ANOMALOUS SOUND DETECTION BY USING LOCAL OUTLIER FACTOR AND GAUSSIAN MIXTURE MODEL

Technical Report

*Kazuki Morita**, *Tomohiko Yano**, *Khai Q. Tran**

Intelligent Systems Laboratory, SECOM CO.,LTD.
{morita-ka,tomo-yano,ku-chan}@secom.co.jp

ABSTRACT

In this report, we introduce our methods and results of the anomalous sound detection in DCASE2020 task2. We attempted to detect anomalous sound without using deep learning methods. Precisely, we first extracted features by applying principal component analysis (PCA) to the log-mel spectrogram of the sound signal. Then we used Local Outlier Factor (LOF) and Gaussian Mixture Model (GMM) as the anomaly detection method. Our experiment showed the proposed method improved the Area Under Curve (AUC) to 0.8706 and the partial Area Under Curve(pAUC) to 0.7403 compared to the baseline system on development dataset.

Index Terms— Anomaly Sound Detection, Local Outlier Factor, Gaussian Mixture Model

1. INTRODUCTION

Recently, deep learning has been showing remarkable performance in various fields such as speech recognition, acoustic event detection and speech synthesis. In fact, the baseline system[1] in DCASE2020 Task2[2] uses an Auto-Encoder algorithm. However, when dealing with this task under limited computing resources and a limited amount of data, it is also very interesting to consider how well conventional anomaly detection methods work. In this report, we use Local Outlier Factor (LOF) and Gaussian Mixture Model (GMM) for anomalous sound detection.

The structure of this report is as below. In section 2, we explain the anomalous sound detection methods we used for this task. In section 3, we show the results of our evaluation experiment and its consideration. In section 4, we summarize our investigation results for this task. In section 5, we describe the model we are submitting.

2. ANOMALOUS SOUND DETECTION METHOD

We show the anomalous sound detection method as Figure 1. In this section, we describe feature extraction, anomaly detection algorithm and scoring.

2.1. Feature Extraction

We use the log-mel spectrogram as the features. We set the parameters as follows.

- frame length is 64ms.

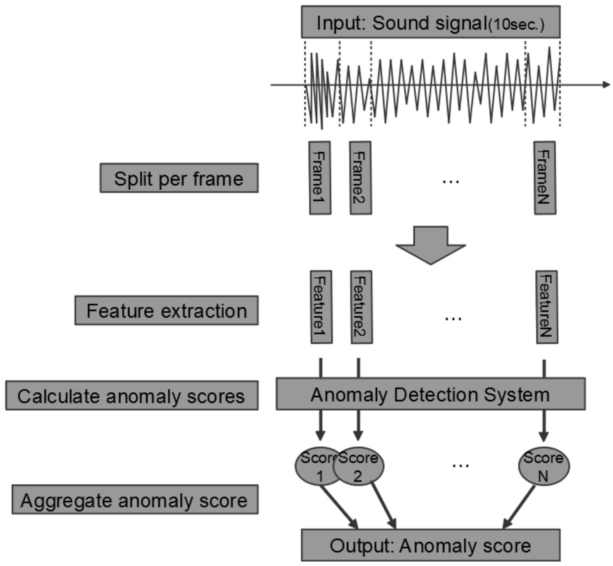

Figure 1: Algorithm Overview

- log mel-band energies (128bands)
- 5 frames are concatenated
- 640 dimensions are input features

In our method, principal component analysis(PCA) was used to reduce the computational cost, and then the dimensionality reduced data were used for anomaly detection. Although there are many kinds of models that could be used for anomaly detection, in this report, two methods, LOF and GMM were considered.

2.2. Anomaly Detection Algorithm

Our proposed approach is LOF and GMM.

Local Outlier Factor(LOF) [3]

Local Outlier Factor is one of the anomaly detection methods we employed. This method is based on local density, which is the density of k-neighboring feature values. When a feature is anomalous, the difference is large between the local density of the anomaly and the neighboring feature. In this report, we use the outputs of LOF as the anomaly score. We set the number of neighbors to 20.

*Equal contribution.

Gaussian Mixture Model(GMM) [4]

Gaussian Mixture Model is the second of the anomaly detection methods we employed. First, we estimate parameters of the gaussian mixture model by using training features. We then calculate the likelihood of the target feature. When a feature is anomalous, the likelihood is small. In this report, we use negative log-likelihood as the anomaly score, and we set number of mixture components to 32, and the co-variance type to full.

2.3. Scoring

After calculating anomaly scores per frame, we aggregate them into an anomaly score of the sound signal. In this report, we use the mean of the frame scores, and the variance of the frame scores. The former shows the amount of the anomaly score, and latter shows the fluctuation in the anomaly score.

3. EXPERIMENTS

3.1. Experimental conditions

10-sec length audio (monaural, 16 kHz) was sampled from machinery sound sources. There are six types of machines (Machine Type); ToyConveyor, ToyCar[1], fan, pump, slider and valve[5]. For each Machine Type, there are several IDs (Machine ID). We constructed the anomalous sound detection model for each Machine ID in Table 1. We used scikit-learn[6] for the implementation.

Table 1: Experimental conditions

Model name	Feature dimensions	Algorithm	Scoring
LOF-mean	40	LOF	mean
LOF-var	40	LOF	variance
GMM-mean	80	GMM	mean
GMM-var	80	GMM	variance

4. RESULTS AND CONSIDERATION

The results are shown in Table 2 and 3. From the results, we can see that the average score approach for calculating anomaly score worked well with the ToyCar, ToyConveyor, fan and pump Machine Types. On the other hand, the variance score approach worked well with the slider and valve Machine Types. It is considered that this difference is due to the characteristic of the anomalous sound sources.

For an anomalous sound of ToyCar, ToyConveyor, fan and pump, because the anomaly score is high for almost every frame of sound signal, we calculated the average score of all frames and used it as the anomaly score of sound signal. For an anomalous sound of slider and valve, because the anomaly score of each frame of sound signal varies dramatically large, we calculated the variance score of all frames and used it as the anomaly score of sound signal.

5. CONCLUSION

In this report, LOF and GMM were used to detect anomalous sounds. The proposed method gave an AUC score of 0.8706 and a pAUC score of 0.7403. Recently, deep learning methods have been seeing widespread use because of their effectiveness. However, in this report, we demonstrated that with limited computing resources and a limited amount of data, non-deep learning methods can also be effective. the effectiveness of non-deep learning methods.

6. SUBMISSIONS

In this report, we submit three anomalous sound detection systems; Morita_SECOM_task2_1, Morita_SECOM_task2_2 and Morita_SECOM_task2_3.

Morita_SECOM_task2_1 uses the ”LOF-mean” model given in Table 1; Morita_SECOM_task2_2 uses the ”GMM-var” model. For Morita_SECOM_task2_3, we used the model that gave the largest AUC and pAUC for each Machine Type. We select the algorithm method from either GMM or LOF, and the scoring method from mean, percentile, or variance. When we select the variance, we also select the range of anomaly scores we use. For example, variance (70%-100%) indicates that only the largest 30% of all anomaly scores in sound signal were used. The conditions of Morita_SECOM_task2_3 are shown in Table 4. The AUC and pAUC results from using these parameters are shown under ”Our best” in Table 2 and 3.

Table 4: Parameters of Morita_SECOM_task2_3

Machine type	Algorithm	Scoring
ToyCar	LOF	percentile(30%)
ToyConveyor	GMM	percentile(10%)
fan	LOF	percentile(30%)
pump	LOF	percentile(30%)
slider	GMM	variance(0 - 100%)
valve	GMM	variance(70 - 100%)

Table 2: AUC Results of Development Dataset (%)

MachineType	MachineID	Baseline	LOF-mean (Submission 1)	LOF-var	GMM-mean	GMM-var (Submission 2)	Our best (Submission 3)
ToyCar	1	81.36	**94.13**	88.38	91.94	75.51	94.48
	2	85.97	**96.57**	92.32	96.04	79.05	97.10
	3	63.30	**90.52**	80.24	85.60	60.54	91.38
	4	84.45	**99.63**	94.66	99.09	76.42	99.78
	Avg.	78.77	**95.21**	88.90	93.17	72.88	95.69
ToyConveyor	1	78.07	79.48	72.21	**80.14**	58.46	82.79
	2	64.16	67.04	61.20	**67.26**	53.13	68.84
	3	75.35	**85.47**	75.46	84.62	58.88	87.34
	Avg.	72.53	77.33	69.62	**77.34**	56.82	79.66
fan	0	54.41	**65.23**	58.91	55.75	49.69	67.40
	2	73.40	**86.57**	82.12	78.53	73.30	87.09
	4	61.61	**77.90**	72.53	61.16	63.23	79.33
	6	73.92	**95.52**	82.36	89.89	78.07	96.19
	Avg.	65.83	**81.30**	73.98	71.33	66.07	82.50
pump	0	67.15	**73.87**	74.26	71.43	65.13	72.46
	2	61.53	**68.28**	54.87	67.64	43.23	70.41
	4	88.33	93.82	74.10	**96.11**	83.44	94.18
	6	74.55	**85.45**	73.35	80.00	66.98	87.09
	Avg.	72.89	**80.36**	69.15	78.80	64.70	81.04
slider	0	96.19	95.92	**98.57**	92.72	97.67	97.67
	2	78.97	**79.11**	78.67	77.94	75.94	75.94
	4	94.30	83.52	81.06	85.56	**96.91**	96.91
	6	69.59	65.31	63.96	59.13	**94.24**	94.24
	Avg.	84.76	80.96	80.56	78.84	**91.19**	91.19
valve	0	68.76	79.19	96.39	61.64	**98.29**	99.35
	2	68.18	56.77	56.18	55.59	**87.36**	91.80
	4	74.30	75.39	81.28	62.35	**91.94**	94.20
	6	53.90	67.65	**81.94**	49.68	77.83	80.74
	Avg.	66.28	69.75	78.95	57.31	**88.85**	91.52
Total	Avg.	73.55	**80.97**	77.17	76.08	74.14	87.25

Table 3: pAUC Results of Development Dataset (%)

MachineType	MachineID	Baseline	LOF-mean (Submission 1)	LOF-var	GMM-mean	GMM-var (Submission 2)	Our best (Submission 3)
ToyCar	1	68.40	**82.58**	70.16	79.81	63.16	83.59
	2	77.72	**92.63**	81.56	89.67	62.84	93.49
	3	55.21	**76.46**	60.57	69.44	51.39	78.82
	4	68.97	**98.46**	81.25	96.12	59.84	98.97
	Avg.	67.58	**87.53**	73.38	83.76	59.31	88.72
ToyConveyor	1	64.25	64.53	55.21	**65.15**	52.93	68.32
	2	**56.01**	53.25	50.47	55.91	50.45	57.15
	3	61.03	65.54	55.36	**66.46**	52.52	69.24
	Avg.	60.43	61.11	53.68	**62.51**	51.97	64.90
fan	0	49.37	**51.66**	49.52	50.34	48.70	52.26
	2	54.81	**65.77**	54.00	59.84	52.37	67.35
	4	53.26	**57.59**	52.86	53.02	51.47	57.74
	6	52.35	**80.81**	53.52	73.49	50.27	81.88
	Avg.	52.45	**63.96**	52.47	59.18	50.70	64.81
pump	0	56.74	58.15	**63.89**	55.36	61.02	55.83
	2	58.10	63.35	54.10	**64.01**	52.02	64.20
	4	67.10	75.16	53.11	**83.11**	68.68	76.53
	6	58.02	**66.41**	56.81	65.17	54.90	66.51
	Avg.	59.99	65.77	56.98	**66.91**	59.16	65.77
slider	0	81.44	80.10	**93.76**	69.56	89.44	89.44
	2	63.68	66.73	**66.94**	64.62	62.59	62.59
	4	71.98	56.24	51.01	63.93	**92.08**	92.08
	6	49.02	50.15	49.20	50.03	**75.40**	75.40
	Avg.	66.53	63.30	65.23	62.03	**79.88**	79.88
valve	0	51.70	51.97	81.65	53.56	**91.91**	96.59
	2	51.83	49.04	51.01	51.23	**65.88**	77.24
	4	51.97	51.23	53.33	50.88	**79.25**	83.99
	6	48.43	49.34	55.13	49.30	**56.97**	58.42
	Avg.	50.98	50.39	60.28	51.24	**73.50**	79.06
Total	Avg.	59.63	**65.53**	60.63	64.35	62.87	74.24

7. REFERENCES

[1] 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. [Online]. Available: https://ieeexplore.ieee.org/document/8937164

[2] 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 *arXiv e-prints: 2006.05822*, June 2020, pp. 1–4. [Online]. Available: https://arxiv.org/abs/2006.05822

[3] M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "Lof: Identifying density-based local outliers," in *PROCEEDINGS OF THE 2000 ACM SIGMOD INTERNATIONAL CONFERENCE ON MANAGEMENT OF DATA*. ACM, 2000, pp. 93–104.

[4] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006.

[5] 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 *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2019 Workshop (DCASE2019)*, November 2019, pp. 209–213. [Online]. Available: http://dcase.community/documents/workshop2019/proceedings/DCASE2019Workshop_Purohit_21.pdf

[6] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.