Пут и саобраћај
Journal of Road and Traffic Engineering
Journal homepage: www.putisaobracaj.rs

Automatic Detection of Pavement Distresses Using Convolutional Neural Networks

Matija Rašović[a*], Nenad Brodić[a], Marko Orešković[a]

[a] *University of Belgrade, Faculty of Civil Engineering. Bulevar kralja Aleksandra 73, Belgrade, Serbia*

ARTICLE INFO

DOI: 10.31075/PIS.70.04.03
Professional paper
Received: 01.12.2024
Accepted: 17.12.2024
Corresponding author:
matija.rasovic2@gmail.com

ORCID ID
Matija Rašović: N.A.
Nenad Brodić: 0000-0002-8757-0991
Marko Orešković: 0000-0001-8978-0354

Keywords
Pavement Distresses
Image analysis
Convolutional Neural Networks
YOLO

ABSTRACT

This paper presents a method for detecting pavement distresses using convolutional neural network architectures. The dataset used as training data was Road Damage Dataset 2022 (RDD2022) which has been adapted for the YOLO (You Only Look Once) architecture. In this study, bounding box detection and classification techniques were employed with several variants of YOLO and RT-DETR (Real-Time Detection Transformer) architecture. The global quality metric was mean Average Precision at 50% (mAP-50%), which varied between 64% and 70.4% depending on the model applied. Inference of the fine-tuned model was conducted on images excluded from the training dataset. All images subjected to detection were later geo-tagged and plotted on OpenStreetMap (OSM). Although this study showed promising results, it should be mentioned that more annotated data is required to achieve more precise results.

1. Introduction

The increasing demand for safe and efficient transportation infrastructure requires timely detection and road maintenance, where a key principle is that current investment reduces future costs. In other words, the worse the condition becomes, the more substantial these costs grow.

Pavement distresses are typically detected using one of three approaches: manual, semi-automatic, or fully automated. The manual method is a traditional method that represents the direct detection of pavement distresses. It relies on visual inspections carried out by trained personnel. Inspectors either walk along the road or drive slowly, documenting visible distresses such as cracks, potholes, bearing capacity issues, etc. This method provides detailed and accurate assessments based on the inspectors' expertise, but it is labor-intensive, time-consuming and potentially hazardous due to exposure to traffic risks. Additionally, results can vary due to inspector subjectivity, making it impractical for large-scale inspections. Semi-automated methods use certain technology to enhance the efficiency and safety of inspections. This approach involves capturing high-resolution photographs or videos with cameras mounted on vehicles, while analysis is supposed to be performed later through post-processing. Although this method reduces inspectors' exposure to traffic risks and speeds up data collection, it still requires significant human intervention for distresses analysis and identification. Accuracy depends on image quality and analyst skill, whereas the time gap between data collection and analysis can delay necessary maintenance responses. Automated methods use advanced sensors, machine learning algorithms, or image processing software for autonomous distress detection and classification.

These systems enable rapid, consistent, and objective road condition assessments with minimal human intervention, greatly improving safety and efficiency. However, high initial costs and the need for technical expertise for development, implementation, and maintenance pose challenges. System efficiency can be reduced by environmental factors, such as lighting and weather conditions, which can impact sensor performance. A fully automated approach is often paired with vehicles equipped with sophisticated and expensive sensors, such as lasers, road profilers, LiDAR, etc. Such vehicles are usually very expensive, and the final price depends mostly on the sensors' setup used in the system. The development of high-resolution digital cameras has made it possible to conduct efficient and cost-effective road inspections using only a medium- to low-cost camera mounted on a vehicle.

Advances in machine learning, particularly Convolutional Neural Networks (CNNs), offer a promising opportunity to automate detection of pavement distresses. This approach enables the identification of various types of road defects—such as cracks, potholes, and other surface distresses—from images captured by smartphones, vehicle-mounted cameras, or drones. Both the accuracy and speed of distresses assessment can be significantly improved by using CNNs, enabling proactive maintenance strategies that ultimately enhance roadway safety and durability. The most commonly used models, such as You Only Look Once (YOLO) [1], Single Shot MultiBox Detector (SSD) [2], and Faster R-CNN [3], are widely utilized for real-time applications due to their ability to quickly and reliably detect objects in images. Accordingly, this paper explores the methodologies, challenges, and implications of implementing CNN-based algorithms for automatic pavement distresses detection, to contribute to smarter infrastructure management.

2. Related Work

There is a large number of studies focused on detecting pavement distresses, using different approaches: sensor-based techniques [4-7], 3D reconstruction – laser-based [8–10] and stereo vision-based [11–14], image processing techniques [15-19], and model-based - machine learning (ML) and deep learning (DL) methods [20-28].

Sensor-based techniques use vibration sensors to detect potholes. These techniques can yield false positives (detecting road joints as potholes) and false negatives (missing potholes in lane centers). These techniques require expensive setups, high computational demands, and potential inaccuracies due to camera misalignment that could impact detection accuracy.

Traditional image processing techniques offer high accuracy but require manual feature extraction and parameter adjustments for various road conditions. Advances in image processing and low-cost camera systems have driven the development of model-based pothole detection methods. Traditional ML models achieve significant accuracy but demand high computational power and need experts for feature extraction. In contrast, DL techniques, utilizing CNNs, automate both feature extraction and classification processes, enhancing efficiency and accuracy.

Among all the mentioned approaches for automatic detection, ML and DL techniques nowadays stand out in terms of accuracy and execution speed. Applying ML and DL techniques to develop trained models is a growing trend for detecting potholes, cracks and other types of roadway degradation in 2D digital images.

3. Dataset Used for Training

The dataset used in this project is the publicly available *Road Damage Dataset* (RDD202) [29], designed for object detection tasks related to road surface distresses (Fig. 1). This dataset includes a total of 47,420 road images from diverse geographic locations — China, Japan, India, Norway, the United States, and the Czech Republic — and covers around 55,000 instances of four primary types of distress: longitudinal cracks (D00), transverse cracks (D10), alligator cracks (D20) and potholes (D40). The diversity in origin and technical characteristics, such as image resolution and perspective, enriches the dataset and enhances the model's ability to generalize across various road infrastructures and conditions.

RDD2022 is structured to support ML tasks for road infrastructure analysis. Each image is paired with annotations in the PASCAL VOC format, which provides structured information for training object detection models. To adapt the dataset for the YOLO model architecture used in this work, annotation files were converted to YOLO format using a Python-based conversion script. This conversion ensured compatibility with the *Ultralytics* platform [30] and standardized the bounding box coordinates for consistent data input.

Images in the dataset offer a variety of perspectives, different image resolutions (Figure 1 - Figure 4) and various environmental contexts which contribute to the robustness of the model. For example, Japanese and Czech images were captured via vehicle-mounted smartphone cameras, providing image resolution of 600x600 pixels, whereas Norway's images were taken using high-resolution cameras in specialized vehicles, yielding 3650x2044-pixel panoramic views. Additionally, images from China include aerial shots taken by Unmanned Aerial Vehicles (UAVs) equipped with stabilizing gimbal, providing vertical perspectives on road surfaces. The geographic and technical diversity in the dataset helps the robustness of the model to perform reliably across different countries and infrastructure types.

Additionally, the independent validation dataset for this work includes 2,052 images of a road in Belgrade, collected using the iSTAR Pulsar+ 360° camera system mounted on a moving vehicle. This data collection was conducted at the end of July 2023 on Ustanička and Ozrenska streets. The vehicle was driven at a maximum speed of 30 km/h, with camera capturing images every 3.5 to 4 meters, with a frame rate of 7fps.This separate dataset provides a realistic, high-resolution testing environment to evaluate the model's accuracy in detecting road distress without prior training data exposure. By testing on local, unannotated images, the project highlights the model's generalization capabilities and identifies areas where additional locally collected data could improve detection precision for road distress.

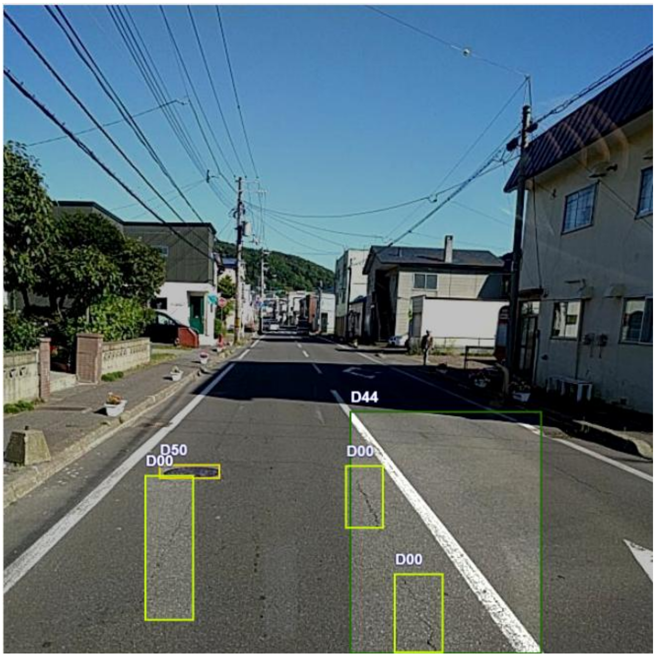

Figure 1. Sample of the Japanese subset with annotations
Source: Dataset RDD2022 used for model training

To illustrate the diversity and relevance of the dataset, the following images were shown in this work: a sample from the Czech subset, demonstrating typical smartphone-collected road images at standard resolutions, a high-resolution panoramic image from Norway's subset, highlighting the detail captured by specialized cameras, and a UAV-collected drone image from China, emphasizing the dataset's technical range and utility for road distress detection.

Figure 2. Samples of "China Drone" and the "USA" subsets
Source: Dataset used in training RDD2022

Figure 3. Samples of "Japan" and the "Czech" subsets
Source: Dataset used in training RDD2022

Figure 4. Samples of "India" and the "China_motorBike" subsets
Source: Dataset used in training RDD2022

4. Data Preparation

Before training, Exploratory Data Analysis (EDA) was conducted to identify potential issues and characteristics of the dataset. This analysis involved visualizing the distribution of different pavement distresses and image resolutions, which were crucial for understanding the structure of the data and fine-tuning the model. The EDA revealed significant variability in image resolution and varying perspectives in the road images, underscoring the normalization importance and parameter adjustment to enable the model to handle this heterogeneity.

The experimental component of this research was conducted using the *Python* programming language, leveraging multiple integrated development environments (IDEs) to support the distinct phases of the project. The workflow, combined coding, testing, and data analysis. *JupyterLab* served as a primary environment for exploratory tasks, enabling interactive code execution and documentation within its cell-based structure. This setup facilitated iterative development and analysis, making it particularly suitable for EDA and prototyping. *PyCharm* was employed during the data formatting phase, where its debugging capabilities proved essential for managing complex scripts and resolving semantic errors. For computationally intensive tasks such as model training, *Google Colab Pro+* was used, providing access to high-performance Graphics Processing Units (GPUs) like NVIDIA A100 and T4.

These cloud resources significantly reduced training time and alleviated the limitations having by using local hardware. Figure 5 shows the flowchart of the technologies used, with the dotted lines representing the knowledge accumulated in said processes rather than actual transfer of data itself.

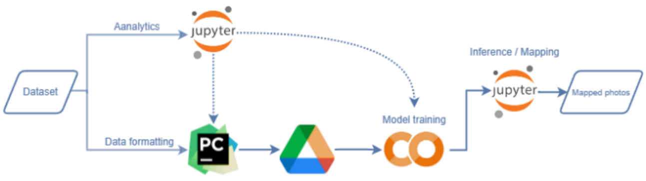

Figure 5. Flowchart of the projects data pipeline

Data preparation constituted a significant portion of the experimental work and was pivotal for the project's success. The raw dataset, sourced from public repositories, consisted of annotated pavement distress images in PASCAL VOC XML format, which were incompatible with the YOLO models used in this study. To address this, a custom Python script was developed to automate the conversion of annotations into YOLO's text-based format. This process ensured compatibility while preserving the dataset's integrity. In addition to annotated images, a subset of unannotated photographs was reserved for validation purposes. These images, collected from real-world conditions in Belgrade, served to test the models' generalization capabilities and robustness when exposed to novel data.

EDA was a cornerstone of the experimental process, offering critical insights into the dataset's structure and quality. The dataset comprised over 50,000 images annotated with four types of distresses: longitudinal cracks, transverse cracks, alligator cracks, and potholes. A detailed distribution analysis revealed class imbalances, with certain types of distress significantly underrepresented. This imbalance informed subsequent augmentation strategies, such as flipping, cropping, and brightness adjustments, designed to artificially expand the representation of minority classes. Statistical graphics and visualizations, developed using tools like *Pandas* and *Matplotlib*, highlighted these imbalances and provided an empirical basis for preprocessing decisions (Figure 6). It is important to note that some of the classes such as D11 and D01, were meant to be subclasses of the different types of longitudinal and transverse cracks, respectively. Due to their extremely low instance count, they were omitted from the training dataset.

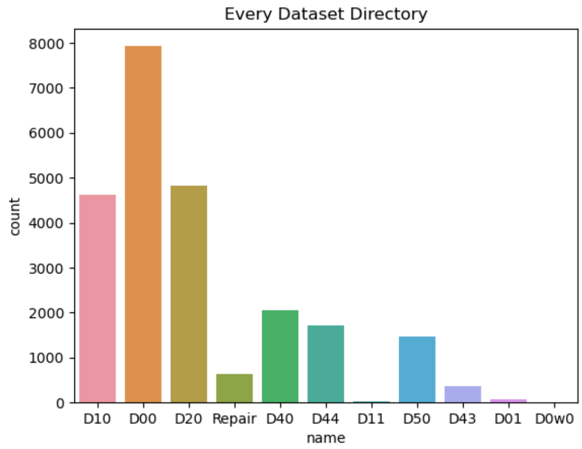

Figure 6. Instance distribution graph of all the annotated classes

Several inconsistencies were identified in some of the subsets, leading to the omission of the "Repair" class and "D0w0.". The first reason is that it was rightly hypothesized that the models would implicitly learn to identify repair patches. The second reason is the presumption that the class itself was an annotation error. Additionally, the dimensions and resolutions of the images varied widely. Resizing and normalization steps were implemented to standardize inputs to the neural networks, ensuring consistent processing.

As shown in the (Figure 7), approximately 33% of the original dataset were null unannotated images. It is advised not to have more than 10% of the combined dataset populated with unannotated images, as after that threshold they do not impact the model's performance but do increase the training time. This subset was crucial for evaluating the models' ability to distinguish between damaged and undamaged road segments. These unannotated images were carefully managed during validation to prevent their inclusion from skewing performance metrics or inducing unintended biases. The dataset used in model training was adjusted according to architecture recommendations, below the 10% rate of unannotated image occurrence. The exploratory analysis identified correlations between distinct types of distresses, including instances of co-occurrence, which proved significant for fine-tuning model thresholds and reducing misclassification.

Spatial distribution and bounding box dimensions were closely analyzed to identify anomalies. Heat-maps displayed in Figure 8 revealed regions where annotations were densely clustered, suggesting potential challenges for object detection models. Simultaneously, bounding box sizes were examined to identify outliers, such as extremely small or disproportionately large annotations.

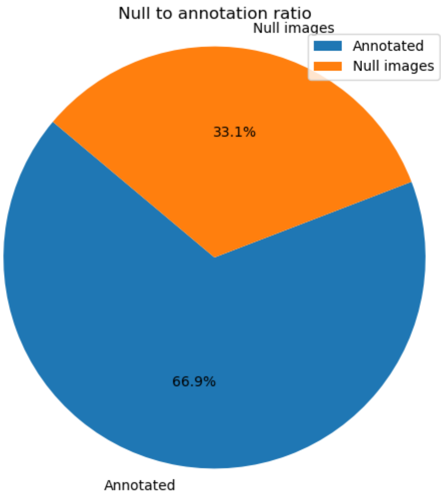

Figure 7. Instance distribution graph of all the annotated classes

These analyses were integral to refining the dataset and ensuring high-quality inputs for training. Representative samples of annotated images were visualized to validate the annotation process and to assess qualitative aspects of the dataset. Adjustments to the preprocessing pipeline were made based on these observations, particularly to address edge cases such as occluded objects or poorly defined distresses.

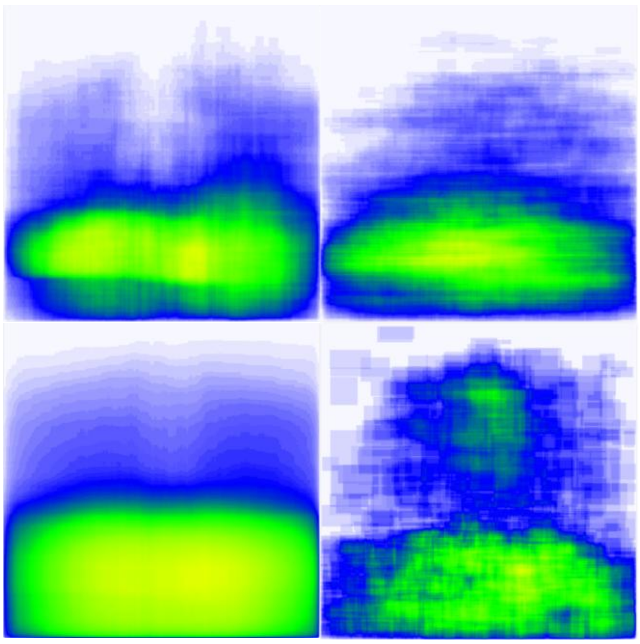

Figure 8. Heat-maps of 4 main types of pavement distresses: longitudinal (upper left), transverse (upper right), alligator cracks (lower left) and potholes (lower right)

5. Performance Metrics

The performance of pavement distress detection models in this study was assessed using metrics that provide a comprehensive assessment of accuracy, precision, and practicality. The primary metric was mean Average Precision (mAP), which evaluates a model's ability to correctly identify and localize distress types within the dataset. Specifically, the mAP@50 metric was applied, using a 50% overlap threshold between predicted and actual bounding boxes. This threshold strikes a balance between precision and tolerance for minor deviations, making it suitable for assessing real-world applications where some variation is acceptable. The mAP metric is particularly valuable for summarizing performance across multiple types of pavement distresses, ensuring that all distress types—such as longitudinal cracks, transverse cracks, alligator cracks, and potholes—are equally considered in the evaluation.

Intersection over Union (IoU) was used to measure the degree of overlap between the predicted bounding boxes and the actual distress areas annotated in the dataset (Figure 9). This metric is crucial for ensuring accurate localization of detected distresses, which is essential for effective maintenance planning. IoU values range from 0 to 1, with higher values indicating better localization. For road maintenance purposes, accurate localization means that repair efforts can be targeted with minimal material waste and disruption to traffic. The IoU metric also allowed the evaluation of model performance under varying levels of confidence, revealing how well the model could handle ambiguous cases or overlapping distresses.

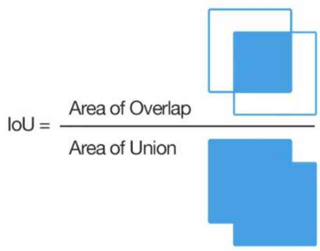

Figure 9. Explanation how IoU is determined

In addition to these accuracy-based metrics, inference speed was considered a critical factor. Road and traffic engineers require solutions that can operate in real-time or near-real-time environments, particularly when integrated into vehicle-mounted systems or mobile devices for field inspections. Models that achieved high detection accuracy but required significant computational resources were noted as less practical for immediate deployment. Instead, models with balanced performance across accuracy and speed were emphasized as the most suitable for real-world use.

While these metrics provided robust quantitative assessments, qualitative observations also contributed to the evaluation. For instance, cases where models struggled with distress types that were underrepresented in the dataset highlighted the importance of balanced training data. Similarly, false positives or poorly localized detections, though infrequent, underscored the challenges posed by visually complex scenes such as overlapping road markings or shadowed areas. These observations complemented the numerical evaluations, offering insights into areas for further improvement.

By combining precision-focused metrics such as mAP and IoU with considerations of inference speed and practical challenges, the study provided a holistic evaluation framework. This approach ensured that the results were both scientifically rigorous and aligned with the practical needs of road and traffic engineering. The findings underscore the potential of these models to enhance road maintenance workflows, while also identifying areas for refinement to maximize their applicability in diverse conditions.

6. Choice of Models, Inference, and Mapping

The selection of models in this study was guided by a balance between accuracy, computational efficiency, and practical usability in the context of road maintenance applications. Several models were considered, with YOLOv8m emerging as the optimal choice due to its ability to provide reliable detection performance while maintaining a manageable level of computational complexity. There were 7 distinct classes that the models were tasked with detecting. The annotations were marked with the convention “D\d{2}” and as such are:

- **D00 - Longitudinal crack**
- **D10 - Lateral crack**
- **D20 - Alligator crack**
- **D40 – Pothole, segregation, rutting**
- **D43 - Cross walk corruption**
- **D44 - White line corruption**
- **D50 - Manhole covers.**

Some of the classes weren’t distresses as such, but were already annotated, and made the model more generalist, without raising complexity by a considerable margin. This model achieved a mean Average Precision (mAP) of over 70% on the validation set, demonstrating a robust capacity to identify and localize various types of road distress. Its performance, when compared to alternatives like YOLOv8L and RT-DETR, indicated a favorable trade-off between detection quality and the resources required for deployment in real-world scenarios. The resulting confusion matrix can be seen as shown in (Figure 10).

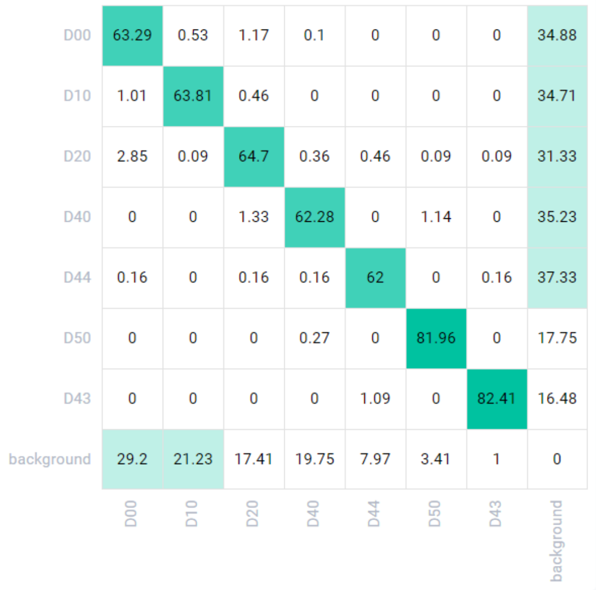

Figure 10. Confusion Matrix for YOLOv8m with class percentage in the respective cells

The inference process involved applying the fine-tuned model to images that were not part of the training dataset. These images were captured using a panoramic camera system mounted on a vehicle, covering urban roads in Belgrade. The inference stage tested the model’s ability to generalize to new environments and data conditions, including varying light levels and road surface types. Some of the results can be seen in (Figure 11).

Results showed consistent performance across most types of distresses, though some challenges were noted in detecting less frequent types, such as transverse cracks. This limitation highlighted the importance of balanced datasets and the potential need for further augmentation or retraining to improve performance in these areas.

Once the detections were processed, the inferred results were mapped using geospatial tools to visualize the locations of the distresses. Each detected instance was georeferenced using metadata from the camera system, which included GPS coordinates. These locations were overlaid on a digital map, providing a visual representation of the road distress across the surveyed area. This mapping process not only validated the model's detections but also demonstrated its practical utility in supporting maintenance workflows. By enabling engineers to pinpoint areas requiring repair, the system offers a scalable and efficient alternative to manual inspections.

Figure 11. Examples of road distress: longitudinal cracks and potholes (upper left and lower right), transversal cracks (upper right) and alligator cracks (lower left)

Mapping the detected inferences involved integrating model outputs with geospatial data to create a comprehensive visual representation of road distress locations. Each detected instance was georeferenced using metadata from the panoramic camera system, which included GPS coordinates embedded during data collection. These detections were plotted on an interactive digital map using geospatial libraries like Folium. This mapping process not only provided a visual verification of the model's detections but also transformed raw detection data into actionable insights. For instance, clusters of distress detections could be analyzed to identify hotspots of road deterioration, enabling road maintenance teams to prioritize repairs efficiently.

By overlaying distress locations on a map, engineers could easily identify critical areas requiring intervention while accounting for the spatial distribution of distress. This capability is particularly beneficial for managing maintenance schedules and optimizing repair sections, minimizing traffic interruption. Furthermore, the mapped data can be integrated into broader infrastructure management systems, offering opportunities for long-term monitoring and planning. This process highlights the transformative potential of combining deep learning models with geospatial analysis to modernize road inspection and maintenance practices.

7. Conclusion

The selection of models in this study was based on a balance between accuracy, computational efficiency, and practical usability for road maintenance applications. Several models were considered, with YOLOv8m emerging as the optimal choice due to its ability to provide reliable detection performance while maintaining a manageable level of computational complexity. This model achieved a mean Average Precision (mAP) of over 70% on the validation set, demonstrating a robust capacity to identify and localize various types of pavement distresses. Its performance, when compared to alternatives like YOLOv8L and RT-DETR, indicated a favorable trade-off between detection quality and the resources required for deployment in real-world scenarios.

The inference process involved applying the fine-tuned model to images that were not part of the training nor validation dataset. These images were captured using a panoramic camera system mounted on a vehicle, covering urban roads in Belgrade. The inference stage tested the model’s ability to generalize to new environments and data conditions, including varying light levels and road surface types. Results showed consistent performance across most pavement distresses, though some challenges were noted in detecting less frequent types of distress, such as transverse cracks. There is a notable discrepancy in the inference results made on the images from Belgrade and the validation metrics made with the test subset. This indicates the inability of the model to generalize to the specificities of the environment it was tasked to operate in. This limitation highlighted the importance of balanced datasets and the potential need for further augmentation or retraining to improve performance in these areas. Models performance can also be improved by adding additional annotated data from the geographically similar location, with whom the model could possibly be able to perform detection more reliably.

Mapping the detected inferences involved integrating model outputs with geospatial data to create a comprehensive visual representation of pavement distress locations. Each detected instance was georeferenced using metadata from the panoramic camera system, which included GPS coordinates embedded during data collection. These detections were plotted on an interactive digital map using geospatial libraries like Folium. This mapping process not only provided a visual verification of the model's detections but also transformed raw detection data into actionable insights. For instance, clusters of distress detections could be analyzed to identify hotspots of road deterioration, enabling road authorities to efficiently perform maintenance strategies.

Improvements can be made in various ways, such as creating a sequential model, one that is optimized for *high recall* and the other for *high precision*. Specializing the models in this way usually provides better results compared to having a single CNN that does both well.

Acknowledgements - The authors are thankful to the MapSoft d.o.o. for providing the datasets used in this research and to the Ministry of Education, Science, and Technological Development of the Republic of Serbia for their support under research project No. 2000092.

References

[1] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.

[2] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A.C. SSD: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016.

[3] Ren, S., He, K., Girshick, R., Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149.

[4] Yu, B.X., Yu, X. (2006) Vibration-based system for pavement condition evaluation. In Proceedings of the 9th International Conference on Applications of Advanced Technology in Transportation (AATT), Chicago, IL, USA, 13–16 August 2006.

[5] Zoysa, K.D., Keppitiyagama, C., Seneviratne, G.P., Shihan, W. (2007) A public transport system based sensor network for road surface condition monitoring. In Proceedings of the Workshop on Networked Systems for Developing Regions, Kyoto, Japan, 27 August 2007.

[6] Tai, Y.-C., Chan, C.-W., Hsu, J.Y.-J. (2010) Automatic road anomaly detection using smart mobile device. In Proceedings of the Conference on Technologies and Applications of Artificial Intelligence (TAAI2010), Hsinchu City, Taiwan, 18

[7] Mednis, A., Strazdins, G., Zviedris, R., Kanonirs, G., Selavo, L. (2011) Real time pothole detection using android smartphones with accelerometers. In Proceedings of the 2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS), Barcelona, Spain, 27–29 June 2011.

[8] Chang, K., Chang, J.R., Liu, J.K. (2005) Detection of pavement distress using 3D laser scanning. In Proceedings of the ASCE International Conference on Computing in Civil Engineering, Cancun, Mexico, 12–15 July 2005.

[9] Hou, Z., Wang, K.C.P., Gong, W. (2007) Experimentation of 3D pavement imaging through stereovision. In Proceedings of the International Conference on Transportation Engineering, Chengdu, China, 22–24 July 2007.

[10] Yu, X., Salari, E. (2011) Pavement pothole detection and severity measurement using laser imaging. In Proceedings of the International Conference on Electro/Information Technology, Mankato, MN, USA, 15–17 May 2011.

[11] Moazzam, I., Kamal, K., Mathavan, S., Usman, S., Rahman, M. (2013) Metrology and visualization of potholes using the Microsoft Kinect sensor. In Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems, The Hague, The Netherlands, 6–9 October 2013.

[12] Staniek, M. (2017) Stereo vision techniques in the road pavement evaluation. Balt. J. Road Bridge Eng. 2017, 12, 38–47.

[13] Zhang, Z., Ai, X., Chan, C.K., Dahnoun, N. (2014) An efficient algorithm for pothole detection using stereo vision. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014.

[14] He, K., Zhang, X., Ren, S., Sun, J. (2016) Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.

[15] Koch, C., Brilakis, I. (2011) Pothole detection in asphalt pavement images. Adv. Eng. Inform. 2011, 25, 507–515.

[16] Koch, C., Jog, G.M., Brilakis, I. (2013) Potholes detection with image processing and spectral clustering. J. Comput. Civ. Eng. 2013, 27, 370–378. DOI: 10.1061/(ASCE)CP.1943-5487.0000232

[17] Jog, G.M., Koch, C., Golparvar-Fard, M., Brilakis, I. (2012) Pothole properties measurement through visual 2D recognition and 3D reconstruction. In Proceedings of the ASCE International Conference on Computing in Civil Engineering, Clearwater Beach, FL, USA, 17–20 June 2012.

[18] Nienaber, S., Booysen, M.T., Kroon, R. (2015) Detecting potholes using simple image processing techniques and real-world footage. In Proceedings of the 34th Southern African Transport Conference (SATC 2015), Pretoria, South Africa, 6–9 July 2015.

[19] Lokeshwor, H., Das, L.K., Sud, S.K. (2013) Method for automated assessment of potholes, cracks and patches from road surface video clips. Procedia Soc. Behav. Sci. 2013, 104, 312–321.

[20] Lin, J., Liu, Y. (2010) Potholes detection based on svm in the pavement distress image. In Proceedings of the 2010 Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science, Hong Kong, China, 10–121 August 2010.

[21] Hoang, N.-D. (2018) An artificial intelligence method for asphalt pavement pothole detection using least squares support vector machine and neural network with steerable filter-based feature exteraction. Adv. Civ. Eng. 2018, 2018, 7419058. DOI: 10.1155/2018/7419058

[22] Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., Omata, a.H. (2018) Road Damage Detection Using Deep Neural Networks with Images Captured through a Smartphone. Comput. Aided Civ. Infrastruct. Eng. 2018, 33, 1127–1141. DOI: 10.1111/mice.12387

[23] Hoang, N.-D., Huynh, T.-C., Tran, V.-D. (2021) Computer Vision-Based Patched and Unpatched Pothole Classification Using Machine Learning Approach Optimized by Forensic-Based Investigation Metaheuristic. Complexity 2021, 2021, 3511375. DOI: 10.1155/2021/3511375

[24] Kotha, M., Chadalavada, M., Karuturi, S.H., Venkataraman, H. (2020) PotSense: Pothole Detection on Indian Roads using Smartphone Sensors. In Proceedings of the 1st ACM Workshop on Autonomous and Intelligent Mobile Systems, Bangalore, India, 11 January 2020.

[25] Baek, K., Byun, Y., Song, H. (2018) Pothole detection using machine learning. Adv. Sci Technol. Lett. 2018, 150, 151–155.

[26] Dharneeshkar, J., Aniruthan, S., Karthika, R., Parameswaran, L. (2020) Deep Learning based Detection of potholes in Indian roads using YOLO. In Proceedings of the 2020 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 26–28 February 2020.

[27] Shaghouri, A.A., Alkhatib, R., Berjaoui, S. (2021) Real-Time Pothole Detection Using Deep Learning. arXiv 2021, DOI: 10.48550/arXiv.2107.06356

[28] Ahmed, K.R. (2021) Smart Pothole Detection Using Deep Learning Based on Dilated Convolution. Sensors 2021, 21, 8406. DOI: 10.3390/s21248406

[29] Arya, D., Maeda, H., Ghosh, S., Toshniwal, D.,Sekimoto, Y. (2022) RDD2022: A multi-national image dataset for automatic Road Damage Detection. DOI: 10.48550/arXiv.2209.08538

[30] Jocher, G., Qiu, J., Ayush, C. (2023). Ultralytics YOLOv8 [Computer software]. URL: https://ultralytics.com