

CMPE 257 - Project Report

**Potholes Detection using Machine Learning**

**Submitted By:**

**Deepali Upadhyay (011809694)**

**Mayank Dhingra(013773942)**

**Shivani Parate(013729612)**

**Under the guidance of:**

**Dr. Prof Jerry Gao**

**Table of Contents**

[**Chapter 1 Introduction**](#_x68pfwdqkqdj) **3**

[1.1 Overview of current scenario](#_4mh7q01i2siy) 3

[1.2 Objectives of our project](#_cxb7ywhlyec7) 4

[1.3 Difference between our work and existing works](#_h192lsm7p097) 4

[1.4 Report Organization](#_brrxfgwbhswm) 4

[**Chapter 2 Related Work**](#_kbrl8376h3m) **4**

[**Chapter 3 Training and Test Data Preparation**](#_d8dem8tmwe75) **5**

[3.1 Dataset Statistics](#_5fmrfqbuqr8o) 5

[**Chapter 4 Machine Learning Models and AI Algorithms**](#_9r04qtobmwqb) **5**

[4.1 Overview](#_nw3nv72gjb6n) 5

[4.2 Model 1](#_jo09cu97rdts) 5

[4.3 Model 2](#_sixpxvoji9jc) 5

[4.4 Model 3](#_nb56xuqhgj5f) 5

[**Chapter 5 Implementation**](#_kr1sw5tr7eeq) **5**

[**Chapter 6 Results**](#_m2u7buq6u9ki) **5**

[**Chapter 7 Comparison of the models used**](#_rn2emjjpjpxh) **5**

[**Chapter 8 References**](#_i6z993f4hb3t) **5**

# 

# 

# 

# Chapter 1 Introduction

## 1.1 Overview of current scenario

Potholes are a structural damage to the road with hollow which is usually caused by poor drainage system on the roads which allows water to go under the soil structure of the road. Due to unpredictable weather these days, the damage to the roads is increasing. According to the statistics, one-third of the accidents are caused by bad road conditions. According to American Automobile association, almost 16 million drivers have suffered some kind of damage due to potholes like tire puncture, bent wheels , damage to suspensions and shockers. The bicyclist face greatest dangers from these as they can easily lose balance and get hit by incoming traffic. So, it's really important to have a reliable pothole detection system to correctly identify the pothole which can later be used to warn the drivers.



*Fig: 1 Example of a pothole*

Current pothole detection methods fall in three important categories:

a.) Vibration Based Methods - They use accelerometers to sense the pothole. The accelerometer senses the force on the sensor. The advantage of this method is low cost maintenance of the equipment and the data captured can be directly used to know the existence of pothole. The disadvantages are the low response time and the need to go through the pothole to get the readings.

b.) 3D Reconstruction - They use 3D laser scanners to create a 3D image of the pothole. The advantage of this methods is that potholes can be detected in real time and the data capured is quite accurate. The only disadvantage is the high cost of the equipment used which is a challenge.

c) Vision Based Methods - These methods use expensive equipment like cameras to capture the images in real time. The images are then processed using complex algorithms to get the desired data. The disadvantage of this method is high maintenance cost of the hardware. The advantage of this method is that characterization of the pothole can be done in addition to its detection. Also there is no need to go over the pothole to capture the data.

## 1.2 Objectives of our project

We plan to develop a system based on vision which can detect the potholes with at least 85% accuracy. Additionally, we plan to measure the depth of the pothole. We also plan to connect our dataset with the google api to capture the latitude and longitude of the potholes which can later be used to know the location of the pothole on the map.

## 1.3 Difference between our work and existing works

Traditionally, the pothole detection project only deals with the identification of the pothole on the street. No studies have been performed to get the statistics of the pothole in terms of size and depth. We plan to include the statistics as well along with the detection of the pothole.

## 1.4 Report Organization

We have organized our report in eight chapters.

Chapter 1 deals with the objective and motivation of our project.

Chapter 2 summarizes the research work already done in this area.

Chapter 3 talks about the dataset used and the statistics of the data.

Chapter 4 summarizes the different machine learning models used to solve the problem. Also tells about the advantages and disadvantages of all the different methods.

Chapter 5 lists down the implementation method of the algorithm and the model. Also it tells about how the performance of the model was improved.

Chapter 6 shows the results achieved through each of the different models.

Chapter 7 compares the different models based on different performance factors.

Chapter 8 lists down the references used to create this project and the report.

# 

# 

# 

# Chapter 2 Related Work

There is a lot of research work done in this area.

There was a proposal to use SVM (Support Vector Machine)

A traditional machine learning method based on SVM or Support Vector Machine was proposed [3] for road pothole detection task. In this experiment they extracted the image region feature based on the histogram and non linear kernel SVM is added up to identify the target. The result shows that in this experiment a pothole can be well and highly recognize.

While in paper [4], a deep learning-based particularly convolutional neural network was used as classifier for crack damage detection from concrete images. They build a classifier which can be less influenced by the noise

caused by illumination, shadow casting and so on. The advantages of this experiment are it automatically learn the feature without conducting any feature extraction process and computation compared to the conventional methods.

Hiroya Maeda, et al. [5] introduced road damage detection using deep neural networks with images captured through a smartphone. They developed a new large-scale dataset for road damage detection and apply the end-to-end object detection method based on deep learning to the road surface damage detection problem, and verify its detection accuracy and processing speed for road damage detection and classification.

A binary classification method using neural network [6] was presented for classifying whether the road images are belonging to crack or normal road images. The network works by feed by the feature of the images before performing the classification.

Another convolutional neural network model called Crack-Net [7] was proposed for crack detection in pixel-level pavement. Different from common CNNs, pooling layers does not include in this model. Experiment result shows this method efficiently work in crack detection task.

A low-cost sensor and deep CNN-based was proposed [8] for automatic crack detection. This experiment presented a CNN model which can automatically learn the feature without any feature extraction process. The input images manually annotated before feed to the model.

Finally, a real - time automatic pavement crack and pothole recognition system for mobile Android-based devices was proposed by A. Tedeshi & F. Benedetto [9].

# 

# Chapter 3 Training and Test Data Preparation

## 3.1 Dataset Statistics

We found a dataset which was compiled at Electrical and Electronic Department, Stelllenbosch University in 2015. The entire dataset consists of two different sets, one was considered to be simple and the other more complex.  
These datasets do share some files and there are a few instances where two different images would have the same name.  
Therefore, the appropriate measures need to be taken if the data is combined into one larger dataset.

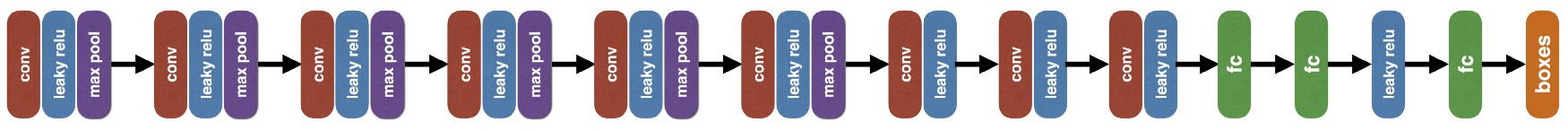
All the potholes are labeled with the help of any labeling tool

# Chapter 4 Machine Learning Models and AI Algorithms

## 4.1 Overview

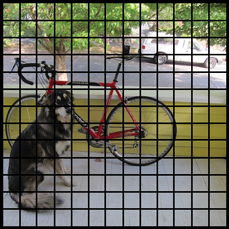
## 4.2 Model 1

YOLO (You Only Look Once), is an object detection algorithm that uses just a single neural network for object detection. The object detection task is divided into two parts. The first part is determining the location of the objects present on the image and the second part is classifying that object. All the previous algorithms like R-CNN use a pipeline to perform these simple tasks in multiple steps which was slow to run and hard to optimize. YOLO made it easier by using a single neural network for both classification and localising the object using bounding boxes to solve this problem. This is the architecture of YOLO :



**Fig: YOLO Architecture**

The end result is a tensor value of 13\*13\*30 where 13\*13 are the number of grids on the image and 30 is the size of a single grid in this case.

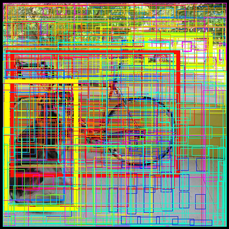


**Fig: Image divided into grids**

Each of these cells is responsible for predicting 5 rectangles that encloses an object. It also gives the confidence score that tells how certain is that the predicted object location of bounding box actually encloses some object. The image after predicting all the boundary boxes may look something like this.

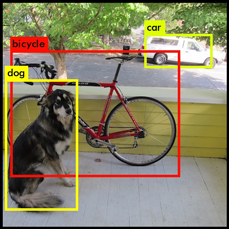


The confidence value can be determined by the thickness of the bounding box. This grids not only predicts the bounding boxes but also predicts a class. A final value is calculated by combining the class prediction and confidence score of the bounding boxes which gives us the probability of the presence of a class label in that bounding box. For example, the big fat yellow box on the left is 85% sure it contains the object “dog”:



Out of all the 169 grid cells or 845 bounding boxes, we get only three bounding boxes if we set the threshold confidence score to 0.3 which means we only get those bounding boxes whose confidence score is greater than 0.3.

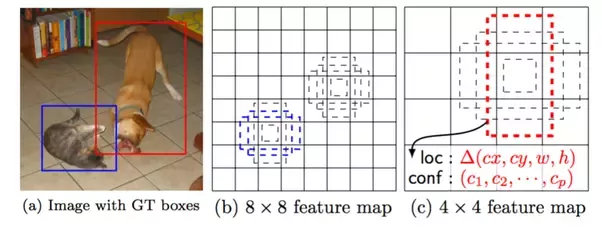
The final bounding boxes are:



## 4.3 Model 2

SSD (Single shot detector)

This algorithm does the localization and classification in single forward pass and is similar to that of YOLO as it divides the image into grids of equal sizes



This above image explains how SSD works. It matches labels with default boxes of different aspects as dashed rectangles. A number is associated with every element of the feature map. Any default box with an [IOU](http://www.pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/) of 0.5 or greater with a ground truth box is considered a match. Two of the 8x8 boxes are matched with the cat (shown in blue), and one of the 4x4 boxes is matched with the dog (shown in red). It is important to note that the boxes in the 8x8 feature map are smaller than those in the 4x4 feature map: SSD has six feature maps in total, each responsible for a different scale of objects, allowing it to identify objects across a large range of scales.

For each default box in each cell, the network outputs:

* A probability vector of length c, where c is the number of classes, representing the probabilities of the box containing an object of each class (including a background class indicating that there is no object in the box).
* An offset vector with 4 entries containing the predicted offsets required to make the default box match the underlying object’s bounding box. They are given in the format (cx, cy, w, h) – centre x, centre y, and width & height offsets, and are only meaningful if there actually is an object contained in the default box.

## 4.4 Model 3

Faster R-CNN ()

Faster R-CNN has two networks: region proposal network (RPN) for generating region proposals and a network using these proposals to detect objects.

How it works-

1) Run the image through a CNN to get a Feature Map  
2) Run the Activation Map through a separate network, called the Region Proposal Network(RPN), that outputs interesting boxes/regions  
3) For the interesting boxes/regions from RPN use several fully connected layer to output class + Bounding Box coordinates.

# Chapter 5 Implementation

# Chapter 6 Results

# Chapter 7 Comparison of the models used

# Chapter 8 References

<https://www.pothole.info/the-facts/>

J. Lin, Y. Liu, "Potholes detection based on SVM in the pavement distress image", *Appl. Bus. Eng. Sci*, pp. 544-547, Aug. 2010.

YoungJin Cha, Wooram Choi, Oral Bykztrk, "Deep LearningBased Crack Damage Detection Using Convolutional Neural Networks", 2017.

Hiroya Maeda, Yoshihide Sekimoto, Toshikazu Seto, Takehiro Kashiyama, Hiroshi Omata, Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone, 4-6-1 Komaba, Tokyo, Japan:University of Tokyo.

Justin Bray, Brijesh Verma, Xue Li, Wade He, "A Neural Network based Technique for Automatic Classification of Road Cracks", *2006 International Joint Conference on Neural Networks Sheraton Vancouver Wall Centre Hotel*, July 16-21, 2006.

Allen Zhang, Kelvin C. P. Wang, Baoxian Li, Enhui Yang, Xianxing Dai, Yi Peng, Yue Fei, Yang Liu, Joshua Q. Li, Cheng Chen, "Automated Pixel - Level Pavement Crack Detection on 3D Asphalt Surfaces Using a Deep-Learning Network", *Computer-Aided Civil and Infrastructure Engineering*, vol. 00, pp. 1-15, 2017.

Lei Zhang, Fan Yang, Yimin Daniel Zhang, Ying Julie Zhu, "Road Crack Detection Using Deep Convolutional Neural Network".

A. Tedeschi, F. Benedetto, "A real-time automatic pavement crack and pothole recognition system for mobile Android-based devices", *Advanced Engineering Informatics*, vol. 32, pp. 11-25, 2017.

<https://medium.com/diaryofawannapreneur/yolo-you-only-look-once-for-object-detection-explained-6f80ea7aaa1e>

<http://machinethink.net/blog/object-detection-with-yolo/>

[Wei Liu](https://arxiv.org/search/cs?searchtype=author&query=Liu%2C+W), [Dragomir Anguelov](https://arxiv.org/search/cs?searchtype=author&query=Anguelov%2C+D), [Dumitru Erhan](https://arxiv.org/search/cs?searchtype=author&query=Erhan%2C+D), [Christian Szegedy](https://arxiv.org/search/cs?searchtype=author&query=Szegedy%2C+C), [Scott Reed](https://arxiv.org/search/cs?searchtype=author&query=Reed%2C+S), [Cheng-Yang Fu](https://arxiv.org/search/cs?searchtype=author&query=Fu%2C+C), [Alexander C. Berg](https://arxiv.org/search/cs?searchtype=author&query=Berg%2C+A+C)*, “SSD: Single Shot MultiBox Detector”(Submitted on 8 Dec 2015 (*[*v1*](https://arxiv.org/abs/1512.02325v1)*), last revised 29 Dec 2016 (this version, v5))*

<https://drive.google.com/drive/folders/1vUmCvdW3-2lMrhsMbXdMWeLcEz__Ocuy>