Article

Mask R-CNN-Based Stone Detection and Segmentation for Underground Pipeline Exploration Robots

Humayun Kabir [1] and Heung-Shik Lee [2,*]

[1] Department of Integrated System Engineering, Inha University, Incheon 22212, Republic of Korea
[2] Department of Smart Mobility Engineering, Joongbu University, Goyang-si 10279, Republic of Korea
* Correspondence: jsheung@joongbu.ac.kr

Abstract: Stones are one of the primary objects that impede the normal activity of underground pipelines. As human intervention is difficult inside a narrow underground pipe, a robot with a machine vision system is required. In order to remove the stones during regular robotic inspections, precise stone detection, segmentation, and measurement of their distance from the robot are needed. We applied Mask R-CNN to perform an instant segmentation of stones. The distance between the robot and the segmented stones was calculated using spatial information obtained from a lidar camera. Artificial light was used for both image acquisition and testing, as natural light is not available inside the underground pipe. ResNet101 was chosen as the foundation of the Mask R-CNN, and transfer learning was utilized to shorten the training time. The experimental results of our model showed that the average detection precision rate reached 92.0; the recall rate was 90.0%; and the F1 score rate reached 91.0%. The distance values were calculated efficiently with an error margin of 11.36 mm. Moreover, the Mask R-CNN-based stone detection model can detect asymmetrically shaped stones in complex background and lighting conditions.

Keywords: Mask R-CNN; deep learning; segmentation; distance measurement; lidar camera; underground pipe

1. Introduction

In recent decades, with the proliferation of cities, tens of thousands of underground pipes have been installed to serve a multitude of functions. It is not common for oil and gas pipelines to carry large solid objects like stones or rocks. But any abnormality or pipeline break due to natural disaster may cause pipeline blockage through solid materials like stones and sand aggregates. At the same time, efficient transportation of gravel and sand aggregates through pipelines is essential for a variety of industries, from construction to mining. In order to ensure a seamless flow of materials, maintaining the integrity of these pipelines is crucial for their continued operation. There is a substantial risk in this domain of stones, rocks, or other solid debris entering into the pipeline, which can lead to disruptions, blockages, and structural damage. Therefore, routine checking and maintenance are essential. It is difficult for workers to explore inside a tiny pipe. Even with large-diameter pipes, workers are at risk of severe and often deadly injuries. Taking all the facts into account, a robotic solution came into place and has been increasingly deployed in extremely difficult environments, including underground pipeline exploration.

A large number of pipeline inspection robots have been developed over the past decades and can be operated in various pipelines such as water, oil, and gas, sewerage systems, or any specific pipeline where regular inspection is inevitable [1–3]. However, for the underground pipeline exploration robot, it is important to understand the pipeline environment and identify potential objects that could block the pipeline and interrupt its usual activities.

Currently, computer vision technology is extensively employed in robot inspection systems to understand an inspection site. Several studies have been conducted on vision-based systems. In a previous work [4], a pipe-inspection robot was developed based on YOLOv3 for defect detection and localization in a sewage pipeline. Also, another study worked on a deep learning-based method for underground sewage pipeline defect classification and location recognition [5]. However, the above studies were conducted for defect detection in pipes, but identifying objects inside the underground pipe is still very challenging, and stones are the most commonly found objects on the internal pipeline surface. Therefore, to recognize stones and their precise positioning along the underground pipeline, a robotic system with stone detection is needed.

The primary challenges in vision-based stone grasping are to develop robust and accurate stone detection algorithms that can effectively detect stones in different background and lighting conditions in an underground pipeline. In the past few years, many researchers have worked on rock detection and segmentation problems. For example, ref. [6] proposed an object detection model that was designed based on a modified U-net to recognize complex rock fragments. Further, a visual system with online rock mass assessment was developed based on semantic segmentation [7]. Rocks located in natural scenes were taken into account for another work [8], and a superpixel segmentation algorithm was developed to detect and locate rocks with an exact border. Furthermore, deep learning methods were utilized in the detection and segmentation processes of rock blast fragments [9].

Numerous research studies have been carried out on segmenting rocks and stones using Mask R-CNN. In reference [10], researchers explored dump particle segmentation and achieved a training accuracy of 97.2%. In another study [11], the author focused on blast fragmentation, achieving a precision score of 92%. In reference [12], a 93.18% detection accuracy was achieved by analyzing thin rock slices. The training dataset influences identification accuracy, particularly when it concerns recognizing small particles. Although the above approaches can detect rocks from daylight images of outdoor environments, when the lighting conditions change, the detection performance is significantly reduced. A CNN-based method for rock fragment classification inside a tunnel using indoor lighting was studied in [13]. However, the diameter of tunnel-like spaces provides sufficient room for object investigation, whereas narrow underground pipelines pose significant challenges for tasks related to object recognition.

For estimating the distance between the camera and segmented stones, and for the grasping action of the explorer robot, spatial knowledge is needed. Object distance measurement has been gradually used in various industrial applications. Several distance measurement methods are available, such as stereo vision image processing techniques [14–16], ultrasonic sensors [17], and depth cameras [18]. However, depth camera technology has gained much attention in recent years, mostly due to its high acquisition rates, long measurement ranges, and cost effectiveness. In this research study, we chose a single RGB-D camera (Intel RealSense L515) to perform both segmentation and distance measurement.

To automatically identify stones and determine their distances from the explorer robot, a stone detection and distance measurement method based on Mask R-CNN is proposed. An end-to-end instance segmentation architecture is applied in the proposed method, which accepts a single picture as input and, without any preprocessing, returns all instances of stones (detection and pixel-level classification). Finally, distance estimation and classification results are merged for use in real applications. In this research study, we focus on detecting the stones on the underground pipeline surface and measuring their distance from the exploration robot. This study does not consider the additional process of weighing the stone and moving it by the robot manipulator.

The contributions of this study are summarized as follows:

- A Mask R-CNN-based system is developed to identify the stones and measure their distances from the robot in the underground pipeline.
- A manually validated and labeled data set is presented for the segmentation work.

- This system offers a precise and fast robotic study of underground pipeline object detection research.

The structure of this paper is organized as follows: Section 2 introduces the Mask R-CNN model, data acquisition, data processing, and evaluation methods. Experimental details and test results are discussed in Section 3. In Section 4, limitations, future directions, and conclusions are presented.

2. Materials and Methods

2.1. Dataset Acquisition

Our research target for this study was to explore inside a narrow pipe to detect stones and measure their distance from the robot. Therefore, a pipe with an inner diameter of 25.0 cm was chosen for this study. Sand and gravel aggregates are typically classified according to their gradation, and the size of these aggregates is determined by the use for which they are intended. Past research [19] shows the gravel size can be 2 mm to 20 mm and more. However, in this study, the stone size was 5 mm to 50 mm. An Intel RealSense L515 camera (manufactured by Intel Corporation, 2200 Mission College Blvd, Santa Clara, CA 95054-1549, USA) mounted on a four-wheeler explorer robot captured all the images, as depicted in Figure 1. The camera system was placed inside the pipe and captured the images from various angles. During image capture, since natural light does not reach the underground pipe, an artificial light with moderate intensity was applied. For this study, stones scattered on divergent backgrounds were considered, such as a wetted pipe, a pipe partially filled with sand, and a regular pipe surface. At the time of capture, the resolution of the RGB frame was set to 1920 × 1080 pixels, and the storage format was JPEG. Three sample images with different backgrounds captured by the Intel RealSense L515 camera are shown in Figure 2. There were 365 captures, and approximately 3500 stone images were obtained. However, this stone dataset was incredibly challenging in terms of the segmentation task, considering the sizes of the stones, the backgrounds, and the light intensities.
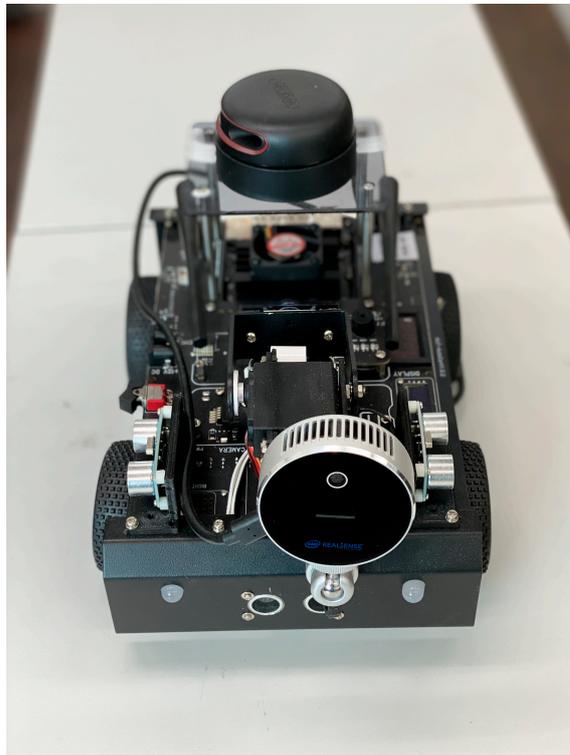

Figure 1. The explorer robot with LiDAR L515 camera.

Figure 2. Sample images.

2.2. Dataset Construction and Annotation

In order to ensure the accuracy of model training, the images were scaled down to 760 × 570 pixels. Out of 365 images overall, 280 were randomly selected for the Mask R-CNN model training, 30 images for verification, and 55 images for testing. The details are shown in Table 1. The VIA Tool (VGG image annotator) was utilized to manually create the annotation [20]. Each image contains multiple polygonal masks of the different shapes of the annotated stones at pixel level. Figure 3 demonstrates an example of manually annotated images with multiple stones.

Table 1. Splitting of training, test, and validation dataset.

Usage	#Images	#Stones
Training	175	2185
Validation	75	937
Testing	30	375
Total	280	3497

Figure 3. Manually annotated images using VIA tool.

2.3. Mask R-CNN

This paper utilized the Mask R-CNN model to implement the stone detection system. The Mask R-CNN [21] model is an extended algorithm based on a fusion of the Faster R-CNN [22] object detection algorithm. Mask R-CNN introduces a novel feature in the form of a 'mask', which can provide pixel-to-pixel estimation of the shape of the detected object. Therefore, the Mask R-CNN model is capable of both object detection and instance segmentation, while the Faster R-CNN model is unable to provide the pixel-to-pixel analysis

at the output level and was only designed for object detection. Mask R-CNN outputs a segmentation mask in each region of interest (ROI) through a fully convolutional network (FCN). Figure 4 shows the architecture of the Mask R-CNN model.
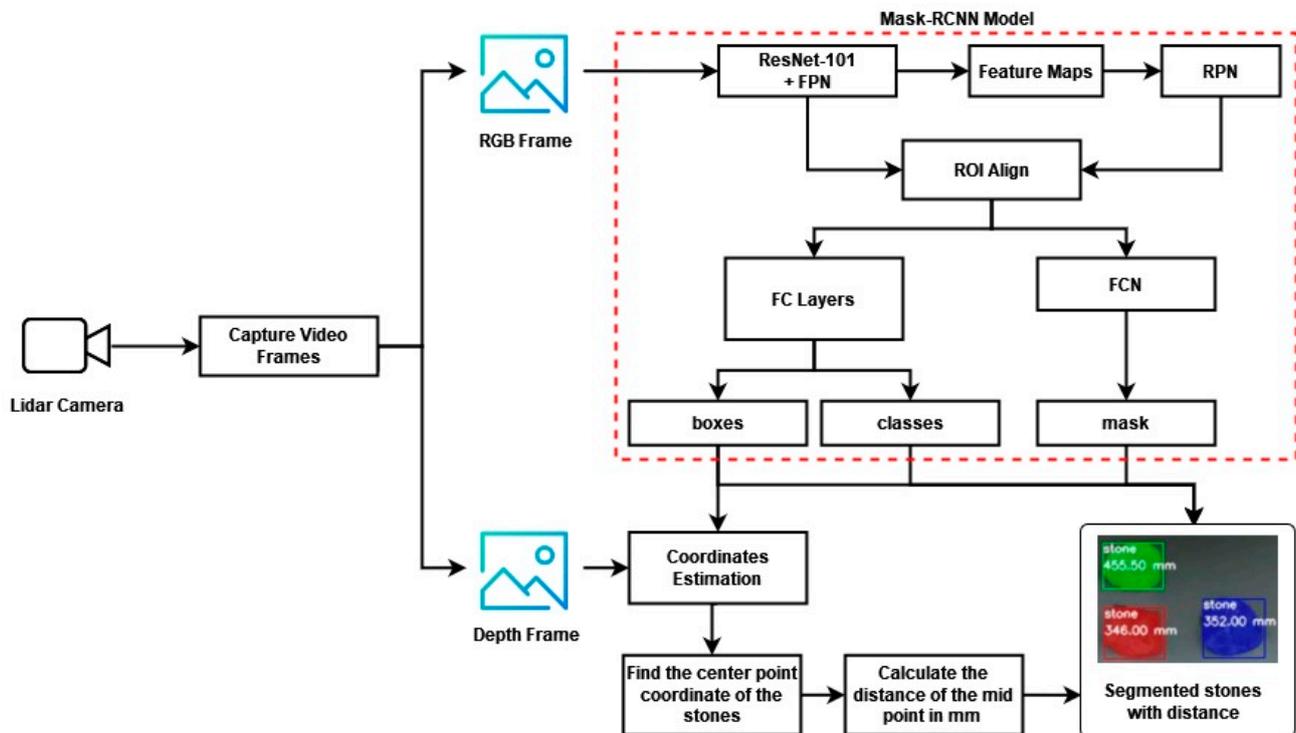

Figure 4. Framework of the Mask R-CNN based stone segmentation and distance measurement system.

2.4. Training and Loss Function

In this study, Mask R-CNN was implemented based on the feature pyramid network (FPN), which extracts low-level features as well as high-level features, and ResNet101 was applied as a backbone. ResNet101 makes use of a residual learning process, which decreases the number of parameters that must be adjusted and minimizes the total computing cost. However, the learning time taken by ResNet101 is too long, as the structure has 101 layers, and every layer requires considering so many things while solving a new problem. Therefore, we applied transfer learning by employing a pretrained model trained with an MS-COCO (Microsoft Common Objects in Context) dataset to lessen the learning time and determine the number of images required to train the whole network [23]. The RPN, classifier, and mask heads of the network were trained up to 100 epochs to fine-tune the weights for detecting stones. The stochastic gradient descent with a momentum of 0.9 was used to train the model, where the learning rate was set to 0.001, and the details of the training specification are shown in Table 2.

Table 2. Specification of the training and testing platform.

Attribute Name	Value
CPU	AMD Ryzen 7 5800H at 3.2 GHz $\times$ 16
Memory	40 GB
GPU	NVDIA RTX3060
OS	Windows 10

The loss function of the Mask R-CNN [21] model is calculated as Equation (1) in the following:

$$L = L_{class} + L_{bbox} + L_{mask} \tag{1}$$

where L represents the total validation loss and is calculated with the combination of L_{class} (classification loss), L_{bbox} (bounding box loss), and L_{mask} (mask loss). Classification loss takes into account the model's confidence in predicting the true class, and the bounding box loss function provides the error between the actual bounding box and the predicted bounding box. Finally, the mask loss function generates a pixel-to-pixel mask for all detected classes.

2.5. Distance Measurment of the Segmented Stones

In this research study, an Intel RealSense LiDAR L515 camera was used, which uses laser scanning technology and is compact in size (61 mm × 26 mm) and specially designed for an indoor environment. The camera depth output resolution is 1024 × 768; the field of view is 70° × 55°; and the measurement range is 0.25 m to 9.0 m. And the RGB frame resolution is 1920 × 1080, and the field of view is 70° × 43°. Moreover, it has the ability to generate 23 million depth points per second, and this advantage was utilized for real-time distance calculation.

The basic idea was to align the RGB and depth frames obtained from the lidar camera. By using the stone detection model based on Mask R-CNN, stones are segmented into the bounding box. The bounding box of the segmented stone is defined by the coordinates of its top-left (x_1, y_1) and bottom-right (x_2, y_2) corners. And we calculated the middle point coordinates of the segmented stone using the following equations:

$$x = \frac{x_1 + x_2}{2} \tag{2}$$

$$y = \frac{y_1 + y_2}{2} \tag{3}$$

We utilized the Intel RealSense camera APIs to align the depth frame. Then, the center points of each stone were estimated from the bounding box coordinates using Equations (1) and (2). The distance value of the center point coordinate was obtained from the aligned depth map. Finally, the distance between the explorer robot and each segmented stone was shown with a label. The algorithmic process of distance measurement is illustrated in Figure 4. A measuring tape with a precision of ±0.5 mm is used to verify the estimated distance from our model.

2.6. Evaluation for the Stone Detection Model

The stone detection model was evaluated by applying three performance matrices: precision, recall, and F1 score. Thirty selected images were used for model evaluation. *True Positive* (*TP*) represents the number of positive cases that are detected as positive. *False Positive* (*FP*) represents the number of negative cases that are detected as positive. *False Negative* (*FN*) represents the number of positive cases that are detected as negative. These parameters are calculated by using the following equations:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{4}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{5}$$

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{6}$$

3. Experimental Results and Analysis

3.1. Evaluation of the Stone Detection Model

The loss function change of Mask R-CNN during training is shown in Figure 5. In the training process of our model, while the learning rate was set to 0.001, the loss value dropped quickly from 0.9 to 0.6 then gradually dropped to less than 0.2. The Precision–Recall curve shown in Figure 6 indicates that the trained model has achieved an adequate detection accuracy.
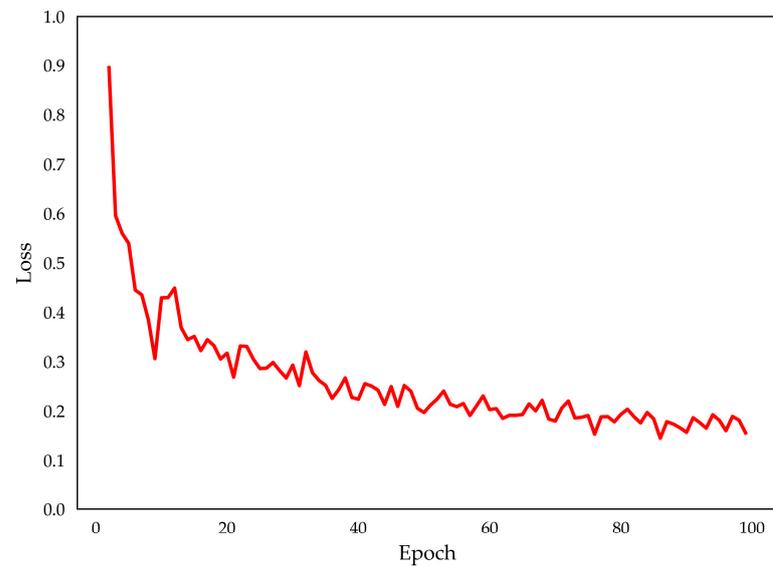

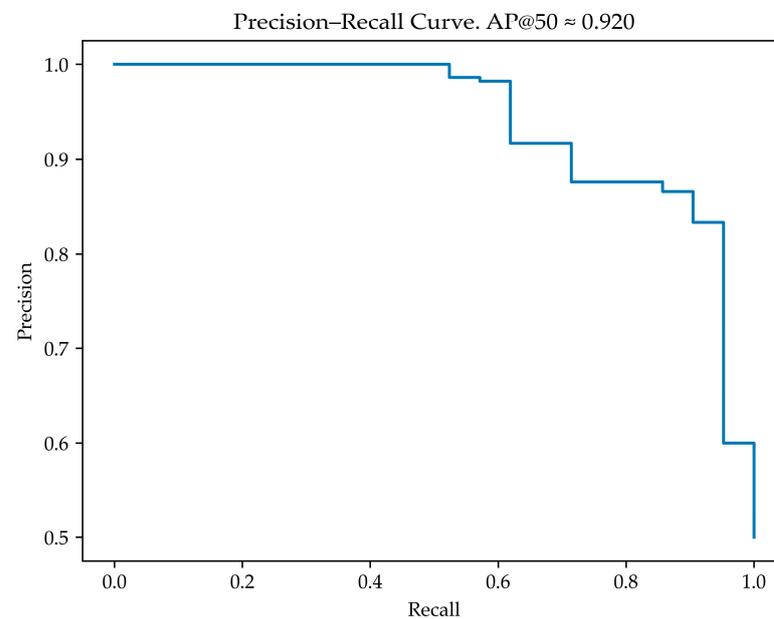**Figure 5.** Training loss of Mask R-CNN.

Figure 6. Precision–Recall curve of Mask R-CNN.

An example of true positive results of the stone detection model is shown in Figure 7, where all the stones are detected and segmented correctly. Figure 7a shows the original image of a regular pipe with low light intensity. In Figure 7b, the detected stones are shown in the square box with the confidence score. And Figure 7c shows the binary image with all the segmented stones.
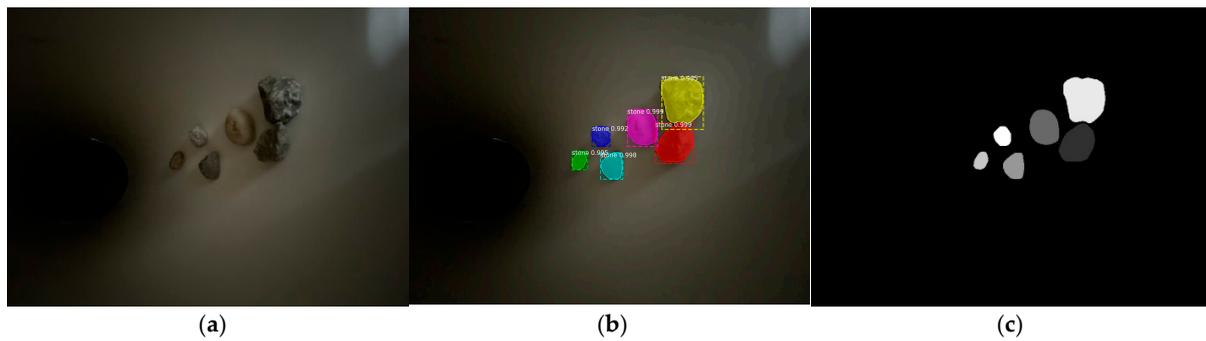
(**a**)   (**b**)   (**c**)

Figure 7. Example of accurate detection result. (**a**) Original image with normal background and low light intensity, (**b**) Output of the stone detection model (***true positive*** case), and (**c**) Binary image with all segmented stones.

While our model successfully identified the stones in the majority of test photos, there were instances where it was unable to identify the objects. To demonstrate the shortcomings of our model, we ran the test for false negative and false positive scenarios. The details of these tests are shown in Figures 8 and 9, respectively.
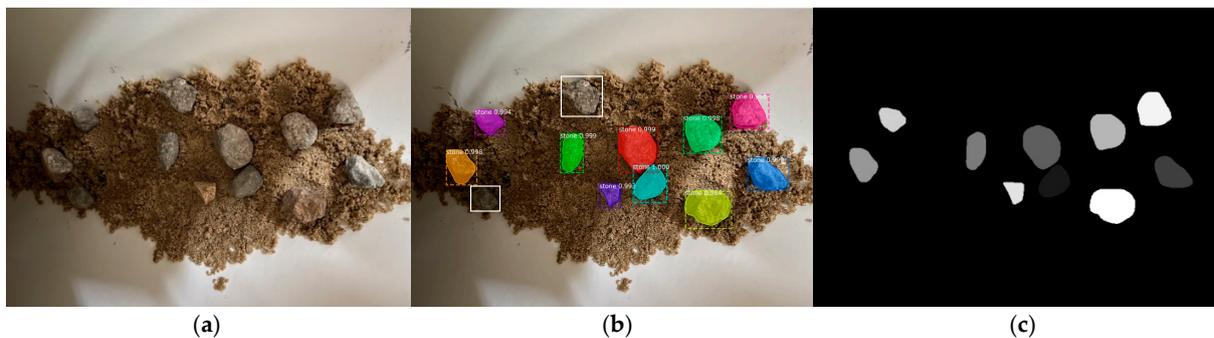

(**a**)   (**b**)   (**c**)

Figure 8. Examples of false detection. (**a**) Original image with sand-filled background and moderate light intensity, (**b**) Output of the stone detection model (***false negative*** case), and (**c**) Binary image of segmented stones.

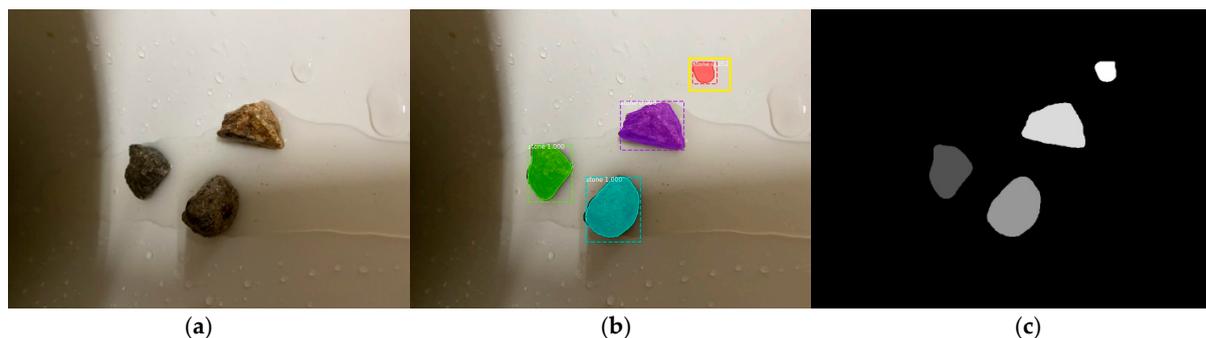(**a**)   (**b**)   (**c**)

Figure 9. Examples of wrong detection. (**a**) Original image with wetted background and moderate light intensity, (**b**) Output of the stone detection model (***false positive*** case), and (**c**) Binary image with false positive segmented stones.

In Figure 8, two objects indicated by white square boxes are expected to be detected by the model but are not detected and thus create false negative cases. There were sufficient training images similar to the Figure 8 backgrounds that are partially filled with sand. The original image of the stones with a sand background in moderate light intensity is shown in Figure 8a. Figure 8b shows the detected stones with a color mask and the undetected stones with square boxes. Figure 8c demonstrates the binary image of the segmented stones.

Finally, an example of a false positive case is shown in Figure 9, where an object marked by a yellow squared box is wrongly detected as a stone by the stone detection model. Figure 9a shows the original picture of a wetted background with stones in moderate light intensity. Figure 9b shows the segmented stones with a false positive case. And Figure 9c shows the binary image of the segmented result.

3.2. Testing of Distance Measurement

Our model is tested in a real environment to measure the distance between the explorer robot and the segmented stones. The explorer robot shown in Figure 1 is used for testing the real-time stone segmentation and distance measurement systems inside the underground pipeline. The robot is equipped with an L515 LiDAR camera, and the major specification details are shown in Table 3. The distance of each masked stone in the RGB frame is calculated through the depth frame by estimating their centroid point. Examples of underground pipeline object distance measurements in various backgrounds along with spatial frames are shown in Figure 10. The class name and measured distance in millimeters is displayed on each segmented stone. Figure 10a shows a regular background of an underground pipe where three stones are segmented by the stone detection model, and their distances are calculated and marked inside the squared boxes; the depth frame shows the clear spatial information. Figure 10b shows a frame partially filled with sand, and all the eight stones are correctly segmented, including their distance values. Figure 10c shows the underground pipeline surface with water where six stones are partially immersed, and the model detects those perfectly and provides distance information.

Table 3. Specification of the robot platform.

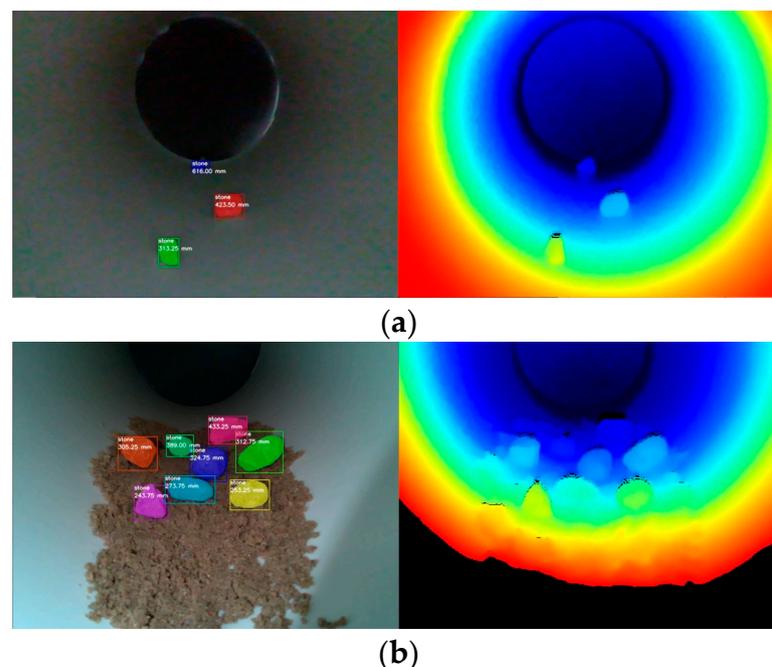| Attribute Name	Value
CPU	Quad-Core ARM A57 @ 1.43 GHz
Memory	4 GB 64-bit LPDDR4 25.6 GB/s
GPU	Maxwell Core 128EA
OS	Ubuntu 20.04

(a)

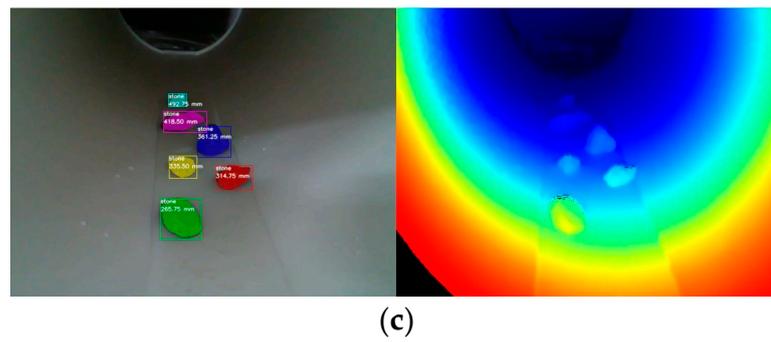

(b)

Figure 10. *Cont.*

Figure 10. Examples of distance measurement with depth image. (**a**) Stones are in an empty pipe; (**b**) Stones are scattered on sand; and (**c**) Stones are dipped in water in the pipe.

In this study, object distance measurement is tested on 40 segmented objects from eight frames, and their values are recorded to estimate the average error. The experimental images with distance information are shown in Figure 11. The error difference between the actual distance and the measured distance is shown in millimeters in Table 4. A measuring tape with a precision of ±0.5 mm is used to verify the estimated distance from our model. The experimental result shows that the average absolute error is 11.36 mm. We developed a GUI using Python to monitor the real-time pipeline inspection results using our model. In Figure 12, the system interface is presented, where segmented stones with their estimated distances are shown.
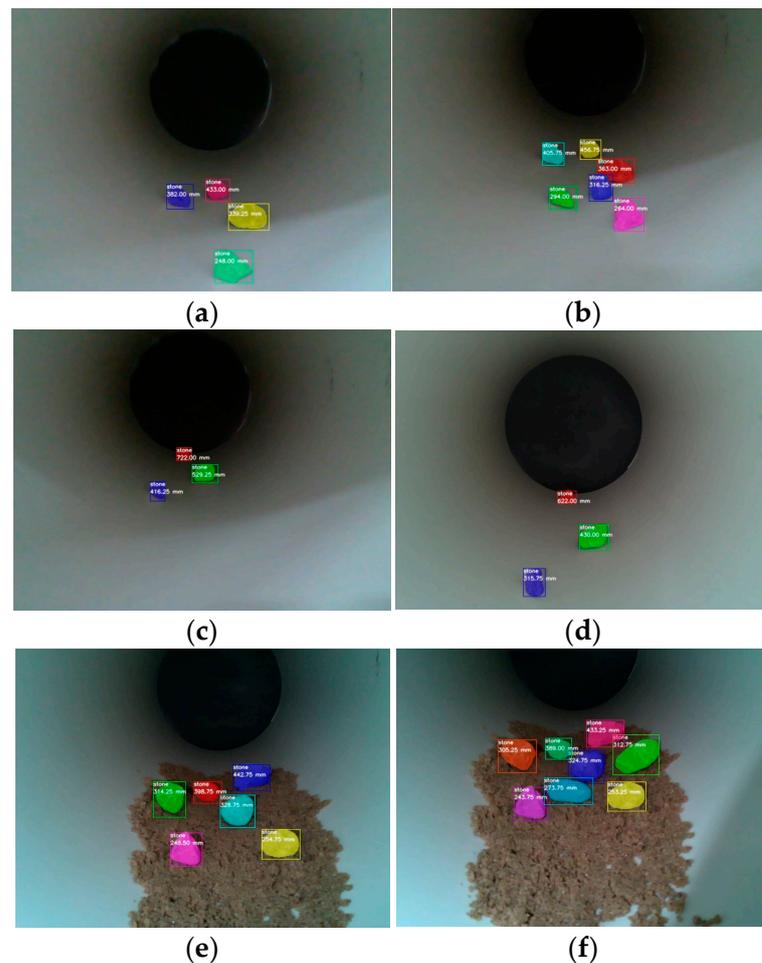

Figure 11. *Cont.*

(**g**)   (**h**)

Figure 11. Measurement data. (**a**) Showing four stones and their distances inside the pipe's hollow space; (**b**) Showing six stones and their distances inside the pipe's hollow space; (**c**) Showing three stones and their distances inside the pipe's hollow space; (**d**) Showing three stones and their distances inside the pipe's hollow space; (**e**) Displaying six stones and their distances in a pipe partially filled with sand; (**f**) Displaying eight stones and their distances in a pipe partially filled with sand; (**g**) Displaying four stones and their distances in a pipe with water; and (**h**) Displaying six stones and their distances in a pipe with water.
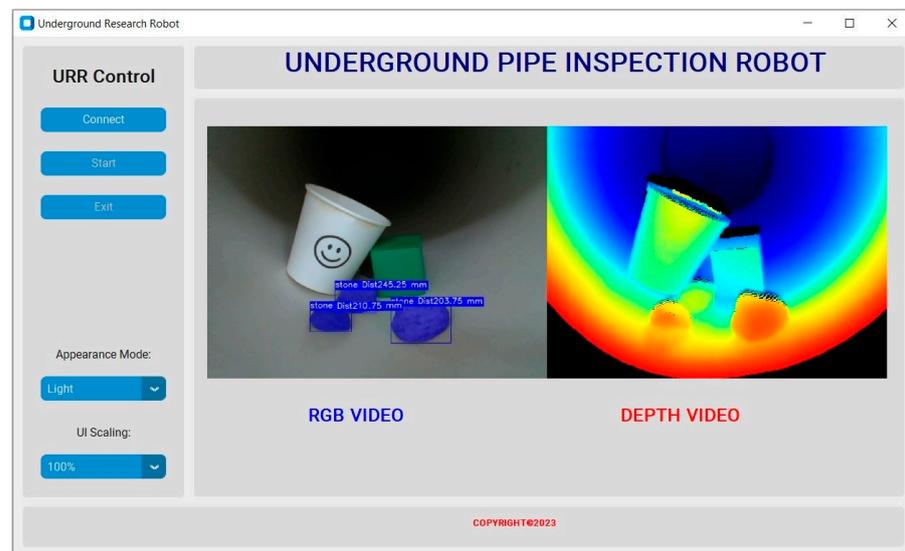

Figure 12. Stone segmentation and distance measurement system interface.

Table 4. Average distance measurement error.

Trail Images Detail			Measurement Unit: mm		
Test Images	**No. of Objects**	**Serial No.**	**Actual Distance**	**Measured Distance**	**Absolute Error**
Figure 11a	4	1	260.50	248.00	12.50
		2	355.00	339.25	15.75
		3	389.50	382.00	7.50
		4	437.50	433.00	4.50
Figure 11b	6	1	270.00	264.00	6.00
		2	305.00	294.00	11.00
		3	326.00	316.25	9.75
		4	357.00	363.00	6.00
		5	418.00	405.75	12.25
		6	460.00	456.75	3.25
Figure 11c	3	1	431.00	416.25	14.75
		2	537.00	529.25	7.75
		3	726.00	722.00	4.00

Table 4. *Cont.*

Trail Images Detail			Measurement Unit: mm		
Test Images	**No. of Objects**	**Serial No.**	**Actual Distance**	**Measured Distance**	**Absolute Error**
Figure 11d	3	1	328.50	315.75	12.75
		2	447.00	430.00	17.00
		3	635.00	622.00	13.00
Figure 11e	6	1	262.00	248.50	13.50
		2	268.50	254.75	13.75
		3	335.00	314.25	20.75
		4	317.00	328.75	11.75
		5	379.00	398.75	19.75
		6	453.00	442.75	10.25
Figure 11f	8	1	260.50	243.75	16.75
		2	267.00	253.25	13.75
		3	280.50	273.75	6.75
		4	318.00	305.25	12.75
		5	320.50	312.75	7.75
		6	313.00	324.75	11.75
		7	402.00	389.00	13.00
		8	441.00	433.25	7.75
Figure 11g	4	1	242.50	230.75	11.75
		2	318.00	310.50	7.50
		3	357.00	342.50	14.50
		4	410.00	393.75	16.25
Figure 11h	6	1	278.00	265.75	12.25
		2	309.00	314.75	5.75
		3	345.00	335.50	9.50
		4	373.50	361.25	12.25
		5	431.00	418.50	12.50
		6	507.00	492.75	14.25
Total	40				454.25
Average absolute error					11.36

4. Conclusions

In this research study, Mask-RCNN was applied to detect stones inside an underground pipeline through an inspection robot. Stones are a very common object that can be found on pipeline surfaces during a regular inspection period. Since natural sunlight does not reach into an underground pipe, artificial lights with moderate intensity are used during the inspection time. Furthermore, because the stones vary in size and color, it is difficult to accurately identify the stones using a straightforward image processing technique. To detect the stones in moderate light, it is crucial to understand the background and the stones. Through a simple image processing approach, it is hard to contend with background noise. Mask R-CNN can reliably detect stones and understand the background of the pipeline, even at low or medium brightness. All the images were collected from inside a pipe with moderate lighting and a variety of backgrounds and stones. The total number of images was 280, which covers various angles and contains approximately 3500 stones. The test results of our stone detection model showed that the precision, recall, and F1 scores were 92.0%, 90.0%, and 91.0%, respectively. Distance measurement algorithms can be used in real time to calculate the distance between a robot and each stone with an error margin of 11.36 mm. The idea of this project can be useful for researchers in this field for any kind of underground pipeline exploration work.

References

1. Kawaguchi, Y.; Yoshida, I.; Kurumatani, H.; Kikuta, T.; Yamada, Y. Internal pipe inspection robot. In Proceedings of the 1995 IEEE International Conference on Robotics and Automation, Nagoya, Japan, 21–27 May 1995; Volume 1, pp. 857–862. [CrossRef]
2. Roh, S.G.; Ryew, S.M.; Yang, J.H.; Choi, H.R. Actively steerable in-pipe inspection robots for underground urban gas pipelines. In Proceedings of the 2001 ICRA, IEEE International Conference on Robotics and Automation (Cat. No.01CH37164), Seoul, Republic of Korea, 21–26 May 2001; Volume 1, pp. 761–766. [CrossRef]
3. Abdellatif, M.; Mohamed, H.; Hesham, M.; Abdelmoneim, A.; Kamal, A.; Khaled, A. Mechatronics Design of an Autonomous Pipe-Inspection Robot. *MATEC Web Conf.* **2018**, *153*, 02002. [CrossRef]
4. Hu, Z.; Zhou, J.; Yang, B.; Chen, A. Design of Pipe-inspection Robot Based on YOLOv3. *J. Phys. Conf. Ser.* **2022**, *2284*, 012023. [CrossRef]
5. Hassan, S.I.; Dang, L.M.; Mehmood, I.; Im, S.; Choi, C.; Kang, J.; Park, Y.-S.; Moon, H. Underground sewer pipe condition assessment based on convolutional neural networks. *Autom. Constr.* **2019**, *106*, 102849. [CrossRef]
6. Qiao, W.; Zhao, Y.; Xu, Y.; Lei, Y.; Wang, Y.; Yu, S.; Li, H. Deep learning-based pixel-level rock fragment recognition during tunnel excavation using instance segmentation model. *Tunn. Undergr. Space Technol.* **2021**, *115*, 104072. [CrossRef]
7. Xue, Z.; Chen, L.; Liu, Z.; Lin, F.; Mao, W. Rock segmentation visual system for assisting driving in TBM construction. *Mach. Vis. Appl.* **2021**, *32*, 77. [CrossRef]
8. Dunlop, H. Automatic Rock Detection and Classification in Natural Scenes. Master's Thesis, Carnegie Mellon University, Pittsburgh, PA, USA, 2006.
9. Bamford, T.; Esmaeili, K.; Schoellig, A.P. A deep learning approach for rock fragmentation analysis. *Int. J. Rock Mech. Min. Sci. Géoméch. Abstr.* **2021**, *145*, 104839. [CrossRef]
10. Shrivastava, S.; Bhattacharjee, S.; Deb, D. Segmentation of mine overburden dump particles from images using Mask R CNN. *Sci. Rep.* **2023**, *13*, 2046. [CrossRef] [PubMed]
11. Tsung-Shiang, H.; Bao, T.; Hoang, Q.V.; Drebenstetd, C.; Van Hoa, P.; Thang, H.H. Measuring blast fragmentation at Nui Phao open-pit mine, Vietnam using the Mask R-CNN deep learning model. *Min. Technol.* **2021**, *130*, 232–243. [CrossRef]
12. Liu, T.; Li, C.; Liu, Z.; Zhang, K.; Liu, F.; Li, D.; Zhang, Y.; Liu, Z.; Liu, L.; Huang, J. Research on Image Identification Method of Rock Thin Slices in Tight Oil Reservoirs Based on Mask R-CNN. *Energies* **2022**, *15*, 5818. [CrossRef]
13. Yang, Z.; He, B.; Liu, Y.; Wang, D.; Zhu, G. Classification of rock fragments produced by tunnel boring machine using convolutional neural networks. *Autom. Constr.* **2021**, *125*, 103612. [CrossRef]
14. Mustafah, Y.M.; Noor, R.; Hasbi, H.; Azma, A.W. Stereo vision images processing for real-time object distance and size measurements. In Proceedings of the 2012 International Conference on Computer and Communication Engineering (ICCCE), Kuala Lumpur, Malaysia, 3–5 July 2012; pp. 659–663. [CrossRef]
15. Zivingy, M.; Melhum, A.; Kochery, F.A. Object distance measurement by stereo vision. *Int. J. Sci. Appl. Inf. Technol.* **2013**, *2*, 5–8.
16. Tsung-Shiang, H.; Hsu, T.-S.; Wang, T.-C. An Improvement Stereo Vision Images Processing for Object Distance Measurement. *Int. J. Autom. Smart Technol.* **2015**, *5*, 85–90. [CrossRef]
17. Zhmud, V.A.; Kondratiev, O.; Kuznetsov, A.K.; Trubin, V.G.; Dimitrov, L.V. Application of ultrasonic sensor for measuring distances in robotics. *J. Phys. Conf. Ser.* **2018**, *1015*, 032189. [CrossRef]
18. Frangez, V.; Salido-Monzu, D.; Wieser, A. Assessment and Improvement of Distance Measurement Accuracy for Time-of-Flight Cameras. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1003511. [CrossRef]
19. Sahmaran, M.; Lachemi, M.; Hossain, K.M.; Ranade, R.; Li, V.C. Influence of aggregate type and size on ductility and mechanical properties of engineered cementitious composites. *ACI Mater. J.* **2009**, *106*, 308.
20. Dutta, A.; Zisserman, A. The VIA Annotation Software for Images, Audio and Video. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 2276–2279. [CrossRef]

21. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 386–397. [CrossRef] [PubMed]
22. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *arXiv* **2015**, arXiv:1506.01497v3. [CrossRef] [PubMed]
23. Lin, T.Y.; Maire, M.; Belongie, S.; Bourdev, L.; Girshick, R.; Hays, J.; Perona, P.; Zitnick, C.L.; Dollár, P. Microsoft COCO: Common Objects in Context. *arXiv* **2015**, arXiv:1405.0312.