INTRODUCTION

Road potholes have been a major cause of the accidents all over the world. Thus detecting them becomes extremely important. The goal is to not only detect the potholes but at the same time access and quantify the severity of them. In this project, we propose and end to end solution which utilizes the 2D and 3D data to detect and access the potholes. We also propose the employment of Internet-Of-Things (IOT) solution which aids in 3D computation and enahnce the accuracy of detection. It also helps in live tracking of pothole locations along with their severity.

METHODOLOGY

1. First we trained a simple YoloV8 model to get the bounding boxes on the road to detect potholes on the road as the starting point. To avoid false pothole detections, we added a pre trained YoloV8 model in pipeline to detect and remove all the unwanted objects (like vehicles) from the video. We also employed pothole segmentation on top of that to calculate the dimensions of potholes. With this approach in pipeline we first tracked the road and only considered those bounding box which were appearing on the road as pothole.

2. To estimate the area of the potholes we make use of segmented pothole images. Using computer vision contour detection technique we find out the area (A) of the pothole. Since the area will vary as the distance of the pothole from the camera changes so we found out the camera distance from the pothole let's call it d. The final 2D severity score for the pothole in the HD images will be the multiplication of the area and the distance from the camera (A*d) which is scaled between 0 and 1 by dividing with largest value of A*d detected. This quantity will be relatively constant as long as the pothole is in the frame signifying 2D severity score. The depth estimation method is elaborated in detail in the next point.

3. We utilise stereo cam images for the 3D construction of the roads along with potholes. This 3D reconstruction takes place on the central server only when 2D severity score obtained from the onboard model of survey vehicle denotes severity of the pothole to be either in high or medium category. The first step is to load the left and right images and acquire the disparity map from the stereo images. We then use it to obtain the 3D point cloud. For this, transform the disparity map so that we can obtain depth information. To do this we use disparity-to-depth matrix calculated using calibration information of the stereo camera. Once we get the point clouds, we fit a linear regression model to determine the plane equation of the road in the 3D point cloud. It is to be noted that the equation of the plane depends on the camera angle in a particular vehicle. We determine the depth by calculating the maximum distance of a point from the calculated plane. We combine area information from the 2D pipeline with depth information to calculate volume which we scale between 0 and 1 by dividing with the volume of the largest pothole detected in the dataset.

4. For detecting the urgency in severity of potholes we are proposing a distributed parallel computing IOT based approach, in which all the Survey and Police Patrol vehicles will be sending the relevant pothole information (described in detail in ppt) along with co-ordinates of the location. These set of informations is processed on the central server. In case of survey vehicles, we utilize the pathole information to classify it in a severity category. If the

severity is High or Medium, we then go for 3D reconstruction. In the case of patrol vehicles, we maintain a counter for each category corresponding to a particular location. These counters represent the number of patrol vehcile reporting the location to be in a particular category. Once the counter corresponding to a category for that location crosses a threshold, we allocate that severity category to that location. In this way we build severity map on the top of the city map and when the number of severe potholes in an active traffic area increase we will mark it Urgent.

5. We also propose the “training on the go” methodology to continuously train the pothole detection YOLO V8 model weights located on the central server using Semi Supervised Learning (SSL). We use images coming from vehicles for the training in which the confidence of bounding boxes (predicted onboard on all vehicles) are more than 0.8. The trained weights of central server model is distributed to all the survey and patrol vehicles weekly. It is to be noted that the YOLO model located on the central server is not used for the 2D inference on the server itself but rather used to update the survey/patrol car model weights. 2D detection is always done on the onboard computers of vehicles.

SUMMARY AND SCOPE:

1. There are several factors that shows potential in our approach of detecting potholes:
    - We use a state of art object detection model for detecting potholes.
    - We use 2 levels of preprocessing to remove the unnecessary objects which can produce defects in our model.
    - To reduce the inference time we are also using TensorRT which reduces the time taken to detect the pothole by 10x.

2.
    - For this we will use NVIDIA state of the art object tracking model Object Flow SDK

3.
    - Utilizing IOT based approach for live tracking and mapping of severity on the map.
    - We have also used contour detection to get the accurate area from the pothole shape as well as we have considered the camera distance from the pothole as a parameter in getting accurate severity score.

4.
    - We use established methods to create 3D reconstruction of the roads from stereo camera images. Thus, we get accurate depth information provided that the image quality is up to mark.
    - However, the proposed solution’s performance may get hampered during the night as the visibility will not be enough in stereo cam images for 3D reconstruction. In such cases, mobile lidar may perform better.

5.
    - The whole project is structured around autonomy, it needs minimal intervention the whole infrastructure can be hosted on AWS as it is built around it.