Using Per-Channel Energy Normalization to Enhance Accuracy in Machine Learning-Based Lung Sound Classification

Rahman, Samer (School: Detroit Country Day School)

Ma, Nathan (School: Detroit Country Day School)

This study investigates various audio preprocessing techniques for classifying lung sounds, particularly Per-Channel Energy Normalization (PCEN). We combined and augmented two primary datasets to create a comprehensive set of labeled audio clips covering a range of respiratory conditions. Using these datasets, we pursued three classification tasks: disease diagnosis; distinguishing between wheezes, crackles, and normal sounds; and differentiating between normal and abnormal lung sounds. Each dataset was processed using several methods, including log-mel spectrograms, Mel-Frequency Cepstral Coefficients (MFCCs), and PCEN spectrograms. PCEN aims to reduce the common background noise in the other processing methods. The datasets were then fed into a convolutional neural network (CNN) for training and evaluation. The CNN architecture utilizes 2D convolutional layers to learn frequency-time features while careful data splitting and augmentation minimize overfitting. Our results support our hypothesis and demonstrate that PCEN consistently outperformed MFCC and log-mel spectrogram methods across all datasets and evaluation metrics, such as loss and accuracy. These results demonstrate PCEN's effectiveness in suppressing background noise and amplifying subtle lung sounds, leading to improved evaluation metrics. The most critical parameters were the smoothing coefficient (T), which controls signal smoothing, and root (r), which balances soft and loud sounds. Despite requiring more computational effort and parameter tuning, PCEN's noise suppression makes it a powerful tool for automated lung sound diagnostics.

Awards Won: