Automated Spatial Mapping and Localised 3D Dimensional Extraction of Pavement Distress Using Deep Learning and GCP-Free UAV Photogrammetry

Sasmito, B.,* Manurung, E. F. J., Mahanani, A. K. and Qoyimah, S.
Department of Geodetic Engineering, Faculty of Engineering, Universitas Diponegoro, Semarang, Central Java, Indonesia, E-mail: bandisasmito@live.undip.ac.id, ORCID ID: https://orcid.org/0000-0001-8637-6727
**Corresponding Author*

DOI: https://doi.org/10.52939/ijg.v22i8.5140

Abstract
Traditional pavement monitoring is labour-intensive, and while automated 2D deep learning approaches improve detection speed, they lack the critical depth information needed for structural severity assessment. Conversely, high-fidelity 3D photogrammetry typically demands labour-intensive Ground Control Points (GCPs). To bridge this gap, this study proposes a lightweight, GCP-free framework integrating UAV-RTK photogrammetry and the YOLOv11m algorithm to automatically detect and geometrically quantify 3D pavement distress. Using the six primary distress categories (ASTM D6433), the trained model was deployed directly into a geographic information system (GIS) environment via the QGIS Deepness plugin. To mitigate the inherent absolute elevation shifts in GCP-free mapping and global road gradients, a novel localised 3D depth extraction algorithm was introduced. This approach dynamically calculates the relative difference between a defect's localised Z-minimum and its surrounding reference plane on the Digital Surface Model (DSM). Validation in two study areas demonstrated overall detection accuracies of 81.74% and 88.79%. Planimetrically, the model achieved a Root Mean Square Error (RMSE) of 0.13 m for potholes and 0.14 m for alligator cracking. Crucially, the localised algorithm extracted vertical distress dimensions with centimetre-level precision, yielding exceptional depth RMSEs of 0.021 m (potholes) and 0.019 m (rutting). These findings demonstrate that integrating GCP-free spatial data with localised AI extraction yields a highly scalable, ready-to-use solution for modern Pavement Management Systems (PMS), while highlighting the need for instance segmentation to address bounding-box limitations for elongated deformations.

Keywords: Pavement Distress Detection, YOLOv11, Unmanned Aerial Vehicle (UAV), Digital Surface Model (DSM), 3D Depth Extraction; Photogrammetry

1. Introduction

Road infrastructure requires continuous monitoring to ensure transportation safety and serviceability. Pavement deterioration, including cracks, potholes, and rutting, progressively reduces structural performance and increases accident risks and vehicle damage [1] and [2]. In addition, delayed detection of pavement distress can significantly increase maintenance costs and reduce road service life. Conventional inspection methods are predominant based on manual visual surveys and vehicle-mounted systems, which are labour-intensive, time-consuming, and prone to subjectivity, leading to inconsistent, non-repeatable assessments [3] and [4]. Furthermore, these methods lack sufficient spatial accuracy for quantitative evaluation of pavement conditions [5]. Traditional inspection approaches and automated vehicle-based systems also face limitations in scalability, cost, and operational safety, particularly in high-traffic environments [6] and [7].

To overcome these limitations, unmanned aerial vehicles have been widely adopted for road inspection due to their flexibility, cost efficiency, and ability to capture high-resolution imagery over large areas [8][9] and [10]. UAV-based photogrammetry enables rapid data acquisition and supports the generation of three-dimensional surface models, allowing detailed geometric analysis of pavement distress [11] and [12]. This capability facilitates accurate mapping of road surface conditions, including in hazardous or difficult-to-access locations [13]. Moreover, UAV systems improve operational efficiency by integrating spatial data acquisition with automated analysis workflows [14] and [15]. Previous studies have demonstrated that

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online)

UAV-derived data can effectively support monitoring of road geometry and surface defects under varying environmental and operational conditions [16] and [17].

Recent advances in deep learning, particularly the YOLO (You Only Look Once) architecture, have significantly accelerated automated pavement distress detection. Convolutional neural networks are capable of extracting complex visual patterns from high-resolution imagery and have demonstrated superior performance compared to traditional image processing techniques [18] and [19]. Previous studies of this architecture have shown different levels of accuracy and computational efficiency. For example, an SSD-based implementation achieved a precision and recall of 71%-75% for multi-class damage detection in smartphone images [20]. Two-stage detectors like Faster R-CNN achieve high detection precision but are slow in inference, which is not suitable for fast and large-scale assessment. On the other hand, YOLO variants always have the best balance. Recent implementations using YOLOv5, modified YOLOv8, and other strong architectures have produced highly competitive mean Average Precision (mAP) values, usually between 80% and over 90%, while maintaining near real-time processing capabilities for both ground-based and UAV inspections [19] [20] and [21] This ultimately justifies the widespread use of these architectures for automated pavement inspection. Among these, the YOLO algorithm is widely adopted due to its real-time detection capability and high accuracy in object localisation and classification [21] and [22]. YOLO-based models enable efficient multi-class detection of pavement defects through bounding-box prediction, making them highly suitable for large-scale road inspection applications [23] [24] [25] [26] and [27].

Despite these advancements, several limitations remain. Most existing approaches rely on two-dimensional image analysis, which only provides planimetric information. For example, while recent mobile mapping systems integrating YOLOv8 and GNSS receivers have successfully achieved real-time road damage detection and precise planimetric geolocation [28], they inherently lack the depth (Z-axis) data. This limitation restricts the ability to evaluate structural severity, particularly for defects such as potholes and rutting, where depth information is essential [29]. Furthermore, bounding-box-based object detection is not optimal for representing elongated or continuous defects, resulting in geometric inaccuracies in dimension estimation [23]. These issues reduce the reliability of automated pavement condition assessment systems. Therefore, a critical gap exists in integrating automated detection with accurate geometric quantification of pavement distress.

Cracks in pavement are categorised by their failure mechanisms and geometric shapes, following the ASTM classification system. Alligator skin cracks exhibit structural fatigue and form interconnected, polygonal patterns. Longitudinal and transverse cracks are linear, primarily resulting from shrinkage caused by temperature fluctuations or the reflection of joints. Conversely, edge cracks develop near the pavement's edges due to insufficient lateral support. Potholes and patches indicate significant localised damage and prior maintenance efforts. Additionally, grooves and rutting signify surface deterioration in the wheel path.

Three-dimensional data obtained from UAV photogrammetry has been explored to address this issue. Digital Surface Models provide detailed elevation information that enables quantitative assessment of vertical deformation and defect depth [30] [31] and [32]. These models allow the extraction of geometric parameters such as depth and surface deformation by analysing elevation differences between damaged and intact pavement areas. Furthermore, integration between image-based detection and 3D spatial data has been explored to improve geometric accuracy and enable metric analysis of defects [7] and [33]. However, the integration between deep learning detection outputs with elevation data for comprehensive, scalable dimensional analysis remains limited. In addition, the accuracy of photogrammetric reconstruction is influenced by factors such as flight parameters, camera configuration, and ground control distribution, which affect the reliability of spatial measurements [34] and [35]. Recent implementations further validate the supremacy of YOLO models for UAV-based pavement imagery, achieving near-real-time processing speeds with robust multi-class accuracy for urban road infrastructure [36] [37] [38] and [39].

However, the current state of the art remains fundamentally constrained by its two-dimensional paradigm. Most existing deep learning frameworks merely output 2D bounding boxes or pixel masks, providing only the planar location of the distress [40] [41] [42] [43] and [44]. While specialised studies have explored 3D pavement assessments using LiDAR sensors and data fusion [35], these high-fidelity systems are often hindered by prohibitive hardware acquisition costs, intensive computational requirements, and complex payload configurations, limiting their large-scale operational feasibility [41]. On the other hand, translating 2D AI detections into actionable 3D engineering metrics using UAV photogrammetry is highly complex. The global gradient and cross-slope (camber) of the road surface inherently distort absolute elevation readings, making direct depth extraction from a Digital Surface Model highly complex [31] and [32].

Furthermore, UAV-based photogrammetric approaches typically rely on numerous physical Ground Control Points (GCPs) to achieve engineering-grade accuracy, which drastically negates the time-saving benefits of drone surveys [35] and [36]. Recent evaluations indicate that automated dimensional extraction from UAV-derived 3D models inherently introduces geometric deviations, with measurement errors often reaching several centimetres [11]. Consequently, a critical gap exists for a lightweight, GCP-free framework that seamlessly integrates AI-driven detection with localised 3D dimensional extraction to mitigate these discrepancies.

To bridge this gap, this study proposes a novel framework that integrates UAV-RTK photogrammetry with the state-of-the-art YOLOv11m algorithm for automated 3D pavement condition assessment. The proposed methodology explicitly targets multi-class distress detection, focusing on translating AI-generated bounding boxes into precise engineering metrics. By deploying the trained inference model directly within a Geographic Information System (GIS) environment via the QGIS Deepness plugin, this research seamlessly couples spatial orthomosaics with deep learning outputs. Furthermore, to address the challenges of global road gradients, this study introduces a localised depth extraction algorithm that dynamically calculates the relative elevation difference between a defect's Z-minimum and its immediate reference plane in the Digital Surface Model (DSM). Ultimately, this research demonstrates the geometric viability of a strictly GCP-free UAV-RTK pipeline, providing a highly scalable, robust, and cost-effective solution tailored for modern Pavement Management Systems (PMS).

2. Methodology

2.1 Study Area

This study was conducted in East Ungaran Subdistrict, located in Semarang Regency, Central Java, Indonesia. Two specific road segments in this region suffer from severe pavement distress. These roads require frequent monitoring and maintenance; therefore, they were selected as the primary experimental subjects for this research. The spatial distribution and location of the study areas are presented in
Figure **1**.

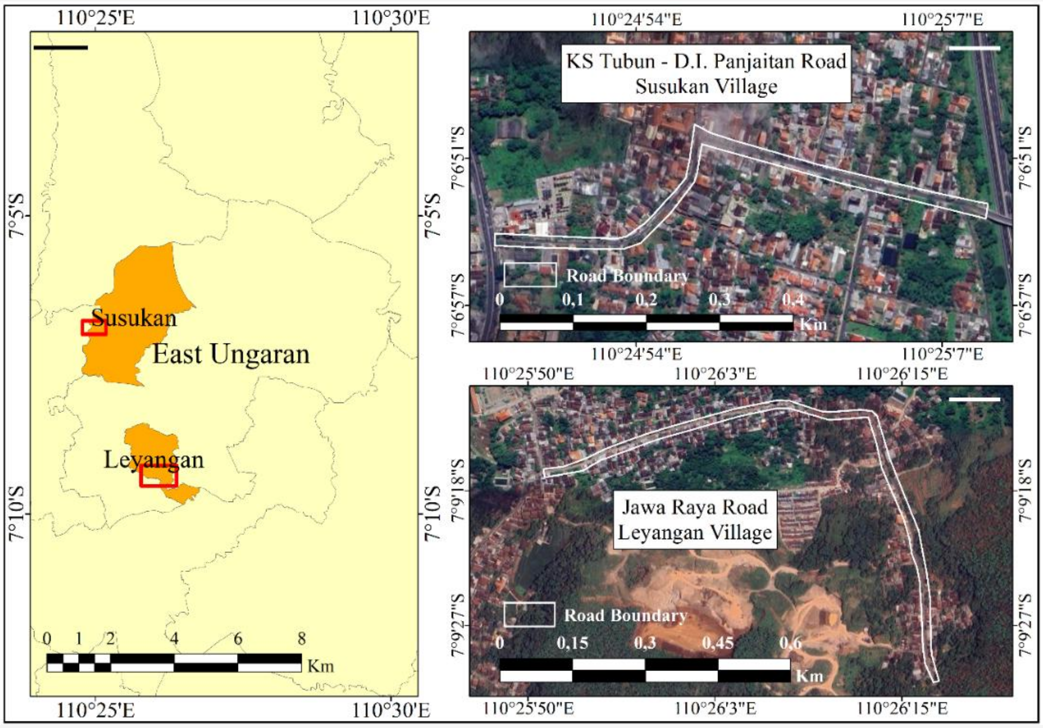

Figure 1: KS Tubun- D.I. Panjaitan Road in Susukan, and Jawa Raya Road in Leyangan

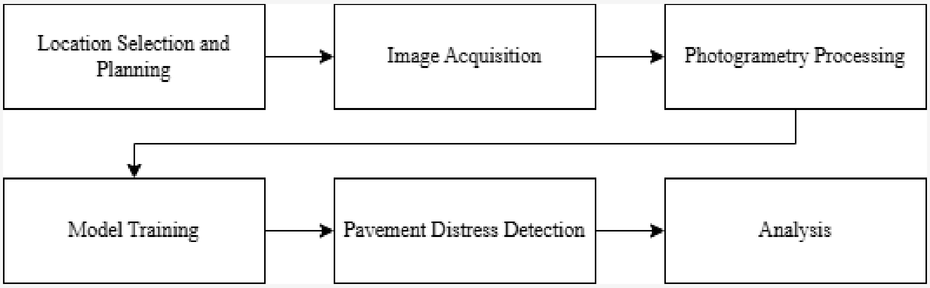

Figure 2: Pavement distress detection using UAV study workflow

The areas of interest in this experiment comprised two separate routes. The first location is Jawa Raya Road, situated in Leyangan Village. The second location encompasses the connected route of KS Tubun Road and D.I. Panjaitan Road, located in Susukan Village. Jawa Raya Road measures approximately 1,300 meters, while the combined segment in Susukan spans approximately 800 meters. These road sections are highly exposed to varying traffic loads and environmental conditions, making them highly prone to surface degradation over time. Their suitability for automated inspection purposes directly influenced the selection of these locations. Both routes present clearly visible surface defects, including extensive cracks, deep potholes, and measurable rutting. Furthermore, these locations provide excellent access for unmanned aerial vehicle (UAV) imaging, enabling safe drone operations and straightforward verification of field data against the aerial imagery.

2.2 Research Design and Workflow

As shown in Figure 2, this study began with site selection and flight planning. A suitable road section with visible pavement distress was identified for monitoring. Real-Time Kinematic (RTK) positioning integrated directly into the unmanned aerial vehicle ensured high georeferencing accuracy. Direct georeferencing using GNSS and inertial systems enables accurate determination of exterior orientation parameters without relying on extensive ground control [42].

This precise positioning method was used to minimise fieldwork duration, deliberately omitting the placement of physical Ground Control Points (GCPs) to evaluate the RTK system's standalone capability for photogrammetric reconstruction. Images were acquired using the unmanned aerial vehicle under planned flight parameters to ensure adequate coverage and image overlap for reliable processing. High image overlap is essential to satisfy stereoscopic principles and ensure accurate 3D reconstruction from photogrammetric datasets [43]. The collected images were subsequently processed using photogrammetry to generate orthomosaics and Digital Surface Models (DSM). These datasets were then analysed using the YOLO deep learning algorithm for automated multi-class distress detection. The model localised and classified six primary pavement distresses based on ASTM D6433 guidelines across the study area. Finally, an assessment evaluated the reliability of the detection results using a confusion matrix. This study also evaluated the accuracy of planar damage dimensions extracted from bounding boxes and the depth extraction accuracy for potholes and rutting based on the DSM.

2.3 Data Acquisition

A reconnaissance survey established the data collection workflow prior to image acquisition. This study utilised a Real-Time Kinematic (RTK)-enabled unmanned aerial vehicle to capture road conditions. This advanced technology ensured high positional accuracy for generating georeferenced orthomosaics and Digital Surface Models. Traditional photogrammetry requires the physical placement of Ground Control Points (GCPs) across the survey area. This research, however, omitted the use of physical ground control to test the system's standalone efficiency, relying instead on recording precise spatial coordinates directly into image files during flight. The aerial system received continuous positioning correction signals to maintain high spatial consistency throughout the image acquisition phase. The World Geodetic System 1984 coordinate system standardised the spatial data for processing in the photogrammetry software. While this direct georeferencing approach significantly reduced fieldwork duration and preserved the dense relative spatial precision required to evaluate local distress dimensions, it acknowledged the potential for absolute elevation shifts. The survey utilised specific positioning parameters to ensure optimal data quality. The observation interval was set to 1 second, and the spatial data collection required a Position Dilution of Precision (PDOP) value below 3 to ensure favourable satellite geometry. These configurations supported accurate photogrammetric reconstruction and prevented severe spatial distortion across the lengthy road segments.

Furthermore, this study utilised 24 Independent Check Points (ICPs), 12 evenly distributed in Leyangan and 12 in Susukan, to validate the horizontal and vertical geometric quality of the output products. The evaluation assessed the horizontal and vertical accuracy of the resulting orthophoto and Digital Surface Model. The field survey acquired these ICPs using a GNSS Tersus David receiver. The data collection employed an RTK NTRIP method. This approach provided reliable ground truth coordinates for spatial verification. Table 1 presents the technical specifications of the aerial and ground receivers. This study established specific flight parameters to accurately evaluate pavement distress. All aerial images were acquired using a DJI Mavic 3 Enterprise aircraft equipped with an RTK module for precise spatial positioning. The module achieves a horizontal accuracy of 1 cm + 1 ppm and a vertical accuracy of 1.5 cm + 1 ppm. The aircraft features a wide-angle camera with a 4/3 CMOS sensor, capturing 20-megapixel images using a 24-millimetre equivalent focal length, while a mechanical shutter prevents motion blur during aerial mapping operations.

The image acquisition was performed in a single line along the selected road segments at a flight altitude of 25 m above the ground level to obtain a sub-centimetre Ground Sampling Distance (GSD) of 7.6-8.5 mm/pixel. Such high spatial resolution is a key scientific need to ensure that fine-scale pavement damage (on the order of millimetres), e.g., narrow alligator cracks, can be visually discerned and reliably detected by deep learning algorithms while maintaining a safe operational distance from nearby urban infrastructure. The unmanned aerial vehicle operated in terrain-following mode throughout the survey, allowing the sensor to maintain a constant altitude relative to the varying road topography. This uniform altitude ensured consistent image resolution and ground sampling distance throughout the study area. Flight height and ground sampling distance are critical parameters that directly influence spatial resolution and accuracy in UAV-based photogrammetry [43].

The camera configuration used a vertical downward (nadir) orientation to capture the planar geometry required for accurate, detailed surface modelling. The data collection strategy also employed a dense 90% image overlap. This substantial overlap is a key geometric requirement to compensate for the absence of ground control since this study examines a photogrammetric workflow that is entirely free of GCPs. This results in a large geometrical redundancy, which enhances the bundle block adjustment and reduces the internal spatial distortions. This high overlap rate yielded a robust tie-point network for subsequent three-dimensional reconstruction, mitigating internal distortion and ensuring highly reliable relative spatial extraction despite the deliberate absence of physical ground control points. Ultimately, these comprehensive flight configurations optimised the spatial data quality necessary for the deep learning detection process.

2.4 Photogrammetric Processing

All acquired image datasets underwent photogrammetric processing using an automated photogrammetric pipeline. The initial step involved image alignment, where the software performed photogrammetric triangulation to align the overlapping images, utilising the embedded RTK spatial coordinates. This process established a precise spatial framework and automatically georeferenced the model before proceeding with the orthomosaic reconstruction and the Digital Surface Model (DSM). The spatial framework was then refined to process only the selected road segments, effectively reducing processing time, conserving computational resources, and excluding surrounding vegetation and residential structures to prevent misdetection. The geometric precision for the reconstruction was set to the highest parameters to ensure the resulting spatial models remained sharply defined and free from blurring artefacts. This clarity is critical for improving visibility of fine surface details, such as cracks and potholes. The detailed parameters and error margins of this photogrammetric processing are summarised in Table 2. Processing parameters of two study locations in the villages of Leyangan and Susukan. It should be noted that the planned flight altitude was 25 m, but the UAV was in terrain-following mode over undulating road gradients. Therefore, the average effective flying heights calculated by the photogrammetric software were 28 m and 25.3 m, respectively, according to the final reconstructed 3D space.

Table 1: GNSS parameters

Aspect	Horizontal Accuracy	Vertical Accuracy
DJI Mavic 3 Enterprise RTK Module	1 cm + 1 ppm	1.5 cm + 1 ppm
GNSS Tersus David	10 mm + 1 ppm	15 mm + 1 ppm

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) |

Table 2: UAV-based photogrammetric model processing parameters

Parameters	Leyangan Study Area	Susukan Study Area
Number of Photos	966	580
Flying Height	28 m	25.3 m
Ground Resolution	8.53 mm/pixel	7.6 mm/pixels
Tie Points	2,611,179	2,019,006
Reprojection Error	0.784 pixels	0.787 pixels
DEM Resolution	1.71 cm/pixels	1.52 cm/pixels
Camera Location Error (X – Error (cm))	36.8401	10.9611
Camera Location Error (Y – Error (cm))	33.2029	10.0919
Camera Location Error (Z – Error (cm))	31.8411	20.8394
Camera Location Error (XY Error (cm))	49.5946	14.8994
Camera Location Error (Total Error (cm))	58.9362	25.6178
ICP Horizontal (XY) RMSE (m)	0.237	0.354
ICP Vertical (Z) RMSE (m)	0.113	0.279

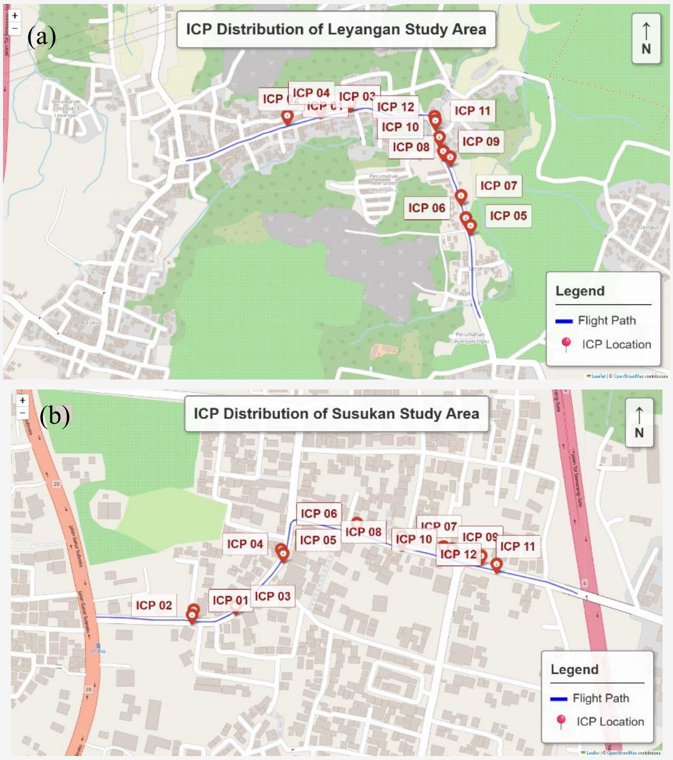

Figure 3: The distribution of independent check points (ICP):
(a) the distribution of ICP in Leyangan and (b) the distribution of ICP in Susukan

The Leyangan dataset had an unexpectedly large total camera location error (58.93 cm), which was different from the best centimetre-level specification of the DJI Mavic 3 Enterprise RTK, due to sporadic degradation of the RTK signal. The short 'float' status was caused by the signal being blocked by surrounding canopy cover and residential structures along the 1.3 km corridor and resulted in deviations from the baseline along the flight path. An Independent Check Point (ICP) validation in the field was performed to explicitly check the absolute accuracy of the resulting photogrammetry model. The evaluation results in Table 2 show the absolute RMSE values of 0.237 m in horizontal (XY) and 0.113 m vertical (Z) directions at the Leyangan site, respectively. The RMSE value of the horizontal in the Susukan location is 0.354 m, while the RMSE value of the vertical is 0.279 m. This absolute vertical shift underlines the intrinsic limits of a fully GCP-free approach but also underlines the need to implement the proposed local relative depth extraction algorithm to neutralise anomalies on the global DSM. Independent Check Points (ICPs) are located at the shoulders and edges of roads shown in Figure 3. The site was selected to enable accurate elevation of the road pavement surface and also to provide safety to the surveyor and minimum disturbance to the traffic in progress. In this study, 12 ICPs were installed at the Leyangan area (1.3 km segment) and another 12 at Susukan (0.8 km segment). The total number of 24 points for linear corridor mapping is above the minimum sample size required to conduct a statistically reliable Root Mean Square Error (RMSE) evaluation of photogrammetric reconstruction results.

2.5 Dataset Preparation and Model Training

Following the generation of the high-resolution orthomosaic, the road damage detection process was initiated using QGIS, Roboflow, and Google Colab. The size of the image pixels is quite large because the orthomosaics cover large road corridors, i.e., 95,133 × 67,829 pixels for the Leyangan site and 87,910 × 22,249 pixels for the Susukan site. These large raster files were then split into 640 × 640 pixel segments using a sliding-window approach to comply with the input requirements of the YOLO algorithm. This grid slicing process theoretically produced a total grid of 20,624 tiles of raw data, which were made up of 15,794 tiles for the Leyangan site and 4,830 tiles for the Susukan site. However, the corridor-shaped UAV flight paths resulted in raw tiles covering most of the non-road surroundings (e.g., residential houses, vegetation, and bare soil). The dataset was subjected to strict filtering and curation to ensure the deep learning model did not learn irrelevant backgrounds and to avoid potential false positives. We filtered out all tiles that did not contain any road features, leaving us with a final dataset of 1,130 high-quality tiles that only contain pavement surfaces to train and evaluate the model. Image labelling was subsequently performed using the Roboflow platform, where bounding boxes were manually drawn to mark the exact locations of pavement distress across every image tile. This stage transforms raw spatial data into structured inputs for object detection.

Data augmentation increased the dataset variation without requiring additional field collection. The applied operations included rotation, flipping, and brightness adjustments, with every original image tile undergoing six distinct augmentation processes. Each image tile underwent multiple augmentation processes. The dataset was then divided into training, validation, and test subsets with ratios of 70%, 15%, and 15%, respectively. The dataset is divided into train, validation, and test sets at ratios of 70%, 15%, and 15%, respectively. A portion of the training (70%) is only used for training the model and updating its weights. A fraction of the validation (15%) is used to track the model's performance in the training phase, to tune the hyperparameters, and to early stop in order to prevent overfitting. Finally, 15% of the tests were used as entirely new data (the model had never seen before) for an unbiased evaluation of the detection accuracy of the final model.

The YOLOv11 model, specifically YOLOv11m, was trained using the prepared dataset through Google Colab. The network accepts images that are normalised to a size of 640 × 640 pixels. Training was conducted for 100 epochs using the Stochastic Gradient Descent optimiser with a batch size of 8. During each epoch, the model was trained on the entire dataset, which consists of 791 augmented image segments, to iteratively optimise its feature extraction capabilities. The training utilised GPU acceleration with a Tesla T4 to improve computational efficiency. A patience parameter of 20 was applied to control early stopping during training. The training and validation process enabled the model to learn object characteristics and optimise detection performance through transfer learning using pre-trained weights. The training process also produced outputs such as execution time per epoch and total training duration for performance analysis.

The trained model was implemented for automated detection using the Deepness plugin in QGIS. Model evaluation was conducted using precision, recall, mean Average Precision, and F1-score, supported by a confusion matrix to provide a comprehensive assessment of detection performance.

2.6 Automated Detection and Spatial Analysis

Finally, automated object detection was executed. The trained YOLOv11m model was implemented using the Deepness plugin in QGIS to identify and localise pavement distress directly on the orthomosaic. The inference process used a tile size of 640 pixels with 50% overlap to minimise missed detections at tile boundaries. A confidence threshold of 0.25 and an Intersection over Union threshold of 0.50 were used to control detection sensitivity and reduce duplicate bounding boxes. The detection results were generated as spatial vector outputs representing the location and distribution of pavement distress. These outputs were then utilised for further spatial analysis, including the calculation of damage density and the projection of detected objects onto the Digital Surface Model to support geometric evaluation.

2.7 Analysis

The ground truth for the detection model assessment was established using manual visual inspection and digital orthophoto analysis. Field data collection involved the physical measurement of the dimensions of pavement damage at the site, providing highly reliable reference data (ground

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

truth). The researcher employed a standard fibreglass tape, which was stretched along the visible cracks and patches to obtain planimetric dimensions (length and width). To measure vertical dimensions, such as the depth of holes and grooves, a rigid straight bar is positioned across an intact section of the pavement above the damaged area. The maximum vertical drop of the bar to the deepest point of deformation is then measured using a steel ruler. Several samples of cracks, potholes, and rutting directly on the road surface are measured. These physical measurements served as a reliable reference for evaluating the model's geometric output. Additionally, the photogrammetrically derived orthophotos verified the overall spatial distribution of the surface damage [14]. This dual approach ensures comprehensive, verifiable distress annotations for evaluating detection performance.

This study utilised a confusion matrix during the validation phase to evaluate the multi-class detection results. The confusion matrix categorises the model predictions into true positives, false positives, true negatives, and false negatives. This research validates six distinct pavement distress classes: alligator cracks, edge cracks, longitudinal cracks, transverse cracks, patching, potholes, and rutting. The selection of these specific classes is guided by the ASTM D6433 standard for Pavement Condition Index (PCI) surveys. While the standard identifies nineteen types of asphalt distress, these six were prioritised because they represent the most critical structural failures frequently observed in the study areas and are visually distinct for extraction from nadir-view UAV orthophotos. This research validated six different categories of pavement damage: alligator skin cracks, edge cracks, longitudinal and transverse cracks, patching, potholes, and rutting. The selection of these specific categories is based on the ASTM D6433 standard for the Pavement Condition Index (PCI) survey, as previously outlined in the study's background. With reference to these standards, this matrix presents a breakdown of the reliable, technically precise classification accuracy for these types of damage. By grounding the classification in this standard, the matrix provides a reliable, technically sound breakdown of classification accuracy across these distress types, serving as the foundation for calculating standard evaluation metrics. Precision, recall, and the F1-score are standard metrics used to evaluate object detection models, particularly when handling imbalanced datasets where distress instances are sparse compared to background pavement pixels [44]. The calculations for precision, recall, and the F1-score are defined in Equations 1 to 3.

$$Precision=\frac{TP}{TP+FP}$$

Equation 1

$$Recall=\frac{TP}{TP+FN}$$

Equation 2

$$F1=2\times\frac{Precision\times Recall}{Presision+Recall}$$

Equation 3

Where TP is true positives, FP is false positives, and FN is false negatives.

Precision quantifies the proportion of detected defects that are actual defects in the imagery, while recall reflects the model's ability to detect actual defects among all existing surface damage. High precision suggests the detected damage instances are mostly accurate, whereas high recall indicates the model successfully detects most existing road damage. The F1-score serves as a balanced indicator of overall detection performance, balancing the identification of all true defects while minimising false alarms. Furthermore, this study evaluated the geometric accuracy of the predicted bounding boxes and the extracted depths. The planar dimensions of the bounding boxes were directly compared with the physical field measurements to assess the reliability of extracting the damage length and width from the object detection output. The deviation between the model-predicted dimensions and the actual physical field measurements was quantified using the Root Mean Square Error (RMSE). The calculation is determined using the following Equation 4.

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(R-R_i\right)^2}$$

Equation 4

Where:

R = Reference value (the actual physical length, width, or depth measured in the field)

R_i = Measurement result (the predicted length, width, or depth extracted from the model)

n = Number of measurements

This $RMSE$ calculation is quantifying, in particular, the 1-dimensional geometric deviation between the bounding box dimensions (or DSM depth estimations) and the true physical dimensions of the distress. To quantify the depth of potholes and rutting, it is important to pay attention to a two-step workflow. Initially, the YOLO algorithm was only used to detect 2D bounding boxes on the RGB orthomosaic; the Digital Surface Model (DSM) was not used in the deep learning inference. Subsequently, the spatial limits of these 2D bounding boxes were projected geographically onto respective DSMs. Then, a localised extraction algorithm was implemented within each projected bounding box area to calculate the relative depth of the distress. A localised extraction algorithm was implemented for each bounding box. The relative depth was calculated by establishing a local reference plane based on the average elevation of the pixels along the bounding box perimeter, which represents the surrounding intact asphalt. The maximum depth of the distress was then determined by subtracting the lowest elevation value (Z-minimum) found within the bounding box area from this local reference plane. This relative calculation ensures that global DSM elevation anomalies do not interfere with the local defect measurement.

To formalise the localised 3D depth extraction algorithm computationally, the process is mathematically defined in three consecutive steps. First, let $B = \{P_{b,1}, P_{b,2, \dots} P_{b,N}\}$ be the set of pixels located strictly along the perimeter boundary of the predicted bounding box, where N represents the total number of perimeter pixels. The localised reference plane elevation Z_{ref}, which represents the surrounding intact asphalt surface, is calculated as the arithmetic average of these boundary elevations using Equation 5.

$$Z_{ref} = \frac{1}{n}\sum_{i=1}^{n} Z_{b,i}$$

Equation 5

Where $Z_{b,i}$ is the absolute elevation of the *i-th* pixel on the boundary set B.

Second, let $A = \{P_{a,1}, P_{a,2, \dots} P_{a,M}\}$ be the set of all pixels located within the interior area of the bounding box. The minimum absolute elevation point (Z_{min}), which corresponds to the maximum vertical depression inside the distress area, is defined using a minimisation function with equation 6.

$$Z_{\min} = \min_{p_j \in A}\left(Z_{a,j}\right)$$

Equation 6

Where $Z_{a,j}$ is the absolute elevation of the *j-th* pixel within the interior area set A.

Finally, the actual maximum relative depth (D_{max})of the pavement distress, fully isolated from global Digital Surface Model (DSM) shifts and local road cross-slopes, is calculated by finding the net vertical distance with Equation 7.

$$D_{max} = Z_{ref} - Z_{min}$$

Equation 7

This three-step mathematical formulation ensures that the vertical distress depth is extracted relative only to its immediate undistressed surroundings, making the framework robust against terrain topography variations.

3. Results and Discussion

3.1 Model Detection Performance

The YOLOv11 model's detection results are presented in Figure 4, showing examples of pavement distress identified from UAV orthophoto imagery. The model successfully detects several damage types, including potholes, patching, alligator cracking, and longitudinal and transversal cracking. Each detected object is represented by a bounding box with an associated confidence score indicating the probability of correct prediction. The model was trained using a dataset of 1,130 annotated images. The dataset consisted of 791 training images, 169 validation images, and 170 test images. Multiple distress instances are frequently detected within a single image, demonstrating the model's ability to simultaneously recognise different pavement damage patterns.

Moreover, multiple distress cases are often detected in one image, which demonstrates the model is able to recognise different patterns of pavement damage simultaneously. On the other hand, the model also demonstrates a strong true-negative performance, as shown in Figure 4(d), by correctly avoiding false detections in intact pavement areas with no structural distress. Detection performance was evaluated using Precision-Recall analysis and mean Average Precision metrics. The model achieved an overall mAP@0.5 value of 0.499. While this metric indicates moderate performance in general object detection, it is highly representative of the inherent complexities of multi-class pavement distress detection, where distress features blend heavily with the background asphalt texture. Patching obtained the highest Average Precision value of 0.768, followed by potholes with 0.701 and alligator cracking with 0.654, indicating that defects with clear boundaries and strong texture contrast are more reliably detected.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

Figure 4: Pavement distress detection results using the YOLO deep learning model from UAV orthophoto imagery: (a) True positive (b) False positive (c) False negative (d) True negative

Lower performance was observed for edge cracking and rutting, with Average Precision values of 0.270 and 0.147, respectively. These distress types exhibit subtle or elongated deformation patterns that are visually similar to surrounding pavement textures in nadir UAV imagery. In addition, class imbalance in the training dataset may limit the model's ability to learn representative features for these categories. The F1 Confidence analysis indicates an optimal detection threshold at a confidence value of approximately 0.25 with an F1 score of about 0.50. These evaluation results provide the basis for further spatial implementation of the trained model in QGIS for automated pavement distress extraction across the study area.

3.2 Automated Pavement Distress Detection

The automated detection of road damage used the Deepness plugin integrated into QGIS, as shown in Table 3. The process used a pre-trained YOLOv11m model converted to the Open Neural Network Exchange format. The model architecture required an input dimension of 640 x 640 pixels across three RGB colour channels. Therefore, channel mapping in QGIS adjusted the four-band orthomosaic imagery to input only the three primary visual channels. Spatial processing at the first location in Leyangan Village successfully extracted 596 distress objects. Patching accounted for 37.42% of the detections, followed by potholes at 34.56%. The second location in Susukan Village recorded 312 damaged objects. Patching again formed the majority at 51.28%, while potholes accounted for 22.76%. The high patching rate indicates that both road segments have historically undergone frequent structural maintenance. Furthermore, when normalising object counts to the surveyed road length, Leyangan exhibits a higher distress density (approximately 458 defects/km) than Susukan (approximately 390 defects/km). The extreme quantity of potholes and alligator cracks in Leyangan demonstrates that structural degradation at this location is proportionally more massive and critical.

Table 3: Automatic detection results with the deepness plugin in QGIS

Type of Damage	KS Tubun Road	Jawa Raya Road
Alligator Crack		
Edge Crack		
Longitudinal & Transversal Crack		
Patching		
Pothole		
Rutting		

3.3 Detection Accuracy Evaluation

The evaluation assessed the model's performance in classifying six distress types and one non-distress class using confusion matrices. The overall validation sample size was determined statistically using Slovin's formula. Subsequently, a proportional stratified sampling technique was applied to ensure the validation samples accurately represented the total automated detection population across each damage class. Testing the Susukan dataset, which comprised 367 samples, yielded an overall accuracy of 81.74%. The accuracy evaluation for this location is depicted in Figure 5, with the detailed metrics presented in Table 4.

The model proved highly reliable in detecting potholes, achieving a recall of 98.08% and an F1-score of 90.27%. However, alligator cracks presented a significant challenge. Despite achieving a perfect recall of 100%, its precision drastically dropped to 45.45%. This severe metric imbalance indicates a strong over-prediction behaviour by the model; it achieved high recall by aggressively misclassifying

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

non-distress features such as rough pavement textures, tree shadows, and old patches as alligator cracks, resulting in an inflated false-positive rate. The model also showed low recall for edge cracks, detecting only 50% of the actual defects. This occurred because surrounding vegetation covered the road edges, and the shoulder boundaries blended with the asphalt in the nadir orthophotos. The performance trends across different classes for Susukan are further illustrated in the precision, recall, and F1-score graph in Figure 5. The Kappa Coefficient for the Susukan location was 0.7584, indicating substantial agreement and confirming good predictive reliability.

The Leyangan dataset evaluation demonstrated superior performance, achieving an overall accuracy of 88.79% across 571 samples. The accuracy evaluation for these results is shown in Figure 6, and the corresponding accuracy metrics are listed in

Table **5**. The model showed significantly better adaptability, with almost all classes recording F1-scores above 83%. Alligator crack precision improved dramatically to 95.59% because the orthophoto at Leyangan displayed clearer texture contrast and crack patterns, preventing confusion with non-distress areas. However, the rutting class had the lowest precision of 66.67% due to severe class imbalance, as the evaluation dataset contained only 2 actual rutting samples. This minority-data constraint made the evaluation metric highly sensitive to a single detection error. The graphical representation of the precision, recall, and F1-score for the Leyangan location is provided in Figure 6.

Table 4: Detection accuracy evaluation metrics

Configuration	Precision %	Recall %	F1 score %
Alligator Crack	45.45	100.00	62.50
Edge Crack	85.71	50.00	63.16
Longitudinal Crack	77.27	87.18	81.93
Patching	81.74	90.38	85.84
Potholes	83.61	98.08	90.27
Rutting	85.71	85.71	85.71
Non-Distress	85.22	72.06	78.09
Accuracy (P_O)		81.74	

Table 5: Detection accuracy evaluation metrics

Configuration	Precision %	Recall %	F1 score %
Alligator Crack	95.59	94.20	94.89
Edge Crack	96.30	81.25	88.14
Longitudinal Crack	79.59	95.12	86.67
Patching	88.19	90.07	89.12
Potholes	87.50	97.54	92.25
Rutting	66.67	100.0	80.00
Non-Distress	89.58	78.66	83.77
Accuracy (P_O)		88.79	

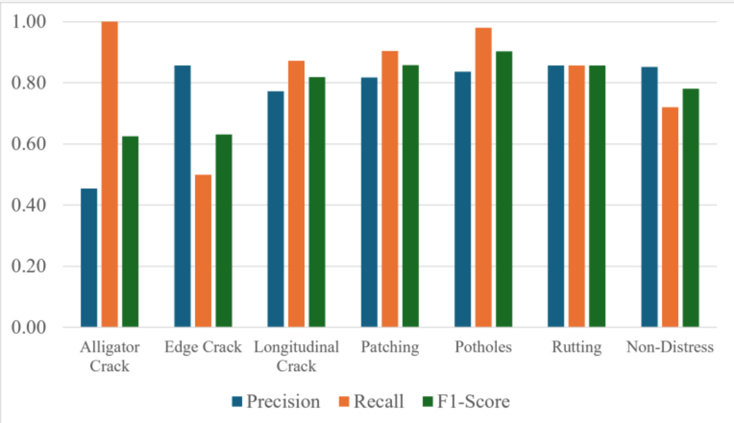

Figure 5: Precision, recall, and F1-score graph

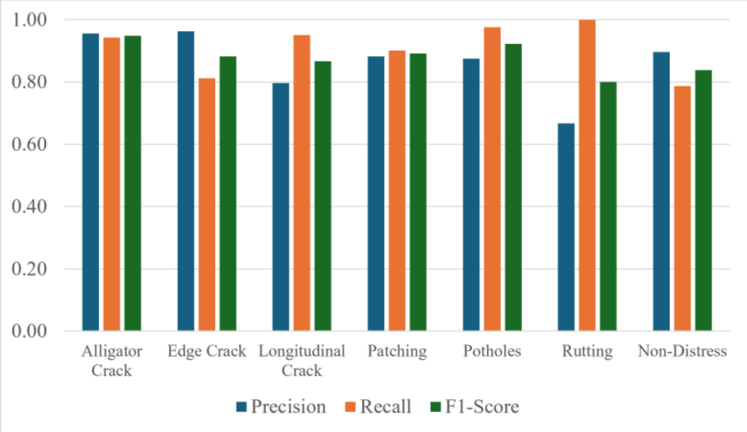

Figure 6: Precision, recall, and F1-score graph

Table 6: Planimetric dimensional accuracy test

Class	RMSE (m)
Alligator Cracking	0.14
Edge Cracking	0.89
Transverse Cracking	1.27
Patching	0.86
Pothole	0.13
Rutting	2.86

3.4 Geometric Validation of Distress Dimensions

The validation of the geometric accuracy was performed by directly comparing the automated spatial dimensions (length, width, and depth) extracted by the model with the physical field measurements (ground truth) obtained during the manual survey. The Root Mean Square Error (RMSE) is used to measure the difference between the model predictions and the real physical conditions. The results of the planimetric accuracy validation are summarised in Table **6**. For planimetric dimensions, alligator cracking achieved excellent accuracy with an RMSE of 0.14 metres, indicating the bounding boxes tightly represented the actual defect areas. Edge cracking showed a slightly larger RMSE of 0.89 metres. This larger deviation resulted from elongated crack shapes blending with road shoulders and vegetation, complicating the precise delineation of polygon boundaries. Longitudinal and transversal cracking had an RMSE of 1.27 metres because the algorithm occasionally split continuous cracks into multiple smaller bounding boxes, leading to measurement discrepancies compared to full tape measurements. Patching demonstrated high accuracy, with an RMSE of 0.86 metres, indicating that the algorithm consistently delineated colour and texture differences between old and new asphalt. The largest planimetric deviation occurred in the rutting class with an RMSE of 2.864 metres. The faint, continuous visual characteristics of wheel ruts led to overlapping detection areas, resulting in significant differences in length estimates.

Planimetric deviations are a serious problem for the standard Pavement Condition Index (PCI) calculations (ASTM D6433). The PCI methodology is heavily dependent on the correct measurement of distress density (total area or length of a defect divided by the sample unit area). The rutting dimension overestimation of 2.86 m, or the transverse cracking overestimation of 1.27 m, is artificially increasing the distress density. The overestimation is yielding negative values that are too high and thus an overly pessimistic, inaccurate PCI score.

This limitation could be addressed by adopting instance segmentation methods such as YOLO-Seg in place of standard bounding-box detection. A bounding box encloses a rectangular region that, for irregular or diagonally aligned defects, includes a substantial area of intact pavement outside the defect boundary. Instance segmentation instead delineates the defect using a pixel-level mask that follows its actual contour so that length and width are derived from the true defect boundary rather than an enclosing rectangle. This approach is expected to reduce the planimetric RMSE, although the extent of improvement would need to be verified empirically in future work.

Furthermore, depth extraction (Z-axis) validation from the Digital Surface Model proved critical for determining distress severity. The vertical accuracy evaluation is presented in Table 7.

Table 7: Depth extraction accuracy test

Class	RMSE (m)
Pothole	0.021
Rutting	0.019

However, it is necessary to distinguish between the absolute vertical accuracy of the photogrammetric model and the relative depth accuracy of the localised distress features. As indicated in Table 2, the overall DSM exhibited absolute Z-errors of 31.84 cm and 20.83 cm due to the absence of physical ground control points. To circumvent this absolute spatial shift, depth evaluation was performed relatively. By calculating the elevation difference between the localised intact pavement at the bounding box perimeter and the lowest point (Z-minimum) within the defect, the depth extraction demonstrated exceptional relative vertical accuracy, achieving an RMSE of 0.021 metres for potholes and 0.019 metres for rutting. Therefore, even a raw DSM without GCPs is affected by absolute elevation shifts and global gradient distortions, while the defect is isolated by local relative depth extraction only within the AI-generated bounding boxes. This targeted processing can effectively cancel the raw global spatial errors to achieve much better centimetre-level precision that is highly reliable for structural severity assessments. This confirms that while RTK-enabled UAVs without GCPs may suffer from absolute elevation shifts, their dense point clouds remain highly reliable for extracting relative localised depressions.

To further verify the geometric stability of the localised depth extraction algorithm across varying scales of defect severity, a statistical correlation analysis was performed, as shown in Figure 7. A scatter plot was generated to directly compare the individual physical field measurements (ground truth) against their model-extracted depth counterparts. The analysis yielded a positive linear relationship with a Coefficient of Determination (R^2) of 0.455 and a trendline equation of y = 0.579x - 0.004. This moderate correlation confirms that the localised 3D extraction method responds proportionately to the true vertical deformation of the distresses. While the variance and trendline slope indicate a tendency for the photogrammetric model to systematically underestimate absolute depth, a common inherent limitation in UAV photogrammetry due to the spatial smoothing of point clouds over sharp vertical drop-offs and narrow distresses, the stable linear trend provides empirical evidence of its practical applicability for relative structural severity assessments, especially given the methodology's strict GCP-free constraints.

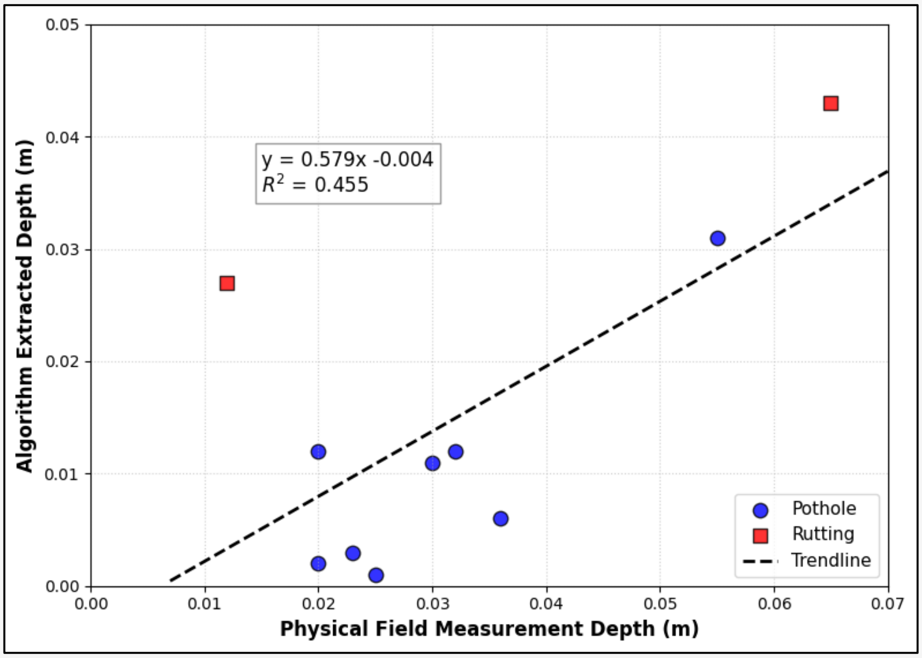

Figure 7: Correlation scatter plot analysis of localised depth extraction

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

4. Conclusions

This study investigated the applicability and reliability of the YOLOv11 deep learning algorithm for automated detection of road surface damage in unmanned aerial vehicle orthomosaics. The results show that the model achieves high classification accuracy when validated under actual field conditions. The algorithm achieved an overall accuracy of 81.74% in Susukan Village and 88.79% in Leyangan Village, with Kappa Coefficient values indicating substantial to almost perfect agreement. The model successfully identified major distress types with clear boundaries, such as potholes and patching. However, detection performance decreased for edge cracking and rutting due to a combination of training data imbalance, subtle visual textures, and limitations in geometric extraction.

Crucially, this research demonstrates the geometric viability of a purely GCP-free UAV-RTK photogrammetric pipeline. By employing a localised 3D depth extraction algorithm that dynamically calculates the relative difference between the localised Z-minimum and the surrounding Z-reference, this method effectively neutralises the absolute elevation anomalies caused by the absence of physical ground control and by global road gradients. The model extracted planimetric dimensions for potholes and alligator cracking, with Root Mean Square Errors (RMSEs) of 0.13 metres and 0.14 metres, respectively. Furthermore, this localised 3D approach enabled the extraction of damage depth for potholes and rutting with centimetre-level accuracy, achieving remarkable vertical RMSEs of 0.021 metres and 0.019 metres, respectively. These findings, coupled with the successful deployment of the model directly within a GIS environment (via the QGIS Deepness plugin), confirm that this lightweight spatial AI framework provides a highly practical and scalable solution for modern Pavement Management Systems (PMS).

The current study also highlights fundamental limitations in applying standard object-detection bounding boxes to continuous, morphologically irregular features. The model produced the largest planimetric deviation for the rutting class with an error of 2.864 metres. This significant deviation is not merely due to faint visual characteristics. However, rectangular bounding boxes are inherently geometrically suboptimal for elongated deformations, often enclosing large areas of un-distressed pavement and unnecessarily overlapping. Future implementations must address this constraint by transitioning from standard object detection to instance segmentation algorithms, such as YOLO-Seg. Utilising polygonal segmentation rather than bounding boxes will provide precise contour delineation, significantly advancing the geometric extraction of continuous structural deformations.

Acknowledgment

The authors would like to express their deepest gratitude to the Department of Geodetic Engineering, Faculty of Engineering, Universitas Diponegoro, Semarang, Indonesia, for the technical facilities, computing, and equipment for this research. Special thanks were also extended to the survey team for their hard work during UAV flight operations, GNSS positioning, and road pavement data sampling.

References

[1] Karballaeezadeh, N., Zaremotekhases, F., Shamshirband, S., Mosavi, A., Nabipour, N., Csiba, P. and Várkonyi-Kóczy, A. R., (2020). Intelligent Road Inspection with Advanced Machine Learning; Hybrid Prediction Models for Smart Mobility and Transportation Maintenance Systems. *Energies,* Vol. 13(7). https://doi.org/10.3390/en13071718.

[2] Wang, J., Wang, K. and Li, K., (2023). Research on Road Damage Recognition and Classification Based on Improved VGG-19. *Mathematical Models in Engineering,* Vol. 9(4); 115–129, https://doi.org/10.21595/mme.2023.23455.

[3] López-González, P. J., Reyes-González, D., Moreno-Vázquez, O., Vivar-Ocampo, R., Zamora-Castro, S. A., Santos Cortés, L. del C., Trujillo-García, B. S. and Sangabriel-Lomelí, J., (2026). Inspection and Evaluation of Urban Pavement Deterioration Using Drones: Review of Methods. *Challenges, and Future Trends, Future Transportation,* Vol. 6(1). https://doi.org/10.3390/futuretransp6010010.

[4] Ranyal, E., Sadhu, A. and Jain, K., (2022). Road Condition Monitoring Using Smart Sensing and Artificial Intelligence: A Review. *Sensors,* Vol. 22(8). https://doi.org/10.3390/s22083044.

[5] Sompie, T. P. F., Makangiras, R. E., Sumajouw, J. A. J. and Hoinbokau, C., (2024). Application of Unmanned Aerial Vehicle and Ground Control Point for Mapping and Road Geometric Review: The Case Study from Pandu to Kima Atas Street. *Engineering, Technology and Applied Science Research,* Vol. 14(4); 15986–15992, https://doi.org/10.48084/etasr.8040.

[6] Bu, T., Zhu, J. and Ma, T., (2022). A UAV Photography–Based Detection Method for Defective Road Marking. *Journal of Performance of Constructed Facilities,* Vol.

36(5); 1–13, https://doi.org/10.1061/(ASCE) CF.1943-550 9.0001748.

[7] Krefft, L. and Hoegner, L., (2025). A Method for Crack Detection and Quantification in Masonry using Neural Network-Based Image Analysis. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* Vol. X-4/W6-202(September); 113–119. https://doi.org/10.5194/isprs-annals-X-4-W6-2025-113-2025.

[8] Tan, Y. and Li, Y., (2019). UAV Photogrammetry-Based 3D Road Distress Detection. *ISPRS International Journal of Geo-Information,* Vol. 8(9). https://doi.org/10.3390/ijgi8090409.

[9] Roberts, R., Inzerillo, L. and Di Mino, G., (2020). Using UAV Based 3D Modelling to Provide Smart Monitoring of Road Pavement Conditions. *Information,* Vol. 11(12). https://doi.org/10.3390/info11120568.

[10] Chathuranga, M., Darwin, N. B., Ahmad, A. B., Sandamali, J. and Gohari, A., (2025). Trends, Advancements, and Future Directions of Unmanned Aerial Vehicle Technology-Based Stockpile Volume Estimation: A Systematic Review. *International Journal of Geoinformatics,* Vol. 21(11); 94–120, https://doi.org/10.52939/ijg.v21i11.4605.

[11] Biçici, S. and Zeybek, M., (2021). An Approach for the Automated Extraction of Road Surface Distress from a UAV-Derived Point Cloud. *Automation in Construction,* Vol. 122. https://doi.org/10.1016/j.autcon.2020.103475.

[12] Darmawan, M. A. H., Mukti, S. N. A. and Tahar, K. N., (2025). Pothole Detection Based on UAV Photogrammetry. *Revue Internationale de Géomatique,* Vol. 34(1); 21–35, https://doi.org/10.32604/rig.2024.057266.

[13] Wróblewska, M. and Grygierek, M., (2025). Uav Observations of Changes in Road Geometry in a Mining Area, Architecture. *Civil Engineering, Environment,* Vol. 18(3); 157–164, https://doi.org/10.2478/acee-2025-0038.

[14] Astor, Y., Nabesima, Y., Utami, R., Sihombing, A. V. R., Adli, M. and Firdaus, M. R., (2023). Unmanned Aerial Vehicle Implementation for Pavement Condition Survey. *Transportation Engineering,* Vol. 12(2). https://doi.org/10.1016/j.treng.2023.100168.

[15] Silva, L. A., Leithardt, V. R. Q., Batista, V. F. L., Villarrubia González, G. and De Paz Santana, J. F., (2023). Automated Road Damage Detection Using UAV Images and Deep Learning Techniques. *IEEE Access,* Vol. 11. 62918–62931. https://doi.org/10.1109/ACCESS.2023.3287770.

[16] Yan, H. and Zhang, J., (2023). UAV-PDD2023: A Benchmark Dataset for Pavement Distress Detection Based on UAV Images. *Data in Brief,* Vol. 51. https://doi.org/10.1016/j.dib.2023.109692.

[17] Zhu, J., Zhong, J., Ma, T., Huang, X., Zhang, W. and Zhou, Y., (2022). Pavement Distress Detection using Convolutional Neural Networks with Images Captured via UAV. *Automation in Construction,* Vol. 133(1). https://doi.org/10.1016/j.autcon.2021.103991.

[18] Cha, Y., Choi, W. and Büyüköztürk, O., (2017). Deep Learning-Based Crack Damage Detection using Convolutional Neural Networks. *Computer-Aided Civil and Infrastructure Engineering,* Vol. 32(5); 361–378. https://doi.org/10.1111/mice.12263.

[19] Hao, Z., Zhu, Y. and Lu, C., (2025). Modified YOLOv8 for Automated Detection and Identification Towards Highway Pavement Distress. *Construction and Building Materials,* Vol. 495(38). https://doi.org/10.1016/j.conbuildmat.2025.143428.

[20] Xu, L., Fu, K., Ma, T., Tang, F. and Fan, J., (2024). Automatic Detection of Urban Pavement Distress and Dropped Objects with a Comprehensive Dataset Collected via Smartphone. *Buildings,* Vol. 14(6). https://doi.org/10.3390/buildings14061546.

[21] Jiang, Y., Yan, H., Zhang, Y., Wu, K., Liu, R. and Lin, C., (2023). RDD-YOLOv5: Road Defect Detection Algorithm with Self-Attention Based on Unmanned Aerial Vehicle Inspection. *Sensors,* Vol. 23(19). https://doi.org/10.3390/s23198241.

[22] Sadhin, A. H., Hashim, S. Z. M., Samma, H. and Khamis, N., (2024). YOLO: A Competitive Analysis of Modern Object Detection Algorithms for Road Defects Detection Using Drone Images. *Baghdad Science Journal,* Vol. 21(6). https://doi.org/10.21123/bsj.2023.9027.

[23] Bao, H., Wu, W., Yu, J., Cao, Y., Liu, Y., Zhang, H., Li, X. and Shi, W., (2024). Application of YOLO Model in Road Defect Detection. *2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT)*. 2280–2284. https://doi.org/10.1109/AINIT61980.2024.10581746.

[24] Ukhwah, E. N., Yuniarno, E. M. and Suprapto, Y. K., (2019). Asphalt Pavement Pothole Detection Using Deep Learning Method Based on YOLO Neural Network. *2019 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 35–40. https://doi.org/10.1109/ISITIA.2019.8937176.

[25] Gopalakrishnan, K., Khaitan, S. K., Choudhary, A. and Agrawal, A., (2017). Deep Convolutional Neural Networks with Transfer Learning for Computer Vision-Based Data-Driven Pavement Distress Detection. *Construction and Building Materials,* Vol. 157(27); 322–330. https://doi.org/10.1016/j.conbuildmat.2017.09.110.

[26] Majidifard, H., Adu-Gyamfi, Y. and Buttlar, W. G., (2020). Deep Machine Learning Approach to Develop a New Asphalt Pavement Condition Index. *Construction and Building Materials,* Vol. 247(17). https://doi.org/10.1016/j.conbuildmat.2020.118513.

[27] Mohand, O. F., Raham, D., Boukerch, I., Mezhoud, S., Saidi, K. and Kariche, I., (2025). A Low-Cost Smartphone-Based Photogram metric Approach for Highway Pavement Defects Detection and Mapping. *International Journal of Geoinformatics,* Vol. 21(11); 59–77. https://doi.org/10.52939/ijg.v2 1i11.4601.

[28] Sasmito, B., Setiadji, B. H. and Isnanto, R. R., (2025). Object Detection: Real-Time Road Damage Detection and Geolocation Using YOLOv8 and GNSS Integration. *Ingénierie des systèmes d'Information*, Vol. 30(9); 2321–2329, https://doi.org/10.18280/isi.300909.

[29] Wang, J., Qin, Q., Zhao, J., Ye, X., Feng, X., Qin, X. and Yang, X., (2015). Knowledge-Based Detection and Assessment of Damaged Roads Using Post-Disaster High-Resolution Remote Sensing Image. *Remote Sensing,* Vol. 7(4); 4948–4967. https://doi.org/10.3390/rs70404948.

[30] Zhu, J., Gao, Y., Huang, S., Bu, T. and Jiang, S., (2023). Measuring Surface Deformation of Asphalt Pavement via Airborne LiDAR: A Pilot Study. *Drones,* Vol. 7(9). https://doi.org/10.3390/drones7090570.

[31] Saad, A. M. and Tahar, K. N., (2019). Identification of Rut and Pothole by Using Multirotor Unmanned Aerial Vehicle (UAV). *Measurement: Journal of the International Measurement Confederation,* Vol. 137(6); 647–654, https://doi.org/10.1016/j.measurement.2019.01.093.

[32] Shukina, O., Muborakov, K., Ruziev, A., Yakubov, G. and Ergashev, M., (2022). Using Digital Photogrammetry to Create Large-Scale Topographic Maps and Plans in Uzbekistan. *International Journal of Geoinformatics,* Vol. 18(1); 37–42, https://doi.org/10.52939/ijg.v18i1.2103.

[33] Esmat, S. Z. S., Zainuddin, K., Rasib, A. W., Majid, Z., Ghazali, M. D., Marzukhi, F., Yusof, N. M. and Aziz, A. F. A., (2026). Crack Detection Using Pretrained Deep Learning under Varying UAV Imaging Configurations. *International Journal of Geoinformatics,* Vol. 22(1); 1–11. https://doi.org/10.52939/ijg.v22i1.4711.

[34] Thammaboribal, P., Tripathi, N., and Lipiloet, S. (2026). Comparative Study of UAV Mapping with and without Ground Control Points (GCPs). *International Journal of Geoinformatic*s, Vol. 22(3), 36–50. https://doi.org/10.52939/ijg.v22i3.4861.

[35] Romero-Chambi, E., Villarroel-Quezada, S., Atencio, E. and Muñoz-La Rivera, F., (2020). Analysis of Optimal Flight Parameters of Unmanned Aerial Vehicles (UAVs) for Detecting Potholes in Pavements. *Applied Sciences,* Vol. 10(12). https://doi.org/10.3390/app10124157.

[36] Liu, C., Li, J., Gao, J., Gao, Z. and Chen, Z., (2022). Combination of Pixel-Wise and Region-Based Deep Learning for Pavement Inspection and Segmentation. *International Journal of Pavement Engineering,* Vol. 23(9); 3011–3023, https://doi.org/10.1080/10298436.2021.1877704.

[37] Aamir, M. S. Bin, Ilyas, M. H., Khalique, F., Bashir, B., Mahmood, S. and Khalil, R. A., (2025). Autonomous Road Defects Segmentation Using Transformer-Based Deep Learning Models with Custom Dataset. *IEEE Access,* Vol. 13(01); 174177–174199. https://doi.org/10.1109/ACCESS.2025.3617336.

[38] Moreh, F., Lyu, H., Beth, C., Strohm, S., Rizvi, Z. H., Wuttke, F. and Renz, M., (2022). Crack Detection and Localization Based on Spatio-Temporal Data Using Residual Networks. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3538712.3538743.

[39] Liu, X., Guo, Y. and He, W., (2024). A New Artificial Intelligence-Based Method for Detecting Pavement Distresses. *ACM International Conference Proceeding Series*. 233–237. https://doi.org/10.1145/3695080.3695121.

[40] Li, D., Gao, M., Guan, X. and Jin, X., (2025). Quantification Method of Concrete Pavement Diseases Based on Rotating Box Annotation. *Science Progress,* Vol. 108(4); 1–23, https://doi.org/10.1177/00368504251399763.

[41] Elamin, A. and El-Rabbany, A., (2023). UAV-Based Image and LiDAR Fusion for Pavement Crack Segmentation. *Sensors,* Vol. 23(23). https://doi.org/10.3390/s23239315.

[42] Heipke, C., Jacobsen, K. and Wegmann, H., (2002). Integrated Sensor Orientation Test Report and Workshop Proceedings. *European Organization for Experimental Photogrammetric Research*, 1–302. https://www.eurosdr.net/sites/default/files/uploaded_files/eurosdr_publication_ndeg_43.pdf.

[43] Atencio, E., Plaza-Muñoz, F., Muñoz-La Rivera, F. and Lozano-Galant, J. A., (2022). Calibration of UAV Flight Parameters for Pavement Pothole Detection Using Orthogonal Arrays. *Automation in Construction*, Vol. 143. https://doi.org/10.1016/j.autcon.2022.104545.

[44] Ren, M., Zhang, X., Chen, X., Zhou, B. and 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,* Vol. 120. https://doi.org/10.1016/j.jag.2023.103335.