Clustering of Hammering Sounds and Identification of Defect Clusters Based on Acoustic Energy per Impact for Detection of Concrete Defects

Koki Shoda[1], Jun Younes Louhi Kasahara[2], Hajime Asama[2], Qi An[1] and Atsushi Yamashita[1]

[1] Department of Human and Engineered Environmental Studies, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8563, Japan ({shoda, an, yamashita}@robot.t.u-tokyo.ac.jp)

[2] Department of Precision Engineering, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyoku, Tokyo, 113-8656, Japan ({louhi, asama}@robot.t.u-tokyo.ac.jp)

Abstract: In this study, we propose a novel method for detecting defects in concrete structures using the sound and force signals from hammering tests. In recent years, the aging of concrete structures has become a significant issue for social infrastructure. Hammering inspection is one of the most widely used non-destructive testing methods for detecting defects in concrete. Although unsupervised methods for defect detection that do not require training machine learning models have been proposed, they essentially only allow for the clustering of hammering sounds, without indicating whether the clusters consist of healthy or defective hammering sounds. In this study, we propose a novel method to calculate the likelihood of defects in each cluster. This method identifies defect clusters based on the characteristic that defect areas generate higher sound energy during hammering tests. Experiments were conducted using multiple concrete specimens, including those with delamination and void, to validate the proposed method. The proposed method was able to identify defect clusters effectively and demonstrated high performance for defect detection.

Keywords: Infrastructure inspections, Hammering test, Machine learning, Unsupervised learning, Force sensor

1. INTRODUCTION

In recent years, the aging of concrete structures, a crucial part of social infrastructure, has become a significant concern [1]. Incidents such as the collapse of the Morandi bridge in Italy [2] and the Sasago tunnel in Japan [3] highlight the urgency of addressing the deterioration of concrete infrastructures. One non-destructive method of inspection widely used to assess the condition of concrete is the hammering test, where concrete is struck and the resulting sound is analyzed to gauge deterioration, as shown in Fig. 2.

The hammering test has become widespread due to its simplicity, yet the efficiency of inspections is increasingly demanded due to the decreasing number of skilled inspectors in developed countries, a consequence of an aging population. Considerable research has been conducted on using machine learning to automatically identify defects in concrete from the sounds produced during hammering tests [4].

Several approaches using supervised learning for defect discrimination have been proposed. These methods require preparing specific features, such as the Fourier spectrum of the hammering sound, and labels that indicate whether the state is healthy or defective. The data is then used to train classifiers, including Deep Neural Networks [5] and Ensemble Models [6]. However, due to variations in concrete composition and curing conditions, the hammering sounds differ [7], necessitating on-site training. Training defect discriminator requires a substantial amount of labeled hammering sounds by skilled inspectors, which significantly increases their workload.

An alternative approach using transfer learning has been suggested [8]. This method involves fine-tuning a model trained under controlled condition at different sites or on concrete specimens to the conditions at the actual site. However, this approach still requires labeled hammering sounds for fine-tuning.

Fig. 1. An inspector conducting a hammering test.

Consequently, an unsupervised approach that does not require training a discriminator has been proposed [9]. This approach utilizes sounds and positions to perform clustering of the hammering sounds. Furthermore, it identifies the cluster with the largest number of data points as the healthy cluster, based on the assumption that the majority of hammering sounds collected during inspections are from healthy concrete. The challenge with this unsupervised method is that it does not inherently reveal the attributes of the grouped data. For instance, if the collected data contains more defect sounds than healthy ones, or if multiple healthy clusters exist, this method [9] may fail to identify defects.

Given these considerations, the development of an unsupervised method for defect discrimination that can identify defects even when there are fewer healthy hammering sounds or multiple healthy clusters, and does not require labeling at the inspection site, is extremely important yet has not been realized. Therefore, the objective of this study is to develop an unsupervised method for defect discrimination capable of identifying defective

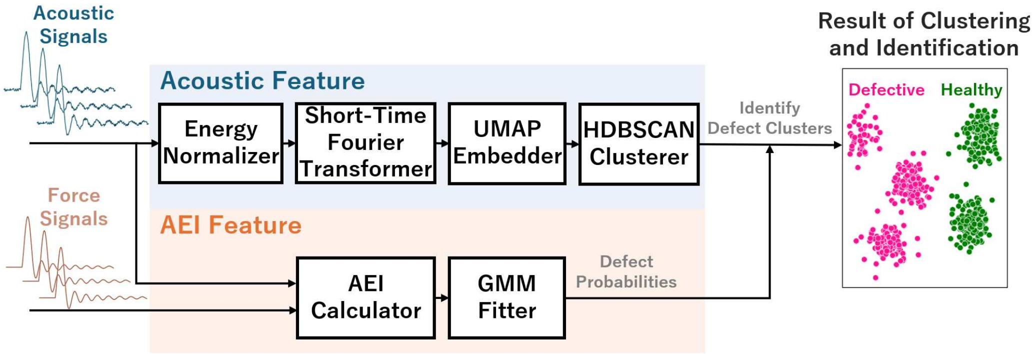


Fig. 2. Overview of the proposed method.

clusters. To address this challenge, we propose a novel method that calculates the likelihood of defects for each cluster based on the characteristic that defect areas produce higher sound energy during hammering tests.

2. PROPOSED METHOD

2.1 Concept

The defects targeted for detection in hammering tests are voids within concrete, such as delamination and spalling, caused by concrete deterioration. When voids exist without being filled, they generate unique vibration modes depending on the size and shape of the void. Therefore, regions with defects are more likely to excite vibrations [10], resulting in larger acoustic energy during hammering compared to healthy regions. In this study, we propose a method that calculates the likelihood of defects for each cluster based on the characteristic that defect areas produce higher sound energy during hammering tests.

The energy of hammering sounds depends not only on the internal state of the concrete but also on the strength of the impact. To address this issue, we use an impact hammer equipped with a force sensor to measure the energy of the impact. In this study, we propose Acoustic Energy per Impact (AEI), normalized by the energy of the impact. On the other hand, compared to the AEI, the frequency features of the hammering sound are more robust and have been widely used in previous research [11], [12]. Therefore, we propose a method that combines AEI with frequency features to accurately perform clustering while enabling the calculation of defect likelihood for each cluster.

Overview of the proposed method is shown in Fig. 2. Initially, through measurements using a microphone and force sensor, sets of sound and force signals are acquired. The proposed method consists of two processes: one for processing acoustic features and another for processing the AEI.

In the processing of acoustic features, clustering is performed based on the characteristics of the hammering sound. In the processing of AEI, AEI is used to calculate the defect probability for each hammering within the clusters formed by the acoustic features. Finally, the clusters are identified as healthy or defective based on the average of defect probability within each cluster.

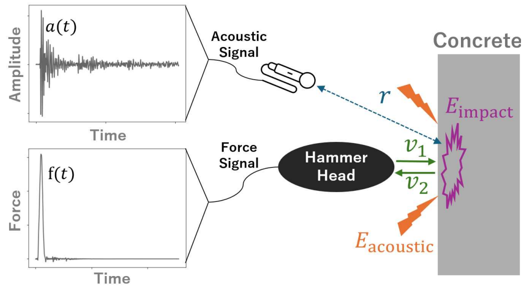


Fig. 3. Diagram of the time series signals of sound and force in hammering tests.

2.2 Acoustic Energy per Impact

Figure 3 shows diagram of the time series signals of sound and force in hammering tests, where v_1 is the velocity just before the hammer head collides, v_2 is the velocity just after the collision, r is the distance from the hammering position to the microphone, $a(t)$ is the amplitude of the sound pressure and $f(t)$ is the force, at time t, respectively. In the hammering test, strikes are generally performed perpendicular to the surface, so the velocity of the hammer is considered only in the component that is perpendicular to the surface.

First, to determine the energy E_{impact} imparted to the concrete by an impact, we can use the law of energy conservation, the law of momentum conservation, and the coefficient of restitution e:

$$\text{Energy conservation: } E_{\text{impact}} = \frac{1}{2}mv_1^2 - \frac{1}{2}mv_2^2, \tag{1}$$

$$\text{Momentum conservation: } \int f(t)dt = m(v_2 - v_1), \tag{2}$$

$$\text{Restitution coefficient: } e = \frac{v_2}{v_1}, \tag{3}$$

where m is the mass of the hammer head.

By eliminating v_1 and v_2 from these equations and approximating e as constant, we find that E_{impact} is proportional to the square of the impulse:

$$E_{\text{impact}} \propto \left(\int f(t)dt\right)^2. \tag{4}$$

Next, we determine the acoustic energy generated by the concrete E_{acoustic} . Appoloximating that the distance between the microphone and the hammering position r is constant, E_{acoustic} is proportional to the acoustic energy picked up by the microphone, $E_{\text{microphone}}$. Furthermore, $E_{\text{microphone}}$ is proportional to the square of the amplitude of the sound pressure $a(t)$:

$$E_{\text{acoustic}} \propto \int a(t)^2 dt. \tag{5}$$

In regions with voids, such as defects, the energy imparted to the concrete by the impact is more easily converted into vibration and the resulting acoustic energy. Therefore, we define AEI as the ratio of the acoustic energy generated by the concrete E_{acoustic} to the energy imparted to the concrete by the impact E_{impact}:

$$\text{AEI} = \frac{E_{\text{acoustic}}}{E_{\text{impact}}} \propto \frac{\int a(t)^2 dt}{\left(\int f(t) dt\right)^2}. \tag{6}$$

2.3 Processing of Acoustic Feature

In the processing of acoustic features, the first step involves eliminating variations in striking force and normalizing the energy of each hammering sound to extract its acoustic characteristics. Next, a Short-Time Fourier Transform (STFT) is performed to extract the time-frequency domain features of each hammering sound. STFT is a common feature in machine learning for audio signals, and the frequency-decomposed signals facilitate feature recognition.

However, the high-dimensional features transformed by STFT can reduce clustering accuracy due to the phenomenon known as the concentration of measure [13]. To address this issue, Uniform Manifold Approximation and Projection (UMAP) [14] is used to embed high-dimensional features into a lower-dimensional space. UMAP is a dimensionality reduction technique noted for its excellent balance in maintaining both local and global structures, and has recently gained attention in the field of machine learning. Another significant advantage of dimensionality reduction is that it greatly facilitates the interpretation of clustering results through visualization, thus simplifying the verification of system integrity and manual tuning of hyperparameters.

On the other hand, in the embedding space of UMAP, clusters of irregular shapes tend to form. Clustering methods with constraints on cluster shapes, such as k-means or Gaussian Mixture Models, struggle to adequately cluster such data groups. Therefore, this study employs Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) [15] for clustering. HDBSCAN is a density-based clustering method that can form clusters of flexible shapes.

Conversely, Clustering merely groups similar hammering sounds and cannot identify which clusters are defective. Therefore, we propose a identification method for defect clusters based on AEI.

2.4 Processing of AEI

AEI is easy to calculate from only force and acoustic signals by Eq. (6), making it highly convenient and suitable for automated inspection scenarios. However, because there are factors that influence the magnitude of the AEI, relying solely on threshold processing of the AEI to discriminate between defects and healthy conditions lacks reliability. These influencing factors include variations and diversity in the angle of hammering, the coefficient of restitution, the distance between the microphone and the hammering position, the environment of sound reverberation, and acoustic noise at inspection sites.

Hence, there are many factors influence the magnitude of AEI, with uncertainties and unknown variations overlapping. This overlapping is termed *process noise*, and assuming it linearly combines, it can be hypothesized from the central limit theorem that *process noise* follows a normal distribution.

Therefore, AEIs are fitted using Gaussian Mixture Model (GMM). In GMM, the probability distribution $p(x)$ of the feature x is represented as a linear combination of normal distributions N:

$$p(x) = \sum_{k=1}^{K} \pi_k N(x|\mu_k, \sigma_k), \tag{7}$$

where, K is the number of normal distributions, and π_k, μ_k, and σ_k represent the mixture coefficient, mean, and standard deviation of the k-th distribution, respectively. The parameters of each normal distribution–π_k, μ_k, and σ_k–are optimized by maximizing the log-likelihood using the EM algorithm. Additionally, since the number of normal distributions K is unknown, the optimal K is determined by using a grid search to minimize the Bayesian Information Criterion [16].

Since a healthy hammering sound has a lower AEI, the normal distribution with the smallest μ_k is designated as the healthy one. Therefore, the probability of a defect for data with feature value x is calculated as follows:

$$\text{Defect Probability} = 1 - \frac{\pi_i N(x|\mu_i, \sigma_i)}{p(x)}, \tag{8}$$

$$i = \arg\min_k \mu_k. \tag{9}$$

Using the defect probability calculated by AEI, identification of defect clusters formed by sound characteristics is conducted. The average defect probability within each cluster is calculated, and clusters where this average exceeds 0.5 are identified as defect clusters.

3. EXPERIMENT

3.1 Experimental Setup

To validate the effectiveness of the proposed method, experiments were conducted. An impact hammer was used to strike both the healthy and defective areas of concrete specimens, and the force and acoustic signals obtained were used for defect discrimination. Figure 4

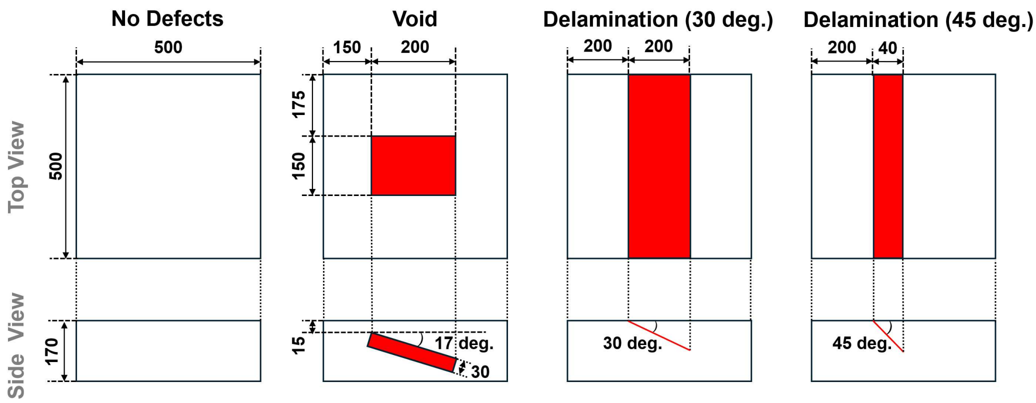


Fig. 4. Schematic of the concrete specimens used in the experiment, with all distances in mm. Their body size are the same and the artificial defects are shown in red.

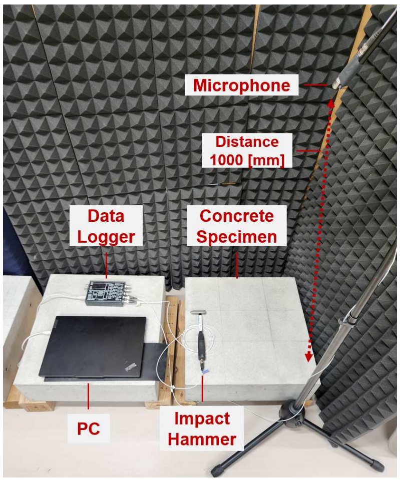


Fig. 5. Setup of experimental equipments.

shows the concrete specimens used in the experiments, which included artificial defects. The concrete was uniformly struck, obtaining 400 samples of healthy sounds and 600 samples of defective sounds from each specimen.

Additionally, the experimental setup was arranged as shown in Fig. 5. A microphone (PCB model 377B02) was mounted on a stand to capture the hammering sounds generated by the impact hammer (PCB model 086C03), which can measure force in only one axis, and the sounds, along with force signals, were recorded simultaneously using a data logger (MC model DT9837B).

3.2 Parameter Setting

The sampling frequency was set at the maximum limit of the data logger, 100 kHz, with the measurement duration for force signals at 5.12 ms (512 samples) and for audio signals at 40.96 ms (4096 samples). The window size for the STFT was 256 samples, with an overlap of 128 samples, and a Hamming window was used as the window function. The dimensionality of UMAP was set to two dimensions. The hyperparameters for HDBSCAN, namely the minimum cluster size and the minimum number of samples S_m, were set to the same value.

HDBSCAN includes a parameter called the minimum number of samples S_m, which implicitly specifies the

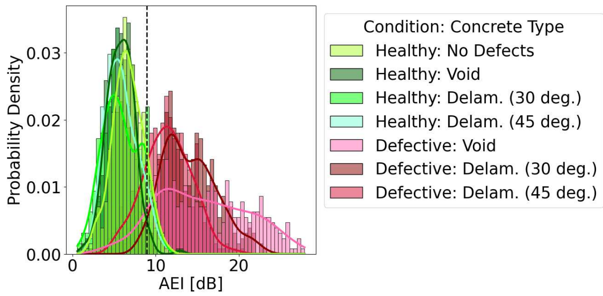


Fig. 6. Histogram and KDE of AEI for ground truth. The dashed black line represents the AEI threshold when the defect probability reaches 0.5. AEI is significant in terms of relative magnitude, therefore, the smallest value in the collected data is set as the baseline 0 dB.

number of clusters. In this study, to simplify the parameter settings, S_m was determined through grid search to maximize the accuracy of defect discriminaton and was set at 64.

4. RESULTS

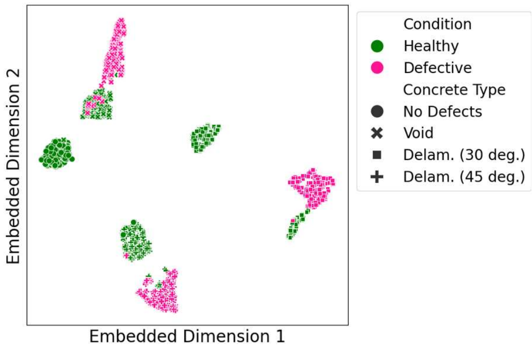
To verify the effectiveness of the proposed method when there are more defect sounds than healthy ones, we collected a total of 3400 samples—400 healthy sounds and 600 defective sounds from each type of concrete, amounting to $400 \times 4 + 600 \times 3$. We then evaluated the performance of defect discrimination based on these 3600 samples.

The histogram and Kernel Density Estimation (KDE) of the AEI calculated from the collected data are shown in Fig. 6. Although there is some overlap between the distributions of healthy and defective samples, they are clearly separated. This demonstrates that the AEI is useful in distinguishing between healthy and defective states in concrete, and it aids in identifying defect clusters among the clustered acoustic features.

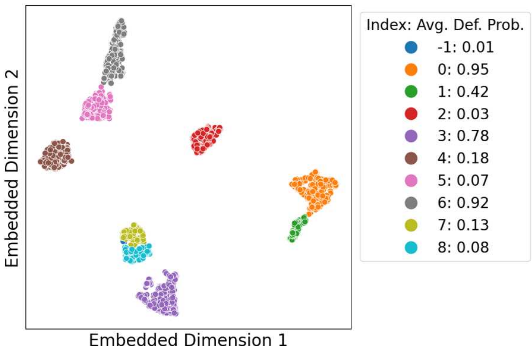
Table 1 shows the performance of defect discrimination when using only the defect probability calculated from AEI versus the proposed method that combines AEI and acoustic features. Given that the methods evaluated

Table 1. Performance metrics with mean $\pm$ standard deviation for each method. The threshold of AEI is set to the value where the defect probability in Eq. (8) reaches 0.5 .

	AEI Threshold	Proposed Method	Proposed Method (w/o STFT)	Proposed Method (w/o UMAP)
Accuracy	0.895 ± 0.003	**0.973** $\pm$ 0.021	0.485 ± 0.136	0.594 ± 0.066
Precision	0.863 ± 0.009	**0.952** $\pm$ 0.044	0.333 ± 0.221	0.535 ± 0.088
Recall	0.897 ± 0.004	**0.988** $\pm$ 0.003	0.358 ± 0.300	0.436 ± 0.138
F1 Score	0.880 ± 0.003	**0.969** $\pm$ 0.023	0.337 ± 0.245	0.471 ± 0.104



(a) Ground truth, where healthy and defective points are plotted in green and pink, respectively, and the style of the points varies according to the type of concrete.



(b) Result of clustering, where the index of each cluster and defect probability are shown in the legend. Points of data identified as noise are assigned a cluster index of -1.

Fig. 7. Scatter plots of the embedded space for hammering sounds.

are sensitive to initial parameter settings, we conducted each method 32 times with varied initial values to ensure a comprehensive evaluation. The proposed method, which combines AEI and acoustic features for defect discrimination, demonstrated superior performance compared to methods using only AEI and versions of the proposed method with some features omitted.

An example of the clustering results obtained with the proposed method is shown in Fig. 7. Even in cases where multiple healthy clusters exist or there are more defect sounds than healthy sounds, the proposed method is able to identify healthy and defective clusters. However, in cases of concrete with void and concrete with 30 degrees delamination, the clusters for defective and healthy conditions are located in close proximity, and there is some mixing between them. This is due to the acoustic features being similar near the boundaries between defective and healthy areas of the concrete where the hammering occurs.

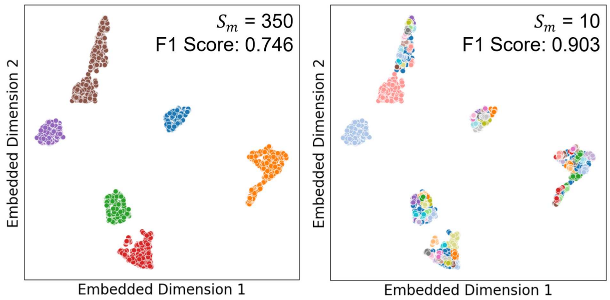


Fig. 8. Clustering results and F1 score representing performance of defect discrimination when changing S_m (the parameter that implicitly specifies the number of clusters for HDBSCAN). The results are plotted with different colors for each cluster. When S_m=10, 81 clusters were formed.

5. DISCUSSION

Figure 8 presents the clustering results and F1 scores when the parameter S_m is varied extremely. When S_m is too large, the clusters for healthy and defective conditions merge into a single cluster, leading to incorrect discrimination of healthy and defective states and resulting in a low F1 score. Conversely, when S_m is too small, although the clusters for healthy and defective conditions are well separated, the averaging of the defect probability calculated by AEI is insufficient, also resulting in a low F1 score.

In clustering methods, there is always a hyperparameter, implicit or explicit, that determines the number of clusters as in the case of S_m in HDBSCAN. This hyperparameter significantly influences the performance of clustering, hence it requires careful adjustment. There are methods for determining the number of clusters using performance metrics such as information criteria or silhouette scores. However, due to the diversity of data and the characteristics of these metrics, their application can be challenging.

Setting parameters to maximize accuracy using ground truth, as in this study, compromises the benefits of unsupervised methods that do not require labeled data. Therefore, it may be necessary to combine this approach with weakly-supervised methods, such as those proposed by Louhi et al. [17], where instead of identifying each hammering sound as defective or healthy, only a simple and low-burden human judgment about the similarity intervenes, based on the representative sound of the cluster. For example, it is conceivable that inspectors could listen to representative hammering sounds from each cluster and adjust the number of clusters by consolidating them.

6. CONCLUSION

In this study, we proposed a novel unsupervised method for defect detection that identifies defect clusters in hammering inspections. Based on the characteristic that striking defect areas produces higher sound energy, we developed a method to calculate the likelihood of defects for each cluster. To validate the applicability of the proposed method, we used four types of concrete specimens, including those with delamination and void defects, in our experiments.

The proposed method demonstrated high performance for defect discrimination and the ability to identify defective clusters. Future work should focus on automating the setting of clustering parameters and enhancing performance by integrating weakly-supervised methods [17]. By involving only a simple and low-burden human judgment in the process, such as listening to representative hammering sounds from each cluster, inspectors could potentially refine the clustering results and improve defect detection accuracy.

ACKNOWLEDGEMENT

This work was supported in part by JSPS KAKENHI under Grant Number JP21K17829.

REFERENCES

[1] K. Breugel, "Societal burden and engineering challenges of ageing infrastructure," *Procedia Engineering*, vol. 171, pp. 53–63, 2017.

[2] M. Morgese, F. Ansari, M. Domaneschi, and G. Cimellaro, "Post-collapse analysis of morandi's polcevera viaduct in genoa italy," *Journal of Civil Structural Health Monitoring*, vol. 10, 2020.

[3] Y. Fujino and D. M. Siringoringo, "Recent research and development programs for infrastructures maintenance, renovation and management in japan," *Structure and Infrastructure Engineering*, vol. 16, no. 1, pp. 3–25, 2020.

[4] P. Chun, J. Dang, S. Hamasaki, R. Yajima, T. Kameda, H. Wada, T. Yamane, S. Izumi, and K. Nagatani, "Utilization of unmanned aerial vehicle, artificial intelligence, and remote measurement technology for bridge inspections," *Journal of Robotics and Mechatronics*, vol. 32, no. 6, pp. 1244–1258, 2020.

[5] T. Fukumura, H. Aratame, A. Ito, M. Koike, K. Hibino, and Y. Kawamura, "A study on hammering test using deep learning," in *Proceedings of the 2020 IEEE 15th International Conference of System of Systems Engineering*, 2020, pp. 261–264.

[6] H. Fujii, A. Yamashita, and H. Asama, "Defect detection with estimation of material condition using ensemble learning for hammering test," in *Proceedings of the 2016 IEEE International Conference on Robotics and Automation*, 2016, pp. 3847–3854.

[7] K. Ushiroda, J. Y. Louhi Kasahara, A. Yamashita, and H. Asama, "Multi-modal classification using domain adaptation for automated defect detection based on the hammering test," in *Proceedings of the 2022 IEEE/SICE International Symposium on System Integration*, 2022, pp. 991–996.

[8] T. Fukumura, H. Aratame, A. Ito, M. Koike, K. Hibino, and Y. Kawamura, "An efficient learning method for sound classification using transfer learning for hammering test," in *Proceedings of the 2020 IEEE SENSORS*, 2020, pp. 1–4.

[9] J. Y. L. Kasahara, H. Fujii, A. Yamashita, and H. Asama, "Fuzzy clustering of spatially relevant acoustic data for defect detection," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2616–2623, 2018.

[10] N. Yasuda, N. Misaki, Y. Shimada, and D. Yamaoka, "Detection and characteristics estimation of defects in concrete structures using laser ablation-induced vibration," *Tunnelling and Underground Space Technology*, vol. 103, p. 103460, 2020.

[11] J. Y. L. Kasahara, H. Fujii, A. Yamashita, and H. Asama, "Clustering of spatially relevant audio data using mel-frequency cepstrum for diagnosis of concrete structure by hammering test," in *Proceedings of the 2017 IEEE/SICE International Symposium on System Integration (SII)*, 2017, pp. 787–792.

[12] H. Fujii, A. Yamashita, and H. Asama, "Defect detection with estimation of material condition using ensemble learning for hammering test," in *Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 3847–3854.

[13] P. Hall, J. S. Marron, and A. Neeman, "Geometric representation of high dimension, low sample size data," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 67, pp. 427–444, 2005.

[14] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:1802.03426*, 2020.

[15] L. McInnes, J. Healy, and S. Astels, "hdbscan: Hierarchical density based clustering," *The Journal of Open Source Software*, vol. 2, 2017.

[16] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, pp. 461–464, 1978.

[17] J. Y. L. Kasahara, A. Yamashita, and H. Asama, "Positive weak supervision quality increase by consolidation for acoustic defect detection in concrete structures," in *Proceedings of the 2021 IEEE/SICE International Symposium on System Integration (SII)*, 2021, pp. 1–5.