologies

In this section, we will discuss some relevant concepts and terminologies using which the problem of pavement distress detection has been solved by most of the studies in the literature survey. The concept of deep learning is the baseline for proposing solutions to the problem. Various architectures, including CNN, region-based CNN, YOLO, etc., have been used by various studies to propose their solutions. These architectures have been explained in brief in this section.

2.1 Deep Learning

This type of learning tries to mimic the workings of the human brain. There are billions of neurons present inside a human brain that operate using electrical impulses [42]. In DL, we try to create a similar network of neurons, which is known as a neural network that consists of multiple layers. The first layer accepts the inputs, the last layer outputs the results, and in between the network may contain hidden layers. As the data is growing at a very fast rate and with the advancements in the computation power of the systems, the capacity of handling such large data has drastically improved. Deep Learning techniques coupled with this improved processing power have achieved significant results in problems involving big data. Deep learning is used in different areas, such as natural language processing, optical character recognition, recommendation systems, etc. It has proven to be extremely effective, particularly in the identification of objects in images, and it is also used to identify road damage using images of defects.

2.2 Convolutional Neural Networks

It is used in a variety of image processing applications. It extracts important information from an image through a four-layered process that includes a convolution layer, a pooling layer, a ReLU layer, and a fully connected layer [43]. This type of neural network works on data that is in grid format. An example of this kind of data is images. Hence, for image processing tasks, convolutional neural networks are preferred.

2.3 Object Detection Algorithms

With the emergence of the concept of the self-driving car, there has been a significant increase in the field of object detection. The main aim of this task is to locate the entity of interest in the given input image. As a conventional CNN cannot be used for the efficient implementation of such tasks, advanced algorithms have been developed. These architectures are briefly described below:

R-CNN: [37] Using the selective search algorithm, a set of 2000 regions is drawn out from the image and the task of finding the objects is performed.

Fast R-CNN: [38] It is quicker than the R-CNN algorithm, as the convolution phase is performed just once for each input and a corresponding feature map is produced from it.

Faster R-CNN: [39] It outperforms the above algorithms by using a different network to forecast the region's proposals instead of using the selective search algorithm.

You Only Look Once (YOLO): [40] This object detection algorithm creates bounding boxes along with the class probabilities using only a single CNN. Then, to find the thing of interest, the box with a likelihood estimate above a certain threshold is chosen.

Single Shot Detection: [41] Similar to the YOLO algorithm, the Single Shot Detection algorithm is also capable of detecting the object of interest in a single scan of the image. It is faster and more accurate than R-CNN's as it gets rid of the region proposals used in the latter.

3 Literature Review

There has been a significant amount of research in the domain of pavement distress detection. In this section, we have summarized different methodologies used by researchers. We have also listed the datasets used by the authors for the implementation process in Table 1. We have explored various machine learning techniques which use advanced detection algorithms for the above problem domain. Table 2 is a summary of the articles we looked at, which brings this section to a close.

Du et al. [9] have used a huge dataset of 45,788 images that were collected using a camera mounted on a dedicated vehicle covering almost 200 km at a rate compared with fewer than 80 km per hour for assessing road damage. They have used the YOLO algorithm to project the position of the fault as well as its category in a given input image. However, they have not used the cost-effective approach of using mobile phone devices for image collection. Instead, they have used advanced high-resolution cameras for this purpose. The YOLO-based model used has a very high level of accuracy and doesn't require any manual work during detection.

Majidifard et al. [10] have proposed a solution using a labelled dataset that consists of images collected from different camera views, which consists of both top-down and wide-view for the classification purpose. They have gathered street-view images using the Google API and have manually labelled them into nine different categories of road damage. They have also collated the outcomes of the Faster R-CNN model with the YOLO-based model, which shows that the latter performs much better. However, they used a small dataset consisting of 7,237 images only. The advantage of their approach is the easy availability of Google street-view images.

The solution presented by Patra et al. [11] uses a CNN based model which achieves the task of pothole detection using a dataset consisting of images collected using the Google API. They have also compared their model with six other conventional models. Their suggested framework delivers a significant accuracy of around 97.6%. The area under the curve value (AUC) is also greater than that of those other models.

According to Goodfellow et al. [12], to achieve sufficiently good results, a dataset consisting of at least 5,000 images of each category must be used for the task of image classification. Although the techniques used for image collection using the Google API by the above authors have the advantage of the images being easily accessible, and free

of cost, this advantage comes with the tedious task of manually labeling each collected image which is very time-consuming.

Smartphones were used for collecting the images in a methodology developed by Maeda et al. [8] for the task of pavement distress detection. This study was a major breakthrough and served as a starting point for future studies. It categorized the road defects in Japan's road network into eight different classes. They developed a dataset which was called the Road Damage Dataset-2018, which was also made open-source for easy accessibility. They also developed a mobile application for instantaneous road damage evaluation. This application was adopted by multiple townships in Japan. The IEEE Big Data Conference, which was held in the USA in 2018, conducted a challenge named Big Data Cup [1], in which several teams submitted solutions to the problem of road damage type detection. Although the solutions provided were novel and better in terms of accuracy than the model proposed by Maeda et al. [8], the dataset used by these teams was the same as developed by [8].

The solution proposed by Alfarrarjeh et al. [13] uses Deep Learning methodologies for performing the task of finding the entity of interest in a given input image. They have used the YOLO object detection algorithm. The model was trained on a dataset that had different categories of damage as defined by the Japan Road Association.

[14] gave a solution that was again a part of the IEEE Big Data Cup challenge organized in 2018 in which they improved the accuracy of the model proposed by [8] by performing certain hyperparameter tuning. They also presented the approaches to enhancing the available dataset. Some modifications were also proposed to the dataset provided by [8], which contained 9,053 images.

The model proposed by Wang et al. [15] was again submitted as a solution to the IEEE Big Data Cup challenge-2018. They used Faster R-CNN and SSD algorithms for the task of object recognition. They used VGG-16 and ImageNet pre-trained ResNet-101 as the underlying basis of these object detection models. For better accuracy, they have used ensemble techniques. Wang et al. [16] have used Faster R-CNN for classifying the types of road damage, training their model on the dataset provided by [8]. The parameters were tuned based on the examination of the area of the location of the damage as well as the aspect ratio. A few data augmentation techniques were also applied before training.

Angulo et al. [17] augmented the dataset provided by [8] by the addition of more images gathered from Mexico and Italy. Their dataset contained 1,803,454 images collected using mobile phone devices. Various publicly available datasets supplemented with crowdsourced images are the sources of this huge dataset. The labelling of the dataset was done manually and the location, damage type, as well as the extent of damage were marked for each image. Both the conventional and deep learning-based models were trained on this dataset, and their performance was compared. The conventional model used was LBP-Cascaded Classifier, and the deep learning models used were RetinaNet and MobileNet.

More images were gathered from Italy by Roberts et al. [18] using the mobile application developed by [8]. They have classified the type of road damage as well as the intensity of the damage. Biçici and Zeybek [19] used UAV photogrammetry to generate point clouds for detecting road distress, which is considered to be a high-accuracy, efficient, automated method for detection. The SFM pipeline was used to generate a

high-density 3D model from UAV images. After that, a vertically based algorithm was used to get rid of unrelated parts of the environment in a 3D point cloud.

Zhang et al. [20] presented a CNN-centered methodology for correctly detecting pavement cracks, which is called CrackNet. It works on over one million parameters during the learning process and has over five layers. It was trained over multiple three-dimensional road images under various conditions and was tested successfully. Using a gradient-based optimization function for training to minimize the cost is the main objective. They have used Mini-batch Gradient Descent in their study. Cracks were detected at a pixel level by their proposed model. However, in this version of CrackNet, it requires considerable processing time and has a lot of difficulties in detecting hairline cracks.

Another study [21] uses a dataset of 500 images to categorize each image using supervised deep CNN for training purposes. The images were collected by a mobile phone. Their proposed model outperforms other models used in this problem domain. The Drop out methodology is used for reducing the overfitting of their proposed model by using a threshold value of 0.5. Silva and Lucena [22] are aimed at detecting cracks on concrete surfaces using a model centered upon machine learning that is intended to increase automation by combining with UAV (unmanned aerial vehicles). As a result, the transfer learning technique was employed. For the development of this model, the opensource model VGG16 was used.

Anand et al. [23] have proposed an approach to detect potholes and cracks using texture and spatial features centered upon Deep Neural Networks. Their proposed model uses the image texture to categorize damaged roads. They have combined two separate datasets, of which the first one consists of grayscale images collected using a mobile phone device and the second one consists of images collected using sensors. They have achieved very good accuracy over other models.

In [24], Fan et al. developed a technique for learning crack structure without altering unprocessed photos to analyze different pavement conditions. The model used two datasets, which consisted of images collected using a mobile phone device. The performance of their model was collated with different other models. Their approach in terms of dealing with different pavement textures shows a better performance. Without any preprocessing, the network is able to learn from the images that are unprocessed.

Zhu et al. [44] came up with a way to find potholes on roads by using three cutting-edge methods: Faster R-CNN, YOLOv3, and YOLOv4. These were trained on a set of data that was made by unmanned aerial vehicles. With an average accuracy of 56.62%, YOLOv3 was better at making predictions than the other methods.

With a cost-efficient video data collection technique, Zhang et al. [45] proposed a pavement distress detection technique using convolutional neural networks. In this work, the detectors are put into different groups. The F1-score for all kinds of flaws was compared, but the score went up when two kinds of cracks were combined into one.

Guerrieri et al. [46] used multiple public datasets to find pavement damage and created a new dataset for stone pavement damage. They then used deep learning and YOLOv3 to train a model to find pavement damage. Different parameters, such as loss, precision, recall, and RSME, are used to measure performance. For different kinds of flaws, the rate of finding them was between 91.0% and 97.3%.

Wen et al. [47] came up with a good deep learning model that was trained on both 2D and 3D images. The model's performance was judged based on a parameter called "interaction over union," which gives a detection accuracy of 83.7%.

Table 1 describes the databases utilized in various investigations. The sizes of these datasets are classified into three categories: less than (<) 1000 images; greater than (>) 1000; less than 5000 images; and greater than 5000 images. In Table 1, NK refers to Not Known, which means that the location from where the dataset was collected is unknown.

Table 1. Datasets available for road damage detection.

Dataset name	Collected from (Location)	Size of the dataset
CrackIT [25]	NK	<1000
CrackTree200 [31]	NK	<1000
SDNET2018 [30]	USA	<1000
Crack500 [33]	USA	<1000
GAPs v1 [3]	Germany	Between 1000–5000
GAPs v2 [32]	Germany	Between 1000–5000
Majidifard et al. [10]	USA	>5000
Maeda et al. [8]	Japan	>5000
Angulo et al. [17]	Italy, Japan, Mexico	>5000
RDD-2020 [35]	India, Japan, Czech Republic	>5000
Du et al. [9]	China	>5000

Table 2. Summary of the works related to the role of Deep Learning and Object detection algorithms for road damage detection.

Author Name	Techniques used	Dataset collected using	Collection procedure	Captured vision	Performance
Oliveira and Correia [25], 2017	KNN algorithm	Ocular equipment	–	Vertical	F score: 0.97, Precision: 0.98, Recall: 0.95
Shi et al. [26], 2016	Random Forest algorithm	Mobile Phone	Static	Vertical	F score: 0.939, Precision: 0.945, Recall: 0.936

(continued)

Table 2. (continued)

Author Name	Techniques used	Dataset collected using	Collection procedure	Captured vision	Performance
Majidifard et al. [10], 2020	Faster R-CNN and YOLOv2	Google API	–	Vertical and Wide	F score: 0.84, Precision: 0.93, Recall: 0.77
Maeda et al. [8], 2018	Single Shot detection algorithm, MobileNet	Mobile Phone	Mobile Application	Wide	Precision and recall greater than 71% and 77% in MobileNet and InceptionV2 respectively
Angulo et al. [17], 2019	RetinaNet	Mobile Phone	–	Wide	mAP: 0.91522
Cui et al. [27], 2015	Random Forest Algorithm	Mobile Phone	Static	Vertical	F score: 0.79, Precision: 0.78, Recall: 0.81
Mei et al. [28], 2020	GAN	GoPro camera device	–	Wide	F score: 0.92, Precision: 0.97, Recall: 0.88
Li et al. [29], 2014	Deep Learning	Mechanized pavement assessor	–	Vertical	Alligator, transverse, and longitudinal crack accuracy rates are 97.5 percent, 100 percent, and 88 percent, respectively
Eisenbach et al. [3], 2017	ASINVOSNet	High-tech Industrial Camera	Mobile Vehicle	Vertical	BER: 0.071, Accuracy: 94%, F score: 0.88
Dorafshan et al. [30], 2020	AlexNet Deep CNN	Camera	–	Vertical	Accuracy: 95.52% using transfer learning with 10 training epochs
Zou et al. [31], 2012	Minimum spanning Trees	Camera	–	Vertical	F score: 0.85, Precision: 0.79, Recall: 0.92

(continued)

Table 2. (continued)

Author Name	Techniques used	Dataset collected using	Collection procedure	Captured vision	Performance
Stricker et al. [32], 2019	ASINVOSNet, ResNet	High-tech Industrial Camera	Mobile Vehicle	Vertical	F score: 0.90, BER: 0.08, GME: 0.91
Yang et al. [33], 2020	Hierarchical Boosting Architecture	Mobile Phone	–	Vertical	AIU: 0.17, ODS: 0.68, OIS: 0.70
Weng et al. [34], 2019	Edge Identification and Morphological Functioning	-	–	Vertical	Reconstruction accuracy: 93.7% and the coefficient of correlation between breadth assessment findings and actual results is 0.97
Zhu et al. [44], 2021	Faster RCNN, YOLOv3 and YOLOv4	Unmanned Aerial Vehicle	–	Vertical	The prediction performance of YOLOv3 outperformed the other techniques with an average precision of 56.62%
Zhang et al. [45], 2021	CNN	GOPRO hero 7	–	–	The F1 score for the crack class increased when two classes of cracks were merged together and tested
Guerrieri et al. [46], 2022	YOLOv3	Digital Camera	–	–	The detection rate for several types of defects was in the range 91.0% to 97.3%

(continued)

Table 2. (continued)

Author Name	Techniques used	Dataset collected using	Collection procedure	Captured vision	Performance
Wen et al. [47], 2022	PDSNet	Camera and lasers	–	–	The model gives a detection accuracy of 83.7%

4 Challenges in This Domain and Contributions of Our Study

In this arena, problems include road distress detection and comparing existing solutions. Effective damage estimation, including more cracks and defect types for detection, installing and mounting developed systems on vehicles, performing efficiently on road images in different lighting conditions, complex crack topology, and including more images of roads from various countries remain challenges in this area. Low contrast between damages and surrounds, inhomogeneity of concentration along fissures, potential reflections identical to cracks, and real-time distress identification also offer obstacles in developing sophisticated pavement fault detection systems. Traditional, semi-automated, and completely autonomous operations are ineffective in terms of cost, time, and labor. The organizations responsible for maintaining decent road conditions still struggle to cope with this problem due to a lack of finances for pricey image-capturing devices and a unified dataset that is not limited to photographs from certain nations and damages in those countries.

This paper reviews studies on cost-effective road defect identification. Several state-of-the-art investigations on this issue domain included using deep learning architectures. These strategies must be explored to reduce road fatalities and preserve their condition because they automate pavement inspection, are cost-effective, need less human intervention, and save time. We’ve also summarized relevant datasets. Future researchers can utilize this information to construct a standard dataset to compare model results. This can also help government agencies employ a holistic database for detection. This study will give scholars a rundown of previous works and their shortcomings. Future researchers can expand on our extensive work. We’ve listed the authors’ datasets, which will be beneficial for future research in this area.

5 Conclusion and Future Scope

This paper has explored the role of various techniques that use advanced models to detect damage on roads using images collected at various geographical locations. The techniques using state-of-the-art object detection models fared better than other approaches in our analysis. Although there are several ways to perform this task using advanced devices, these techniques are not cost-effective as many countries lack the finance required to

implement such sophisticated systems. Therefore, to overcome the above constraint, researchers have come up with alternative cost-effective strategies that mostly deal with the data collection phase. As the problem of road accidents is universal in nature, the dataset used must be made more inclusive with images from the world over. A video input would also help to improve the functionality of existing techniques. To expand the coverage of a model, other sensor data from mobile phones can be integrated. We have surveyed these approaches along with their advantages, which will be handy for any future research in this field.

References

1. Road Accidents in India: Ministry of Road Transport and Highways, Transport Research Wing, Govt. of India (2018). https://morth.nic.in/sites/default/files/Road_Accidednt.pdf
2. McGhee, K.: Automated Pavement Distress Collection Techniques, National Cooperative Highway Research Program, NCHRP Synthesis 334 Report, Transportation Research Board, Washington DC (2004). <https://doi.org/10.17226/23348>
3. Eisenbach, M.: Gross, How to get pavement distress detection ready for deep learning? A systematic approach. In: International Joint Conference on Neural Networks (IJCNN), pp. 2039–2047 (2017). <https://doi.org/10.1109/IJCNN.2017.7966101>
4. Zhang, D., et al.: Automatic pavement defect detection using 3D laser profiling technology. *Autom. Constr.* **96**, 350–365 (2018). <https://doi.org/10.1016/j.autcon.2018.09.019>
5. Guan, J., Yang, X., Ding, L., Cheng, X., Lee, V.C., Jin, C.: Automated pixel-level pavement distress detection based on stereo vision and deep learning. *Autom. Constr.* **129**, 103788 (2021). <https://doi.org/10.1016/j.autcon.2021.103788>
6. Mertz, C., Varadharajan, S., Jose, S., Sharma, K., Wander, L., Wang, J.: City-wide road distress monitoring with smartphones. In: Proceedings of ITS World Congress, pp. 1–9 (2014). https://www.ri.cmu.edu/pub_files/2014/9/road_monitor_mertz_final.pdf. Accessed 28 July 2021
7. Avellaneda, D.A.C., López-Parra, J.F.: Detection and localization of potholes in roadways using smartphones. *DYNA* **83**(195), 156–162 (2016). <https://doi.org/10.15446/dyna.v83n195.44919>
8. Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., Omata, H.: Road damage detection and classification using deep neural networks with smartphone images. *Comput. Aided Civ. Infrastruct. Eng.* **33**, 1127–1141 (2018). <https://doi.org/10.1111/mice.12387>
9. Yuchuan, D., Pan, N., Zihao, X., Fuwen Deng, Y., Shen, H.K.: Pavement distress detection and classification based on YOLO network. *Int. J. Pavement Eng.* **22**(13), 1659–1672 (2020). <https://doi.org/10.1080/10298436.2020.1714047>
10. Majidifard, H., Jin, P., Adu-Gyamfi, Y., Buttlar, W.G.: Pavement image datasets: a new benchmark dataset to classify and densify pavement distresses. *Transp. Res. Rec.* **2674**, 328–339 (2020). <https://doi.org/10.1177/0361198120907283>
11. Patra, S., Middya, A.I., Roy, S.: PotSpot: Participatory sensing based monitoring system for pothole detection using deep learning. *Multimedia Tools Appl.* **80**(16), 25171–25195 (2021). <https://doi.org/10.1007/s11042-021-10874-4>
12. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT press, (2016). <http://www.deeplearningbook.org>
13. Alfarrarjeh, A., Trivedi, D., Kim, S.H., Shahabi, C.: A deep learning approach for road damage detection from smartphone images. In: 2018 IEEE International Conference on Big Data (Big Data), pp. 5201–5204. IEEE (2018). <https://doi.org/10.1109/BigData.2018.8621899>

14. Kluger, F., et al.: Region-based cycle-consistent data augmentation for object detection. In: 2018 IEEE International Conference on Big Data (Big Data), , pp. 5205–5211. IEEE (2018). <https://doi.org/10.1109/BigData.2018.8622318>
15. Wang, Y.J., Ding, M., Kan, S., Zhang, S., Lu, C.: Deep proposal and detection networks for road damage detection and classification. In: 2018 IEEE International Conference on Big Data (Big Data), pp. 5224–5227. IEEE (2018). <https://doi.org/10.1109/BigData.2018.8622599>
16. Wang, W., Wu, B., Yang, S., Wang, Z.: Road damage detection and classification with faster R-CNN. In: 2018 IEEE International Conference on Big Data (Big Data), pp. 5220–5223. IEEE (2018). <https://doi.org/10.1109/BigData.2018.8622354>
17. Angulo, A., Vega-Fernández, J.A., Aguilar-Lobo, L.M., Natraj, S., Ochoa-Ruiz, G.: Road damage detection acquisition system based on deep neural networks for physical asset management. In: Martínez-Villaseñor, L., Batyrshin, I., Marín-Hernández, A. (eds.) MICAI 2019. LNCS (LNAI), vol. 11835, pp. 3–14. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33749-0_1
18. Roberts, R., Giancontieri, G., Inzerillo, L., Di Mino, G.: Towards low-cost pavement condition health monitoring and analysis using deep learning. *Appl. Sci.* **10**, 319 (2020). <https://doi.org/10.3390/app10010319>
19. Biçici, S., Zeybek, M.: An approach for the automated extraction of road surface distress from a UAV-derived point cloud. *Autom. Constr.* **122**, 103475 (2021). <https://doi.org/10.1016/j.autcon.2020.103475>
20. Zhang, A., et al.: Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network. *Comput. Aided Civ. Infrastruct. Eng.* **32**, 805–819 (2017). <https://doi.org/10.1111/mice.12297>
21. Zhang, L., Yang, F., Zhang, Y.D., Zhu, Y.J.: Road crack detection using deep convolutional neural network. In: 2016 IEEE International Conference on Image Processing (ICIP), pp. 3708–3712. IEEE (2016). <https://doi.org/10.1109/ICIP.2016.7533052>
22. Silva, W.R.L.D., Lucena, D.S.D.: Concrete cracks detection based on deep learning image classification. In: Multidisciplinary Digital Publishing Institute Proceedings, vol. 2, p. 489 (2018). <https://doi.org/10.3390/ICEM18-05387>
23. Anand, S., Gupta, S., Darbari, V., Kohli, S.: Crack-pot: autonomous road crack and pothole detection. In: 2018 Digital Image Computing: Techniques and Applications (DICTA), pp. 1–6. IEEE (2018). <https://doi.org/10.1109/DICTA.2018.8615819>
24. Fan, Z., Wu, Y., Lu, J., Li, W.: Automatic Pavement Crack Detection Based on Structured Prediction with the Convolutional Neural Network. *arXiv preprint arXiv:1802.02208* (2018)
25. Oliveira, H., Correia, P.L.: CrackIT—an image processing toolbox for crack detection and characterization. In: 2014 IEEE International Conference on Image Processing (ICIP), pp. 798–802. IEEE (2014). <https://doi.org/10.1109/ICIP.2014.7025160>
26. Shi, Y., Cui, L., Qi, Z., Meng, F., Chen, Z.: Automatic road crack detection using random structured forests. *IEEE Trans. Transp. Syst.* **17**, 3434–3445 (2016). <https://doi.org/10.1109/TITS.2016.2552248>
27. Cui, L., Qi, Z., Chen, Z., Meng, F., Shi, Y.: Pavement distress detection using random decision forests. In: Zhang, C., et al. (eds.) ICDS 2015. LNCS, vol. 9208, pp. 95–102. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24474-7_14
28. Mei, Q., Gül, M.: A cost effective solution for pavement crack inspection using cameras and deep neural networks. *Constr. Build. Mater.* **256**, 119397 (2020). <https://doi.org/10.1016/j.conbuildmat.2020.119397>
29. Li, L., Sun, L., Ning, G., Tan, S.: Automatic pavement crack recognition based on BP neural network. *PROMET-Traffic Transp.* **26**, 11–22 (2014). <https://doi.org/10.7307/ptt.v26i1.1477>
30. Dorafshan, S., Thomas, R.J., Maguire, M.: SDNET2018: an annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks. *Data Brief.* **21**, 1664–1668 (2018). <https://doi.org/10.1016/j.dib.2018.11.015>

31. Qin Zou, Y., Cao, Q.L., Mao, Q., Wang, S.: CrackTree: automatic crack detection from pavement images. *Pattern Recogn Let* **33**(3), 227–238 (2012). <https://doi.org/10.1016/j.patrec.2011.11.004>
32. Stricker, R., Eisenbach, M., Sesselmann, M., Debes, K., Gross, H.-M.: Improving visual road condition assessment by extensive experiments on the extended gaps dataset. In: *International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8 (2019). <https://doi.org/10.1109/IJCNN.2019.8852257>
33. Yang, F., Zhang, L., Yu, S., Prokhorov, D., Mei, X., Ling, H.: Feature pyramid and hierarchical boosting network for pavement crack detection. *IEEE Trans. Intell. Transp. Syst.* **21**, 1525–1535 (2020). <https://doi.org/10.1109/TITS.2019.2910595>
34. Weng, X., Huang, Y., Wang, W.: Segment-based pavement crack quantification. *Autom. Constr.* **105**, 102819 (2019). <https://doi.org/10.1016/j.autcon.2019.04.014>
35. Arya, D., et al.: Deep learning-based road damage detection and classification for multiple countries. *Autom. Construct.* **132**, 103935 (2021). ISSN 0926-5805, <https://doi.org/10.1016/j.autcon.2021.103935>
36. Hatmoko, J., Setiadji, B., Wibowo, M.: Investigating causal factors of road damage: a case study. *MATEC Web Conf.* **258**, 02007 (2019). <https://doi.org/10.1051/mateconf/201925802007>
37. Ross, B.G., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587 (2014)
38. Ross, B.G.: Fast R-CNN. In: *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1440–1448 (2015)
39. Ren, S., He, K., Girshick, R.B., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 1137–1149 (2015)
40. Joseph, R., Divvala, S.K., Girshick R.B., Farhadi, A.: You only look once: unified, real-time object detection. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788 (2016)
41. Liu, W., et al.: SSD: single shot multibox detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) *ECCV 2016. LNCS*, vol. 9905, pp. 21–37. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_2
42. Schmidhuber, J.: Deep learning in neural networks: an overview. *Neural Netw.* **61**, 85–117 (2015)
43. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) *Computer Vision – ECCV 2014. ECCV 2014. LNCS*, vol. 8689. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10590-1_53
44. Zhu, J., Zhong, J., Ma, T., Huang, X., Zhang, W., Zhou, Y.: Pavement distress detection using convolutional neural networks with images captured via UAV. *Autom. Constr.* **133**, 103991 (2022). <https://doi.org/10.1016/j.autcon.2021.103991>
45. Zhang, C., Nateghinia, E., Miranda-Moreno, L.F., Sun, L.: Pavement distress detection using convolutional neural network (CNN): a case study in Montreal, Canada. *Int. J. Transp. Sci. Technol.* **11**(2), 298–309 (2022). <https://doi.org/10.1016/j.ijtst.2021.04.008>
46. Guerrieri, M., Parla, G.: Flexible and stone pavements distress detection and measurement by deep learning and low-cost detection devices. *Eng. Failure Anal.* **141**, 106714 (2022). <https://doi.org/10.1016/j.engfailanal.2022.106714>
47. Wen, T., et al.: Automated pavement distress segmentation on asphalt surfaces using a deep learning network. *Int. J. Pavem. Eng.* 1–14 (2022). <https://doi.org/10.1080/10298436.2022.2027414>