Automated Lightweight Model for Asthma Detection Using Respiratory and Cough Sound Signals

Shuting Xu [1,2] , Ravinesh C. Deo [1] , Oliver Faust [3,*] , Prabal D. Barua [1,2] , Jeffrey Soar [4]
and Rajendra Acharya [1]

[1] Artificial Intelligence Applications Laboratory, School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, QLD 4300, Australia; ravinesh.deo@unisq.edu.au (R.C.D.); prabal.barua@unisq.edu.au (P.D.B.); rajendra.acharya@unisq.edu.au (R.A.)
[2] Cogninet Australia, Sydney, NSW 2010, Australia
[3] School of Computing and Information Science Research, Anglia Ruskin University Cambridge Campus, East Rd, Cambridge CB1 1PT, UK
[4] School of Business, University of Southern Queensland, Springfield, QLD 4350, Australia; jeffrey.soar@unisq.edu.au
[*] Correspondence: oliver.faust@aru.ac.uk

Abstract: Background and objective: Chronic respiratory diseases, such as asthma and COPD, pose significant challenges to human health and global healthcare systems. This pioneering study utilises AI analysis and modelling of cough and respiratory sound signals to classify and differentiate between asthma, COPD, and healthy subjects. The aim is to develop an AI-based diagnostic system capable of accurately distinguishing these conditions, thereby enhancing early detection and clinical management. Our study, therefore, presents the first AI system that leverages dual acoustic signals to enhance the diagnostic ACC of asthma using automated, lightweight deep learning models. **Methods**: To build an automated, lightweight model for asthma detection, tested separately with respiratory and cough sounds to assess their suitability for detecting asthma and COPD, the proposed AI models integrate the following ML algorithms: RF, SVM, DT, NN, and KNN, with an overall aim to demonstrate the efficacy of the proposed method for future clinical use. Model training and validation were performed using 5-fold cross-validation, wherein the dataset was randomly divided into five folds and the models were trained and tested iteratively to ensure robust performance. We evaluated the model outcomes with several performance metrics: ACC, precision, recall, F1 score, and area under the AUC. Additionally, a majority voting ensemble technique was employed to aggregate the predictions of the various classifiers for improved diagnostic reliability. We applied Gabor time–frequency transformation for feature extraction and NCA) for feature selection to optimise predictive accuracy. Independent comparative experiments were conducted, where cough-sound subsets were used to evaluate asthma detection capabilities, and respiratory-sound subsets were used to evaluate COPD detection capabilities, allowing for targeted model assessment. **Results**: The proposed ensemble approach, facilitated by a majority voting approach for model efficacy evaluation, achieved acceptable ACC values of 94.05% and 83.31% for differentiating between asthma and normal cases utilising separate respiratory sounds and cough sounds, respectively. The results highlight a substantial benefit in integrating multiple classifier models and sound modalities while demonstrating an unprecedented level of ACC and robustness for future diagnostic predictions of the disease. **Conclusions**: The present study sets a new benchmark in AI-based detection of respiratory diseases by integrating cough and respiratory sound signals for future diagnostics. The successful implementation of a dual-sound analysis approach promises advancements in the early detection and management of asthma and COPD. We conclude that the proposed model

holds strong potential to transform asthma diagnostic practices and support clinicians in their respiratory healthcare practices.

1. Introduction

Asthma affects approximately 339 million people worldwide, while COPD is projected to become the third leading cause of death globally by 2030 [1]. These statistics underscore the urgent need for effective disease management strategies. The economic burden of these diseases is significant, with costs associated with medical treatment, loss of productivity, and premature death. Asthma and COPD are linked to both genetic predispositions and environmental factors. For instance, asthma is often associated with allergen exposure, which leads to airway hyper-responsiveness [2]. This condition is further exacerbated by factors such as pollution, respiratory infections, and occupational exposures. On the other hand, COPD is commonly related to long-term smoking and environmental pollutants [3]. It is characterised by airflow obstruction and is often accompanied by chronic bronchitis and emphysema. Common symptoms include shortness of breath, wheezing, chest tightness, and coughing. COPD is characterised by progressive breathlessness and chronic cough with sputum, while asthma is marked by episodic symptoms triggered by specific factors such as allergens, exercise, or cold air. If unmanaged, these diseases can lead to severe complications, including respiratory failure, pulmonary hypertension, heart complications, and death [4]. The morbidity and mortality associated with these diseases necessitate advancements in diagnostic and therapeutic approaches.

Recent advancements have shown promise in using AI for the automated detection of asthma and COPD through the analysis of cough and respiratory sound signals [5,6]. AI and ML techniques offer a novel approach to diagnosing these conditions with high ACC. Previous studies have demonstrated high accuracy rates using DL models, particularly CNN. These models have achieved diagnostic ACC rates as high as 92% when distinguishing between healthy individuals and those with asthma [1]. These models have shown high sensitivity (up to 89%) and specificity (up to 90%), with an area under the curve (AUC) exceeding 95%, indicating superior performance and reliability in clinical diagnostics compared to traditional machine learning ML models such as support vector machine (SVM) and random forest (RF) [7].

Despite these advancements, gaps remain in the utilisation of lightweight AI models for the analysis of small datasets. While DL techniques have achieved good performance, their computational complexity and reliance on extensive training and testing data make them unsuitable for exploratory studies or resource-constrained environments. Classical ML approaches, which incorporate algorithmic feature engineering to extract information from the signals, often limit the information available to the classifier. Hence, the performance of such systems was suboptimal, especially when they were presented with unseen data in practical inference scenarios, like detecting asthma from respiratory sounds. Lightweight models offer an advantage in environments with restricted processing capabilities. One such environment is asthma detection at the edge, where real-time decision support is provided close to the signal source [8].
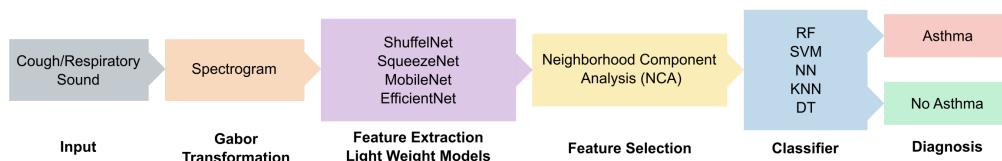
This study aims to address these research gaps by developing and validating an automated, lightweight AI model that has the ability to enhance diagnostic precision, particularly for asthma and COPD detection using cough and respiratory sound analysis.

Our approach integrates ML algorithms, including RF, SVM, DT, NN, and KNN, as well as feature extraction techniques that are optimised by Gabor transformation and NCA. The proposed model employs a majority voting system to refine its diagnostic reliability. Overall, the results demonstrate that the ensemble approach can achieve a high ACC in differentiating between asthma and COPD. Hence, the study highlights the benefits of integrating multiple classifier models and sound modalities. Furthermore, it sets a new benchmark for the AI-based detection of respiratory diseases. This can lead to significant advancements in the early detection and management of asthma and COPD. By combining cough and respiratory sounds for diagnostic purposes, this study also has the potential to improve diagnostic practices in respiratory healthcare.

2. Materials and Methods

This section introduces materials and methods used to create a hybrid model to discriminate between asthma and normal sound signals. The workflow, illustrated in Figure 1, outlines a streamlined process for detecting asthma from audio datasets using a hybrid model. Further details are provided with the pseudocode reproduced in Appendix A. The workflow comprises the following steps:

- **Audio input**: The process begins by collecting audio data that capture cough or respiratory sounds. The data are input to the proposed system.
- **Time–frequency transformation**: Utilises STFT to generate spectrograms, facilitating time-resolved frequency analysis of audio signals [9].
- **Lightweight DL feature extraction**: Uses the most efficient CNN architectures such as ShuffleNet, SqueezeNet, MobileNet, and EfficientNet to extract the most relevant features from the spectrograms.
- **Feature selection**: Selects the most informative features to reduce the dimensionality of the data and, therefore, focuses on the classification of the best features in the relevant data [10]. ML classification: Employs classical ML algorithms (i.e., DT, SVM, NN, RF, and KNN) to classify the audio data and detect the patterns that are indicative of asthma [11].

Figure 1. Block diagram of the proposed asthma detection system using respiratory and cough sound signals.

2.1. Dataset

For this specific study, we have used a dedicated private dataset from Firat University hospital (Turkey) provided by the project industry partner, Cogninet Australia Pty Ltd, Level 5/29-37 Bellevue St, Surry Hills NSW 2010. The primary dataset includes both cough and respiratory sounds for asthma and COPD, respectively, as well as a healthy comparison group. Table 1 provides the details of these dataset.

The dataset consists of respiratory and cough sounds collected from 1326 individuals, including 511 asthmatic and 815 healthy participants. The asthmatic group comprises 380 females and 131 males, with an age range of 19 to 75 years, while the healthy group includes 306 females and 509 males, aged 19 to 80 years. Most healthy participants were university students without underlying health conditions, whereas heart disease was the most commonly observed comorbidity among asthmatic participants.

Table 1. Characteristics of cough and respiratory sound dataset used to develop an automated lightweight model for asthma detection.

Characteristics	Details
Duration	Duration of each audio file used to determine the range, the average, and the distribution of data length
File size	Assessing the file size to determine the dataset storage requirements
Sampling rate	The number of snapshots taken to recreate the original sound wave
Bit depth	The number of amplitude values in each snapshot to determine the audio resolution with the sampling rate

The dataset includes 1875 cough recordings, split into 994 and 881 segments after preprocessing, where non-cough segments, speech, and background noise were removed to ensure consistency. The duration of cough recordings ranged from 0.23 to 6.59 s and was standardised into equal-length segments for model training.

Table 2 shows significant differences in the average duration of cough sounds (≈1–2 s) and respiratory sounds (≈13–30 s). It is envisaged that a shorter cough sound may be focused on capturing the intensity and the quality of the cough, while a longer respiratory sound could provide a more comprehensive representation of the breathing patterns.

Table 2. Statistical analysis of the audio data used to build an automated lightweight model for asthma detection.

Metrics	Cough-Cough	Cough-Control	Resp-Asthma	Resp-Control
Avg duration (s)	1.77	1.60	17.07	29.93
Avg size (kB)	174	162	136	239
Sampling rate (Hz)	48,000	48,000	4000	4000
Bit depth	16	16	16	16

As for the resampling rate, the cough sound files have a high sampling rate of 48,000 Hz, which is generally typical of high-quality audio and has the capability to capture a wide range of frequencies. By contrast, the respiratory sounds are sampled at 4000 Hz, which is much lower and may therefore miss finer details in the recorded audio [12]. This could be representative of different recording equipment or purposes, with cough sounds potentially requiring higher fidelity for accurate analysis.
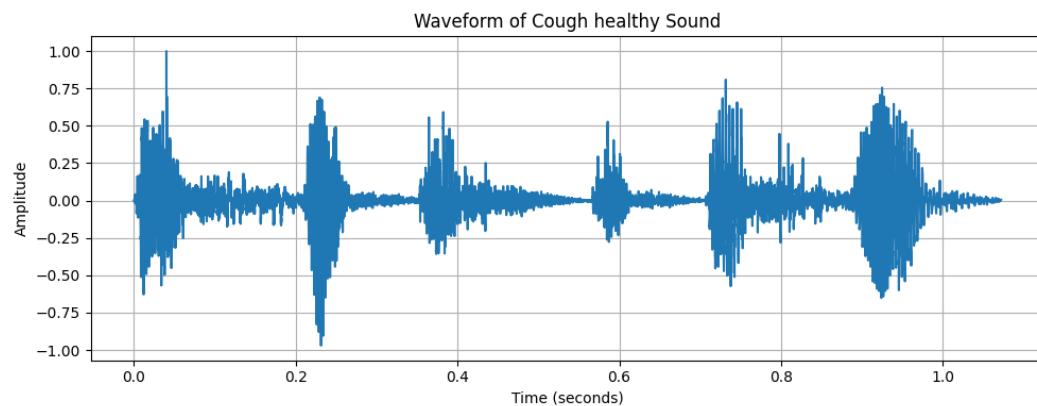
2.2. Data Analysis

In this study, we have adopted waveform processing as a fundamental analysis tool for audio files. In particular, waveform analysis provides details of the visual representation of the audio signal's amplitude over the passage of time. It may also help in identifying the loudness, the duration of the sound segments, and the presence of silence and/or noise. This is particularly useful in understanding the dynamics of the audio, such as the variations in sound intensity and temporal patterns, which can represent specific disease features for different asthma patients [13].
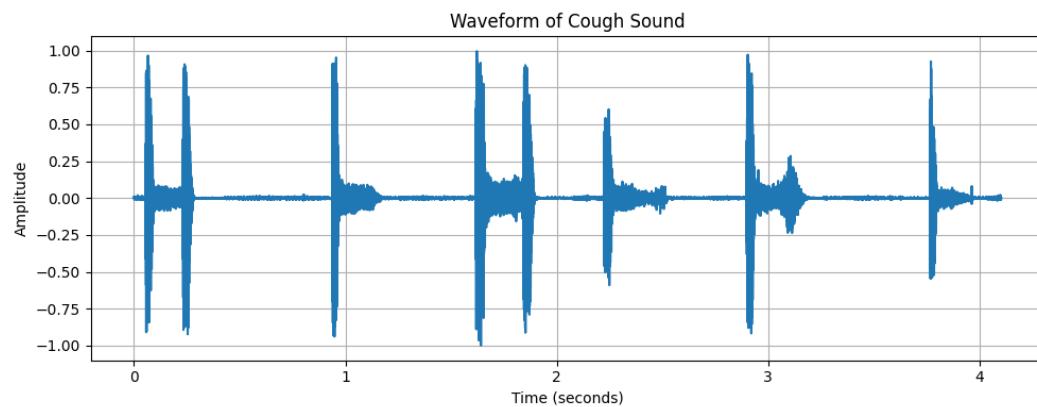
To enhance interpretability, the waveform plots have been updated to present the *x*-axis in seconds instead of sample points. This change provides clearer insight into the temporal characteristics of the sound signals.

The dynamic range of Figure 2, with a maximum amplitude of +32,767 and minimum amplitude of 24,217, reflects the varying intensity of breathing sounds. The peaks in the waveform represent louder respiratory phases, likely corresponding to inhaling or exhaling. However, quieter segments could indicate softer breathing or pauses between breaths. Figures 3–5 depict the waveform analysis of sample audio from the respective dataset. In
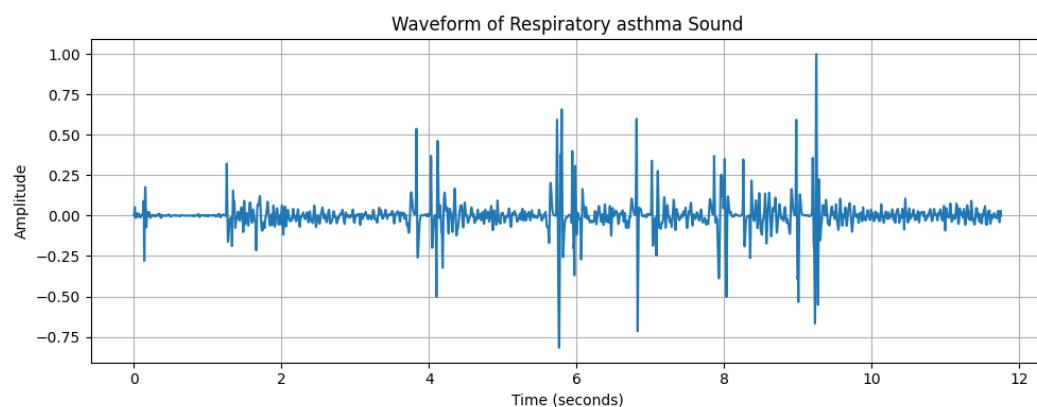
summary, the audio recording exhibits a broad dynamic range and a low-frequency profile characteristic of respiratory sounds. The variation in amplitude may reflect different phases of breathing, while the low-frequency dominance is typical of such sounds. This type of analysis can be valuable in medical contexts, including monitoring respiratory health, diagnosing conditions, and assessing treatment efficacy [14].
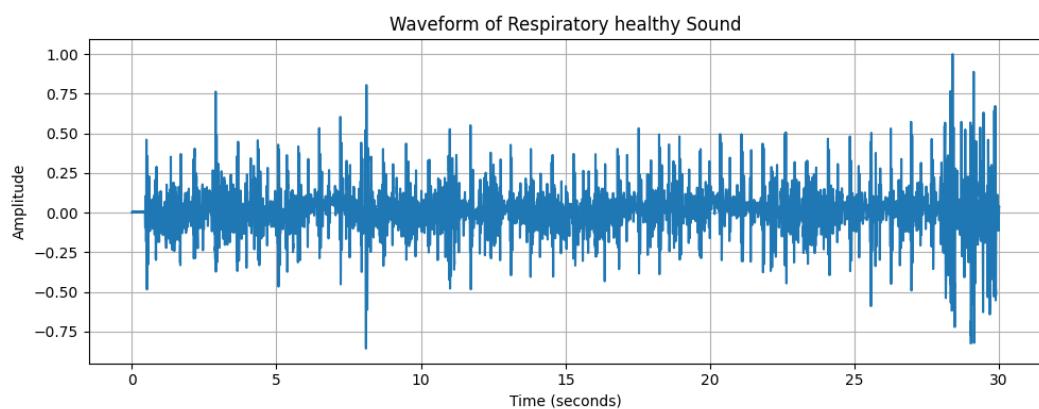
Figure 2. Waveform of a cough sound from an asthma subject, with the *x*-axis representing time (s) and the *y*-axis representing amplitude.

Figure 3. Waveform of a cough sound from a healthy subject, with the *x*-axis representing time (s) and the *y*-axis representing amplitude.

Figure 4. Waveform of a respiratory sound from an asthmatic subject, with the *x*-axis representing time (s) and the *y*-axis representing amplitude.

Figure 5. Waveform of a respiratory sound from a healthy subject, with the *x*-axis representing time (s) and the *y*-axis representing amplitude.

### 2.3. Signal Transformation Approaches

To develop an automated, lightweight model for asthma detection with respiratory and cough sounds, the audio is passed through a time–frequency transformation process utilising a spectrogram as a visual way to represent the audio's frequency information over the passage of time. In general, a spectrogram is a visual representation of the frequency spectrum of a sound signal, depicting three dimensions—time, frequency, and amplitude. It is expected that patients *with* and *without* asthma are likely to demonstrate different spectrum features within the prescribed respiratory and cough sounds and, therefore, a spectrum analysis can enable the model to determine its efficacy based on any of the two specific data inputs.
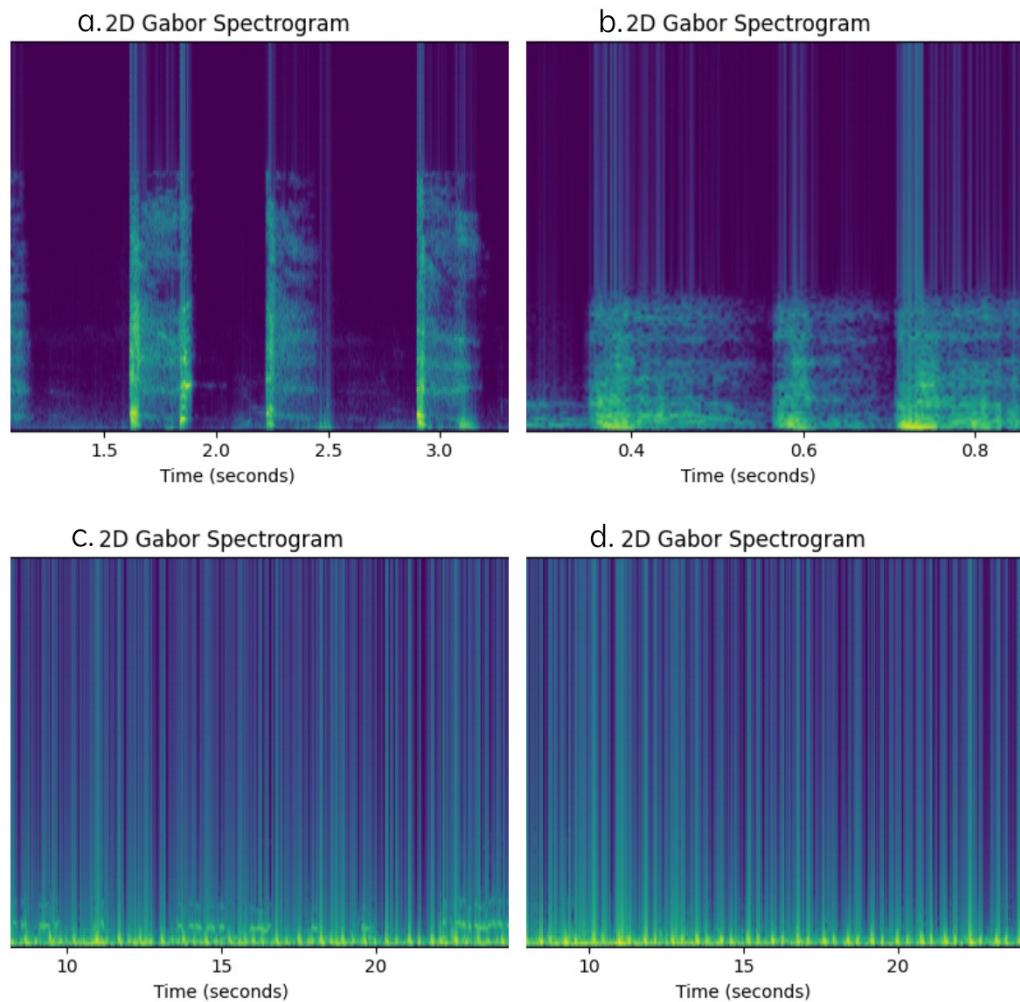
Figure 6 shows the 2D spectrogram of a cough sound using the Gabor transform. Note that the *x*-axis represents time (seconds), the *y*-axis represents frequency (Hz), and the colour intensity indicates the signal amplitude in decibels (dB).

It is noteworthy that this research utilises three primary techniques for interpreting these non-auditory signals: the STFT, Gabor, and the CWT. All of these approaches are designed to transform the auditory signals into two-dimensional images or characteristic representations that are compatible with their integration into various models [15].

Firstly, STFT is a mathematical technique used in signal processing to determine the sinusoidal frequency and the phase content of the local sections of a time-series signal. The primary feature of this signal comprises a series of Fourier transforms of a windowed signal with the window being shifted along the time axis for each calculation stage [16]. The STFT is defined in Equation (1) [17,18].

$$STFT\{x(t)\}(f, \tau) = \int_{-\infty}^{+\infty} x(t)\, w(t - \tau)\, e^{-j2\pi ft}\, dt \tag{1}$$

where $x(t)$ is the input signal, $w(t - \tau)$ is the window function, centred at time τ. f is the frequency variable. $e^{-j2\pi ft}$ is the complex exponential function, representing a sinusoid. The integral is taken over time. In this equation, the window function $w(t - \tau)$ is critical as it localises the signal in time, allowing the Fourier transform to provide frequency information specific to the time segment around τ. The result is a function of both time (τ) and frequency (f), providing a time–frequency representation of the original signal $x(t)$.

Figure 6. Gabor transform of the cough and respiratory signals, representing the spectra of ((**a**): **Top Left**) asthma cough audio, ((**b**): **Top Right**) healthy cough audio, ((**c**): **Bottom Left**), asthma respiratory audio, and ((**d**): **Bottom Right**) healthy respiratory audio.

Secondly, CWT is a specific analysis method used in signal processing to extract the frequency information from the time series signal [19]. CWT aims to employ wavelets, that is, small waves that grow and decay within a confined duration. These wavelets are scaled and translated versions of a mother wavelet, offering a more flexible approach to analysing signals with varying frequency content over time. The CWT of a continuous signal $x(t)$ is defined in Equation (2):

$$CWT(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t)\, \psi^* \left(\frac{t - \tau}{s} \right) dt \qquad (2)$$

where $x(t)$ is the input signal. $\psi(t)$ is the mother wavelet. $\psi^*(t)$ is the complex conjugate of the mother wavelet. τ is the translation parameter. s is the scale parameter. The integral is taken over time. Note that the function $\psi\left(\frac{t-\tau}{s}\right)$ represents the wavelet function scaled by s and translated by τ.

Thirdly, **Gabor** analysis is a variant of STFT where the key difference lies in the use of a specific window function, the Gaussian window, which is a bell-shaped curve. This window is applied to the signal to isolate small sections for analysis. The Gaussian window is known for its minimal spread in both the time and frequency domains, allowing for a

more precise time–frequency representation [4]. The Gabor transform of a signal $x(t)$ is defined in Equation (3):

$$G(t, f) = \int_{-\infty}^{\infty} x(\tau) \, e^{-\pi(\tau-t)^2} \, e^{-j2\pi f \tau} \, d\tau \tag{3}$$

where $x(t)$ is the input signal. $e^{-\pi(\tau-t)^2}$ represents the Gaussian window function, at time t. $e^{-j2\pi f \tau}$ is the complex exponential function, representing a sinusoid of frequency f. The integral is taken over time τ.

It is imperative to mention that this equation applies a Gaussian window to the signal and then computes the Fourier transform of the windowed signal. The result $G(t, f)$ represents the time–frequency distribution of the signal, whereby we have the following:

x-axis: Represents time in seconds, indicating the progression of the signal over a 4-s duration.

y-axis: Denotes frequency in Hertz (Hz), ranging from 0 Hz to 25,000 Hz, which captures the frequency components of the signal.

z-axis: Indicates the amplitude or intensity of the signal at each time–frequency point, measured in decibels (dB).

Finally, the colour map provides a visual representation of the amplitude. In particular, the warmer colours (yellow) correspond to higher intensity levels and the cooler colours (purple) to lower levels, which could be representative of the distinct features of the disease under detection.

2.4. Feature Extraction

In this study, we have adopted lightweight models, namely, ShuffleNet, SqueezeNet, MobileNet and EfficientNet, to extract features from the spectrogram.

It is especially noted that these models were trained on the ImageNet platform, commonly used in several other studies [10,20–22]. In particular, ImageNet is a large-scale, diverse dataset that has over 14 million labelled images widely used as a benchmark for the training and evaluation of deep learning models in the computer vision area. Its extensive coverage and pre-trained models can enable an effective transfer learning process, making it a foundational resource for advancing research in real-world applications. These lightweight models were also selected based on their computational efficiency, which is beneficial for mobile and edge devices with limited processing power. The specific importance of each lightweight model used in designing an automated system for asthma detection using respiratory and cough sound signals is as follows:

- **ShuffleNet** is a type of neural network architecture that is ideal for devices with limited computational resources, like smartphones or other mobile devices. The data flow through channels in layers. In ShuffleNet, the data from these channels are mixed up or "shuffled". This mixing helps ShuffleNet learn better from the data without needing more power or memory [10]. ShuffleNet organises its data channels into groups. Within these groups, it does convolutions (filtering data) that help it learn from images or other inputs. By using groups, it can do many operations at the same time, which saves time and energy.

- **EfficientNet** is an innovative neural network architecture that sets new benchmarks for both efficiency and ACC. EfficientNet uses compound scaling to uniformly scale the network's depth, width, and resolution using a set of fixed scaling coefficients. This approach not only increases the depth and improves the resolution but also ensures that the foundations become wider, so everything stays in proportion and works well together [23].

- **SqueezeNet** is designed to be both small in size and fast in performance, while still maintaining a high level of accuracy in tasks, requiring less computational power and storage. SqueezeNet allows it to do more with fewer parameters. The architecture uses squeezed layers that reduce the number of parameters, followed by expanded layers that increase them again. The building blocks of SqueezeNet are called fire modules, which are made up of squeezed and expanded layers. These modules are designed to keep the network efficient [24].
- **MobileNet** is designed with a focus on mobile and embedded vision applications, providing high-performance model architectures that can run efficiently on smartphones and other devices with limited computational resources. The core ideas in MobileNet involve the use of separable convolutions [25]. It breaks down the usual complex operations into simpler, smaller operations that are easier to compute. This results in a dramatic reduction in the number of computations and the model size. The architecture is modular, which means it is made up of building blocks that can be mixed and matched to create networks that suit different needs and capacities.

It is expected that each CNN model will extract 500 features concatenated with other models, so the features may be up to 2000 in all of the net combinations. And it processes its respective time–frequency representation to extract a set of features that are particularly suited to the characteristics captured by the transformation [26], resulting in a rich, diverse feature set for subsequent analysis. There are four CNN models, resulting in a total of fifteen combinations for feature extraction. Two kinds of feature selectors are provided, so there are 450 model trials shown in Table 3 in each dataset.

Table 3. The signal transformation method, feature extractor, feature selector, and the final classification model used to develop an automated lightweight model for asthma detection with respiratory and cough sounds.

Transformation	Feature Extractor	Feature Selector	Classifier
Gabor	ShuffleNet	Relief	SVM
	ShuffleNet+SqueezeNet		
	ShuffleNet+MobileNet		
	ShuffleNet+EfficientNet		
	ShuffleNet+SqueezeNet+ MobileNet		RF
CWT	ShuffleNet+SqueezeNet+EfficientNet		
	SqueezeNet"2"		DT
	SqueezeNet+MobileNet		
	SqueezeNet+EfficientNet	NCA	
	SqueezeNet+MobileNet+EfficientNet		
STFT	ShuffleNet+MobileNet+EfficientNet "134"		NN
	MobileNet		
	MobileNet+EfficientNet		
	EfficientNet		KNN
	All Nets		

2.5. Feature Selection

This study has adopted the Relief method as an algorithm designed to assess the quality of the features in a dataset based on how well their values distinguish between instances that are near each other. Essentially, it works by repeatedly sampling an instance

from the dataset and then finding its nearest neighbour from the same class (nearest hit) and from a different class (nearest miss) [27]. The relevance of each feature is updated based on how well it distinguishes the sampled instance from its nearest hit and nearest miss.

This study has also adopted the NCA method as a sophisticated feature selector in ML that aims to enhance classification performance by learning a distance metric to organise data more effectively. It prioritises crucial features by assigning weights, which helps in clustering similar instances closer together and distancing dissimilar ones. NCA focuses on maximising a function that quantifies the probability of correct classification based on nearest neighbours within the learned metric. The method involves using gradient descent to iteratively adjust feature weights to achieve optimal performance [28]. By emphasising important features and reducing dimensionality, NCA not only speeds up the training process but also improves the model's ACC by mitigating the risk of overfitting. Although powerful, NCA requires significant computational resources and careful parameter tuning, making it ideal for applications with complex datasets where the relationships between features and outcomes are not straightforward.

In this case, 500 features are ranked in feature selection steps, and the original feature size is from 500 to 2000, depending on the feature extraction used. Feature selection helped to reduce computation in modelling and avoid over-fitting.

2.6. Classification Methods

The selected features were then fed into classical ML classifiers based on SVM, NN, RF, KNN, and DT algorithms. A summary and brief justification for the use of each classifier method are outlined in the following:

- **RF** is an ensemble learning method that operates by constructing multiple decision trees during training and outputting the class that is the mode of the classes of the individual trees [29].
- **SVM** is a classifier that finds the optimal boundary to separate different classes by maximising the margin between support vectors, which are the data points closest to the boundary [30].
- **NNs** are a set of algorithms, modelled loosely after the human brain, which are designed to recognise patterns and perform tasks like classification by learning from examples [31].
- **DT** is a flowchart-like tree structure where an internal node represents a feature(or attribute), the branch represents a decision rule, and each leaf node represents the outcome [32]. These classifiers are trained to detect the presence of respiratory diseases by recognising patterns and signatures in the cough and respiratory sounds that are characteristic of the condition.
- **KNN** is a simple, instance-based learning algorithm that classifies a sample based on the majority vote of its k nearest neighbours in the feature space [33].

Overall, the present study utilises five different classifier models that provide their own merits in developing an automated model for the detection of asthma.

2.7. Parameter Tuning

In this study, we use several ML models, including RF, NN, SVM, DT, and KNN. The choice of model parameters for each of these models is, therefore, paramount. These key model parameters were optimised through a grid search approach to achieve an optimal performance of ML models, and the range of model parameters is outlined below.

- RF: Criterion: ['gini', 'entropy', 'log_loss']. N Estimators: [100, 200, 300] to stabilise predictions. Max Depth: [10, 20, 30, 40, 50]. Max Features: 'sqrt', to improve model diversity.

- SVM: C: [0.1, 1, 10, 100] to control regularisation strength. Kernel: ['linear', 'poly', 'rbf'] for handling both linear and non-linear relationships.
- NN: Activation: ['identity', 'logistic', 'tanh', 'relu'] to introduce non-linearity. Solver: ['sgd', 'adam'] to explore different optimisation techniques. Hidden Layer Sizes: [(32), (64), (128), (32, 32), (64, 64), (128, 128)] for different network depths. Alpha: [0.0001, 0.001, 0.01, 0.1, 1.0] as a regularisation parameter to prevent over-fitting.
- DT: Criterion: ['gini', 'entropy', 'log_loss'] as splitting criteria. Max Depth: [10, 20, 30, 40, 50] to control tree depth and prevent over-fitting. Min Samples Split: 10, to ensure each split has sufficient samples. Min Samples Leaf: 10, to avoid overly small leaves.
- KNN: N neighbours: [3, 5, ..., 135] (odd numbers), to find the optimal number of neighbours.

2.8. Model Evaluation

Model performance for automated asthma detection was evaluated using 5-fold cross-validation and a set of key metrics, as follows: ACC, F1 score, precision, recall, and AUC. These metrics were also used for tuning model parameters. The mathematical definitions of these metrics are presented below.

- **ACC**: This measures the proportion of true results (both true positives and true negatives) in the total number of cases examined [34]. It is defined in Equation (4).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where we have the following:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

- **Precision**: This is also known as PPV; quantifies the proportion of correctly predicted positive cases among all predicted positives [34]. Mathematically, it is defined as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5}$$

- **Recall**: This is also known as sensitivity or the true positive rate; it measures the proportion of actual positives correctly identified [7]. It is defined as follows:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6}$$

- **F1 score**: This is the harmonic mean of precision and recall, offering a balance between them [7]. It is defined as follows:
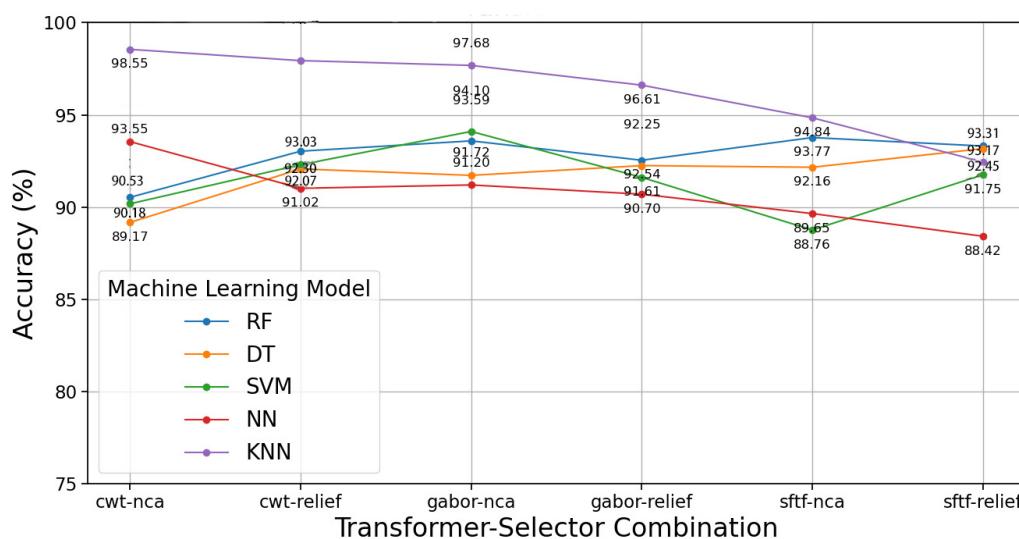
$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

As a final evaluation method of the proposed asthma detection system, we employ the ROC curve [35] and the AUC metric [36]. The ROC curve plots the true positive rate against the false positive rate across various classification thresholds, providing insights into the model's discrimination ability. The AUC quantifies the overall performance, with values closer to 1.0 indicating superior classification capability. Given the use of multiple classifiers (DT, SVM, NN, RF, KNN), AUC comparisons help us to determine the most effective model for distinguishing asthma-related respiratory and cough sounds from non-asthmatic cases.

3. Results

3.1. Proposed Model Results with Respiratory Sound

Figure 7 compares the average performance of the five proposed ML classifiers applied across these different feature extraction methods, where the values range from about 82% to 98%. It is, therefore, notable that several feature extractors seem to yield relatively high ACC across the different transformations applied. This suggests that all lightweight CNN net combinations are effective in detecting asthma. However, CWT is consistently among the highest-performing transformations across various feature extraction methods. The transformation of CWT with the NCA selector does not consistently improve the average performance of the classifiers. For some feature extraction methods, the performance with the NCA selector is slightly lower or comparable to the performance with the Relief selector. The increased performance variability seen with Gabor and Gabor-NCA transformations may be attributed to their inherent sensitivity to the specific characteristics of the input features.

Figure 7. Average ACC obtained using classifiers and majority voting for the case of respiratory sound.

The results of Figure 7 highlight several noteworthy patterns with respect to model performance across different transformation–selector combinations. KNN stands out with exceptionally high ACC, reaching 98.55%, particularly when paired with STFT-Relief. However, this uniformity in performance, achieved with a fixed parameter setup (n_neighbours = 3), raises concerns about over-fitting or over-dependence on specific feature extraction and selection methods. In contrast, SVM and RF models exhibit more consistent performance across combinations, with peak accuracies of 94.10% and 93.77%, respectively, under Gabor-NCA and STFT-NCA. These results suggest that these models may offer better robustness and adaptability to different feature engineering strategies.

The role of feature extraction and selection methods is also apparent. For example, Gabor-NCA appears to be a strong combination, yielding high ACC across multiple models, including SVM, RF, and DT. On the other hand, NN and DT exhibit more variability and generally lower performance, with their best results at 93.55% and 93.17%, respectively. These observations underscore the importance of selecting suitable feature engineering techniques and hyperparameters for optimal model performance. Further analysis should focus on validating the reliability of KNN's exceptional performance, exploring the impact of feature engineering choices, and refining the configurations of models like SVM and RF to improve their adaptability and generalizability.

Table 4 reveals that—in a respiratory sound dataset—while KNN demonstrates a high ACC (e.g., 98.55% with Gabor-NCA), its other evaluation metrics, such as F1 score, precision, recall, and AUC, are significantly less balanced. For instance, despite the high ACC, the F1 score and recall for Gabor-NCA + KNN are only 66.46% and 57.30%, respectively, indicating that the model struggles with certain classes or edge cases. Moreover, models like CWT and CWT-Relief exhibit lower accuracies (e.g., 96.02% and 94.52%) without compensating improvements in other metrics, further highlighting KNN's limitations in achieving balanced performance.

Table 4. The performance metrics (%) of the proposed SVM and KNN models based on the respiratory sound signal. N = n_neighbour, EXT represents the number of features used, and SEL represents the feature selection method.

Transform	EXT	SEL	Model	ACC	Parameter	F1	Precision	Recall	AUC
Gabor	124	NCA	KNN	98.55	n_neighbour = 3	66.46	63.43	57.30	69.86
Gabor	1234	NCA	KNN	97.12	n_neighbour = 3	66.05	64.04	65.30	67.28
Gabor	12	NCA	KNN	97.10	n_neighbour = 3	58.45	53.94	70.56	60.09
CWT	12	Relief	KNN	96.02	n_neighbour = 3	61.93	57.13	72.86	61.66
CWT	13	Relief	KNN	94.52	n_neighbour = 3	59.13	58.41	64.17	63.13
Gabor	34	NCA	SVM	94.12	C = 1/kernel linear	82.11	83.36	77.73	85.54
Gabor	23	NCA	SVM	94.10	C = 100/kernel RBF	74.73	73.11	75.69	77.69
Gabor	24	NCA	SVM	94.05	C = 100/kernel poly	92.49	91.04	91.85	93.06

On the other hand, SVM demonstrates better overall stability and balance across all metrics (see Tables 4 and 5). Further, the Gabor + NCA + SVM model achieves an F1 score of 92.49%, precision of 91.04%, recall of 91.85%, and AUC of 93.06%, indicating robust and reliable performance. Even with different parameter configurations (e.g., C = 1, kernel = linear or C = 100, kernel = poly), SVM maintains consistent and competitive results. Therefore, while KNN may deliver high ACC, its lack of balance across other critical metrics makes SVM a more suitable and reliable choice for this dataset. Furthermore, Table 5 presents the performance metrics with 95% confidence intervals for the respiratory sound dataset. It is evident that the SVM model registered the highest accuracy, recision, recall and F1 score compared to the random forest, KNN, decision tree and neural network models.

Table 5. The performance metrics with the 95% confidence interval (CI) for the respiratory sound dataset. The best model is boldfaced.

Model	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F1 Score (95% CI)
SVM	94.05 ± 1.2%	93.8 ± 1.3%	93.6 ± 1.2%	93.7 ± 1.2%
Random Forest	92.70 ± 1.4%	92.5 ± 1.5%	92.3 ± 1.4%	92.4 ± 1.3%
KNN	91.80 ± 1.6%	91.5 ± 1.5%	91.2 ± 1.6%	91.4 ± 1.5%
Decision Tree	89.50 ± 1.9%	89.3 ± 1.8%	89.0 ± 1.9%	89.1 ± 1.8%
Neural Network	93.10 ± 1.3%	92.9 ± 1.3%	92.7 ± 1.2%	92.8 ± 1.3%
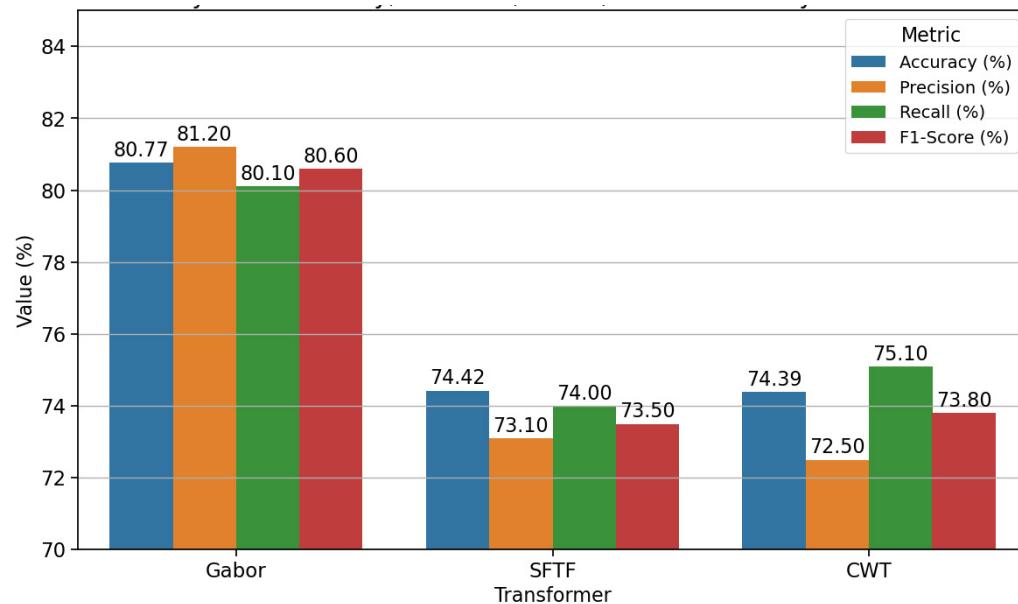
3.2. Proposed Model Results with Cough Sound

The cough sound dataset consists of 994 and 881 recordings; the recordings were collected from 511 asthmatic individuals (103 males and 408 females; mean age 55.23 ± 14.97 years, age range 10–82 years) and 815 non-asthmatic subjects (509 males and 306 females, primarily healthy university students without a history of asthma), respectively. This dataset was collected using a mobile phone's microphone, specifically a basic mobile phone with a sampling frequency of 48 kHz. For asthmatic participants, the duration of the recordings ranged from a minimum of 0.5 s to a maximum of 6.59 s. In

contrast, the durations of the recordings from non-asthmatic subjects ranged from 0.23 s to 5.42 s. This dataset is a valuable resource for studying the acoustic characteristics of cough sounds and their medical applications in distinguishing between asthmatic and non-asthmatic participants.

Based on aggregated ACC for each transformation method, the analysis revealed the best and the highest average ACC attained using the Gabor transformation to be 80.77% and 80.95%, respectively. When using the CWT transformation method, we obtained an average ACC of 74.39% and the highest ACC of 79.74%; using STFT resulted in an average ACC of 74.42% and the highest ACC of 74.84%.

Figure 8 compares the performance of the three transformations—Gabor, STFT, and CWT—on a cough sound dataset using the following four key metrics: accuracy, precision, recall, and F1 score. Among these, Gabor demonstrates the highest performance with an accuracy of 80.77%, precision has an accuracy of 81.20%, recall has an accuracy of 80.10%, and the F1 score has an accuracy of 80.60%, making it the most effective transformation for this dataset.

Figure 8. Model accuracy in terms of the transformation method and feature selection combinations for the model developed using cough sound signals.

It is interesting to note that both STFT and CWT also deliver comparable results, with STFT achieving an accuracy of 74.42%, precision of 73.10%, recall of 74.00%, and an F1 score of 73.50%. Similarly, CWT records an accuracy of 74.39%, precision of 72.50%, recall of 75.10%, and an F1 score of 73.80%. While their overall performances are consistent, CWT's higher recall suggests it identifies positive cases slightly better than STFT but at the expense of lower precision.
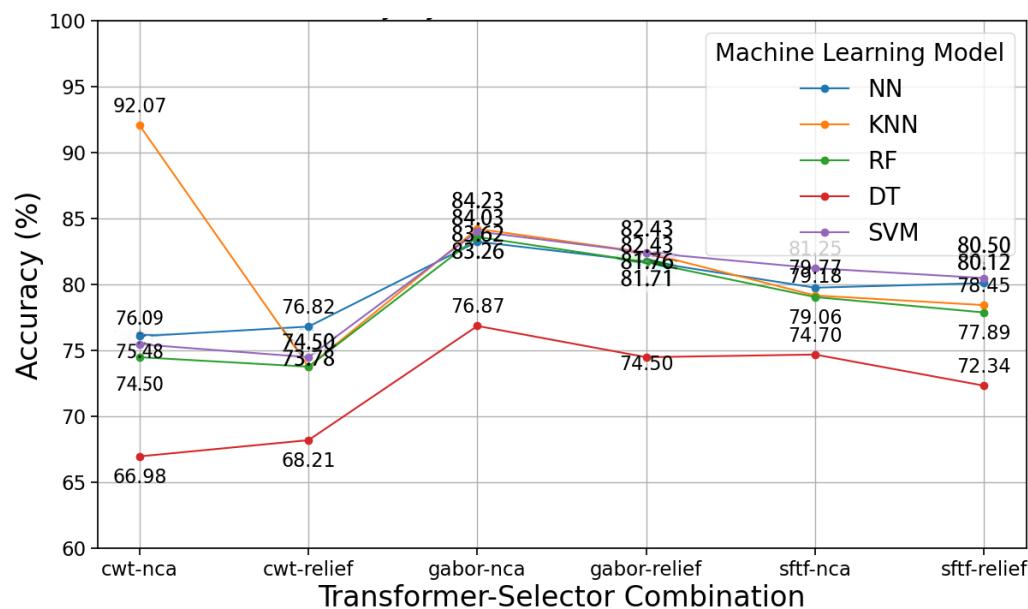
Overall, we show that the Gabor transformation method stands out as the most suitable transformation for this dataset, significantly outperforming STFT and CWT in all metrics. This suggests Gabor's robustness in capturing meaningful features from cough sound signals. Further optimisation of STFT and CWT could potentially narrow the performance gap and enhance their suitability for similar tasks.

Table 6 highlights the performance of different transformation and feature selection combinations on the cough sound dataset, measured by average and highest ACC. The Gabor-Relief combination stands out with the highest average ACC (80.79%) and peak ACC (80.95%), closely followed by Gabor-NCA. In contrast, combinations like CWT-Relief

and STFT-Relief show relatively lower performance. Overall, the Gabor transformation paired with either Relief or NCA consistently delivers the best results.

Table 6. The average and the highest accuracy obtained by the transformation method and feature selection.

Transformation-Feature Selection	Average ACC (%)	Highest ACC (%)
CWT-NCA	74.65	79.74
CWT-Relief	69.82	71.94
Gabor-NCA	80.79	80.95
Gabor-Relief	80.18	80.68
STFT-NCA	73.63	74.84
STFT-Relief	71.64	74.15

Figure 9. Model ACC in terms of the transformation method and feature selection combinations for the case of cough sound signals.

As presented in Figure 9, we now evaluate the ACC (%) values of five ML models (NN, KNN, RF, DT, and SVM) applied across six transformation-feature selection combinations (CWT-NCA, CWT-Relief, Gabor-NCA, Gabor-Relief, STFT-NCA, STFT-Relief). Among these, KNN achieves the highest ACC of 92.07% for the CWT-NCA combination, but its performance is inconsistent, dropping to 78.45% for STFT-Relief. NN and RF perform consistently, with Gabor-NCA yielding the best results for both (83.26% and 83.62%, respectively). DT, however, shows the weakest performance overall, with ACC peaking at only 76.87% for Gabor-NCA, making it the least suitable model.

In contrast, SVM demonstrates the most balanced and robust performance, achieving high ACC across all combinations, with its best result at 84.03% for Gabor-NCA and maintaining strong results for STFT-Relief (80.50%). Combinations involving the Gabor transformation (e.g., Gabor-NCA, Gabor-Relief) consistently yield the highest ACC for most models, indicating their suitability for this dataset. Based on this analysis, SVM paired with Gabor-NCA emerges as the optimal choice for achieving reliable and consistent results.

Table 7 reveals a significant disparity between the ACC and other metrics for KNN models. While the best-performing KNN configuration achieves an ACC of 92.07% with

the CWT transformation, NCA selector, and feature extractor 34, its other metrics are much lower—F1 score (59.81%), precision (59.82%), and recall (62.85%). This pattern persists in other KNN configurations, such as one with 89.62% ACC, where the F1 score drops to 42.32%, precision to 54.51%, and recall to 38.73%, or configuration with 85.76% ACC, which records an F1 score of only 30.45%, precision of 51.98%, and recall of 26.92%. These results highlight a systemic issue with KNN, where high ACC often comes at the expense of other critical metrics, making it unsuitable for tasks requiring balanced predictions.

Table 7. The performance metrics (%) of the proposed SVM and KNN models based on cough sound signals. N = n_neighbour, EXT represents the feature extractor method, and SEL represents the feature selector method.

Transform	EXT	SEL	Model	ACC	Parameter	F1	Precision	Recall	AUC
CWT	NCA	134	KNN	85.76	n_neighbour = 3	30.45	51.98	26.92	54.86
CWT	NCA	23	KNN	88.60	n_neighbour = 3	43.09	54.98	40.34	57.26
CWT	NCA	123	KNN	89.62	n_neighbour = 3	42.32	54.51	38.73	57.76
CWT	NCA	124	KNN	90.12	n_neighbour = 3	48.81	58.81	45.04	63.87
CWT	NCA	24	KNN	90.12	n_neighbour = 3	32.69	54.12	30.24	56.99
CWT	NCA	34	KNN	91.12	n_neighbour = 3	48.45	56.82	47.06	60.83
STFT	Relief	34	KNN	92.07	n_neighbour = 3	59.81	59.82	62.85	65.77
Gabor	NCA	24	SVM	83.31	C = 10, kernel = rbf	80.07	75.93	80.41	75.90
Gabor	NCA	2	SVM	84.03	C = 10, kernel = rbf	79.84	76.66	78.57	76.63
Gabor	NCA	124	SVM	84.03	C = 10, kernel = rbf	80.23	76.85	79.13	76.97

In contrast, SVM models paired with the Gabor transformation and NCA selector exhibit much more balanced performance across all metrics. For example, the SVM configuration with feature extractor 24 achieves an ACC of 83.31%, alongside a strong F1 score (80.07%), precision (75.93%), and recall (80.41%). Similarly, the SVM configuration with feature extractor 124 maintains an ACC of 84.03%, with an F1 score of 80.23%, precision of 76.85%, and recall of 79.13%. These balanced results demonstrate SVM's ability to generalise effectively while avoiding the trade-offs seen in KNN. Given its consistent and reliable performance across all metrics, SVM with the Gabor transformation and NCA selector is the preferred choice for this dataset, making it better suited for real-world tasks requiring balanced and dependable predictions. This result has also been reinforced in Table 8 where the SVM model demonstrates the highest accuracy, precision, recall and F1 score.

Table 8. The performance metrics with the 95% confidence interval (CI) for the cough sound dataset. The best model is boldfaced.

Model	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F1 Score (95% CI)
SVM	83.31 ± 1.5%	83.0 ± 1.6%	82.8 ± 1.5%	82.9 ± 1.4%
Random Forest	82.10 ± 1.6%	81.8 ± 1.5%	81.5 ± 1.6%	81.6 ± 1.5%
KNN	81.05 ± 1.8%	80.7 ± 1.7%	80.5 ± 1.8%	80.6 ± 1.7%
Decision Tree	79.50 ± 2.0%	79.1 ± 1.9%	78.8 ± 2.0%	78.9 ± 1.9%
Neural Network	82.75 ± 1.4%	82.4 ± 1.5%	82.2 ± 1.4%	82.3 ± 1.4%

4. Discussion

To the best of the authors' knowledge, this is the first work to propose a unified model for asthma detection using both cough and respiratory sounds. Previous studies have focused on either cough-based or respiratory sound-based analysis, requiring separate models and potentially even pre-processing steps. In contrast, our approach eliminates the

need to distinguish between sound types prior to analysis. This enables continuous input of audio data without the computational overhead of sound-type classification, making the method more efficient and cost-effective. Moreover, by removing this intermediate decision step, we reduce the potential for error propagation, which might enhance the system's overall robustness and reliability.

For respiratory sound data, Xu et al. [37] achieved a significant performance improvement, reporting an ACC of 94.05%, compared to the performance reported by Haider et al. [38] and Nabi et al. [39]. For cough sound data, the model achieves an ACC of 83.31% on a significantly larger dataset (1943 samples), compared to work by Infante et al. [40] and Balamurali et al. [41], which reported higher AUC values of 94% and 91.2%, respectively, but on much smaller datasets. This highlights the robustness of the model in handling large-scale data, despite its slightly lower performance metrics compared to models using logistic regression or pDNN. Xi et al. [42] explored exhaled aerosol patterns for non-invasive lung disease diagnosis. While disease differentiation remains challenging, their approach—integrating fractal analysis with SVM—achieved 100% ACC on a small dataset and 99.1% on a larger one, demonstrating strong potential for precise, non-invasive asthma detection.

4.1. Key Findings

It is important to highlight that the results from both datasets—cough sound and respiratory sound—indicate some common trends. While KNN models exhibit high ACC in some configurations, their other metrics, such as F1 score, precision, and recall, remain unbalanced, making them less suitable for reliable and robust classification across both datasets. For example, in the respiratory sound dataset, KNN models show high ACC (e.g., 98.55% with Gabor-NCA), but their F1 scores and recall values are significantly lower, demonstrating poor performance on certain classes. Similarly, in the cough sound dataset, KNN models with various configurations also fail to provide balanced results, despite achieving high ACC.

Table 9 presents a comprehensive performance comparison between our proposed lightweight hybrid model and several state-of-the-art methods previously reported in the literature. Specifically, for respiratory sound data, our framework achieves an accuracy of 94.05%, which notably exceeds the reported AUC of approximately 83.6% from earlier studies employing SVM and ensemble classifiers [38,39]. This improvement in accuracy clearly highlights the efficacy of integrating lightweight CNN architectures with classical ML methods alongside optimised feature selection NCA, demonstrating superior diagnostic capability compared to existing techniques.

In cough sound-based asthma detection, our method achieved an accuracy of 83.31%. While this is slightly lower than the best-reported performances—91.2% accuracy by Balamurali et al. [41] using deep neural networks and 94.0% AUC by Infante et al. [40] using logistic regression—it is crucial to emphasise that our results were validated on a significantly larger dataset, comprising 1943 cough segments from 1326 individuals. This much larger and more heterogeneous dataset offers a realistic evaluation scenario, confirming the robustness and generalizability of our proposed model in practical clinical settings. Additionally, our lightweight approach substantially reduces computational complexity and resource requirements, making it especially suitable for deployment in resource-constrained or edge-computing environments.

In summary, our hybrid model not only surpasses existing respiratory sound analysis methods but also achieves competitive performance on cough-based diagnosis under significantly more demanding conditions. These findings validate the potential of our lightweight hybrid approach for practical and efficient asthma diagnostics.

Table 9. Comparison of the proposed framework with existing methods on cough and respiratory sound analysis.

Author, Year	Subtype	Classifier		Performance	Data	Data Size
		Type	Algo.			
Balamurali et al. 2019 [41]	Cough sound	DL	DNN	ACC = 91.2%	Private	51 subjects
Infante et al. 2017 [40]	Cough sound	ML	LR	AUC = 94	Private	54 patients
Xu et al. 2024 [37]	Cough sound	ML	SVM	ACC = 83.31%	Private	1943 cough sounds
Haider et al. 2019 [38]	Respiratory sounds	ML	SVM	AUC = 83.6	Private	30 COPD; 25 healthy subjects
Nabi et al. 2019 [39]	Respiratory sounds	ML	ensemble	AUC = 83.6	Private	55 asthmatic patients
Xu et al. 2024 [37]	Respiratory sounds	ML	SVM	ACC = 94.05%	Private	205 respiratory sounds
Proposed work	Cough sounds	Hybrid model for Normal vs. asthma	Feature extractor: Gabor-lightweight model Classifier: SVM	ACC = 83.31%	Private	1326 participants (511 asthma, 815 healthy) 1943 cough segments (511 asthma, 815 healthy)
	Respiratory sounds			ACC = 94.05%		

4.2. Limitations

Despite these promising results, several limitations should be acknowledged. This study relied on a limited dataset, which may have affected the generalizability of the findings to a broader population. Additionally, while ensemble feature extraction and majority voting show effectiveness, this approach increases computational complexity and may not be feasible in real-time diagnostic settings. Moreover, the reliance on specific feature extractors such as SqueezeNet, MobileNet, and EfficientNet means that the results are influenced by the inherent strengths and weaknesses of these models, potentially limiting adaptability to other types of sound data.

4.3. Future Work

In future work, we will aim to explore ensemble methods, such as majority voting, to enhance the overall ACC and robustness of the classification model across both cough sound and respiratory sound datasets. By combining predictions from multiple models—each trained with different configurations of feature extractors, feature selectors, and classifiers—majority voting can help mitigate the weaknesses of individual models and improve decision consistency.

For instance, leveraging models with complementary strengths, such as KNN for high ACC and SVM for balanced performance across metrics, could provide a more reliable classification system. This approach has the potential to further improve the model's generalizability and ensure that the system performs well under diverse real-world scenarios. Additionally, future research could investigate dynamic weighting in the voting mechanism to prioritise models with higher reliability for specific datasets or sound types.

The promising outcomes of this research show advancements in the early detection and management of asthma and COPD. Future research should focus on enhancing AI models to more effectively detect the early stages of these diseases [43]. This involves integrating larger and more diverse datasets that not only encompass a broader demographic but also include varied environmental exposures to improve the models' robustness and diagnostic ACC across different populations.

To address the complexities of these respiratory conditions, it is crucial to expand the datasets to include early-stage asthma and COPD cases. Such datasets will enable AI models to learn the subtle variations and early markers of these diseases, which are critical for timely and accurate diagnosis [44].

Further development could explore hybrid models that combine different ML techniques, including advanced DL algorithms. These models would leverage the strengths of various approaches to improve the sensitivity and specificity of diagnostics, making them more effective in distinguishing between asthma, COPD, and other respiratory conditions.

Refining the majority voting system by implementing weighted voting strategies—where weights are assigned based on each model's training performance—could enhance the overall diagnostic ACC. This approach would make the AI systems more reliable and robust, providing clinicians with tools that support better decision-making in clinical settings [45].

Moreover, the incorporation of real-time data processing capabilities could revolutionise diagnostic procedures by enabling instant and on-site disease detection. This is especially important for managing conditions like asthma and COPD, where early detection can significantly alter disease outcomes and improve patient management.

Additionally, embedding real-time data processing capabilities would revolutionise diagnostic procedures by facilitating immediate and on-site disease detection. This is particularly crucial for managing asthma, where the early detection of stage progression can significantly alter therapeutic approaches and improve patient outcomes [46].

5. Conclusions

This study marks a significant advancement in the AI-driven detection of respiratory diseases by combining cough and respiratory sound analysis. The proposed framework employs advanced machine learning techniques, including the SVM model with SqueezeNet and EfficientNet as a combined feature extractor (represented as "24") and NCA for feature selection. This configuration delivers a well-balanced and reliable diagnostic performance across key evaluation metrics, making it an optimal solution for analysing both cough and respiratory sound datasets. The approach ensures robustness and adaptability, addressing the diverse characteristics of the two datasets while maintaining consistency in performance.

The integration of dual-sound analysis and majority voting ensemble techniques has proven highly effective, achieving precise and consistent diagnostic outcomes. Among the tested configurations, the SVM model with the SqueezeNet + EfficientNet ("24") feature extractor emerges as the best choice for compatibility across datasets. It achieves a balanced performance in ACC, F1 score, precision, recall, and AUC, underscoring the model's versatility and reliability in diagnosing conditions such as asthma and AUC.

Future work should focus on expanding datasets to include broader demographic and environmental variations and refining real-time diagnostic capabilities. The continued development of ensemble methods, such as majority voting with dynamic weighting, could further enhance the ACC and robustness of the system. These advancements have the potential to revolutionise respiratory healthcare practices, enabling earlier detection and better management of asthma, COPD, and other respiratory conditions on a global scale.

Author Contributions: Conceptualization, S.X., R.A., O.F. and R.C.D.; methodology, S.X. and R.A.; software, S.X. and R.A.; validation, S.X., R.A., O.F. and R.C.D.; formal analysis, S.X. and R.A. investigation, S.X. and R.A.; resources, R.A., O.F., R.C.D., P.D.B. and J.S.; data curation, S.X.; writing—original draft preparation, S.X.; writing—review and editing, O.F., R.C.D. and J.S.; visualization, R.A. and O.F.; supervision, R.C.D., R.A., J.S. and P.D.B.; project administration, P.D.B., R.C.D., J.S. and R.A.; funding acquisition, P.D.B. All authors have read and agreed to the published version of the manuscript.

Abbreviations

The following abbreviations are used in this manuscript:

3G	International Mobile Telecommunications-2000
4G	successor to 3G
TRP	true positive rate
FRP	false positive rate
ACC	accuracy
A-V	atrioventricular
ACR	American College of Radiology
ADA	American Diabetes Association
ADASYN	adaptive synthetic sampling
AE	autoencoder
AED	automated external defibrillator
AES	advanced encryption standard
AFL	atrial flutter
AFIB	atrial fibrillation
AFDB	atrial fibrillation database
ApEn	approximate entropy
ALLSTAR	Allostatic State Mapping by Ambulatory ECG Repository
ALN	axillary lymph node
ALND	axillary lymph node dissection
ANFIS	adaptive network-based fuzzy inference system
ANS	autonomic nervous system
ANN	artificial neural network
ANOVA	analysis of variance
ANSI	American National Standards Institute
ANT	anterior left atrial free wall
AHE	adaptive histogram equalisation
AHI	Apnea–Hypopnea Index
AI	artificial intelligence
ASM	active shape model
API	application programming interface
AR	autoregressive

ARF	acoustic radiation force
ARMA	autoregressive moving average
ARMAX	autoregressive moving average with exogenous inputs
AT	atrial tachycardia
ATM	automated teller machine
AUC	area under the curve
AVN	atrioventricular node
ATS/ERS	American Thoracic Society/European Respiratory Society
QCT	quantitative CT
BA	Bayesian averaging
BCI	brain–computer interface
BEMD	bidimensional empirical mode decomposition
BERT	bidirectional encoder representations from transformers
MTS	Mahalanobis–Taguchi System
BLE	Bluetooth Low Energy
BMI	body mass index
BPA	Backpropagation Algorithm
BPSO	binary particle swarm optimisation
BPM	beats per minute
BSN	biomedical sensor network
CAD	computer-aided diagnosis
CaD	capacity dimension
CAN	cardiovascular autonomic neuropathy
CART	classification and regression tree
CD	celiac disease
CEUS	contrast-enhanced ultrasound
CI	confidence interval
CHF	congestive heart failure
CHS	community health centres
DL	deep learning
CLAHE	contrast limited adaptive histogram equalisation
CLDA	clustering linear discriminant analysis algorithm
CNN	convolutional neural network
CM	clustered microcalcification
COPD	chronic obstructive pulmonary disease
CSP	communicating sequential process
CS	compressed sampling
CT	computed tomography
CPA	communicating process architecture
CPC	cardiopulmonary coupling
CPU	central process unit
CSA	central sleep apnea
CSME	clinically significant macular oedema
CSR	Cambridge silicon radio
CVD	cardiovascular disease
CWT	continuous wavelet transform
CTA	tomographic pulmonary angiography
D2H2	distributed diagnosis and home healthcare
DAGSVM	directed acyclic graph support vector machine
DB	database
DBN	deep belief network
DCN	deep convolutional network
DCT	discrete cosine transform
DET	determinism

DFT	discrete Fourier transform
DII	Diabetic Integrated Index
DL	deep learning
DLLE	diagonal line length entropy
DM	diabetes mellitus
DN	diabetic neuropathy
DNN	deep neural network
DR	diabetic retinopathy
DT	decision tree
DWT	discrete wavelet transform
E	external
E-M	expectation–maximisation
ECG	electrocardiogram
EDR	ECG-derived respiration
ECHONET	EchoCardiographic Healthcare Online Networking Expertise in Tasmania
EEG	electroencephalogram
ELM	extreme learning machine
EM	electromagnetic
EMD	empirical mode decomposition
EMG	electromyogram
EOG	electrooculogram
FAQ	frequently asked questions
FB	fusion beat
FD	fractal dimension
FDDI	fibre-distributed data interface
FEn	fuzzy entropy
FES	functional electrical stimulation
FCM	fuzzy C-means
EHR	electronic health record
FIR	finite impulse response
EMR	electronic medical record
FN	false negative
FNA	fine-needle aspiration
FNN	false nearest neighbour
FP	false positive
FSC	fuzzy Sugeno classifier
FT	Fourier transform
GA	genetic algorithm
GB	gradient boosting
GLCM	grey-level co-occurrence matrix
GMM	Gaussian mixture model
GOE	Global Observatory for eHealth
GPC	Gaussian process classification
GPRS	General Packet Radio Service
GPU	graphics processing unit
GSM	Global System for Mobile communication
GUI	graphical user interface
H	Hurst exponent
HDNTs	hospital digital networking technologies
HIHM	home-integrated health monitor
HIPAA	Health Insurance Portability and Accountability Act
HMA	haemorrhages and microaneurysms
HMM	hidden Markov model
HOG	histogram of gradient

HOS	higher-order spectrum
HP	Helicobacter Pylori
HR	heart rate
HRUS	high-resolution ultrasound
HRV	heart rate variability
HRVAS	heart rate variability analysis software
HSDPA	high-speed downlink packet access
I	internal
IBD	inflammatory bowel disease
ICA	independent component analysis
ICU	intensive care unit
IEEE	Institute of Electrical and Electronic Engineers
IMF	intrinsic mode function
IoT	Internet of Things
IIR	infinite impulse response
IR	infrared
IT	information technology
IVUS	intravascular ultrasound
JMMB	*Journal of Mechanics in Medicine and Biology*
KS	Kolmogorov–Sinai
KNN	K-nearest neighbour
LA	left atrium
LAM	laminarity
LBP	local binary pattern
LC	linear classifier
LCP	local configuration pattern
LDA	linear discriminant analysis
LFP	local field potential
LIP	left inferior pulmonary
LLE	largest Lyapunov exponent
LMNN	Levenberg–Marquardt neural network
LP	lamina propria
LR	logistic regression
LSP	left superior pulmonary
LSTM	long short-term memory
LTAFDB	long-term AF database
LTE	long-term evolution
LR	logistic regression
LRNC	lipid-rich necrotic core
LV	left ventricular
MA	moving average
MC	clustered microcalcifications
MCA	multiple correspondence analysis
NCSME	non-clinically significant macular oedema
MCT	mobile cardiac telemetry
MDA	multiple discriminant analysis
NDDF	normal density discriminant function
MI	myocardial infarction
MIL	multiple instance learning
ML	machine learning
MLP	multilayer perceptron
MMSE	modified multiscale entropy
MOH	Ministry Of Health
MP	mobile phone

MR	magnetic resonance
MRI	magnetic resonance imaging
MSE	multiscale entropy
MTS	multivariate time series
MVA	multiple variable analysis
NASA	National Aeronautics and Space Administration
NB	Naïve Bayes
NCA	neighbourhood component analysis
nB	number of beats
NCD	non-celiac duodenitis
NEB	non-ectopic beat
NEWS	National Early Warning Score
NF	neuro-fuzzy
NN	neural network
NoF	number of features
nP	number of patients
NPDR	non-proliferative diabetic retinopathy
NSR	normal sinus rhythm
OFDM	orthogonal frequency-division multiplexing
OSA	obstructive sleep apnea
PAF	paroxysmal atrial fibrillation
PCA	principal component analysis
PDR	proliferative diabetic retinopathy
PET	positron emission tomography
PET-MRI	positron emission tomography-magnetic resonance imaging
PFT	pulmonary function testing
PHC	primary health centre
PHOG	pyramid histogram of gradient
PLS-DA	partial least squares discriminant analysis
PNN	probabilistic neural network
PPG	photoplethysmogram
POS	posterior left atrial free wall
PPV	positive predictive value
PRE	precision
PRISMA	preferred reporting items for systematic reviews and meta-analyses
PSD	power spectral density
PSG	polysomnography
PVC	premature ventricular contraction
QDA	quadratic discriminant analysis
QoS	quality of service
qCT	quantified computed tomography
RAM	random access memory
RBF	radial basis function
RBFNN	radial basis function neural network
RBM	restricted Boltzmann machine
REC	recurrence rate
ReLU	rectified linear unit
ResNet	residual neural network
RF	random forest
RHN	regional health network
RIP	right inferior pulmonary
RNA	ribonucleic acid
RNN	recurrent neural network
ROC	receiver operating characteristic

ROI	region of interest
RQA	recurrence quantification analysis
RP	recurrence plot
RSA	respiratory sinus arrhythmia
RSP	right superior pulmonary
RTF	recurrent neural network
RVM	relevance vector machine
SWT	stationary wavelet transform
SA	sinoatrial
SAE	sparse autoencoder
SampEn	sample entropy
SAN	sinoatrial node
SC	statistical classifier
SD	standard deviation
SCD	sudden cardiac death
SDNN	standard deviation of normal-to-normal intervals
SEN	sensitivity
SE	standard error
SpO2	oxygen saturation
ShanEn	Shannon entropy
SLNB	sentinel lymph node biopsy
SMDS	switched multi-megabit data service
SNN	spiking neural network
SOM	self-organising map
SPE	specificity
STFT	short-time Fourier transform
SWE	shear wave elastography
SWT	stationary wavelet transform
SQL	structured query language
SVEB	supraventricular ectopic beat
SVM	support vector machine
SNPs	single-nucleotide polymorphisms
MTS	Mahalanobis–Taguchi system
T	thoracic
TIA	transient ischaemic attack
TEE	transesophageal echocardiogram
TN	true negative
TP	true positive
TT	trapping time
TTE	transthoracic echocardiography
TVCF	time-varying coherence function
UB	unclassified beat
UIT	urinary tract infection
ULP	ultra-low power
UMMC	University Malaya Medical Centre
UMTS	Universal Mobile Telecommunications System
US	ultrasound
USG	ultrasonography
UWB	ultra-wideband
xAI	explainable AI
VDSNet	VGG data STN with CNN
VEB	ventricular ectopic beat
VF	ventricular fibrillation
VFL	ventricular flutter

VT	ventricular tachycardia
WBAN	wireless body area network
WEP	wired equivalent privacy
WHO	World Health Organization
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	wireless local area network
WMAN	wireless metropolitan area network
WPA	Wi-Fi-protected access
WPAN	wireless personal area network
WPD	wavelet packet decomposition
WWAN	wireless wide area network

Appendix A. Pseudocode

The pseudocode below describes the way in which the methods, described in Section 2, were used to establish the hybrid model.

```
LOAD dataset
CHECK data imbalance and distribution
APPLY data cleaning AND augmentation IF needed

FOR each candidate model IN [SVM, RandomForest, XGBoost, MLP, CNN]:
    PERFORM feature selection OR extraction (PCA/LDA/autoencoder)
    FOR each hyperparameter combination IN parameter_grid:
        PERFORM cross-validation (e.g., 5-fold CV)
        RECORD performance (accuracy, precision, recall, F1)

SELECT best-performing model AND parameters
RETRAIN best model ON entire training dataset
EVALUATE final model ON independent test dataset
```

References

1. Alzghoul, B.N.; Reddy, R.; Chizinga, M.; Innabi, A.; Zou, B.; Papierniak, E.S.; Faruqi, I. Pulmonary embolism in acute asthma exacerbation: Clinical characteristics, prediction model and hospital outcomes. *Lung* **2020**, *198*, 661–669. [CrossRef] [PubMed]
2. Lanclus, M.; Clukers, J.; Van Holsbeke, C.; Vos, W.; Leemans, G.; Holbrechts, B.; Barboza, K.; De Backer, W.; De Backer, J. Machine learning algorithms utilizing functional respiratory imaging may predict COPD exacerbations. *Acad. Radiol.* **2019**, *26*, 1191–1199. [CrossRef] [PubMed]
3. Abadi, M. TensorFlow: Learning functions at scale. In Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming, Nara, Japan, 18–24 September 2016; p. 1.
4. Chen, M.; Yu, L.; Zhi, C.; Sun, R.; Zhu, S.; Gao, Z.; Ke, Z.; Zhu, M.; Zhang, Y. Improved faster R-CNN for fabric defect detection based on Gabor filter with Genetic Algorithm optimization. *Comput. Ind.* **2022**, *134*, 103551. [CrossRef]
5. Lin, S.; Zhang, Q.; Chen, F.; Luo, L.; Chen, L.; Zhang, W. Smooth Bayesian network model for the prediction of future high-cost patients with COPD. *Int. J. Med. Inform.* **2019**, *126*, 147–155. [CrossRef] [PubMed]
6. Barua, P.D.; Keles, T.; Kuluozturk, M.; Kobat, M.A.; Dogan, S.; Baygin, M.; Tuncer, T.; Tan, R.S.; Acharya, U.R. Automated asthma detection in a 1326-subject cohort using a one-dimensional attractive-and-repulsive center-symmetric local binary pattern technique with cough sounds. *Neural Comput. Appl.* **2024**, *36*, 16857–16871. [CrossRef]
7. Lee, C.S.; Parise, C.; Burleson, J.; Seidenwurm, D. Assessing the recall rate for screening mammography: Comparing the medicare hospital compare dataset with the national mammography database. *Am. J. Roentgenol.* **2018**, *211*, 127–132. [CrossRef]
8. Peng, J.; Chen, C.; Zhou, M.; Xie, X.; Zhou, Y.; Luo, C.H. A machine-learning approach to forecast aggravation risk in patients with acute exacerbation of chronic obstructive pulmonary disease with clinical indicators. *Sci. Rep.* **2020**, *10*, 1–9. [CrossRef]
9. Wang, W.; Zhang, G.; Yang, L.; Balaji, V.; Elamaran, V.; Arunkumar, N. Revisiting signal processing with spectrogram analysis on EEG, ECG and speech signals. *Future Gener. Comput. Syst.* **2019**, *98*, 227–232. [CrossRef]

10. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 6848–6856.

11. Amin, H.U.; Malik, A.S.; Ahmad, R.F.; Badruddin, N.; Kamel, N.; Hussain, M.; Chooi, W.T. Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques. *Australas. Phys. Eng. Sci. Med.* **2015**, *38*, 139–149. [CrossRef]

12. Binson, V.A.; Subramoniam, M.; Mathew, L. Detection of COPD and Lung Cancer with electronic nose using ensemble learning methods. *Clin. Chim. Acta* **2021**, *523*, 231–238.

13. Giannakopoulos, T.; Pikrakis, A. *Introduction to Audio Analysis: A MATLAB® Approach*; Academic Press: Cambridge, MA, USA, 2014.

14. Purwins, H.; Li, B.; Virtanen, T.; Schlüter, J.; Chang, S.Y.; Sainath, T. Deep learning for audio signal processing. *IEEE J. Sel. Top. Signal Process.* **2019**, *13*, 206–219. [CrossRef]

15. Goto, T.; Jo, T.; Matsui, H.; Fushimi, K.; Hayashi, H.; Yasunaga, H. Machine learning-based prediction models for 30-day readmission after hospitalization for chronic obstructive pulmonary disease. *COPD J. Chronic Obstr. Pulm. Dis.* **2019**, *16*, 338–343. [CrossRef] [PubMed]

16. Burriel-Valencia, J.; Puche-Panadero, R.; Martinez-Roman, J.; Sapena-Bano, A.; Pineda-Sanchez, M. Short-frequency Fourier transform for fault diagnosis of induction machines working in transient regime. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 432–440. [CrossRef]

17. Mateo, C.; Talavera, J.A. Short-time Fourier transform with the window size fixed in the frequency domain. *Digit. Signal Process.* **2018**, *77*, 13–21. [CrossRef]

18. Dinh, P.H. Combining gabor energy with equilibrium optimizer algorithm for multi-modality medical image fusion. *Biomed. Signal Process. Control* **2021**, *68*, 102696. [CrossRef]

19. Aguiar-Conraria, L.; Soares, M.J. The continuous wavelet transform: Moving beyond uni-and bivariate analysis. *J. Econ. Surv.* **2014**, *28*, 344–375. [CrossRef]

20. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning. PMLR, Long Beach, CA, USA, 9–15 June2019; pp. 6105–6114.

21. Iandola, F.N. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv* **2016**, arXiv:1602.07360.

22. Howard, A.G. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.

23. Koonce, B.; Koonce, B. EfficientNet. In *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 109–123.

24. Koonce, B.; Koonce, B. SqueezeNet. In *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 73–85.

25. Sinha, D.; El-Sharkawy, M. Thin mobilenet: An enhanced mobilenet architecture. In Proceedings of the 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 10–12 October 2019; pp. 280–285.

26. Petmezas, G.; Cheimariotis, G.A.; Stefanopoulos, L.; Rocha, B.; Paiva, R.P.; Katsaggelos, A.K.; Maglaveras, N. Automated Lung Sound Classification Using a Hybrid CNN-LSTM Network and Focal Loss Function. *Sensors* **2022**, *22*, 1232. [CrossRef]

27. Urbanowicz, R.J.; Meeker, M.; La Cava, W.; Olson, R.S.; Moore, J.H. Relief-based feature selection: Introduction and review. *J. Biomed. Inform.* **2018**, *85*, 189–203. [CrossRef]

28. Amankwaa-Kyeremeh, B.; Greet, C.; Zanin, M.; Skinner, W.; Asamoah, R. Selecting key predictor parameters for regression analysis using modified Neighbourhood Component Analysis (NCA) Algorithm. In Proceedings of the 6th UMaT Biennial International Mining and Mineral Conference, Tarkwa, Ghana, 5–6 August 2020; pp. 320–325.

29. Rigatti, S.J. Random forest. *J. Insur. Med.* **2017**, *47*, 31–39. [CrossRef]

30. Binson, V.; Subramoniam, M.; Sunny, Y.; Mathew, L. Prediction of pulmonary diseases with electronic nose using SVM and XGBoost. *IEEE Sens. J.* **2021**, *21*, 20886–20895. [CrossRef]

31. Abiyev, R.H.; Ma'aitaH, M.K.S. Deep convolutional neural networks for chest diseases detection. *J. Healthc. Eng.* **2018**, *2018*. [CrossRef]

32. Rai, K.; Devi, M.S.; Guleria, A. Decision tree based algorithm for intrusion detection. *Int. J. Adv. Netw. Appl.* **2016**, *7*, 2828.

33. Zhang, S.; Cheng, D.; Deng, Z.; Zong, M.; Deng, X. A novel kNN algorithm with data-driven k parameter computation. *Pattern Recognit. Lett.* **2018**, *109*, 44–54. [CrossRef]

34. Faust, O.; Barika, R.; Shenfield, A.; Ciaccio, E.J.; Acharya, U.R. Accurate detection of sleep apnea with long short-term memory network based on RR interval signals. *Knowl.-Based Syst.* **2021**, *212*, 106591. [CrossRef]

35. Nahm, F.S. Receiver operating characteristic curve: Overview and practical use for clinicians. *Korean J. Anesthesiol.* **2022**, *75*, 25–36. [CrossRef]

36. De Hond, A.A.; Steyerberg, E.W.; Van Calster, B. Interpreting area under the receiver operating characteristic curve. *Lancet Digit. Health* **2022**, *4*, e853–e855. [CrossRef]

37. Xu, S.; Deo, R.C.; Soar, J.; Barua, P.D.; Faust, O.; Homaira, N.; Jaffe, A.; Kabir, A.L.; Acharya, U.R. Automated detection of airflow obstructive diseases: A systematic review of the last decade (2013–2022). *Comput. Methods Programs Biomed.* **2023**, *241*, 107746. [CrossRef]

38. Haider, N.S.; Singh, B.K.; Periyasamy, R.; Behera, A.K. Respiratory sound based classification of chronic obstructive pulmonary disease: A risk stratification approach in machine learning paradigm. *J. Med. Syst.* **2019**, *43*, 1–13. [CrossRef]

39. Nabi, F.G.; Sundaraj, K.; Lam, C.K.; Palaniappan, R. Characterization and classification of asthmatic wheeze sounds according to severity level using spectral integrated features. *Comput. Biol. Med.* **2019**, *104*, 52–61. [CrossRef] [PubMed]

40. Infante, C.; Chamberlain, D.B.; Kodgule, R.; Fletcher, R.R. Classification of voluntary coughs applied to the screening of respiratory disease. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Republic of Korea, 11–15 July 2017; pp. 1413–1416.

41. Balamurali, B.; Hee, H.I.; Kapoor, S.; Teoh, O.H.; Teng, S.S.; Lee, K.P.; Herremans, D.; Chen, J.M. Deep neural network-based respiratory pathology classification using cough sounds. *Sensors* **2021**, *21*, 5555. [CrossRef] [PubMed]

42. Xi, J.; Zhao, W.; Yuan, J.E.; Kim, J.; Si, X.; Xu, X. Detecting lung diseases from exhaled aerosols: Non-invasive lung diagnosis using fractal analysis and SVM classification. *PLoS ONE* **2015**, *10*, e0139511. [CrossRef] [PubMed]

43. Schroeder, J.D.; Lanfredi, R.B.; Li, T.; Chan, J.; Vachet, C.; Paine III, R.; Srikumar, V.; Tasdizen, T. Prediction of obstructive lung disease from chest radiographs via deep learning trained on pulmonary function data. *Int. J. Chronic Obstr. Pulm. Dis.* **2020**, *15*, 3455. [CrossRef]

44. Ullah, R.; Khan, S.; Ali, H.; Chaudhary, I.I.; Bilal, M.; Ahmad, I. A comparative study of machine learning classifiers for risk prediction of asthma disease. *Photodiagnosis Photodyn. Ther.* **2019**, *28*, 292–296. [CrossRef]

45. Luo, G.; Stone, B.L.; Fassl, B.; Maloney, C.G.; Gesteland, P.H.; Yerram, S.R.; Nkoy, F.L. Predicting asthma control deterioration in children. *BMC Med. Inform. Decis. Mak.* **2015**, *15*, 1–8. [CrossRef]

46. Kor, C.T.; Li, Y.R.; Lin, P.R.; Lin, S.H.; Wang, B.Y.; Lin, C.H. Explainable Machine Learning Model for Predicting First-Time Acute Exacerbation in Patients with Chronic Obstructive Pulmonary Disease. *J. Pers. Med.* **2022**, *12*, 228. [CrossRef]