A DEEP LEARNING APPROACH FOR POTHOLE DETECTION

A V S RADHIKA[1], G SAHITHI[2], M SAI SRI HARSHITHA[3], J SANDHYA RANI[4], V VIJAYA LAKSHMI[5]

ASSISTANT PROFESSOR[1], UG SCHOLAR [2, 3, 4 &5]

DEPARTMENT OF CSE, BHOJ REDDY ENGINEERING COLLEGE FOR WOMEN, VINAY NAGAR, HYDERABAD-59

ABSTRACT — Roads are connecting lines between different places and are used in our daily life. Roads' periodic maintenance keeps them safe and functional. However, asphalt pavement distresses cause potholes which may increase the number of accidents. Detecting and reporting the existence of potholes to responsible departments can save the roads from getting worse. This study deployed and tested different deep learning architectures to detect the presence of potholes. First, several images of potholes are captured by a cellphone mounted on the car windshield. Then pothole images downloaded from the internet increased the size and the variability of our database (1087 images with more than 2000 potholes). Second, various object detection algorithms are employed to detect potholes in the road images. Real-time Deep Learning algorithms with several configurations like SSDTensorFlow, YOLOv3-Darknet53, and YOLOv4-CSPDarknet53 are used to compare their performances on pothole detection. YOLOv4 achieved a high recall of 81%, high precision of 85% and 85.39% mean Average Precision (mAP). The speed of YOLOv4 processing is recorded around 20 frames per second (FPS) at an image resolution of 832*832. Furthermore, the proposed system detected potholes at distances reaching a hundred meters. Compared with other state-of-the-art methods, our result demonstrated superior performance in real-time. The proposed method can help in reporting road potholes to government agencies, increasing the safety of drivers by detecting potholes time ahead, and improving the performance of self-driving cars to ensure safe trips for passengers in the future.

Index Terms — Pothole detection, Deep Learning Architectures, YOLO, Darknet, SSD, TensorFlow, Real-time.

INTRODUCTION ROADS form a basis for people transportation and joining between different places. The size of roads varies based on their functionality. For instance, highways are large enough to contain many lanes designed for massive traffic. However, roads inside towns are constructed to be smaller and made up of one or two lanes. Roads are vital in people's daily life, so periodic maintenance shall be made to keep them functional and safe. The many roads that exist within a given country make it difficult to have a continuous assessment of roads; therefore, one can't predict the formation of potholes. Pavement distress is the main cause of defects of roads. Pavement distress can be classified into three classes [1]: Pavement distortion (shoving, corrugation, and rutting), fracture (fatiguing, spalling, and cracking), and disintegration (raveling and stripping). This work focuses on potholes which are considered the worst pavement distress, and their creation is unpredictable. The main reason behind such distortions can be related to a combination of environmental conditions and traffic pavement stresses. Potholes are a worldwide

problem asthey cost governments and citizens billions of dollars yearly [2, 3]. 1.25 million people die each year because of road traffic accidents, 34% of which are related to road potholes [4]. Pothole detection can be categorized into three approaches [5]: Vibration-technique approach [6, 7, 8], 3D reconstructiontechnique approach (with laser scanner method, stereo vision method, and Kinect sensor method) [9, 10, 11], and Visiontechnique approach [12, 13, 14]. Table 1 summarizes and compares different pothole detection approaches based on technology used, response and sense time, processing, cost, pothole characterization, and accuracy of detection [15]. Traditionally, a group of employees was performing the pothole detection through reviewing recorded digital videos captured from roads. This method is costly and time-consuming [16]. The edge detection method has been used in [17]. The authors used Sobel operators on original images and the vector of edges is then detected using the Fuzzy Inference System (FIS). The usage of FIS type 1 remarkably improves results with a neural network. FIS type 2 improved the training cost oneural network. [18] provides a deep review and comparison of the existing pothole detection methods. The vibration method, which is based on an accelerometer sensor, can't predict potholes ahead of time. This is because the vehicle should pass over the pothole for detection. This method cannot differentiate between potholes and other artifacts on the road like bridge joints and road reflectors. Laser scanning systems are classified among the 3D reconstruction methods. This system can detect potholes in real-time. However, such a type is costly and has a short range of detection. Likewise, stereo vision methods are used in pothole detection. The drawback of this method is the high computational effort for pavement surface reconstruction, vibration sensitivity, and the need for perfect alignment of cameras. Kinect sensor, based on infrared technology, is developed by Microsoft. The sensor is regarded to be costly and necessitates a close distance to the pothole. Finally, the visionbased approaches that make use of monocular camera along with some object detection algorithms showed an improvement in real-time pothole detection. A pothole detection system made up of Raspberry Pi, camera, 3G mode, and Micro SD card was proposed to be affordable and simple [19]. The authors made use of OpenCV and the system is installed in a stationary manner to report potholes in real-time. The method is based on collecting different frames and convert the picture into a blurring grayscale image. Finally, and after applying some morphological functions, edge detection is employed to define the contours which are fed to Hough transform. Following those few steps, the author was able to detect the presence of a pothole. The system sends an automatic email containing the output image and its location to the transportation department in case a pothole is detected. Another research was able to detect potholes using morphological methods on videos [20]. The authors used digital image processing methods to find the shapes of different objects. The results achieved are recoded to be 93% for accuracy, 93% for precision, and 100% for recall. In another work [21], the authors claimed that a stereo vision system can work in real-time by generating a disparity map for the road with high accuracy and by generating a quadratic fitting in the world coordinate system point cloud. The paper presented the pothole detection successfully without showing any metric of evaluation. An algorithm that can work for the black-box camera is proposed in [22] to create a specialized vision-based system for pothole

detection. They claimed that their solution is a low-cost solution and can work in real-time [23, 24]. The authors' work is divided into 3 stages, pre-processing, candidate extraction, and cascade detector. In the first step, the authors extracted the dark areas from a grayscale image, then the lanes of the road are used to find the vanishing point to create virtual lanes, and finally, the pothole region is extracted using some threshold values. They achieved 71% for recall and 88% for precision. This work is limited due to its total dependence on digital image processing techniques. The limitation is in distinguishing potholes from other objects in front of the camera. Moreover, one paper made use of removing unwanted information from the images' border, like foliage and plants on the border of the street which can lead to wrong readings [25]. Then the convex hull algorithm is used to build a convex contour over the furthest points for different points of interest. The detection of the pothole is based on differences in the color where a large dark shadow area is considered as a pothole. The performance of the system was tested on 53 images with 97 potholes. The results were acceptable with 81.8% precision and 74.4% recall. This technique was able to reject vehicles that appear in the image by an 80% success rate. Furthermore, it was possible to increase the accuracy by more dilation; but this can affect negatively and leads to loss of pothole detection. The system works from a distance of 20 meters down to 2 meters. The computational time of the algorithms is recorded to be 0.148 sec to extract the road from the image and 0.037 sec to find potholes from the image frame. A new technique based on a thermal IR sensor or camera for pothole detection is proposed in [15]. The method is based on the temperature difference of objects from the surrounding. The captured thermal images of the road are fed into a Convolutional Neural Network (CNN). The paper tested different models of CNN: Self-built CNN models and CNNbased ResNet models. The system with ResNet152 achieved an accuracy of 97.08%. Another deep learning algorithm is used to detect potholes [26]. 900 images with potholes were used to train Fast-RCNN. After detecting a pothole, the image of the pothole is saved along with the pothole GPS and the time of its detection. The collected data are saved in a cloud server. The authors claimed to achieve an average precision exceeding 93%. However, the images of the potholes were taken from a close view of the pavement. A dashboard camera along with a CNN is used also for pothole detection in another work [27]. The neural network consisted of four convolutional and pooling layers and one fully connected layer. The team captured images from different places and had variant conditions, like dry, wet, and shady potholes. The images collected were resized to 200x200 pixels and cropped to remove unwanted parts of the image, they only kept the pothole. After data augmentation is used to get a larger dataset for training with 13244 images, 3250 images for validation, and 500 images for testing. The system achieved 99.8% accuracy, 100% precision, and 99.6% recall. The authors highlighted that their model performed better than the Support Vector Machine (SVM). Furthermore, [28] presented a research on detecting different road damages. This paper studied variant types of road distortion such as white line blur, crosswalk blur, rutting, bump, pothole, alligator crack, construction joint part, wheel mark part, etc. The dataset collected is made up of 9,053 road damage images. The authors adopted object detection methods using CNN in training their system and detecting road distortions. The authors also

tested their object detection on smartphones. They achieved 71% and 77% for recall and precision respectively. A combination of vision and vibration methods for pothole detection has been proposed in [29]. The authors used an accelerometer and camera of mobile phone for this task. Based on SSD with MobileNet, the system was able to detect potholes with 55% accuracy for vibration-based method, and 60% for vision-based method. Although the system was able to detect the pothole in real-time, this would be only done at a close distance to the pothole. A further work proposed a system to detect transverse cracks, longitudinal cracks, and potholes [30]. The authors used mobile phones to get images for those road damages, then sent the data with their GPS location to an online server to evaluate the road damage severity. YOLOv2-tiny is recommended among the tested different deep learning object detectors. Another research made use of digital image processing with spectral clustering to find potholes [31]. The system achieved 81% accuracy, but with the images being cropped around the pothole. An evaluation of the vision-based system using YOLOv2, YOLOv3, and YOLOv3 Tiny has been conducted [32]. The system is trained on 80% of 1500 images of Indian roads. The paper recommended the YOLOv3 Tiny with 72.12 mean Average Precision at 25% Intersection Over Union threshold, 76% for precision, and 40% for recall. However, the speed of processing is not presented. YOLO with ResNet-50 is trained to detect potholes and bumps based on 3399 images for normal roads, 1337 images for potholes, and 547 images for bump [33]. 80% of the dataset was used for training and the rest for validation. The system achieved 88.9% true positive detections. Another research used 448 images, 50% of the images for the training, to train YOLO neural network [34]. YOLOv3, YOLOv3 Tiny, and YOLOv3 SPP achieved mean average precision of 83.43%, 79.33%, and 88.93% respectively, and accuracy was 64.45%, 53.26%, and 72.10% respectively. The database was built by scanning the road from the top view, using a camera mounted on the rear side of an SUV car. The processing of each image is 4ms. On the other side, the gyroscope and accelerometer of a smartphone are used for pothole detection [35]. Inception V3 as “Transfer Learning”, is used for feature extraction in CNN. The authors used 70% of the data collected to train CNN and 30% to test it. Their results exhibited a 100% correct classification rate after testing the trained model. In [36] the accelerometer data collected from mobile phones are normalized by Euler angle computation which is fed to a combination algorithm of Z-THRESH and G-ZERO approaches. Then, spatial interpolation is used to locate the pothole. Results revealed a 100 % accuracy in detecting potholes without false positives. Another work presented a realtime system for inspecting and detecting road distresses [37]. The system used a high-speed 3D transverse scanning method. Structured light triangulation formed the base of the characterization and dynamic generation of the 3D pavement profile. The detection system is made up of a GigE digital camera and an infrared laser (810 nm) line projector. The system is mounted on the rear side of the vehicle. To make the laser stripe covers a full lane of pavement transversely, an 80◦ fan angle laser projector has been used. The camera catches continuous images for the lines of the laser to compute 3D transverse profile. Based on the triangulation principle, the elevation of a specific point can be found. Another work employed laser imaging for pavement distress inspection [38]. Several features are then captured including the total

number of distress tiles and the depth index which are given to a threelayer neural network for classifying the type of the crack and to estimate its severity. Another research achieved a high recognition rate of potholes based on texture measures from the histogram as features to train the Support Vector Machine (SVM) [39]. To conclude, the vibration method is a low cost method and can evaluate pothole severity. However, this method could be harmful to the car and it can't differentiate between potholes and other artifacts on the road. In the 3D reconstruction technique, the laser scanner method can estimate pothole size and its severity, but this system is expensive and has a short range of detection. The stereo vision method has an average price and can estimate pothole size, but it can't evaluate pothole severity. The alignment of cameras must be perfect and has a short range of detection. The Kinect method is still a new way in detecting potholes and in evaluating the pothole size and severity, but this method can't operate under the sunlight and it has a short range of detection. For vision technique, it's cheaper than most of the previously mentioned methods and can detect potholes time ahead. It can estimate the pothole size but can't evaluate pothole severity. To build an efficient pothole detection model, a pothole dataset is created for Lebanese and other countries' roads. The model of detection is improved to work under real-life conditions. The system has high processing capability and achieved high frame/second (FPS) to detect potholes time ahead. Our system is robust with acceptable precision and sensitivity/recall. This work can be used to report potholes from roads in realtime to responsible agencies, increase the safety of drivers by helping them detect potholes time ahead, and the proposed system could improve the performance of self-driving cars to ensure safe trips for passengers in the future.

EXISTING SYSTEM

The Existing pothole detection system has three techniques:

• Sensor-based techniques

• 3D reconstruction techniques

• Image processing techniques

DISADVANTAGES

• They are not suitable to be implemented on devices with limited hardware

• Sometimes, false positives and false negatives could be detected as potholes they cannot detect potholes.

• Require a high computational effort to reconstruct pavement surface.

PROPOSED SYSTEM

• The proposed system is automatic pothole detection system equipped with various efficient and accurate deep learning algorithms.

• It not only detects the potholes automatically but also alerts the driver, so that the driver can take necessary actions thus avoiding potential accidents .

• This system is a major milestone in self-driving cars where, in case of pothole detection the vehicle suspension can be adjusted and the impact on the vehicle can be reduced

ADVANTAGES

• Accidents due to pothole can be avoided.

• Driver will be intimated about potholes.

• In thr case of self driving cars, automatic speed can be controlled if pothole detected

DATA PREPROCESSING TECHNIQUES

• IMAGE CROPPING

Image cropping allows to explore target objects or concentrate on a single target region. It gives the freedom to pick selected region of interest within the image for the analysis.

• GRAYSCALE CONVERSION

Grayscale is simply converting images from colored to black and white. It is normally used to reduce computation complexity in deep learning algorithms.

• THRESHOLDING

Thresholding is a type of image segmentation, where we change pixels of an image, so as to select areas of interest of an image, while ignoring parts we are not concerned.

CONCLUSION The potholes detector using SSD-TensorFlow achieved 73% and 37.5% for the precision and recall respectively, and 32.5% for the mAP. YOLOv3 achieved high recall 78%, high precision 84%, and 83.62% mAP. YOLOv4 achieved high recall 81%, high precision 85%, and 85.39% mAP. The speed of processing for SSD was low and can't be used for real-time applications. YOLOv3 and YOLOv4 have a speed of processing of around 20 FPS and this can be considered high enough for our real-time application. Therefore, our potholes detector using YOLOv4 can be considered as a robust and realtime system that can be used in real-life scenarios. Our future work will include a larger dataset, more than 2000 images, for training and it has more images for potholes from different roads, with several severities and different lighting and weather conditions. Alexey AB [53] and many other researchers emphasized using a large dataset with more than 2000 images for training to get a robust object detector that can work under any conditions and circumstances. Moreover, in our future work, we will include

manholes in the training of our system. Manholes and potholes have common features and this important improvement to be done into our current system to differentiate between manholes as a pothole. Knowing that the current system was able to differentiate between manholes and potholes in most of the cases. Moreover, we aim to deploy our system into several cars to analyze the road condition in a live and real-time manner by adding GPS to get the coordinates of the pothole for maintenance. **REFERENCES**

[1] O. o. I. R. a. Development, "Distress Identification Manual for the Long-Term Pavement Performance Project," U.S Department of Transportation Federal Highway Administration, 2014. [2] I. G. V. P. Heggie, "Commercial Management and Financing of Roads," World Bank, Washington, 1998.

[3] A. M. Legreid, "Potholes and Strategies on the Road to Campus Internationalization," International Research and Review: Journal of Phi Beta Delta Honor Society for International Scholars, vol. 6, no. 1, 2016.

[4] A. F. Rita Justo-Silva, "Pavement maintenance considering traffic accident costs," International Journal of Pavement Research and Technology, 2019.

[5] S.-K. R. Taehyeong Kim, "A Guideline for Pothole Classification," International Journal of Engineering and Technology, vol. 4, 2014.

[6] B. X. Yu and X. Yu, "Vibration-Based System for Pavement Condition Evaluation," in Applications of Advanced Technology in Transportation, 2006.

[7] K. D. Zoysa, G. P. Seneviratne, W. W. A. T. Shihan and C. Keppitiyagama, "A Public Transport System Based Sensor Network for Road Surface Condition Monitoring," in SIGCOMM07: ACM SIGCOMM 2007 Conference, Kyoto, 2007.

[8] L. G. B. H. R. N. S. M. H. B. Jakob Eriksson, "The Pothole Patrol: Using a Mobile Sensor Network for Road Surface Monitoring," in Mobisys08: The 6th International Conference on Mobile Systems, Applications, and Services, Breckenridge, 2008.

[9] K. C. P. Wang, "Challenges and Feasibility for Comprehensive Automated Survey of Pavement Conditions," in Eighth International Conference on Applications of Advanced Technologies in Transportation Engineering (AATTE), Beijing, 2004.

[10] K. T. Chang, J. R. Chang and J. K. Liu, "Detection of Pavement Distresses Using 3D Laser Scanning Technology," in International Conference on Computing in Civil Engineering 2005, Cancun, 2005