PAPER • OPEN ACCESS

A Real-time Pothole Detection Based on Deep Learning Approach

To cite this article: Yeoh Keng Yik *et al* 2021 *J. Phys.: Conf. Ser.* **1828** 012001

View the article online for updates and enhancements.

IOP ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection−download the first chapter of every title for free.

A Real-time Pothole Detection Based on Deep Learning Approach

Yeoh Keng Yik, Nurul Ezaila Alias, Yusmeeraz Yusof and Suhaila Isaak[*]

Division of Electronic and Computer Engineering and Environmental Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia

*Email: suhailaisaak@utm.my

Abstract. Today, the number of vehicles using the road including highways and single carriage way is increasing. road structure safety monitoring system that is safe for road users and also important to ensure long-term vehicle safety and prevent accidents due to road damage such as potholes, landslides and uneven roads. Most news reports of road accidents are also caused by potholes that are almost 10-30 cm deep, coupled with heavy rainfall that reduces visibility among drivers, significant damage to the suspension system to the vehicle or unnecessary traffic congestion. In this paper, deep learning detection with YOLOv3 algorithm is proposed apart from researches ranging from accelerometer detection, image processing or machine learning based detection as it is easier to develop and provide more accurate results. After pothole has been detected in real-time webcam, the location will be logged and displayed using Google Maps API for visualization. a total of 330 sets of data were sampled for the implementation of the pothole detection training model. As the results, the model provided 65.05 mAP and 0.9 % precision rate and 0.41 recall rate. The limitation of YOLOv3 algorithm detection can be improve further using GPU with higher specification performances and can sample 1000 to 10,000 datasets. The proposed algorithm provides acceptably high precision and efficient pothole monitoring solution under different scenarios for the users and may benefit the public and the government to monitor pothole in real-time.

1. Introduction

In most of the developing country, paved road has been covering most of the surface area of the country and facilitated the connectivity of people from place to place in their daily routine. Hence that, the road is considered as one of the important infrastructure systems provided by the government for the transportation activities. Therefore, the conditions of road asset are important to ensure the user safety and bringing the excellent experience, which in turn reflecting the image of the country. However, the climate change will be affected the transportation system due to the increase in temperature, precipitation, sea level and storm surges will be resulting as road crack, potholes or other defects, thus risky the road users [1]. In addition, flood and rains may contribute to difficult driving conditions, increase in maintenance requirement to the pavement causing the traffic disruption impacts. Among the various types of road defects, potholes are the most dangerous one and bringing destructive side effect to the user. It is a common type of pavement surface distresses, which initiated by soil erosion and creating a hollow space below the asphalt layer and after frequently pressured by the overweight vehicle, the bowl-shaped hole is created after the asphalt layer are being broken up. Figure 1 shows the pothole on the road that may cause life threatening to the traffic user. The pothole can be worsened without

proper maintenance for a period of time. As a result, it will directly lead to several consequences such as a fatal road accident to motorcyclist, significant damage to the suspension system to the vehicle or lead to unnecessary traffic congestion if the potholes are not solved immediately.

According to the analysis of Malaysia road traffic death conducted, there are about 11.25 % of overall fatal road accident cases caused by a road defect, meanwhile 11.2 % of overall road accident is caused by the present of potholes on the road [2]. This problem is also included the problem with the paved road in rural area, which is significantly less monitored and therefore the probability of occurring potholes is usually higher than the road in urban area. As a result, the number of deaths caused by road traffic in rural area contributed 66 % compared to the cases in urban areas which account for the remaining 34 %. This scenario reflects that the road health and safety monitoring play an important role to control the issue of pothole.

Figure 1. Pothole on the road.

Traditional effort consists of relying government's Ground Penetrating Radar vehicle [1] or public effort to upload a picture on road abnormalities or report on maps navigation service platform such as Waze or i-Tegur [2,3] is no longer efficient since the previous method is limited by number of vehicles and do not covered on small road and latter is cumbersome to the publics. Therefore, effortless and less expensive solution [4] are required to ease the road structure monitoring which is not limited to pothole condition. As reported by other researcher, accelerometer detection is easy to set-up [5], but false detection is often happened on hump then may damage the vehicle suspension system [6]. Image processing-based detection offers less intensive computation hardware and acceptable result, but the algorithms is vulnerable to lighting variation and almost similar texture of minor pothole and road surface [7]. The machine learning based detection is the improvisation of previous solution [8], but it often required extensive time and knowledge to develop on feature extractors and somehow the accuracy can be still be improved [9]. Therefore, deep learning has been proposed to accomplish the easily developed with high accuracy detection which in turns can be installed into the embedded system of publics vehicle so that these transportations can help to monitor the pothole at real-time at minimal efforts. Many researches have been implemented on deep learning-based detection, which contributed extensive application on computer vision. Deep learning replicated the interconnection of biological neurons in human brains which consists of many interconnection layers, when certain raw data are feed in as input [10], some perceptron's or neurons will be triggered and the interconnected layer will lead to fully connected layer to produce the output. During the training, the weights and the biasing input at each perceptron are adjusted to minimize the function loss.

Convolutional Neural Network (CNN) is the most popular network structure that have been used in deep learning for computer vision. A typical CNN structure is built based on four main processes: convolution, pooling, non-linearity and classification [11]. Convolution process is repeated for all low-level features including edge, curves, straight line etc. until each of them become an independent complete feature map. Then, these features maps are combined to form convolutional layer in the first

layer which is used for feature extraction later. For pooling process, it is used to reduce the dimension of image data while maintaining important information on them so that overfitting problem can be solved eventually. Rectified Linear Units (ReLU) is another layer, which introduces nonlinearity process. It is an important process that replace negative value in feature maps into zero since most data are non-linear in real life. At the final stage, the classification process is done by the last layer which is fully connected layer. The fully connected layers will calculate the product between the weights of different classes with the previous layer to get the probability for each class [12]. Figure 2 illustrates an example of architecture of CNN that consists of multiple layers of convolution, pooling, ReLU and staked fully connected layers used to detect object.
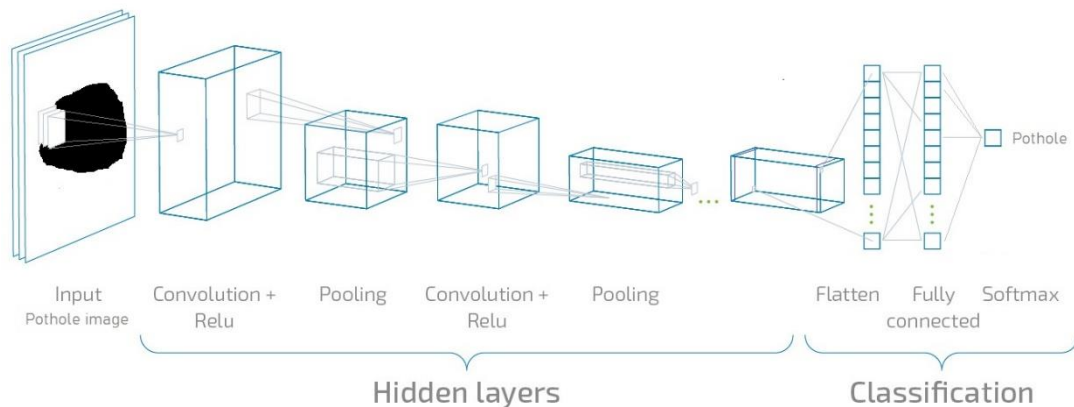

Figure 2. The CNN architecture [12].

The complexity of layer often determining the speed of inference and the result accuracy. You Only Look Once (YOLO) is used extensively as real time object detection system due to its extremely fast as compared to top detection systems such as Recurrent convolutional neural network (R-CNN). R-CNN use region proposal method which run detection on proposed region only after generates potential bounding boxes in an image. After that, post-processing is required to remove unwanted bounding boxes and rescore the bounding boxes based on other objects in the image. Each process needs to be trained separately increase the complexity and slow down the overall detection. In this paper, implementation of You Only Look Once version 3 (YOLOv3) with DarkNet-53 backbones using 53 layers of convolutional layer in the interconnected layer is reported [13]. The proposed system is used to locate the user the detected pothole on the road. Unlike recurrent convolution neural network (R-CNN), YOLOv3 pipeline each convolutional network on the entire image, thus resulting as fast and efficient data mapping. Then, it predicts the class probabilities and the bounding boxes concurrently if exceed threshold value and intersection over Union (IoU) [14].

2. Measurement System Architecture
The process of developing this pothole detection system is divided into three main part including the pothole detection model, data logger and the data visualization. The overall block diagram of pothole detection system is shown in Figure 3.
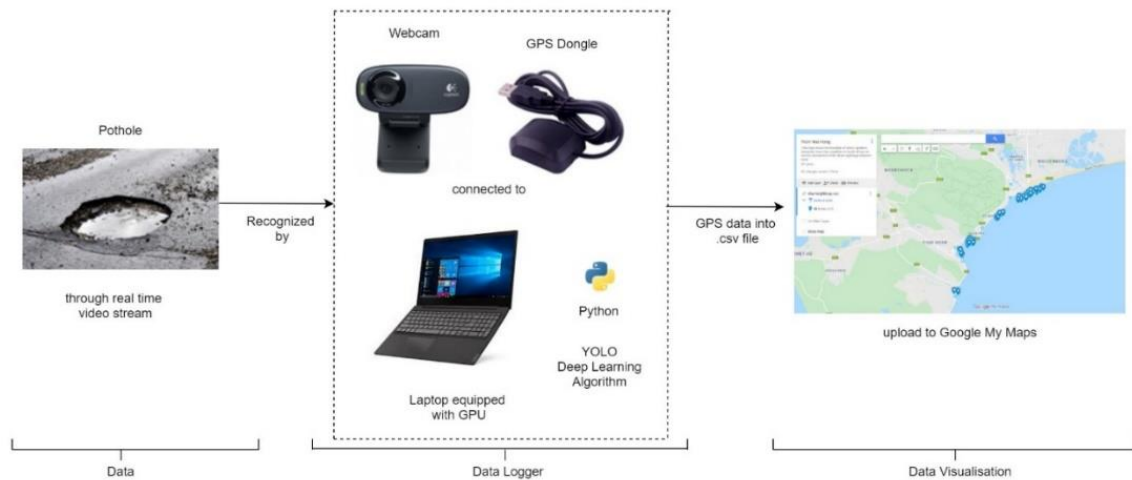
Figure 3. Block diagram of the proposed pothole detector module.

The image for the dataset is strictly limited to 416×416 pixel resolution so that it fulfil the input criteria of Darknet architecture. The raw image taken can be resized through OpenCV which is the open-source library for image processing in Python Environment. The pretrained model darknet53.conv11 is trained with 330 Jpeg formatted and pixel value of 416×416 datasets containing pothole raw images replicate in real-life, where 66 images is used as test dataset and 264 will be trained the dataset in generating the detection model. The dataset contains various size of pothole taken under different scenario including raining, low light and even slightly blurred condition. The training was implemented using Intel i7 laptop armed with Graphic Processor Unit of Nvidia GeForce GTX960M of computing ability of 5.0. The other open-source library like OpenCV 3.4 and Nvidia CUDA 10 is installed in 18.04LTS Ubuntu as the extension for training.

Before performing training, each dataset needs to perform annotation to draw the bounding box on pothole through labelling library and exported as .txt file placed together with jpeg files. Next the configuration on training was fixed at 2000 maximum epoch with learning rate of 0.001. Each batch of training consists of 32 images with subdivision of 2 to compromise with the 4 GB ram of GTX960M. The overall algorithm of data logger is depicted in Figure 4. After the detection model has been obtained, the weight file and the configuration file are exported for implementation. The 3MP PlayStation Eye camera is mounted below the car plate at 25° angle of depression to capture frame and transmit to laptop for real-time processing. The frequency of frame taken by camera is at 25 fps, meanwhile the inference is set at 4 FPS or 0.25s interval to reach time constrain of inference. At the same time, the remaining frame was not performed with inferencing. The multiprocessing library are utilising 3 cores processors which constantly read serial GPS data from VK-162 module, inferencing and write data into comma-separated value (CSV) file. When the pothole is detected, it will be bounded and cropped as single jpeg file. The .CSV file is updated whenever the pothole is detected. Therefore, bokeh library and pandas are called to translate the location logged and visualize on maps with Google Maps API. The value such as location (latitude and longitude) can be remove from the .CSV file when the pothole is patched. After the annotation are completed for all image, the original image with its respective annotated text file are loaded into the YOLOv3 network for training as shown in Figure 5.
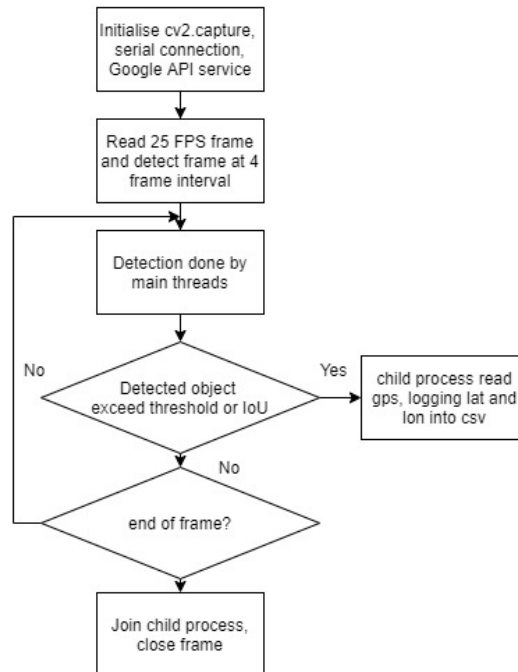
Figure 4. Flow chart of developing custom object detection model.

Figure 5. Annotation done on raw images using Labelimg.

3. Measurement Results and Discussions

The score of overall precision, and recall rate of this model are presented in Table 1 and Table 2. The controlling factors are the detection threshold value and the IoU. According to Table 1, when the detecting threshold is kept constant, increasing the IoU threshold will reduce ratio of correctly detected pothole to the total number of detections, thus lower IoU threshold is favoured.

Table 1. Map result for various IoU and detection threshold.

	Threshold			
IoU	0.20	0.25	0.30	0.35
0.20	65.05%	65.05%	65.05%	65.05%
0.25	64.27%	64.27%	64.27%	64.27%
0.30	63.49%	63.49%	63.49%	63.49%
0.35	62.05%	62.05%	62 05%	62.05%
0.40	58.35%	58.35%	58.35%	58.35%
0.50	49.32%	49.32%	49.32%	49.32%
0.75	6.72%	6.72%	6.72%	6.72%

Table 2. Recall rate for various threshold and IoU input.

	Threshold			
IoU	0.20	0.25	0.30	0.35
0.20	51%	45%	41%	36%
0.25	51%	45%	41%	36%
0.30	50%	45%	41%	36%
0.35	49%	45%	40%	35%
0.40	46%	42%	39%	34%
0.50	42%	39%	35%	32%
0.75	15%	15%	13%	12%

On another hand, while keeping IoU threshold constant, any increase to threshold value will increase the precision too. The lower IoU and higher threshold is needed for higher precision. However, the incremented threshold value indicates the reversed effect to the recall rate. Since the recall rate is the number of potholes detected over total number of pothole images fed, therefore lower threshold values are favoured. By keeping threshold value at constant lebvel, meanwhile increase the IoU threshold will reduce the recall rate, hence lower threshold value with lower IoU is preferred. In order to incorporate the precision rate with recall rate, the optimised threshold and IoU for the model is required. Therefore, the threshold between 0.25 to 0.3 and IoU between 0.2 to 0.25 is favoured to give the best performance on this detection model. The overall average precision also shows that this model able to obtain at least at 60 %, which is good enough to detect irregular shape of pothole with limited number of datasets. After setting the threshold and IoU, the width and height pixel value of webcam is fixed at 416×416, which is the best requirement for this model, any larger pixel frame will consume more hardware source. The frame at 0.25 s interval is only been processed even though the fps is taken by the webcam is 25 since the processing of frame is around 0.02 to 0.2, thus the processing is possible for real-time. Other than that, the vehicle is assumed will move at around 20 km/h to 60 km/h when encountered the road with potholes, thus following the simultaneous equation given below, detections are performed for every 1.14 m to 4.17 m. Any improved GPU will lead to lower processing time and give higher frame rate (with lower processing time), thus yielding more coverage. Optimization for the performance also included utilizing multiprocessing library to reduce the time spent in main processors for recording .CSV file and I/O requesting for GPS data. When program is terminated for each travel, the .CSV file will be logged with the detected pothole location. Visualization are done by calling to bokeh and pandas library and able to performance with operation such as zooming and panning.

4. Conclusion

A novel for real-time pothole detection using YOLOv3 deep learning algorithm has been studied and proven successful with its 65.65 mAP, 0.9 % precision rate and 0.45 % recall rate. As a conclusion, this setup can be installed in public transport or taxi with higher coverage and frequency of travelling. Thus, would help to detect the pothole and reduce the accidents due to pothole problem in Malaysia with high efficiency and accuracy system.

References
[1] Benedetto A, Tosti F, Ciampoli L B, and D'amico F 2017 An overview of ground-penetrating radar signal processing techniques for road inspections *Signal Processing*, **132** 201-209
[2] Khor S 2018 *You Can Now Report Potholes & Broken Traffic Lights To Local Authorities Via A Mobile App.* Available: https://says.com/my/tech/govt-launches-i-tegur-official-app-to-file-complaints-about-potholes
[3] Mednis A, Strazdins G, Zviedris R, Kanonirs G, and Selavo L 2011 Real time pothole detection using android smartphones with accelerometers *in 2011 International conference on distributed computing in sensor systems and workshops (DCOSS)*, IEEE pp. 1-6.
[4] Ateş H 2019 Pothole detection in asphalt images using convolutional neural networks (Middle East Technical University)
[5] Mednis A, Strazdins G, Zviedris R, Kanonirs G, and Selavo L 2011 Real time pothole detection using android smartphones with accelerometers *in 2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS)* IEEE pp 1-6
[6] Mohan P, Padmanabhan V N, and Ramjee R 2008 Nericell: using mobile smartphones for rich monitoring of road and traffic conditions *in Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems* pp 357-358
[7] Srivastava S, Sharma A, and Balot H 2018 Analysis and Improvements on Current Pothole Detection Techniques *in 2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*, IEEE pp 1-4
[8] Zapata J, Vilar R, and Ruiz R 2011 Performance evaluation of an automatic inspection system of weld defects in radiographic images based on neuro-classifiers *Expert Systems with Applications*, **38**(7) 8812-8824
[9] Valavanis I and Kosmopoulos D 2010 Multiclass defect detection and classification in weld radiographic images using geometric and texture features *Expert Systems with Applications* **37**(12) 7606-7614
[10] LeCun Y, Bengio Y, and Hinton G 2015 Deep learning *Nature* **521**(7553) 436-444
[11] Rosenberg I, Sicard G, and David E 2018 End-to-end deep neural networks and transfer learning for automatic analysis of nation-state malware *Entropy*, **20**(5) 390
[12] Pereira V, Tamura S, Hayamizu S, and Fukai H 2018 A Deep Learning-Based Approach for Road Pothole Detection in Timor Leste *in 2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)*, IEEE pp 279-284
[13] Redmon J, Divvala S, Girshick R, and Farhadi A 2016 You only look once: Unified, real-time object detection *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016*, pp 779-788
[14] Redmon J and Farhadi A 2017 YOLO9000: better, faster, stronger *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* pp 7263-7271