2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Role of breath phase and breath boundaries for the classification between asthmatic and healthy subjects

Shivani Yadav[1], Dipanjan Gope[2], Uma Maheswari K.[3], Prasanta Kumar Ghosh[4]

Abstract— **Asthma is an inflammatory disease of the airways which causes cough, chest tightness, wheezing and other distinct sounds during breathing. Spirometry is a golden standard lung function test, is used to monitor and diagnose asthma. Spirometry is very time-consuming and requires a lot of patient's efforts. Therefore, an alternate method which can overcome spirometry limitations is required. Sound based method can be one such alternative as it is less tedious, less time consuming and suitable for patients of all ages. It has been shown in the past that breath, among other vocal sounds, performs the best for an asthma vs healthy subject classification task. Breath consists of two phases, namely, inhale and exhale. Experiments in this work show, exhale performs better for classification task compared to the entire breath cycle as well as the inhale. However, this requires manual marking of the breath boundaries, which is a very time-consuming task. We, in this work, investigate how critical are the breath cycle and breath phase boundaries for the classification task. Experiments with chunks of random duration shows that they perform on par or better than the exhale. However, a segment comprising the second and third quarters of a breath cycle results in the highest classification accuracy of 80.64%. This suggests that, while breath phase boundaries may not be important, breath cycle boundaries could benefit in the classification task.**

I. INTRODUCTION

Asthma is an obstructive lung disease that causes cough, wheeze, chest tightness, and other peculiar symptoms [1]. Around 339 million people are suffering from asthma worldwide, including 1000 deaths every day [2]. Spirometry is a golden standard lung function test used to monitor and diagnose asthma. During the test, a subject has to take deep inhalation followed by forced exhalation for at least 6 seconds into a mouthpiece attached with the spirometer. Throughout the test, the subject's nose is closed with the nose clip. Spirometer measures forced vital capacity (FVC), forced expiratory volume in 1 sec (FEV1) and FEV1/FVC, and other lung function parameters. Measured values of these variables are compared with the reference values of spirometry variables. Reference values are calculated based on the age, height, weight, and gender of the subject-discrepancies between reference and measured values of spirometry variables are used to diagnose and determine the severity of asthma. However, spirometry is very tedious and time-consuming, especially for older people and kids [3]. Another technique to monitor asthma is the peak flow meter (PFM), which is simpler, easy, and less time-consuming than spirometry. PFM measures peak expiratory flow rate from major airways but fails to do so through minor airways, which can also get affected during asthma [4]. Therefore, the need of the time is an alternate method for asthma diagnosis and monitoring. The sound-based method can be one such method because it is less strenuous and time-consuming.

In the literature, wheeze sound has been used for the classification and monitoring of asthma. Wheeze is a whistling sound produced during breathing due to obstruction of the airways. Wheeze can occur during inhalation, exhalation, or during entire breathing. Wisniewski et al. [5] have used tonality index and spectral envelope of wheeze sound for asthma monitoring. Few works have used duration of the breath [6], dominant frequency range [7], pitch [8], wavelet features [9] for asthma classification. One of our research interests is sound-based monitoring and diagnosis of asthma. To address this problem, we have, previously, reported few related results [10], [11] and [4]. Rao et al. [4] have predicted spirometry variables using statistical spectrum descriptors to determine asthma severity using cough and wheeze. Yadav et al. [10], [11] performed asthma and healthy subject classification using wheeze, cough, and sustained phonations. Experiments [10] have shown that the wheeze (referred to as breath in this paper) is the best sound for the classification between asthmatic and healthy subjects. However, the role of inhale and exhale (referred to as breath phases) has not been separately analyzed for the classification task in [10]. In all our earlier works [10], [11], [4], breath boundaries have been marked manually by listening and visual inspection of spectrogram of breath sound, which is a very time-consuming task. Therefore, the motivation of our current work is to address two key questions: 1. Do breath phases contribute equally towards the classification between asthmatic and healthy subjects? 2. In a recording of continuous breath cycles are exact locations of breath boundaries required for the classification task? In this work, we answer these questions using a classification experiment setup identical to that in [10]. Experiments with 48 healthy subjects and 45 asthmatic patients reveal that exhale is better than both inhale and complete breath cycle, with an average classification accuracy of 75.15%($\pm$ 9.29%). Experiments with random chunks reveal that five chunks of duration 2.2 secs picked randomly from a recording of the continuous breath cycles

[1]Shivani Yadav is with BioSystems Science and Engineering, Indian Institute of Science (IISc), Bangalore-560012, India shivaniyadav@iisc.ac.in

[2] Dipanjan Gope is a faculty of Electrical Communication Engineering, Indian Institute of Science (IISc), Bangalore-560012, India dipanjan@iisc.ac.in

[3] Uma Maheswari is a doctor in Pulmonary Medicine, St. Johns National Academy of Health Sciences, Bangalore-560034, India umamohan99@gmail.com

[4] Prasanta Kumar Ghosh is a faculty in Electrical Engineering, Indian Institute of Science (IISc), Bangalore-560012, India prasantg@iisc.ac.in

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

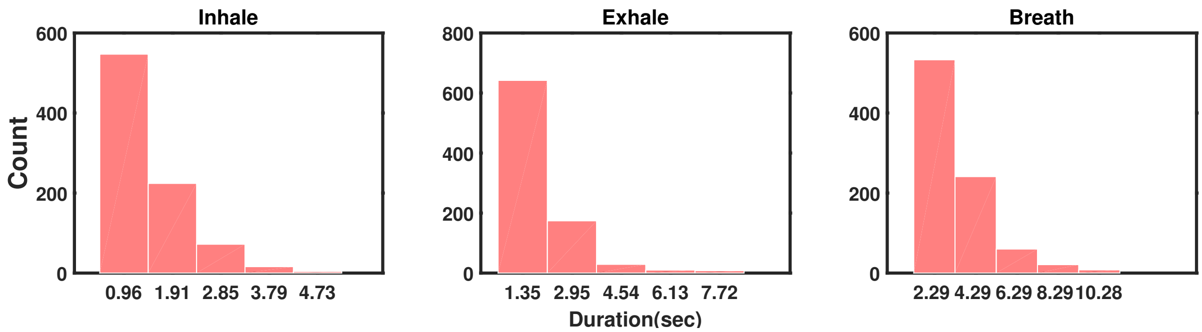

Fig. 1. Histogram of manually annotated inhale, exhale and entire breath cycle. X-axis indicates duration in secs and Y-axis denotes the total count.

are good enough for achieving a classification performance comparable to that using exhale. This suggests that exact breath phase boundaries are not critical for the classification task. However, a segment comprising the second and third quarters of a breath cycle results in the highest classification accuracy of 80.64%, indicating that the breath cycle boundaries, unlike breath phase boundaries, could benefit in the classification task.

II. Dataset

A total of 93 subjects is used in this work, out of which 45 (26 male, 19 female) are asthmatic patients, and 48 (24 female, 24 male) are healthy controls. All data have been recorded in the St. Johns medical college hospital, Bangalore, India under the guidance of the doctor. The maximum and minimum age of the patients is 71 years and 15 years, respectively, with an average age of 43.42($\pm$ 13.86) years. Similarly, the average age of controls is 36.18($\pm$ 11.22) years, with minimum and maximum age being 19 years and 60 years, respectively. From each subject, breath signals are recorded in a typical noisy environment of the hospital. Each subject signs a consent form before the recording. The breath signal is recorded with a ZOOM H6 handy recorder at a sampling rate of 44.1kHz. Subjects are instructed to take deep breaths while sitting. During the recording, microphone is kept at around 5 cm in front of the mouth. Throughout the recording, a nose clip is used to close the nose of a subject so they can breathe up to their total capacity through the mouth only. On average, nine breath cycles per subject are recorded. Total 863 breath cycles are recorded, out of which 450 are from healthy subjects and 413 are from patients. Histograms of inhale, exhale and breath duration are given in Fig. 1. For inhale, the average duration is 1.45($\pm$ 0.68) secs with a maximum and minimum duration of 5.20 secs to 0.49 secs as shown in Fig. 1. Average duration of exhale is 1.86($\pm$ 1.06) secs with a range of duration being 0.55 secs to 8.513 secs. The average breath cycle duration is 3.37($\pm$ 1.54) secs with a duration range of 1.28 secs to 11.28 secs as shown in Fig. 1. Breath phase boundaries are manually marked through visual inspection of the spectrogram and listening of the breath sound signal waveform through Audacity [12].

III. Methodology

To the best of our knowledge, no other works in the literature, except by Yadav et al. [10], reported classification of asthmatic and healthy subjects using vocal breath sounds. Hence, in this work, we have used the features, classifier, and evaluation measure used in [10]. During our previous work reported in [11], [4], [10], it has been observed that manual marking of breath phase boundaries is a tedious task. An alternate approach could be the development of automatic breath phase boundary segmentation methods. However, we need to examine how critical is the requirement of exact breath phase and breath boundaries for the classification task at hand. That, in turn, would determine the required accuracy for a breath segmentation approach. For this purpose, we investigate the significance of breath and breath phase boundaries for the classification between asthmatic patients and healthy subjects. For comparing the entire breath cycle and breath phases, support vector machine (SVM) classifiers are trained for each of them, and classification performance is reported. We, in this work, assume that the breath sounds (referred to as $B[n]$, where n denotes sample index) are continuously recorded without any pause or silence. From $B[n]$, n_r many chunks of duration, d_r are selected randomly. For each chunk of duration d_r, a feature vector comprising features given in [10] is computed. Therefore for every $B[n]$, n_r feature vectors are computed. A support vector machine is used for the classification task.

TABLE I

Mean(standard deviation)(%) of classification accuracy, recall and specificity using inhale, exhale and breath using manually annotated boundaries.

	Inhale	Exhale	Breath
Mean(SD)	60.18(5.02)	75.15(9.29)	71.93(6.45)
Mean Recall(SD)	60.2(14.0)	70.4(17.4)	68.4(13.3)
Mean Specificity(SD)	60.0(6.0)	80.0(4.9)	75.5(9.3)

IV. Experimental Setup

12 MFCC coefficients have been calculated by using a window length of 20 msec with 10 msec shift. Six statistics of MFCC, namely, mean, median, mode, variance, standard deviation, and root mean square error, are calculated. Feature extraction is done using Voicebox toolkit [13] and SVM

TABLE II
MEAN(SD)(%) OF CLASSIFICATION ACCURACY USING SEGMENTS OF A BREATH CYCLE, NAMELY B_{0-25}, B_{25-75} AND B_{75-100}.

	B_{0-25}	B_{25-75}	B_{75-100}
Mean(SD)	62.34(10.67)	80.64(8.82)	76.37(8.69)

is used from LIBSVM-3.21 toolkit [14]. Feature extraction, SVM training, and testing are done in MATLAB. Another experiment is done to understand which region of breath provides more information for the classification task apart from inhale and exhale. For this task, each breath cycle is divided into 3 parts, which are referred as B_{0-25}, B_{25-75} and B_{75-100}. B_{0-25} indicate initial 25% of the breath cycle, B_{25-75} denote next 25% to 75% of the breath cycle duration and B_{75-100} denotes last 25% duration of the breath. Classification task is performed separately for each B_{0-25}, B_{25-75} and B_{75-100} and performance is compared. d_r varies from 1 sec to 5 sec with a step size of 0.4 sec, including average breath, inhale and exhale duration as shown in Fig. 1. This range is selected based on the mean and standard deviation of breath cycle, inhale and exhale duration as shown in 1. d_r also set to frame duration corresponding to the minimum and maximum breathing rate, which are 15 breaths per minute and 20 breaths per minute [15] that is 4 secs and 3 secs, respectively. n_r varies from 5 to 30 with a step size of 5. $n_r = 1$ is also used in this work. Five-fold setup have been used. Each fold contains 9 patients. Each of two among five folds contains 9 controls, and each of the remaining three folds has 10 controls. SVM hyper-parameters are tuned using a grid search using a five-fold cross-validation within the training set using radial basis function. Grid search is performed for $\log_2$C and $\log_2\gamma$ in the range of -6 to 8 and -1 to 20, respectively, with the step size of 1. Mean classification accuracy, precision, and recall are used as the evaluation metrics [16]. As n_r number of chunks are picked randomly, classification is highly dependent on the locations of these chunks in the continuous recordings of the breath cycles. Certain choices of the locations may results in high classification accuracy while other choices may not. Therefore for every combination of n_r and d_r, an experiment is repeated eight times to obtain a mean accuracy. By eight-time repetition, each fold will have eight accuracy values, whose average is computed. The mean and standard deviation of mean accuracy across eight repetitions is shown in Table III.

V. RESULTS

Results are presented in 3 sub-sections. The first sub-section describes the classification performance using inhale, exhale, and entire breath cycle. In the second sub-section, the classification experiment is done using B_{0-25}, B_{25-75}, and B_{75-100}. Last sub-section describes the classification performance by different combinations of n_r and d_r. The last two sub-sections help in understanding the role of breath boundary for the classification task.

A. Comparison of breath and breath phase for classification

Mean classification accuracy, mean recall and mean specificity using inhale, exhale, and entire breath cycle are given in Table I. From the first row of Table I, it can be observed that exhale performs the best in terms of the mean classification accuracy, which is (75.15%) followed by the entire breath cycle, which results in a mean classification accuracy of 71.93%. However, the SD (6.45%) using the whole breath cycle is lower than that (9.29%) using the exhale. Inhale performs the worst among all. It can be observed that specificity is maximum in exhale, 80%(± 4.9%), whereas, using the entire breath signal, it is 75.5%(± 9.3%). Similarly, mean recall using exhale and entire breath cycle are 70.4%(± 17.4%) and 68.4%(± 13.3%), respectively. Interestingly, the mean recall is lower in inhale, exhale, and entire breath than specificity. It shows proposed features are better in predicting asthmatic patients compared to healthy controls.

B. Significance of B_{0-25}, B_{25-75} and B_{75-100} in a breath

From the Table II, we observed that B_{25-75} performs the best among all three with a mean(SD) of 80.64%(± 8.82%), whereas B_{0-25} is performed worst among all with a mean(SD) of 62.34%(± 10.67%). Interestingly, as inhale duration is generally shorter than that of the exhale, major part of B_{25-75} would be contributed by exhale compared to inhale. Nevertheless, even performance of B_{25-75} is better than exhale, which is 75.15%(± 9.29%) (as given in Table I), which indicates a transition from inhale to exhale provides most of the cues for the classification task. It is also interesting to observe that not all parts of the exhale contribute equally to the classification task. B_{0-25} performs the worst among three parts considered. Interestingly, B_{0-25} is contributed mostly by the inhale phase of the breathing cycle. From this experiment, it can be concluded that 2^{nd} to 3^{rd} quarter of the entire breath cycle carries most of the information for the classification.

C. Role of random d_r and n_r on the classification accuracy

This experiment compares the classification performance obtained from breath cycle segments obtained using manually annotated boundaries, e.g. , inhale, exhale, and complete breath cycle vs. randomly chosen chunks from the breath recording. Results of this experiment are shown in Table III. From Table III, it can be seen that for every n_r (except for $n_r = 1$) and d_r, performance is always greater than 71.54%, identical to the accuracy obtained using manually annotated breath sound which is 71.93%. For $n_r = 1$, underfitting is observed. Underfitting occurs due to a very less number of representative data points from both asthmatic and healthy subject class. In this case, all subjects are classified as patients which leads to as mean classification accuracy of 48%. The best performance of 77.21(± 7.56)% is observed with $d_r = 2.2$ secs and $n_r = 5$. We observed that for a n_r between 5 to 25, high accuracy is obtained for d_r, 220 to 300 frames (highlighted as the blue region in the Table III), which is greater than the average exhale duration but less than the average breath duration (shown in colour in Table

TABLE III

MEAN(SD)(%) OF CLASSIFICATION ACCURACY BETWEEN ASTHMATIC AND HEALTHY SUBJECTS USING VARIOUS d_r AND n_r. IN TABLE d_r CORRESPONDING TO THE AVERAGE DURATION OF INHALE, EXHALE AND BREATH CYCLE IS SHOWN IN COLOUR WHICH IS 145, 186 AND 337 FRAMES, RESPECTIVELY. d_r ESTIMATED FROM THE MAXIMUM AND MINIMUM BREATHING RATE IS SHOWN IN COLOUR WHICH IS 300 AND 400 FRAMES, RESPECTIVELY.

Number of chunks	**Breath duration, d_r (frames)**														
n_r	**100**	**140**	**145**	**180**	**186**	**220**	**260**	**300**	**337**	**340**	**380**	**400**	**420**	**460**	**500**
1	48.29 1.58	48.42 1.44	48.42 1.44	48.42 1.44	48.42 1.44	48.42 1.44	48.42 1.44	48.42 1.44	48.42 1.44	48.42 1.44	48.42 1.44	48.42 1.44	48.42 1.44	48.42 1.44	48.42 1.44
5	71.54 8.35	74.71 7.36	72.67 6.86	73.95 7.2	72.8 6.75	**77.21** **7.56**	73.08 5.91	75.6 4.41	75.71 7	74.99 4.74	72.41 6.1	73.16 7.94	74.04 3.9	73.61 6.85	72.65 4.1
10	73.27 8.21	74.7 8.24	73.49 8.56	74.01 8.03	**75.07** **7.94**	74.76 10.03	**76.38** **9.7**	73.31 7.74	74.81 8.42	75.11 3.94	74.81 5.25	74.4 5.29	74.82 5.87	75.77 5.58	75.2 6.13
15	73.25 7.23	74.59 6.61	75.45 8.61	73.07 7.74	74.31 7.55	75.88 7.63	75.47 8.38	76.16 5.62	75.36 5.13	**75.49** **4.49**	73.06 5.73	73.31 5.61	75.07 3.44	74.29 4.26	74.72 4.31
20	**74.57** **8.3**	**75.48** **7.47**	73.22 9.58	**75.04** **9.75**	74.5 7.6	76.39 7.75	76.02 8.56	75.64 6.54	**76.28** **4.35**	74.81 5.35	**74.82** **5.72**	**76.04** **5.31**	74.7 5.24	**76.01** **6.03**	74.42 4.02
25	74.29 6.83	74.02 8.93	**75.61** **7.16**	73.06 7.53	74.39 8.11	73.46 8.49	72.89 8.89	**76.54** **7.69**	74.96 6.29	73.71 5.25	73.6 4.82	74.13 5.47	**75.1** **5.39**	74.85 2.69	74.7 3.99
30	73.79 7.12	72.39 8.1	74.23 8.1	72.35 10.25	73.39 6.95	74.12 8.52	74.25 10.09	75.31 9.04	74.14 4.87	74.79 7.06	74.8 5.89	75.07 4.9	74.57 4.55	75.21 6.21	**76.7** **2.76**

III). Best performance is observed the maximum number of times for $n_r = 20$ (6 out of 15) with varying d_r. From this experiment, it has been concluded that the exact locations of breath cycles are required, but the exact location of breath phases is not critical for the classification task.

VI. CONCLUSIONS

In this work, we carry out asthmatic vs health subject classification experiments using parts of breath cycles chosen in three different manners: 1) breath phases, i.e., inhale and exhale, which requires manually marked breath phase boundaries, 2) chunks of random duration chosen at random locations from recordings of continuous breath sounds, which does not require any boundary marking 3) first 25%, middle 50% and last 25% of a breath cycle, which requires manual marking of breath cycle boundaries. Interestingly, the classification accuracy with breath cycle parts chosen in the second manner does not perform worse than that using those chosen in the first manner. However, the highest classification performance is achieved when the parts of a breath cycle are chosen in the third manner, in particular the middle 50% of a breath cycle is found to be the most informative for discrimination between asthmatic and healthy subjects. These experiments indicate that, while breath phase boundaries may not be critical, having breath cycle boundaries would benefit the classification task. As parts of our future works, we plan to develop robust automatic breath segmentation algorithms and investigate how errors in automatic segmentation could impact the classification accuracy.

REFERENCES

[1] Fernando D Martinez, "Genes, environments, development and asthma: a reappraisal," *European Respiratory Journal*, vol. 29, no. 1, pp. 179–184, 2007.

[2] Global Asthma Network, "The global asthma report 2018.," .

[3] National Asthma Education, Prevention Program (National Heart, Lung, Blood Institute). Second Expert Panel on the Management of Asthma, National Heart, Lung, and Blood Institute. National Asthma Education Program. Expert Panel on the Management of Asthma, *Expert panel report 2: guidelines for the diagnosis and management of asthma*, National Institutes of Health, National Heart, Lung, and Blood Institute, 1998.

[4] MV Achuth Rao, NK Kausthubha, Shivani Yadav, Dipanjan Gope, Uma Maheswari Krishnaswamy, and Prasanta Kumar Ghosh, "Automatic prediction of spirometry readings from cough and wheeze for monitoring of asthma severity," in *25th European Signal Processing Conference (EUSIPCO)*. IEEE, 2017, pp. 41–45.

[5] Marcin Wiśniewski and Tomasz P Zieliński, "Joint application of audio spectral envelope and tonality index in an e-asthma monitoring system," *IEEE journal of biomedical and health informatics*, vol. 19, no. 3, pp. 1009–1018, 2014.

[6] Sandra Reichert, Raymond Gass, Christian Brandt, and Emmanuel Andrès, "Analysis of respiratory sounds: state of the art," *Clinical medicine. Circulatory, respiratory and pulmonary medicine*, vol. 2, pp. CCRPM–S530, 2008.

[7] T Richard Fenton, Hans Pasterkamp, A Tal, and Victor Chernick, "Automated spectral characterization of wheezing in asthmatic children," *IEEE transactions on biomedical engineering*, , no. 1, pp. 50–55, 1985.

[8] Gérard Charbonneau, J Racineux, M Sudraud, and E Tuchais, "Digital processing techniques of breath sounds for objective assistance of asthma diagnosis," in *International Conference on Acoustics, Speech, and Signal Processing*. IEEE, 1982, vol. 7, pp. 736–738.

[9] Fatma Z Göğüş, Bekir Karlik, and Güneş Harman, "Classification of asthmatic breath sounds by using wavelet transforms and neural networks," *International Journal of Signal Processing Systems*, vol. 3, no. 2, pp. 106–111, 2015.

[10] Shivani Yadav, NK Kausthubha, Dipanjan Gope, Uma Maheswari Krishnaswamy, and Prasanta Kumar Ghosh, "Comparison of cough, wheeze and sustained phonations for automatic classification between healthy subjects and asthmatic patients," in *40th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 1400–1403.

[11] Shivani Yadav, Merugu Keerthana, Dipanjan Gope, Prasanta Kumar Ghosh, et al., "Analysis of acoustic features for speech sound based classification of asthmatic and healthy subjects," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 6789–6793.

[12] D Mazzoni and R Dannenberg, "Audacity [software]. pittsburg," 2000.

[13] Mike Brookes et al., "Voicebox: Speech processing toolbox for matlab," *Software, available [Mar. 2011] from www. ee. ic. ac. uk/hp/staff/dmb/voicebox/voicebox. html*, vol. 47, 1997.

[14] Chih-Chung Chang and Chih-Jen Lin, "Libsvm: a library for support vector machines," *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, no. 3, pp. 1–27, 2011.

[15] Charilaos Chourpiliadis and Abhishek Bhardwaj, "Physiology, respiratory rate," *StatPearls [Internet]*, 2019.

[16] David MW Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020.