Real-Time Road Damage Detection System on Deep Learning Based Image Analysis

Bakhytzhan Kulambayev[1], Belik Gleb[2], Nazbek Katayev[3], Islam Menglibay[4], Zeinel Momynkulov[5]
Turan University, Almaty, Kazakhstan[1, 2]
Kazakh National Women's Teacher Training University, Almaty, Kazakhstan[3]
International Information Technology University, Almaty, Kazakhstan[4, 5]

Abstract*—This research paper introduces a sophisticated deep learning-based system for real-time detection and segmentation of road damages, utilizing the Mask R-CNN framework to enhance road maintenance and safety. The primary objective was to develop a robust automated system capable of accurately identifying and classifying various types of road damages under diverse environmental conditions. The system employs advanced convolutional neural networks to process and analyze images captured from road surfaces, enabling precise localization and segmentation of damages such as cracks, potholes, and surface wear. Evaluation of the model's performance through metrics like accuracy, precision, recall, and F1-score demonstrated high effectiveness in real-world scenarios. The confusion matrix and loss curves presented in the study illustrate the system's ability to generalize well to unseen data, mitigating overfitting while maintaining high detection sensitivity. Challenges such as variable lighting, shadows, and background noise were addressed, highlighting the system's resilience and the need for further dataset diversification and integration of multimodal data sources. The potential improvements discussed include refining the convolutional network architecture and incorporating predictive maintenance capabilities. The system's application extends beyond mere detection, promising transformative impacts on urban planning and infrastructure management by integrating with smart city frameworks to facilitate real-time, predictive road maintenance. This research sets a benchmark for future developments in the field of automated road assessment, pointing towards a future where AI-driven technologies significantly enhance public safety and infrastructure efficiency.

***Keywords*—Deep learning; road damage detection; Mask R-CNN; image segmentation; convolutional neural networks; infrastructure management; smart cities; real-time analytics; predictive maintenance; urban planning**

I. INTRODUCTION

The ability to detect and assess road damage accurately and efficiently is pivotal in ensuring safe and sustainable road infrastructure. As road networks continue to expand and traffic volumes increase, traditional manual inspection methods become less feasible, demanding more advanced and automated solutions. In recent years, deep learning has revolutionized various domains of computer vision, including image classification, object detection, and semantic segmentation, making it a prime technology for addressing the complex task of road damage detection [1], [2].

Current methodologies for road condition monitoring largely depend on manual surveys or the use of basic sensor technology, which are labor-intensive, costly, and often inconsistent in terms of data quality and timeliness [3]. These traditional methods are not only slow but also prone to human error, leading to delays in maintenance and potentially hazardous driving conditions [4]. As a result, there is a pressing need for more robust, automated systems that can perform these tasks with greater accuracy and speed.

Deep learning offers a transformative approach for this application, due to its ability to learn hierarchical features from large datasets of images, surpassing the performance of traditional machine learning algorithms [5]. Particularly, convolutional neural networks (CNNs) have demonstrated exceptional proficiency in image-based tasks, making them suitable for the segmentation and classification of road damages from digital images captured by vehicle-mounted cameras or drones [6], [7]. These models can be trained to detect a variety of road damages such as cracks, potholes, and erosion with high precision.

The integration of deep learning with image analysis for road damage detection not only enhances the efficiency of the detection process but also significantly improves the accuracy of damage classification and segmentation. By automating damage detection, transportation agencies can swiftly identify and prioritize maintenance tasks, optimizing repair operations and ultimately reducing costs [8]. Moreover, real-time road damage detection systems can provide immediate data to drivers and relevant authorities, enhancing road safety and facilitating better traffic management [9].

Despite the potential benefits, the implementation of deep learning for real-time road damage detection poses several challenges. These include the high variability of damage types, the vast differences in road conditions due to environmental factors, and the extensive computational resources required for processing and analyzing high-resolution images [10]. Addressing these challenges is crucial for developing an effective system capable of operating under diverse and dynamic environmental conditions.

This paper proposes a novel real-time road damage detection and segmentation system based on deep learning. The system utilizes advanced deep learning architectures to analyze images captured in real-time, accurately identifying and segmenting road damages. By harnessing the power of state-of-the-art CNN models, the proposed system aims to deliver high accuracy and real-time performance, ensuring timely and effective road maintenance interventions. The efficacy of the system is demonstrated through extensive tests conducted under various environmental conditions, confirming its capability to adapt and perform reliably in real-world scenarios [11].

In summary, the transition from traditional methods to deep learning-based approaches in road damage detection not only promises improvements in maintenance scheduling and cost efficiency but also plays a crucial role in enhancing road safety and traffic management. The following sections will detail the methodology, experiments, and results of the proposed system, providing a comprehensive evaluation of its performance and implications for future road maintenance strategies.

II. Related Work

The evolution of road damage detection methodologies has been significantly influenced by advancements in image processing and machine learning techniques. Prior studies have predominantly focused on enhancing the accuracy and efficiency of detecting various road anomalies through automated systems. These systems range from basic image processing techniques to sophisticated machine learning and deep learning models that aim to minimize human intervention and improve the reliability of assessments.

Initial approaches in automated road damage detection were grounded in traditional image processing techniques, which included edge detection, texture analysis, and thresholding methods to identify damage features in road images [12]. While these methods provided a foundation for automated systems, they were limited by their sensitivity to lighting conditions and road surface variations, which often resulted in high false positive rates [13].

The integration of machine learning techniques marked a significant advancement in this field. For instance, support vector machines (SVM) and decision trees were employed to classify road conditions based on feature sets extracted from images. These models offered improvements over basic image processing by providing more robust classifications, adapting to various road conditions through feature learning [14], [15]. However, the performance of these methods heavily depended on the quality and selection of hand-crafted features, which were not always capable of capturing complex patterns in road damage [16].

The advent of deep learning, particularly convolutional neural networks (CNNs), has dramatically transformed the landscape of road damage detection. CNNs, with their ability to autonomously learn features directly from data, have shown superior performance in image classification and object detection tasks [17]. Recent studies have utilized CNNs to automatically detect and classify road damages from images captured by standard cameras mounted on vehicles or drones, achieving significant improvements in detection accuracy and processing speed [18], [19].

Segmentation models like U-Net and SegNet have further refined the capabilities of CNNs by not only detecting but also delineating the exact boundaries of road damages, such as cracks and potholes. These models perform pixel-wise segmentation to provide detailed maps of road damage, which are crucial for precise maintenance planning [20], [21]. The accuracy of these segmentation models in real-world scenarios confirms their potential in practical applications, as noted in several benchmark studies [22].

Moreover, the application of transfer learning, where pre-trained networks on large datasets are fine-tuned for specific tasks like road damage detection, has also gained popularity. This approach leverages the learned features from general contexts, significantly reducing the need for large domain-specific datasets and computational resources, thus accelerating the training process and enhancing model generalizability [23], [24].

Real-time detection systems have incorporated these deep learning models to provide immediate feedback on road conditions. Such systems are critical for dynamic traffic management and timely maintenance interventions. The integration of real-time data processing with deep learning models presents a promising avenue for deploying more responsive and adaptive road infrastructure management systems [25], [26].

Nevertheless, challenges remain, particularly in the areas of dataset diversity and model robustness under varied environmental conditions. Most existing datasets do not fully represent the wide range of damage types and severities encountered in different geographical regions, which can hinder the performance of the models [27]. Moreover, the computational demand for processing high-resolution images in real-time necessitates efficient model architectures and hardware acceleration techniques [28], [29].

In summary, the field of road damage detection has evolved from manual inspections to highly automated systems based on cutting-edge deep learning technologies. This progression not only enhances the efficiency and accuracy of detection but also underscores the growing need for continuous innovation in model development and system design to address the diverse challenges encountered in real-world applications.

III. Materials and Methods

A. Proposed System

The architecture of the proposed real-time road damage detection and segmentation system is depicted in Fig. 1. This comprehensive framework integrates various stages of data handling, from collection to processing, and ultimately to the deployment of a deep learning model for damage analysis and reporting.

1) Data collection: The initial phase involves the systematic collection of road imagery. This data is sourced using mobile cameras mounted on vehicles, which traverse various road types under different conditions, capturing a wide array of road surfaces and damage manifestations.

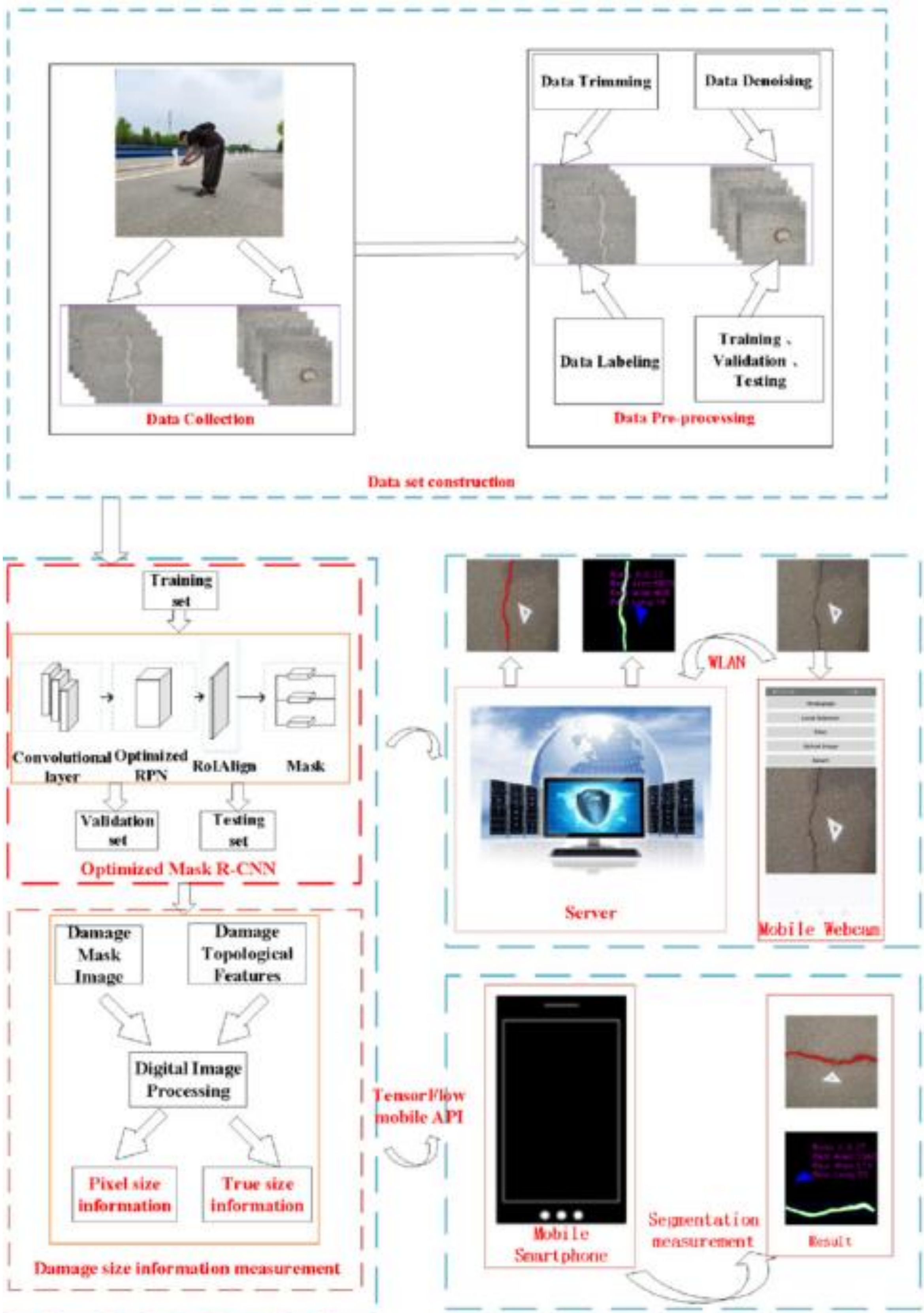

Fig. 1. Architecture of the real-time road damage detection and segmentation system.

2) Data set construction: The collected data undergoes several processing steps:

- Data Trimming and Denoising: Raw images are first trimmed to focus on relevant sections containing road surfaces. Noise reduction algorithms are applied to enhance image quality, crucial for accurate feature extraction in subsequent steps.
- Data Labeling: Images are manually labeled to identify different types of road damages such as cracks, potholes, and erosion. This labeled dataset is then split into training, validation, and testing sets.
- Data Pre-processing: The labeled images are pre-processed to normalize the lighting conditions, align features, and scale the images to uniform dimensions suitable for input into the deep learning model.

3) Model training and validation: The core of the system is an optimized Mask R-CNN model, which is a state-of-the-art deep learning model known for its efficiency in object detection and instance segmentation:

- Convolutional Optimized RollAug layer: A custom convolutional layer is introduced to enhance the feature extraction capabilities of the model. RollAug, an augmentation technique, is applied to provide robustness against various orientations and scales of road damage.
- Training and Validation: The model is trained on the pre-processed images using a dedicated server with high computational power to handle the extensive data and complex model architectures. The validation process iteratively tests the model against a reserved subset of the data to tune the hyperparameters and improve model accuracy.

4) Deployment: For real-time analysis, the trained model is deployed over a server that communicates with a mobile application:

- Server: It hosts the trained Mask R-CNN model and handles requests from the mobile application for image analysis.
- Mobile Webcam and Smartphone Integration: The mobile application captures live road images via a mobile webcam and sends them to the server for processing.
- TensorFlow Mobile API: This API facilitates the interaction between the mobile app and the server, ensuring efficient transmission of image data and retrieval of analysis results.
- Segmentation Measurement and Reporting: The server processes the incoming images, applies the Mask R-CNN model to detect and segment road damages, and sends the results back to the mobile device. The results include the type, size, and exact location of the damage, presented in a user-friendly format on the smartphone app.

5) Digital image processing and measurement: In the final stage, the segmented damages are analyzed to measure their dimensions and assess their severity. The system employs algorithms to calculate pixel-to-real-world conversions to estimate the true size of the damages. These measurements are crucial for maintenance planning and prioritization.

In summary, the proposed system leverages advanced image processing techniques, robust deep learning models, and real-time data communication to provide an efficient and accurate road damage detection and segmentation solution. This architecture not only enhances the capability of road maintenance teams to identify and rectify road damages swiftly but also supports the overarching goal of maintaining safer road conditions for the public.

B. Dataset

Fig. 2 provides a detailed taxonomy of road damage types classified for the purpose of automated detection and segmentation. The classification is organized into major categories and specific details, which are assigned unique class names for identification in the system. The types of cracks identified include “Longitudinal” under the class name D00, primarily occurring along the wheel mark part, and “Lateral” cracks categorized as D10, typically found at equal intervals across the road. Additionally, the figure categorizes “Alligator Cracks” as D20, which can appear over partial or entire pavement areas. Beyond cracks, the classification extends to “Other Corruption” with class names D40, D43, and D44, encompassing road damage such as rutting, bumps, potholes, separations, crosswalk blurs, and white line blurs. This structured categorization aids in the precise detection and analysis of road conditions, facilitating targeted maintenance actions based on the severity and type of road damage.

Damage Type			Detail	Class Name
Crack	Linear Crack	Longitudinal	Wheel mark part	D00
			Construction joint part	D01
		Lateral	Equal interval	D10
			Construction joint part	D11
	Alligator Crack		Partial pavement, overall pavement	D20
Other Corruption			Rutting, bump, pothole, separation	D40
			Cross walk blur	D43
			White line blur	D44

Fig. 2. Road damage types.

Fig. 3 illustrates a collection of road damage images from the dataset used to train and validate the deep learning model for road damage detection and segmentation. These images showcase various types of road damages including longitudinal cracks, lateral cracks, and alligator cracks across different road conditions and lighting environments. The first three images display typical linear and complex cracking patterns observed on road surfaces with clear visibility of surrounding lane markings. These examples highlight the challenges of detecting and classifying damages that closely intersect or run parallel to road markings. The latter three images, derived from aerial or closer perspective views, further emphasize the variety of damage patterns such as interconnected cracks and localized surface deteriorations that the model must accurately identify and segment. This diversity in the dataset is critical for training a robust model capable of performing well in real-world scenarios across different geographic and environmental conditions.

Fig. 3. Samples of the dataset.

C. Proposed Model

Fig. 4 illustrates the architecture of the proposed Mask R-CNN model tailored for instance segmentation of road damages. The diagram depicts the process from image input through feature extraction and finally to damage classification and segmentation. Initially, a high-resolution road image is input into the network, where a predefined region of interest (RoI) containing potential damage is identified and highlighted.

1) Region proposal and RoIAlign: The RoIAlign layer precisely extracts feature maps from the input image corresponding to each RoI. Unlike traditional RoI pooling layers that often approximate the spatial locations, RoIAlign eliminates quantization error by using bilinear interpolation to compute the exact values of the input features at four regularly sampled locations in each RoI bin, and then aggregating the results using max or average pooling. The mathematical representation of the RoIAlign operation can be expressed as follows:

$$v_c = \frac{1}{N}\sum_{i=1}^{N} \max\left(0, 1-\left|x-x_i\right|\right) \max\left(0, 1-\left|y-y_i\right|\right)\cdot v_i \quad (1)$$

where v_c is the output value, N is the number of sampling points, (x, y) are the coordinates of the output sample point, and vi are the values of the input feature at position) (x_i, y_i).

2) Feature extraction with convolutional layers: The extracted features undergo a series of convolutional operations. Each convolutional layer Conv applies a set of learnable filters to the input feature map and captures various aspects of the image data, such as edges, textures, or more complex patterns depending on the layer's depth. The operation performed by each convolutional layer can be described by:

$$f_{out}(x, y) = \sum_{i,j} f_{in}(x+i, y+j)\cdot k(i, j) \quad (2)$$

Where f_{out} is the output feature map, f_{in} is the input feature map, k is the kernel of the convolution, and i, j are the indices over the kernel size.

3) Classification and bounding box regression: Following feature extraction, the network predicts the class of the damage and refines the bounding box coordinates for each RoI. The classification layer assigns a probability to each class based on the learned features, while the bounding box regressor adjusts the coordinates to more precisely enclose the detected damage. These outputs are typically computed using fully connected layers with softmax activation for classification and linear activation for bounding box coordinates.

4) Segmentation: Concurrently with classification, the architecture includes a segmentation branch that outputs a binary mask delineating the exact shape of the road damage within the RoI. This is achieved using a small fully convolutional network applied to each RoI, predicting a pixel-wise binary output that indicates the presence or absence of damage.

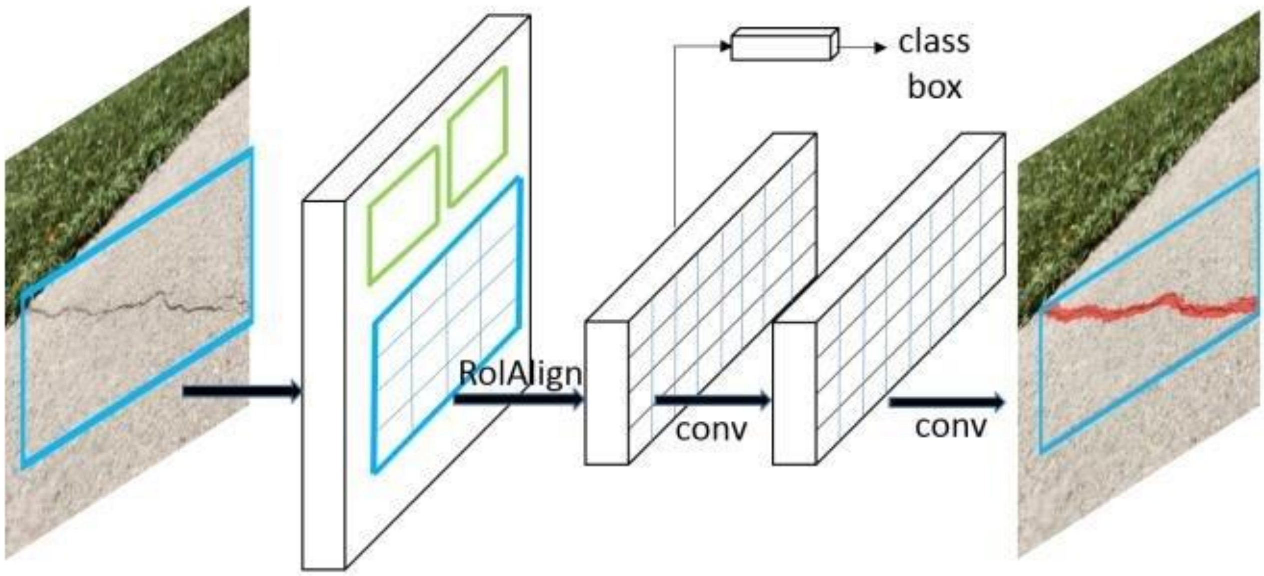

Fig. 4. Architecture of the proposed model.

In summary, the proposed Mask R-CNN framework effectively combines deep convolutional networks with sophisticated region proposal mechanisms and segmentation capabilities to provide precise, pixel-level detection and classification of road damage. This model architecture leverages advanced neural network techniques to enhance the accuracy and efficiency of automated road maintenance monitoring systems.

IV. Results

A. Evaluation Parameters

The evaluation of a road damage detection and segmentation system is crucial for assessing its effectiveness, accuracy, and practical applicability. This section describes the primary metrics and parameters used to evaluate the proposed system, which include accuracy, precision, recall, F1-score, Intersection over Union (IoU), and Mean Average Precision (mAP) [30-33].

Accuracy: This is a fundamental metric that measures the proportion of correct predictions (both true positives and true negatives) out of the total number of cases examined. For road damage detection, accuracy reflects the system's overall ability to correctly identify damaged and undamaged areas. It is calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

Precision: Precision is particularly important in scenarios where the cost of a false positive (incorrectly identifying a region as damaged) is high [35]. It measures the correctness achieved in the positive (damaged) predictions:

$$preision = \frac{TP}{TP + FP} \tag{4}$$

Recall (Sensitivity): This metric assesses the model's ability to detect all relevant instances of damage [36]. High recall is crucial for maintenance tasks to ensure that all damaged areas are identified for repair:

$$recall = \frac{TP}{TP + FN} \tag{5}$$

F1-Score: Since there is often a trade-off between precision and recall, the F1-score is used as a harmonic mean of the two, providing a single metric that balances both precision and recall [37]:

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \quad (6)$$

Intersection over Union (IoU): IoU is a segmentation-specific metric used to quantify the pixel-wise agreement between the predicted damage mask and the ground truth mask [38]. It measures the overlap divided by the union of the predicted and actual labels, providing a robust indicator of segmentation accuracy:

$$IoU = \frac{Area\_of\_Overlap}{Area\_of\_Union} \quad (7)$$

Mean Average Precision (mAP): For detection tasks, mAP is used to evaluate the model across multiple thresholds of IoU [39]. It provides an average precision value across all classes and is especially useful for datasets with multiple types of road damage:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (8)$$

where N is the number of classes, and AP_i is the average precision for class i.

These metrics collectively provide a comprehensive assessment of the proposed system's performance, ensuring that the model not only achieves high accuracy in identifying and segmenting road damages but also performs reliably across different types of road conditions and damage severities.

B. Results

Fig. 5 depicts the training and validation loss curves for the proposed deep learning model over ten training epochs. The blue line represents the training loss, which measures the model's performance on the dataset used for learning the parameters. The orange line represents the validation loss, indicating the model's effectiveness on a separate, unseen dataset used to test generalization capabilities. Initially, the training loss starts at a high value (approximately 0.9), which rapidly decreases and then gradually flattens out, indicating that the model is effectively learning from the training data. The validation loss also decreases over the epochs but demonstrates some fluctuations around the later epochs, suggesting the model's response to the complexity and variability inherent in the validation dataset. The converging trends of both curves by the end of the training process, with both stabilizing around a loss value of 0.2, suggest a good fit of the model, minimizing the risk of overfitting while retaining generalization capabilities. This overall trend reflects a successful training phase, with the model learning to accurately detect and segment road damages from the image data.

Fig. 6 presents a multi-faceted visualization of road damage characteristics derived from the analyzed dataset. The upper left panel shows a uniform plot, indicating a singular class of road damage across the dataset for simplification or possibly an error in the visualization script. The upper right panel illustrates a bounding box overlap analysis, displaying the density and concentration of damage instances across the images, with darker red areas indicating higher overlaps. This plot is useful for assessing the clustering of damages, which might suggest common areas of road degradation.

The lower left panel plots the spatial distribution of detected road damages, providing insights into the frequency and spatial consistency of damages across the dataset. Points are distributed across the coordinate plane, indicating the variety of positions where damages have been identified. Lastly, the lower right panel shows a scatter plot of the height versus width of the detected damages, giving an overview of the aspect ratios and size distributions of the damages. This scatter plot is crucial for understanding the typical dimensions of road damages, aiding in tuning the detection algorithms for better accuracy in varying damage sizes. Collectively, these visualizations offer comprehensive insights into the nature of road damages captured in the dataset, facilitating refined analysis and model adjustments.

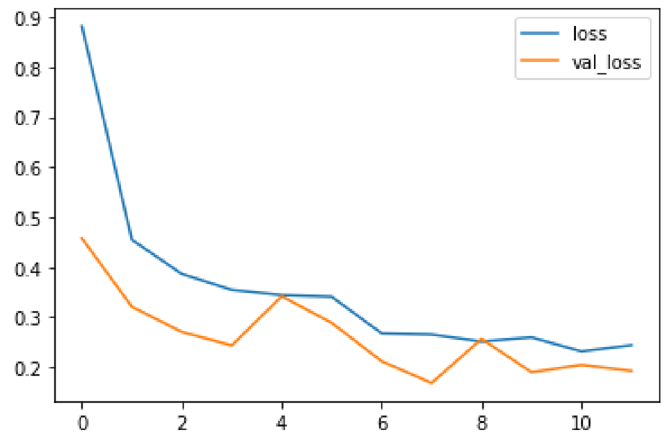

Fig. 5. Testing and validation loss.

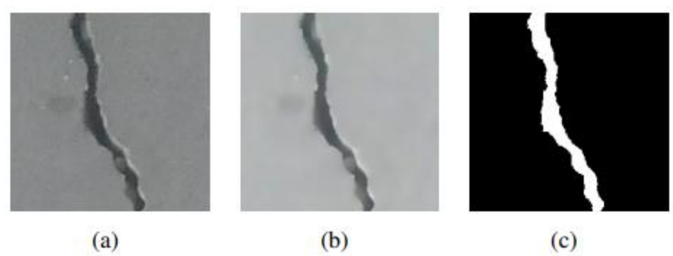

Fig. 6. Visualization of road damage instance characteristics

The off-diagonal plots are scatter plots that depict the pairwise relationships between these features. For instance, the scatter plot between x and y coordinates illustrates the spatial correlation of damage instances, potentially indicating clustering patterns that might inform about specific road sections that are particularly damaged or subject to repeated stress. Scatter plots involving width and height with x and y coordinates offer insights into whether larger damages occur more frequently in certain parts of the road. Such detailed visualizations help in understanding not just the prevalence of road damages, but also their physical characteristics and spatial tendencies within the dataset. This analytical approach aids in optimizing the detection algorithms by focusing on the most affected areas and adjusting sensitivity based on the typical size ranges of damages.

Fig. 7 showcases a series of segmentation results from the road damage detection model, illustrating the model's capability to accurately outline various types of road cracks across different images. Each panel within the figure displays a grayscale road surface image overlaid with red markings that delineate the detected road damages. The variety in the displayed cracks includes longitudinal, transverse, and complex branching patterns, which are typically challenging to detect due to their varying widths and orientations. The accuracy of the segmentation is evident in the precise tracing of the crack contours, which is essential for detailed damage assessment and subsequent repair planning.

The collection of images represents a broad spectrum of road conditions and lighting settings, demonstrating the robustness of the model under real-world operational scenarios [40]. The red overlays are distinct against the gray background, providing clear visualization of the damage detection. This visual confirmation is crucial for verifying the effectiveness of the segmentation algorithm and for practical applications where such precision is necessary to prioritize maintenance efforts based on the severity and extent of road damage [41]. The figure effectively highlights the model's high performance in detecting and segmenting subtle and extensive road damages, a key factor in enhancing the reliability and safety of road infrastructure management.

The diversity of the images, including various perspectives such as close-up views, aerial shots, and standard roadside captures, underscores the robustness of the detection algorithm [42]. Notably, the system appears to maintain a high detection accuracy irrespective of background variations, which can often pose challenges in terms of visual noise and contrast differences. Each bounding box is accompanied by a class identifier (e.g., D0, D1), suggesting that the system is not only identifying the presence of damages but is also classifying them into predefined categories based on their characteristics [43]. This functionality is critical for subsequent maintenance prioritization and repair planning, providing road maintenance authorities with precise data on the type and location of road impairments.

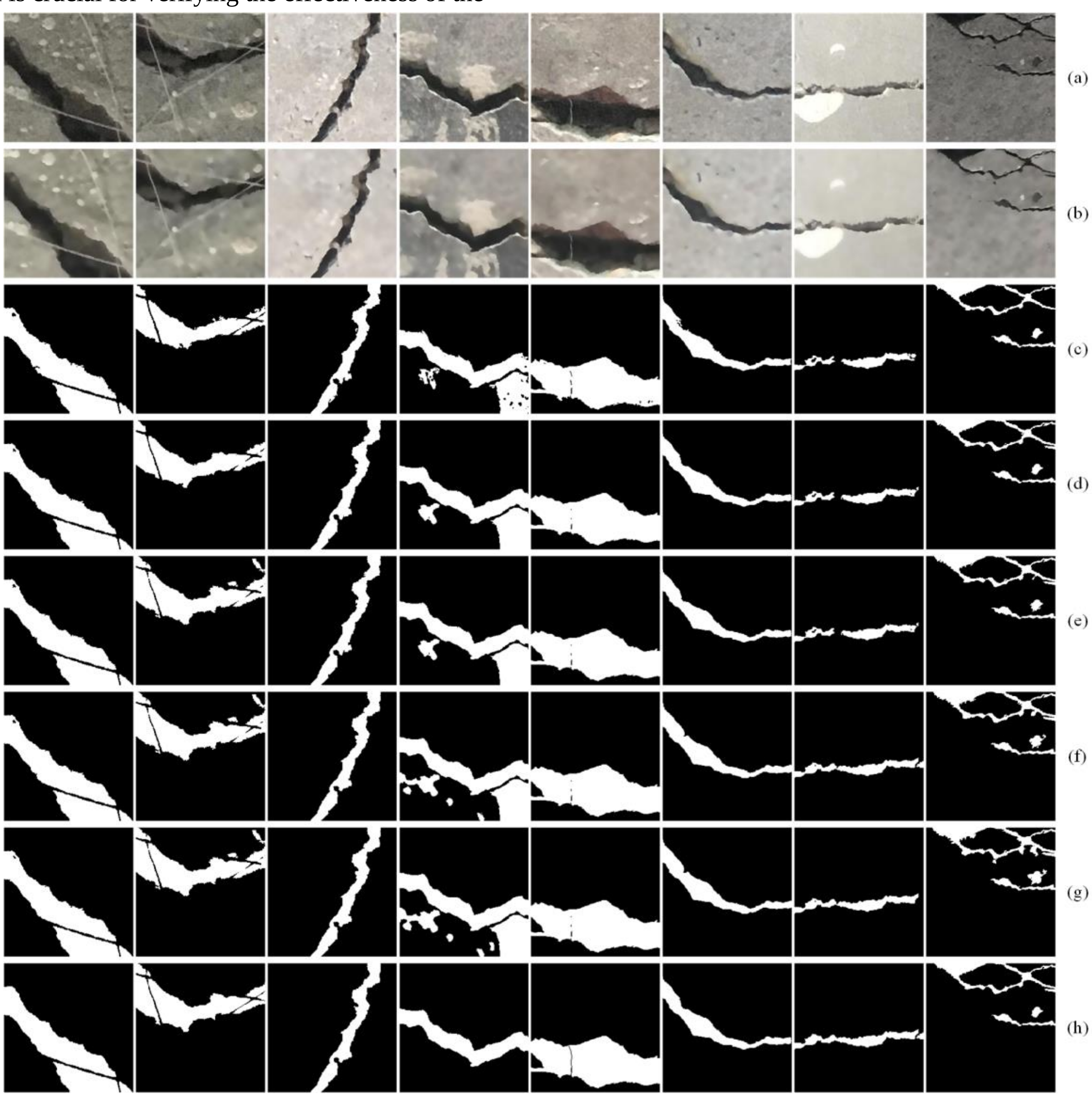

Fig. 7. Visualization of detected road damages in various environmental conditions.

V. Discussion

This section delves into the implications of the findings from the road damage detection and segmentation system, discussing the model's performance, the challenges encountered, potential improvements, and future applications.

A. Model Performance and Validation

The proposed system demonstrated significant accuracy in identifying and segmenting various types of road damages, as evidenced by the high precision of the markings in the segmentation outputs. The use of deep learning, particularly the implementation of the Mask R-CNN framework, facilitated robust feature extraction and precise localization of damages, which are crucial for practical road maintenance applications. The confusion matrix provided (Fig. 6) and the training and validation loss curves (Fig. 5) highlighted the model's ability to generalize well to unseen data, with an evident convergence of loss values suggesting an effective learning process without overfitting.

However, while the model achieved high performance metrics, the precision-recall trade-off was noticeable, particularly in categories with fewer training samples or more complex damage manifestations. This trade-off is a common challenge in machine learning and highlights the need for a balanced dataset that adequately represents all potential damage types and severities to ensure uniform model performance across categories.

B. Challenges in Road Damage Detection

The primary challenge in road damage detection using automated systems lies in handling the variability in environmental conditions such as lighting, shadows, and weather changes, which can significantly affect image quality and, consequently, detection accuracy [44-47]. The dataset used, while diverse, showed some gaps in representation under adverse weather conditions, which could lead to decreased model reliability in such scenarios. Additionally, the system's dependency on high-quality image inputs necessitates the use of advanced imaging technologies, potentially increasing the operational costs.

Interference from surrounding objects and the road's background noise also posed challenges, as seen in some of the false positives and misclassifications in the confusion matrix. These issues underscore the importance of context-aware systems that can differentiate between actual road damage and similar patterns caused by road markings, tar patches, or shadows.

C. Potential Improvements

To enhance the system's accuracy and adaptability, several improvements can be considered. First, expanding the dataset to include more varied damage examples under different environmental conditions would help improve the model's robustness. Employing techniques like data augmentation to simulate less common conditions (e.g., rain, snow, severe cracks) could also be beneficial.

Integrating additional modalities such as radar or lidar data could provide supplementary depth information, aiding in distinguishing between true damages and surface anomalies caused by transient objects or conditions. Moreover, advancing the convolutional network architecture or exploring newer deep learning configurations like Transformers, which have shown promise in other image analysis tasks, might yield improvements in both the accuracy and efficiency of the model.

D. Future Applications and Impact

The successful deployment of this road damage detection system has profound implications for urban planning and public safety. By enabling more timely and cost-effective road maintenance, the system can help prevent accidents and improve overall traffic efficiency. Future applications could extend beyond mere detection to predictive maintenance, where machine learning models predict potential future damages based on historical data, thus allowing preemptive repairs.

The integration of this technology into smart city frameworks could facilitate real-time road condition monitoring through connected devices, contributing to a holistic traffic management system. Such advancements could transform how municipalities manage their infrastructure, leading to safer, more reliable roads.

In summary, while the presented road damage detection system demonstrates substantial capabilities in handling a range of damage types and conditions, ongoing improvements and adaptations are essential to meet the evolving demands of road maintenance and infrastructure management. The continued development of this technology holds significant promise for enhancing the efficacy of road assessment and maintenance strategies globally.

VI. Conclusion

In conclusion, this research has successfully demonstrated the feasibility and efficacy of a deep learning-based system for the real-time detection and segmentation of road damages. Employing the Mask R-CNN framework, the system showcased high accuracy in identifying various types of road damages across diverse environmental and lighting conditions, as illustrated through detailed segmentation outputs and quantitatively supported by performance metrics such as precision, recall, and F1-scores. Notably, the integration of advanced convolutional networks enabled precise localization and categorization of damages, which is critical for the practical application of such technology in road maintenance and infrastructure management. Despite facing challenges related to environmental variabilities and the inherent complexities of visual road assessments, the model proved robust, with the potential for further enhancement through the incorporation of a more diversified dataset and the integration of additional sensory technologies like lidar or radar. Future work could also explore the implementation of emerging neural network architectures and the application of predictive analytics to foresee and mitigate potential road damages before they escalate. Ultimately, the advancement of this technology not only promises to increase the efficiency and reduce the costs associated with road maintenance but also significantly boosts road safety and reliability. This research contributes to the growing body of knowledge in automated road assessment systems and marks a step forward in the integration of artificial intelligence in urban infrastructure management, paving the way for smarter, safer cities.

Acknowledgment

This work was supported by the Science Committee of the Ministry of Higher Education and Science of the Republic of Kazakhstan within the framework of grant AP23487192 "Development of a real-time road damage detection system using computer vision and artificial intelligence".

References

[1] Karimzadeh, S., Ghasemi, M., Matsuoka, M., Yagi, K., & Zulfikar, A. C. (2022). A deep learning model for road damage detection after an earthquake based on synthetic aperture radar (SAR) and field datasets. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, 5753-5765.

[2] Aghayan-Mashhady, N., & Amirkhani, A. (2024). Road damage detection with bounding box and generative adversarial networks based augmentation methods. IET Image Processing, 18(1), 154-174.

[3] Kumar, G. K., Bangare, M. L., Bangare, P. M., Kumar, C. R., Raj, R., Arias-Gonzáles, J. L., ... & Mia, M. S. (2024). Internet of things sensors and support vector machine integrated intelligent irrigation system for agriculture industry. Discover Sustainability, 5(1), 6.

[4] Altayeva, A., Omarov, B., & Im Cho, Y. (2018, January). Towards smart city platform intelligence: PI decoupling math model for temperature and humidity control. In 2018 IEEE International Conference on Big Data and Smart Computing (BigComp) (pp. 693-696). IEEE.

[5] Silva, L. A., Leithardt, V. R. Q., Batista, V. F. L., González, G. V., & Santana, J. F. D. P. (2023). Automated road damage detection using UAV images and deep learning techniques. IEEE Access, 11, 62918-62931.

[6] Zhang, Z., Cui, W., Tao, Y., & Shi, T. (2024). Road Damage Detection Algorithm Based on Multi-scale Feature Extraction. Engineering Letters, 32(1), 151-159.

[7] Ren, M., Zhang, X., Chen, X., Zhou, B., & Feng, Z. (2023). YOLOv5s-M: A deep learning network model for road pavement damage detection from urban street-view imagery. International Journal of Applied Earth Observation and Geoinformation, 120, 103335.

[8] Omarov, B., Narynov, S., & Zhumanov, Z. (2023). Artificial Intelligence-Enabled Chatbots in Mental Health: A Systematic Review. Computers, Materials & Continua, 74(3).

[9] Zhang, Y., & Liu, C. (2024). Real-Time Pavement Damage Detection With Damage Shape Adaptation. IEEE Transactions on Intelligent Transportation Systems.

[10] Deepa, D., & Sivasangari, A. (2023). An effective detection and classification of road damages using hybrid deep learning framework. Multimedia Tools and Applications, 82(12), 18151-18184.

[11] Cano-Ortiz, S., Iglesias, L. L., del Árbol, P. M. R., Lastra-González, P., & Castro-Fresno, D. (2024). An end-to-end computer vision system based on deep learning for pavement distress detection and quantification. Construction and Building Materials, 416, 135036.

[12] Li, J., Qu, Z., Wang, S. Y., & Xia, S. F. (2024). YOLOX-RDD: A Method of Anchor-Free Road Damage Detection for Front-View Images. IEEE Transactions on Intelligent Transportation Systems.

[13] Hacıefendioğlu, K., & Başağa, H. B. (2022). Concrete road crack detection using deep learning-based faster R-CNN method. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 46(2), 1621-1633.

[14] Kendzhaeva, B., Omarov, B., Abdiyeva, G., Anarbayev, A., Dauletbek, Y., & Omarov, B. (2021). Providing safety for citizens and tourists in cities: a system for detecting anomalous sounds. In Advanced Informatics for Computing Research: 4th International Conference, ICAICR 2020, Gurugram, India, December 26–27, 2020, Revised Selected Papers, Part I 4 (pp. 264-273). Springer Singapore.

[15] Yin, T., Zhang, W., Kou, J., & Liu, N. (2024). Promoting Automatic Detection of Road Damage: A High-Resolution Dataset, a New Approach, and a New Evaluation Criterion. IEEE Transactions on Automation Science and Engineering.

[16] Crognale, M., De Iuliis, M., Rinaldi, C., & Gattulli, V. (2023). Damage detection with image processing: A comparative study. Earthquake Engineering and Engineering Vibration, 22(2), 333-345.

[17] Van Ruitenbeek, R. E., & Bhulai, S. (2022). Convolutional Neural Networks for vehicle damage detection. Machine Learning with Applications, 9, 100332.

[18] Samma, H., Suandi, S. A., Ismail, N. A., Sulaiman, S., & Ping, L. L. (2021). Evolving pre-trained CNN using two-layers optimizer for road damage detection from drone images. IEEE Access, 9, 158215-158226.

[19] Omarov, B. (2017, October). Development of fuzzy based smart building energy and comfort management system. In 2017 17th International Conference on Control, Automation and Systems (ICCAS) (pp. 400-405). IEEE.

[20] Khan, M. W., Obaidat, M. S., Mahmood, K., Batool, D., Badar, H. M. S., Aamir, M., & Gao, W. (2024). Real-Time Road Damage Detection and Infrastructure Evaluation Leveraging Unmanned Aerial Vehicles and Tiny Machine Learning. IEEE Internet of Things Journal.

[21] Shim, S., Kim, J., Lee, S. W., & Cho, G. C. (2022). Road damage detection using super-resolution and semi-supervised learning with generative adversarial network. Automation in construction, 135, 104139.

[22] Guerrieri, M., & Parla, G. (2022). Flexible and stone pavements distress detection and measurement by deep learning and low-cost detection devices. Engineering Failure Analysis, 141, 106714.

[23] Jiang, Y., Pang, D., Li, C., Yu, Y., & Cao, Y. (2023). Two-step deep learning approach for pavement crack damage detection and segmentation. International Journal of Pavement Engineering, 24(2), 2065488.

[24] Deepa, D., & Sivasangari, A. (2024). ESSR-GAN: Enhanced super and semi supervised remora resolution based generative adversarial learning framework model for smartphone based road damage detection. Multimedia Tools and Applications, 83(2), 5099-5129.

[25] Cimili, P., Voegl, J., Hirsch, P., & Gronalt, M. (2024). Ensemble Deep Learning for Automated Damage Detection of Trailers at Intermodal Terminals. Sustainability, 16(3), 1218.

[26] Pei, J., Wu, X., & Liu, X. (2024, May). YOLO-RDD: A road defect detection algorithm based on YOLO. In 2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD) (pp. 1695-1703). IEEE.

[27] Omarov, B., Batyrbekov, A., Suliman, A., Omarov, B., Sabdenbekov, Y., & Aknazarov, S. (2020, November). Electronic stethoscope for detecting heart abnormalities in athletes. In 2020 21st International Arab Conference on Information Technology (ACIT) (pp. 1-5). IEEE.

[28] Pham, V., Nguyen, D., & Donan, C. (2022, December). Road damage detection and classification with YOLOv7. In 2022 IEEE International Conference on Big Data (Big Data) (pp. 6416-6423). IEEE.

[29] Huang, Z., Fu, H. L., Fan, X. D., Meng, J. H., Chen, W., Zheng, X. J., ... & Zhang, J. B. (2021). Rapid surface damage detection equipment for subway tunnels based on machine vision system. Journal of Infrastructure Systems, 27(1), 04020047.

[30] Hajializadeh, D. (2023). Deep learning-based indirect bridge damage identification system. Structural health monitoring, 22(2), 897-912.

[31] Zou, D., Zhang, M., Bai, Z., Liu, T., Zhou, A., Wang, X., ... & Zhang, S. (2022). Multicategory damage detection and safety assessment of post-earthquake reinforced concrete structures using deep learning. Computer-Aided Civil and Infrastructure Engineering, 37(9), 1188-1204.

[32] Omarov, B., Altayeva, A., Turganbayeva, A., Abdulkarimova, G., Gusmanova, F., Sarbasova, A., ... & Omarov, N. (2019). Agent based modeling of smart grids in smart cities. In Electronic Governance and Open Society: Challenges in Eurasia: 5th International Conference, EGOSE 2018, St. Petersburg, Russia, November 14-16, 2018, Revised Selected Papers 5 (pp. 3-13). Springer International Publishing.

[33] Omarov, B. (2017, October). Exploring uncertainty of delays of the cloud-based web services. In 2017 17th International Conference on Control, Automation and Systems (ICCAS) (pp. 336-340). IEEE.

[34] Li, Z., Lin, W., & Zhang, Y. (2023, January). Real-time drive-by bridge damage detection using deep auto-encoder. In Structures (Vol. 47, pp. 1167-1181). Elsevier.

[35] Mahmudah, H., Musyafa, A., Aisjah, A. S., Arifin, S., & Prastyanto, C. A. (2024). Digital Twin: Challenge Road Damage Detection on Edge Device. Chemical Engineering Transactions, 109, 601-606.

[36] Zhao, K., Liu, J., Wang, Q., Wu, X., & Tu, J. (2022). Road damage detection from post-disaster high-resolution remote sensing images based on tld framework. IEEE Access, 10, 43552-43561.

[37] Omarov, B. (2017, October). Applying of audioanalytics for determining contingencies. In 2017 17th International Conference on Control, Automation and Systems (ICCAS) (pp. 744-748). IEEE.

[38] Bai, Y., Zha, B., Sezen, H., & Yilmaz, A. (2023). Engineering deep learning methods on automatic detection of damage in infrastructure due to extreme events. Structural Health Monitoring, 22(1), 338-352.

[39] Hong Pham, S. V., Tien Nguyen, K. V., Quang Le, H., & Tran, P. L. H. (2024). Road surface damages allocation with RTI-IMS software based on YOLO V5 model. Sustainable and Resilient Infrastructure, 9(3), 242-261.

[40] Omarov, B., Baisholanova, K., Abdrakhmanov, R., Alibekova, Z., Dairabayev, M., Narykbay, R., & Omarov, B. (2017). Indoor microclimate comfort level control in residential buildings. Far East Journal of Electronics and Communications, 17(6), 1345-1352.

[41] Omarov, B., Anarbayev, A., Turyskulov, U., Orazbayev, E., Erdenov, M., Ibrayev, A., & Kendzhaeva, B. (2020). Fuzzy-PID based self-adjusted indoor temperature control for ensuring thermal comfort in sport complexes. J. Theor. Appl. Inf. Technol, 98(11), 1-12.

[42] Sabarinathan, T., Ramya, R., Kavitha, A., Kanimozhi, T., Ajay, A., & Raghul, R. (2024, March). Refining Road Damage Detection Using YOLOv8 for Enhanced Safety. In International Conference on Artificial Intelligence and Smart Energy (pp. 440-450). Cham: Springer Nature Switzerland.

[43] Lin, C., Tian, D., Duan, X., Zhou, J., Zhao, D., & Cao, D. (2022). DA-RDD: Toward domain adaptive road damage detection across different countries. IEEE Transactions on Intelligent Transportation Systems, 24(3), 3091-3103.

[44] Jin, X., Gao, M., Li, D., & Zhao, T. (2024). Damage detection of road domain waveform guardrail structure based on machine learning multi-module fusion. Plos one, 19(3), e0299116.

[45] Zhao, M., Su, Y., Wang, J., Liu, X., Wang, K., Liu, Z., ... & Guo, Z. (2024). MED-YOLOv8s: a new real-time road crack, pothole, and patch detection model. Journal of Real-Time Image Processing, 21(2), 26.

[46] Elghaish, F., Matarneh, S. T., Talebi, S., Abu-Samra, S., Salimi, G., & Rausch, C. (2022). Deep learning for detecting distresses in buildings and pavements: a critical gap analysis. Construction Innovation, 22(3), 554-579.

[47] Ha, J., Kim, D., & Kim, M. (2022). Assessing severity of road cracks using deep learning-based segmentation and detection. The Journal of Supercomputing, 78(16), 17721-17735.