ASEAN Engineering Journal

Full Paper

AUTONOMOUS DRONE-BASED IMAGING SYSTEM FOR DETECTION OF POTHOLES IN RURAL ROAD

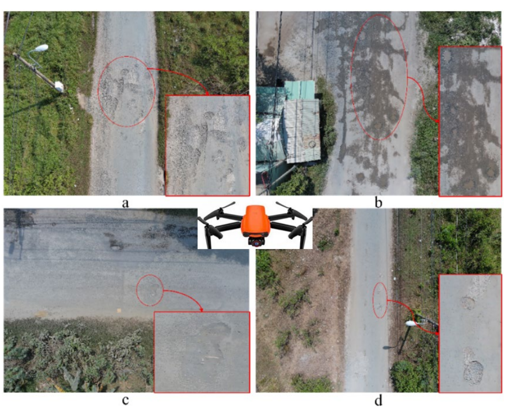
Nguyen Kim Ngan Luu[a,b], Chan Huy Quan[a,b], Khanh Hieu Ngo[a,b,c]

[a]Department of Aerospace Engineering, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam
[b]Vietnam National University Ho Chi Minh City (VNU-HCM), Linh Trung ward, Thu Duc District, Ho Chi Minh City, Viet Nam
[c]VNU-HCM Key Lab. for Internal Combustion Engine, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam

Article history
Received
04 November 2024
Received in revised form
05 March 2025
Accepted
25 June 2025
Published online
28 February 2026

*Corresponding author
ngokhanhhieu@hcmut.edu.vn

Graphical abstract



Abstract

As a country with a high percentage of motorbikes, the appearance of potholes has a serious impact on people. Today, detecting and repairing potholes takes a long time. Therefore, this research paper aims to provide solutions for inspectors to save time in detecting potholes. In this study, the authors used drones to col-lect data combined with the common computer vision model YOLOv5 to pro-duce pothole detection models for images taken from drones. From the identifi-cation results, we calculate the size and location of the potholes so that the inspectors can identify the danger range and come up with a reasonable repair plan. The current model reached the precision, recall, and F1 score at 0.946, 0.961, and 0.95, respectively.

Keywords: YOLO, drone-based imaging, road pothole detection, road inspection, ultra-light drone

1.0 INTRODUCTION

A pothole is a depression on a road, which is the result of traffic or weather conditions. This kind of road damage can cause accidents for motorbikes, even death, so it needs to be discovered and repaired soon. However, the maintenance process starts with the inspectors receiving damage information from the people or after the periodic inspection, then planning the repair. As such, it is time-consuming and inefficient. Moreover, inspectors can also be at risk of injury due to hard-to-detect potholes. Therefore, to shorten the process and protect the inspectors, the authors suggest using a drone for automatic pothole detection. With the assistance of the UAVs, the inspection process can be done in much faster and safer ways, allowing to greatly reduce the workloads on the inspectors and responding rapidly to high-risk defects that can cause significant traffic accidents.

Automated pothole detection technologies have been studied for decades. There are 3 popular methods including vision-based [1,2], vibration-based [3,4,5] and 3D reconstruction-based method [6,7]. The first method applies road surface processing and deep learning. It requires 2D images or videos input, easily obtained by devices such as phones, or devices with cameras such as drones. It is suitable for determining the potholes as well as their shape and size, but it is not able to export depth. The second one uses data from acceleration sensors in smartphones to detect potholes and their depth, except for the shape and size. The final method measures the shape along with the depth. After analysis and comparison, [8] indicated that the 3D reconstruction-based is the most expensive method while the vibration-based is the most cost-effective one. Table 1 summarizes these methods. This study aims to detect the 2D size of potholes with low cost, so the vision-based method is the most suitable. Thanks to the simple input requirements, the equipment used for data acquisition also has a wide range of options. [9] from Iraq applied Deep Learning with YOLOv4's algorithm to train 5,300 images collected from phones and drones. Then use this model to identify potholes from the digitized image. The obtained research results prove the feasibility of using images taken from unmanned aerial vehicles (UAVs) to create orthogonal images and identify potholes from those images using deep learning techniques. This method is safe, economical, and faster than the traditional method.

Table 1 Comparison of vision-based, vibration-based and 3D reconstruction-based methods

Method	Strengths	Weaknesses
Vision-based	Determine the 2D shapes and the number of potholes 2D images/ video as input	Cannot determine the depth
Vibration-based	Detect depth potholes Require small storage Allow real-time application	Cannot determine the shape and size
3D reconstruction-based	Measure accurately the shape	The most expensive method

This research is an integrative solution for pothole detection and management utilizing UAV for data acquisition and deep learning for pothole detection. The authors propose the use of UAV to collect high-resolution images of the roads where there is the presence of potholes varying in appearance, then create a cus-tom dataset to train YOLOv5 model for pothole detection. The acquired images and the trained model can be used in future studies. In addition, we also developed a simple tool in which the information about the detected potholes from the images such as 2D sizes and locations is extracted, summarized on a table, and presented on a map.

2.0 METHODOLOGY

2.1 General Workflow

Figure 1 illustrates the general workflow in this project. There are 3 steps to this process including collecting data by drone, and training to identify and analyze the results. This workflow is simple so that it can be widely performed by anyone that has the necessary tools.

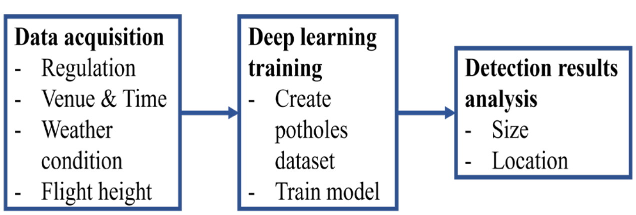


Figure 1 Workflow for this research

2.2 Data Acquisition

There are a lot of pothole datasets publicly available on the internet, however, they are not from the perspective of UAV. There are a lot of pothole datasets publicly available on the internet [10], however, they are not from the perspective of UAV. This is the reason for making our own datasets.

The dataset is an important part of the training object detection model and pothole is the specific object with tools for capturing pictures, so there are some requirements for acquiring images to get the best data:

- Regulations: obtaining the flight license for flying a drone if needed.
- Venue: choosing a location to fly is the first thing you need to do. This location must not be in a no-fly or restricted-fly zone.
- Weather: check the weather forecast before deciding the day to fly because this drone cannot fly in rain and wet lenses affect the quality of images.
- Time: most of the roads with potholes are caused by too many containers, trucks, and other vehicles passing during the day, so choosing the time to fly is important. The flight time slot must ensure few cars, and daytime for the best quality images.
- Others: depending on the location of the potholes, there may be many obstacles such as wires or tree shadows, with the main goal of data diversity, the pilot can flexibly in the angle captured and altitude to see the pothole. In addition, collecting data on potholes with trees shadows, or wires across also increases the ability to identify.

The general rule for image quality is that at the same image resolution, the lower the drone hovers, the more detailed the image is. Flying too high will result in the pothole shrinking from a few to no pixels. The altitude can be calculated based on the parameters of the camera, the requirement for the smallest pixel that can be detected, and the criteria for assessing pothole damage.

The formula Equation 1 calculates the required altitude from the ground sampling distance (GSD) value.

$$\mathrm{GSD}\left(\frac{cm}{pixel}\right) = \frac{\text{flight altitude (m)}}{\text{focal length (m)}} \times \frac{\text{sensor}\,\frac{\text{width}}{\text{height(cm)}}}{\text{image}\,\frac{\text{width}}{\text{height(pixel)}}} \qquad (1)$$

Based on the pothole damage risk assessment Table 2 provided by the Faculty of Civil Engineering, and the minimum object size that can be detected is 2 pixels [11], the required altitude for the pothole to become detectable is 25 meters. However, to get more detailed information the drone should fly as low as possible.

Table 2 Criteria for assessing the risk of pothole damage

Maximum depth (mm)	Average diameter (mm)		
	100 – 200	200 – 450	450 – 750
13 – 25	Low	Low	Moderate
25 – 50	Low	Moderate	High

2.3 Deep Learning Model for The Detection of The Pothole

The model used in this paper is YOLOv5. This is widely used for object detection models, famous for its accuracy and speed, which is suitable for real-time applications. YOLOv5 [12-14] is a single-stage detector model, so its accuracy may be lower than some two-stage detector models because it focuses on balancing be-tween inference time and accuracy. However, thanks to its single-stage architecture, the model is light in size, easy to train, and fast to infer on each image. This model has successfully recognized many objects from easy to complex scenarios, such as fruits [15, 16], potholes taken with a phone [17, 18] and so on. YOLOv5 has five types of models in total, including: nano (n), small (s), medium (m), large (l), and extra-large (x). The small version YOLOv5s takes less time while the biggest version YOLOv5x6 takes the longest time but is the most accurate.

2.4 Detection Results Analysis

After detection, each detected image returned a label file concluding center coordinates and size in two dimensions in pixels of potholes. From that information, it is possible to convert to width and height in centimeter and calculate the average diameter. The results will be compared with the standard in Table 2 and given the risk range. To evaluate the above dimensions, we need to compare them with the real dimensions. However, potholes are located on the street with many vehicles that are difficult to measure normally, for safety purposes, we chose the method of photographing the ruler with potholes by drone Figure 2a, then from the actual size of the ruler, we find the ratio scale to measure pothole size in AutoCAD Figure 2b.

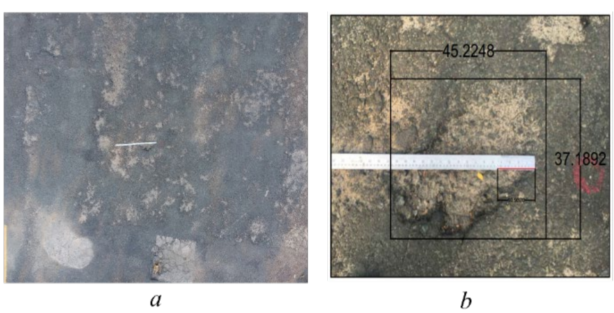


Figure 2 (a) image including pothole and ruler captured by drone, (b) dimension measured in AutoCAD

3.0 CASE STUDY

3.1 Area Survey and Data Acquisition

This research is carried out in the city, so the drone needs to be ensured that the flight regulations are issued on noise and allowed locations. Besides, the selected drone should be compact for easy carrying and can take off and land in narrow spaces, battery capacity, and signal transmission are good enough to fly a road. After considering the above requirements, Autel Evo Nano in Figure 3 is the main UAV collecting pothole images to diversify datasets with the parameters in Table 3.

Figure 3 Autel Evo Nano [19]

Table 3 Configuration of Autel Evo Nano [19]

Take off weight	Max flight time	Camera sensor	Effective pixels	Max Flight Distance (no wind)	Opera -ting Tempera-ture
249 g	28 min	CMOS: ½ inch (6.4x4.8 mm)	48 MP (8000x6000)	16.8 km	0°C~40°C

Thanks to the support from Thu Duc City, the authors captured potholes to diversify the data. According to section 2.2, this dataset is collected at an altitude of about 10 to 15 m depending on the area with obstacles or not. The sole goal of this step is to create a dataset of potholes. These captured images have a resolution of 8000x6000 pixels, with good lighting, and are clear in high detail. Figure 4 shows a variety of potholes with dry and water-filled potholes, from very small to relatively large, with diverse shapes. Some potholes are covered by power lines and trees. The road surface is wet in some places, making it feel like potholes. This dataset is unique because potholes datasets are often captured by camera in cars.

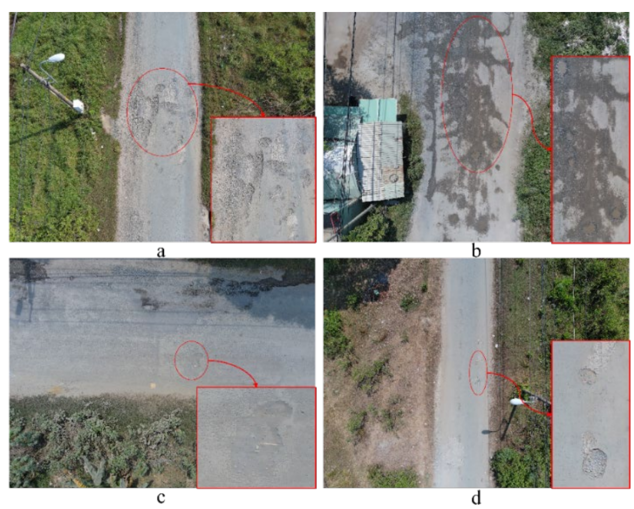


Figure 4 Some images from data acquisition (a) dry potholes with rocks, (b) water-filled potholes, (c) dry pothole hidden in the road surface, (d) dry potholes.

3.2 Training Implementation

About 450 captured images with a variety in size, shape, and properties of potholes were labelled and brought to train on an Intel UHD Graphics 630 desktop computer running the Ubuntu operating system with 32 gigabytes of total memory with

YOLOv5x6 model for the highest accuracy. It took about 65 hours to complete 100 epochs. The main problem in this step is the training is too long, but this is characteristic of YOLOv5x6. To solve this issue, the personal workstation needs to be upgraded for more graphics processing units. This study used the default of model YOLOv5x6 as shown in Table 4.

Table 4 Training configuration parameters

Training Parameter	Value
Epoch	100
Batch size	4
Image size	1280
Model	YOLOv5x6

The training reached the optimal value at epoch 66 with the parameters shown in Table 5. mAP@50 and mAP@50-95 are 0.979 and 0.66 respectively. mAP@50 is very high while mAP@50-95 is just higher than average. The reason is that the shape of potholes sometimes it is difficult to be labeled fit well, so when the requirement of IoU increases, the mAP decreases. However, these values are acceptable for this model.

Table 5 Optimal value

Box – precision	Box – recall	Box – mAP@0.5	Box – mAP@0.5:0.95
0.946	0.961	0.979	0.65

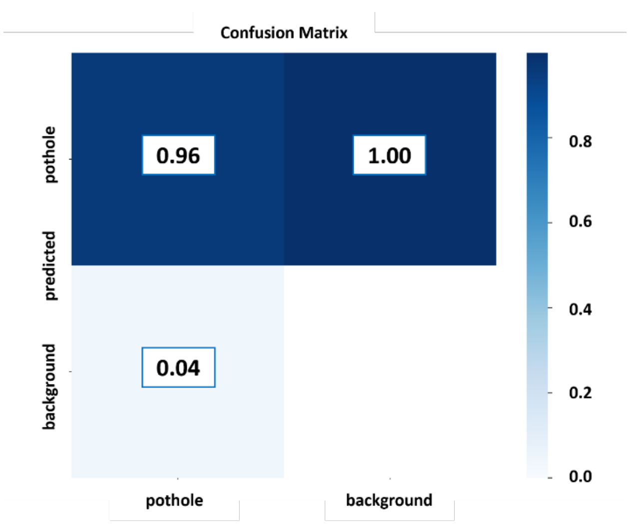


Figure 5 Confusion Matrix

As can be seen in Figure 5, there are 96% real potholes that can be detected. However, the background can be mistaken with the pothole because the learned data was not diverse enough. The training results also showed that the precision, recall, and F1 score of the proposed method were 0.946, 0.961 and 0.95 respectively. These values are adequate, showing that the trained model has a decent ability for detecting potholes.

The trained model was used for detecting potholes on a test case of a specific road. This is Lo Lu Street in Thu Duc City with a lot of kinds of potholes such as potholes having power lines through them, and potholes with rocks. The results show that this new model correctly detected many objects with confidence of more than 60% in Figure 6a, however, there are also some mistakes in Figure 6c and omitted in Figure 6b.

Figure 6 Detection results (a) correct, (b) omission, (c) mistaken

3.3 Detection Results Analysis

The reasons for misidentification and omission can be mentioned as:

- The dataset is still not diverse enough for potholes, the pothole image is quite uniform for the adjacent road surface. For wet road surfaces, dry gaps are similar to the difference between potholes and pavement. Potholes are learned not to be cut by electrical wires if viewed from above.
- The labeling step is done manually, so it is easy to make mistakes. The labelers may sometimes not recognize the object to label or determine whether the pothole size is insufficient to match.

Further analysis from the information about the size of the bounding box, the authors easily calculate the average diameter of the pothole and then show them in a table as Figure 7. After comparing it with the remeasured size after normalizing it with AutoCAD, the results show that the error is in the range of less than 30%.

class	x_center	y_center	width	height	conf	Average Diameter (cm)	Image
0	0.634	0.097667	0.05275	0.195333	0.662301	239.0997	LoLu_MAX_0255.txt
0	0.634313	0.636583	0.053375	0.222167	0.632901	264.0003	LoLu_MAX_0257.txt
0	0.932	0.110417	0.08525	0.091167	0.878895	184.35003	LoLu_MAX_0257.txt
0	0.61375	0.889	0.05775	0.217667	0.626025	265.2003	LoLu_MAX_0258.txt
0	0.911375	0.372167	0.09575	0.097	0.745815	202.2	LoLu_MAX_0258.txt
0	0.92225	0.58175	0.088	0.098167	0.591549	193.95003	LoLu_MAX_0259.txt
0	0.602125	0.05175	0.02425	0.0355	0.824215	61.05	LoLu_MAX_0260.txt
0	0.601375	0.575833	0.02375	0.036667	0.637587	61.50003	LoLu_MAX_0262.txt
0	0.607938	0.872917	0.024375	0.0365	0.809461	62.1	LoLu_MAX_0263.txt
0	0.270813	0.618917	0.031875	0.067833	0.604777	99.29997	LoLu_MAX_0264.txt
0	0.307625	0.690583	0.03575	0.0465	0.641148	84.75	LoLu_MAX_0267.txt

Figure 7 Average diameter results

Thanks to the GPS information of the drone, we can determine the area having the potholes. The accuracy of this information is from 0.5 meters to 1.5 meters. In Figure 8, we present this information on a map with an image of the pothole at each location that contains it for easier tracking.

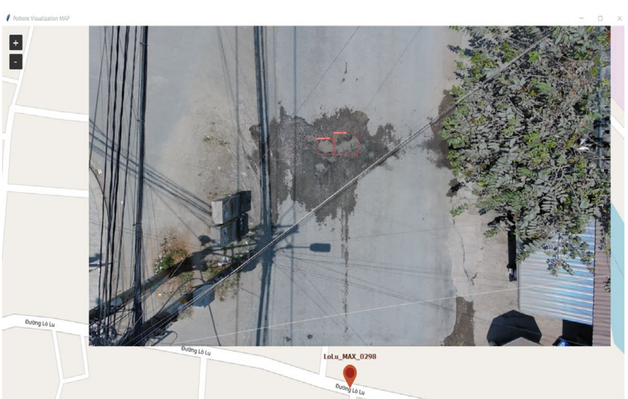


Figure 8 Location has detected potholes

4.0 CONCLUSION

In summary, the project has built a model to identify potholes from images taken by drones. With the assistance of drones, road surface inspection can be more effectively conducted. The results of size and position help the inspector to detect damage and danger early to make a proper repair schedule. Furthermore, these acquired datasets and results can be extended to identify additional road surface defects. The process is easy to perform, fast and convenient so it can be widely applied by anyone, especially authority figures. The authors believe that with this outcome, if the government apply this method to detect road potholes, the road maintenance process will be shortened, and potholes will also be repaired as soon as possible to give the passer-by the best road. To assess the risk of potholes more accurately, depth determination is essential. However, the current method cannot perform depth measurement, so additional research is needed. In future work, other types of defects will be also taken into account.

Acknowledgement

This research is funded by VNU-HCM under grant number C2023-20-25. We acknowledge the support of time and facilities from Ho Chi Minh City University of Technology (HCMUT), VNU-HCM for this study.

Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper

References

[1] V. P. a. S. T. a. S. H. a. H. Fukai, 2018. "A Deep Learning-Based Approach for Road Pothole Detection in Timor Leste," *2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)*, 279-284, DOI : https://doi.org/10.1109/SOLI.2018.8476795

[2] E. B. a. S. O. A. Akagic, 2017. "Pothole detection: An efficient vision based method using RGB color space image segmentation," in *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* DOI: https://doi.org/10.23919/MIPRO.2017.7973589

[3] A. A. a. A. K. a. T. A. a. A. Ammar, 2017. "RoadSense: Smartphone Application to Estimate Road Conditions Using Accelerometer and Gyroscope," *IEEE Sensors Journal,* 17: 4231-4238. DOI: https://doi.org/10.1109/JSEN.2017.2702739

[4] J. M. a. H. T. a. H. X. a. J. S. a. J. Cao, 2024. "Reconstruction-based anomaly detection for multivariate time series using contrastive generative adversarial networks," *Information Processing & Management,* 61: 103569, DOI: https://doi.org/10.1016/j.ipm.2023.103569

[5] J. M. a. H. T. a. H. X. a. J. S. a. J. Cao, 2024. "Reconstruction-based anomaly detection for multivariate time series using contrastive generative adversarial networks," Information Processing & Management. 61: 103569, DOI : https://doi.org/10.1109/ICASSP48485.2024.10446766

[6] Young-Mok Kim, Young-Gil Kim, Seungjoo Son, S. Lim, Bong-Yeol Choi, Doo-Hyun Choi, 2022. Review of Recent Automated Pothole-Detection Methods. *Applied Science* 12: 5320. DOI: https://doi.org/10.3390/app12115320

[7] Akhila Daniel Dana Mohammed Ali & Haval A. Sadeq, 2022. Road Pothole Detection Using Unmanned Aerial Vehicle Imagery and Deep Learning Technique. *ZANCO Journal of Pure and Applied Sciences,* 34: 107–115. DOI: https://doi.org/10.21271/zjpas.34.6.12

[8] S. K. M. H. Y. M. O. U. A. A. Muhammad Haroon Asad, 2024. "Pothole Detection Using Deep Learning: A Real-Time and AI-on-the-Edge Perspective," in *Proceedings of the 12th International Conference on Soft Computing for Problem Solving. SocProS 2023. Lecture Notes in Networks and Systems*, DOI: https://doi.org/10.1007/978-981-97-3180-0_34

[9] Ultralytics, 2022. *Ultralytics.* [Online] Available at: https://community.ultralytics.com/t/minimum-maximum-object-bounding-box-size/126

[10] D. Thuan, 2021."Evolution of YOLO algorithm and YOLOv5: The state-of-the-art object detection algorithm," [Online] Available at: https://www.theseus.fi/handle/10024/452552

[11] Abhishek Ranjan & Rajendra Machavaram, 2022. Detection and Localisation of Farm Mangoes using YOLOv5 Deep Learning Technique. *2022 IEEE 7th International conference for Convergence in Technology,* 1-5. DOI : https://doi.org/10.1109/I2CT54291.2022.9825078

[12] Ashit Patel, Lei Huang, Gustavo Vejerano, 2022. Pothole Detection from Dash Camera Images using YOLOv5. *26th International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV'22).* DOI: https://doi.org/10.1016/j.trpro.2023.11.255

[13] Y. H. Y. Y. ,. M. G. a. R. L. Ruoli Yang, 2017. "Fruit Target Detection Based on BCo-YOLOv5 Model," *Mobile Information Systems,* DOI : https://doi.org/10.1155/2022/8457173

[14] Abhishek Ranjan & Rajendra Machavaram, 2022. Detection and Localisation of Farm Mangoes using YOLOv5 Deep Learning Technique. *2022 IEEE 7th International conference for Convergence in Technology,* 1-5. DOI : https://doi.org/10.1109/I2CT54291.2022.9825078

[15] Ashit Patel, Lei Huang, Gustavo Vejerano, 2022. Pothole Detection from Dash Camera Images using YOLOv5. *26th International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV'22).* DOI: https://doi.org/10.1016/j.trpro.2023.11.255

[16] A. E. K. A. B. &. S. K. Prabhat Singh, 2024. "Road pothole detection from smartphone sensor data using improved LSTM," *Multimed Tools Application*, 83: 26009–26030. DOI : https://doi.org/10.1007/s11042-023-16177-0

[17] Robotics, A., 2021. *Autel Robotics.* [Online] Available at: https://www.autelrobotics.com/productdetail/evo-nano-series-drones.html#jsgg