Early Detection of Pulmonary Fibrosis Using Digital Stethoscope Audio and Signal Processing

Devin Von Arx, Emer Doheny, PhD, and Sinead O'Connor, MSc
University College Dublin, Rice University and Trinity College Dublin

What is Pulmonary Fibrosis?

- Pulmonary fibrosis is a progressive disease that causes scarring (fibrosis) in the lung tissue, making it harder to breathe.
- Early diagnosis is difficult because symptoms overlap with more common conditions, and crackles are often too subtle to hear.
- Average life expectancy after diagnosis is 3 to 5 years.

Goal

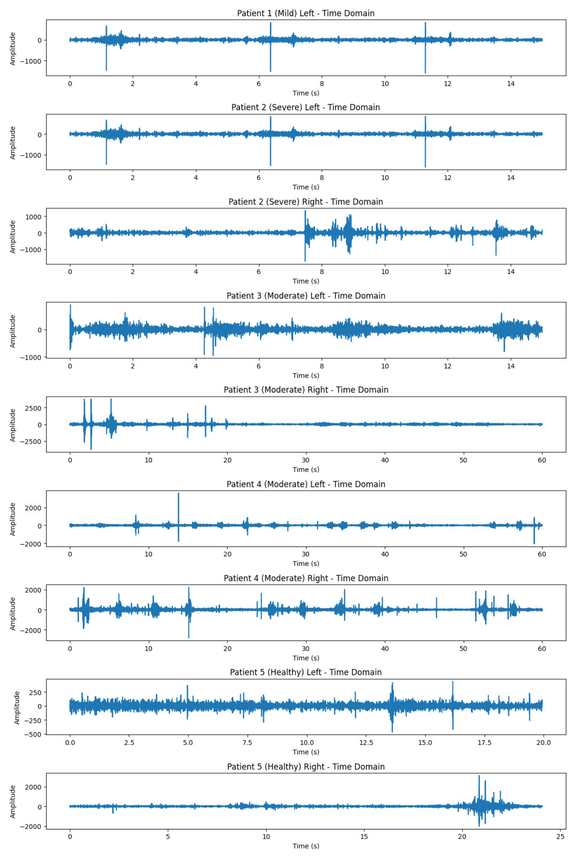
- Digital stethoscopes capture inaudible sound frequencies, offering a new opportunity for screening.
- Project investigates whether audio recordings from a digital stethoscope can be used to detect pulmonary fibrosis earlier using signal processing and machine learning techniques.

Data Collection

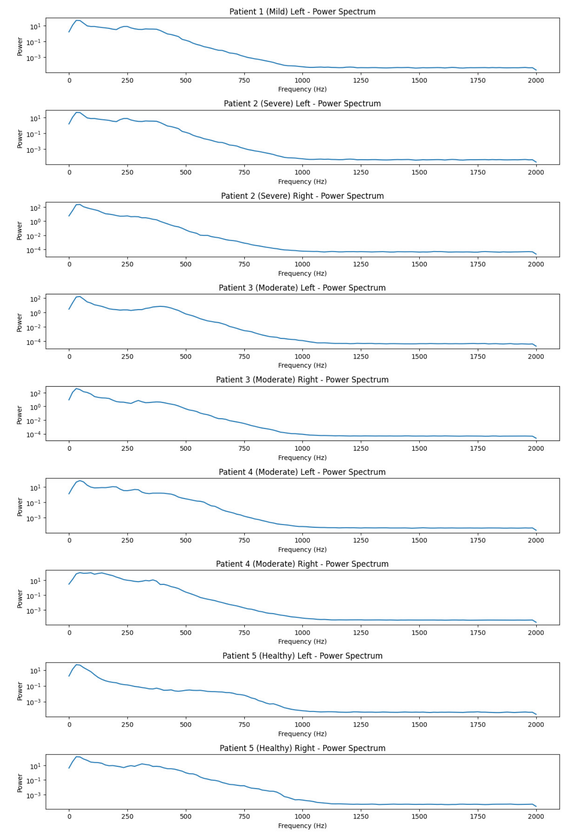
- Breath sounds were recorded from 5 patients (4 with pulmonary fibrosis, 1 healthy) using a 3M Littmann CORE Digital Stethoscope (Eko platform).
- For each patient, lung sounds were recorded from both the left and right lower posterior regions.

Data

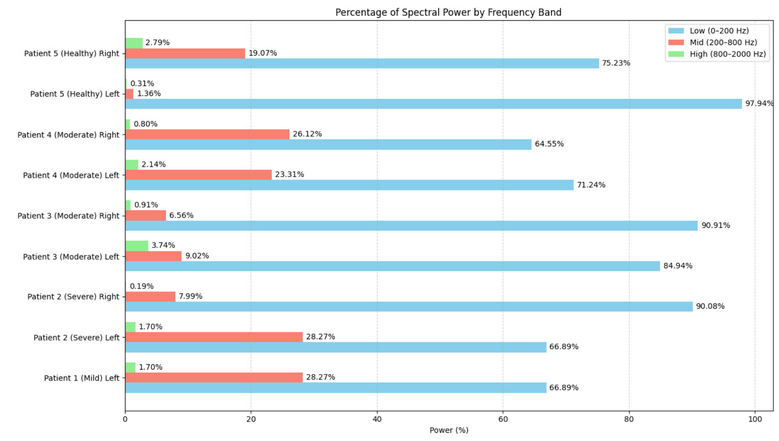
Time Domain



Power Spectrum

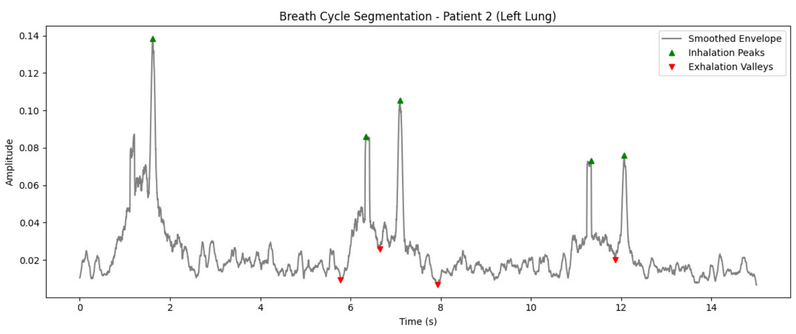


Power Spectrum



Wavelet Scalogram

Breath Cycle



Analysis

- Healthy patients had >95% of power in the low frequency band. Fibrotic patients, especially moderate/severe cases, showed increased mid-band power (up to almost 30%) and slightly elevated high-band power.
- Asymmetries observed in several patients with some having higher mid-frequency power on one side.
- Wavelet analysis shows time-localized bursts of mid-scale energy consistent with crackles. These bursts occur periodically and align with inhalation.

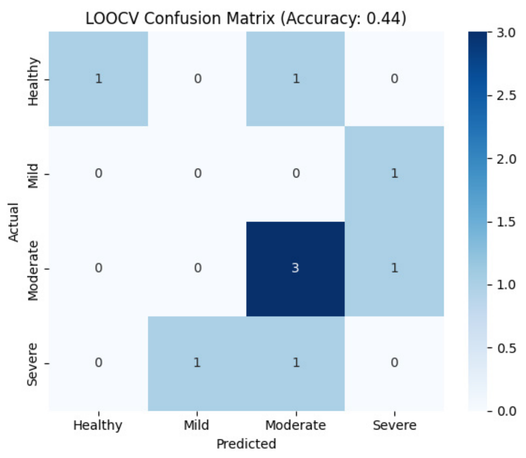
Attempt at Supervised Classification

Applied Leave-One-Out Cross Validation (LOOCV) with a Random Forest Classifier. As expected, five patients was not enough data to have accurate results.

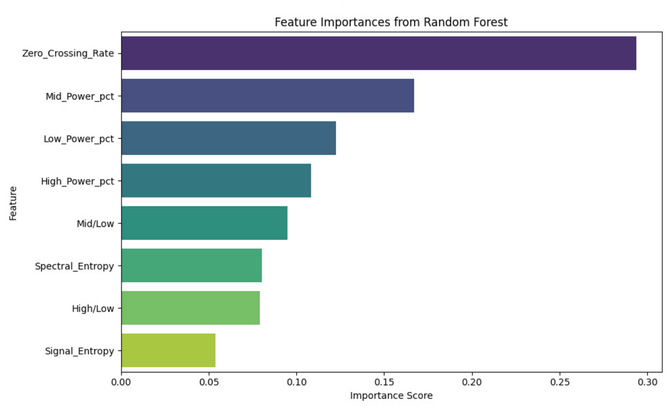
Feature Extraction

- Band Power %: Low (0–150 Hz), Mid (150–600 Hz), High (600–1200 Hz)
- Ratios: Mid/Low, High/Low
- Zero-Crossing Rate
- Signal Entropy
- Spectral Entropy

Confusion Matrix



Feature Importances



Next Steps

Expand the Dataset

- Increase the number and diversity of breath sound recordings by collecting data from more patients across a wider range of fibrosis severities.
- A larger dataset will improve the generalizability of machine learning models.

Compare Pulmonary Fibrosis to Other Respiratory Conditions

- Extend the current analysis by including patients with other respiratory diseases such as asthma and COPD.
- Necessary to reduce false positives in real-world screening.

Conclusion

This early-stage research hints at the potential of digital stethoscope audio and signal processing to detect acoustic patterns associated with pulmonary fibrosis. However, further data collection is critical for robust machine learning classification and generalization. Future work will focus on expanding the dataset and eventually comparing pulmonary fibrosis to other respiratory conditions.

References

Pal, R., Umapathy, K., & Ghoraani, B. (2024). Automated detection of interstitial pulmonary fibrosis using respiratory sounds. medRxiv,
Elizalde, B., Deshmukh, S., Al Ismail, M., & Wang, H. (2023). CLAP: Learning audio concepts from natural language supervision. ICASSP 2023.
Lee, C. S., Li, M., Lou, Y., Abbasi, Q. H., & Imran, M. (2024). An acoustic system of sound acquisition and image generation for frequent and reliable lung function assessment. IEEE Sensors Journal, 24(3), 3731–3741.
Muhammad, M., Li, M., Lou, Y., & Lee, C. S. (2024). Continual monitoring of respiratory disorders via real-time lung sound imaging in telemedicine. Electronics, 13(9), 1669.
Baur, S., Nabulsi, Z., Weng, W. H., Garrison, J., & Shetty, S. (2024). HeAR: Health Acoustic Representations. arXiv preprint arXiv:2403.02522.

Acknowledgments

- Dr. Seamas Donnelly
- Dr. Madeleine Lowery
- Dr. Emer Doheny
- Sinead O'Connor