MeLoDicA AI- Machine Learning Based Detection of Asthma via Vocal Audio Analysis

Zhi Qing Looi
School of Electrical and Electronic Engineering
Singapore Polytechnic
Singapore, Singapore
zhiqinglooi12@gmail.com

Zi Heng Ng
School of Electrical and Electronic Engineering
Singapore Polytechnic
Singapore, Singapore
xgmefixg@gmail.com

Ren Xiang Yak
School of Electrical and Electronic Engineering
Singapore Polytechnic
Singapore, Singapore
yakrenxiang@gmail.com

Oren Rosen
Electrical Engineering Faculty
Technion – Israel Institute of Technology
Haifa,Israel
orenren87@yahoo.com

Arun Kumar
School of Electrical and Electronic Engineering
Singapore Polytechnic
Singapore, Singapore
Arun_KUMAR@ichat.sp.edu.sg

Abstract*— Mild asthma symptoms can be easily mistaken for those of other health conditions, e.g., allergy and obesity-related symptoms. This might lead to delayed diagnosis and wrong treatment, risking complications and even permanent lung damage. In this study, we introduce a pioneering method for asthma detection. By integrating machine learning (ML) algorithms and advanced audio analysis techniques, the proposed method rapidly and accurately detects asthmatic patients from audio clips containing vocal recordings without the need for dedicated medical devices. By utilising various ML algorithms, the examination of audio clips analyses diverse asthma-related airflow control and respiratory aspects. To strengthen the reliability and accuracy of the model, feature selection algorithms are integrated into the proposed method to address overfitting concerns and retain the most informative features. The model evaluation employs various performance metrics, e.g., accuracy and confusion matrix, to ensure a comprehensive and quantitative assessment of asthma detection. Another novelty of the proposed method is the inclusion of established data analytics tools, e.g., Spotfire, H2O and Sklearn. This approach marks a paradigm shift compared to existing methods, which results in a simplified and more effective asthma detection method. Examining the overall results, vowel pronunciation with its full set of 14 features using the Random Forest (RF) algorithm on the H2O tool yielded the highest accuracy of 92%. Moreover, Support Vector Machine (SVM) and K-Nearest Neighbours (KNN) with correlation-based feature selection (CFS) and customised hyperparameters yielded accuracy of 85% respectively. The proposed method contributes to the advancement of ML theory in healthcare applications, providing a practical framework for the implementation of asthma detection while demonstrating superior performance compared to existing.

Keywords*— *Asthma, Machine Learning, Respiratory audio, Acoustic signals, Random Forest, H2O, Spotfire©, K-Nearest Neighbours, Healthcare, Support Vector Machine, Correlation-based feature selection

I. Introduction

Asthma is a common lung disease characterised by inflammation of the airways [1]. Asthma is a major global health problem, affecting millions of people worldwide [2]. If left untreated, it can lead to severe respiratory complications and in extreme cases, it can even be fatal [1]. Asthma detection refers to the process of identifying or diagnosing asthma. Its accuracy is essential since nearly half of asthma-related deaths occur in patients initially misclassified as moderate or less severe [3].

Accurate detection of asthma is crucial as asthma patients may present with respiratory symptoms that are also present in other conditions, e.g., allergies, chronic obstructive pulmonary disease (COPD) and obesity, increasing the risk of misdiagnosis and delayed treatment [1]. Traditional methods, e.g., spirometry, require unnatural breathing manoeuvres and are dependent on patient compliance, making accurate asthma detection a difficult task [4]. In addition, spirometry alone has limited sensitivity for detecting asthma, with a recall rate of only 29% [5].

Recently, the focus of research has shifted to non-invasive asthma detection techniques, with a notable advance being the integration of ML algorithms [6-9]. The main reason for ML utilisation lies in its ability to significantly improve detection accuracy by recognising complex patterns in voice and breath data. Unlike traditional methods, ML algorithms have self-learning capabilities and the ability to adapt to different patterns and nuances, overcoming the limitations of rigid algorithms used in traditional approaches. Several ML-based asthma detection methods use spectral and temporal features, e.g., spectral contrast and mean amplitude (Mean Signal Strength) [8]. This approach utilises the RF algorithm, a binary classification model for predicting lung function from recorded voice sounds, which can be used in real-time asthma management applications. It was selected thanks to its low probability of overfitting. Even though it presents 85% accuracy and an 84% F1-score, the recall rate achieved is 44%, which was influenced by an uneven distribution of the sample and the low dimensionality of the dataset. Another approach used a different set of temporal features, upward expiration and derivatives of the entire expiratory phase via carbon dioxide (CO_2) signals [9]. SVM was selected due to the use of an optimum hyperplane to differentiate the asthmatic and non-asthmatic groups by maximising the distance from the decision boundaries. This approach obtained an accuracy of 94.52%, a recall of 97.67%, and a specificity of 90.0%. However, the signals are extracted from specialised hardware (HW), which is not widely accessible.

This paper presents a novel method for asthma detection using a ML algorithm, by analysing patients' audio clips. The method employs customised feature selection, using recursive feature elimination (RFE) and CFS techniques. RF, Logistics Regression (LR), KNN, SVM and Decision Tree (DT) ML models were created and evaluated on the selected features. The proposed method achieved performance superior to other

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00166

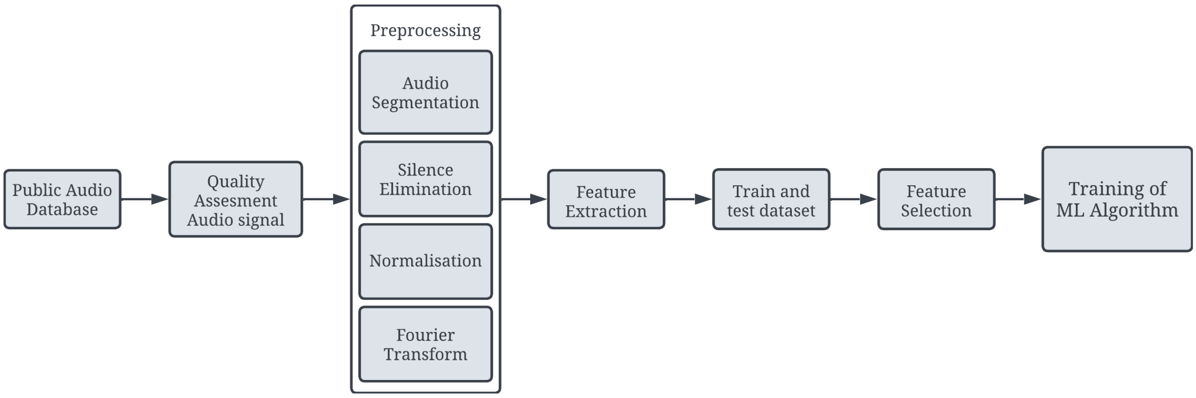

Fig. 1. The proposed algorithm's block diagram

methods: 92% accuracy and 83% recall with the RF algorithm. The H2O ML tool was employed for further fine-tuning of the model's hyperparameters and capturing intricate patterns in the data, making asthma detection more accurate. The proposed method can efficiently detect asthma even via smartphone audio without the need for specialised HW. This makes the proposed method accessible and scalable.

This paper is organised as follows: In Section II, we elaborate on the proposed method and introduce the database along with the evaluation criteria. The experimental setup and performance of the asthma detection algorithm are presented in Section III. Section IV discusses the performance of the proposed algorithm compared with other studies and possible improvements. Finally, Section V concludes this paper.

II. Materials and methods

This section discusses the proposed method's building blocks, i.e., dataset curation, signal preprocessing, feature extraction, feature selection, and ML model training.

The audio signal quality was assessed and preprocessed prior to the feature extraction stage. Following the extraction of 14 meaningful features from each audio type, the data were randomly divided into two sets. 80% of the data were used for training and validation, while 20% of the data were used for testing. Feature selection and hyperparameter optimisation techniques were applied to reduce both computational complexity and the likelihood of algorithm overfitting. Finally, the chosen features were employed in training ML algorithms by utilising data analytics platforms, e.g., Sklearn, Spotfire©, and H2O. Fig. 1 describes the overall block diagram of the proposed algorithm.

A. *Public Audio Database*

The Coswara database, a prominent resource in this field, is crucial for the development of a diagnostic tool to detect COVID-19 from audio recordings [10]. It comprises 2746 participants, including 1984 healthy individuals and 134 individuals with asthma.

B. *Quality Assessment Audio Signal*

Due to the limited number of subjects with asthma available, a meticulous selection of 49 participants from each asthmatic and healthy subgroup was made based on audio quality and annotation files, with a specific focus on the audio clips that were rated as excellent or good. The assessment included a detailed examination of various parameters, i.e., clarity, background noise, language and overall recording quality. In particular, this study considers audio clips with Vowel O (VO), Breathing Shallow (BS), Cough Heavy (CH), and Counting Normal (CN) recorded at a sampling frequency of 48 kHz.

C. *Audio Segmentation*

The raw audio signals were prepared through different preprocessing stages before feature extraction. Audio segmentation is an initial step in the preprocessing pipeline for raw audio signals. It serves as the foundation for subsequent feature extractions. It involves breaking the continuous audio signal into specific segments, each corresponding to distinct sound events. This includes a 3-second representation of the VO and a 3-second segment; the former provides insights into voice and airway functionality, while the latter serves to identify respiratory diseases from CH. Furthermore, a 5-second segment was used to detect shallow breathing patterns indicative of respiratory distress in BS. Additionally, a segment involving controlled vocalisation, specifically counting from 1 to 5, offered a structured assessment of vocal and respiratory skills. These segments collectively contributed to a nuanced analysis, thereby enriching the diagnostic potential of the model by capturing diverse aspects of respiratory health through targeted audio signal examination.

D. *Silence Elimination*

Following audio segmentation, the silence clip was eliminated at the beginning of audio clips by isolating background noises and unvoiced segments [11]. This was achieved by using the PyDub library, NumPy, and SciPy. The method identified the starting point of non-silent audio signals, thereby setting a threshold for absolute amplitude by trimming audio data. This enhanced the overall quality of the audio data for subsequent feature extraction.

E. *Normalisation*

Following silence elimination, the normalisation process was implemented to bring the audio clip to a consistent level of audio gain [12]. The PyDub library aided the process by manipulating audio data to create an "audio segment" from the audio clip. This segment serves to identify the maximum volume and adjust the overall audio proportionally by increasing the volume of the quieter sections.

F. *Fourier Transform*

Fast Fourier transform (FFT) was used to distinguish frequency bands in asthmatics, focusing on 0-200 Hz and 200-

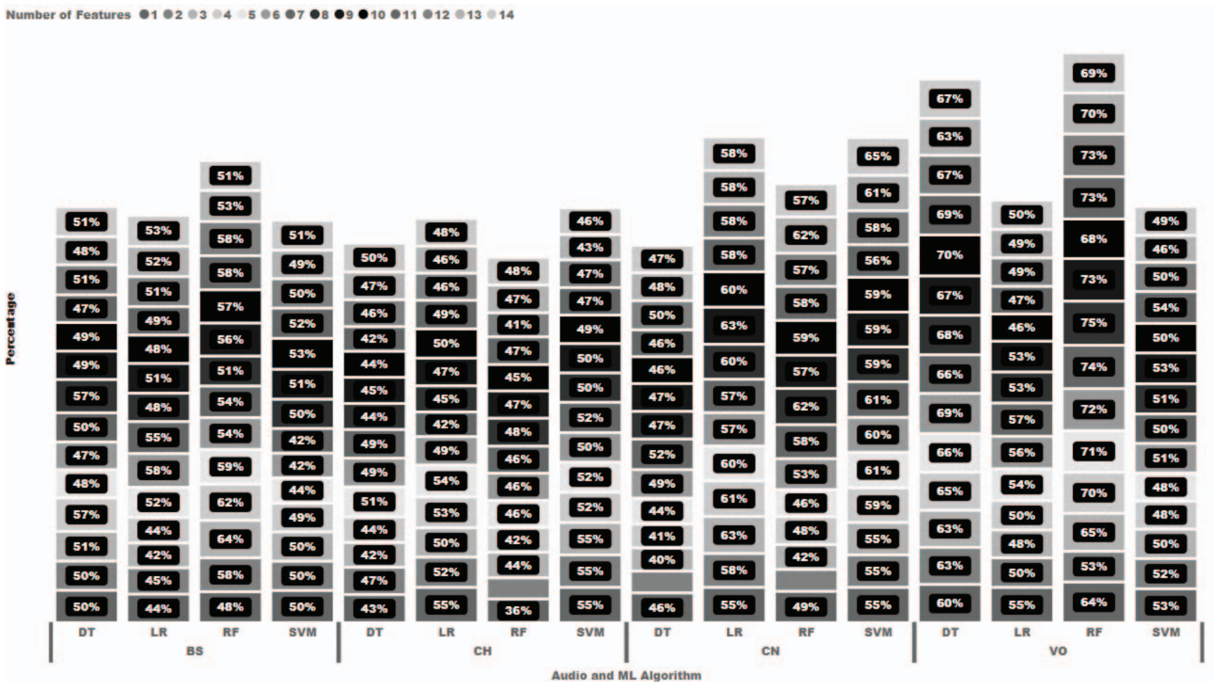

Fig. 2. RFE cross-validation accuracy for different algorithms and audio

400 Hz [13]. This shows that the signal strength is increased in asthma patients compared to healthy individuals. In addition, the Short Time Fourier transform (STFT) helped in feature by dividing the signal into short, overlapping windows to improve the detection of spectral variations in audio signals [14].

G. *Feature Extraction*

A total of 14 features were extracted from the audio signal, the STFT and the magnitude spectrum of the STFT.

The Spectral roll-off (SRO), Root Mean Square (RMS) energy, Zero Crossing Rate (ZCR), and Spectral Centroid (SC) were extracted from the audio signal using the librosa.feature library. SRO indicates the frequency below a certain percentage of the total spectral energy and is calculated at 25, 50, and 75 percentiles of 1% and 99% roll-off percentages. This method aims to characterise the spectral distribution at different energy levels and thus gain insight into the energy distribution within the signal, which is central to asthma-related patterns [8]. RMS energy quantifies the total energy content of the signals, providing valuable information about the overall signal strength. RMS is particularly important in asthma analysis, where variations in energy content may indicate changes in breathing patterns or symptom intensity [8]. ZCR indicates the rate of polarity change in the signal, which contributes to our understanding of signal dynamics that can be related to the respiratory fluctuations observed in asthma patients [8]. SC reveals the centroid of the spectral distribution, which helps to characterise the overall frequency distribution in the signal. This provides additional insight into potential asthma-related variations in the frequency domain [8].

The features derived from the STFT are the Mean Signal Strength and the derivative of the spectral slope (D-Slope). The Mean Signal Strength is calculated for two different frequency bands, which were explained in Section II F. It provides a detailed overview of the signal strength in specific frequency segments and highlights possible variations associated with asthma. D-Slope represents the rate of change of the spectral slope and provides information on how quickly the frequency components change, which helps to detect subtle variations in asthma-related patterns [15].

Finally, the band energy (BE) ratio, extracted from the magnitude spectrum of the STFT, was analysed in the two frequency bands. This feature provides information on the distribution of energy within these segments and provides valuable information on how the energy is distributed across different frequency ranges in relation to asthma characteristics [15].

H. *Train and Test block*

Following preprocessing and feature extraction, the dataset were split into a training set (80%) and a test set (20%). The training dataset were used to train ML algorithms to learn patterns and relationships indicative of asthma. The test set evaluated the model's ability. A random state of 42 ensured consistent data distribution and reproducibility.

I. *Feature Selection algorithms*

Feature selection is used to reduce the risk of overfitting the algorithms. Two feature selection methods were used in the proposed method: RFE and CFS. RFE is a feature selection algorithm that reduces dimensionality by removing unwanted features from the model. It is combined with cross-validation, which aids in determining the best number of features to retain [16]. CFS is employed to assess whether a feature is highly correlated with the class, i.e., asthma, and not with any of the other features [17]. These algorithms reduce redundancy and improve the overall efficiency of ML models, which provides a strong foundation for the proposed methods and goals.

J. ML algorithms

TABLE I. FEATURES SELECTED BY RFE AND CFS FOR DIFFERENT ML ALGORITHMS AND AUDIO

Feature Selection Algorithms	ML Alogrithms	VO	BS	CN	CH
RFE	RF	VO (200-400Hz) VO ZCR VO SC VO 25th SRO 99% VO 50th SRO 99% VO BE 0-200 VO BE 200-400 VO D-Slope	BS RMS Energy BS BE 0-200 BS D-Slope	CN (0-200Hz) CN (200-400Hz) CN RMS Energy CN ZCR CN SC CN 25th SRO 99% CN 75th SRO 99% CN D-Slope	CH BE 0-200 CH BE 200-400
	SVM	VO (0-200Hz) VO (200-400Hz) VO RMS Enegry VO ZCR VO 75th SRO 99% VO 25th SRO 1% VO 50th SRO 1% VO 75th SRO 1% VO BE 0-200 VO BE 200-400 VO D-Slope	BS (0-200Hz) BS (200-400Hz) BS RMS Energy BS ZCR BS 25th SRO 1% BS 50th SRO 1% BS 75th SRO 1% BS BE 0-200 BS BE 200-400 BS D-Slope	CN (0-200Hz) CN (200-400Hz) CN RMS Energy CN ZCR CN SC CN 25th SRO 99% CN 50th SRO 99% CN 75th SRO 99% CN 25th SRO 1% CN 50th SRO 1% CN 75th SRO 1% CN BE 0-200 CN BE 200-400 CN D-Slope	CH ZCR
	LR	VO (200-400Hz) VO (0-200Hz) VO RMS Energy VO BE 200-400 VO BE 0-200 VO 75% SRO 1% VO 50% SRO 1%	BS (200-400Hz) BS 50% SRO 1% BS SC BS BE 0-200 BS (0-200Hz) BS 25% SRO 1%	CN (200-400Hz) CN BE 0-200 CN 50% SRO 1% CN SC CN 75% SRO 1% CN 25% SRO 1% CN BE 200-400 CN RMS Energy CN (0-200Hz)	CH BE 200-400
	DT	VO (200-400Hz) VO SC VO 25th SRO 99% VO 50th SRO 99% VO 25th SRO 1% VO 50th SRO 1% VO 75th SRO 1% VO BE 0-200 VO BE 200-400 VO D-Slope	BS RMS Energy BS 50th SRO 99% BS BE 0-200 BS D-Slope	CN (0-200Hz) CN RMS Energy CN ZCR CN SC CN 25th SRO 99% CN 75th SRO 99% CN D-Slope	CH (0-200Hz) CH ZCR CH 25th SRO 99% CH BE 200-400 CH D-Slope
CFS	RF SVM LR DT KNN	VO D-Slope	BS (200-400Hz) BS RMS Energy BS BE 0-200 BS BE 200-400 BS D-Slope	CN 75th SRO 1% CN (200-400Hz) CN BE 0-200 CN 75% SRO 99% CN D-Slope	CH RMS Energy CH 75th SRO 99% CH 75th SRO 1% CN BE 200-400 CN D-Slope CH BE 0-200

Following feature selection, the features were trained with ML algorithms from Sklearn, H2O, and Spotfire© for asthma detection on the training set. Sklearn is a Python library for ML and predictive modelling, while H2O is a tool in Python for building ML models with a focus on scalability and performance. Spotfire© is integrated with ML tools for predictive analytics [18]. Five algorithms, i.e., RF, LR, KNN, SVM and DT, were tested using 10-fold cross-validation and grid search as hyperparameter optimizers to improve accuracy [19]. They contribute by recognising patterns in the data and accurately classifying samples as asthmatic or healthy. Each algorithm capitalises on different methodologies to complete this classification task. RF, known for its effectiveness in processing complicated patterns, excels at combining multiple classifiers to accurately recognise asthma features [20]. LR is used for two-class classification and facilitates accurate identification of asthmatic and non-asthmatic samples based on characteristic features [21]. KNN, a non-parametric method, contributes to asthma detection by identifying the k-nearest neighbours for a given data point and improving prediction accuracy by using majority classes [22]. SVM defines a hyperplane to maximise the separation between asthmatic and healthy samples and ensure a clear class distinction for accurate detection [22]. DT identifies key attributes indicative of asthma by recursively subdividing the data based on features and iteratively refining the classification process for accurate asthma detection [22].

By assessing the accuracy of the chosen algorithm and audio during each iteration of the recursive elimination in RFE, the significance of the number of features to be selected is evaluated. The process is illustrated in Fig. 2.

The RFE method ultimately identifies a set of features, which are presented in Table I. This table serves as a comprehensive list, encapsulating the features selected by the RFE method to achieve optimal results with specific ML algorithms and audio. For instance, when employing the RFE method with the SVM ML algorithm, it can eliminate up to 13 less significant features to achieve a more reliable model using CH audio. However, when the same method and algorithm is employed to CN audio, all 14 features are utilised. Therefore, the RFE method identifies up to 13 less significant features across different settings from audio signal types and ML algorithm used.

K. Evaluation Criteria

In evaluating the performance of ML algorithms in asthma detection on the testing set, 6 criteria were adhered to [23]. Accuracy measures the ability of the algorithm to correctly identify both asthmatic and non-asthmatic cases. Precision measures the correctness of identifying true positive cases among all positive predictions. Recall indicates the effectiveness of the model in detecting as many true positive instances as possible, minimising the risk of overlooking asthmatic individuals. Specificity evaluates the algorithm's ability to accurately identify non-asthma cases among all true negative cases. The F1 score, a harmonic mean of Precision and Recall, indicates a balanced compromise between the two values and ensures a robust and reliable algorithm for processing unseen data. Finally, the Area under the Curve (AUC) quantifies the model's ability to distinguish between positive and negative instances. It provides insight into the overall discriminative power and predictive performance of the model at different thresholds, improving our understanding of its effectiveness in classifying asthma cases.

III. RESULTS

This section provides a comprehensive overview of the performance of 21 ML algorithms in all four audio clips. Out of which there are 5 ML algorithms with Sklearn using the CFS algorithm, 4 ML algorithms with Sklearn using the RFE algorithm, 5 ML algorithms with Sklearn, 3 ML algorithms with H20 and 4 ML algorithms with Spotfire© using the full set of 14 features.

TABLE II. MEAN EVALUATION MATRIX FOR MODEL USING FOUR AUDIO SIGNALS

Audio	Accuracy	Precision	Recall	Specificity	F1 Score	AUC
BS	58.67%	52.44%	62.62%	60.52%	52.32%	62.18%
CH	49.29%	68.12%	44.28%	64.49%	45.08%	50.82%
CN	44.24%	67.21%	38.22%	61.09%	42.48%	46.14%
VO	68.76%	74.87%	52.19%	80.90%	59.21%	75.91%

A. Audio performance

Table II shows the mean performance matrix for four audio signals across various ML algorithms and tools. Notably, VO consistently outperformed other audio types across most evaluation metrics. With an accuracy of 68.76% and precision of 74.87%, VO demonstrated reliability in predicting unseen data. Additionally, its AUC of 75.91% indicates high predictive performance, as discussed in Section II K which can be attributed to its monosyllabic nature. However, recall achieved the second-highest score at 52.19%.

The primary variation in VO lies in the decibel levels among different subjects, resulting in consistently high accuracy. In contrast, the other audio clips are influenced by two additional factors: the rate of speech ,i.e., fast or slow, and the presence of silence, in addition to the decibel levels.

B. Performance of feature selection algorithms

TABLE III. MEAN EVALUATION MATRIX FOR MODEL USING RFE AND CFS FROM SKLEARN

Features Selection Algorithm	Accuracy	Precision	Recall	Specificity	F1 Score	AUC
CFS	55.75%	77.95%	48.09%	70.45%	56.45%	64.09%
RFE	48.13%	59.19%	40.08%	56.01%	43.63%	53.49%

When evaluating the RFE and CFS feature selection algorithms based on mean performance in Table III, CFS consistently outperformed RFE by at least 7%. Which can be explained by the recursive elimination of RFE, which inadvertently excludes potentially significant features early in the process, preventing an improvement in model performance. The sensitivity of RFE to the order of feature elimination is evident in its tendency to discard important features prematurely. While CFS proved useful specifically in our study, its effectiveness may not be generalisable to other studies with a different set of features.

C. Tools Review

TABLE IV. MEAN EVALUATION MATRIX FOR MODEL USING H2O, SKLEARN AND SPOTFIRE©

Platform	Accuracy	Precision	Recall	Specificity	F1 Score	AUC
H2O	58.75%	71.58%	62.75%	83.79%	48.91%	59.15%
SKLearn	51.88%	66.64%	43.87%	63.04%	49.03%	58.68%
Spotfire©	64.38%	57.81%	58.38%	66.94%	53.00%	

Table IV shows that H2O outperformed both Spotfire© and Sklearn, especially in precision (71.58%) and specificity of 83.79%. Although the accuracy and F1 score are slightly lower compared to Spotfire© and Sklearn by at most 5%, H2O achieved high precision and specificity by at most 16%. High precision is crucial in medical diagnosis where false positives can lead to unnecessary treatments or interventions. This shows that H2O is best suited for our selected features. The platform's accessibility also benefits users with different ML skills. It enables efficient exploration of different models and facilitates the identification of the best-performing algorithms without extensive manual fine-tuning. This reduces the time needed to identify effective ML algorithms before customising them to specific requirements.

D. Top-Performing Models

Three models stand out as the best performers when comparing audio performance, feature selection algorithms and tools. The best model, which used VO audio with all 14 features and employed the RF algorithm in the H2O tool, achieved recall rates of 83%, precision rates of 95% and F1 scores of 73%. The SVM and KNN algorithms in the Sklearn tool using CFS achieved recall rates of 76% and 71% with precision rates of 83% and F1 scores of 77%, respectively.

The three most important features, used in the best model, were 75th SRO 99%, BE 200-400 and 50th SRO 99%, with importance scores of 26%, 8% and 7% respectively. These features in the frequency band are aligned with conditions such as airway narrowing and inflammation in asthma patients, exhibiting a high correlation [19]. This emphasises the importance of analysing specific frequency components, especially in medical applications where the identified features can provide valuable insights into physiological processes. This can help in the diagnosis and monitoring of asthma.

On the other hand, the follow-up best models used only one feature, D-slope. Therefore, there is no comparison of the importance of the features in this model. Moreover, KNN algorithm has no feature importance as it treats all features equally in terms of their contribution to the decision process and relies on measuring the distance between data points to make predictions.

IV. DISCUSSION AND LIMITATIONS

A. Discussion

The proposed method was compared to two other prominent asthma detection studies, as presented in Table V.

The proposed method achieved a recall rate of 83%, which is 39% higher than RF [8], although similar spectral and temporal features as well as RF algorithm were used. This is due to the balanced distribution of the classes and the feature selection methods used. As explained in Section II B, we balanced the dataset to prevent a bias in favour of the majority class and to improve the model's ability to detect instances of the minority class. This is crucial when dealing with underrepresented target classes. Additionally, the proposed method applied feature selection methods to reduce feature redundancy in the model, despite having 9 fewer features than in [8]. Therefore, the notable patterns were captured and the performance was improved, especially in terms of recall.

Furthermore, the proposed method achieved an accuracy of 92%, which is 2.52% lower than [9]. This can be explained due to the feature types and method. As explained in Section II G, we extracted the spectral and temporal features from audio signals which can be readily extracted from smartphones. This approach eliminates the need for dedicated HW, making it accessible. In addition, the proposed RF algorithm provides a more effective way in handling complex datasets compared to using a decision boundary in SVM [9], yet both approaches serve the common purpose of detecting asthmatic patients. Despite this diversity, the proposed method demonstrated that the collection of relevant data for asthma detection is practical and easily accessible.

In summary, the proposed method prioritises efficiency, scalability, and user-friendliness. It utilises a ML approach using recorded audio data from smartphones, which is specifically tailored to detect asthma with minimal individual effort. Importantly, our system does not require any specialised HW, opting instead for a scalable, straightforward solution that uses the convenience of a simple smartphone while delivering superior performance.

TABLE V. COMPARISON OF ASTHMA DETECTION PERFORMANCE

Reference	Accuracy	Precision	Recall	Specificity	F1 Score	AUC	Features Type	# Features	ML Algorithms
This Study	92.00%	95.00%	83.00%	95.00%	73.00%	92.31%	Spectral, Temporal	14	RF
[8]	85.00%		44.00%		84.00%	88.00%	Spectral, Temporal	23	RF
[9]	94.52%	94.60%	97.76%	90.00%			Physiological	3	SVM

B. Limitations

The data used to report on the performance of the different algorithms are not consistent across the various studies. Differences in signal acquisition methods, recording conditions and types of signals affect performance. In addition, the use of different types of datasets addressing the same asthma detection problem makes direct comparison difficult. Despite these challenges, the results reported in our study show that the proposed method can be used for asthma detection with minimal effort. In addition, the dependence on the Coswara database, which mainly contains participants of Indian origin and the gaps in the training data, especially the lack of subjects mimicking asthma conditions, need to be considered. As our next step, we plan to diversify the dataset by including databases from different ethnic backgrounds, including the Coswara COVID-19 dataset, to ensure a more representative training set. We would also like to improve the exploration of breathing patterns by including different vowel sounds and audio types in the Coswara dataset.

V. CONCLUSION

A novel method for asthma detection based on ML and patient audio clip analysis is presented in this paper. The main novelty of the proposed method lies in its use of various performance metrics, ensuring a comprehensive and quantitative assessment of asthma detection. Another innovative aspect is the integration of existing data analytics tools, resulting in improved efficiency and effectiveness of the proposed model. The proposed approach achieves an accuracy of 92% with the RF algorithm, superior compared to existing methods. Furthermore, by balancing the dataset and selecting features, the proposed method achieves a recall rate of 83%. The user-friendly, HW-free design positions it as a scalable and accessible solution for smartphone-based asthma detection. This study represents a significant advancement in the field of health applications, particularly in non-invasive asthma diagnosis, as it recognises various limitations and outlines plans for dataset diversification. Future expansion involves developing a mobile application that would be accessible and serves as an early detection mechanism to prevent serious health complications through timely intervention. This comprehensive approach not only addresses the current limitations in asthma diagnosis, but also sets the stage for practical, accessible and proactive healthcare solutions.

REFERENCES

[1] Cleveland Clinic, “Asthma,” Cleveland Clinic, 2022. [Online]. Available: https://my.clevelandclinic.org/health/diseases/6424-asthma.

[2] M. R. Partridge, “Examining the Unmet Need in Adults with Severe Asthma,” European Respiratory Review, vol. 16, no. 104, pp. 67–72, Sep. 2007.

[3] J. Kavanagh, D. J. Jackson, and B. D. Kent, “Over- and under-diagnosis in Asthma,” Breathe, vol. 15, no. 1, pp. e20–e27, Mar. 2019.

[4] M. A. Almeshari, J. Stockley, and E. Sapey, “The Diagnosis of asthma. Can Physiological Tests of Small Airways Function help?,” Chronic Respiratory Disease, vol. 18, p. 147997312110533, Jan. 2021.

[5] Schneider et al., “Diagnostic Accuracy of Spirometry in Primary Care,” BMC Pulmonary Medicine, vol. 9, no. 1, Jul. 2009.

[6] Md. A. Awal et al., “An Early Detection of Asthma Using BOMLA Detector,” IEEE Access, vol. 9, pp. 58403–58420, 2021.

[7] H. Joo, D. Lee, Sang Haak Lee, Young Kyoon Kim, and Chin Kook Rhee, “Increasing the Accuracy of the Asthma Diagnosis Using an Operational Definition for Asthma and a Machine Learning Method,” BMC Pulmonary Medicine, vol. 23, no. 1, Jun. 2023.

[8] Md. Z. Alam et al., “Predicting Pulmonary Function from the Analysis of Voice: a Machine Learning Approach,” Frontiers in Digital Health, vol. 4, Feb. 2022.

[9] P. Singh, R. Palaniappan, and M. Malarvili, “Automatic Quantitative Analysis of Human Respired Carbon Dioxide Waveform for Asthma and Non-Asthma Classification Using Support Vector Machine,” IEEE Access, vol. 6, pp. 55245–55256, 2018.

[10] N. Sharma et al., “Coswara -- A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis,” Interspeech 2020, pp. 4811–4815, Oct. 2020.

[11] Tushar Ranjan Sahoo and S. Patra, “Silence Removal and Endpoint Detection of Speech Signal for Text Independent Speaker Identification,” International Journal of Image, Graphics and Signal Processing, vol. 6, no. 6, p. 27, 2014, Accessed: Dec. 09, 2023. [Online]. Available: https://www.mecs-press.org/ijigsp/ijigsp-v6-n6/v6n6-4.html.

[12] Tshephisho Sefara, “The Effects of Normalisation Methods on Speech Emotion Recognition,” Nov. 2019.

[13] T. Shimoda et al., “Lung Sound Analysis Helps Localize Airway Inflammation in Patients with Bronchial Asthma,” Journal of Asthma and Allergy, vol. Volume10, pp. 99–108, Mar. 2017.

[14] N. Kehtarnavaz, “Frequency Domain Processing,” Digital Signal Processing System Design, pp. 175–196, 2008.

[15] E.0 Babaee, N. B. Anuar, A. W. Abdul Wahab, S. Shamshirband, and A. T. Chronopoulos, “An Overview of Audio Event Detection Methods from Feature Extraction to Classification,” Applied Artificial Intelligence, vol. 31, no. 9–10, pp. 661–714, Nov. 2017.

[16] G. Sharma, K. Umapathy, and S. Krishnan, “Trends in Audio Signal Feature Extraction Methods,” Applied Acoustics, vol. 158, p. 107020, Jan. 2020.

[17] M. Hall, “Feature Selection for Discrete and Numeric Class Machine Learning,” Apr. 1999.

[18] “Connected Intelligence Cloud,” *TIBCO Cloud*^{TM}. https://cloud.tibco.com Accessed: Jan. 06, 2024.

[19] M. Adnan, A. A. S. Alarood, M. I. Uddin, and I. ur Rehman, “Utilizing grid search cross-validation with adaptive boosting for augmenting performance of machine learning models,” PeerJ Computer Science, vol. 8, p. e803, Feb. 2022.

[20]] H. Dabiri, V. Farhangi, M. J. Moradi, M. Zadehmohamad, and M. Karakouzian, “Applications of Decision Tree and Random Forest as Tree-Based Machine Learning Techniques for Analyzing the Ultimate Strain of Spliced and Non-Spliced Reinforcement Bars,” Applied Sciences, vol. 12, no. 10, p. 4851, May 2022.

[21] M. B. Korsós, R. Erdélyi, J. Liu, and H. Morgan, “Testing and Validating Two Morphological Flare Predictors by Logistic Regression Machine Learning,” arXiv (Cornell University), Dc. 2020.

[22]] S. Wiyono, D. S. Wibowo, M. F. Hidayatullah, and D. Dairoh, “Comparative Study of KNN, SVM and Decision Tree Algorithm for Student’s Performance Prediction,” International Journal of Computing Science and Applied Mathematics, vol. 6, no. 2, p. 50, Aug. 2020.

[23] B. Anderson, “The CCB-ID approach to tree species mapping with airborne imaging spectroscopy,” PeerJ, vol. 6, p. e5666, Oct. 2020.