Machine Learning-Based Anomaly Detection in Concrete Structures: Influence of Training Data Set and On-Site Validation

Ruben Vargas[1], Katsuya Ikeno[2], Hideki Naito[3] and Tomoyuki Kimoto[4]

[1] Member of JSCE, Institute of Technology, Penta-Ocean Construction Co., Ltd.
(1534-1 Yonku-cho, Nasushiobara, Tochigi, Japan)
E-mail: ruben.vargas@mail.penta-ocean.co.jp (Corresponding Author)
[2] Member of JSCE, Institute of Technology, Penta-Ocean Construction Co., Ltd.
(1534-1 Yonku-cho, Nasushiobara, Tochigi, Japan)
E-mail: katsuya.ikeno@mail.penta-ocean.co.jp
[3] Member of JSCE, Associate Professor, Tohoku University
(6-6-06, Aramaki Aza Aoba, Aoba-ku,Sendai, Miyagi, Japan)
E-mail: hideki.naito.c2@tohoku.ac.jp
[4] Professor, National Institute of Technology, Oita College
(1666 Maki, Oita City, Oita Prefecture, Japan)
E-mail: kimoto@oita-ct.ac.jp

A machine learning-based anomaly detection method for concrete structures, using localized vibration measurements and transfer learning, has been previously proposed by the authors. The method detects anomalies based on changes in the frequency response function obtained from localized vibrations measurements. An autoencoder model is first trained using sound-state data from an arbitrary-structure to create a general-purpose model, which is then customized for a target structure using transfer learning with a small amount of sound-state data from that target. This paper investigates the influence of the arbitrary structure's structural type on the model's performance and provides an on-site validation. To this end, three general-purpose autoencoder models were developed using data from full-scale beam and slab specimens: a "beam-trained model", a "slab-trained model", and a "mixed-data model". The validity of these models was confirmed by using them to identify load-induced damage in the specimens. The models were then evaluated through an on-site inspection of a concrete pier's main beam and deck slab. Results indicate that while transfer learning provides a robust baseline for anomaly detection even with mismatched structural types, accuracy is significantly improved when the training data's source structural type matches the target structure. Furthermore, a model trained on a mixed-data set exhibits superior generalization and robustness, suggesting that for practical field applications, the development of a versatile model would be beneficial.

Key Words : *RC structure, anomaly detection, local vibration test, transfer learning, on-site inspection*

1. INTRODUCTION

Starting in the 1950s, Japan's period of rapid economic growth led to intensive infrastructure development. Consequently, the proportion of structures now exceeding 50 years of service is rapidly increasing, and this ratio is expected to grow significantly in the following 20 years[1)]. The proactive maintenance and inspection of these aging structures is paramount to ensure public safety and economic stability. While visual inspections[2)] and conventional non-destructive testing (NDT) methods such as ultrasonic testing[3)], thermography[4)], etc. play a vital role, they have limitations in assessing the internal integrity of a structural member, especially when complex damage patterns such as internal cracking are present.

To address these challenges, the localized vibration method has been proposed as an effective technique for evaluating the internal condition of concrete members[5),6)]. As shown in **Fig.1**, the method involves applying a wide-band input vibration with a portable shaker and capturing the resulting acceleration response. A Fast Fourier

Transform (FFT) is then applied to the acceleration records to estimate the localized Frequency Response Function (FRF).

Originally[5], the anomaly identification methodology focused on identifying the primary resonant frequency of the member and comparing this peak-frequency value to its corresponding sound-state baseline, which can be approximated using Equation (1a).

When multiple-reflected waves are induced in the thickness direction of the target, the first-order resonance frequency in sound state can be approximately estimated by Equations (1a) and (1b).

$$f_0 = \frac{c}{2L} \tag{1a}$$

$$c = \sqrt{\frac{E_d}{\rho}} \tag{1b}$$

Where f_0 is the first-order resonance frequency, c is the apparent speed of sound within concrete, L is the vibration length (thickness), E_d is the dynamic Young's Modulus, and ρ is the element's density. A reduction in the estimated f_0 is interpreted as an indication of internal damage, such as cracking, which reduces the member's stiffness.

This conventional approach faces two practical challenges: (1) a reliable healthy-state baseline is often unavailable for existing infrastructure; and (2) complex on-site conditions, and diverse damage patterns can result in an FRF with multiple peaks, making it difficult to reliably identify the primary resonant frequency, consequently, leading to potentially inaccurate assessments.

To overcome these challenges, the authors have proposed a machine-learning based approach, that analyzes the entire shape of the FRF using an autoencoder[7), 8)]. As schematically shown in **Fig.2**, an autoencoder extracts the features of input data through dimensional reduction in intermediate layers, and then reconstruct the original data from those features. By training the model exclusively with sound-state FRF datasets, it learns to accurately reconstruct "sound-state" vibration responses only. When presented with data from a damaged location, the model fails to reconstruct it accurately, resulting in relatively large reconstruction errors, which are used as an index to identify potential damage.

A significant challenge for applying this method in the field is the need for a relatively large set of sound-state data from the target structure itself, which is often difficult to acquire for aged structures that are already partially damaged. Therefore, to make the method practical for field use, the use of transfer learning has been proposed[8)]. In this scheme, a "general-purpose model" is first trained on a large dataset from arbitrary structures. This model is then leveraged for a specific target by customizing it with a very small set of sound-state data collected from the target structure.

While a validation of the transfer learning approach on a pre-stressed slab subjected to fatigue damage was performed by the authors[8)], the influence of the training data used to build the

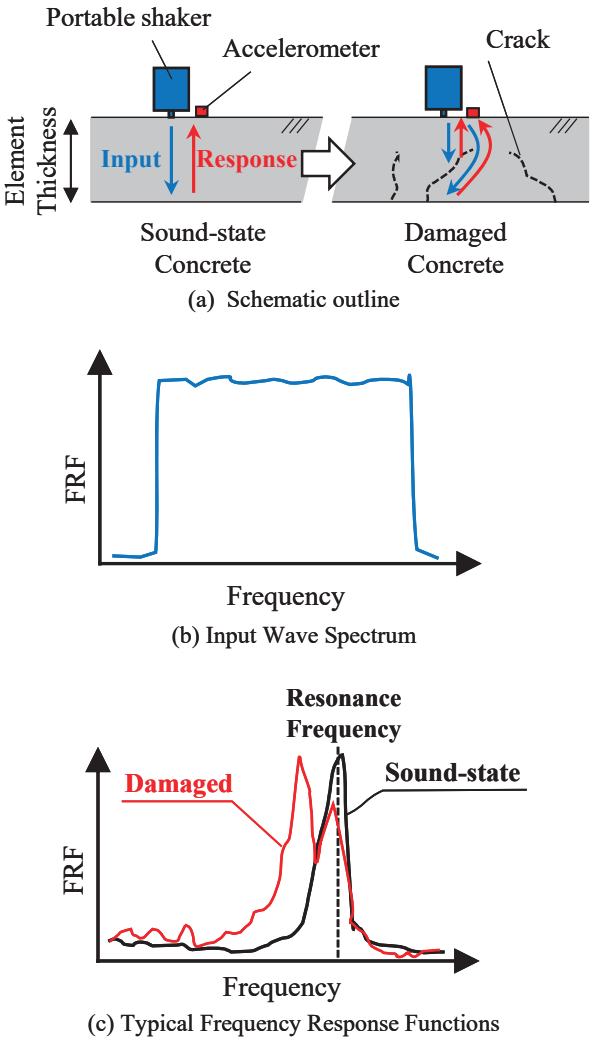


Fig.1 Schematic outline of the localized vibration method

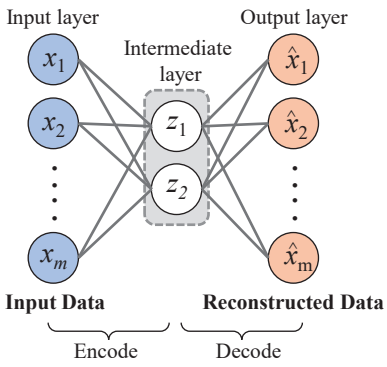


Fig.2 Autoencoder image

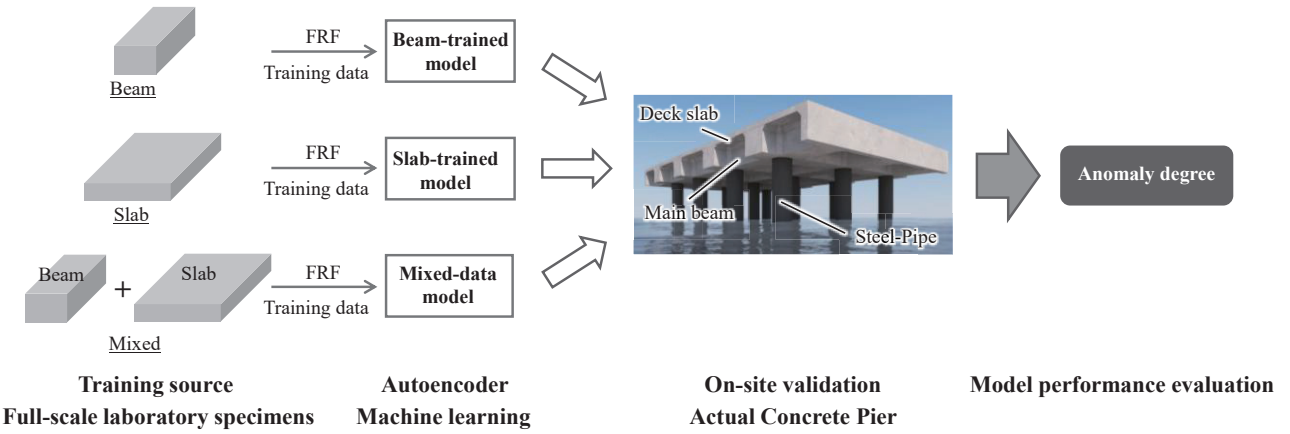


Fig.3 Research flow outline

general-purpose model was not clarified. This is a critical issue for inspecting complex structures like piers, which are composed of different structural types (e.g., beams and slabs). It remains unconfirmed whether the structural type of the training data significantly impacts inspection accuracy.

This paper, therefore, investigates two key questions: first, whether structural-type similarity between the training source and the inspection target is critical for accuracy, and second, whether a versatile model trained on mixed data can be effective. As shown in **Fig.3**, to address these questions, three distinct general-purpose models were established using data from full-scale laboratory specimens: a "beam-trained model" a "slab-trained model" and a "mixed-data model". The validity of each model in detecting damage was then confirmed through loading tests on the same specimens. Finally, the performance of these models was evaluated in a real-world scenario through an on-site, transfer learning-based inspection of the main beam and deck slab of an existing concrete pier.

2. ANOMALY DETECTION METHOD

As previously introduced, the proposed anomaly detection method is a two-stage process that first establishes a general-purpose model using an autoencoder, which is then customized for a specific target structure via transfer learning.

(1) Autoencoder – General Purpose Model

An autoencoder is a type of unsupervised neural network trained to extract the features of input data through dimensional reduction (encoding), and then reconstruct the original data by inverse processing (decoding)[9]. As illustrated in **Fig.4(a)**, the autoencoder architecture consists of an input layer, three encoding layers that progressively compress the data, and three decoding layers that reconstruct the data back to its original dimension.

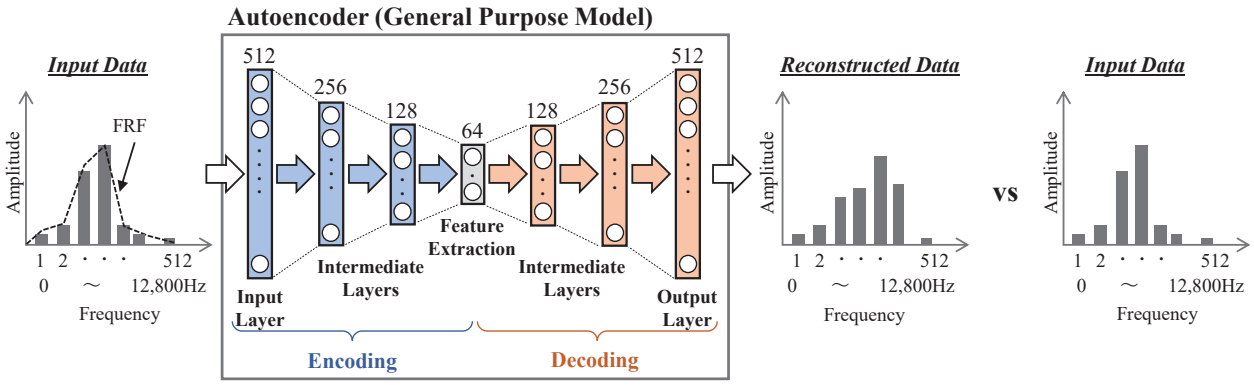
The input for the model is the FRF from the local vibration test, treated as a one-dimensional vector with 512 parameters, covering a frequency range of 0 to 12,800 Hz (at a resolution of 25 Hz). To ensure consistency and robustness, the data undergoes pre-processing steps prior to training. First, to reduce variability from measurement conditions, the amplitude of each FRF is normalized. Second, to enable the comparison of members with different thicknesses and physical properties (stiffness and density), the frequency axis is normalized by scaling the theoretical resonant frequency (estimated through Equation (1a) and (1b)) to a standard frequency, arbitrarily set to 5,000 Hz, which is approximately the mid-point between the theoretical resonance frequency of typical beams and slabs in pier structures. Third, the training dataset is augmented five-fold by shifting each FRF along the frequency axis[8].

The model is trained using the Adam optimization algorithm with a learning rate of 0.001 and a mini-batch size of 32 for a maximum of 500 epochs.

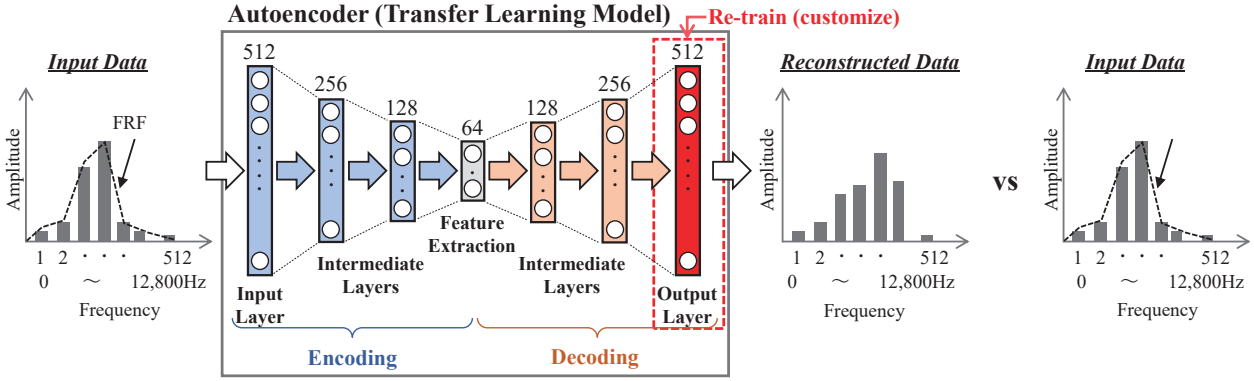
By the above processing, the input frequency response function was set to $f_i(1 \le i \le 512)$, the output reconstructed frequency response function was set to $g_i(1 \le i \le 512)$, and the reconstruction error was evaluated as the anomaly degree e using equation (2).

$$e = \frac{\sum_{i=1}^{512} |f_i - g_i|}{\sum_{i=1}^{512} g_i} \tag{2}$$

In addition, since the frequency response function has a steep multi-peak, the accumulation of small differences between the frequency response function input to the autoencoder and the reconstructed



(a) Autoencoder – General-purpose model



(b) Autoencoder – Transfer learning model

Fig.4 Anomaly detection schematic outline

frequency response may be evaluated as a large reconstruction error. Therefore, to estimate the anomaly degree, the data is smoothed using a 10-point moving average function, and noise is reduced by zeroing out any components below 10% of the peak amplitude.

(2) Autoencoder - Transfer Learning Model

The underlying principle of transfer learning is to leverage the knowledge a model has gained from a general dataset and apply it to a new, specific one[10]. In this context, once a model is trained on a general-purpose dataset, its intermediate and output layers can be updated through re-training, customizing it into a prediction model that utilizes the newly learned features.

In practice, RC structures even at stages with significant deterioration (such as curing cracks, rebar corrosion, delamination, etc.) rarely experience uniform, severe damage. It is therefore often possible to obtain measurement data from locations that are still relatively sound. This small amount of sound-state data from the target can be used to customize the pre-trained autoencoder via transfer learning. While this assumption forms the basis for the method's applicability, it also defines its primary limitation: the approach may not be suitable for structures that have undergone uniform, advanced deterioration, where a representative "sound-state" baseline cannot be established. The proposed method is therefore positioned as a tool for the early detection of localized internal damage, a stage at which conventional methods like visual inspection may be insufficient, but a baseline of relatively sound data is typically still available.

When inspecting a structure, since it is difficult to identify the "sound-state" measurement points directly, an initial screening using the pre-trained general-purpose model is performed to estimate preliminary anomaly degrees of all measurement data. A small subset of the data corresponding to the lowest anomaly indices is then selected and designated as the "sound-state" dataset for the model customization. Regarding the required amount of sound-state data for customization, this study uses a fixed number of 10 points. A sensitivity analysis based on the on-site validation results justifying this choice is presented in Chapter 4. However, a comprehensive methodology for determining the optimal data requirement for effective customization remains a subject for future investigation.

As shown in **Fig.4(b)**, this small, target-specific

dataset is used to customize the general-purpose model. Specifically, only the weights of the final output layer of the decoder are re-trained, customizing the model to the unique characteristics of the target structure, creating a specialized detector with minimal on-site data.

3. GENERAL-PURPOSE MODELS CONSTRUCTION AND VALIDATION

Full-scale RC beam and slab specimens were fabricated, and sound-state data was collected to build the general-purpose models. This section details the creation of the general-purpose models and their initial validation.

(1) Beam and Slab Test Specimens

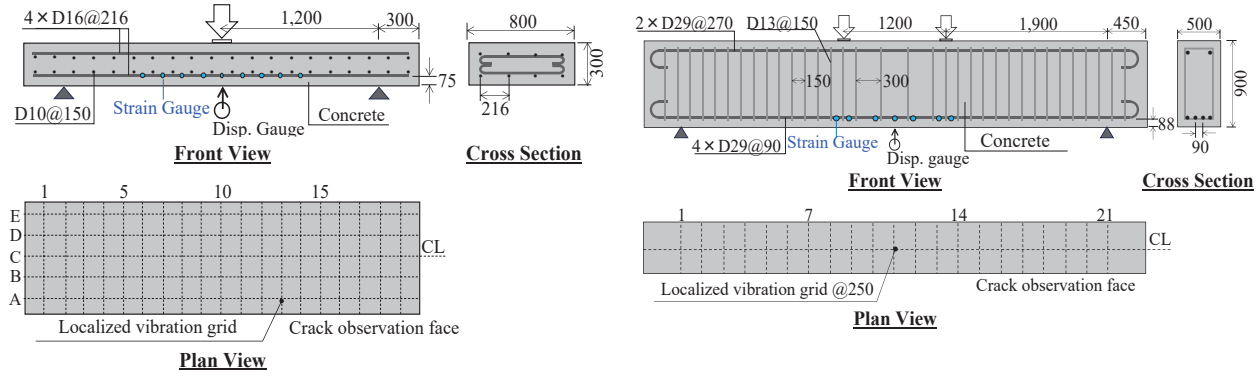
The outlines of the full-scale slab and beam specimens are shown in **Fig.5**. The slab specimen had dimensions of 3,000 mm in length, 300 mm in thickness, and 800 mm in width. It was reinforced with four D16 rebars at the top and bottom in the longitudinal direction (@216 mm) and D10 rebars on the short side at a 150 mm pitch. The beam specimen measured 5,900 mm in length, 900 mm in height, and 500 mm in width. It was reinforced with four D29 rebars at the bottom (@90 mm), two D29 rebars at the top, and D13 stirrups at spacings between 150 to 300 mm.

The target concrete compressive strength for both specimens was 30 N/mm², and the reinforcing steel was grade SD345. The mechanical properties of the rebar and the concrete are summarized in **Table 1**. Prior the loading process, an extensive set of localized vibration measurements was performed across the surface of each specimen to generate the sound-state datasets for model training.

(2) Loading Tests and Damage Induction

The slab and beam specimens were subjected to static loading aiming to induce damage in a controlled manner. As shown in **Fig.5**, the slab specimen was loaded under three-point load bending, with the load applied at its mid-span. The beam specimen was tested under four-point load bending, creating a 1,200 mm zone of pure bending. The load was applied incrementally using a hydraulic jack, and un-loading was performed at key stages: initial crack, $0.7P_{yd}$, $0.9P_{yd}$, P_{yd}, $1\delta_y$, and $2\delta_y$; where P_{yd} is the main-rebar design yield load, and $1\delta_y$ corresponds to the displacement at the actual main rebar yield point.

The measured load-displacement behavior is shown in **Fig.6**. For the slab, the first flexural crack appeared at approximately 60 kN, followed by the development of vertical cracks in the central portion, typical of three-point bending tests. For the beam, initial flexural cracks appeared in the pure bending zone at 270 kN, followed by the



(a) Slab Specimen (b) Beam Specimen

Fig.5 Schematic diagram of the specimens (unit: mm)

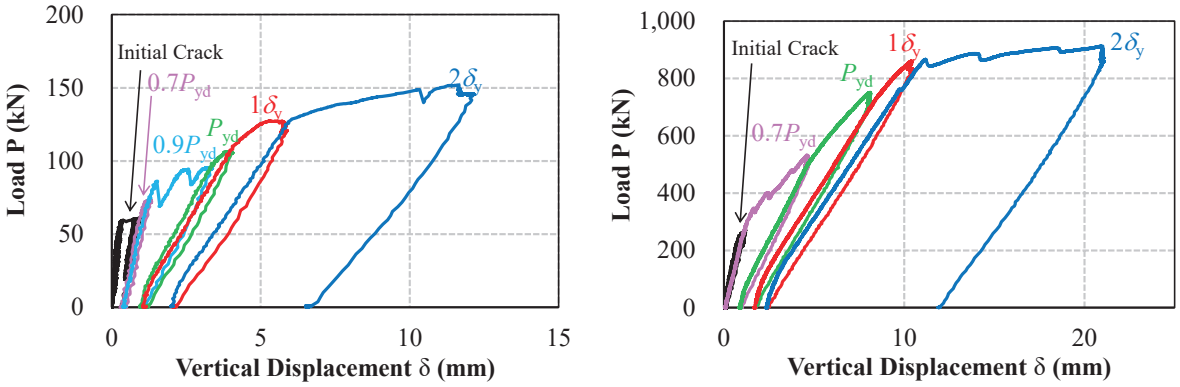
Table 1 Mechanical properties of Rebar and Concrete

(a) Mechanical Properties – Concrete

Specimen	Compressive Strength f_c [N/mm²]	Split tensile strength f_t [N/mm²]	Speed of sound c [m/s]	Density ρ [kg/m³]	Young's Modulus E_d [kN/mm²]
Slab	32.4	2.8	3,860	2,330	34.7
Beam	37.5	3.4	3,820	2,320	33.8

(b) Mechanical Properties - Rebar

Specimen	Rebar D [mm]	Yield Stress f_y [N/mm²]	Young's Modulus E [kN/mm²]
Slab	10	370	187
Beam	13	389	191
Slab	16	389	186
Beam	29	384	200



(a) Slab Specimen (b) Beam Specimen

Fig.6 Loading Test – Load-Displacement graph

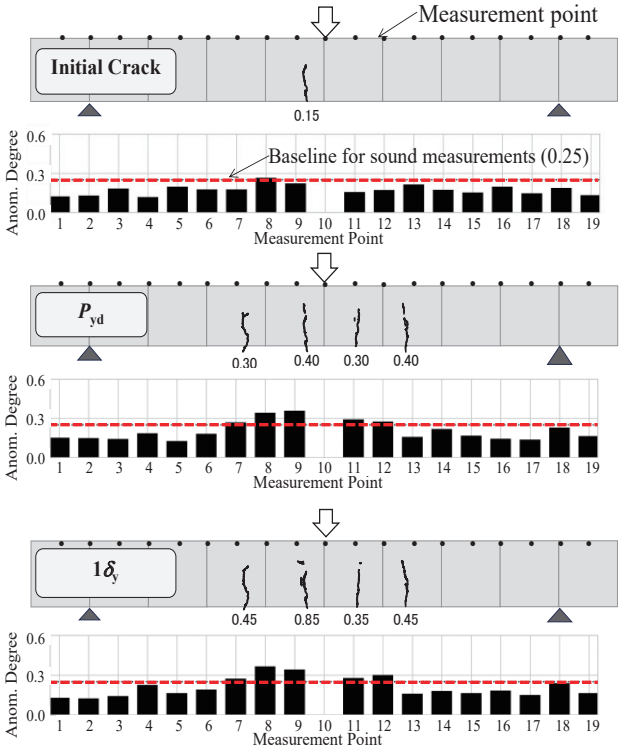


Fig.7 Slab specimen's crack sketch (Line C) and anomaly degree distribution using the "Slab-trained model" (during loading)

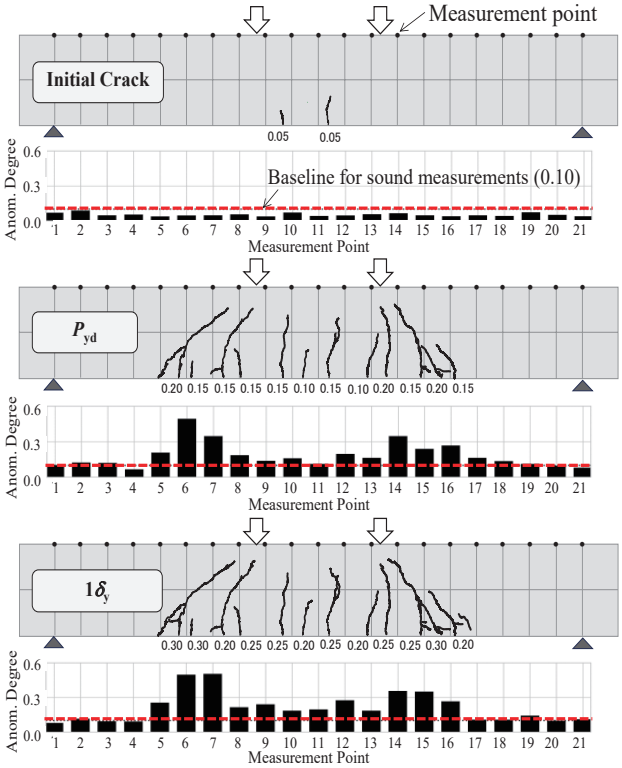


Fig.8 Beam specimen's crack sketch and anomaly degree distribution using the "Beam-trained model" (after unloading)

development of diagonal shear cracks in the shear spans as the load increased. The crack patterns at representative loading steps are shown in **Fig.7** and **Fig.8**. A complete set of localized vibration measurements was taken on the key stages shown on **Fig.6**, during loading, and un-loading processes.

(3) Build and Validation of General-Purpose Models

Following the procedure in **Section 2(1)**, the sound-state FRF data collected from the undamaged specimens were used to train three distinct general-purpose models: a "beam-trained model" from 193 data points, a "slab-trained model" from 476 data points, and a "mixed-data model" utilizing the combined dataset.

The validity of these models was confirmed by comparing the anomaly degree distributions estimated for damaged states with the observed crack patterns, at the Initial Crack, P_{yd}, and $1\delta_y$ stages. **Fig.7** presents this comparison for the slab specimen using the "slab-trained model", and **Fig.8** shows the results for the beam specimen using the "beam-trained model". It is noted that for the slab, measurements shown on **Fig.7** correspond to those taken during loading (Row 10 was inaccessible due to the presence of the loading plate), as the cracks were relatively small after unloading. For the beam, measurements shown on **Fig.8** correspond to those taken after unloading to show the data beneath the loading frame (Rows 8 to 14).

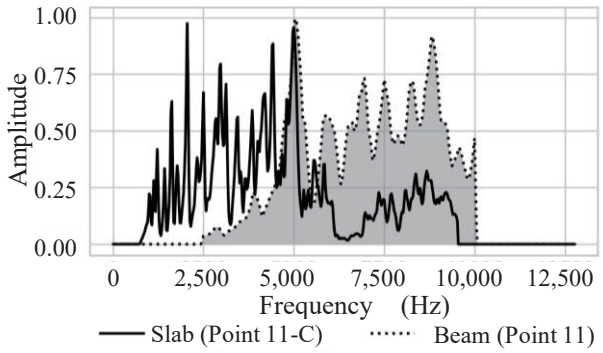
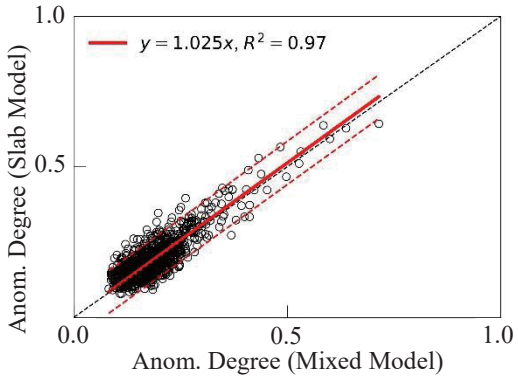


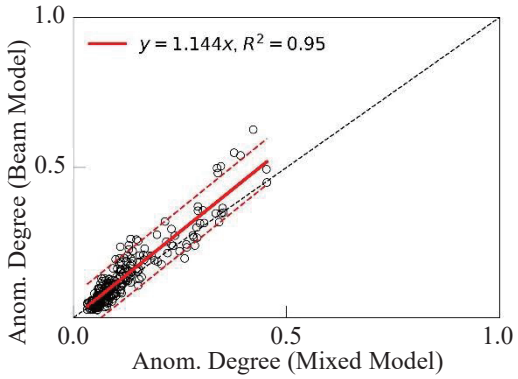
Fig.9 Typical sound-state FRFs for the slab and beam specimens (after normalization in the Amplitude and Frequency axis)

A visual correlation was found between areas with high anomaly indices and the locations of significant cracking, particularly for diagonal shear cracks, as these cause larger impact on the FRF due to wave diffraction around diagonal cracks. This confirmed that the models, could effectively identify and locate areas of damage. As shown in **Fig.7** and **Fig.8**, the baseline anomaly degree for sound areas was found to be approximately 0.25 for the slab specimen, and 0.10 for the beam specimen. Among other effects, this difference on baselines arises because the anomaly index is normalized by the reconstructed FRF's area (Eq. (2)). As the sample sound-state FRFs in **Fig.9** illustrate, the beam specimen has a larger typical spectral area (approximately 1.4 times in the example shown), resulting in a smaller normalized anomaly value for the same absolute reconstruction error.

Furthermore, **Fig.10** compares the predictions of the "slab-trained model" against the "mixed-data model" when applied to the slab specimen, and the "beam-trained model" against the "mixed-data model" when applied to the beam specimen, for the full measurement data set. For both cases, a good correlation, represented by the R^2 value was found. For the slab specimen, the predictions of the "slab-trained model" and "mixed-data model" were found to be nearly equivalent, with a first-order coefficient close to 1.0. For the beam specimen, however, the ratio between the "beam-trained model" and the "mixed-data model" was found to be approximately 1.14. As will be discussed in **Chapter 4**, this difference may be attributed to the superior generalization of the "mixed-data model", which was trained on a larger and more diverse dataset. This effect is particularly evident when compared to the "beam-trained model", which was developed from a relatively small dataset.



(a) Slab-trained model vs Mixed-data model



(b) Beam-trained model vs Mixed-data model

Fig.10 General Purpose Model – Anomaly Degree Comparison

4. ON-SITE VALIDATION

To evaluate the performance of the proposed model in an actual structure, and investigate the influence of the general-purpose models, an on-site survey was conducted on the superstructure of an existing concrete pier supported on steel-pipe piles, with the three validated general-purpose models.

(1) Target Measurement Overview

The target structure was a relatively sound concrete pier, from which a main beam and a section of the deck slab were selected as a target for inspection. The main beam had a height of 1,400 mm, while the deck slab had a thickness of 300 mm. Localized vibration measurements were performed on a grid of points established on both members, as shown in the plan view and cross-section in **Fig.11**.

The localized vibration measurements were obtained on the actual structure, and a visual inspection of the target areas to correlate the model predictions with actual structural conditions was performed at the time the pier was dismantled. The post-dismantling inspection revealed localized damage on the target members. On the main beam, a 0.70mm width crack around the cross section was found around measurement point B13 (see Photo 1), and a 0.60mm width crack between points B16 and B17 (see Photo 2); existing markings of a previous inspection indicate that the identified cracks existed

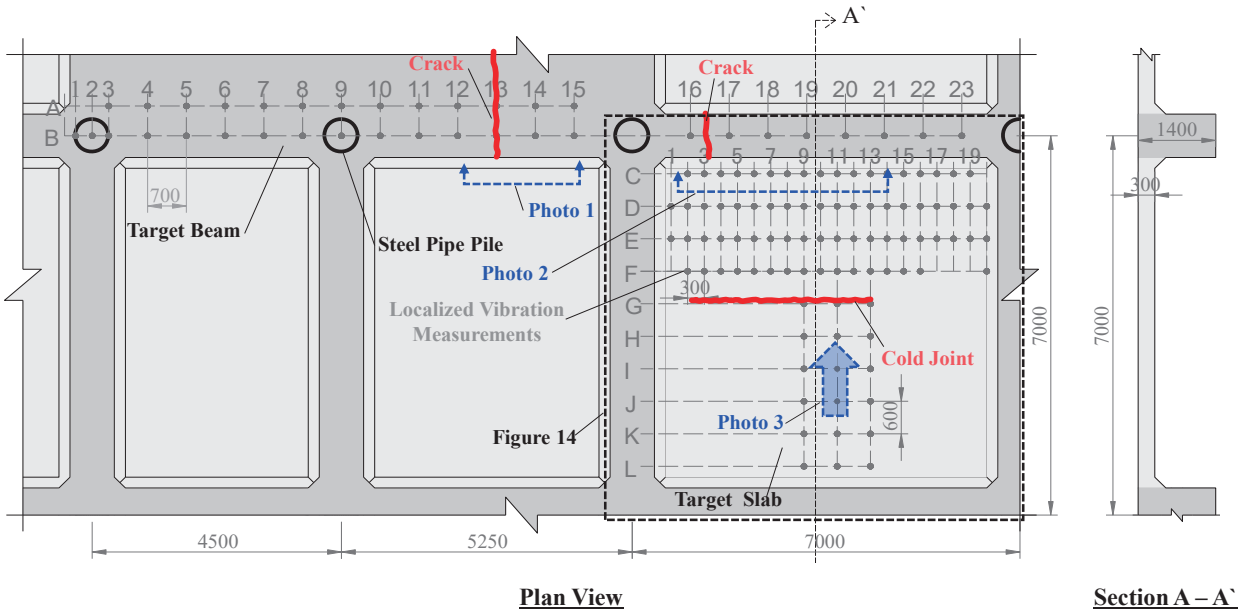


Fig.11 Plan view and cross-section of the target Concrete Pier Structure (unit: mm)

before the dismantling process. On the deck slab, where safety constraints limited the inspection to the upper surface and the cross-sections of the dismantled segments, a cold joint was identified along Grid G (see Photo 3). These known areas of damage served as benchmarks for evaluating the performance of the anomaly detection models.

(2) Determination of Data Points for Transfer Learning

To determine the number of "sound-state" data points, hereafter referred to as Dp, required to customize the general-purpose models, a sensitivity analysis was conducted using the on-site pier data. First, an initial screening was performed by processing the field-measured FRF data with the general-purpose models (the “Beam-trained model” for beam measurements, and the “Slab-trained model” for slab measurements). The transfer learning procedure described in **Section 2(2)** was then applied using datasets composed of Dp = 5, 8, 10, 13, 15, 18, and 20 measurements corresponding to those with the lowest initial anomaly degrees estimated in the initial screening.

Fig. 12 illustrates the influence of varying the Dp on the resulting anomaly degree distributions. It also shows the “sequential R^2” value, a metric representing the correlation between results from one Dp increment and the next (e.g., the value at Dp=5 represents the R^2 between the Dp=5 and Dp=8 datasets).

As shown for the pier beam (**Fig. 12a**), the data distribution and the sequential R^2 value both stabilize after Dp=10, indicating that using additional data does not significantly alter the model's predictions. For the pier slab (**Fig. 12b**),

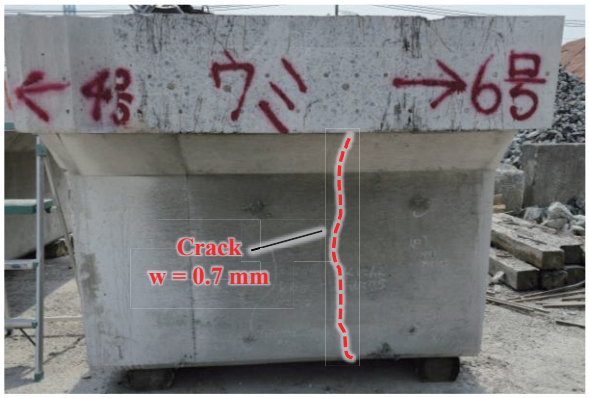


Photo 1 Crack observed near point B13 on the dismantled pier beam

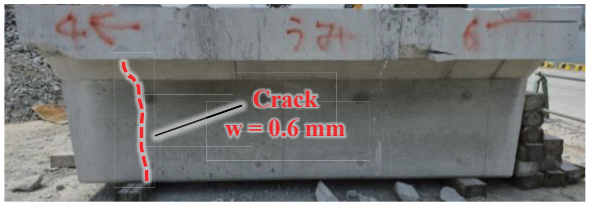


Photo 2 Crack observed between points B16 and B17 on the dismantled pier beam

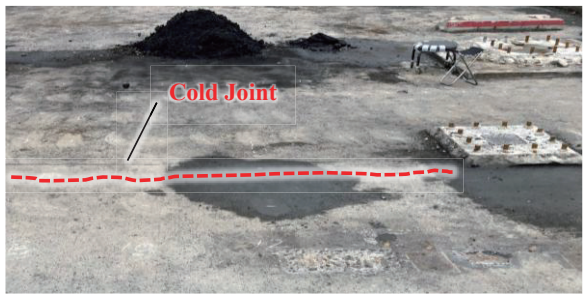
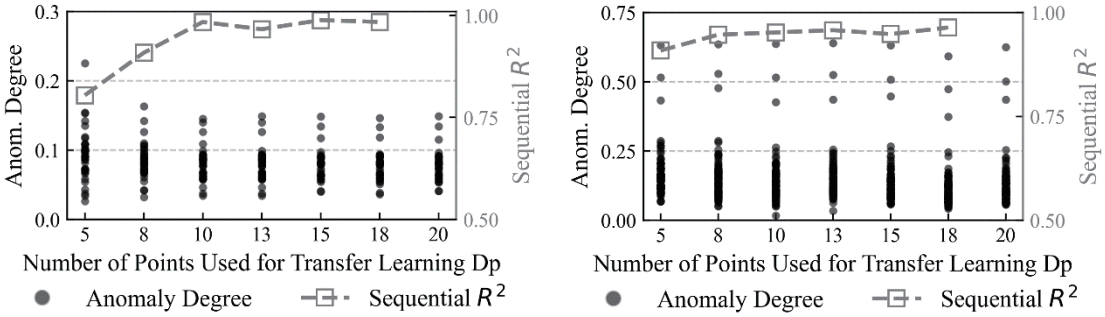
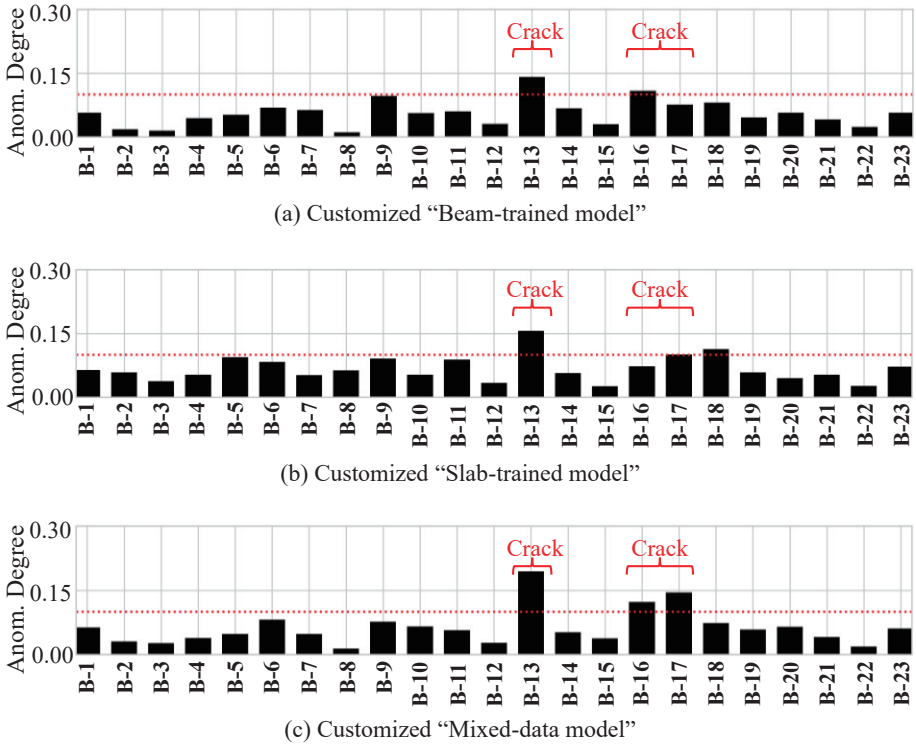


Photo 3 Cold joint observed along Grid G on the pier slab



(a) Pier beam (b) Pier slab

Fig.12 Sensitivity analysis for the number of data points Dp used in transfer learning



(a) Customized "Beam-trained model"

(b) Customized "Slab-trained model"

(c) Customized "Mixed-data model"

Fig.13 Anomaly inspection results for the pier beam using the three customized models

which exhibits more uniform structural characteristics, this stability is achieved with as few as Dp=8. Based on these findings, 10 points was adopted for all subsequent analyses.

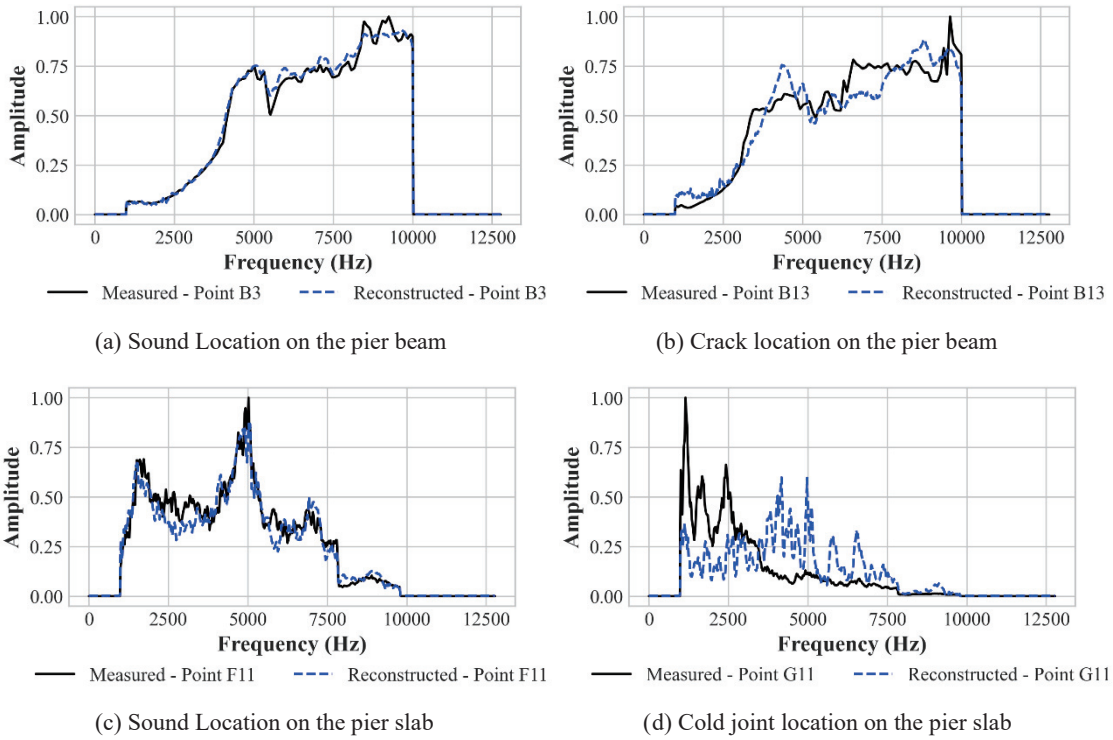
It is noteworthy that this fixed number represents 24.4% of the beam's measurement points, but only 9.2% of the slab's. The higher relative data requirement for the beam is likely attributable to the larger variability in the measured FRFs, stemming from more complex boundary conditions such as connections with piles and adjacent members.

(3) Anomaly Inspection of the Pier Beam

The three general-purpose models were customized for the pier beam inspection by applying the transfer learning procedure with the established Dp=10 dataset. The final anomaly degrees estimated by each of the three customized models are shown in **Fig.13**.

When using the "beam-trained model" as a general-purpose model, it successfully identifies the known damage, with anomaly peaks above the baseline at points B13 and B16, corresponding to the crack locations. The "slab-trained model" also identifies a relatively large anomaly at point B13, indicating it can detect the most significant crack. However, the anomaly estimated at point B16 is shown below the baseline, and incorrectly flags point B18, as anomalous, demonstrating reduced reliability.

Notably, the "mixed-data model" performs best, correctly identifying anomalies at points B13 and B16, detecting also the damage at B17. Furthermore, the contrast between the anomalous points and the sound-state baseline is larger for this model. This suggests that the mixed-data approach



(a) Sound Location on the pier beam

(b) Crack location on the pier beam

(c) Sound Location on the pier slab

(d) Cold joint location on the pier slab

Fig.14 Examples of measured and reconstructed FRFs from the on-site validation (frequency axis normalized)

allows the model to better generalize the characteristics of a sound-state FRF, leading to a more robust and sensitive damage detection capability.

To visually demonstrate the mechanism behind these results, **Fig. 14a** and **14b** present examples of the measured on-site FRFs alongside their reconstructed counterparts. For a sound location on the beam (point B-3), the customized "beam-trained model" reconstructs the measured FRF with high fidelity, resulting in a small reconstruction error. In contrast, for the cracked location (point B-13), the model fails to reproduce the spectral shape of the measured data. This discrepancy results in a relatively high reconstruction error, which is identified as an anomaly.

(4) Anomaly Inspection of the Pier Slab

A similar analysis was performed for the pier deck slab, applying the same transfer learning methodology with Dp=10 for model customization. The final anomaly degrees estimated by the three customized models are shown in **Fig.15**.

The "slab-trained model" performs effectively, identifying all points (with exception of C-2) outside the identified cold joint as being in a sound state (i.e., below the baseline), and it correctly flags the points along the cold joint at Grid G.

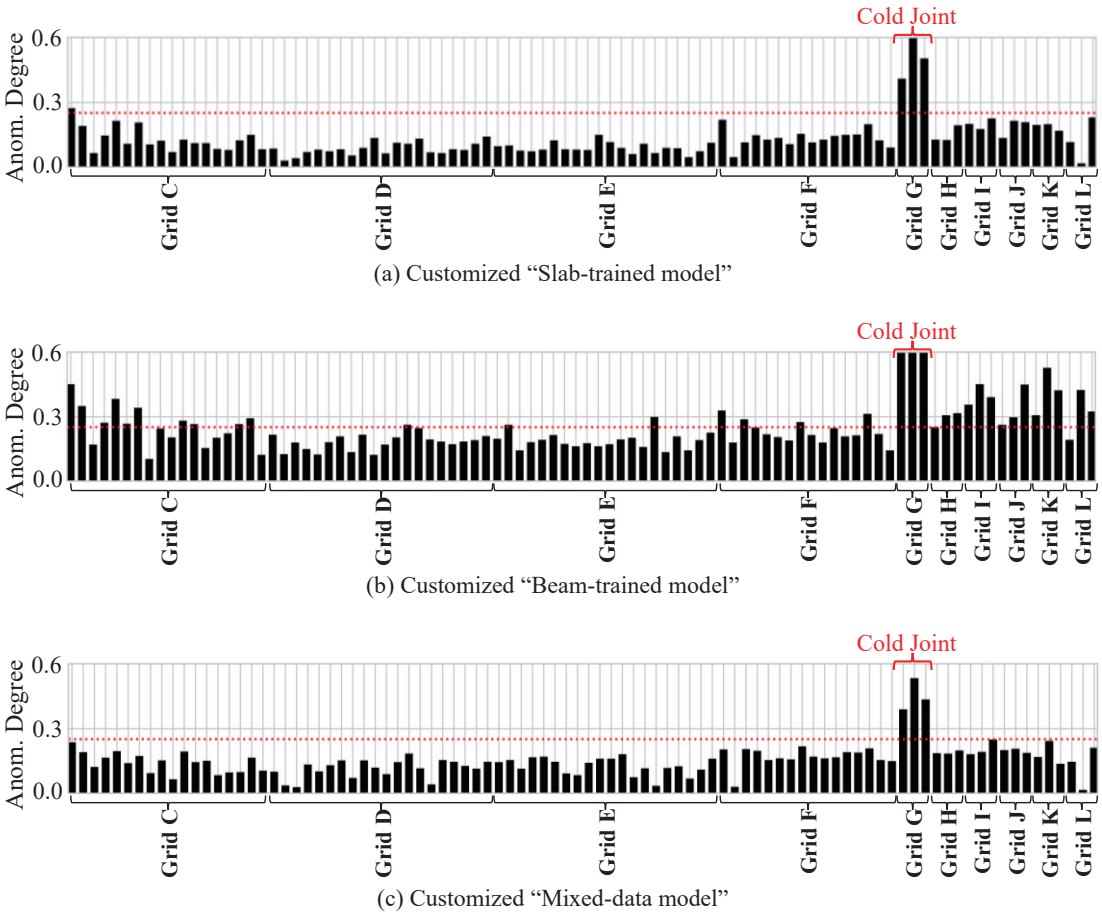
The "beam-trained model" shows the same general trend, creating a clear contrast between the sound points and the ones nearby the cold joint at Grid G; however, its performance is less precise, with several sound points registering anomaly scores that are above the baseline. Finally, the "mixed-data model" shows an improved accuracy, properly distinguishing all sound-points under the baseline, and showing a contrast on the ones nearby the cold joint at Grid G. As shown in the contour plot in **Fig.16**, the anomaly (cold joint) is clearly identified.

Examples of the measured on site FRFs alongside their reconstructed counterparts are show in **Fig. 14(c)** and **14(d)**. These plots respectively show the proper reconstruction by the customized “slab-trained model” for a sound location (point F-11), versus the large reconstruction error at the cold joint (point G-11).

(5) Discussion on the Influence of the General-Purpose Model

The on-site measurement results provide insights into the influence of the training data's structural type on the performance of the transfer learning-based anomaly detection method. When the structural type of the general-purpose model's training data matched the target structure (i.e., "beam-trained model" for the pier beam and "slab-trained model" for the pier slab), the detection accuracy was high.

Conversely, a mismatch in structural type (e.g., using the "slab-trained model" on the beam) led to a decrease in accuracy, including the incorrect



(a) Customized "Slab-trained model"

(b) Customized "Beam-trained model"

(c) Customized "Mixed-data model"

Fig.15 Anomaly inspection results for the pier slab using the three customized models

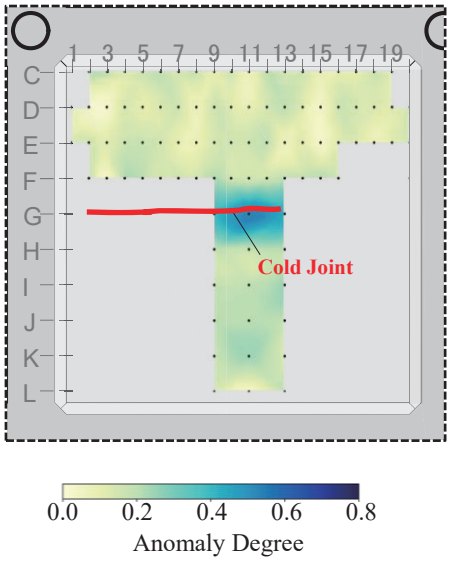


Fig.16 Target Pier Slab – Anomaly degree distribution (Customized "Mixed-data model")

identification of sound areas as anomalous. This confirms that structural similarity between the source and target is an important factor for optimal performance.

As for the "mixed-data model", for the pier beam inspection, it was more sensitive than the "beam-trained model" successfully identifying an additional damaged location (B17) while maintaining a clear distinction between sound and damaged states. For the slab inspection, it slightly improved the performance of the specialized "slab-trained model".

This suggests that a model trained on a more diverse dataset, incorporating multiple geometries, may be able to better generalize the fundamental characteristics of a "sound" FRF. This improved generalization makes the model more robust and less susceptible to variations, which becomes advantageous for practical field applications where a wide variety of structural elements is found. Therefore, while using a structural-type-specific model is a valid and reliable approach, developing a general-purpose model from a diverse library of sound-state data appears to be a promising strategy for creating a truly versatile and reliable inspection tool.

5. CONCLUSIONS

This study investigated the influence of training

data structural type on a transfer learning-based anomaly detection method, with validation performed through both full-scale test specimens, and on-site measurements. The main findings are as follows:

1. The proposed machine learning method was successfully validated on full scale laboratory specimens, where models trained on sound-state data from beam and slab specimens effectively identified and located areas of load-induced damage.
2. The on-site validation confirmed that general-purpose models can be effectively customized using transfer learning to detect damage in real-world structures. When the structural type of the training data matches the target member a high accuracy was found; while mismatched geometries may still allow for the identification of major anomalies, the reliability of the assessment is reduced.
3. A "mixed-data model" trained on a combined dataset of both beam and slab data demonstrated the most robust and versatile performance. This suggests that a more diverse training set improves the model's ability to generalize sound-state characteristics, making it more effective in practice.

The results presented represent a first milestone in confirming the practical viability of the proposed method. For widespread field application, the development of a comprehensive general-purpose model, trained on a diverse library of sound-state data from various structural shapes, appears to be a promising strategy for creating a truly versatile and reliable inspection tool. To that end, additional efforts to further increase the training data, and validation targets are required. Furthermore, although the built of a general-purpose model composed of a comprehensive dataset taken directly from different structural type specimens would be desirable, significant cost and time efforts would make it less practical. To overcome this, the feasibility of using data generated from advanced numerical simulations will be investigated. This approach could involve modeling wave propagation using methods like the Finite-Difference Time-Domain (FDTD) or Finite Element Method (FEM), coupled with fracture-analysis models such as the Rigid-Body Spring Model[11] to represent damage patterns.

REFERENCES

1) Ministry of Land, Infrastructure, Transport and Tourism, Japan: Portal site for information on measures against aging of social capital, "Current status and future of aging of social capital", https://www.mlit.go.jp/sogoseisaku/maintenance/02research/02_01.html , (accessed June 13, 2025).
2) Graybeal, B. A., Rolander, D. D., Phares, B. M., Moore, M. E. and Washer, G. A.: Reliability and Accuracy of In-Depth Inspection of Highway Bridges, *Transportation Research Record*, Vol. 1749, Issue 1, 2001.
3) Kawashima, K.: Current Status of Ultrasonic Nondestructive Inspection and Material Evaluation, *Journal of the Acoustical Society of Japan*, Vol.76, No.11, pp.640-647, 2020. (in Japanese).
4) Kanamitsu, J., Yanai, M. and Kida, T: Relationship of Internal Flaws and Temperature Distribution Changes in Slabs Based on the Thermography Method Using Asphalt Pavement Heat, *Concrete Research and Technology*, Vol.16, No.1, 2005. (in Japanese).
5) Naito, H., Kobayashi, S., Tsuchiya, Y., Sugiyama, R., Yamaguchi, K., Hayasaka, Y., Yasukawa, Y. and Suzuki, M.: Damage identification for RC bridge deck based on local vibration test, *Journal of JSCE E2*, Vol. 73, Issue 2, pp. 133-149, 2017. (in Japanese).
6) Naito H, Sugiyama R, Bolander JE: Local Vibration Testing and Damage Evaluation for RC Bridge Decks. *Journal of Structural Engineering*, Vol.146, Issue 9, 2020.
7) Naito, H., Kimoto, T., Fujioka, H., Fujikura, H. and Unjoh, S.: Seismic damage detection method for con-crete structures using vibration measurement and auto-encoder, *Journal of AI and Data Science JSCE*, Vol. 3, No. 2, pp. 134-144, 2022. (in Japanese).
8) Vargas, R., Ikeno, K., Naito, H. and Kimoto, T.: Anomaly detection in concrete structures using autoencoder models customized by transfer learning. *Journal of Structural Engineering JSCE*, Vol. 71 A, pp.592-602, 2025. (in Japanese).
9) Hinton, G. E., and Salakhutdinov, R. R.: Reducing the Dimensionality of Data with Neural Networks. *Science*, Vol. 313, Issue 5786, pp. 504–507, 2006.
10) Pan, S. J. and Yang, Q.: A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering, Vol. 22, Issue 10, pp. 1345–1359, 2009.
11) Bolander, J.E. and Saito, S.: Fracture analyses using spring networks with random geometry, Vol. 61, No. 5-6, pp. 569-591, 1998.

(Received June 30, 2025)
(Accepted August 31, 2025)