THE STUDY OF ANOMALOUS MACHINE SOUND DETECTION BASED ON CYCLOSTATIONARITY MODEL

Technical Report

Vladimir Klychnikov
Bauman Moscow State Technical University
2-nd Baumanskaya St., 5, bld 1, Moscow, 105005, Russia
klychnikovvv@student.bmstu.ru

Dmitriy Lapin*
Bauman Moscow State Technical University
2-nd Baumanskaya St., 5, bld 1, Moscow, 105005, Russia
lapindv@bmstu.ru

Mark Hubbatulin
Bauman Moscow State Technical University
2-nd Baumanskaya St., 5, bld 1, Moscow, 105005, Russia
khmea15u755@student.bmstu.ru

ABSTRACT

In industrial predictive maintenance, one of the most important direction in Industry 4.0, machine monitoring and diagnostics is critical part of its operation. Non-contact acoustic data gathering is particular interest because of high ergonomics and low costs. The general method for processing such type of data is anomalous sound detection. This method allows express diagnostics of machines and units with minimum integrations. Based on DCASE 2020 Challenge, the study of the proposed method was presented. Problem description with physical interpretation and model elements review was conducted. Model based on Winger-Ville transform with architecture improvement and ensemble score calculation was developed. Model results on provided development dataset were calculated. Discussion of model results with assumptions for further research and development was shown. Conclusion about present study and future work was received.

Index Terms — anomalous sound detection, machine condition monitoring, cyclostationarity processes, pseudo Winger-Ville transform, deep learning, autoencoder

1. INTRODUCTION

One of the most important direction in Industry 4.0 is predictive maintenance based on monitoring and diagnostic of machines and tools [1, 2]. Non-contact data mining [3] is particular interest because it provides non-destructive testing on important industrial units. Non-contact vibroacoustic data gathering is particular interest because of high ergonomics, low cost and ability to collect data from one or more objects simultaneously [4-10]. The general method for processing such type of data is anomalous sound detection [11-13], that allows express diagnostics of machines and units with minimum integrations and without complete knowledge of how the units work. Methods of machine health monitoring are presented on figure 1.

* Thanks to LCC «Ctrl2GO» for funding.

Figure 1: Methods of machine health monitoring.

The goal of this study, based on DCASE 2020 Challenge [14], is to analyze and synthesize anomalous machine sound detection model on provided datasets [15-17]. To achieve this goal, the following tasks need to be accomplished:

• to analyze problem description;
• to detail model elements and merge them;
• to calculate results of model;
• to discuss calculated result for suggestions;
• to draw conclusions about future work.

2. PROBLEM DESCRIPTION

Challenge task was explained in detail in the enclosed materials [14]. Provided datasets were described in accompanying papers [15-17]. This analysis is aimed at the physical interpretation of data mining methods and model elements review, that partially entered in the final model.

2.1. Physical interpretation

Based on the description of the experiments and our experience in diagnostics on machine sound data mining, a number of peculiarities are identified that determined the model concept.

Mechanical systems in this challenge — e.g. fan, slider, valve, pump, toy car, toy conveyer and their parts — sustain periodic motion of their components which in turn periodically modulate the vibration or noise they radiate. The cyclostationary property covers a rich statistical typology of signals including periodic signals and stationary random signals as particular cases. In particular,
because it is intuitive and physically meaningful, the approach relies as much as possible on the key concept of the energy conveyed by a signal: this will make easier the introduction of the various types of densities that characterize cyclostationary signals in the time, the frequency and the cyclic domains [18].

However, in order to use the cyclostationarity theory, it is necessary to know more about experiments on recording datasets, in particular the microphone parameters and technical specification of machines and their modes of operation.

### 2.2. Model elements review

The following hypotheses are considered individually and together according to numbered elements of general pipeline, shown on figure 2.

![General pipeline of anomalous sound detection](image)

Figure 2: General pipeline of anomalous sound detection.

Data augmentation for expansion datasets on machines and on id (1 and 2 fig. 1). Since the operation sounds are presented for different machines, it may not be correct to create additional spectrograms in different modes of operation. Different augmentations showed completely different results for different machines and id. And since the estimated dataset will have an id that the model has not yet seen, it has been decided to reject the augmentation.

Features for more information from signal (3 fig. 1). An attempt to use signal features, such as - min, max, mean, 25 and 75 quartiles and fundamental frequency estimation [19] – does not improve the quality of anomaly detection, but on the contrary, worsens it for the similar cause as for signal features.

Different types of autoencoder anomaly detectors (8 fig. 1). Convolutional, variational and denoising autoencoders, pre-train encoder VGGish [25] were used. But in the current study these solutions did not allow to increase the final result.

After analysis, the highest result was achieved by window pseudo Winger-Ville transform together with mel-scaled spectrogram in concatenated frames on reshape baseline autoencoder.

Next, the selected elements and their merge will be discussed in more detail.

### 3. MODEL ARCHITECTURE

#### 3.1. Feature extraction

For this study of anomalous machine sound detection of cyclostationarity processes it is important to have high resolution on spectrogram, both in time and frequency. Therefore, the Window Pseudo Winger-Ville distribution (WPWVD) [26] was chosen for the basic transformation. WPWVD spectrograms are visually distinctly different than SFFT spectrograms. WPDV spectrograms are too slow for streaming audio compared to SFFT ones: they take about 50 times longer to compute. WPWVD is a better choice than SFFT when studying audio in one detail, where the highest quality TF graph is required. To generate a sample-accurate (1024 band) WPWVD spectrogram in real time would require about 16 cores. We did not have enough computing resources to calculate the full spectrograms. So, we had to downsample it to 2 kHz. For this reason, in the final model, we also used Mel-scaled spectrogram to find defects at frequencies above 2 kHz. Final configurations of acoustic features is shown on table 1.

![Example of spectrograms before and after pre-processing on FAN machine](image)

Table 1: Configurations of acoustic features.

<table>
<thead>
<tr>
<th>Feature</th>
<th>WPWVD</th>
<th>Mel spectrogram</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window length</td>
<td>20</td>
<td>1024</td>
</tr>
<tr>
<td>Hop length</td>
<td></td>
<td>512</td>
</tr>
<tr>
<td>Low Frequency</td>
<td>0 Hz</td>
<td>0 Hz</td>
</tr>
<tr>
<td>High Frequency</td>
<td>2 kHz</td>
<td>8 kHz</td>
</tr>
<tr>
<td>Feature dim</td>
<td>900</td>
<td>650</td>
</tr>
</tbody>
</table>

#### 3.2. Architecture improvement

To improve the overall quality of the model it is necessary to perform a tuning of baseline autoencoder.

Experimentally, a rational number of neurons in layers has been established. For example, the dependence of average AUC on the number of neurons in dense and bottleneck layers on figure 4 is presented.
Values of reshape layers, presented on figure 5, is chosen for current study.

Figure 5: Reshape autoencoders for: a) WPWVD; b) Mel spectrogram

3.3. Score calculation

The idea of the ensemble is to use two models that individually give common results, but together can improve it. The scheme is shown on figure 4.

Table 2: Score of ensemble model.

<table>
<thead>
<tr>
<th>Machine</th>
<th>Mel spectrogram</th>
<th>Mel spectrogram</th>
<th>Final score</th>
</tr>
</thead>
<tbody>
<tr>
<td>FUND</td>
<td>0.66042</td>
<td>0.51133</td>
<td>0.41270</td>
</tr>
<tr>
<td>VALUE</td>
<td>0.58613</td>
<td>0.53877</td>
<td>0.51148</td>
</tr>
<tr>
<td>WP</td>
<td>0.56159</td>
<td>0.53067</td>
<td>0.62054</td>
</tr>
<tr>
<td>FAN</td>
<td>0.53182</td>
<td>0.52318</td>
<td>0.41159</td>
</tr>
<tr>
<td>SLIDER</td>
<td>0.53222</td>
<td>0.51063</td>
<td>0.49799</td>
</tr>
<tr>
<td>TDYTOY</td>
<td>0.53177</td>
<td>0.51059</td>
<td>0.62054</td>
</tr>
<tr>
<td>CONVER</td>
<td>0.52783</td>
<td>0.50937</td>
<td>0.62054</td>
</tr>
</tbody>
</table>

4. RESULTS

Based on the model developed, the following results are obtained, as presented in Table 2.
5. DISCUSSION

A number of the following assumptions can be made based on the results. These assumptions will make it possible to further improve the concept.

5.1. Harmonics distribution as definition of anomalies

The increase in model quality by high resolution spectral picture together with reshape of layers means the efficiency of cyclostationary theory application. It also means criticality of recovery process from input data of harmonic distribution. Combinations of linear dependent harmonics on fundamental frequency for selected machine parts give more information about anomalies. This allows a significant increase in the interpretability of anomaly detection. However, it imposes a number of requirements for data collection experiments and additional machine research.

5.2. Approach to weak classification

Based on previous assumption, the transition from anomaly detection to a weak classification model should be considered. In the framework of the conducted research it became obvious that it is difficult to generalize the problem of detecting anomalies from different machines. Therefore, in further research the concept will be modified to solve the problem of weak classification.

6. CONCLUSION

The following conclusions can be drawn from the study:

- the problem description made it possible to establish a basic approach in the model and to carry out an effective review of the methods;
- the model architecture allowed, with limited computing resources, to implement the required calculation methods;
- the results shown that the model have also high scores on development dataset;
- the discussion led to the following steps in the development of the solution.

Future work will be aimed on improvement of anomalous sound detection for machine condition monitoring by using harmonics distribution for weak anomaly classification.

7. ACKNOWLEDGMENT

Our team thanks LCC «Ctrl2GO» for funding the current study.

8. REFERENCES


[17] Harsh Purohit, Ryo Tanabe, Takeshi Ichige, Takashi Endo, Yuki Nikaido, Kaori Suefusa, and Yohei Kawaguchi,


