A new lung disorder detection model based on graphene pattern using respiratory sounds

Speech Communication 181 (2026) 103414

Contents lists available at ScienceDirect

Speech Communication

journal homepage: www.elsevier.com/locate/specom

A new lung disorder detection model based on graphene pattern using respiratory sounds

Prabal Datta Barua [a], Omer Faruk Goktas [b], Sengul Dogan [c], Nursena Baygin [d], Mehmet Baygin [d], Massimo Salvi [e,*], Turker Tuncer [c], Ru-San Tan [f,g], U.R. Acharya [h]

[a] School of Business (Information System), University of Southern Queensland, Toowoomba, QLD 4350, Australia
[b] Department of Electronics and Automation, Technical Sciences Vocational School, Ankara Yildirim Beyazit University, Ankara, Türkiye
[c] Department of Digital Forensics Engineering, College of Technology, Firat University, Elazig, Türkiye
[d] Department of Computer Engineering, College of Engineering, Erzurum Technical University, Erzurum, Türkiye
[e] Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy
[f] Department of Cardiology, National Heart Centre Singapore, Singapore
[g] Duke-NUS Medical School, Singapore
[h] School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, Australia

ARTICLE INFO

Keywords:
Digital stethoscope
Lung sound
Signal classification
Bidirectional wavelet transformation

ABSTRACT

Background and purpose: Auscultatory sounds acquired using a stethoscope can offer clinical clues to the presence of cardiorespiratory diseases. In this work, we aimed to develop an accurate and lightweight model for disease detection using lung sounds.

Method: Our model comprises: (1) signal decomposition using a multilevel bidirectional wavelet transformation; (2) multilevel feature generation using a novel lattice-based graphene pattern to create minimum- and maximum directed graphs to extract textural features; (3) feature selection using iterative neighborhood component analysis; (4) classification using a standard shallow k-nearest neighbor function. We tested the model on a public 336-subject eight-class lung sound dataset. Model performance was reported for eight- and three-class diagnostic classification.

Results: Our model achieved accuracy rates exceeding 94 % for all classification tasks. The maximum distance path through the graphene pattern consistently outperformed the minimum distance path, indicating that significant amplitude transitions in respiratory sounds contain more discriminative information than regions of relative uniformity. Elements of the input signal and wavelet decomposition bands that contributed most to the selected feature vector were visualized, which enhanced model explainability and revealed that low-pass filtered wavelet coefficients, particularly the L3 band, were most informative for classification.

Conclusion: Our handcrafted computationally lightweight model yielded accurate and explainable results. These attributes facilitate potential integration into digital stethoscopes for point-of-care screening of respiratory diseases.

1. Introduction

Cardiorespiratory diseases are among the leading causes of death globally (Labaki and Han, 2020), with cardiovascular disease, cancer, and chronic obstructive pulmonary disease constituting the top three causes (WHO, 2022). Auscultatory sounds acquired using a stethoscope can offer clinical clues to the presence of heart and lung diseases (Kim et al., 2022) and are important for disease screening. With the widespread adoption of digital stethoscopes (Leng et al., 2015), heart and respiratory sounds can be easily recorded, digitized, and archived for later/remote manual review by experts or analysis using artificial intelligence-enabled algorithms (Alsuliman et al., 2020).

Indeed, automated cardiorespiratory disease detection using digitized lung sounds has emerged as a popular research field (Muthusamy

* Corresponding author.
E-mail addresses: prabal.barua@usq.edu.au (P.D. Barua), ofgoktas@aybu.edu.tr (O.F. Goktas), sdogan@firat.edu.tr (S. Dogan), nursena.baygin@erzurum.edu.tr (N. Baygin), mehmet.baygin@erzurum.edu.tr (M. Baygin), massimo.salvi@polito.it (M. Salvi), turkertuncer@firat.edu.tr (T. Tuncer), tanrsnhc@gmail.com (R.-S. Tan), Rajendra.Acharya@usq.edu.au (U.R. Acharya).

https://doi.org/10.1016/j.specom.2026.103414
Received 4 September 2025; Received in revised form 31 January 2026; Accepted 11 May 2026
Available online 12 May 2026

et al., 2020), enabling applications ranging from routine clinical from routine clinical screening to specialized use cases such as rapid triage of Covid-19 patients (Soltan et al., 2022). The acoustic characteristics of respiratory sounds contain valuable diagnostic information that, when properly analyzed, can distinguish between various pathological conditions affecting the lungs.

In recent years, numerous approaches have been developed for lung sound classification. In (A.M. Alqudah et al., 2022), convolutional neural networks were used for classifying seven lung sounds–normal, coarse crackle, fine crackle, monophonic wheeze, polyphonic wheeze, squawk, and stridor-based on two-dimensional spectrogram images generated from one-dimensional audio files. Messner et al. (Messner et al., 2020) developed a convolutional recurrent neural network for classifying multi-channel lung sounds and attained an F1-score value of 92 % for the binary classification of lung sound signals into healthy vs. pathological classes. Jayalakshmy and Sudha (Jayalakshmy and Sudha, 2021) combined empirical mode decomposition with bidirectional long short term memory in their four-class lung sound classification model. Tasar et al. developed a model that used a novel piccolo pattern-based feature extractor to generate features from lung sounds (Tasar et al., 2022). Their model attained over 99 % classification success. Wu et al. built an electronic stethoscope that could acquire heart and lung sounds (Wu et al., 2022). Using mel frequency cepstral coefficients to process the acquired sounds and statistical feature extraction, their method attained 86.9 % and 73.3 % accuracy rates for heart and lung sound classification, respectively.

More recently, advanced deep learning architectures have been proposed for respiratory sound classification. Roy and Satija developed RDLINet, a lightweight inception network achieving 96.6 % accuracy for seven-class respiratory disease classification (A. Roy and Satija, 2023). The same research group later proposed Pulmo-TS2ONN, a triple-scale self-operational neural network that achieved 98.88 % accuracy for pulmonary disorder detection (Roy et al., 2024). For COPD-specific detection, Roy et al. introduced a multi-head self-organized operational neural network (MHSONN) utilizing triplet time-frequency representations (Roy and Satija, 2024), and VGAResNet, a visibility graph adjacency matrix-based residual network (Roy et al., 2023). Asthma detection has also received attention, with AsthmaSCELNet employing supervised contrastive embedding learning (A. Roy and Satija, 2023) and AsTFSONN using time-frequency domain self-operational neural networks (Roy and Satija, 2023). Additionally, mel-spectrogram snippet representation learning has been applied for COPD severity detection (A. Roy and Satija, 2023), and automated severity detection approaches have been explored (Roy and Satija, 2022).

1.1. Motivation and contributions

Most of the published studies on lung sound classification used deep models, which typically have high computational complexity (Fraiwan et al., 2022; Petmezas et al., 2022). Lastly, many studies have focused on distinguishing the different types of respiratory sounds (e.g., crackle, wheeze, etc.) rather than actual disease diagnosis, which is clinically more relevant (Palaniappan et al., 2013; Soni et al., 2022).

In recognition of these limitations, we were motivated to develop a computationally efficient handcrafted lung sound detection model that could discriminate diagnostic classes accurately. We were inspired by the molecular structure of the graphene molecule to design a lattice-based pattern, graphene pattern, for feature extraction. This was combined with a new bidirectional wavelet transformation, which decomposed input signals into wavelet bands for downstream multilevel feature generation and a simple but effective parametric feature selector and classifier. Our main contributions in this paper include:

- Development of a novel signal processing methodology that combines multilevel bidirectional wavelet transformation with a lattice-based feature extraction approach (graphene pattern) inspired by molecular structures. This combination allows for effective signal decomposition and extraction of discriminative features from respiratory sounds.
- Design of a computationally lightweight model architecture that achieves high-performance classification (>94 % accuracy) for both eight-class and three-class diagnostic problems without requiring specialized computing hardware. The model integrates feature selection through iterative neighborhood component analysis (INCA) and classification using a standard k-nearest neighbor approach.
- Emphasis on clinical applicability, which provides a more realistic assessment of model performance by keeping all samples from the same subject together during validation. This approach, combined with the model's explainability features that visualize which input signal components contribute most to classification decisions, enhances the potential for clinical implementation.

2. Material and method

2.1. Dataset

We employed a publicly available lung sound database with 336 samples (sampling rate of 4 kHz) that were acquired through an electronic stethoscope and divided into eight categories: “asthma”, “bronchitis”, “chronic obstructive pulmonary disease” (“COPD”), “heart failure” (“HF”), „normal", pleural effusion, pneumonia and lung fibrosis (Fraiwan et al., 2021). In the present study we cut the sounds into two-second pieces (8000 data -Table 1).

In this study, two classification systems were assessed. In Case 1, all eight diagnostic classes were employed separately, which corresponded to the overall multiclass diagnosis among different respiratory diseases. For Case 2, we classified the eight classes into three clinically meaningful categories: chronic obstructive pulmonary disease (COPD), normal, and non-chronic obstructive pulmonary disease (made of the remaining six classes-asthma, bronchitis, heart failure pleural effusion pneumonia lung fibrosis). This three-class method yielded 222 COPD segments, 853 normal segments and 1647 non-COPD segments. Case 2 is focused on a classification problem that brings to attention the case of COPD detection, since there are millions of cases worldwide where this disease was only detected when it was too late (WHO, 2022). It also represents a more balanced point of view with regard to the ability of the model to detect/save this very severe pathology while properly recognizing normal and other pathological situations.

2.2. Time-frequency characteristics of respiratory sounds

To gain insight into specific acoustic characteristics associated with each physiological condition, we conducted TF analysis of lung sound signals in all 8 disease categories. This analysis sheds light on the biological foundation of our feature extraction method.

The time-domain waveforms of example samples from each class are shown in Fig. 1. Clear breath sounds are associated with relatively

Table 1
Details of the dataset used for both classification schemes.

Disorder	Number of records	Number of observations	
		Case 1 (8-class problem)	Case 2 (3-class problem)
Normal	105	853	853
COPD	27	222	222
Asthma	96	765	1647
Bronchitis	9	54	
Heart failure	63	501	
Pleural effusion	6	75	
Pneumonia	15	141	
Lung fibrosis	15	111	
Total	*336*	*2722*	*2722*

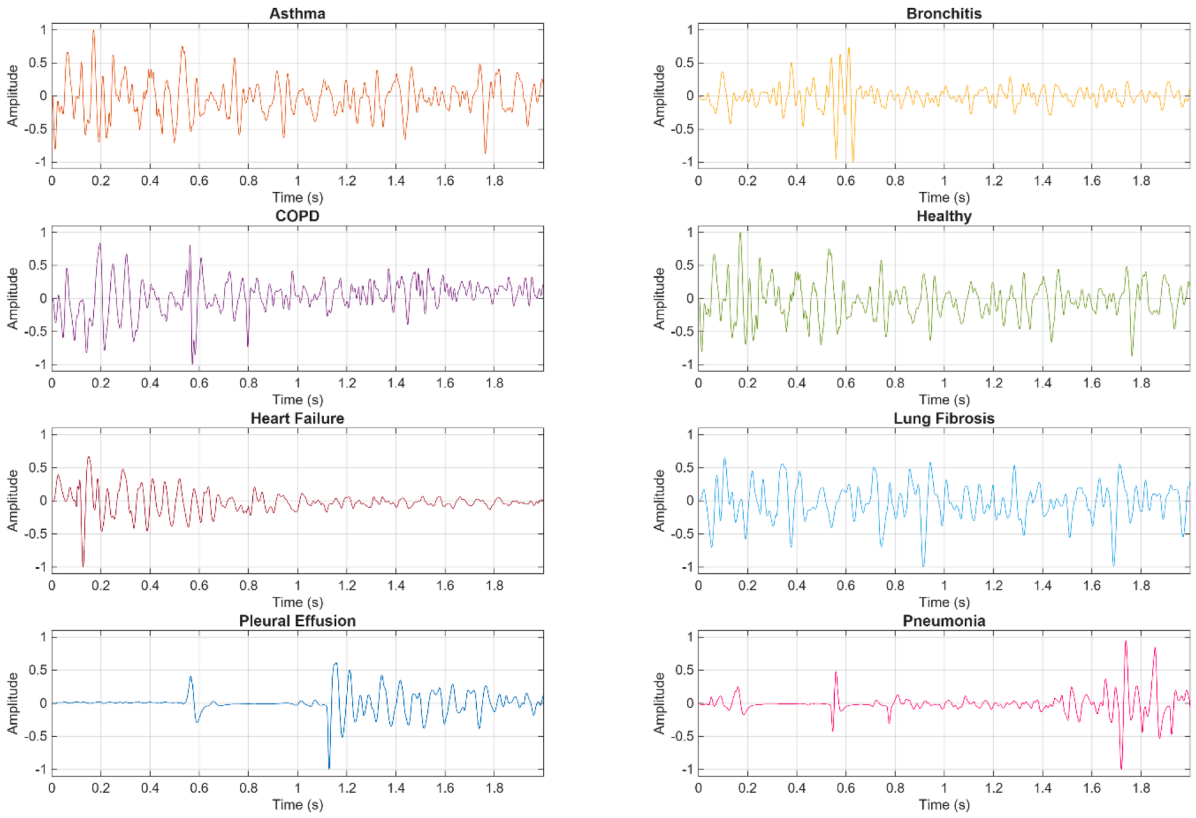

Fig. 1. Time-domain representation of respiratory sounds for all eight disease classes. Each subplot displays a 2-second segment normalized to unit amplitude. Healthy sounds show regular breathing patterns, while pathological conditions exhibit characteristic irregularities.

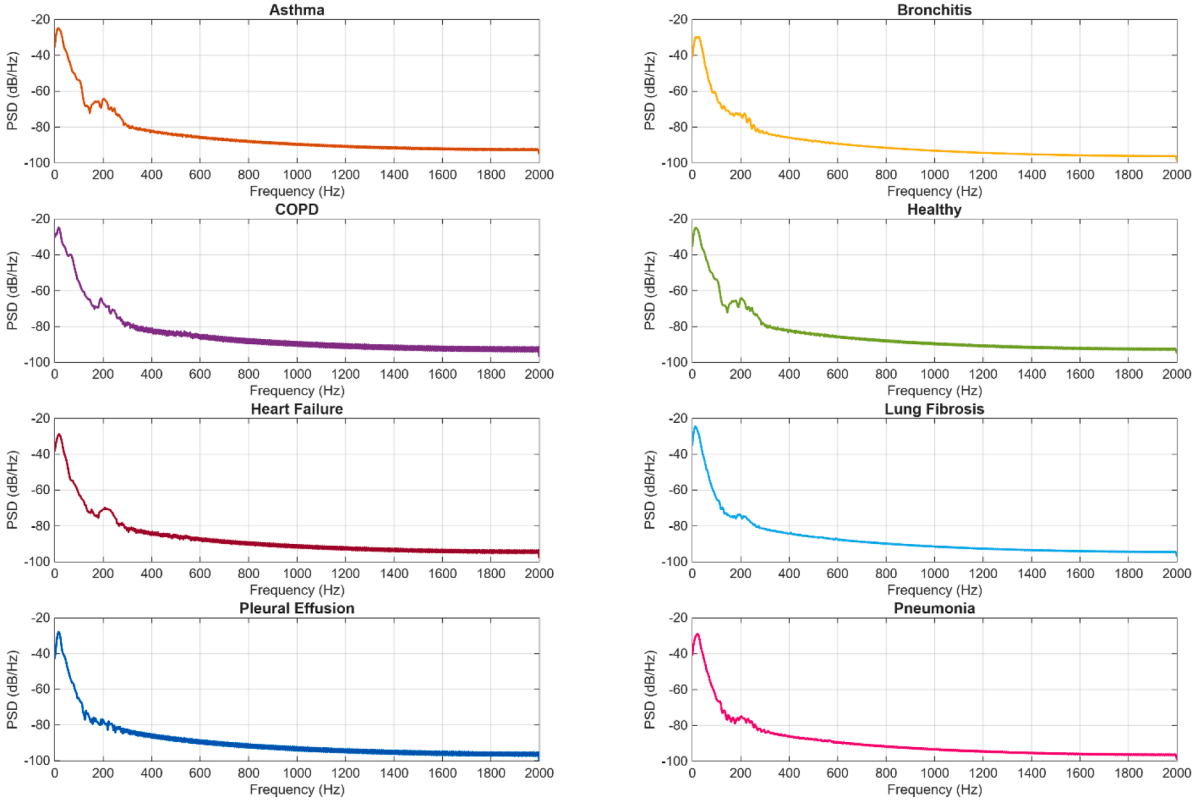

Fig. 2. Power spectral density (PSD) analysis of respiratory sounds across disease classes computed using Welch's method. The frequency distribution patterns reveal class-specific acoustic signatures that form the basis for discrimination.

regular, sinusoidal breathing patterns of consistent amplitude and resembling unobstructed airflow. In contrast, pathological conditions have characteristic deviations: asthma and COPD show oscillatory expiration due to airway constriction, pneumonia and lung fibrosis display high amplitude transients corresponding to crackles and pleural effusion has lower overall amplitude due to the presence of fluid.

As we analyzed the power spectral density (PSD) in Fig. 2, each condition has different frequency signatures. Wheeze-related diseases (asthma, COPD) have a strong energy focus in the 200–500 Hz band, known as the characteristic frequency of wheezing sounds generated through turbulent airflow from narrowing airways. Conditions related with crackles (pneumonia, lung fibrosis) have more extended spectral content (>500 Hz) due to the impulsive nature of crackle sounds. Heart failure shows low-frequency dominance below 200 Hz, and is sometimes associated with fine crackles due to pulmonary edema. Pleural effusion exhibits overall decreased spectral energy levels at all frequency bands, which results from damping by pleural fluid.

The waveform representations are depicted in Fig. 3 to visually demonstrate the time evolution of frequency content, thus offering a time-frequency depiction of the respiratory sounds. Healthy sounds show uniform, smooth distribution of energy across the breathing cycle. In the asthma and COPD spectrograms, harmonic structures are observed during expiration that look similar to musical wheeze sounds. Pneumonia and lung fibrosis exhibit vertical lines in the spectrogram, which are impulsive crackles with a very broadband frequency content. HF has both low-frequency cardiac interference and fine crackle patterns. Bronchitis manifest with more continuous low-frequency bands representing rhonchi from airway secretions.

These class-specific time–frequency domain (TFD) discrepancies provide a physiological rationale for our feature extraction strategy. In particular, the consistently superior performance of the maximum distance path can be explained by the prevalence of abrupt amplitude transitions in pathological respiratory sounds, such as crackles and wheezes, which carry more discriminative information than slowly varying or quasi-stationary signal components.

2.3. The proposed graphene pattern

Inspired by the molecular structure of graphene, we designed a lattice pattern that could dynamically generate directed graphs (fitness functions) for extracting textural features from input signal blocks using the signum function (Fig. 4). The hexagonal structure of graphene provides a framework for creating interconnected paths that can capture complex patterns in respiratory sounds. Steps involved in feature generation are detailed below.

2.3.1. Signal blocking and vertex population

The process begins by dividing the input respiratory sound signal into overlapping blocks that match the graphene lattice structure. Each block populates the vertices of our graphene pattern:

$$v^i(j) = S(i+j-1), i \in \{1,2,\ldots,\mathscr{L}-43\}, j \in \{1,2,\ldots,44\} \tag{1}$$

where v^i represents the i^{th} overlapping block, S is input one-dimensional signal, and $\mathscr{L}$ is length of signal. Each overlapping block effectively maps a segment of the signal to the graphene pattern vertices, allowing us to analyze local signal characteristics through the lens of this structural framework.

2.3.2. Distance calculation and path generation

After populating the vertices, we calculate distances between connected vertices using the L1-norm (Manhattan distance). For instance, where d represents the distance matrix, $d(5,9) = |v^i(5) - v^i(9)|$ would generate a distance matrix of size 43×44 for this lattice. From this distance matrix, we generate patterns with minimum and maximum distances by identifying optimal starting points:

$$\begin{aligned} fw^1 &= argmin(d(1:4,:)) \\ p^1(1) &= fw^1 - 4 \\ fw^2 &= argmax(d(1:4,:)) \\ p^2(1) &= fw^2 - 4 \end{aligned} \tag{2}$$

where p^1 and p^2represent patterns 1 and 2 with the minimum and

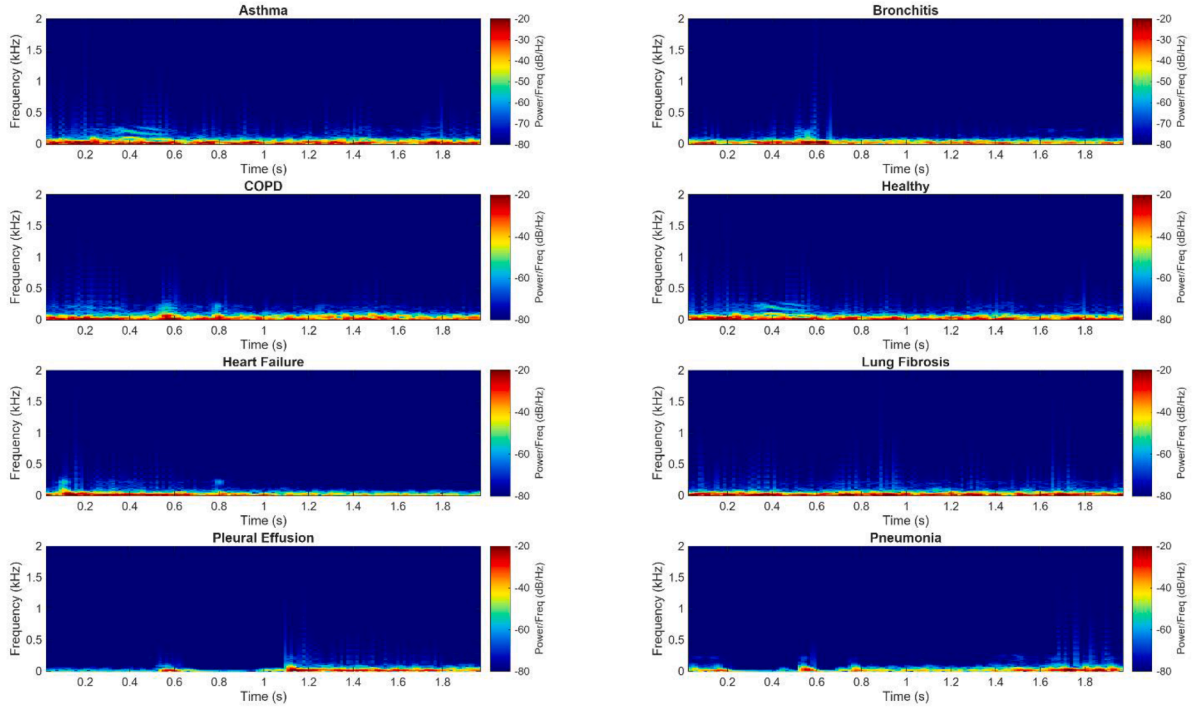

Fig. 3. Spectrogram (time-frequency) visualization of respiratory sounds for each disease class. Color intensity represents power spectral density in dB/Hz. The temporal patterns reveal the distinct acoustic signatures of continuous sounds (wheezes, rhonchi) versus discontinuous sounds (crackles).

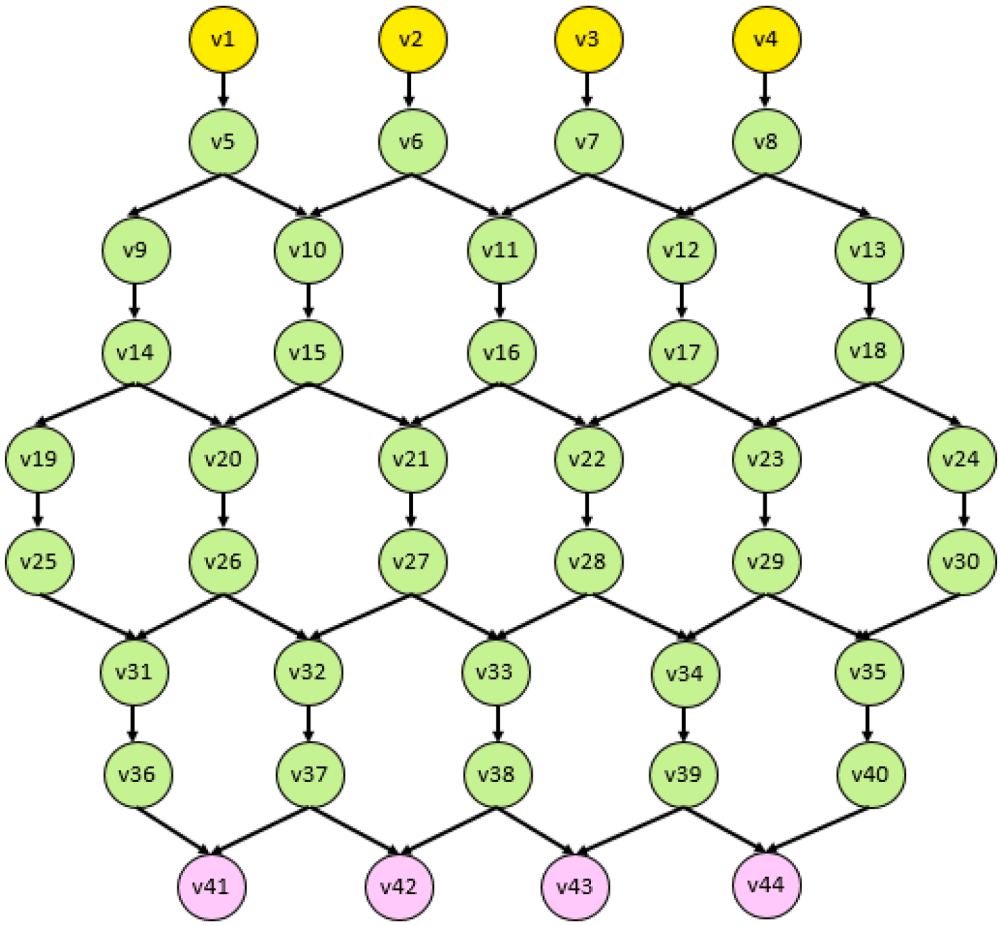

Fig. 4. Architecture of the proposed graphene pattern. It consists of a nine-tiered lattice with 44 vertexes connected by 56 directed edges (20 vertical and 36 obliques; all pointing down). As the numbered vertexes become populated by serial values within each input signal block, two walking paths with the maximum and minimum L1-norm distances were traced starting at v1, v2, v3 or v4, and terminating at v41, v42, v43 or v44. These paths serve as directed graphs for feature extraction using signum function of that specific signal input.

maximum distances, fw^1 and fw^2, respectively. Here, the minimum and maximum distances of the first four values of the graphene pattern (v1, v2, v3, and v4) are used to determine the first values of and p^2. The key innovation here is identifying two distinct paths through the lattice - one that follows connections with minimum distance (representing subtle signal variations) and one that follows connections with maximum distance (representing sharp signal transitions).

We then generate eight additional path nodes according to the first values by iteratively selecting connections with minimum or maximum distances:

$$\begin{aligned} w^1 &= argmin\left(d\left(p^1(k),:\right)\right), k \in \{1,2,\ldots,8\} \\ p^1(k+1) &= w^1 \\ w^2 &= argmax\left(d\left(p^2(k),:\right)\right) \\ p^2(k+1) &= w^2 \end{aligned} \tag{3}$$

where w^1 and w^2 represent the indices of vertices with minimum and maximum distance values relative to the current positions in the two created patterns. This step creates two distinct "walking paths" through the graphene structure that capture different aspects of the signal's local behavior - one path follows areas of signal similarity, while the other highlights areas of significant signal change.

2.3.3. Feature extraction using signum function

Once the paths are established, we extract binary features using these paths together with the signum function:

$$\begin{aligned} bit_1^i(q) &= \phi\left(v^i\left(p^1(q)\right), v^i\left(p^1(q+1)\right)\right),\ q \in \{1,2,\ldots,8\} \\ bit_1^i(q) &= \phi\left(v^i\left(p^2(q)\right), v^i\left(p^2(q+1)\right)\right) \\ \phi\left(v^i\left(p^1(q)\right), v^i\left(p^1(q+1)\right)\right) &= \begin{cases} 0, & v^i\left(p^1(q)\right) - v^i\left(p^1(q+1)\right) < 0 \\ 1, & v^i\left(p^1(q)\right) - v^i\left(p^1(q+1)\right) \geq 0 \end{cases} \end{aligned} \tag{4}$$

where bit_1^i and bit_2^i represent the first and second-bit vectors, each of length 8, of the i^{th} block; and $\phi(.)$ is the signum function that encodes the relative magnitude relationship between consecutive vertices in each path. The signum function effectively transforms the signal magnitude differences into a binary pattern, capturing the directional changes (increases or decreases) along each path through the graphene lattice.

2.3.4. Feature vector generation

From the binary patterns extracted along our graphene-inspired lattice paths, we create two map signals by converting the binary patterns to decimal values. This conversion transforms the sequential binary representations into compact numerical features that capture the signal's structural characteristics:

$$\begin{aligned} m^1(i) &= \sum_{n=1}^{8} bit_1^i(n)\ \times 2^{n-1} \\ m^2(i) &= \sum_{n=1}^{8} bit_2^i(n)\ \times 2^{n-1} \end{aligned} \tag{5}$$

where m^1 and m^2 represent the first and second map signals, respec-

tively. Each 8-bit binary pattern (from the minimum and maximum paths) is converted to a decimal value between 0 and 255, creating a one-dimensional signal that encodes the directional changes observed along the paths through the graphene lattice. The final step involves generating feature vectors by extracting histograms from these map signals:

$$\begin{aligned} f^1 &= \lambda(m^1) \\ f^2 &= \lambda(m^2) \end{aligned} \tag{6}$$

where f^1 and f^2 represent feature vectors, each of length 256; and $\lambda(.)$ is the histogram extraction function that counts the frequency of each possible decimal value (0–255) in the map signals.

These histogram-based feature vectors provide a statistical representation of the signal's behavior throughout the entire respiratory sound segment. The feature vectors capture how frequently certain patterns occur in the signal, creating a robust representation that is invariant to time shifts while preserving the essential textural characteristics of the respiratory sounds. The two complementary feature vectors—derived from minimum and maximum distance paths—offer different perspectives on the signal dynamics:

- The minimum distance path f^1 captures subtle variations and gradual transitions in the signal, highlighting regions of relative stability.
- The maximum distance path f^2 emphasizes abrupt changes and significant transitions, detecting distinctive acoustic events that may be characteristic of specific respiratory conditions.

Together, these feature vectors provide a comprehensive representation of the respiratory sound's structural properties, enabling effective discrimination between different pathological conditions. The dimensionality of these features ($256 \times 2 = 512$) strikes a balance between expressiveness and computational efficiency, making our approach suitable for real-time applications while maintaining high classification performance.

2.4. Proposed model

Starting from the graphene pattern feature extraction described in Section 2.2, we developed a model for respiratory sound classification that integrates several complementary signal processing and machine learning components (Fig. 5). The model architecture was designed to maximize classification performance while maintaining computational efficiency. The graphene pattern-based feature engineering method has the following components: (1) signal decomposition using a multilevel bidirectional wavelet transformation; (2) multilevel feature generation; (3) feature selection using iterative neighborhood component analysis (INCA) (Tuncer et al., 2020; Tuncer et al., 2021); (4) classification using shallow k-nearest neighbor (kNN) function (Peterson, 2009) with LORO (CV). Details of the individual phases are described in the following section.

2.4.1. Signal preprocessing and wavelet decomposition

After segmenting the respiratory recordings into 2-second intervals, we applied a novel multilevel bidirectional wavelet transformation (MDWT) (Ye and Hou, 2015) to decompose each segment into multiple frequency bands. Unlike conventional wavelet approaches, our bidirectional MDWT combines aspects of discrete wavelet transformation with wavelet packet decomposition, analyzing both high and low-frequency components:

$$\begin{aligned} [l^1, h^1] &= \psi(S) \\ [l^{t+1}, h^{t+1}] &= \psi(l^t), t \in \{1, 2, \ldots, lv-1\} \\ [l^{lv+1}, h^{lv+1}] &= \psi(h^1) \\ [l^{q+1}, h^{q+1}] &= \psi(h^q), q \in \{lv+1, lv+2, \ldots, 2lv-2\} \end{aligned} \tag{7}$$

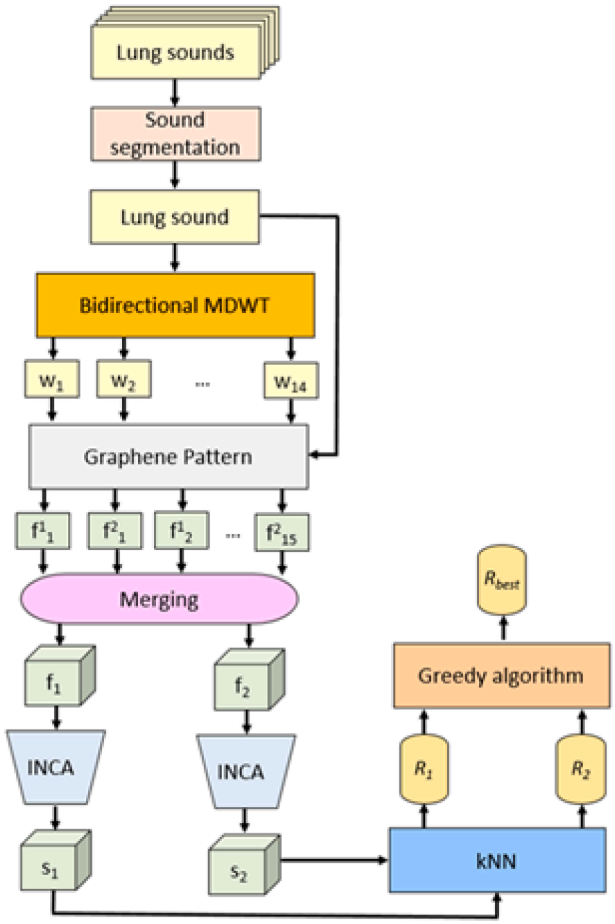

Fig. 5. The pipeline flows from signal preprocessing through wavelet decomposition, feature extraction, selection, and classification. F: feature vector; R: results; s: selected feature vector; w: wavelet bands.

where l and h represent the low- and high-pass filtered wavelet coefficients, respectively; $\psi(.)$ denotes the bidirectional wavelet transformation function (implemented using symlet 4 mother wavelet function for its effectiveness in noise elimination); S is the input signal segment; $\mathscr{L}$ represents the signal segment length; and lv is the number of decomposition levels.

For our 2-second segments (8000 samples at 4 kHz), we calculated the optimal number of levels as:

$$lv = \left\lceil \log_2\left(\frac{\mathscr{L}}{44}\right) \times \frac{1}{2} \right\rceil = 4 \tag{8}$$

The Symlet-4 (sym4) mother wavelet was selected due to its near-symmetric properties and effectiveness in biomedical signal analysis, particularly for preserving both time and frequency localization. This 4-level bidirectional MDWT generates 14 wavelet bands ($2^4 - 2$), providing a rich multi-resolution representation of the respiratory sounds.

2.4.2. Multilevel feature generation

We applied the graphene pattern feature extraction method (described in Section 2.2) to both the original signal and all generated wavelet bands:

$$\begin{aligned} F^1 &= \omega(f_1^1, f_2^1, \ldots, f_{15}^1) \\ F^2 &= \omega(f_1^2, f_2^2, \ldots, f_{15}^2) \end{aligned} \tag{9}$$

where $\omega(.)$ is the feature merging function; and F^1 and F^2 represents the concatenated feature vectors formed by merging feature vectors gener-

ated by the minimum- and maximum-distance directed graphs) with 3840 (=256 × 15) dimensions each.

2.4.3. Feature selection with iterative neighborhood component analysis

To reduce dimensionality and improve classification efficiency, we employed an iterative neighborhood component analysis (INCA) approach (Tuncer et al., 2020; Tuncer et al., 2021). Unlike standard NCA, our iterative version automatically identifies the optimal subset of features through a loop-based optimization process. The iteration range of 100–1000 features was set based on the total feature dimension (3840), where the lower bound ensures sufficient discriminative capacity and the upper bound prevents overfitting. The optimal feature counts were automatically identified by minimizing classification loss.

$$\begin{aligned} id^1 &= \mu(F^1, y) \\ id^2 &= \mu(F^2, y) \\ x_1^{t-sv+1}(d,j) &= F^1\left(d, id^1(j)\right), t \in \{sv, sv+1, \ldots, fv\}, \\ d &\in \{1,2,\ldots,N\}, j \in \{1,2,\ldots,t\} \\ x_2^{t-sv+1}(d,j) &= F^1\left(d, id^2(j)\right) \\ \ell^1(t-sv+1) &= \mathscr{C}\left(x_1^{t-sv+1}, y\right) \\ \ell^2(t-sv+1) &= \mathscr{C}\left(x_2^{t-sv+1}, y\right) \end{aligned} \tag{10}$$

where $\mu(.)$ is the NCA function; *id* contains the qualified index generated by NCA for each feature vector; *x* represents selected features; *sv* and *fv* are the starting and final values of the iteration loop, respectively; *N* is the number of lung sounds; ℓ is the loss value for each selected feature calculated by the classifier function $\mathscr{C}(.)$. The selected feature vector containing an optimal number (between *sv* and *fv*) of the most discriminative feature vectors that yielded the minimum loss value was selected as defined below:

$$\begin{aligned} in^u &= argmin(\ell^u),\ u \in \{1,2\}, \\ s^u(d,j) &= F^u(d, id^u(j)),\ j \in \{1,2,\ldots,in^u + sv - 1\} \end{aligned} \tag{11}$$

where *in* represents the index of the minimum loss value, and *s*, is the selected feature vector.

2.4.4. Classification using kNN

For classification, we employed the k-nearest neighbor (kNN) algorithm with a Leave-One-Recording-Out (LORO) cross-validation strategy:

$$R^u = \kappa(s^u, y) \tag{12}$$

where *R* represents the predicted labels; and $\kappa(.)$ is the kNN classifier function with hyperparameters: $k = 1$, Euclidean distance metric, and stochastic gradient descent optimization.

Finally, we implemented a greedy algorithm to automatically select the best classification result between the two feature paths:

$$\begin{aligned} acc^u &= \theta(R^u, y) \\ b &= argmax(acc) \\ p &= R^b \end{aligned} \tag{13}$$

where acc represents calculated accuracy; θ(.) is the accuracy calculation function; b is the index of the most accurate predicted label vector; and p is the last predicted label vector of the proposed model. This self-optimizing approach eliminates the need for manual parameter tuning and selects the most effective feature representation (minimum or maximum distance) for each classification task. The proposed model has minimal hyperparameters requiring manual tuning, as key parameters are either mathematically derived from signal properties or automatically optimized during the INCA process.

3. Experimental results

3.1. Implementation details

The proposed method was implemented on a standalone personal computer equipped with 32 GB of RAM, an Intel i7–7700 processor operating at 4.20 GHz, and running Windows 11. The implementation does not rely on any graphical or tensor processing unit (GPU). The computational cost of the proposed model was evaluated on this platform. On average, processing a single 2-second respiratory sound segment required 0.12 s, including approximately 0.03 s for signal decomposition (MDWT), 0.05 s for feature extraction (graphene pattern), 0.02 s for feature selection (INCA, performed once and stored for subsequent use), and 0.02 s for classification using kNN.

The complete training process on the entire dataset (2722 segments), using the LORO cross-validation protocol, required approximately 8.5 min. By comparison, deep learning-based approaches typically require 30–60 min of training time on GPU-based systems. Given its lightweight design, CPU-only implementation, and small memory footprint (approximately 50 MB of RAM), the proposed model is well suited for deployment in digital stethoscopes and other resource-constrained point-of-care devices.

The main parameters applied for each stage of the model are shown in Table 2. In the feature extraction stage, a bidirectional MDWT with symlet 4 mother wavelet and 4 decomposition levels was used. The graphene pattern was configured with an overlapping block length of 44 and L1-distance for graph generation. The feature selection phase utilized INCA with an iteration loop range from 100 to 1000, while the classification phase employed kNN with $k = 1$, Euclidean distance, and LORO cross-validation.

For INCA-based feature selection, we observed that the minimum loss values occurred at feature vector lengths of 319 and 380 for the 8-class and 3-class classification problems, respectively (Fig. 6). This demonstrates how our iterative approach automatically identifies the optimal dimensionality for each classification task without manual parameter tuning.

3.2. Classification performance

Both the 8-class (Case 1) and 3-class (Case 2) classification problems were evaluated using standard performance metrics for the minimum and maximum directed graph functions. The confusion matrices for both classification problems are presented in Fig. 7. For the 8-class problem (Case 1), we observe that most misclassifications occurred in the

Table 2
Transition table of the proposed bidirectional MDWT and graphene pattern-based model.

Phase	Method	Parameters
Feature extraction	Bidirectional MDWT	Mother wavelet function: sym4 Number of levels: 4
	Graphene pattern	The length of the overlapping block: 44, Graph generation function: L1-distance, Feature extraction kernel: Signum function, Feature vectors: Two feature vectors are generated, and the length of each feature vector is 256.
	Feature generation	Two feature vectors have been generated, and the length of each feature vector is 3840 (=256 × 15).
Feature selection	INCA	Loop range: from 100 to 1000, loss function: kNN, optimizer: stochastic gradient descend (sgd)
Classification	kNN	k: 1, distance: Euclidean, voting: None, validation: LORO CV.
Best result selection	Greedy algorithm	A maximum accurate predicted labels vector has been selected.

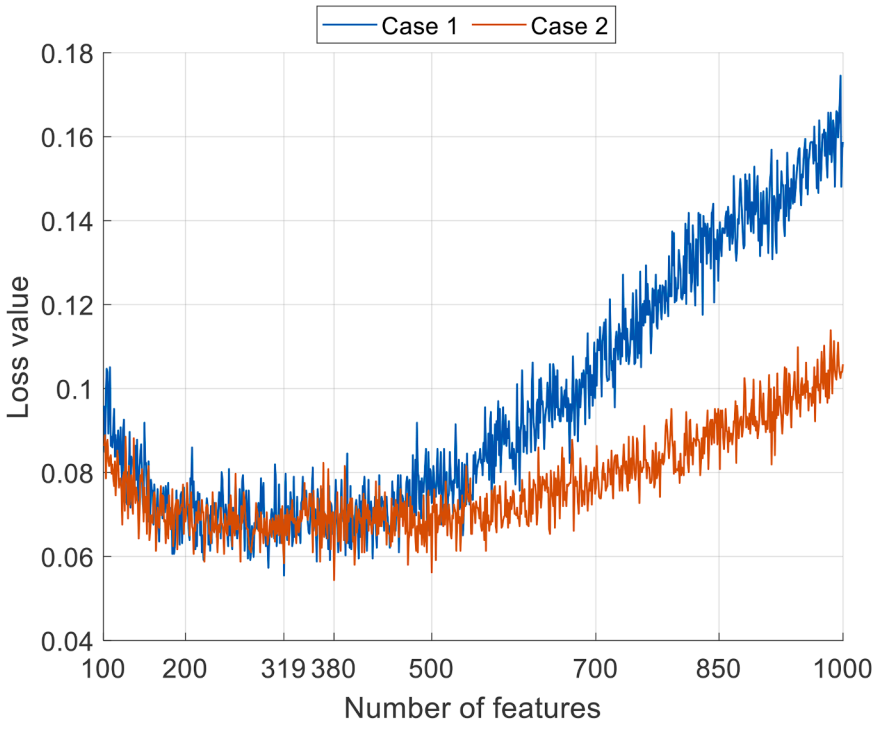

Fig. 6. INCA iterative feature selection process for both Case 1 (8-class) and Case 2 (3-class) classifications. The minimum loss values occur at 319 and 380 features, respectively.

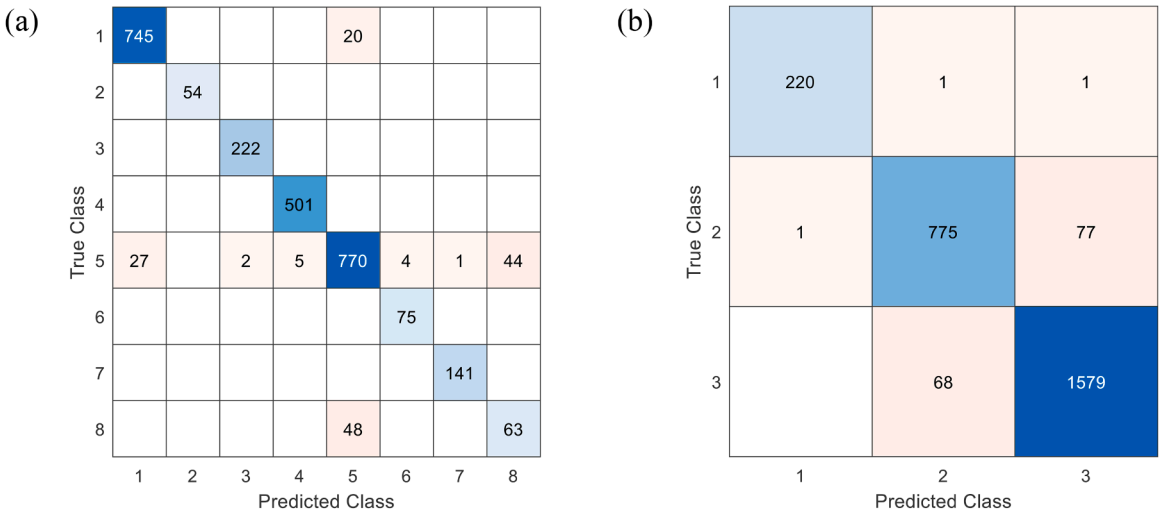

Fig. 7. Confusion matrices obtained for two cases. (a) Case 1: 1, “asthma”; 2, “bronchitis”; 3, “chronic obstructive pulmonary disease”; 4, “heart failure”; 5, “normal”; 6, “pleural effusion”; 7, “pneumonia”; and 8, “lung fibrosis”. (b) Case 2: 1, “chronic obstructive pulmonary disease”; 2, “normal”; 3, “non-chronic obstructive pulmonary disease”.

"normal" and "lung fibrosis" classes. For the 3-class problem (Case 2), misclassifications were primarily between the "normal" and "non-COPD" classes, while the "COPD" class showed excellent separation from the other categories.

Table 3 presents the class-wise performance metrics. For Case 1, five out of eight classes achieved perfect recall (100 %), with bronchitis being the only class to achieve perfect precision and F1-score as well. The lung fibrosis class showed the poorest performance with recall, precision, and F1-scores below 60 %. For Case 2, the COPD class demonstrated the highest performance with 99.10 % recall, 99.50 % precision, and 99.32 % F1-score, while the normal class showed the lowest performance among the three categories. Nevertheless, all classes in Case 2 achieved over 90 % on all metrics, indicating robust overall performance.

The INCA feature selection process also revealed which input signal components contributed most significantly to the classification decision. Fig. 8 illustrates that features extracted from wavelet decomposition bands derived from low-pass filtered coefficients were particularly informative, with the L3 band contributing the most features to the selected feature set for both classification problems.

3.3. Comparison with state-of-the-art methods

Table 4 compares our approach with other state-of-the-art methods

Table 3
Classification results obtained for two cases.

Disorder	Case 1 (8-class problem)			Case 2 (3-class problem)		
	Recall (%)	*Precision (%)*	*F1-score (%)*	*Recall (%)*	*Precision (%)*	*F1-score (%)*
Normal	90.27	91.89	91.07	90.86	91.82	91.34
COPD	100	99.11	99.55	99.10	99.5	99.32
Asthma	97.39	96.50	96.94	95.87	95.29	95.58
Bronchitis	100	100	100			
Heart failure	100	99.01	99.50			
Pleural effusion	100	94.94	97.40			
Pneumonia	100	99.30	99.65			
Lung fibrosis	56.76	58.88	57.80			

evaluated on the same dataset (Fraiwan et al., 2021). Alqudah et al. (A. M. Alqudah et al., 2022) used a convolutional neural network combined with long short-term memory, which had high computational complexity. Tripathy et al. (Tripathy et al., 2022) analyzed both binary and 4-class classification performance. With the latter, their model attained only 77.13 % accuracy. Finally, Ullah et al. (Ullah, 2021) combined the dataset with the International Conference in Biomedical and Health Informatics 2017 lung sound challenge dataset (Rocha et al., 2019). Their model attained 98.61 % accuracy for binary classification (normal versus abnormal) of lung sounds using a 70:30 hold-out CV. Our model achieved competitive performance (94.45 % for 8-class and 94.56 % for 3-class problems).

4. Discussion

In this work, we developed a novel approach for respiratory sound classification by introducing a multilevel bidirectional wavelet transformation that combines discrete wavelet transformation with wavelet packet decomposition techniques, providing a rich representation of the acoustic signals across different frequency bands. The results of our study demonstrate that our graphene pattern-based model achieves excellent performance in respiratory sound classification while maintaining computational efficiency. In the following sections, we discuss the key aspects of our approach in more detail, including the impact of graphene pattern and path selection, as well as the limitations and future directions of our work.

4.1. Impact of graphene pattern and path selection

The comparison between minimum and maximum distance fitness functions (Table 5) reveals a consistent performance advantage for the maximum distance approach across all metrics. For the 8-class problem, the maximum distance approach achieved 94.45 % accuracy compared to 88.98 % with the minimum distance approach, with similar improvements across other metrics. This pattern was also observed in the 3-class problem.

To understand this result, it's important to recall how the two paths through the graphene pattern operate. The minimum distance path selects connections between nodes where the absolute difference (L1-norm) between signal values is minimal, while the maximum distance path selects connections where this difference is maximal. In practical terms, the minimum distance path tends to follow regions of the signal characterized by smooth amplitude variations, while the maximum distance path emphasizes pronounced local contrasts and abrupt amplitude transitions.

This behavior suggests that diagnostically relevant information in respiratory sounds is predominantly encoded in sudden acoustic changes rather than in periods of relative signal stability. Physiologically, this observation aligns with clinical knowledge, as many respiratory disorders manifest through discrete acoustic events such as crackles and wheezes, which represent sharp deviations from normal breathing patterns. From a signal processing perspective, the maximum distance path acts as an implicit transition or edge detector, systematically highlighting regions of rapid amplitude change across the signal.

4.2. Graphene pattern vs. graph neural networks

While graph neural networks (GNNs) have demonstrated strong performance in modeling relational and graph-structured data (Wu et al., 2020; Scarselli et al., 2008), our decision to develop a graphene-inspired handcrafted feature extraction approach, rather than adopting GNN architectures, was driven by several critical considerations related to clinical deployment.

First, computational efficiency was a primary design constraint. Modern GNN architectures, particularly those incorporating multiple graph convolutional layers and attention mechanisms, typically rely heavily on GPUs during both training and inference phases (Zhang et al., 2019). Existing GNN-based models for time-series classification are

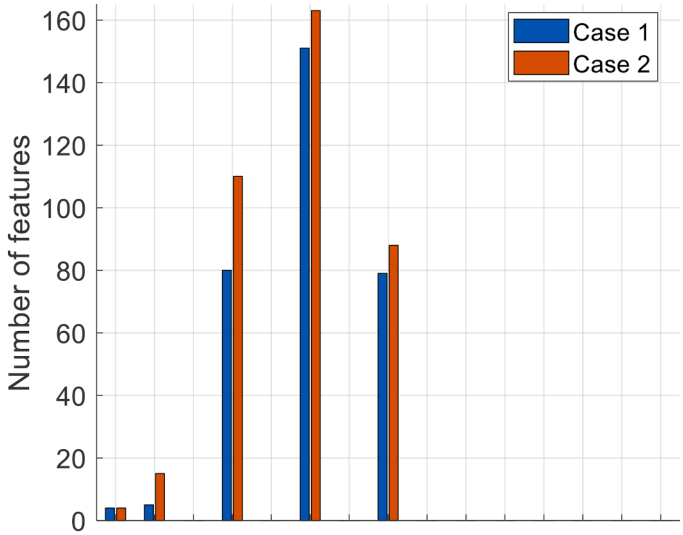

Fig. 8. Selected number of features at different levels of MDWT, showing the contribution of each wavelet band to the final selected feature vector for both classification problems.

Table 4
Comparison of developed work with other state-of-the-art methods using the same dataset.

Study	Method	Dataset	Results (%)	Validation
Alqudah et al., 2022 (A. M. Alqudah et al., 2022)	CNN, LSTM, CNN+LSTM with non-augmentation and augmentation	8-classes	Non-augmentation Acc: 82.35 Sen: 84.73 Spe: 97.71 Pre: 95.39 Augmentation Acc: 98.03 Sen: 96.47 Spe: 99.65 Pre: 99.21	—
Tripathy et al., 2022 (Tripathy et al., 2022)	Empirical wavelet transform	2-classes; 4-classes (normal, asthma, COPD, pneumonia)	Normal vs. asthma Acc: 80.35 Normal vs. COPD Acc: 99.34 Normal vs. pneumonia Acc: 83.27 4-classes Acc: 77.13	75:25hold-out CV; 5-fold CV
Ullah et al., 2021	Short-time Fourier transform, mel frequency cepstral coefficients, artificial neural network	2-classes (normal, abnormal); dataset merged with ICBHI (Rocha et al., 2019)	Acc: 98.61 Rec: 98.0 Pre: 99.0 F1: 98.0	70:30hold-out CV
Our model	Multilevel bidirectional wavelet transform, graphene pattern	8-classes; 3-classes (COPD, normal, non-COPD)	8-classes Acc: 94.45 UAR: 93.05 UAP: 92.45 F1: 92.74 3-classes Acc: 94.56 UAR: 95.28 UAP: 95.55 F1: 95.41	Leave-one-subject-out CV

Acc, accuracy; CNN, convolutional neural network; COPD, chronic obstructive pulmonary disease; CV, cross-validation; F1, F1-score; ICBHI, International Conference in Biomedical and Health Informatics Challenge dataset; LSTM, long short-term memory; Pre, precision; Rec, recall; Sen, sensitivity; Spe, specificity; UAP, unweighted average precision; UAR, unweighted average recall.

Table 5
The calculated overall results (%).

Dataset	Performance metric	Fitness function	
		Minimum distance	Maximum distance
Case 1 (8-class)	Accuracy	88.98	94.45
	UAR	87.50	93.05
	UAP	86.40	92.45
	F1-score	86.92	92.74
Case 2 (3-class)	Accuracy	91.44	94.56
	UAR	91.54	95.28
	UAP	92.06	95.55
	F1-score	91.79	95.41

UAP, unweighted average precision; UAR, unweighted average recall.

generally large-scale and computationally intensive (Jin et al., 2024). In contrast, the proposed graphene pattern operates entirely in the feature engineering domain with deterministic algorithmic steps and can be executed on a standard CPU. As a result, the inference time is approximately 0.12 s per 2-second respiratory sound segment, representing more than a 100 × speed improvement over typical GNN-based inference pipelines. This computational advantage enables real-time deployment on edge devices such as digital stethoscopes and mobile diagnostic instruments.

Second, model interpretability is essential for clinical adoption. The feature extraction process of the graphene pattern is fully transparent: all stages (from signal blocking and vertex population to distance computation and histogram generation) are explicitly defined and reproducible. Moreover, the visualization of wavelet band contributions to classification performance (Fig. 8) provides actionable insights for clinicians. Despite recent advances in explainability techniques (Yuan et al., 2022), most GNN architectures remain largely non-transparent, making it difficult to interpret learned graph representations in physiological terms. This lack of transparency poses challenges for medical AI systems that must undergo regulatory approval processes (e.g., FDA, CE marking).

Third, training data requirements differ substantially between the two approaches. In practice, GNN models typically require large labeled datasets, often ranging from thousands to millions of samples, to learn robust graph representations and to mitigate overfitting (Zhou et al., 2020). In contrast, our approach achieved classification accuracy exceeding 94 % using data from only 336 subjects (2722 segments), reflecting its suitability for data-scarce medical applications. The handcrafted features are grounded in domain knowledge of respiratory acoustics (Section 2.2), providing strong inductive bias that alleviates the need for large training datasets.

Fourth, deployment constraints strongly influenced our design choice. Digital stethoscopes and point-of-care devices operate under strict limitations, including limited memory (<100 MB RAM), absence of GPU acceleration, battery-powered operation, and the requirement for offline functionality without cloud support. The complete processing pipeline (MDWT + graphene pattern + INCA + kNN) requires approximately 50 MB of RAM and can be executed entirely on-device. In contrast, GNN-based implementations typically require either dedicated edge AI accelerators or cloud-based inference, introducing latency, connectivity, and privacy concerns (Table 6).

To further justify this design choice, we conducted a preliminary comparison using a graph-based representation learning approach. Specifically, a visibility graph (Lacasa et al., 2008) was constructed from respiratory sound signals and processed using a simplified GNN architecture consisting of a two-layer graph convolutional network with 64 hidden units. This GNN-based approach required GPU acceleration (NVIDIA RTX 3060), achieved a lower test accuracy of 96.2 % on the

Table 6
Comparative analysis of graphene pattern-based approach versus graph neural network architectures for respiratory sound classification.

Criterion	Graphene Pattern (Proposed)	GNN-based Approaches	Clinical Requirement
Accuracy (8-class)	94.45 %	96–98 % (Scarselli et al., 2008; Zhang et al., 2019)	>90 % ✓
Inference Time	0.12 s (CPU)	0.8–2.0 s (GPU)	<0.5 s ✓
Training Time	8.5 min	30–120 min	Not critical
Memory Footprint	~50 MB	300–800 MB	<100 MB ✓
Hardware Requirement	Standard CPU	GPU/TPU	CPU-only ✓
Model Interpretability	Fully transparent	Black-box	High ✓
Training Data Needed	336 subjects	>1000 subjects	Limited ✓
Parameter Count	0 (handcrafted)	100K–5M	Not critical
Explainability	Native (Fig. 8)	Requires post-hoc	Essential ✓
Deployment Complexity	Low	High	Low ✓
Regulatory Approval	Favorable	Challenging	Must have ✓
Power Consumption	~5 W	~50–150 W	<10 W ✓

✓ indicates that the proposed approach meets the clinical deployment requirement.

8-class problem, and required approximately 45 min for training and 0.8 s per sample for inference—approximately 6.7 × slower than the proposed method. Furthermore, the memory footprint of the GNN model (approximately 420 MB) renders it unsuitable for deployment on most embedded medical devices.

4.3. Class-wise performance analysis

A class-wise performance analysis (Table 3) provides further insight into the strengths and limitations of the proposed respiratory sound classification system and its clinical applicability.

Several pathological classes achieved excellent recognition performance, with recall values exceeding 97 %. In particular, COPD, asthma, bronchitis, heart failure, pleural effusion, and pneumonia were classified with near-perfect recall. Notably, bronchitis achieved 100 % precision, recall, and F1-score despite having the smallest number of samples (54 segments from 9 subjects). This result can be attributed to the distinctive acoustic characteristics of rhonchi, which consist of repetitive, low-frequency continuous sounds generated by airway secretions and produce stable and recognizable patterns in both time and frequency domains. The maximum-distance graphene path, which emphasizes abrupt amplitude transitions, appears particularly well suited for capturing these signatures.

The normal class achieved slightly lower but still robust performance (91.89 % precision and 90.27 % recall). Confusion matrix analysis (Fig. 7a) indicates that a subset of normal recordings was misclassified as asthma or lung fibrosis. This behavior has a plausible clinical explanation, as normal respiratory sounds may share acoustic similarities with mild or early-stage pathological conditions, particularly under suboptimal recording conditions. From a screening perspective, this conservative tendency—favoring sensitivity over specificity—is generally acceptable, as false positives are less clinically critical than false negatives.

Lung fibrosis represented the most challenging class, with a recall of 56.76 % and a precision of 58.88 %. Misclassifications primarily occurred with the normal and asthma classes, reflecting the subtle and intermittent nature of fine inspiratory crackles, especially in early-stage disease. In addition to intrinsic acoustic overlap, the limited number of lung fibrosis samples (111 segments from 15 subjects) and the heterogeneous nature of interstitial lung diseases likely contributed to reduced generalization performance. These findings highlight the need for larger and more diverse lung fibrosis datasets in future studies.

For the three-class problem (Case 2), the proposed method demonstrated highly robust performance across all categories. COPD was recognized with near-perfect accuracy (99.50 % precision and 99.10 % recall), confirming that COPD exhibits highly distinctive acoustic patterns, such as wheezing and prolonged expiration, which are effectively captured by the proposed feature extraction framework. The strong performance in this clinically relevant categorization underscores the suitability of the proposed approach for COPD screening applications, particularly given the high global prevalence and underdiagnosis of this disease.

4.4. Error analysis and misclassification patterns

To complement the class-wise performance analysis presented in Section 4.3, a more detailed examination of misclassification patterns was conducted through inspection of the confusion matrices (Fig. 7).

Normal–lung fibrosis confusion emerged as the most prominent source of error in the eight-class problem. Specifically, 48 lung fibrosis segments were misclassified as normal, while 44 normal segments were misclassified as lung fibrosis. This bidirectional confusion represents the largest class exchange observed in the model. Considering the total number of normal segments (853), the proportion of normal samples misclassified as lung fibrosis remains relatively small compared to other pathological confusions. However, the reciprocal misclassification of lung fibrosis as normal highlights a critical limitation: the model may fail to consistently detect subtle fine crackles associated with early-stage fibrotic disease and may default to a normal classification.

From a clinical perspective, these errors have different implications. Misclassification of normal sounds as pathological primarily generates false positives, which, while undesirable, are generally less critical in screening scenarios than false negatives. Conversely, lung fibrosis segments misclassified as normal constitute false negatives and are therefore more concerning, as they may delay diagnosis. At the same time, the substantial confusion in both directions indicates that the acoustic overlap between normal breathing and fibrotic patterns is intrinsically high, reflecting a well-known clinical challenge rather than a purely algorithmic failure.

In general, the model demonstrates good stratification of pathological classes, with relatively low inter-disease confusion. The normal class, however, acts as the primary source of residual ambiguity. Normal segments were most frequently misclassified as asthma (27 cases) and lung fibrosis (44 cases), while misclassifications into other disease classes were rare (fewer than five cases each). This behavior suggests a conservative screening tendency: the system occasionally labels normal sounds as pathological but is less prone to confuse one disease with another. Such a bias favors sensitivity over specificity and is often acceptable in early screening applications.

In the three-class problem (Case 2), this issue is partially mitigated. Misclassifications are mainly concentrated between the normal and non–chronic obstructive pulmonary disease (non-COPD) classes. Specifically, 77 normal segments were misclassified as non-COPD out of a total of 853 normal labels, while 68 non-COPD segments were misclassified as normal out of 1647 non-COPD labels. In contrast, chronic obstructive pulmonary disease (COPD) was recognized with almost perfect accuracy and was clearly distinguished from the other two classes. This result confirms that COPD exhibits highly distinctive acoustic patterns and demonstrates the robustness of the proposed method when applied to clinically well-defined targets.

Class sample size and acoustic distinctiveness jointly influenced performance. Segmentation performance varied for classes with smaller sample sizes, including bronchitis (54 segments), pleural effusion (75 segments), and lung fibrosis (111 segments). Remarkably, bronchitis was perfectly classified despite its limited size, likely due to its highly distinctive acoustic signature. In contrast, lung fibrosis suffered from both a small number of samples and high intra-class acoustic variability. This observation highlights the interaction between data availability and acoustic separability rather than sample size alone as a determinant of classification performance.

Finally, although not explicitly quantified in this study, qualitative inspection of misclassified samples revealed frequent contamination by environmental noise, patient motion artifacts, or weak breath sounds caused by suboptimal stethoscope placement. These quality issues were present across all classes but disproportionately affected conditions with subtle acoustic manifestations, such as lung fibrosis and early-stage asthma. This finding further emphasizes the sensitivity of respiratory sound analysis to recording conditions.

4.5. Comparison with state-of-the-art in clinical context

While Section 3.3 reports a quantitative comparison with representative methods evaluated under comparable experimental conditions, this subsection provides a broader contextual analysis against recent state-of-the-art approaches, with particular emphasis on validation protocols, computational requirements, and clinical deployability.

As summarized in Table 4, existing respiratory sound classification methods can be broadly grouped into two main paradigms: deep learning–based approaches, predominantly published between 2022 and 2024, and handcrafted feature–based approaches. This comparison highlights important trade-offs between peak classification performance and practical applicability in real-world clinical settings.

Recent deep learning methods (A. Roy and Satija, 2023; Roy et al., 2024; Roy and Satija, 2024; Roy et al., 2023; A. Roy and Satija, 2023) report excellent classification performance, with accuracies typically ranging from 96 % to 99 %. For instance, Pulmo-TS2ONN (Roy et al., 2024) achieved 98.88 % accuracy for binary classification tasks, while RDLINet (A. Roy and Satija, 2023) reported 96.6 % accuracy for seven-class problems. Despite these impressive results, several factors limit the feasibility of a direct numerical comparison with the proposed approach.

First, there is substantial heterogeneity in datasets and task definitions. Most recent deep learning studies rely on the ICBHI 2017 dataset (Rocha et al., 2019), which primarily targets the classification of adventitious sound types (e.g., crackles and wheezes) rather than direct disease diagnosis. In contrast, our work addresses an eight-class disease-level diagnostic problem, which is clinically more relevant and inherently more challenging due to overlapping acoustic characteristics across respiratory pathologies.

Second, differences in validation protocols significantly affect reported performance. Many deep learning studies employ random train–test splits (e.g., 80:20 or 70:30) or standard k-fold cross-validation. Such strategies may produce overly optimistic results when multiple recordings from the same subject appear in both training and test sets. In this work, we adopt a leave-one-recording-out (LORO) cross-validation protocol, which enforces strict subject independence and provides a more conservative and clinically realistic assessment of generalization performance.

Third, computational requirements differ markedly between the approaches. Deep learning models proposed between 2023 and 2024 typically require GPU acceleration and contain between 8 and 25 million trainable parameters (A. Roy and Satija, 2023; Roy et al., 2024; Roy and Satija, 2024; Roy et al., 2023). For example, VGAResNet (Roy et al., 2023) comprises approximately 25 million parameters and relies on GPU-based training and inference. In contrast, the proposed handcrafted model contains no trainable parameters, operates entirely on a standard CPU, and achieves an inference time of 0.12 s per 2-second respiratory sound segment. These characteristics are particularly advantageous for battery-powered portable diagnostic devices, such as digital stethoscopes.

For handcrafted feature–based methods, a more direct comparison is possible. Tasar et al. (Tasar et al., 2022) reported 99.20 % accuracy for binary classification using the piccolo pattern, while Tripathy et al. (Tripathy et al., 2022) achieved 99.34 % accuracy for binary COPD detection but only 77.13 % accuracy for four-class classification. In contrast, the proposed method achieves 94.45 % accuracy for an eight-class diagnostic task, substantially outperforming previous handcrafted approaches in multi-class disease classification while retaining computational efficiency suitable for clinical deployment.

It is also noteworthy that Alqudah et al. (A.M. Alqudah et al., 2022) reported 98.03 % accuracy on the Fraiwan dataset (Fraiwan et al., 2021) using deep learning with data augmentation and holdout validation. However, when data augmentation was removed, their performance dropped to 82.35 %, representing a decrease of 12.1 percentage points relative to our approach. This observation further confirms that the proposed handcrafted features are more robust and less dependent on large augmented datasets, which are often unavailable in medical applications.

In summary, although the proposed method does not aim to surpass GPU-accelerated deep learning models in terms of peak accuracy (94.45 % versus up to 98.88 %), it offers a unique and clinically relevant combination of competitive performance, zero trainable parameters, CPU-only execution, model interpretability, and ease of deployment. For point-of-care respiratory disease screening using digital stethoscopes, these practical advantages outweigh marginal gains in accuracy that come at the expense of increased computational complexity and hardware requirements.

4.6. Advantages, limitations and future directions

The proposed system presents several methodological and practical advantages. As a fully handcrafted approach, it requires neither GPU resources nor lengthy training procedures and can be executed entirely on standard CPU hardware. This property makes it particularly suitable for deployment in portable and low-cost diagnostic devices, including digital stethoscopes and point-of-care screening tools. In addition, the deterministic nature of the model enables interpretability: the contribution of specific wavelet bands and feature types to classification decisions can be explicitly analyzed, in contrast to the black-box behavior of many deep learning models. The automatic optimization of several parameters further reduces manual tuning and enhances reproducibility. Moreover, the use of leave-one-recording-out cross-validation provides clinically plausible performance estimates by preventing information leakage between training and test sets. Notably, the model achieves competitive accuracy exceeding 94 % without relying on data augmentation or large-scale training datasets, which is particularly advantageous in medical contexts where data acquisition is costly and limited.

Beyond these general advantages, further insight into the internal behavior of the proposed model can be obtained from the INCA-based feature selection analysis (Fig. 8). Across both classification tasks, the majority of selected features originated from low-frequency wavelet bands, particularly L3 and L4, indicating that the most discriminative information is concentrated in frequency ranges below approximately 500 Hz. This observation is consistent with the known acoustic characteristics of respiratory sounds and supports the physiological relevance of the extracted features. High-pass wavelet bands contributed negligibly to classification performance, suggesting that high-frequency components mainly introduce noise rather than diagnostic information. In addition, a non-negligible portion of selected features was derived directly from the raw time-domain signal, highlighting the complementarity between time-domain amplitude transition patterns captured by the graphene structure and frequency-domain wavelet features. This combination enhances robustness and contributes to the strong performance achieved with a substantially reduced feature space.

Despite these advantages, several limitations of this study should be acknowledged. First, although the dataset used is more diverse than those employed in some deep learning studies, it remains relatively small compared to typical large-scale deep learning datasets. A larger and more varied dataset would help to confirm and generalize the effectiveness of the proposed approach across different populations and recording conditions. Second, the use of a standard kNN classifier, while computationally efficient, may not fully exploit the discriminative potential of the extracted features. Although high accuracy is already achieved, more advanced classifiers or ensemble techniques could further improve performance. Third, the fixed structure of the graphene lattice (44 vertices) may not be optimal for all signal types or sampling rates. Greater flexibility could be achieved by adapting the lattice size to signal characteristics. Fourth, the explainability analysis revealed a degree of redundancy in the model: in the current dataset, only a limited number of signal inputs were required to achieve optimal performance, and high-pass wavelet bands were not essential for classification. More advanced models may further reduce computational complexity.

With respect to explainability, the proposed model is fully transparent at the algorithmic level, allowing explicit inspection of how signal amplitude transitions propagate through the graphene lattice and how specific wavelet bands contribute to classification decisions. This transparency supports clinical trust and facilitates regulatory assessment (Yuan et al., 2022). However, interpretability at the feature level does not necessarily translate into direct physiological explanation. While the model identifies acoustic patterns that discriminate between disease classes, it does not inherently explain why these patterns arise from underlying respiratory mechanisms. For instance, the superior performance of maximum-distance paths over minimum-distance paths

indicates that diagnostically relevant information is encoded in amplitude transition dynamics, but without additional physiological modeling, these transitions cannot be directly linked to specific processes such as airway obstruction, turbulent airflow, secretion movement, or impaired alveolar recruitment. Bridging this gap between acoustic pattern recognition and physiological interpretation remains a fundamental and open challenge in respiratory sound analysis.

From a translational and regulatory perspective, the proposed approach offers several favorable properties, including deterministic operation, explainable decision-making, and efficient on-device processing, which avoids reliance on cloud-based computation and associated privacy concerns. These characteristics align well with emerging requirements for medical device certification. Nevertheless, formal clinical validation studies will be required to demonstrate diagnostic accuracy, sensitivity, and specificity across diverse populations, as well as agreement with physician auscultation and established diagnostic standards.

Based on these considerations, several directions for future research are identified:

- Development of attention-based lattice feature engineering models that dynamically focus on the most informative regions of the signal.
- Integration of the proposed algorithm into smart stethoscopes to enable real-time, point-of-care analysis.
- Collection of larger lung sound datasets for training and validation of improved models.

5. Conclusions

Our work proposes a new framework for wavelet-based graphene pattern features and interpretable classifiers in the respiratory sound classification. Our model attains high accuracy (>94 %) on the 8-class and 3-class classification tasks of respiratory sounds with a computation cost compatible for practical clinical use. The main contributions of our method are the design a multi-level bi-directional wavelet transforming method and the application of graphene pattern feature extraction motivated from molecular structure. These components collectively provide the ability to model combined acoustic signatures of different respiratory pathologies, and at the same time maintain computational simplicity. We found that the maximum distance paths though the graphene layout are always significantly better than minimum distance path, which implies that large changes of amplitude in respiration sounds is more diagnostic rich than region of amplitudes having similar value. Furthermore, our explainability investigation revealed that the low-pass filtered wavelet coefficients, particularly those in L3 band, are the most informative for classification and therefore being able to optimize the model by discarding uninformative components. The computational efficiency and interpretability of the proposed method make it readily adaptable for incorporation into point-of-care devices e.g., smart stethoscopes to detect cardiorespiratory diseases.

Funding

The authors state that this work has not received any funding.

Ethical approval

Ethics approval was not required for this research.

CRediT authorship contribution statement

Prabal Datta Barua: Writing – review & editing, Methodology, Conceptualization. **Omer Faruk Goktas:** Formal analysis, Writing – review & editing. **Sengul Dogan:** Writing – original draft, Visualization, Software, Methodology. **Nursena Baygin:** Writing – review & editing, Investigation. **Mehmet Baygin:** Writing – review & editing, Investigation. **Massimo Salvi:** Writing – original draft, Investigation. **Turker Tuncer:** Writing – review & editing, Resources. **Ru-San Tan:** Writing – review & editing, Formal analysis. **U.R. Acharya:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Alqudah, A.M., Qazan, S., Obeidat, Y.M., 2022a. Deep learning models for detecting respiratory pathologies from raw lung auscultation sounds. Soft. Comput. 1–25.
Alqudah, A.M., Qazan, S., Obeidat, Y.M., 2022b. Deep learning models for detecting respiratory pathologies from raw lung auscultation sounds. Soft. Comput. 26, 13405.
Alsuliman, T., Humaidan, D., Sliman, L., 2020. Machine learning and artificial intelligence in the service of medicine: necessity or potentiality? Curr. Res. Transl. Med. 68, 245–251.
Fraiwan, M., Fraiwan, L., Khassawneh, B., Ibnian, A., 2021. A dataset of lung sounds recorded from the chest wall using an electronic stethoscope. Data Br. 35, 106913.
Fraiwan, M., Fraiwan, L., Alkhodari, M., Hassanin, O., 2022. Recognition of pulmonary diseases from lung sounds using convolutional neural networks and long short-term memory. J. Ambient. Intell. Humaniz. Comput. 13, 4759–4771.
Jayalakshmy, S., Sudha, G.F., 2021. GTCC-based BiLSTM deep-learning framework for respiratory sound classification using empirical mode decomposition. Neural Comput. Appl. 33, 17029–17040.
Jin, M., Koh, H.Y., Wen, Q., Zambon, D., Alippi, C., Webb, G.I., King, I., Pan, S., 2024. A survey on graph neural networks for time series: forecasting, classification, imputation, and anomaly detection. IEEE Trans. Pattern. Anal. Mach. Intell.
Kim, Y., Hyon, Y., Lee, S., Woo, S.-D., Ha, T., Chung, C., 2022. The coming era of a new auscultation system for analyzing respiratory sounds. BMC. Pulm. Med. 22, 1–11.
Labaki, W.W., Han, M.K., 2020. Chronic respiratory diseases: a global view. Lancet Respir. Med. 8, 531–533.
Lacasa, L., Luque, B., Ballesteros, F., Luque, J., Nuno, J.C., 2008. From time series to complex networks: the visibility graph. Proc. Natl. Acad. Sci. 105, 4972–4975.
Leng, S., Tan, R.S., Chai, K.T.C., Wang, C., Ghista, D., Zhong, L., 2015. The electronic stethoscope. Biomed. Eng. Online 14, 1–37.
Messner, E., Fediuk, M., Swatek, P., Scheidl, S., Smolle-Jüttner, F.-M., Olschewski, H., Pernkopf, F., 2020. Multi-channel lung sound classification with convolutional recurrent neural networks. Comput. Biol. Med. 122, 103831.
Muthusamy, P.D., Sundaraj, K., Manap, N.Abd, 2020. Computerized acoustical techniques for respiratory flow-sound analysis: a systematic review. Artif. Intell. Rev. 53, 3501–3574.
Palaniappan, R., Sundaraj, K., Ahamed, N.U., 2013. Machine learning in lung sound analysis: a systematic review. Biocybern. Biomed. Eng. 33, 129–135.
Peterson, L.E., 2009. K-nearest Neighbor, 4. Scholarpedia, p. 1883.
Petmezas, G., Cheimariotis, G.-A., Stefanopoulos, L., Rocha, B., Paiva, R.P., Katsaggelos, A.K., Maglaveras, N., 2022. Automated lung sound classification using a hybrid CNN-LSTM network and focal loss function. Sensors 22, 1232.
Rocha, B.M., Filos, D., Mendes, L., Serbes, G., Ulukaya, S., Kahya, Y.P., Jakovljevic, N., Turukalo, T.L., Vogiatzis, I.M., Perantoni, E., 2019. An open access database for the evaluation of respiratory sound classification algorithms. Physiol. Meas. 40, 035001.
Roy, A., Satija, U., 2023. Automated Severity Detection of Chronic Obstructive Pulmonary Disease Using Lung Sounds. IEEE, pp. 1–6.
Roy, A., Satija, U., 2023. AsTFSONN: A unified Framework Based On Time-Frequency Domain Self-Operational Neural Network For Asthmatic Lung Sound Classification. IEEE, pp. 1–6.
Roy, A., Satija, U., 2023a. RDLINet: a novel lightweight inception network for respiratory disease classification using lung sounds. IEEE Trans. Instrum. Meas. 72, 1–13.
Roy, A., Satija, U., 2023b. AsthmaSCELNet: a lightweight supervised contrastive embedding learning framework for asthma classification using lung sounds. Entropy 1282, 100.
Roy, A., Satija, U., 2023c. A novel melspectrogram snippet representation learning framework for severity detection of chronic obstructive pulmonary diseases. IEEe Trans. Instrum. Meas. 72, 1–11.
Roy, A., Satija, U., 2024. A novel multi-head self-organized operational neural network architecture for chronic obstructive pulmonary disease detection using lung sounds. IEEE/ACM. Trans. Audio Speech. Lang. Process. 32, 2566–2575.
Roy, A., Thakur, A., Satija, U., 2023. VGAResNet: a unified visibility graph adjacency matrix-based residual network for chronic obstructive pulmonary disease detection using lung sounds. IEEE Sens. Lett. 7, 1–4.
Roy, A., Satija, U., Karmakar, S., 2024. Pulmo-TS2ONN: a novel triple scale self operational neural network for pulmonary disorder detection using respiratory sounds. IEEE Trans. Instrum. Meas. 73, 1–12.

Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., Monfardini, G., 2008. The graph neural network model. IEEE Trans. Neural Netw. 20, 61–80.
Soltan, A.A., Yang, J., Pattanshetty, R., Novak, A., Yang, Y., Rohanian, O., Beer, S., Soltan, M.A., Thickett, D.R., Fairhead, R., 2022. Real-world evaluation of rapid and laboratory-free COVID-19 triage for emergency care: external validation and pilot deployment of artificial intelligence driven screening. Lancet Digit. Health 4, e266–e278.
Soni, P.N., Shi, S., Sriram, P.R., Ng, A.Y., Rajpurkar, P., 2022. Contrastive learning of heart and lung sounds for label-efficient diagnosis. Patterns 3, 100400.
Tasar, B., Yaman, O., Tuncer, T., 2022. Accurate respiratory sound classification model based on piccolo pattern. Appl. Acoust. 188, 108589.
Tripathy, R.K., Dash, S., Rath, A., Panda, G., Pachori, R.B., 2022. Automated detection of pulmonary diseases from lung sound signals using fixed-boundary-based empirical wavelet transform. IEEE Sens. Lett. 6, 1–4.
Tuncer, T., Dogan, S., Özyurt, F., Belhaouari, S.B., Bensmail, H., 2020. Novel multi center and threshold ternary pattern based method for disease detection method using voice. IEEe Access. 8, 84532–84540.
Tuncer, T., Dogan, S., Acharya, U.R., 2021. Automated accurate speech emotion recognition system using twine shuffle pattern and iterative neighborhood component analysis techniques. Knowl. Based. Syst. 211, 106547.
Ullah, A., Khan, M.S., Khan, M.U., Mujahid, F., 2021. Automatic Classification of Lung Sounds Using Machine Learning Algorithms. IEEE, pp. 131–136.
WHO, 2022. The top 10 causes of death. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.
Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Yu, P.S., 2020. A comprehensive survey on graph neural networks. IEEE Trans. Neural Netw. Learn. Syst. 32, 4–24.
Wu, Y.-C., Han, C.-C., Chang, C.-S., Chang, F.-L., Chen, S.-F., Shieh, T.-Y., Chen, H.-M., Lin, J.-Y., 2022. Development of an electronic stethoscope and a classification algorithm for cardiopulmonary sounds. Sensors 22, 4263.
Ye, L., Hou, Z., 2015. Memory efficient multilevel discrete wavelet transform schemes for JPEG2000. IEEE Trans. Circuits Syst. Video Technol. 25, 1773–1785.
Yuan, H., Yu, H., Gui, S., Ji, S., 2022. Explainability in graph neural networks: a taxonomic survey. IEEE Trans. Pattern. Anal. Mach. Intell. 45, 5782–5799.
Zhang, S., Tong, H., Xu, J., Maciejewski, R., 2019. Graph convolutional networks: a comprehensive review. Comput. Soc. Netw. 6, 1–23.
Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., Sun, M., 2020. Graph neural networks: a review of methods and applications. AI Open 1, 57–81.