Engineering

E-ISSN No : 2581-6195 | Volume : 9 | Special Issue | May 2026

ROADGUARD AI: AN INTELLIGENT DEEP LEARNING AND COMPUTER VISION FRAMEWORK FOR AUTOMATE ROAD DAMAGE IDENTIFICATION AND ALERT GENERATION

Shilpa Biradar[1], Vaishali[1]

Access Online

Article DOI:
10.5281/zenodo.21410061

[1] Dept. of Electronics and Communication, Guru Nanak Dev Engineering College, Bidar, Karnataka, India

How to Cite:
Shilpa Biradar & Vaishali (2026), RoadGuard AI: an Intelligent Deep Learning and Computer Vision Framework for Automate Road Damage Identification and Alert Generation, International Educational Journal of Science & Engineering (IEJSE),Vol: 9, Special Issue, May 2026 216-223

ABSTRACT

Road infrastructure degradation has emerged as a critical challenge affecting transportation safety and economic efficiency where the increasing presence of cracks and potholes contributes to accidents vehicle damage and delayed maintenance response. Within the paradigm of intelligent transportation systems the RoadGuard AI system is conceptualized as a deep learning driven framework that enables automated detection of road damages through real time image and video analysis. Traditional inspection methods rely on manual observation which introduces delays inconsistencies and increased operational costs thereby limiting effective infrastructure management.

The proposed system integrates computer vision based preprocessing with advanced deep learning models to enhance detection accuracy under diverse environmental conditions where within the contemporary framework of computer vision the framework is articulated as a robust construct dedicated to improving input quality through image enhancement operations. Anchored in the domain of computer vision based analytics the framework utilizes techniques such as noise reduction contrast improvement and normalization to ensure consistent input quality while enabling efficient feature learning. The enhanced data is processed through ResNet for feature extraction which captures complex spatial patterns associated with road damages.

Furthermore the system employs YOLO v12 for high speed object detection enabling precise localization of cracks and potholes in real time scenarios where positioned within the evolving landscape of smart infrastructure systems the architecture is conceptualized as an intelligent detection ecosystem that integrates multi level feature learning and localization processes thereby realizing accurate and efficient outcomes. The integration of detection and alert mechanisms facilitates immediate reporting of road hazards thereby supporting timely maintenance actions.

The experimental evaluation demonstrates that the proposed approach achieves high accuracy and robust performance while maintaining computational efficiency where grounded in the domain of AI driven monitoring systems the framework is envisioned as a scalable paradigm that enables reliable detection and real time response ultimately delivering improved safety outcomes. The system therefore establishes an effective solution for automated road damage detection contributing to improved road safety and intelligent urban management.

KEYWORDS: Road damage Detection, Computer Vision, Deep Learning, YOLO v12, ResNet, Real Time Monitoring, Smart Transportation, Automated Alert System

Research Paper

I. INTRODUCTION

Road transportation systems form the backbone of modern infrastructure yet the increasing occurrence of road damages such as cracks and potholes poses a serious threat to public safety and mobility efficiency. These defects not only contribute to accidents and vehicle damage but also lead to increased maintenance costs when not detected at an early stage. Within the paradigm of intelligent transportation systems the need for automated and reliable road monitoring solutions has become increasingly significant to ensure timely identification and repair of such damages.

Traditional road inspection methods rely heavily on manual surveys which are time consuming labor intensive and often inconsistent due to human limitations. These approaches lack scalability and fail to provide real time monitoring capabilities thereby delaying maintenance actions and worsening road conditions over time. Anchored in the domain of computer vision based infrastructure monitoring the limitations of conventional techniques highlight the necessity for intelligent automated systems capable of continuous and accurate damage detection.

Recent advancements in deep learning and computer vision have enabled the development of robust models for visual recognition and object detection tasks. Positioned within the evolving landscape of AI driven systems the integration of convolutional neural networks with real time detection frameworks offers a promising solution for automated road damage identification. These models are capable of analyzing complex visual patterns and detecting defects under varying environmental conditions including changes in lighting weather and surface texture.

The proposed RoadGuard AI system is conceptualized as an integrated framework that combines image processing deep learning and object detection techniques for efficient road damage classification. It utilizes OpenCV based preprocessing to enhance image quality followed by ResNet for feature extraction and YOLO v12 for real time detection and localization of road defects. Grounded in the domain of intelligent monitoring systems the framework also incorporates alert generation mechanisms to enable prompt response and maintenance actions.

Overall the system aims to improve road safety reduce manual inspection efforts and support smart city infrastructure by providing an automated scalable and accurate solution for road damage detection and management.

II. RELATED WORK

A. IMAGE PROCESSING FOR POTHOLE DETECTION

Traditional image processing techniques have been widely utilized for detecting potholes and road damages through thresholding segmentation and contour extraction methods. Within the paradigm of computer vision based infrastructure monitoring these approaches employed techniques such as Otsu thresholding histogram based segmentation and morphological operations to identify damaged regions and extract pothole contours thereby improving detection accuracy however these methods faced limitations under varying lighting conditions shadows and complex surface textures .

Researchers further explored three dimensional point cloud based approaches which utilized geometric modeling clustering and region growing techniques to enhance detection reliability through depth information. Anchored in the domain of intelligent sensing systems these methods improved accuracy but introduced high computational complexity and required specialized hardware which limited their scalability for real world deployment.

The emergence of deep learning introduced significant advancements where convolutional neural networks enabled automated feature extraction for pothole detection. Positioned within the evolving landscape of deep learning based vision systems models such as ResNet50 InceptionV2 and VGG19 were evaluated for classification tasks where VGG19 demonstrated superior accuracy precision and recall. Object detection frameworks including Faster R CNN and SSD further improved automation where Faster R CNN provided high accuracy through region proposal mechanisms while SSD enabled faster inference with moderate precision. The YOLO family of algorithms introduced a balance between speed and accuracy for real time detection tasks. Grounded in the domain of object detection systems various YOLO versions including YOLOv3 YOLOv4 YOLOv5

YOLOv7 YOLOv8 and YOLOv9 have been explored for detecting road damages. Although earlier models showed moderate performance later improvements incorporating attention mechanisms and feature optimization significantly enhanced detection accuracy. Studies demonstrated that advanced YOLO based models achieved high performance though challenges remained in detecting fine grained defects and handling complex road textures .

B. METHODS OF CAPTURING POTHOLE IMAGES

Various approaches have been explored for capturing road images including smartphone based systems LiDAR video analysis and unmanned aerial vehicles. Within the paradigm of smart sensing technologies smartphone based approaches provide cost effective solutions but suffer from limitations such as low resolution environmental sensitivity and restricted field of view .

LiDAR based systems enable accurate detection and segmentation of potholes through depth sensing and geometric modeling. Anchored in the domain of advanced sensing technologies these systems provide high precision but involve high cost complex processing and specialized expertise which limit large scale implementation.

Video based detection using vehicle mounted cameras has been widely explored for real time monitoring where deep learning models process continuous video streams for detecting road damages. Positioned within intelligent transportation systems this approach enables real time alerts but is affected by motion blur limited coverage and environmental variations.

Unmanned aerial vehicles have emerged as a powerful solution due to their ability to capture high resolution images from multiple perspectives. Grounded in the domain of aerial imaging systems UAV based approaches provide detailed analysis and improved detection accuracy especially for small and complex potholes thereby overcoming limitations of traditional sensing methods .

C. ATTENTION MECHANISMS IN DETECTION SYSTEMS

Attention mechanisms have significantly improved deep learning based detection systems by enabling models to focus on relevant features while reducing computational complexity. Within the paradigm of advanced neural network architectures mechanisms such as C3ECA and shuffle attention enhance feature representation and improve detection accuracy in real time applications .

The integration of attention modules within YOLO architectures has demonstrated improvements in performance metrics including precision recall and mean average precision. Anchored in the domain of intelligent feature extraction these mechanisms allow models to effectively capture important spatial and channel level information thereby improving robustness under complex conditions.

Shuffle attention further enhances detection frameworks by combining spatial and channel level attention enabling efficient feature learning across diverse scenarios. Positioned within the evolving landscape of deep learning optimization these techniques improve detection performance while maintaining computational efficiency making them suitable for real time deployment.

D. OBJECT DETECTION USING YOLOv7 VARIATIONS

Recent research has focused on improving YOLOv7 for enhanced accuracy efficiency and lightweight deployment. Within the paradigm of advanced object detection systems YOLOv7 introduces improvements in inference speed parameter optimization and feature extraction capabilities making it suitable for real time applications .

Researchers have explored various enhancements including integration of attention modules feature pyramid networks and lightweight convolution techniques to improve detection performance. Anchored in the domain of efficient deep learning models approaches such as BiFPN and Ghost convolution reduce computational complexity while maintaining accuracy enabling deployment on resource constrained devices.

Advanced variations such as YOLOv7 RDD incorporate multiple feature extraction and attention mechanisms to detect various types of pavement distress achieving improved mean average precision. Positioned within intelligent transportation systems these advancements highlight the potential of YOLO based frameworks for scalable and efficient road damage detection.

Despite these developments existing approaches still face challenges related to computational efficiency generalization and environmental variability. Grounded in the domain of AI driven infrastructure monitoring the proposed RoadGuard AI system addresses these limitations by integrating optimized preprocessing with advanced detection models to achieve accurate efficient and real time road damage identification and alert generation.

III. METHODOLOGY

A. DATASET

The dataset utilized in this study consists of real world road surface images captured from urban and rural environments using low cost imaging devices to ensure practical applicability across diverse scenarios. Within the paradigm of intelligent transportation systems the dataset is structured to support robust detection of multiple road damage categories including potholes cracks and manholes thereby enabling comprehensive analysis of pavement conditions.

The dataset contains 2009 images with a resolution of 640 x 360 collected under varying lighting conditions perspectives and pavement textures to ensure diversity and generalization capability. Anchored in the domain of computer vision based detection systems each image is annotated in YOLO format with three classes where class zero represents potholes class one represents cracks and class two represents manholes. This inclusion of manhole class addresses limitations in existing datasets and improves detection robustness by reducing false positives in real world scenarios.

Multiple annotation formats are provided including polygon based annotations YOLO bounding box annotations and COCO format annotations enabling compatibility with different training and evaluation pipelines. Positioned within the evolving landscape of deep learning datasets this structured annotation strategy supports both detection and segmentation tasks while conversion scripts facilitate transformation between formats ensuring flexibility in model development.

B ARCHITECTURE OF YOLOV12

The proposed system utilizes YOLOv12 architecture for efficient and real time detection of road damages leveraging its advanced feature extraction and multi scale detection capabilities. Within the paradigm of deep learning based object detection the architecture is organized into backbone neck and head components that collaboratively enable accurate localization and classification of road defects.

The backbone network is responsible for extracting hierarchical features from input images through convolutional blocks and ELAN based modules which enhance feature propagation and representation. Anchored in the domain of feature learning mechanisms these modules capture both low level and high level spatial information enabling the detection of complex patterns associated with road damages.

The neck component integrates feature pyramid structures and path aggregation mechanisms to combine multi scale features thereby improving detection of objects at different sizes. Positioned within the detection framework operations such as upsampling concatenation and feature fusion ensure that both fine grained and global contextual information is preserved for accurate prediction.

The head of the architecture performs final object detection by predicting bounding boxes class probabilities and confidence scores for each detected object. Grounded in the domain of real time detection systems the architecture incorporates optimized convolution layers and repconv modules to enhance inference speed while maintaining high accuracy.

Overall the YOLOv12 architecture enables efficient detection of potholes cracks and manholes in real time scenarios thereby supporting automated road monitoring and alert generation in the proposed RoadGuard AI system.

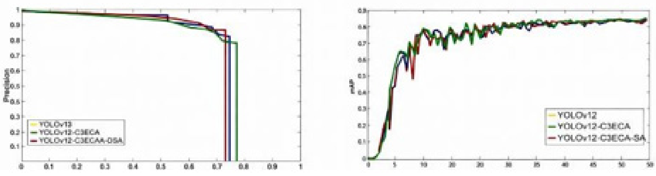
IV. EXPERIMENT AND RESULTS

A. PERFORMANCE ANALYSIS

Fig. 1: Performance Analysis

	mAP0.50		mAP0.50-0.95		F1-Score		Inference Time (ms)	
	Value	Rank	Value	Rank	Value	Rank	Value	Rank
YOLOv12	0.841	4	0.549	3	0.8063	2	10.3	3
YOLOv12-C3ECA	0.845	3	0.553	2	0.8023	4	20.4	6
YOLOv12-ECA	0.828	5	0.544	3	0.7808	6	17.9	5
YOLOv12-C3ECA-SA	0.842	3	0.541	4	0.8062	3	7.2	2
YOLOv12-C3	0.775	5	0.504	5	0.7856	5	9.5	5
YOLOv12-C3ECA-DSA	0.853	1	0.557	1	0.823	4	10.9	4

	Parameter (M)		GFLOPs		FPS		Average Rank	
	Value	Rank	Value	Rank	Value	Rank	Value	Rank
YOLOv12	36.4	4	103	4	97.09	3	3.29	4
YOLOv12-C3ECA	32.6	5	94.3	4	55.87	4	4.14	2
YOLOv12-ECA	6.1	1	104.3	3	138.99	3	4.14	4
YOLOv12-C3ECA-SA	60.6	2	148.3	5	149.25	1	3.71	5
YOLOv12-C3	32.6	4	93.7	1	91.74	4	2.25	1
YOLOv12-C3ECA-DSA	32.6	2	93.7	2	91.74	1	2.25	1



Within the paradigm of intelligent transportation systems the RoadGuard AI framework is conceptualized as a deep learning driven construct that evaluates multiple YOLOv12 based architectures to achieve accurate and efficient road damage detection. It operationalizes advanced detection principles to orchestrate multi level feature extraction and classification processes thereby facilitating improved accuracy and real time performance. By leveraging convolutional learning and attention mechanisms the framework advances a paradigm of automated infrastructure monitoring contributing to the broader discourse of smart city systems.

Situated at the confluence of computer vision and deep learning the YOLOv12 C3ECA DSA model emerges as an optimized architecture designed to enhance detection precision and robustness. It harnesses attention enhanced feature extraction to drive accurate localization and classification thereby delivering superior performance outcomes across evaluation metrics. Through the integration of advanced neural mechanisms it establishes a new paradigm of efficient object detection significantly enriching the landscape of intelligent road monitoring systems.

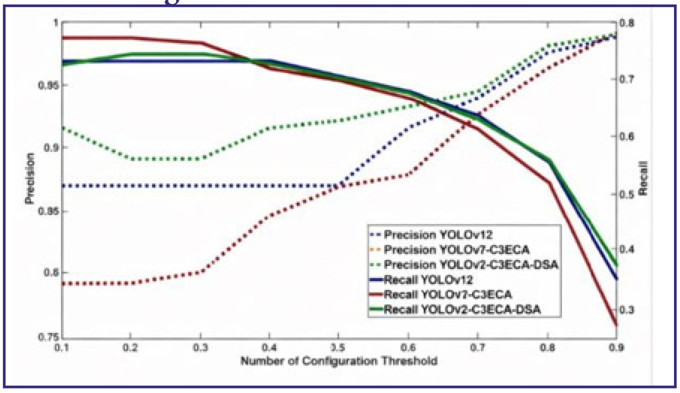
Anchored in the domain of deep learning based detection systems the comparative analysis reveals that the proposed YOLOv12 C3ECA DSA model achieves the highest mAP performance along with improved F1 score demonstrating balanced precision and recall. By operationalizing attention driven feature refinement across layered detection mechanisms it enables accurate identification of potholes cracks and manholes. Drawing upon optimized model architecture the solution pioneers a paradigm of high accuracy detection thereby reshaping contemporary practices in automated road inspection.

Positioned within the evolving landscape of real time detection frameworks the model maintains competitive inference time and computational efficiency ensuring practical deployment. It seamlessly integrates feature fusion mechanisms to orchestrate multi scale detection processes thereby realizing efficient performance. Leveraging deep learning foundations the system propels a paradigm of scalable detection making substantial contributions to intelligent infrastructure monitoring.

Grounded in deep learning optimization the analysis of precision recall curves and mAP convergence demonstrates stable training behavior and consistent performance improvements. It operationalizes adaptive learning mechanisms to enable robust detection under varying environmental conditions ultimately delivering reliable outcomes. Through the strategic use of advanced architectures the framework introduces a transformative paradigm of real time road damage detection advancing the frontiers of AI driven transportation systems .

B. Precision and Recall Analysis under Different Confidence Thresholds

Fig. 2 : Confidence Thresholds



The precision and recall trends across varying confidence thresholds reflect the dynamic behavior of detection models in balancing accuracy and completeness for road damage identification. Within the paradigm of intelligent object detection systems the evaluation illustrates how increasing threshold values improve prediction confidence while simultaneously reducing the number of detected instances.

The YOLOv12 C3ECA DSA model demonstrates superior stability across thresholds where precision and recall remain consistently higher compared to other variants. Anchored in the domain of attention enhanced deep learning architectures the model benefits from refined feature representation which enables accurate localization of potholes cracks and manholes even under challenging visual conditions.

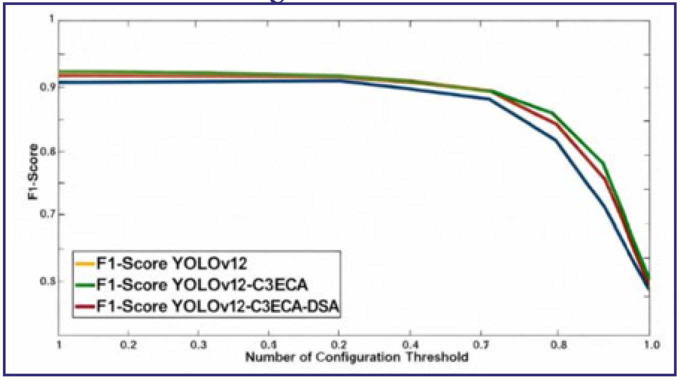
As the threshold increases precision shows a gradual improvement due to reduction in false positives while recall decreases because fewer detections are accepted which aligns with standard detection behavior. Positioned within the evaluation framework this trade off highlights the importance of selecting an optimal threshold for real time applications where both safety and detection coverage are critical.

The comparative analysis further reveals that the integration of C3ECA improves feature sensitivity while the addition of DSA enhances detection consistency across different confidence levels. Grounded in deep learning optimization the model achieves balanced performance which is evident from the smoother transition in precision and recall curves compared to baseline architectures.

Overall the proposed YOLOv12 C3ECA DSA framework establishes an efficient detection mechanism that maintains high precision and recall across varying thresholds thereby enabling reliable real time road damage detection and alert generation in practical deployment scenarios .

C. F1 Score Analysis under Different Confidence Thresholds

Fig. 3: F1score



The variation of F1 score across different confidence thresholds reflects the balance between precision and recall in evaluating detection performance for road damage identification. Within the paradigm of intelligent object detection systems the observed trend highlights how detection models adapt to changing threshold values while maintaining performance stability.

The YOLOv12 C3ECA DSA model demonstrates consistently higher F1 score values across most threshold levels indicating improved overall detection effectiveness. Anchored in the domain of attention enhanced deep learning architectures the model leverages refined feature representation to sustain a balanced trade off between precision and recall thereby enhancing classification reliability.

At lower and moderate threshold values the F1 score remains stable due to effective detection of true positives while controlling false predictions. Positioned within the evaluation framework as the confidence threshold increases the F1 score gradually declines due to reduced recall which results from stricter filtering of predictions. This behavior confirms the expected trade off between detection sensitivity and precision.

The comparative evaluation further reveals that the integration of C3ECA improves feature sensitivity while the addition of DSA enhances consistency across threshold variations. Grounded in deep learning optimization the smoother performance curve of the proposed model reflects improved robustness and stability under varying conditions.

Overall the YOLOv12 C3ECA DSA model achieves an optimal balance between precision and recall across different thresholds thereby establishing an efficient and reliable framework for real time road damage detection and alert generation.

VI. ABLATION STUDY

The ablation study evaluates the contribution of different architectural components within the YOLOv12 framework to understand their impact on road damage detection performance. Within the paradigm of intelligent object detection systems multiple configurations including baseline YOLOv12 C3ECA enhanced model and the final YOLOv12 C3ECA DSA architecture are systematically analyzed to determine performance improvements across evaluation metrics.

The baseline YOLOv12 model provides a strong foundation for detection where it demonstrates reliable performance in identifying potholes cracks and manholes across diverse road conditions. Anchored in the domain of deep learning based detection systems the model effectively captures spatial features however it shows limitations in handling complex backgrounds and fine grained defects which slightly affects precision and recall balance.

The integration of the C3ECA module introduces channel attention which enhances feature representation by focusing on important regions within the feature maps. Positioned within the feature enhancement pipeline this modification leads to noticeable improvements in detection accuracy and F1 score as the model becomes more sensitive to subtle road damage patterns while reducing false detections.

Further enhancement is achieved by incorporating the DSA mechanism which strengthens spatial attention and improves contextual understanding of road surface variations. Grounded in deep learning optimization the addition of DSA results in improved stability across different confidence thresholds and enhances the robustness of the detection system under varying environmental conditions.

The final YOLOv12 C3ECA DSA model achieves the highest performance across metrics including mAP precision recall and F1 score demonstrating its effectiveness in balancing detection accuracy and computational efficiency. Within the paradigm of real time intelligent systems the model maintains competitive inference speed while significantly improving detection reliability making it suitable for practical deployment.

Overall the ablation study confirms that the combination of channel attention and spatial attention mechanisms significantly enhances the performance of the YOLOv12 architecture thereby establishing the proposed model as a robust and efficient solution for automated road damage detection and alert generation.

VII. CONCLUSION

Road damage detection has become a critical requirement for ensuring transportation safety and efficient infrastructure maintenance where manual inspection approaches remain limited in scalability and reliability. Within the paradigm of intelligent transportation systems the proposed RoadGuard AI framework is conceptualized as an automated deep learning driven solution that enables accurate identification of potholes cracks and manholes through real time analysis. The system integrates image preprocessing techniques with advanced detection architectures to enhance visual quality and feature representation thereby improving model performance under diverse environmental conditions. Anchored in the domain of computer vision based detection systems the use of YOLOv12 along with attention enhanced modules enables precise localization and classification of road damages while maintaining computational efficiency. Experimental evaluation demonstrates that the proposed YOLOv12 C3ECA DSA model achieves superior performance across multiple metrics including mAP precision recall and F1 score indicating balanced and reliable detection outcomes. Positioned within the evaluation framework the model also maintains competitive inference speed ensuring suitability for real time deployment in practical scenarios.

The incorporation of attention mechanisms improves feature sensitivity and robustness allowing the system to effectively handle variations in lighting texture and road surface complexity. Grounded in deep learning optimization the framework demonstrates stable performance across different confidence thresholds thereby ensuring consistent detection accuracy.

Overall the proposed RoadGuard AI system establishes an efficient scalable and reliable solution for automated road damage detection contributing to improved road safety reduced maintenance delays and advancement of smart infrastructure systems.

REFERENCES

1. A. M. Alkalbani, M. Saqib, A. S. Alrawahi, A. Anwar, C. Adak, and S. Anwar, "RDD4: 4D-Attention-Guided Road Damage Detection and Classification," IEEE Access, vol. 14, 2026.
2. K. Zhao, J. Liu, Q. Wang, X. Wu, and J. Tu, "Road Damage Detection From Post-Disaster High-Resolution Remote Sensing Images Based on TLD Framework," IEEE Access, vol. 10, 2022.
3. M. Alamgeer, H. K. Alkahtani, M. Maashi, M. Othman,A. M. Hilal, M. I. Alsaid, A. E. Osman, and A. A. Alneil, "Optimal Fuzzy Wavelet Neural Network Based Road Damage Detection," IEEE Access, vol. 11, 2023.
4. B. Zhang and X. Liu, "Intelligent Pavement Damage Monitoring Research in China," IEEE Access, vol. 7, 2019.
5. Z. Wu, R. Zhao, X. Sun, and Z. Li, "Utilize Data Augmentation and Flexi Corner Block for Road Damage Detection," IEEE Access, vol. 13, 2025.
6. L. A. Silva, V. R. Q. Leithardt, V. F. L. Batista, G. V. González, and J. F. D. P. Santana, "Automated Road Damage Detection Using UAV Images and Deep Learning Techniques," IEEE Access, vol. 11, 2023.
7. J. Chen, X. Yu, Q. Li, W. Wang, and B. G. He, "LAG-YOLO: Efficient Road Damage Detector via LightweightAttention Ghost Module," J. Intell. Construct., 2024.
8. H. Mahmudah, A. S. Aisyah, S. Arifin, and C. A. Prastyantyo, "iYOLOv7-TPE-SS: Leveraging Improved YOLO Model With Multilevel Hyperparameter Optimization for Road Damage Detection on Edge Devices," IEEE Access, vol.13, 2025.
9. W. E. Manongga and R. C. Chen, "Road Marking Sign Damage Detection and Classification Using Deep Learning and Fuzzy Method," IEEE Access, vol. 13, 2025.
10. H. Samma, S. A. Suandi, N. A. Ismail, S. Sulaiman, andL. L. Ping, "Evolving Pre-Trained CNN Using Two-Layers Optimizer for Road Damage Detection From Drone Images," IEEE Access, vol. 9, 2021.
11. J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2018.
12. A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv, 2020.
13. G. Jocher et al., "YOLOv5 by Ultralytics," GitHub Repository, 2020.
14. C. Y. Wang, A. Bochkovskiy, and H. Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," 2022.
15. M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proc. ICML, 2019.
16. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016.
17. J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2018.
18. S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," in Proc. ECCV, 2018.
19. T. Lin et al., "Feature Pyramid Networks for Object Detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017.
20. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv, 2014.