TOYADMOS2#: YET ANOTHER DATASET FOR THE DCASE2024 CHALLENGE TASK 2 FIRST-SHOT ANOMALOUS SOUND DETECTION

Daisuke Niizumi, Noboru Harada, Yasunori Ohishi, Daiki Takeuchi, Masahiro Yasuda

NTT Corporation, Japan
daisuke.niizumi@ntt.com

ABSTRACT

First-shot anomalous sound detection (ASD) is a task designed to challenge a system's applicability to new data based on the needs of real-world application scenarios. This paper describes new ToyADMOS2 data to evaluate the first-shot compliant systems for the DCASE2024 Challenge Task 2, First-Shot Unsupervised Anomalous Sound Detection for Machine Condition Monitoring. The new data is designed to differ from the previous in the new machine sounds, including HoveringDrone, HairDryer, ToyCircuit, and ToothBrush, as well as in that each sound has a different background noise. The HairDryer and ToothBrush sounds are also designed as examples of ASD application scenarios in factory pre-shipment inspections, and we confirm their potential in the evaluation. We detail these data and show the baseline performance for reference in future studies.

Index Terms— DCASE 2024 Challenge Task 2, First-Shot Anomalous Sound Detection, ToyADMOS dataset

1. INTRODUCTION

Anomalous sound detection (ASD), which uses sound as a cue to detect anomalies, has been actively studied for applications such as factory automation. To facilitate the research, the annual Detection and Classification of Acoustic Scenes and Events (DCASE) Challenge hosts an ASD task that has been drawing the attention of various participants.

The ASD challenges (the DCASE 2020-24 Challenge Task 2) [1, 2, 3, 4, 5] take a task setting that provides only normal samples for training while using normal and anomalous samples to evaluate the detection systems at test time (unsupervised ASD). This setting reflects the real-world situation where anomalous samples are hardly available or the available data cannot cover the distribution of the possible anomalies.

As we found new directions, the focus of the challenges transitioned from the ASD problem itself in 2020 [1] to a domain-shift condition in 2021 [2], a domain generalization in 2022 [3], and a first-shot condition in 2023 [4]. The first-shot condition reflects the demand for the rapid deployment of ASD systems for new machine sounds. While the conventional ASD setting uses the same machine sounds for both the development and evaluation phases, it uses only the new machine sounds for evaluation. In addition, it limits the use of training data other than the target machine sounds. Therefore, it forces the detection systems to do the first-shot solution on the new data.

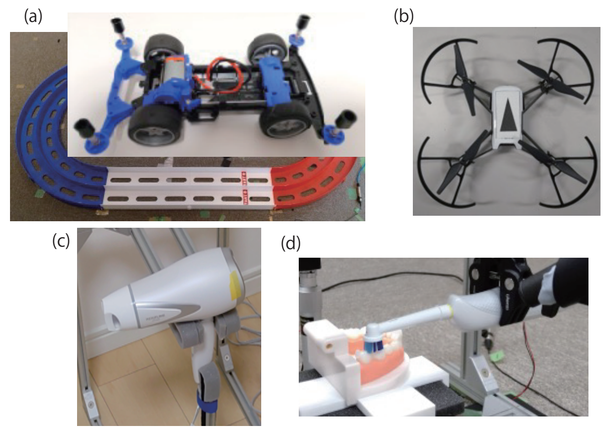

Figure 1: ToyADMOS2# includes two toys (a) ToyCircuit and (b) HoveringDrone and two home electrical appliances, (c) HairDryer, and (d) ToothBrush, bringing the evaluation setting closer to the real-world problem.

The 2024 challenge (DCASE2024T2) [5] extends the first-shot condition and limits the availability of the sample attribute information for some machines. This reflects the trend in ASD solutions where the outlier exposure (OE) approach [6, 7, 8] typically perform better. The OE approach uses sounds from different machines or attributes as anomalies; however, in real-world scenarios, we cannot always obtain machine attribute information.

For the 2020–2023 challenges, the developed datasets include: ToyADMOS [9], MIMII [10], ToyADMOS2 [11], MIMII DUE [12], MIMII DG [13], and ToyADMOS2+ [14]. This paper introduces new data for the ToyADMOS family, ToyADMOS2# (*two sharp*), that meets the first shot requirement of the DCASE2024T2.

The previous ASD sounds are mainly targeted at detecting failures in operating machines, such as production ma-

Table 1: ToyADMOS data history.

Machine type	DCASE usage	Product type	Mobility	Recorded sounds
(i) ToyADMOS/ToyADMOS2/ToyADMOS2+				
ToyCar	2020-24 dev. & 2020-22 eval.	Toy	Fixed	Running while fixed at a stand
ToyTrain	2020-24 dev. & 2020-22 eval.	Toy	Mobile	Circling on a railroad track
ToyConveyor	2020 dev. & eval.	Toy	Fixed	Conveying small cargo
Vacuum	2023 evaluation	Appliance	Fixed	Vacuuming at a fixed point
ToyTank	2023 evaluation	Toy	Mobile	Moving back and forth
ToyNscale	2023 evaluation	Toy	Mobile	Circling on a railroad track
ToyDrone	2023 evaluation	Toy	Mobile	Taking off and landing
(ii) ToyADMOS2#				
ToyCircuit	2024 evaluation	Toy	Mobile	Circling on a circuit track
HoveringDrone	2024 evaluation	Toy	Stationary	Hovering with rotation at a fixed point
HairDryer	2024 evaluation	Appliance	Fixed	Blowing at a fixed point
ToothBrush	2024 evaluation	Appliance	Fixed	Brushing at a fixed point

chines in a factory. In contrast, to show the possibility of an ASD application to the production process in a factory, we have two sound settings of typical home electrical appliances as examples of pre-shipment inspection. ToyADMOS2# provides the additional training and evaluation datasets of the DCASE2024T2 with four machine sounds described in Fig. 1. The dataset is available at Zenodo [5, 15, 16].

2. PREVIOUS TOYADMOS DATASETS

The ToyADMOS family has released three datasets, and Table 1(i) lists all the data they provided to the past DCASE Challenge Task 2. The first release, ToyADMOS [9], contains three miniature machine (toy) sounds with various anomalous sounds. It uses toys as machine sound sources and simulates anomalies by breaking parts of them to address the difficulty of collecting anomalous sounds. ToyADMOS2 [11], released in 2021, contains a wider variety of sounds of two toys and enables the generation of datasets that simulate domain shift conditions. ToyADMOS2+ [14] recorded the sounds of new machines to enable first-shot ASD and provided them as an evaluation set in the DCASE 2023 Challenge Task 2 [17]. It also features a home electrical appliance as one of the machines, introducing a new ASD setting of the everyday sounds around us.

3. TOYADMOS2#: NEW DATA FOR THE DCASE2024 CHALLENGE TASK 2

ToyADMOS2# adds data for the four machines shown in Table 1(ii) for evaluation under the new first-shot ASD condition. To distinguish them from the previous data, we used the sounds of two home electrical appliances and different background noises for all of them. We specifically designed HairDryer and ToothBrush as example scenarios of ASD applications in product pre-shipment inspections of home electrical appliances.

ToyCircuit: The sound of the ToyCar (TAMIYA Mini 4WD) driving on a circuit track, characterized by the sound of friction with the track lanes and the change in distance from the microphone.

HoveringDrone: The sound of the drone (DJI Tello) hovering at one point and rotating. Unlike the ToyDrone last year, we made the distance from the microphone almost constant.

HairDryer: The sound of the dryer (Panasonic/Koizumi) airflow to evaluate the detection of anomalous sound when airflow is obstructed, such as when unintended foreign matter adheres to the dryer during the production process.

ToothBrush: The sound of the electric toothbrush (Brown DB5510) brushing teeth to evaluate the detection of anomalies, such as brushes manufactured with defects. We maintained constant pressure between the brush against the teeth to avoid changes in sound due to pressure differences.

Table 2 summarizes the details of each machine, especially the speed/mode and background noise characterizing the differences between them. Domain shift settings commonly changed from source to target: ID from A and B to C, and microphone from 1 to 2. We changed the speed and mode for each machine and basically assigned the unused values in the source to the target.

Table 3 details the anomalous conditions for each machine. ToyCircuit differs from ToyCar in that the anomalous sounds are also produced by friction with the running surface. HoveringDrone assumes that the adhesion of foreign objects occurs during use. Anomalous conditions for HairDryer and ToothBrush also assume adhesion while further assuming that defects in the manufacturing process of these products cause anomalous sounds and that future ASD systems detect them in product pre-shipment inspections. For example, future ASD systems could be combined with optical inspection (e.g., toothbrushes in the MVTec AD dataset [18]) to improve factory product inspection performance. Fig. 2 showcases the anomalous condition examples of ToothBrush used in the recordings.

3.1. Recording control details

The sounds were recorded in a controlled environment by following the recording layouts and microphone arrange-

Table 2: ToyADMOS2# data details.

	Settings and parameter variations				Domain shift settings	Samples	
Machine type	ID	Dur.	Speed/mode	Background noise	Source→target	Training	Eval.
(a) ToyCircuit	A, B, C	8 s	1: 1.3 V, 2: 1.4 V, 3: 1.5 V, 4: 1.6 V	Large air conditioner outdoor unit outlet noise	ID: A,B→C, Mic: 1→2 Speed: 1, 2, 3→1, 4	Src. 990 Trg. 10	Src. 100 Trg. 100
(b) HoveringDrone†	A, B, C	8 s	1: Rotate CW 180° and CCW 180°, 2: Rotate CW 180° and CW 180°, 3: Rotate CCW 180° and CCW 180°	City noise near a river under a highway bridge	ID: A,B→C, Mic: 1→2 Mode: 1, 2→3	Src. 990 Trg. 10	Src. 100 Trg. 100
(c) HairDryer	A, B, C	7 s	1: 92 V, 2: 96 V, 3: 100 V, 4: 104 V	Running water sound in a drainage ditch in a park	ID: A,B→C, Mic: 1→2 Speed: 2,3→1,4	Src. 990 Trg. 10	Src. 100 Trg. 100
(d) ToothBrush†	A, B, C	6 s	1: Lower teeth/ 2.7 V, 2: Lower teeth/ 2.8 V, 3: Upper teeth/ 2.8 V, 4: Lower teeth/ 2.9 V, 5: Upper teeth/ 2.9 V, 6: Lower teeth/ 3.0 V	Home air purifier outlet noise	ID: A,B→C, Mic: 1→2 Mode: 2, 3, 4, 5→1, 6	Src. 990 Trg. 10	Src. 100 Trg. 100

†The actual parameters (sample attributes) were not provided in the data files following the focus of the DCASE2024T2.

Table 3: Anomaly conditions for each machine type.

(a) ToyCircuit		(b) HoveringDrone	
Part	Condition	Part	Condition
Tire	Foreign objects	Arm	Foreign object
	Scratches	Propeller	Foreign object/one side
Shaft	No grease		Foreign object/two sides
Gear	Locked gear	Body	Offset weight

(c) HairDryer		(d) ToothBrush	
Part	Condition	Part	Condition
Outlet	Foreign object	Brush hair	Damaged brush hair
Inlet	Foreign object		Foreign object stuck
Vane	Foreign object		Partially missing brush hair
	Chipped vane	Brush head	Half-insertion of brush head

ments shown in Figs. 3 and 4 and by automating the machines' controls. The system used optical sensors to manage the laps of the ToyCircuit, image recognition to control the HoveringDrone, and automatic control of the main power supply for the appliances. The resulting sounds should reflect the differences in hardware, actual mechanical movements, and course and drive/flight conditions.

While in the controlled sound recording environment, we limited the number of samples obtained in a single recording to make the data distribution closer to the intrinsic nature of the machine's data distribution. In particular, the recording of the AC-powered HairDryer can continuously provide many samples at once; however, it ends up with many similar samples and cannot cover the data distribution gained by the differences in installation, assembly, time, and natural degradation. Therefore, we avoided continuous operation and switched to recording under physically different conditions for no more than 30 samples.

3.2. Data sample details

All the operating sound and noise samples were recorded with 48-kHz sampling, 24 bits for each channel, and then downsampled to 16-kHz, 16 bits, monaural in the final data samples. Sample duration varied from 6 s to 8 s, depending on the machine type, as shown in Table 2.

The training data (Additional training dataset) for each machine type has 1000 normal samples, 990 from the source domain and 10 from the target domain. The evaluation data (Evaluation dataset) for each machine type consists of 50 normal and 50 anomaly samples from each source and target domain, for a total of 200 samples. The total of these data provides 4800 samples with 580 minutes of recordings. The data are available at the Zenodo links [15, 16] under the Creative Commons Attribution 4.0 International Public License [19].

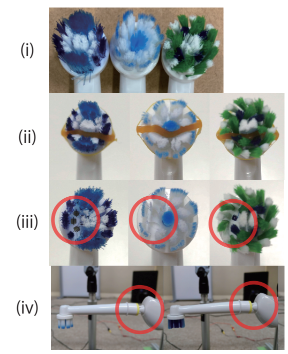

Figure 2: Anomalous condition examples of ToothBrush: (i) Damaged brush hair, (ii) foreign object stuck in the brush, (iii) partially missing brush hair, and (iv) half-insertion of the brush head.

4. BENCHMARKS

We show the evaluation results obtained using the DCASE2024 Challenge Task 2 baseline system in Table 4. The baseline is a reconstruction-based ASD system using Autoencoder and has two operating modes: First-shot-compliant Simple Autoencoder mode and Selective Mahalanobis Autoencoder mode. The former calculates the distance between the input sample and the reconstruction using MSE (mean squared error), while the latter does based on Mahalanobis's distance [20]. The results are the area under the receiver operating characteristic curve (AUC) and partial AUC (pAUC), where the pAUC measures performance in a

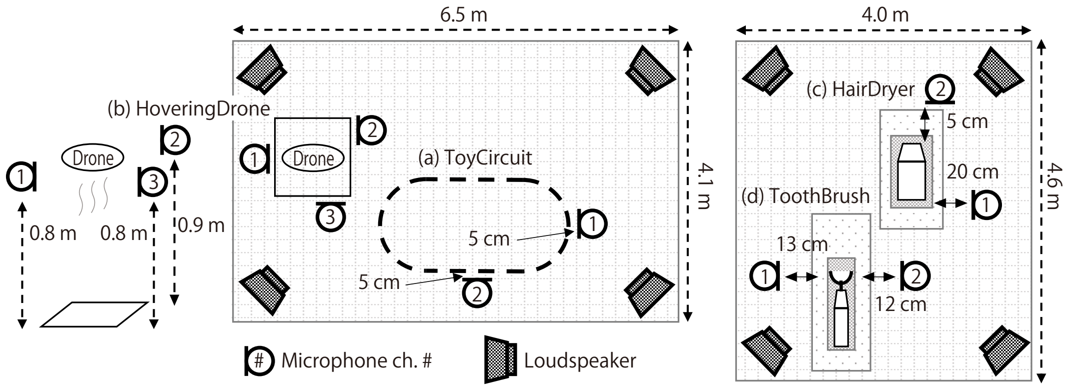

Figure 3: Recording-room layouts and microphone arrangements: (a) ToyCircuit, (b) HoveringDrone, (c) HairDryer, and (d) ToothBrush.

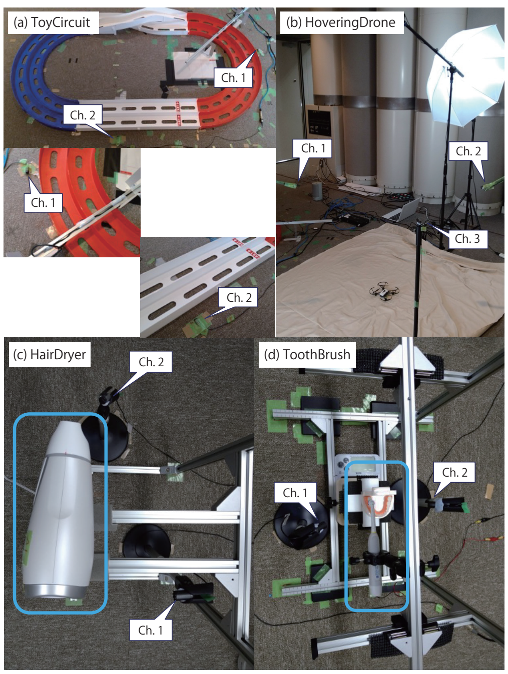

Figure 4: Microphone arrangements: (a) ToyCircuit, (b) HoveringDrone, (c) HairDryer, and (d) ToothBrush.

low false-positive rate (FPR) range $[0, p]$ with a p of 0.1. For the details, see [20, 21].

The results show that for all machine types, the baseline performs well in the source domain while generalization to the target domain is difficult, a trend similar to that for data through 2023. The exception for ToothBrush is that the baseline also performs well on the target domain data, suggesting that the degree of domain shift is small. The performance of the machines simulating a product pre-shipment inspection scenario (HairDryer and ToothBrush) shows a similar trend to that of the other two machines, implying the potential for future ASD applications in the scenario.

Table 4: Benchmark results

Machine type	AUC [%]		pAUC [%]
	Source	Target	
(i) First-shot-compliant baseline: Simple Autoencoder mode			
ToyCircuit	77.50 ± 0.82	51.25 ± 1.22	50.12 ± 0.17
HoveringDrone	85.93 ± 0.77	47.87 ± 4.30	51.05 ± 1.46
HairDryer	64.94 ± 3.40	43.75 ± 2.09	50.56 ± 1.01
ToothBrush	73.80 ± 1.08	70.14 ± 4.92	54.19 ± 1.73
(ii) First-shot-compliant baseline: Selective Mahalanobis Autoencoder mode			
ToyCircuit	69.67 ± 1.72	42.34 ± 2.11	49.23 ± 0.03
HoveringDrone	84.07 ± 1.10	48.50 ± 3.64	58.95 ± 2.76
HairDryer	64.23 ± 3.44	56.71 ± 1.97	55.12 ± 0.71
ToothBrush	63.17 ± 2.43	57.55 ± 3.59	49.81 ± 1.29

5. CONCLUSION

This paper introduced new ToyADMOS2 data to evaluate the first-shot compliant systems for the DCASE2024 Challenge Task 2, First-Shot Unsupervised Anomalous Sound Detection for Machine Condition Monitoring. The first-shot anomalous sound detection (ASD) is a task designed to challenge a system's applicability to new data based on the needs of real-world application scenarios. The new sounds include HoveringDrone, HairDryer, ToyCircuit, and ToothBrush, and they are mixed with four different environmental noises to enhance their differences from the previous sounds. We specifically designed two sounds, HairDryer and ToothBrush, as example scenarios of ASD applications in product pre-shipment inspections of home electrical appliances and confirmed their potential in the evaluation. The ToyADMOS2# dataset (DCASE 2024 Challenge Task 2 Additional Training Dataset and Evaluation Dataset) is available at [5, 15, 16] with the Creative Commons Attribution 4.0 International Public License [19].

6. ACKNOWLEDGMENT

The authors thank Hideaki Sanada, Akito Karasawa, Keisuke Hasegawa, Azusa Takahashi, and Yuri Musashijima for their hard work implementing the systems and conducting experiments.

7. REFERENCES

[1] Y. Koizumi, Y. Kawaguchi, K. Imoto, T. Nakamura, Y. Nikaido, R. Tanabe, H. Purohit, K. Suefusa, T. Endo, M. Yasuda, and N. Harada, "Description and discussion on DCASE2020 challenge task2: Unsupervised anomalous sound detection for machine condition monitoring," in *DCASE Workshop*, November 2020, pp. 81–85.

[2] Y. Kawaguchi, K. Imoto, Y. Koizumi, N. Harada, D. Niizumi, K. Dohi, R. Tanabe, H. Purohit, and T. Endo, "Description and discussion on dcase 2021 challenge task 2: Unsupervised anomalous detection for machine condition monitoring under domain shifted conditions," in *DCASE Workshop*, Barcelona, Spain, November 2021, pp. 186–190.

[3] K. Dohi, K. Imoto, N. Harada, D. Niizumi, Y. Koizumi, T. Nishida, H. Purohit, R. Tanabe, T. Endo, M. Yamamoto, and Y. Kawaguchi, "Description and discussion on dcase 2022 challenge task 2: Unsupervised anomalous sound detection for machine condition monitoring applying domain generalization techniques," in *DCASE Workshop*, Nancy, France, November 2022.

[4] K. Dohi, K. Imoto, N. Harada, D. Niizumi, Y. Koizumi, T. Nishida, H. Purohit, R. Tanabe, T. Endo, and Y. Kawaguchi, "Description and discussion on DCASE 2023 challenge task 2: First-shot unsupervised anomalous sound detection for machine condition monitoring," in *DCASE Workshop*, Tampere, Finland, September 2023, pp. 31–35.

[5] "DCASE 2024 challenge task 2: First-shot unsupervised anomalous sound detection for machine condition monitoring," 2024, https://dcase.community/challenge2024/task-first-shot-unsupervised-anomalous-sound-detection-for-machine-condition-monitoring.

[6] D. Hendrycks, M. Mazeika, and T. Dietterich, "Deep anomaly detection with outlier exposure," in *International Conference on Learning Representations*, 2019. [Online]. Available: https://openreview.net/forum?id=HyxCxhRcY7

[7] R. Giri, S. V. Tenneti, F. Cheng, K. Helwani, U. Isik, and A. Krishnaswamy, "Self-supervised classification for detecting anomalous sounds," in *DCASE Workshop*, 2020, pp. 46–50.

[8] K. Wilkinghoff and F. Kurth, "Why do angular margin losses work well for semi-supervised anomalous sound detection?" *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 32, pp. 608–622, 2024.

[9] Y. Koizumi, S. Saito, H. Uematsu, N. Harada, and K. Imoto, "ToyADMOS: A dataset of miniature-machine operating sounds for anomalous sound detection," in *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, November 2019, pp. 308–312.

[10] 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," in *DCASE Workshop*, November 2019, pp. 209–213.

[11] N. Harada, D. Niizumi, D. Takeuchi, Y. Ohishi, M. Yasuda, and S. Saito, "ToyADMOS2: Another dataset of miniature-machine operating sounds for anomalous sound detection under domain shift conditions," in *DCASE Workshop*, Barcelona, Spain, November 2021, pp. 1–5.

[12] R. Tanabe, H. Purohit, K. Dohi, T. Endo, Y. Nikaido, T. Nakamura, and Y. Kawaguchi, "MIMII DUE: Sound dataset for malfunctioning industrial machine investigation and inspection with domain shifts due to changes in operational and environmental conditions," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 21–25, 2021.

[13] K. Dohi, T. Nishida, H. Purohit, R. Tanabe, T. Endo, M. Yamamoto, Y. Nikaido, and Y. Kawaguchi, "MIMII DG: Sound dataset for malfunctioning industrial machine investigation and inspection for domain generalization task," in *DCASE Workshop*, Nancy, France, November 2022.

[14] N. Harada, D. Niizumi, D. Takeuchi, Y. Ohishi, and M. Yasuda, "ToyADMOS2+: New toyadmos data and benchmark results of the first-shot anomalous sound event detection baseline," in *DCASE Workshop*, Tampere, Finland, September 2023, pp. 41–45.

[15] "DCASE 2024 Challenge Task 2 Additional Training Dataset," 2024, https://zenodo.org/records/11259435.

[16] "DCASE 2024 Challenge Task 2 Evaluation Dataset," 2024, https://zenodo.org/records/11363076.

[17] "DCASE 2023 challenge task 2: First-shot unsupervised anomalous sound detection for machine condition monitoring," 2023, https://dcase.community/challenge2023/task-first-shot-unsupervised-anomalous-sound-detection-for-machine-condition-monitoring.

[18] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger, "MVTec AD — a comprehensive real-world dataset for unsupervised anomaly detection," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 9584–9592.

[19] "Creative commons attribution 4.0 international public license," https://creativecommons.org/licenses/by/4.0/legalcode.

[20] N. Harada, D. Niizumi, Y. Ohishi, D. Takeuchi, and M. Yasuda, "First-shot anomaly sound detection for machine condition monitoring: A domain generalization baseline," in *2023 31st European Signal Processing Conference (EUSIPCO)*, 2023, pp. 191–195.

[21] N. Harada, Y. Musashijima, and D. Niizumi, "dcase2023_task2_baseline_ae," 2023, https://github.com/nttcslab/dcase2023_task2_baseline_ae.