Developing an Automated System for Pothole Detection and Management Using Deep Learning

P. D. S. S. Lakshmi Kumari[1(✉)], Gidugu Srinija Sivasatya Ramacharanteja[1], S. Suresh Kumar[2], Gorrela Bhuvana Sri[1], Gottumukkala Sai Naga Jyotsna[1], and Aki Hari Keerthi Naga Safalya[1]

[1] Department of Information Technology, S.R.K.R. Engineering College (A), Bhimavaram, AP, India
divyapannasa@gmail.com

[2] CSE Department S.R.K.R. Engineering College (A), Bhimavaram, AP, India

Abstract. The economy greatly benefits from the use of roads as a platform for mobility. Roadway potholes are one of the primary problems with transportation infrastructure. Accidents frequently occur because of these potholes. Many research have advocated using computer vision methods to automate pothole identification, including various image processing and object detection algorithms. The system must be easy to use, economical to set up, and capable of automating pothole identification quickly and accurately. In this study, we constructed effective deep learning convolution neural networks (CNNs) to identify potholes accurately and quickly and to enhance training results and lower computational costs. In addition to that, the performance of YOLOv7 and Faster R-CNN with ResNet50 (FPN) backbone is also contrasted in this paper. According to the trial findings, the YOLO's speed makes it more useful for real-time pothole detection.

Keywords: YOLOV7 · Faster-RCNN · ResNet50 · PyTorch Pothole Detection

1 Introduction

Roads are a nation's primary source of mobility when it comes to offering nationwide commuting options. Road infrastructure makes it possible to link people and carry commodities, enhancing commercial prospects, job access, economic growth, and the nation's healthcare system. Although high-quality roads boost the nation's GDP, poor road infrastructure can be deadly for the wellbeing of drivers and passengers as well as the condition of vehicles. This poor road infrastructure is caused due to many reasons like heavy traffic, poor maintenance, bad weather conditions and wear and tear etc. One of the main anomalies of poor road infrastructure is potholes. Potholes are basically concave depressions in the road surface that need to be repaired since they cause terrible events like accidents, unpleasant driving experiences, and malfunctioning automobiles. Sometimes these potholes may lead to fatal accidents.

R. N. Shaw et al. (Eds.): ICACIS 2023, CCIS 1921, pp. 12–22, 2023.
https://doi.org/10.1007/978-3-031-45124-9_2

A few initiatives are being made utilizing various strategies to automate the pothole identification system on roads. Few are based on sensors. Some are based on 3-D reconstruction Processing of images, and some are model-based. Pothole detection methods based on sensors use vibration sensors. The vibration sensor may mistake joints, cracks in the road as potholes or fail to detect potholes in the middle of a lane, which can lead to false readings that may be false positives or false negatives which can compromise the accuracy of detecting potholes. In three-dimensional reconstruction techniques, the 3-dimensional road data captured by using lasers and cameras are used to detect potholes. The clarity of data can be compromised by the camera misalignment that could reduce detection accuracy and require expensive design and computing work to recreate the pavement surface.

Even though standard image processing techniques for pothole detection offer a high degree of accuracy, they still require difficult tasks like manually bringing out features and modifying the image processing variables and stages for various road conditions. The development of model-based pothole identification and recognition algorithms has been driven by the advancement of image processing techniques and the accessibility of inexpensive camera equipment. Using conventional ML techniques, a trained model for spotting potholes in 2-dimensional digital images was created. They use a lot of computation power and still manage to attain great accuracy. Specialists are required to manually bringing out the features in order to increase the pothole detection accuracy of ML algorithms. Convolutional neural network (CNN) methods, which are capable of concurrently automating the functioning of feature extraction and classification, were employed in deep learning (DL) approaches.

We proposed a pothole detection system using YOLO (YOU ONLY LOOK ONCE) and Faster-RCNN. We are using the current version of YOLO (YOLOv7) for this study, and we also want to contrast the performance of single stage detector YOLO and two stage detector Faster-RCNN.

2 Literature Survey

Seung-Ki Ryu, Taehyeong Kim, Young-Ro Kim, 2015 [1]. The authors proposed the Image-Based Pothole Detection System for the ITS Service and Road Management System in 2015. The three steps of this system are as follows: Segmentation, candidate region extraction, and decision-making. To begin, dark areas for pothole identification are extracted using a histogram and the morphological filter's closure operation. Next, prospective regions of a pothole are retrieved using a variety of criteria, including size and compactness. Lastly, by comparing pothole and background features, it is determined if the candidate regions are potholes or not.

Hsiu-Wen Wang, Chi-Hua Chen, Ding-Yuan Cheng, Chun-Hao Lin, and Chi-Chun Lo, 2015 [2]. An approach for real-time pothole detection for intelligent transportation systems was put forth by the authors in 2015. Based on mobile sensing, the authors provide a strategy for locating potholes. The pothole information is obtained by applying the Euler angle mechanisms to the data of accelerometer to normalise it. In addition, the interpolation techniques like spatial interpolation approach is employed to minimise locality mistakes in GPS data. The results of the trials demonstrate that the suggested approach can accurately identify classes with no false positives, and the recommended approach performs with a greater degree of accuracy.

K. Vigneshwar, B. Hema Kumar, 2016 [3]**.** Utilizing image processing techniques, the authors suggested Detection and Counting of Potholes in 2016. To achieve better results, the authors suggested a system where image processing is based on different Gaussian-Filtering and clustering based image segmentation approaches. According to the findings, segmentation based on edge- detection is favoured for its specificity and segmentation based on K-Means clustering was preferred for its quickest computation time.

Jeong-Joo Kim, Choi Soo-il, 2017 [4]**.** In 2017, Implementation of Pothole Detection System Using 2d Lidar was the idea put out by the authors. By using information from cameras and 2D lidar, the scientists have suggested a method for detecting potholes. These lidar cameras' major application is their capacity to capture a large region with fine detail. After the data is captured, a number of techniques are applied, including filtering, classification, line extraction, and gradient of data function. The performance of pothole recognition from video data and the combination with 2D lidar was improved.
Asif Ahmed, Samiul Islam, Amitabha Chakrabarty, 2019 [5]**.** In 2019, The identification and comparison of potholes using image processing algorithms is suggested by the authors. To detect every form of pothole, the topic chosen employs 4 different image segmentation approaches. The approaches Image Thresholding, Canny Edge Detection, K-Means clustering, and Fuzzy C-Means clustering were all worked on. After that, several scenarios were examined to see how well the various image segmentation approaches worked. In terms of accuracy and precision results were produced. Also, the outcomes were contrasted with one another in order to assess their viability.

Aparna, Yukti Bhatia, Rachna Rai, Varun Gupta, Naveen Aggarwal, Aparna Akula, 2019 [6]**.** The authors have presented Convolutional Neural Networks based Potholes Detection Using Thermal Imaging in 2019. This technique is able to identify potholes from thermal imagery. Real-time pothole detection is an option. A deep learning-based strategy has been used to this. Thermal photographs of roads with and without potholes are used as input in a model built on a convolutional neural network (CNN). The model recognizes and identifies if the feeded in image contains a pothole or not after being trained on this data.

Z. Hasan, S. N. Shampa, T. R. Shahidi and S. Siddique 2020 [7]**.** In 2020, The authors suggested using smartphone cameras and CNN's to detect potholes and speed bumps. In this study, we have created a model that uses computer vision and machine learning methods to identify undesirable potholes, deep ridges, and speed bumps. In order to train their machine learning algorithms, they have created a unique dataset, which they've named Bumpy. In their study, they offer a method for identifying speed bumps, deep ridges, and potholes using a pre-trained Tensorflow model.

Surekha Arjapure, D.R. Kalbande, 2020 [8]**.** In 2020, Road Pothole Detection Using Deep Learning Classifiers is the solution put out by the authors. A deep learning method was suggested by the authors. It is suggested to use a Mask Region-Based Convolutional Neural Network to precisely detect and segment such potholes in order to anticipate and determine their extent. The 291 photos in the database were painstakingly gathered on Mumbai's city streets and close-by motorways. The manual VGG Image Annotator tool is used to manually annotate the dataset. Potholes are identified as a zone of interest using Mask Region-Based Convolutional Neural Network (Mask RCNN).

Bucko, Boris, Eva Lieskovská, Katarína Zábovská, and Michal Zábovský, 2022 [9]**.** In this paper You Look Only Once version 3, often known as Yolo v3, a computer vision model library, is used to automatically detect potholes. Driving in low light or bad weather inherently limits our ability to see road damage. Visual object detection performance is likewise negatively impacted by such unfavourable circumstances.

B, M.P., K.C, S., 2022 [10]. The main objective of this paper is to train and evaluate the YOLOX model for pothole identification. YOLOX is an object detection technique. The YOLOX model is trained using a dataset of potholes, and the results are examined by assessing the model's accuracy, recall, and size, which are then contrasted with those of other YOLO algorithms.

3 Methodology

From the related work, we have observed some drawbacks like low accuracy, high computation power needed, high cost setup and etc. To fill this research gap, we proposed Pothole detection system using YOLO and Faster-RCNN

A. YOLOv7

YOLO is shortly named for 'You Only Look Once'. This is an algorithm that identifies and recognizes various objects in a image. In YOLO, an object is detected as a regression problem and lay outs the class probabilities for all detected images. YOLO is a single stage object detector which passes all components of the object detection into a single neural network. The CNN is used to predict various class probabilities and bounding boxes simultaneously. The YOLO algorithm consists of various variants. In this study we are using the version 7 of YOLO (Base YOLOv7) (Fig. 1).

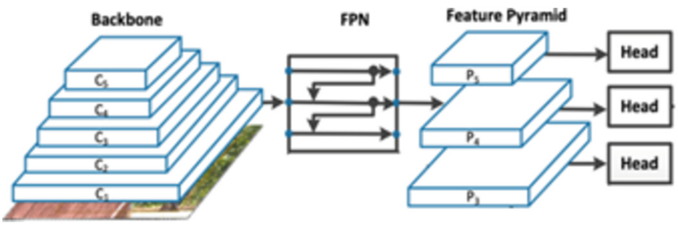

Fig. 1. Architecture of YOLOv7

The image data is feeded to backbone. Backbone is stack of several pre-trained weights. The extracted features from backbone is feeded to sequential network FPN. FPN is the neck of the yolo architecture. Here in FPN, convolution and pooling is done and produces feature pyramid. The feature pyramid is feeded to head where all the computation is done. Finally the head will predict the class label of corresponding feature

These are the steps needed to follow:

a) Data Collection.

We have created a custom datasets with the help of images downloaded from different sources like Kaggle, Roboflow and etc.

b) Data Annotation.

Annotating or labelling the images in the dataset using any data annotation tool. We have used labelImg for image annotation.

iii) Data hierarchy.

Split the images and labels into train and val folders.

iv) Training.

- We trained our custom pothole dataset with the help of official YOLOv7 GitHub repository.
- We have made few changes to some of the files in the repository like changing the number of classes and class names.
- After all changes made, we trained,validated our custom dataset on google colab with pre-trained weights respectively

v) Testing.

We have tested a video of potholes with our custom model weights and got predictions (Fig. 2).

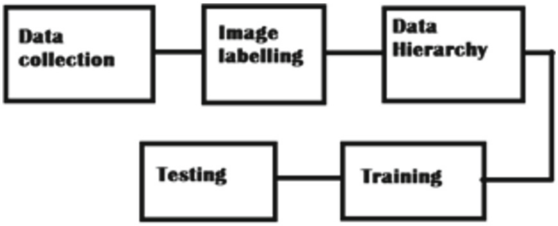

Fig. 2. Simple overview of pothole detection using YOLOv7

B. FASTER-RCNN

Faster – RCNN is short term for Faster Region based convolutional neural network, it is updated version of Fast-RCNN which is updated version of R-CNN.

Faster-RCNN contains two modules:

- **RPN**: For generating featured maps of anchor boxes
- **ROI pooling**: It is used for making all feature maps of anchor boxes to same size (Fig. 3)

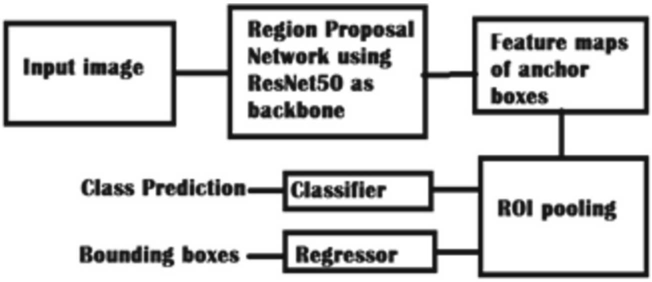

Fig. 3. Architecture of Faster-RCNN

The image is passed to region proposal network. In region proposal network, the CNN produces the feature maps of anchor boxes of foreground regions. Anchor boxes are nothing but bounding boxes used for region proposals which are of different dimensions. Foreground regions are the regions which contains the object. The output of the RPN is feature maps of anchor boxes. This output will be feeded to ROI pooling layer which makes all feature maps equal in size. Then the classifier identifies the object for the respective feature map and regressor will give a bounding box around the object.

These are the steps needed to follow:

a) Data Collection.

- we have created a custom data from internet sources like Kaggle.
- This dataset was annotated.

b) Directory setup.

- We have six python scripts, input folder, checkpoints folder, Test predictions folder
- The project directory consists of following details.
- The input folder contains the images and annotations required for training and also contains images for testing
- The savedmodel folder contains the model saved after the training
- Test_predictions folder contains the results produced after testing

c) Training.

•Run the train.py script

d) Testing.

- Run the test.py script (Fig. 4)

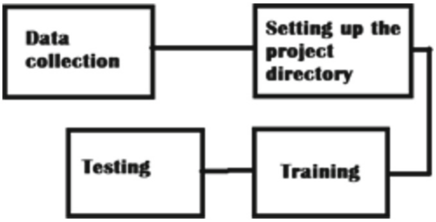

Fig. 4. Simple overview of pothole detection using Faster-RCNN

4 Experiments and Results

After training the YOLOv7 model with our custom dataset, we got batches of trained images. We can see them below (Fig. 5).

Fig. 5. Batches of Trained Images of YOLOv7

In training, for some set of epochs, a trained model will be saved in weights folder. At last, the best model out of all epochs will be saved. This will look like below (Fig. 6).

Fig. 6. Best Model saved after training the YOLOv7

Using this best model we can test our custom videos/images. Below is the frame of pothole prediction from a video (Fig. 7).

Fig. 7. YOLOv7 Prediction frames from an input video

For Faster-RCNN training, we set up the number of epochs to 8 (from epoch 0 to epoch 7) (Table 1).

Table 1. Loss values of each epochs

Epoch No	Loss Value	Time taken for each epoch
0	0.2209340971360504	~84 min
1	0.15503743677671766	~83 min
2	0.13977767626044735	~166 min
3	0.13120384020374815	~82 min
4	0.12274352340981623	~83 min
5	0.11972583682694185	~ 82 min
6	0.11648200763393216	~83 min
7	0.11067249564457707	~82 min

The predictions of Faster-RCNN is as follows (Fig. 8):

Fig. 8. Predictions of Faster-RCNN

5 Evaluation Metrics

True positives: Model predicted "yes" and the prediction is correct
True negatives: Model predicted "no" and the prediction is correct
False positives: Model predicted "yes" and the prediction is wrong (actually it was "no")
False negatives: Model predicted "no" and the prediction is wrong (actually it was "yes")
Precision: It is the proportion of correctly predicted forecasts to all predicted predictions that were predicted as "true."
Recall: It is the proportion of correctly predicted forecasts to all true positives
From the obtained graphs, we can see our YOLO model has precision and recall about 90% and mAP@0.5 is somewhat near to 0.9. Since it is a object detection, accuracy is termed as the ratio of number of correctly predicted potholes to all the pothole images (Table 2).

Table 2. Comparison with related work

Evaluation metrics	Proposed Method	[9]	[10]
mAP@0.5	0.9	0.747	0.85.6
Methodology	FASTER RCNN	SPARSE RCNN	Darknet-53

From the above confusion matrix, it is clearly means that our model predicts 97% correctly. i.e., out of 100 actual potholes it can clearly predict 97 potholes correctly (Figs. 9 and 10).

This means we have 97% accuracy for the custom yolo model.

For Faster-RCNN, we have tested 100 images of potholes and we got 85% accuracy that means it predicts 85 images correctly (Table 3).

Although YOLO has higher accuracy, it didn't predict the smaller potholes as potholes and Faster-RCNN predicts tiny potholes as potholes but it has lower accuracy.

From all these metrics, we can say our YOLOv7 model is best enough for real time application in aspect of speed and accuracy.

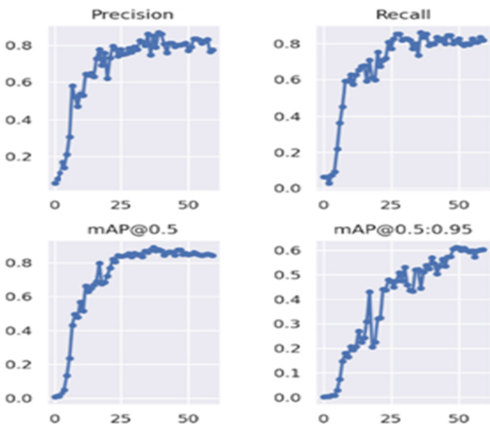

Fig. 9. Results of YOLOv7

Fig. 10. Confusion Matrix

Table 3. YOLOv7 vs Faster-RCNN

PARAMETERS	YOLOv7	Faster-RCNN
Image size	640	800
Batch size	16	2
No. of Epochs	60	8
Training time per epoch	28.56 s	4920 s
Training time	0.476 h	11.4 h
Average Frames per second	25	0.312
Accuracy	97%	85%
Able to detect tiny potholes	No	Yes
Able to detect from Larger distances	Yes	No

6 Conclusion

This research created effective CNN models taking into account the requirements to reliably and quickly identify potholes in roadways. The trials carried out in this study made use of a dataset that included photographs of potholes that were taken under various lighting situations, on various types of roads, and with various forms and sizes. Two distinct CNNs were used to train the pothole dataset: YOLOv7 and Faster-RCNN.

Our main aim is to tradeoff between the single stage object detector YOLOv7 and two stage object detector Faster-RCNN for real time application of pothole detection. We can conclude that YOLOv7 is mostly applicable to pothole detection in real time because of its speed and accuracy.

7 Future Work

There is a scope for combining this project with IOT to get more realistic pothole detection system. If YOLOv7 is deployed along with IOT it will be more useful as there is a chance to alert the driver with a beep sound.

References

1. Ryu, S.-K., Kim, T., Kim, Y.-R.: Proposed image-based pothole detection system for its service and road management system (2015)
2. Wang, H.-W., Chen, C.-H., Cheng, D.-Y., Lin, C.-H., Lo, C.-C.: Proposed a real-time pothole detection approach for intelligent transportation system (2015)
3. Vigneshwar, K., Hema Kumar, B.: Proposed detection and counting of pothole using image processing techniques (2016)
4. Kim, J.-J., Soo-il, C.: Proposed implementation of pothole detection system using 2d lidar (2017)
5. Soni, A., Dharmacharya, D., Pal, A., Srivastava, V.K., Shaw, R.N., Ghosh, A.: Design of a machine learning-based self-driving car. In: Bianchini, M., Simic, M., Ghosh, A., Shaw, R.N. (eds.) Machine Learning for Robotics Applications. SCI, vol. 960, pp. 139–151. Springer, Singapore (2021). https://doi.org/10.1007/978-981-16-0598-7_11
6. Ahmed, A., Islam, S., Chakrabarty, A.: Proposed identification and comparative analysis of potholes using image processing algorithms (2019)
7. Hasan, Z., Shampa, S.N., Shahidi, T.R., Siddique, S.: Proposed pothole and speed breaker detection using smartphone cameras and convolutional neural networks (2020)
8. Arjapure, S., Kalbande, D.R.: Proposed Road Pothole Detection Using Deep Learning Classifiers (2020)
9. Bucko, B., Lieskovská, E., Zábovská, K., Zábovský, M.: Computer vision based pothole detection under challenging conditions. Sensors **22**, 8878 (2022). https://doi.org/10.3390/s22228878
10. Mohan Prakash, B., Sriharipriya, K.C.: Enhanced pothole detection system using YOLOX algorithm. Auton. Intell. Syst. **2**, 22 (2022). https://doi.org/10.1007/s43684-022-00037-z