ANOMALY DETECTION WITH AUTOENCODER FOR CONCRETE DEFORMATION EXTRACTION FROM IMAGES

Tetsu Yamaguchi, Hina Matoba, Masafumi Nakagawa
Shibaura Institute of Technology, 3-7-5, Toyosu, Koto-ku, Tokyo
Email: ah19012@shibaura-it.ac.jp

KEY WORDS: Image processing, Image classification, Anomaly detection, Concrete deformation extraction

ABSTRACT: The safe and secure maintenance and management of infrastructure require inspection. Infrastructure inspection covers a vast and extensive area and requires much time and labor. The infrastructure inspection works. Therefore, the efficiency of inspection works has been improved with various approaches, such as laser scanning and image sensing. In image sensing, detection and crack detection are the major topics in concrete structure inspection work. However, various technical issues remain such as parameter setting in image processing to extract all types of deformations in various environments. Moreover, a deep learning algorithm is expected as a robust approach to image-based inspection, but large-scale training data preparation remains a technical issue to be solved. In this study, we propose a method for detecting concrete deformations with an autoencoder using an anomaly detection approach that uses normal images of concrete surfaces, such as bridge abutments and revetments, to simplify training data preparation. The proposed method makes use of the structural similarity index measure to evaluate local structures and minute anomalies in images. The proposed method consists of a training step to generate an optimal model using normal data and an inference step to construct a model to evaluate test data in a feature space represented with the features of normal images. We evaluated the model using the confusion matrix and four indices, namely, the result of region segmentation of an anomalous image, anomaly scores, the receiver-operating characteristic curve, and the area under the curve. In our experiment, we confirmed that our proposed methodology can detect anomalous images with approximately 90% accuracy. The obtained results suggest that automated anomaly detection can be applied to the image-based screening process in infrastructure inspection.

1. INTRODUCTION

Infrastructure supports the safe and secure lives of the populations and serves as the foundation of social and economic activities. After construction, it requires appropriate maintenance and management to ensure proper function. In Japan, many civil engineering structures were constructed during the period of high economic growth between the 1950s and 1970s. In recent years, maintenance management for these buildings become more important, as the structures have aged. As of July 2014, there is a requirement for a periodical close-up visual inspection of bridges, tunnels, and other structures every five years. However, the structures to be inspected and surveyed are huge in size and scope, and such inspections require much time and effort. In addition, the huge volume of surveys makes it difficult to conduct the high-frequency inspections and surveys required for prevention. Furthermore, even though it is necessary for proper maintenance and management, the number of efficient repair workers is decreasing due to the shortage of human resources in the construction industry. Numerous attempts have been made to apply image measurement to detect the location of cracks and extract deformations in concrete because the use of image measurement allows for more effective location detection of cracks and to extract deformations. However, it is not easy to construct image processing algorithms and set parameters that can be suitable for various deformations in diverse measurement environments because images of the subject concrete may contain shadows or have different colors. compared with conventional machine learning, it is necessary to prepare a huge amount of normal and anomaly images as training data. This is supervised learning, where both normal and anomaly data are required. In this study, we propose a method for detecting concrete deformations using an anomaly detection approach that uses only normal images as training data.

2. METHODOLOGY

The proposed process (Figure 1) consists of a learning phase and an inference phase. In the learning phase, an anomaly detection model is constructed by learning from normal data features. In the inference phase, the model for anomaly detection constructed in the learning phase is used to detect anomalies by evaluating test data in a feature space that reflects the features of normal images. In this work, an autoencoder was applied to anomaly detection. When only normal images are used in the learning phase, the autoencoder process reconstructs the original input normal image, into an image like the input normal image and outputs it. However, it also outputs an image like the normal image

for the anomaly image. Using this characteristic, the problem of anomaly detection can be replaced with the problem of finding the location of the deformation.

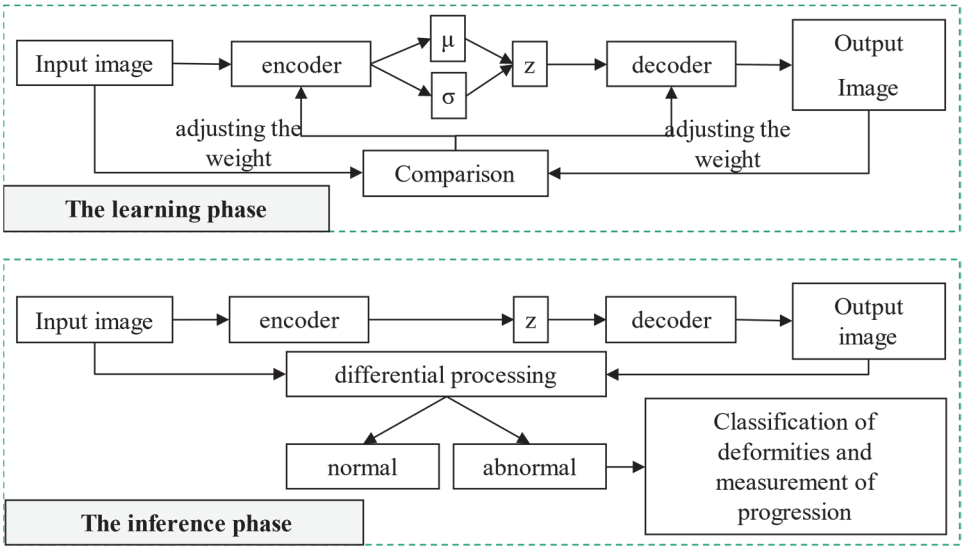

Figure 1. Flowchart of autoencoder processing

2.1 Autoencoder

An autoencoder (self-encoder) consists of an encoder and a decoder. The encoder encodes and compresses the input data and performs downsampling to convert it into a latent variable z, a variable that is subjective and cannot be objectively measured. Next, the decoder performs upsampling to restore the image converted to the latent variable z by the encoder to its original form. The amount of the difference between the output image and the input image data is visualized as a heat map. The autoencoder is a method of dimensionality reduction and has been used to reduce the amount of computation without losing data features. Autoencoders mainly use l2 methodology and are based on structural similarity index measure (SSIM) methodology and each has its own characteristics. The l2 methodology of autoencoders uses the squared error as the loss function. It calculates the difference by simply comparing pixels in two images and obtaining the difference image. Therefore, it does not consider the relationship between neighboring pixels. It causes blurring of the output image and may fail to detect anomalies in anomaly detection due to a large restoration error. By contrast, the SSIM methodology uses structural similarity. It measures the similarity between two images and calculates the similarity based on luminance, contrast, and structural information. It is characterized by its ability to correctly evaluate local structures in the image that cannot be evaluated by the squared error, making it easy to detect subtle anomalies.

2.2 The learning phase

In the learning phase, the encoder compresses the features of the input normal image and outputs the mean μ and variance σ of an n-dimensional Gaussian distribution. Next, the variable compressed latent feature z is calculated from the output mean μ and variance σ using the formula $z = \mu + \sigma$. The learning process is then repeated with normal data, adjusting the weight of each edge to reduce the restoration error between the input and output images.

2.3 The inference phase

In the inference phase, the model obtained in the training phase is first used to obtain a feature space that reflects the features of the normal image. On this feature space, an image subtraction process is performed between the output image and the image for evaluation, and the difference between the features of the two images is calculated as a score. If the score is less than a threshold value, the image is considered normal, and if the score is more than a threshold value, the image is considered an anomaly.

2.4 Model evaluation indicators

In this work, region segmentation of the anomalous image, anomaly scores, the receiver-operating characteristic (ROC) curve, the area under the curve (AUC), and the confusion matrix are used to evaluate the model. Anomaly region segmentation is performed to visualize the location of the detected anomaly, and the difference image is

obtained by calculating the difference between the input image and output image. Image differencing is performed between the output image and the image for evaluation on the feature space reflecting the features of the normal image, and the sum of the absolute values of the differences between the input and the output images is used as the anomaly score. To evaluate whether the machine learning model judges as anomaly what should be judged as an anomaly, on the ROC curve the true positive rate is pointed on the vertical axis and the false positive rate on the horizontal axis. The area under the curve is defined as the AUC. In general, when the AUC is over 0.9, the performance of machine learning is sufficient, and the model can be evaluated as highly accurate. This shows that there is little tendency to infer anomalies from those that should be inferred as normal, and there is a large tendency to infer anomalies from what should be inferred as an anomaly. The mixture matrix is used to evaluate the accuracy of the classification. The harmonic mean of the correctness rate, the fit rate, and the recall rate, is used for setting F1 score.

3. EXPERIMENTS

3.1 Image gathering of concrete surfaces

Concrete surfaces such as bridge abutments and revetments were targeted for photography, and images were taken on the ground (Urawa, Saitama-shintoshin, Monzennakacho, Toyosu, Minuma) and from a boat (Kanda River, Nihonbashi River). The photography photographs were taken from 20/10/2021 to 03/12/2021. A total of 1,393 images were collected. A hand-held digital camera (Cybershot DSC-HX60V, Sony) was used for image collection.

3.2 Creation of training datasets and model building

In generating the dataset for training, a 50 × 50-pixel region was first cut out from the collected images. Next, the images were visually classified into normal images for training (109,878 images), images for validation (30,772 images), normal images for evaluation (1,360 images), and anomaly images for evaluation (1,400 images) (Figure 2). The batch size was set to 32, and the number of training epochs was set to 100, 300, and 500.

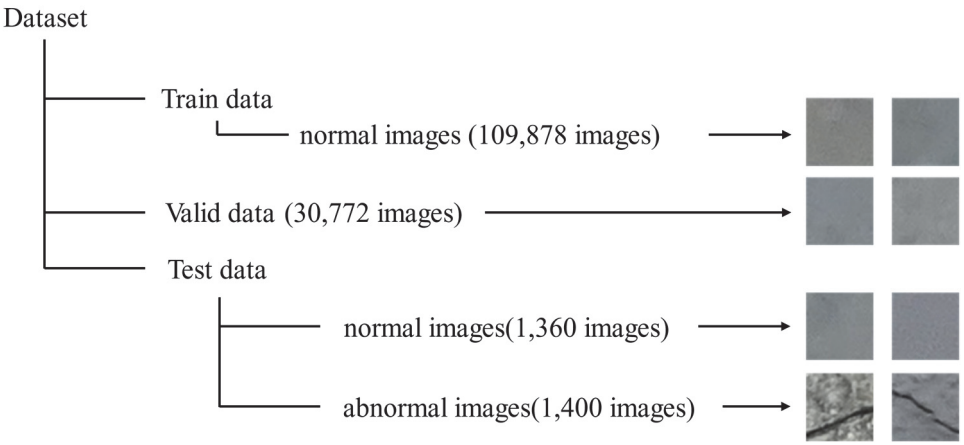

Figure 2. Overview of datasets

4. RESULTS

Three images were randomly selected for evaluation and segmented with the epoch numbers of 100, 300, and 500 to visualize the location of a detected anomaly (deformities). The results are shown in Figure 3. The upper row shows the input image, the middle row shows the output image, and the lower row shows the difference between the input and output images.

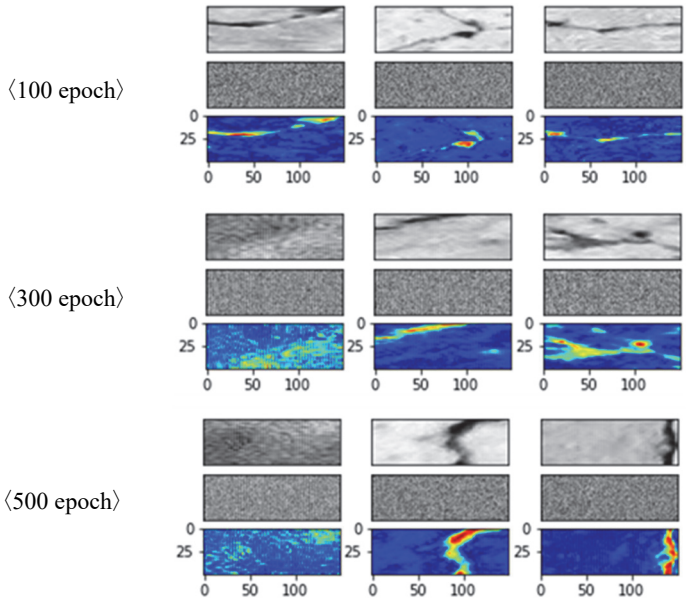

Figure 3. Segmentation of anomaly regions
(Upper row: input image, middle row: output image, lower row: segmentation result)

The numerical values of the differences were calculated, and anomaly scores were calculated for epochs 100, 300, and 500, and the results are presented in a histogram in Figure 4.

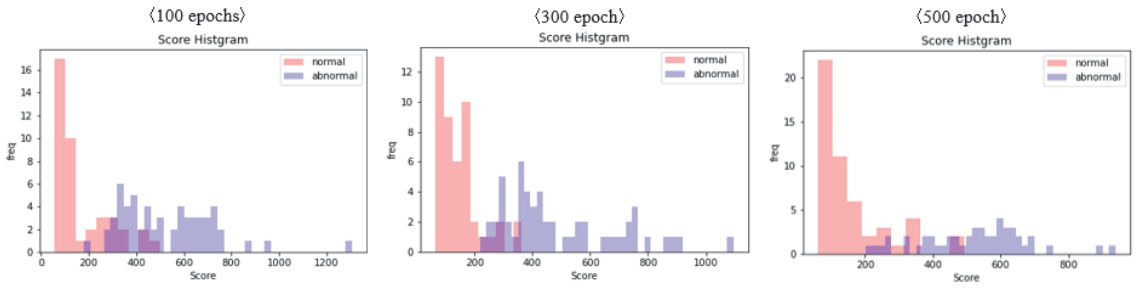

Figure 4. Histogram of scores for each epoch
(Normal: when the test image is "normal". Anomaly: when the test image is " abnormal ".)

The ROC curve obtained in this study is shown in Figure 5. The AUC was 0.93.

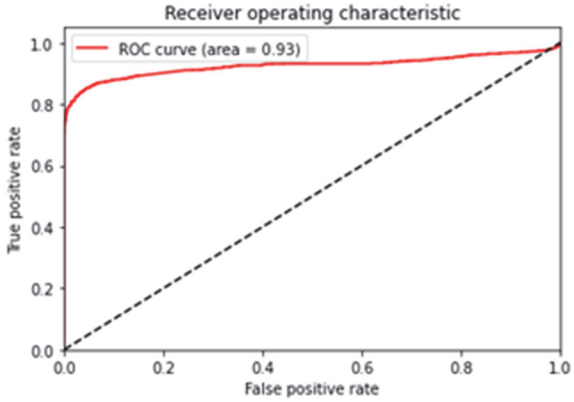

Figure5. ROC curve

Table 1 shows the confusion matrix resulting from the processing of the 300 epochs with the best accuracy. From these results, the performance of the model was estimated as 90.1% correct, 96.6% good fit, 83.4% repeatability, and the F1 score was 89.5%. The correctness rate was the percentage of correct predictions out of the total predictions, the fit rate was the percentage of the predicted anomaly to the actual anomaly, the reproducibility rate was the percentage of predicted anomaly out of the actual anomaly, and the F1 score was the harmonic average of the fit and reproducibility rates.

Table 1 Confusion matrix (300 epochs)

		Predictted class	
		Positive	Negative
Correct class	Positive	True Positive 1,168	False Negative (omission) 232
	Negative	False Positive (error detection) 41	True Negative 1,319

5. DISCUSSION

From the verification experiment, automated anomaly detection can be applied to image deciphering in the screening process when the processing parameters are set to avoid overlooking all deformities. The segmentation of anomaly areas s estimated clear results when input image includes clear cracks (Figure 6, left). However, when the input image included unclear and wide deformities, such as peeling (Figure 6, right), the results were unclear. The failure reason was that the decoder was not able sufficiently to restore the normal image as a normal image, and the difference between the input image and the output image was unclear.

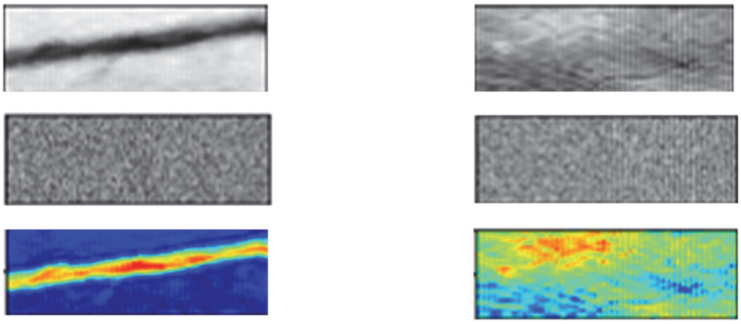

Figure6. Segmentation of cracks and delamination

Furthermore, a comparison of the different results shows a difference in segmentation accuracy depending on epoch number. When the number of epochs was 100, some cracks ware not completely segmented and were considered normal (Figure 3, upper row). By contrast, when the number of epochs was 300, the cracks were generally segmented compared with the case when the number of epochs was 100 (middle row of Figure 3). When the number of epochs was 500, the normal image could be correctly identified as such because the area showing the normal state was wider (lower row of Figure 3). Therefore, there is a possibility that the difference between the image and the input images could be made clearer by increasing the number of epochs in the segmentation of anomaly regions. The following is a comparison of the difference in anomaly scores depending on the number of epochs. When the number of epochs was 100, some images had a score of over 400 despite being normal, and some images had a score of less than 200 despite being abnormal (Figure 4, upper row). When the number of epochs was 300, the scores of normal and abnormal images were mixed, but the score of normal images was below 400 and the score of abnormal images was above 200 (Figure 4, middle row). When the number of epochs was 500, the scores of normal and abnormal images were mixed, such that some images had a score of 400 or more despite being normal (Figure 4, lower row). Therefore, the most accurate score comparison was performed when the number of epochs was 300. In the calculation of the anomaly score, the scores of normal and abnormal images were mixed for all epochs. This is because the training

data included abnormal images that resembled normal ones because deformities, including cracks, were not clearly captured, and normal images that appeared to have captured deformities, when images were collected under various shooting conditions including shooting positions and shooting environments to avoid overlearning. This is thought to be because the decoder was unable to determine which images were normal and which were abnormal. If the decoder cannot sufficiently recover a normal image as a such, more accurate anomaly detection can be achieved by applying GAN or AnoGAN, which have higher image recovery performance than autoencoders.

6. CONCLUSION

In this study, to simplify the training data maintenance, we proposed a method to apply an anomaly detection approach to concrete deformation detection, in which only normal images of concrete surfaces such as bridge abutments and revetments are used for training autoencoders. In a validation experiment, we evaluated the model using four indices: anomaly region segmentation, anomaly score calculation, ROC curve, and AUC. The evaluation confirmed that the proposed method can detect anomalies with an accuracy of approximately 90%. The reason for the accuracy not being 100% was that the images were collected under various shooting conditions, such as shooting positions and shooting environments, to avoid overearning, and among the collected images, there were images in which deformities, including cracks, were not clearly captured and normal images in which deformities appeared to have been captured. This is the case. More accurate anomaly detection may be achieved by applying GAN and AnoGAN, which have higher image restoration performance than autoencoders.

ACKNOWLEDGMENTS

This research was supported by the MEXT Coordination Funds for Promoting AeroSpace Utilization; Grant Number JPJ21460038.

REFERNCES

Paul Bergmann, Kilian Batzner, Michael Fauser, David Sattlegger, Carsten Steger, 2021. The MVTec Anomaly Detection Dataset: A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection, International Journal of Computer Vision. Pages 1038-1059.

Paul Bergmann, Sindy Löwe, Michael Fauser, David Sattlegger, Carsten Steger,2019. Improving Unsupervised Defect Segmentation by Applying Structural Similarity to Autoencoders, arXiv:1807.02011,2019.

Kosuke AOSHIMA, Satoshi NAKANO, Kohei TOKUNAGA, Hideaki NAKAMURA, 2019. Damage detection of concrete surface usinganomaly detection method by deep learning, Journal of Japan Society of Civil Engineers, Volume 75 Issue 2 Pages I_559-I_570

Yasutoshi NOMURA, Saki MURAO, Yukihiro SAKAGUCHI, Hitoshi FURUTA, 2017. Crack detection system for concrete surface based on deep convolution neural network, Journal of Japan Society of Civil Engineers, Volume 73 Issue 2 Pages I_189-I_198

Yasutoshi NOMURA, Kazuki SHIGEMURA, 2019. Development of Real-Time Screening System for Structural Surface Damage Using Object Detection and Generative Model Based on Deep Learning, Journal of Japan Society of Civil Engineers, Volume 68 Issue 3 Pages 250-257.

Schlegl, T., Seebock, P., Waldstein, S. M., Schmidt-Erfurth, U. and Langs, G, 2017. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery, International Conference on Information Processing in Medical Imaging (IPMI), pp. 146-157.

Ministry of Land, Infrastructure, Transport and Tourism, WHITE PAPER ON LAND, INFRASTRUCTURE, TRANSPORT AND TOURISM IN JAPAN, 2021. p45.

Ministry of Land, Infrastructure, Transport and Tourism, WHITE PAPER ON LAND, INFRASTRUCTURE, TRANSPORT AND TOURISM IN JAPAN, 2020. p114 p115.