National Institute of Technology Rourkela

Progress Seminar

Seminar Title	: A Lightweight Deep Learning Framework for Respiratory Sound Denoising and Classification on Edge Devices
Speaker	: Subhasikta Behera (Rollno : 522cs1001)
Supervisor	: Puneet Kumar Jain
Venue	: Convention Hall (CS)
Date and Time	: 16 Jul 2025 (05.00 PM)
Abstract	: Respiratory diseases are progressive conditions affecting the lungs and various parts of the respiratory system, significantly impacting the quality of human life. Specifically in rural India, lack of spirometry, misdiagnosis, and stigma prevent effective disease management. Therefore, developing a portable system to detect respiratory disorders at an early stage is crucial to curtail their burden in rural and remote areas. Respiratory sounds are the acoustic signals produced by the respiratory system that healthcare practitioners hear to perform the diagnosis. Various deep learning frameworks have been proposed for the automatic analysis of the respiratory sounds. However, it is often challenged by noise contamination and model limitations like poor generalization, and slow inference speed while deploying on edge devices. This study focuses on a lightweight, portable deep learning framework capable of respiratory sound denoising and classification, tailored for deployment on edge devices in resource-constrained settings. For the denoising task, we introduced a novel adaptive thresholding approach of the DWT coefficients using a UNet model for respiratory sound denoising. Combining the multi-resolution capabilities of DWT with UNet&rsquos multi-scale feature extraction, the method demonstrates its robustness and efficacy on two publicly available datasets with various noise types and levels, outperforming traditional methods. The method achieved a 2.0 dB and 3.0 dB higher SNR than the second-best result in SPRS-23 and ICBHI-17 datasets, respectively, under real-life noise conditions. For the Classification task, we propose a knowledge distillation-based framework leveraging a multi-frequency representation of the signal. The framework, first, employs a sliding window augmentation (SWA) strategy using triangular window-based overlap fusion (TWOF) to address the class imbalance issue and signal enhancement. Then, the multiple time-frequency representations, including enhanced generalized S-transform (EGST), continuous wavelet transform (CWT), and Log-Mel spectrograms, are used to capture diverse spectral-temporal characteristics of lung sounds. These representations are fed into a knowledge distillation framework, where a teacher-student scheme is used to achieve a trade-off between model performance and complexity. In this framework, a Triplet-integrated DenseNet-201 model is used as a teacher model that extracts deep features from each representation and independently transfers them to a lightweight DenseNet-121 student model. Feature outputs from all student models are fused and classified using a parameter-efficient artificial neural network (ANN). The experimental results indicate that the proposed classification method reaches an accuracy score of 0.8073 on seven-category classification tasks in the SPRsound-22 dataset and 0.5074 on four-category classification tasks in the ICBHI-17 dataset. By integrating denoising and classification into a unified, lightweight framework, the system enables it to deliver rapid inference, low computational overhead, and high diagnostic performance, making it ideal for real-time respiratory screening on edge devices in rural and remote regions.