Development of a Real-Time Analysis System for Concrete Compaction Positioning and a Construction Site Trial

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

[1] Shimizu Corporation, Tokyo, Japan;
[2] Create-C Corporation, Tokyo, Japan; Email: j-chujo@create-c.co.jp (J.C.)
[3] HOSEI University, Tokyo, Japan; Email: imai@hosei.ac.jp (R.I.)
*Corresponding author: uno@shimz.co.jp (M.U.)

Abstract—The environment surrounding the construction industry in Japan is drastically changing with advances in IT (Information Technology). The central reason for this change is the development of AI (Artificial Intelligence) technologies such as deep learning. On the other hand, in concrete works, which are often used in the construction industry, the system of carrying fresh concrete produced at a concrete plant to a construction site using an agitator car, placing concrete in a form, compacting with a vibrator, finishing the surface, and curing the finished concrete surface has hardly changed in more than 30 years. It was confirmed in this study that this system can be used to carry out more effective compaction management by shifting from qualitative management using individual know-how based on the experience of conventional workers to quantitative management based on AI analysis using images.

Keywords—concrete site, compaction work, image processing, deep learning, convolutional neural network

I. INTRODUCTION

The environment surrounding the construction industry in Japan is changing dramatically with the advancement of IT (Information Technology). The central reason for this change is the development of AI (Artificial Intelligence) technologies such as deep learning.

On the other hand, the flow of concrete work often used in the construction industry is as follows: (1) manufacturing in a concrete plant, (2) bringing fresh concrete to a site using an agitator vehicle, (3) putting concrete into a form, (4) compaction using a vibrator, (5) finishing the surface, and (6) curing the concrete surface, and this system has hardly changed in more than 30 years.

The quality control of concrete works is often based on the know-how and experience of the workers. Once trouble occurs, countermeasures are discussed considering the age and cost and whether reinforcement works are required. In the worst case, if the concrete breaks, it may be necessary to rebuild, requiring a lot of extra time and cost. Focusing on compaction work in concrete construction, the cause of trouble is often attributed to the insufficient compaction of concrete. Therefore, by understanding the compaction point at a construction site in real time and detecting insufficient compaction points before the concrete hardens, re-compaction can lead to the solution of this problem.

In a previous study, Sumaga *et al.* [1] compared the compaction works of skilled and junior engineers using a wearable camera and found that skilled engineers compacted more effectively, but the position, time, and depth of insertion of the vibrator by skilled engineers were not measured. In Imai *et al.* [2, 3], we devised a measurement method using image analysis and processing that takes pictures of the insertion position, time, and depth of the vibrator with a wearable camera and also using AI. In addition, measurement experiments have been carried out on simulated sites to confirm their usefulness, but methods for obtaining analysis results in real time have not been discussed.

In this study, focusing on the vibration position of vibrators in concrete compaction, a system was developed to measure precisely in three dimensions where compaction was completed during construction, and to quantitatively and real-time detect where and when compaction was done in comparison with the construction time. Data acquisition was performed using moving images taken by a camera mounted on a worker's helmet, and the moving images were transferred to a cloud server, and the processing of the moving images using AI was carried out and fed back, so that it could be confirmed at the construction site.

As for the concrete compaction position, the problem was that the vibrator was buried in the concrete, and it was not directly visible during construction, and it was not quantitatively detected. To detect the compaction position, it is necessary to measure the planar 'position' and insertion 'depth'.

Therefore, in this study, in order to determine the planar "position", an AR marker was placed on-site and it was determined from the moving images. At the actual site, it is common that concrete is scattered, and the markers are dirty. Therefore, by utilizing the inference of

the AI model, it is possible to detect markers even if they are dirty.

In order to detect the "depth" of insertion, the vibrator was colored, the color was discriminated from the moving image, and the insertion depth was determined.

With this system, it is possible to shift from qualitative management using know-how based on the experience of conventional workers to quantitative management based on AI analysis using moving images, which enables reliable compaction management and eliminates problems caused by construction.

II. METHOD OF DETERMINING COMPACTION POSITION

In this study, we basically adopted the method devised by Imai *et al.* [2, 3] for determining the insertion position and depth (= compaction position) of a vibrator, and improved the following items and constructed new functions to ensure the real-time display of the analysis results.

A. Wearable Cameras for Workers

In the previous study [2], a GoPro Hero 6 was adopted as the camera mounted on the compaction worker's helmet, but in this study, we decided to use the built-in camera of a smartphone, which can transmit the moving images taken in real time to the cloud (Fig. 1). The smartphone used was a SHARP AQUOS R2 compact and AQUOS R5G, and an app has been created to capture work situations running on the smartphone. The shooting application was designed to start shooting after entering the number of the construction site and a code that identifies the worker, and the shooting image quality could be set to 5 options (SD, HD, 2K, 4K, 8K) and the frame rate was set to 10 FPS. The videos captured were sent to the cloud once a minute.

B. Positioning AR Markers to Indicate Plane Position

In order to understand the position of the vibrator, AR markers were attached to the concrete placing forms and scaffolds (Fig. 2). The AR markers were printed on 50 mm square paper and laminated and placed on-site at 500 mm intervals. The ID values of the individual AR markers and their on-site coordinate values were registered in the analysis system described later.

Figure 1.   Installation of the wearable camera

Figure 2.   Positioning of AR markers showing planar coordinates

C. Coloring on Vibrator Hoses

In order to determine the insertion depth of the vibrator, the hose was colored (Fig. 3). The coloring form was made to be another color inside the outer color referring to the research of Imai *et al.* [3], and 1 double color marker was 50 mm. As a further improvement in this study, 25 mm was left between the color markers, and this part was also colored according to the distance from the hose tip.

A total of 6 colors were used: red, purple, green, blue, yellow, and white, and these were placed on the outside and inside of the marker and in the middle of the marker as shown in Fig. 4. Thus, by using a combination of the color markers (50 mm) and neutral colors (25 mm), the distance from the hose tip (insertion depth) could be uniquely identified.

Figure 3.   Vibrator hose coloring

Figure 4.   Color combination patterns

D. Building System Configuration Using the Cloud

In this study, a system configuration utilizing the cloud was constructed for making real-time the process from photographing to analysis processing and result display (Fig. 5). The shaded part of Fig. 5 represents a cloud server. When a moving image taken by a smartphone is

uploaded and saved on Server 1, the subsequent automatic processing is activated. Automatic processing is terminated in the DB that stores the analysis results of Server 3 via the analysis processing server of Server 2. By going to the cloud, the machine specifications of the analysis processing server can be enhanced, and the processing speed is increased. In addition, since the analysis results can be confirmed anywhere, there is an advantage that they can be confirmed on-site, at an office, or in a remote place.
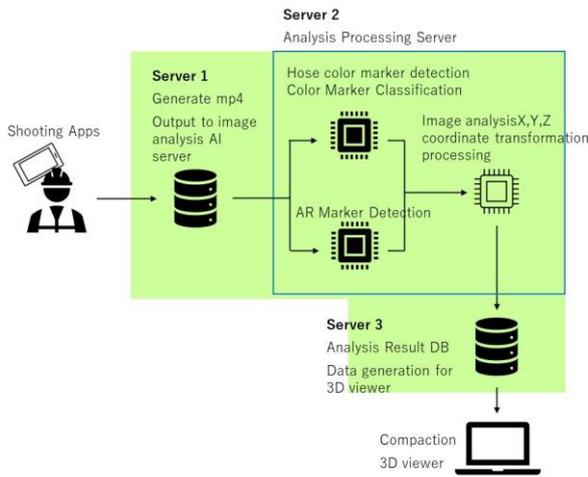
Figure 5. System configuration

E. New Construction of a Visualization System for Compaction Position

We constructed a system that automatically visualizes the calculated compaction position using AI and video analysis (Fig. 6). In the visualization system, the position and depth of compaction by one insertion of the vibrator were expressed by one sphere. The size of the sphere was considered to be an appropriately compacted range, and specifically, the diameter of the vibrator was set at about 10 times the manufacturer's recommended value [4]. If all the insertion points of multiple workers can be analyzed, the interior of the form will be filled sequentially by the filling of the spheres and the concrete placing work. On the other hand, if there is compaction unevenness, a space without a sphere exists. In the construction site trial described in the next chapter, compaction is performed while checking the system shown in Fig. 6 with a tablet (i-Pad).
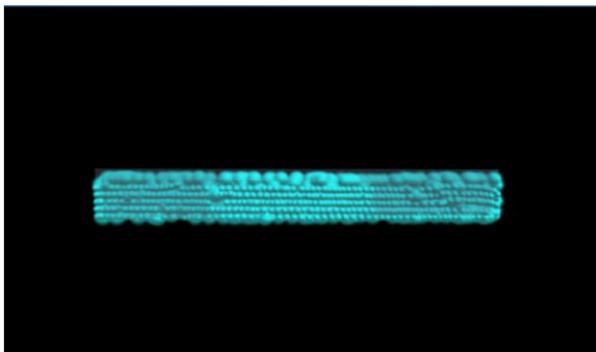
Figure 6. Compaction position visualization 3D system

III. OVERVIEW OF CONSTRUCTION SITE TRIALS

A. Construction Summary

The construction to test the technology of this study was "Upstream of Omono River: Construction of a new Osawa River Sluice Gate (Ordered by the Ministry of Land, Infrastructure, Transport and Tourism, Tohoku Regional Development Bureau, Contractor: Shimizu Corporation)" (hereinafter referred to as the "covered works"). The target construction is part of the Omono River Emergency Flood Control and River Extreme Disaster Countermeasures Special Emergency Section in July 2017, and the construction was to build a bank and install a sluice gate in the lowest part of Osawa River to prevent the backflow of river water from Omono River to Osawa River in case of flood. The outline of the target construction is shown in Table I and the completed image is shown in Fig. 7.

TABLE I.   OUTLINE OF THE TARGET WORKS

Item	Content
River Earthworks excavation (ICT)	Sediment 38,100m³ Soft rock 20,050m³ Road fill (ICT) 21,100m³
Sluice gate	Sluice gate height 15.5m, width34.2m, 4-ply BOX structure Box and culvert height 6.3m width29.4m 2m 4 box structure
Extension	63.8m
Concrete volume	8,230m³ include building
Temporary closing construction	Part 1 Part 2 Split Construction
The embankment revetment Fill (ICT)	39,200m³
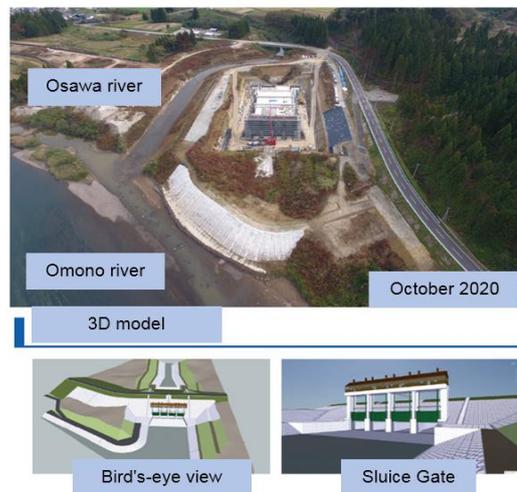
Figure 7. Aerial view as of October 2020 (top) and completed 3D model (bottom)

B. Building a Local 5G Communications Network

In the target construction, a proprietary local 5G communication network was constructed at the site for various ICT construction purposes. 5G (5th generation mobile communication system) is a communication standard in which 4G (LTE) is advanced, and it is a

network which has been in full operation in Japan since the spring of 2020. 5G is a communication standard with three main features: High-capacity communication, multiple connections, and low latency and high reliability.

Local 5G is a dedicated network that allows local and corporate entities to build, operate, and use their own 5G networks in specific areas, such as in their own buildings or premises, as opposed to the 5G services provided by national carriers (Fig. 8).

Figure 8.   Local 5G antenna installed at the site.

IV.   COMPACTION POSITION DETERMINATION ALGORITHM USING AI MODEL

A. Overall Data Processing Flow

The processing flow from the step of acquiring and transmitting the moving images to the step of outputting the analysis result of the camera mounted on the helmet of the compaction worker is shown in Fig. 9. In the processing flow, (1) is performed by a smartphone application, and (2) to (12) are performed by a data analysis server in the cloud. The moving image transmission in (1) is used as a trigger, and all subsequent processing is automated.
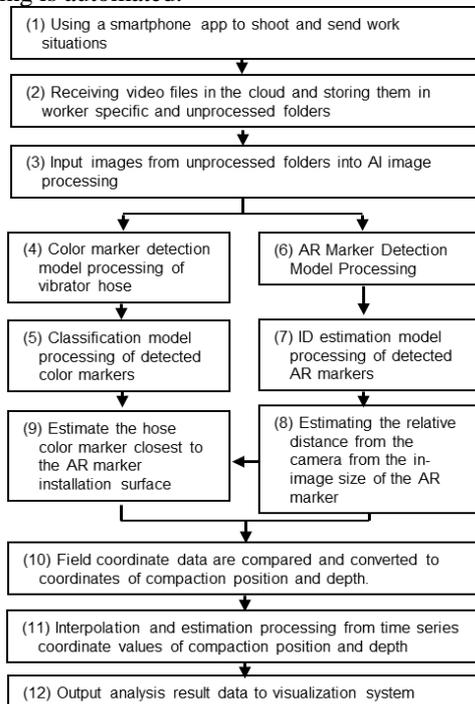
Figure 9.   Processing flow for compaction location analysis
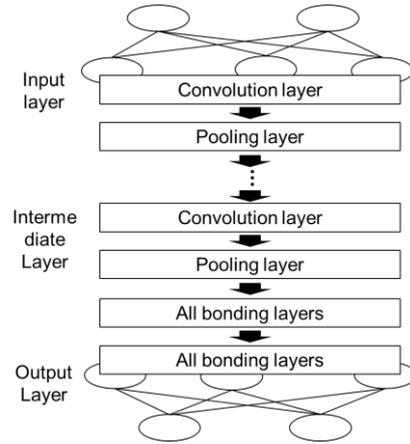
Figure 10. Example of a CNN network structure [5]
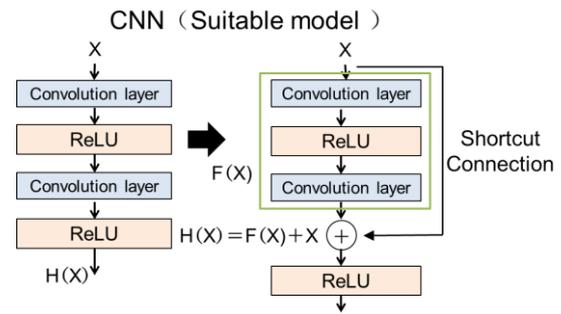
Figure 11. Portions of a typical CNN network and ResNet [5]

B. Application of AI Models

AI models for image detection include CNN (Convolutional Neural Network), a type of deep learning [5]–[12]. 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. In deep learning, CNN has a network with a convolution layer, a pooling layer, and all bonding layers (Fig. 10).

In recent years, CNN research has shown that deepening the layers improves performance, but it has also been reported that simply deepening the layers deteriorates the performance [13].

The deterioration problem is the phenomenon that the improvement of training error in learning a model with deep layers plateaus earlier than a model with shallow layers. Focusing on this problem, ResNet has been devised as a network architecture that can learn even in deep layers.

The difference between ResNet and ordinary CNN is that it learns the residual function with reference to the input of the layer. A part of a typical network and ResNet is shown in Fig. 11. In the case of the function of H(x), in ResNet, in two consecutive convolutional layers, the input x to the output two layers ahead, skips to connect. At this time, we can input x difference from F(x) to become Eq. (1), which is modified and redefined to learn Eq. (2).

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

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

Thus, the optimal function F is replaced by the problem of estimating the optimal residual function. Even when the identity map is optimal, it can be easily approximated if the Shortcut Connection acts as a detour to add up the layer's input values with the network's output before the activation function. Since the Shortcut Connection flows the input information as it is, it flows the gradient as it is even during reverse propagation. Therefore, there is no worry that the slope may be small or too large, and a significant slope 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 layers is called a Residual Network. ResNet is a network model that adds an input layer and an output layer to this. There are 5 types of ResNet that have differences in the number of layers and the number of parameters that can be learned: ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152 [5].

In this study, we use the ResNet50 model and ResNet34 to detect vibrators. RetinaNet is used to detect AR markers. RetinaNet is an object detection model proposed by Facebook AI Research (FAIR) in their August 2017 paper "Focal Loss for Dense Object Detection." As pointed out in the motivation for the development of the paper, many of the accurate object detection models before RetinaNet were built on the R-CNN-based two-stage object detector, but RetinaNet was improved to make it faster.

C. Building AI Models

Of the processing flows in Fig. 9, AI models were constructed for Fig. 9{(4), (5), and (6)}. In the selection of the AI model, assuming that the accuracy of the analysis is ensured, the inference time should be as short as possible in this study because the analysis results must be output before concrete hardening. The ResNet50 model was adopted to detect the color marker of the vibrator, although the difficulty is high and an inference time is required to detect the vibrator from various objects in the moving images. As for the classification of the color markers, we adopted ResNet34, which has a fast inference time, because moving images after the vibrator is detected by the aforementioned model are subject to analysis. For the detection of AR markers, the RetinaNet model, which is characterized by high-speed processing, was adopted because it was confirmed at the examination stage that the detection accuracy could be ensured.

Fig. 9{(4)}'s vibrator hose color marker detection exploits the ResNet50 model and returns a bounding box for the location of the color marker. The color marker classification model in Fig. 9{(5)} judges three colors, the outer color, inner color, and border color, for the image in the bounding box extracted in Fig. 9{(4)} above. The model used a custom classification model from ResNet34. The AR marker detection in Fig. 9 {(6)} utilizes the RetinaNet model and returns a bounding box for the location of the AR marker.

In order to construct these artificial intelligence models, teacher data were acquired and analyzed on the dates shown in Table II among the concrete placing dates from June to December 2020 in the target construction.

We verified the results, and the characteristics of this study are that all the work days were permanent construction, and various teacher data were acquired and models were constructed during the actual construction site work.

The number of teacher data images for the AI model is shown in Table III. In total, more than 20,000 teacher data images were used for learning. The large number of teacher data values in the color marker classification model is due to the fact that there are 12 kinds of color markers and 6 neutral colors, and learning was carried out to make them classifiable.

TABLE II. 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 III. NUMBER OF IMAGES OF TEACHER DATA

	Model Type	Number of teacher data	Number of evaluation data
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 IV. 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	ResNet 50	83	96	0.89
AR Marker Detection	Retina-Net	86	70	0.77

TABLE V. JUDGMENT ACCURACY OF COLOR MARKER DISCRIMINATION MODEL

	model type	Outside color	Inside color	Fit ratio (%)	Reproducibility (%)	F-Value
Classification 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

D. Decision Accuracy of AI Models

The accuracy of the AI model constructed in this study is shown in Tables IV and V. The color marker detection of the vibrator is evaluated to have high evaluation performance with high accuracy and reproducibility.

An example of a slightly low recall rate of 70% in AR marker detection is because some AR markers are hidden by obstacles such as rebars and hoses (Fig. 12).

Since many images have more than 3–4 AR markers in one image at a normal shooting angle of view, this model was also evaluated as an effective level considering the requirement to detect at least 2 AR markers.

The discriminant model of color markers has a high accuracy rate and high recall rate, with an F-value of 0.93 or more, and can be judged to have high evaluation performance, even when viewed with a total of 12 color schemes.
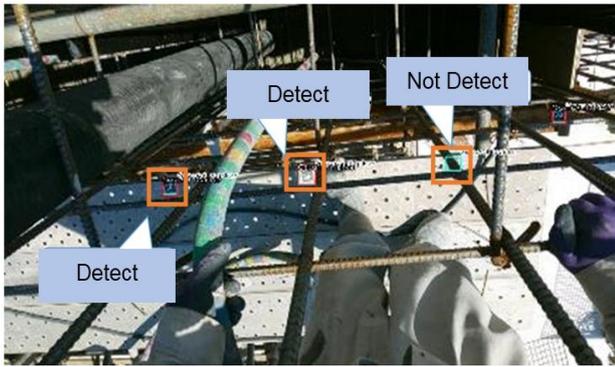
Figure 12. Success and failure cases of AR marker detection

V. CONSTRUCTION SITE TRIAL RESULTS

A. Summary of the Trial Results

In the target construction, we tried photographing with a worker's wearable camera for analysis with the AI model constructed in this study. Fig. 13 shows an example of the analysis results, in which the compaction effect range (about 10 times the diameter of the vibrator) centered on the insertion point of the vibrator is represented by a sphere. It was confirmed that the spheres were displayed by placing layers of concrete according to the work procedure. At the construction site, trials were conducted from September 18 to December 18, as shown in Table II.
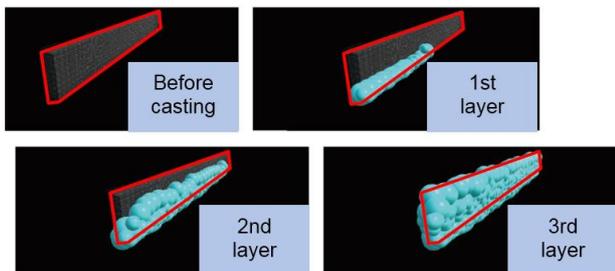
Figure 13. Results of on-site construction on September 18, 2020

B. Real-time Verification

1) Wearable camera image quality and transmission time

Since the target construction site was located about 500 m away from the neighboring village, the actual communication speed was about 1 Mbps on the 4G line (docomo) available with a common smartphone. In addition, since it was possible to construct and use a proprietary local 5G line in the target construction, moving images captured by the wearable camera were actually transmitted using this network and compared with the above 4G lines (Table VI).

It has been confirmed that the video data capacity for 1 minute of shooting increases as the number of pixels increases, reaching about 934MB at 8K. The time required to upload these moving image files to a cloud server on the Internet (transmission time) is generally two minutes or less, considering the time required for the post-process moving image processing. In the target construction, as shown in Table VI, it became clear that the transmission time would be within 2 minutes even at 8K image quality on a local 5G line, but the number of pixels should be less than HD image quality on a general 4G line.

TABLE VI. DATA CAPACITY AND TRANSMISSION TIME FOR 1 MINUTE OF PHOTOGRAPHY BY NUMBER OF PIXELS (RESULTS OF MEASUREMENTS AT THE CONSTRUCTION SITE)

	Number of pixels	Data capacity for 1minute of shooting (MB)	Local 5G send time (seconds)	4G send time (seconds)
8K	7680×4320	934	103	827
4K	3840×2160	403	44	357
2K	1920×1080	204	23	181
HD	1280×720	94	10	83
SD	640×480	60	7	53

2) Analysis time of AI model

For analysis of the AI model on the cloud side, we prepared multiple servers with machine specifications of 16GB GPU, 8 core CPU, and 61GB memory, and performed parallel processing. Of the flows shown in Fig. 9, the analysis processing time from Fig. 9{(3)–(11)} was about 4 to 6 minutes.

3) Total time feedback result

In the arrangement up to the preceding paragraph, it took up to 2 minutes to transmit data from the wearable camera, up to 6 minutes to analyze the AI model on the cloud side, and up to 1 minute to display the visualization system, so the total was 9 minutes, and the results of the analysis could be fed back within 10 minutes of the shooting. In this study, the goal was to detect under-compacted areas before the concrete hardens so that re-vibration can be performed. As such, result feedback within 15 minutes was initially the target value, and we were able to obtain a result which achieved this.

C. Challenges in Construction Site Applications

As a result of the trial of this technology on-site, the following problems were realized and issues with a

system utilizing an AI model are described for construction site application.

1) Response to weather, temperature, etc.

As a result of applying this technology to concrete construction from June to December, the thermal runaway of devices (equipment) in summer became a problem. Wearable cameras at work often shut down due to heat, and we replaced the devices each time.

Second, in winter, when the temperature is low, the battery life of the device decreases, which also necessitates frequent replacement of the wearable camera. When these wearable devices and mobile communication devices are used at the construction site, it is important to maintain a sufficient reserve for construction.

On the other hand, the detection situation of the AI model showed a certain accuracy regardless of the season, and stable function operation could be carried out. The reason for this is that there was no night construction in the target construction, and we could take pictures under sunlight during any season.

2) Adjusting the angle of view of the wearable camera

Wearable cameras worn on the helmets of workers made light contact with reinforcing bars when the workers performed actions such as putting their heads between the bars to check the lower layers, and the angle of the cameras changed, making it difficult to photograph the objects. In this study, as shown in Fig. 1, although in the existing helmet the camera is fixed to the outside with a band and tape, the camera is mounted in a protruding manner, which can cause contact with the rebars.

On the construction site, the camera angle was adjusted each time the camera was out of alignment, but the fundamental challenge is to make the device even smaller and lighter, and to develop a mounting method that does not disturb the wearer.

3) Ensuring an on-site communication environment

In order to shorten the data transmission time from the wearable camera, it is important to secure the communication environment at the construction site. When the communication speed of a general mobile line is not high enough, as in the case of the subject construction, the improvement of local 5G and outdoor Wi-Fi environment becomes a problem.

VI. Summary

In this study, we developed a system to quantitatively examine compaction points in concrete construction using moving image analysis with an AI model and to feed back the results in real time. The required function of this system is to feed back the compaction position quickly and accurately from the moving images, but in actual concrete work, there are obstacles such as dirt on the markers and overhanging reinforcing. The feature of this study is that the data acquired under actual site conditions were utilized as teacher images, and the accurate judgment was verified by the AI model that had been trained.

In the construction site trial, it was confirmed that if the analysis results were fed back in real time, it would be possible to detect compaction deficiencies early enough

to enable re-vibration, etc. In addition, this system is considered highly useful in shifting from conventional qualitative management that relies on experience and know-how to quantitative and objective compaction management.

On the other hand, there are also challenges in operating the system at outdoor sites, and it is necessary to continuously examine these countermeasures in the future.

In addition, future development utilizing this technology is considered to include the function of automatically determining insufficient compaction points, informing of the points that need to be re-vibrated, guiding the workers, and the application to education and guidance for inexperienced workers.

Conflict of Interest

The authors declare no conflicts of interest with this manuscript.

Author Contributions

All authors made substantial contributions to the study concept or the data analysis or interpretation; drafted the manuscript or revised it critically for important intellectual content; approved the final version of the manuscript to be published; and agreed to be accountable for all aspects of the work.

Acknowledgment

References

[1] S. Sunaga, Y. Tsujita, K. Futamura, J. Murakami, and T. 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] R. Imai, T. Kurihara, T. Taniguchi, M. Ito, T. 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.

[3] R. Imai, T. Kurihara, and T. 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.

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

[5] Y. Tsuzuki, P. J. Chun, and T. Yamane, "Automatic detection of cracks in asphalt pavement using weak type improved deep learning," *AI and Data Science Theses*, vol. 1, no. 1, pp. 168-179, 2020.

[6] P. J. Chun, S. Izumi, and T. Yamane, "Automatic detection method of cracks from concrete surface imagery using two‐step light gradient boosting machine," *Computer-Aided Civil and Infrastructure Engineering*, 2020.

[7] P. J. Chun, S. Izumi, T. Yamane, and N. Kuramoto, "Development of a machine learning-based damage identification method using multi-point simultaneous acceleration measurement results," *Sensors*, vol. 20, no. 10, p. 2780, 2020.

[8] P. J. Chun, T. Yamane, S. Izumi, and T. Kameda, "Eval-uation of tensile performance of steel members by analysis of corroded steel surface using deep learning," *Metals*, vol. 9, no. 12, p. 1259. 2019.

[9] P. J. Chun, K. Ujike, K. Mishima, M. Kusumoto, and S. Okazaki, "Random forest-based evaluation technique for internal damage in reinforced concrete featuring multiple nondestructive testing results," *Construction & Building Materials*, vol. 253, no. 30, 119238, 2020.

[10] P. J. Chun, A. Igo, "Detection of concrete surface cracks by random forest," *Journal of the Japanese Society of Civil Engineers,* vol. 71, no. 2, pp. I _ 1-I _ 8, 2015.

[11] P. Chun, Y. Shimamoto, K. Okubo, C. Miwa, and M. Ohga, "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.

[12] Y. Lecun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackle, "Back-propagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541-551, 1989.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770-778.