|  |  |  |
| --- | --- | --- |
| **Machine Learning for Noise Reduction in old audio records** | | |
| ACRONYM |  | ML4NR |
| PROPOSED BY |  | Carl Haber for the Lawrence Berkeley National Laboratory |
| STUDENT-S |  | Benoît Ruffray |
| PROFESSOR-S |  | Frédéric Bapst, Jean Hennebert |
| EXPERT-S |  | Vito Grisanti, Emeka Mosanya |
| No |  | B19I19 |
| TYPE |  | Bachelor Thesis |
| CONTACT |  | benoit.ruffray@hotmail.com |

**Context**

Audio recording has a long history. It started in 1877 with Thomas Edison’s phonograph, able to physically record sounds on cylinders made out of wax or tinfoil and to play them back. Those materials are brittle, and they wear out with each play back. Over the years, multiple improvement were made on the concept, changing the records’ shape and material, until the disc we know and use. However, it is those old brittle records that interest us.

The Lawrence Berkeley National Laboratory is home of an important historical project : the preservation of old audio records. These cylinders and discs contain historical data such as presidential speeches, native American interviews, traditional songs, and so on. Since they wear out when played physically, they had to find a way to read them without touching them.

IRENE is an imaging machine able to digitize any audio support with high resolution 2D and 3D cameras. Having high fidelity pictures of the discs and cylinders assures preservation. They then started to search for ways to play back those records using only the images. 

Figure : Part of an imaged disc. White parts are flat, black parts are slopped.

**Current state**

The team at the Laboratory uses a deterministic algorithm to detect the groove edges, simulate the needle following it, and compute the resulting sound wave. However defects, scratches, dust, and even the texture add a noise on the image that is also detected by their algorithm. This results in a high-pitched background noise when simulating a play back.

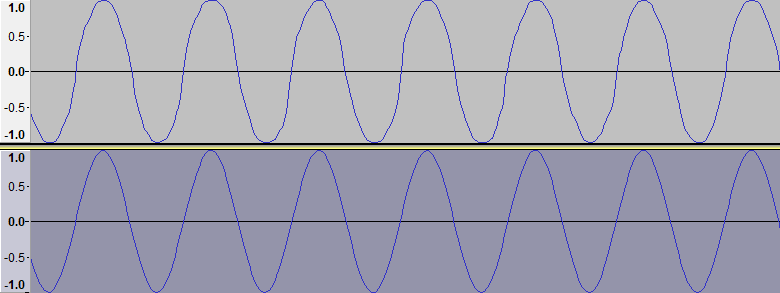
**Machine Learning to the rescue**

In recent years, machine learning has shown impressive results in a large variety of tasks, and can certainly help here. The idea of ML4NR project is to train a model to generate a sound wave by only looking at the disc image. This is developed in multiple steps :

* Step 1 : train on clean groove images artificially generated with clear sound as goal.
* Step 2 : train on noisy artificial groove images, still using the clear sound as a goal.
* Step 3 : adapt to the disc images, train more if necessary.

A big part of the project is also the auto-formation in machine learning, as we haven’t practiced it a lot.

**Developement**

****We mainly focused on one type of neural network in this project : convolutional neural networks. Their advantage is the ability to learn shapes and features at different scales, which seems good for our problem. The main challenge is finding the right parameters for training : what size should the convolution filters be, how many layers of them, do we use dropout or not, etc. Deep learning is still a dark art, and the only way to find the right combination of parameters is to try.

Top : generated sound from noisy image

Bottom : perfect audio

**Results**

Our convolutional network shows promising results in steps 1 and 2, generating a sound wave close to the truth even when the image is noisy (black rectangles hidding things, Gaussian noise). However, we couldn’t find the right combination of parameters for actual disc images, and the generated audio is mainly background noise. This still open doors for future work in the same direction, as the input data of step 2 isn’t so different from step 3.