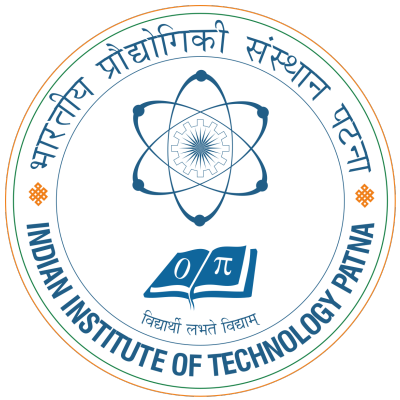
AsTFSONN: A Unified Framework Based on Time-Frequency Domain Self-Operational Neural Network for Asthmatic Lung Sound Classification

by:

Arka Roy, and Dr. Udit Satija

Department of Electrical Engineering

Indian Institute of Technology Patna, Bihar, 801106



Presentation Layout

- Lung sounds
- Asthma
- Description of database
- Proposed framework
- Simulation results
- Conclusion
- References

Lung sounds

- Produced due to turbulent airflow in the trachea-bronchial tree
- Can be heard during the course of chest auscultation
- Linked to structural faults of lungs due to diseases

Table 1: Description of different lung sounds

Lung sound	Frequency range	Inspiratory/ expiratory	Associated disease
Normal	below 1000Hz	Both	Healthy
Wheeze	200-2500Hz	Expiration or both in severe condition	COPD, asthma
Crackle	100-500Hz	Inspiratory	Pneumonia, pulmonary edema

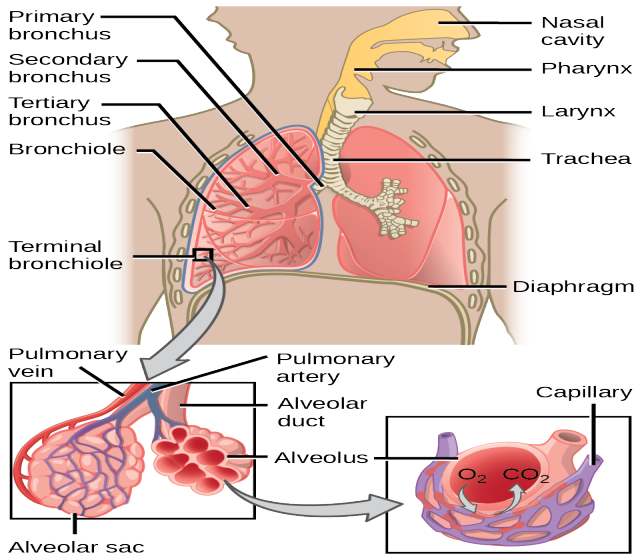


Figure 1: Human respiratory system [1]

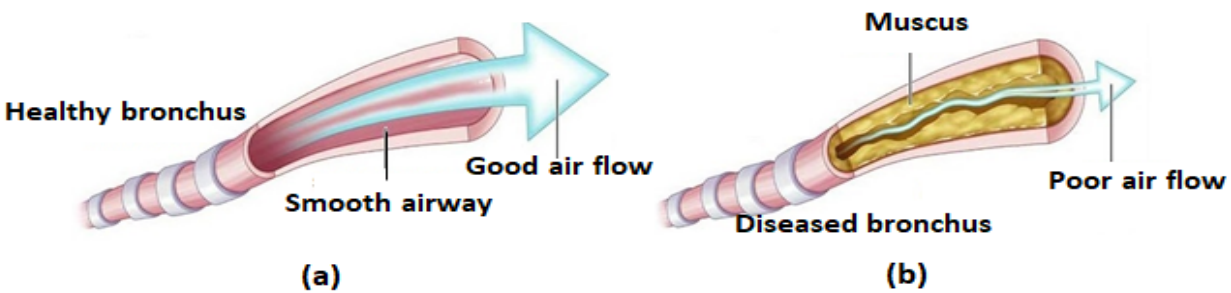


Figure 2: Cross-sectional view of (a) healthy bronchus and (b) Asthma diseased bronchus [2]

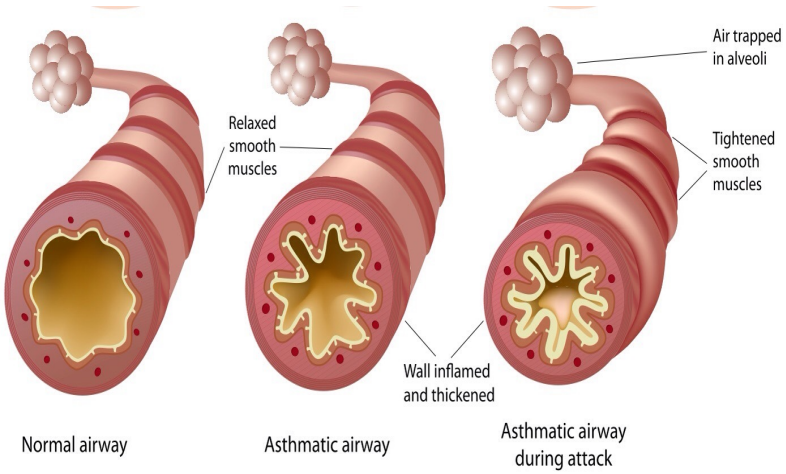
[1] https://www.yaclass.in/anatomy-of-the-respiratory-system

[2] https://my.clevelandclinic.org/health/diseases

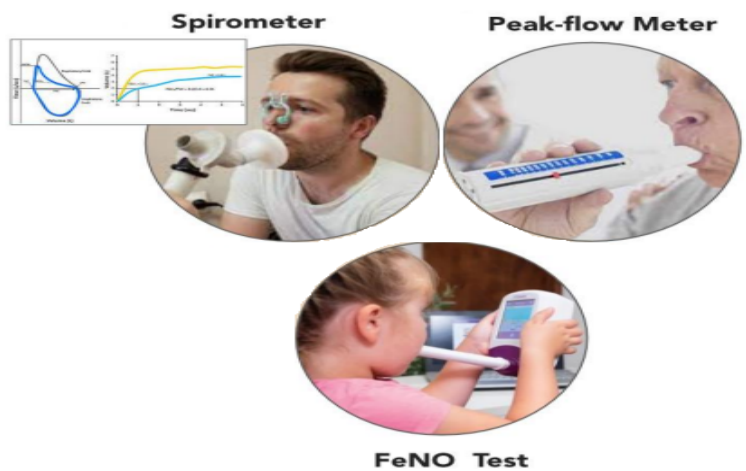
Asthma: a respiratory threat

- A noncommunicable obstructive respiratory disease in which the airways become inflamed, narrowed, produce extra mucus
- Symptoms: coughing, wheezing, shortness of breath
- The global asthma study reported over 339 million people suffer from asthma [1]

Disadvantage existing modalities: Laborious, dependent on patient efforts, and very costly.



a



b

Figure 3: (a) Airway obstructions in Asthma [3], (b) Different diagnostic modalities for asthma [4]

Why lung sounds: Observable wheezing sounds, low-cost non-invasive modality.

[3] https://www.saisivahospital.com/asthma-in-india/, [4] https://www.nhlbi.nih.gov/health/asthma/diagnosis

Database

Database used: Chest wall lung sound database (CWLSD) [2]

Table 2: Description of Chest Wall Lung Sound Database (CWLSD)

Items	Details
Acquisition device	Littmann3200M digital stethoscope
Sampling rate	4000 Hz
Acquisition sites	Both posterior and and anterior side of the body
Length of recordings	Uneven, from 10 sec- 50 sec
Number of recordings in each class	Asthma: 96, bronchiectasis (BRON): 9, pneumonia: 14, heart failure: 56, chronic obstructive pulmonary disease (COPD): 38, fibrosis: 12, pleural effusion: 6, and healthy: 105

Proposed framework: AsTFSONN

Objective: To develop a novel deep learning framework that can provide accurate distinct feature representation of both healthy and asthmatic lung sound signals

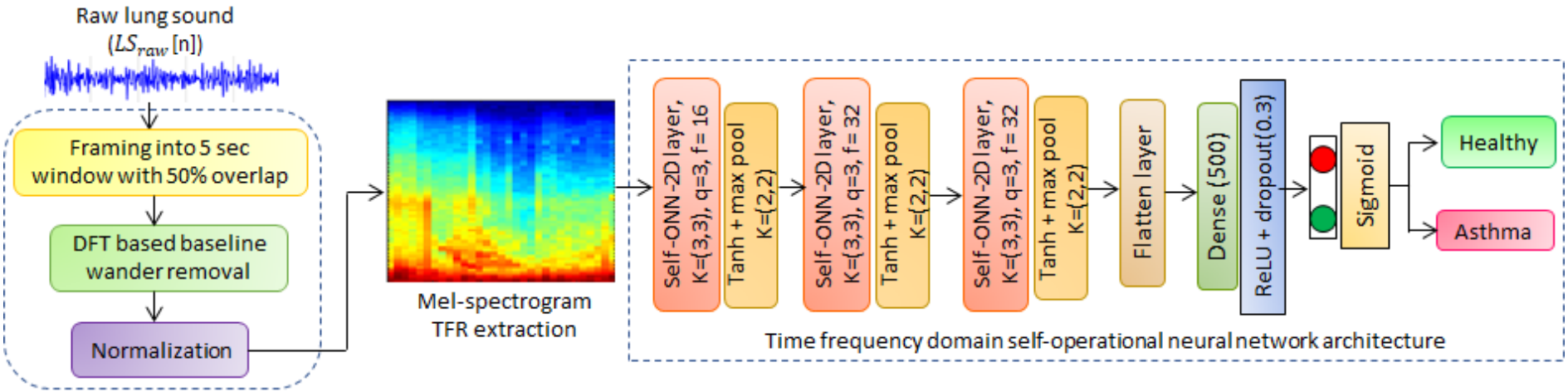


Figure 4: Block diagram of proposed AsTFSONN for asthmatic lung sound classification

Pre-processing

- **Framing:**
  Let, $LS_{raw}[n]$ be the raw lung sound signal, being segmented into 5 sec length (f) frames ($LS^t_{fm}[n]$) keeping 50% overlap (r) with adjacent frame.

$$LS^t_{fm}[n] = LS_{raw}[f \cdot t \cdot r + n], \qquad t = 1, 2, ...U \tag{1}$$

- **DFT based baseline wander (BW) component removal [3]:**
  DFT of the t^{th} frame ($LS^t_{fm}[n]$) is calculated as:

$$LS^t_{fm}[k] = \sum_{n=0}^{N-1} LS^t_{fm}[n] e^{\frac{-j2\pi nk}{N}} \tag{2}$$

  Frequency range of BW component is 0-1 Hz. Thereby, removing DFT coefficients that are smaller than 1 Hz. DFT coefficient index for the f Hz: $k = \lceil \frac{fN}{f_s} \rceil$ where f_s denotes the sampling rate of lung sound. Thresholded DFT coefficient:
  $\tilde{LS}^t_{fm}[k] = [0, 0, ..., 0, LS^t_{fm}[k+1], ..., LS^t_{fm}[N-k-1], 0, ..., 0, 0]$
  Baseline wander removed signal:

$$LS^t_{bwr}[n] = \frac{1}{N} \sum_{k=0}^{N-1} \tilde{LS}^t_{fm}[k] e^{\frac{j2\pi nk}{N}} \tag{3}$$

- **Normalization:**

$$LS^t_{norm}[n] = \frac{LS^t_{bwr}[n]}{\max |LS^t_{bwr}[n]|} \tag{4}$$

Pre-processing (Contd.)

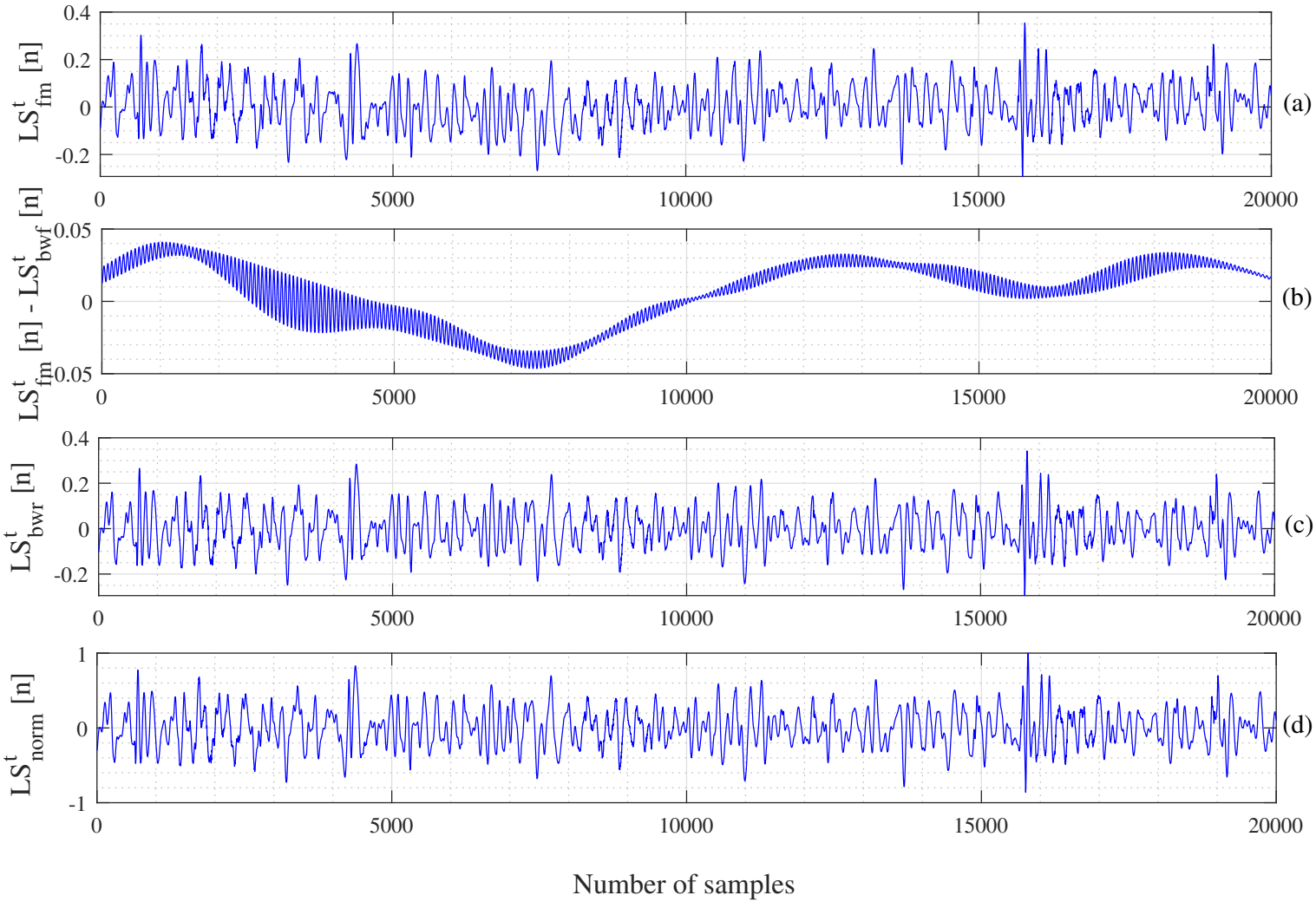


Figure 5: (a) Raw lung sound frame, (b) BW component, (c) BW removed signal, (d) Normalised lung sound frame.

Mel-spectrogram generation

- **Spectrogram extraction:**

$$ST^t_{ls}[l, f] = |\sum_{n=0}^{N-1} LS^t_{norm}[n] \cdot \mathcal{W}[n - l\mathcal{H}] \cdot e^{-j\frac{2\pi nf}{N}}|^2 \quad (5)$$

$$\mathcal{W}[n] = 0.54 - 0.46 \cdot \cos(\frac{2\pi n}{N-1}) \quad (6)$$

where $ST_{ls}[l, f]$, and $\mathcal{W}[n]$ refer to the short-time Fourier transform-based spectrogram, and the Hanning window, with an overlap-length ($\mathcal{H}$).

- **Map the Hertz frequency components to mel frequency scale:**
Convert the lowest and highest Hertz frequency to mel $\{f^L, f^H\} \rightarrow \{f^L_{mel}, f^H_{mel}\}$ using the following equation [4]

$$f_{mel} = 2595 \cdot \log(1 + \frac{f}{700}) \quad (7)$$

- **Create triangular mel-filter banks:** $linspace\{f^L_{mel}, f^H_{mel}, B\}$ where B is number of mel filters. In this paper, we have considered $B = 64$
- **Mel-spectrogram = Spectrogram $\times$ triangular mel filter banks**

Visualization

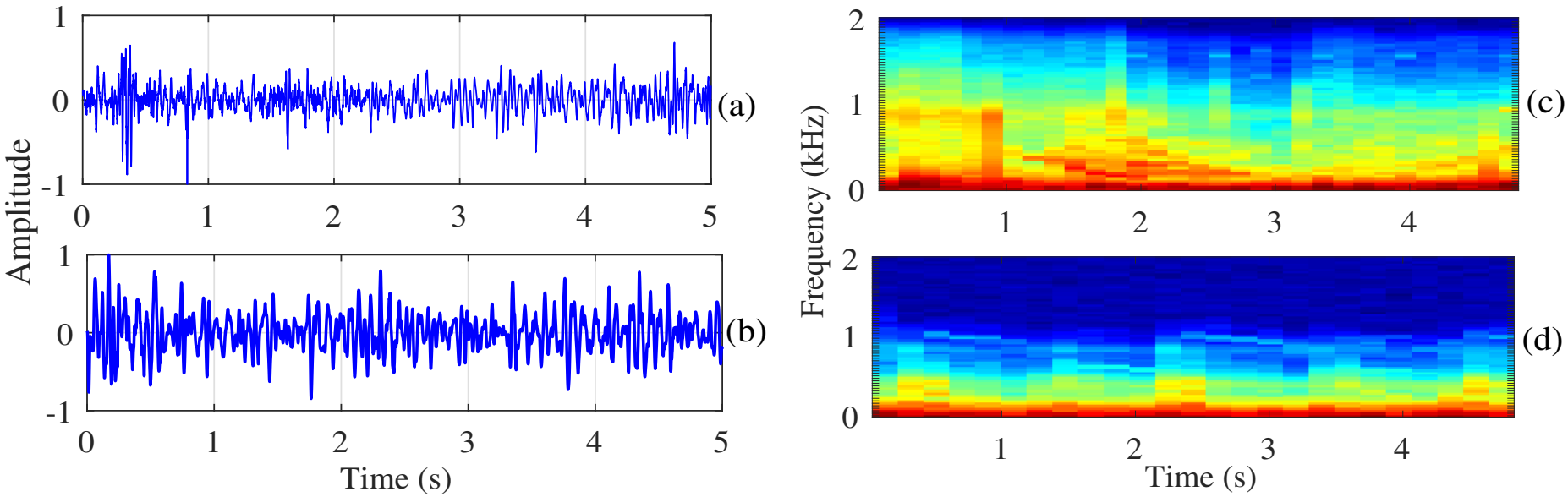


Figure 6: (a), (b) The normalized time-domain visualization of asthmatic and healthy lung sound, and (c), (d) The mel-spectrogram representation of asthmatic and healthy lung sound respectively.

Self-operation neural network (SONN)

Let, $\mathscr{X} \in \mathscr{R}^{L \times B \times D_{in}}$ denotes an input tensor, which has height L, width of B, and depth of D_{in}, and $\mathscr{K}_{Dout}$ indicates the D_{out}^{th} kernel of a layer. And output is $\mathscr{Y}_{Dout} \in \mathscr{R}^{L \times B \times D_{out}}$.

- **Working of linear neurones of CNN:**

$$\begin{aligned} \mathscr{Y}_{Dout} &= \sum_{i,j,d_{in}=1}^{L,B,D_{in}} \mathscr{K}_{Dout}(i,j,d_{in})\mathscr{X}(i,j,d_{in}) + \mathscr{B}_{Dout} \\ &= \mathscr{K}_{Dout}^{T}\mathscr{X} + \mathscr{B}_{Dout} \end{aligned} \tag{8}$$

- **Working of operational neuron (ONN) [5]:**

$$\mathscr{Y}_{Dout} = \phi\{\mathscr{K}_{Dout}, \mathscr{X}\} + \mathscr{B}_{Dout} \tag{9}$$

where ϕ is a learnable nodal operator: a combination of several distinct functions [5]. However, it is a very cumbersome search process leading to larger training time [5].
Remark: $\phi\{\cdot\}$ is found to be the dot product, then ONN layer becomes a typical CNN layer.

- **Working of SONN [6]:**

$$\begin{aligned} \phi(\mathscr{K}_{Dout,1}, \mathscr{K}_{Dout,2} \ldots, \mathscr{K}_{Dout,q}, \mathscr{X}) &= \sum_{e=1}^{q} \mathscr{K}_{Dout,e}^{T} \cdot \mathscr{X}^{e} \\ &= \mathscr{K}_{Dout,1}^{T}\mathscr{X} + \mathscr{K}_{Dout,2}^{T}\mathscr{X}^{2} + \ldots + \mathscr{K}_{Dout,q}^{T}\mathscr{X}^{q} \end{aligned} \tag{10}$$

where $\mathscr{X}^{q}$ is q^{th} power of the original tensor $\mathscr{X}$ and $\mathscr{K}_{Dout,q}$ denotes the learnable

Model architecture and simulation results

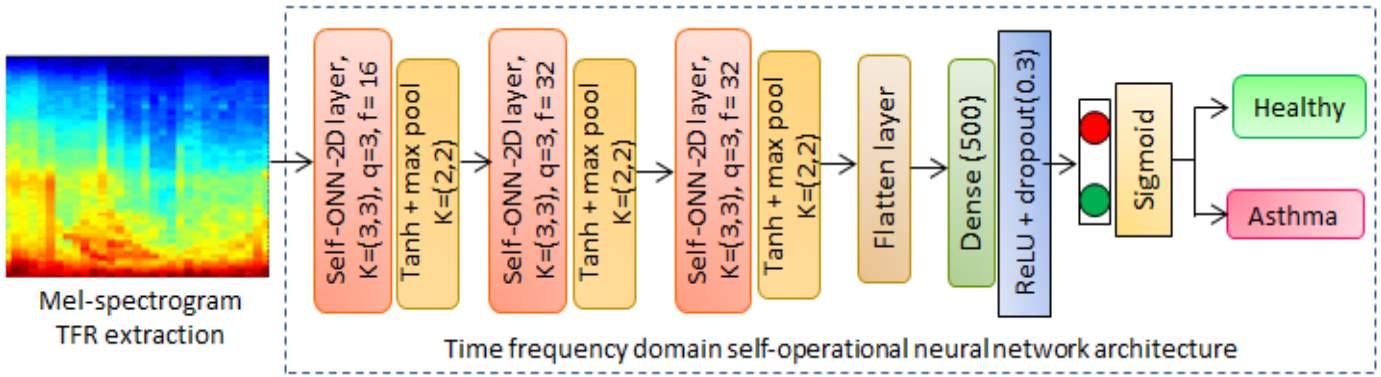


Figure 7: Architecture of proposed AsTFSONN model

Table 3: Model Simulation Parameters for AsTFSONN

Simulation parameter	Details
Learning rate	0.001
Epochs	300
Optimizer	Adam
Batch size	128
Loss function	Binary cross entropy

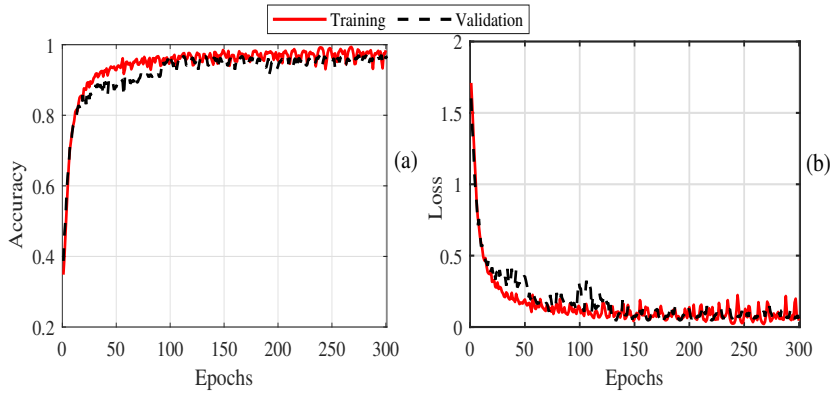


Figure 8: (a-b) Classification accuracy and loss curves of our proposed AsTFSONN.

Simulation results (contd.)

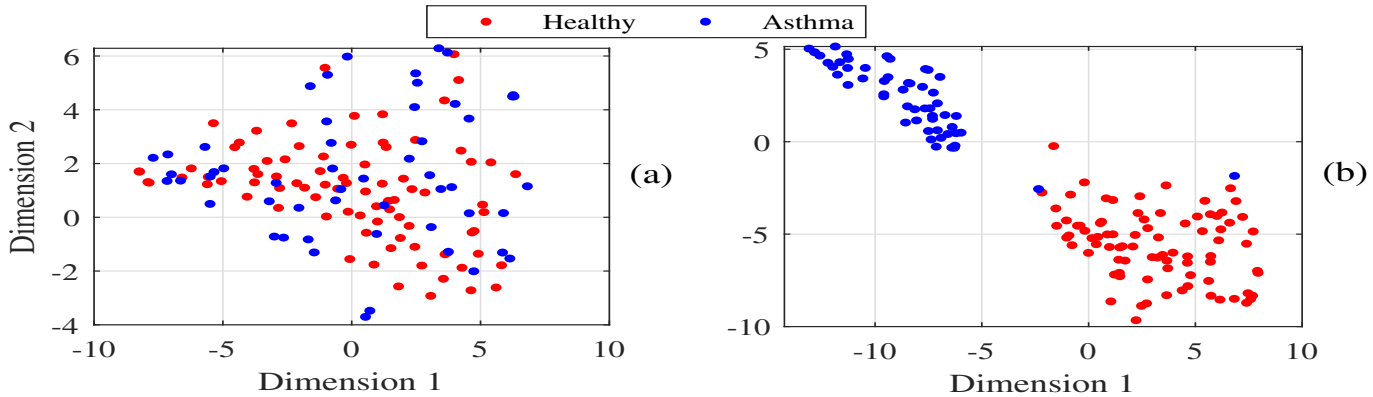


Figure 9: 2D t-SNE visualization for (a) input raw lung sound signal, and (b) feature embeddings of the classified lung sound signals, taken from the dense layer of the proposed AsTFSONN.

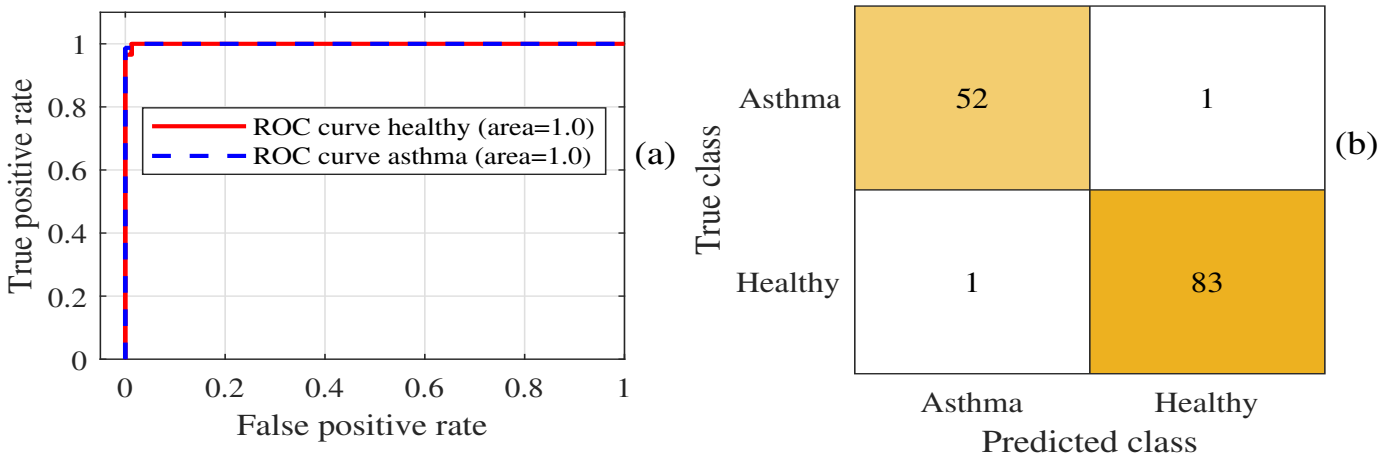


Figure 10: (a) Illustrates the ROC curve of healthy and asthmatic lung sounds, and (b) the confusion matrix of the test signals from both classes using AsTFSONN.

Simulation results (contd.)

Performance metrics: Accuracy (ac), specificity (sp), sensitivity (sn) or recall (rc), precision (pr), F1-score, and ICBHI-score.

$$ac = \frac{t_p+t_n}{t_p+t_n+f_n+f_p}, \quad sn/rc = \frac{t_p}{t_p+f_n}, \quad sp = \frac{t_n}{t_n+f_p}, \quad pr = \frac{t_p}{t_p+f_p},$$
$$F1-score = \frac{2\times pr\times rc}{pr+rc}, \quad ICBHI-score = \frac{sn+sp}{2}$$

Table 4: Performance Comparison of AsTFSONN with Existing Works

Reference	Data type (database)	Methodology	Results (%)			
			ac	sn	sp	ICBHI score
Yadav et al. [7]	Speech (own)	MFCC features, SVM	75.40	–	–	–
Altan et al. [8]	Lung sound (own)	Filtering, HHT, statistical features, DBN	84.61	85.83	77.11	81.47
Islam et al.[9]	Lung sound (own)	ANN, SVM	89.20	90.00	86.74	88.36
Tripathy et al. [10]	Lung sound (CWLSD)	EWT, temporal-spectral features, ML classifiers	80.35	84.88	75.23	80.05
Proposed framework	**Lung sound (CWLSD)**	**Melspectrogram driven AsTFSONN**	**98.50**	**98.11**	**98.80**	**98.46**

Simulation results (contd.)

Table 5: Classification Performance Comparison with Respect to Different Deep Learning Architectures

Deep learning architecture topology	**Trainable parameter**	**Performance metrics (%)**					
		ac	**sn**	**sp**	**pr**	**F1 score**	**ICBHI score**
CNN (3 layers, K=(3*3), f= 16,32,32)	5,91,550	81.02	92.45	73.80	69.01	79.02	83.13
CNN (4 layers, K=(3*3), f= 16,32,32,64)	9,06,110	79.56	62.96	90.36	80.95	70.83	76.66
CNN (5 layers, K=(3*3), f= 16,32,32,64,128)	12,33,918	80.43	90.06	71.94	67.44	77.12	81.00
AsTFSONN baseline version (3 layers, K=(3*3), f= 16,32,32)	**6,19,486**	**98.50**	**98.11**	**98.80**	**98.11**	**98.11**	**98.46**

Conclusion

- Establishes the fact that lung sounds play an important role in asthma detection
- Outperform all the traditional ML approaches and reduces the burden of manual feature extraction
- The framework demonstrates the effectiveness of employing SONN over simple CNN for lung sound-based asthma classification

References I

S. Salvi, G. A. Kumar, R. Dhaliwal, K. Paulson, A. Agrawal, P. A. Koul, P. Mahesh, S. Nair, V. Singh, A. N. Aggarwal, "The burden of chronic respiratory diseases and their heterogeneity across the states of India: the Global Burden of Disease Study 1990–2016." *The Lancet Global Health* 6.12 (2018): e1363-e1374.

M. Fraiwan, L. Fraiwan, B. Khassawneh, and A. Ibnian, "A dataset of lung sounds recorded from the chest wall using an electronic stethoscope," *Data in Brief,* vol. 35, p. 106913, 2021.

U. Satija, B. Ramkumar, and M. Sabarimalai Manikandan, "Real-time signal quality-aware ECG telemetry system for IoT-based health care monitoring," *IEEE Internet of Things Journal,* vol. 4, no. 3, pp. 815–823, 2017.

O. Ilyas, "Pseudo-colored rate map representation for speech emotion recognition," *Biomedical Signal Processing and Control,* vol. 66, p. 102502, 2021.

S. Kiranyaz, T. Ince, A. Iosifidis, and M. Gabbouj, "Operational neural networks," *Neural Computing and Applications,* vol. 32, pp. 6645–6668, 2020.

S. Kiranyaz, J. Malik, H. B. Abdallah, T. Ince, A. Iosifidis, and M. Gabbouj, "Self-organized operational neural networks with generative neurons," Neural Networks, vol. 140, pp. 294–308, 2021.

S. Yadav, M. Keerthana, D. Gope, P. K. Ghosh, "Analysis of acoustic features for speech sound based classification of asthmatic and healthy subjects," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, 2020, pp. 6789–6793.

G. Altan, Y. Kutlu, A. Ö. Pekmezci, and S. Nural, "The diagnosis of asthma using hilbert-huang transform and deep learning on lung sounds," *arXiv preprint* arXiv:2101.08288, 2021.

M. A. Islam, I. Bandyopadhyaya, P. Bhattacharyya, and G. Saha, "Multichannel lung sound analysis for asthma detection," *Computer Methods and Programs in Biomedicine,* vol. 159, pp. 111–123, 2018.

R. K. Tripathy, S. Dash, A. Rath, G. Panda and R. B. Pachori, "Automated detection of pulmonary diseases from lung sound signals using fixed-boundary-based empirical wavelet transform," in *IEEE Sensors Letters*, vol. 6, no. 5, pp. 1-4, 2022.

Thank You