INTEGRATION OF ANOMALY MACHINE SOUND DETECTION INTO ACTIVE NOISE CONTROL TO SHAPE THE RESIDUAL SOUND

Chuang Shi, Mengjie Huang, Huitian Jiang, Huiyong Li

School of Information and Communication Engineering
University of Electronic Science and Technology of China
2006 Xiyuan Avenue, Chengdu, China

ABSTRACT

An active noise control (ANC) system generates a secondary sound to destructively interfere with the undesirable noise. Existing ANC algorithms are mainly designed to minimize the power of the residual sound, with few considerations to the listening experience. This results in a pressing issue in practice whereby the residual sound is perceived to be different from the undesirable noise. When the ANC system is deployed to reduce the noise level in a factory environment, workers may feel strange because they are used to detecting the anomaly machine sound by their auditory perception. In order to solve this problem, this paper proposes to integrate anomaly sound detection (ASD) into the ANC system in order for the residual sound to represent the same machine status as the original machine noise. The ASD module is used to simulate human judgement. A homothety constrained ANC algorithm is developed to synchronously reduce the sample-wise power and keep the segment-wise machine status of the residual sound. The experiment results validate the effectiveness of the homothety constrained ANC algorithm in noise reduction, and the subjective test results show that the ASD-integrated ANC system results in less confusing perceptions of the residual sound.

Index Terms— Active noise control, anomaly machine sound detection, residual sound, homothety constrained FxLMS

1. INTRODUCTION

Noise pollution is an inescapable issue due to the inhabition of human beings and the environment. Prolonged exposure to loud noises may cause health problems for people, such as hearing loss, cardiovascular disease, sleep deprivation, and physiological stress [1]. Passive noise control (PNC) measures are developed to isolate the noise source, block the noise propagation path, and shield the receiver. They are efficient in abating high-frequency noise, but become bulky and costly when noise frequencies are low [2]. For example, in a factory, machines can be placed in an enclosed space to control their noise at source, but the complexity of ventilation and maintenance increases [3]. Alternatively, workers can wear hearing protection earbuds, which consequently halts all verbal communications during their working hours [4].

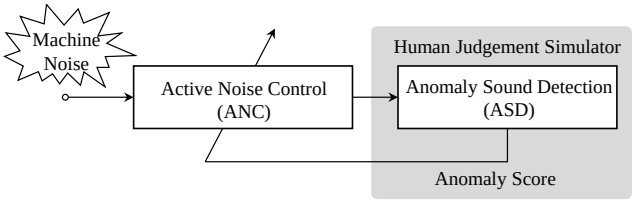

Fig. 1. Block diagram of the ANC system integrated with an ASD module to simulate human judgement.

Due to several restrictions of PNC, active noise control (ANC) gained popularity in the past decade [5]. Noise canceling headphones have become the most eye-catching product in consumer electronics [6]. An ANC system generates a secondary sound to destructively interfere with the undesirable noise. There are two fundamental ANC structures. They are the feedforward and feedback structures. The feedforward ANC structure consists of reference, control and error signals, while the feedback structure consists of only error and control signals. Owing to the reference signals, the feedforward structure is preferable for broadband noise control applications [7, 8].

The generation of the secondary sound in an ANC system can be implemented by adaptive filtering algorithms [9, 10]. Existing ANC algorithms are mainly designed to minimize the power of the error signal, in order for quiet zones to be formed around error or virtual error microphones [11]. However, as human ears have frequency-dependant sensitivities, the perceptual loudness is not proportional to the sound pressure level (SPL). Therefore, psychoacoustic ANC algorithms are investigated to achieve noise equalization, sound profiling, or sound quality control that make the residual sound seem softer without further reducing its SPL [12, 13, 14].

This manuscript is prepared based on the research work supported by the National Natural Science Foundation of China and the Civil Aviation Administration of China (Joint Grant No. U1933127).

When an ANC system is deployed to reduce the noise level, the residual sound is often perceived to be drastically different from the undesirable noise. This causes a pressing issue in a factory environment as workers may feel strange because they are used to detecting the anomaly machine sound using their auditory perception. The ANC system protects hearings of the workers, but also confuses their judgments wtih regards to the machine status. In order to solve this problem, this paper integrates an anomaly sound detection (ASD) module into the ANC system to simulate human judgement, as shown in Fig. 1. A homothety constrained FxLMS algorithm is thereafter proposed to synchronously reduce the sample-wise power and keep the segment-wise machine status of the residual sound, in order for the residual sound to represent the same machine status as the original machine noise.

The work presented here has focused on the formulation of the homothety constrained FxLMS algorithm and the design of an ASD-integrated ANC system, which takes advantage of the emerging field of machine hearing [15]. Recent studies have demonstrated that machine learning algorithms substantially outperformed human listeners in the task of sound classification [16, 17]. Previous works using machine learning algorithms in ANC consider either selection or generation of the control filter by neural networks [18, 19]. While the present study is related to recent attempts in substituting human listeners with machine hearing systems, it feedbacks new information for the control filter to update, which was not considered in earlier studies.

2. HOMOTHETY CONSTRAINED FXLMS ALGORITHM

As illustrated in Fig. 2, a single channel feedforward ANC system receives the reference signal $x(n)$ and generates the control signal $y(n)$ to minimize the power of the error signal $e(n)$. The primary path and the secondary path to the error microphone are denoted as $\mathbf{p}(n)$ and $\mathbf{s}(n)$, respectively. The objective function of the classic FxLMS algorithm is given by

$$J(n) = E\left[e^2(n)\right]. \tag{1}$$

The update equation of the control filter obtained on this basis is written as

$$\mathbf{w}(n+1) = \mathbf{w}(n) - 2\mu e(n)\hat{\mathbf{r}}(n), \tag{2}$$

where $\mathbf{w}(n)$ is the weight coefficient vector; μ is the step size; and $\hat{\mathbf{r}}(n)$ is the filtered reference signal vector. $\hat{\mathbf{r}}(n)$ is further written as

$$\hat{\mathbf{r}}(n) = [\hat{r}(n), \hat{r}(n-1), \cdots, \hat{r}(n-L_w+1)]^T, \tag{3}$$

where

$$\hat{r}(n) = [x(n), x(n-1), \cdots, x(n-L_s+1)]^T \cdot \hat{\mathbf{s}}(\mathbf{n}); \tag{4}$$

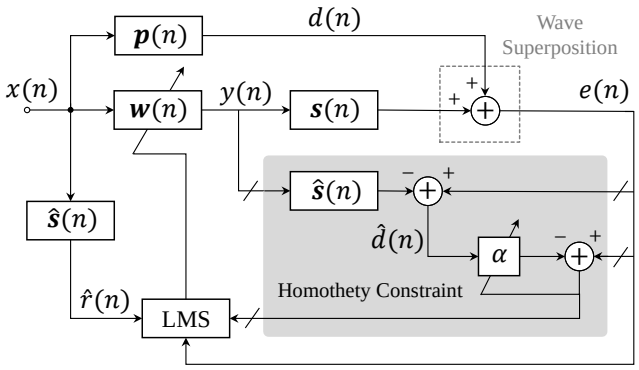

Fig. 2. Block diagram of the homothety-constrained FxLMS algorithm.

$\hat{\mathbf{s}}(\mathbf{n})$ is the secondary path model; L_w and L_s are the length of the control filter and the length of the secondary path model, respectively.

The classic FxLMS only reduces the sample-wise power of the error signal. In order to keep the segment-wise machine status of the residual sound at the same time, the objective function of the homothety constrained ANC algorithm is modified to

$$J(n) = E\left\{e^2(n) + \lambda\left[e(n) - \alpha \cdot \hat{d}(n)\right]^2\right\}, \tag{5}$$

where $\hat{d}(n)$ is the estimate of the disturbance signal $d(n)$; α and λ are adaptive hyper-parameters, namely the homothety ratio and the Lagrange multiplier, respectively. The additional term in (5), as compared to (1), is called the homothety constraint. It is integrated into the block diagram of the classic FxLMS algorithm to form the homothety constrained FxLMS (HC-FxLMS) algorithm, as shown in Fig. 2.

The update equation of the control filter is thereafter written as

$$\mathbf{w}(n+1) = \mathbf{w}(n) - 2\mu\left[e(n) - \frac{\lambda}{1+\lambda}\alpha \cdot \hat{d}(n)\right]\hat{\mathbf{r}}(n), \tag{6}$$

where both α and λ are updated in a segment-wise way, catering for the ASD module.

In the k-th segment, the homothety ratio is firstly estimated by

$$\alpha^{(k)} = \arg\min_{\alpha}\left\{\left\|\mathbf{e}^{(k)} - \alpha\hat{\mathbf{d}}^{(k)}\right\|_2\right\}, \tag{7}$$

where $\mathbf{d}^{(k)}$ is the disturbance signal estimate vector, *i.e.*

$$\hat{\mathbf{d}}^{(k)} = [\hat{d}(n), \hat{d}(n-1), \cdots, \hat{d}(n-L_f+1)]^T; \tag{8}$$

$\mathbf{e}^{(k)}$ is the error signal vector, *i.e.*

$$\mathbf{e}^{(k)} = [e(n), e(n-1), \cdots, e(n-L_f+1)]^T; \tag{9}$$

and L_f is the segment size. Therefore, the segment index is associated with the time index by

$$k = \left\lfloor \frac{n}{L_f} \right\rfloor - 1. \tag{10}$$

The Lagrange multiplier is subsequently calculated by

$$\lambda^{(k+1)} = \beta h^{(k)}, \tag{11}$$

where β is a scaling factor; and $h^{(k)}$ is the minimum in (7), given by

$$h^{(k)} = \left\| \mathbf{e}^{(k)} - \alpha^{(k)} \hat{\mathbf{d}}^{(k)} \right\|_2. \tag{12}$$

The Lagrange multiplier determines the noise reduction performance of the HC-FxLMS algorithm. Adopting a variable Lagrange multiplier and an early stopping strategy allows the HC-FxLMS algorithm to converge to a pre-defined noise reduction level.

3. ANOMALY SOUND DETECTION MODULE

The ASD module used in this paper belongs to the category of unsupervised anomaly detection in sound, where "anomaly" is defined as the patterns in data that do not conform to expected "normal" behavior [20]. It is dedicated to the detection of unknown anomalous machine sound without training data of anomalous machine sound. In order to achieve this target, an auxiliary task of machine type classification is carried out by a residual neural network (ResNet) [21, 22].

The architecture of the ASD module is shown in Fig. 3. The log-mel spectrum of one segment of the residual sound is the input of the ResNet. After 1 convolution layer, 9 residual blocks, global pooling and softmax, a machine type prediction is made. The anomaly score is calculated by

$$Anomaly\ Score = \sqrt{\sum_{i=1}^{n} \left(\mathbf{T}^{(truth)}(i) - \mathbf{T}^{(predict)}(i)\right)^2}, \tag{13}$$

where $\mathbf{T}^{(predict)}$ denotes the machine type prediction; $\mathbf{T}^{(truth)}$ denotes the ground truth; and n denotes the dimension of the machine type label, which is given by the number of machine types in the training dataset plus one augmented "unknown" type. The spectrum distortion method is adopted to transform an original sound segment into the "unknown" type of machine sound.

The performance of the ASD module is provided by its receiver operating characteristic (ROC) curve, which exhibits the trade-off relationship between the true positive rate (TPR) and false positive rate (FPR). The TPR is the proportion of anomaly sound segments that are correctly identified, and the FPR is the proportion of normal sound segments that are mistaken as anomalies. Conventionally, the ROC curve is quantized by the area under the curve (AUC), which ranges from 0.5 (neutral) to 1 (best).

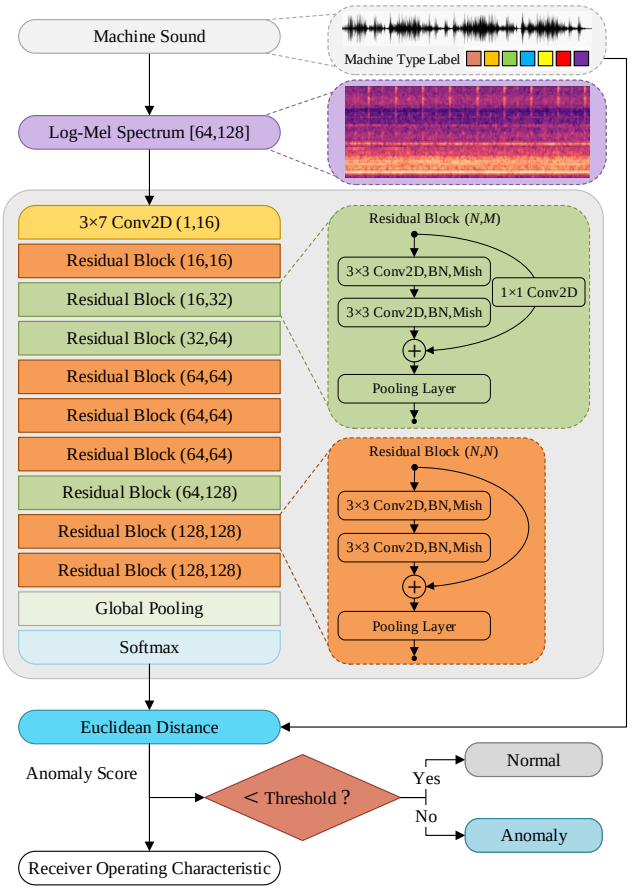

Fig. 3. Architecture of the ASD module based on auxiliary machine type classification.

4. EXPERIMENTAL RESULTS

MIMII and ToyADMOS datasets are used to train the ASD module and validate the effectiveness of the ASD-integrated ANC system [23, 24]. In total, there are six machine types. They are the fan, pump, slide rail, valve, toy car, and toy conveyor. There are 4000 sound clips of each type of machine in the training subset and 800 sound clips of each type of machine in the evaluation subset. The implementation of the ASD module is carried out with the adaptive moment estimation optimizer and the binary cross entropy loss.

The primary path and secondary path for ANC simulations are measured in a laboratory environment at a sampling rate of 16 kHz, which is in consistency with the datasets. The length of the control filter is set to 800 taps. Figure 4 shows the homothety ratio curves of the FxLMS and HC-FxLMS algorithms on six types of machine noise with respect to noise reduction (NR). The markers indicate the measurement data and the solid lines present least-squares fitting results. In theory, there is an upper bound, given by $\alpha = 10^{-(NR/20)}$. Figure 4 demonstrates that the HC-FxLMS algorithm obtains

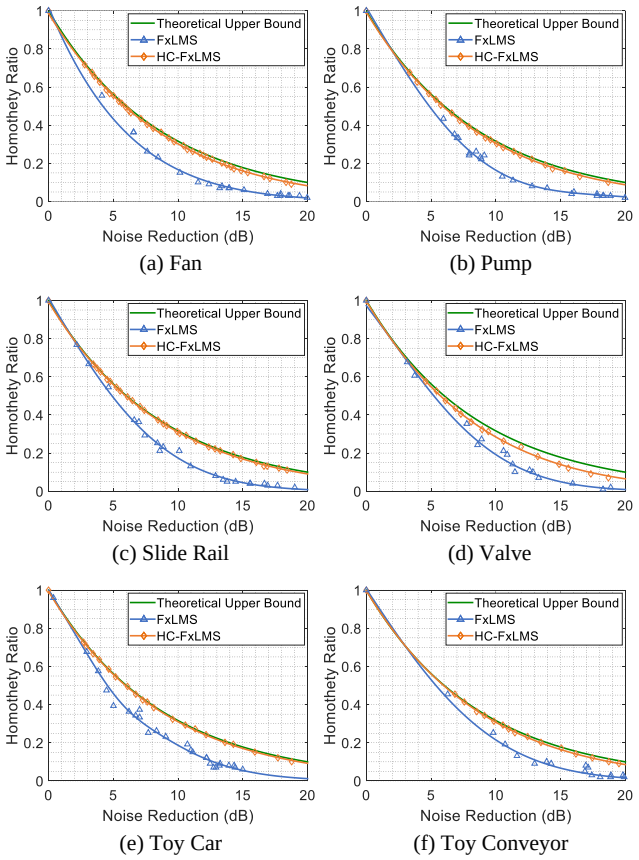

Fig. 4. Homothety ratio curves of the FxLMS and HC-FxLMS algorithms on six types of machine noise.

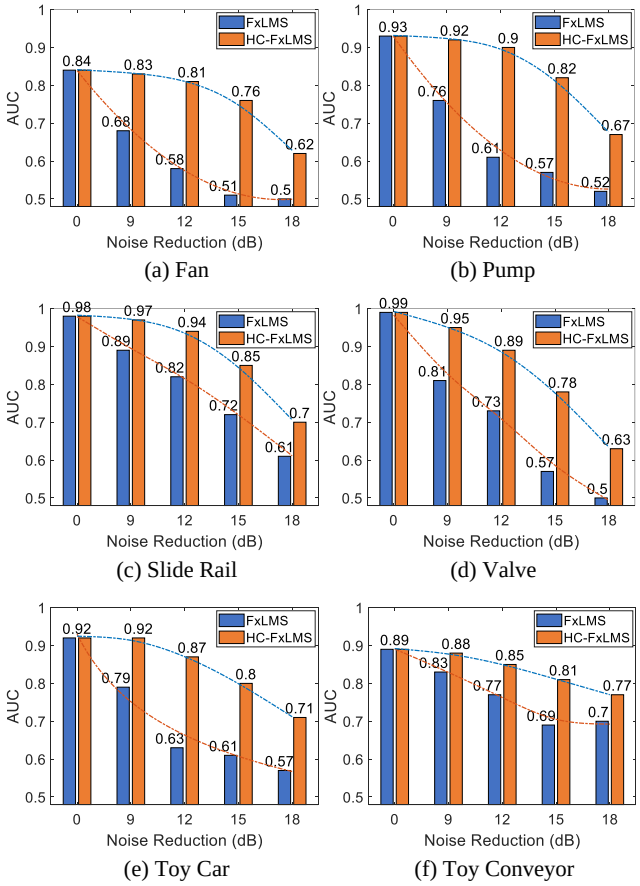

Fig. 5. AUC values of the ASD module with respect to NR.

closer results to the theoretical upper bound, as compared to the FxLMS algorithm. The improvement achieved by the homothety constraint is prominent when NR ranges from 5 dB to 15 dB.

Figure 5 shows AUC values of the ASD module when its input is provided by the residual sound of the ANC system. It is observed that the AUC value always decreases when NR increases. However, the HC-FxLMS algorithm leads to a much slower AUC degradation as compared to the FxLMS algorithm. When NR is 12 dB, the HC-FxLMS algorithm significantly outperforms the FxLMS algorithm for three machine types, *i.e.* fan, pump and toy car.

Subjective tests are carried out to evaluate the similarity of the residual sound and the original machine noise. There are 22 participants, consisting of 15 males and 7 females. Two sound clips are chosen for each type of machine under different noise reduction levels. The participants are asked to rate the similarity from 1 (bad) to 5 (excellent). The test results are presented in the average score and the 95% confidence interval with respect to the noise reduction level in Table 1. The HC-FxLMS algorithm is proven to result in less confusing perceptions.

Table 1. Subjective test scores under different noise reduction levels.

Noise Reduction	9dB	12dB	15dB	18dB
FxLMS	3.3±0.1	3.1±0.09	2.7±0.1	2.5±0.1
HC-FxLMS	4.0±0.07	3.9±0.08	3.5±0.09	2.9±0.1

5. CONCLUSIONS

This paper tackles the pressing issue of deploying ANC systems in a factory environment where the residual sound confuses the workers' judgements on the machine status. An unsupervised ASD module is trained by the auxiliary machine type classification task based on the ResNet architecture. It is integrated into the ANC system to simulate human judgements. The HC-FxLMS algorithm is proposed and results in homothety ratio curves that are very close to the theoretical upper bound. Both the AUC and subjective tests validate that the residual sound of the HC-FxLMS algorithm reduces instances of confusing perceptions. Leveraging a machine hearing system to simulate human judgments for ANC systems may bring forth a new approach of integrating ANC and machine learning.

6. REFERENCES

[1] B. Lam, C. Shi, D. Shi, and W. S. Gan, "Active control of sound through full-sized open windows," *Building Environ.*, vol. 141, pp. 16–27, 2018.

[2] S. M. Kuo and D. R. Morgan, "Active noise control: a tutorial review," *Proc. IEEE*, vol. 87, no. 6, pp. 943–973, 1999.

[3] K. Mazur, S. Wrona, and M. Pawelczyk, "Design and implementation of multichannel global active structural acoustic control for a device casing," *Mech. Syst. Signal Process.*, vol. 98, pp. 877–889, 2018.

[4] N. Miyazaki and Y. Kajikawa, "Head-mounted active noise control system with virtual sensing technique," *J. Vib. Acoust.*, vol. 339, pp. 65–83, 2015.

[5] Y. Kajikawa, W. S. Gan, and S. M. Kuo, "Recent advances on active noise control: Open issues and innovative applications," *APSIPA Trans. Signal Inf. Process.*, vol. 1, no. e3, pp. 1–21, 2012.

[6] S. Shen, N. Roy, J. Guan, H. Hassanieh, and R. R. Choudhury, "MUTE: Bringing IoT to noise cancellation," *Proc. 2018 SIGCOMM*, Budapest, Hungary, 2018.

[7] J. Pan, R. Paurobally, and X. Qiu, "Active noise control in workplaces," *Acoust. Aust.*, vol. 44, no. 1, pp. 45–50, 2015.

[8] J. Cheer and S. J. Elliott, "Active noise control of a diesel generator in a luxury yacht," *Appl. Acoust.*, vol. 105, pp. 209–214, 2016.

[9] F. Yang, J. Guo, and J. Yang, "Stochastic analysis of the filtered-x LMS algorithm for active noise control," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 28, pp. 2252-2266, 2020.

[10] B. Huang, Y. Xiao, J. Sun, and G. Wei, "A variable step-size FXLMS algorithm for narrowband active noise control," *IEEE Trans. Audio, Speech, Language Process.*, vol. 21, no. 2, pp. 301-312, 2012.

[11] C. Shi, Z. Jia, R. Xie, and H. Li, "An active noise control casing using the multi-channel feedforward control system and the relative path based virtual sensing method," *Mech. Syst. Signal Process.*, vol. 144, 2020.

[12] H. Bao and I. M. S. Panahi, "Psychoacoustic active noise control based on delayless subband adaptive filtering," *Proc. 35th IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Dallas, TX, 2010.

[13] J. Liu and X. Chen, "Adaptive compensation of mis-equalization in narrowband active noise equalizer systems," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 24, no. 12, pp. 2390-2399, 2016.

[14] V. Belyi and W. S. Gan, "Integrated psychoacoustic active noise control and masking," *Appl. Acoust.*, vol. 145, pp. 339–348, 2019.

[15] R. F. Lyon, "Machine hearing: An emerging field," *IEEE Signal Process. Mag.*, vol. 27, no. 5, pp. 131-139, 2010.

[16] A. Mesaros, T. Heittola and T. Virtanen, "Assessment of human and machine performance in acoustic scene classification: Dcase 2016 case study," *Proc. 2017 IEEE Workshop Appl. Signal Process. Audio Acoust. (WASPAA)*, New Paltz, NY, 2017.

[17] S. Zielinski, H. Lee, P. Antoniuk, and O. Dadan, "A comparison of human against machine-classification of spatial audio scenes in binaural recordings of music", *Appl. Sci.*, vol. 10, no. 17, pp. 1-24, 2020.

[18] D. Shi, W. S. Gan, B. Lam and S. Wen, "Feedforward selective Fixed-Filter active noise control: Algorithm and implementation," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 28, pp. 1479-1492, 2020.

[19] H. Zhang and D. L. Wang, "Deep ANC: A deep learning approach to active noise control," *Neural Netw.*, vol. 141, pp. 1–10, 2021.

[20] Y. Koizumi, S. Saito, H. Uematsu, Y. Kawachi, and N. Harada, "Unsupervised detection of anomalous sound based on deep learning and the Neyman–Pearson lemma," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 27, no. 1, pp. 212-224, Jan. 2019.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. 2016 IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, Nevada, 2016.

[22] H. Wu, H. Jiang, H. Wen, and C. Shi, "Abnormal drone noise detection system based on the microphone array and self-supervised learning," *Proc. 50th Int. Congr. Expo Noise Control Eng. (INTERNOISE)*, Washington, DC, 2021.

[23] H. Purohit, R. Tanabe, T. Ichige, T. Endo, Y. Nikaido, K. Suefusa, and Y. Kawaguchi, "MIMII Dataset: Sound dataset for malfunctioning industrial machine investigation and inspection," *Proc. 2019 DCASE Workshop*, New York City, NY, 2019.

[24] Y. Koizumi, S. Saito, H. Uematsu, N. Harada, and K. Imoto, "ToyADMOS: A dataset of miniature-machine operating sounds for anomalous sound detection," *Proc. 2019 IEEE Workshop Appl. Signal Process. Audio Acoust. (WASPAA)*, New Paltz, NY, 2019.