REFRAMING UNSUPERVISED MACHINE CONDITION MONITORING AS A SUPERVISED CLASSIFICATION TASK WITH OUTLIER-EXPOSED CLASSIFIERS

Technical Report

Paul Primus

Institute of Computational Perception (CP-JKU)
Johannes Kepler University, Austria

ABSTRACT

This technical report contains a detailed summary of our submissions to the *Unsupervised Detection of Anomalous Sounds for Machine Condition Monitoring* (MCM) Task of the *IEEE AASP Challenge on Detection and Classification of Acoustic Scenes and Events 2020* (DCASE). The goal of acoustic MCM is to identify whether a sound emitted from a machine is normal or anomalous. In contrast to the task coordinator's claim that 'this task cannot be solved as a simple classification problem,' we show that a simple binary classifier substantially outperforms the provided unsupervised Autoencoder baseline across all machine types and instances, if *outliers* i.e., various other recordings, are available. In addition to this technical description, we release our implementation for reproducibility.

Index Terms— Unsupervised Anomaly Detection, Outlier-Exposed Classifiers, Machine Condition Monitoring, DCASE2020

1. INTRODUCTION

- anomaly detection as a problem class
- explain the difficulty of obtaining anomalous samples - mostly because destroying stuff that works is expensive and anomalies are scarce.
- one instance of this problem is *Machine Condition Monitoring*

2. RELATED WORK

- Taxonomy of Anomaly Detection Methods (TODO: find a survey paper)
- Challenge Baseline Paper [1]
- Very complicated MCM approach [2]
- AUC, pAUC [3]
- AUC Loss Equation [2]
- Rethinking Assumptions in Anomaly Detection [4]

3. OUTLIER-EXPOSED CLASSIFIERS

Using random outliers somehow improves classification results – still not sure why. Anyhow, we investigate under which conditions we can use outlier exposed classifiers for unsupervised MCM.

An outlier expose-classifier is a binary classifier trained in a one-vs-everything fashion.

4. EXPERIMENTS

4.1. Experimental Setup

4.1.1. Dataset & Pre-Processing

DataSets

- ToyADMOS [1]
- MIMII [5]

Same pre-processing as baseline:

4.1.2. Network Architecture

ResNet			
Type	#K	KS 1	KS 2
Conv	$c \cdot 2^0$	5	
BN	-	-	
RB	$c \cdot 2^0$	3	1
Max Pool	-	2	-
RB	$c \cdot 2^0$	3	3
Max Pool	-	2	-
RB	$c \cdot 2^0$	3	a
RB	$c \cdot 2^0$	3	b
Max Pool	-	2	-
RB	$c \cdot 2^1$	1	1
RB	$c \cdot 2^2$	1	1
RB	$c \cdot 2^2$	1	1
Conv	1	1	-
BN	-	-	-
GAP	-	-	-

Residual Block (RB)	
Type	KS
Conv	KS 1
BN	
Conv	KS 2
BN	
Add Input	

Table 1: Model Architecture by [?] for experiments with the acoustic scenes dataset. #K and KS are the number of kernels and kernel size, respectively. Residual Blocks (RB) consist of two Convolutional (Conv) layers with #K kernels, each followed by a Batch Normalization (BN) layer. GAP is a Global Average Pooling Layer.

ID	a	b	c	LR	loss
0	1	1	64	10^{-4}	BCE
1	1	1	64	10^{-4}	BCE

Table 2: Network Variants

[6]

4.1.3. Training

4.2. Results

Baseline [3]
Fancy table or bar plot containing test results on development set for all submissions.

5. CONCLUSION

not fancy post processing no machine-type specific feature engineering no-ensembling

OEC *easily* beat the unsupervised baseline.

6. 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, 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 Lang. Process.*, vol. 27, no. 1, pp. 212–224, 2019. [Online]. Available: https://doi.org/10.1109/TASLP.2018.2877258

[3] 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

[4] L. Ruff, R. A. Vandermeulen, B. J. Franks, K. Müller, and M. Kloft, "Rethinking assumptions in deep anomaly detection," *CoRR*, vol. abs/2006.00339, 2020. [Online]. Available: https://arxiv.org/abs/2006.00339

[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] K. Koutini, H. Eghbal-zadeh, and G. Widmer, "Receptive-field-regularized CNN variants for acoustic scene classification," *CoRR*, vol. abs/1909.02859, 2019. [Online]. Available: http://arxiv.org/abs/1909.02859