

# Detection of Deadly Potholes using YOLO Model

Vineet Kaushik<sup>1</sup>, Gurmeet Singh<sup>2</sup> and Pulkit Jain<sup>3</sup>

<sup>1-3</sup>Chandigarh University, Mohali, India

Email: vineetkaushik184@gmail.com, gurmeet.mech@cumail.in, pulkit.mech@cumail.in

**Abstract**—Road management requires huge manpower and still, there is no guarantee that how many days the road will last without any cracks. Even a small crack on road can lead to a big accident. But along with these cracks, the major problem comes with the pavements which are commonly known as potholes. Due to these deadly potholes lots of people have to lose their lives without their mistakes. We also know that India has one of the largest road transportation in the world. Even in terms of population India has the second largest population in the world. Road accidents is a major topic of concern not only for the government but also for the society. Potholes have a major contribution in the road accident. The government has taken initiative by forming several agencies and through different programs which majorly focus on developing a solution for detecting these potholes which can easily detect these potholes and send the data to road management authority. Although several algorithms are used and even several techniques are used, some of them are object detection technique. But if we are using this method in real time than it's not easy to get clear data using low spec hardware. Even the YOLO algorithms require very high spec hardware to perform proper function. But we came up with a solution by using YOLOv4 tiny algorithm which was made by combination of SPP and FPN with CSPdarknet53 tiny [1]. Thousands of raw data were obtained in the form of datasets with the help of data augmentation, namely gamma regulation, scaling, and horizontal flip to make up for the data lacks, which were further divided for training, testing, and validation. We have compared several YOLO algorithms with includes YOLOv3, YOLOv4, YOLOv3 tiny, YOLOv4 tiny, and YOLOv4 tiny –SPFPN [2]. In conclusion we get YOLOv4 tiny-SPFPN shows approx. 3-5% better performance than others in the mean average precision and also we developed real time algorithm which detect potholes in real time.

**Index Terms**— Feature pyramid network, neural network, object detection, spatial pyramid pooling.

## I. INTRODUCTION

Potholes are round shape or we can say oval shaped structure which have a minimum depth and that is approx. 150mm. Potholes are generally caused in rainy season because water gets accumulated on the roads and due to that cracks were formed due to water and sunlight. Sometimes some constructors also use low grade materials for their greed but they don't know that their greed can cause someone's life in danger. These potholes cause accidents majorly in two seasons namely winter season and rainy season because in rainy season the water gets accumulated on the roads due to which the drivers doesn't get any idea about the pothole and such case leads to casualties. Similar is the case in winter season when there is too much fog and the visibility is low. Thousands of people lose their lives due to these potholes not only in India but in the world. More than 15000 accidents

happened due to Potholes in India in last three to four years. One can imagine, how big the number will be if we consider the whole world. If we compare state wise results than Uttar Pradesh tops the list with more than 1000 casualties [3]. Also nowadays the number of potholes are rapidly increasing due to the irregularities caused in weather conditions due to global warming. The government of India is contributing a lot towards road transportation and they ensure best road conditions but roads need timely maintenance as they develop cracks and potholes. These Potholes need to be detected immediately without manpower because it's not possible to cover every road with the help of manpower. It will be very time consuming and even costly. So several methods were developed to detect these potholes in real time and send data to road management server. Some of them are 3D reconstruction, 2D image based pothole detection system and vibration based detection system.

The technology used in 3D reconstruction method are stereo vision based cameras and laser scanners. They can analyse the depth and width of the pavement with the help of these cameras and scanners by converting the road surface into 3D image [4]. But the drawback with this method is that these scanners and cameras are very expensive and also they are very sensitive even to a single vibration caused due to road surface to the vehicles and due to this the accuracy produced is very low.

The vibration based system used the gyroscope or accelerometer for the detection of potholes. This method can find the depth and position of the pavement by the impact which was created on the vehicle due to pothole and also with the help of the motion of the vehicle and this can be done using accelerometer [3]. But we cannot use this method in real time detection because this method does not predict the presence of pothole in advance. It can predict only when the vehicle pass through it. Also this method does not detect the potholes which are in other lane or which are under the vehicle. Further we can't distinguish between the pothole, manhole and the speed breakers because it can detect all of them.

The 2 dimensional image based pothole detecting method estimate the texture and properties of potholes with the help of image of the surface of road captured through camera. We can use this method for real time detection as it can detect the pothole well in advance which can prevent casualties in advance [5]. This method uses cameras which are very cheaper as compared to the laser scanners. The image captured here under goes several image processing steps to detect whether the given image contains a pothole or not. Further several 2D image based detection techniques were developed which works with high accuracy and even the cost is very less [6].

We all know that deep learning techniques are rapidly growing, several neural networks are developed that can detect the object by automatic learning. The algorithm used for object detection generally categorises the objects by creating square shaped boxes around them which are used for indicating their locations. The algorithms which is famous nowadays and which is said to be the fastest algorithm for detecting objects is YOLO (You Only Look at Once) algorithm. In YOLO also several version were developed. Even though YOLO has the fastest computational speed but it can't be used in real time detection with embedded hardware which has low spec. as it require high performance GPU. To counter this drawback several YOLO tiny algorithm series are developed which can be used with low spec hardware for real time detection. Although, YOLO tiny has lesser exactness than the basic YOLO algorithms this is due to the size of network is reduced to improve the computational speed.

In order to solve the above problem, we came up with a solution and that is a tiny algorithm namely YOLOv4 tiny-SPFPN algorithm. We concluded that it has greater exactness than the basic YOLOv4 algorithm and also it has better result while using in real time as compared to already used YOLOv4 tiny algorithm. The SPP and FPN are further added to the CSPdarknet53-tiny, and it is the spine of YOLOv4 tiny algorithm. While comparing the working of detection, YOLOv4tiny- SPFPN has best functionality in the MaP, IOU, FPS (Frames per second) and F1-score.

## II. LITERATURE SURVEY

Before the evolution of machine learning algorithm, the object detection is done using several techniques, some of them are vibration based method, 3D-reconstruction and 2D image extraction. These all we have already discussed. But with the advancement of the machine learning, firstly 2D pothole detection methods were used which uses machine learning based classification algorithms. But these machine learning classification algorithms only differentiate the road surface whether it is affected or not affected and this is done by feature classifiers which were explained by the user. The algorithms which uses appearance based feature classifiers detects the darkened part as the pothole or the object what is defined by the user. The researcher named Powell used pothole detection algorithm which is based on machine learning by combining it with shadow removal algorithms to detect the potholes in view of their shapes to be elliptical [2]. This technique has a limitation that it can only detect the pothole with the help of images but it is not capable of providing the location of the pothole.

After the expansion of the deep learning techniques, the restrictions or we can say that the constraint of machine learning based detection technique conquered with the help of object detecting methods in CV. The object detecting methods which uses computer vision are further divided into two stage detectors and one stage detectors. In two-stage detector we can take out several candidates in the area where there is object and further back to back processes the categorization and localization processes, due to this the exactness achieved is high but the computational speed decreases. But in one stage detectors the classification and localization processes are done at the same time, and here the exactness achieved is low but the computational speed is high. Among the one-stage object detector algorithms YOLO algorithms have the highest computational speed. Further several researchers are developing 2D pothole detection algorithms based on YOLO architecture.

The researcher named Ukhwah performed the experiment by fixing the camera on the rear side of the vehicle for detecting the road surface through the captured images of the surface of the road [7]. He trained YOLOv3, YOLOv3 SPP and YOLOv3 tiny with the help of 448 images. In return the result he got after training various YOLO models is that he is able to achieve the map of 83.43%, 88.93% and 79.33% respectively. But with this consequence, the researchers declared that the YOLOv3 SPP has got the highest accuracy by combining with SPP (Spatial pyramid pooling). Another researcher Bucko collected 1052 images under distinct climatic conditions likely clear, rainy, sunset, evening and night and further trained them [8]. In clear weather conditions, he achieved the map of 77%, 79.1% and 72.6% for YOLOv3, YOLOv3 SPP and Sparse R- CNN respectively. Park et al. collected a dataset of 665 images and then split it into validation, training and test datasets and selected 20%, 70%, and 10% of total samples, respectively [9]. He further trained a dataset of pothole images which has 598 images to YOLOv4, YOLOv4 tiny and YOLOv5s models. Dharneshkar et al. had taken 1500 road images with potholes and he collected the images from Coimbatore, Idukki, and kymily and trained them with YOLOv2, YOLOv3, and YOLOv3 tiny models, and through this he achieved the map of 45.33%, 38.41%, and 49.71% respectively [10]. Another researcher named Asad performed the real time experiment and he took a dataset of 665 images and trained with YOLOv4, YOLOv4 tiny and YOLOv5 models [11]. He used a low spec embedded hardware namely Raspberry Pi as the main board and Oak-D camera for detection of potholes. He got that only YOLOv4 tiny is able to detect the pothole in real time with 31.76 FPS among the other YOLO models used.

We know that the computational speed of the YOLOv4 tiny architecture is high, but this algorithm needs to take out substantial object characteristic that need to be located. In this paper, we mainly focussed on improving the accuracy YOLOv4 tiny model by combining some multi-scale characteristic networks, those are SPP and FPN. These networks are combined with YOLOv4 tiny model in order to extract substantial object features. These feature networks helps to compensate the spatial information loss during the max pooling. The spatial pyramid pooling is used for extracting features of pothole of different sizes with the help of max pool filters and this is used for compensating the structural details loss of the image which occurs during max pooling and this pooling is the part of CNN process. Similarly, the feature pyramid network is used for compensating the spatial information loss with the help of combination of high-level characteristic and low-level characteristic that are generated in deep convolutional layer and shallow convolutional layers respectively. The overall production of the deep learning model or we can say that YOLO models is improved with the help of the combination of these feature networks as they make use of the structural details of the image and further extract the substantial characteristic of the image.

### III. METHODOLOGY

In this paper, we have developed YOLOv4 tiny – SPFPN algorithm which was developed by attaching multi-scale characteristic networks, which includes SPP (spatial pyramid pooling) and FPN (feature pyramid network), with YOLOv4 tiny model. This whole system make this model suitable for real time detection and can be used with embedded systems.

The very first step is to collect the dataset which have images and annotations. Then this is used for training different deep learning (DL) models. Then several data augmentation method such as gamma regulation, horizontal flip, and scaling were tried with the data randomly, in order to get two thousand six hundred and sixty five datasets which would be required for compensating the insufficient datasets. Eventually, from 2665 datasets, 2400 which are training and validation datasets are used for training the models. The model performance evaluation was done with the help of the 265 datasets which were remaining and this will be done when training will be completed [12]. In order to support the presented algorithm given in this study, we have compared the performance with YOLOv2, YOLOv3, YOLOv4 tiny models. We have measured the Map at (Iou=0.5), recall,

exactness, and F1-score in order to verify the improvement in the accuracy. Further, to confirm the performance of real time detection, we have compared the FPS values without a GPU and also on low spec hardware.

#### A. YOLOv4 tiny – SPFPN Development

This section describes the significance and performance of algorithm used for detection of potholes, which is based on the multi-scale characteristic network which is used for the development of proposed YOLOv4 tiny SPFPN model. We have also shown the coupling of SPP and FPN with the YOLOv4 tiny model which is already existing.

#### B. Feature Networks

The overall performance of the pothole detection models mainly depends on how much key information of the objects is present in the convolution neural network. The main function of these feature networks is that they permit the individualistic detection of characteristic maps of various sizes that fuses between high-level and low-level characteristic map which creates a new network map which has better information on objects. Nowadays the pothole detection method which uses multi scale feature detection is the area of research as it provides enhanced performance. Further this method is divided into 2 categories, first is that which uses several layers of network for independently detecting of feature maps [13].

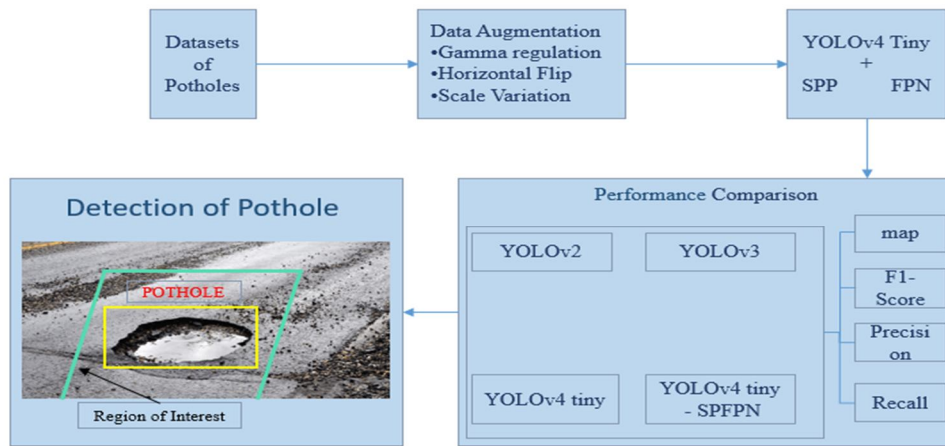


Figure 1. Performance Comparison and Pothole Detection System

Second is the combination of feature maps that were taken out from distinct layers of the network. During the first method as we have already discussed that data is applied to single shot detection network in order to detect the object in independent feature map extracted from distinct layers of convolution neural network and as a result better accuracy is achieved if small objects were detected from characteristic maps which are attached with high-level and low-level rather than the small objects detected from characteristic maps gathered from low-level only. Several researchers have shown interest in the second case. Li et al. has given a performance report of pedestrian detection in which he used the scale aware mechanism which was acquired from combination and addition of output characteristic maps with small and big networks for the combination of characteristic map [14]. Another researcher showed that the performance of detection can be enhanced nevertheless of the size of the input image as they form feature map using max pool filters and spatial pyramid pooling. Lin et al. integrated information with high-level and low-level characteristic maps with combination of local area feature of distinct size for developing FPN and a top-down side linked structure [15].

The literature survey showed that for detecting small objects, multi scale feature are very successful and also it achieve high accuracy when combination technology is used in which high level and low level feature maps are connected. Keeping all this in mind we have made the algorithm which uses the YOLOv4 tiny architecture as the base but some additional feature are merged with it which are spatial pyramid pooling and feature pyramid network so that it can easily detect the small potholes which are mostly neglected by the drivers but they cause huge damage.

#### C. YOLOv4 tiny (SPFPN)

We found that many government organizations and researchers are working with YOLO models as they have fast computational speed and high accuracy in the group of one stage detector type pothole detection system.

YOLO is generally used for predicting and detecting with the help of prediction bounding box and hence have the highest confidence score and further divide the input image in  $M \times M$  grid cells. The clarified version of YOLOv4 model is YOLOv4 tiny and this YOLOv4 tiny model had shown the high rise performance of 43.5% average precision. The detection speed or we can say that the computational speed of YOLOv4 tiny is faster than YOLOv4 but it has lower accuracy because YOLOv4 tiny has smaller convolution neural network structure than YOLOv4. We know that YOLOv4 tiny has lower accuracy but it shows good result when used for real time detection with low spec embedded hardware's like Raspberry Pi. Due to this advantage, we have used YOLOv4 tiny in our study as it is the most suitable deep learning method for detection of potholes.

How YOLOv4 tiny detects the object? The answer to this question is given here. Basically, YOLOv4 tiny detects the pothole based on CSPdarknet53-tiny, and it further uses CSP Net and additional feature pyramid network and it is a type of multi scale network. The complete architecture of SPFPN YOLOv4 tiny is shown in fig. 2. The area marked with red dotted line shows the CSPdarknet53-tiny which consist of three convolution layers and three CSP blocks, this is also considered as the backbone of YOLOv4 tiny.

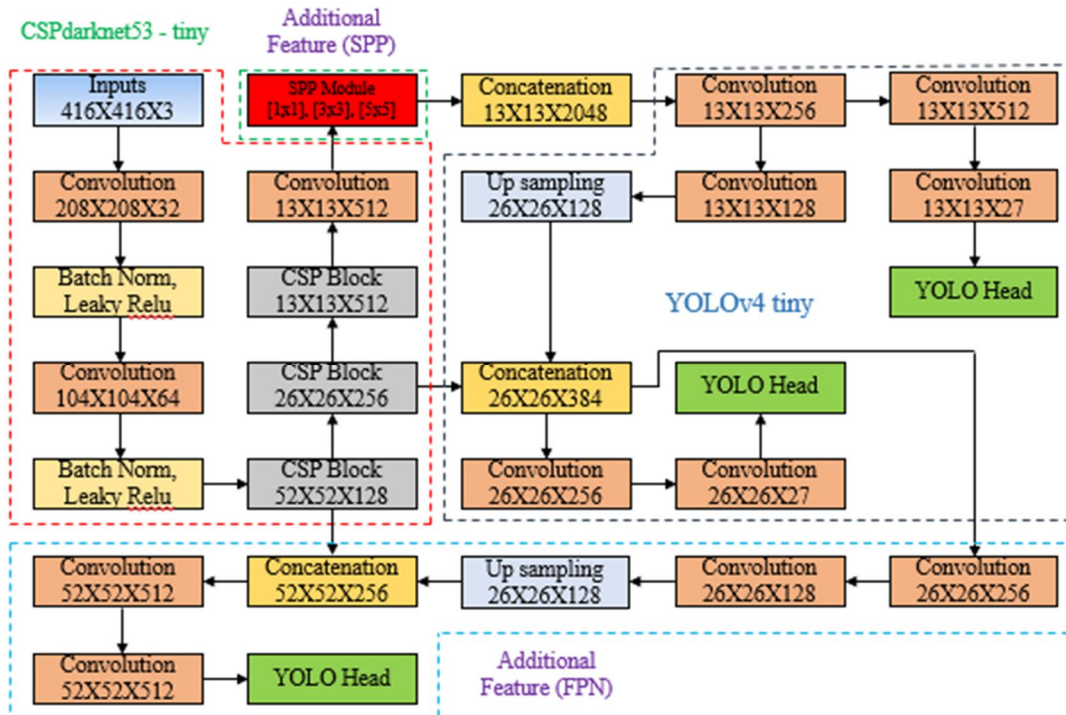


Figure 2. Architecture of SPFPN YOLOv4 tiny

The additional feature FPN shown with light blue dotted line consists of four convolution layer and an up sampling layer and they were used to connect 13x13 feature map and 26x26 feature maps. The dark blue dotted line is the area which covers YOLOv4 tiny and object that is detected will be in feature maps of sizes 13x13 and 26x26.

The 13x13 feature maps contains the all in depth information of the object as they are high level feature maps. As compared to 13x13, 26x26 contains less intensive information as they are middle level feature maps [16]. The shape of the small size potholes gets twisted as the size of the feature maps gets decreased due to this difficulties arises in detection of small size potholes. Due to this additional feature map known as FPN is added as it can utilize the low level feature maps of size 52x52 to enhance the detection accuracy of small potholes. In addition to this SPP (spatial pyramid pooling) is connected after the third CSP block and it has high expectation of creation of small twisted objects [17]. The main function of this SPP is to apply max pool filters so that the image can be divided in many images and after that merging them which results in less loss of information as compared to the conventional max pooling process. The total of 13x13x2048 feature maps gets generated after passing through this whole network as additional [1x1], [3x3], and [5x5] max pool filter were also applied in the added SPP network. The feature maps are first attached to high level characteristics map of 13x13 before passing the third CSP block which is connected with SPP. The high level characteristic map is back to back attached with low level and middle level characteristic map with the help of two FPN. In order to train the low level

features such as shape, edges, size, and texture, the low level feature map of 52x52 which is in charge for small object characteristics is further attached with 13x13 and 26x26 which finally shows key information features of potholes.

#### D. Performance Evaluation Metrics

In order to check the performance evaluation of SPFPN YOLOv4 tiny we need to calculate the precision or map, recall, F1 score and AP and these are calculated using confusing matrix index, and it is the most widely used performance evaluation technique in computer vision.

Precision is calculated with the help of True Positive (TP) divide by the sum of True Positive and False Positive (FP).

$$\text{Precision} = \frac{TP}{TP+FP} \quad (a)$$

Similarly, recall is (TP) by the sum of TP and False Negative (FN).

$$\text{Recall} = \frac{TP}{TP+FN} \quad (b)$$

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N AP_i \quad (c)$$

Mean average precision (mAP) is denoted by above-mentioned equation. It is the mean value of each class of all the indices which were used for evaluating object detection performance.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (d)$$

F1 score can be defined as the harmonic mean of recall and precision.

True positive (TP) is basically the right recognition of the object which is present in the image, similarly False Positive (FP) is wrong recognition i.e. the function marks the objects which is not present in the image. False negative (FN) is an object which is not detected by the network but is present in the image and True Negative (TN) is the recognition of the object which is not present in the image.

If we see mathematically, then both exactness and recollection parameters of two tiny parts that have the same numerators but dissimilar denominators and they both are non-negative numerals and less than or equal to one. Higher the exactness value means higher perfection of founding objects. High recall means the rate of (TP) is high and the chance of lost positive objects is low [17].

The evaluation of the model is done by changing the threshold value and observing the precision and recall values. The calculation of exactness and recall by assuming an N threshold and each of these thresholds is a pair of precision ( $P_n$ ) and recall ( $R_n$ ) ( $n=1, 2, 3, \dots, N$ ) [18].

TABLE I. UN-NORMALIZED CONFUSION MATRIX

| Actual/Prediction | Predicted as Positive | Predicted as Negative |
|-------------------|-----------------------|-----------------------|
| Positive          | True Positive (TP)    | False Negative (FN)   |
| Negative          | False Positive (FP)   | True Negative (TN)    |

#### IV. RESULTS

Darknet repository, which is a framework for neural network and was developed by Joseph Redmond, was used for performing the development and training of deep learning models. The training environment used was: firstly the computer environment was google colab, along with jetson nano as it provide high performance specifications. Table 2. shows the hardware specification used in our study.

Different teaching criteria's were set such as the teaching rate of 0.9 momentum, batch size was of 32, subdivision of 16 and decay of 0.0005. The size of the loaded image was fixed at 416x416, and in between the training phase the mix up augmentation was fixed to true, due to that size of loaded image can be changed arbitrarily in between the phase of training [19]. The trusted value and the intersection over union threshold values were set as 0.25 and 0.5 respectively for the proposed model and also for other YOLO models [20]. While

comparing the training map of each models we got that the YOLOv4 tiny SPFPN has achieved highest map value but its training time is little higher than YOLOv4 but if we compare it with YOLOv2 and YOLOv3 it is much lesser. Table 3. shows the comparison result of different YOLO architectures at MaP(IoU=0.5).

TABLE II. HARDWARE SPECIFICATION

| Environment | Version                         |
|-------------|---------------------------------|
| CPU         | Intel(R) Core(TM) CPU @ 2.30GHz |
| IDE         | Google Colab Pro                |
| GPU         | Jetson Nano (4GB)               |
| CUDA        | 10.2                            |
| Python      | 3.8                             |

TABLE III. COMPARISON RESULT OF DIFFERENT YOLO MODELS

| Model                       | maP (Iou=0.5) | Training Hours |
|-----------------------------|---------------|----------------|
| YOLOv2                      | 74.80%        | 8.8            |
| YOLOv3                      | 77.80%        | 9.4            |
| YOLOv4 tiny                 | 72.70%        | 4.2            |
| Presented SPFPN YOLOv4 tiny | 79.60%        | 5.8            |

TABLE IV. COMPARISON RESULT OF PERFORMANCE OF DATASET OF POTHOLES

| Model                       | Precision | Recall | F1-Score | FPS |
|-----------------------------|-----------|--------|----------|-----|
| YOLOv2                      | 83%       | 74%    | 78.20%   | 26  |
| YOLOv3                      | 88%       | 72%    | 79.20%   | 28  |
| YOLOv4 tiny                 | 74%       | 76%    | 72.50%   | 56  |
| Presented SPFPN YOLOv4 tiny | 89%       | 84%    | 86.40%   | 38  |

Table 4. shown above shows the performance result of the dataset. The precision, recall, and F1 score value are highest for YOLOv4 tiny SPFPN model as compared with other models. The FPS achieved by YOLOv4 tiny SPFPN model is slightly higher than the standard used for real time detection. The standard is of 30 FPS and the achieved value is 38 FPS.

Below mentioned figures shows the training result achieved by various models. The basic information available in the figure is the map value, losses, training hours, and iterations. The line shown with blue colour in the fig. 3 shows the loss and the line shown with red colour shows the achieved Map (IoU = 0.5) per hundred times from thousands iterations. Fig 3. shows the training graph of YOLOv4 tiny. The convergence loss in YOLOv4 tiny is stable and goes to zeroes at 2000 iterations and map value here is 72.7%. Fig 4. shows the training graph of proposed model. Although the convergence loss is not so stable in this model but still the map value achieved is the highest among all YOLO models discussed. The iteration for each model was set to 2000.

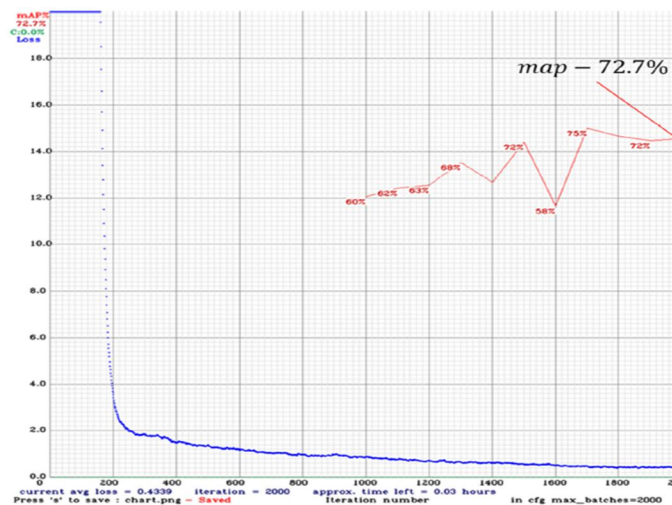


Figure 3. Training graph of YOLOv4 tiny

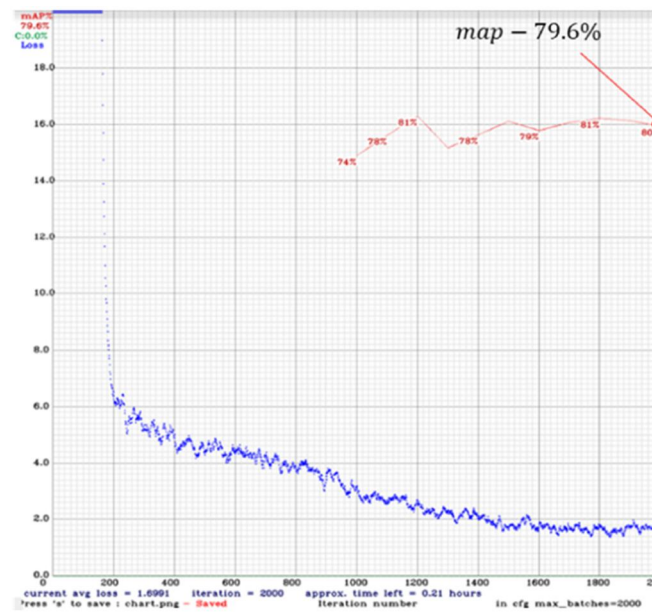


Figure 4. Training graph of SPFPN YOLOv4 tiny

## V. CONCLUSION

The detection algorithm we have proposed in this study was for low spec embedded system. In order to get rich information from images we have developed SPFPN YOLOv4 tiny model by combination of CSPdarknet53 tiny with spatial and feature networks.

Several data augmentation methods were used which includes horizontal flip, gamma regulations, and scaling for compensation of insufficient datasets. We have compared various YOLO models and the training result of each models were shown. The mean average precision (IOU=0.5) achieved by YOLOversion2, YOLOversion3, YOLOv4 tiny and YOLOv4 tiny Spatial feature networks are 74.8%, 77.8%, 72.7% and 79.6% respectively. This shows that the proposed model has the highest value achieved. Even the exactness, recall, F1 score of the presented model is highest among all YOLO models. The real time detection frames per second is also above the standard level.

This shows we have developed an algorithm which is suitable with low grade hardware environment and has good result in real time detection.

## VI. FUTURE SCOPE

By achieving good result in the detection model. We are looking forward for creating a novel idea in which the risk of damage or the impact of the pothole on the vehicle will be calculated well in advance before the vehicle cross the pothole. Also it will detect the impact the pothole have on the tire of the vehicle. This will help the drivers to even recognize the unrecognized pothole which will surely reduce the number of accidents.

Further real time detection should be focussed more in which we can get the clear scenario about the detection of the pothole. A survey or a trail can be arranged by the government in which few vehicle should be arranged for selected cities to capture the real time results. We can also give alarming system such as shivering system, LED alerts for the driver so that they can be alerted in advance.

## REFERENCE

- [1] D. H. Heo, J. Y. Choi, S. B. Kim, T. O. Tak, and S. P. Zhang, "Image-Based Pothole Detection Using Multi-Scale Feature Network and Risk Assessment," *Electron.*, vol. 12, no. 4, 2023, doi: 10.3390/electronics12040826.
- [2] L. Powell, "Automated Road Distress Detection," 2016.
- [3] V. Kaushik and B. S. Kalyan, "Pothole Detection System: A Review of Different Methods Used for Detection," 2022 *2nd Int. Conf. Comput. Sci. Eng. Appl. ICCSEA 2022*, 2022, doi: 10.1109/ICCSEA54677.2022.9936360.
- [4] K. T. Chang, J. R. Chang, and J. K. Liu, "Detection of pavement distresses using 3D laser scanning technology," *Proc. 2005 ASCE Int. Conf. Comput. Civ. Eng.*, no. 2005, pp. 1085–1095, 2005, doi: 10.1061/40794(179)103.

- [5] S. K. Ryu, T. Kim, and Y. R. Kim, "Image-Based Pothole Detection System for ITS Service and Road Management System," *Math. Probl. Eng.*, vol. 2015, 2015, doi: 10.1155/2015/968361.
- [6] E. Buza, S. Omanovic, and A. Huseinovic, "Pothole Detection with Image Processing and Spectral Clustering," *Recent Adv. Comput. Sci. Netw. Pothole*, pp. 2–7, 2013, [Online]. Available: <http://www.wseas.us/e-library/conferences/2013/Antalya/ITCN/ITCN-05.pdf>.
- [7] E. N. Ukhwah, E. M. Yuniarno, and Y. K. Suprpto, "Asphalt Pavement Pothole Detection using Deep learning method based on YOLO Neural Network," *Proc. - 2019 Int. Semin. Intell. Technol. Its Appl. ISITIA 2019*, pp. 35–40, 2019, doi: 10.1109/ISITIA.2019.8937176.
- [8] B. Bučko, E. Lieskovská, K. Zábovská, and M. Zábovský, "Computer Vision Based Pothole Detection under Challenging Conditions," *Sensors*, vol. 22, no. 22, 2022, doi: 10.3390/s22228878.
- [9] S. S. Park, V. T. Tran, and D. E. Lee, "Application of various yolo models for computer vision-based real-time pothole detection," *Appl. Sci.*, vol. 11, no. 23, 2021, doi: 10.3390/app112311229.
- [10] J. Dharneeshkar, S. Dhakshana, S. Aniruthan, R. Karthika, and L. Parameswaran, "Deep Learning based Detection of potholes in Indian roads using YOLO," *Proc. 5th Int. Conf. Inven. Comput. Technol. ICICT 2020*, pp. 381–385, 2020, doi: 10.1109/ICICT48043.2020.9112424.
- [11] M. H. Asad, S. Khaliq, M. H. Yousaf, M. O. Ullah, and A. Ahmad, "Pothole Detection Using Deep Learning: A Real-Time and AI-on-the-Edge Perspective," *Adv. Civ. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/9221211.
- [12] M. Jakubec, E. Lieskovská, B. Bučko, and K. Zábovská, "Comparison of CNN-Based Models for Pothole Detection in Real-World Adverse Conditions: Overview and Evaluation," *Appl. Sci.*, vol. 13, no. 9, 2023, doi: 10.3390/app13095810.
- [13] P. Ping, X. Yang, and Z. Gao, "A Deep Learning Approach for Street Pothole Detection," *Proc. - 2020 IEEE 6th Int. Conf. Big Data Comput. Serv. Appl. BigDataService 2020*, pp. 198–204, 2020, doi: 10.1109/BigDataService49289.2020.00039.
- [14] J. Li, X. Liang, S. Shen, T. Xu, J. Feng, and S. Yan, "Scale-Aware Fast R-CNN for Pedestrian Detection," *IEEE Trans. Multimed.*, vol. 20, no. 4, pp. 985–996, 2018, doi: 10.1109/TMM.2017.2759508.
- [15] Y. Zhao, R. Han, and Y. Rao, "A new feature pyramid network for object detection," *Proc. - 2019 Int. Conf. Virtual Real. Intell. Syst. ICVRIS 2019*, pp. 428–431, 2019, doi: 10.1109/ICVRIS.2019.00110.
- [16] R. A. S. B. Anas Al-Shaghouri, "Real-Time Pothole Detection Using Deep Learning," *Comput. Vis. Pattern Recognit.*, pp. 1–10, 2021.
- [17] A. Kumar, Chakrapani, D. J. Kalita, and V. P. Singh, "A modern pothole detection technique using deep learning," *2nd Int. Conf. Data, Eng. Appl. IDEA 2020*, 2020, doi: 10.1109/IDEA49133.2020.9170705.
- [18] Z. Fan, Y. Wu, J. Lu, and W. Li, "Automatic Pavement Crack Detection Based on Structured Prediction with the Convolutional Neural Network," *2019 IEEE Int. Conf. Consum. Electron.*, vol. 1, no. c, pp. 1–2, 2018, [Online]. Available: <http://arxiv.org/abs/1802.02208>.
- [19] P. A. Chitale, K. Y. Kekre, H. R. Shenai, R. Karani, and J. P. Gala, "Pothole Detection and Dimension Estimation System using Deep Learning (YOLO) and Image Processing," *Int. Conf. Image Vis. Comput. New Zeal.*, vol. 2020-Novem, 2020, doi: 10.1109/IVCNZ51579.2020.9290547.
- [20] N. Ma *et al.*, "Computer vision for road imaging and pothole detection: a state-of-the-art review of systems and algorithms," *Transp. Saf. Environ.*, vol. 4, no. 4, pp. 1–16, 2022, doi: 10.1093/tse/tdac026.