WSEAS TRANSACTIONS on SIGNAL PROCESSING
DOI: 10.37394/232014.2025.21.5

An Automated and Integrated Sensing System for Road Monitoring using UAV Images and an Optimized R-CNN

VINCENZO BARRILE, FRANCESCO SCOPELLITI, EMANUELA GENOVESE,
GIOVANNI LEONARDI, GIUSEPPE MARIA MEDURI
Department of Civil Engineering, Energy, Environment and Materials (DICEAM),
"Mediterranea" University of Reggio Calabria,
Via Zehender (loc. Feo di Vito), 89122, Reggio Calabria,
ITALY

**Corresponding Author*

Abstract: - One of the most relevant, but at the same time most time-consuming and costly, aspects of the infrastructure system is the monitoring of road infrastructures, often subject to deterioration that compromises their use. Current monitoring systems consist of individual reports or the use of human resources that, through equipped vehicles, have the purpose of carrying out a reconnaissance process, which is often characterized by errors and uncertainties. In this context, the aim of this work was to experiment and implement an experimental and innovative Automated and Integrated Sensing System (AISS) for the monitoring of road infrastructures. This system, starting from Remote Sensing images from Unmanned Aerial Vehicles (UAVs), uses a Mask R-CNN neural network to identify road cracks. This information, together with other information, is included in a database, which is then used in a Geographical Information System (GIS) for relative visualization. This work therefore proposes a methodology for the implementation of a system that helps policy makers in determining the most urgent interventions. In fact, a categorization of the severity of degradation and a user-friendly visualization, allow us to make decisions based on data.

Key-Words: - Road monitoring, Remote Sensing, UAV, CNN, Deep Learning, Geographical Information System.

Received: April 14, 2024. Revised: November 17, 2024. Accepted: March 6, 2025. Published: April 7, 2025.

1 Introduction

Road infrastructure is a key component of the urban infrastructure system. It plays a fundamental role in the proper functioning of the city and in sustainable development. Road pavements, more than any other civil engineering structure, are subject to deterioration over time due to various stresses (e.g., continuous traffic) and atmospheric phenomena. Therefore, it is essential to design the structures appropriately and conduct regular maintenance operations. The initial phase of road surface inspection involves data collection to perform routine assessments, and identify, and delineate areas of discomfort and deterioration. The duration of mapping operations for data collection is critical to minimizing overall maintenance and repair costs. Until the early 1980s, data regarding the condition of road surfaces were obtained through "visual inspections," in which personnel drove along the road to collect both qualitative and quantitative information on the pavement by examining cracks. This investigative method was unreliable due to its subjective nature, frequent inconsistencies, and the need for time and skilled technicians to perform the procedure. It was also dangerous, as it put the safety of the operators at risk by exposing them to vehicular traffic. Inspections were conducted using cameras mounted on vehicles; nonetheless, despite the rapid acquisition of information, this technology proved to incur high costs and produced low-accuracy photos attributable to outside and environmental factors.

Consequently, in recent decades, progress has been made in designing methods for the automated acquisition of imagery and the extraction of reliable and important data regarding pavement conditions. At the same time, the acquisition of huge amounts of images has made precise and rapid analysis difficult. This problem has been partly overcome thanks to the advent of AI (Artificial Intelligence), and in particular deep learning, which facilitates the administration of substantial quantities of data and the automated extraction of relevant information from them. In relation to this aspect, the literature contains several interesting contributions that address the problem of infrastructure monitoring and the maintenance of road traffic and safety using Machine Learning and Artificial Intelligence techniques. An interesting contribution by [1] describes a novel method based on a modified recurrent neural network for automated monitoring of road pavement aging conditions using high-resolution satellite images.

In this contribution, the proposed method is based on a neural network called the Bi-directional Gated Recurrent Unit (Bi-GRU), which is optimized to improve the classification of various aging stages of asphalt and unpaved objects. Bi-GRU is a variant of traditional recurrent neural networks that uses a particular gradient to predict sequences of data such as text or time signals. In addition, to further improve the network's performance, a spectral augmentation method has been developed by the authors to enhance the spectral details of road images. Scientific literature is, in fact, focusing on the use of images no longer taken by humans with cameras mounted on equipped vehicles, but rather on Remote Sensing from satellites and drones. In the contribution, very high-resolution images are cited, such as the WorldView-2, which guarantees a resolution of 16,360 x 7,728 pixels. Obviously, satellite imagery, especially privately distributed high-resolution imagery, provides an efficient method for large-scale road health monitoring and could be a key component of automated road pavement aging monitoring. To achieve excellent resolutions at a significantly lower cost, the use of UAVs (Unmanned Aerial Vehicles) is also becoming increasingly popular in this field. A recent study by [2] discusses a drone-based solution for automating road segmentation by addressing the practical challenges of using such technologies. The paper highlights a recent World Bank report, which states that 80% of a road’s life-cycle costs are spent on maintenance, which, of course, includes monitoring and repair processes. The study presents a dedicated road monitoring platform that enables the drone to detect road boundaries, locate vanishing points, identify edge lines, and determine the center of the road. The process includes segmentation and labeling of the road with temporal and geographical information from the drone's IMU (Inertial Measurement Unit). There are also several reviews on this topic, particularly the review by [3], which highlights important conclusions in this field:

- The growing use of drones, particularly in China, which is a leader in this field.
- The preference for rotor drones for data collection, especially quadcopters.
- The increasing trend of using deep learning-based techniques for vehicle detection and tracking, as well as for extracting their trajectories.

In this field, neural networks have emerged as optimized methods for road extraction. The article [4] describes various techniques and approaches for automatic road network extraction. Neural networks and genetic algorithms are considered optimized techniques. Special attention is given to a general processing framework for urban road network extraction in Synthetic Aperture Radar (SAR) images, utilizing a Markov Random Field (MRF) description, Support Vector Machines, and mathematical morphology. SAR involves acquiring detailed radar images of the Earth's surface by exploiting the movement of the satellite to synthesize an antenna aperture that is longer than its actual physical size. The article also describes an approach that combines the extended Kalman filter with a special particle filter and automatic intersection detection using raster and vector information for classification and segmentation.

Other interesting aspects regarding the state of the art in these technologies are the control systems and flight techniques, [5], [6], [7]. Missions and tasks can vary significantly, requiring specific flight plans and approaches to flight system control for each situation. Aircraft control can range from full manual control, stabilized control, and planned control to fully automated flight plans without direct control of the flight path Additionally, flight systems can be organized with multiple devices, creating an interconnected network for information collection. The level of automation in the flight mission depends on several factors, such as the type of mission, the number and repetition of required movements, the proximity of the devices, as well as the resolution, accuracy, and quality of the survey, [8], [9], [10], [11], [12].

Today's most advanced techniques merge these models with Computer Vision systems, a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos, and other visual inputs, and take actions or issue warnings based on that information. To function properly, computer vision systems must be trained on a large volume of data, which will constitute the dataset that can make the algorithm truly intelligent.

An additional advancement in this domain pertains to Structure from Motion (SfM) methods [13], [14], which are used in contemporary photogrammetry. Structure from Motion (SfM) algorithms are a photogrammetric imaging method used for predicting three-dimensional structures, which can be integrated with local motion signals. Structure from Motion (SfM) is used in several applications, including three-dimensional scanning, augmented reality, and visual Simultaneous Localization and Mapping (vSLAM). The procedure involves estimating the three-dimensional structure of a scene from a sequence of two-dimensional photographs. To align images, feature points are represented as corners (edges exhibiting gradients in various directions) from one image to another. After identifying feature points in all photos, they are subsequently compared. Structure from Motion (SfM) can be computed using various methodologies. The approach is contingent upon various elements, including the quantity and type of cameras used and the organization of the images. Notable photogrammetry tools that utilize computer vision and Structure from Motion (SfM) techniques include VisualSfM, Meshroom, OpenMVG, and Regard 3D. Agisoft Metashape is a highly adaptable software that facilitates digital image processing and produces 3D spatial data applicable in photogrammetry, GIS applications, cultural heritage documentation, and indirect measurements of objects of varying dimensions. The program is capable of processing images from both RGB and multispectral cameras, as well as multi-camera systems, [15], [16], [17], [18].

Consequently, it is essential to create an integrated system that employs Remote Sensing imagery and Neural Networks for road surveillance. This research proposes an integrated and automated method for road monitoring using Remote Sensing photos from a drone and a neural network (R-CNN) for the automatic detection of road deterioration and cracks. A prototype and experimental system have been created to automate road monitoring using UAVs and to facilitate the analytical processing of the gathered data. The primary aim of this research is to enhance the automation of data processing, augment the precision of autonomous data selection, and ultimately transfer delocalized data to a GIS system. The Automated and Integrated Sensing System (AISS) elucidates the procedure for executing updates by integrating existing methodologies and frameworks with conventional operational processes. Moreover, for overall urban efficiency, the AISS is receptive to integrating additional related digital infrastructure technologies for joint enhancement.

2 Materials and Methods

This research work focused on the construction and implementation of an integrated prototype and experimental system (AISS) that, using images acquired by an automated drone and a structural network specifically optimized for this work, identifies road surface irregularities.

Big Data [19] forms the basis of the systems involved in road surface management. It is necessary to collect, analyze, and use a large quantity of data on road conditions. This data can come from a variety of sources, including roadside sensors, connected vehicles, UAVs, satellite images, and historical data. Typically, the ways in which Big Data is used for road surface management are:

- Road surface inspection: Data collected on road pavement conditions is analyzed to identify areas requiring maintenance or repair.
- Maintenance planning: Data analysis can help predict road pavement lifetimes and plan preventive maintenance. For example, predictive models can be used to evaluate when work will be needed.
- Resource optimization: Big data can help optimize the use of available road pavement resources, prioritize work, and allocate resources effectively to maximize efficiency and minimize costs.
- Evaluation of the impact of road conditions: Collected data can be used to assess the impact of road conditions on safety and traffic efficiency. For example, the analysis of traffic data and road status can help identify the causes of accidents and take corrective measures to improve road safety.

The methodology used in this work includes several phases: the identification of road surfaces and cracks with relative 3D reconstruction, data analysis and processing, and finally, integration into the system to support decision-making. The different phases of the methodology are represented in Figure 1.

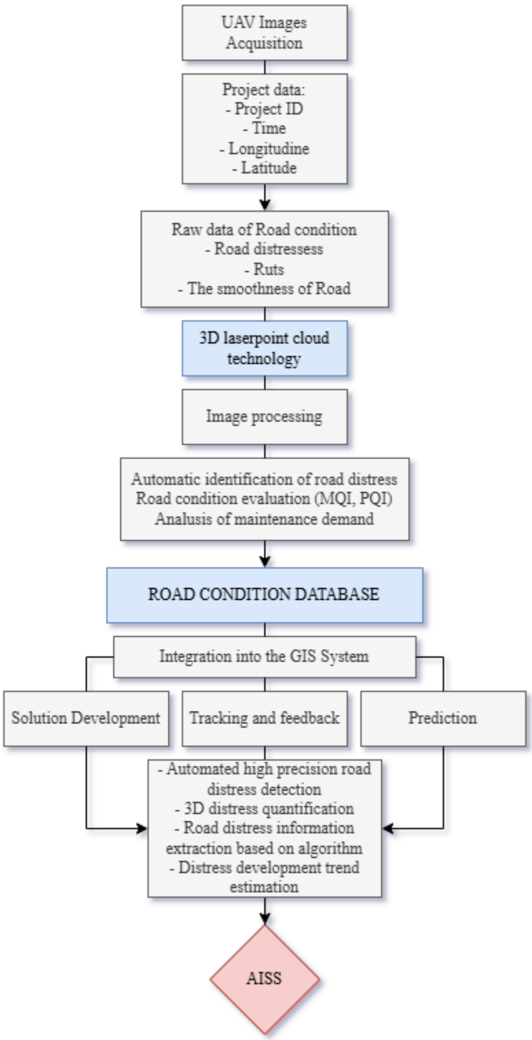

Fig. 1: Methodology of the experimental proposed system for road monitoring

2.1 UAV Images Acquisition

In particular, the first phase was dedicated to the acquisition of images with a drone. The drone available and used by the Geomatics Laboratory of the Mediterranea University of Reggio Calabria is a DJI Matrice 300 RTK (Figure 2), equipped with a LiDAR camera, [20]. The DJI Matrice 300 RTK is a high-end professional drone, known for its precision and versatility. Equipped with an RTK system, this drone is suitable for industrial applications such as mapping, precision agriculture, control, and inspection. Additionally, it also features advanced autonomous flight capabilities. The drone's maximum payload reaches 2.7kg and supports various types of payloads. It has a flight time of 55 minutes and a long-lasting lithium battery. It has an IP45 resistance rating and supports not only LiDAR but also the thermal wide-angle camera.

Fig. 2: DJI Matrice 300 RTK

For this study, a UAV was used to collect high-definition image data. The image acquisition process primarily involves flight planning, pixel calibration, and image quality assessment.

The image must have strong contrast in the crack area and have pixels no smaller than 560x540 (the clearer, the better) in order to meet the data criteria. To make later training processes easier, every image should have the same format. Once collected, the images were subsequently pre-processed using Pix4D software, [21]. The software was used both in flight planning and as processing software. For the flight plan, an altitude of 100 m and a speed of 10 m/s were chosen.

Along with the images, important information was collected thanks to the advanced positioning system of the drone. In fact, information about time and geographical coordinates (latitude and longitude, expressed in a specific reference system) was obtained. This raw data allowed for the definition of road damage, its regularity, surrounding structures, and detailed information about the road status. All this was used for an initial visualization through 3D laser point cloud technology, [22], [23], [24], [25].

Based on this technology, which has been used in civil engineering in the last few decades, a 3D model of road conditions was implemented, including all the necessary information. In this way, the 3D model can help achieve real-time monitoring and visualization of road deterioration.

2.2 Image Processing

The second phase, on the other hand, focused on data analysis, which included image processing for the automatic identification of road distress. This data, along with the assessment of road conditions through specific indices and the analysis of road maintenance requests, allowed for the creation of a road condition database to be used in the AISS system, [26]. In this phase, the focus was initially on data processing and, in particular, on crack labeling. Image labeling plays a fundamental role in computer vision. The goal of image labeling is related to the specific activity that is intended to be identified. To optimize the functionality of machine learning, it is essential to increase the precision of the training data and correctly label the images in order to obtain accurate results. The collected images were labeled using the software Labelling, [27].

Mask-RCNN, [28], [29], [30], a neural network-based model, was then created and refined. The ROI-Align module, the Region Proposal Network (RPN), and the backbone network are components that make up the Mask-RCNN instance segmentation algorithm. This network's primary objective is to detect and analyze issues with road pavements, including cracks, by using object detection, instance segmentation, and keypoint identification. The model may achieve detailed semantic segmentation by labeling each individual pixel, in addition to detecting damage in the image.

The process begins with the use of a convolutional network to generate a Region Proposal Network (RPN) [31] that can propose regions in the image that could contain objects. These regions are then classified using a convolutional network and assigned to a specific class. What distinguishes Mask R-CNN from other convolutional networks is the addition of a mask generation branch for each detected object. After identifying the position of the object, the network constructs an object-specific binary mask, which allows for precise delineation of its shape. For each region proposal, in addition to classification, a "mask head" is used to predict the mask for each object, with an output that matches the extent of the suggested region. This method ensures detailed and high-quality segmentation that is not limited to object classification but also provides precise delineation of edges.

In addition to this domain, Mask R-CNN is applied in various areas, such as medical image recognition, robotic vision, and security monitoring. The initial training weight of this model uses the parameters of the Asphalt Cracked and Uncracked Image Dataset, which was used for training, [32].

The Asphalt Cracked and Uncracked Image Dataset, which contains images of asphalt pavements with uncracked areas and cracks recorded, was used in this investigation. The dataset provides ground truth data for applications such as crack recognition and pavement condition assessments. It contains more than 2,000 images with crack annotations in various shapes and severity levels to train models for object detection and segmentation in road infrastructure monitoring.

The initial tactic was to focus on adjusting the parameters of the primary neural network component while freezing other network components during training. After the primary component was sufficiently trained, all frozen parameters were unlocked for global training. To guarantee thoroughness, this training procedure was conducted over 40 epochs with 200 steps per epoch.

To enhance the model's performance and generalization, an additional 500 new image data points were incorporated into the training set. The dataset's quality was further optimized by carefully screening factors that could impact the final results, including pixel quality, background variations, and the different types of cracks (e.g., surface, deep, wide).

Regarding the problem of model superparameters that did not reach the optimal solution, this paper studies the adjustment of superparameters and optimizes the model structure. Upon marking and processing 500 image data points, the Mask-RCNN model was constructed. Of these, 450 image data points were allocated for training the model, while the remaining 50 image data points were used for testing and model evaluation.

After adding the new images, the optimized neural network showed significantly better performance compared to the first configuration, highlighting that the optimization was successful. Compared to previous studies, the neural network with the new configuration shows advantages in relation to faster labeling of the dataset followed by better accuracy in classification. It is important to highlight that the photos captured in situ are certainly influenced by some disturbances such as noise and shadows that negatively affect the identification of deterioration on the asphalt.

With the exception of classification, the severity of the road difficulties, e.g. the depth of the ruts, will also be assessed. After this step, detailed information on all road difficulties is collected. The assessment of road conditions also simplifies the decision-making process. There are several road condition assessment criteria. These criteria are integrated into the system and will be used to classify road pavement problems, also according to priority, in order to streamline the road maintenance phase. In this section, a database containing all useful data for service optimization is also generated over time. For example, the temporal analysis of road maintenance data.

One of the indicators used is the Road Maintenance Quality Indicator. This is a common indicator for assessing the quality of road infrastructure elements, such as highway surfaces, sidewalks, bridge and tunnel structures, and roadside structures. Closely related to this indicator is the Pavement Quality Index, which together helps define the actual conditions of the road surface.

As is known, in large cities, consisting mainly of a dense network of roads, it is often necessary to conduct numerous and complex monitoring that often results in numerous requests for intervention. It is therefore essential to assign priorities to each of these requests, based on the degree of severity of the deterioration identified. Thanks to this classification, decision-makers can understand which intervention is more urgent compared to the hundreds of others present on a large scale.

This system not only allows you to identify the most damaged sections of the road, allowing you to act promptly and in order of severity, but also allows you to minimize the inconveniences that could arise from on-site interventions, both for motorists and pedestrians.

Obviously, the data related to the vehicle fleet grows exponentially, which allows for more in-depth and temporal analyses regarding the maintenance of the road surface. The data collected, processed, and cataloged need to be stored in a structured and dedicated database. The information resulting from the road surveys will be used to justify the need for specific maintenance activities and to plan further maintenance interventions.

2.3 AISS system

The AISS system, appropriately implemented and described in the flowchart of Figure 1 and described in detail in its various components in paragraph 2, is developed to automatically identify the defects of the road surface through drone surveys, displaying them automatically on a computerized georeferenced system (GIS). In this final phase, the data collected and entered into the database were used to allow informed decisions on the development of solutions, feedback, and possible prediction of some susceptible areas. All this is to obtain a system that identifies distress automatically, with high resolution and precision, the quantification of distress in 3D, the extraction of information on road distress based on algorithms, and the estimation of the development trend of distress, [33], [34], [35], [36], [37], [38], [39], [40], [41].

Road problems can be categorized into pavement deformation and pavement degradation. Pavement deformation is further divided into surface uniformity and surface rutting, while pavement degradation includes surface cracking and surface loosening. After measuring the texture depth, the abnormal road sections are identified and assessed. All issues are analyzed using the same method.

Once the road problem is identified, additional road attributes, such as operating life, inspection logs, and maintenance logs, can be quickly and easily retrieved from the database. After collecting all the necessary data on the road issues, an integrated information system compiles and presents this data in a clearer and more intuitive format.

The GIS system can visually display basic route information, road performance assessments, road difficulties, the maintenance planning scheme, and the patrolling track on an electronic map. The integrated information platform supports decision-making processes. It increases management preferences and simplifies workflow.

In conclusion, the proposed system has the following characteristics:

The process of identifying, classifying, and ascertaining road problems (cracks) is automated. A more intuitive presentation of partial and detailed road information.

Quickly and accurately identifying important information on a specific section of the road and easily displaying it. Simplifying decision-making and improving management performance.

3 Results

With the aim of creating and implementing the Automated and Integrated Sensing System (AISS), several factors were initially considered, including the uniformity of the surface, the structures along the road, and the particular texture of the road surface. This process is far more precise than human intervention and manual investigation (Figure 3). The designed system can collect data in a consistent and uniform manner, allowing for high precision. In fact, small variations in uniformity or minor details that may emerge from this analysis can easily be overlooked by a human operator. Furthermore, the reduced environmental impact of the entire process should not be underestimated: by using drones and specific sensors, it is possible to minimize carbon dioxide emissions.

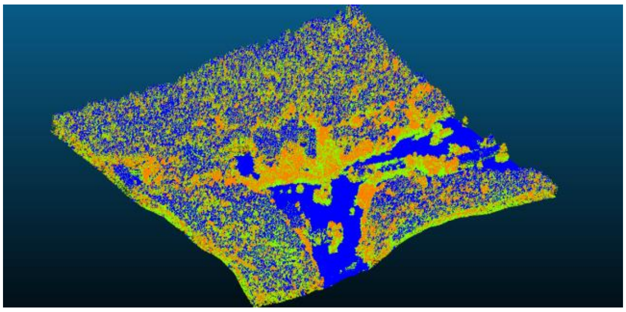

Fig. 3: Point Cloud sample

Regarding data labeling, Figure 4 illustrates some of the results obtained using the methodology outlined in Section 2.

Fig. 4: Data labelling

As a result of the neural network's performance, Figure 5 reproduces the same crack that was correctly detected and segmented by the neural network implemented by the authors.

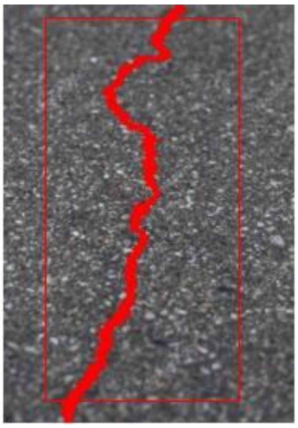

Fig. 5: Results of crack detection

As a final step, the information is stored in the relational database, including all details related to the final crack detection, such as its location, severity level, and associated image, for integration into the GIS system (Figure 6).

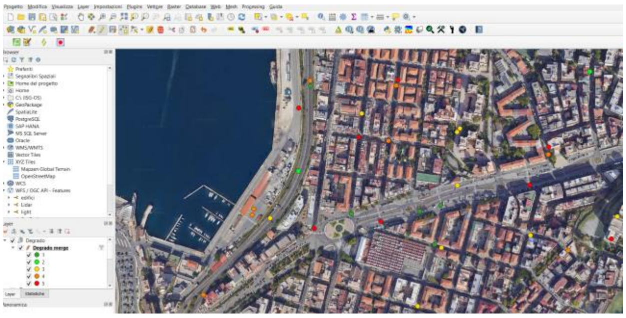

Fig. 6: Automated and Integrated Sensing System (AISS)

4 Conclusion

In conclusion, the system focuses on implementing an innovative and experimental solution that enables the acquisition and subsequent classification of road deterioration and cracks through a dedicated neural network, with the goal of displaying the results within a GIS to plan potential maintenance interventions. This system addresses the need for competent authorities to have a categorized and organized system that allows them to efficiently manage intervention operations.

The main advantages of this technology lie in the low cost of drone images compared to satellite images, the speed of execution and investigation compared to in situ inspections, and the possibility of integrating the system with existing infrastructure. The primary disadvantages may include the relatively high initial cost of adopting the system and the need for multidisciplinary operators.

This methodology can be integrated into a broader context of infrastructure management, as it can be easily inserted into existing tools and ensures the possibility of planning and organizing maintenance interventions of city infrastructures. Furthermore, the system is scalable and adaptable, allowing it to adapt to both small urban cities and large metropolitan cities. Furthermore, it is an approach that takes into account sustainable practices, both in the context of civil engineering and urban planning.

Possible future developments of this research mainly focus on the creation of a WebGIS, allowing users to access and utilize these techniques via the web. Additionally, the system could be integrated into warning systems and predictive models to identify potential traffic disruptions caused by road surface damage. The application details of the different sections of the system will need to be explored in future studies.

Road infrastructure safety is of fundamental importance, and tools like AISS are well-suited to the needs of municipalities, experts, and professionals in the sector to improve and plan interventions in cities, making them smarter, more efficient, and safer.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

The authors wrote, reviewed and edited the content as needed and they have not utilised artificial intelligence (AI) tools. The authors take full responsibility for the content of the publication.

References:

[1] Chen, X., Zhang, X., Li, J., Ren, M., & Zhou, B. (2022). A new method for automated monitoring of road pavement aging conditions based on recurrent neural networks. *IEEE Transactions on Intelligent Transportation Systems*, 23(12), 24510-24523, DOI: 10.1109/TITS.2022.3204334.

[2] Ranjbar, H., Forsythe, P., Fini, A. A. F., Maghrebi, M., & Waller, T. S. (2023). Addressing the practical challenge of using autopilot drone for asphalt surface monitoring: Road detection, segmentation, and following. *Results in Engineering*, 18, 101130, https://doi.org/10.1016/j.rineng.2023.101130

[3] Butilă, E. V., & Boboc, R. G. (2022). Urban traffic monitoring and analysis using unmanned aerial vehicles (UAVs): A systematic literature review. *Remote Sensing*, 14(3), 620, https://doi.org/10.3390/rs14030620.

[4] Kirthika, A., & Mookambiga, A. (2011, June). Automated road network extraction using artificial neural network. *In 2011 International Conference on Recent Trends in Information Technology (ICRTIT)* (pp. 1061-1065). IEEE, Chennai, India, 03-05 June 2011. DOI: 10.1109/ICRTIT.2011.5972323.

[5] Barrile, V., Cotroneo, F., & Genovese, E. (2023, December). Platform Prototype for the Prediction of Landslide Susceptibility Through a 4D WebGIS Equipped with Cellular Automata and Neural Networks. *In Italian Conference on Geomatics and Geospatial Technologies* (pp. 81-95). Cham: Springer Nature Switzerland. Virtual Event, December 18–20, 2023. https://doi.org/10.1007/978-3-031-59925-5_7.

[6] Yang, L., Li, S., Li, C., Zhang, A., & Zhang, X. (2023). A survey of unmanned aerial vehicle flight data anomaly detection: Technologies, applications, and future directions. *Science China Technological Sciences*, 66(4), 901-919. https://doi.org/10.1007/s11431-022-2213-8.

[7] Zhang, J., Shi, Z., Zhang, A., Yang, Q., Shi, G., & Wu, Y. (2023). UAV trajectory prediction based on flight state recognition. *IEEE Transactions on Aerospace and Electronic Systems*, 60(3), 2629-2641. DOI: 10.1109/TAES.2023.3303854.

[8] Kelner, J. M., Burzynski, W., & Stecz, W. (2024). Modeling UAV swarm flight trajectories using Rapidly-exploring Random Tree algorithm. *Journal of King Saud University-Computer and Information Sciences*, 36(1), 101909. https://doi.org/10.1016/j.jksuci.2023.101909.

[9] Boucher, P. B., Hockridge, E. G., Singh, J., & Davies, A. B. (2023). Flying high: Sampling savanna vegetation with UAV-lidar. *Methods in Ecology and Evolution*, 14(7), 1668-1686. https://doi.org/10.1111/2041-210X.14081.

[10] Vacca, G., & Vecchi, E. (2024). UAV Photogrammetric Surveys for Tree Height Estimation. *Drones*, 8(3), 106. https://doi.org/10.3390/drones8030106.

[11] Sonugür, G. (2023). A Review of quadrotor UAV: Control and SLAM methodologies ranging from conventional to innovative approaches. *Robotics and Autonomous Systems*, 161, 104342. https://doi.org/10.1016/j.robot.2022.104342.

[12] Ma, B., Liu, Z., Dang, Q., Zhao, W., Wang, J., Cheng, Y., & Yuan, Z. (2023). Deep reinforcement learning of UAV tracking control under wind disturbances environments. *IEEE Transactions on Instrumentation and Measurement*, 72, 1-13. DOI: 10.1109/TIM.2023.3265741.

[13] Dai, W., Zheng, G., Antoniazza, G., Zhao, F., Chen, K., Lu, W., & Lane, S. N. (2023). Improving UAV- SfM photogrammetry for modelling high- relief terrain: Image collection strategies and ground control quantity. *Earth Surface Processes and Landforms*, 48(14), 2884-2899. https://doi.org/10.1002/esp.5665.

[14] Howland, M. D., Tamberino, A., Liritzis, I., & Levy, T. E. (2022). Digital deforestation: Comparing automated approaches to the production of digital terrain models (DTMs)

in agisoft metashape. *Quaternary*, 5(1), 5. https://doi.org/10.3390/quat5010005.

[15] Sestras, P., Roșca, S., Bilașco, Ș., Șoimoșan, T. M., & Nedevschi, S. (2023). The use of budget UAV systems and GIS spatial analysis in cadastral and construction surveying for building planning. *Frontiers in Built Environment*, 9, 1206947. https://doi.org/10.3389/fbuil.2023.1206947.

[16] Mohren, J., & Schulze, M. (2024). Automated cleaning of tie point clouds following USGS guidelines in Agisoft Metashape professional (ver. 2.1. 0). *MethodsX*, 12, 102679. https://doi.org/10.1016/j.mex.2024.102679.

[17] Zhang, S., Yang, J., Qi, Z., & Zhang, Y. (2023). Toward Meta-shape based Multi-view 3D Point Cloud Registration: An Evaluation. *IEEE Transactions on Circuits and Systems for Video Technology*. vol. 34, no. 7, pp. 5361-5375, July 2024, doi: 10.1109/TCSVT.2023.3341622.

[18] Saputra, H., Okcandra, W., Sutikno, S., & Lubis, M. Z. (2023). Application of Fixed-Wing UAVs to Develop Digital Terrain Model on Coastal Peatland Bengkalis Island. *Journal of Applied Geospatial Information*, 7(2), 867-874. DOI:10.30871/jagi.v7i2.4594.

[19] Himeur, Y., Elnour, M., Fadli, F., Meskin, N., Petri, I., Rezgui, Y., ... & Amira, A. (2023). AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives. *Artificial Intelligence Review*, 56(6), 4929-5021. https://doi.org/10.1007/s10462-022-10286-2.

[20] Czyża, S., Szuniewicz, K., Kowalczyk, K., Dumalski, A., Ogrodniczak, M., & Zieleniewicz, Ł. (2023). Assessment of accuracy in unmanned aerial vehicle (uav) pose estimation with the real-time kinematic (rtk) method on the example of dji matrice 300 rtk. *Sensors*, 23(4), 2092. https://doi.org/10.3390/s23042092.

[21] Barrile, V., & Genovese, E. (2024). Gis-Like Environments And Hbim Integration For Ancient Villages Management And Dissemination. The International Archives of the Photogrammetry, *Remote Sensing and Spatial Information Sciences*, 48, 41-47. https://doi.org/10.5194/isprs-archives-XLVIII-2-W4-2024-41-2024.

[22] Huo, L., Liu, Y., Yang, Y., Zhuang, Z., & Sun, M. (2023). Research on product surface quality inspection technology based on 3D point cloud. *Advances in Mechanical Engineering*, 15(3), 16878132231159523. https://doi.org/10.1177/16878132231159523.

[23] Chen, J., Fang, Q., Zhang, D., & Huang, H. (2023). A critical review of automated extraction of rock mass parameters using 3D point cloud data. *Intelligent Transportation Infrastructure*, 2, liad005. https://doi.org/10.1093/iti/liad005.

[24] Xie, W. (2024, February). Development and Application Analysis of 3D Point Cloud Technology Based on Computer Vision in the Field of Welding Robots. *In 2023 International Conference on Data Science, Advanced Algorithm and Intelligent Computing* (DAI 2023) (pp. 121-128). Atlantis Press. Melaka Malaysia, DOI: 10.2991/978-94-6463-370-2_14.

[25] Xin, L., Zhang, J., & Zhang, N. (2024, August). Application of BIM model based on Lidar point cloud technology in construction of power substations. In Seventh International Conference on Advanced Electronic Materials, Computers, and Software Engineering (AEMCSE 2024) (Vol. 13229, pp. 410-415). SPIE. Nanchang, China. https://doi.org/10.1117/12.3038394

[26] Zhang, H., Fu, X., & Gong, M. (2023). Comparative analysis of condition assessment standards for roads and bridges based on fuzzy comprehensive evaluation. *Australian Journal of Civil Engineering*, 1-11. https://doi.org/10.1080/14488353.2023.2197275.

[27] Tanioka, H., Miura, T., Matsuura, K., & Karungaru, S. (2023, July). Development of semi-automatic image annotation using object recognition. *In 2023 14th IIAI International Congress on Advanced Applied Informatics* (IIAI-AAI) (pp. 546-549). IEEE. Koriyama, Japan. DOI: 10.1109/IIAI-AAI59060.2023.00109.

[28] Sapkota, R., Ahmed, D., & Karkee, M. (2024). Comparing YOLOv8 and Mask R-CNN for instance segmentation in complex orchard environments. *Artificial Intelligence in Agriculture*, 13, 84-99. https://doi.org/10.1016/j.aiia.2024.07.001.

[29] Du, L., Lu, X., & Li, H. (2023). Automatic fracture detection from the images of electrical image logs using Mask R-CNN. *Fuel*, 351, 128992. https://doi.org/10.1016/j.fuel.2023.128992.

[30] Wang, J., Chen, G., Jaboyedoff, M., Derron, M. H., Fei, L., Li, H., & Luo, X. (2023). Loess landslides detection via a partially supervised learning and improved Mask-RCNN with multi-source remote sensing data. *Catena*, 231, 107371. https://doi.org/10.1016/j.catena.2023.107371

[31] Taye, M. M. (2023). Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. *Computation*, 11(3), 52. https://doi.org/10.3390/computation11030052.

[32] Khan, M. A., & Sultana, M. (2024). Real-Time Object Identification Using Coco Dataset. *Journal of Cyber Security, Privacy Issues and Challenges*, 3(2), 8-15, [Online]. https://matjournals.net/engineering/index.php/JCSPIC/article/view/538 (Accessed Date: November 23, 2024).

[33] Liu, C., Liu, Y., Chen, Y., Zhao, C., Qiu, J., Wu, D., & Tang, K. (2023). A state-of-the-practice review of three-dimensional laser scanning technology for tunnel distress monitoring. *Journal of Performance of Constructed Facilities*, 37(2), 03123001. https://doi.org/10.1061/JPCFEV.CFENG-4205.

[34] Ragnoli, A., De Blasiis, M. R., & Di Benedetto, A. (2018). Pavement distress detection methods: A review. *Infrastructures*, 3(4), 58. https://doi.org/10.3390/infrastructures3040058.

[35] Ruseruka, C., Mwakalonge, J., Comert, G., Siuhi, S., & Perkins, J. (2023). Road condition monitoring using vehicle built-in cameras and gps sensors: a deep learning approach. *Vehicles*, 5(3), 931-948. https://doi.org/10.3390/vehicles5030051.

[36] Valipour, P. S., Golroo, A., Kheirati, A., Fahmani, M., & Amani, M. J. (2024). Automatic pavement distress severity detection using deep learning. *Road Materials and Pavement Design*, 25(8), 1830-1846. https://doi.org/10.1080/14680629.2023.2276422.

[37] Skraparlis, A. N., Ntalianis, K. S., Ntaliani, M. S., Ntalianis, F. S., & Mastorakis, N. E. (2024). Detecting Indoor Tiny Autonomous Malicious Drones within Critical Infrastructures: An Innovative Algorithm based on Harmonic Radar-Equipped Mini-Drones. *WSEAS Transactions on Information Science and Applications*, 21, 466-479. https://doi.org/10.37394/23209.2024.21.42.

[38] Onaolapo, A. K., & Abe, B. T. (2024). Innovative Feature Analysis of Electric Vehicle Technology, Charging Infrastructure, Power management, and Control Methods. *WSEAS Transactions on Power Systems*, 19, 350-359. https://doi.org/10.37394/232016.2024.19.30.

[39] Barrile, V., Bilotta, G., Fotia, A., & Bernardo, E. (2020). Road extraction for emergencies from satellite imagery. In *Computational Science and Its Applications–ICCSA 2020: 20th International Conference, Cagliari, Italy, July 1–4, 2020, Proceedings, Part IV 20* (pp. 767-781). Springer International Publishing. Cagliari, Italy. https://doi.org/10.1007/978-3-030-58811-3_55.

[40] Ballesteros, J. R., Sanchez-Torres, G., & Branch, J. W. (2021, August). Automatic road extraction in small urban areas of developing countries using drone imagery and Image Translation. In *2021 2nd Sustainable Cities Latin America Conference (SCLA)* (pp. 1-6). IEEE. Medellin, Colombia. DOI: 10.1109/SCLA53004.2021.9540111.

[41] Barrile, V., & Bilotta, G. (2014). Self-localization by laser scanner and GPS in automated surveys. *Computational Problems in Engineering*, 293-311. https://doi.org/10.1007/978-3-319-03967-1_23.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare.