Pothole Detection Approach Based on Deep Learning Algorithms

Y. Aneesh Chowdary, V. Sai Teja, V. Vamsi Krishna, N. Venkaiah Naidu, and R. Karthika

1 Introduction

Potholes on streets are noticeable constructional catastrophes within the plane of a street, caused due to the activity and the awful climate. It has presently been to be one of the major defects of our Indian roads. Potholes not only make streets seem unattractive but also create a risk to the security of individuals traveling on streets. Potholes are a boon to vehicles, causing significant injuries to anyone who is caught in them. They are hazardous for drivers, together with the people walking on footpaths, bicyclists, and street specialists as well. Potholes may cause injury to everyone who travels on the street.

As of the most recent estimates from a few state administrations, potholes cause roughly 30 deaths per day on the streets. In comparison to 2016, the mortality rate in 2017 climbed by more than 50%., i.e., 3597 [1] instances every year. It is critical to correctly detect and fix potholes in order to reduce the number of accidents as well as other losses. However, manually detecting potholes is not recommended because it is expensive and time-consuming. As a result, extensive research was conducted in order to develop a system that can detect potholes, which would have been a significant step forward in increasing the efficacy of survey, and the quality of pavement through inspection, observation, and quick reaction.

Human error is one cause for road accidents, since drivers are unable to spot potholes and make rash judgments. To combat it, the advanced driver assistance system (ADAS) assists drivers in spotting possible hazards and dangers ahead of time. ADAS is also in charge of managing the vehicle's control, balance, and movement in dangerous situations. Through a careful interface, ADAS has increased automobile

Y. A. Chowdary · V. S. Teja · V. V. Krishna · N. V. Naidu · R. Karthika (✉)
Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: r_karthika@cb.amrita.edu

V. S. Reddy et al. (eds.), *Soft Computing and Signal Processing*,
Smart Innovation, Systems and Technologies 313,
https://doi.org/10.1007/978-981-19-8669-7_53

and road safety [2]. To increase safety qualities and obtain a clear steer on road deaths, technology such as warning the driver of possible dangers, providing protections, or even taking control of cars during a snag has been developed.

Potholes cause a variety of obstacles brought on by such disasters. In such instances, it appears that the only option is to contact or protest to the appropriate authorities. These forces are not readily available or quick enough, perhaps worsening the problem. As a result, an automated solution is required to assist responsible authorities in effectively managing pothole issues. A real-time application is provided to meet the requirement by automating the monitoring process using latest deep learning algorithms and image processing techniques [3]. Using deep neural networks, it was feasible to reliably identify potholes [4–8].

2 Literature Survey

This paper describes creating a two-stage process in which the model tries to localize the regions that are more likely to have potholes, then increases the resolution of these regions and concentrates upon segments that differentiate between areas with potholes and areas without potholes [9]. Aparna et al. [10] A CNN-based 2D vision-based method for identifying potholes includes two primary networks: a localization convolution network and part classification network (PCNN). To locate the region most likely to have potholes, LCNN is used which has a high recall metric. The PCNN uses classification to forecast potholes in the region. Separately, the LCNN and PCNN are trained: With 100 epochs, the image was scaled to 352 $\times$ 224.

The major objective of the paper [11] is to construct a profound neural network to identify images with potholes and to create an end-to-end framework, and develop an android application, and after that, send the demonstration into it. Images of streets were taken where the pictures are gathered using versatile cameras in sunshine along most streets and contract paths and a dataset was created out of them. For pothole location, a CNN model is presented and tested with images. The Google Maps Application Programming Interface (API) is used to generate real-time pothole-marked maps [5, 12–14]. The pothole application system take care of the pothole detection and also helps in finding the location of the image in the map using the framework.

The paper describes how to develop a real-time automated system application that will aid authorities in effectively addressing the problems of waste, open manholes, and potholes [3]. To automate the monitoring procedure, powerful artificial intelligence and image processing techniques are used. The dataset was constructed by obtaining 700 photographs from Google Images for all the classes for the model which are manholes, potholes, and as well as garbage. Users may contribute images of potholes and garbage and are allowed to post the images with the exact location of them through the Android application. For picture validation, a classifier is created with F-RCNN. This information is then recorded in a database.

The model used is a F-RCNN with Inception-v2 network [4]. The performance, accuracy, detection time, and differences between F-RCNN using SSD and the algorithm YOLO are compared in the paper. The suggested approach uses You Only Look Once version 2 (YOLO v2) and a convolutional neural network to detect potholes [13]. The characteristics of testing and images are extracted using the preconfigured CNN resnet50.

In this study [15], a mask region-based CNN is presented as a deep learning approach for properly detecting and segmenting such potholes in order to compute their area. Images comprising 291 that were personally gathered in Mumbai streets, neighboring roads. The dataset is manually labeled with VGG tool, which is publicly accessible. Potholes were recognized as ROI using the Mask Region-Based CNN (Mask RCNN). The area of a pothole is then determined based on the created zone of interest. After that, the calculated area and accurate samples are set side by side. This research [16] proposes a deep learning-based approach for detecting potholes somewhat faster with photos, lowering risk and trouble. Deep learning and Faster-RCNN with Inceptionv2 were the main components of this model.

3 Methodology

3.1 R-CNN

The faster region-based convolutional neural network (F-RCNN) comprises three networks, namely: Feature, detection, and region proposal networks. Feature network in F-RCNN is responsible for the extraction of features from the images. This process of generating features is usually carried out by using a feature extractor, such as ResNet50 or Inception v3. Once the features are generated by the feature network, the execution is passed on to the region proposal network to generate object proposals. This layer estimates the areas in the image where objects are most likely to exist. The RPN is a three-layered convolutional network, where one layer is used to classify the images with classes and the remaining performs regression. The regression provides the bounding box coordinates. Eventually, the class of the object detected within the bounding box is produced by the detection network [1, 10, 17–19].

R-CNN model is trained using the Resnet101 feature extractor with the augmented dataset collected and sent for testing and validation by splitting the dataset. In each epoch, the precision and recall values are printed and stored while the model is trained. Figure 1 shows the block diagram of pothole detection for R-CNN.

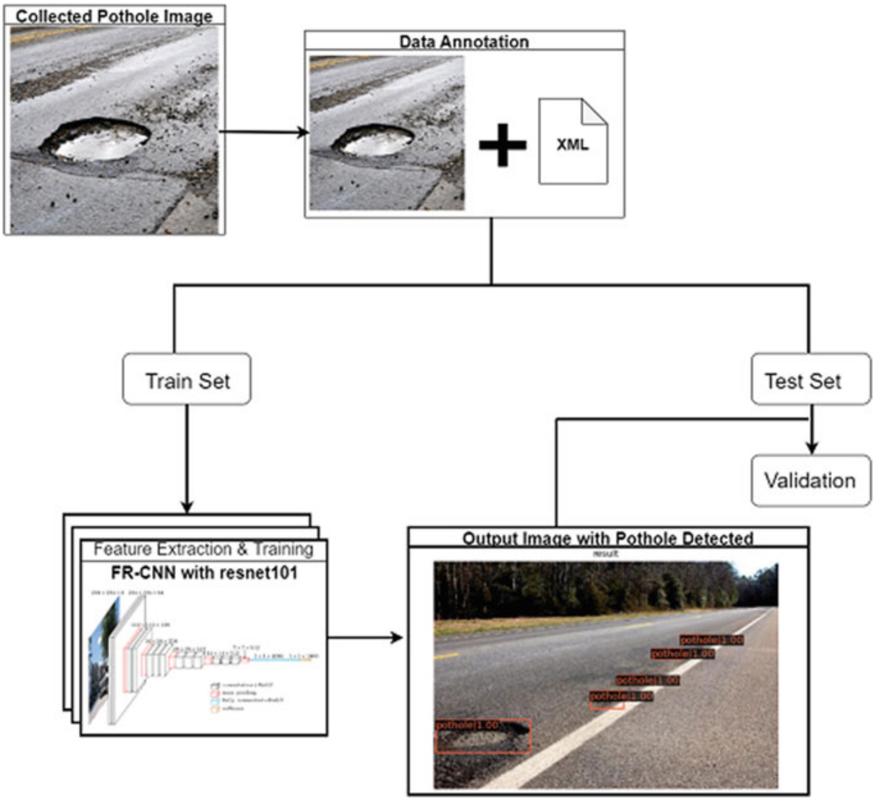

Fig. 1 Block diagram of pothole detection for R-CNN

3.2 YOLO

YOLO is a real-time object detection system that detects a variety of things in a fraction of time. It also recognizes items faster and more precisely than previous recognition systems. YOLO was designed to develop time taking two staged object detector like F-RCNN work better. Even when running on a GPU, R-CNNs are fast, yet they're slow. Single-staged detectors, such as YOLO, are relatively fast and can attain hyper-real-time efficiency on a GPU. Since YOLO is trained to do both classification and regression at single time, it takes less time for prediction [17].

The YOLOv5 consists of two parts: Model Backbone and Model Head. The Model Backbone extracts the important features present in the given input image and reduces the number of network parameters to prevent overfitting, it extracts the informative features of the image and reduces the amount of captured features. The Model head generates probabilities of classes, localization vectors, and the performance metrics of detected target object. Following this, an activation function in the final detection layer will be used to activate neurons based on its weight. To automatically find the optimum bounding box for the dataset and apply them during training, YOLO v5 employs a k-means clustering technique with various k values.

YOLO-v5 is the updated version of YOLO-v4 which is identical to YOLOv4 in terms of performance, implementation, and design. Furthermore, the models in YOLOv5 are substantially smaller, faster to train, and more practical to utilize in a real-world application. For their model, YOLOv5 provides four distinct scales: S, M, L, and X, which stands for small, medium, large, and Xlarge, respectively. Each of these scales multiplies the depth and breadth of the model by a different factor, so the general structure of the model remains the same, but the size of each model is scaled. In this paper, we had implemented the models YOLOv5s and YOLOv5m and compared their performance. The YOLOv5 implementation is compatible with the metrics of the Microsoft Common Objects in Context (COCO) API at three distinct object sizes (bounding box areas) and Intersection over Unions (IOU), which are most important terms for this study. The method for calculating values at specified scales can provide a good indicator of the model's performance, although it may be somewhat erroneous in extreme circumstances, which will not be an issue for most of the cases.

3.3 Dataset

Indian potholes are unique from those found in other countries. As a result, there is a need to create a new dataset that accurately depicts contemporary Indian road conditions. Our crew acquired 665 photographs from a variety of perspectives, distances, and lighting situations. Photographs for the dataset are collected in different weather circumstances such as during rain and cloudy sunny. The annotation is done using LabelImg because it can label pictures in YOLO format. Sample images of the dataset are shown in Fig. 2.

Data augmentation is a strategy for extending the quantity of data available by generating new dataset from current dataset. It improves the dataset by creating a variety of images with different filters applied.

The following data enhancements were performed: Rotation of the images: We rotated the image clockwise by 15° and 90°. Horizontal flip is used for photographs that, when flipped horizontally, seem almost similar. For data augmentation, Gaussian noise is injected into the picture. The process of blurring a picture entails averaging nearby pixels. This blurs the image and lowers the amount of information. The dataset size grows larger, and the model's training improves. The 665 picture dataset was turned into 1995 image datasets using the aforesaid data augmentation approaches, and object identification algorithms were employed to find the potholes.

The dataset is divided into the train, test sets, and validation sets at random. The train set comprises 70% of the dataset, 20% for validation, and 10% for testing.

Fig. 2 Sample images of the dataset

Fig. 3 Detection of potholes by using R-CNN model

4 Results

4.1 R-CNN

The results are obtained based on mean Average Precision. Faster-RCNN with ResNet101 feature extractor achieved a precision 63.9% and a Recall 75%. Figure 3 shows the detection of potholes by using R-CNN model.

4.2 YOLO

YOLOv5s algorithm had an accuracy of 82% and a Recall of 71.5%, whereas Faster-RCNN had an average precision of 63.9% and an average Recall of 75%. The below Fig. 4 detected of potholes by using YOLO model.

Fig. 4 Detected of potholes by using YOLO model

Above Fig. 5 shows the loss graph and the mAP graph after testing the trained YOLOv5 model, and the average loss and mean average precision after validation of the dataset are obtained.

Table 1 shows the metrics table for F-RCNN and YOLOv5 algorithm The above metrics are calculated using mathematical formulas and equations. Precision and Recall are obtained after training the model and testing the model which is trained. Average Precision and Average Recall are the average of precision and recall values, respectively. F1 score is calculated by using average precision and average recall values given by the formula,

$$F1 = 2 * (P * R)/(P + R) \tag{1}$$

Mean average precision is calculated by using precision values and given by the formula,

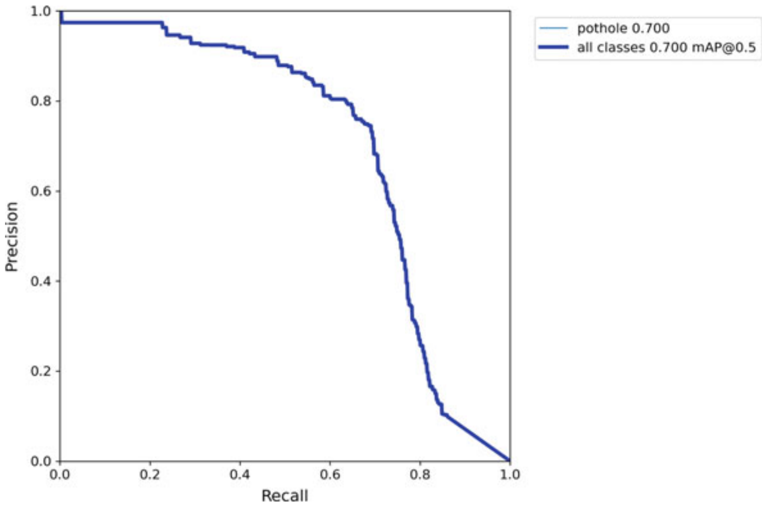

Fig. 5 Precision versus recall graph for YOLOv5 algorithm

Table 1 Metrics table for F-RCNN and YOLOv5 algorithm: Metrics table

Metrics	F-RCNN with ResNet101	YOLOv5s	YOLOv5m
Average precision	0.78	0.82	0.84
Mean average precision (mAP)	0.51	0.74	0.75
Average recall	0.70	0.72	0.74
$F1$ score	0.69	0.77	0.79
Loss	0.04	0.03	0.03
Time	1.98 s	0.01 s	0.02 s

$$mAP = (1/N)\Sigma(Pi)\text{—(where } i \text{ ranges from 1 to } n) \tag{2}$$

Loss is a statistic that represents how accurate the model is, or to put it another way, a faulty prediction results in a loss. A loss of 0 indicates that the model is perfect, whereas a loss of more than 0 suggests that the model is inefficient. Time is defined as the average time taken to detect the image after training the model.

5 Conclusion

Pothole detection accomplished by object identification techniques with increased speed and accuracy plays a significant role in the prevention of disasters and troubles happening on roads. Due to this, a model that employs the convolutional neural networks to detect the potholes was developed. This paper proposes four distinct algorithms, which are then compared.

The algorithms Faster R-CNN and YOLOv5 are implemented. The size of the dataset was increased from 665 to 1995 using data augmentation techniques. ResNet101 is used as the feature extractor for Faster R-CNN, and the performance are compared. The pre-trained weight for convolutional layers in YOLOv5 is the reduced configuration of Darknet-53. YOLOv5s and YOLOv5m are the algorithms implemented and compared. Although each algorithm had its own strengths and issues, YOLOv5 m outperform other designs in terms of accuracy.

Using the live front-view camera that is employed by Advanced Driver-Assistance Systems, this pothole detection model can be used in alerting the driver through its real-time detection. Potholes, on the other hand, may also go unnoticed and lead to false negatives due to various factors such as insufficient lighting, water-covered potholes, and high-speed vehicle movement. The model can also lead to false positives as a result of different kinds of shadows and the various shapes in which potholes are being deployed. To overcome these challenges, additional cameras can be installed, and features more specific to the potholes can be included in the recommended model. The data obtained by the system can also be given to a database. This may be linked with an application that pins pothole locations or help Google Maps in generating directions and routes with fewer potholes.

References

1. Dunna, S., Nair, B.B., Panda, M.K.: A deep learning based system for fast detection of obstacles using rear-view camera under parking scenarios. In: 2021 IEEE International Power and Renewable Energy Conference (IPRECON), pp. 1–7. IEEE (2021)
2. Dharneeshkar, J., Aniruthan, S.A., Karthika, R., Parameswaran, L.: Deep learning based detection of potholes in Indian roads using YOLO. In: 2020 International Conference on Inventive Computation Technologies (ICICT), pp. 381–385 (2020). https://doi.org/10.1109/ICICT48043.2020.9112424
3. Sayyad, S., Parmar, S., Jadhav, M., Khadayate, K.: Real-time garbage, potholes and manholes monitoring system using deep learning techniques. In: 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI) (48184), pp. 826–831 (2020). https://doi.org/10.1109/ICOEI48184.2020.9143030
4. Varona, B., Monteserin, A., Teyseyre, A.: A deep learning approach to automatic road surface monitoring and pothole detection. Pers Ubiquit Comput. **24**, 519–534 (2020). https://doi.org/10.1007/s00779-019-01234-z
5. Carneiro, T., Medeiros Da NóBrega, R.V., Nepomuceno, T., Bian, G., De Albuquerque, V.H.C., Filho, P.P.R.: Performance analysis of google colaboratory as a tool for accelerating deep learning applications. IEEE Access **6**, 61677–61685 (2018). https://doi.org/10.1109/ACCESS.2018.2874767
6. CC, G.B.R., SB Rao, M., SM, K.E., SJ: Deep learning based pothole detection and reporting system. In: 2020 7th International Conference on Smart Structures and Systems (ICSSS), pp. 1–6 (2020). https://doi.org/10.1109/ICSSS49621.2020.9202061
7. Rohitaa, R., Shreya, S., Amutha, R.: Intelligent deep learning based pothole detection and reporting system. In: 2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT), pp. 1–5 (2021). https://doi.org/10.1109/ICECCT52121.2021.9616703
8. Yik, Y., Alias, N., Yusof, Y., Isaak, S.: A real-time pothole detection based on deep learning approach. J. Phys: Conf. Ser. **1828**, 012001 (2021). https://doi.org/10.1088/1742-6596/1828/1/012001
9. Chen, H., Yao, M., Gu, Q.: Pothole detection using location-aware convolutional neural networks. Int. J. Mach. Learn. Cyber. **11**, 899–911 (2020). https://doi.org/10.1007/s13042-020-01078-7
10. Aparna, Y.B., Rai, R., Gupta, V., Aggarwal, N., Akula, A.: Convolutional neural networks based potholes detection using thermal imaging. J. King Saud Univ. Comput. Inf. Sci.
11. Patra, S., Middya, A., Roy, S. : PotSpot: Participatory sensing based monitoring system for pothole detection using deep learning. Multimedia Tools Appl. **80** (2021). https://doi.org/10.1007/s11042-021-10874-4
12. Sumalatha, R., Rao, R.V., Devi, S.M.R.: Pothole detection using YOLOv2 object detection network and convolutional neural network. In: Iyer, B., Ghosh, D., Balas, V.E. (eds) Applied Information Processing Systems . Advances in Intelligent Systems and Computing, vol. 1354. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-2008-9_28
13. Omar, M., Kumar, P.: Detection of Roads Potholes using YOLOv4. Int. Conf. Inform. Sci. Commun. Technol. (ICISCT) **2020**, 1–6 (2020). https://doi.org/10.1109/ICISCT50599.2020.9351373
14. Arjapure, S., Kalbande, D.R.: Deep learning model for pothole detection and area computation. Int. Conf. Commun. Inf. Comput. Technol. (ICCICT) **2021**, 1–6 (2021). https://doi.org/10.1109/ICCICT50803.2021.9510073

15. Kumar, A., Chakrapani, Kalita, D., Singh, V.: A modern pothole detection technique using deep learning 1–5 (2020). https://doi.org/10.1109/IDEA49133.2020.9170705
16. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2017)
17. Dharneeshkar, J., Aniruthan, S.A., Karthika, R., Parameswaran, L.: Deep learning based detection of potholes in Indian roads using YOLO. In: 2020 International Conference on Inventive Computation Technologies (ICICT), pp. 381–385. IEEE (2020)
18. Hemaanand, M., Rakesh Chowdary, P., Darshan, S., Jagadeeswaran, S., Karthika, R., Parameswaran, L.: Advanced driver assistance system using computer vision and IOT. In: International Conference on Computational Vision and Bio Inspired Computing, pp. 768–778. Springer, Cham (2019)
19. Haritha, H., Thangavel, S.K.: A modified deep learning architecture for vehicle detection in traffic monitoring system. Int. J. Comput. Appl. 1–10 (2019)