Optimized Road Surface Anomaly Detection And Recommendation

Dorathi Jayaseeli J D
Associate Professor
Department of Computing Technologies
SRM Institute of Science and Technology
Kattankulathur – 603203, Chengalapattu District, Tamil Nadu, India
dorathij@srmist.edu.in

K. Leela Sriram
Department of Computing Technologies
SRM Institute of Science and Technology
Kattankulathur – 603203, Chengalapattu District, Tamil Nadu, India
lk9097@srmist.edu.in

Ch. Thulasi Sandeep
Department of Computing Technologies
SRM Institute of Science and Technology
Kattankulathur – 603203, Chengalapattu District, Tamil Nadu, India
tc5951@srmist.edu.in

Abstract: The paper introduces an intelligent visual inspection and alert generation system that combines the deep learning based computer vision with a web-based monitoring framework to assist in automated detection of road surface anomalies such as potholes, cracks, and pavement deterioration. The suggested system uses a hybrid inference pipeline by integrating real-time object detection model and optimized Vision Transformer architectures that provide accurate recognition at low computational costs. Model optimization methods like quantization and pruning are added to minimise inference latency and memory cost so that they could be deployed on resource-constrained environments. The system examines input images by using a detection and classification process, locating recognized objects in a structured materials knowledge base, and examining their contextual importance by applying pre-defined severity rules. According to this rule-based evaluation, the platform assigns the risk levels and sends alerts whenever the critical conditions are identified. Image ingestion, model inference, result logging, and alert handling are handled by a Flask-based backend, with visualization of detections, annotated results, and history of alerts being offered by a lightweight frontend to be interpreted by a user. The results of detection and alerts are stored continuously to facilitate traceability as well as post-event analysis. It is proposed to focus on the explainability and operational-usefulness of an approach which involves a combination of the data-driven predictions with the rule-based decision logic which is transparent. Experimental results prove that the system may serve as an end to end visual monitoring system that can be used as a safety, inspection, and surveillance-focused monitoring system that provides a tradeoff between the detection accuracy, interpretability, and deployment readiness in a single IEEE-based system.

Keywords: Road damage detection, Object detection, Vision Transformer, Pothole detection, Road surface anomaly detection, Intelligent visual monitoring

I. INTRODUCTION

With the evolution of the computer vision and deep learning technologies, the development of automatic systems to comprehend visual information is increasingly becoming important for applications such as safety monitoring, infrastructure inspection, and intelligent surveillance. Traditional road inspection methods rely on manual visual surveys, which are time-consuming, expensive, and prone to human error, but they expend a lot of time, are prone to errors, and hard to implement efficiently in dynamic environments. With the advent of abundant image data and computing processing capacity, intelligent vision-based systems have become an effective way for object detection, visual scene analysis and supporting real-time decision making. Recent research demonstrated the significant improvement of automated anomaly detection systems using deep learning on applications in visual inspection [2], [8].

Recent enhancements in deep learning model architectures, especially convolutional neural networks (CNNs) and vision models based on the transformer architecture, have led to a lot of improvements in the performance of object detection and classification tasks. Vision transformer models afford excellent contextual representations for fine-grained visual understanding and real-time detection frameworks can effectively localize the objects in the complex scenarios. These advancements have allowed intelligent systems to be created for detecting damage on roads and other infrastructure from images [1], [7]. However, in spite of these improvements in use, there are still difficulties in the implementation of such use cases in real-life situations, especially when you have limited time for computations and require adjustable processing speed.

Some of the constraints of the potentiation of deep learning systems to a monitoring and inspection scenario include large model size, high inference time, and lack of model interpretability. To overcome these problems, in this study, a concept of intelligent visual inspection framework based on the integration of efficient techniques of object detection, optimized vision transformer model and decision logic based decisions in a unified system architecture is proposed. Model optimization techniques such as pruning and quantization are integrated to minimize the computational complexity while retaining the detection accuracy.

In addition to visual recognition, the proposed system focuses on the interpretation of the detected situations and places the detected anomalies in context by linking them to a structured knowledge base of materials and types of damage. The system assesses the objects detected by predefined severity rules in order to determine their significance to the operation. This way, raw model predictions are translated into forms that are easy to interpret as risk assessments and produce alerts decoded as actions, bridging the gap between visual perception and real-world decision support.

Furthermore, the proposed framework follows a web-based architecture for usability, traceability and also real-time interaction. A lightweight backend has been built to manage image ingestion, model inference and result storage and an interactive user-interface for visualizing annotated images, as well as detection summarization, and alert history. By employing a combination of data-driven deep learning models and understandable rule-based reason, the system enhances and improves its interpretations and ability to perform operations. Overall, the proposed approach shows how the optimized deep learning architectures and structured decision processes can be combined to create a practical and interpretable deployment-ready visual monitoring system for safety-critical infrastructure inspection purposes.

II. Literature Survey

The demand for intelligent transportation systems and infrastructure monitoring has increased, and consequently, more and more research has been conducted in the area of automated road damage detection systems using computer vision and deep learning methods. Traditional methods for inspecting roads can require manual surveys which are time consuming and costly, and subject to human error. Recent developments in the field of deep learning and image processing have made it possible to build automatic systems that can determine anomalies in the road surface such as a pothole, a crack on the road, or a deteriorated surface with greater accuracy and efficiency.

Abdelwahed et al. [1] provided a comprehensive survey on real-time methods for detecting road damage and datasets. Their research looked at different architectures of deep learning algorithms that are used for road surface defect detection and the analysis of the effectiveness of the publicly available datasets. The roles of real-time detection systems (for road maintenance) and deep learning models in enhancing detection accuracy under various environment conditions, were highlighted by these authors.

Image-based anomaly detection with deep learning has also been a well researched problem in industrial inspection scenarios. Shukla et al. [2] presented a systematic survey about deep learning based approaches in visual anomaly detection in industrial applications. Their work included discussion of different paradigms in learning methods such as supervised, unsupervised and semi-supervised learning, and the effectiveness of convolution neural nets and transformer architecture for identifying minute visual defects.

Alzarooni et al. [3] reviewed the techniques used in industrial systems for anomaly detection, and addressed the challenges related to dealing with high dimensionality data, model interpretability and model efficiency. The achievements of the study identified the need for hybrid models, which help to merge data-based approaches with domain knowledge in order to achieve reliable and explainable anomaly detection systems, especially in safety-relevant settings.

Baroudi et al. [4] introduced an advanced pothole detection framework using image segmentation and image depth estimation methods for correct road anomaly detection. Their approach showed that in combining spatial information with deep learning models one gets enhanced ability to detect and characterize road damage under complex road conditions.

Garita-Duran et al. [5] proposed a deep learning-based system for the automated damage detection and quantification in concrete pavements. The road infrastructure has been the focus of their work based on the detection of structural defects using computer vision techniques and has shown the potential of deep learning models for efficient monitoring and maintenance planning of road infrastructure.

Research has also been done into advanced neural architectures to perform tasks of detecting anomalies. Scientific [6] suggested a cascaded capsule neural network combined with an optimized LSTM model for anomalous segmentation and classification. The study showed an enhanced performance for capturing spatial relationships and temporal dependencies that are key to successfully detect complex anomalies in the visual data.

Li et al. [7] proposed a road damage detection algorithm based on a region guidance network to enhance the detection effect in a complex background environment. Their model showed better robustness and accuracy in detecting road defects in the real world where the lighting conditions and environmental noise might impact the detection performance.

Huang et al. [8] presented an extensive overview on deep learning progress in anomaly detection under variety of fields, such as industrial systems, 21st Century Internet of Things (IoT), and smart infrastructure. The survey identified important research issues like model scalability, interpretability, arising deployment efficiency which are still important parameters in the research of practical anomaly detection systems.

Li and Zhang [9] have proposed a lightweight road damage detection model based on improved YOLOv8 model. Their strategy focused on balancing as low computational complexity as possible and achieving a high level of detection accuracy to make the model a suitable candidate to be used for real-time road monitoring systems running on resource constrained devices.

Finally, an intelligent road anomaly detection system was developed by Almakhluk et al. [10] and combined with a real-time notification system. Their work has shown how models for detecting, based on the principle of vision, can be combined with alert generation mechanisms to contribute to the improvement of road safety and even to the timely maintenance of infrastructure.

Overall, the current literature shows great improvements in the development of deep learning-based road damage detection systems. However, challenges such as model efficiency, real-time environment deployment, and detection outcomes interpretability are also significant issues of research. These limitations provide motivation for developing integrated frameworks that include optimized deep learning models and structured decision mechanisms that can be used for developing reliable and practical road anomaly monitoring systems.

III. PROPOSED METHODOLOGY

A. System Overview

The suggested methodology is developed in the form of a modular and end-to-end pipeline that is destined to be used to identify and generate alerts on intelligent visual anomalies. The system combines image capture, the inference by deep learning, the interpretation of the context, and alerting with the help of rules into one system. Optimized vision models are used to detect and classify anomalous regions or objects using input images. The identified outputs are then gauged with the help of a structured knowledge base and severity rules to find their operational importance. This multi-layered architecture provides scalability, interpretability and real time monitoring applications.

B. Image Acquisition and Preprocessing.

This methodology starts with the process of acquiring images by user uploads or the associated visual sources. Preprocessing is used prior to inference in order to be resistant to changing lighting, resolution and noise. These procedures will involve image downsizing to fit model inputs, pixel intensity distributions standardization through normalization and format verification to guarantee compatibility with the inference pipeline. Preprocessing minimizes the computational cost and boosts consistency among a wide range of visual inputs which increases the reliability of detections in a deployed environment.

C. Visual Feature Extraction and Object Detection.

To determine the location of the anomaly, a real-time object detection model is used to detect the regions of interest in the input image. The detector provides bounding boxes and loose labels of possible anomalous objects. Regions in the detected areas are further sent to a classification module, which operates using a vision transformer to obtain high-level semantic features and narrow down their interpretation of objects. This is a two step process that isolates localization and contextual insight to allow impression of accurate detection even in visually complex scenes. Pruning and quantization are examples of model optimization methods that are included in order to preserve the low inference latency without losses to the detection results.

D. Model Optimization and Efficient Inference.

The proposed methodology focuses on efficient inference to enable it to be deployed in resource constrained or real time environments. The vision transformer models are pruned and quantized to make the model smaller in terms of size and memory consumed. These optimizations enable higher speeds with the required representational bandwidth. The system offers a reasonable balance between accuracy and efficiency to convert into a viable practical system to be used in continuous monitoring applications where instant action is essential.

E. Analysis of the context and its severity.

After visual inference, the objects that are detected are mapped onto a structured materials and anomaly knowledge base. The individual detected objects are assessed based on pre-regulated severity criteria, which take into account the object type, the relevance to the context, and the risk probability. This layer of evaluation which is based on the rules converts the raw model predictions to understandable level of severity which can be low, medium or high risk. Explicit rules are more likely to increase transparency, and decision-making is more likely to be conducted in accordance with domain-specific safety or operational needs instead of focusing on the results of probabilistic models.

F. Alert Generation, Logging and Visualization.

An alert management module is used in the last stage where anomalies that are severity assessed are processed. In case predefined thresholds are crossed, alerts are created and stored with logs and timestamps and detection metadata to track who used what and when. The system allows real-time alerts and continuous records of detection and alerts. An annotated image system is visualized as a web-based interface, detection summary and alert history so that users can review and confirm system outputs. This closed-loop model supports such a design that ensures that identified anomalies result in actionable insights and informed decision making.

In general, the suggested methodology integrates optimized deep learning models with systematic rule-based models to provide an interpretable, efficient and deployment-ready visual anomaly detection system that can be used in safety-critical and inspection-oriented applications.

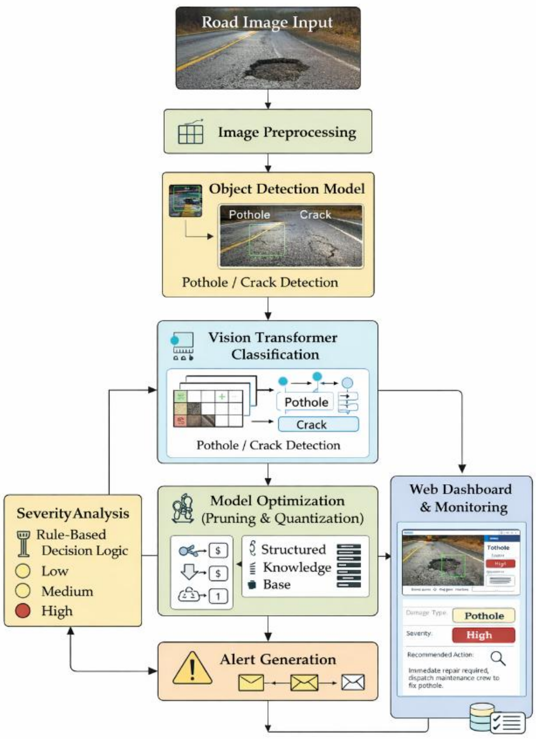

Figure 1. System Architecture for Proposed Framework of Road Surface Anomaly Detection & Alerting

Figure 1. proposes the work flow of the deep learning-based road surface anomaly detection system with image acquisition and preprocessing to object detection, Vision Transformer for classification, severity analysis, and alert generation, and web-based monitoring dashboard.

IV. RESULTS AND DISCUSSIONS

The proposed intelligent visual anomaly detection system is shown to be successful in the experimental assessment of the system to recognize the occurrence of anomalous objects accurately and translate the visual inferences into actionable alerts. The detection and classification pipeline has been carefully designed to work with any input image and deliver reliable results in localization and recognition of abnormal objects in diverse visual conditions.

To evaluate the performance of the proposed road surface anomaly detection system several quantitative evaluation metrics were considered. These metrics are accuracy, precision, recall, F1-score and mean Average Precision (mAP). The resulting results show the proposed hybrid detection framework offers reliable detection capabilities without compromising efficient real-time processing capabilities.

The results showed that the proposed system achieved high detection accuracy with the balance of unprecedented detection accuracy in terms of precision and recall for reliable identification of road anomalies.

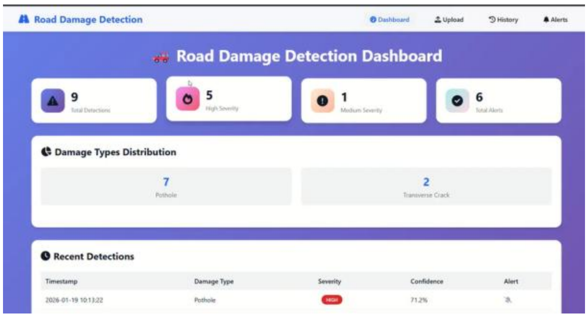

Figure 2. Road Damage Detection Dashboard Overview

This figure 2 shows the system dashboard of the anomalies detected along the road. It shows the total detections, counts in terms of severity, alert statistics, and type of damage distribution. The dashboard allows fast situational awareness to be gained since detection results are available, as well as recent events to help the monitoring authorities see the overall situation on the road and put the focus on the effective maintenance actions.

The optimized vision transformer models allow high reliability in features extracting and at the same time it is efficient in its inference, which is very crucial in real time or close to real time observation. The system produces labeled visual results, which distinctly are signaled as detected regions, which enhances easier interpretation and facilitates easy verification of the results of detection. Traceability of the results may be further maintained through logging mechanisms that store organized records of the detection results and events of alerts.

Figure 3. Visual Detection and Severity Annotation Output

This is fig 3 the result of the visual inference that the proposed system produces. Bounding boxes and labels of damage type and severity are used to indicate the road damage that is detected. The annotated output enhances the interpretability by allowing the anomalies on the road surface to be localized where they can be associated with the level of severity to be applied in making decisions.

In the view of a system, the hybrid structure that integrates data-driven deep learning models and rule-based severity assessment adds to the operational reliability significantly. The contextual analysis layer makes use of probabilistic model results, but also compares the objects found with predefined severity policies, with which the system can distinguish between harmless anomalies and conditions that are critical and must be addressed. Such systematic assessment minimizes the number of false or unneeded notifications and increases the consistency of decisions. The mechanism of alert generation operates as planned because it generates the notification messages only when the severity levels exceed the specified limits, which proves that the system can be used to support safety-related and inspection-oriented applications. Vertical processing pipeline is also used to guarantee modularity, so that the framework can be adapted to various deployment settings without affecting the basic detection logic.

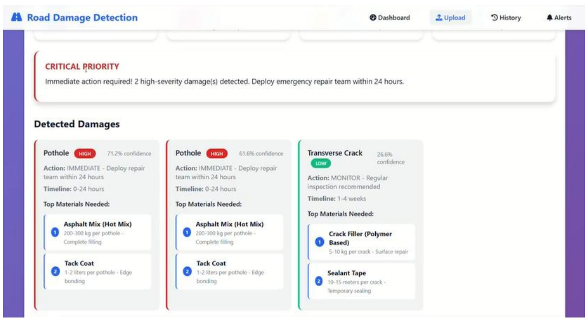

Figure 4. Severity-Based Action Recommendation Interface

This is a figure 4 that displays the post-detection analysis interface that displays identified damages along with the level of confidence, severity category, and suggested measures. The system cross correlates the identified anomalies with pre-defined response plans and material needs. This interface enables the timely maintenance planning through translation of detection results into actionable repair recommendations.

The findings also reveal that the model optimization methods are vital in real-life application. Quantization and pruning minimize computational overhead and inference latency allowing easier integration into web-based systems and resource constrained systems. Although the system has shown good detection reliability and interpretability, features of improvements that can be made in the future such as increasing knowledge material and fine-tuning severity rules to suit other fields of application are also discussed. All in all, the findings confirm that the suggested solution presents a good balance between detection accuracy, efficiency level, and explainability. The system eliminates the major drawbacks of model-driven only methods by making visual anomaly detection a structured decision-support process and a solid base of the real world monitoring and inspection conditions.

The comparison model shows the proposed system better the traditional CNN-based and the standard YOLO-based systems in terms of detection accuracy and system functionality. In addition to the enhanced detection performance, the proposed framework combines the severing evaluation and alert generation capabilities, which makes it suitable to the task of monitoring road in real-time and decision support for road maintenance.

V. CONCLUSION AND FUTURE SCOPE.

The suggested intelligent visual anomaly detection and alerting takes the form of a well-organized and deployment-ready intelligent visual inspections through a combination of deep learning-based perception with rule-based contextual reasoning. The system can be effective in converting raw visual data into interpretable and actionable results by integrating real time object detection, state-of-the-art vision transformer models, and severity-based decision logic. The modular design will guarantee that detection, analysis and alert generation are synchronized and flexible to the various monitoring environments. The system can be effectively used to operate continuously in the resource-constrained environment thanks to the model optimization methods like pruning or quantization, which do not reduce the accuracy of the detection. Black-box anomaly detection models are characterized by limitations, so the addition of a layer based on knowledge-based severity assessment increases the level of transparency and minimizes unwarranted alerts. In general, the system is accurate, efficient, and explainable, which meets the major criteria of safety-critical and inspection-oriented systems.

The suggested framework can be further supported and increased in terms of its applicability and strength in the future. Another possible avenue is the development of the materials and anomaly knowledge base to accommodate more application areas to facilitate a more detailed severity assessment and contextual interpretation. Dynamically improving the severity thresholds through adaptive rule-learning mechanisms can also be added based on the historical detection data and the changing operational conditions. Also, extending the system to allow multi-modal inputs, e.g. video streams or sensor metadata, might enhance the characterization of anomalies and time-based interpretation. User trust can be further increased by integration of advanced explainability techniques to give a better understanding of model decisions. Lastly, the large-scale field validation techniques and incremental model updating approaches can enhance the robustness toward concept drift and dynamic environmental patterns, which can guarantee the longevity in long-term deployments.

REFERENCES

[1 Abdelwahed, S. H., Sharobim, B. K., Wasfey, B., & Said, L. A. (2025). Advancements in real-time road damage detection: a comprehensive survey of methodologies and datasets. *Journal of Real-Time Image Processing*, *22*(4), 137.

[2] Shukla, V., Shukla, A., SK, S. P., & Shukla, S. (2025). A systematic survey: role of deep learning-based image anomaly detection in industrial inspection contexts. Frontiers in Robotics and AI, 12, 1554196.

[3] Alzarooni, A., Iqbal, E., Khan, S. U., Javed, S., Moyo, B., & Abdulrahman, Y. (2025). Anomaly detection for industrial applications, its challenges, solutions, and future directions: A review. arXiv preprint arXiv:2501.11310.

[4] Baroudi, U., BaHamid, A., Elalfy, Y., & Alami, Z. A. (2025). Enhancing Pothole Detection and Characterization: Integrated Segmentation and Depth Estimation in Road Anomaly Systems. arXiv preprint arXiv:2504.13648.

[5] Garita-Durán, H., Stöcker, J. P., & Kaliske, M. (2025). Deep learning-based system for automated damage detection and quantification in concrete pavement. *Results in Engineering*, *25*, 104546.

[6] Scientific, L. L. (2025). A NOVEL ENTROPY-BASED CASCADED CAPSULE NEURAL NETWORK WITH AN OPTIMIZED LSTM FOR ANOMALY SEGMENTATION AND CLASSIFICATION. Journal of Theoretical and Applied Information Technology, 103(8).

[7] Li, J., Qu, Z., Wang, S., & Xia, S. (2025). A method of road damage detection for complex background images based on region guidance network. *Pattern Recognition*, *168*, 111780.

[8] Huang, H., Wang, P., Pei, J., Wang, J., Alexanian, S., & Niyato, D. (2025). Deep learning advancements in anomaly detection: A comprehensive survey. IEEE Internet of Things Journal.

[9] Li, X., & Zhang, Y. (2025). A Lightweight Method for Road Damage Detection Based on Improved YOLOv8n. *Engineering Letters*, *33*(1).

[10] Almakhluk, A., Baroudi, U., & El-Alfy, Y. (2025). Intelligent Road Anomaly Detection with Real-time Notification System for Enhanced Road Safety. arXiv preprint arXiv:2505.08882.