ISBN-978-989-9121-29-4
ISTANBUL 25th International Conference on "Innovations in Engineering, Technology & Health Sciences" (IIETH-23)
Sept. 5-7, 2023 Istanbul (Turkiye)

Clinical Decision Supporting Algorithm for Screening COVID-19 Using Exhaled Breath

M. B. Malarvili[1,2*], Tjokorda Gde Tirta Nindhia[1*], Santheraleka Ramanathan[2], Nadhira Dahari[2], Nabiel Rafa Angel Bhaswara[1], Thalitakum Hutajulu[1], Rishya Manikam[3], Ary Esta Dewi Wirastuti[4]

*Corresponding authors: malarvili@biomedical.utm.my, tirta.nindhia@me.unud.ac.id

[1]Study Program of Mechanical Engineering, Engineering Faculty, Udayana University Jimbaran, Bali, Indonesia.
[2] School of Biomedical and Health Science Engineering, Universiti Teknologi Malaysia, 81310, Skudai, Johor, Malaysia
[3]Faculty of Medicine, University Malaya Medical Centre (UMMC), Kuala Lumpur 59100, Malaysia
[4]Department of Electrical Engineering, Faculty of Engineering, Udayana University

Abstract: *An interpretable algorithm is developed to identify the COVID-19 infection using exhaled breath profiling and its features. The preliminary study was conducted with 100 subjects, where 50 with positive COVID-19 category 2 condition. A two minutes of exhaled breath pattern was recorded using a nasal cannula sampling tube, resulting in the collection of exhaled breath waveforms from each subject. The developed algorithm is utilized to evaluate the valid exhaled breath waveforms and compute the features classified to distinguish COVID and non-COVID condition. The significant features showed good predictive power for compliance with p-value analysis with area under the receiver operating characteristic (ROC) curve of 0.500 – 0.550. Exhaled breath profiling with developed algorithm is an alternative approach of detecting COVID-19 infection at clinical practice and expected to replace the invasive sampling of the virus infection with further validation in future studies.*

Keywords*: Exhaled breath; COVID-19; Algorithmic approach; Clinical decision support system; SARS-CoV-2*

1. Introduction

COVID-19 disease is defined as a respiratory infection, which leads to severe pulmonary complications and multi-organ infection. The historical evidence caused by pneumonias such as SARS and MERS have been recapitulated by the severe impact caused by SARS-CoV-2 virus spreading, resulting in unmanageable infectivity rate and huge number in death [1]. The wide spread of COVID-19 global pandemic forced recent medical experts to focus on the disease virology, physiology, and immunology for clinical practices to curb

the pandemic. The infected COVID-19 subjects reported to exhibit an acute abnormal breathing pattern both during and after their infection [2]. Breath based detection gained huge attention in clinical practices as it highly reduces the workload and cost associated with invasive sampling such as nasal swabs, and other sophisticated sampling techniques such as RT-PCR, isothermal amplification techniques, and enzymatic assays [3]. Breath sampling is rapid and repeatable. It shows huge potential for primary COVID-19 screening upon patient admission, which would help to initiate the preventive actions at clinical practices. Experts in clinical practice have recommended initiating invasive ventilation as soon as the patients show sudden surge in respiratory symptoms as critical care providers become more familiar with the harmful course of the COVID-19. Exhaled breath in this content involves non-invasive measurement of the concentration of carbon dioxide (CO_2) [4]. Exhaled breath contains useful information for the diagnosis of respiratory airway diseases including asthma and chronic obstructive pulmonary disease (COPD) [5,6]. In this study, a novel algorithmic approach for detecting COVID-19 condition using exhaled breath is presented. This study aims to evaluate the CO_2 partial pressure variation in a COVID-19 patient. Fig 1 shows the schematic illustration of the proposed approach of identifying COVID-19 using exhaled CO_2 gas and the developed algorithm for exhaled breath cycle profiling.

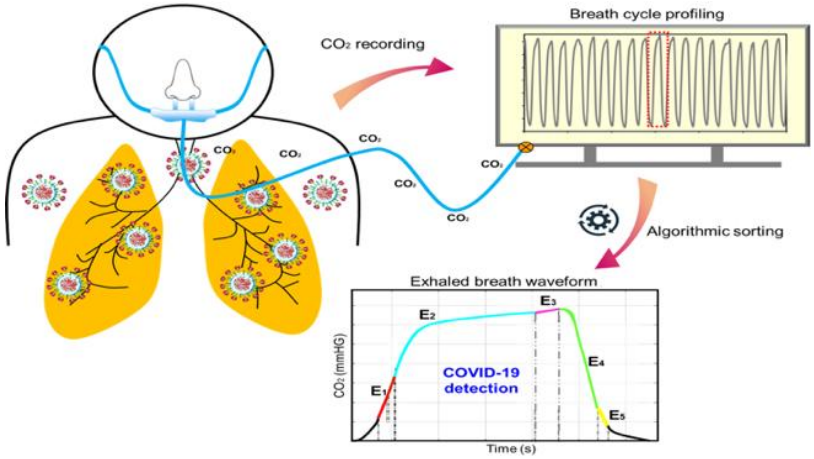

Fig 1: Illustration of the utilizing exhaled breaths for identifying COVID-19 using proposed algorithm.

2. Methodology

2.1 Clinical Data Sources

Clinical data collection was conducted at a public hospital in Malaysia between February to March 2022. The clinical study is registered in the National Medical Research Register (NMRR) with the reference number (NMRR-21-763-59692). Adult COVID-19 patients between 18 to 60 years of age range were targeted for this study. Patients with only COVID-19 category 2 (CAT 2) infection stage were approached to participate in this study. Patients with asthma, COPD, and pulmonary edema history are excluded to ensure the CO_2 exhaled breaths recording is not affected by other respiratory complications. The CO_2 breath analysis was carried out with 100 participants, where the confirmed COVID-19 CAT 2 participants, n_P is 50. The study was evaluated with 50 non-COVID patients, n_N, who were negative COVID-19 participants. COVID-19 with CAT 2 infection stage were validated with RTK test kit upon arrival.

2.2 Feature Extraction Algorithms

A sorting hierarchical algorithm as shown in Figure 2 is developed to analyze the exhaled breath recorded. The algorithm begins by implementing digital filter to remove all unwanted noise. A first order FIR low pass filter with the cut-off frequency of 10 Hz and a moving average filter is used for this purpose [8]. Later, the algorithm for valid breath cycle selection is developed by assessing the minimum end-tidal carbon dioxide ($EtCO_2$) value, peak-to-peak distance, and the pathologic properties of exhaled breath waveform. The selection criteria outlier exhalation breath cycles in the waveform pattern, which are designated as the valid CO_2 signal recorded from a patient. As each valid breath cycle is identified, the algorithm sends information to segmentize CO_2 signal waveform into five epochs. Epoch segmentation is the crux of the presented study. Each valid exhaled breath cycle extracted is segmented into five epochs by employing a CO_2 partial pressure threshold value. These epochs are the original proposal of the study, which contrasts with the visual analysis of capnogram waveform that is known in capnography technology [7], where the alveolar plateau and $EtCO_2$ peak value were highly prioritized. These segments were generated by considering the average CO_2 partial pressure in a breath cycle. The partial pressures in the expiration curve and inspiration curve are concentrated in generating the epochs. E_1 and E_2 represent the expiration region, where E_4 and E_5 represent the inspiration region of a waveform. E_3 shows the end-tidal point of a waveform, where the epoch is measured 0.25s from the $EtCO_2$ and to the $EtCO_2$ peak. The epoch is developed as such due to the trend-line based analysis performed in this study [10]. This information is embedded in the developed algorithm, which identifies each valid exhaled breath waveform and segmentize it into five epochs. As established in previous studies, the slope and area indices of each epoch were computed using Equation 2 and 3, respectively [11]. The least squares linear fitting method was used to compute the slopes of each sub-cycle. By lowering the residue in accordance with (3), it is possible to include the entire CO_2 signal while still computing the intercept and slope of the CO_2 waveform.

$$Slope\ (S_j) = \frac{1}{C} \sum_{j=0}^{C-1} b_j\ (M_j - S_j)^2 \tag{2}$$

$$AR_i = \frac{dt}{5} \sum_{j=0}^{i} (\ R_{j-1}(t) + 4R_j(t) + R_{j+1}(t)) \tag{3}$$

Where, slope (S) length is defined by C, S_j defines the jth element, the best fit and weight of jth element are defined by M_j and b_j, respectively. The CO_2 signal and the sampling interval are defined by $R(t)$ and dt in area (AR_i) computational formula. In addition, the Hjorth parameters of each epoch were extracted. Equation 4 defines the computational formula for Hjorth activity (variance) [12].

$$\sigma_y^2 = \frac{1}{L} \sum_{l=0}^{L-1} (y(l) - \bar{y})^2 \tag{4}$$

Where, y(l) defines the CO_2 signal, and the L indicates the data length. Mobility is the second Hjorth parameter, which indicates the ratio of the standard deviation of the signal's first derivative to that of the signal itself.. The computational formula for mobility is as below, where y'(l) refers to the first derivative of the signal [13].

$$Mobility[y(l)] = \frac{\sigma_{y\prime}}{\sigma_y} \tag{5}$$

Equation 6 represents the formula to calculate Hjorth complexity [14].

$$Complexity[y(l)] = \frac{mobility\ [y'(l)]}{mobility\ [y(l)]} = \frac{\sigma_{y\prime\prime}/\sigma_{y\prime}}{\sigma_{y\prime}/\sigma_y} \tag{6}$$

The y′(l) and y′′(l) represent the first and second derivative of CO_2 signal. The following equations define the expression for the derivations.

$$y'(l) = \frac{1}{T_s}[y(l+1) - y(l)] \tag{7}$$

$$y''(l) = \frac{1}{T_s}[y'(l+1) - y'(l)] \quad (8)$$

Where Ts defines the sampling time interval [12]. In this study, an approximate 2-minute data sampling resulted in an approximate of 35 valid exhaled breath waveforms from a single patient. About 1700 valid exhaled breath waveforms resulted in feature computation and interpretation for high resolution prediction and the ability to detect even minute waveform changes. The $EtCO_2$ value and respiratory rate of a patient is computed from the exhaled breath signal. The inspiratory to expiratory ratio is one of the key indices analyzed in the present study. The corresponding slopes and areas of the two phases were defined in ratios.

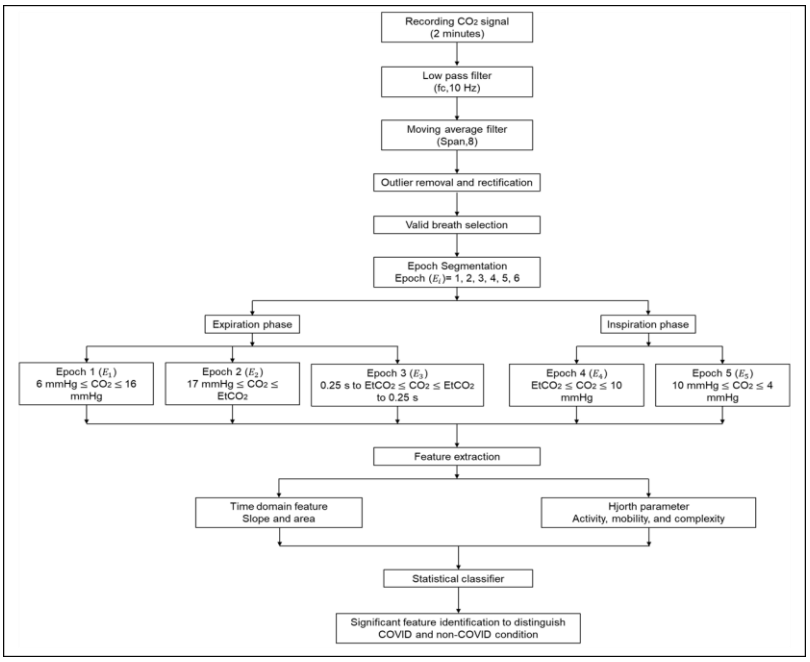

Fig 2: Overview of the presented study of detecting COVID-19 from exhaled breath profile analysis.

2.3 Data Analysis

A descriptive statistical analysis was performed to present the baseline data acquired in this study. The significance of feature characteristics in defining COVID and non-COVID condition was evaluated using significance value from paired sample T-test. Receiver operating characteristic (ROC) curves were generated for exhaled breath features and indices, where the sensitivity and specificity were quantified through area under curve (AUC) and its 95% confidence interval [9].

3. Results

In this study, a new approach of exhaled breath segmentation is introduced. A valid exhaled breath cycle is segmented into five epochs. The sorting algorithm developed in this study identifies each valid breath cycles and segmentize each cycle into the five epochs. The area and slope indices of each epoch were computed using the developed algorithm [15]. As the Hjorth parameters are highly approached in differentiating asthmatic condition, the parameters are included in the present study for evaluating the waveform of COVID and non-COVID subject. The entire computational were performed using the sorting algorithm developed in this study. Fig 3 shows the output of a valid exhaled breath waveform for COVID and non-COVID condition. Fig 3a-3c shows the exhaled breath waveform extracted from a CAT 2 COVID

subject. The shape of waveform deviates from the standard shape. Deviation plays the key role in computing the algorithmic output of each epoch. Such variation is analyzed to determine the presence of COVID and non-COVID state of a patient based on algorithmic approach. Fig 3d-3f shows the exhaled breath waveform of a non-COVID subject. The shape of waveform is close to the standard capnography, which indicates the absence of any deviation in the subject breath cycle.

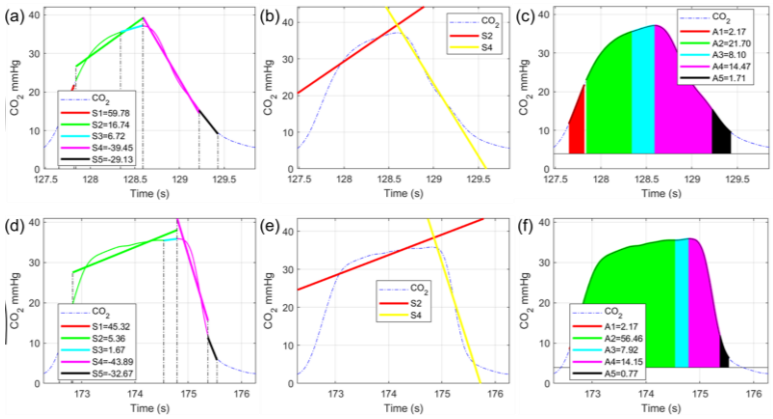

Fig 3: Algorithmic approach of feature extraction. (a) – (b) show exhaled breath waveform indicating slope, angle, and area of each epoch segmented from a COVID-19 subject breath profile, whereas (d) – (f) indicate segmented features for non-COVID subject.

The primary analysis of features in each epoch were presented as the mean with standard deviation (SD). The $EtCO_2$ mean for COVID condition is 36.36 with SD of 3.53, where the non-COVID condition showed 34.71 mean with 3.46 SD. COVID subjects showed 22.16 mean respiratory rare (RR) with 6.05 SD. Non-COVID subjects 19.22 mean RR with 4.33 SD. The paired T-test analysis for both $EtCO_2$ and RR values showed no significant difference with *p* values equal to 0.817 and 0.274, respectively. The mean and SD of features for each epoch were computed. Slope, area, activity, mobility, and complexity of five epochs in a valid breath cycle, resulted in about 25 different features for evaluating COVID and non-COVID condition. Table 1 summarizes the significant features out of the 25 features. The slope of E_2 showed significant variation between COVID and non-COVID condition with 0.016 *p*-value. The area of E_2 and the activity of E_4 showed high significance with $p = 0.000$. Mobility of E_2 showed $p = 0.001$, indicating significant variation to be considered for evaluating COVID conditions. The complexity feature showed significant variation for only E_3, with $p = 0.042$. The inspiratory to expiratory ratio (E_4/ E_2) computed for slope and area showed significant differences with $p = 0.005$. The rest of features of each epochs indicated insignificant variation between COVID and non-COVID condition. The T-test outputs manifest good correlation between COVID and non-COVID condition with the *p* value exhibited by the significance features, which is highly approached for preliminary clinical decision making upon COVID-19 diagnosis.

Receiver operating characteristics (ROC) revealed good predictive values in comparing the significance of exhaled features for predicting COVID and non-COVID condition. The predictive values for $EtCO_2$ and RR were moderately higher than the other significant features, with area under curve (AUC) values of 0.617 and 0.641, respectively. The slope of E_2, area of E_2, and significant Hjorth parameters showed fair compliance with COVID and non-COVID condition with AUC values between 0.500 and 0.550 (Table 1, Fig 4). The slope and area of E_4/ E_2 ratio showed 79.8% and 80.6% of sensitivity and a specificity of 81.1% and 80.4%, respectively for prediction good compliance with COVID and non-COVID condition with a maximum cut-off value of 1%.

In this study, a two-tailed paired sample T-test was performed as the primary statistical analysis to identify the features, which show significant differences between COVID and non-COVID condition. The algorithmic simulation showed that slope and area of E_2 is significant. As the developed system sensitively detects the variation in the exhaled CO_2 gas, it is highly expected that significant differences between the slope and area of E_2 as the E_2 indicates the exhalation region of the waveform. These indicate that the slope and area feature for E_2 are significant to be used in distinguishing COVID and non-COVID condition. The mobility feature of E_2 showed significant differences. Mobility, which indicates the second derivate of the SD proportional to the signal power spectrum can differentiate COVID conditions. Based on *p*-value analysis, the mobility of E_2 can be used in discriminating COVID and non-COVID condition. Besides, the complexity of E_3 and activity of E_4 showed significant differences. E_3 reveals the $EtCO_2$ peak variation between COVID and non-COVID condition, whereas E_4 reveals the inspiratory region of the waveform. In the presented study, E_4/E_2 ratio for slope and area feature were statistically computed through the developed algorithm. It is justified to be significant in differentiating COVID and non-COVID condition. The significant features are further analyzed using ROC analysis. The area under the ROC curve is moderately greater than the 0.500 for all significant features. The slope and area of E_4/E_2 ratio showed a higher AUC than the rest, which emphasizes the importance of computing the inspiratory to expiratory ratio in identifying COVID and non-COVID condition. The logistic regression coefficient showed good correlation between $EtCO_2$ and RR. It revealed that $EtCO_2$ and RR computational analysis is insufficient to identify COVID condition. Based on the statistical analyses, Slope of E_2, area of E_2, E_4/E_2 ratio for slope and area are identified as the promising features to be chosen in discriminating COVID and non-COVID condition.

TABLE I: Statistical analysis performed for the extracted significant features between COVID and non-COVID patients.

Features	Mean ± SD		*P* value	AUC	95% Confidence interval	
	COVID	Non-COVID			Lower limit	Upper limit
$EtCO_2$	36.36 ± 3.53	34.71 ± 3.46	0.817	0.617	0.50	0.73
RR	22.16 ± 6.05	19.22 ± 4.33	0.274	0.641	0.53	0.75
Slope of E_2	14.34 ± 7.66	13.98 ± 6.52	0.016	0.502	0.0005	0.89
Area of E_2	25.56 ± 17.22	24.31 ± 14.49	0.000	0.515	0.24	2.38
Activity of E_4	40.52 ± 16.40	42.27 ± 15.50	0.000	0.530	-2.08	-0.11
Mobility of E_2	0.038 ± 0.017	0.038 ± 0.015	0.001	0.530	-0.0007	0.003
Complexity of E_3	8.64 ± 7.34	9.85 ± 8.22	0.042	0.526	-1.56	-0.59
Slope ratio of E_2	3.35 ± 2.25	3.58 ± 2.49	0.005	0.542	0.09	0.40
Area ratio of E_2	0.721 ± 0.468	0.649 ± 0.328	0.005	0.546	0.04	0.099

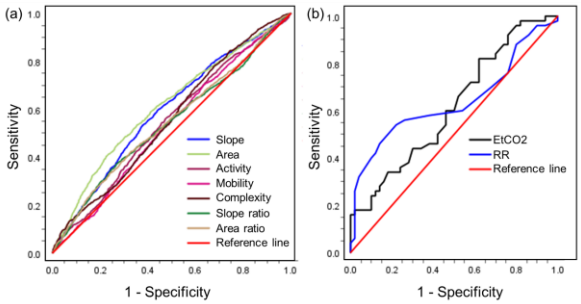

Fig 4: Receiver operating characteristics (ROC) curves for the (a) significant features computed from the developed algorithm and (b) for the measured $EtCO_2$ and RR as predictors of good agreement for differentiating COVID-19 condition.

4. Conclusion

This work has supported the alternative approach of detecting COVID-19 infection through the exhaled breath profiling using a developed algorithm and emphasized that capnography technology is not only aids in ventilation tidal volume monitoring, but also helps in demonstrating the deviation between COVID-19 infected condition and non-infected condition. The interpretable algorithm delivered desired output by segmenting a valid exhaled breath waveform into five epochs and computing its features readings. The statistical analyses revealed that slope of E_2, area of E_2, E_4/E_2 ratio for slope and area are the promising features identified for discriminating COVID and non-COVID condition. The outcome of this work aims in implementing in clinical practice as the primary disease detection technique, which will help to clinicians to initiative preventive actions and saves many workloads associated with invasive sampling methods.

5. Acknowledgement

The authors express their gratitude to the Universiti Teknologi Malaysia (UTM) Health Centre and Policlinic Siva, Malaysia for their contributions to clinical data collection. This study is supported by UTM GUP with grant number R.J130000.7351.4B746 and UNISERF 2023 of Udayana University, Indonesia.

6. References

[1] Xu R, Liu P, Zhang T, Wu Q, Zeng M, Ma Y, et al. Progressive deterioration of the upper respiratory tract and the gut microbiomes in children during the early infection stages of COVID-19. Journal of Genetics and Genomics. Elsevier Limited and Science Press; 2021;48:803–14. https://doi.org/10.1016/j.jgg.2021.05.004

[2] Taleghani N, Taghipour F. Diagnosis of COVID-19 for controlling the pandemic: A review of the state-of-the-art. Biosens Bioelectron. Elsevier B.V.; 2021;174:112830. https://doi.org/10.1016/j.bios.2020.112830

[3] Basin S, Valentin S, Maurac A, Poussel M, Pequignot B, Brindel A, et al. Progression to a severe form of COVID-19 among patients with chronic respiratory diseases. Respir Med Res. 2022;81. https://doi.org/10.1016/j.resmer.2021.100880

[4] Anderson MR. Capnography: Considerations for Its Use in the Emergency Department. J Emerg Nurs. 2006;32:149–53. https://doi.org/10.1016/j.jen.2005.12.025

[5] Peveling-Oberhag J, Michael F, Tal A, Welsch C, Vermehren J, Farnik H, et al. Capnography monitoring of non-anesthesiologist provided sedation during percutaneous endoscopic gastrostomy placement: A prospective, controlled, randomized trial. Journal of Gastroenterology and Hepatology (Australia). 2020;35:401–7. https://doi.org/10.1111/jgh.14760

[6] Bouazza B, Hadj-Said D, Pescatore KA, Chahed R. Are patients with asthma and chronic obstructive pulmonary disease preferred targets of COVID-19? Tuberc Respir Dis (Seoul). 2021;84:22–34. https://doi.org/10.4046/trd.2020.0101

[7] Yang J, An K, Wang B, Wang L. New mainstream double-end carbon dioxide capnograph for human respiration. J Biomed Opt. 2010;15:065007. https://doi.org/10.1117/1.3523620

[8] Singh OP, Howe TA, Malarvili MB. Real-time human respiration carbon dioxide measurement device for cardiorespiratory assessment. J Breath Res. Institute of Physics Publishing; 2018;12. https://doi.org/10.1088/1752-7163/aa8dbd

[9] Balakrishnan M, Kazemi M, Mahmood NH, Malarvili MB, Humaimi Mahmood N. Investigating Frequency Contents of Capnogram using Fast Fourier Transform (FFT) and Autoregressive Modeling (AR) Heart Rate

Monitoring System View project Investigating Frequency Contents of Capnogram using Fast Fourier Transform (FFT) and Autoregressive Modeling (AR) [Internet]. Available from: http://www.biomedical.utm.my

[10] Howe TA, Jaalam K, Ahmad R, Sheng CK, Nik Ab Rahman NH. The use of end-tidal capnography to monitor non-intubated patients presenting with acute exacerbation of asthma in the emergency department. Journal of Emergency Medicine [Internet]. Elsevier Inc.; 2011;41:581–9. Available from: http://dx.doi.org/10.1016/j.jemermed.2008.10.017

[11] Malarvili MB, Alexie M, Dahari N, Kamarudin A. On Analyzing Capnogram as a Novel Method for Screening COVID-19 : A Review on Assessment Methods for COVID-19. Life Journal. 2021;1–26. https://doi.org/10.3390/life11101101

[12] El-Badawy IM, Omar Z, Singh OP. An Effective Machine Learning Approach for Classifying Artefact-Free and Distorted Capnogram Segments Using Simple Time-Domain Features. IEEE Access. Institute of Electrical and Electronics Engineers Inc.; 2022;10:8767–78. https://doi.org/10.1109/ACCESS.2022.3143617

[13] Oh S-H, Lee Y-R, Kim H-N. A Novel EEG Feature Extraction Method using Hjorth Parameter.

[14] Hjorth BO. TECHNICAL EEG ANALYSIS BASED CONTRIBUTIONS ON TIME DOMAIN PROPERTIES.

[15] OMPS Singh, MB Malarvili. Review of Infrared Carbon-Dioxide Sensors and Capnogram Features for Developing Asthma-Monitoring Device. Journal of Clinical & Diagnostic Research 12 (10).