Deep Learning

Idan Tobis

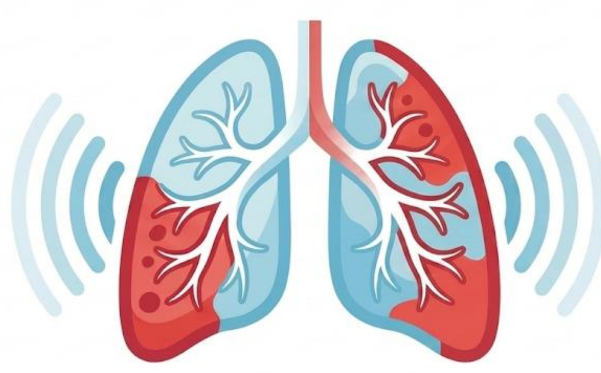
RespiNet

Afik Haviv, Shir Molakandove

סמסטר ב, תשפ״ו

מכון טכנולוגי חולון
Holon Institute of Technology

Project Overview



Introduction

Respiratory diseases like COPD and Pneumonia are a leading cause of mortality worldwide. Traditional diagnosis relies on manual auscultation (listening with a stethoscope), which is subjective, prone to human error, and difficult in noisy environments.

The Goal

To develop an automated, deep-learning-based system capable of classifying respiratory sounds as either Healthy or Pathological with high sensitivity.

The Challenge

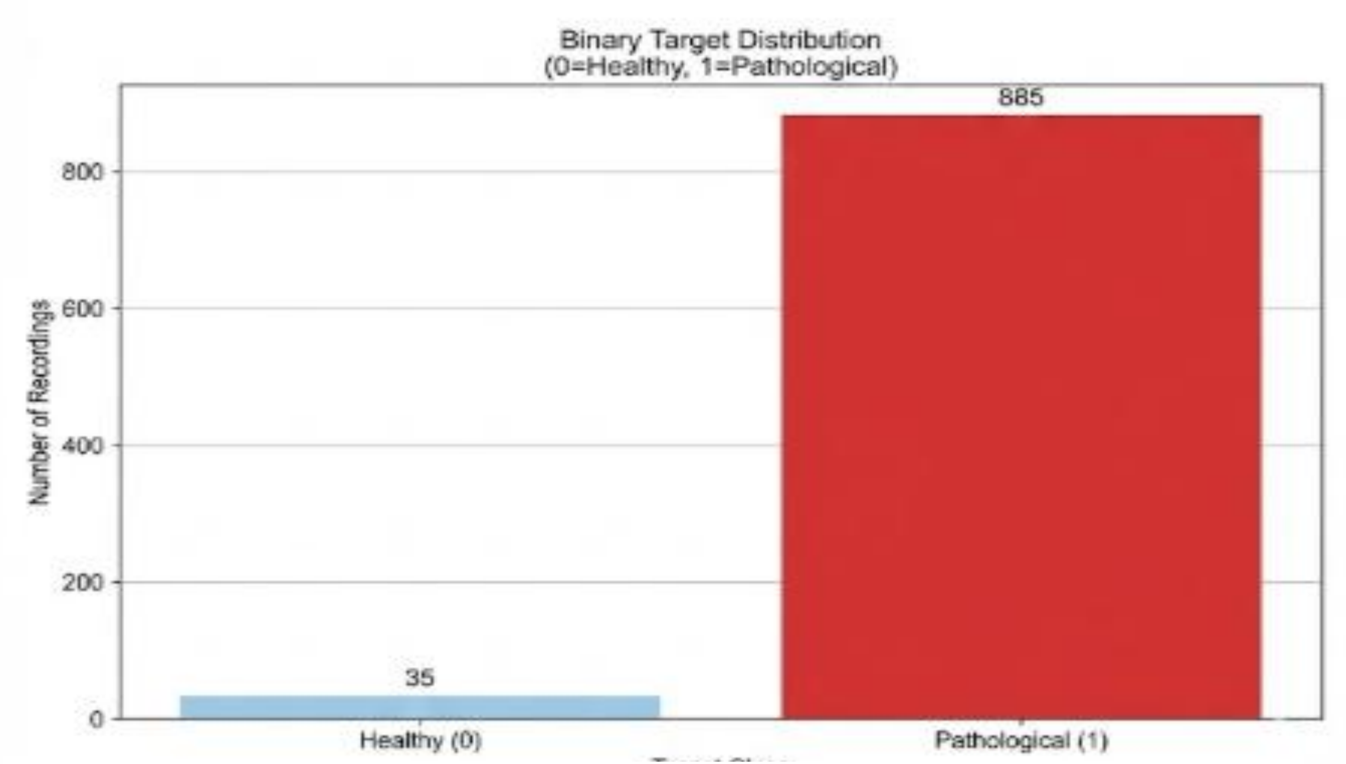
Standard Convolutional Neural Networks (CNNs) often fail to capture long-duration auditory features (like continuous wheezing) because of their limited "Receptive Field." Our project addresses this by introducing Dilated Convolutions to "hear" the whole picture.

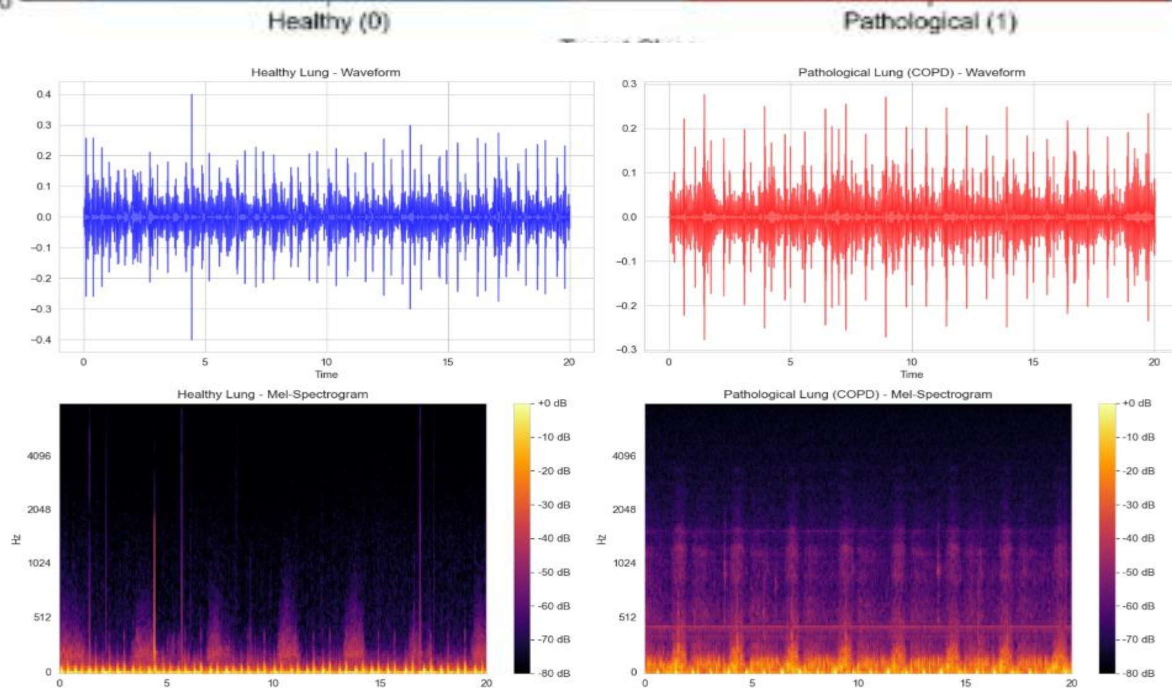
Methodology: The Dilated Approach and Balancing

We processed the ICBHI 2017 Respiratory Sound Database using a novel pipeline for noisy, imbalanced medical data.

Data Preprocessing & Ambiguity Filtering

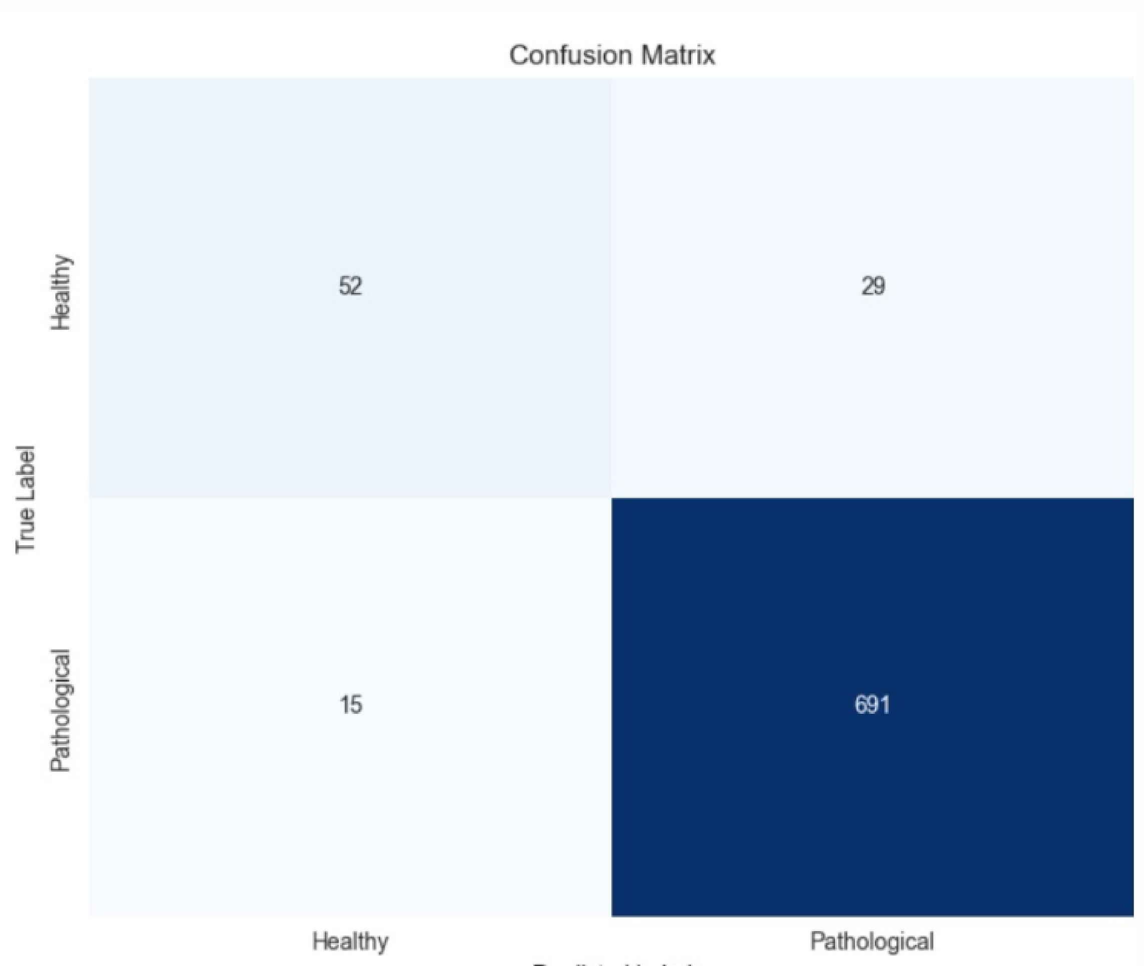
Signal Processing: Raw audio was sliced into breath cycles and converted to Mel-Spectrograms.
The "Ambiguity" Fix: To address labeling issues, we filtered the data, training only on strictly Healthy patients vs. clearly Pathological cycles (Crackles/Wheezes), removing "clean breaths" from sick patients.
Balancing: We used a Weighted Random Sampler for a 50/50 split to prevent majority class bias.





Model Architecture (RespiNet)

We designed a custom CNN with Dilated Layers,
Standard Conv Layers: Capture fine details (high frequencies).Dilated Conv Layers (Dilation=2, 4): Expand the receptive field without losing resolution, allowing detection of long-duration patterns like wheezes.



Conclusion

Architecture Matters: Dilated Convolutions are superior to standard CNNs for respiratory sound analysis because they capture temporal context (duration of wheezes).

Data Quality is King: The biggest performance jump (from 60% to 95%) came from removing ambiguous data ("clean" breaths labeled as "sick") and using Weighted Sampling.

High Sensitivity: The model achieved a 98% detection rate for pathologies, making it a viable tool for initial medical screening.

Discussion

Future Development:

Multi-Class Classification: Expanding the model to distinguish between COPD, Pneumonia, and Bronchiectasis (currently limited by data scarcity).
Edge Deployment: Porting the model to a mobile application for real-time analysis on smartphones.
Noise Cancellation: Integrating active noise suppression to handle non-clinical recording environments.