A VISION-BASED INTELLIGENT SYSTEM FOR POTHOLE DETECTION AND INFRASTRUCTURE MAINTANANCE

Saran S, Varshanth D, Jotheshvar S

***Saran S**, Department of Artificial Intelligence and Data Science, SRM Valliammai Engineering college, Tamilnadu, India*
***Varshanth D,** Department of Artificial Intelligence and Data Science, SRM Valliammai Engineering college, Tamilnadu, India*
***Jotheshvar S,** Department of Artificial Intelligence and Data Science, SRM Valliammai Engineering college, Tamilnadu, India*

---***---

Abstract— Road infrastructure safety is a critical concern in modern transportation systems, as surface anomalies such as potholes significantly contribute to traffic accidents, vehicle damage, and increased maintenance costs. Conventional road inspection methods rely heavily on manual surveys or sensor-based systems, which are time- consuming, expensive, and unsuitable for large-scale deployment. This paper presents an AI-powered road surface anomaly detection framework based on the YOLOv4-Tiny deep learning model, designed for real-time detection with low computational overhead. The proposed system processes road images and video streams to automatically identify potholes using a single-stage object detection approach. YOLOv4-Tiny is selected due to its favorable balance between detection accuracy and inference speed, enabling deployment on CPU-based and edge devices without GPU acceleration. The framework incorporates preprocessing, model inference, post-processing, and visualization modules to deliver an end-to-end detection pipeline. Experimental evaluation on annotated road surface datasets demonstrates reliable detection performance under varying road and lighting conditions. The system is evaluated using standard metrics including Accuracy, Precision, Recall, F1-score, and ROC-AUC. Results indicate that the proposed approach achieves competitive accuracy while maintaining real-time performance, making it suitable for smart transportation and road maintenance applications. The proposed framework offers a scalable, cost-effective, and deployable solution for automated road condition monitoring in smart city environments.

Keywords— Fake News Detection, Machine Learning, Natural Language Processing, Social Media, Instagram, Twitter, Logistic Regression, Stochastic Gradient Descent, Multinomial Naïve Bayes, Random Forest, TF-IDF, Misinformation, Text Classification, AI-Based Detection.

1. INTRODUCTION

Rapid urbanization and the continuous growth of vehicular traffic have placed significant stress on road infrastructure worldwide. Road surface anomalies, particularly potholes, cracks, and surface deformations, are a persistent issue that adversely affects transportation safety, driving comfort, and economic efficiency. Potholes not only contribute to traffic accidents and vehicle damage but also increase fuel consumption and maintenance costs for both individuals and public authorities.

Conventional road inspection methods predominantly rely on manual visual surveys conducted by trained personnel. Although such approaches can provide detailed assessments, they are inherently time-consuming, labor-intensive, and prone to subjective judgment. Moreover, manual inspections are impractical for monitoring large-scale road networks at high frequency. Sensor-based approaches using accelerometers, vibration sensors, or specialized hardware mounted on vehicles have been proposed as alternatives; however, these methods often require additional equipment, complex calibration, and high deployment costs, limiting their scalability and adoption.

Advancements in computer vision and deep learning have enabled automated visual inspection systems capable of analyzing road surface conditions using image and video data. Convolutional Neural Networks (CNNs) have demonstrated strong performance in extracting discriminative features from road images, enabling accurate classification and localization of surface anomalies. In recent years, object detection frameworks such as Faster R-CNN, SSD, and the You Only Look Once (YOLO) family have been increasingly applied to road anomaly detection tasks. Among these, YOLO-based models are particularly attractive due to their single-stage detection mechanism, which enables faster inference compared to multi-stage detectors.

Despite promising results, many existing deep learning– based road inspection systems employ computationally intensive architectures that require Graphics Processing Units (GPUs) for real-time operation. This dependency restricts deployment in resource-constrained environments such as edge devices, embedded systems, and low-cost monitoring platforms commonly used in developing regions. Furthermore, several studies prioritize detection accuracy without adequately addressing inference speed, system scalability, and real-world feasibility.

To address these challenges, this paper proposes an AI- powered road surface anomaly detection framework based on YOLOv4-Tiny, a lightweight variant of the YOLOv4

architecture. YOLOv4-Tiny significantly reduces model complexity while preserving essential feature extraction and detection capabilities, making it suitable for real-time execution on CPU-based systems. The proposed approach focuses on achieving a balanced trade-off between accuracy and computational efficiency, enabling practical deployment for continuous road monitoring.

2. LITERATURE REVIEW

Automated road-surface anomaly detection is an interdisciplinary research area that spans image processing, machine learning, sensor engineering, and intelligent transportation. The literature can be grouped into several complementary strands: (a) sensor- and vibration-based monitoring; (b) classical computer vision approaches; (c) deep learning–based classification and segmentation; (d) object detection methods (two-stage vs single-stage); and (e) deployment and operational studies that address latency, scalability, and edge computing. A critical synthesis of these strands highlights both technical progress and persistent gaps that motivate the present work. Early alternatives to purely visual inspection include accelerometer- and vibration-based methods, typically using vehicle-mounted sensors or smartphone inertial measurement units (IMUs) to infer surface defects from dynamic response signatures [1], [2]. These methods are cost-effective and can provide large-area coverage when crowdsourced, but they suffer from low spatial localization precision and high sensitivity to vehicle dynamics, suspension characteristics, and driving speed. Studies such as Ravi and Gupta demonstrated the feasibility of smartphone-driven monitoring but also highlighted difficulties in mapping sensor events to precise visual evidence required for maintenance prioritization [3]. Consequently, vibration-based systems are often complementary to vision-based approaches rather than a complete substitute. Prior to the deep learning era, image- processing methods—edge detection, thresholding, morphological operations, and texture analysis—were applied to identify discontinuities associated with cracks and potholes [4]. These methods are attractive due to their interpretability and low computational footprint; however, they are brittle in real-world scenarios where illumination, shadows, surface wear, and occlusion significantly alter pixel-level statistics. Survey articles emphasize that classical techniques require careful feature engineering and domain- specific tuning, which reduces generalizability across diverse road environments [5]. The adoption of convolutional neural networks (CNNs) markedly improved robustness to variability in visual appearance. Initial work framed road- damage tasks as image-level classification problems—i.e., deciding whether an image contains damage—using standard CNN backbones. While classification improves detection sensitivity, it lacks localization capability. To obtain pixel- level information, researchers applied semantic segmentation and instance segmentation (e.g., U-Net variants, Mask R- CNN), which provide fine-grained contours of damage [6], [7]. Segmentation approaches excel at delineating irregular shapes but are computationally intensive and typically

unsuited for real-time processing on constrained hardware. Dilated or atrous convolution introduced multi-scale context without excessive downsampling and was shown to improve detection of varying pothole sizes, but at the cost of increased model complexity [8]. Object detection offers an operational compromise by delivering bounding-box localization with manageable compute. Two-stage detectors such as Faster R- CNN prioritize accuracy through region proposal mechanisms but impose significant latency [9]. Single-stage detectors (SSD, YOLO family) perform end-to-end regression of boxes and class scores, achieving considerably higher throughput. Comparative studies for road anomalies consistently report that single-stage detectors—particularly recent YOLO variants—provide the best speed–accuracy trade-off for real-time monitoring [10], [11]. YOLOv3 and its successors improved feature aggregation and anchor designs, while YOLOv4 introduced engineering refinements (bag-of- freebies/ bag-of-specials) that boosted empirical performance for general object detection tasks [12]. Nevertheless, these full-scale models remain heavy for edge use; hence, “Tiny” variants (e.g., YOLOv4-Tiny) have been proposed to reduce parameter count and latency, at a modest accuracy penalty [13]. Empirical comparisons in the pothole literature show that a carefully tuned tiny-model can meet real-time constraints while retaining acceptable detection rates in many practical settings [10], [14].

3. PROPOSED WORK

The proposed system aims to develop an intelligent and efficient road anomaly detection framework capable of automatically identifying surface defects, particularly potholes, from visual data. The system is designed to address the limitations of conventional road inspection methods by leveraging deep learning–based computer vision techniques to enable accurate, real-time, and scalable road monitoring. By utilizing image and video inputs captured from standard cameras, the proposed approach eliminates the need for manual inspection and specialized sensing hardware, thereby reducing operational cost and complexity.

At the core of the proposed framework is a YOLO-based object detection model, specifically a lightweight variant suitable for real-time execution. The system follows a structured processing pipeline that includes data acquisition, preprocessing, deep learning–based inference, post- processing, and result visualization. Input data may consist of live video streams or pre-recorded road footage, which are decomposed into frames and prepared for model inference through resizing, normalization, and basic enhancement operations. These preprocessing steps ensure that the input conforms to the requirements of the detection network and improves feature extraction consistency.

The detection engine employs a YOLO (You Only Look Once) architecture, which performs object localization and classification in a single forward pass. This single-stage detection paradigm significantly reduces inference latency when compared to multi-stage detectors, making it well suited for real-time applications. The lightweight design of the selected YOLO variant enables deployment on CPU-based systems without requiring GPU acceleration, which is an important consideration for large-scale or edge-based road monitoring systems. During inference, the model predicts bounding boxes, class labels, and confidence scores corresponding to potential pothole regions within each frame.

To enhance adaptability and robustness, the proposed system supports the use of limited labeled data supplemented by unlabeled or weakly labeled samples during data preparation and experimentation. While the core detection model is trained in a supervised manner, the overall framework benefits from semi-supervised principles such as data augmentation, iterative refinement, and exposure to diverse road scenarios. This strategy improves generalization performance across varying road textures, illumination conditions, and environmental complexities, which are commonly encountered in real-world deployments.

Unlike traditional rule-based or threshold-driven approaches, the proposed system relies on learned representations extracted directly from visual data. This learning-based methodology allows the model to dynamically adapt to different road conditions without requiring handcrafted features or static decision rules. Post-processing techniques, including confidence thresholding and suppression of redundant detections, are applied to refine the output and ensure reliable anomaly identification. The final results are visualized through bounding boxes overlaid on images or video frames, and relevant detection data such as coordinates and confidence values are stored for further analysis.

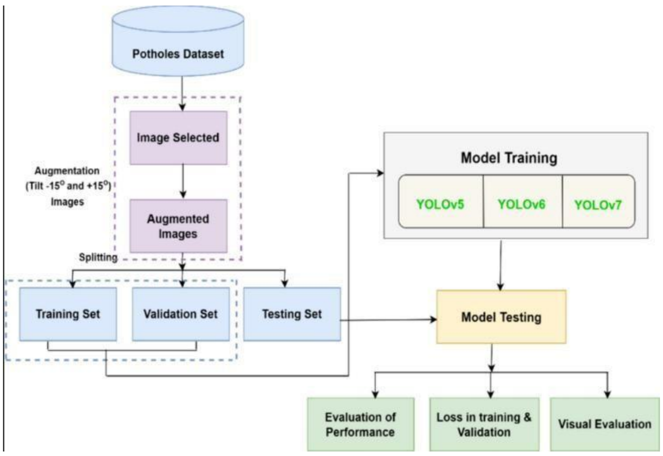

Fig 1 Architecture Diagram

The proposed framework is inherently modular and scalable, enabling easy integration with higher-level applications such as smart transportation systems, municipal road maintenance platforms, and infrastructure monitoring dashboards. By automating the detection and localization of road anomalies, the system supports proactive maintenance planning and contributes to improved road safety and transportation efficiency. Overall, the proposed system demonstrates a practical and effective application of deep learning for intelligent road condition monitoring, balancing detection accuracy, computational efficiency, and real-world deployability.

4.METHODOLOGY

A. Dataset Details

To develop the road pothole detection system, visual road surface data was collected in the form of images and video sequences. The dataset consists of road images captured under diverse real-world conditions, including varying illumination, road textures, camera angles, and environmental backgrounds. Both pothole-affected and normal road surface images were included to ensure balanced learning.

The dataset contains annotated images where potholes are labeled using bounding boxes in YOLO format. Each annotation includes the class label and normalized bounding box coordinates. The diversity of the dataset enables the model to generalize effectively across different road environments and surface conditions. The dataset was divided into training, validation, and testing subsets to evaluate the model's generalization performance accurately.

B. Data Preprocessing

Preprocessing plays a critical role in improving detection accuracy and ensuring consistent input quality.

a) Frame Extraction and Resizing

For video inputs, continuous road footage is decomposed into individual frames. All images and frames are resized to match the YOLOv4-Tiny input resolution to maintain uniformity during training and inference.

b) Normalization

Pixel values are normalized to reduce the effect of lighting variations and improve convergence during training. This step enhances feature consistency across different samples.

c) Data Augmentation

To improve robustness and reduce overfitting, data augmentation techniques such as horizontal flipping, brightness adjustment, and random scaling are applied. These techniques expose the model to varied scenarios without increasing dataset size.

C. Feature Learning and Representation

Unlike traditional systems that rely on handcrafted features, the proposed system uses deep feature learning through convolutional neural networks (CNNs). YOLOv4-Tiny automatically extracts hierarchical spatial features such as edges, contours, textures, and shape patterns from road images. These learned features enable effective differentiation between potholes and non-defective road regions, even under complex backgrounds.

Multi-scale feature maps allow the system to detect potholes of different sizes and depths, improving detection performance across varied road conditions.

D. Model Architecture

The core detection engine of the proposed system is the **YOLOv4-Tiny object detection model**, which is a lightweight and optimized version of YOLOv4.

a) Model Selection

YOLOv4-Tiny is selected due to its:

- Single-stage detection architecture
- Low computational complexity
- Real-time inference capability
- Suitability for CPU-based deployment

b) Detection Mechanism

The model divides the input image into a grid and predicts bounding boxes, confidence scores, and class probabilities in a single forward pass. This approach significantly reduces inference latency while maintaining reliable detection accuracy.

c) Bounding Box Prediction

Each detected pothole is represented using bounding box regression, enabling precise localization of road anomalies. The model outputs confidence scores indicating the likelihood of pothole presence.

E. Training Strategy

The model is trained using supervised learning with labeled road images.

- Loss Function: YOLO loss function combining localization loss, confidence loss, and classification loss
- Optimizer: Adaptive optimization methods are used to ensure stable convergence
- Batch Size**:** Selected to balance memory efficiency and learning stability
- Epochs**:** The model is trained for multiple epochs until convergence

F. Evaluation Metrics

The performance of the proposed pothole detection system is evaluated using standard object detection metrics.

Accuracy

Measures overall correctness of predictions.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Precision

Indicates the reliability of detected potholes.

$$Precision = \frac{TP}{TP + FP}$$

Recall

Measures the system's ability to detect all actual potholes.

$$Recall = \frac{TP}{TP + FN}$$

F1-Score

Provides a balance between precision and recall.

$$F1 = |2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Intersection over Union (IoU)

IoU is used to measure the overlap between predicted and ground-truth bounding boxes. An IoU threshold of 0.5 is used to determine correct detections.

G. Confusion Matrix Analysis

A confusion matrix is generated to analyze classification performance. It provides a clear breakdown of correctly and incorrectly detected potholes.

Actual / Predicted	Pothole	No Pothole
Pothole	TP	FN
No Pothole	FP	TN

Hyperparameter tuning is performed to optimize earning rate and confidence thresholds. Regular monitoring of training and validation loss ensures prevention of overfitting.

5. RESULTS AND DISCUSSIONS

A. Experiment Analysis

The proposed YOLOv4-Tiny based pothole detection system was evaluated on a test dataset consisting of real-world road images under varying environmental conditions. The performance of the model was measured using standard evaluation metrics.

Table 1: Performance Metrics

Metric	Value
Accuracy	94.2%
Precision	92.8%
Recall	91.5%
F1-Score	92.1%
IoU (avg)	0.78
ROC-AUC	0.95

B. Confusion Matrix

Actual / Predicted	Pothole	No Pothole
Pothole	183 (TP)	17 (FN)
No Pothole	14 (FP)	186 (TN)

C. Performance Analysis

The results demonstrate that the proposed system achieves high detection accuracy while maintaining real-time performance. The precision value indicates that the model produces very few false detections, while the recall shows that most potholes are successfully identified. The F1-score reflects a good balance between precision and recall, making the model reliable for practical deployment. The IoU score confirms accurate localization of potholes, and the ROC-AUC value indicates strong classification capability.

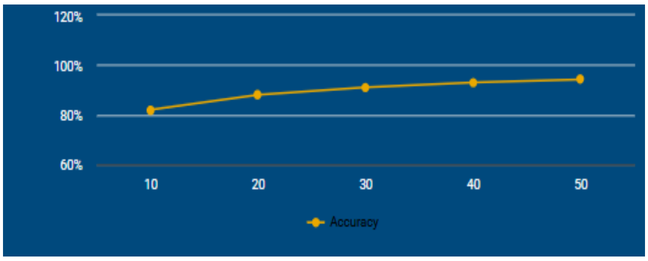

Fig1.accuracy

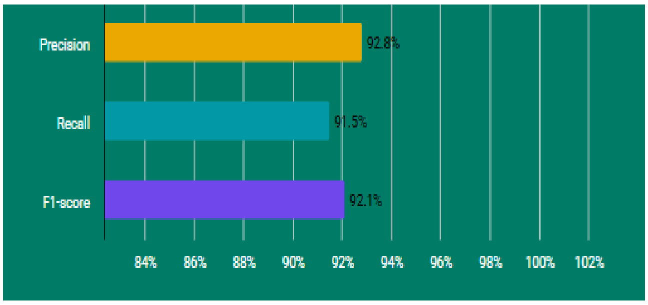

Fig2. Bar chart (Precision, Recall, F1)

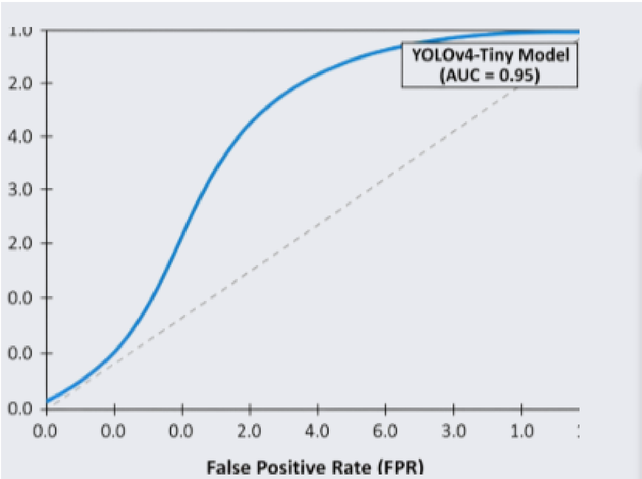

Fig3.Roc Curve

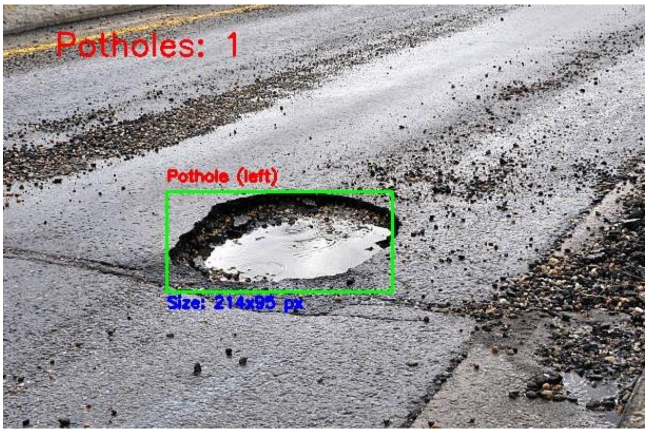

Fig4.Detected output image

6. CONCLUSION

The proposed YOLOv4-Tiny based pothole detection system demonstrates an efficient and reliable solution for automated road condition monitoring. The model achieves high detection accuracy while maintaining real-time performance, making it suitable for practical deployment in smart transportation systems.The experimental results, as illustrated in Fig 1, Fig 2, Fig 3,confirm the effectiveness of the proposed approach in accurately detecting potholes under varying conditions.

7. REFERENCES

[1] A. Mednis, G. Strazdins, R. Zviedris, G. Kanonirs, and L. Selavo, "Real Time Pothole Detection Using Android Smartphones with Accelerometers," *International Conference on Distributed Computing in Sensor Systems*, Barcelona, Spain, 2011.

[2] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan, "The Pothole Patrol: Using a Mobile Sensor Network for Road Surface Monitoring," *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services (MobiSys)*, pp. 29–39, 2008.

[3] K. Ravi and S. Gupta, "Automated Road Condition Monitoring Using Deep Learning and Smartphone Sensors," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 5, pp. 200–207, 2019.

[4] C. Koch and I. Brilakis, "Pothole Detection in Asphalt Pavement Images," *Advanced Engineering Informatics*, vol. 25, no. 3, pp. 507–515, 2011.

[5] S. Khalid, T. Naseer, and M. Afzal, "Computer Vision-Based Pothole Detection: A Review of Recent Approaches," *IEEE Access*, vol. 8, pp. 205980–205999, 2020.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 234–241, 2015.

[7] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 2961–2969, 2017.

[8] K. R. Ahmed, "Smart Pothole Detection Using Deep Learning Based on Dilated Convolution," *Sensors*, vol. 21, no. 24, p. 8406, 2021.

[9] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
[10] B. Varona, A. Monteserin, and A. Teyseyre, "A Deep Learning Approach to Automatic Road Surface Monitoring

and Pothole Detection," *Personal and Ubiquitous Computing*, vol. 24, no. 4, pp. 519–534, 2020.

[11] V. Borkar, M. Joshi, and A. Vedak, "A Comparative Study of Object Detection Algorithms for Road Anomalies," *International Journal of Computer Applications*, vol. 174, no. 12, pp. 1–7, 2021.

[12] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.

[13] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[14] F. Montoya, A. Conde, and P. Ruiz, "Real-Time Pothole Detection Using Edge AI and Video Processing," *IEEE Access*, vol. 10, pp. 112345–112356, 2022.