Preprints.org

Article

Pavement Crack Detection of UAV Autonomous Inspection System for Mountainous Roads

Chen Xinbao [*] , Wang Chenxi [*] , Liu Chang [*] , Xiaodong Zhu , Zhang Yaohui , Luo Tianxiang , Zhang Junhao

Article

Pavement Crack Detection of UAV Autonomous Inspection System for Mountainous Roads

Xinbao Chen *, Chenxi Wang *, Chang Liu *, Xiaodong Zhu, Yaohui Zhang, Tianxian Luo and JunHao Zhang

School of Earth Sciences and Spatial Information Engineering, Hunan University of Sciences and Technology, Xiangtan 411201, China; zhuxiaodong2302@gmail.com (X.Z.); eminem@mail.hnust.edu.cn (Y.Z.); 2101080206@mail.hnust.edu.cn (T.L.); 2001080232@mail.hnust.edu.cn (J.Z.)

* Correspondence: xchen@hnust.edu.cn (X.C.); Tel.: +86-18670925719; 2101080215@mail.hnust.edu.cn (C.W.); liuchang@mail.hnust.edu.cn (C.L.)

Abstract: Road cracks significantly affect the serviceability and safety of roadways, especially in mountainous terrain. Traditional inspection methods, such as manual detection, are excessively time-consuming, labor-intensive, and inefficient. Additionally, multi-function detection vehicles equipped with various sensors are expensive and not suitable for mountainous roads due to poor conditions. To address these challenges, this study proposes a customized Unmanned Aerial Vehicles (UAV) inspection system designed for automatic crack detection. The system focuses on enhancing autonomous capabilities in mountainous terrains by incorporating embedded algorithms for route planning, autonomous navigation, and automatic crack detection. The slide window method is proposed to enhance the autonomous navigation of UAV flights by generating path planning in mountainous roads. This method compensates for GPS/IMU positioning errors, particularly in GPS-denied or GPS-drift scenarios. Moreover, the YOLOv8 algorithm is introduced to conduct autonomous crack detection from UAV imagery in the offboard module. To validate our UAV inspection system's performance, we conducted several experiments to assess its accuracy, robustness, and efficiency. The experimental results demonstrate that our UAS inspection system accurately estimates the flight trajectory along the mountainous road for autonomous navigation. Furthermore, the experimental detection results also demonstrate that the YOLOv8 algorithm achieves higher recognition accuracy compared to the YOLOv5, and YOLOv7-tiny algorithms, exceeding 3.1%, and 6.8% of mAP, respectively. In summary, the UAV inspection system proposed in this study can provide a useful tool and technological guidance for the regular inspection of mountainous roads.

Keywords: Concrete crack detection; UAV inspection system; YOLOv8; mountainous road

1. Introduction

Road crack detection is crucial in monitoring the health of road pavement to prevent further deterioration and potential collapse of infrastructure (Jiang, Y. et al., 2023[1]). With the increase in the number of vehicles and the continuous growth in road usage, urban roads, rural roads, mountain roads, and highways are experiencing varying degrees of deterioration. Meanwhile, road maintenance poses a significant challenge in the field of road transportation, particularly in China, where a large number of roads require routine maintenance. Cracks are one of the most challenging problems faced in road maintenance. Therefore, timely detection and treatment of road defects can significantly reduce maintenance costs. Traditional methods, such as manual detection, are excessively time-consuming, labor-intensive, and inefficient. On the other hand, current multi-function detection vehicles equipped with various sensors are expensive and suitable for flat terrain such as urban road,or highways, but they are not widely available for rural roads due to poor conditions, especially in the very high or or narrow mountain areas. As shown in Figure 1, mountain roads typically present difficulties in ground access, steep and narrow characteristics, and bad roadbed hazards, which makes such inspection mission extremely difficult and time-consuming using the existing manual inspection method or vehicle-mounted inspection.

Figure 1. Mountainous road, the mountain road is narrow and winding (for example).

Nowadays, road inspection technologies have been developed from manual measurements to automatic inspections assisted by computer vision and deep learning algorithms. These automatic inspection technologies have gained popularity owing to their efficiency and accuracy (Zhu, J.etc. 2021[2]). In current study, commonly road inspection vehicles equipped with various sensors have demonstrated higher levels of accuracy. However, their high costs and limited accessibility hinder widespread adoption. In recent years,Unmanned Aerial Vehicles (UAV) that can carry a variety of micro-sensors have been widely applied in various inspection missions (Yu,C etc.,2022[3]) due to due to their compact size, affordability, flexibility, mobility. For instance, in the inspection of power transmission, using UAV inspection system instead of manual inspection can greatly reduce the risk of daily inspection (Ma, Y etc.,2021[4]); Meanwhile, UAVs have gained increasing popularity in road detection and monitoring, which can greatly reduce the inspection time and extend to unknown complex environments. Liu et al.(2023[5]) proposed a systematic solution to automatic crack detection for UAV inspection with equipment of hardware and software,but the UAV lacks of autonomous ability. In addition, UAVs are gaining significance in the field of structural health performance inspections for civil engineering infrastructure, such pipeline inspection (Y C. etc.,2022[3];Chen X etc.,2023[6]) or bridge inspection (Metni, N etc.,2007.[7]).

Currently, there is a lack of developed UAVs specifically designed for crack detection on mountainous roads. Most existing UAV-based inspection systems primarily focus on image acquisition, transmission, and some pre-processing. Inspectors still rely on visual inspection of the images or partly computer vision methods to identify cracks (Zhou Q etc., 2022[8]). For instance, some researchers have utilized UAV imagery to study deep learning-based methods for road crack object detection, achieving impressive accuracy results [8-15]. However, the complexities of mountainous terrain and uneven roads pose challenges in manipulating flight routes and maintaining consistent vertical altitudes, making it difficult to obtain high-quality imagery. Additionally, the limited load capacity, computing ability, and endurance of small UAVs make it challenging to carry multiple sensors and fly for extended periods (Yu,C etc.,2022[3]). For these reasons, existing commercial UAVs cannot be directly applied for this mission. Therefore, there is a need to design a high-performance, multi-functional, and reliable UAV inspection system with autonomous capabilities specifically tailored for mountainous roads and detecting defects such as pavement cracks.

In this study, it aims to propose a novel UAV inspection system for examining pavement cracks on mountainous roads. The UAV can carry out its flight mission by autonomously following a designated path and consistent vertical altitude using onboard vision and GPS/IMU sensors. Additionally, the onboard computer processes the images captured by the vision sensor, enabling the

detection and recognition of objects. Engineers can access the detection results and raw image data in real-time on an offboard computer, allowing them to identify various cracks and their locations for maintenance purposes. To conduct experiments, a benchmark dataset of cracks found on concrete pavements of mountainous roads was created and made publicly available for the research community. This dataset also serves as a valuable supplement to existing crack databases. Subsequently, an comparative analysis was performed to evaluate the accuracy of pavement crack detection using the YOLOv8 model. The experimental results demonstrate that the YOLOv8 outperforms other methods in terms of recognition performance during UAV inspection missions.

The main contributions of our works can be summarized as follows: (i).An overall integrated architecture for UAV inspections is proposed, including UAVs hardware system and software system, as well as advanced algorithms for autonomous cracks detection,such as YOLOv8. (ii).We have developed a visual sliding window method to extract the flight routes along the mountainous road for path planning and UAV autonomous navigation during the GPS-denied or GPS-drift periods. Furthermore, we have conducted fusion strategy method to compare with the GPS trajectories. (iii). The benchmark dataset of concrete cracks from mountainous roads has been created and made publicly available for the research community, serving as a valuable supplement to existing crack databases.

The paper is structured as follows: Section 2 provides a detailed description of the main hardware components of the UAV inspection system, including their parameters, and presents the software architecture along with the data transfer process. Section 3 explains the crack detection procedure using deep learning and introduces the theoretical background of YOLOv8. Section 4 covers the experimental scenarios and results, while Section 5 provides discussions on the experimental findings. Finally, Section 6 summarizes the key findings and conclusions of the research.

2. Architecture of UAV Autonomous Inspection System

2.1. Hardware System

To complete the pavement cracks inspection of mountainous road, the UAV needs to carry a variety of sensors to maintain flight stability and carry out other tasks.The overall hardware architecture of our UAV inspection system is shown in Figure 2. The entire UAV hardware system comprises a mini-computer, two CMOS sensors, a motion control system, an Infrared Range Finder, and a rack-mount. The UAV platform utilizes a folding carbon fiber rack equipped with an open-source Pixhawk autopilot. The motion controller, composed of a Pixhawk (v2.4.8), four step motors, a GPS/IMU component, and a Remote Control, is responsible for controlling the autonomous flight of the UAV. The IR Range Finder is an infrared sensor used to detect distance from the underground pavement and avoid obstacles such as trees or power lines. All sensors and controllers are connected to the mini-computer, which is responsible for all computational tasks including processing pose data from the GPS/IMU, RGB video data from the mono-camera, and encoder data from the four step motors. These final results can be transmitted to a WIFI-connected laptop computer for display and further process. The entire system is powered by a 12V lithium battery.
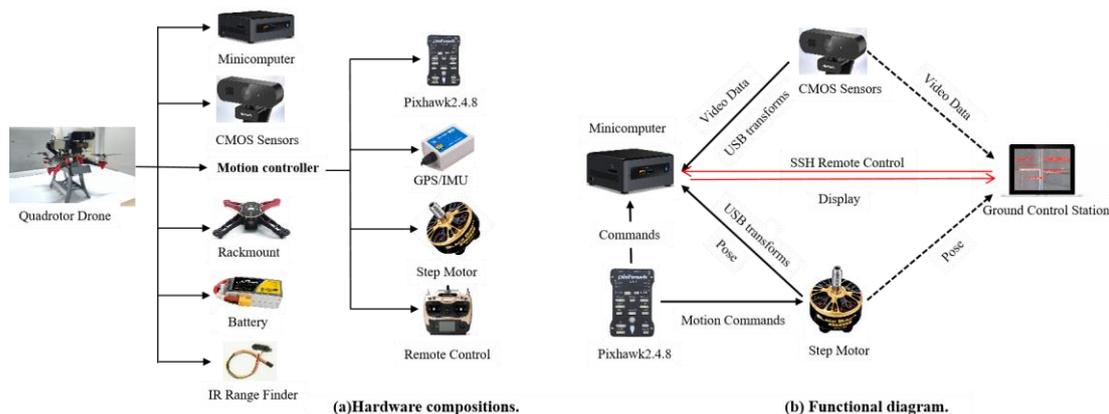

(a)Hardware compositions.   (b) Functional diagram.

4

Figure 2. Hardware architecture of the UAV inspection system(Modified by [Chen.et.al,2023][6]).

2.2. Software Systems

To achieve comprehensive integration of hardware, algorithms, and communication with different sensor devices, we utilize the open source Robot Operating System (ROS, v1.0). The infrared data, GPS/IMU data and mono-camera inputs are processed by a minicomputer to obtain the source data stream. The integrated software architecture of autonomous UAV crack inspections for mountainous road is shown in the **Figure 3**.
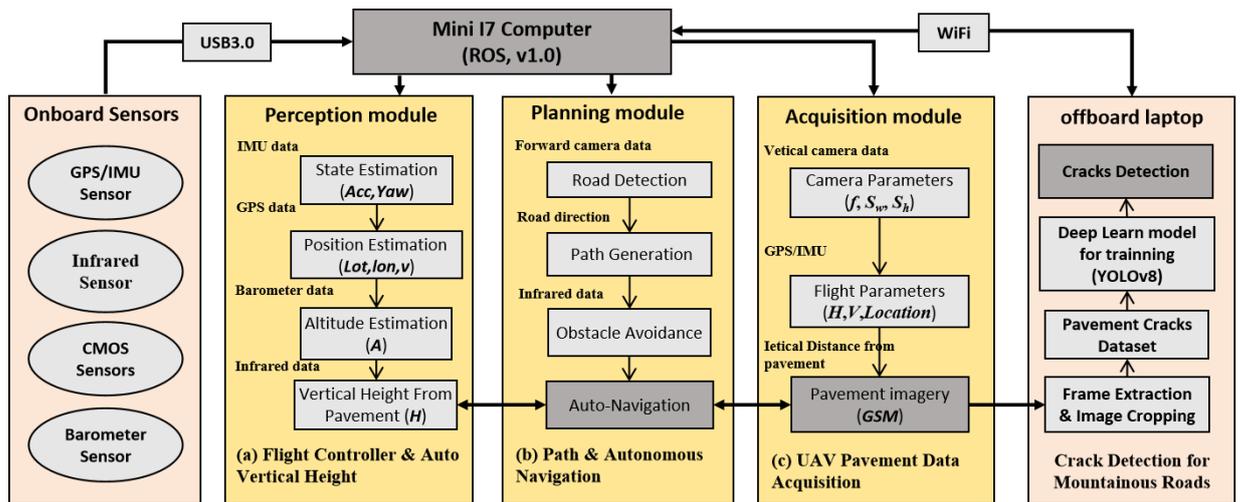

Figure 3. Software architecture of the UAV inspection system.

2.2.1. Flight Controller and Auto Vertical Height

Simultaneously, we employ the GPS/IMU for Pixhawk2 position estimation and utilize TSDF's 6 DOF for pose estimation as input. In **Figure 4**, a typical control method is illustrated, which involves a cascade control loop addressing the inner loop control (the attitude loop) and the outer loop control (the position loop).we utilized a fusion strategy called graded Kalman filter (Suwandi, B., et.al.,2017 [16]) to integrate IMU and GPS positioning information, and then estimate current position and state estimation of UAV. Among the control approaches we have explored, the composite nonlinear feedback (CNF) method for attitude stabilization and robust perfect tracking (RPT) method(Li et.al.,2014[17])for trajectory tracking have shown the best performance. We have implemented the flight control laws using a PX4-based flight controller.
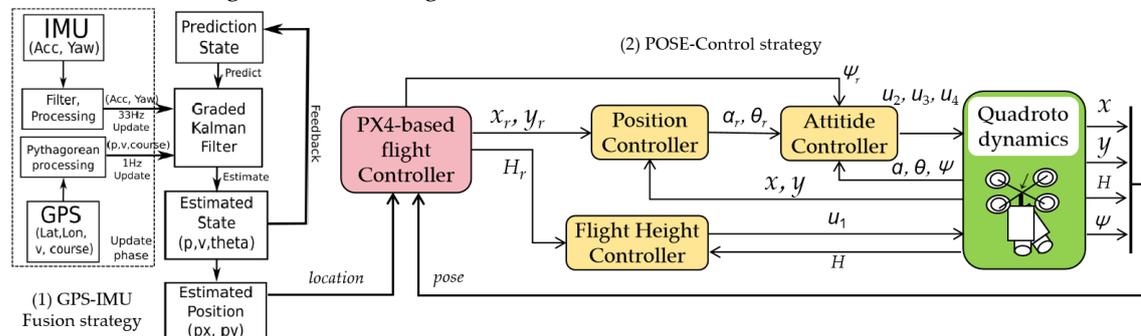

Figure 4. Cascaded control scheme of the quadrotor drones.(Modified from [16,17];).

2.2.2. The UAV Data Acquisition for Mountainous Roads

The acquisition of data imagery in our UAV inspection system is divided into two parts: local data imagery for road pavement and global data imagery for mountainous terrains. **Figure 5** illustrates the diagram of the UAV data acquisition process. The forward camera is mainly used to

capture the global mountainous terrain, assisting the drone in tracking ground road feature points and landmarks for autonomous positioning and navigation. It is a compact camera that provides a comprehensive view of the scene. On the other hand, the vertical camera is a high-definition camera that focuses on the local field of view scene and is primarily used for identifying and detecting road crack targets.Additionally, to determine the optimal altitude of the UAV and ensure efficient flight, the following considerations should be taken into account: (i) The UAV camera view should cover the full width of the road that needs to be inspected;(ii) It is important to avoid any interference from auxiliary facilities such as road trees and street lights during the flight; (iii)To minimize image distortion, it is crucial to maintain a constant vertical distance from pavement, consistent speed, and capture vertical imagery.
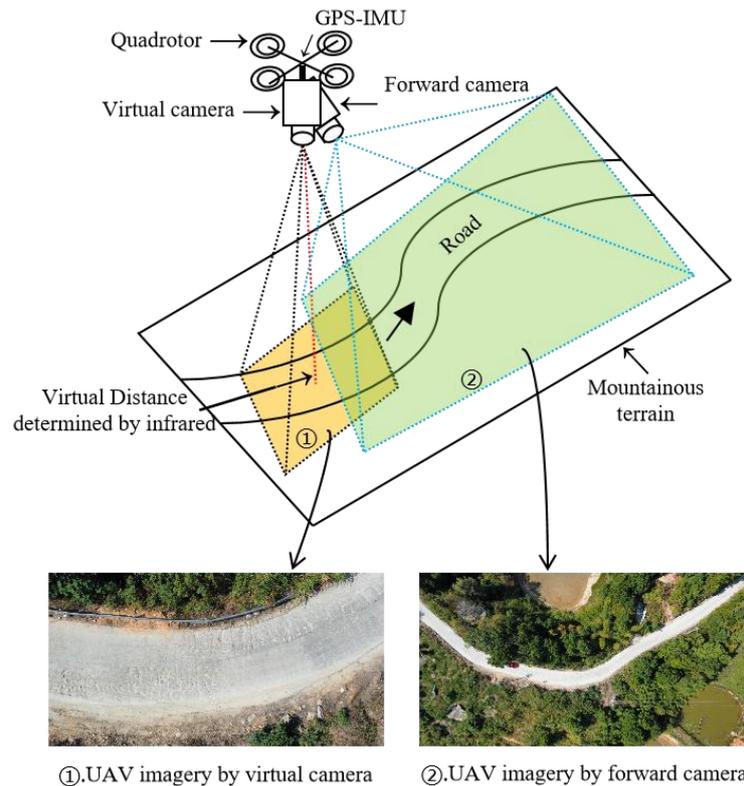

Figure 5. The illustration diagram of UAV data acquisition process.

6

2.2.3. Path Extraction and Autonomous Navigation

In this study, GPS/IMU based sensor fusion is directly employed for state estimation and also for autonomous flying of UAVs, which demonstrates its accuracy performances. However, it is known to suffer from inaccuracies and drift in the GPS signal. This poses a challenge in areas with compromised GPS reception, particularly in our experimental region covered with forests or valleys surrounded by mountains (Lee et al., 2016[18]; Rizk et al., 2020[19]). To tackle this issue, a novel approach is proposed for localizing UAVs and enabling route navigation in mountainous terrains with limited GNSS coverage. The proposed approach involves utilizing the visual sliding window algorithm to calculate the angle and travel distance between two consecutive captured images. Furthermore, we utilize the altitude of the UAV, determined by the integrated barometer sensor in the autopilot, to accurately measure the actual distance using the ground sample distance (GSD) method.

The algorithm workflow of the visual sliding window method is depicted in Figure 6. Firstly, parameters such as the size of the initial sliding window ($l{\times}w$) and the minimum number of non-zero pixels in the sliding window (minPix) are set. The initial sliding window should encompass the entire width of the road, with the length of the sliding window (l) indicating the direction of travel, and the center of the sliding window serving as the initial position. Subsequently, the multiple proposed sliding windows (W_{t+1}^{i}) are generated by rotating (θ) the specific sliding window (W_{t+1}^{0}) around the center point of the previous sliding window (W_t), acting as the axis. The ***Sobel operation*** is then applied to these specific sliding windows to extract road features, followed by counting the number of non-zero pixels and calculating the average value of the horizontal coordinate of all non-zero pixels. The optimal sliding window is determined by selecting the sliding window with the lowest number of non-zero pixels and the lowest average value of the horizontal coordinate. Furthermore, the actual latitude and longitude coordinates (***Lat***, ***Lon***) of the current sliding window (W_{t+1}) are computed based on the central pixel coordinates of the optimal sliding window, the rotation angle (θ), the offset (***D***) of the adjacent sliding window, the ground sampling distance (***GSD***) of the known camera, and the current altitude (***H***) of the UAV. The altitude (***H***) is determined by barometer sensor embedded in the autopilot. Subsequently, following the completion of sliding window detection, GNSS coordinate data files with the NMEA format are generated for route planning of mountain roads and autonomous navigation of UAV. The autopilot is equipped with a port to receive NMEA data from GPS receiver. Hence, the computed coordinates are converted to the standard NMEA format and fed to the autopilot through its GPS input port (Rizk et.al.,2020 [19]). Finally, all connected sliding windows are superimposed with the original images to extract the entire mountain road from UAV imagery.
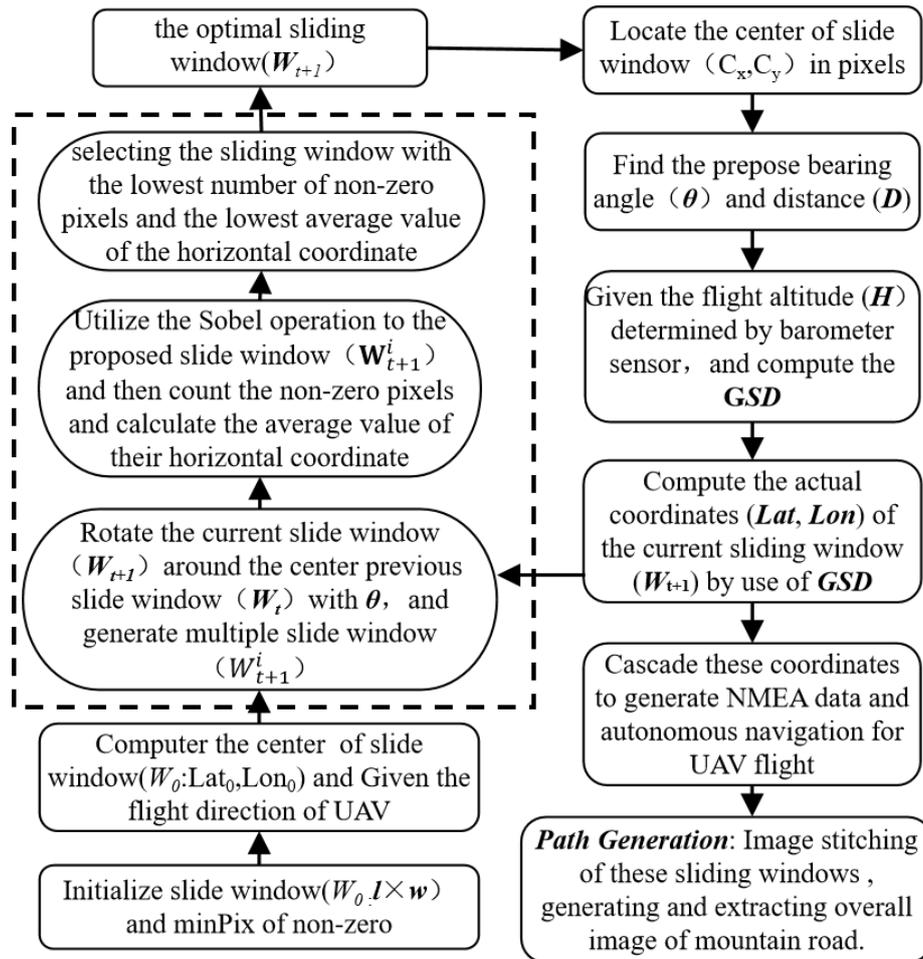
Figure 6. The workflow of sliding window method for route generation.

3. Pavement Autonomous Crack Detection with YOLOv8

3.1. Data Set Collection and Labeling

The dataset utilized in this research was obtained from authentic images of road cracks captured in real-world settings. The data was collected using a Huawei cellphone equipped with a 1/2.3 CMOS camera with a resolution of 12 megapixels. The images were taken at a low height range of 1~3 m, resulting in a total of 658 road crack photos with a resolution of 3840×2160. Various real-world road environments were selected for photography, including rural and mountainous roads commonly surfaced with concrete materials in China. As shown in **Table 1**, the photos captured a diverse range of crack morphologies from concrete pavement, including Horizontal Cracks (HC), Vertical Cracks (VC), Oblique Cracks (OC), Pothole cracks (PC), block cracks (BC),and Expansion Joints(EJ) . It is important to note that expansion joints are incorporated in concrete to prevent expansive cracks caused by temperature changes, which are not classified as crack types with damages.

Table 1. Classification and description of road cracks**.**

Having damage with cracks (Cracks)					No damages	
Horizontal Cracks (HC)	Vertical Cracks (VC)	Oblique Cracks (OC)	Block Crack (BC)	Pothole Cracks (PC)	Expansion Joint (EJ)	No-Cracks (NC)

To optimize parallel batch computing speed, deep learning models require specific requirements for training image datasets. For instance, images are often cropped to 640×640 or 320×320 pixels to improve model efficiency during training. The original images extracted from UAV imagery should be cropped, resized, and standardized to ensure consistent specifications. In this experiment, extracted frame images were used as the initial dataset and cropped into 640×640 pixels images. Data enhancement techniques such as augmentation, translation, flipping, and rotation were employed to increase the number of samples for crack categories. The Labelimg tool was utilized to decode, annotate, and categorize various types of cracks in the UAV imagery cracks dataset from mountainous roads. In this study, the dataset consisted of a total of 413 images, including 57 HC images, 62 VC images, 56 OC images, 76 BC images, 64 PC images, 48 EJ images, and 50 non-crack images. The dataset was divided into training, validation, and test sets with ratios of 80%, 10%, and 10%, respectively, to evaluate the effectiveness of the deep learning model.

3.2. Detection Method Based on YOLOv8

The YOLO (You Only Look Once) algorithm is a fast one-stage object detection. Here, YOLOv8 (Ultralytics,2023 [20]) is the latest version of the YOLO series, which offers significant improvements in both detection accuracy and speed (Wang X.,2023[21];Wang, G.,2023[22]). Additionally, YOLOv8 has a smaller weight file, being 90% smaller than YOLOv5. This makes it suitable for real-time detection on embedded devices. Compared to previous versions of YOLO, YOLOv8 excels in detection accuracy, lightweight characteristics, and fast detection time (Chen x. etc.,2023 [23]). In the context of UAV-based pavement crack detection, accuracy and efficiency are crucial due to the limited computing ability. Therefore, we decided to choice the model using the latest YOLOv8s framework and then further compare the algorithm's accuracy with other deep learning models.

Figure 7 illustrates the basic structure of the algorithm, which comprises an input segment, a backbone, a neck, and an output segment. The input segment performs mosaic data augmentation, adaptive anchor calculation, and adaptive grayscale padding on the input image. The backbone network and neck module form the central structures in the YOLOv8 network. The input image is processed by multiple Conv and C2f modules to extract feature maps at different scales. The C2f module is an improved version of the original C3 module and functions as the primary residual learning module. It incorporates the benefits of the ELAN structure in YOLOv7 (Wong,K.Y. etc.,2023[24]), reducing one standard convolutional layer and making full use of the Bottleneck module to enhance the gradient branch. This approach not only preserves the lightweight characteristics but also captures more abundant gradient flow information. Figure 1 depicts the basic structure of the algorithm. The output feature maps are processed by the SPPF module, which employs pooling with varying kernel sizes to combine the feature maps, and then the results are passed to the neck layer.
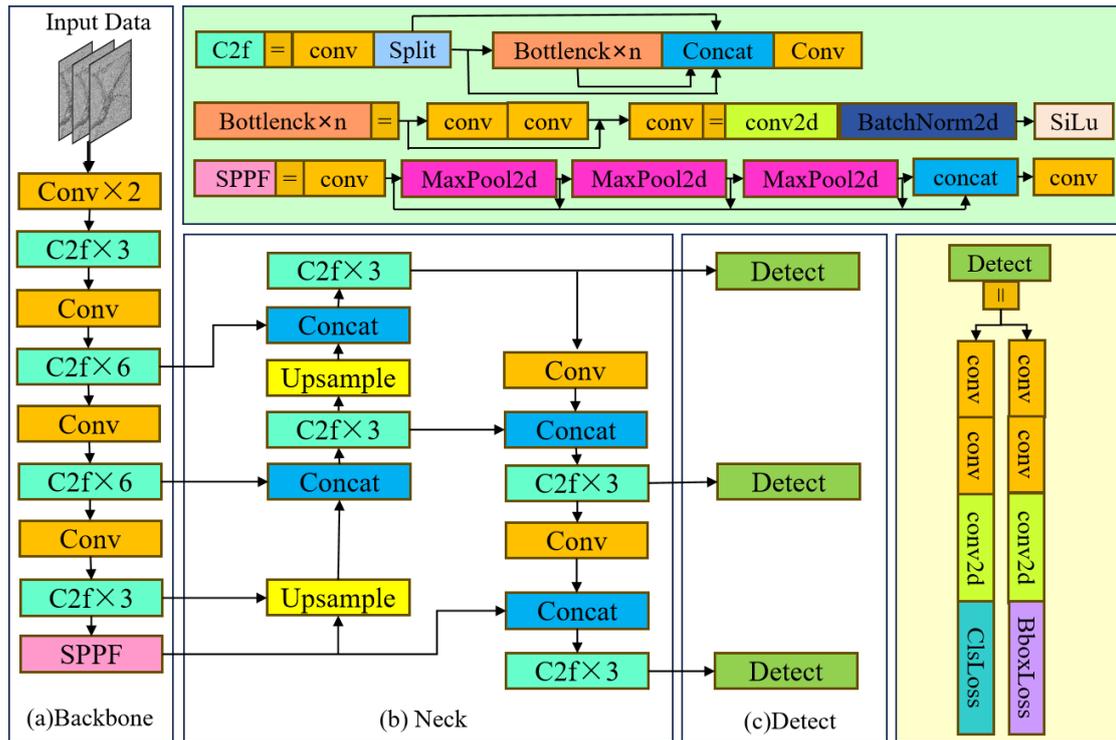
Figure 7. The network structure of YOLOv8.

3.3. Training Process of YOLOv8

The study utilized deep learning algorithms on a consistent configuration, as outlined in **Table 2**. To compare YOLO8 performance, pretrained models like YOLOv5, and YOLOv7 were employed, with code implementation in Python using the Pytorch framework. The experiments ran on a computer with an Intel Xeon(R) E5-2640 v2 processor and 64 GB memory. Hyperparameters were set with a unified input image size of 640px×640px, 200 training iterations, a batch size of 16, and a learning rate of 0.001. The learning rate decay followed the Cosine Annealing algorithm, with the optimizer being SGD and Momentum set to 0.937.

Table 2. Configuration of the experimental environment in ground station.

Software	Configures	Matrix	Versions
Operating system	Windows10	Python	3.9
CPU	Intel Xeon（R）E5-2640 v2	PyTorch	2.0
GPU	NVIDIA GeoForce GTX 4080Ti 8G	CUDA	11.8
Memory	Samsung 64GB		

4. Experiments and Results

4.1. Experimental Scenarios

In this experiment, the flight mission was conducted on Xiangsi Mountain Road in Yueyang County, Hunan Province, China. Xiangsi Mountain Road is a concrete pavement with a width of 4m. The road reaches a maximum height of 957.8 meters above sea level and a minimum of 785.2 meters, resulting in a height difference of 172.6 meters. The road features a maximum bend of 45 degrees. The UAV aerial photography covered a distance of 2.4km during the experiment. Constructed and

10

opened to traffic in 2015, the road has shown visible damage over the past 7 years, including transverse cracks, longitudinal cracks, and alligator cracks. The experiment took place on October 10th, 2023, at 10:00 a.m. under sunny conditions with light traffic. Utilizing UAV flight parameters and relevant formulas, the initial flight height (H) was set at 21.5m, and the initial flight velocity (v) was set at 3.5m/s. The forward camera, using a 2x focal length, captured imagery of the mountainous road from a global perspective, while the vertical camera, with a 4x focal length, captured imagery of the pavement. **Figure 8** displays partial imagery results from Xiangsi Mountain Road.
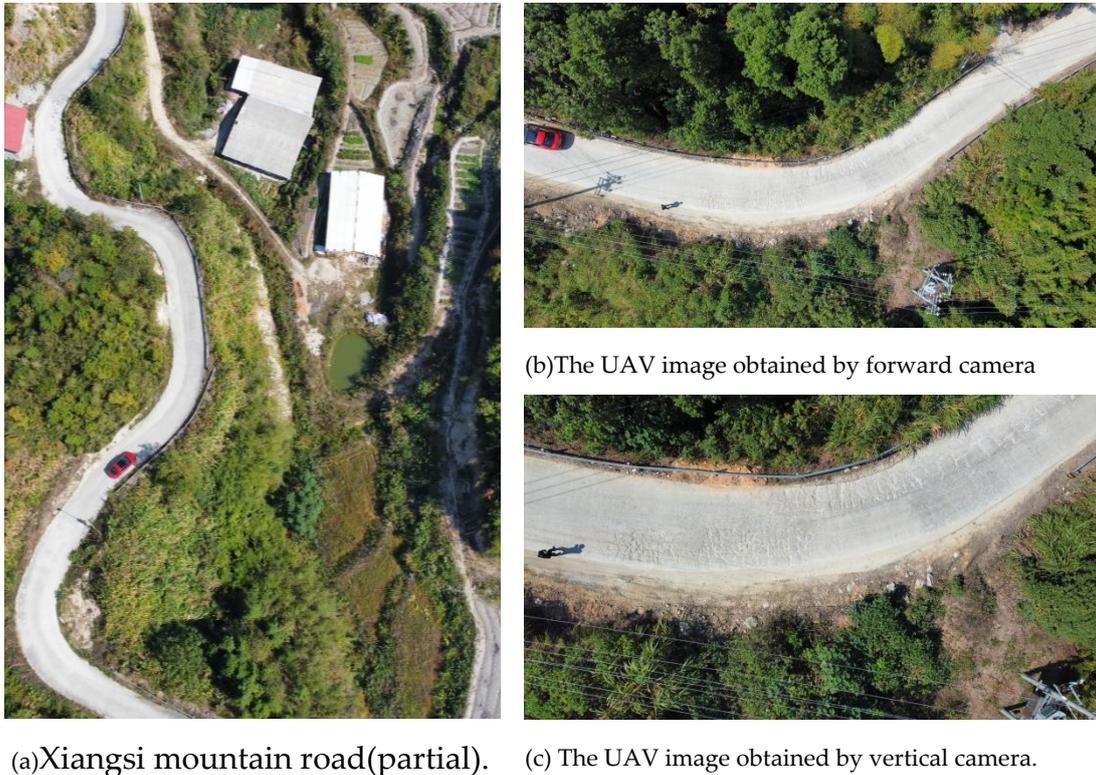

(b)The UAV image obtained by forward camera

(a)Xiangsi mountain road(partial). (c) The UAV image obtained by vertical camera.

Figure 8. Xiangsi mountainous road in this experiments.

4.2. Experimental Results

4.2.1. Results for Autonomous Navigation Using UAV Forward Imagery

(1) Three strategy of visual window for route generation

In order to accurately extract the flight route of UAV in a mountainous road, this study conducted several implementations of route generation from UAV imagery obtained by a forward camera. **Figure 9** illustrates three strategies of the sliding window method used to demonstrate the overall process of route generation. (a).The first strategy involves adjacent sliding windows with a forward superposition degree of 20~30%. The forward sliding window rotates at an arbitrary angle to extract the road through frame selection. This strategy has the fastest implementation algorithm and achieves more accurate road extraction. However, the generated route does not align well with the road curvature. (b).In the second strategy, adjacent sliding windows have a forward superposition degree of 50%, and the rotation is directly canceled. Instead, the forward sliding window is used for lateral translation to select and extract the road. This strategy has a slower implementation algorithm and poor road extraction effect, resulting in an incomplete extracted road. Additionally, the generated route has the worst fit with the road curvature. (c).In the third strategy, adjacent sliding windows have a forward superposition degree of 50%, and the rotation angle of the forward sliding window is arbitrary. The road is extracted through frame selection. Although the implementation algorithm of this strategy runs slowly, it achieves optimal road extraction with a complete extracted road. Moreover, the generated route exhibits the highest consistency with the

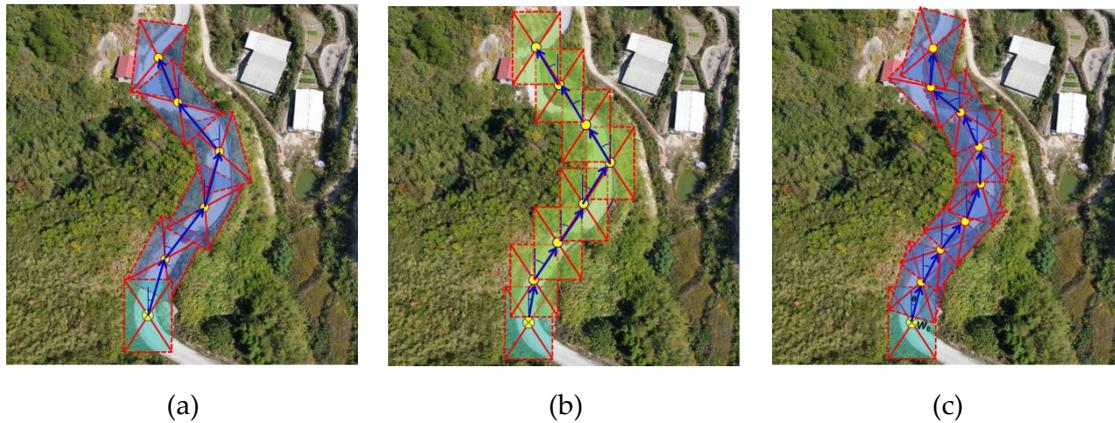road curvature. Therefore, this experiment utilizes the third strategy to generate and extract road routes.

(a)   (b)   (c)

Figure 9. Three strategy of sliding window method: (a)the first strategy;(b)the second strategy;(c) the third strategy.

(2)   Comparative results of route generation

To evaluate the effectiveness of our proposed method, this study conducted comparative analysis of route generation using the customized UAV inspection system. The experimental scenario involved simulated data and and analyzed with the graded Kalman filter (Suwandi, B., et.al.,2017[16]). **Figure 10** illustrates the comparative results of route generation, highlighting the potential of the proposed sliding window method to compensate for GPS/IMU positioning errors and even replace GPS/IMU data, particularly in GPS-denied or GPS-drift mountainous roads during UAV flights. In this figure, the fusion GPS positions (represented by the red dotted line) obtained from the GPS/IMU sensor were plotted for each second of received information. The simulated positioning was also calculated using the proposed sliding window method, and the results are shown through the blue arrow line. A general comparison with the ground truth (represented by the black dotted line) reveals that both routes experienced some positioning errors, especially the GPS/IMU sensor with GPS-drift at max 4m visible error. Furthermore, a fusion method was employed to integrate the GPS/IMU data and the simulated data. The results demonstrate that the fusion route closely aligns with the road centerline, namely the ground truth.
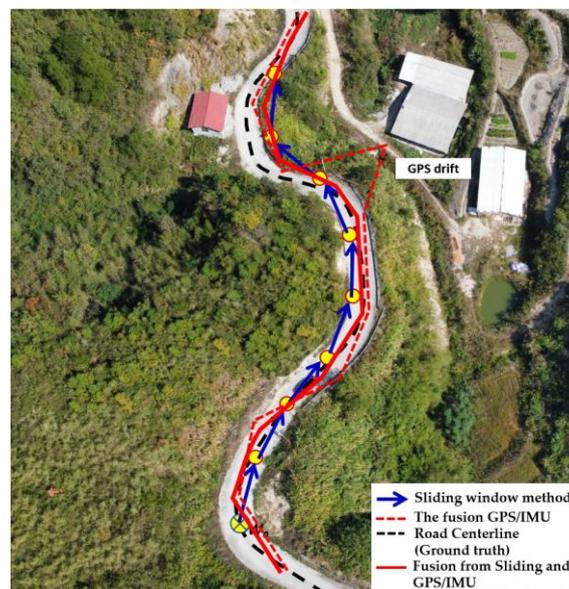

Figure 10. Comparative result of routs generation.

We also calculate the error using Root Mean Square Error(RMSE[Suwandi, B,2017][16]) method to get an exact number on how much error generated from each method compared to the benchmark. We have done this by comparing positioning results from Sliding Window Method, the fusion of GPS/IMU, and the fused trajectory with the ground truth.RMSE formula is given as follows:

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}$$ (1)

Table 3 presents a comparison of RMSE values among three methods for an experimental flight distance of 1.5 kilometers. The results indicate that the SWM shows better GPS accuracy compared to the X and Y axes positions, albeit slightly lower than GPS/IMU. The fusion strategy primarily employed the Graded Kalman filter in this experiment. The fusion of GPS/IMU outperformed standalone GPS in terms of accuracy, with slightly better RMSE values for sensor positioning. The Y axis exhibited higher RMSE values than the X axis due to the predominant flight direction movements along the Y axis. In terms of the H axis, GPS/IMU (RMSE 0.323) performed better than GPS alone (RMSE 0.571).Despite GPS/IMU offering better accuracy in total distance calculations, the final fusion method also utilizing the Graded Kalman filter surpassed their performance with an RMSE value of 0.339, which is an improvement over GPS/IMU's RMSE of 0.427. Overall, the final fusion results demonstrated superior accuracy in the X axis, Y axis, and H elevation.

Table 3. Comparative result form RMSE(Unit:m).

RMSE	X Axis RMSE	Y Axis RMSE	H Axis RMSE	Total Distance RMSE
The Sliding Window Method(SWM)	0.58	1.64	---	0.431
GPS only	0.51	1.95	0.571	0.544
The fusion of GPS/IMU	0.46	1.67	0.323	0.427
The final fusion results	0.46	1.52	0.219	0.339

4.2.2. Results for Autonomous Crack Detection Using UAV Vertical Imagery

(1) The running performance of autonomous crack detection

To evaluate the performance of the lightweight YOLOv8, this study conducted a comparison with the YOLOv5 and YOLOv7 algorithms. The operational performance of the three algorithms is summarized in **Table 4**. The YOLO series models, known for being single-stage algorithms, exhibit notably faster running performance. Among them, the YOLOv8 model stands out with the highest number of parameters (10.37×10^6) and minimal memory requirements, while achieving superior frame rates(132.38**f·s^{-1}**) and training speed(2.6**h**) compared to the other two algorithms. When examining the same datasets, it is evident that the YOLOv8 model offers superior running performance and requires specific environment configurations. As a result, the YOLOv8 algorithms are well-suited for lightweight deployment in real-time detection tasks on UAV platforms.

Table 4. The results of running performance with various models.

Models	Number of Parameters($\times10^6$)	Training Duration(h)	Memory consumption(MB)	Video Memory Usage(GB)	*FPS*(f·s^{-1})
YOLO v5s	8.01	2.8	11.31	3.7	130.73
YOLO v7-tiny	6.78	2.3	12.72	2.4	97.16
YOLO v8s	10.37	2.6	15.63	2.7	132.38

(2) The identification accuracy of autonomous crack detection

The identification accuracy of YOLOv8 was also assessed using key metrics including Precision (P), Recall (R), F1-Score, Average Precision (AP), and Mean Average Precision (mAP). Results in **Table 5** indicate that YOLOv8 also exhibited the highest accuracy compared to other algorithms, surpassing the performance of the remaining YOLO models. Specifically, YOLOv8 achieved values of 85.7% for Precision, 68.4% for Recall, 75.2% for F1-Score, and 78.1% for mAP. Among the YOLO models, YOLOv7-tiny displayed the lowest precision overall with figures of 77.1% for Precision, 70.5% for Recall, 69.6% for F1-Score, and 71.3% for mAP.

Table 5. Results of overall accuracy with various models (%).

Models	*Precision*	*Recall*	*F1-score*	*mAP*
YOLO v5s	82.6	73.9	73.8	75.0
YOLO v7-tiny	77.1	70.5	69.6	71.3
YOLO v8s	85.7	68.4	75.2	78.1

A comparative analysis was further conducted to highlight the discrepancies in model recognition accuracy across different crack categories. Seven types were considered: Horizontal Cracks (HC), Vertical Cracks (VC), Oblique Cracks (OC), Block Cracks (BC), Pothole Cracks (PC), Expansion Joints (EJ), and No Cracks (NC). The results, presented in **Table 6** and **Figure 11**, show that the YOLOv8 algorithm outperformed the other two algorithms in most categories. For example, in HC (AP 82.8%, F1 81.2%), VC (AP 84.1%, F1 82.7%), OC (AP 79.3%), BC (AP 90.2%, F1 88.7%), PC (AP 91.5%, F1 90.5%), EJ (F1 75.1%), and NC (AP 94.0%, F1 90.2%). Notably, EJ exhibited the lowest accuracy, possibly due to misclassification with HC or VC types.

Table 6. Results of detection accuracy with three models under seven types.

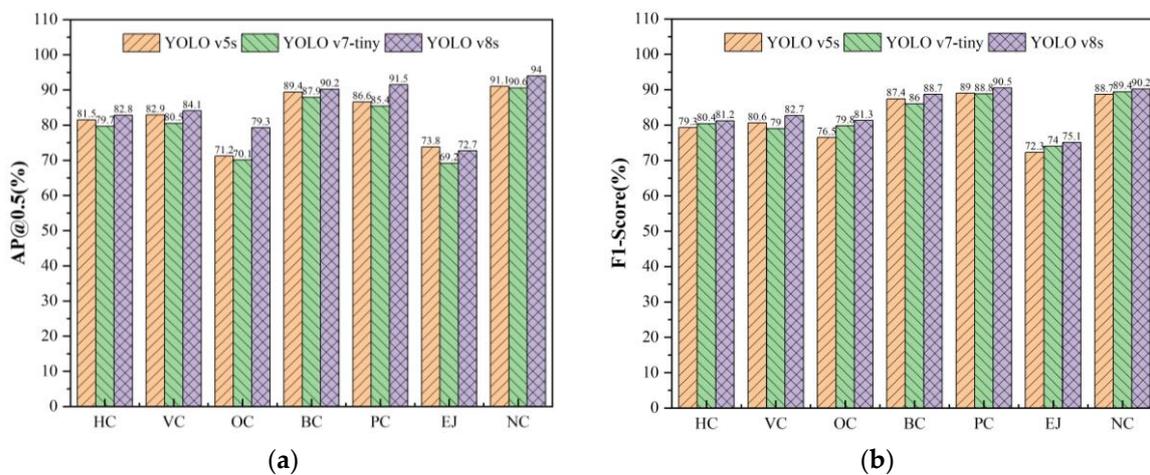| Models	AP@0.5(%)							F1-Score (%)						
	HC	VC	OC	BC	PC	EJ	NC	HC	VC	OC	BC	PC	EJ	NC
YOLOv5s	81.5	82.9	71.2	89.4	86.6	73.8	91.1	79.3	80.6	76.5	87.4	89.0	72.3	88.7
YOLOv7-tiny	79.7	80.5	70.1	87.9	85.4	69.2	90.6	80.4	79.0	79.8	86.0	88.8	74.0	89.4
YOLOv8s	**82.8**	**84.1**	**79.3**	**90.2**	**91.5**	72.7	**94.0**	**81.2**	**82.7**	79.3	**88.7**	**90.5**	**75.1**	90.2

Figure 11. Comparative results of identification accuracy from three algorithms for seven types: (**a**) AP@0.5(%); (**b**) F1-Score(%).

(3) The visual detection results based on YOLOv8

In summary,in comparison to YOLOv5 and YOLOv7-tiny algorithms, YOLOv8 demonstrates higher detection accuracy and faster detection time. In this study, we further present partial visual results of cracks detection using YOLOv8 for concrete pavement from UAV vertical imagery, as illustrated in Figure 12. The YOLOv8 model showed excellent performance in identifying BC types (Confidence Coefficient,namely CC > 0.91), EJ type (CC,0.97), and OC types (CC,0.95). It is worth noting that the YOLOv8 model successfully identified the majority of BC types and their locations, while a few were only partially identified (indicated by the red arrow marker in Figure 12). This discrepancy may be attributed to variations in crack scales.

Figure 12. Partial visual results of cracks detection based on YOLOv8 for concrete pavement from our UAV vertical imagery in this experiment.

5. Discussions

Road cracks have a significant impact on the serviceability and safety of roadways, particularly in mountainous areas. Traditional inspection methods, such as manual detection, are time-consuming, labor-intensive, and inefficient. Additionally, multi-function detection vehicles equipped with various sensors are expensive and not suitable for mountainous roads due to poor conditions.

To address these challenges, this study proposes a low-cost customized UAV inspection system with high autonomous capabilities specifically designed for mountainous terrains. The system integrates autonomous navigation methods and automatic pavement crack detection based on a deep learning model. It consists of an onboard mini-computer, a perception module, a path planning module, an acquisition module, and an offboard detection module. The perception module incorporates two monocular cameras, a GPS/IMU, and an Infrared Range Finder to serve as the basic sensor suite. Its primary function is to obtain target information and transmit processed data to the path planning module. The path planning module generates a flight path based on a pre-set GPS/IMU trajectory along the mountainous road, even in GPS-drift or GPS-denied environments, using the sliding window method of the UAV forward camera. During the autonomous flight, the acquisition module captures footage or takes photos of the pavement using the UAV's vertical camera.Finally,during the detection phase, the offboard module generates a real-time spatial trajectory and visual results of cracks detection based on YOLOv8.

The system also has the capability to store geographical information, including mountain road directions, and can extract GPS trajectory information for preliminary autonomous path planning of UAV. This is a common approach used for autonomous drones in commercial applications. However, relying solely on geographic information (GIS) to plan the path of UAVs on mountain roads presents several key challenges that hinder true autonomous navigation. The main challenges are as follows: (i). GIS provides planar geographic information, which is suitable for vehicle navigation, but insufficient for drones that require three-dimensional spatial information.(ii). The drone's flying area should maintain a safe distance from the road and consider factors such as wind direction, requiring special attention to safety.(iii). When the drone flies autonomously, it needs to avoid trees, which cannot be reflected in GIS and requires on-site measurements. To address these challenges, this study

also explores the potential solution of utilizing infrared scanning technology and our proposed sliding window method to be helpful for automatic planning and navigation.

The sliding window method can take effect in the context of mountainous roads that need further explanation. Mountainous roads have unique geographical characteristics, such as being single, narrow, curved, and steep, and are often surrounded by forest vegetation and rocks. This makes the appearance and structure of mountain roads significantly different from the surrounding environment and vegetation. Computer vision techniques can utilize these features to identify, locate, and extract mountain roads from UAV imagery. To achieve autonomous positioning and navigation of the UAV, the visual sliding window method is employed to detect the road direction based on the color and zigzag structure of the mountain road. This study explores three strategies for generating and extracting routes using the sliding window method for UAV autonomous navigation along the desired road. Experimental results demonstrate that our fusion method can perform real-time route planning in GPS-denied mountainous environments, with an average localization drift of 2.75% and a per-point local scanning error of 0.33m over a travelling distance of 1.5 km.

Meanwhile, a deep learning model based on YOLOv8 in this work is applied to the UAV autonomous crack-detecting scenario for mountainous roads.YOLOv8 further optimizes the model structure and training strategy of YOLOv7 to improve both detection speed and accuracy. One notable improvement is the incorporation of Extended-ELAN (E-ELAN), a more efficient long-range attention network that enhances the model's feature extraction capability. YOLOv8 also introduces new loss functions, such as VFL Loss and CIOU Loss+DFL (Distribution Focal Loss), to improve localization accuracy and category differentiation ability. Moreover, YOLOv8 employs new data enhancement methods, including Mosaic + MixUp, to enhance the model's generalization and robustness. The offboard module of the YOLOv8 model accurately identifies the position of pavement cracks using pictures obtained from the UAV inspection system, offering competitive computational speed and accuracy. Experimental results demonstrate that the YOLOv8 algorithm achieves higher recognition accuracy compared to the YOLOv5 and YOLOv7-tiny algorithms, exceeding 3.1%,and 6.8% of mAP, respectively. In summary, YOLOv8 is a lightweight automatic crack-detecting method in real-time that can be integrated into the UAV inspection system, significantly reducing the human resources required for regular distress inspection of mountainous roads.

6. Conclusions

In this study, we propose a novel UAV inspection system designed to examine pavement cracks on mountainous roads. Our system focuses on enhancing autonomous capabilities in mountainous terrains by incorporating advanced embedded algorithms for route planning, autonomous navigation, and automatic crack detection. We provide detailed information about the hardware design and introduce a real-time motion estimation algorithm that generates route planning and enables autonomous navigation in mountainous terrains. Simultaneously, the off-board computer conducts pavement crack detection in the mountainous road environment. The UAV inspection system offers advantages such as mechanical simplicity, mobility, low weight, low cost, and real-time estimation. Therefore, our customized UAV inspection system is suitable for autonomous navigation in complex mountainous terrains as well as pavement crack detection. To validate the system's performance, we conducted several experiments to assess its accuracy, robustness, and efficiency. The experimental results demonstrate that our system accurately estimates the flight trajectory along the mountainous road for autonomous navigation and conducts the autonomous crack detection using YOLOv8 for pavement ,yielding more satisfactory results. In addition, a benchmark dataset for mountainous road is also established and open-sourced for the community.

In general, our UAV inspection system can provide significant assistance to traffic management authorities, rural planners, and infrastructure maintenance companies. It allows them to respond quickly to road issues and take timely measures, thereby improving road availability and safety. Despite the broad and diverse development of UAV detection, there is still a need for research on tiny and more complex road distress detection models. In order to establish a comprehensive UAV

road defect detection system, our future research will focus on multi-category defect detection systems, including various road issues such as the collapsed road. We also plan to develop more advanced CNN architectures to enhance the performance of crack detection and segmentation. Furthermore, we acknowledge that real-world applications may be influenced by varying road conditions and environmental factors, which can affect detection performance. Therefore, we will explore optimization and refinement of our model in diverse road scenarios of mountainous terrains.

Author Contributions: Conceptualization, X.C., C.W., and C.L.; methodology, X.C., C.W., and C.L.; software, C.L. and C.W.; validation, C.W., and C.L.; formal analysis, X.C. and Y.Z.; investigation, L.C., C.W., T.L. and Y.Z.; resources, C.W., C.L., J.Z.and T.L.; data curation, C.W., X.L., T.L., and J.Z.; writing—original draft preparation, X.C., C.L., and Y.Z.; writing—review and editing, X.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jiang, Y.T; Yan, H.T; Zhang, Y.R; Wu, K.Q; Liu, R.Y; Lin, C.Y. RDD-YOLOv5: Road Defect Detection Algorithm with Self-Attention Based on Unmanned Aerial Vehicle Inspection. Sensors 2023, 23, 8241. https://doi.org/10.3390/s23198241.
2. Zhu,J.;Zhong,J.;Ma,T., et al. Pavement distress detection using convolutional neural networks with images captured via UAV[J/OL].Automation in Construction.2022,103991.http://dx.doi.org/10.1016/j.autcon.2021.103991.
3. Yu, C.;Yang, Y.;Cheng, Y.;Wang, Z.;Shi, M.;Yao,Z. UAV-based pipeline inspection system with Swin Transformer for the EAST. Fusion Engineering and Design.2022,184,1-6. https://doi.org/10.1016/j.fusengdes.2022.113277.
4. Ma, Y.;Li, Q.;Chu, L.;Zhou, Y.;Xu, C.Real-Time Detection and Spatial Localization of Insulators for UAV Inspection Based on Binocular Stereo Vision. Remote Sensing.2021,230. https://doi.org/10.3390/rs13020230.
5. Liu,K.Learning-based defect recognitions for autonomous uav inspections. arXiv preprint arXiv 2023, 2302.06093. https://doi.org/arXiv:2302.06093v1.
6. Chen, X.; Zhu, X.; Liu, C.Real-Time 3D Reconstruction of UAV Acquisition System for the Urban Pipe Based on RTAB-Map. Applied Sciences.2023, 13, 13182. https://doi.org/10.3390/app132413182.
7. Metni, N.;Hamel, T.A UAV for bridge inspection: Visual servoing control law with orientation limits. Automation in Construction.2007,17(1), 3–10. https://doi.org/10.1016/j.autcon.2006.12.010
8. Zhou,Q.;Ding,S.;Qing,G., et al. UAV vision detection method for crane surface cracks based on Faster R-CNN and image segmentation[J/OL].Journal of Civil Structural Health Monitoring, 2022: 845-855. http://dx.doi.org/10.1007/s13349-022-00577-1.
9. Yokoyama, S.; Matsumoto, T. Development of an automatic detector of cracks in concrete using machine learning. Procedia engineering.2017, 171, 1250-1255. https://doi.org/10.1016/j.proeng.2017.01.418.
10. Jiang, Y.T; Yan, H.T; Zhang, Y.R; Wu, K.Q; Liu, R.Y; Lin, C.Y. RDD-YOLOv5: Road Defect Detection Algorithm with Self-Attention Based on Unmanned Aerial Vehicle Inspection. Sensors 2023, 23, 8241. https://doi.org/10.3390/s23198241.
11. Zhang, Y.; Zuo, Z.; Xu, X.; Wu, J.; Zhu, J.; Zhang, H.; Wang, J.; Tian, Y. Road damage detection using UAV images based on multi-level attention mechanism. Automation in Construction 2022, 144, 104613. https://doi.org/10.1016/j.autcon.2022.104613.
12. Xiang, X.; Hu, H.; Ding, Y.; Zheng, Y.; Wu, S.J.A.S. GC-YOLOv5s: A Lightweight Detector for UAV Road Crack Detection. Applied Sciences 2023, 13, 11030. https://doi.org/10.3390/app131911030 .
13. Omoebamije, O.; Omoniyi, T.M.; Musa, A.; Duna, S. An improved deep learning convolutional neural network for crack detection based on UAV images. Innovative Infrastructure Solutions 2023, 8, 236. https://doi.org/10.1007/s41062-023-01209-3 .

14. Zhao, Y.; Zhou, L.; Wang, X.; Wang, F.; Shi, G. Highway Crack Detection and Classification Using UAV Remote Sensing Images Based on CrackNet and CrackClassification. Applied Sciences 2023, 13, 7269. https://doi.org/10.3390/app13127269 .

15. Samadzadegan, F.; Dadrass Javan, F.; Hasanlou, M.; Gholamshahi, M.; Ashtari Mahini, F. Automatic Road Crack Recognition Based on Deep Learning Networks from UAV Imagery.ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., X-4/W1-2022, 685–690. https://doi.org/10.5194/isprs-annals-X-4-W1-2022-685-2023, 2023.

16. Suwandi, B.; Kitasuka, T.;Aritsugi, M. Low-cost IMU and GPS fusion strategy for apron vehicle positioning. TENCON 2017.2017 IEEE Region 10 Conference, Penang, Malaysia, 2017, pp. 449-454, doi: 10.1109/TENCON.2017.8227906.

17. Li, D.; Li, Q.; Cheng, N.; Song, J. Sampling-Based Real-Time Motion Planning under State Uncertainty for Autonomous Micro-Aerial Vehicles in GPS-Denied Environments. Sensors. 2014,14,21791-21825. https://doi.org/10.3390/s141121791.

18. Lee,Y.;Yoon,J.;Yang,H.;Kim,C.;Lee,D.Camera-GPS-IMU sensor fusion for autonomous flying.2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN), Vienna, Austria, 2016, pp. 85-88, doi: 10.1109/ICUFN.2016.7536988.

19. Rizk, M.;Mroue, A.;Farran, M.;Charara, J.Real-Time SLAM Based on Image Stitching for Autonomous Navigation of UAVs in GNSS-Denied Regions. 2020. 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS), 301-304.DOI:10.1109/aicas48895.2020.9073793.

20. Ultralytics.YOLOv8[CP]. 2023.Available online. https://github.com/ultralytics/ultralytics. (Accessed on 26 January 2023).

21. Wang, X.; Gao, H.; Jia, Z.; Li, Z. BL-YOLOv8: An Improved Road Defect Detection Model Based on YOLOv8. Sensors.2023, 23, 8361. https://doi.org/10.3390/s23208361.

22. Wang, G.; Chen, Y.; An, P.; Hong, H.; Hu, J.; Huang, T. UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios.Sensors.2023, 23, 7190. https://doi.org/10.3390/s23167190.

23. Chen, X.; Liu, C.; Chen, L.; Zhu, X.; Zhang, Y.; Wang, C. A Pavement Crack Detection and Evaluation Framework for a UAV Inspection System Based on Deep Learning. Appl. Sci. 2024, 14, 1157. https://doi.org/10.3390/app14031157.

24. Wong,K.Y.YOLOv7[CP]. 2023.Available online. https://github.com/WongKinYiu/yolov7.(Accessed on 26 January 2023).