MEKATRONIKA
e-ISSN: 2637-0883
VOL. 5, ISSUE 2, 87 – 99
DOI: https://doi.org/10.15282/mekatronika.v5i2.9813

ORIGINAL ARTICLE

Investigation on a Vision-Based Approch For Smart Pothole Detection Using Deep learning Based on Fast CNN

Ashraf Khaled Alsharafi[1], Muhammed Nafis Osman Zahid[1,*]

Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pahang, Malaysia.

ABSTRACT – The quality of road these days are important and roads always dangerous since its filled with potholes and damages which cause a lot of incident and numbers gets more increased in crowded area , this article investigates and compare the performance metrics of different object detection models that utilized the Fast CNN structure in it's backbones , Four processes make up the standard method of pothole detection: data acquisition, data pre-processing, feature extraction, and pothole classification. for the task of pothole detection. The study focuses on the evaluation of YOLOv6n, YOLOv8n, YOLOv5n, and YOLOv7 models using a dataset of road images containing pothole instances. The performance metrics analyzed include precision (P), recall (R), mean average precision at 50% IoU (mAP@.5), and mean average precision from 50% to 95% IoU (mAP@.5:.95) . The findings indicate that YOLOv8n demonstrates the highest overall performance, achieving significant precision and recall rates. These results provide valuable insights into the effectiveness of object detection models for pothole detection, contributing to the field of road maintenance and safety. The outcomes of this study can assist in the development of intelligent systems for automated pothole detection and maintenance planning.

ARTICLE HISTORY
Received: 2nd Oct 2023
Revised: 3rd Nov 2023
Accepted: 10th Dec 2023
Published: 17th Dec 2023

KEYWORDS
Yolo models
Mean Average Precision
Smart Pothole Detection
Vision-Based Approch

1.0 INTRODUCTION

Potholes can result in damage like flat tyres and damaged wheels, vehicle collisions, and serious accidents; this is a frightening concern in the modern world. The effort on this subject entails finding potholes on roadways and keeping a database of the coordinates of that location. Creating a gadget that is integrated into the vehicle is frequently used to accomplish this. The gadget monitors the road ahead using its infrared and ultrasonic sensors, warning the driver as soon as a pothole is about to appear. This technique is helpful during the rainy season when the roads are clogged with water from the downpours and vision is reduced due to fogging up over the winter [1].

The quality of road these days are important and roads always dangerous since its filled with potholes and damages which cause a lot of incident and numbers gets more increased in crowded area , this article investigates the common methods scientists and engineers has been developed that helps identifies the road damage to avoid pothole through vision-based approaches and other sophisticated approaches which will be discussed briefly and focus more on vision based approaches and other methods which will be discussed in details later in this article, also the purpose of these methods are not only to assist intelligence driving to detect damages in front of the drivers, it's purpose goes beyond that and helps the operations for road maintenance by drawing all the possible pothole location in a specific area by using an approach called vibration-based methods that can generate a map to predict potholes, each method has its own various way to implement as the figure 1.1 shows [2].

*CORRESPONDING AUTHOR | Muhammed Nafis Osman Zahid | nafis@umpsa.edu.my

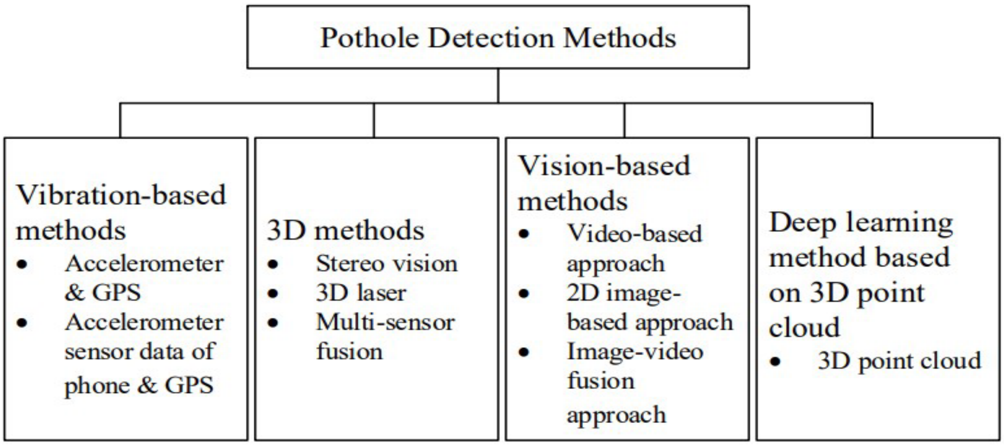

Figure 1. Classification of Pothole Detection Method

Road potholes feature a great deal of uncertainty in their perimeter and depth, making it challenging to recognise items with well-defined forms like automobiles, pedestrians, cats, and dogs. Road pothole recognition, however, still needs more in-depth study. Examples include feasible area detection, positive obstacle detection and traffic island identification [3]. J. Eriksson et al. [4] were the first to introduce a pothole detection system that relies on vibrations and utilizes GPS in conjunction with a three-axis accelerometer. Subsequently, Kang Chen et al. [5] enhanced this system by incorporating a gyroscope, enabling it to identify both manholes and potholes. Other solutions, such as Nericell [6] and Traffic Sense [7], employ mobile smartphones as sensor devices, utilizing accelerometers to detect potholes. Z. Hou et al. [8] proposed an approach based on stereo vision, which involves three-dimensional reconstruction from two images taken in different directions to identify potholes. E. Salari and E. Chou [9] improved the accuracy of their detection results by processing two-dimensional images. Additionally, K. T. Chang et al. [10] and Lei Yang [11] directly employ lidar technology to scan the road surface, collecting point cloud information and analyzing data to identify potholes.

Four processes make up the standard method of pothole detection: data acquisition, data pre-processing, feature extraction, and pothole classification. The process of creating a dataset to learn about or analyse potholes by gathering raw data is known as the data-acquisition stage. For comprehending and assessing the problem, it is essential to collect sufficient data during the data-acquisition phase. In the data-pre-processing step, the dataset produced in the previous stage is utilised to clean up the data for rapid learning or analysis of potholes.

Several signal-processing methods, including filtering and masking, are used in the data-pre-processing stage to clean up the raw data. Finding a factor to discriminate between potholes and non-potholes in the pre-processed data is the process of feature extraction. Finding a factor to discriminate between potholes and non-potholes in the pre-processed data is the process of feature extraction. When looking for parameters to classify potholes during the feature-extraction stage, it is crucial to take the kind and properties of the pre-processed data into account. By using a pothole-detection algorithm based on the characteristics, the pothole classification step establishes the presence of potholes. Various techniques are used in the pothole classification stage to utilised to increase the pothole classification's accuracy, and in the process, an ideal solution is discovered [12].

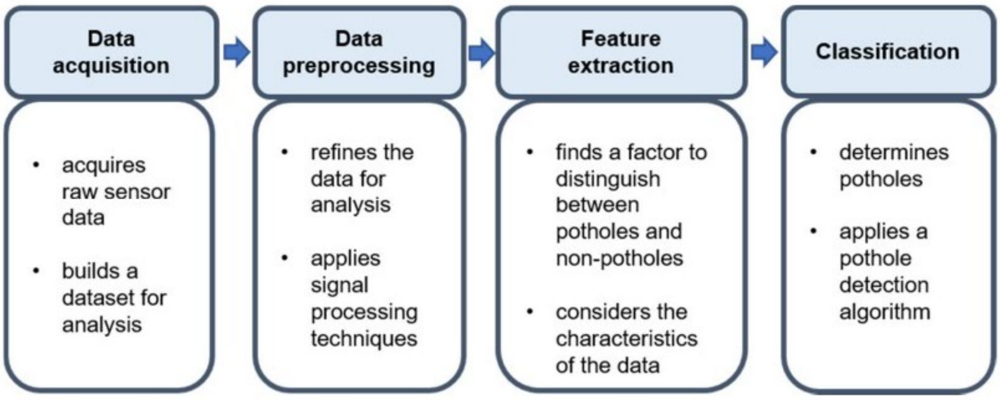

Figure 2. The General Prosecc for Pothole Detection

2.0 METHODOLOGY

The preparation of the YOLOV7 mathematical model that has been utilized to train data that later is used to detect potholes. anaconda navigator is used to set the environments for YOLOV7 and the tool used in this method along with python.py script for prepossessing, training, and detecting. On the other hand, Google-Collab is used to train the data and send the best value to google drive, other tools used to label images and prepare them such as labelling. The Datasets used to train the pothole detection system is a collection of 300+ untrained unlabelled images that are obtained using Google search and the Kaggle website.

This study used CNN (Convolutional neural network) to train the data since the architecture of yolo7v is using CNN that each layer consists of input data (our data set) filter detector consider as a feature detector that moves across the image and perform a down operation between the input and the value of the filter to produce an output which is a feature map (feature extraction). We obtained the feature map for the feature extraction using labelling. The feature map produced by the convolution is a multidimensional array it must be reshaped (flattened) into a victor unit to be able to feed the CNN input.

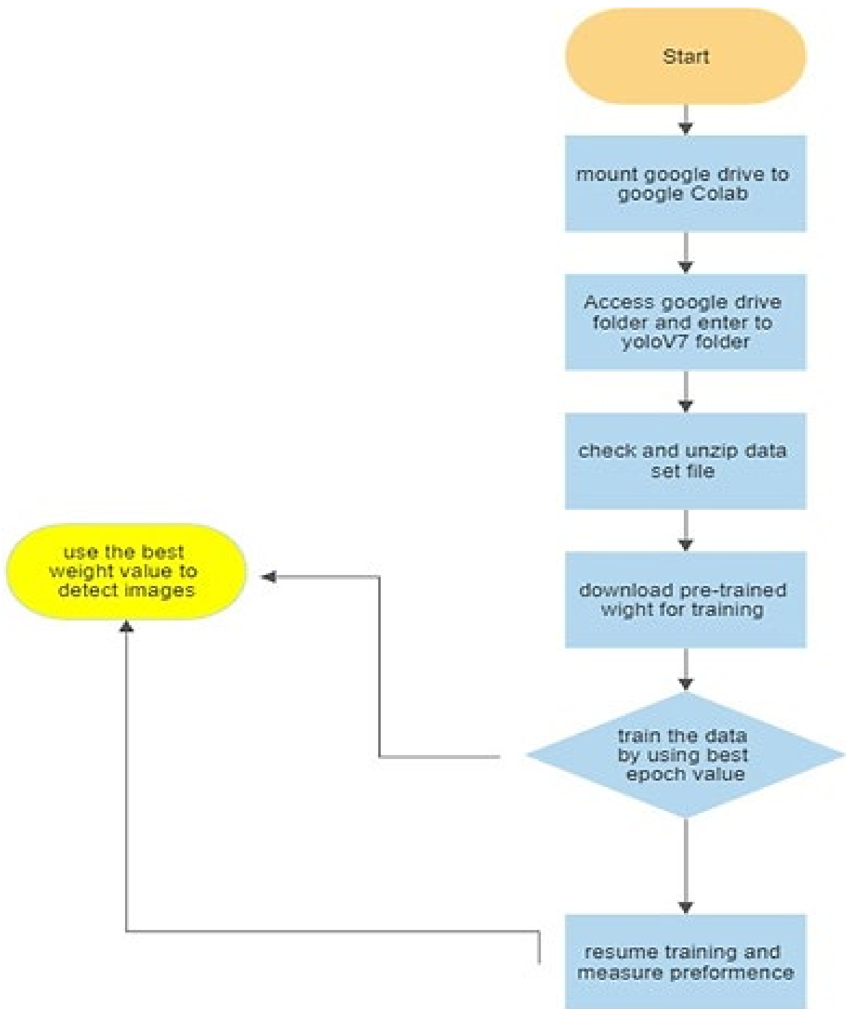

Figure 3 Flow chart for training data using google collab

2.1 Vision Detection Method

Vision detection is the method of processing images done by many mathematical models in our case we use YOLO model v7 version to detect our data set, the model is able to detect a lot of things since it includes a huge library but it can't detect potholes unless the data is trained and preprocessed, The flowchart shows the mothed setup

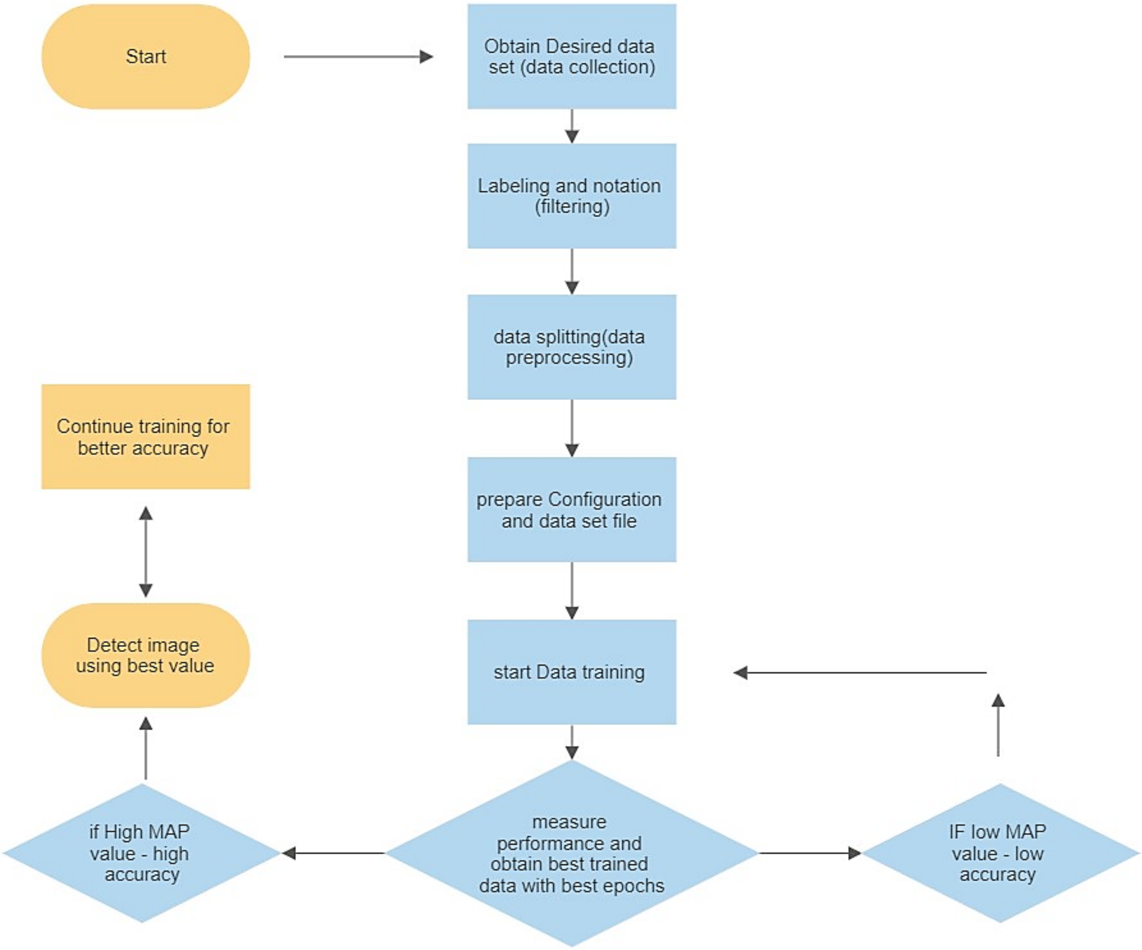

Figure 4 Flow chart for preforming vision detection

2.2 Data preprocessing

To perform data pre-processing, an annotation and labelling for the image are used with labelling environment

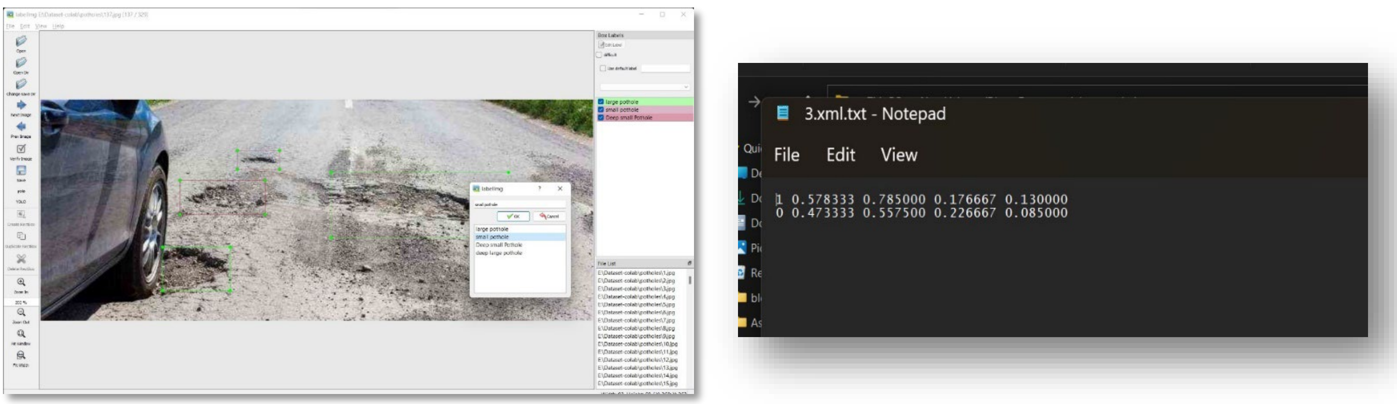

Figure 5. Labelling environment and anotation file

2.3 Data splitting

to prevent overfitting the test set which is used to evaluate the performance of the trained model and the trained set, the reason of splitting is to prevent overfitting, overfitting occurs when a model performs well on the training set and poorly on data that has never seen, data is required to test the tunning hyper parameter, because the data used is not from the training set, an additional section (validation set is required, so in total, the data set is split into 3 sections: 75% training set, 15 % test set, 15 % validation set)

```
custom_dataset
├── images
│   ├── train
│   │   ├── train0.jpg
│   │   └── train1.jpg
│   ├── val
│   │   ├── val0.jpg
│   │   └── val1.jpg
│   └── test
│       ├── test0.jpg
│       └── test1.jpg
└── labels
    ├── train
    │   ├── train0.txt
    │   └── train1.txt
    ├── val
    │   ├── val0.txt
    │   └── val1.txt
    └── test
        ├── test0.txt
        └── test1.txt
```

Figure 6. Data splitting architecture

2.4 Data Trining

The training of data for the pothole detection model is performed on Google-Colab website by using yolo7v.py, this library uses various of arguments that call scripts for training, installing pretraining weights, continuing training, measuring performance and obtaining the accuracy graph using Tensor board to obtain the best epoch value(epoch) Table 1 and Table 2 shows the arguments and parameters used for train and detect.

Table 1 Arguments and parameters used for trainings and testing preformence

Argument	Description	Example
--workers	The number of processes that generates batches in parallel	--worker 0
--batch-size	The number of images processed before updating the model	--batch-size 640
--device	CUDA device	--device 0
--data	Data file	--data data/pothole.yaml
--img	Image size	--img 512 512
--cfg	Configuration file	--cfg (directory)/yolo-pothole.yaml
--weights	Calling the weight for data training	--weights yolov7_training.pt
--epochs	The number of times the learning algorithm will work to process the entire data set	--epochs 250
--conf-thres	Object confidence threshold	--conf-thres 0.5

Table 2 Parameters and parameters used for trainings and testing preformence

Parameter	Description
True Positive [TP]	The number of correctly classified positive instances
False Positive [FP]	The number of falsely classified positive instances
False Negative [FN]	The number of falsely classified negative instances
Precision [P]	The ratio of true positive predictions to all actual positive predictions
Recall [R]	The ratio of true positive predictions to all actual positive instances
F1 Score [F1]	The harmonic mean of precision and recall, representing the balance between the two metrics
Threshold [IOU]	The confidence threshold used to determine positive predictions
Mean Average Precision (MAP)	The average precision across all classes at a given confidence threshold

3.0 RESULT AND DISCUSSION

After the models (Yolov7-t ,Yolov8n,Yolov5n,Yolov6n) were trained ,in this section the performance of Yolov7 in comparison with other models is measured and compared after the preformence matrices for each model is obtained such as (confusion matrices , F1 score curve , precision-recall curves and Mean average precision).

3.1 Confusion matrices results

Table 3 shows the classification performance of different models for each class , its clearly can be seen that Yolov8n appears to perform consistently well across different classes, demonstrating high TP rates and competitive FP rates

Table 3. Results of confusion matrices for each model

Class	Yolov7-tiny	Yolov5n	Yolov6n	Yolov8n
Large pothole	TP: 87%	TP: 84%	TP: 55%	TP: 82%
	FP: 2%	FP: 10%	FP: 31%	FP: 14%
	FN: 9%	FN: 6%	FN: 14%	FN: 4%
Small pothole	TP: 78%	TP: 82%	TP: 75%	TP: 85%
	FP: 22%	FP: 3%	FP: 5%	FP: ---
	FN: ---	FN: 16%	FN: 20%	FN: 15%
Deep small pothole	TP: 85%	TP: 67%	TP: 47%	TP: 73%
	FP: --	FP: 27%	FP: 40%	FP: 0.27%
	FN: 15%	FN: 7%	FN: 13%	FN: ---
Deep large pothole	TP:80%	TP:80%	TP:59%	TP: 80%
	FP: 10%	FP:10%	FP: 41%	FP: 20%
	FN: 10%	FN: 10%	FN: ---	FN: ---

Some classes is detected beter under yolov5n model like (Large pothole class) is shown better TP postive intances ,over all the confusion matrices showed how each model detect each class for a data set of 300+ images inder a threshold 0.5 .

3.2 F1 score curves

the F1 score curve showed which class is more favourable to the model the more area under the curve of each class the more favourable the model is to that class, in the Figure 7 and Figure 8 shows the F1 curve for each model:

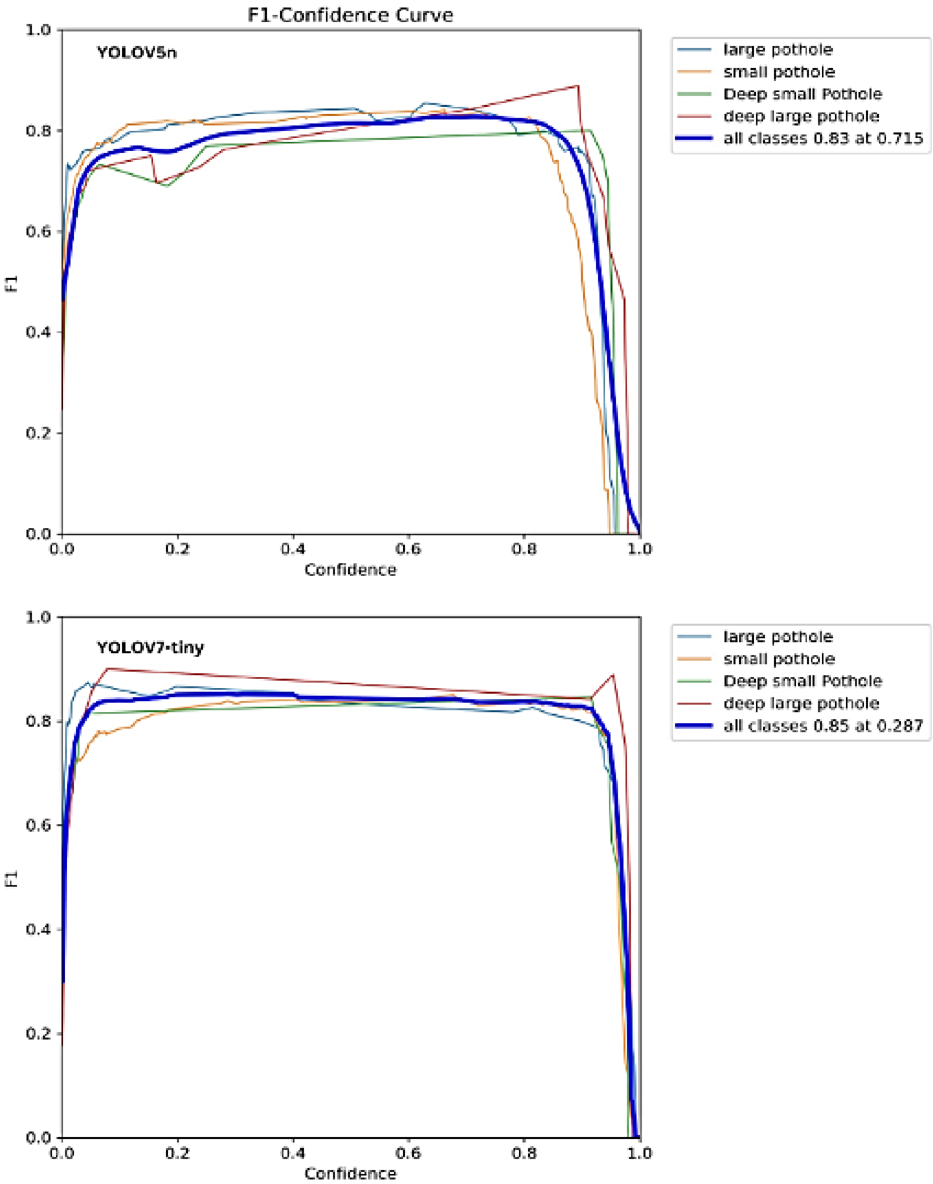

Figure 7. F1 score curve for each model for Yolo5n and YoloV7-tiny

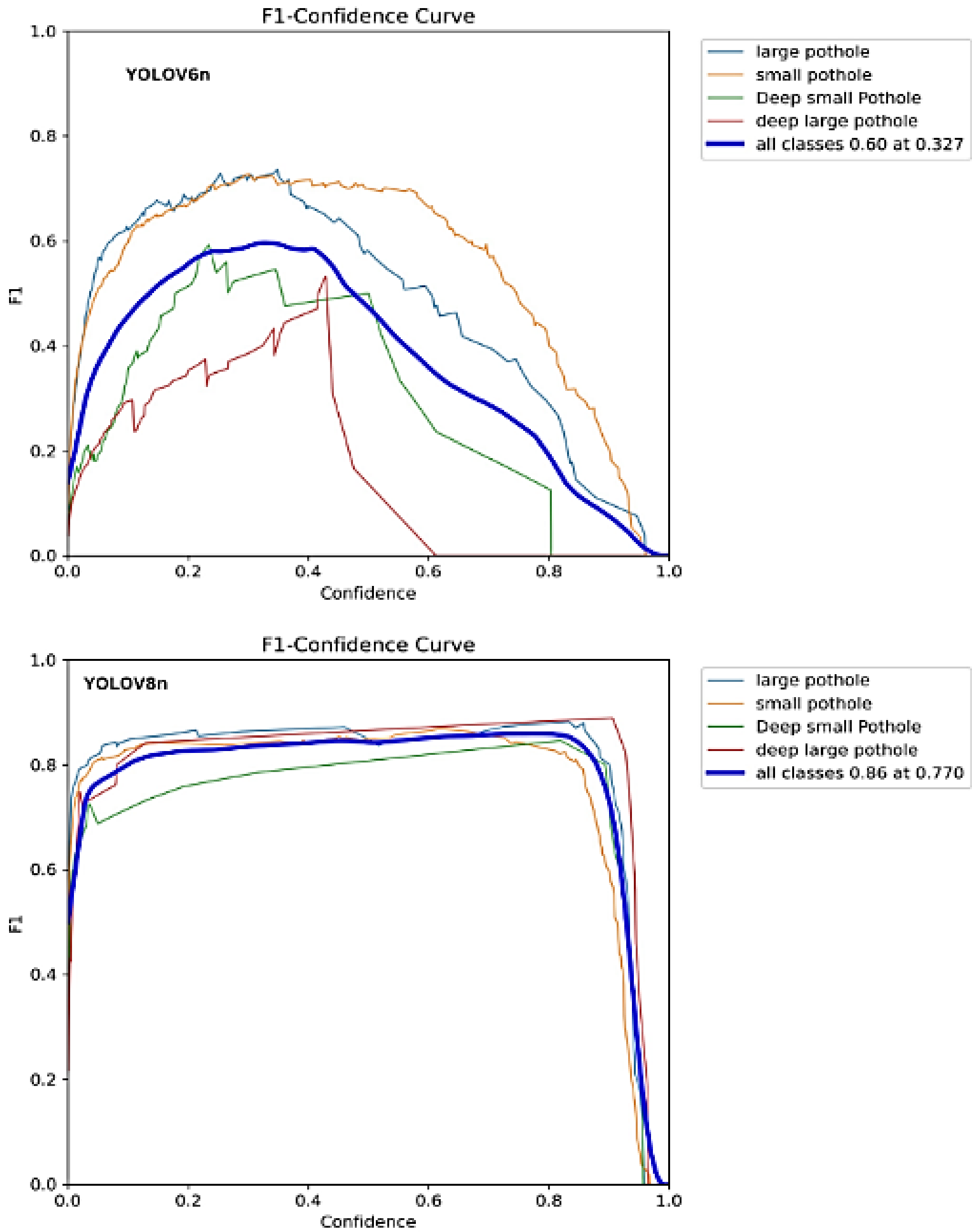

Figure 8.. F1 score curve for Yolov6n and Yolov8n

Table 4. F1 Score for all classes at different cobfidence threshold

Model	F1 score	Confidence threshold
Yolov7-tiny	0.85	0.287
Yolov5n	0.83	0.715
Yolov6n	0.60	0.327
Yolov8n	0.86	0.770

The F1 score considers both precision and recall. It is a harmonic mean of these two metrics and provides a single value that represents the overall performance of the model as in Table 4 in terms of correctly identifying positive instances while minimizing false positives and false negatives.

By analyzing the F1 confidence curve, results determined the optimal confidence threshold that balances precision and recall for your specific task. Setting a higher confidence threshold can increase precision but may result in a lower recall, while setting a lower confidence threshold can increase recall but may result in false positives.

3.3 Precision-Recall curves

The precision-recall (P-R) curves provide valuable insights into the performance of a model across different levels of precision and recall as in Figure 9 and Figure 10. Unlike the F1 score curve, which focuses on a single metric that balances precision and recall, the P-R curves allow for a more comprehensive analysis of the model's performance. Comparing P-R curves across multiple models enables to assess their relative performance. A model with a higher curve, which indicates higher precision and recall values across various thresholds, generally outperforms models with lower curves.

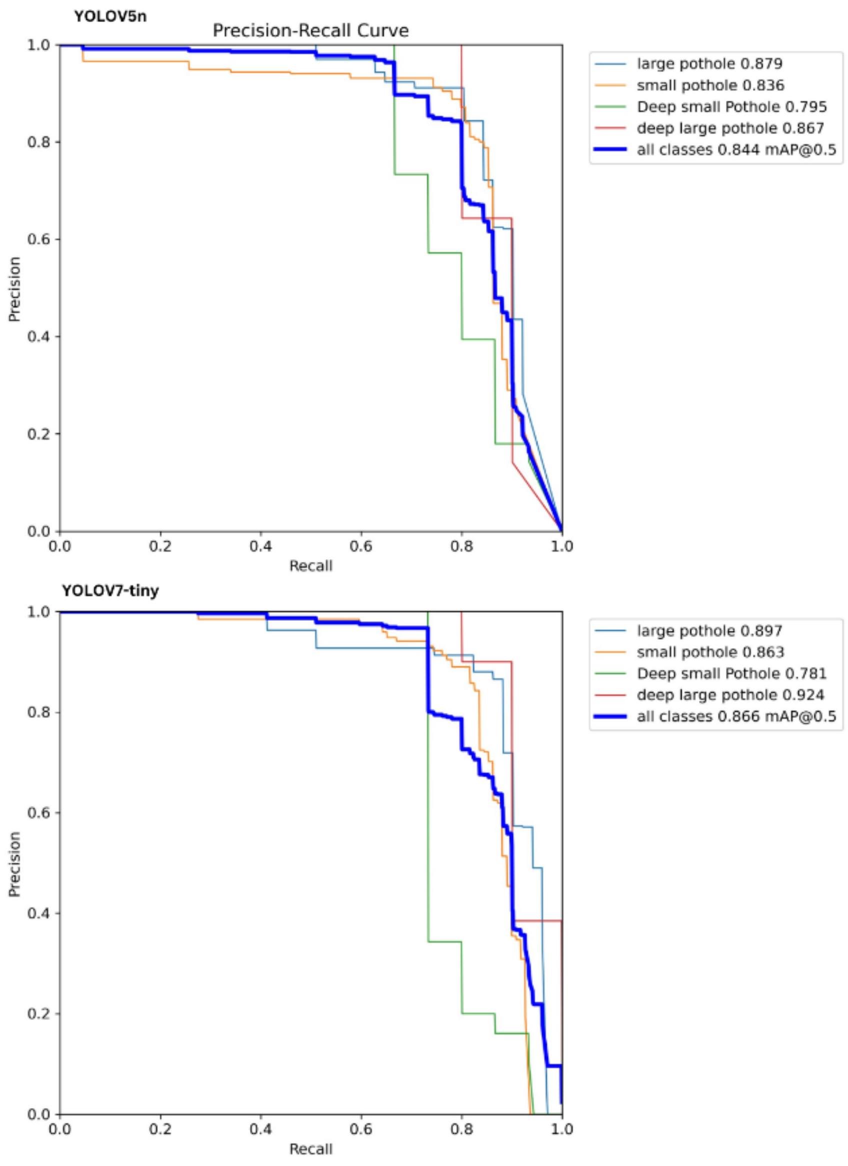

Figure 9. P_R curve for Yolo5n and Yolov7-tiny

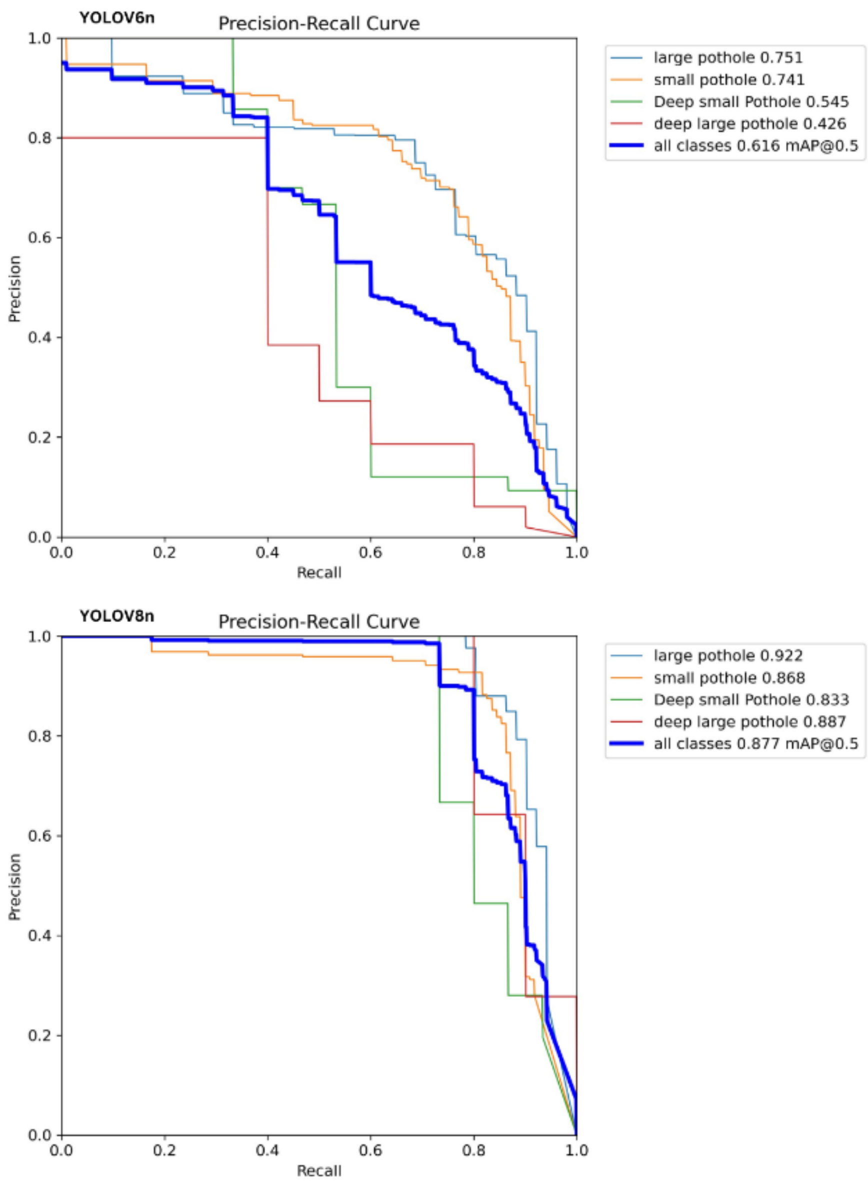

Figure 10. P_R curve for YoloV6nand YoloV8n

3.4 Mean avergae precision

Figure 11 displays the precision (B), recall (B), mean average precision (MAP) at IOU threshold 0.5, and MAP at IOU thresholds 0.5-0.95 curves for each model. The precision (B) curve represents the model's ability to correctly classify positive instances, while the recall (B) curve showcases the model's sensitivity in capturing all positive instances. The MAP at IOU threshold 0.5 indicates the model's overall performance in object detection, considering a moderate IOU threshold. Lastly, the MAP at IOU thresholds 0.5-0.95 provides a broader evaluation of the model's object detection capabilities across various IOU thresholds as in Table 5.

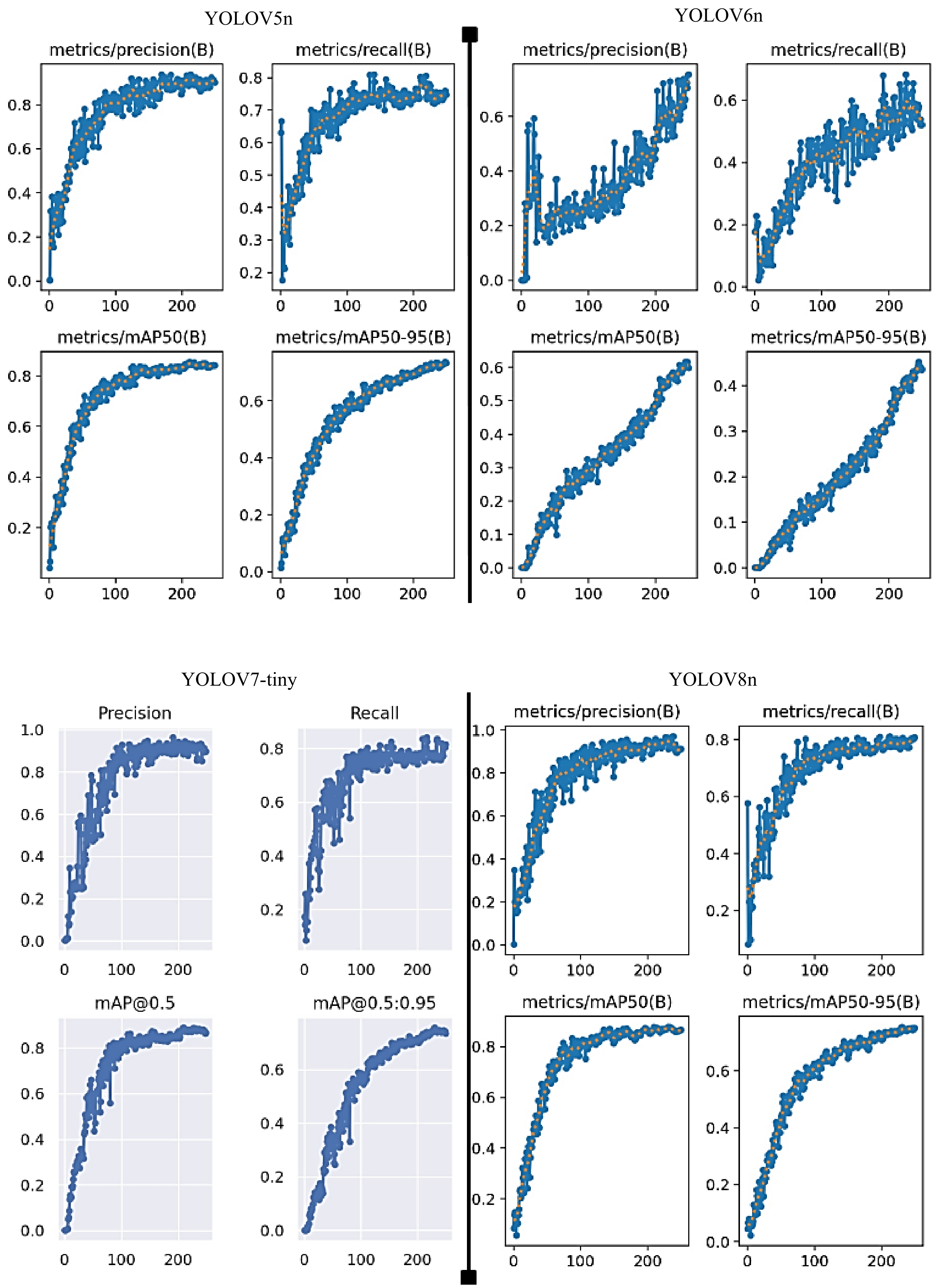

Figure 11. Training and validation results

Table 5. Mean Average preceion results for a traning under 250 epochs

Model	Precision (P)	Recall(R)	MAP@0.5	MAP@0.5:0.95
Yolov6n	0.664	0.58	0.616	0.451
Yolov8n	0.964	0.775	0.877	0.754
Yolov5n	0.912	0.759	0.844	0.738
Yolov7-t	0.898	0.814	0.866	0.736

4.0 CONCLUSION

The analysis of confusion matrices revealed that Yolov8n achieved the highest true positive rates for all classes, indicating superior accuracy in identifying large, small, deep small, and deep large potholes. Yolov7-tiny also performed well, particularly in detecting large potholes. Yolov5n showed promising results for small and deep small pothole detection, while Yolov6n exhibited relatively lower performance across all classes.The F1 curves demonstrated that Yolov7-tiny and Yolov8n consistently achieved higher F1 scores, indicating their ability to balance precision and recall effectively. Yolov5n showed competitive performance, while Yolov6n lagged in achieving high F1 scores. The precision-recall curves further confirmed that Yolov7-tiny and Yolov8n exhibited better precision-recall trade-offs compared to Yolov5n and Yolov6n, indicating their capability to achieve high precision while maintaining reasonable recall rates.

The mean average precision (MAP) analysis at different confidence thresholds consistently favored Yolov8n as the best-performing model, followed by Yolov7-tiny. Yolov5n demonstrated competitive results, while Yolov6n showed relatively lower performance. In conclusion, Yolov8n and Yolov7-tiny are the most effective models for pothole detection, with higher accuracy, precision, and recall rates across all classes. Yolov5n also showed promising performance, while Yolov6n exhibited lower overall performance. For future work, its recommended a dataset expansion to increase the size and diversity of training data, fine-tuning and hyperparameter optimization to further improve model performance, exploration of transfer learning and ensemble models for enhanced detection capabilities, addressing class imbalance issues, and real-time implementation and testing in real-world scenarios to assess practical feasibility and performance. By considering these recommendations, future research can build upon the findings of this study and advance pothole detection techniques, contributing to more accurate and efficient road maintenance and safety measures.

5.0 ACKNOWLEDGEMENT

The authors would like to thank Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah for the chances and support on this project.

REFERENCES

[1] Anandhi, R. J., Baswaraju, S., Silpa, S., Nandagopalan, S. P., & Rao, S. J. (2022). Survey on IOT Based Pothole Detection. 2022 IEEE Delhi Section Conference, DELCON 2022.

[2] Gunawan, F. E., Yanfi, & Soewito, B. (2015). A vibratory-based method for road damage classification. 2015 International Seminar on Intelligent Technology and Its Applications, ISITIA 2015 - Proceeding, 1–4.

[3] Zhang, F., & Hamdulla, A. (2022). Research on Pothole Detection Method for Intelligent Driving Vehicle. 124–130.

[4] Montoya-Alcaraz, M., Mungaray-Moctezuma, A., Calderón-Ramírez, J., García, L., & Martinez-Lazcano, C. (2020). Road safety analysis of high-risk roads: case study in baja California, méxico. Safety, 6(4), 45.

[5] Kuthyar, S. R., Roopashree, S., Rasika, V., Manjesh, S., Girimaji, R., Arjun, R. D., & Priyadarshini, S. (2021, December). An Intelligent Pothole Detection and Alerting System using Mobile Sensors and Deep Learning. In 2021 IEEE 18th India Council International Conference (INDICON) (pp. 1-6). IEEE.

[6] Road Accidents in India 2021, [online] Available: https://morth.nic.in/sites/default/files/RA_2021_Compressed.pdf.

[7] 1481 deaths 3103 injured: Pothole-related accidents on the rise says report, [online] Available: https://english.mathrubhumi.com/news/kerala/1-481-dead-3-103-injured-union-govt-report-shows-pothole-related-accidents-on-increase-18187782.

[8] Ryu, S. K., Kim, T., & Kim, Y. R. (2015). Image-based pothole detection system for ITS service and road management system. Mathematical Problems in Engineering, 2015, 1-10.

[9] Ukhwah, E. N., Yuniarno, E. M., & Suprapto, Y. K. (2019, August). Asphalt pavement pothole detection using deep learning method based on YOLO neural network. In 2019 International Seminar on Intelligent Technology and Its Applications (ISITIA) (pp. 35-40). IEEE.

[10] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 7464-7475).

[11] Pothole Object Detection Dataset-raw Chitholian, [online] Available: https://public.roboflow.com/object-detection/pothole/1.

[12] Kim, Y. M., Kim, Y. G., Son, S. Y., Lim, S. Y., Choi, B. Y., & Choi, D. H. (2022). Review of Recent Automated Pothole-Detection Methods. In Applied Sciences (Switzerland) (Vol. 12, Issue 11). MDPI.