Audio Processing

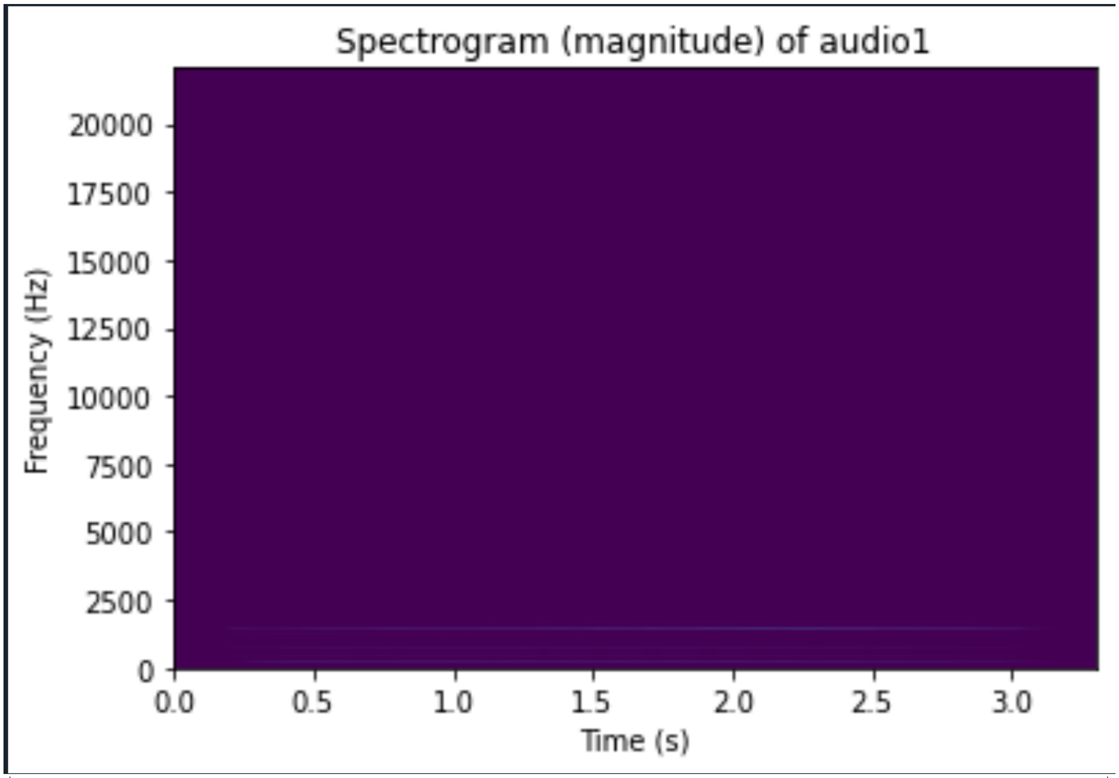
Ex02 – Report

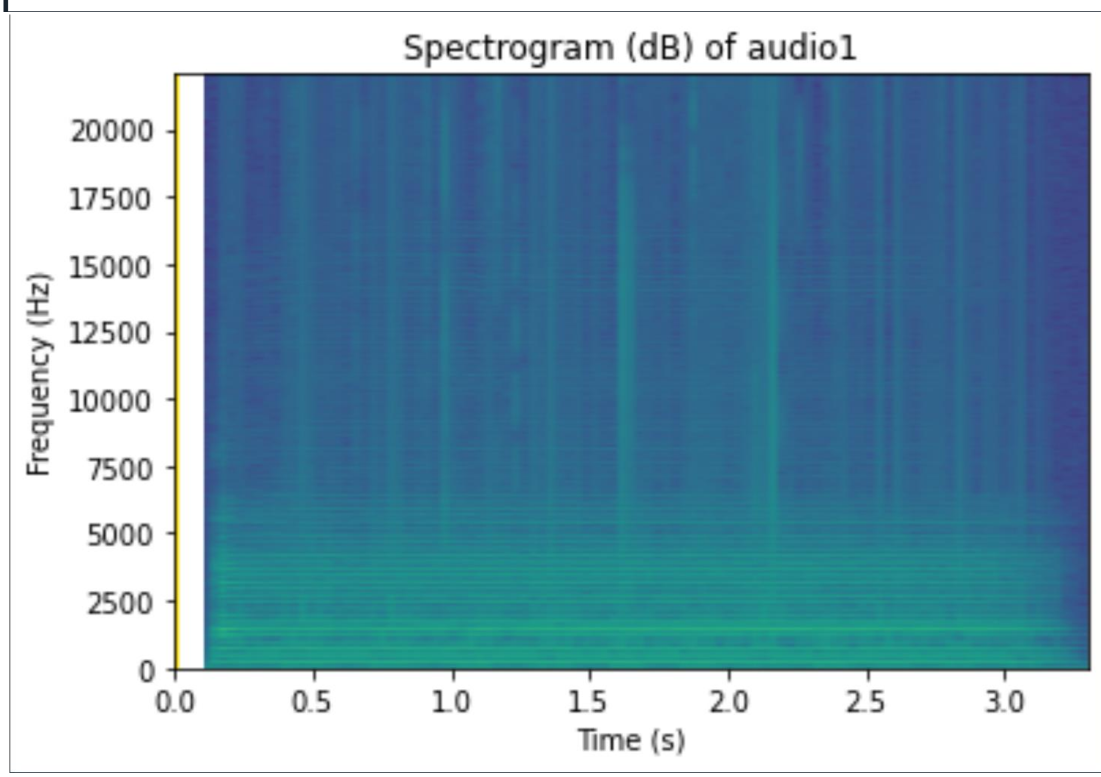
Student: Anh Huy Bui

ID: 293257

Problem 1:

c.

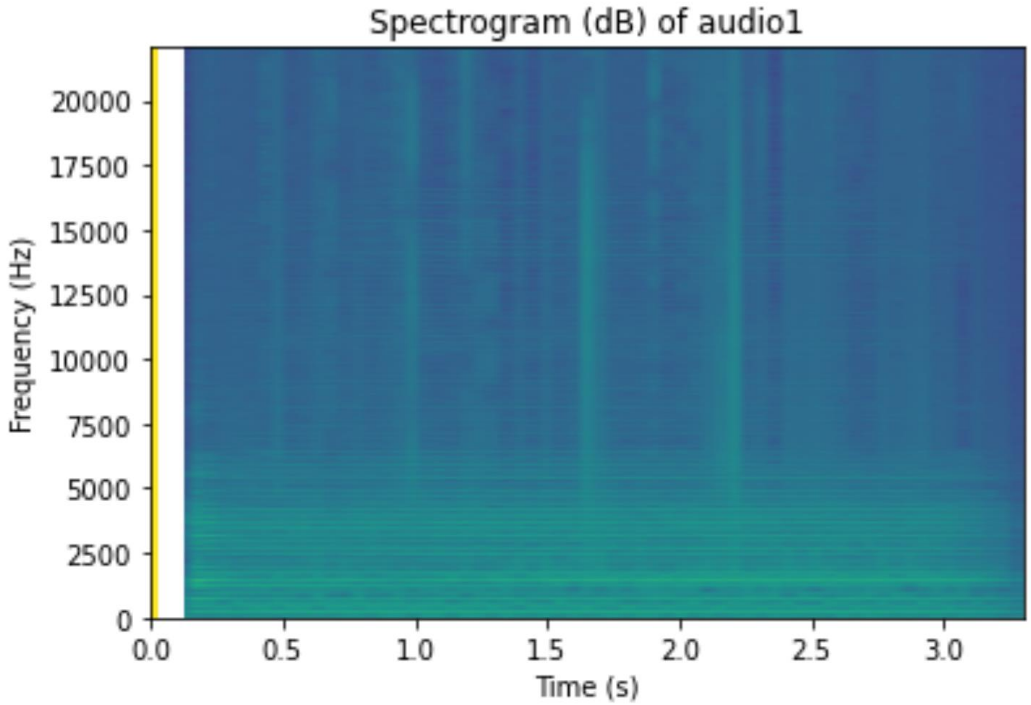


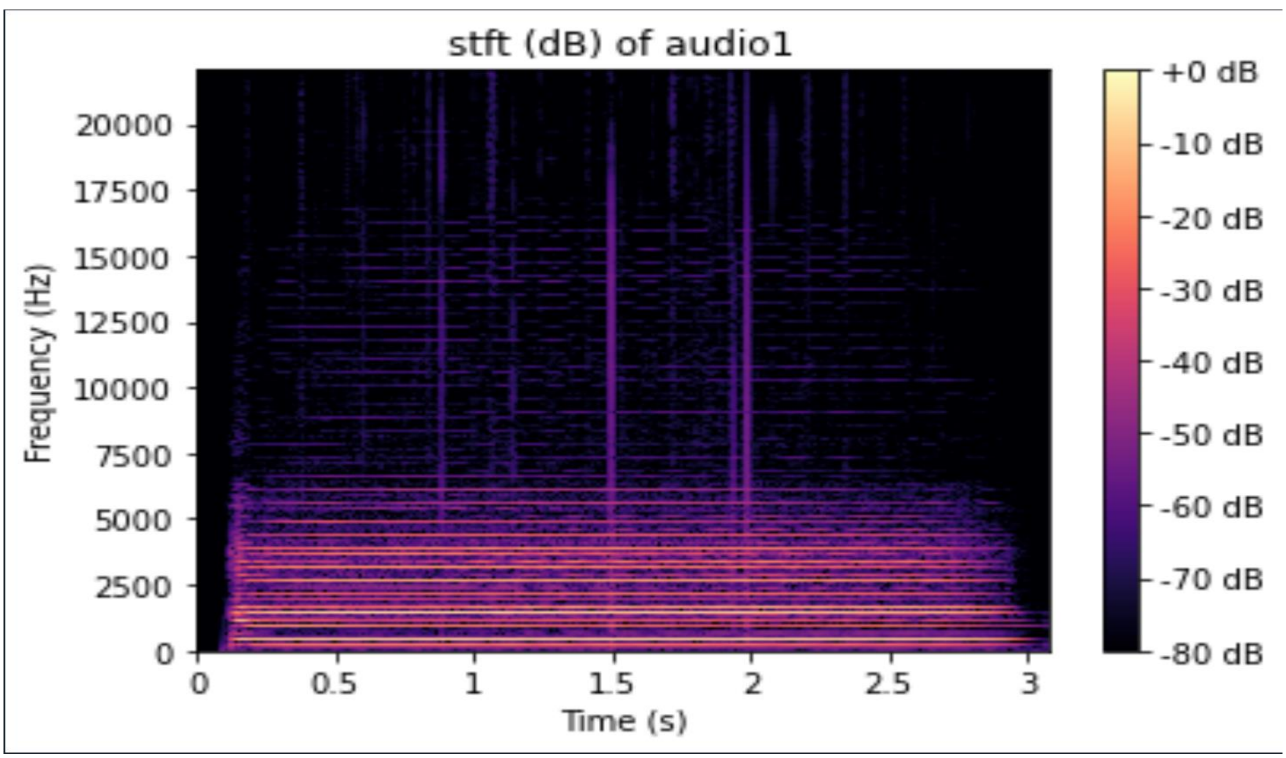


It can be seen form the graphs that spectrogram in logarithm scale is much easier to observe. This is because the magnitude of origin signal is too low, while logarithm scale is a relatively scale based on the relation among all frames of the audio.

Problem 2:

a.





I use librosa.stft to calculate. Both are configured with a window length as 100ms. However, librosa Fs is default 22050, so I did some calculation to get correct frame length.
As a brief look, the 2 graphs display quite similarly, although there is a blank column in my own spectrogram. This is because I did not solve the log10(0) problem. Beside that, it seems ok in my opinion.

b.

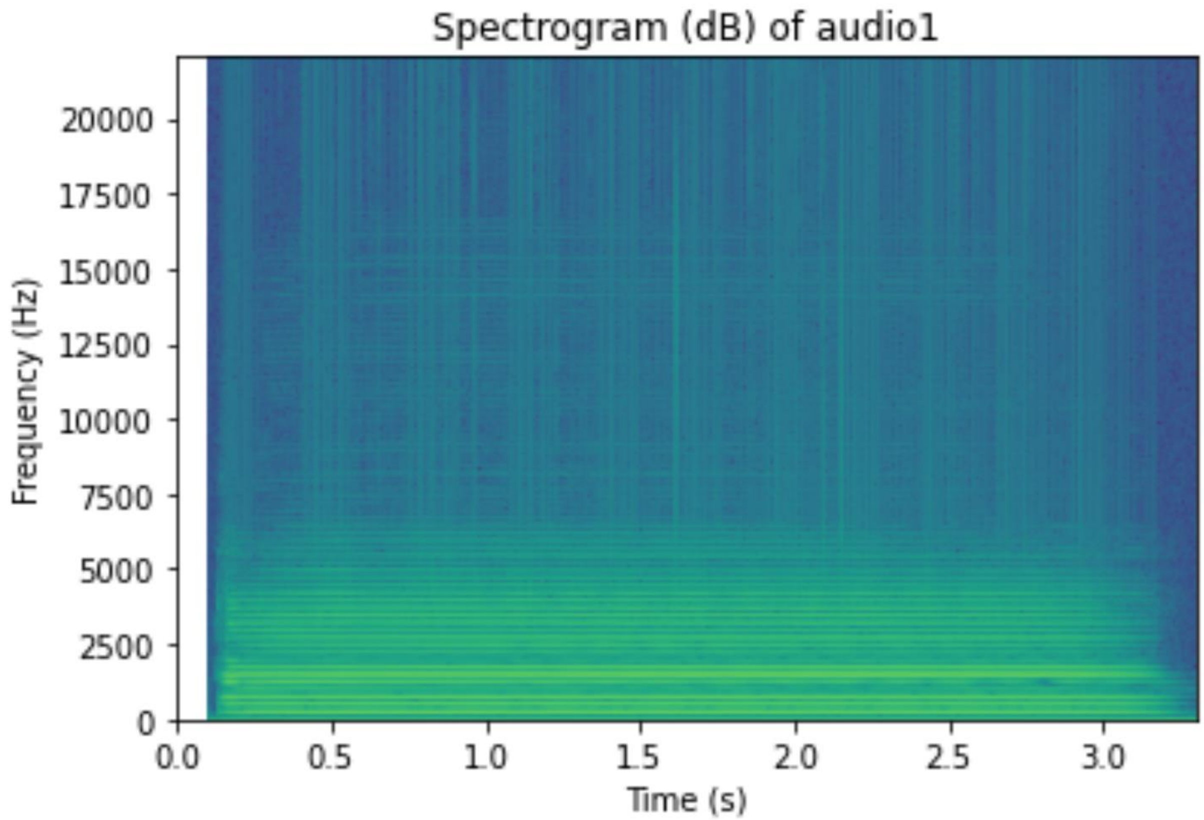


Figure 2b.1. Spectrogram of audio1 with window length = 16ms

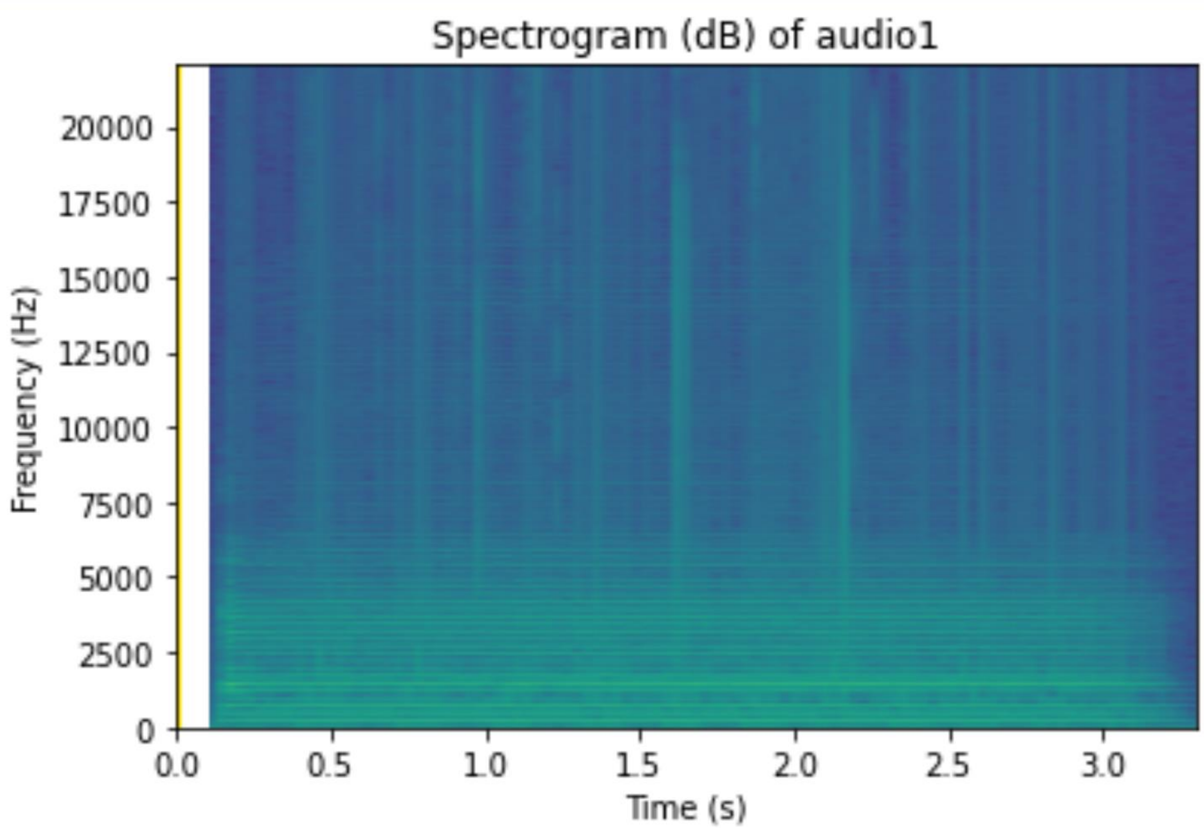


Figure 2b.2. Spectrogram of audio1 with window length = 64ms

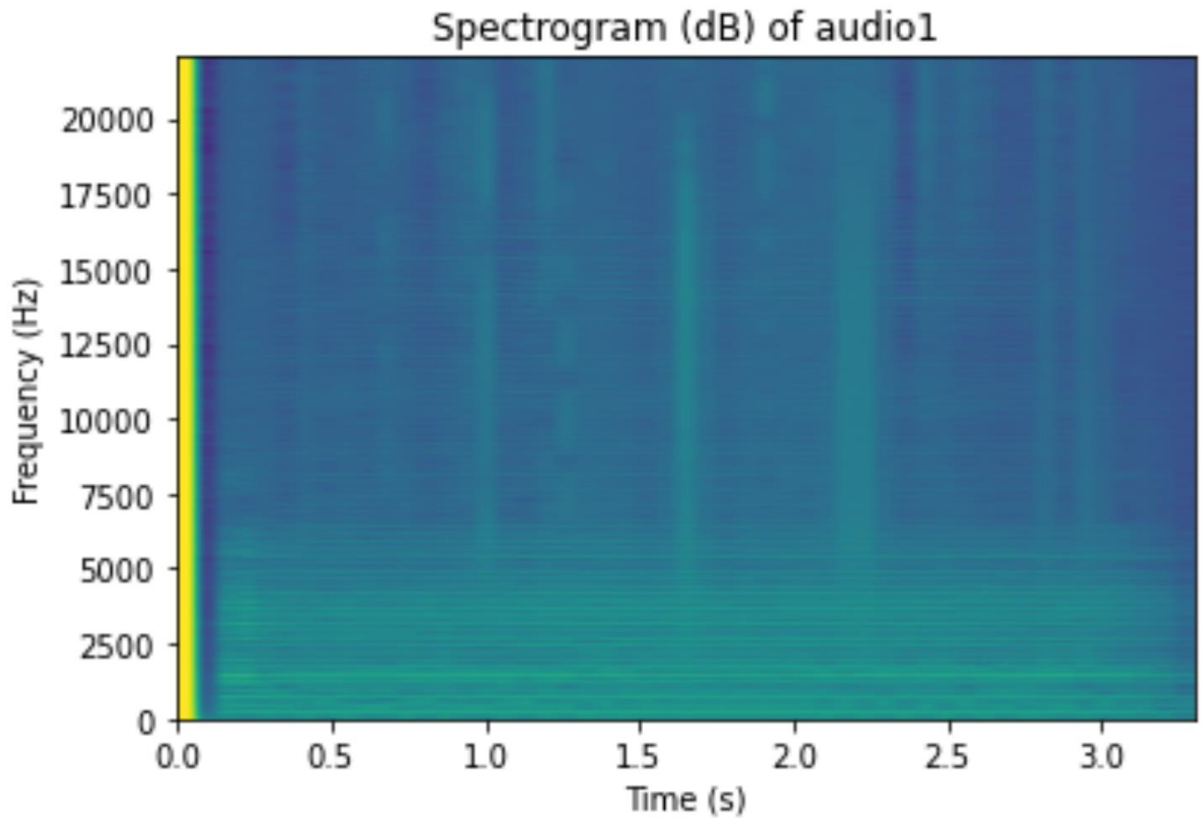
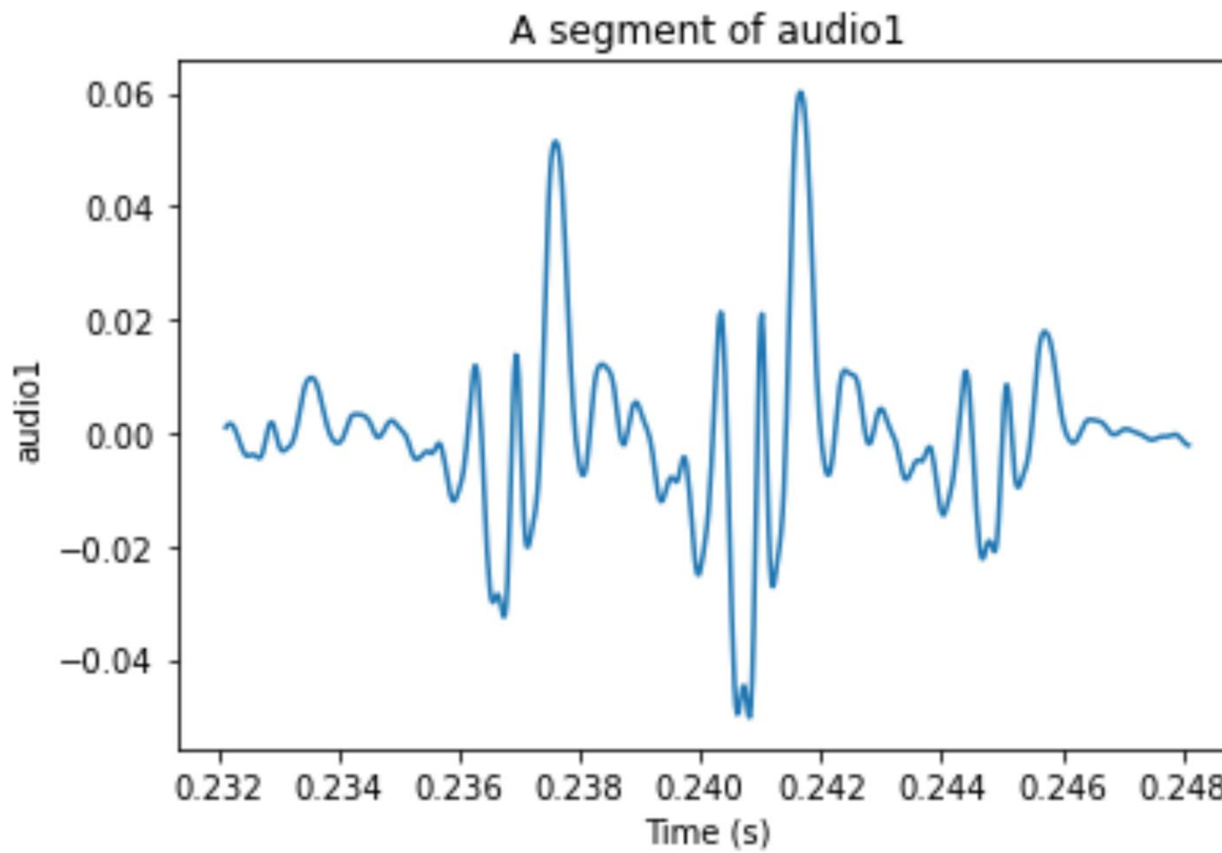


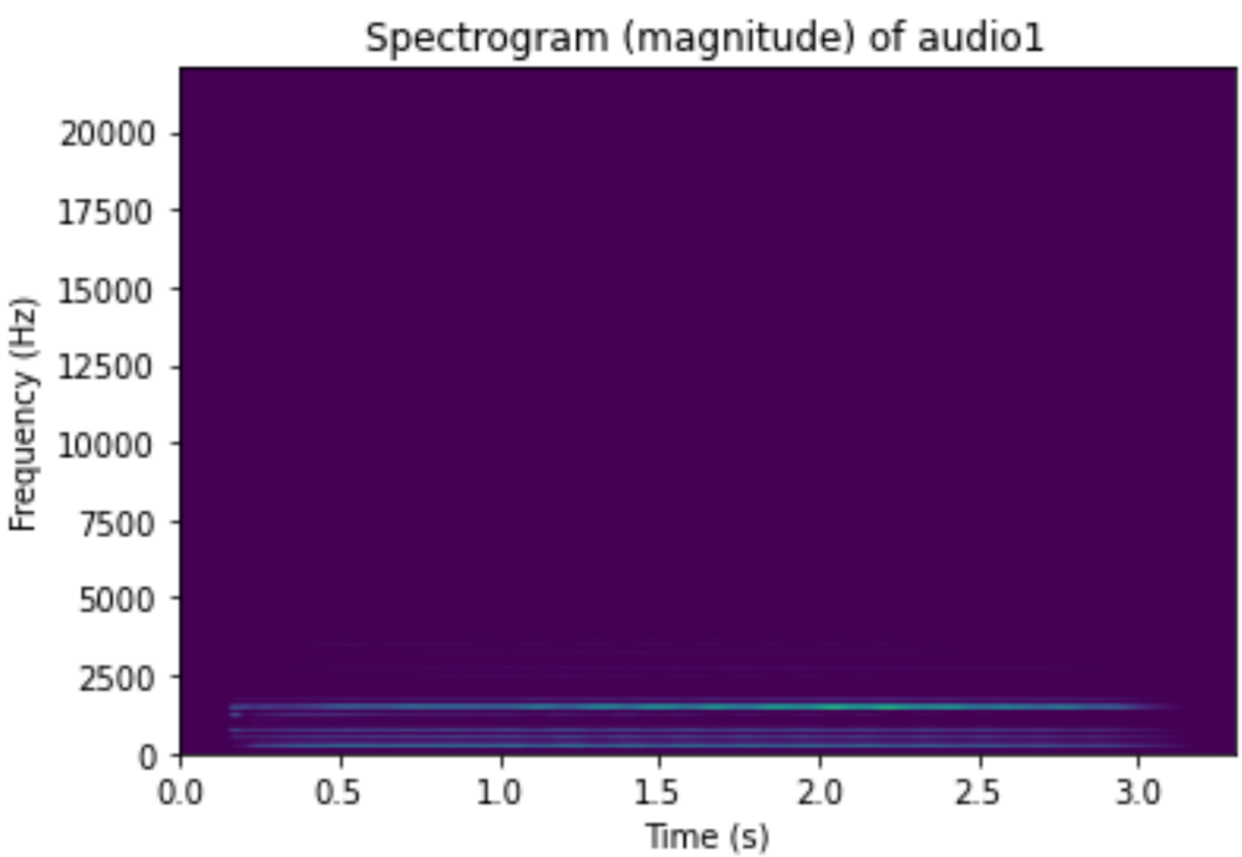
Figure 2b.3. Spectrogram of audio1 with window length = 128ms

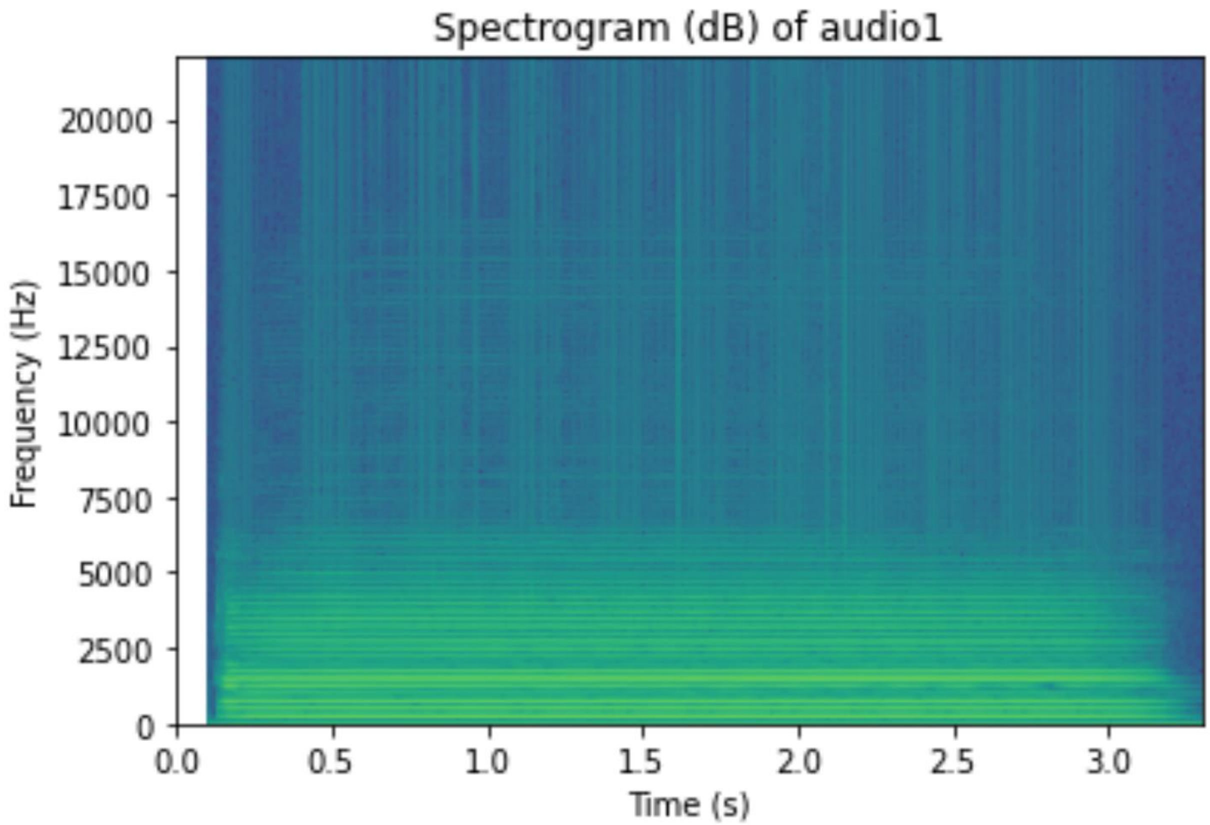
c. From my observation, spectrogram window length 128 has low resolution in time axis, while 16ms one is better.
Because every segment of 16ms one is much more detailed. However, spectrogram window length 16ms lost some data in frequency. In my opinion capturing a 16ms segment (or 62.5Hz in other words) may cause data lost in some low frequency.

All output figures from code: *(Signal – Window length)*

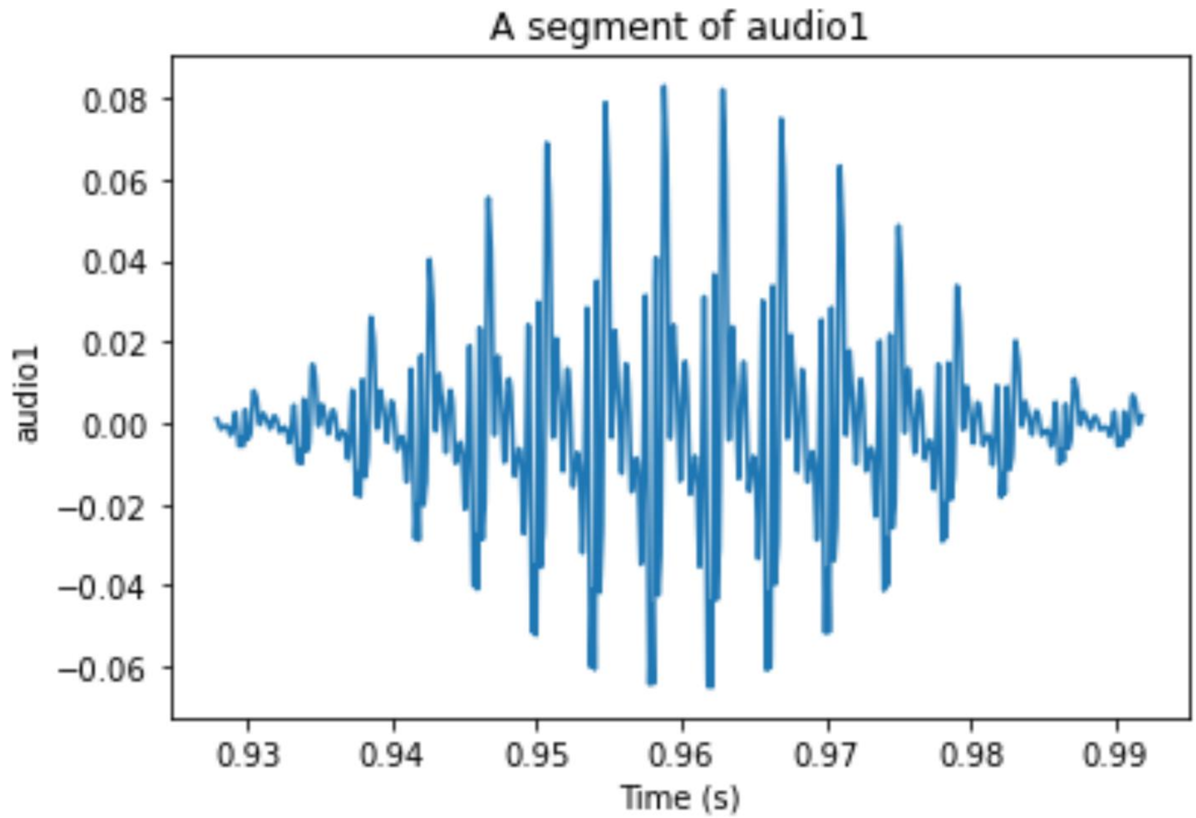
Audio1 - 16ms

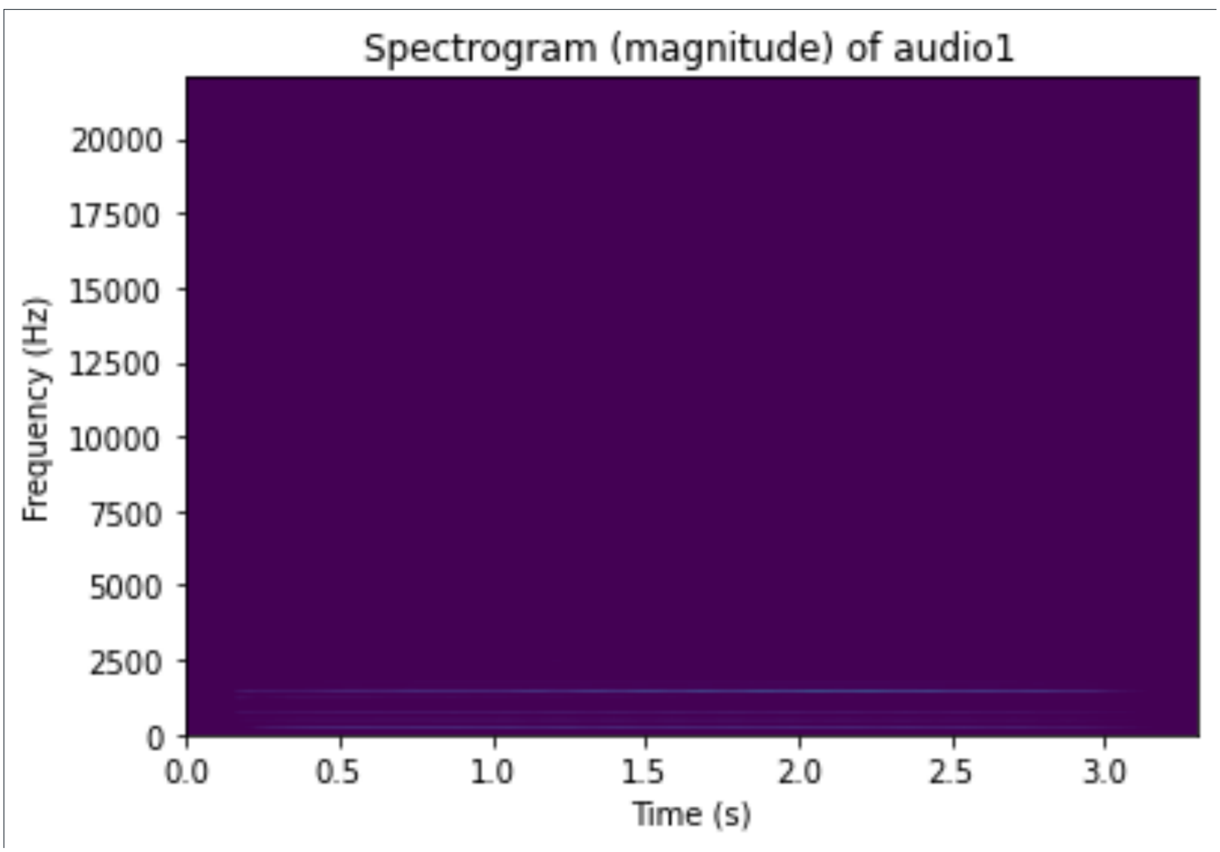


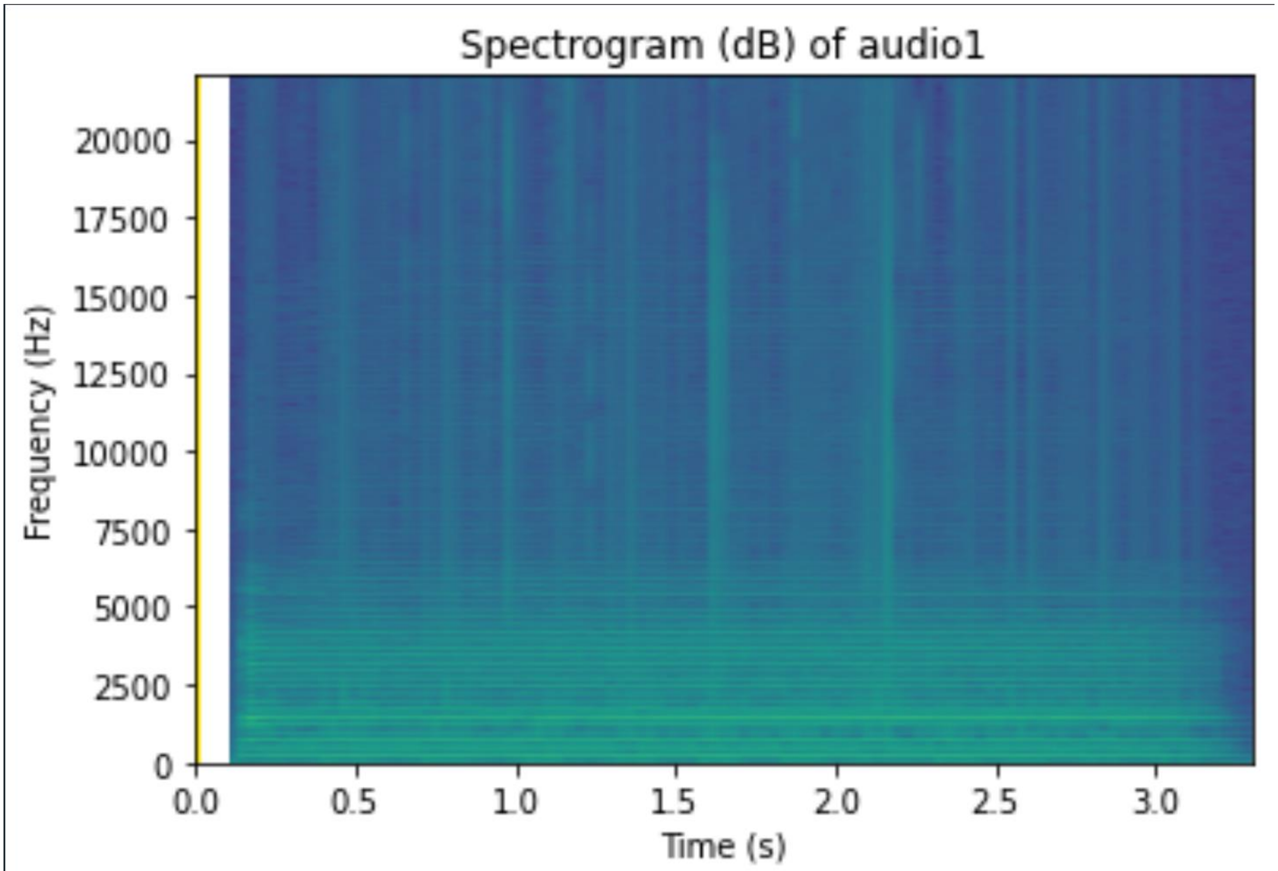



Spectrogram (dB) of audio1
Frequency (Hz)
20000
17500
15000
12500
10000
7500
5000
2500
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Time (s)

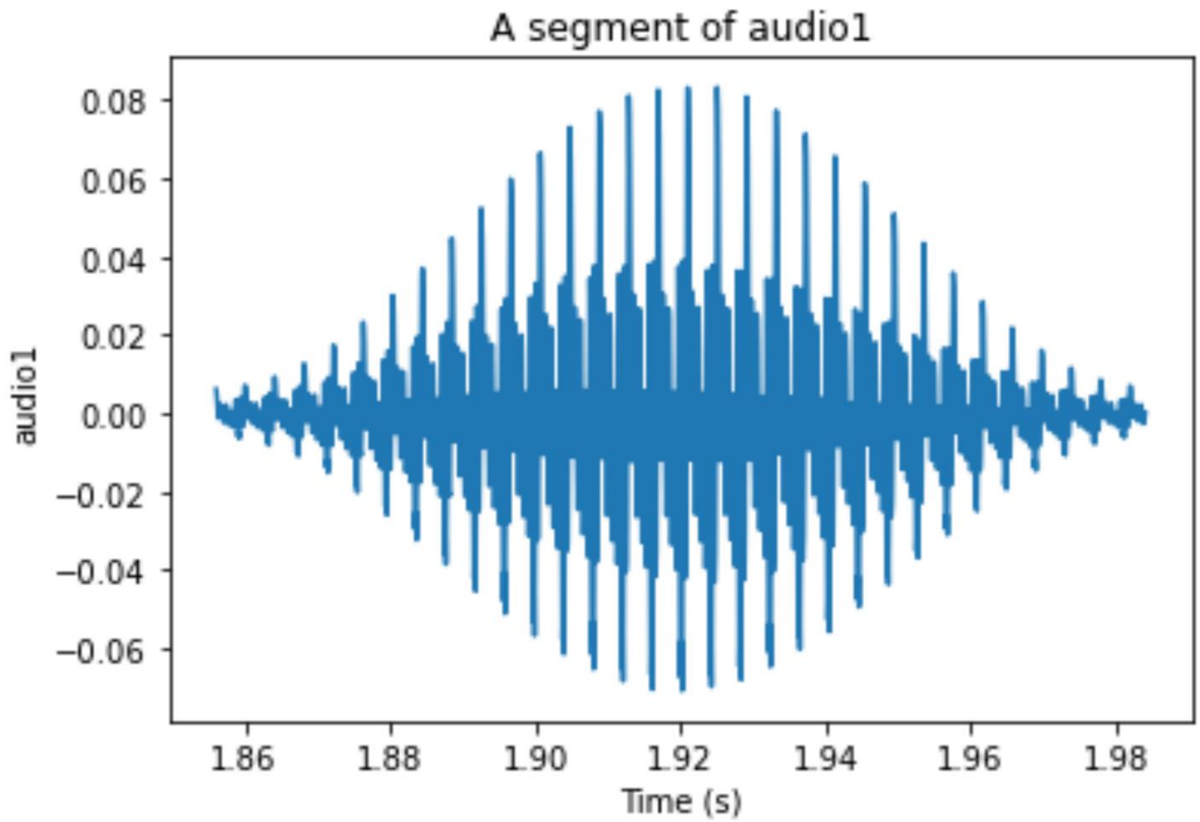
Audio1 - 64ms

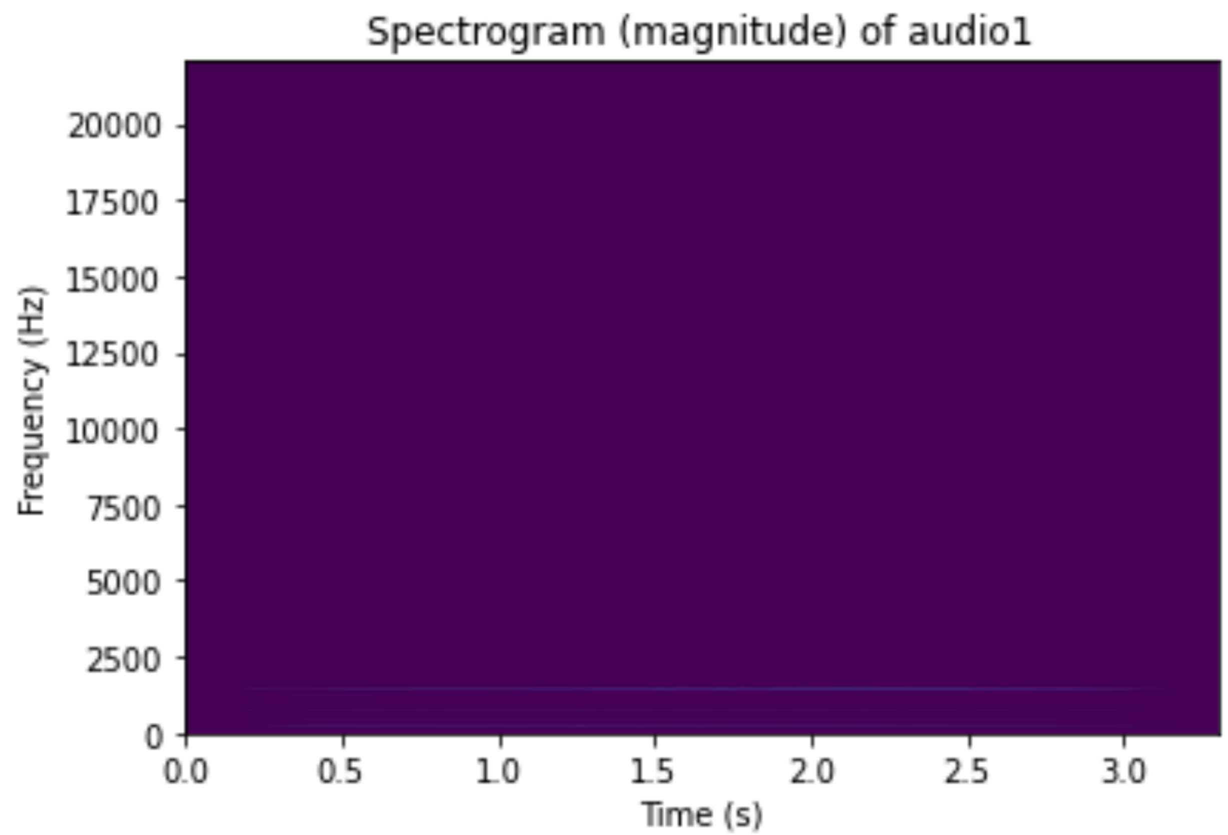

A segment of audio1
0.08
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06
audio1
0.93
0.94
0.95
0.96
0.97
0.98
0.99
Time (s)


Spectrogram (magnitude) of audio1
20000
17500
15000
12500
10000
7500
5000
2500
0
Frequency (Hz)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Time (s)

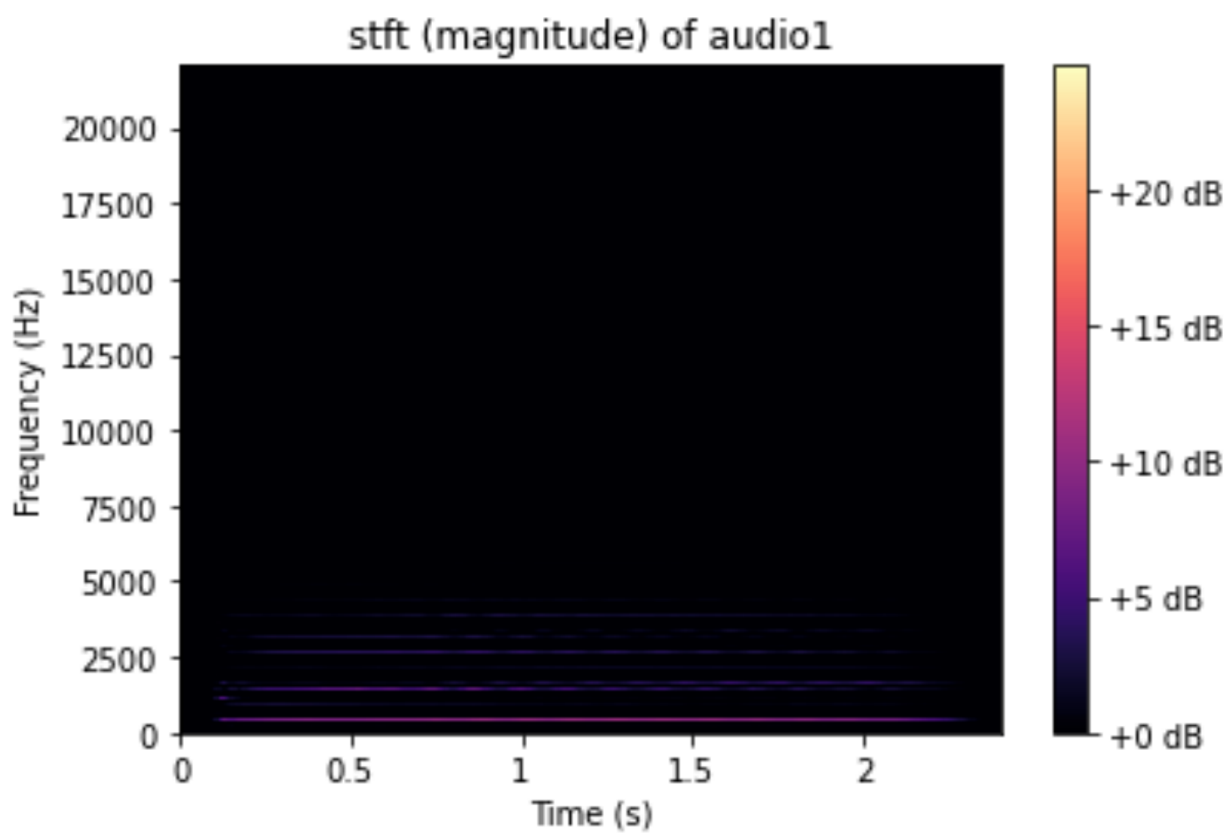

Spectrogram (dB) of audio1
Frequency (Hz)
20000
17500
15000
12500
10000
7500
5000
2500
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Time (s)

Audio1 - 128ms

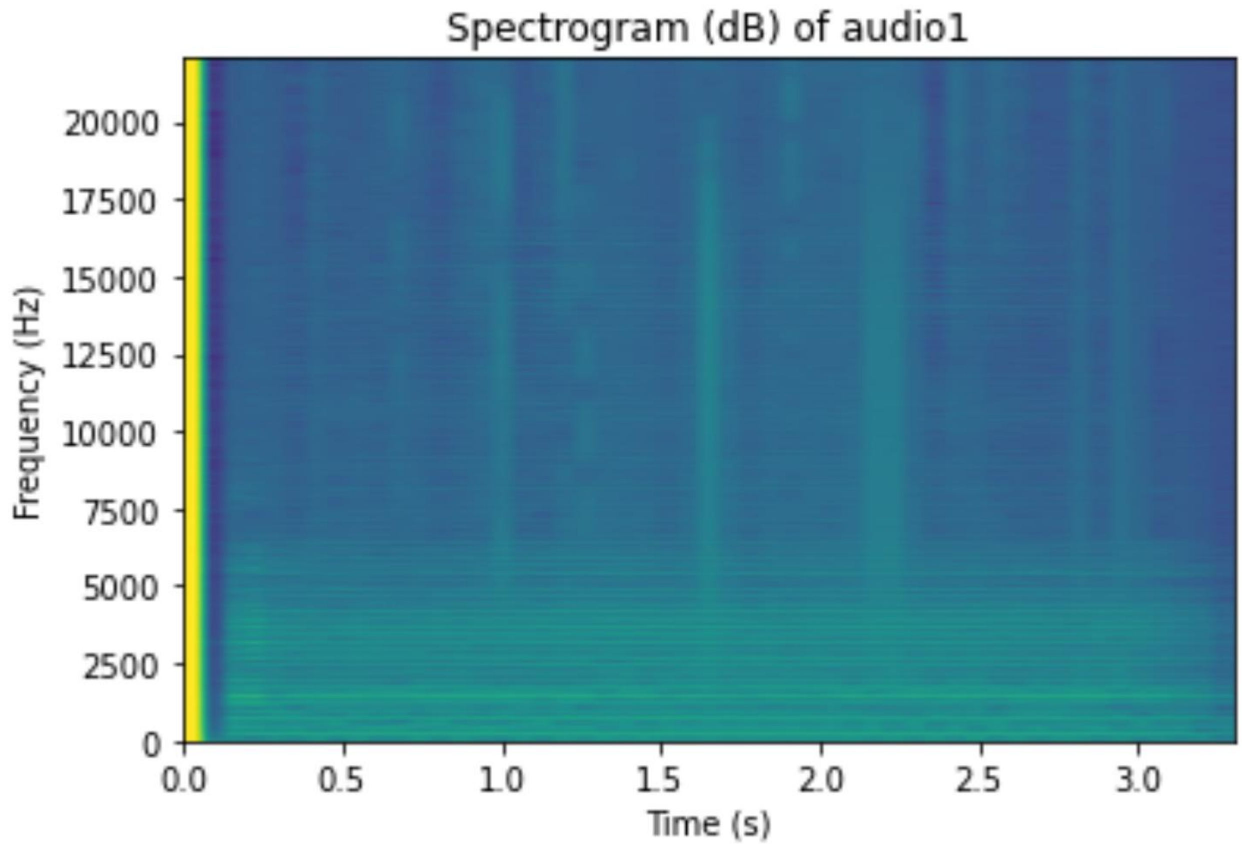


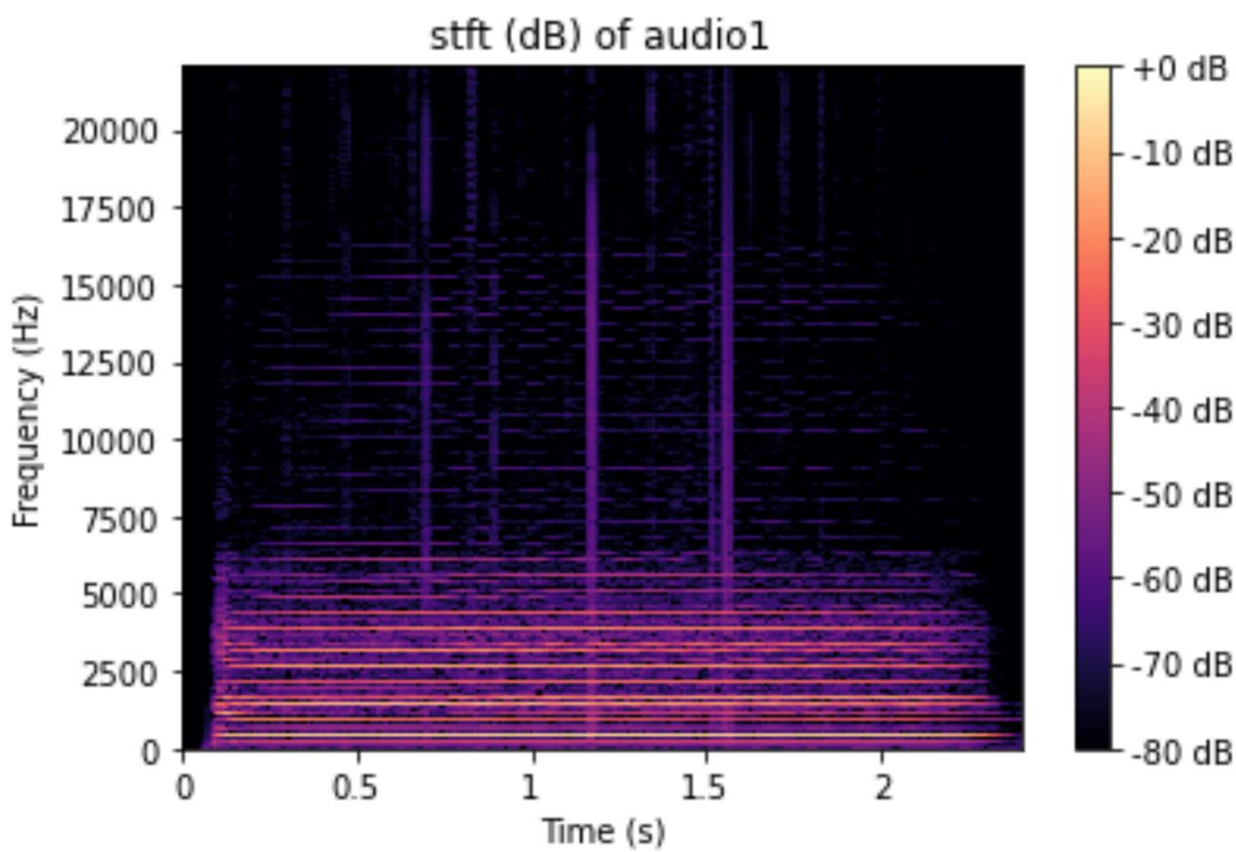


Spectrogram (magnitude) of audio1
Frequency (Hz)
20000
17500
15000
12500
10000
7500
5000
2500
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Time (s)

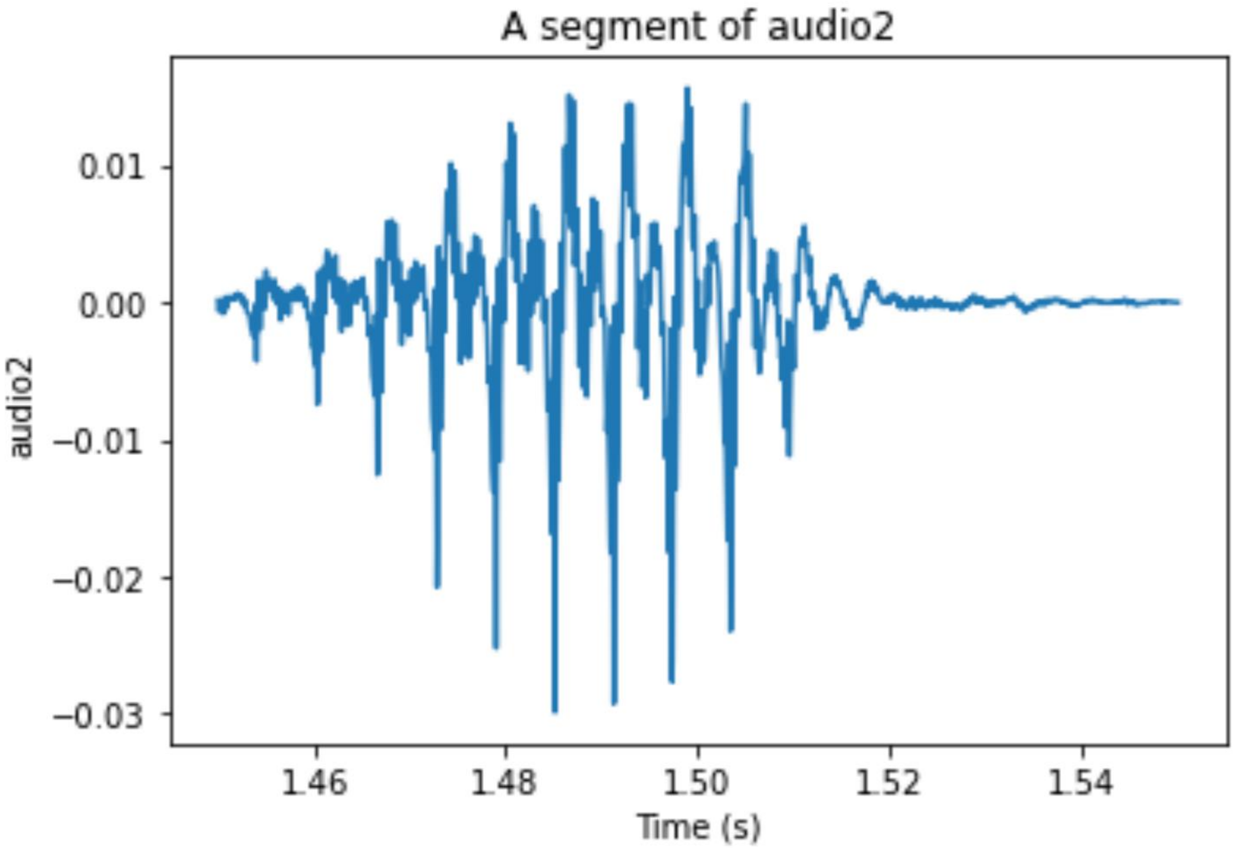


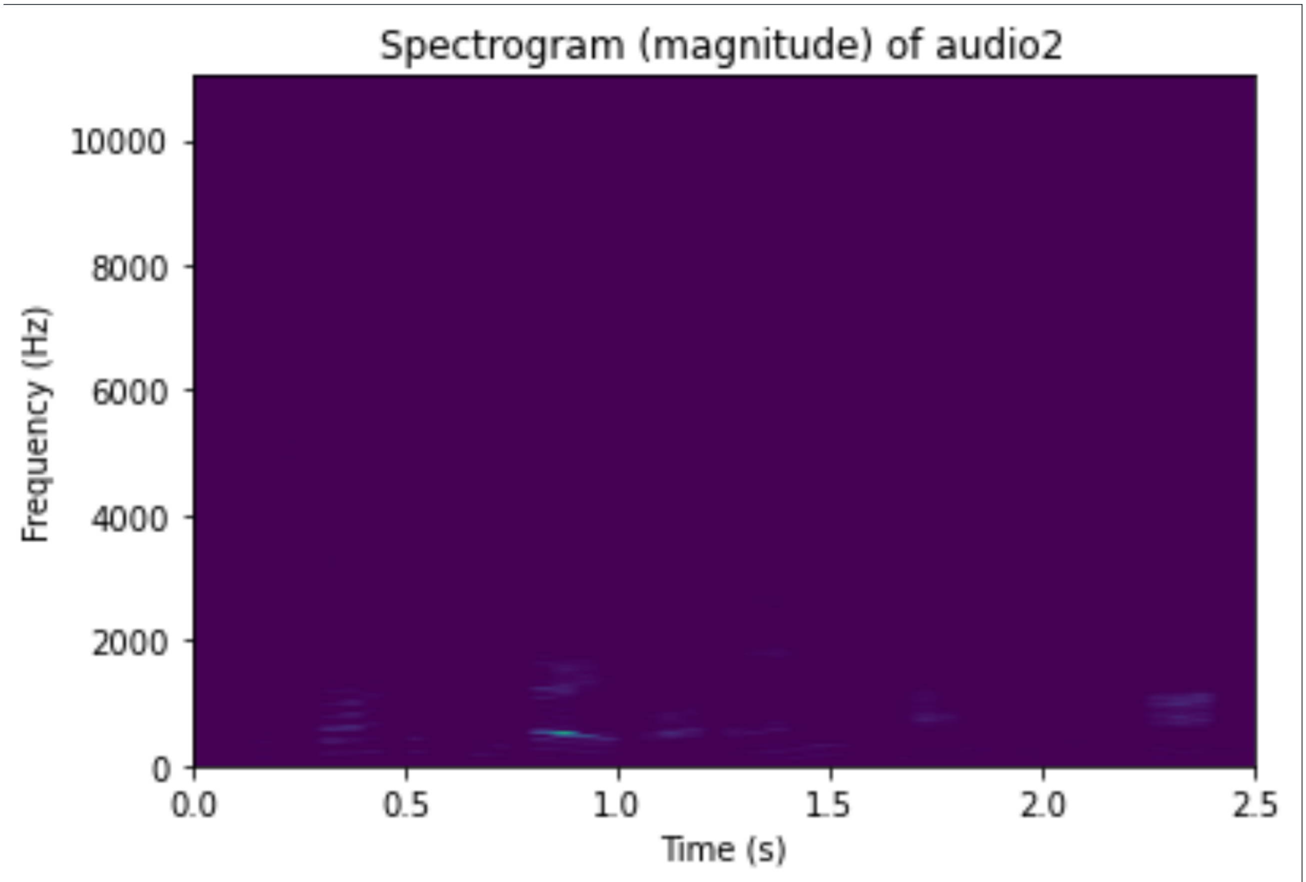
stft (magnitude) of audio1
Frequency (Hz)
20000
17500
15000
12500
10000
7500
5000
2500
0
0
0.5
1
1.5
2
Time (s)
+20 dB
+15 dB
+10 dB
+5 dB
+0 dB

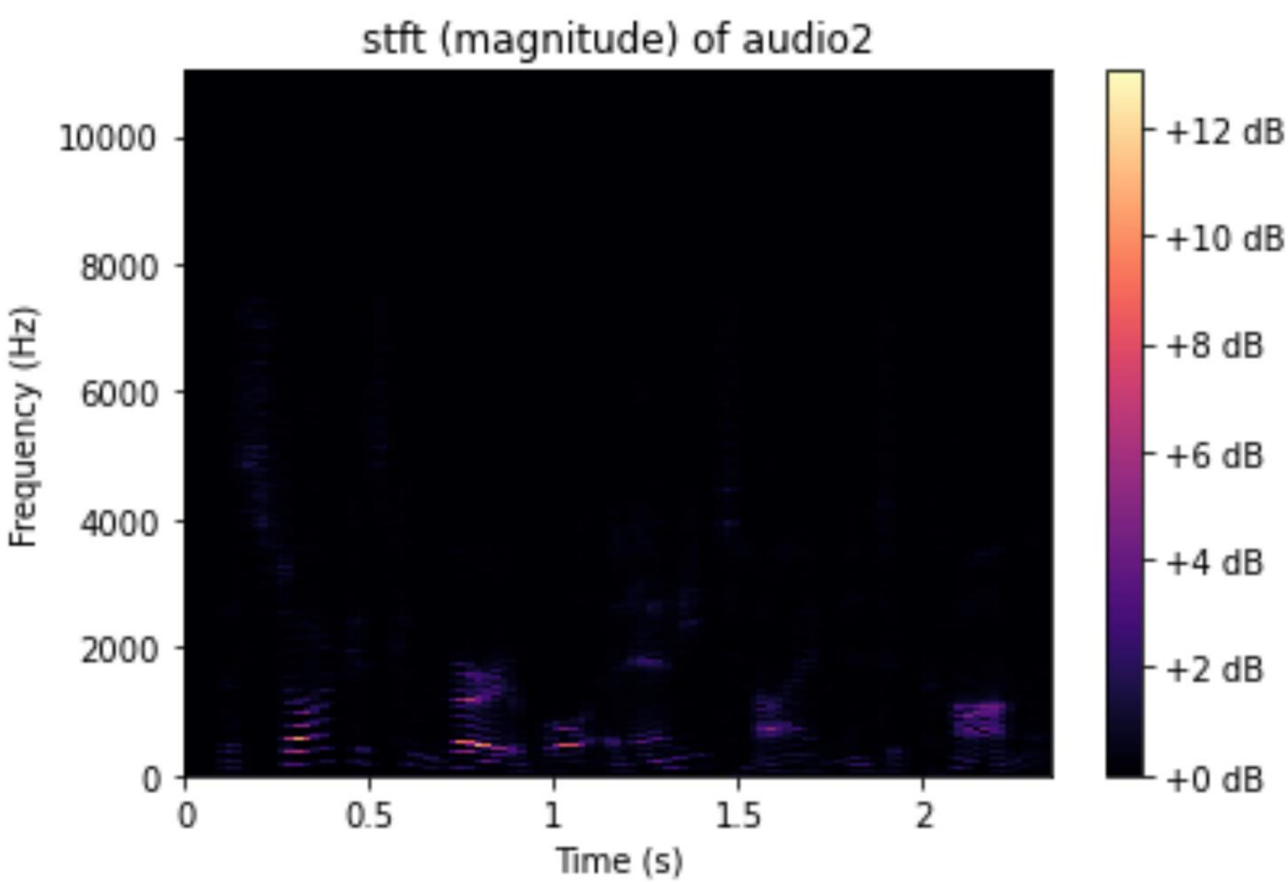

Spectrogram (dB) of audio1
Frequency (Hz)
20000
17500
15000
12500
10000
7500
5000
2500
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Time (s)

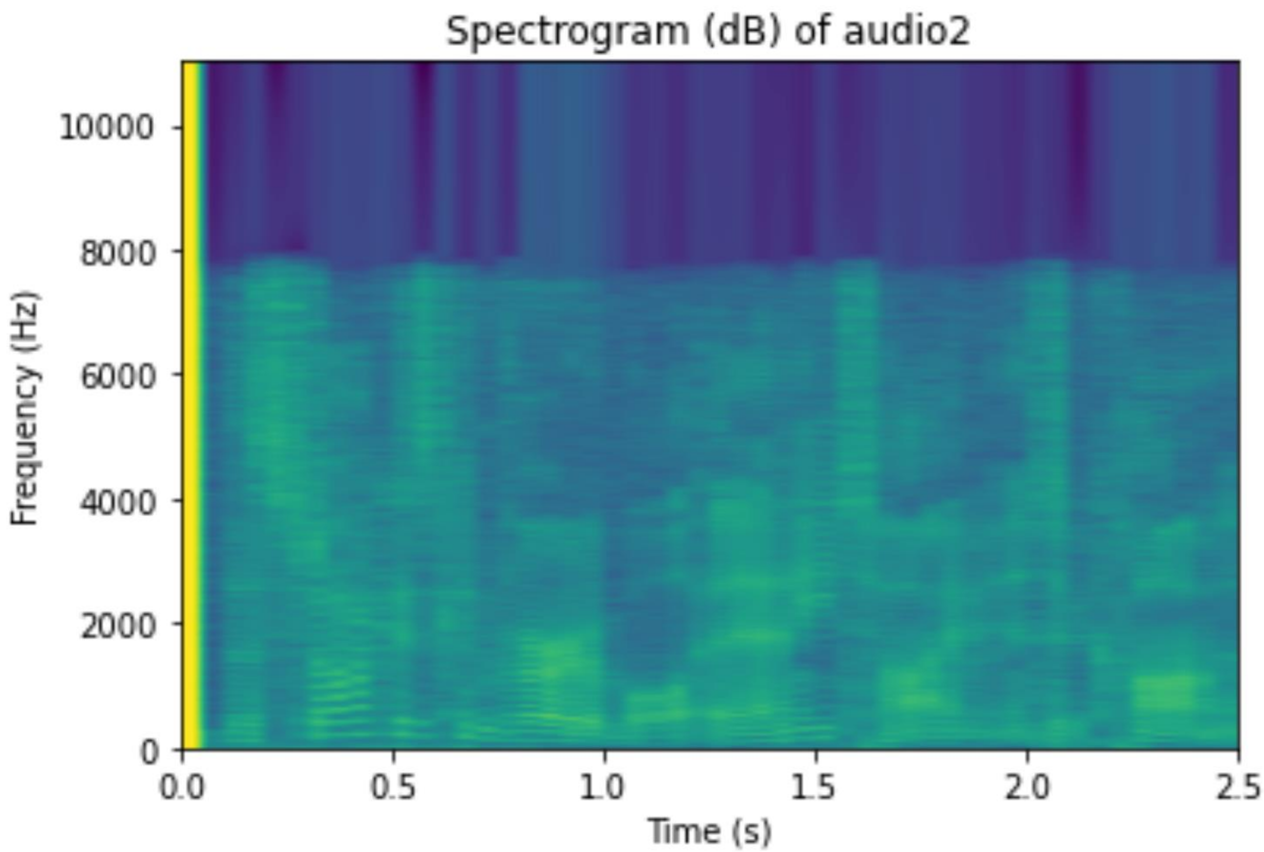

stft (dB) of audio1
Frequency (Hz)
20000
17500
15000
12500
10000
7500
5000
2500
0
0
0.5
1
1.5
2
Time (s)
+0 dB
-10 dB
-20 dB
-30 dB
-40 dB
-50 dB
-60 dB
-70 dB
-80 dB

Audio2 - 100ms

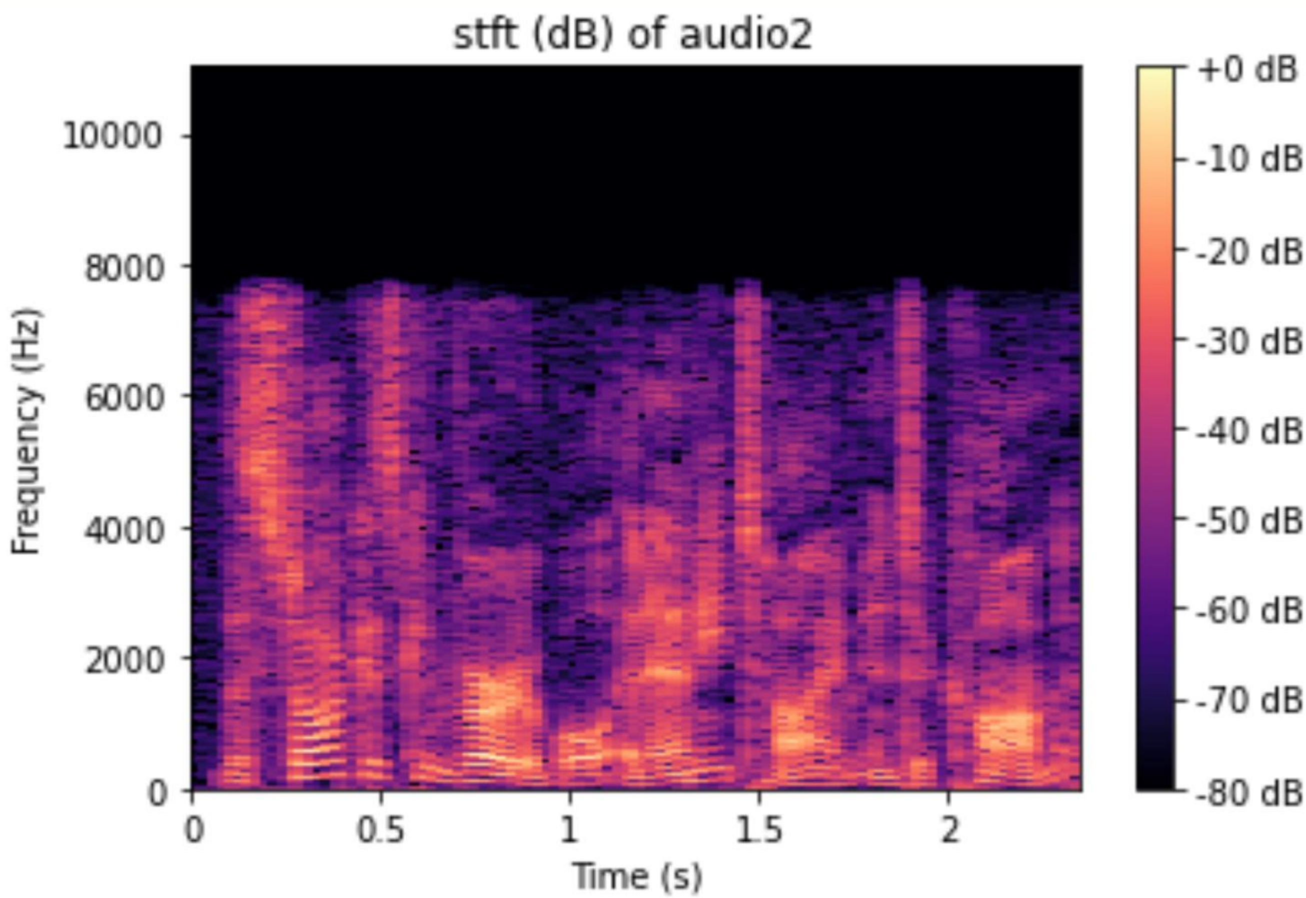



Spectrogram (magnitude) of audio2
Frequency (Hz)
Time (s)
0
2000
4000
6000
8000
10000
0.0
0.5
1.0
1.5
2.0
2.5


stft (magnitude) of audio2
Frequency (Hz)
Time (s)
0
2000
4000
6000
8000
10000
0
0.5
1
1.5
2
+12 dB
+10 dB
+8 dB
+6 dB
+4 dB
+2 dB
+0 dB

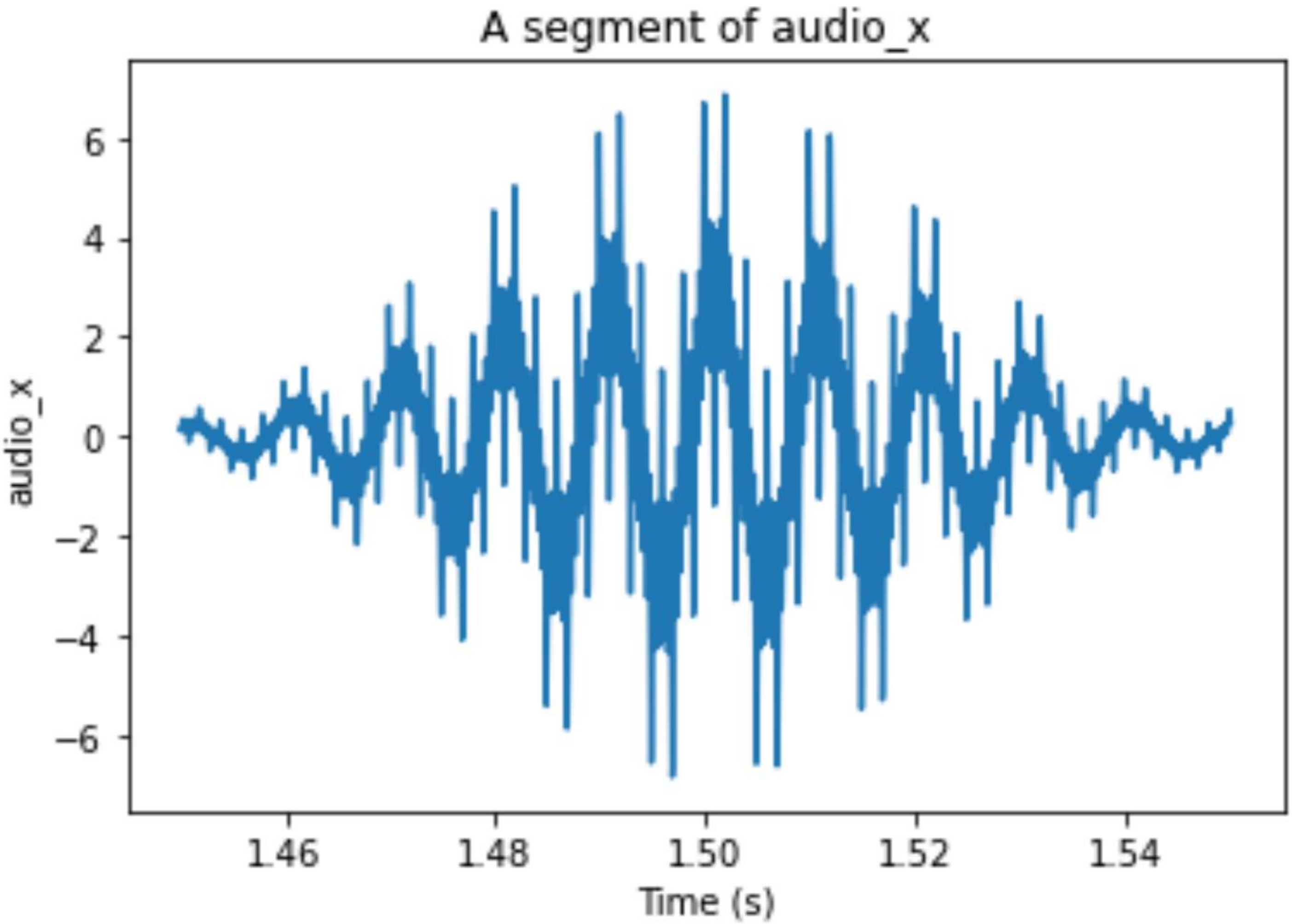


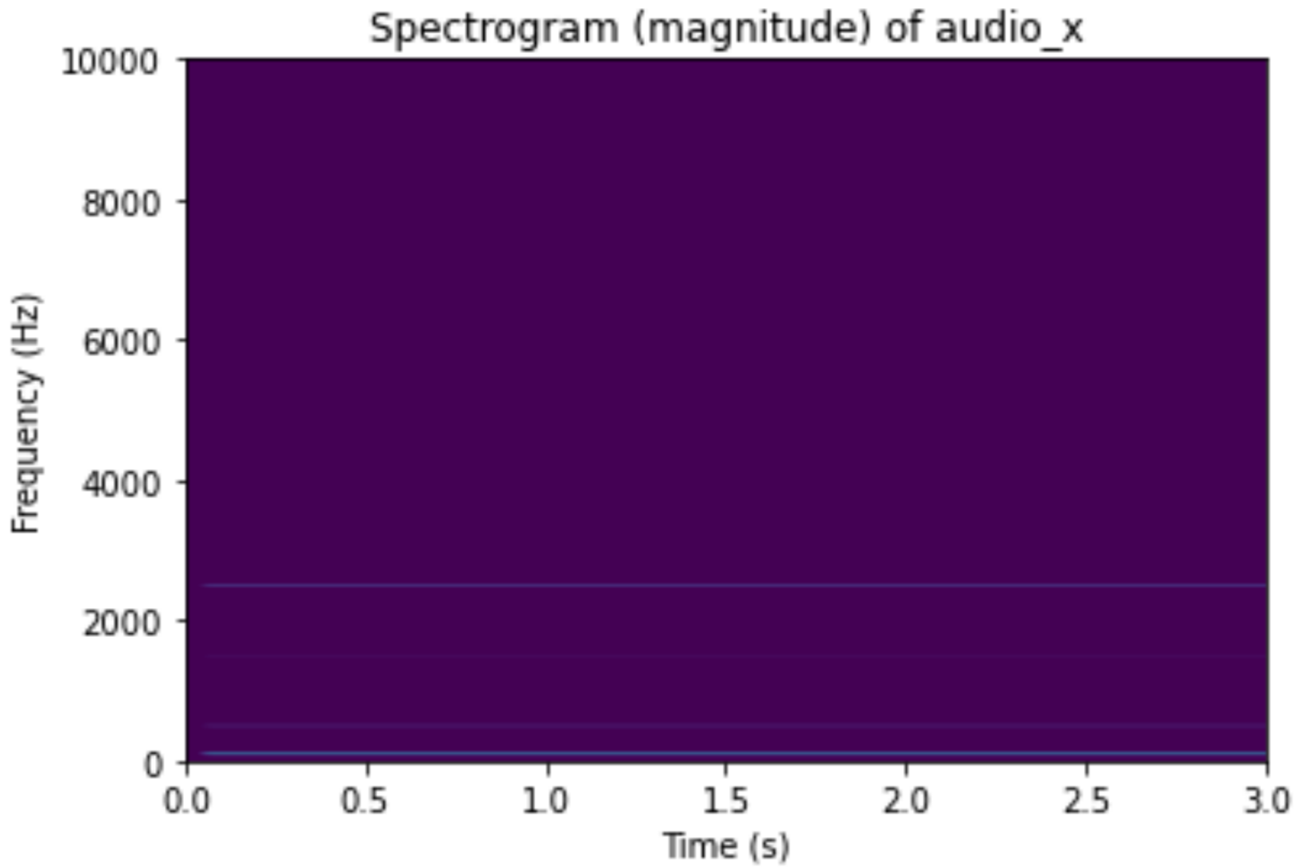
Spectrogram (dB) of audio2
Frequency (Hz)
Time (s)
0
2000
4000
6000
8000
10000
0.0
0.5
1.0
1.5
2.0
2.5



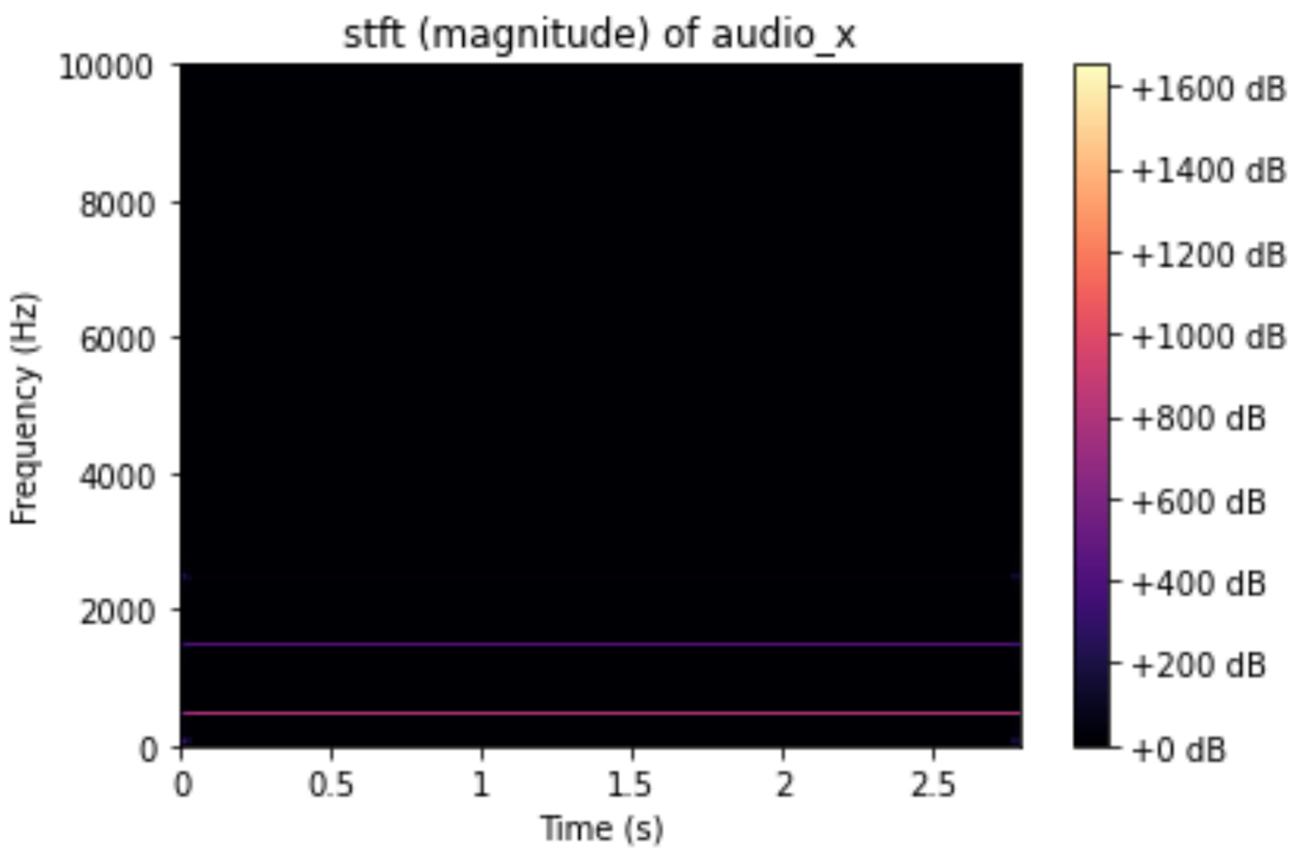
stft (dB) of audio2
Frequency (Hz)
Time (s)
0
2000
4000
6000
8000
10000
0
0.5
1
1.5
2
+0 dB
-10 dB
-20 dB
-30 dB
-40 dB
-50 dB
-60 dB
-70 dB
-80 dB

Sinusoids - 100ms

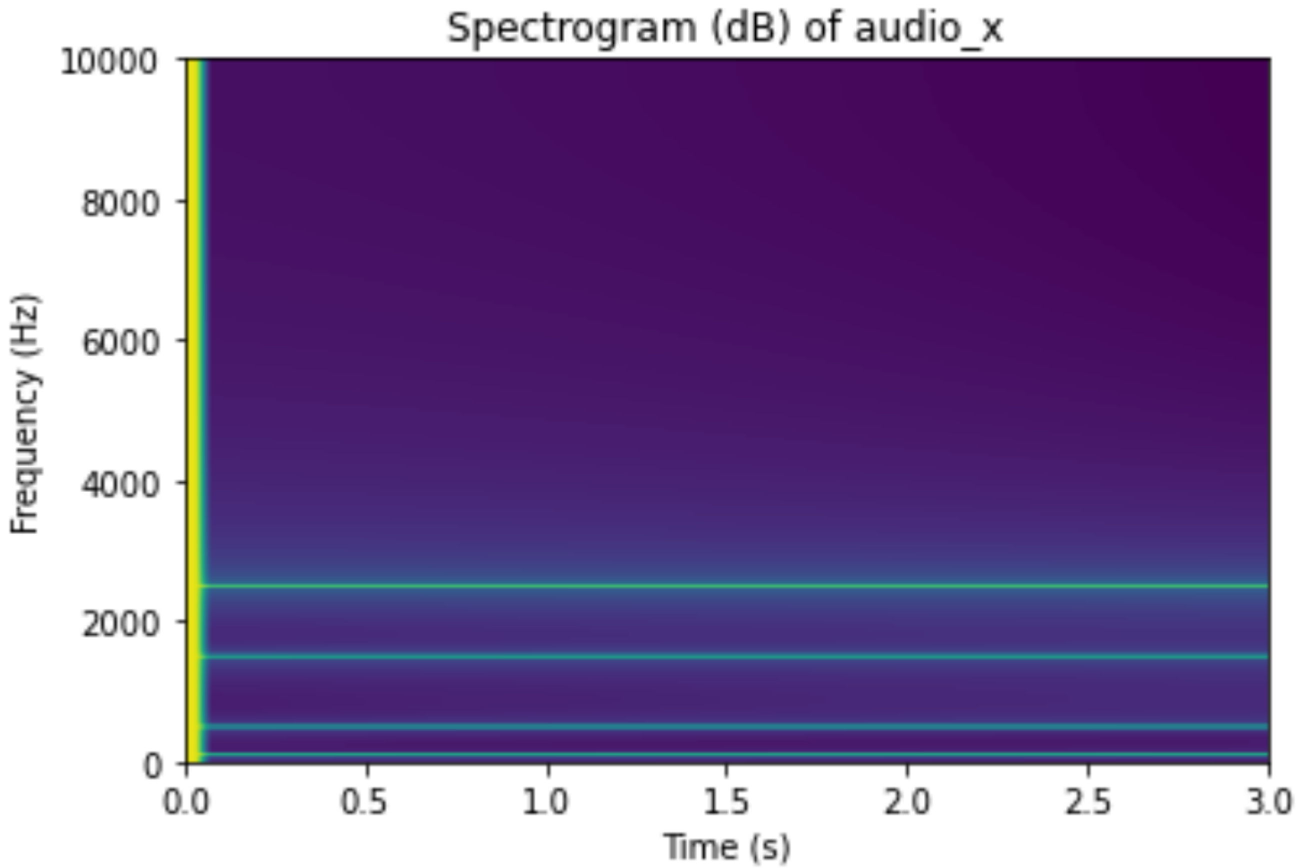




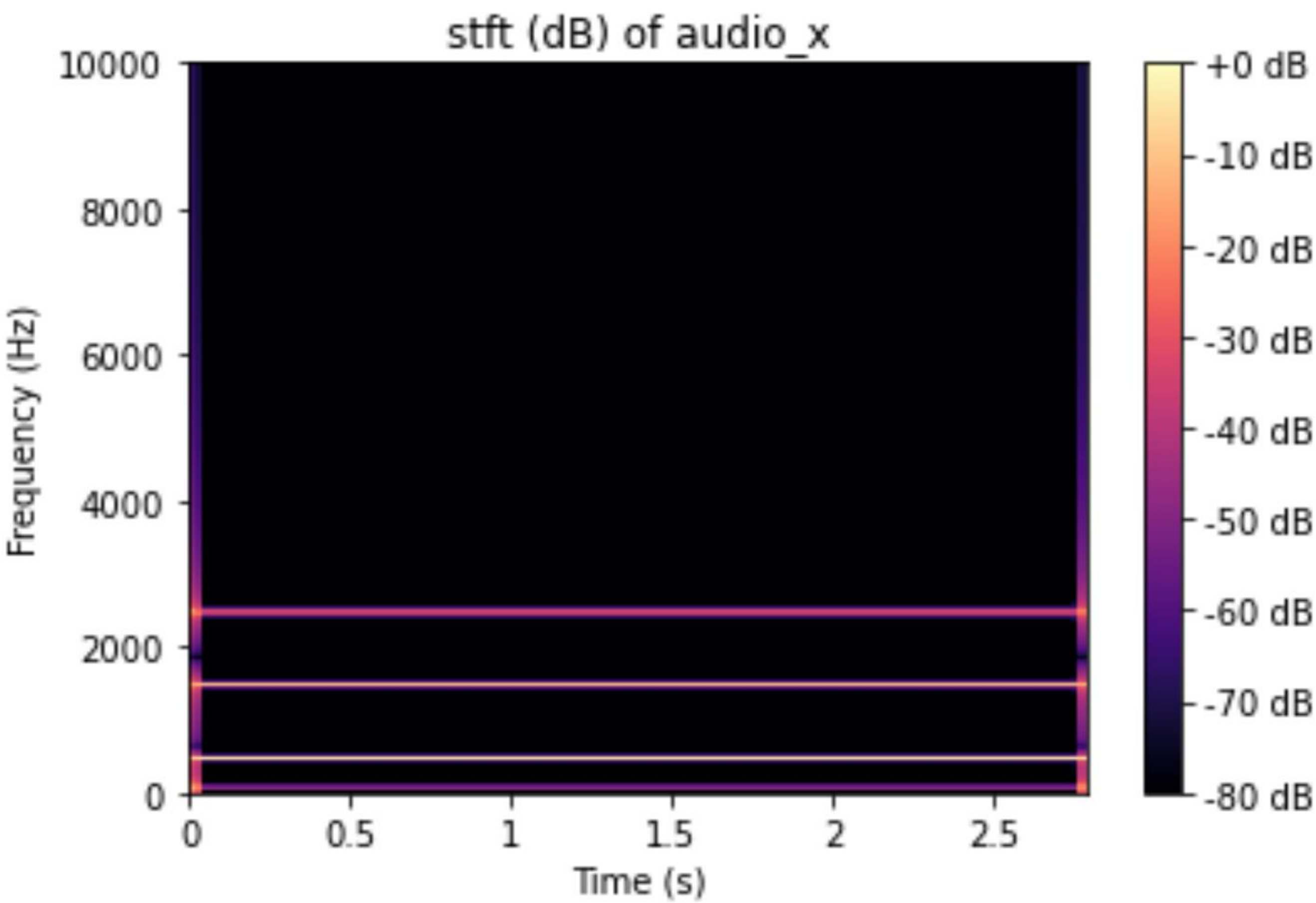
Spectrogram (magnitude) of audio_x
Frequency (Hz)
10000
8000
6000
4000
2000
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Time (s)



stft (magnitude) of audio_x
Frequency (Hz)
10000
8000
6000
4000
2000
0
0
0.5
1
1.5
2
2.5
Time (s)
+1600 dB
+1400 dB
+1200 dB
+1000 dB
+800 dB
+600 dB
+400 dB
+200 dB
+0 dB



Spectrogram (dB) of audio_x
Frequency (Hz)
Time (s)
0
2000
4000
6000
8000
10000
0.0
0.5
1.0
1.5
2.0
2.5
3.0



stft (dB) of audio_x
Frequency (Hz)
Time (s)
0
2000
4000
6000
8000
10000
0
0.5
1
1.5
2
2.5
+0 dB
-10 dB
-20 dB
-30 dB
-40 dB
-50 dB
-60 dB
-70 dB
-80 dB