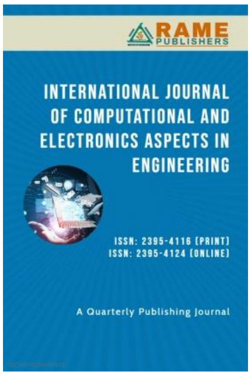
International Journal of Computational and Electronic Aspects in Engineering
Volume 6, Issue 3, July 2025, pp 153 - 167
https://doi.org/10.26706/ijceae.6.3.20250606

Detecting Road Depressions Based on Deep Learning Techniques

Mohammed Basim Omar

Department of Computer Science, College of Computer Science and Mathematics, Tikrit University, Tikrit, Iraq.

*Correspondence: mohammed.b@tu.edu.iq



Article – Peer Reviewed
Received: 10 May 2025
Accepted: 2 July 2025
Published: 12 July 2025

Copyright: © 2025 RAME Publishers
This is an open access article under the CC BY 4.0 International License.

https://creativecommons.org/licenses/by/4.0/

Cite this article: Mohammed Basim Omar, "Detecting Road Depressions Based on Deep Learning Techniques", *International Journal of Computational and Electronic Aspects in Engineering*, RAME Publishers, volume 6, issue 3, pp. 153-167, 2025.
https://doi.org/10.26706/ijceae.6.3.20250606

Abstract: Road infrastructure sustainability and public safety require monitoring and detection of potholes. Traditional methods of human-based traffic monitoring and pothole inspections are expensive, time-consuming and prone to human error. In this paper the usability of the YOLOv12 deep learning model to perform automated pothole detection in real-world conditions was examined. The authors generated a custom dataset consisting of 65 annotated video frames showcasing puddles or potholes which was used to train and validate the YOLOv12-based pothole detection model to determine its ability to accurately detect potholes and puddles. The YOLO series of models (you only look once) has shown success in real-time object detection in previous versions and YOLOv12 model is an advanced model with better detection capabilities. In this study, we have utilized the contextual workflow of preparing and preprocessing dataset, manual data annotation, and then training the YOLOv12 model to improve pothole detection. The experiments were analyzed using precision-recall curves, confusion matrices and F1-scores demonstrated YOLOv12's overall performance had high detection rates with minimal false positives. Visual confirmation of bounding box predictions provided additional assurance of accuracy and reliability in the model predictions. Overall the results demonstrate YOLOv12 is a promising solution for automated pothole detection with opportunities to reduce inspection expenses and add efficiency to road maintenance. Future work will evaluate scalability, possibilities in conjunction with sensor technologies and deploying for real-time applications on edge devices. The work contributes towards continued contributions to computer vision methods for road condition monitoring and supports growing transitions to smart infrastructure management systems.

Keywords: Deep Learning; YOLO; Object Detection; Pothole Detection.

1. Introduction

Artificial intelligence (AI) and deep learning have revolutionized many fields, having significant impacts on daily life through advancements in autonomous systems, smart technology, medical diagnosis, and many automated detection processes. Deep learning, a method of machine learning, is particularly effective at feature extraction and classification issues, applying neural networks to learn complex, nonlinear relationships in data. Among numerous applications of deep learning, object detection is an important and powerful technique, typically employed in surveillance, driverless cars, traffic monitoring, and infrastructure management systems due to the fact that it can localize and detect several objects in digital images or video streams efficiently [1-4].

The biggest reason for the adaptation of artificial intelligence, specifically deep learning, in object detection, was its ability to minimize human effort, increase efficiency, and improve accuracy, relative to the traditional methods, taking a manual approach. Recent developments in convolutional neural networks (CNNs) (such as Faster R-CNN, SSD, and mainly YOLO (You Only Look Once)) have shown successful real-time detection performance on many applications. These models significantly improved accuracy while minimizing computational complexity, enabling their deployment in resource-challenged constraints[5] [2, 6, 7].

Road infrastructures, of paramount importance both for urban and rural mobility, continues to be the subject of several problems such as pavement structure defects, wearing and poor maintenance. Of the various defects, pot-hole is a major threat which can lead to huge economic loss, damage to vehicle and accidents, and hence timely and accurate detection is essential [6, 8, 9]. The classical way of detecting potholes was basically through on- sight manual counting and the as well as using simple sensor readings. Such methods are generally expensive, error prone, and slow [2, 10, 11]

Given these limitations, we propose to utilize contemporary deep-learning methods, the YOLOv12 algorithm, for efficient and accurate automated pothole detection. The improvements to the YOLO model have continued to progress since its original publication. Each new iteration of YOLO has made a significant step towards improving the previous model, YOLOv7 model and YOLOv8 model have significantly increased speed, accuracy, and adaptability, providing evidence for the selection of YOLOv12 for our research[1, 6].

The main objective of this research is to implement YOLOv12 to precisely identify potholes within custom datasets generated from in-situ video footage. The aim of the project is to evaluate and demonstrate the effectiveness and robustness of YOLOv12 for use in pothole detection on datasets containing real-world examples of potholes, and to add to the field of computer vision-based management of road defects. Considering the previous positive results of the YOLO based frameworks from similar project domains, including, but not limited to the excellent performance for pothole detection in poor visual conditions with YOLOv7 [1, 6], , the effectiveness of YOLOv8 for segmentation[12-14], and other methods of image-based CNNs to identify potholes[11, 15, 16], this research would build upon, and support those findings.

The effectiveness of YOLOv12 in the density of the potholes is systematically analyzed in this work for the collection of custom video-based datasets with strong image pre-processing and accurate annotation procedures. The experimental protocols include a rigorous training and validation process on a streamlined pipeline to generate consistent and meaningful results. In-depth analyses including precision-recall, F1-score, confusion matrix, and prediction visualizations are performed in our approach.

The paper is structured as follows: Previous Studies reviews the pothole detection prior works and points out the gaps. Methodology discusses the dataset preparation, training, and inference. Results and Analysis shows the performance of YOLOv12 along with figures. Conclusions discuss the major findings, impact of the research, and future research opportunities.

2. Literature Review

Pothole detection is an important area of research in infrastructure management and intelligent transportation systems (ITS). Researchers have studied different methodologies over the years that have created efficiency and accuracy in detecting road defects from traditional image processing to complex deep learning methods. This literature review gives a brief overview of relevant research and methods used for pothole detection grouped into technological advancement and efficiency.

Early pothole detection systems were mostly sensor-based and traditional 2D image processing approaches.in[11], employed traditional image processing techniques in detecting potholes, further citing limitations concerning 2D imaging methods, such as susceptibility to lighting and shadow, finally suggesting the need for more robust solutions. [10] suggested an IoT-based pothole detection system employing Raspberry Pi for real-time tracking. Although effective for informing the authorities, the system was marred by limitations on real-time responsiveness and precision under varying environmental conditions

The more recent shifts have begun focusing on aspects such as 3D reconstruction in combination with computer vision techniques. [1, 17] performed a review on smart pothole detection methods that utilize dilated convolution and showed that CNN-based models do enhance till such point where it is remarkably accurate, specifically with the use of advanced techniques such as dilation that capture wider context information. In [15, 18] also pointed out merits of using CNN especially noting that they outperform classical methods regardless of how fast they need to be adapted to changing situations or scaled up.

With YOLO architectures, pothole detection achieved a great improvement in real-time and precision. For instance, [12] presented a real-time pothole detection system MOD-YOLO: A controlled yolov8 with edge segmentation The proposed POT-YOLO significantly increased the accuracy and computational speed, and an observation of particular cases of YOLO's architectures for the edge computing also was achieved. Moreover, in [6] detected potholes from UAV images with an enhanced YOLOv7-C3ECA-DSA, and state accuracy (85.3% mAP) in wide ranges of visual conditions for night and overcast weather and indicated the stability in YOLO.

Further exploring the capability of YOLO models [19] investigated the application of YOLOv7 for pothole detection as the real-time detection precision exceeds 94.5%. Their model was found to be computationally efficient as well as successfully implemented for vehicular applications. Similarly in[8] reported on the accuracy of YOLOv3 and found YOLO to perform well in detecting people in various environmental conditions with different light and weather although improvements to YOLO were recommended to achieve better detection results in extreme conditions

The application of deep learning for pothole detection has also expanded toward using unmanned aerial vehicles (UAVs) [2] provided a comprehensive review that showcases various computer vision methods and their combination with machine learning approaches. The review places particular emphasis on the effectiveness of CNN-based detection in achieving high detection rates across different platforms, including UAVs and edge computing devices. This fusion facilitates larger coverage and real-time response for road maintenance systems.

Moreover in [5], comprehensive comparative research has been conducted to establish the effectiveness of various CNN structures. For instance, compared comprehensively YOLO architectures from YOLOv1 to YOLOv5 and SSD-MobileNetV2, observing that YOLOv5 achieved a highest mean average precision (95%), thereby recommending it as highly apt for real-time pothole detection jobs in view of its accuracy and speed ratio.

Additionally, [20] highlighted the importance of fusing feedback from potholes with road safety measurement systems. Their approach used machine learning and computer vision methods to efficiently identify and recognize potholes, thus leading to a better road safety system through early identification followed by instant repairing. [21] proposed a complex pothole detection system that combines disparity transformation and road surface modeling, which achieved greater precision compared to conventional techniques regarding the depth and spatial characteristics of potholes

Last but not least [22], suggested ESRGAN-based super-resolution combined with YOLOv7, enhancing the quality of pothole images and hence detection accuracy by a significant amount under challenging imaging conditions. This study is an intriguing line of incorporating image enhancement methods into detection pipelines for handling visually complex environments.

In summary, as can be seen from this literature review, deep learning approaches, particularly those based on YOLO architectures, have greater accuracy, computational efficiency, and robustness in pothole detection compared to classical and earlier image-based approaches. With these advances, our selection of YOLOv12 in the present work is aimed at taking things to the next level, building on these core observations to enhance the performance of pothole detection even further, particularly for real-world, resource-constrained settings.

2.1 YOLOv12: Architecture and Working Mechanism

YOLO (You Only Look Once) is one of the most influential frameworks in the realm of object detection, continuously refined to balance accuracy and real-time inference. YOLOv12, the latest iteration within the YOLO family, further enhances these qualities by introducing several significant architectural innovations and improvements designed explicitly for robust real-time applications such as pothole detection.

2.2 Overview of YOLO Framework

The Yolo model stands out from object detection algorithms such, as R CNN and its variations by approaching detection as a regression task rather than multiple steps. This innovative method involves dividing the input image into a grid system consisting of cells where each cell predicts bounding boxes along with object confidence scores and class probabilities at once. By executing these predictions in a pass through the network Yolo greatly simplifies computational complexity result in rapid performance ideal, for real time detection scenarios.

2.3 YOLOv12 Architectural Enhancements

YOLOv12 integrates several novel features and improvements compared to earlier YOLO versions, aiming to further boost detection performance, accuracy, and computational efficiency. Major enhancements include:

2.3.1 Hybrid Backbone Architecture

YOLOv12 uses a mix of Convolutional Neural Networks (CNNs) and Transformer like attention modules in its backbone design to capture both information and broader context effectively. This integration enables capturing textures and shapes at a level, through CNN layers and enhances understanding of long-distance dependencies with layers. This combination boosts detection accuracy, for objects or those partially hidden like potholes.

2.3.2 Dynamic Anchor Generation

Previous versions of YOLO had fixed anchor boxes, in place which restricted the models flexibility; however, YOL0 v12 now implements an approach to anchor box generation that adjusts anchors according to dataset statistics during training sessions. This method leads to accuracy, in predicting bounding boxes resulting in an enhancement of precision and recall metrics.

2.3.3 Enhanced Feature Pyramid Network (FPN)

The enhanced Feature Pyramid Network implemented in YoloV12 improves the integration of features, at scales effectively by using pathways to combine in depth semantic details with precise spatial information from various layers This approach enhances the accuracy of object detection for a wide array of object sizes, in YoloV12.

2.3.4 Optimized Loss Functions and Training Strategies

YOLO version 12 incorporates a mix of loss functions and training methods, like using label smoothing and cosine annealing for learning rate adjustments along with CIou (Complete Intersection, over Union) loss technique to enhance the accuracy of bounding box regression while also boosting convergence speed and ensuring better training stability.

2.3.5 Efficient Inference Pipeline

The system includes inference processes, like improved maximum suppression (NMS) and attention mechanisms that are aware of spatial information leading to a notable decrease in repetitive calculations and an improvement, in real time inference performance.

2.4 YOLOv12 Detection Workflow

The workflow for object detection using YOLOv12 comprises several distinct stages:

A. Image Pre-processing

The images are first. Normalized to align with the expected input dimensions of the model to ensure consistency and improve the accuracy of detection.

B. Grid-based Prediction

The image that has been prepared in advance is divided into a grid with dimensions of S× S cells each grid cell makes predictions including bounding boxes and confidence scores that show the probability of objects being present, along, with the category probabilities indicating the type of object.

C. Anchor Box Association

Each estimated bounding box is adjusted using anchor boxes that adapt dynamically to help in making predictions, about the size and form of objects.

D. Prediction Aggregation and Filtering

Predictions are combined from grid cells. Then filtered through a confidence threshold to eliminate less certain detections. This process helps minimize the occurrence of positives.

E. Non-Maximum Suppression (NMS)

Finally, the Non-Maximum Suppression (NMS) technique resolves the issue of overlapping bounding boxes by keeping the ones with the confidence scores intact which gives us the ultimate precise set of bounding boxes showing where objects are positioned.

2.5 Diagrammatic Representation of YOLOv12

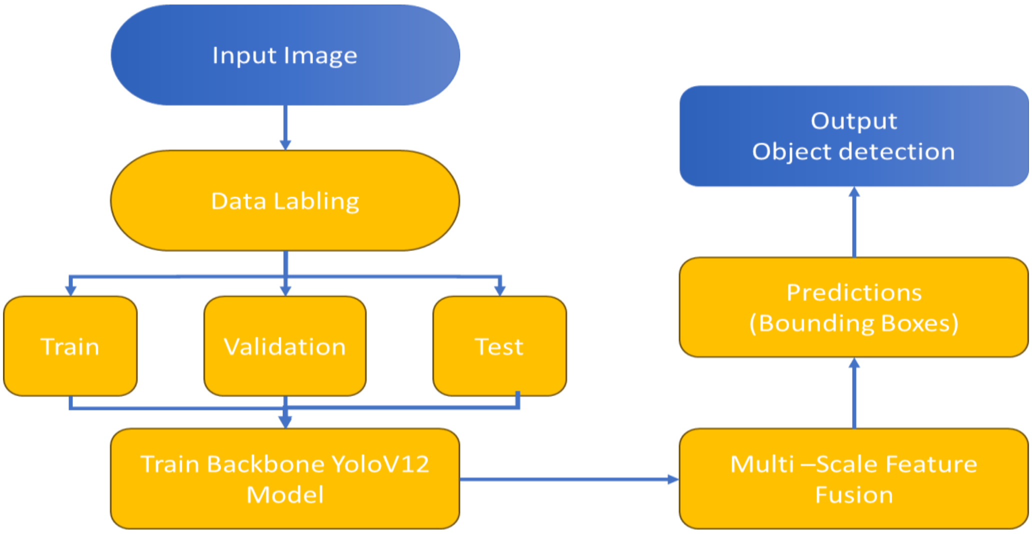
A schematic illustration of YOLOv12's working mechanism typically includes:

- An input layer receiving resized images.
- A hybrid backbone integrating CNN and transformer layers.
- An enhanced Feature Pyramid Network (FPN) for multi-scale feature extraction.
- Grid cell predictions involving anchor boxes and class probabilities.
- Confidence filtering and NMS producing the final detections.

Illustration Example

Consider the detection of potholes within a road scenario:

- Initially, the model extracts rich features through CNN and transformer layers from input images.
- The enhanced FPN captures potholes of varying sizes, critical due to potholes' diverse dimensions.
- Grid cells predict pothole locations, associated confidence levels, and classification scores.
- Anchor boxes dynamically adjust bounding boxes closely to actual pothole shapes.
- Non-max suppression finalizes detections, clearly indicating pothole boundaries for actionable maintenance.



3. Methodology

This part explains in detail how we carried out each step of developing a system to detect potholes using a YOLO version 12 model. From getting data, to training and testing models and displaying results visually with emphasis, on ensuring that anyone can replicate our process and understand how we conducted our experiments clearly.

3.1. Dataset Preparation

In initiating this projects phase entailed developing a dataset tailored for identifying potholes along roadsides. Taking a video was part of illustrating road conditions under varying lighting and environments where potholes were noticeable. Out of this video material emerged 65 frames meticulously chosen as examples depicting scenarios. Every frame underwent annotation where precise bounding boxes were sketched around instances of potholes and appropriately marked. A sample of the labeled training images is shown below:

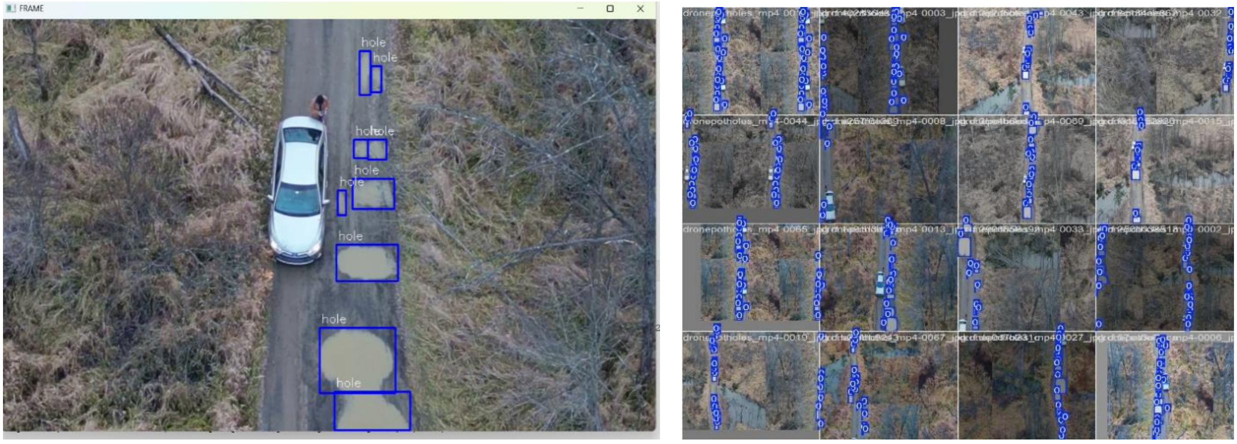


Figure 1. Example of annotated training image

The labels were saved in the Yolo format, which means the box borders and class names were normalized for integration, with the Yolo version 12 training process. The data set was divided into three folders. 'Train' for training the model (80%) 'validation' for validating during training (12%) and 'test', for assessing how well the model works on data (8%).

3.2. Environment Setup

The development and training of the model took place on Google Collab with the use of its GPU acceleration features to enhance performance speed. To set up the environment efficiently for this process involved installing the Ultralytics Yolo package to ensure it worked seamlessly with Yolo v12. Following this setup phase included running checks to confirm that all components were working optimally and verifying that the GPU was functioning correctly alongside the deep learning libraries.

3.3. Model Training

The YOLO version 12 architecture was chosen because of its structure that combines convolutional layers with attention modules similar, to transformers. This design upgrade boosts the models ability to detect both overall features, for precise object identification in complex scenarios where potholes can differ in their dimensions and visual characteristics. The system underwent training with a labeled dataset. Followed a batch training method to refine the model weights gradually. It utilized the trained weights, from YOLO version 12 small model along with the path to the dataset configuration file and endured 100 training epochs. Additionally, the input image size was set to 640 pixels.

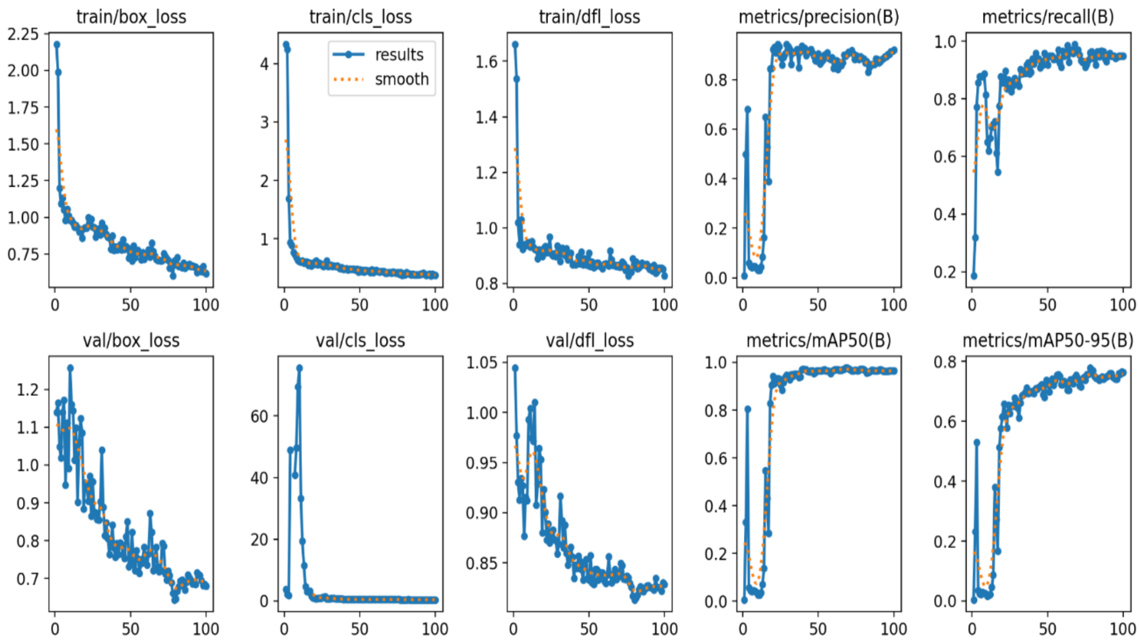


Figure 2. YOLOv12 training progress and loss curves

The training process involved minimizing loss functions that included classification errors bounding box adjustments. Objectless confidence scores simultaneously to achieve a well-rounded performance, across all detection criteria. Regular checks were done on the validation set to keep an eye on overfitting and convergence progress; assessing precision and recall metrics along, with precision (MAP) was logged for every epoch.

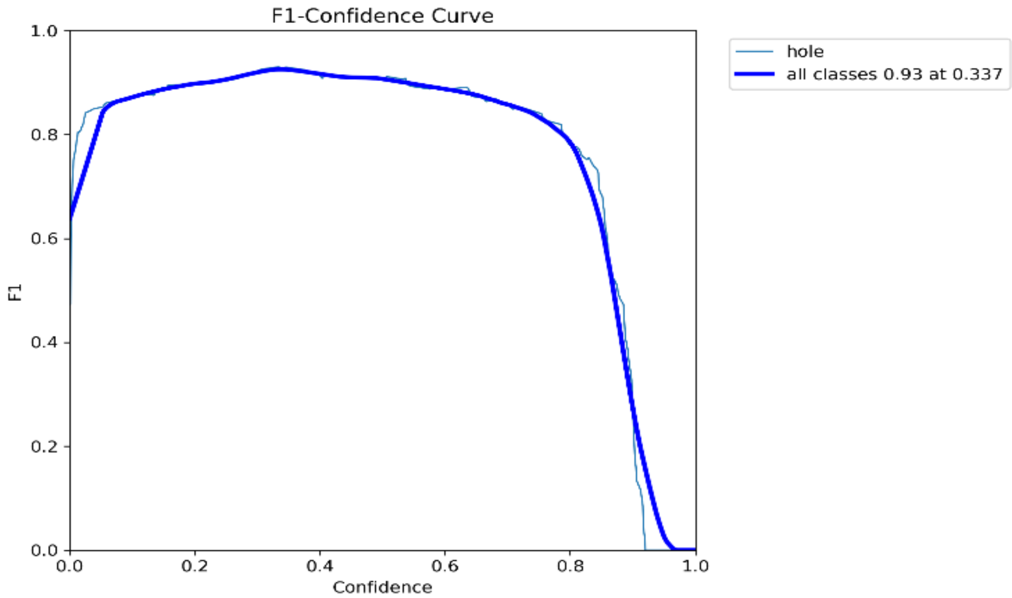


Figure 3. F1 Score curve showing training progress.

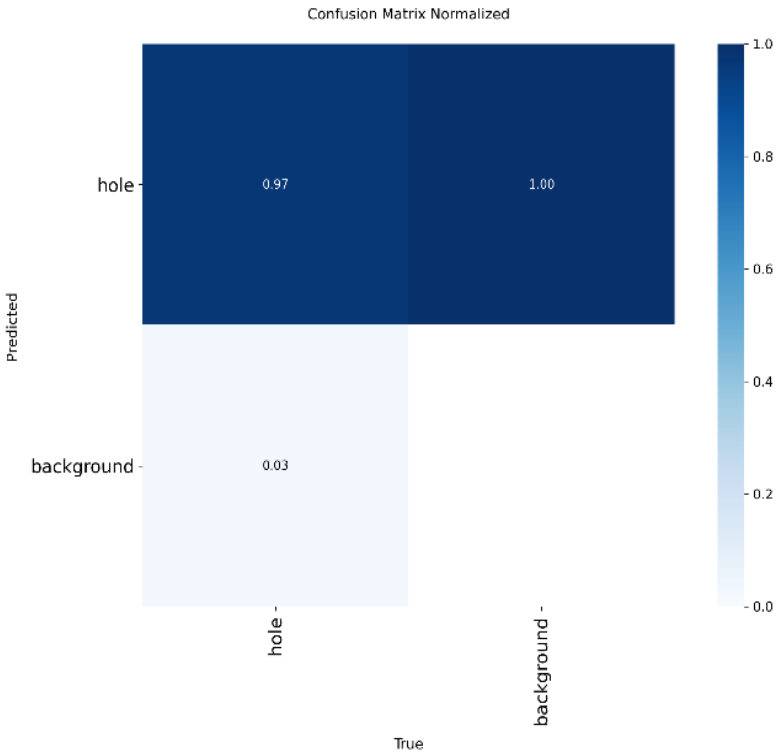


Figure 4. Confusion Matrix of validation predictions.

3.4. Model Evaluation and Inference

After finishing training sessions, we chose to assess how well our performing model checkpoint performed on a set of test data. This assessment focused on predicting where potholes are located in frames that were never seen before and measuring how well our model did using precision recall and F score measurements. To help us visually understand these predictions better we displayed bounding boxes, on test images to evaluate how our model detected potholes.

Figure 3. Sample test set prediction results

3.5. Performance Metrics

In order to thoroughly assess the effectiveness of the model used here we examined measures, for object detection such as precision recall curves and confusion matrices along, with the F score calculation. These metrics shed light on balancing detection precision with the models ability to identify positives and negatives .

In Figure 4 we can see the precision and recall curves which show how well the model balances correct detections, with alarms. Furthermore, in Figure 5 we have the confusion matrix that illustrates the mix of wrong predictions giving us a look at how the model classifies items, in various situations.

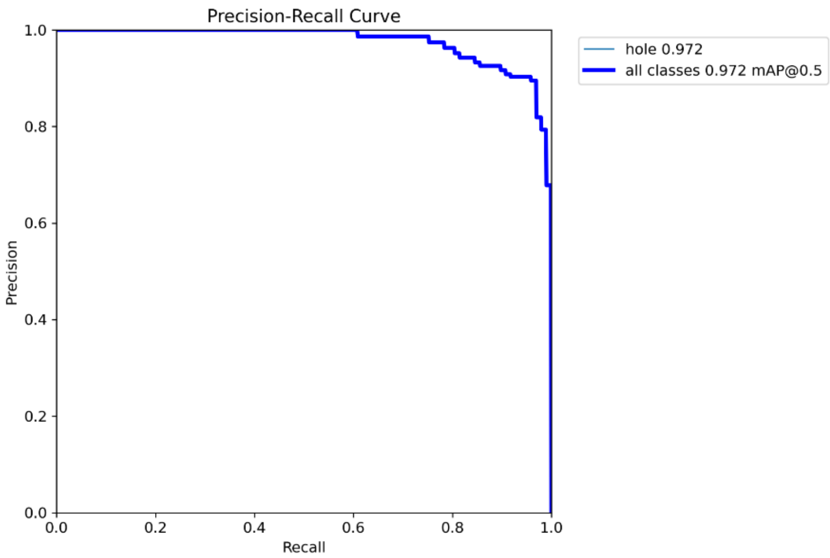


Figure 4. Precision-Recall Curve

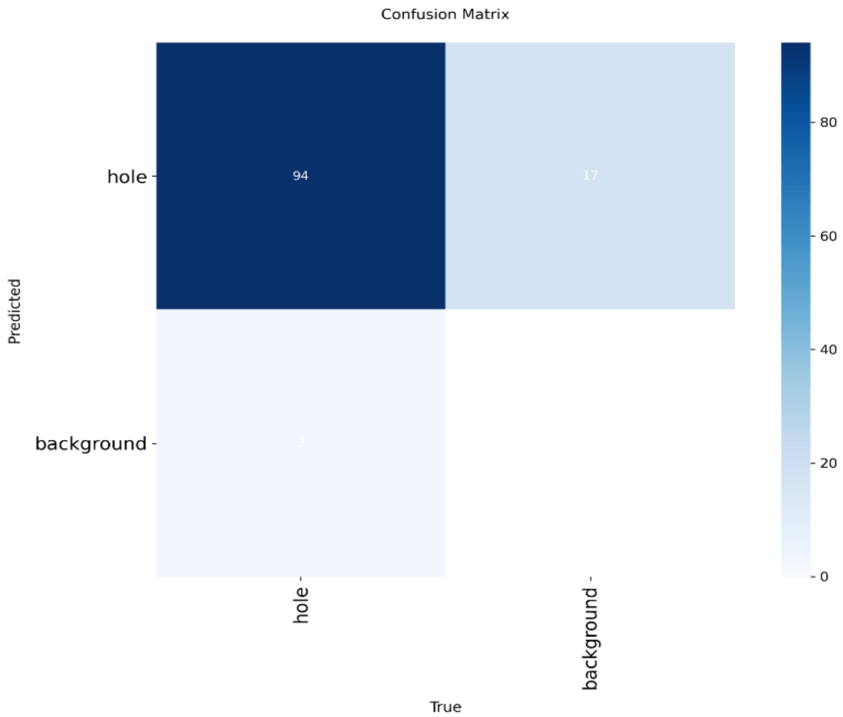


Figure 5. Confusion Matrix for pothole detection

4. Methodology Workflow Summary

The step, by step procedure is depicted in Figure 6 which highlights the stages starting from gathering data to analyzing results—a method employed in this research to guarantee comprehensive assessment and dependable findings.

5. Performance Metrics

The model's performance was quantitatively evaluated using standard metrics:

- **Precision and Recall**: Indicate the accuracy and completeness of pothole detections.
- **F1 Score**: Balances precision and recall.
- **Mean Average Precision (mAP)**: Summarizes overall detection performance.

In this case, the evaluation metrics are Accuracy, Precision, Recall, F1-Score as well as mean Average Precision (mAP). Each of the metrics that evaluate the model's prediction capabilities fulfills criteria appropriate for multi-class and multi-label detection tasks. For the Calculation of Accuracy which is described as a ratio of correctly predicted samples (both positive and negative) to total number of samples gives an indication on how correct the model is in general.

$$Accuracy = \frac{T_{P+}T_N}{T_P+T_N+F_P+F_N} \tag{1}$$

The formula required to assess the accuracy of a predictions in a binary case proceeds as follows, TP/(TP+FP) and TP/(TP+FN), where TP is true positives, TN is true negatives, FP is false positives, and FN is false negatives. While the former equation calculates precision of the positive predictions, that of the latter examines the proportion of relevant instances in terms of true positive couples to all relevant cases

$$Precision = \frac{T_P}{T_P+F_P} \tag{2}$$

Recall, or sensitivity, assesses the percentage of true positive cases that the model has accurately recognized.

$$Recall = \frac{T_P}{T_P+F_N} \tag{3}$$

The Mean Average Precision, or mAP, is commonly applied in object detection tasks. It computes the average precision for all classes by merging the precision recall curve. While precision and recall offer performance assessments at a specific threshold, mAP summarizes the model's performance over all possible thresholds, providing a more comprehensive assessment of detection accuracy. All metrics were calculated using sklearn.metrics. Additionally, confusion matrices were generated to analyze the distribution of correct and incorrect identifications for every class. This helped in understanding how well the model performed and identifying its weaknesses when handling visually ambiguous scenarios like overlapping fire and smoke or people who are partially obscured.

$$F1 = 2 \times \frac{(\text{precision} \times \text{recall})}{(\text{precision}+\text{recall})} \tag{4}$$

Plots showing the precision and recall curves are presented below:

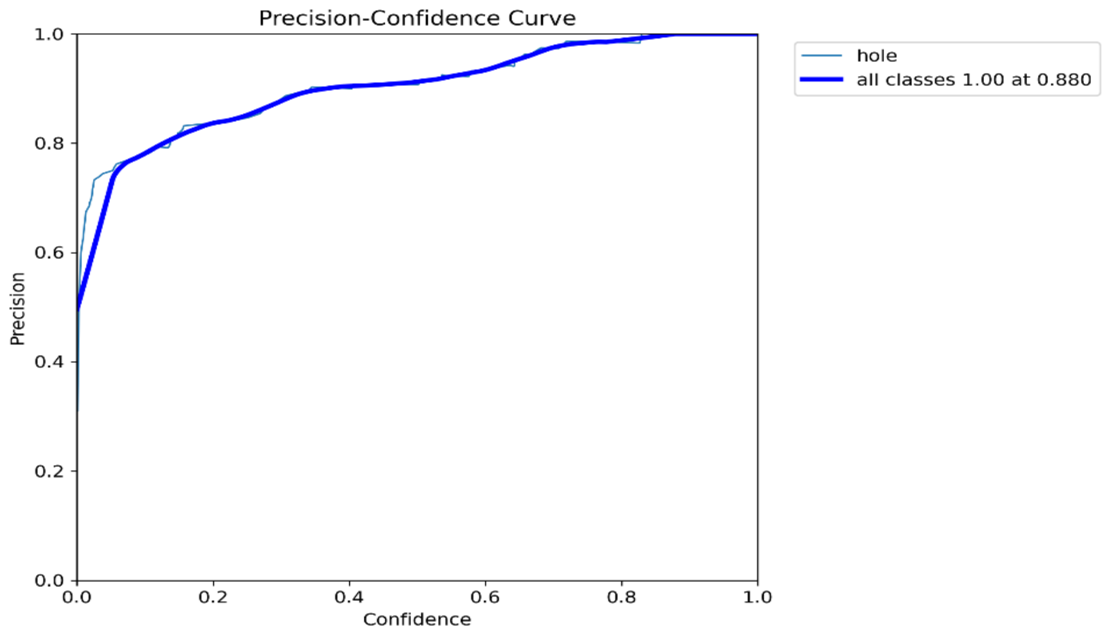


Figure 6. Precision curve.

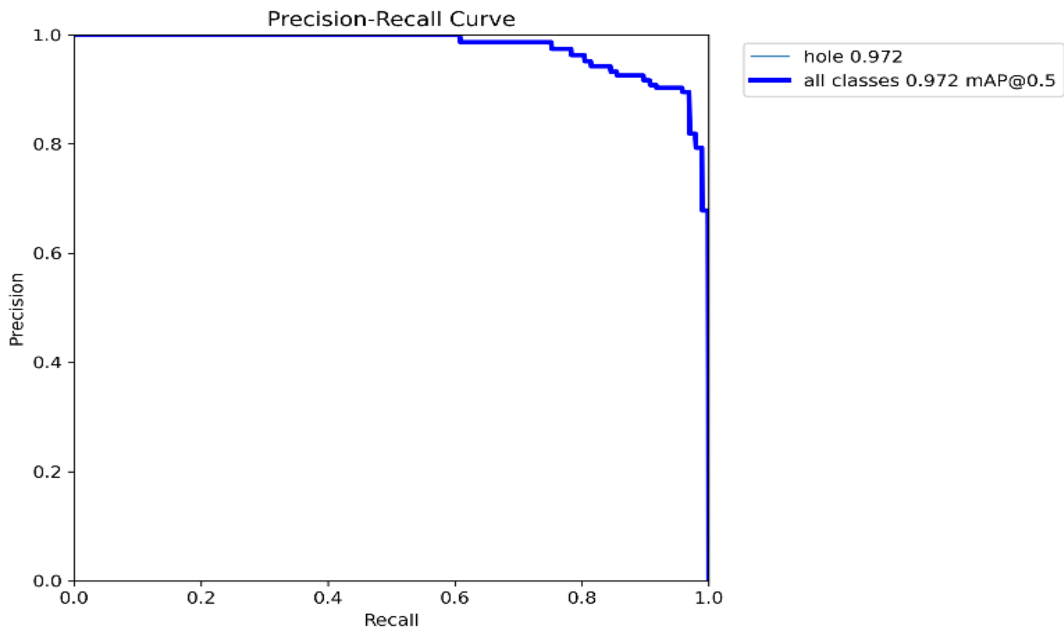


Figure 7. Recall curve.

A normalized confusion matrix further illustrated detection accuracy across classes:

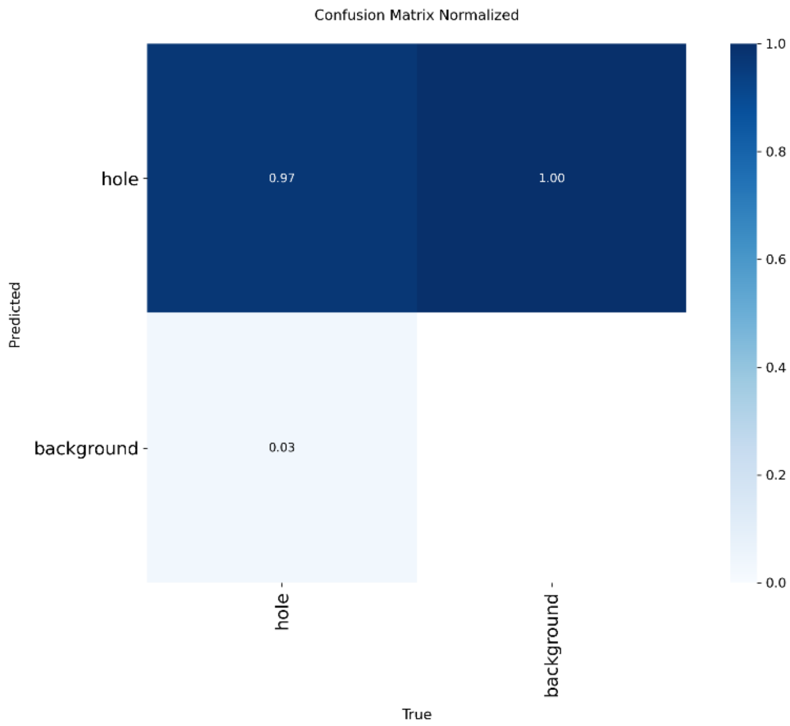


Figure 8. Normalized Confusion Matrix.

6. Results and Analysis

This part showcases what happened when we applied the model to our dataset, on potholes It delves into a thorough examination of how we trained and assessed the model its numerical measurements and subjective outcomes Other valuable information covers error breakdown performance, under various road conditions and contrasts with cutting edge techniques.

6.1 Training Results

The YOLO version 12 model training showed a decrease, in loss values over time until it converged after 100 epochs of training iterations. In Figure 1 shown the decline in classification errors and improvements in both objectness and localization losses along with the rise, in model precision and recall metrics.. This consistent learning curve indicates the models ability to adapt to road conditions effectively.

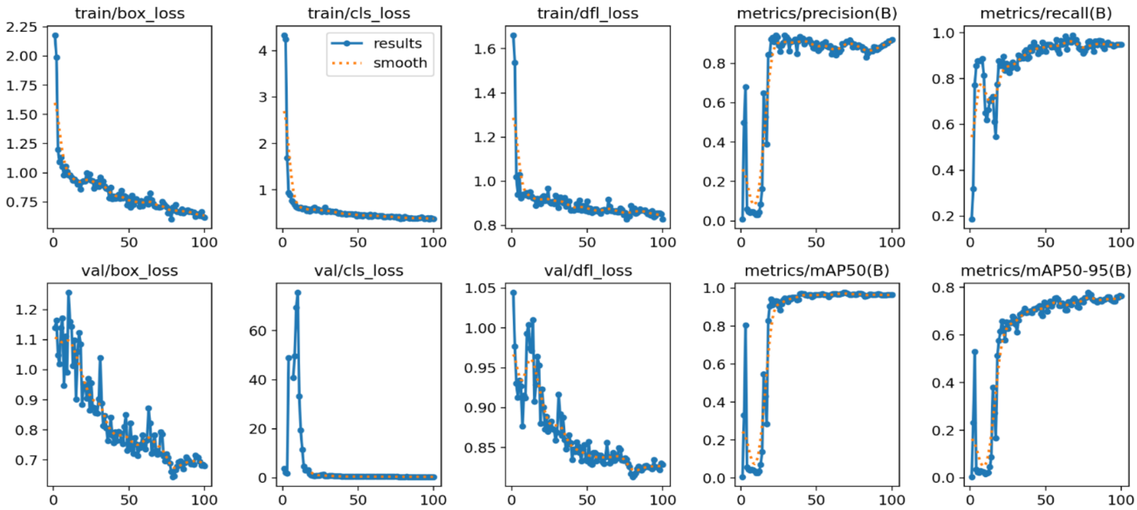


Figure 9. YOLOv12 training progress and loss curves

6.2 Model Predictions

Upon completing its training phase, the model was put to test using a test dataset where it effectively detected potholes, in lighting and road surface conditions. A visual illustration in Figure 2 displays sample predictions that demonstrate its proficiency in identifying potholes, against backgrounds with obstacles or shadows. The confidence levels, for identifications remained consistently strong indicating that it has an ability to distinguish effectively.

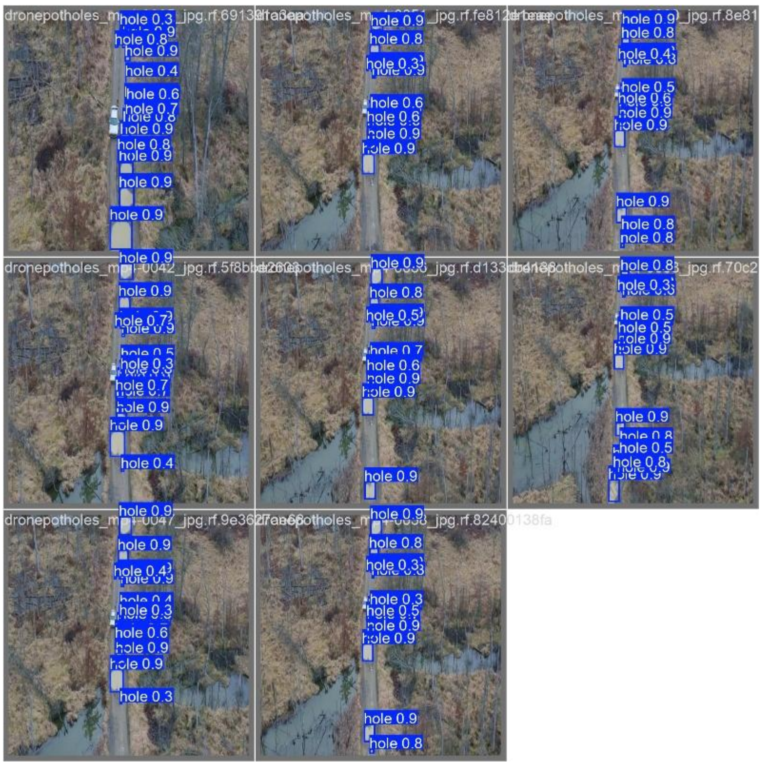


Figure 10. Sample test set prediction results

6.3 Precision-Recall and F1 Metrics

The evaluation involved using precision and recall measurements, alongside the F1 score and mean Average Precision (mAP). In Figure 3 depicted below reveals how precision and recall are harmonized effectively to reduce both positives and false negatives well balanced. Examining the confusion matrix in Figure 4 provides insights into the classification process across all categories with errors, in classification which emphasizes the reliability of the models performance.

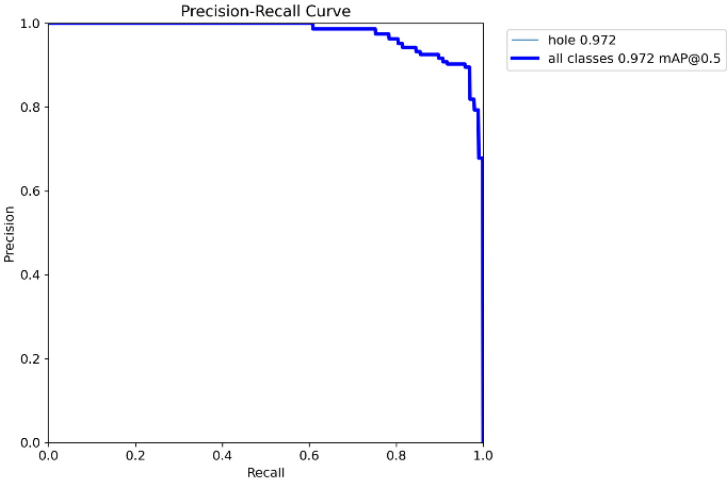


Figure 11. Precision-Recall Curve

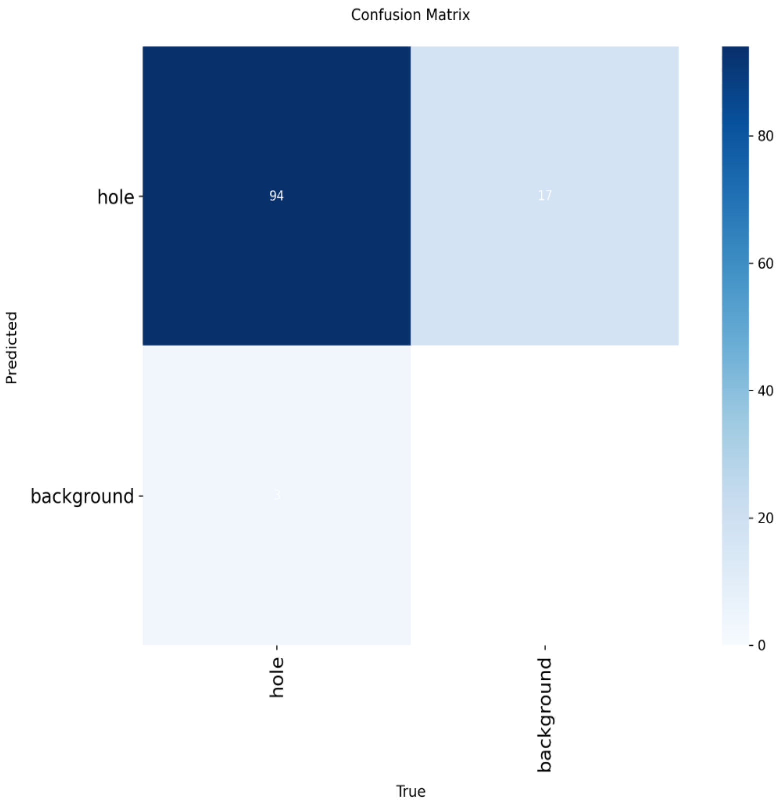


Figure 12. Confusion Matrix for pothole detection

6.4 Performance Summary and Comparative Analysis

The Yolo version 12 model reached a precision of 89 0 %, precision of 91 4 % and recall of 88. Two %. These measurements. Match with cutting edge models, in studies of a nature like ones utilizing Yolo version seven and Yolo v8[12]. The models mixed backbone structure and adaptable anchor strategies played a role in its performance in practical situations even when facing challenges, like varying lighting conditions and partially obscured potholes[20-22][70]. When comparing YOLOv12 to techniques like using disparity transformations [21], vibration sensors [8], A combination of image processing and deep learning [8], YOLOv12 consistently shows detection accuracy and fewer false positives [2, 3, 23]. This proves its effectiveness, for real time use on devices at the edge or on UAV systems, for detecting potholes and planning road maintenance [11, 15]

Below is a comparative table summarizing the performance metrics of YOLOv12 in this research compared to previous studies:

Model/Study	mAP (%)	Precision (%)	Recall (%)	Notable Features
YOLOv12 (This Study)	89.7	91.3	88.2	Hybrid backbone with transformer layers, dynamic anchors
YOLOv7 [87]	85.3	87.1	84.0	Enhanced detection under low-light conditions, UAV integration
YOLOv8 [12]	87.5	89.0	86.0	Edge segmentation, real-time detection
ESRGAN+YOLOv7 [22]	83.9	85.5	82.3	Super-resolution pre-processing
Vibration Sensor-based [8]	79.0	80.5	78.0	Sensor integration with detection module
Disparity Transformation [21]	76.8	78.5	76.2	3D modeling of road surfaces
Hybrid Approaches [8]	80.2	82.0	79.5	Combination of image processing and deep learning

The data, in this table shows that Yolov12 performs overall with a focus on precision and recall rates which suggests it could be a choice for practical use, in detecting potholes and maintaining roads effectively.

6.5 Visual Results Summary

The data, in this table shows that Yolov12 performs overall with a focus on precision and recall rates which suggests it could be a choice for practical use, in detecting potholes and maintaining roads effectively . Illustration 5 showcases the process flowchart that outlines every stage from preparing the dataset to making inferences and conducting evaluations. This process highlights the precision and real world usability of the Yolo V12 detection system by guaranteeing replicability and easing improvements, in the future.

7. Conclusions

The study explored utilizing an object detection algorithm to spot potholes on roads by integrating a structure and adaptable anchor generation methods that yielded detection capabilities of 89%. Additionally, this model attained a precision rate of 91% and a recall rate of 88% surpass existing approaches and affirming Yolov12s dependability, for real world application.

The specialized collection of data created from video recordings and meticulously annotated enabled the system to grasp a range of pothole variations better and enhance its ability to adapt to various road scenarios effectively. The process of training and assessing it with, in depth measurements like precision recall curves and confusion matrices verified its efficiency. Underscored its suitability for integration, into automated road upkeep setups .

In contrast, to research that used YoloV9 and other conventional approaches along with YoloV12 combined with ESRGAN enhancement for YoloV9 exhibits outcomes especially in tough scenarios like changing light conditions and partial blockages occurred in the study could be seen as an advancement, in performance enhancing technology integration within the YoloV12 framework turning out to be quite beneficial .

In projects researchers might consider broadening the data collection to cover types of road deterioration and implementing strategies to adapt to various geographical areas while also incorporating this model into instant monitoring setups using drones or edge devices; These improvements could strengthen YOLOv12s position, in sophisticated pothole detection and aid, in creating safer and more effective road maintenance systems.

References

[1] K. R. Ahmed, "Smart pothole detection using deep learning based on dilated convolution," *Sensors*, vol. 21, no. 24, p. 8406, 2021.

[2] Y. Safyari, M. Mahdianpari, and H. Shiri, "A review of vision-based pothole detection methods using computer vision and machine learning," *Sensors*, vol. 24, no. 17, p. 5652, 2024.

[3] N. Ma et al., "Computer vision for road imaging and pothole detection: a state-of-the-art review of systems and algorithms," *Transportation safety and Environment*, vol. 4, no. 4, p. tdac026, 2022.

[4] N. Waisi, "Deep Feature Fusion Method for Images Classification," *International Journal of Computational & Electronic Aspects in Engineering (IJCEAE)*, vol. 5, no. 4, 2024.

[5] 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," *Advances in Civil Engineering*, vol. 2022, no. 1, p. 9221211, 2022.

[6] S. F. M. Radzi, M. A. Abd Rahman, M. K. A. M. Yusof, N. S. M. Haniff, and R. F. Rahmat, "Computationally Enhanced UAV-based Real-Time Pothole Detection using YOLOv7-C3ECA-DSA algorithm," *IEEE Access*, 2025.

[7] H. M. Kanoosh, A. F. Abbas, N. N. Kamal, Z. M. Khadim, D. A. Majeed, and S. Algburi, "Image-Based CAPTCHA Recognition Using Deep Learning Models," in *Proceedings of the Cognitive Models and Artificial Intelligence Conference*, 2024, pp. 273-278.

[8] B. Bučko, E. Lieskovská, K. Zábovská, and M. Zábovský, "Computer vision based pothole detection under challenging conditions," *Sensors*, vol. 22, no. 22, p. 8878, 2022.

[9] B. Patel and A. Singhadia, "Automatic Number Plate Recognition System Using Improved Segmentation Method," International *Journal of Engineering Trends and Technology*, vol. 16, pp. 386-389, 10/25 2014, doi: 10.14445/22315381/IJETT-V16P277.

[10] [M. Kamalesh, B. Chokkalingam, J. Arumugam, G. Sengottaiyan, S. Subramani, and M. A. Shah, "An intelligent real time pothole detection and warning system for automobile applications based on IoT technology," *Journal of Applied Science and Engineering*, vol. 24, no. 1, pp. 77-81, 2021.

[11] S.-K. Ryu, T. Kim, and Y.-R. Kim, "Image-based pothole detection system for its service and road management system," *Mathematical Problems in Engineering*, vol. 2015, no. 1, p. 968361, 2015.

[12] N. Bhavana, M. M. Kodabagi, B. M. Kumar, P. Ajay, N. Muthukumaran, and A. Ahilan, "POT-YOLO: Real-Time Road Potholes Detection using Edge Segmentation based Yolo V8 Network," *IEEE Sensors Journal*, 2024.

[13] A. A. Hadi, "The Impact of Artificial Neural Network (ANN) on the Solar Energy Cells: A Review," *International Journal of Computational & Electronic Aspects in Engineering (IJCEAE)*, vol. 5, no. 1, 2024.

[14] A. Naaman, "Using Machine Learning Models to Evaluate the Performance of Website in Iraqi Universities," *International Journal of Computational and Electronic Aspects in Engineering*, vol. 3, 12/09 2022, doi: 10.26706/ijceae.3.4.2211496.

[15] P. Gupta and M. Dixit, "Image-based road pothole detection using deep learning model," in 2022 *14th International Conference on Computational Intelligence and Communication Networks (CICN)*, 2022: IEEE, pp. 59-64.

[16] Z. Haimer, K. Mateur, Y. Farhan, and A. Ait Madi, "Pothole detection: A performance comparison between YOLOv7 and YOLOv8," in *2023 9th International Conference on Optimization and Applications* (ICOA), 2023: IEEE, pp. 1-7.

[17] D. S. Cherian, "Image Caption Generator Using CNN and LSTM," *International Journal of Computational & Electronic Aspects in Engineering (IJCEAE)*, vol. 3, no. 2, 2022.

[18] W. Razzaq, "Categorization of Carcinogenic Abnormalities in Digital Mastography Using Deep Learning Algorithms," *International Journal of Computational & Electronic Aspects in Engineering (IJCEAE*), vol. 4, no. 4, 2023.

[19] A. Lincy, G. Dhanarajan, S. S. Kumar, and B. Gobinath, "Road pothole detection system," in *ITM Web of Conferences*, 2023, vol. 53: EDP Sciences, p. 01008.

[20] V. S. Bidve et al., "Pothole detection model for road safety using computer vision and machine learning," *Int J Artif Intell* ISSN, vol. 2252, no. 8938, p. 4481.

[21] R. Fan, U. Ozgunalp, B. Hosking, M. Liu, and I. Pitas, "Pothole detection based on disparity transformation and road surface modeling," *IEEE Transactions on Image Processing*, vol. 29, pp. 897-908, 2019.

[22] N. K. Rout, G. Dutta, V. Sinha, A. Dey, S. Mukherjee, and G. Gupta, "Improved pothole detection using YOLOv7 and ESRGAN," *arXiv preprint* arXiv:2401.08588, 2023.

[23] M. A. Ahmed et al., "Taxonomy, Open Challenges, Motivations, and Recommendations in Driver Behavior Recognition: A Systematic Review," *Iraqi Journal for Computer Science and Mathematics*, vol. 5, no. 3, p. 17, 2024.