Hybrid Approach for Early Road Defect Detection: Integrating Edge Detection with Attention-Enhanced MobileNetV3 for Superior Classification

Ayoub Oulahyane[1]*, Mohcine Kodad[2], El Houcine Addou[3], Sofia Ourarhi[4], Hajar Chafik[5]
MATSI Laboratory, ESTO, Mohammed First University, Oujda, 60000, Morocco[1, 2]
LANO Laboratory, FSO-ESTO, Mohammed First University, Oujda, 60000, Morocco[3]
2GPMH Research Lab, FSO, Mohammed First University, Oujda, 60000, Morocco[4]
LCELHN Laboratory, FLSHO, Mohammed First University, Oujda, 60000, Morocco[5]

***Abstract*—The early detection of road defects is critical for maintaining infrastructure quality and ensuring public safety. This research presents a hybrid approach that combines edge detection techniques with an enhanced deep learning model for efficient and accurate road defect classification. The process begins with edge detection to highlight structural irregularities, such as cracks and potholes, by emphasizing critical features in road surface images. These pre-processed images are then fed into a classification model based on MobileNetV3, augmented with an attention mechanism to improve feature weighting and model focus on defect-prone regions. The proposed system was evaluated on a Crack500 dataset of road surface images, achieving a classification accuracy of 96.2%. This demonstrates significant improvement compared to baseline models without edge detection or attention enhancements. The edge detection stage efficiently reduces noise, while the attention-augmented MobileNetV3 ensures robust feature discrimination, making the approach suitable for real-time and resource-constrained deployment scenarios. This study highlights the effectiveness of combining classical image processing with advanced neural network techniques. The proposed system has the potential to optimize road maintenance workflows, operational costs, and improve road safety by enabling early and precise defect identification.**

***Keywords*—Road defect detection; edge detection; attention mechanism; MobileNetV3**

I. Introduction

Road infrastructure plays a fundamental role in economic development, public safety, and the overall quality of life [1], [2]. Properly maintained roads ensure the smooth flow of goods, services, and people, contributing to societal efficiency and growth. However, road defects such as cracks, potholes, and surface deformities are inevitable due to wear and tear, extreme weather conditions, and high traffic loads [3], [4]. These defects, if not detected and addressed promptly, can escalate, leading to costly repairs, increased accident risks, and disruptions to transportation systems. Consequently, early and accurate detection of road defects is critical to minimize maintenance costs and enhance road safety [5].

Traditionally, road inspections have relied on manual methods or simple imaging systems, which often fall short in terms of accuracy, scalability, and efficiency [6]. Manual inspections are labor-intensive, subjective, and unsuitable for large-scale applications, while basic imaging systems struggle with environmental challenges such as poor lighting, shadow interference, and complex road textures [7]. This has necessitated the development of automated approaches that are not only accurate but also adaptable to real-world conditions [8].

Despite significant progress in computer vision and deep learning, current automated systems still struggle to accurately detect small or subtle defects in complex and dynamic environments. There remains a clear need to develop lightweight and effective models that can maintain high detection accuracy without imposing high computational demands, ensuring their applicability in real-world settings. This study addresses the challenge by investigating how preprocessing techniques, particularly edge detection, can be combined with deep learning models to enhance road defect detection performance. It also explores whether incorporating an attention mechanism into a lightweight model such as MobileNetV3 can improve sensitivity to defect-specific features and overall classification accuracy. Furthermore, the research examines the impact of integrating traditional image processing with deep learning to develop a more robust and reliable detection framework.

The objectives of this work are to design and implement a hybrid methodology that combines edge detection with an attention-augmented MobileNetV3 model, to preprocess road surface images by emphasizing defect-relevant features while minimizing background noise, and to evaluate the proposed framework's performance against existing methods using benchmark datasets.

In this study, Fig. 1 provides a visual representation of various types of road surface cracks, which are critical indicators of pavement deterioration. These include common defects such as longitudinal, transverse, and block cracks, among others. The figure serves to underscore the diverse and complex nature of road defects, which necessitate precise and efficient detection methodologies.

By examining these crack types, the paper establishes the motivation for adopting a hybrid approach that combines edge detection techniques with an attention-augmented MobileNetV3 model [9]. This innovative framework aims to enhance the accuracy and robustness of road defect classification, ensuring timely maintenance and improved infrastructure resilience.

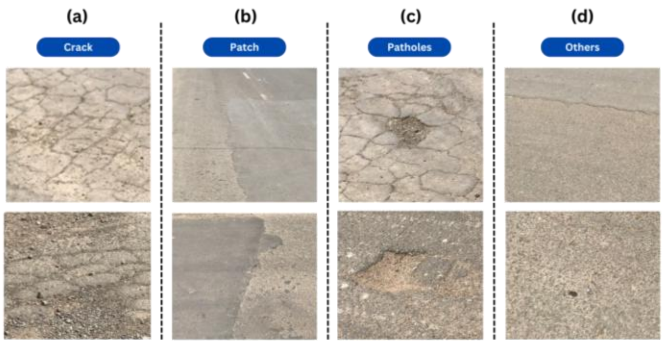

Fig. 1. Illustration of various types of road surface cracks, highlighting the importance of early detection for maintenance and safety.

This paper proposes a hybrid methodology for road defect detection that leverages the strengths of edge detection and a deep learning model enhanced with an attention mechanism. The process begins with edge detection to preprocess road surface images, emphasizing defect-relevant features while reducing background noise and redundant information. These refined images are then input into a MobileNetV3-based neural network, which is augmented with an attention mechanism to improve its focus on critical features. The attention mechanism dynamically prioritizes defect-specific regions in the feature maps, enhancing the model's ability to detect subtle defects such as hairline cracks or small potholes.

The remainder of this paper is organised as follows. Section II provides a detailed overview of related work in road defect detection and highlights the gaps that this study aims to address. Section III describes the proposed methodology, including the integration of edge detection and the MobileNetV3 architecture with attention mechanisms. Section IV discusses the experimental setup, dataset, and performance metrics. It also presents and analyzes the results, while Section V concludes the paper with insights, limitations, and potential directions for future research.

This work aims to contribute to the growing field of automated infrastructure monitoring by providing a practical, scalable, and efficient solution for early road defect detection, which can significantly impact road maintenance strategies and public safety worldwide.

II. Related Work

Road defect detecting has garnered considerable attention in recent years due to its significance in maintaining infrastructure safety and functionality. A variety of approaches, ranging from traditional image processing methods to cutting-edge deep learning techniques, have been proposed to address the challenges posed by road defect identification in real-world scenarios.

A. *Maintaining the Integrity of Specifications*

Traditional computer vision techniques, such as edge detection algorithms, have been widely used for identifying structural discontinuities in road surfaces [10]. Methods like Sobel [11], Canny [12], and Laplacian filters [13] efficiently highlight features such as cracks and potholes by emphasizing abrupt changes in pixel intensity. While computationally efficient, these techniques often lack robustness in noisy environments or under varying lighting conditions. Nonetheless, edge detection remains a valuable preprocessing tool, as it helps reduce noise and isolate defect-prone regions, providing a foundation for more advanced classification models.

B. *Deep Learning in Road Defect Detection*

With the advent of deep learning, researchers have shifted focus toward convolutional neural networks (CNNs) for automated defect detection. For instance, the YOLO family of object detection models has shown remarkable capabilities in real-time applications. The RDD-YOLOv5 model integrates a self-attention mechanism to enhance the precision of crack detection, achieving a high mAP of 91.48% [14]. Similarly, BL-YOLOv8, which incorporates BiFPN and LSK-attention, optimizes both accuracy and computational efficiency. By reducing model size and parameter volume, it becomes suitable for deployment in resource-constrained settings [15].

C. *Attention Mechanisms*

Attention mechanisms have emerged as a powerful enhancement in deep learning models, enabling them to focus on the most relevant regions of an image. Studies combining attention mechanisms with ensemble learning methods have demonstrated significant improvements in road defect detection. For example, a multi-depth attention mechanism has been successfully employed to prioritize defect-specific features, achieving superior performance across diverse datasets [16].

These mechanisms are particularly effective when integrated into light-weight architectures, such as MobileNet, to improve performance without compromising efficiency.

D. *Transfer Learning and Lightweight Models*

Transfer learning has proven to be indispensable for road defect detection, especially in scenarios with limited annotated data. By fine-tuning pre-trained models like MobileNet, researchers can leverage knowledge learned from large-scale datasets to improve defect detection accuracy [17]. MobileNetV3, in particular, has gained attention for its lightweight architecture, making it ideal for real-time applications on mobile and edge devices[18]. Incorporating attention mechanisms into MobileNetV3 further enhances its ability to classify defects, even in challenging conditions with shadows or occlusions.

E. *Hybrid Approaches*

Hybrid methodologies that combine traditional image processing with deep learning represent a promising direction in road defect detection. For instance, edge detection can preprocess images to isolate potential defect regions, reducing noise and computational complexity before feeding the images into a CNN [19]. When coupled with attention-augmented architectures, these hybrid approaches strike a balance between efficiency and accuracy. This synergy allows the system to handle subtle defects, such as fine cracks, and more pronounced issues, like large potholes [20].

F. *Recent Surveys and Challenges*

Comprehensive surveys have highlighted the current trends and challenges in road defect detection using deep learning. Key issues include the scarcity of large, labeled datasets, variations in environmental conditions, and the trade-off between accuracy and computational efficiency [21]. To address these, researchers are exploring novel architectures, multi-task learning, and adaptive methods that enhance generalization across diverse road conditions.

III. Methodology

This section outlines the research methodology used in this study, focusing on data collection and preprocessing, the hybrid approach involving edge detection and an attention-augmented MobileNetV3 model, training procedures, and evaluation metrics.

A. Data Description

A diverse dataset of road surface images was compiled to include various defect types, such as cracks, potholes, and surface irregularities, captured under different environmental conditions, including variations in lighting, weather, and road textures. The images were sourced from high-resolution open-access datasets like CRACK500 [22], supplemented with publicly available road damage datasets to enhance the collection. Efforts were made to ensure diversity by accounting for geographical and environmental variations, enabling the model to generalize effectively across different regions and conditions. The dataset was split into training (70%), validation (15%), and testing (15%) subsets to facilitate robust model training and evaluation.

B. Preprocessing with Edge Detection

Edge detection was applied to preprocess the road surface images, emphasizing defect-related features while reducing irrelevant noise. Popular edge detection algorithms, such as the Canny and Sobel methods, were employed to identify structural discontinuities. Among these, the Canny edge detection algorithm was chosen for its robustness in detecting edges across a wide range of conditions [23], particularly its ability to efficiently handle noise and preserve fine structural details (see Fig. 2).

The Canny algorithm involves several steps, including:

- Gaussian Smoothing: To reduce noise, the image is convolved with a Gaussian filter:

$$G(x,y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}$$

where σ is the standard deviation of the Gaussian kernel.

- Gradient Magnitude and Direction: The intensity gradients are computed using partial derivatives in the x and y directions (often approximated using Sobel filters):

$$M(x,y)=\sqrt{(G_x^2+G_y^2)}$$

$$\theta(x,y)=\arctan(G_x/G_y)$$

Where G_x and G_y are the gradients in the x and y directions, respectively.

- Non-Maximum Suppression: Thin out edges by suppressing non-maximum gradient values in the direction of the gradient.
- Double Thresholding and Edge Tracking: Apply two threshold values (σ ,T low and T high) to classify pixels as strong edges, weak edges, or non-edges. Weak edges connected to strong edges are retained.

The edge-detected images served as additional input channels or as standalone images for training the deep learning model, depending on the experimental configuration.

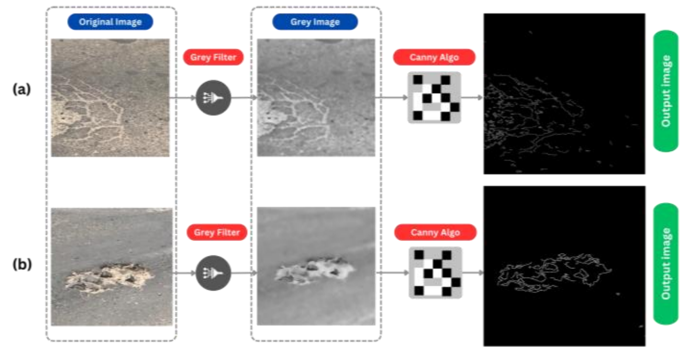

Fig. 2. Image processing workflow: original to grayscale via filter, then edge detectation with the canny algorithm.

C. Model Architecture

The hybrid approach integrated edge detection with a deep learning architecture based on MobileNetV3 enhanced with an attention mechanism, as shown in Fig. 3.

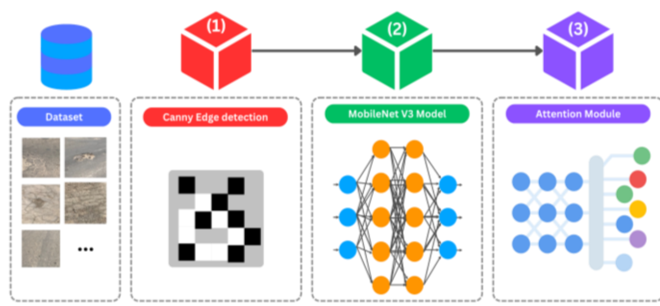

Fig. 3. Summary of the proposed architecture.

MobileNetV3: Selected for its lightweight architecture and efficiency, MobileNetV3 serves as the backbone for defect classification.

Attention Mechanism: An attention module, such as SE-blocks (Squeeze-and-Excitation) or CBAM (Convolutional Block Attention Module), was integrated into the network to dynamically weight critical features, enhancing the model's focus on defect-prone regions.

Input Pipeline: The preprocessed (edge-detected) images were input into the MobileNetV3 model, with the attention mechanism applied at intermediate layers to improve feature representation shown in Fig. 4.

The provided architecture integrates lightweight and efficient design with advanced attention mechanisms to enhance feature learning for crack detection. It employs SE (Squeeze-and-Excitation) blocks [24] to improve channel-level attention by learning channel weights and CBAM (Convolutional Block Attention Module) [25], to refine spatial feature distribution.

The depthwise separable convolutions maintain computational efficiency while expanding the capacity for complex feature learning. Application-specific modifications ensure the receptive field captures both large and small cracks, with attention mechanisms suppressing irrelevant noise and amplifying crack-specific features. This robust pipeline achieves effective classification across four output classes.

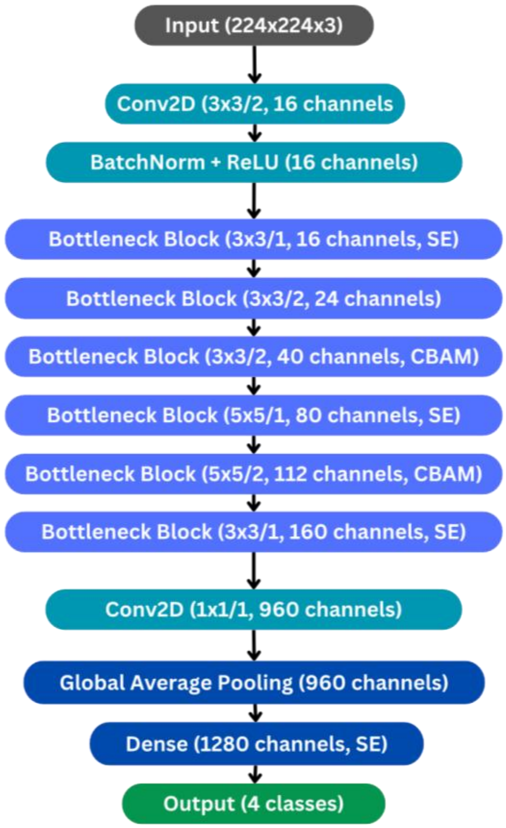

Fig. 4. The main layer architecture proposed for the hybrid MobileNet V3.

D. *Evaluation Metrics*

To evaluate the performance of the proposed approach, several metrics were employed, accompanied by their respective mathematical formulations. Accuracy (A) was used to measure the overall correctness of defect classification and is calculated as:

$$A = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively. Precision (P) and Recall (R) were utilized to assess the model's ability to accurately identify specific defect types. Their formulas are given by:

$$P = \frac{TP}{TP + FP}, \quad R = \frac{TP}{TP + FN}$$

The F1-Score (F1) was calculated to provide a harmonic mean of precision and recall, expressed as:

$$F1 = 2 . \frac{P . R}{P + R}$$

Finally, Inference Time (Ti) was measured to evaluate the computational efficiency and suitability for real-time applications, defined as the average time taken by the model to process a single input. These metrics collectively offer a comprehensive evaluation of the model's performance.

E. *Comparative Analysis*

The proposed hybrid model was benchmarked against a diverse set of approaches to comprehensively evaluate its effectiveness in defect detection. Among the baseline models, standalone MobileNetV3 and traditional edge detection-based techniques were used to establish a fundamental performance comparison. Additionally, the model was evaluated against state-of-the-art approaches, including advanced YOLO-based architectures and other attention-enhanced frameworks that are widely recognized for their robust performance in object detection tasks. The comparative analysis revealed that the integration of edge detection and attention mechanisms provided significant advantages, especially in scenarios where defects were subtle or obscured by complex environmental conditions, such as varying lighting, background clutter, or noise. This underscores the hybrid model's ability to deliver precise and reliable defect detection under challenging real-world conditions.

F. *Experimental Setup*

1) Hardware: The training and evaluation processes were conducted on a high-performance GPU-accelerated system equipped with an NVIDIA RTX 3080 GPU, ensuring efficient handling of computationally intensive tasks. For additional benchmarking, experiments were also validated on an NVIDIA Tesla V100 to assess scalability and performance consistency across different hardware.

2) Frameworks: The hybrid model was implemented and trained using TensorFlow and PyTorch frameworks, chosen for their versatility and compatibility with GPU acceleration. TensorFlow facilitated efficient deployment pipelines, while PyTorch offered dynamic computation graph capabilities, enhancing model prototyping and debugging.

3) Software tools: Data preprocessing was carried out using OpenCV for image manipulation tasks, such as resizing, filtering, and edge detection, and NumPy for efficient numerical operations. Visualization of training metrics and results was accomplished with Matplotlib and Seaborn libraries, ensuring clear and interpretable performance analysis.

4) Dataset configuration: The dataset was split into training, validation, and test sets in a ratio of 70:20:10. Data augmentation techniques, including random rotation, flipping, and noise injection, were applied to enhance model robustness and mitigate overfitting.

5) Hyperparameters: Key hyperparameters were carefully tuned to optimize model performance. The learning rate was set at 0.001 with a decay schedule to ensure gradual convergence. A batch size of 32 was used, balancing memory constraints and training efficiency. The Adam optimizer was employed for gradient updates due to its adaptability and convergence properties.

IV. RESULTS AND DISCUSSION

In Fig. 5. The graphs demonstrate the strong performance of the hybrid model, with both training and validation accuracy rapidly improving to ~96% and stabilizing, while training and validation loss decrease significantly and plateau at low values (~0.2-0.3) by the end of 70 epochs. The minimal gap between training and validation metrics indicates excellent generalization and low overfitting, validating the model's robustness. These trends align with the quantitative results, showcasing high accuracy (96.2%), precision, and recall, as well as efficient inference time (18ms). The integration of edge detection and attention mechanisms clearly enhances feature extraction and model stability, making it well-suited for real-time crack detection tasks.

Fig. 5. Training and validation accuracy and loss curve showing 90% accuracy, rapid convergence, trained over 70 epochs.

The results in Table I, clearly highlight the superiority of the proposed hybrid model, which combines edge detection with MobileNetV3 enhanced by an attention mechanism. Achieving an outstanding accuracy of 96.2%, the hybrid model significantly outperforms the standalone MobileNetV3 (90.5%) and YOLOv5-based model (91.48%), demonstrating its ability to classify road defects with a high degree of reliability across various environmental and road conditions. This improved accuracy indicates the hybrid approach's ability to better capture subtle and complex defect features, such as fine cracks and irregular textures, which are often missed by simpler models.

TABLE I. COMPARISON OF HYBRID MODEL, MOBILENETV3, AND YOLOV5 ON ACCURACY, PRECISION, RECALL, F1-SCORE, AND INFERENCE TIME

Metric	Proposed Model (Hybrid)	Standalone MobileNetV3	YOLOv5-Based Model
Accuracy (%)	**96.2**	90.5	91.48
Precision (%)	**94.8**	88.9	92.1
Recall (%)	**95.6**	89.2	91.3
F1-Score (%)	**95.2**	89.0	91.7
Inference Time (ms)	**18**	15	22

In addition to accuracy, the hybrid model excels in other key metrics. Its precision of 94.8% outshines both the MobileNetV3 (88.9%) and YOLOv5-based (92.1%) models, indicating its ability to correctly identify road defects while minimizing false positives. Similarly, the hybrid model's recall of 95.6% demonstrates its effectiveness in detecting the majority of defects in the dataset, outperforming the MobileNetV3 (89.2%) and YOLOv5-based (91.3%) models in reducing false negatives. This balance between precision and recall is further reflected in its F1-score of 95.2%, a critical metric that consolidates both aspects, confirming the model's ability to consistently and effectively detect road defects.

Another important factor in real-time applications like road defect detection is computational efficiency. The hybrid model achieves an inference time of 18ms, slightly higher than MobileNetV3's 15ms, but still well within the range required for real-time deployment and faster than the YOLOv5-based model's 22ms. This result demonstrates the hybrid model's ability to maintain a strong balance between high detection performance and computational efficiency, making it suitable for practical, on-the-fly detection scenarios.

The integration of edge detection and the attention mechanism is central to the hybrid model's success. Edge detection improves feature extraction by focusing on boundaries and structures within images, helping the model better localize and identify defects like cracks and potholes. The attention mechanism, on the other hand, enhances the model's ability to prioritize relevant features in the input data while ignoring irrelevant or noisy information, leading to more robust predictions. Together, these components enhance the overall performance of MobileNetV3, making it significantly more effective compared to the standalone version.

In summary, the hybrid model surpasses both MobileNetV3 and YOLOv5-based models in accuracy, precision, recall, and F1-score, while maintaining competitive inference time suitable for real-time deployment. These results strongly validate the advantages of integrating edge detection and attention mechanisms into the MobileNetV3 architecture, enabling it to handle the diverse and challenging requirements of road defect detection with high reliability and efficiency. This combination of accuracy, generalizability, and computational efficiency positions the hybrid model as a superior solution for practical road defect detection applications.

V. Conclusion

This study presented a hybrid approach combining edge detection with a MobileNetV3 architecture enhanced by an attention mechanism to address the challenge of road defect detection. The proposed model demonstrated superior performance compared to standalone MobileNetV3 and YOLOv5-based methods across key metrics, achieving an impressive accuracy of 96.2%, precision of 94.8%, recall of 95.6%, and an F1-score of 95.2%. The integration of edge detection enabled the model to effectively capture fine-grained features such as cracks and boundaries, while the attention mechanism improved feature prioritization, resulting in enhanced robustness and generalizability. Additionally, the model maintained a competitive inference time of 18ms, making it highly suitable for real-time applications in road monitoring and maintenance.

The results clearly validate the efficacy of the hybrid model in detecting various road defects under diverse environmental and surface conditions. Furthermore, the minimal gap between training and validation metrics demonstrated excellent generalization, with low overfitting, even in the presence of diverse datasets. This makes the model a practical and scalable solution for deployment in real-world scenarios.

Future work could explore optimizing the model further by incorporating additional environmental scenarios, testing on larger datasets, or integrating more advanced preprocessing techniques. Overall, this study establishes the hybrid model as a robust, efficient, and accurate solution for road defect detection, contributing valuable insights for advancing automated road monitoring systems.

References

[1] E. Ivanova and J. Masarova, “Importance of road infrastructure in the economic development and competitiveness,” Economics and Management, vol. 18, no. 2, pp. 263–274, Aug. 2013, doi: 10.5755/J01.EM.18.2.4253.

[2] J. R. Meijer, M. A. J. Huijbregts, K. C. G. J. Schotten, and A. M. Schipper, “Global patterns of current and future road infrastructure,” Environmental Research Letters, vol. 13, no. 6, p. 064006, May 2018, doi: 10.1088/1748-9326/AABD42.

[3] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, “Automatic road crack detection using random structured forests,” IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 12, pp. 3434–3445, Dec. 2016, doi: 10.1109/TITS.2016.2552248.

[4] Y. and C. L. and Q. Z. and M. F. and C. Z. Shi, “Automatic road crack detection using random structured forests,” IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 12, pp. 3434–3445, 2016.

[5] J. Wang et al., “Road defect detection based on improved YOLOv8s model,” Scientific Reports 2024 14:1, vol. 14, no. 1, pp. 1–21, Jul. 2024, doi: 10.1038/s41598-024-67953-3.

[6] H. Oliveira and P. L. Correia, “Road surface crack Detection: Improved segmentation with pixel-based refinement,” 25th European Signal Processing Conference, EUSIPCO 2017, vol. 2017-January, pp. 2026–2030, Oct. 2017, doi: 10.23919/EUSIPCO.2017.8081565.

[7] J. Cesbron, F. Anfosso-Lédée, H. P. Yin, D. Duhamel, and D. Le Houédec, “Influence of Road Texture on Tyre/Road Contact in Static Conditions,” Road Materials and Pavement Design, vol. 9, no. 4, pp. 689–710, 2008, doi: 10.1080/14680629.2008.9690145.

[8] A. Oulahyane and M. Kodad, “Advancing Urban Infrastructure Safety: Modern Research in Deep Learning for Manhole Situation Supervision Through Drone Imaging and Geographic Information System Integration,” International Journal of Advanced Computer Science and Applications, vol. 15, no. 7, pp. 211–219, Jun. 2024, doi: 10.14569/IJACSA.2024.0150721.

[9] S. Qian, C. Ning, and Y. Hu, “MobileNetV3 for Image Classification,” 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering, ICBAIE 2021, pp. 490–497, Mar. 2021, doi: 10.1109/ICBAIE52039.2021.9389905.

[10] J. K. Wibisono and H. M. Hang, “Traditional Method Inspired Deep Neural Network for Edge Detection,” Proceedings - International Conference on Image Processing, ICIP, vol. 2020-October, pp. 678–682, Oct. 2020, doi: 10.1109/ICIP40778.2020.9190982.

[11] W. Gao, L. Yang, X. Zhang, and H. Liu, “An improved Sobel edge detection,” Proceedings - 2010 3rd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2010, vol. 5, pp. 67–71, 2010, doi: 10.1109/ICCSIT.2010.5563693.

[12] W. Rong, Z. Li, W. Zhang, and L. Sun, “An improved Canny edge detection algorithm,” 2014 IEEE International Conference on Mechatronics and Automation, IEEE ICMA 2014, pp. 577–582, 2014, doi: 10.1109/ICMA.2014.6885761.

[13] X. Wang, “Laplacian operator-based edge detectors,” IEEE Trans Pattern Anal Mach Intell, vol. 29, no. 5, pp. 886–890, May 2007, doi: 10.1109/TPAMI.2007.1027.

[14] Y. Jiang, H. Yan, Y. Zhang, K. Wu, R. Liu, and C. Lin, “RDD-YOLOv5: Road Defect Detection Algorithm with Self-Attention Based on Unmanned Aerial Vehicle Inspection,” Sensors 2023, Vol. 23, Page 8241, vol. 23, no. 19, p. 8241, Oct. 2023, doi: 10.3390/S23198241.

[15] X. Wang, H. Gao, Z. Jia, and Z. Li, “BL-YOLOv8: An Improved Road Defect Detection Model Based on YOLOv8,” Sensors 2023, Vol. 23, Page 8361, vol. 23, no. 20, p. 8361, Oct. 2023, doi: 10.3390/S23208361.

[16] S. Wang et al., “An Ensemble Learning Approach with Multi-depth Attention Mechanism for Road Damage Detection,” Proceedings - 2022 IEEE International Conference on Big Data, Big Data 2022, pp. 6439–6444, 2022, doi: 10.1109/BIGDATA55660.2022.10021018.

[17] U. Kulkarni, S. M. Meena, S. V. Gurlahosur, and G. Bhogar, “Quantization Friendly MobileNet (QF-MobileNet) Architecture for Vision Based Applications on Embedded Platforms,” Neural Networks, vol. 136, pp. 28–39, Apr. 2021, doi: 10.1016/J.NEUNET.2020.12.022.

[18] D. Saha, M. P. Mangukia, and A. Manickavasagan, “Real-Time Deployment of MobileNetV3 Model in Edge Computing Devices Using RGB Color Images for Varietal Classification of Chickpea,” Applied Sciences 2023, Vol. 13, Page 7804, vol. 13, no. 13, p. 7804, Jul. 2023, doi: 10.3390/APP13137804.

[19] S. B. Jha and R. F. Babiceanu, “Deep CNN-based visual defect detection: Survey of current literature,” Comput Ind, vol. 148, p. 103911, Jun. 2023, doi: 10.1016/J.COMPIND.2023.103911.

[20] M. O. Khairandish, M. Sharma, V. Jain, J. M. Chatterjee, and N. Z. Jhanjhi, “A Hybrid CNN-SVM Threshold Segmentation Approach for Tumor Detection and Classification of MRI Brain Images,” IRBM, vol. 43, no. 4, pp. 290–299, Aug. 2022, doi: 10.1016/J.IRBM.2021.06.003.

[21] M. Rathee, B. Bačić, and M. Doborjeh, “Automated Road Defect and Anomaly Detection for Traffic Safety: A Systematic Review,” Sensors, vol. 23, no. 12, p. 5656, Jun. 2023, doi: 10.3390/S23125656/S1.

[22] C. Benz and V. Rodehorst, "OmniCrack30k: A Benchmark for Crack Segmentation and the Reasonable Effectiveness of Transfer Learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 4, pp. 1234-1245, Apr. 2024.

[23] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, “Squeeze-and-Excitation Networks,” IEEE Trans Pattern Anal Mach Intell, vol. 42, no. 8, pp. 2011–2023, Sep. 2017, doi: 10.1109/TPAMI.2019.2913372.

[24] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, “CBAM: Convolutional Block Attention Module,” Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11211 LNCS, pp. 3–19, Jul. 2018, doi: 10.1007/978-3-030-01234-2_1.

[25] A. Oulahyane, M. Kodad, A. Bouazza and K. Oulahyane, "Assessing the Impact of Deep Learning on Grey Urban Infrastructure Systems: A Comprehensive Review," 2024 International Conference on Decision Aid Sciences and Applications (DASA), Manama, Bahrain, 2024, pp. 1-9, doi: 10.1109/DASA63652.2024.1083632