International Journal of Communication and Information Technology

E-ISSN: 2707-6628
P-ISSN: 2707-661X
www.computersciencejournals.com/ijcit
IJCIT 2024; 5(2): 11-16
Received: 06-05-2024
Accepted: 11-06-2024

U Satchithanantham
Assistant Professor, Department of Computer Science with Data Analytics Ajk College of Arts and Science, Coimbatore Tamil Nadu, India

Corresponding Author:
U Satchithanantham
Assistant Professor, Department of Computer Science with Data Analytics Ajk College of Arts and Science, Coimbatore Tamil Nadu, India

Advanced pothole detection using nural network model-VGG16

U Satchithanantham

DOI: https://doi.org/10.33545/2707661X.2024.v5.i2a.86

Abstract
Potholes and unevenness in road surfaces tend to be accident-prone. Potholes disrupt the flow of traffic as drivers navigate around them cautiously for the reason that they pose hazards to drivers and pedestrians. Addressing these hazards promptly and efficiently is crucial to ensure the safety of road users. Therefore, periodic maintenance of roads should be made to keep roads safe and sustainable for the people. However, manual road inspection leads to several challenges such as vast road networks, limited manpower, outdated equipment and budget constraints.

Eventually, automating the process of pothole detection is suggested to improve quick road maintenance and to avoid accidents. Automation provides a faster assessment of road conditions compared to manual inspections and prevents major repairs through early detection.

Ultimately, this research work aims to improve the detection process using deep learning technique by refining the VGG16 convolutional neural network model. Modifications are made to the VGG16 network by removing some convolution layers and using different filter sizes. Additionally, this paper utilizes the YOLOv8 model for comparison of results with those obtained from VGG16. The models are trained using a pothole dataset taken from kaggle, which includes both normal road conditions and images containing potholes. The paper analyses the performance of pre-trained VGG16, VGG16 architecture trained without using pre-existing weights, modified VGG16 and YOLOv8 models on the dataset. This work compares the performance of models by evaluating accuracy, model size and inference time. The focus of this study is to achieve a balance between accuracy and speed in pothole detection, thereby enhancing road maintenance processes.

Keywords: Pre-trained, model, pedestrians, computation

Introduction
Roads have been a major factor in getting from one place to another. They connect different places for people to move to places without any trouble and enable vehicles to traverse roads for commuting, transporting goods, commercial activities, and personal travel. A well-connected road network has significantly boosted trade and commerce in India [1].

As modern societies undergo continuous development and expansion, the necessity for efficient road and bridge infrastructure intensifies [2]. However, the continuous use of roads, combined with environmental factors such as weathering, traffic volume, leads to wear and tear over time. Regular maintenance of road is required to address deteriorations such as pavements, drains, shoulders, and edges is vital for preserving the integrity of roads and providing superior service to road users. Potholes, in particular, presents a major safety hazard to road users.

Potholes are areas of road surface that have cracked, eroded that eventually form a hole. These depressions in the road surface are a common issue in road infrastructure and tend to pose safety risk to drivers and pedestrians. Potholes cause vehicles to lose control, leading to accidents, especially at higher speeds or in heavy traffic areas. Moreover, potholes can cause significant damage to vehicles.

Manual road inspection on busy roads presents significant challenges and has limitations in terms of efficiency, accuracy, and coverage compared to automated or technology-driven approaches.

Therefore, there is a need to automate the process of detecting potholes with improved speed [3]. Various approaches such as vibration-technique, 3D reconstruction technique approach, stereo vision method, and Vision based approach facilitate pothole detection. Recent studies present automated pothole detection with advancements such as deep learning (DL) and machine learning (ML) techniques.

The two distinct approaches to object detection are one-stage and two-stage detectors. One-stage detectors, the YOLO series requires only a single pass through the neural network to directly locate and classify objects. That is much faster but less accurate in detection process. Two-stage detectors, like Faster-RCNN, utilize a CNN to extract features and use those features to locate and classify objects. They are more accurate but are slower than one-stage detectors.

For real-time application of pothole detection, it is necessary to address the trade-off between accuracy and detection speed. The objective of this study is to obtain highest accuracy in less inference time and less computation. As such, this research analyses the feasibility and accuracy of refined VGG16 along with pre- trained VGG16, VGG16 architecture trained without using pre-existing weights. This paper presents the architecture of VGG16 with adjustments. The effectiveness of proposed model is compared in terms of accuracy, inference time, size of the model with YOLOv8 and VGG16.

Literature Review

Various studies have been conducted to address the challenge of pothole detection and classification accurately with significant speed, utilizing a range of deep learning models and techniques [4], present an integrated approach for pothole detection and classification using YOLOv7, SVM, and segmentation techniques. This study utilizes various detection techniques such as segmentation to highlight the pothole area, canny edge detection algorithm to highlight the road's edges, YOLOv7 model and SVM to classify potholes. In [5], researcher proposed YOLOv8, a deep learning model capable of real-time object detection with high accuracy. This study achieves average accuracy of 56.8%. The model is suggested to be applied in practical settings with suitable modifications. In the study [6], the researchers Projects Convolutional Neural Network (CNN) and YOLOV5 for pothole detection. YOLOv5 achieves an accuracy of 95% which is more than another model. In [7], a comparison for selection of object detection algorithm has been done in this paper. Several object detection algorithms such as HOG, R-CNN, Faster R-CNN, SSD, YOLO are discussed. The paper presents YOLO v7 for pothole detection, which is suitable for real-time scenario. In [8], the researchers focused on developing a road damage detection system that utilizes a CNN- based image and video processing model that trigger alerts to notify relevant authorities. This study aims to assess various object detection methods and develop a generalizable model for road damage detection. In [9], the researcher developed a system aims to recognize potholes on muddy roads and highway roads pictures. It utilizes three pre-trained models Resnet50, InceptionV2, VGG19 and analyses their performance. In [10], the focus was on developing a system for detecting potholes that assists drivers in avoiding potholes on the road using YOLOV7 and CNN. This paper presents a comparative analysis of two deep learning- based algorithms, CNN and YOLOv7. In [11], the researchers focused on developing deep learning-based model to detect potholes is developed using F-RCNN and Inception-V2. In [12], the proposed system utilizes Faster Region-based Convolutional Neural Network (F-RCNN) and You Only Look Once Version 3(YOLO V3). It provides a comparative analysis of different deep learning algorithms for their performance based on accuracy. In [13], the study aims locate the object and not just detect it. The paper uses SSD for object detection. In [14], the paper presents an approach of generalized learning model for pothole detection under various environment conditions. Utilizing faster-RCNN, the model predicts if the input image is of a pothole or non-pothole. In [15], goal is to train and analyse this YOLOX algorithm pothole detection. The results obtained are analysed using compared to other models. In [16], the researchers developed a pothole detection system that detects potholes using Transfer Learning technique provided in VGG16. In [17], the researchers focused on developing a generalized learning model for detecting potholes captured in different illumination conditions, with different camera types, camera angles and resolutions. This study utilizes Faster R-CNN object detection model with Inception-V2 architecture as a backbone, and the trained model is applied on 4 datasets. In [18], the paper proposes a real-time pothole detection system using Deep Learning. A comparison deep learning algorithms SSD Tensor Flow, YOLOv3-Darknet53, and YOLOv4-CSP Darknet 53 were made. In [3], A Modified VGG16 (MVGG16) developed to reduce the computational cost and improve the training results. Other algorithms Faster R- CNN with ResNet 50 (FPN), VGG16, MobileNetV2, InceptionV3, and MVGG16 backbones, YOLOv5 (Large (Yl), Medium (Ym), and Small (Ys) are used. In [19], proposed a mechanism for automatic detection of road anomalies by autonomous vehicles and providing road information to upcoming vehicles based on Edge AI and VANET. The techniques used in this study are Residual Convolutional Neural Network (ResNet-18) and Visual Geometry Group (VGG-11). In [20], the focus is to develop a system to detect potholes in real-time with adequate accuracy and speed, and that can be implemented with ease and low setup cost. It compares the performance of YOLOv5 Large (Yl), YOLOv5 Medium (Ym) and YOLOv5 Small (Ys) with ResNet101 backbone and Faster R-CNN with ResNet 50 (FPN), VGG16 and MobileNetV2 backbone. In [21], proposed an android application for detecting potholes in real time. It uses Transfer Learning to train Deep Learning based Single Shot Detector (SSD)model for object detection to detect potholes of all shapes with good accuracy. In [22], Faster R-CNN, SSD and YOLOv3 were trained on the custom dataset containing images of potholes for real-time pothole detection system.

Materials and Method

Experimental Setup

Although the training machine runs on Windows 10 Pro and is equipped with an Intel(R) Core i7 CPU, leveraging Google Colab's T_4 GPU can significantly accelerate deep learning model training. Google Colab provides a cloud-based platform with access to powerful GPUs like the T_4, which are optimized for parallel processing tasks like training neural networks. Packages of python TensorFlow, visualkeras were installed.

Dataset preparation

The two-datasets named Pothole Detection Dataset [23] and Pothole Detection [24] sourced from Kaggle are combined to have a total of 1064 images. These datasets hold both pothole images and normal images. Each pothole image depicts scenes with varying road conditions, weather, and pothole shapes. The dataset is structured with labelled images, indicating the presence or absence of potholes within each scene. The images are in the formats of JPEG or PNG. A sample image of pothole is as shown in figure 1.

Fig 1: Sample Pothole image

Data augmentation

Data augmentation is done to increase the diversity of existing training data [25]. The augmentation techniques zoom, rotation and horizontal flip were applied to the training data.

Deep learning techniques

Deep learning is a method that trains computers to process information in a way that mimics human brain [26]. Deep learning models are files trained to recognize more complicated patterns in text, images, or sounds [26]. Deep learning models make use of several algorithms such as Convolutional Neural Networks (CNNs), Long Short Term Memory Networks (LSTMs) and Recurrent Neural Networks (RNNs) [26]. VGG16, Visual Geometry Group 16 is a convolutional neural network (CNN) architecture popular for its effectiveness in image classification tasks.
This section describes the methods of pothole detection employed in this study.

VGG16

VGG16 convolutional neural network model consists of 16 layers and organized into blocks, where each block consists of multiple convolutional layers followed by a pooling layer. The VGG-16 consists of 13 convolutional layers and three fully connected layers [27]. VGG16 comprises five blocks of convolutional layers, with additional fully connected layers at the end. The model receives an input of image with dimension (224, 224, 3). Convolution layer in block 1 has 64 number of filters, Convolution layer in block 2 has 128 filters, Convolution layer in block 3 has 256 filters, Convolution layer in block 4 has and 5 has 512 filters. The convolution layers are of 3x3 filter with stride 1 and used the same padding and maxpool layer of 2x2 filter of stride 2. There are two dense layers with 4096 units and last dense layer with 1000 units (for imagenet dataset). The VGG16 architecture is as shown in Figure 2 [27].

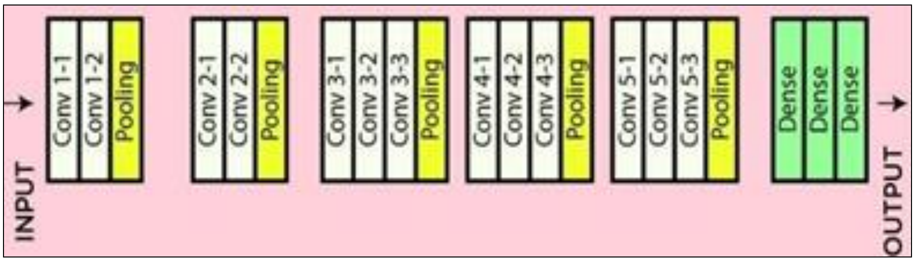

Fig 2: VGG16 architecture

VGG16 pre-trained model

The weights of the pre-trained VGG16 are being reused here. The weights of layers are frozen to prevent them being updated. Only the dense layers are being fine-tuned. The dense layers are followed by an average pooling layer. The dense layer has been substituted with two new dense layers, one layer with 128 units and another with 2 units. The VGG16 with pre- trained weights is trained twice: once with images of resolution 128 x 128 and again with images of resolution 224 x 224.

Building VGG16 architecture

The VGG16 model architecture consists of 16 layers composed of convolutional, pooling, activation, dropout, and fully connected layers. The layers, filters, kernels, strides, and padding used are the same as in VGG16 architecture except for the output layer. The output layer of vgg16 architecture has been modified to consist of 2 units with a softmax activation function. The VGG16 model architecture is trained twice: once with images of resolution 128 x 128 and again with images of resolution 224 x 224.

Modified VGG16

The proposed VGG16 has 5 blocks including 8 convolutional layers and pooling layers. It has a total of 35, 74, 690 trainable parameters. The first convolution layer takes input of size 128 x
128. All convolution layers use 3 × 3 kernel size and dilation rate of 3 x 3. The filter sizes are 32, 64, 64,128, 512 for each blocks respectively. A global max-polling layer follows the 5th layer and finally two dense layers of one with 128 units, relu activation function and another with
2 units, Soft Max activation function. Ultimately, the proposed VGG16, removes one convolution layer from blocks of 3, 4, 5. The proposed VGG16 architecture is as shown in Figure 3.

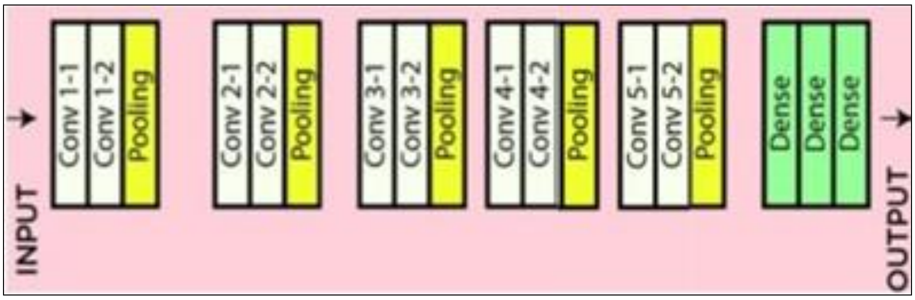

Fig 3: Modified VGG16 architecture

In order to perform compilation, all the models described above use the Adam optimizer with a learning rate of 0.0001. The pre-processed training data are fed to the model and the learning process takes place for 15 epochs. The values of training accuracy, training loss, validation accuracy, validation loss are visualized to identify signs of over fitting and under fitting. The accuracy and loss plots are shown in figure 5 and figure 6. Finally, the test data has been used to evaluate the model's performance. The proposed VGG16 is trained with images of resolution 128 x 128.

YOLOv8

YOLOv8 is the latest version of the YOLO architecture for has been in real-time applications for its speed in detection process. The pre-trained weights of yolov8n.pt is downloaded from the official website ultralytics [28]. The dataset images and annotations paths and configurations are set in a yaml file. The model trains on the annotated dataset for 200 epochs with default configurations of YOLOv8.

Results and Discussion

The metrics used for evaluating models performance are Precision, Recall and F1- score. F1 score is the harmonic mean of the precision and recall scores [29]. F1 score is given in Eq.1.

$$\text{F1 Score} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The comparison of results achieved by the VGG16 with pre-trained weights, built VGG16, Modified VGG16 are as shown in Table 1. Evaluation metrics of YOLOv8 is as shown in Table 2. From the experiment result, it is observed that the smallest model of size 13.64 MB is achieved in Modified VGG16 followed by YOLOv8 which is 22.5 MB. The highest training accuracy and least training loss is achieved in pre-trained VGG16 followed by Modified VGG16. The model with highest accuracy is obtained from pre-trained VGG16 model trained on images of resolution 224 x 224. The training accuracy of VGG16 with pre-trained weights, built VGG16, Modified VGG16 is shown in figure 4, and the loss is as shown in figure 5. Figure 6 shows the Classification loss graph of YOLOv8. The highest precision and recall values were obtained in pre-trained VGG16 trained on image of resolution 224 x 224. The Modified VGG16 model shows better detection confidence than other models. The highest F1- Score is obtained by VGG16 from pre-trained weights trained on image of resolution 224 x 224 which is 97%. It is followed by model trained on image of resolution 128 x 128 with F1 score of 96%. The Modified VGG16 and VGG16 architecture has F1-score of 92%. YOLOv8 produced F1-Score of 79%. While comparing the F1-Scores, Modified VGG16 value is 4% less than the highest value as obtained in Pre-trained VGG16, indicating that there is not much difference between the two. However, for YOLOv8, the variation is 13% less than that of the modified VGG16, indicating a more significant difference compared to Modified VGG16.

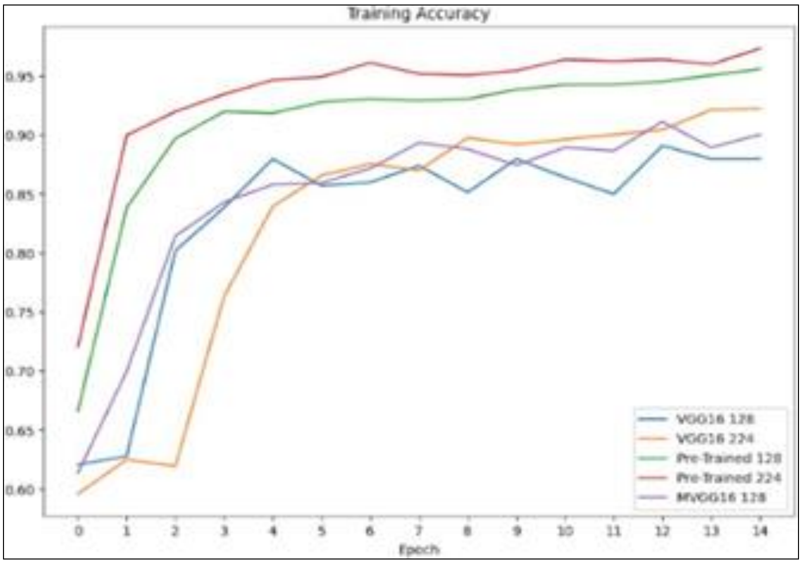

Fig 4: Comparison of training accuracy of VGG16 with pre-trained weights, Built VGG16, Modified VGG16

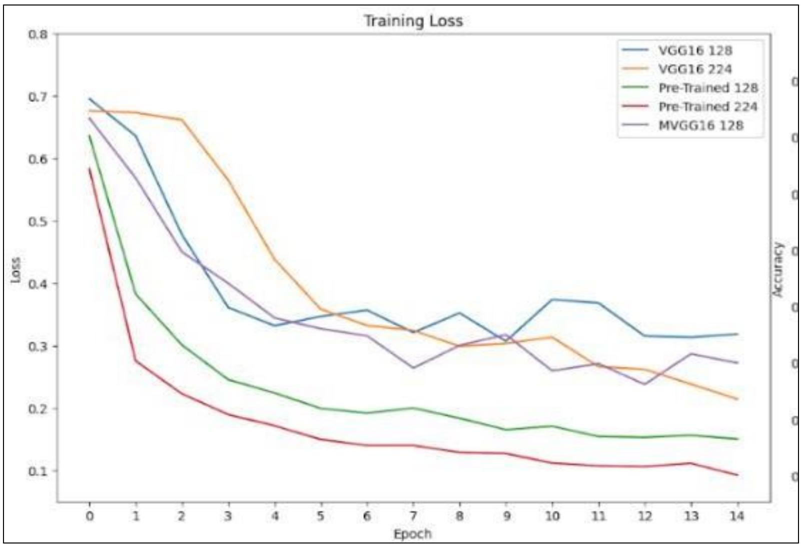

Fig 5: Comparison of training loss values of VGG16 with pre-trained weights, Built VGG16 and Modified VGG16

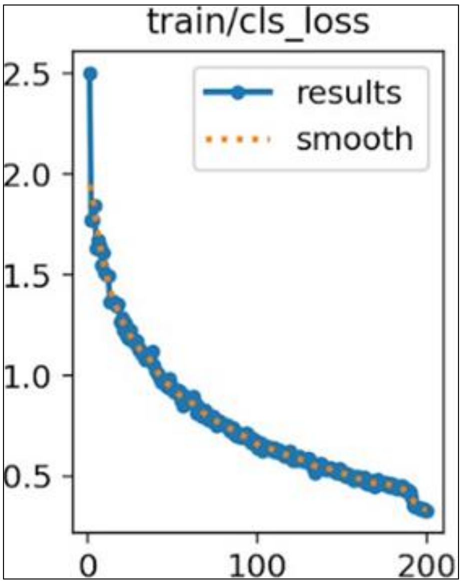

Fig 6: Classification loss of YOLOv8

Conclusion and future work: The proposed Modified VGG16 performed well when compared to the YOLOv8 model. The proposed model predicts potholes with high confident score. The F1-Score for Modified VGG16 was 4% less than the highest value obtained in Pre-trained VGG16, indicating the minimal difference between the two models. Furthermore, the F1-score of Modified VGG16 was 13% higher than YOLOv8, indicating a more significant difference between the two models.

In future analyses, the parameters for comparison can be further refined with additional metrics such as inference time and computational resource utilization. The pothole detection process can be enhanced by training the deep learning model with pothole images captured at various angles various size and weather conditions. The accuracy of models can be increased by training the models with large dataset. Techniques to estimate depth and severity of pothole can be employed.

Table 1: Comparison of models' performance

Model	Training image resolution	Model size	Parameters	Training loss	Test accuracy	Test loss	Average Inference time (milliseconds)	Precision (P)	Recall (R)	F1-score
VGG 16	128 x	57.13	1497721	0.1505	95.57	0.11	417	95%	97%	96%
pre-	128	MB	8		%					
Trained weights	224 x 224	58.38 MB	15304898	0.0929	97.47%	0.09	430	97%	98%	97%
Built	128 x	248.19	6506272	0.3071	89.87	0.40	396	88%	96%	92%
VGG16	128	MB	2		%					
architect ure	224 x 224	512.19 MB	134268738	0.2761	87.97%	0.29	310	88%	96%	92%
Propose	128 x	13.64	3574690	0.2380	89.87	0.22	392	92%	94%	92%
d	128	MB			%					
VGG16										

Table 2: Performance metrics of YOLOv8

Model	Training image resolution	Model size	Training loss	Average Inference time (milliseconds)	Precision (P)	Recall (R)	F1-score
YOLOv8	640 x 640	22.5M B	0.3267	48.8	88%	72%	79%

References

1. India's record road construction gives fillip to its economic growth. Organiser. July 16, 2023. Available from: https://organiser.org/2023/07/16/183938/bharat/indias-record-road-construction-gives-fillip-to-its-economic-growth/
2. The importance of structural integrity in road and bridge maintenance. Utilities One. Available from: https://utilitiesone.com/the-importance-of-structural-integrity-in-road-and-bridge-maintenance
3. Prakash M, Sriharipriya KC. Enhanced pothole detection system using YOLOX algorithm; c2022.
4. Madhvi M, Kondilkar R, Bhakare D, Kaslikar B. Pothole detection using ML and DL algorithms. 2023.
5. Doan Van D. Application of advanced deep convolutional neural networks for the recognition of road surface anomalies; c2023.
6. Upadhyay A, Mishra B, Singh A, Pal A. Comparative study of pothole recognition using CNN and YOLOV5; c2023.
7. Goswami K, Chattopadhyay S, Kundu A. A deep learning based approach to detect potholes using YOLO version 7. 2023.
8. Nalawade K, Bhujbal S, Lahase C, Mhasade S, Tambe K. Smart road damage detection and alarms. 2023.
9. Saisree C, Kumaran U. Pothole detection using deep learning classification method; c2023.
10. Lincy A, Dhanarajan G, Kumar S, Gobinath B. Road pothole detection system; c2023. Reddy KK, Bhavana T, Tejaswini B, Mallesh P, Sharath Chandra A. A deep learning approach for pothole detection; c2023.
11. Seetha M, Prasanna K. Intelligent deep learning based pothole detection and alerting system; c2023.
12. Mishra C, Patel U, Pratap Singh U, Kumar V. Identification of potholes and speed bumps using SSD algorithm; c2022.
13. Chavan P, Mane D, Chavan D. Pothole detection using deep learning; c2022.
14. Mohd Rum SN, Rajaratinam N. A pothole detection using VGG16; c2022.
15. Hassan SI, O'Sullivan D, Mckeever S. Pothole detection under diverse conditions using object detection; c2021.
16. Al-Shaghouri A, Alkhatib R, Berjaoui S. Real-time pothole detection using deep learning; c2021.
17. Ahmed KR. Smart pothole detection using deep learning based on dilated convolution; 2021.
18. Bibi R, Saeed Y, Zeb A, Ghazal TM, Rahman T, Said RA, Abbas S, Ahmad M, Khan MA. Edge AI-based automated detection and classification of road anomalies in VANET using deep learning; c2021.
19. Ahmed KR, Kharel S. Potholes detection using deep learning and area estimation using image processing; c2021.
20. Jain D, Gavali A, Kamble S, Pande A. Spothole - A real-time object detection application using modern deep learning technique; c2021.
21. Gajjar K, van Niekerk T, Wilm T, Mercorelli P. Vision-based deep learning algorithm for detecting potholes; c2021.
22. Pothole detection dataset [dataset]. Available from: https://www.kaggle.com/datasets/atulyakumar98/pothole-detection-dataset
23. Pothole detection [dataset]. Available from: https://www.kaggle.com/datasets/andrewmvd/pothole-detection
24. Data augmentation. Available from: https://research.aimultiple.com/data-augmentation/
25. What is deep learning? Coursera. Available from: https://www.coursera.org/articles/what-is-deep-learning
26. VGG-16 CNN model. Available from: https://www.geeksforgeeks.org/vgg-16-cnn-model/
27. Ultralytics YOLOv8 documentation. Available from: https://docs.ultralytics.com/8
28. F1 score guide. V7 Labs. Available from: https://www.v7labs.com/blog/f1-score-guide