The Quad Chips

WEEK - 3

NORMALIZATION
NOISE REDUCTION
PATTERN MATCHING

Lung sound recorder for portable healthcare devices

Team-38

The Quad Chips

Sai Praneeth - 2022101097 - Pattern matching

Venkata Jahnavi - 2022101118 - Classification

Sai Divya - 2022101090 - Noise reduction , loudness normalization

Vaishnavi Priya - 2022101108 - Normalization, Noise cancellation

Associated Professor: Abhishek Srivastava
Associated TA: Santhoshini Thota
Team No: 38, **Group:** 4

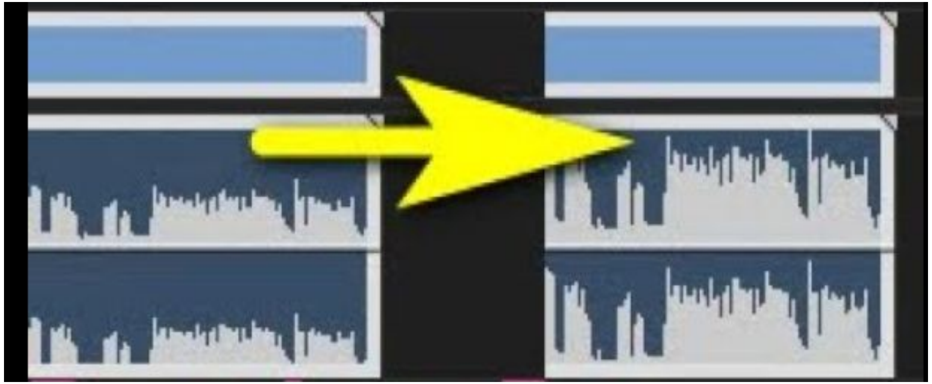
How are lung sounds effected?

Lung sound recordings can be affected by various factors, including both internal and external elements

Normalizing primarily addresses internal factors related to the recorded audio signal itself, specifically its amplitude or volume level

Noise Reduction deals with the external factors.

Normalization of Audio



On normalization

What is normalization?

It is the application of constant amount of gain to audio to maintain amplitude at a target level.

(or)

Process of adjusting audio, so that the audio signal stays the same throughout

Why do we need it?

It creates a consistent volume throughout, without having loud bits and quiet bits.

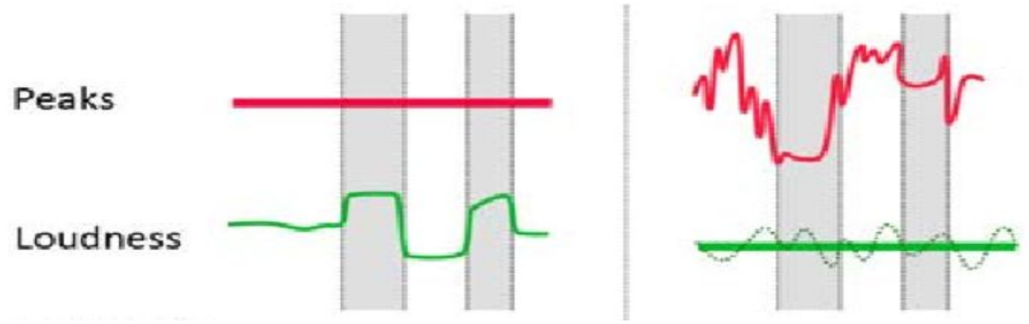
Lets go deep into it…

Two principle methods:

>Peak Normalization: Changes it such that loudest sample falls at a specific level(usually peaks at 0 dbms)

>Loudness Normalization: Changes it such that the average loudness falls at target level.

For this project, we are going to use both of them, Peak to prevent clipping & Loudness to ensure consistency.



Peak Normalization   Loudness Normalization

Pros & Cons

Pros:

1.Volume consistency

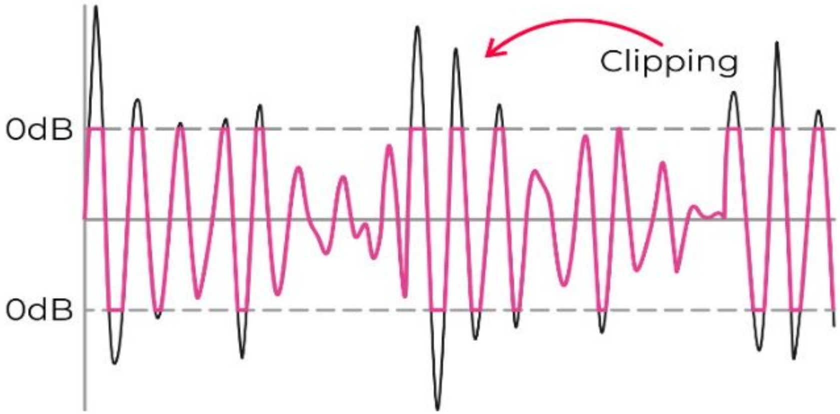
2.Avoids Clipping/distortion

Cons:

1.We can’t undo the process.(destroys the original one, unless u make a copy)

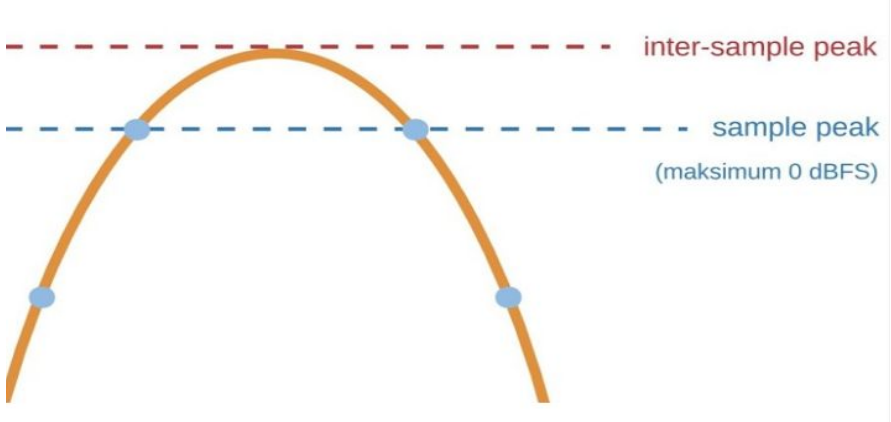
2.Inter-sample peaks

3.Alters the original intent of recording.



Clipping:Its a waveform distortion that occurs when amplifier attempts to deliver an output beyond its maximum capability,so we chop it off.

Inter-Sample Peak:This occurs because the digital device doesn’t actually “see” information between the samples; it just creates a smooth visualization for reference. However, when the digital signal is converted back to analog (D/A), the curve that creates peaks between samples is accurately reconstructed, potentially creating distortion that wasn’t present before the D/A conversion.



Noise Reduction

What is noise reduction?

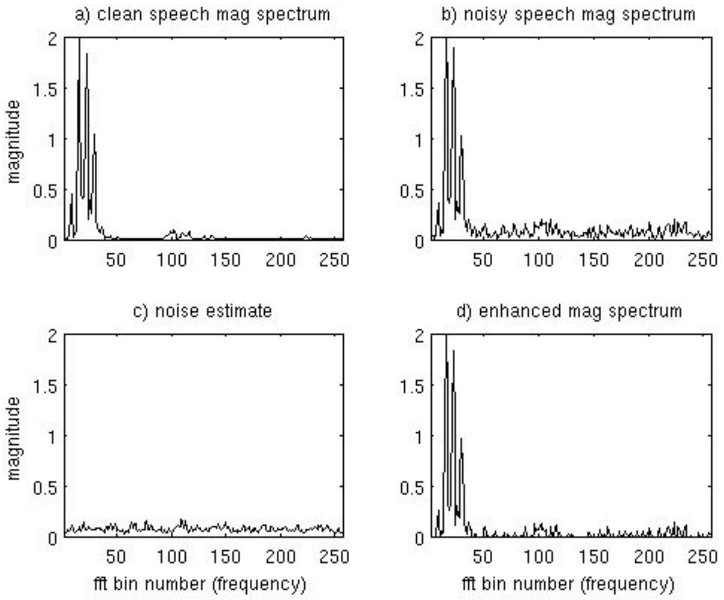
Process of removing noise

This technique focuses on reducing or eliminating unwanted background noise from a recording.

Why do we need it?

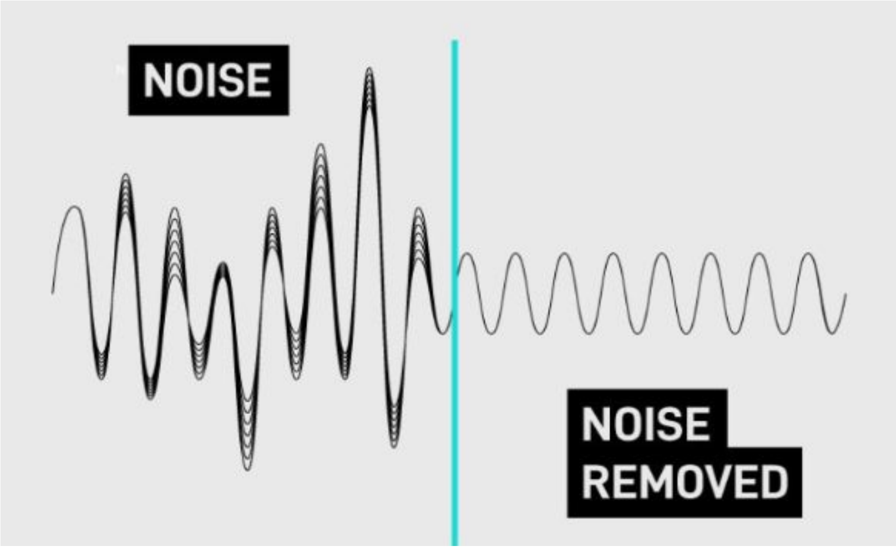
To improve the signal on an image by reducing the inherent imperfections of the imaging system

Spectral subtraction can be effective for reducing stationary background noise. In the case of lung sounds, prior knowledge of the frequency characteristics of normal and abnormal lung sounds can be used to refine the estimation of the noise spectrum.



Reference:
http://practicalcryptography.com/miscellaneous/machine-learning/tutorial-spectral-subraction/

Lets go deep into it…



Its types:

Active Noise Reduction:

They harness the power source of microphones and tiny speakers to create a constant anti-noise signal that is equal and opposite to that of unwanted ambient sounds. The result is increased clarity and volume control

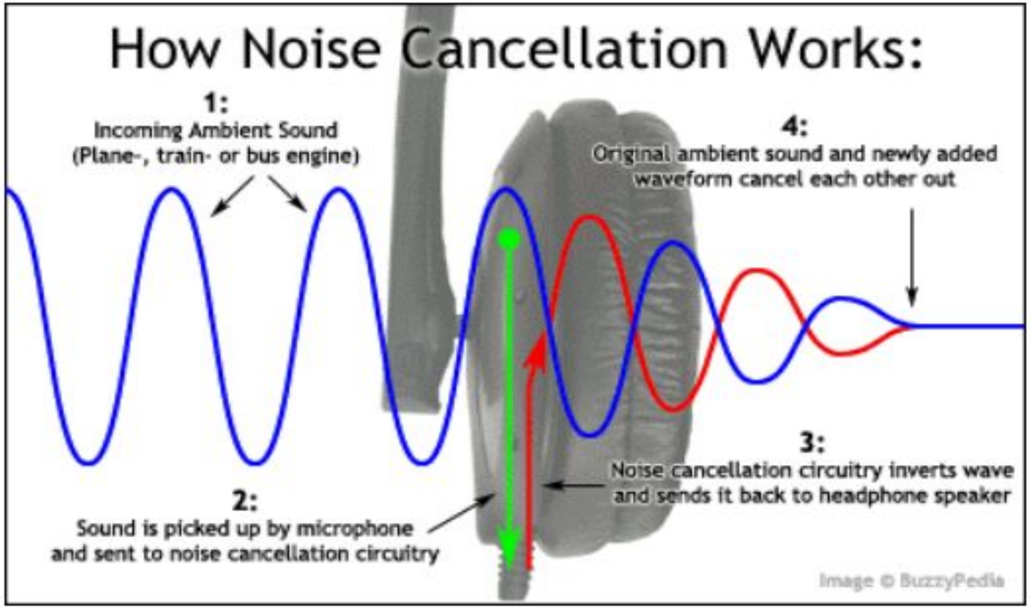
Passive Noise Reduction:

it reduces the loudness without blocking out conversations or music. passive noise cancellation occurs before sounds reach your ear canal, it reduces the actual loudness instead of just blocking specific sounds.

While we are in this, we'll also look into Noise Cancelling.

Noise Cancellation:

blocks out a noisy environment.
They do this by using a microphone to detect outside noise, and then they create an opposite noise to counteract it.



Noise Reduction vs Noise Cancellation

applied after audio recording or during post-processing uses digital processing to analyze and modify existing audio signals,	operates in real-time during audio playback. actively generates anti-noise to counteract ambient sounds

For lung sound recorder, we will use Noise Reduction.

https://www.plant-tours.com/blog/difference-between-noise-cancellation-and-noise-reduction-headphones/

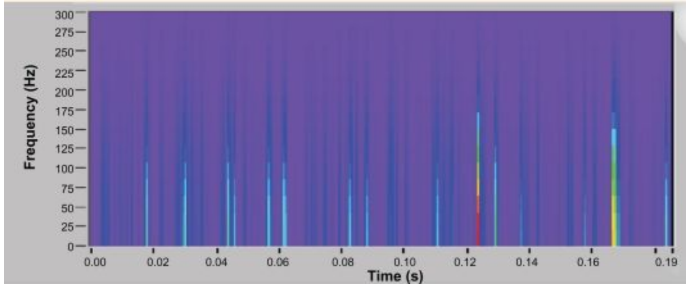
Pattern Matching & Classification

As we reach this stage, we assume that we have clear audio extracted, containing the lung sounds after pre-processing through the Noise Reduction, FFT, etc.

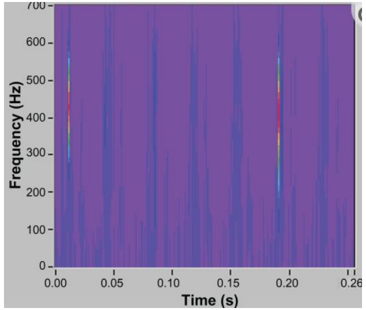
The sounds can be classified into five classes:

(Wheeze, Crackle, Stridor, Rhonchi, Normal)

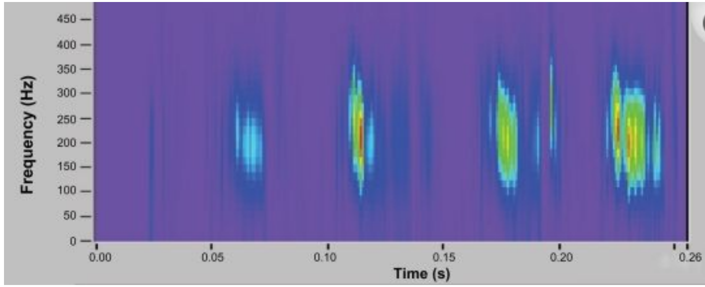
Spectrograms for different lung sounds



Normal



Spectrogram of a respiratory cycle containing fine crackles. The vertical lines in the figure indicate the presence of fine crackles.



Spectrogram of respiratory cycle with wheezing.

source:https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3157943/

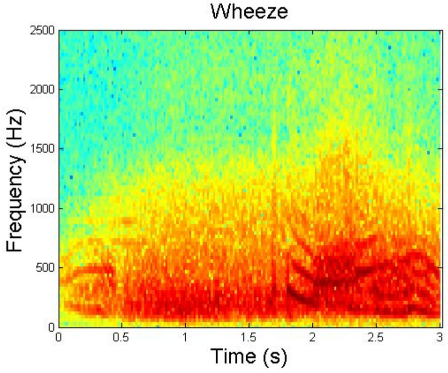
MEL Spectrograms

In order to process the audio and its patterns, first convert the audio into spectrograms.

Why MEL?

We implement MEL spectrogram in this case, because the Mel scale is designed to better align with human auditory perception, and many other factors making it well-suited for tasks like audio classification.

When performing processing to obtain MEL, we use sliding window method.



X-axis: time (frame index)

Y-axis: Mel frequency bins.

The color intensity or brightness at each point in the spectrogram corresponds to the magnitude or power of the filterbank energy at that time and frequency

Patterns or features are extracted from the spectrogram to capture relevant information for the specific audio processing task. One of such approach is the MFCCs
(Mel Frequency cepstral coefficients)

Reference Spectrograms

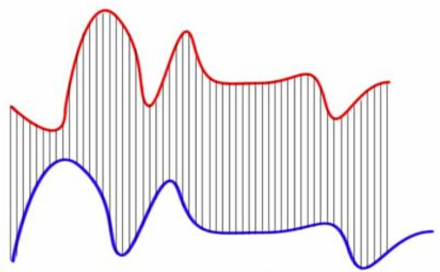
For pattern matching, need to prepare reference spectrograms from the specific audios of different lung sounds.

We then, compute the similarity between the input audios’ spectrogram and the reference lung sounds’

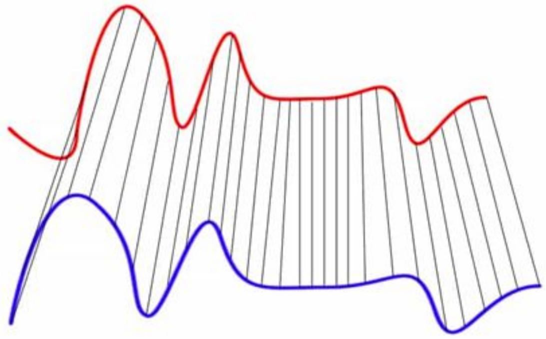
How do you compute Similarity?

Similarity metrics / distance measures (between data points)

We will use methods like Dynamic Time Warping. DTW aligns the frames or segments of two audio signals, and finds the best temporal alignment that minimizes the distance metric between the data points.



Euclidean Matching



Dynamic Time Warping Matching

Matching the patterns / features

We obtain the MFCCs (Mel-frequency cepstral coefficients) from the spectrogram. Using these data points, we run the similarity matching with the reference patterns. This is the point where we need to calibrate the threshold for the accurate, precise result.

We match the similarity with all the 4 reference audios and output the most similar possibility. (algorithms like DTW, cross-relation, etc.)

*RESOURCE: https://chat.openai.com/share/ccd81875-f38d-4619-ae7c-3aadd4d34d48

Next Week's Plan

- Testing the components.
- Building the hardware component and testing it.