ANOMALOUS SOUND DETECTION SYSTEM WITH SELF-CHALLENGE AND METRIC EVALUATION FOR DCASE2022 CHALLENGE TASK 2

Technical Report

Yuming Wei[1], *Jian Guan*[1]*, *Haiyan Lan*[1], *and Wenwu Wang*[2]

[1]Group of Intelligent Signal Processing, College of Computer Science and Technology, Harbin Engineering University, Harbin, China
[2]Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, UK

ABSTRACT

This technical report describes our submission for DCASE2022 Challenge Task 2. To solve the domain generalization problem in anomalous sound detection (ASD), we present an ensemble system with two proposed unsupervised anomalous sound detection methods, i.e., a self-supervised classifier with a self-challenge strategy, and a distance metric evaluation based method. Experiments conducted show that our ensemble system can achieve an average of 87.07% in harmonic mean AUC score under the source domain (h-mean AUC-s), and an average of 76.22% in harmonic mean AUC score under the target domain (h-mean AUC-t), and an average of 66.76% in harmonic mean pAUC (h-mean pAUC) score.

Index Terms— Anomalous sound detection, self-challenge, distance metric, domain generalization

1. INTRODUCTION

Unsupervised anomalous sound detection aims to automatically identify machines' status (normal/abnormal) from learning the emitted normal sound data [1, 2]. The main focus of DCASE2022 Challenge Task 2 [3] is to identify normal and anomalous sound under domain shift conditions. Compared with previous tasks (i.e., DCASE2020 [1] and DCASE2021 task 2 [2]), the main challenge of DCASE2022 Task 2 is the domain generalization problem. Specifically, the samples unaffected by domain shifts (source domain data) and those affected by domain shifts (target domain data) are mixed in the test dataset. Moreover, the source/target domain is not specified for each sample. Therefore, it requires the system to detect the anomalies with the same threshold value regardless of the domain.

Our submission is an ensemble system, which is composed of a self-challenge classification (SCC) system and a clustering-based system with metric evaluation (CBM). An ensemble learning strategy [4] is applied to integrate these two methods.

2. PROPOSED SYSTEM

The baseline systems for DCASE2022 Challenge Task 2 are an autoencoder (AE) system and a MobileNetV2 system. The performance of these two baseline systems is provided respectively in Table 1 and Table 2. Our ensemble system is the integration of two proposed systems, SCC and CBM.

Corresponding author.

This work was supported by the Natural Science Foundation of Heilongjiang Province under Grant No. Y2020F010.

Table 1: Performance of the AE baseline method (%)

	ToyCar	ToyTrain	Bearing	Fan	Gearbox	Slider	Valve	Average
h-mean AUC-s	90.41	76.32	54.42	78.59	68.93	77.95	52.01	71.23
h-mean AUC-t	34.81	23.35	58.38	47.18	62.64	47.67	49.46	46.21
h-mean pAUC	52.74	50.48	51.98	57.52	58.49	55.78	50.36	53.91

Table 2: Performance of the MobileNetV2 baseline method (%)

	ToyCar	ToyTrain	Bearing	Fan	Gearbox	Slider	Valve	Average
h-mean AUC-s	59.12	57.26	60.58	70.75	69.21	65.15	67.09	64.17
h-mean AUC-t	51.96	45.90	59.94	48.22	56.19	38.23	57.22	51.09
h-mean pAUC	52.27	51.52	57.14	56.90	56.03	54.67	62.42	55.85

2.1. Self-Challenge Classification System

We adopt our previous work, i.e., STgram-MFN [5], as the backbone, and introduce the self-challenge strategy [6] in our SCC system. The model is trained with the asymmetric loss [7] to predict section ID. The negative log probability is used as the anomaly score for detection.

2.2. Clustering-based System with Metric Evaluation

Regarding the CBM system, the log-Mel spectrograms of the sounds from the source and target domains are clustered to represent the two domains' centroid. The anomaly degree of unknown samples is calculated based on the distance metric, i.e., Mahalanobis distance. As the domain information of the test data is unavailable, the results obtained via distance metric calculation for each cluster are integrated as the anomaly score, using the ensemble learning strategy [4].

2.3. Ensemble System

The above two methods are integrated using an ensemble learning strategy [4]. The anomaly scores obtained by these two methods are normalized with zero mean and standard deviation. The final anomaly score is the average of the normalized scores of these two systems. We set the decision threshold as 0, i.e., normal when the anomaly score is smaller than 0, and abnormal when it is greater than 0.

3. EXPERIMENTS AND RESULTS

3.1. Experimental Setup

We adopt the development dataset and additional training dataset from the ToyADMOS dataset [8] and MIMII DG dataset [9] for model training and testing according to DCASE2022 Task 2. The dataset consists of the normal/anomalous operating sounds of seven types of toy/real machines, i.e., ToyCar, ToyTrain, fan, gearbox, bearing, slider, and valve.

Each machine includes six sections, and each section is divided into two domains, i.e., source and target. The audio signals are single-channel with around 10 seconds in length, and the sampling rate is 16 kHz. There are 990 training samples from the source domain and 10 training samples from the target domain. 200 samples from both domains are used as the testing set, i.e., 100 from the source domain and 100 from the target domain, whose domain information is not specified for testing.

In all our systems, the audio segment was loaded with the original sampling rate (16 kHz). The logarithm is taken through the Mel filter to obtain the 128 dimensional log-Mel spectrogram as the input, the length of the window (nFFT) is 1024 samples, and the hop length is set as 512 samples.

3.2. Performance Evaluation

Table 3: Performance of the proposed SCC system (%)

	ToyCar	ToyTrain	Bearing	Fan	Gearbox	Slider	Valve	Average
h-mean AUC-s	33.74	59.55	77.66	65.89	82.47	94.10	76.63	70.01
h-mean AUC-t	63.50	44.64	79.50	62.95	69.47	64.81	64.97	64.26
h-mean pAUC	49.91	49.62	64.28	60.95	64.21	63.22	63.32	59.36

Table 4: Performance of the proposed CBM system (%)

	ToyCar	ToyTrain	Bearing	Fan	Gearbox	Slider	Valve	Average
h-mean AUC-s	88.47	86.90	56.87	80.36	81.86	95.93	95.58	83.71
h-mean AUC-t	77.02	51.20	82.07	59.05	80.18	83.48	90.97	74.85
h-mean pAUC	60.14	59.37	54.25	59.59	65.29	76.08	72.92	63.95

The performance of the ensemble system is given in Table 5, where the weights for integrating the two systems are also provided.

Table 5: Performance of the ensemble system (%)

	ToyCar	ToyTrain	Bearing	Fan	Gearbox	Slider	Valve	Average
SCC weight	21.00	0.00	100.00	47.00	24.00	16.00	13.00	-
CBM weight	79.00	100.00	0.00	53.00	76.00	84.00	87.00	-
h-mean AUC-s	87.86	86.90	77.66	81.34	83.68	96.31	95.76	87.07
h-mean AUC-t	79.72	51.20	79.50	66.97	80.99	83.79	91.35	76.22
h-mean pAUC	61.99	59.37	64.28	63.39	66.99	77.09	74.19	66.76

In addition, we provide the performance comparison (i.e., average AUC-s/AUC-t/pAUC) among our systems and the baseline systems. The results are shown in Table 6. From this table, we can see that the ensemble system can obtain better performance as compared to the baseline systems and the proposed SCC system and CBM system.

Table 6: Performance comparison among our systems and the baseline systems (%)

	Baseline(AE)	Baseline(MobileNetV2)	SCC	CBM	**Ensemble system**
h-mean AUC-s	71.23	64.17	70.01	83.71	**87.07**
h-mean AUC-t	46.21	51.09	64.26	74.85	**76.22**
h-mean pAUC	53.91	55.85	59.36	63.95	**66.76**

4. CONCLUSION

In this technical report, we have presented our system submitted for DCASE2022 challenge Task 2, which is an ensemble system that integrates two proposed systems. The individual and ensembled systems are evaluated on the official dataset of DCASE2022 challenge Task 2. The results show that all our systems can outperform the baseline systems, and our ensemble system achieves the best detection performance among those methods.

5. REFERENCES

[1] Y. Kawaguchi, K. Imoto, Y. Koizumi, N. Harada, D. Niizumi, K. Dohi, R. Tanabe, H. Purohit, and T. Endo, "Description and discussion on DCASE2021 challenge task 2: Unsupervised anomalous sound detection for machine condition monitoring under domain shifted conditions," *arXiv preprint arXiv:2106.04492*, 2021.

[2] Y. Koizumi, Y. Kawaguchi, K. Imoto, T. Nakamura, Y. Nikaido, R. Tanabe, H. Purohit, K. Suefusa, T. Endo, M. Yasuda, *et al.*, "Description and discussion on DCASE2020 challenge task2: Unsupervised anomalous sound detection for machine condition monitoring," *arXiv preprint arXiv:2006.05822*, 2020.

[3] K. Dohi, K. Imoto, N. Harada, D. Niizumi, Y. Koizumi, T. Nishida, H. Purohit, 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 arXiv e-prints: 2206.05876*, 2022.

[4] O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 5, p. e1249, 2018.

[5] Y. Liu, J. Guan, Q. Zhu, and W. Wang, "Anomalous sound detection using spectral-temporal information fusion," in *Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 816–820.

[6] Z. Huang, H. Wang, E. P. Xing, and D. Huang, "Self-challenging improves cross-domain generalization," in *Proceedings of European Conference on Computer Vision (ECCV)*. Springer, 2020, pp. 124–140.

[7] T. Ridnik, E. Ben-Baruch, N. Zamir, A. Noy, I. Friedman, M. Protter, and L. Zelnik-Manor, "Asymmetric loss for multi-label classification," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021, pp. 82–91.

[8] 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 *Proceedings of the 6th Detection and Classification of Acoustic Scenes and Events (DCASE) Workshop*, Barcelona, Spain, November 2021, pp. 1–5.

[9] 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," *arXiv e-prints: 2205.13879*, 2022.