Detecting COVID-19 with Audio Analysis and Machine Learning

BY YENYU CHEN AND DARREN ELDREDGE

Introduction and Data Selection

Medical conditions can alter an individual's voice by affecting organs critical to speech, such as the lungs, brain, muscles, or vocal folds. Vocal signals can be analyzed, features extracted, and Machine Learning methods applied to identify vocal biomarkers that can be used for diagnosis, classification, and remote monitoring of patients with a variety of clinical conditions such as Parkinson's Disease, Alzheimer's Disease, and Multiple Sclerosis. This project focuses on applying audio analysis and machine learning methods to COVID-19 detection for two reasons:

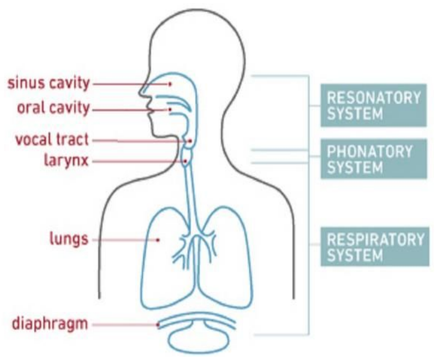
(1) impact on global health and security

(2) success of other researchers in this area.

All analysis was conducted using samples from the COUGHVID open dataset. To learn more, please scan the QR code to view our GitHub repo, README, and Intro to Literature.

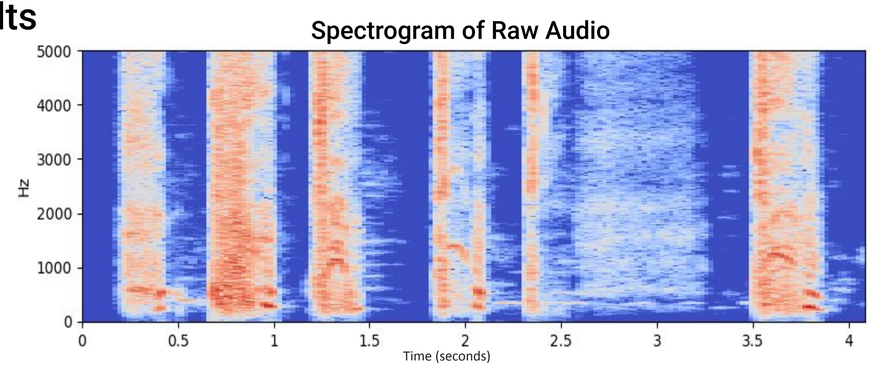
Methodology

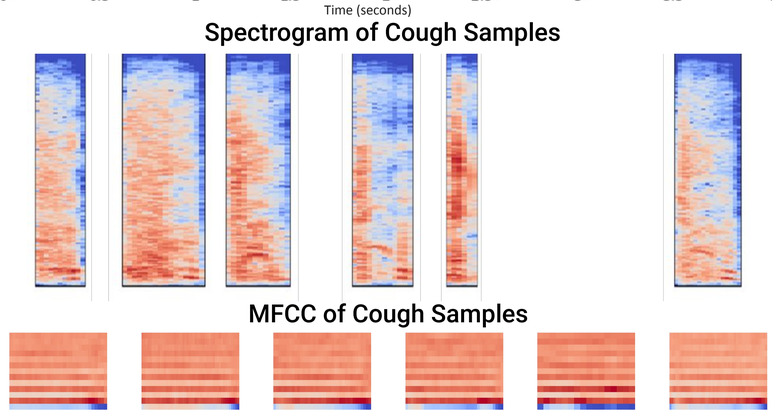
- Use **cough audio samples** to diagnose COVID-19
- **Extract cough** from audio sample (xgb classifier, trim, split)
- **Short time Fourier Transform** on cough
- **Spectrogram** captures how sounds are made from respiratory system, phonatory system, and resonatory system. But the spectrograms contain too much detail.
- **MFCC feature extraction** accurately represents the spectrogram envelope and mimics how the human ear processes sound, making our system work like a human ear (hearing more low frequency than high frequency)
- MFCC also makes data size smaller.
- **Machine Learning** helps classify Healthy and COVID

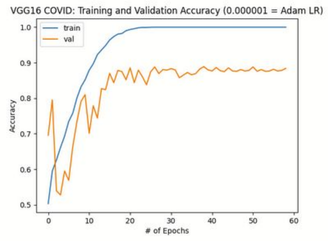
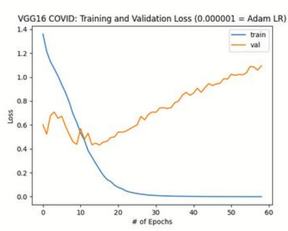
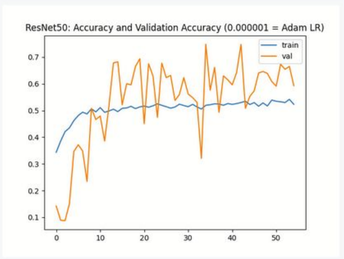
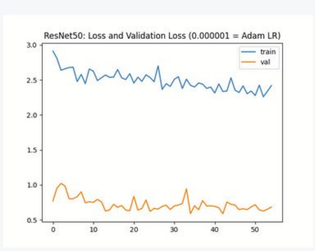
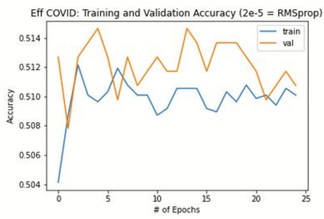
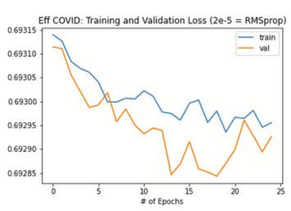


https://www.templehealth.org/about/blog/how-does-my-voice-work

Results





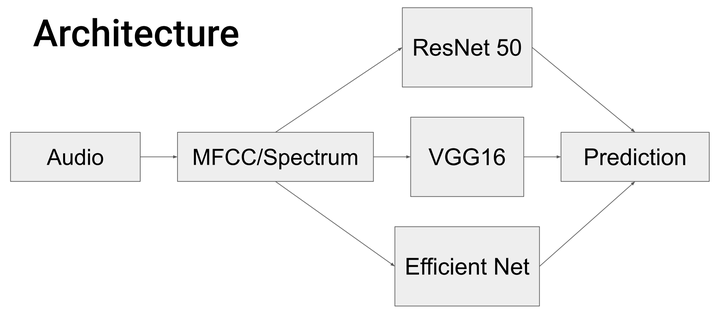
ML Model	Accuracy	Loss
VGG16		
ResNet50		
EfficientNetB4		

Confusion Matrix			
True label	COVID	3	505
	Healthy	0	532
		COVID	Healthy
		Predicted label	

Confusion Matrix			
True label	COVID	508	0
	Healthy	532	0
		COVID	Healthy
		Predicted label	

Confusion Matrix			
True label	COVID	0	508
	Healthy	0	532
		COVID	Healthy
		Predicted label	

Architecture



Discussion

- Distortion of MFCC image when input into models
- Data Bias from healthy
- Fine tune the model
- Combined more data from different database

Suggestions for Future Work

- Combine speech analysis with cough analysis
- Try different feature extraction eg. muscle fatigue
- Optimize models for portable applications

Conclusions

- Analysis of audio samples using VGG16, ResNet50, and efficientnetB4 is difficult because the feature image is resized and information can be lost.
- Data bias can be strong when using the entire dataset to train the model because the available data isn't representative of the population.
- High levels of sensitivity and specificity can be arbitrarily obtained by skewing the data distribution (see accuracy and loss chart of VGG16).
- Success will require larger, cleaner open datasets.

GITHUB