Final Degree Thesis on Audiovisual Systems Engineering

Dept. of Information and Communication Technologies
Universitat Pompeu Fabra

Real-time audio classification to detect scientific hardware malfunction.

Genís Plaja i Roglans

Supervisor: Dr. Xavier Serra i Casals

June 2020

Contents

Dedication

(Optional, if used placed on a right page next to an empty left page)

I would like to dedicate this work to...

Acknowledgement

(Optional, if used placed on a right page next to an empty left page)

I would like to express my sincere gratitude to:

- My supervisor
- My co-supervisor
- My family

Abstract

Automatic audio recognition has a wide range of applications, however using it for scientific research purposes is still a barely explored field, and this fact becomes more apparent when talking about scientific hardware malfunction detection. This work is a deep analysis on how to use automatic audio classification to detect and prevent failures on scientific hardware, focusing the efforts on a specific use case and performing a machine learning process until the very beginning, analyzing and discussing each step to make it robust and enough scalable to extrapolate the results to different use cases. We want to make strong emphasis on the fact that each step of the machine learning procedure has been carefully worked out, from finding a potential problem to developing a suitable dataset and training and evaluating the model, so that at the end of the process we can define a handy methodology that works for every use case that fits in the proposed scope. The last purpose of this project is to prove that a physical implementation of the application can be functional and versatile, adapting the machine learning algorithm to be performed in real-time, and thinking about a prototype that is suitable in terms of the market. This dissemination is a technical version of Sounds of Science (November 2018 - January 2020), an outreach project developed under the supervision of the Music Technology Group (MTG) and supported by the Maria de Maeztu Strategic Research Program which consists on explaining science through the matter of sound and rising an online community where researchers can browse, discover and share sounds from their scientific works done in their departments.

Keywords: Audio classification; Hardware malfunction; Machine learning

Chapter 1

Introduction

This project is a step-by-step work on a machine learning problem to solve a use case

1.1 Motivation and scope

Since November 2018, I have taken part of Sounds of Science **CITAR WEB**, an outreach project that aims on explaining and promoting science through the matter of sound. To achieve this flux of scientific research information through its sounds, an online community based on Freesound was created, where researchers from all over the world are now able to browse and discover international research groups and projects by their soundscapes, and upload sounds recorded in their labs, together with a brief explanation of their department and its related works. To introduce a sound into the Sounds of Science community, it is as easy as adding *soundsofscience* tag to the sound.

Sound event detection for applications that fit in the scope of scientific research sound environments is still a barely explored field. Moreover, this project focuses on hardware malfunction detection, which is a practise that presents a clear real world need (for more information see section XXX). **APROFUNDIR MÉS!!**

As a final point, nowadays is difficult to find datasets with data from certain lab-

oratory elements although for some disciplines are essential. This project pretends to focus on a use case that presents this shortage of related works and material and to contribute, if possible, to datasets under the Creative Commons License such as the Freesound Datasets (FSD).

In terms of personal motivation for this topic, from a wide range of options that the combination of audio and technology has, the one which I am most interested in is audio signal processing and retrieval for machine and deep learning applications, so I decided to find an interesting idea related to that field. After spending one year and a half working on the Sounds of Science project, I learned a lot about this topic and I found it very interesting, moreover, I obtained a huge background of it. That is why this Thesis Project can be seen as a technical fork of Sounds of Science. Besides, audio and technology are my main areas of interest, that is also the reason why I have chosen MTG for the Final Degree Thesis project and I took the opportunity to be a student candidate for an internship in this department.

1.2 Objectives

The overall aim of this work is to solve a machine learning problem paying close attention to carefully analyze and discuss each of the involved steps in the process. Therefore, first of all we want to obtain a deeper understanding of the use case and choosing correctly the classes to represent it. Developing a suitable dataset that involves enough cases and scenarios and is accurately labelled is also an important objective of this dissemination, as we need to provide the algorithm with a great robustness. Regarding the coding part of the project, we intend to implement a versatile, robust and functional audio recognition system that is capable to solve with great results the proposed problem. We expect the accuracy of the machine learning algorithm to be very close to 100%, above 99% if possible, because for a lower accuracy, it is hard to display it as a market product. Hence, to achieve these results, an accurate analysis of the machine learning mathematical and programming content is required. We can set up the machine learning algorithm in several different ways, but the purpose is to use the most precise and efficient one.

The final objective of our work is to implement a real-time audio classification system, hence we aim on finding an optimized and effective implementation that can detect hardware malfunction.

Finally, regarding each step of the process, we want to design a complete methodology from the very beginning that is adaptable to several different use cases within the chosen range of scope.

Personally, I want to obtain as much knowledge as I can on the technical aspects of my project: machine and deep learning, music and sound information retrieval, building and managing datasets. I want to become familiar with open-source libraries and sets of algorithms related to my work, such as Tensorflow and Essentia, and understand the theoretical background is behind them.

1.3 Structure of the Report

i. **Methods.** We cover the procedures followed to work on each step of the proposed machine learning problem. First of all, we review the state of the art related to our dissemination. Secondly, we advance through the designed methodology from the very beginning to explain in detail the used methods along the project timeline.

   To know about the methods used in this project recall to Chapter 2.

ii. **Results.** We state the obtained results of our research. The results are presented in a chronological order, in order to preserve the followed methodology.

   All the obtained results together with their details are stated at Chapter 3.

iii. **Discussion.**

Chapter 2

Methods

Some introduction here...

2.1 Literature review

The most important factor to identify when implementing a real life based product is a necessity. Hence, finding a valuable state of the art is a basic step, because if there is someone working on our topic of concern, it would mean that a clear necessity exists.

Regarding our range of scope, a Czech company called Neuron Soundwave, founded in 2016 by Pavel Konečný in Praga. Their slogan is literally *"our powerful AI analyzes sound patterns to provide unparalleled insight into how mechanical systems operate, so potential failures can be detected early".*

2.2 Use case description

Design and Modelling of Transition Metal Catalyzed Reactions (DIMOCAT) is an interdisciplinary research group of the Institute of Computational Chemistry and Catalysis (IQCC) at the University of Girona made up of organic and computational chemists. DIMOCAT's laboratories are located in the Science Faculty at the Montilivi Campus in Girona city.

One of the group's research lines is the development of new methodologies for the formation of highly complex organic molecules. On the way to synthesising the new compounds, analytical techniques are required in order to facilitate the synthetic pathway. For example, the efficient separation of multicomponent mixtures of volatile organics as well as the identification of the new compounds is essential in the research group. One of these techniques is gas chromatography–mass spectrometry (GC-MS). GC-MS is an analytical technique that couples the separation capabilities of GC with the detection properties of MS to provide greater efficiency in analyzing samples. While GC can separate volatile components in a sample, MS helps fragment the components and identify them on the basis of their mass.

2.2.1 Problem identification

To perform the GC-MS technique the researchers from DIMOCAT have a GC-MS System. This machine needs to be permanently on, as for the optimal functioning of the mass spectrometer it is important to have a high quality vacuum system, which is usually achieved with an oil vacuum pump. Furthermore, to perform the experiments and measures, the machine reaches really precise high temperatures.

In addition, this kind of instrumentation is best suited to a clean laboratory that is not in contact with chemicals and solvents. For this reason, GC-MS systems are usually installed in laboratories apart from day-to-day workplace laboratories. In the DIMOCAT group, a GC-MS was acquired in 2016 at a cost of 35,000 EUR and installed in a laboratory on the same floor as the rest of the DIMOCAT laboratories but separated by a distance of 300 meters.

In contrast to the advantages of having the GC-MS in a separate laboratory, there are also some disadvantages:

i. If GC-MS stops working, air flows inside the cabin so the vacuum is no longer valid, and the temperature goes drastically down. Re-achieving the ideal conditions is a high expensive and lasting task. What is more, if there are chemical materials inside the cabin and the system stops, they get damaged and the ex-

periments are no longer valid. As a final point and the more important one, if the machine is not prepared correctly to be turn it on again or it suffers from sudden power changes, it can get highly damaged. For this reasons, is very important to achieve a fast response to this problem.

ii. It is important for the user to know when a sample run finishes and the instrument becomes available to use again, to achieve the best performance of the machine and prevent the experiments to fail.

2.2.2 Proposed solution

Taking into account the problem description, if the researchers from DIMOCAT could count with a system to receive an alarm when the GC-MS stops working, the consequences of this failure would be highly reduced, as they would be able to attend sooner to the lab where the GC-MS is and lessen the damage. For the case (ii), if the lab members could know exactly when the GC-MS system is cooling, they would be able to attend in time to the experiments to make sure that the result is the optimal, optimizing the usage of the machine and their own working hours. Moreover, they would be able to calculate the right time to move from one laboratory to the other without leaving any experiment unsupervised, taking into account that researcher from DIMOCAT usually work with inflammable, poisonous or explosive material.

2.3 Materials

Still in development...

Chapter 3

Results

This is an example paragraph. As you can see, the main text uses a font size of 12 pt and a line spacing of 1.5. Neither the paragraphs nor the first lines of paragraphs should be indented.

There is no very strict page limit. Your number of pages will be strongly influenced by the size and total number of your figures and tables. It is recommended staying within 30-50 pages. Do not try to fill as many pages as you can. Longer theses are not necessarily of higher quality and of more non-redundant content than shorter theses. Certainly, a master thesis of 15 pages is too short, and a master thesis of 100 pages is too long.

3.1 Tables and graphics

This is an example paragraph. As you can see, the main text uses a font size of 12 pt and a line spacing of 1.5. Neither the paragraphs nor the first lines of paragraphs should be indented.

There is no very strict page limit. Your number of pages will be strongly influenced by the size and total number of your figures and tables. It is recommended staying within 30-50 pages. Do not try to fill as many pages as you can. Longer theses are not necessarily of higher quality and of more non-redundant content than shorter

theses. Certainly, a master thesis of 15 pages is too short, and a master thesis of 100 pages is too long.

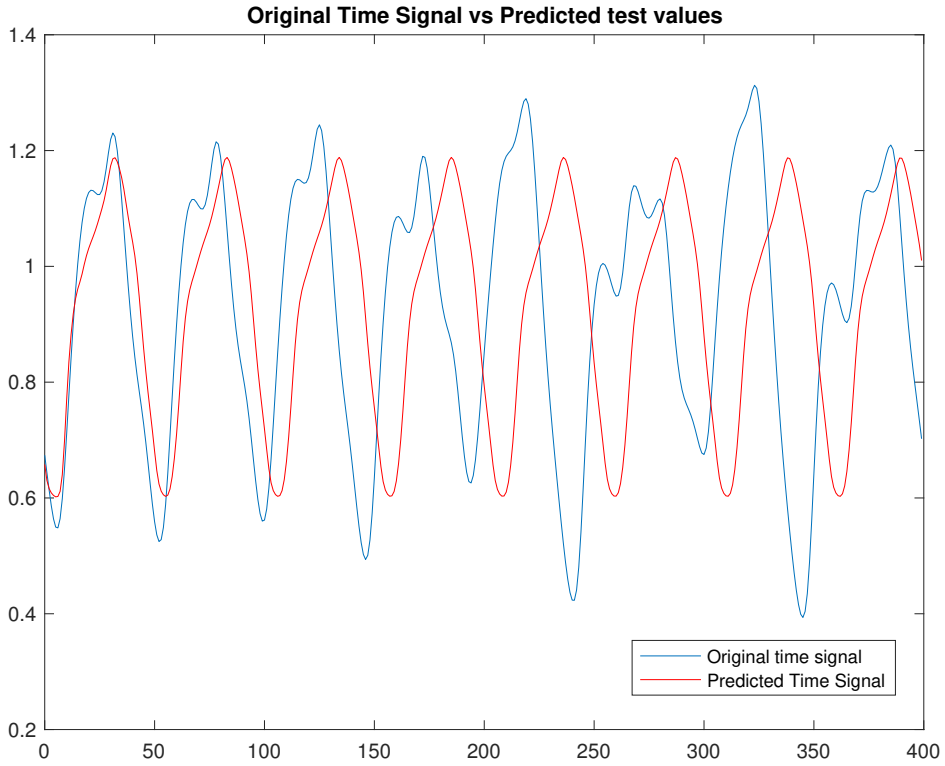

Figure 1: This is an example of a figure and its caption.

Table 1: This is an example of a table and its caption.

PCA	**Residual mean (in absolute values)**
Original PCA	0.1267
PCA on Centroid 1	0.1249
PCA on Centroid 2	0.1214

Chapter 4

Discussion

This is an example paragraph. As you can see, the main text uses a font size of 12 pt and a line spacing of 1.5. Neither the paragraphs nor the first lines of paragraphs should be indented.

There is no very strict page limit. Your number of pages will be strongly influenced by the size and total number of your figures and tables. It is recommended staying within 30-50 pages. Do not try to fill as many pages as you can. Longer theses are not necessarily of higher quality and of more non-redundant content than shorter theses. Certainly, a master thesis of 15 pages is too short, and a master thesis of 100 pages is too long.

4.1 Discussion

4.2 Conclusions

List of Figures

List of Tables

Bibliography

Appendix A

First Appendix

This is an example paragraph. As you can see, the main text uses a font size of 12 pt and a line spacing of 1.5. Neither the paragraphs nor the first lines of paragraphs should be indented.

There is no very strict page limit. Your number of pages will be strongly influenced by the size and total number of your figures and tables. It is recommended staying within 30-50 pages. Do not try to fill as many pages as you can. Longer theses are not necessarily of higher quality and of more non-redundant content than shorter theses. Certainly, a master thesis of 15 pages is too short, and a master thesis of 100 pages is too long.

Appendix B

Second Appendix