2021-05-03

Pothole Detection under Diverse Conditions using Object Detection Models

Ibrahim Hassan Syed
Technological University Dublin, ibrahim.h.syed@mytudublin.ie

Dympna O'Sullivan
Technological University Dublin

Susan Mckeever
Technological University Dublin, susan.mckeever@tudublin.ie

Recommended Citation

Hassan S., O'Sullivan D. and Mckeever S. (2021). Pothole Detection under Diverse Conditions using Object Detection Models. In Proceedings of the International Conference on Image Processing and Vision Engineering - Volume 1: IMPROVE, ISBN 978-989-758-511-1, pages 128-136. DOI: 10.5220/0010463701280136

Pothole Detection under Diverse Conditions using Object Detection Models

Syed Ibrahim Hassan[a], Dympna O'Sullivan[b] and Susan Mckeever[c]
Department of Computer Science, Technological University Dublin, Ireland

Keywords: Object Detection, Pavement Inspection, Deep Learning, Machine Learning.

Abstract: One of the most important tasks in road maintenance is the detection of potholes. This process is usually done through manual visual inspection, where certified engineers assess recorded images of pavements acquired using cameras or professional road assessment vehicles. Machine learning techniques are now being applied to this problem, with models trained to automatically identify road conditions. However, approaching this real-world problem with machine learning techniques presents the classic problem of how to produce generalisable models. Images and videos may be captured in different illumination conditions, with different camera types, camera angles and resolutions. In this paper we present our approach to building a generalized learning model for pothole detection. We apply four datasets that contain a range of image and environment conditions. Using the Faster RCNN object detection model, we demonstrate the extent to which pothole detection models can generalise across various conditions. Our work is a contribution to bringing automated road maintenance techniques from the research lab into the real-world.

1 INTRODUCTION

The assessment of road surface (termed road pavement) condition is a crucial task to ensure their usability and provide maximum safety for the public. The costs involved in maintaining pavements are significant – both to road users (over 60% of Irish people have had their chosen mode of transport damaged as a result of striking a pothole according to recent research) (ALDWORTH, 2018). The UK, councils allocate 75% funds for the maintenance of the local road condition and 25% for construction (Radopoulou and Brilakis, 2015). The two most common surface materials for road pavement are concrete and asphalt. Concrete roads are highly durable when compared to asphalt roads. Although concrete road surfaces last longer, repairing them is more complex. Holes or cracks cannot simply be patched–instead, entire slabs must be replaced. Asphalt paving is cheaper compared to concrete paving. It also creates a smoother drive, and provides better safety due to better traction and skid resistance. Asphalt is ideal for rural road pavements due to the ease of maintenance and repair, patching is simpler and faster than replacing entire slabs of roadways on less heavily trafficked areas such as country roads. But with only a 10-year lifespan, asphalt must be re-laid or repaired on a much more regular basis than concrete.

Pavement defects vary depending on the pavement surface. Pavement defects include cracking caused by failure of the surface layer. Surface deformation such as rutting that results from weakness in one or more layers of the pavement. Disintegration such as potholes caused by progressive breaking up of pavement into small loose pieces and surface defects, such as ravelling caused by errors during construction such as insufficient adhesion between the asphalt and aggregate particulate materials.

In this paper we focus on the detection and localization of potholes which are a common defect on both asphalt and concrete pavement. Potholes are a common cause of accidents and therefore require frequent inspection and timely repair. Pavement inspection usually consists of three main steps: 1) data collection, 2) defect identification and 3) defect assessment. The first step is largely automatic, carried out by specially adapted vehicles for surface surveying. However, the other two steps are largely manual. Images of road pavements are visually inspected by structural engineers or certified inspectors who assess

[a] https://orcid.org/0000-0002-0480-989X
[b] https://orcid.org/0000-0003-2841-9738
[c] https://orcid.org/0000-0003-1766-2441

Hassan, S., O'Sullivan, D. and Mckeever, S.
Pothole Detection under Diverse Conditions using Object Detection Models.
DOI: 10.5220/0010463701280136
In *Proceedings of the International Conference on Image Processing and Vision Engineering (IMPROVE 2021)*, pages 128-136
ISBN: 978-989-758-511-1

the road condition against the Pavement Condition Index (PCI). PCI is widely used in transportation civil engineering across the world and many authorities use it to measure the performance of their road infrastructure. It provides a numerical index between 0 and 100 which is used to specify general condition of pavements.

An automated defect detection and localization system could be a valuable tool for improving the performance and accuracy of the pavement inspection and assessment process as well as reduce the manual overhead of the current process. Such a system could be used to evaluate images or videos to assess pavement condition data. Additionally, an automated pothole detection system could be integrated into existing road inspection tools to support the inspection process by detecting road potholes from images acquired during the pavement inspection process.
This paper proposes a pothole detection method to detect and localize potholes on road surface. We describe the development of a Faster RCNN (Ren et al., 2016) model for pothole detection which is trained on a public potholes dataset (Kaggle, 2019) and tested for its generalizability on a number of other pavement image dataset(s). The contribution of our work is threefold: (1) We present an object detection model that can achieve accuracies of between 70% and 90%, for the task of detecting potholes in images; (2) We measure the impact of various real-world conditions on the accuracy of pothole detection models,and;(3) We publish three re-labelled public dataset(s) and contribute a labelled pothole image dataset to the field. The rest of this paper is organized as follows. In the next section related work on image processing and machine learning for pothole detection has been reviewed. In section 3 we describe experimental work to develop a trained model for pothole detection and the results of testing of this model on several dataset(s). In Section 4 we present our results and conclude with a discussion in Section 5.

2 RELATED WORK

With recent advances in deep learning, computer vision and image processing, research work has been carried out on pothole detection (Dhiman and Klette, 2019). Pothole detection methods are divided into three broad areas: vision based (Sawalakhe and Prakash, 2018) (Koch and Brilakis, 2011) (Ryu et al., 2015), vibration based (Yu and Yu, 2006) and 3D reconstruction methods (Hou et al., 2007) (Cao et al., 2020). Vision based methods rely on image or video data. This approach divides into two main approaches: image processing and machine learning. For machine learning, traditional algorithms that have been applied to this task have relied heavily on hand crafted features (Daniel and Preeja, 2014) (Hoang, 2018). To overcome these challenges, deep learning models, with their capability to extract visual features automatically from images rather than relying on hand crafted features, have become more popular. For example, object detection models are trained which can perform pothole detection by drawing a bounding box around the potholes. (Bhatia et al., 2019) investigate the feasibility and accuracy of thermal imaging in pothole detection. The proposed approach consists of 3 main steps: (1) data acquisition under various lightning condition (2) data augmentation, to increase the size of dataset (3) training a convolutional neural network (CNN) model. They compare the result of a self-built CNN model with pre-trained CNN models, with the pre-trained model achieving an overall detection accuracy of 97.8%. The main objective of this work is to find an efficient CNN model for pothole detection using thermal imaging. However, the drawback of their system is that it can only distinguish potholes and non-potholes on thermal images which are rarely used in road inspection.

(Ping et al., 2020) developed a pothole detection system by training on a pre-processed pothole dataset with four different models and compared the accuracy across each model. The method achieved 82% accuracy on YoloV3 (You Look Only Once: Version 3) (Farhadi and Redmon, 2018). However, the approach has not been tested on real world examples and their trained YoloV3 model does not accurately detect pothole on new images. The authors compare results of four different object detection models and found that SSD (Single Shot Multi-box Detector) gives higher accuracy but lower speed in comparison to the YoloV3 model. YoloV3 provides higher speed but lower detection accuracy and fails to detect small potholes. (Gupta et al., 2020) propose a pothole detection and localization system which can detect potholes by drawing a bounding box around potholes. The research utilizes thermal images and modified ResNet50 model and achieves 91.15% average precision in detecting potholes. However, their method can only work with thermal images which are rarely used by pavement inspection companies. In addition, the method can only detect potholes in images where the camera distance is relatively close to the pothole. In our work, we examine the problem of generalizability of trained models in the domain of road inspection. From the road inspection prospective, it is important to use models that consider the variety of conditions such as different road surfaces, pothole sizes, distance of pothole from imaging device. (Dharneeshkar et al., 2020) propose a pothole detection system on their own collected pothole dataset. The authors trained different versions of YOLO object detection model on 1500 images collected using a smartphone camera with a resolution of 1024 x 768. The proposed method can detect and localize potholes at different angles and on road surface. Similarly, (Ukhwah et al., 2019) train different versions of YOLO object detection model on images from a highway survey vehicle in Indonesia. To train the YOLO model, the authors use 448 grey scale potholes images. The method demonstrates that it can detect potholes from images that are acquired by using highway survey vehicles. The overall average precision achieved by the various models was 83.43%, 79.33% and 88.93% respectively.

From a commercial perspective, many vehicles are now being adapted to include automatic pothole detection system into their autonomous driving modules. For example, Jaguar Land Rover are researching automatic road pothole detection in which their vehicles not only detect potholes but can also identify the location as well as the severity of the pothole (Jaguar, 2015). Their proposed system can also send warning messages to the driver. FORD are also developing a pothole detection system where the system can warn drivers about the location of potholes (FORD, 2018). Another commercial application of automatic pothole detection was proposed by (Bansal et al., 2020). Their approach is based on vibration-based pothole detection which combines GPS, Internet of Things (IoT) sensors including accelerometers and gyroscope and machine learning. The IoT dataset was trained using a Support Vector Machine (SVM), Logistic Regression, Naïve Bayes, Random Forest and K-Nearest Neighbours (KNN). Random forest achieved highest accuracy in detecting potholes 86.8%. The objectives of the systems are to reduce injuries and deaths, alert drivers about potholes before driving over them, share pothole location data with government and civil authorities to repair pothole in timely manner and to build a real time map which updates according to latest road conditions.

The task of pothole detection has many different variations – the variety of potholes, lighting levels, distance of the pothole from camera, shot angle, imaging device and weather conditions. Therefore, a useful road maintenance prediction model should be able to generalise. In our approach we used a publicly available Kaggle dataset (Kaggle, 2019) to train a prediction model for pothole detection which was subsequently tested on a number of other dataset(s) that represent a variety of real-world variations. Our experimental set up and results are presented in the next sections.

3 METHODOLOGY

This paper proposes a method for automatically detecting the presence and location of potholes in an image, whilst also considering the variety of real world conditions that can occur during the automatic pavement assessment process. Using supervised machine learning, we train an object detection model using an image dataset containing potholes. We control the number of each variation in training dataset and then do controlled testing of these conditions using test sets because we wish to check how well a model can generalize to other datasets so it is necessary to control the variation of training and testing samples.

This section describes the details of preparing training and testing data, object detection model architecture and training and testing steps. The training dataset used in this research was the Kaggle pothole dataset, a publicly available dataset with one positive class (pothole present). The dataset contains 618 images. Due to the duplication of images in the dataset, 280 unique images were selected for training. Each image was labelled by drawing a bounding around the pothole using the LabelImg tool (LabelImg, 2015). The model was trained under controlled settings and tested using 4 other datasets described in detail below. Table 1 shows details for the training dataset, including four characteristics of training images which could challenge model generalizability: acquisition device, distance, distance from camera and lighting level. The values or categories for each is provided (e.g. close distance, medium distance and far distance which means pothole is either close from camera angle or it is too far from camera angle or it is near (medium) from camera angle). Figure 1 shows the example of each camera angle.

Table 1: Breakdown of the Kaggle Pothole Dataset used for Training.

Index Name	Description
Training Data	Kaggle Pothole
Acquisition Device	Smartphone/Digital Camera
Image Size	500 x 500
Distance	Close = 1 meter
	Medium = 2∼3 meters
	Far = 5∼10 meters
Lighting Level	Normal: 228, Low: 52
Total Images	280

Figure 1: Example of different camera angles: Top left: Close, Top right: Medium, Bottom: Far.

3.1 Data Preparation

Training images for our model require the location of the pothole(s) in each image to be annotated. This was manually applied to the 280 training images by creating a bounding box around the area of interest i.e. a pothole. Below are the steps that were performed for the training data preparation:

Step 1: Create a dataset using labelImg, generating XML files that contain information about object coordinates, image name, image width and height and object name.

Step 2: Convert these XML files into CSV file which store each XML file details.

Step 3: Convert CSV file to Tensorflow's own binary storage format tf-record i.e. train-record and test-record. TensorFlow's object detection API requires the data to be in the 'tfrecord' format. The tfrecord format enables splitting, creating batches, shuffling data and providing a uniform format across network architectures and systems.

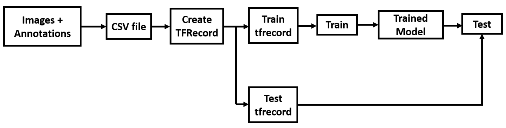

Figure 2: Implementation pipeline for pothole detection using Tensorflow object detection.

3.2 Testing Datasets

The focus of this research is to determine how an object detection model trained on one dataset (with a known range of variations) can perform against a variety of conditions via a number of testing dataset(s). First, we test if the image can distinguish between positive (contains pothole) and negative (does not contain a pothole) images. Then we test for a number of conditions ranging from different image sizes, different image types (stereo images) and different lighting conditions.

The details of the 4 test dataset(s) used for testing are described below. Three dataset (s) (Negative Images (Saxena, 2019), Cranfield (Alzoubi, 2018), Pothole-600 (Fan et al., 2020)) are publicly available. The fourth dataset is our own data, collected by acquiring the pothole images from different streets of Dublin. In each test dataset bounding box labels are not included, therefore we manually labelled each image by drawing a bounding box around the region of interest using the LabelImg tool. Table 2 shows the details of 3 testing dataset(s).

***A)* Negative Images:** This dataset (Saxena, 2019) provides a number of negative images. Due to the presence of other objects in images such as vehicles, trees or people, the pavement surface has been cropped from each image.

***B)* Cranfield Pothole Dataset** (Alzoubi, 2018)**:** This dataset provides images of potholes on asphalt pavement. The reason for choosing this dataset for testing is its small image size i.e. 300 x 300 pixels, as the model was trained on images of size 500 x 500 pixels.

***C)* Pothole-600 Dataset** (Fan et al., 2020)**:** This dataset provides pothole images which were acquired using a stereo camera. The reason for choosing this dataset for testing is that the camera source and image type are different from the training set, as all images in the training set are acquired using smartphone camera or digital camera.

***D)* Dublin Road Dataset:** This dataset was acquired from pavements in Dublin using a smartphone camera under daylight conditions. In total we collected 40 images with a resolution of 3648 x 2736. The reason for collecting this dataset is to test images with normal and low lighting condition. All images in this dataset were collected during daylight hours, as acquiring road images is not an easy or practical task in darker environments. Therefore, an artificial low lighting effect have been applied on the same testing set.

Table 2: Testing dataset details.

Test Dataset	Image Size	Total Images
Cranfield (Alzoubi, 2018)	300 x 300	50
Cranfield	400 x 400	50
Pothole-600 (Fan et al., 2020)	400 x 400	50
Dublin Roads	3648 x 2736	40

Figure 3: Example images of each testing dataset: Top left: Pothole-600, Top right: Cranfield Pothole Dataset, Bottom: Dublin Road Dataset.

3.3 Pothole Detection with Faster RCNN

To address the object detection problem for pothole detection we have trained a state-of-the-art object detection model Faster RCNN (Ren et al., 2016). Faster R-CNN has two stages for detection. In the first stage, images are processed using a feature extractor (e.g. VGG, MobileNet) called the Region Proposal Network (RPN) and simultaneously, intermediate level layers (e.g.," conv5") are used to predict class bounding box proposals. In the second stage, these box proposals are used to crop features from the same intermediate feature map, which are subsequently input to the remainder of the feature extractor in order to predict a class label and its bounding box modification for each proposal. We also use Inception-V2 architecture as a backbone of Faster RCNN model. Inception architecture has yielded better results than a conventional CNN architecture. Additionally, Faster R-CNN model combined with Inception CNN architecture shows an improvement in detection accuracy. The choice of this network was motivated by the fact that it achieves good results on different dataset(s) such as Microsoft Common Object Context (MS COCO) (Lin et al., 2014) and PASCAL VOC (Everingham et al., 2010). Furthermore, it offers a structure that can be modified according to specific task needs. Table 1 shows the details of the Kaggle dataset which is used to train the object detection model. After training the model, we tested the trained model using 4 dataset(s) to investigate various parameters – negative images, image sizes, images acquired from multiple sources and lighting levels. In our experiments model training and testing is done using Python, and the Tensorflow object detection API. In this work we did not use data augmentation. For training, a NVIDIA GeForce RTX 2070 GPU was used. All experiments are performed under Windows 10 on Intel Core i7-9750 with 16GB of DDR4 RAM. The model was trained with Adam optimizer and L2 regularization, using an initial learning rate to 0.00001. The modelling process was done in approximately 2 hours with 20k iterations. During the training process we observed that after the 20k iteration, the training loss did not decrease substantially, so we stopped the training and saved the model parameters for the testing purpose.

3.4 Evaluation Protocol

The performance of the developed model was evaluated on 4 datasets as described in Section 3.2. For each testing set, results generated from the Faster RCNN model were compared with the actual ground truth. Several researchers have proposed different evaluation methods for the object detection task (Padilla et al., 2020) (Zhao et al., 2019). In this paper, Intersection over Union (IoU) (also known as the Jaccard index), precision and recall are used to evaluate trained models. IoU measures the overlap between the actual ground truth bounding box and the predicted bounding box. We defined an IoU threshold of 0.5 which means if the overlap between an actual and predicted bounding box is $<0.50\%$ the model will consider it as false positive whereas, if the overlap between actual and predicted bounding box is $>0.50\%$ the model will consider it as true positive. In this way precision and recall are calculated at 0.5 IoU thresholds. Increasing IoU threshold results in higher precision but lower recall. Conversely, decreasing IoU threshold gives higher recall. For example, if we IoU threshold set to 0.9 then we get higher precision which means model can detect potholes if the overlap between actual and predicted bounding box is $>=0.9\%$ the model will consider it as true positive. Figure 4 shows an example of an actual and predicted bounding box used for IoU calculation. A high IoU threshold is not required as the exact placement of the pothole relative to predicted area just needs to be enough to say that a pothole exists in the area. The task does not require precise pothole perimeter discovery. Figure 5 shows precision and recall results from experiment 4 at IoU thresholds from 0.5 - 0.9. While precision does not vary greatly at different IoU thresholds, relaxing the IoU threshold results in higher recall values.

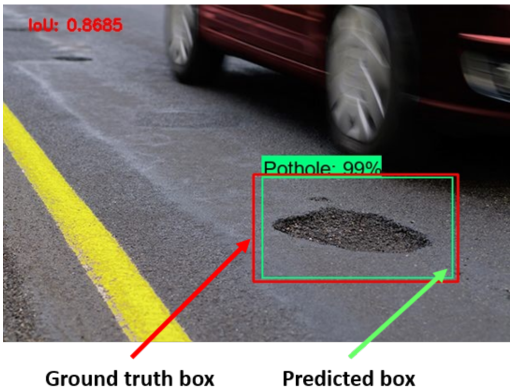

Figure 4: Example of Intersection over Union (IoU).

4 EXPERIMENTAL RESULTS

In this section we describe a series of experiments conducted on a variety of dataset(s). We implemented four experiments - within each, we focus on specific variable conditions for are encountered on the pothole detection task: distinguishing positive and negative images, variety of image size, images captured with different devices and variations in lighting level.

4.1 Experiment 1: (Positive and Negative Images)

This experiment shows the performance of the trained model on negative images i.e. no pothole on the pavement surface. 50 images were passed to the trained model and 45 of those images were predicted correctly. In order to check the reason of wrong detection we check all images manually and found that the errors were due to the appearance of shadows on the road surface where the model considered these dark patches as potholes.

Table 3: Results of pothole detection on negative images.

Non-pothole	Pothole	Accuracy
45	5	90%

4.2 Experiment 2: (Smaller Images)

In this experiment we used the Cranfield dataset and tested with 50 images of size 300 x 300 pixels and 50 images of 400 x 400 pixels. As our model is trained on images of size 500 x 500, in this experiment we are investigating how well our trained model can generalize to detect potholes on smaller images. Initially we have tested images of size 500 x 500 in order to check the model performance on the same image size as those of the training set. We then resize images to test smaller size images. Table 4 shows that model achieve almost same results on both smaller size images, with just a slight improvement on the (trained) 500 x 500 image size.

Table 4: Results of pothole detection on small image.

Model	Backbone	Image Size	Precision	Recall
Faster RCNN	Inception V2	500 x 500	79%	94%
		400 x 400	80%	92%
		300 x 300	79%	92%

4.3 Experiment 3: (Images Captured by Stereo Camera)

The purpose of this experiment is to test the model on a different image type: stereo images. However the images in this dataset were varied in terms of the distance of the pothole from the camera so we also had to check for any effect caused by this feature. Therefore, we conducted 2 experiments. In the first experiment we randomly selected 50 images from the dataset and achieved the results in Table 5.

Table 5: Results of pothole detection on stereo images.

Model	Backbone	Image Size	Precision	Recall
Faster RCNN	Inception V2	400 x 400	77%	55%

Analysing the results in Table 5, the detection performance of the model is low. To understand this further, we conducted a second experiment where we divided the 50 images into two test sets based on distance - 25 images with a close distance between the camera and pothole and 25 images with a medium distance between the camera and pothole. We compared the results to a test set with 50 non-stereo images divided into two testing sets of 25 images with close and medium distance to the camera.

Table 6: Results of pothole detection on stereo and non-stereo images with close and medium distance.

Model	Backbone	Testing Dataset	Distance	No. Images	Precision	Recall
Faster RCNN	Inception V2	Stereo	Medium	25	95%	84%
			Close	25	65%	52%
		Non Stereo	Medium	25	83%	100%
			Close	25	63%	48%

Results in table 6 shows that the image type or source does not affect detection accuracy, rather the distance from the pothole to the camera affects the detection performance. In both experiments it can be seen that images with medium distance have high detection accuracy compared to those at close distance. The main reason for this is that our training set contains only a small sample of images where the distance from the camera to the pothole is close.

4.4 Experiment 4: (Images with Different Lighting Conditions)

This experiment uses images that were collected across Dublin city centre in daylight. A total of 40 images were collected with normal lighting levels. In order to test different lighting conditions, we first tested the original 40 images collected in daylight and then applied an artificial low lighting effect to the same 40 images and retested on those. Table 7 shows the results of experiment 4. Analyzing the results in table 7 the detection performance of the model on normal images is slightly lower than on low lighting level images. The reason for low accuracy on normal lighting images is because the model is unable to detect potholes on two images in the normal lighting dataset. However, in the low lighting dataset, the model correctly detected pothole on those two images.

Table 7: Results of pothole detection on Dublin roads dataset.

Model	Backbone	Testing Dataset	Precision	Recall
Faster RCNN	Inception V2	Dublin (Normal Light)	78%	68%
		Dublin (Low Light)	78%	73%

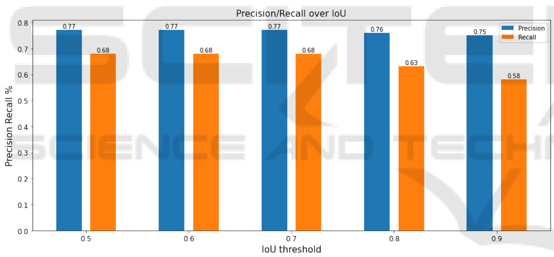

Figure 5: Comparison of Precision and Recall at different IoU threshold values.

5 DISCUSSION

The process of pothole detection for road maintenance is still largely done though manual visual inspection of images or videos acquired using cameras or professional road assessment vehicles. This is a time consuming and expensive task. The task is also complicated by variations in images such as different image types and sizes, camera types, lighting levels and distance of the pothole from the camera.

From experiment 1 we found that shadows and manhole covers on the pavement surface may be identified as false positives (i.e. the model falsely identifies then as potholes) and as such this adversely affects the model's performance. From the second experiment, we conclude that image size is not a major factor which could affect the detection performance as results are similar across all sizes. From the third experiment we conclude that the camera source is not a major factor affecting model performance. However, the distance of a pothole from the camera device has an impact on the detection rate. From experiment 3 we conclude that the model has difficultly in identifying potholes that are close to the camera. From the 4th experiment we notice that lighting effect is not having a major impact on the detection rate as long as the distance of the pothole from camera device is not close.

We believe that our results could be improved by more labelled training data and more balanced training samples, particularly with respect to images taken at different distances from the image capture device.

6 CONCLUSIONS

In this paper we investigated some common issues which are likely affect the generalizability of any model for automated pavement assessment. We determine the variations of images in terms of image size, distance of pothole from camera angle and lighting effect.

We trained an object detection model using the Kaggle pothole dataset. We used Faster RCNN with Inception V2 as a backbone model for object detection. To check model generalizability we used a variety of conditions including small image sizes, different image types and lighting effects. We have attempted to identify factors that may impact a generalizable model for pothole detection. These factors include image size, camera source, lighting levels and distract from the camera. To investigate these factors, we conducted four experiments. In each experiment we explore one condition.

We conclude that distance of a pothole from a camera device plays an important role when aiming to create a generalizable model for pothole detection. A further problem for the pothole detection task is the presence of other objects in the images (e.g. manhole covers) that may be falsely identified as potholes.

Our work is currently limited to detecting a single pavement defect- potholes. This can be extended to detect multiple pavement defects such as cracks, patches, and ruts. The trained model could be deployed in a number of ways - for offline detection on batches of images or on a smartphone or other hardware such as Nvidia Jetson Nano for real time pothole detection. Regardless of the mode of implementation, an automatic pothole detection method may help to speed up and lower the cost of the pavement inspection process which currently relies heavily on manual human expertise.

In future work we will train with larger dataset(s). In particular, we will train with more samples of images containing potholes that are close and far from the camera. We will also use images that includes potholes as well other objects including shadows, manhole as well as other common objects in pavement imagery. In this work we have focused on open-source images, but there are specified images that are collected through commercial road inspection vehicles and provide more consistent images of pavements which can help to build more robust model specially for the task of automatic road inspection. Recently other sources such as drone shots and vehicle windscreen cameras are being used to collect pavement data. Such images often contain a multitude of objects such as vehicles, trees, traffic signs and or/people. Such data would require pre-processing to extract these objects before training for the pothole detection task. Training object detection models with larger dataset(s) will require very high computational power and need more training time. We will also experiment with tuning the hyper parameters and training the model with other feature extraction networks.

ACKNOWLEDGEMENT

This work was funded by Science Foundation Ireland through the SFI Centre for Research Training in Machine Learning (18/CRT/6183).

REFERENCES

ALDWORTH, B. (2018). Road users bearing the cost of damaged surfaces. https://www.theaa.ie/blog/road-users-bearing-cost-damaged-surfaces/.

Alzoubi, A. (2018). Potdataset. https://figshare.com/articles/PotDataset/5999699/1.

Bansal, K., Mittal, K., Ahuja, G., Singh, A., and Gill, S. S. (2020). Deepbus: Machine learning based real time pothole detection system for smart transportation using iot. *Internet Technology Letters*, 3(3):e156.

Bhatia, Y., Rai, R., Gupta, V., Aggarwal, N., Akula, A., et al. (2019). Convolutional neural networks based potholes detection using thermal imaging. *Journal of King Saud University-Computer and Information Sciences*.

Cao, W., Liu, Q., and He, Z. (2020). Review of pavement defect detection methods. *IEEE Access*, 8:14531–14544.

Daniel, A. and Preeja, V. (2014). Automatic road distress detection and analysis. *International Journal of Computer Applications*, 101(10).

Dharneeshkar, J., Aniruthan, S., Karthika, R., Parameswaran, L., et al. (2020). Deep learning based detection of potholes in indian roads using yolo. In *2020 International Conference on Inventive Computation Technologies (ICICT)*, pages 381–385. IEEE.

Dhiman, A. and Klette, R. (2019). Pothole detection using computer vision and learning. *IEEE Transactions on Intelligent Transportation Systems*, 21(8):3536–3550.

Everingham, M., Van Gool, L., Williams, C. K., Winn, J., and Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88(2):303–338.

Fan, R., Wang, H., Bocus, M. J., and Liu, M. (2020). We learn better road pothole detection: from attention aggregation to adversarial domain adaptation. In *European Conference on Computer Vision*, pages 285–300. Springer.

Farhadi, A. and Redmon, J. (2018). Yolov3: An incremental improvement. *Computer Vision and Pattern Recognition, cite as*.

FORD (2018). Ford pothole detection system. https://www.ford.co.uk/shop/research/technology/safety-and-security/pothole-detection-system.

Gupta, S., Sharma, P., Sharma, D., Gupta, V., and Sambyal, N. (2020). Detection and localization of potholes in thermal images using deep neural networks. *Multimedia Tools and Applications*, 79(35):26265–26284.

Hoang, N.-D. (2018). An artificial intelligence method for asphalt pavement pothole detection using least squares support vector machine and neural network with steerable filter-based feature extraction. *Advances in Civil Engineering*, 2018.

Hou, Z., Wang, K. C., and Gong, W. (2007). Experimentation of 3d pavement imaging through stereovision. In *International Conference on Transportation Engineering 2007*, pages 376–381.

Jaguar (2015). Pothole detection technology by jaguar land rover. https://www.landrover.com/experiences/news/pothole-detection.html.

Kaggle (2019). Web scrapped road images for pothole detection. https://www.kaggle.com/sachinpatel21/pothole-image-dataset.

Koch, C. and Brilakis, I. (2011). Pothole detection in asphalt pavement images. *Advanced Engineering Informatics*, 25(3):507–515.

LabelImg (2015). Tzutalin.labelimg.gitcode. https://github.com/tzutalin/labelImg.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer.

Padilla, R., Netto, S. L., and da Silva, E. A. (2020). A survey on performance metrics for object-detection algorithms. In *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 237–242. IEEE.

Ping, P., Yang, X., and Gao, Z. (2020). A deep learning approach for street pothole detection. In *2020 IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService)*, pages 198–204. IEEE.

Radopoulou, S. C. and Brilakis, I. (2015). Patch detection for pavement assessment. *Automation in Construction*, 53:95–104.

Ren, S., He, K., Girshick, R., and Sun, J. (2016). Faster r-cnn: towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6):1137–1149.

Ryu, S.-K., Kim, T., and Kim, Y.-R. (2015). Image-based pothole detection system for its service and road management system. *Mathematical Problems in Engineering*, 2015.

Sawalakhe, H. and Prakash, R. (2018). Development of roads pothole detection system using image processing. In *Intelligent Embedded Systems*, pages 187–195. Springer.

Saxena, S. (2019). Pothole detection. https://github.com/shubhank-saxena/Pothole-Detection.

Ukhwah, E. N., Yuniarno, E. M., and Suprapto, Y. K. (2019). Asphalt pavement pothole detection using deep learning method based on yolo neural network. In *2019 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, pages 35–40. IEEE.

Yu, B. X. and Yu, X. (2006). Vibration-based system for pavement condition evaluation. In *Applications of Advanced Technology in Transportation*, pages 183–189.

Zhao, Z.-Q., Zheng, P., Xu, S.-t., and Wu, X. (2019). Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, 30(11):3212–3232.