



ToyADMOs2: Another dataset of miniature-machine operating sounds for anomalous sound detection under domain shift conditions



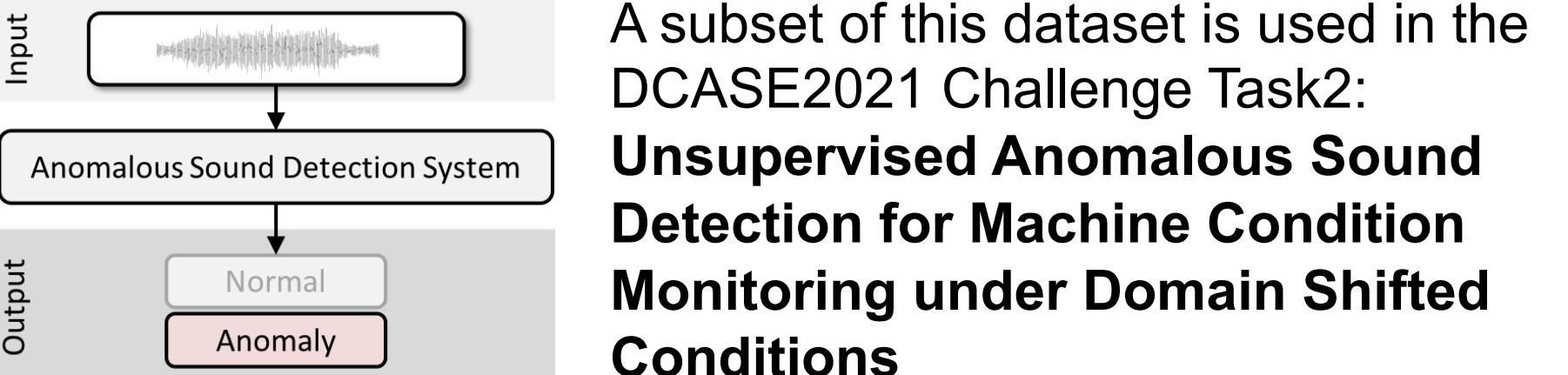
Noboru Harada, Daisuke Niizumi, Daiki Takeuchi, Yasuhiro Ohishi, Masahiro Yasuda, and Shoichiro Saito

NTT Corporation, Japan

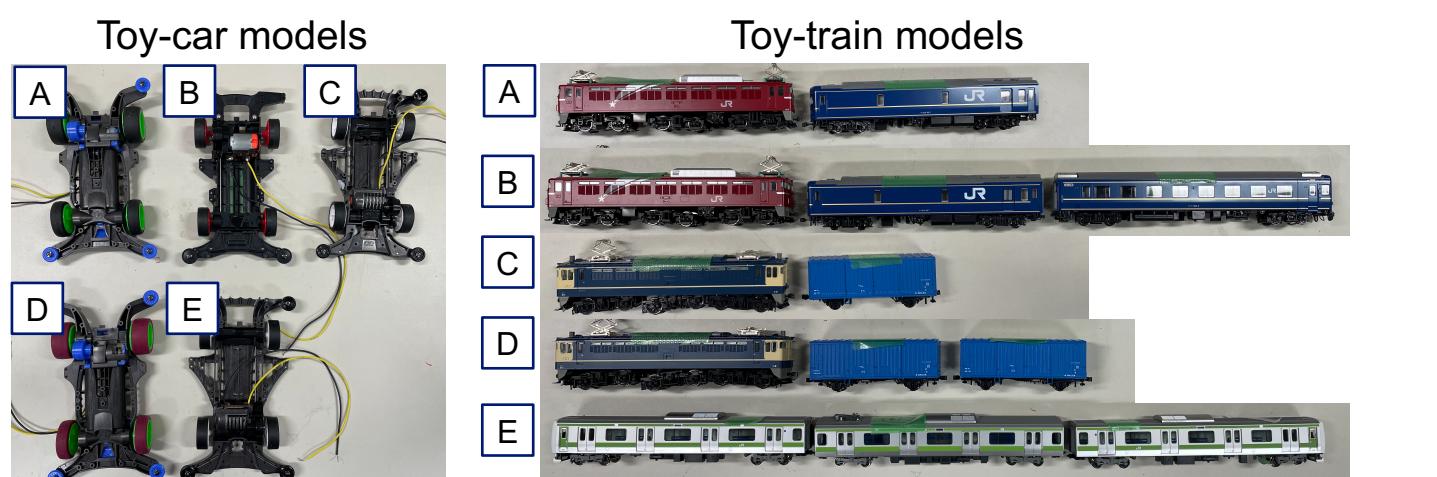
What is ToyADMOs2 ?

- “ToyADMOs2” is a new large-scale dataset for anomaly detection in machine operating sound
- Designed for evaluating anomaly detection systems under Domain-shift conditions

The dataset is freely available at
<https://github.com/nttcslab/ToyADMOs2-dataset>



What is ToyADMOs2 ?



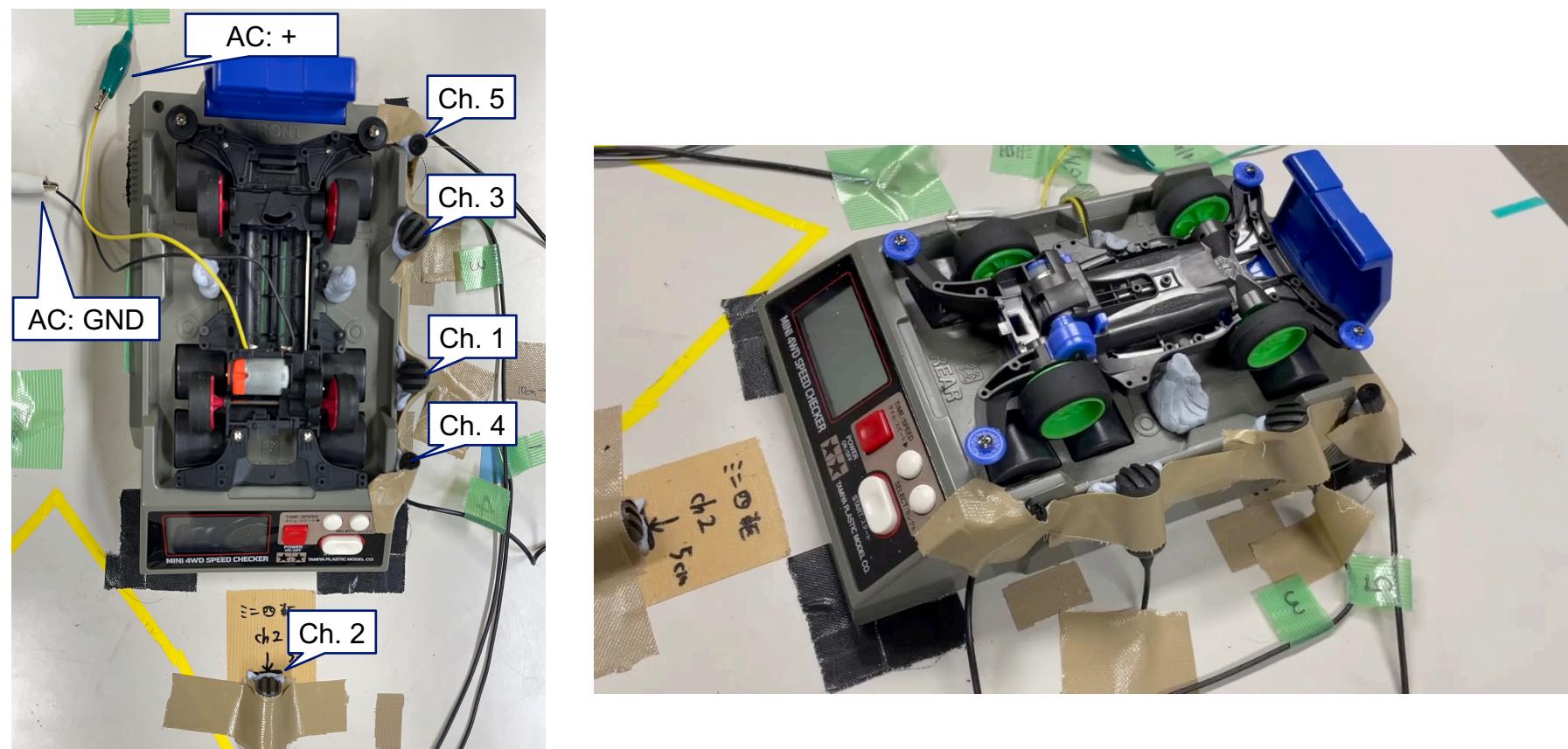
- Collected operating sound of miniature-machines (toys) under normal and anomaly conditions by deliberately damaging them
- with a controlled depth of damages (3 levels) in the anomaly samples

Design choice for domain-shift conditions

Variation settings of sub-datasets

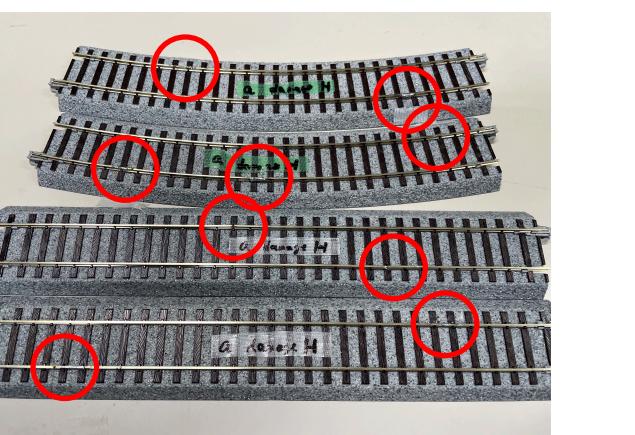
	Toy car	Toy train
Model variations	Five (A, B, C, D, E)	Five (A, B, C, D, E)
Speed levels	Five (1, 2, 3, 4, 5)	Five (1, 2, 3, 4, 5)
Anomaly conditions	Four (a, b, c, d) x 3 damage levels	
Mic. type and channel config.	Dynamic Condenser	Ch. 1-3 Ch. 4, 5
Noise type		Four recordings
Normal samples	1,094 samples x 5 models x 5 speed levels 91 hours x 5 ch	91 hours x 2 ch-sets
Anomaly samples	324 samples x 5 models x 5 speed levels 27 hours x 5 ch	27 hours x 2 ch-sets
Noise samples	Four types (24 hours) per a channel	

Toy car: Microphone arrangement

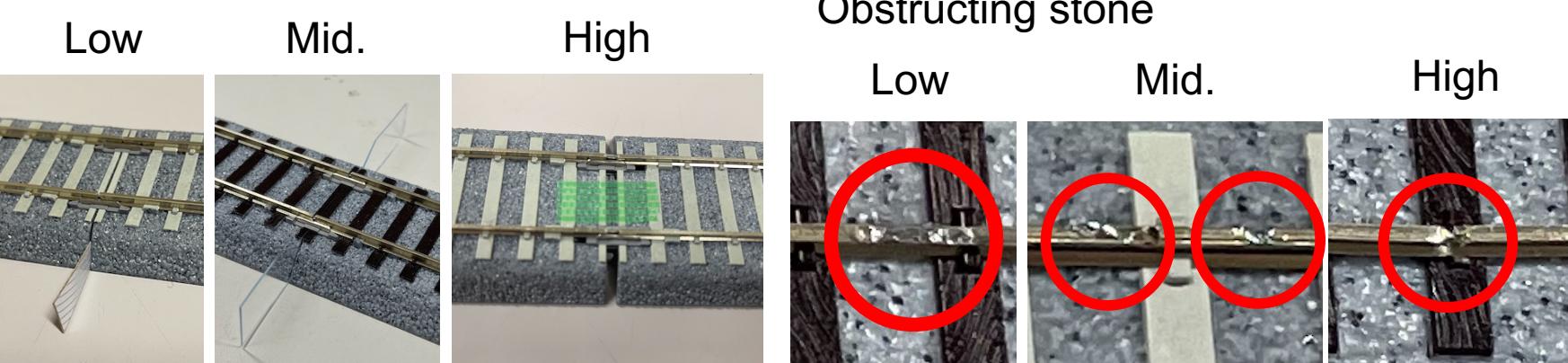


Toy train: Anomaly conditions

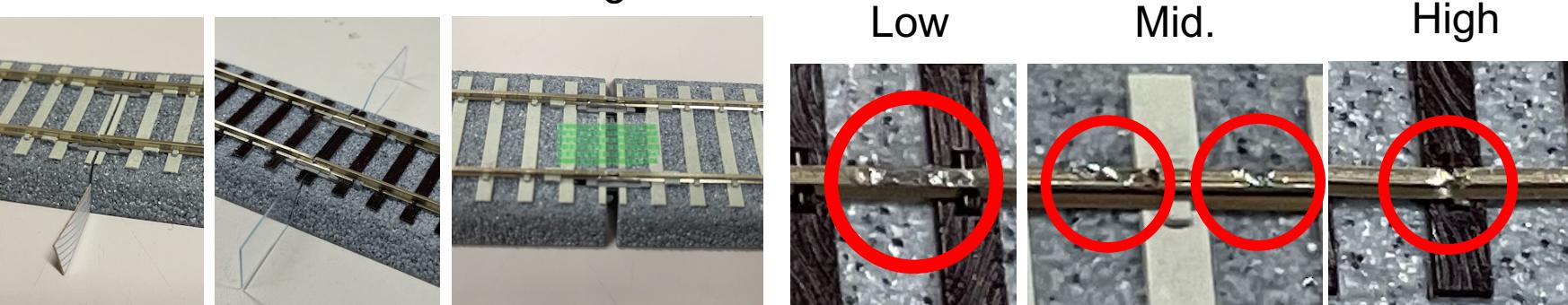
Toy car	Toy train
a Bent shaft	Obstructing stone
b Deformed gears	Disjointed railway
c Melted gears	Broken shaft
d Damaged wheels	Flat tire



Disjointed railway



Obstructing stone

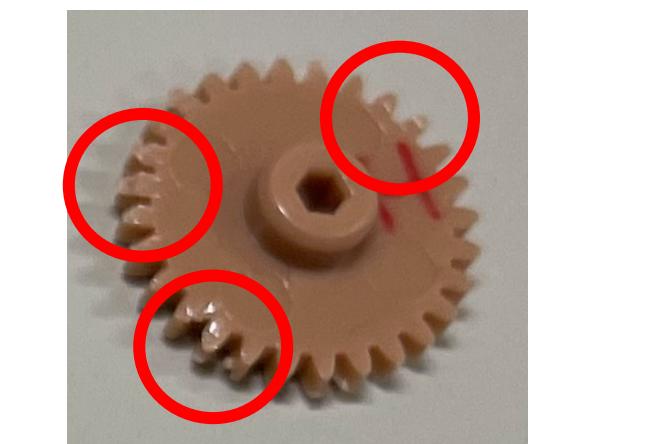


Toy car: Anomaly conditions

Toy car	Toy train
a Bent shaft	Obstructing stone
b Deformed gears	Disjointed railway
c Melted gears	Broken shaft
d Damaged wheels	Flat tire



Deformed gears (High)



Bent shaft (High)

Toy train: Anomaly conditions

Toy car	Toy train
a Bent shaft	Obstructing stone
b Deformed gears	Disjointed railway
c Melted gears	Broken shaft
d Damaged wheels	Flat tire



Broken shaft (High)



10

Toy car: Anomaly conditions

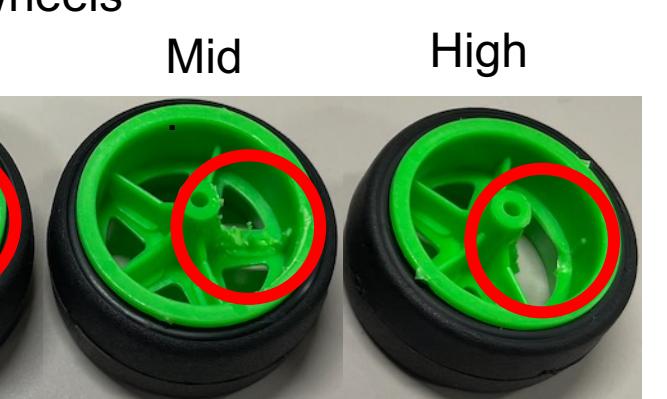
Toy car	Toy train
a Bent shaft	Obstructing stone
b Deformed gears	Disjointed railway
c Melted gears	Broken shaft
d Damaged wheels	Flat tire



Melted gears



Damaged wheels



Recipe (easy template for generating test conditions)

<https://github.com/nttcslab/ToyADMOs2-dataset>

Example: Making Example Dataset

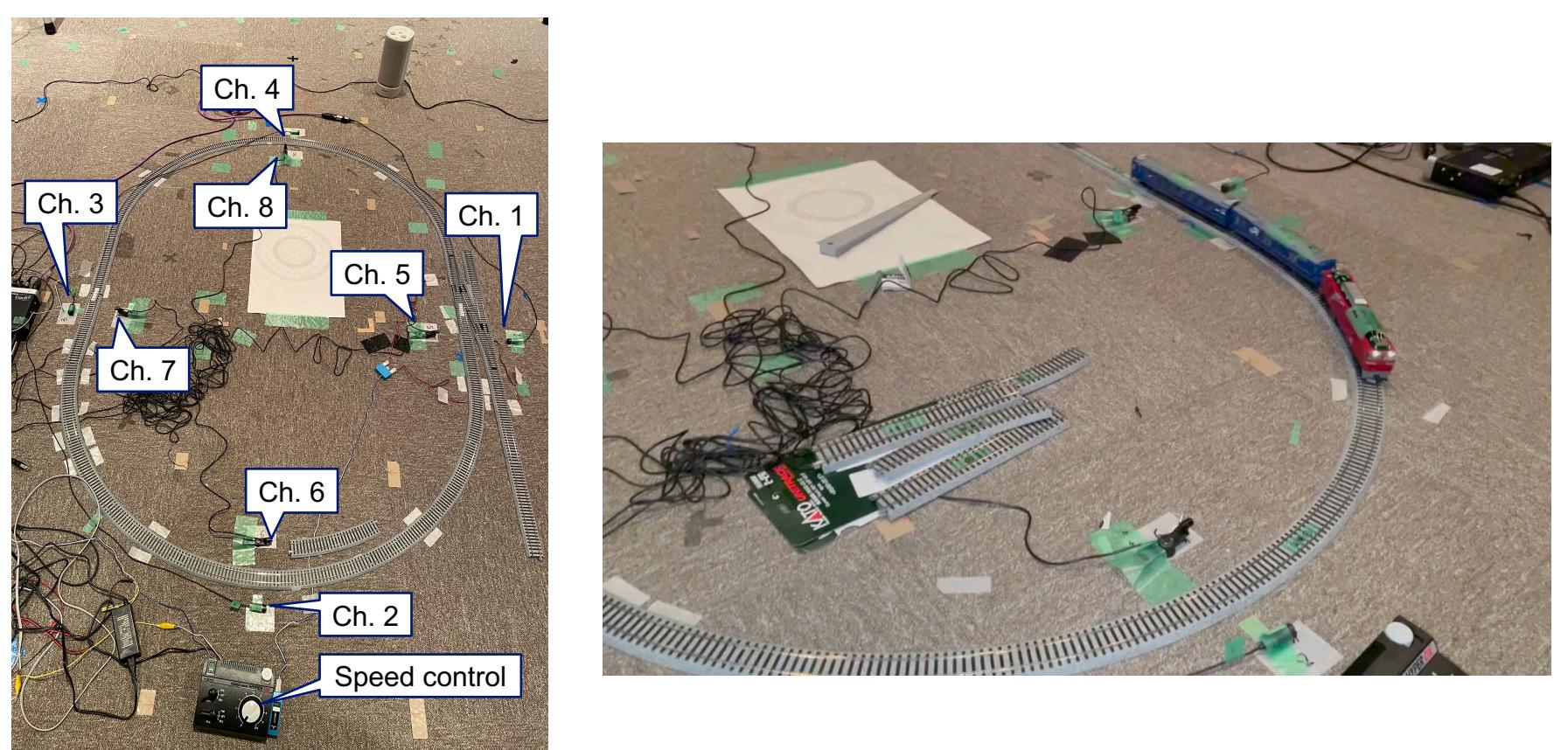
Run the following will create the equivalent benchmark dataset evaluated in the Table 3 of the paper, which is a compatible file-folder structure with the DCASE2021 challenge task 2. This will create dataset folder `your_new_dataset`. This will take about an hour.

```
# This creates `clean` dataset.
python mixer.py /path/to/ToyADMOs2 your_new_dataset recipe_benchmark.xlsx clean
# This creates SNR=6dB dataset.
python mixer.py /path/to/ToyADMOs2 your_new_dataset recipe_benchmark.xlsx 6
```

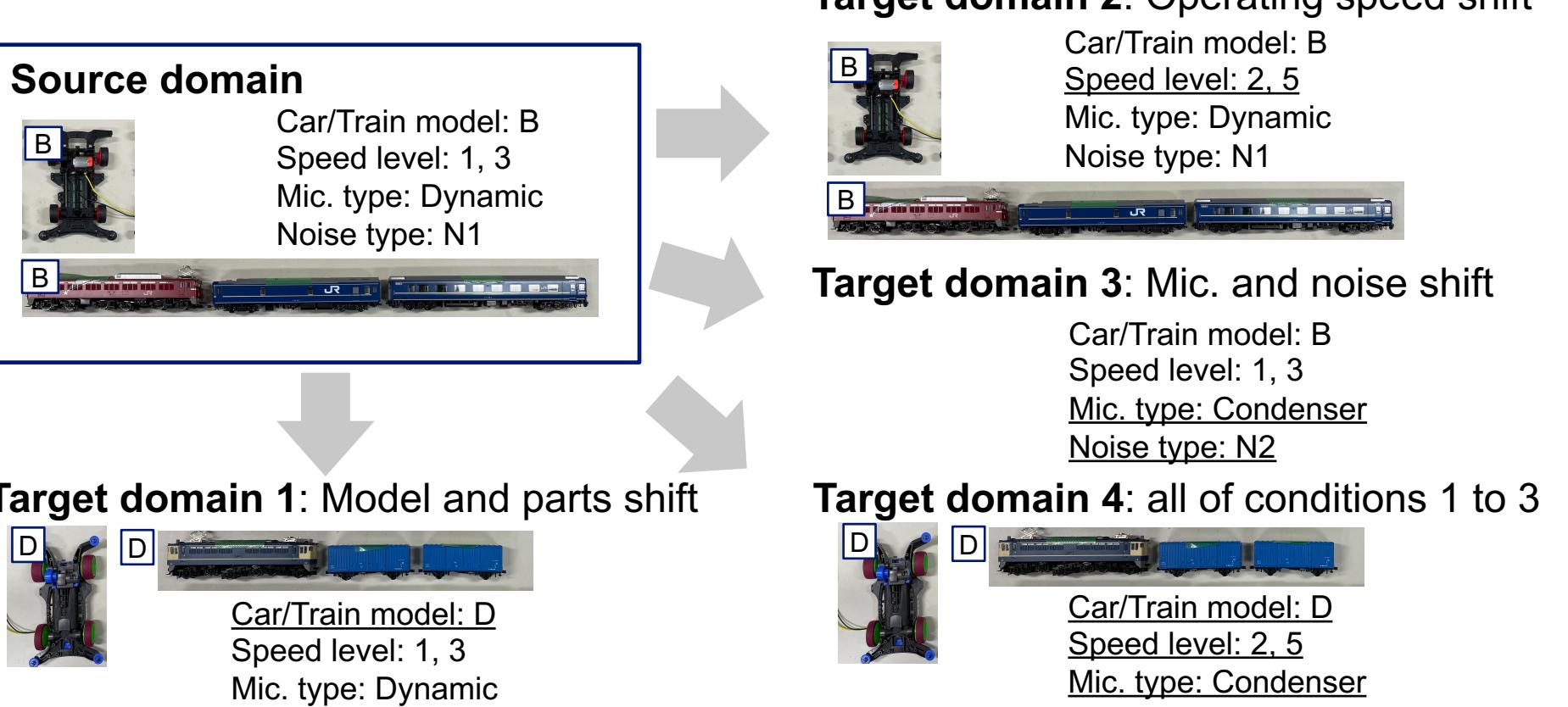
- `recipe_example_car_shift.xlsx` is also another example.
- `recipe_template` is a template, as well as one more example.

11

Toy train: Microphone arrangement



Example domain-shift condition



Recipe (easy template for generating sub-dataset)

You can simply edit the recipe template and run `mixer.py` to generate a test dataset

A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	No.	Folder	FileID	Model	CarID	Speed	Defect	D.Level	# of Rec.	r0_pat	r0_mics	r0_nz	r0_cty
2	1	ToyCar/normal	CN001-carA1-speed1	A	A1	1	-	-	260	ToyCar/train/section_00_source_train_normal_?????.wav	[1,2,3,4,5]	1	30
3	2	ToyCar/normal	CN002-carA1-speed1	A	A1	1	-	-	260	ToyCar/train/section_00_source_train_normal_?????.wav	[1,2,3,4,5]	1	30
4	3	ToyCar/normal	CN003-carA2-speed1	A	A2	1	-	-	260	ToyCar/train/section_00_source_train_normal_?????.wav	[1,2,3,4,5]	1	30
5	4	ToyCar/normal	CN004-carA2-speed1	A	A2	1	-	-	0	0	0	0	0
6	5	ToyCar/normal	CN005-carA1-speed1	A	A1	1	-	-	0	r1_pat	r1_mics	r1_nz	r1_cty
7	6	ToyCar/normal	CN006-carA1-speed1	A	A1	1	-	-	0	0	0	0	0
8	7	ToyCar/normal	CN121-carE1-speed1	A	A1	1	-	-	0	0	0	0	0
9	8	ToyCar/normal	CN122-carE1-speed1	A	A1	1	-	-	0	0	0	0	0
10	9	ToyCar/normal	CN123-carE2-speed1	A	A2	1	-	-	0	0	0	0	0
11	10	ToyCar/normal	CN124-carE2-speed1	A	A2	1	-	-	0	0	0	0	0
12	11	ToyCar/normal	CN125-carE1-speed1	A	A1	1	-	-	0	0	0	0	0
13	12	ToyCar/normal	CN126-carE2-speed1	A	A2	1	-	-	0	0	0	0	0
14	13	ToyCar/normal	CN127-carE1-speed2	A	A1	1	-	-	0	0	0	0	0
15	14	ToyCar/normal	CN128-carE2-speed2	A	A2	1	-	-	0	0	0	0	0
16	15	ToyCar/normal	CN129-carF1-speed2	A	A1	1	-	-	0	0	0	0	0

\$ python mixer.py ToyADMOs2 car_shift recipe_example_car_shift.xlsx -6

Conclusion

- A new dataset for anomaly sound detection “ToyADMOs2” is proposed.
- Consists of two sub-datasets (Toy car and Toy train) for machine condition inspection.
- Domain shifts are represented by introducing several differences in operating conditions, such as different machine models and parts configurations, operating speeds, microphone arrangements and background noise types.
- The dataset can be used to assess anomaly detection systems. It is desired to develop anomaly detection systems working well even under domain-shift conditions.

DCASE2021 Challenge T2 results on ToyADMOs2

Rank	System	Source domain results		Target domain results	
		ToyCar	ToyTrain	ToyCar	ToyTrain
1	Lopez_IL_task				