Development and site application of an AI-based concrete compaction management system

Masatoshi Uno[*1], Jin Chujo[*2], Ryuichi Imai[*3]

1. Introduction

The environment surrounding Japan's construction industry is changing dramatically due to advances in IT. The main change is due to the advancement of artificial intelligence technology such as deep learning. On the other hand, the process of concrete construction has remained almost unchanged for over 30 years.

The pouring of concrete often depends on the know-how of the workers, and the concrete filling in the formwork is generally good, but in rare cases, problems, such as bean shingles, may occur. Once a problem occurs, countermeasures must be considered in consultation with the client, and repair and reinforcement work, for example, may be required. In the worst case, the structure will have to be demolished and rebuilt, which will require a large amount of time and cost. In this context, when focusing on the compaction work in concrete construction, the cause of problems is often insufficient compaction of the concrete. Therefore, the solution to the problem is to recognize the compaction points in on-site work in real time, detect the insufficiently compacted points before the concrete sets, and re-compact the relevant points.

In a previous study, Sunaga et al.[1)] compared the compaction actions of experienced and junior technicians using a wearable camera and found that the experienced technicians were more effective at compacting. However, the insertion position, time, and depth of the vibrator inserted by the experienced technicians were not measured. Therefore, the authors focused on compaction in concrete work and developed technology[2)] that enables reliable compaction management by quantitatively recognizing the compaction points, shifting from conventional qualitative management that relied on the individual know-how of workers based on their experience to quantitative management. The artificial intelligence (AI) model of the video images uses a convolutional neural network (CNN), a type of deep learning, to set the optimal number of repetitions, achieving high-precision and high-speed analysis.

This report introduces the developed technology and describes the results of its application at an actual site.

2. Overview of compaction management technology

2.1 Overview of compaction position identification

As shown in Figure 1, the method for identifying the compaction position is fundamentally based on one devised by Imai et al.[3),4)], with improvements and new functions added to ensure real-time display of the analysis results. Specifically, the site conditions are photographed with a wearable camera attached to the helmet of the compaction worker (Figure 2). The video images are analyzed using an AI model to detect the AR marker, the tip of the vibrator, and the coloring of the vibrator hose, and the compaction position is identified from the relative position to the AR marker, and the depth is identified from the coloring of the vibrator hose (Figures 3 and 4).

＊1 Masatoshi Uno／Shimizu Corporation
＊2 Jin Chujo／Create-C Corporation
＊3 Ryuichi Imai／HOSEI University

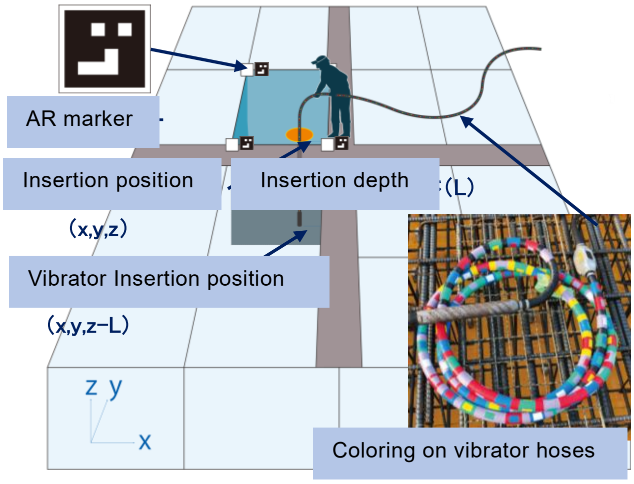

Figure 1　System overview

Figure 2　Installation of the wearable camera

Figure 3　Positioning of AR markers showing planar coordinates

Figure4　Vibrator hose coloring

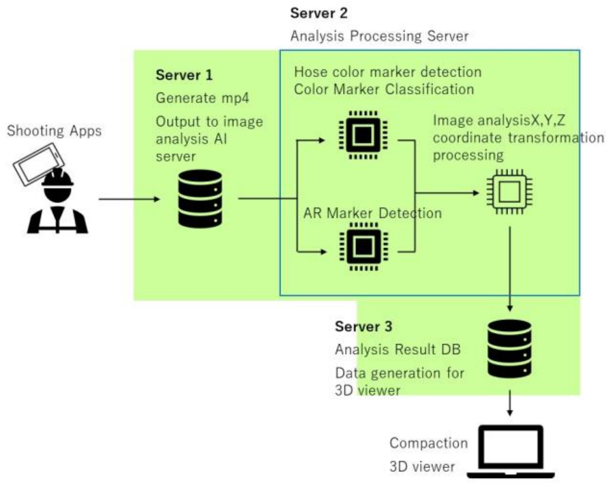

Figure 5　System configuration

2.2 System Overview

One of the goals of this system is to realize real-time processing from shooting to analysis and display the results, so a cloud-based system was constructed (Figure 5). The shaded area is configured within the cloud server. The video images captured by the wearable camera are uploaded and saved to Server 1, which triggers the subsequent automatic processing. The data pass through the analysis processing server on Server 2, and the automatic processing ends in the database that stores the analysis results on Server 3. Cloud computing makes it easier to upgrade the machine specifications of the analysis processing server, and can speed up processing. In addition, the analysis results can be checked anywhere, so there is also the advantage that they can be checked on-site, in the office, or at a remote location. In a system that visualizes the calculated compaction position through video image analysis using artificial intelligence (AI), the compaction position and depth caused by one insertion of the vibrator are expressed as one cylinder (Figure 6). The size of the cylinder is considered to be the range of influence of appropriate compaction, and specifically, it is set to about 10 times the vibrator diameter, which is the manufacturer's recommended value[5]. If it is possible to analyze all the insertion points made by multiple workers, the formwork will be filled with cylinders one by one as the cylinders are filled in. On the other hand, if compaction is missed, there will be spaces that are not filled in by the cylinders.

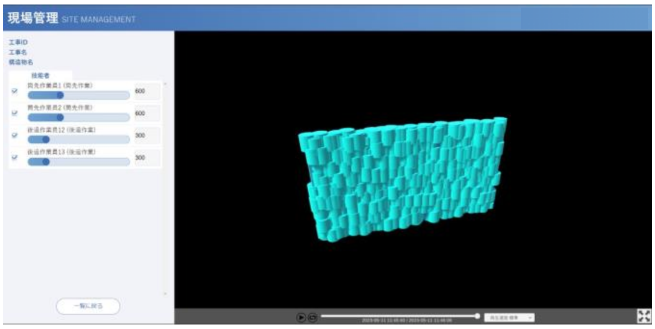

Figure 6 Compaction position visualization 3D system

2.3 Algorithm for determining compaction position using AI

(1) Overall data processing flow

Figure 7 shows the processing flow from the step of acquiring and transmitting video images using a camera attached to the helmet of the compaction worker to the step of outputting the analysis results. In the processing flow, step (1) is performed by the shooting app, and steps (2) to (12) are performed by the data analysis server on the cloud. The video image transmission in step (1) is used as a trigger to automate all subsequent processes.

(2) Application of artificial intelligence model

One of the artificial intelligence models for image detection is the convolutional neural network (CNN), a type of deep learning[6)-13)]. CNN models the receptive field in the visual cortex of the brain and is known to have high performance in the field of image recognition. Among deep learning, CNN has a structure with a network that has a convolutional layer, a pooling layer, and a fully connected layer (Figure 8).

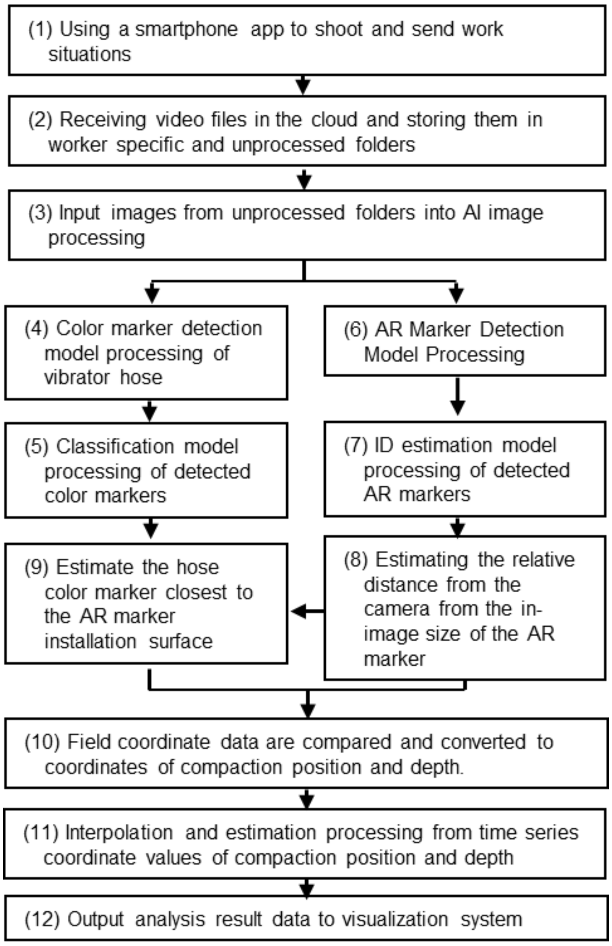

Figure 7 Processing flow for compaction location analysis

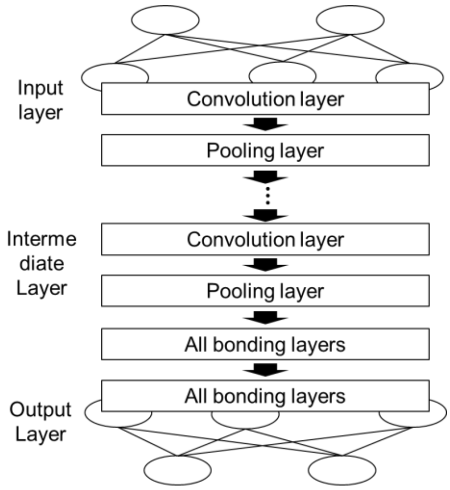

Figure 8 Example of a CNN network structure

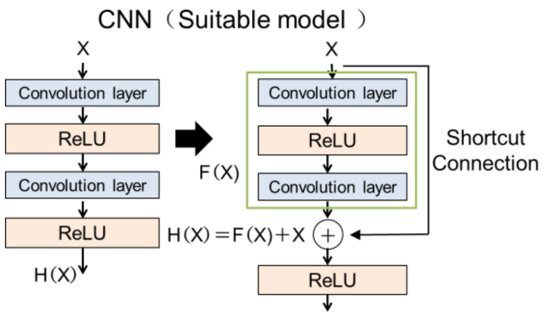

Figure 9 Portions of a typical CNN network and ResNet

In recent years, it has been thought that performance can be improved by making the layers deeper in CNN research, but it has been reported that simply making the layers deeper can actually worsen performance14). The degradation problem is a phenomenon in which the improvement of training error in learning a model with deep layers plateaus at an earlier stage than in models with shallow layers. Focusing on this problem, ResNet (Residual Network) has been devised as a network architecture that can learn even in deep layers.

The difference between ResNet and normal CNN is that it learns a residual function that references the input of the layer. Figure 9 shows a general network and a part of ResNet. Consider the case where the function to be learned is H(x). In ResNet, in two consecutive convolutional layers, the input x is connected by skipping to the output of the next layer. At this time, the difference F(x) with respect to the input x is expressed as Equation (1), which is transformed and redefined to learn Equation (2).

$$F(x) = H(x) - x \qquad (1)$$

$$H(x) = F(x) + x \qquad (2)$$

This replaces the problem of estimating the optimal function H with the problem of estimating the optimal residual function F. Even if the identity mapping H(z) = z is optimal, it can be easily approximated by setting F(z) = 0. The shortcut connection acts as a detour to add the input value of the layer to the output of the network before the activation function. Since the shortcut connection passes the input information as is, it passes the gradient as is during backpropagation. Therefore, there is no need to worry about the gradient becoming too small or too large, and a significant gradient is maintained.

A block consisting of such a shortcut connection and several convolutional layers is called a residual block, and a network consisting of multiple stacked shortcut connections, called ResNet, is a network model that adds an input layer and an output layer to this. Five types of ResNet have been proposed: ResNet 18, ResNet 34, ResNet 50, ResNet 101, and ResNet 152, which differ in the number of layers and the number of learnable parameters[6].

In this system, the ResNet50 model and ResNet 34 are used to detect vibrators. RetinaNet is used to detect AR markers. RetinaNet is an object detection model proposed in the paper "Focal Loss for Dense Object Detection" published by Facebook AI Research (FAIR) in August 2017. As pointed out in the motivation for developing the paper, many of the highly accurate object detection models prior to RetinaNet were configured as an R-CNN-based twostage object detector, but RetinaNet was improved to achieve faster speeds.

(3) Construction of artificial intelligence models

In the processing flow of Figure 7, artificial intelligence models were constructed for three cases: (4) Vibrator hose color marker detection model processing, (5) Classification model processing of detected color markers, and (6) AR marker detection model processing. In selecting the AI model, we considered models with as short an inference time as possible, since this system needs to output the analysis results before the concrete sets, while ensuring analysis accuracy. For vibrator color marker detection, the ResNet 50 model was adopted, although it is difficult and requires inference time because the vibrator is detected from a wide variety of objects that appear in the video. For color marker classification, ResNet 34, which has a fast inference time, was adopted because the video image after the vibrator is detected by the aforementioned model is the subject of analysis. For AR marker detection, we adopted the RetinaNet model, which is characterized by high-speed processing, because it was confirmed at the consideration stage that detection accuracy could be ensured.

For (4) Vibrator hose color marker detection, the ResNet 50 model is used to return a bounding box for the location of the color marker. The color marker classification model (5) determines the three colors of the outer color, inner color, and boundary color for the image inside the bounding box extracted in (4) above. This model used a custom classification model of ResNet 34. For AR marker detection (6), a RetinaNet model was used to return a bounding box for the location of the AR marker.

To build these artificial intelligence models, training data were acquired, and the analysis results were verified at an actual concrete construction site on the dates shown in Table 1 for concrete pouring days from June to December 2020. Note that all work days were for actual construction work, and the system was characterized by the fact that various training data were acquired at the actual construction site and models were built.

The number of training data values for the artificial intelligence model is shown in Table 2. Learning was performed using a total of more than 20,000 training data images. The reason why the color marker classification model has a large amount of training data is that there were 12 types of color markers and 6 intermediate colors, and training was performed to make them classifiable.

Table 1 Teacher Data Acquisition and Data Validation Dates

Date	Location data	Content
6/4	3SP Bottom plate	teacher data image
6/18	3SP Outside wall, Inside wall	teacher data image
6/25	1SP Exterior wall	teacher data image
7/2	4SP Interior wall	teacher data image
7/9	2SP interior wall	teacher data image
8/4	Top plate	teacher data image
9/18	1SP column model	validation
10/8	Color top model	validation
11/9	Back wing wall	Real-time verification
12/2	Wing wall bottom plate	Real-time verification
12/10	Wing wall bottom plate	Real-time verification
12/18	Front wing wall	Real-time verification

Table 2 Number of images of teacher data

	Model Type	Number of teacher data	Number of evaluation data values
Color marker detection of vibrator hose	ResNet50	1,025	264
Color marker classification of vibrator hose	ResNet34	19,348	4,849
AR Marker Detection	RetinaNet	798	198
Total		21,171	5,311

Table 3 Determination accuracy of hose color marker detection and AR marker detection models

	Model Type	Fit ratio (%)	Reproducibility (%)	F-Value
Color marker detection of vibrator hose	ResNet50	83	96	0.89
AR Marker Detection	Retina-Net	86	70	0.77

Table 4 Judgment accuracy of color marker discrimination model

	Model type	Outside color	Inside color	Fit ratio (%)	Reprodu-cibility (%)	F-Value
Classifi-cation of color markers	ResNet34	Red	Yellow	98	98	0.98
		Purple	Red	97	96	0.96
		Green	Purple	98	95	0.96
		Blue	White	98	97	0.97
		Red	Blue	95	96	0.95
		Purple	Green	93	96	0.94
		Green	Yellow	97	99	0.98
		Blue	Red	99	96	0.97
		Red	Purple	94	95	0.94
		Purple	White	96	93	0.94
		Green	Blue	94	92	0.93
		Blue	Green	94	97	0.95

(4) Accuracy of the AI model

The accuracy of the AI model developed in this study is shown in Tables 3 and 4. The accuracy and recall of the vibrator color marker detection was high, and it can be said to have high evaluation performance.

3. Application to a new site

3.1 Construction site in an area where mobile phone service is unavailable

This time, the compaction management system was applied to a part of the "Asuwa River Dam Replacement Prefectural Road No. 4 Bridge and Other Works," a prefectural road replacement project associated with the Asuwa River Dam construction work, which is a project under the jurisdiction of the Kinki Regional Development Bureau (Figure.

10). This bridge is approximately 44 m long, 6 m wide, and 3 m deep, and is located in an area where mobile phone service is unavailable, including the surrounding area. It was confirmed that the system worked reliably on the wall side of the bridge (Figure. 11).

Figure 10　Location of target construction

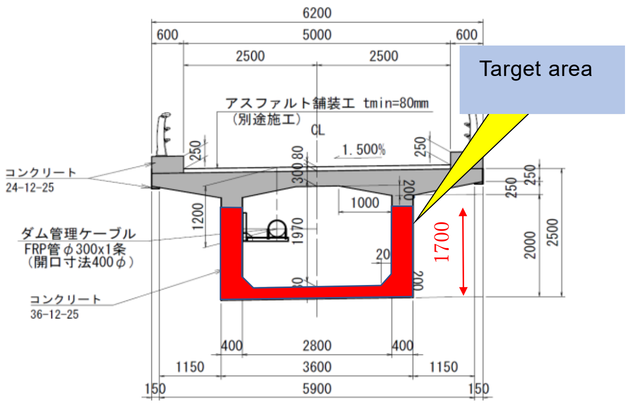

Figure 11　System application location

3.2 Ingenuity when applying the system

(1) Use of satellite communication equipment

Since the target construction work was in an area where mobile phone service is unavailable, it was decided to use satellite communication equipment to apply the system. The satellite communication equipment used was "Starlink," developed by SpaceX, which is provided by several thousand low-earth orbit satellites and achieves data communication with significantly higher speeds and lower latency than conventional satellite communication services.

(2) Communication performance of satellite communication equipment:

When the equipment was installed at the actual site and the data communication status was confirmed, a good communication environment was ensured, with download speeds of 30.4 to 50.2 Mbps and upload speeds of 6.25 to 12.4 Mbps (Figure 12). This system uploads video images taken with a smartphone to a server in 1-minute intervals (Figure 13). It was confirmed that the images were saved to the server with a time lag of about 1.2 minutes, and that there was almost no time lag due to the communication environment.

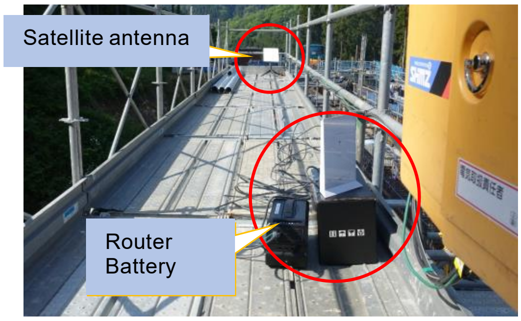

Figure 12　Equipment installation status at an actual site

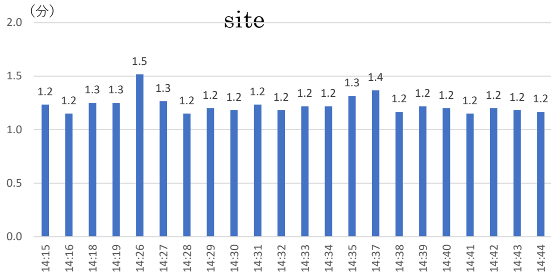

Figure 13　Time from data acquisition to server storage

3.3 System application results

As a result of the demonstration, this system displayed the results of compaction visualization in about 3 minutes and made it possible to recognize the compacted areas with an error of less than 10 cm. Therefore, it was confirmed that reliable compaction can be performed on structures (Figure 14).

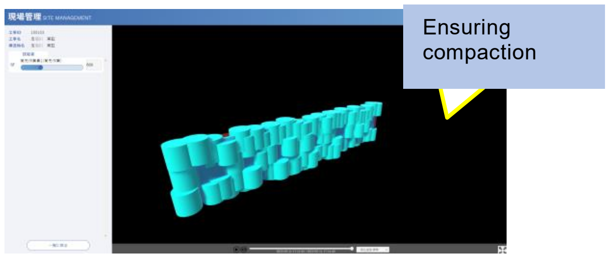

Figure 14　Compaction visualization results

4. Technical improvements for on-site application

4.1 Building a website to check the shooting conditions

In order to check and determine the appropriateness of the shooting conditions using the wearable camera, we built a website where the captured video can be viewed (Figure 15). On this site, when you select the construction location, date, and worker ID, a list of videos is displayed by time, and you can play and view the video by clicking the time you want to check. The video is the data captured by the wearable camera before it is sent to the cloud and processed by AI, so it can be viewed 1 to 5 minutes after shooting (depending on the communication environment). Furthermore, since it is created in a web environment, the wearable camera video for each worker can be viewed on a tablet at the casting site, and the worker can check the current shooting conditions, and if an abnormality in the angle of view occurs, the camera angle can be corrected, etc.

Figure 15　Worker video viewer

4.2 Improvement of the visualization system

(1) Improvement of the registration method for various structures

In order to be able to handle not only general rectangular parallelepiped structures but also structures of various shapes, a mechanism was introduced that automatically generates a 3D model by importing consecutive coordinate points. Table 5 shows an example of input coordinate points when the structure is a cylinder. When the data are read, a structure model is automatically generated (Figure 16). Even if the shape is not a cylinder, a 3D model can be easily generated by creating the data of the outer coordinates.

(2) Improvement of the method for checking areas where compaction is insufficient

In this system, AR markers for analyzing the compaction position are registered at the same time as the 3D model is registered (Table 6). By making the AR marker number appear on the 3D model at the time of registration, it is possible to immediately determine where the compaction is insufficient during construction. As a result, we believe that reliable compaction can be achieved.

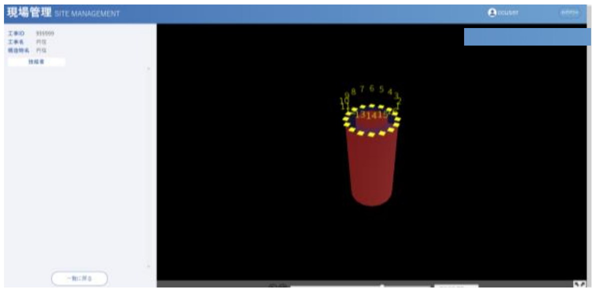

Figure 16　Example of cylindrical structure 3D model and AR marker display

Table 5　Example of coordinate point input values

連番	X	Y	Z
1	1000	500	2000
2	962	309	2000
3	854	146	2000
4	691	38	2000
5	500	0	2000
6	309	38	2000
7	146	146	2000
8	38	309	2000
9	0	500	2000
10	38	691	2000
⋮	⋮	⋮	⋮

Table 6　AR marker registration

マーカーID	マーカー名	X	Y
1	1	980	500
2	2	943	684
3	3	839	839
4	4	684	943
5	5	500	980
6	6	316	943
7	7	161	839
8	8	57	684
9	9	20	500
10	10	57	316

5. Conclusion

We developed a system that quantitatively recognizes the compaction points in concrete work and visualizes the results in real time, and applied it to various sites. For example, for the quality control of concrete, it is extremely important to put into practical use real-time analysis of the concrete compaction position, and the authors have been working on its development. However, in order to put it into practical use, it is necessary to solve the following issues: 1) communication environment issues in areas where mobile phones are not available, and 2) visualization system issues. We devised solutions to these issues and demonstrated them on site, and found that they are practical.

The main feature of this technology is the use of CNN, a type of deep learning artificial intelligence model (AI), from video images of the compaction work. At this stage, we believe that this is the optimal method for obtaining accurate information on the compaction position before the concrete sets. In the future, if more optimal analysis methods become available due to advances in AI technology, we believe that it will be necessary to immediately proceed with application and improve the system. Approaching real time is the need in the field, and it is necessary to be able to take measures such as reliable re-compaction. This technology won the Excellent Technology Award at the 2022 Infrastructure DX Competition hosted by the Kinki Regional Development Bureau of the Ministry of Land, Infrastructure, Transport and Tourism15). As a result, it was demonstrated and evaluated at the Asuwa River Dam Replacement Prefectural Road No. 4 Bridge and other construction sites. We believe this shows the great expectations for this technology. In the future, we aim to apply and deploy this technology to a variety of sites, not only to confirm the validity and effectiveness of this technology, but also to improve its operability and feasibility. In addition, our immediate goal is to add a function to automatically determine areas where compaction is insufficient. If we can automatically determine areas where compaction is insufficient, we can also consider building a function to notify of areas that require re-compaction and a function to guide workers. Based on this, it will be possible to use it for education and guidance for inexperienced workers.

References

1) Shinsuke Sunaga, Yoichiro Tsujita, Kentaro Futamura, Junna Murakami and Takeshi Iyoda: Comparative experiments on concrete compaction technology by skilled and young engineers using a wearable camera, the Society of Civil Engineers Annual Lecture Summary, Vol. 70, pp. 765 -766, 2015

2) Masatoshi Uno, Jin Chujo, Ryuichi Imai: Development of a real-time analysis system for concrete compaction position and a construction site trial, AI and Data Science Papers, Vol. 2 J2, PP. 757-764, 2021

3) Ryuichi Imai, Tetsuhiko Kurihara, Tosun Taniguchi, Makoto Ito, Takuya Yokota: Development of a program for measuring the insertion position of vibrators using moving images of wearable cameras, Journal of the Japan Society of Civil Engineers, F 3 (Civil Informatics), Vol. 74, No. 2, pp. I _ 102 -I _ 112, 2018.

4) Ryuichi Imai, Tetsuhiko Kurihara and Tetsuya Yokota: Development of a method for measuring the insertion depth of vibrators in concrete placing, Journal of the Japan Society of Civil Engineers F3 (Civil Informatics), Vol. 75, No. 2, pp. I _ 12 -I _ 21, 2019.

5) Exen Corporation: General Catalog of Construction Machinery (2019 to 2020 edition), p. 7.

6) Yukino Tsuzuki, Pang-jo Chun, Tatsuro Yamane: Automatic detection of cracks in asphalt pavement using weak type improved deep learning, AI and Data Science Theses, Vol. 1, J No. 1, pp. 168 -179, 2020.

7) Chun, P.J., Izumi, S., and Yamane, T.: Automatic detection method of cracks from concrete surface imagery using two - step light gradient boosting machine, Computer-Aided Civil and Infrastructure Engineering, 2020.

8) Chun, P.J., Izumi, S., Yamane, T., and Kuramoto, N.: Development of a Machine Learning-based Damage Identification Method using Multi-point Simultaneous Acceleration Measurement Results, Sensors 20(10) 2780, 2020.

9) Chun, P. J., Yamane, T., Izumi, S., and Kameda, T.: Evaluation of tensile performance of steel members by analysis of corroded steel surface using deep learning, Metals, 9(12), 1259. 2019.

10) Chun, P. J., Ujike, K., Mishima, K., Kusumoto, M., and Okazaki, S.: Random Forest-based evaluation technique for internal damage in reinforced concrete featuring multiple nondestructive testing results, Construction & Building Materials, 253(30) 119238, 2020.

11) Pang-jo Chun, Atsushi Igo: Detection of Concrete Surface Cracks by Random Forest, Journal of the Japanese Society

of Civil Engineers F3 Vol. 71 No. 2, pp. I _ 1-I _ 8, 2015.
12) Chun, P., Shimamoto, Y., Okubo, K., Miwa, C., Ohga, M.: Deep learning and Random Forest based crack detection from an image of concrete surface, Journal of Japan Society of Civil Engineers, Ser F3 (Civil Engineering Informatics), Vol.73, No.2, pp.I_297-I_307, 2017.
13) Lecun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard W., and Jackle L.D.: Backpropagation applied to handwritten zip code recognition, Neural Computation, Vol.1, No.4, pp.541-551, 1989.
14) He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.770-778, 2016.
15) Ministry of Land, Infrastructure, Transport and Tourism Kinki Regional Development Bureau: About the 2022 Infrastructure DX Competition (Excellent Technology Award and Construction Site Trial Implementation Status) , <https://www.kkr.mlit.go.jp/kingi/infradx-center/dx/urkb1n00000005u5-att/230710_DXCOMPE_1.pdf>, 2023