2021 18th International Conference on Ubiquitous Robots (UR)
July 12-14, 2021. Gangneung-si, Gangwon-do, Korea

Detection of Rust from Images in Pipes Using Deep Learning

Akira Oyama, Hiroto Sato, Kaito Kosuge, Kosuke Uchiyama, Taro Nakamura and Kazunori Umeda

2021 18th International Conference on Ubiquitous Robots (UR) | 978-1-6654-3899-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/UR52253.2021.9494700

Abstract— **This paper proposes a method for detecting rust inside pipes using deep learning. In recent years, the number of pipes that have passed their useful life has been increasing, and earthworm-type robots have been developed to perform regularly inspections of sewage pipes. The images of the sewage pipe taken by the robot are trained on a Variational Auto Encoder, which is an unsupervised learning model, to detect abnormalities by taking the difference between the input image and the output image. In addition, the trained Residual Network is used to estimate the location of anomalies.**

I. INTRODUCTION

Sewage pipes are an important infrastructure in environmental protection and sanitation maintenance. However, in recent years, aged pipes that have passed 50 years have increased. Since such pipes cause water leakage and road collapses, regular inspection of the sewage pipe is required. Therefore, earthworm-type robots for pipe inspection as shown in Fig. 1 have been developed [1][2]. This robot is equipped with a camera on its head, which enables it to acquire images of the inside of sewage pipes.

In this paper, we focus on pressure pipes, which account for 5 to 10% of all sewage pipes. Compared with conventional sewage pipes, pressure pipes are less restricted by topography and can be configured relatively freely. In addition, the defects of pressure pipes differ depending on their materials: polyvinyl chloride pipes are deformed by soil pressure, and cast iron pipes have rust on the inner surface. It is necessary to detect these deformations and rusts from the image and determine their positions on the image.

Due to the recent development of deep learning, its practical application to the field of anomaly detection, such as the inspection of industrial products, has been actively developed [3]. In anomaly detection problems, it is difficult to perform supervised learning because the number of abnormal data is much smaller than that of normal data. For this reason, unsupervised learning using only normal data is often used, and any deviation from normal is considered as an abnormality. Therefore, it is effective to solve this problem as an anomaly detection problem where the defective part is considered as an anomaly in the inspection of the inside of a pipe.

Based on the above, this paper focuses on the detection of rust adhering to the cast iron pipe in the pressure pipe using images. In recent years, there has been a lot of research in the field of image processing on anomaly detection using Auto Encoder and object recognition using Residual Network (ResNet)[4], etc. In this research, we will detect rust using Variational Auto Encoder(VAE)[5] and estimate the location of anomalies using ResNet.

Fig. 1 Earthworm Robot [1][2]

II. PROPOSED METHOD

A. *Outline of Proposal Method*

The flow of the proposed method is as follows. First, the VAE learns the normal images (images without rust) from the images taken by the earthworm robot. Next, the trained VAE is used to compute a score that indicates the degree of abnormality of the input image, and then the abnormality is detected. In addition, the input and output images of VAE are input to ResNet, and its middle layer is extracted to estimate the anomalies.

When detecting anomalies using VAE, the reconstruction of the anomalous image may be distorted, and it may be difficult to identify the location of the anomaly in the image. Therefore, we use ResNet as well as VAE to detect abnormalities.

B. *Rust Detection Using VAE*

In this paper, we use VAE to detect rust on the inside of pipes. VAE is a network that adopts the framework of Bayesian inference to the usual Auto Encoder (AE), as shown in Fig. 2 [4].

AE is a neural network that obtains the features (latent variables) representing the input images by dimensionality reduction using Encoder, and reconstructs the images using the latent variables by Decoder. On the other hand, VAE outputs the mean and variance of the latent variables by assuming a probability distribution for the latent variables obtained by the encoder. Using these means and variances, the Decoder restores the image.

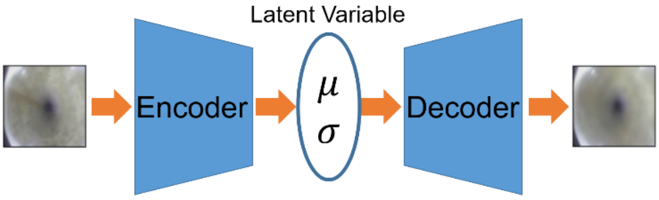

Fig. 2 Variational AutoEncoder

In this paper, anomaly detection is performed by taking the difference between the input image and the output image obtained by VAE.

C. *Estimation of Anomaly Location Using ResNet*

ResNet is a convolutional neural network for general object recognition, which is characterized by its very deep layers [5]. In general, the deeper the layers of a neural network, the higher the accuracy, but when the layers are extremely multi-layered, the product of the derivatives becomes too small, and the gradient vanishing problem occurs. ResNet is a model that eliminates this vanishing gradient problem by providing input shortcuts and calculating the residuals from the conventional output, thus enabling efficient training of very deep neural networks.

In this paper, we input each of the input and output images of the VAE into ResNet50, a trained 50-layer ResNet, and extract the first layer of each intermediate layer. The first layer of each intermediate layer is extracted, and the anomalies are estimated by taking the difference between the input and output intermediate layers. By extracting the first layer, we can find out where on the image contributed to the output of ResNet.

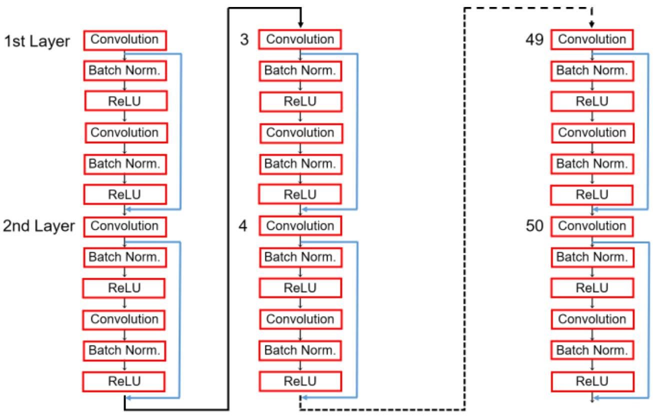

Fig. 3 Residual Network

III. Generation of Data That Imitates Rust Inside a Pipe

In order to learn and conduct experiments for anomaly detection, images of the inside of a pressure pipe are necessary. However, it is difficult to collect data of the inside of a pumping pipe at present. In this paper, two types of data that mimic a pumping pipe are generated and used for the experiment together with a 10-minute video of the inside of the pumping pipe that is currently available.

The first data set consists of images of round stickers of various colors attached to the inner surface of a commercial polyvinyl chloride pipe (Dataset A). This data set consists of two datasets: one with stickers and the other without stickers on the inner surface. Examples of the generated data are shown in Fig. 4 and 5. The normal data is considered to be relatively easy to detect abnormality because the tube has no pattern.

The second type of data set is images of a transparent pipe with a newspaper wrapped around it and stickers placed on top of it (Dataset B). As in Dataset A, the data with stickers is considered abnormal and the data without stickers is considered normal. Fig. 6 and Fig. 7 show examples of the data set B. Compared to the first set, the patterns are more complicated, and the difficulty of detecting abnormalities is considered to be higher.

Finally, the actual images in the pressure pipe is designated as Dataset C. However, since there is no cast iron pipe in this dataset, and it is not known whether there are any specific defects or not, the experiment is conducted assuming that joints between pipes and holes are defects. Therefore, we consider joints between tubes and holes as defects. Examples of normal and abnormal data are shown in Fig. 8 and 9.

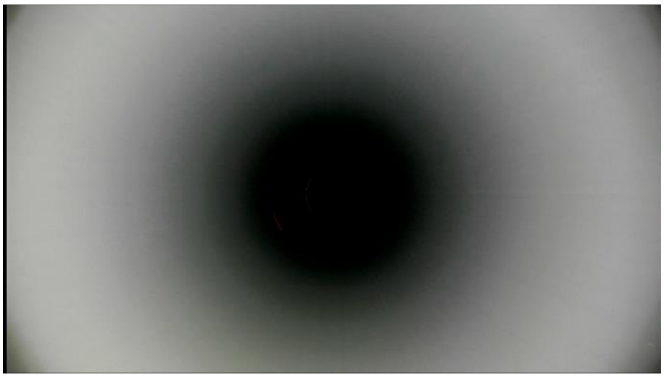

Fig. 4 Image example (Dataset A, normal data)

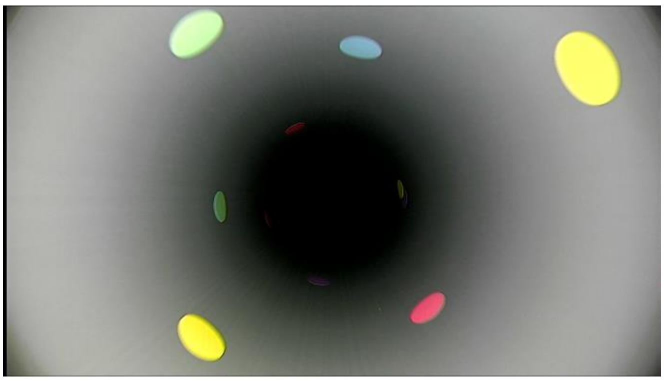

Fig. 5 Image example (Dataset A, anomaly data)

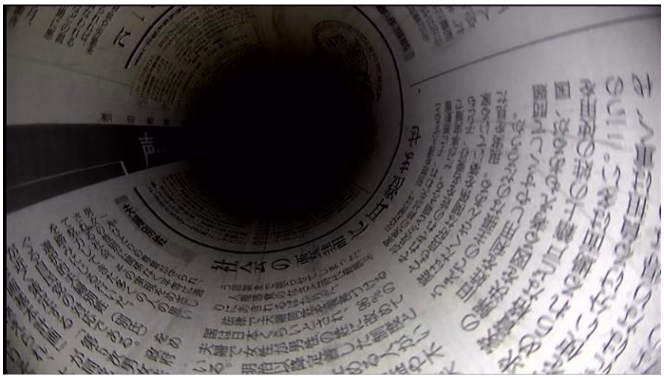

Fig. 6 Image example (Dataset B, normal data)

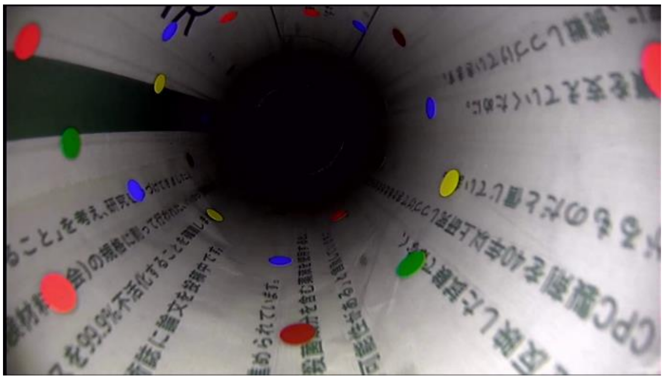

Fig. 7 Image example (Dataset B, anomaly data)

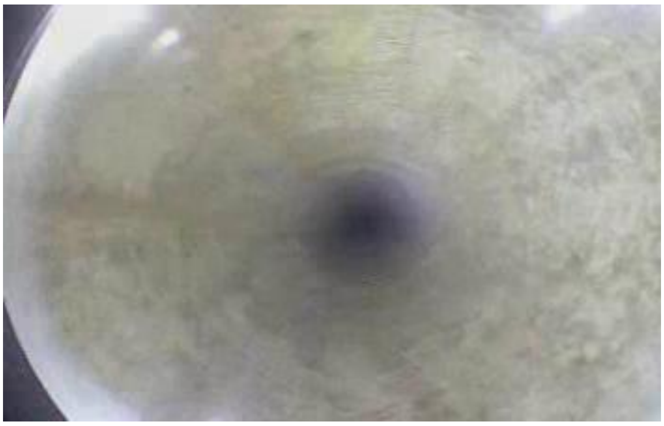

Fig. 8 Image example (Dataset C, normal data)

Fig. 9 Example image (Dataset C, anomaly data (joints are considered anomaly))

IV. EXPERIMENTS OF ANOMALY DETECTION AND DEFECT LOCATION ESTIMATION INSIDE PIPES

A. Experimental Conditions

Using the proposed method, we conducted experiments to detect and estimate the location of abnormalities in pipes. We used the three data sets described in section 3. The number of training data is about 3,000 for data set A, 10,000 for data set B, and 18,000 for data set C. The number of epochs is set to 30 and 10. The dimensions of the latent variables were set to 32, and the batch size was set to 16 for dataset A and 32 for datasets B and C. For ResNet50, the trained model of Pytorch [6] was used. The following equation (1) was used as the loss function, and Adam [7] was used for optimization.

$$L_{total} = L_{recon} + L_{KL} \tag{1}$$

$$L_{recon} = \sum \{-p \log q - (1-p)\log(1-q)\} \tag{2}$$

$$L_{KL} = \frac{1}{2}\sum(-2\sigma + \mu^2 + \sigma^2 - 1) \tag{3}$$

L_{total}: loss function, L_{recon}: reconstruction loss, L_{KL}: KL divergence, p: pixel value of input image, q: pixel value of output image, μ: mean vector of latent variables, σ: standard deviation vector of latent variables

B. Experimental Results

Fig. 10, 11, and 12 show the experimental results of the VAE input image, output image, their differences, and the estimation of anomalies by ResNet50. The images are, in order from the first row, the input image, the output image, the difference image, and the middle layer of ResNet. The value of the intermediate layer of the difference image and ResNet is smaller when the color is blue, and higher when the color is closer to red. The score calculated by summing the pixel values of the difference image is shown in Tables 1, 2, and 3. The numbers in Tables 1, 2, and 3 are ones counted from the left in Fig. 10, 11, and 12.

From Fig. 10, it can be seen that there are images in which the reconstruction of the VAE is particularly broken in the first, third, and fourth images from the left. Table 1 also shows that the anomaly level of these three images is high. This may be due to the lack of training data or the insufficient adjustment of hyperparameters. However, if we look at the middle layer of ResNet, we can see that even if the reconstruction of the VAE is broken, the areas with stickers are highlighted in red. Therefore, it can be said that it is possible to estimate the anomalies using the middle layer of ResNet.

Looking at Fig. 11, we can see that the input and output difference images for the first, second, third, and eighth normal data are generally blue. The difference images for the fourth, fifth, sixth, and seventh images show that the reconstructions near the stickers are broken and are highlighted in red. This suggests that the VAE performed better than dataset A. This is because dataset B has a larger number of training data than dataset A. In addition, the extraction results of the middle layer of ResNet show that the areas where stickers are applied are shown in red, but some areas are misidentified or not emphasized, such as the images No. 5 and No. 7. This is thought to be due to the fact that some parts of the VAE reconstruction have been corrupted. In addition, when the sticker is located far away from the camera, as shown in No. 7, the image becomes darker and it is difficult to identify the sticker with dark color.

Fig. 12 shows that the reconstruction of the VAE is relatively good due to the large number of training data. Table 3 shows that the minimum score of the abnormal data is 195 and the maximum score of the normal data is 173, suggesting that the thresholding can discriminate between abnormal and normal data.

Based on these results, it can be concluded that VAE can be used to detect anomalies even in complex patterns in pipes, if a sufficient amount of data is collected by threshold processing. In addition, it ca be said that the feature extraction of the anomalies could be sufficiently performed by the convolution process of the trained ResNet.

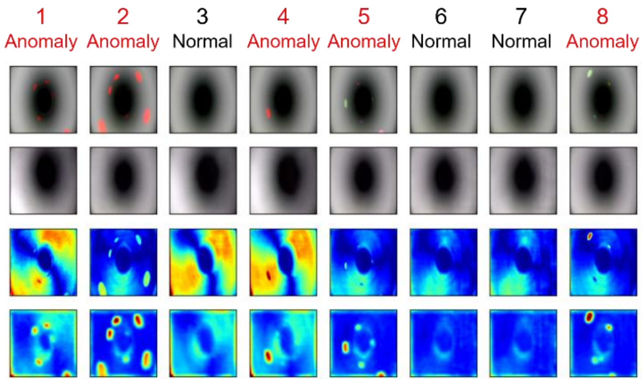

Fig. 10 Experimental results of VAE and ResNet (Dataset A)

Table 1 Anomaly score (Dataset A)

Number	1	2	3	4
score	654	269	760	796
Number	5	6	7	8
score	294	285	304	273

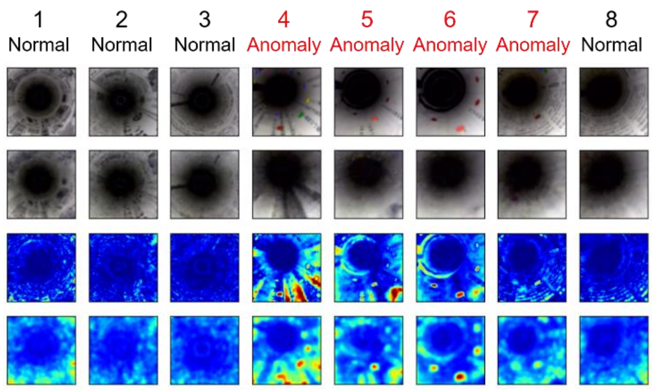

Fig. 11 Experimental results of VAE and ResNet (Dataset B)

Table 2 Anomaly score (Dataset B)

Number	1	2	3	4
score	172	145	118	435
Number	5	6	7	8
score	364	325	179	137

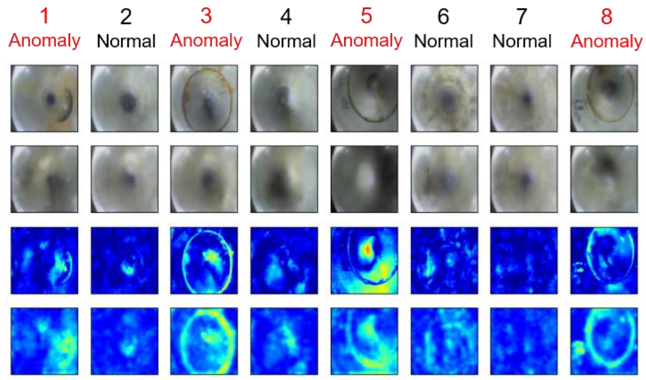

Fig. 12 Experimental results of VAE and ResNet (Dataset C)

Table 3Anomaly score (Dataset C)

Number	1	2	3	4
score	195	122	270	162
Number	5	6	7	8
score	418	173	135	238

V. Conclusion

In this paper, we proposed a method for detecting and estimating the location of rust adhering to the inside of a pipe, detecting the anomaly by taking the difference between the input and output images of VAE, and estimating the location of the anomaly using ResNet50. The results showed that the threshold treatment was sufficient to detect anomalies, and the defect location in the image could be emphasized. In the future, we would like to generate data that more closely resembles the rust of real pressure pipes and conduct experiments to improve the accuracy of the anomaly detection and the anomaly location estimation.

Acknowledgment

This research was partially funded by Sewerage Applied Research Project of the Ministry of Land, Infrastructure, Transport and Tourism.

References

[1] Yuki Mano et al.: “Development of High-speed Type Peristaltic Crawling Robot for Long-distance and Complex-line Sewer Pipe Inspection,” Proc. of IROS, 2018.
[2] Hiroto Sato et al.: “Proposal for Pipeline-Shape Measurement Method Based on Highly Accurate Pipeline Length Measurement by IMU Sensor Using Peristaltic Motion Characteristics,” Proc. of AIM, 2020.
[3] Yaxiang Fan, et al.: “Video Anomaly Detection and Localization via Gaussian Mixture Fully Convolutional Variational Autoencoder,” arXiv, 2018.
[4] Diederik P. Kingma et al.: “Auto-Encoding Variational Bayes,” Proc. of ICLR, 2014.
[5] Kaiming He et al.: “Deep Residual Learning for Image Recognition,” Proc. of CVPR, 2016.
[6] Adam Paszke et al.: “Pytorch: An Imperative Style, High-Performance Deep Learning Library,” Proc. of NeurIPS, pp.8024-8035, 2019.
[7] Diederik P. Kingma et al.: “Adam: A Metod for Stochastic Optimization,” arXiv, 2015.