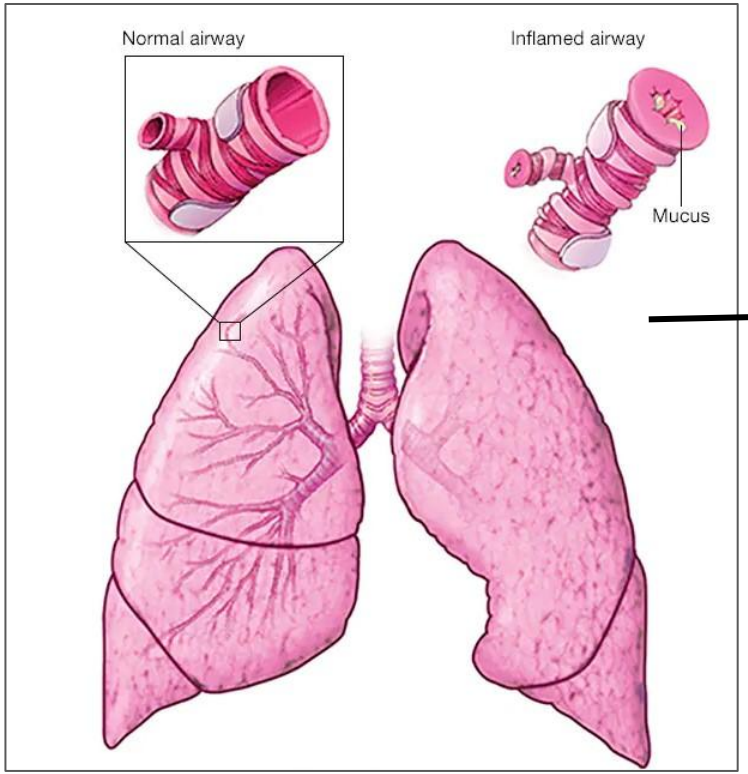
Natural Speech to Predict Asthma using Vowel Inspiration in Machine Learning

By

Anik Bhaumik

ASTHMA



Asthma is a condition in which your airways get narrow and swell and may produce extra mucus. This can make breathing difficult and trigger coughing, a whistling sound (wheezing) when you breathe out and shortness of breath.

For some people, asthma is a minor nuisance. For others, it can be a major problem that interferes with daily activities and may lead to a life-threatening asthma attack.

WHAT CAUSES ASTHMA?

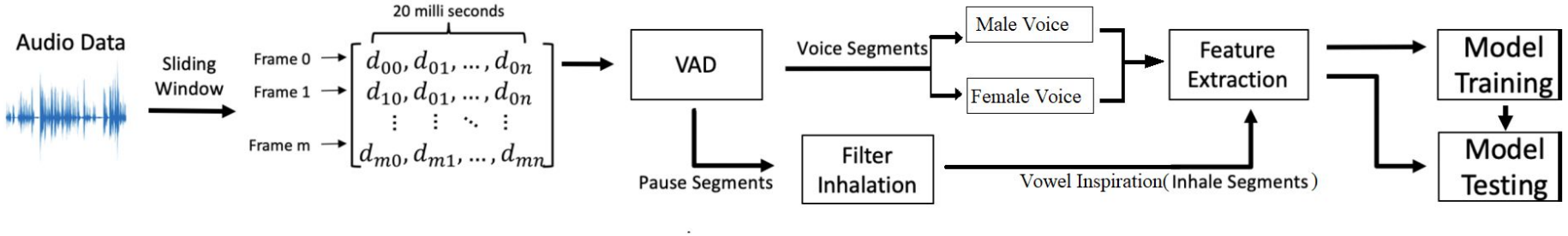
Researchers don’t know why some people have asthma while others don’t. But certain factors present a higher risk:

- **Allergies**: Having allergies can raise your risk of developing asthma.
- **Environmental factors**: People can develop asthma after exposure to things that irritate the airways. These substances include allergens, toxins, fumes and second- or third-hand smoke. These can be especially harmful to infants and young children whose immune systems haven’t finished developing.
- **Genetics**: If your family has a history of asthma or allergic diseases, you have a higher risk of developing the disease.
- **Respiratory infections**: Certain respiratory infections, such as respiratory syncytial virus (RSV), can damage young children’s developing lungs.

HOW IS ASTHMA DIAGNOSED?

- Your healthcare provider will review your medical history, including information about your parents and siblings. Your provider will also ask you about your symptoms. Your provider will need to know any history of allergies, eczema (a bumpy rash caused by allergies) and other lung diseases.
- Your provider may order spirometry. This test measures airflow through your lungs and is used to diagnose and monitor your progress with treatment. Your healthcare provider may order a chest X-ray, blood test or skin test.

PROPOSED METHOD



VOICE ACTIVITY DETECTION

- Speech data is a mixture of voiced segments and pause segments. This is because people take pauses to inhale between each voiced segment.
- Prior to extracting features from the audio file, we identified voiced segments and pause segments from the speech audio data using the LTSD-based algorithm.
- This algorithm looks at the long-term divergence between the noise spectra and potential voice spectra.
- First we'll work with voice segment and then we'll distinguish between male and female voice as males generally have low pitched voices while females have high pitched voices.
- Second we'll work on pause segment which is used later for inhale segment.

INSPIRATION FILTERING

- The act of vocalization process involves continuous exhalation. To compensate for the loss of air during the voice activity, there are often sharp inspiration periods right before or after long vocalization periods. In order to extract the inspiration segments, we further processed the pause segments.
- Inspiration sound is generally independent of the vocal fold movements, but rather more influenced by the obstruction in the airway and the rapidness of inspiration
- In the first step, the potential inspiration frames particularly the vowel segments were identified by computing the energy in the given pause segment and selecting the frames whose energy falls in between the 60th and 95th percentiles.
- When pronouncing a vowel we need to exhale and in order to fill up gap you need to inhale more air then the other alphabets. So, it’ll tend to more inspiration
- In the second step, the selected frames are clustered using Density-Based Spatial Clustering of Applications with Noise (DBSCAN). The extracted inspiration frames were input to the feature extraction block along with the voiced frames.

FEATURE EXTRACTION

Two types of features can be extracted from the voice segment: acoustic features and pulmonary features. Here, we are going to work with only pulmonary features.

- **Pause Frequency**:we captured the information related to the shortness of breath by computing the average number of pauses per minute in each user's speech.
- **Vocalization to Inhalation Ratio**: Since pulmonary disease can affect the ability to speak for long periods and the amount of air inhaled at the end of the speaking period, we considered the inhalation duration during the pause for breath relative to the immediately preceding vocalization. This feature is hereby referred to vocalization to inhalation ratio.
- **Average Phonation Time**: we computed the average phonation time from the voice segments and utilized it as one of the features
- **Inspiration Sound Energy**: we looked at the frequency band ranging from 7,800 Hz to 8,000 Hz. In this band, the maximum frequency component from the healthy inspiratory sound and pathological inspiratory sound showed meaningful differences, with a larger maximum frequency component in pathological inspiratory sounds.

MODEL BUILDING AND EVALUATION

We developed two kinds of models from the collected data sets: classification model and regression model. The classification model was designed to detect participants with pathological conditions, and the regression model was built to predict the participant’s FEV1-to-FVC ratio, which is one of the standard measures for pulmonary function.

REFERENCES

- Towards Passive Assessment of Pulmonary Function from Natural Speech Recorded Using a Mobile Phone by Keum San Chun, Viswam Nathan, Korosh Vatanparvar, Ebrahim Nemati, Md Mahbubur Rahman, Erin Blackstock, Jilong Kuang

  https://ieeexplore.ieee.org/abstract/document/9127380/

- Speech Analysis and Synthesis by Linear Prediction of the Speech Wave by B. S. Atal and Suzanne L. Hanauer

  https://asa.scitation.org/doi/abs/10.1121/1.1912679

- https://www.mayoclinic.org/diseases-conditions/asthma/symptoms-causes/syc-20369653#:~:text=Asthma%20is%20a%20condition%20in,asthma%20is%20a%20minor%20nuisance.
- https://my.clevelandclinic.org/health/diseases/6424-asthma
- https://www.urmc.rochester.edu/encyclopedia/content.aspx?contenttypeid=197&contentid=28587