Non-Contact Supervision of COVID-19 Breathing Behaviour With FMCW Radar and Stacked Ensemble Learning Model in Real-Time

Ariana Tulus Purnomo, Kokoy Siti Komariah, *Graduate Student Member, IEEE*, Ding-Bing Lin, *Senior Member, IEEE*, Willy Fitra Hendria, Bong-Kee Sin, *Member, IEEE*, and Nur Ahmadi, *Member, IEEE*

***Abstract*—A respiratory disorder that attacks COVID-19 patients requires intensive supervision of medical practitioners during the isolation period. A non-contact monitoring device will be a suitable solution for reducing the spread risk of the virus while monitoring the COVID-19 patient. This study uses Frequency-Modulated Continuous Wave (FMCW) radar and Machine Learning (ML) to obtain respiratory information and analyze respiratory signals, respectively. Multiple subjects in a room can be detected simultaneously by calculating the Angle of Arrival (AoA) of the received signal and utilizing the Multiple Input Multiple Output (MIMO) of FMCW radar. Fast Fourier Transform (FFT) and some signal processing are implemented to obtain a breathing waveform. ML helps the system to analyze the respiratory signals automatically. This paper also compares the performance of several ML algorithms such as Multinomial Logistic Regression (MLR), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), eXtreme Gradient Boosting (XGB), Light Gradient Boosting Machine (LGBM), CatBoosting (CB) Classifier, Multilayer Perceptron (MLP), and three proposed stacked ensemble models, namely Stacked Ensemble Classifier (SEC), Boosting Tree-based Stacked Classifier (BTSC), and Neural Stacked Ensemble Model (NSEM) to obtain the best ML model. The results show that the NSEM algorithm achieves the best performance with 97.1% accuracy. In the real-time implementation, the system could simultaneously detect several objects with different breathing characteristics and classify the respiratory signals into five different classes.**

***Index Terms*—COVID-19, ensemble model, FMCW, machine learning, stacking, vital signs.**

Manuscript received 21 January 2022; revised 30 March 2022 and 19 May 2022; accepted 24 June 2022. Date of publication 19 July 2022; date of current version 10 October 2022. This work was supported by ITB-NTUST joint research program 2022 under Grant ITB-NTUST-111-02. This paper was recommended by Associate Editor Y. Li. *(Ariana Tulus Purnomo and Kokoy Siti Komariah contributed equally to this work.) (Corresponding author: Ding-Bing Lin.)*

Ariana Tulus Purnomo and Ding-Bing Lin are with the Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, Taipei 10607, Taiwan (e-mail: d10602808@mail.ntust.edu.tw; dblin@mail.ntust.edu.tw).

Kokoy Siti Komariah and Bong-Kee Sin are with the Department of AI Convergence and the Division of Computer Engineering (respectively), Pukyong National University, Busan 48513, Republic of Korea (e-mail: kokoysk@pukyong.ac.kr; bkshin@pknu.ac.kr).

Willy Fitra Hendria is with the Department of Intelligent Mechatronics Engineering, Sejong University, Seoul 05006, Republic of Korea (e-mail: willyfitrahendria@sju.ac.kr).

Nur Ahmadi is with the Center for Artificial Intelligence (U-CoE AI-VLB), School of Electrical Engineering and Informatics, Bandung Institute of Technology, Bandung 40132, Indonesia (e-mail: nahmadi@itb.ac.id).

This work involved human subjects or animals in its research. The author(s) confirm(s) that all human/animal subject research procedures and protocols are exempt from review board approval.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TBCAS.2022.3192359.

Digital Object Identifier 10.1109/TBCAS.2022.3192359

I. Introduction

COVID-19 causes most sufferers to experience respiratory problems. Usually, COVID-19 patients will suffer from coughing and shortness of breath in an unknown period of time [1], [2]. In the worst case, the patient might be unable to breathe and subsequently lose their life. Therefore, supervision of COVID-19 patients is very crucial to reduce the death rate [3], [4].

Direct contact with COVID-19 patients is very dangerous since the virus can spread quickly through the air [5], [6]. Therefore, a non-contact technology that can be controlled from a central room is a helpful solution to reduce the direct contact between COVID-19 patients and medical practitioners.

Previous studies have been conducted to detect respiratory signals. They used ECG sensors in [7], and gyroscope and accelerometer sensors in [8] to classify some respiratory signals. Unfortunately, these wearable sensors were less comfortable since they required physical contact with the device. However, it is better to observe COVID-19 patients without any physical contact to reduce the spread of the virus. Therefore, non-contact medical devices become a research concern.

Non-contact technology such as described in [9] used camera technology to observe chest displacement. Unfortunately, the detection with a camera had its drawbacks in terms of light and privacy. In [10], detecting vital signs using radar was used to classify the subject's age by using single range-FFT image. However, their classification method did not observe the characteristic of the breathing waveform. Patients suffering from respiratory disorders or COVID-19 have a unique breathing characteristic that cannot be represented by heart and breath rates. Furthermore, the system in [10] had not been implemented to perform a detection in real-time. Thus, based on this background, the challenge is to develop a non-contact monitoring device that can classify breathing waveform based on the characteristics of the signal in real-time.

One of the non-contact technologies that does not need a good lighting for object detection is radar. There are three types of radar: Continuous Wave (CW) radar [11], [12], Ultra-Wide Band (UWB) radar [13], [14] and Frequency-Modulated Continuous Wave (FMCW) radar [15]–[17], which can be used to detect human vital signs [18], [19]. FMCW is able to perform small multi-object detection with a reasonable low transmit power [20]. Compared to camera, FMCW radar as a non-contact sensor has an advantage in terms of light and privacy [21]. With its frequency modulation, FMCW radar has broader range and higher resolution detection so that it can detect subtle movement such as chest displacement, which contains breathing information [22], [23].

To answer the challenge mentioned earlier, we used FMCW radar to build a non-contact monitoring device that detects the breathing disorder of a multiple subjects in a room. They can be detected simultaneously by calculating the Angle of Arrival (AoA) of the received signal and utilizing the Multiple Input Multiple Output (MIMO) of FMCW radar. Fast Fourier Transform (FFT) and some signal processing are implemented to obtain the breathing waveform.

Adding Machine Learning (ML) will help the system detect and analyze different breathing disorder behaviour. We conducted the experiment using several machine learning algorithms to choose a suitable ML model for our system. We proposed a stacked ensemble model to further enhance the performance of our system in classifying breathing waveform data. We compared the performance of several classifiers that have been proven to be effective in similar cases such as Multinomial Logistic Regression (MLR) [24], Decision Tree (DT) [25], [26], Random Forest (RF) [27], Support Vector Machine (SVM) [28], [29], eXtreme Gradient Boosting (XGB) [30], [31], Light Gradient Boosting Machine (LGBM), CatBoosting Classifier (CB), Multilayer Perceptron (MLP) [32], [33] and three proposed stacked ensemble models.

The stacked ensemble model was first introduced by [34] and proven to be better than any single model in terms of prediction performance [35]. This study built stacked ensemble models using a stacking technique to combine multiple classification or regression models. There are numerous ensemble-modeling methods [36], [37], the most well-known of which are bagging and boosting [38]. Bagging calculates the average of multiple similar models with high variance to reduce the spread of data. Boosting creates multiple incremental models to reduce bias while keeping the variance low. However, stacking works differently, and it can combine various models or learners and has a concept called meta learner. This classifier uses the predictions as features to make a final prediction among all of the predictions.

Three different stacked ensemble classifiers, namely Stacked Ensemble Classifier (SEC), Boosting Tree-based Stacked Classifier (BTSC), and Neural Stacked Ensemble Model (NSEM), were proposed and evaluated. Based on the results, the NSEM algorithm achieves the best performance with a mean accuracy of 97.1% using 10-fold CV. In the real-time implementation, multiple subjects with different breathing characteristics could be identified simultaneously, and the system was able to identify five different types of classes.

Our previous study [30] successfully detected the breathing waveform using FMCW radar with one transmitter and one receiver antenna. The range FFT and phase change information was used to provide chest displacement information. The help of the MFCC Feature Extraction (FE) and XGBoost algorithm has successfully achieved 87.375% accuracy in classifying five breathing classes. In this study, the system was developed in terms of radar and machine learning. The challenge of this research is to develop the detection of multiple persons and increase the accuracy in detecting breathing classes.

Our contributions are listed as follows.

- Two-transmitter and four-receiver antennas are used to determine each subject's Angle-of-Arrival (AoA) estimation. This AoA estimation helps the system plot the range-azimuth Map to estimate and track the position of objects.
- Multi-person breathing waveforms can be recognized simultaneously based on the range FFT, 2D-FFT, AoA, and phase change.
- The combination of MFCC and statistical FE helps the system distinguish different breathing classes with a better performance than [30].
- Proposing a stacked ensemble learning approach for system performance improvement and showing that the proposed model outperforms previous research [30].

The rest of the chapter is summarized as follows: Section 2 explains the proposed method, Section 3 demonstrates the experiment and result, and Section 4 concludes the study.

II. Proposed System

In this study, we developed five types of detection of respiratory signals with FMCW radar which can be used to monitor breathing conditions of several subjects simultaneously without physical contact. Five types of respiratory signals that can be detected are normal breathing, deep and quick breathing, deep breathing, quick breathing, and hold breathing.

Normal breathing signals have a breathing rate at about 12-20 breaths per minute with an amplitude ranging from 1-12 mm [39], [40]. Deep and quick breathing signals have a big amplitude with a high breathing rate. It usually occur in hyperventilation [41], Kussmaul [39]and biot patients [42]. Deep breathing signals have a big amplitude with a low breathing rate. It often attacks patients with hyperpnea symptoms. Quick breathing signals have a small amplitude with a high breathing rate. Asthma [43], [44], COVID-19 and tachypnoea patients [29] are often attacked with shortness of breath or quick breathing. Hold breathing signals have a tiny amplitude that is almost imperceptible. Respiration of bradypnea patients has the same characteristics as hold breathing signals [45].

Our proposed system, shown in Fig. 1, consists of 4 main modules: FMCW radar module, object tracking module, breathing extraction module, and ML module. FMCW radar produces $s(t)$ signal and transmits it through Tx transmitter. The reflected signal of the object is received by the receiver antenna as $r(t)$. A breathing waveform can be obtained by passing the received

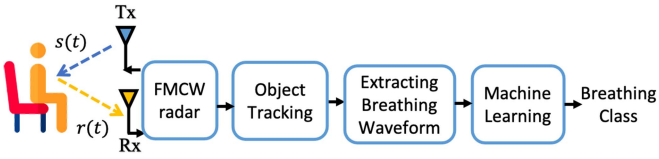

Fig. 1. Four main modules of the proposed system.

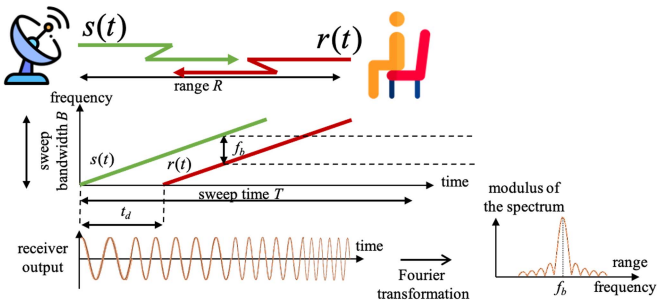

Fig. 2. The basic concept of FMCW radar.

signal through several signal processing blocks. The addition of the ML model will help the system to classify the breathing disorder. The following subsection will explain the proposed system in detail.

A. MIMO FMCW Radar

In this subsection, we briefly describe the basic concept of MIMO FMCW radar [46]. FMCW is a continuous signal with linearly-changing frequency with time, as shown in Fig. 2.

An FMCW signal is transmitted to an object at time index t as

$$s(t) = A\cos\left(2\pi f_c t + \pi\frac{B}{T}t^2 + \phi(t)\right), \tag{1}$$

where A is the transmitted signal power, T is the sweeping time, f_c is the starting frequency of the chirp, B is the sweeping bandwidth, and $\phi(t)$ is the phase. The object reflects the transmitted FMCW signal with the same frequency change, and the reflected signal is received by the receiver at time t_d as

$$r(t) = \alpha A\cos\left[2\pi f_c(t-t_d) + \pi\frac{B}{T}(t-t_d)^2 + \phi(t-t_d)\right], \tag{2}$$

where α is the resized scale. The received signal is mixed with the transmitted signal and passed through the low pass filter (LPF) to produce a beat signal

$$\begin{aligned} b(t) &= LPF(s(t)r(t)) \\ &= \frac{\alpha(A)^2}{2}\cos\left(2\pi\frac{Bt_d}{T}t + 2\pi f_c t_d - \pi\frac{B}{T}(t_d)^2 + \Delta\phi(t)\right), \end{aligned} \tag{3}$$

where $\Delta\phi(t)$ is the residual phase noise, $\Delta\phi(t) = \phi(t) - \phi(t)(t - \frac{2R(t)}{c})$. The chest movement can be expressed as $R(t) = R_{const} + R_l$, where R_{cons} is a constant range from the radar and R_l is the relative movement of the object. We have $t_d = \frac{2R(t)}{c}$,and $f_c = \frac{c}{\lambda}$, where λ is the wavelength and c is the

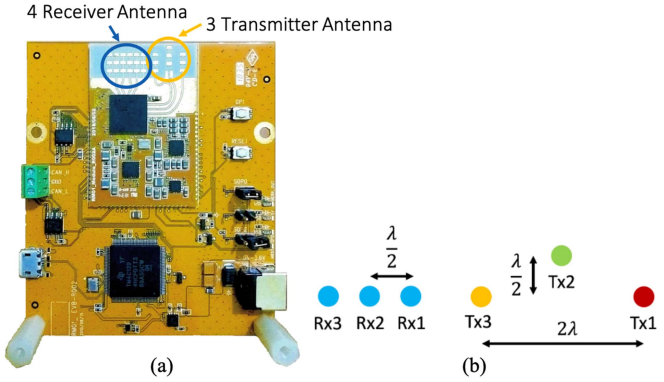

Fig. 3. (a) Texas Instrument - IWR1443 board with three-transmitter and four-receiver antennas, (b) IWR1443 antenna configuration.

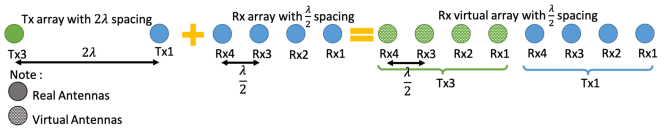

Fig. 4. Two-transmitter and four-receiver antennas form eight virtual array antennas.

speed of light. Note that $f_b = \frac{2BR(t)}{cT}$. Thus, we have

$$\begin{aligned} b(t) &\approx \frac{\alpha(A)^2}{2}\cos\left(2\pi\frac{2BR(t)}{cT}t + \frac{4\pi R(t)}{\lambda} + \Delta\phi(t)\right) \\ &= \frac{\alpha(A)^2}{2}\cos\left(2\pi f_b t + \Phi_b(t) + \Delta\phi(t)\right), \end{aligned} \tag{4}$$

where $\Phi_b(t)$ is the phase of the beat signal. The signal in (4) can be expressed as a complex signal as

$$b(t) = \alpha A e^{j(2\pi f_b t + \Phi_b(t) + \Delta\phi(t))}. \tag{5}$$

In this study, we used a Texas Instruments - IWR1443 board which has three-transmitter and four-receiver patch antennas as shown in Fig. 3(a) [47]. This antenna has a configuration as shown in Fig. 3(b). For our proposed system, we used a linear antenna configuration. The linear arrangement of antennas with the MIMO concept can provide a higher resolution result with fewer antenna elements. Thus, we picked two-transmitter antennas, $Tx1$ and $Tx3$, and four-receiver antennas, $Rx1$, $Rx2$, $Rx3$, and $Rx4$, antennas of the IWR1443 board.

To process the signal received by the radar, a Uniform Linear Array (ULA) is formed from two transmitters and four receivers. Two transmitters with a distance of 2λ and four receivers spaced by $\frac{\lambda}{2}$ produce four additional receiver antennas spaced by $\frac{\lambda}{2}$. Thus, the system has eight virtual array antennas where each element is spaced by $\frac{\lambda}{2}$ as shown in Fig. 4. This additional antenna exists because each receiver antenna receives two reflected object signals from the first and second transmitters, as shown in Fig. 5(a).

For multiple receiver antennas, the signal received by each receiver increases the distance Δd to the reference point as shown in Fig. 6 and described as

$$\Delta d_n = d_n \sin\theta, \tag{6}$$

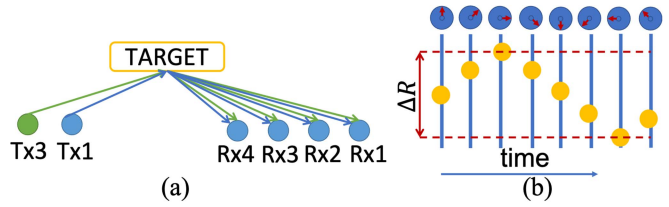

Fig. 5. (a) Transmitting and receiving FMCW signals with the MIMO concept, (b) The vibration of an object with different phase value.

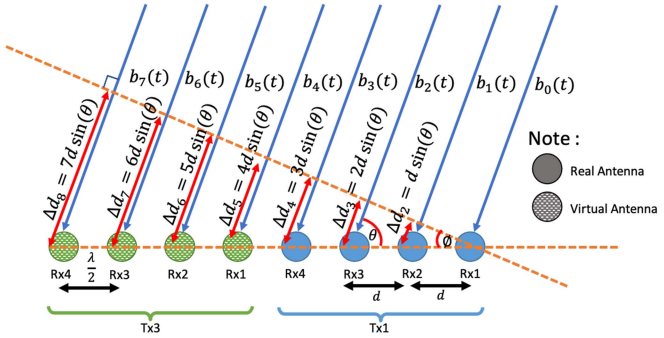

Fig. 6. Multiple receiver antenna concept.

where θ is the AoA and d_n as the antenna distance of receiver n. Therefore, the phase change for nth antenna becomes

$$\delta_n = \frac{2\pi}{\lambda} d_n \sin\theta, \tag{7}$$

and such as, AoA obtained is

$$\theta = \sin^{-1}\left(\frac{\delta_n \lambda}{2\pi d_n}\right). \tag{8}$$

This azimuth angle is obtained from the real-time delay of each virtual array element.

Detection of multiple objects in front of the radar is processed using 2D-FFT, AoA, and object tracking. Multiple objects in front of the radar will produce multiple peaks in the spectrum analysis. 2D-FFT and AoA analysis provide the position and angle of the object. If the object's position has been obtained, the detection of chest motion is focused on the tracking point. Furthermore, the extraction of breathing waveform at that location was carried out.

Based on the concept as mentioned earlier, for MIMO FMCW with N receiver antenna, there is an additional phase shift for the nth receiver antenna, which is arranged linearly. The received signal for the nth receiver antenna is given by

$$b_n(t) = \alpha A e^{j(2\pi f_b t + \Phi_b(t) + \Delta\phi(t) + \frac{2\pi}{\lambda} d_n \sin(\theta))}, \tag{9}$$

where $1 \leq n \leq N$ and d_n is the relative distance between the reference object and the nth receiver antenna. The basic idea for detecting a vibrating object such as Fig. 5(b) is to send multiple chirps separated by time. Each received chirp will have a different arrival delay. Because chest displacement has a tiny movement, the FFT range of each chirp will have peaks with the same location but with different phases. The phase difference of the signal ($\Delta\Phi_b$) represents chest displacement ΔR, which is caused by lung activities [15]. It can be calculated by the

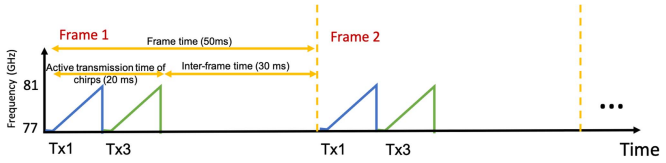

Fig. 7. FMCW chirps configuration consists of a sequence of chirps followed by idle time.

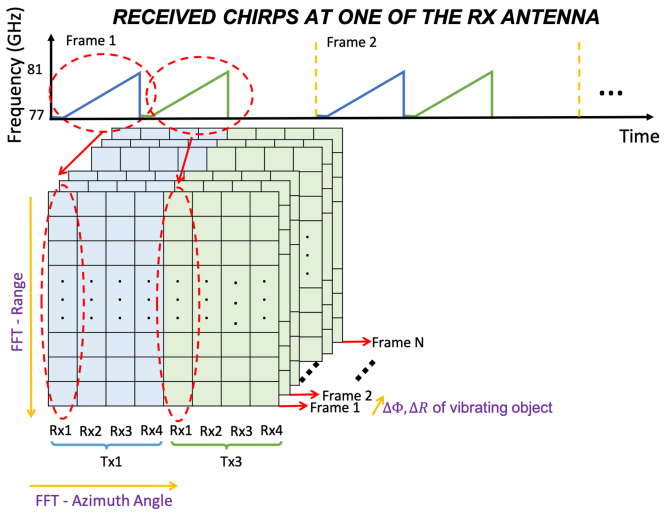

Fig. 8. The received chirps at one of the receiver antennas and the radar cube data.

equation

$$\Delta\Phi_b = \frac{4\pi\Delta R}{\lambda}, \tag{10}$$

where $\Delta\Phi_b$ is the phase change of the beat signal and ΔR is the position change. From this information, the phase shift can be obtained and used to analyze the signal in the frequency domain. To synthesize the total number of MIMO channels, the system must be able to distinguish the Tx channels based on the original antenna. One of the easiest ways is to use the time division technique. Using this method, the transmitting antenna alternately transmits its signal at different times. A chirp of two transmitters, transmitter one and transmitter three, is sent alternately using the Time Division Multiplexing (TDM) MIMO method as shown in Fig. 7. Each received chirp data is stored in each column of the radar cube presented in Fig. 8. The FFT on each sweep informs the range of the beat signal spectrum. FFT on the receiving channel provides information on the angular spectrum for each distance so that an azimuth-range plot can be obtained. The azimuth angle of the object will have a wide angular resolution of $\approx \frac{2}{N}$ radians.

Since we used TDM MIMO FMCW, the MIMO calibration is implemented with a conventional beamformer on the receiver side [48]. Each signal received by the receiver has a different arrival time. The signal received by all antennas is processed by Fourier transform followed by multiplication of the steering vector. The signal output is the combination of the sum of all sensors.

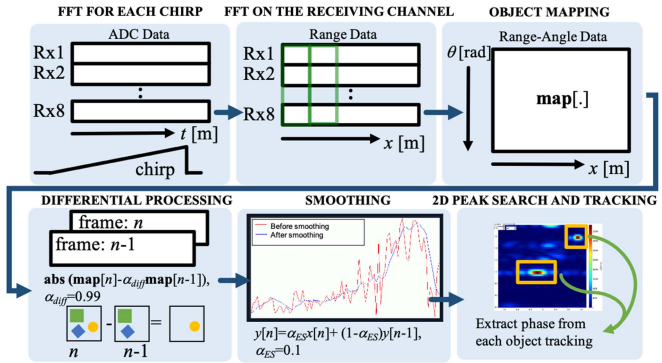

Fig. 9. Object tracking steps.

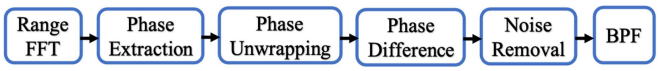

Fig. 10. Signal processing step for extracting the breathing waveform.

B. Object Tracking

In our previous study [30], 1D-FFT was used to detect an object in front of the radar. An object in front of the radar will produce a peak in the frequency domain analysis, as shown in Fig. 2. In 1D-FFT, the highest peak represents the object information. Detection of an object's breathing signal can be obtained only by doing once FFT process. If there are two objects in front of the radar, there will be two peaks in a received FMCW signal. The detected peaks through the 1D-FFT fluctuate in terms of frequency and magnitude. This will not be a problem in detecting one subject because the highest peaks can be directly identified as our object. If there are multiple subjects in front of the radar, there will be several peaks that need to be identified. Thus, the challenge in detecting multiple subjects is to separate and identify the information for each subject in a received FMCW signal so that each piece of information is not swapped with each other. Therefore, in this study, 2D-FFT and FMCW MIMO were used to track each object in a certain range and angle. With 2D-FFT, the position and angle of each subject can be obtained from the range-azimuth map. Besides, in this study, differential processing is used to cancel static objects, and smoothing processing is used to remove the high-frequency signals. If the subject has been identified, then the phase extraction is carried out on each detected object. Furthermore, the process of extracting the respiratory signal can be carried out for each subject. The object tracking steps is briefly described in Fig. 9.

C. Extracting Breathing Waveform

In this study, several signal processing techniques as shown in Fig. 10 are used to identify the breathing signals. Changes of the signal in each process can be seen in Fig. 11. The details of each block will be explained in the following parts.

1) RANGE - Fast Fourier Transformation (FFT): As mentioned earlier, the received signal $r(t)$ passed through the mixer and a Band Pass Filter (BPF) to produce a beat signal $b(t)$. This beat signal is sampled with an Analog to Digital Converter (ADC) and transformed to the frequency domain using FFT to produce a signal spectrum. The object distance to the radar can be determined by detecting the peak value of the signal spectrum, as shown in Fig. 2.

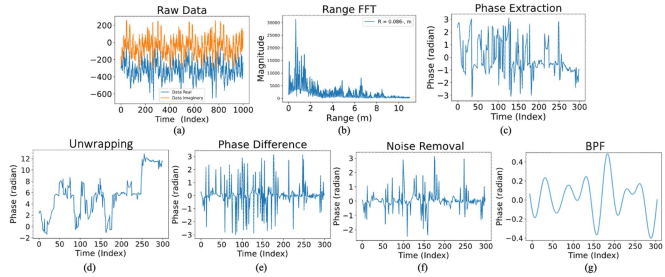

Fig. 11. Extracting breathing waveform: (a) raw data (beat signal) of FMCW in real and imaginary form, (b) output of range FFT that show the spectrum of the signal, (c) phase extraction output, (d) unwrapping output, (e) phase difference output and, (f) noise removal and BPF output.

Chest movement with a small amplitude <12 mm can be observed by applying slow time FFT [15]. To detect a small movement, the displacement distance is obtained from the phase change across multiple chirps, as shown in Fig. 8.

2) Extracting and Unwrapping the Signal: The phase value from chirp $\hat{m}$ is obtained by extracting and unwrapping the signal as

$$\Phi_b^{\hat{m}} = unwrap\left[\tan^{-1}\left(\frac{I}{Q}\right),\right] \tag{11}$$

where I and Q is imaginary and real part of the measured signals respectively. In this step, phase ambiguity should be prevented while unwrapping the signal. This phenomenon may occur because the operator $\tan^{-1}$ changes the radian phase into a phase with an interval of 2π under the $[-\pi, \pi]$ domain. The output of unwrapping should be in the range of $]-\pi, \pi]$, which is a zero centered, 2π length interval with the modulo-2π operator [49], with no loss of generality by subtracting π radian as

$$\Phi(t-1,t) = \begin{cases} \Delta\Phi_t, & \text{if } |\Delta\Phi(t)| < \pi \\ \Delta\Phi_t + 2\pi, & \text{if } \Delta\Phi(t) \leq -\pi \\ \Delta\Phi_t - 2\pi, & \text{if } \Delta\Phi(t) \geq -\pi, \end{cases} \tag{12}$$

where $\Phi(t)$ is the phase at the current time index t and $\Phi(t-1)$ is the phase at the previous time index. After extracting and unwrapping the signal in (9), the phase difference in (10) is obtained, and ΔR can be calculated.

3) Noise Removal: The unwrapped differential phase close to $-\pi$ or π suffers from a noise-induced phase wrapping error. This error can be eliminated by calculating the forward and backward phase difference for each time index t. Let $a(t)$ be the unwrapped differential phase at time index t. The forward differential phase is defined as $a(t) - a(t+1)$ and backward differential phase is defined as $a(t) - a(t-1)$. If the forward or backward differential phase exceeds the threshold then $a(t)$ will be replaced by the interpolated value. The output of phase difference BPF block without noise removal and output of phase difference BPF block with noise removal are presented in Fig. 12.

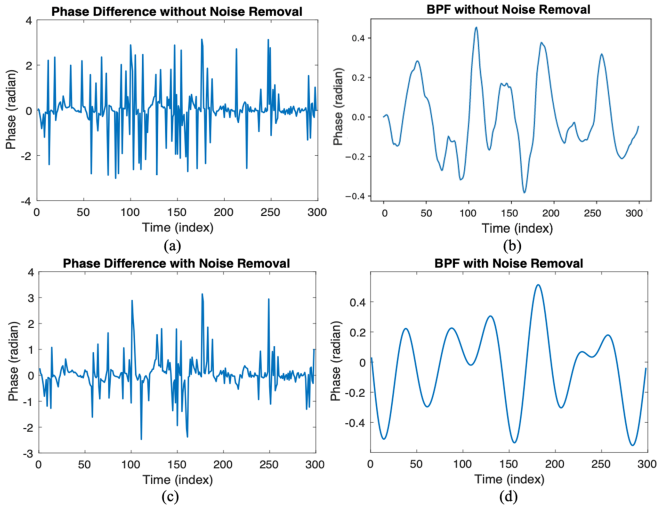

Fig. 12. Output of (a) phase difference, (b) BPF block without noise removal and output of (c) phase difference, (d) BPF block with noise removal.

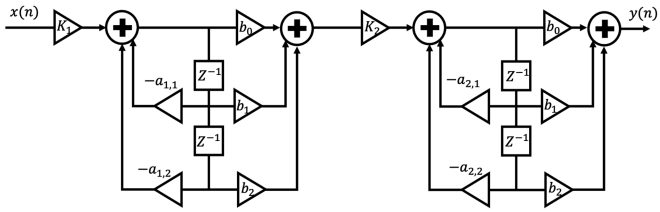

Fig. 13. Second order of IIR BPF using cascaded bi-quad.

4) 2-nd Order of IIR Cascaded Bi-Quad BPF: Human respiration causes a chest displacement in the range of $0.1-0.5$ Hz with an amplitude of $0-12$ mm [83]. To obtain this respiration signal, a BPF filter, which passes signals with a frequency between 0.1 and 0.5 Hz, is used. The *2-nd* order IIR cascade bi-quad BPF, as shown in Fig. 13, can be used to obtain the breathing signal [30]. The denominator coefficient $\tilde{a}$, numerator coefficient $\tilde{b}$, and the gain $\tilde{K}$ follow

$$\tilde{a} = \begin{bmatrix} 1 & \tilde{a}_{1,1} & \tilde{a}_{1,2} \\ 1 & \tilde{a}_{2,1} & \tilde{a}_{2,2} \end{bmatrix} = \begin{bmatrix} 1 & -1,963 & 0,964 \\ 1 & -1,850 & 0,868 \end{bmatrix},$$

$$\tilde{b} = \begin{bmatrix} \tilde{b}_0 & \tilde{b}_1 & \tilde{b}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}, \tilde{K} = \begin{bmatrix} 0,116 & 0,031 \end{bmatrix}$$

D. Machine Learning Development

This section presents the detailed machine learning process, such as data preparation steps, FE approach, classification techniques, hyperparameter optimization, and evaluation metrics in our research experiment.

1) Data Preparation: In this study, we used a breathing waveform data set that was manually collected and labeled. Each data in the data set is recorded for 5 seconds. The data set consists of 4000 data, divided into 80:20 percentage portions for training and testing respectively. It is approximately 3200 data for training and 800 data for testing the system. See Table I for the details. Most ML algorithms work best with an equal number of samples in each class since most algorithms are designed to maximize accuracy while minimizing errors. Hence, if the class distribution is not balanced, ML algorithms will perform poorly and simply predict the majority class in all cases [50], [51]. This case will lead to a phenomenon called "*accuracy paradox* [52]," the situation where a high accuracy value does not correspond to a high-quality model because the model is biased towards the majority class and can obscure the obtained results. However, to avoid the issue, we assumed that we had a balanced data set to train the model.

TABLE I
TRAINING AND TESTING DATA SET

Class	Training Data	Testing Data
Deep breathing	640	160
Deep and quick breathing	640	160
Hold breathing	640	160
Normal breathing	640	160
Quick breathing	640	160
Total	**3200**	**800**

Our data set contains five classes of breathing waveform as described in Table I. The reason for classifying into these five classes is related to the breathing pattern of several disease symptoms as described in the beginning of Section 2 and in [30].

Our focus is mainly to detect the symptoms of COVID-19 patients that mostly appear with quick and short breathing symptoms called tachypnoea. In this condition, the breathing rate is higher than the normal breathing, where the normal breathing rate for an average adult is 12 to 20 breaths per minute [53], [54] and can be higher in children. This critical breathing behavior needs supervised care and quick treatment. If it is not handled correctly, it can lead to a crucial situation of respiratory distress and cause a life-threatening issue for the patient.

Before applying the breathing waveform data set to our model, we need to normalize the data by converting them to a standard unified scale. The purpose of normalization is to reduce the unwanted variation between the data sets and allow the data on different scales to be compared.

In this study, we used z-score standardization to ensure that the distribution has the mean $(\mu) = 0$ and the standard deviation $(\sigma) = 1$. The formula for calculating the z-score of a point, x, is described as $z = \frac{x-\mu}{\sigma}$, where, x is the original feature vector. For this study, we set each data to be limited to 5 seconds with 85 step sizes.

2) Features Extraction (FE): Several FEs are performed to reduce the number of features in our data set. New features were created and selected from the original data set. Selected features represent some of the most important information from the original features. This study used statistical and Mel-Frequency Cepstral Coefficient (MFCC) FE described in the following paragraphs.

1) *Statistical:* Statistical-based FE method [55], [56] uses a statistical approach to determine the relationship between each input variable and the target variable. Besides, it selects the input variable with the strongest relationship with the target variable to be evaluated. This study uses statistical FE to calculate the basic statistical feature of respiratory signals such as mean, median, variance, standard deviation, absolute deviation, maximum, kurtosis, and

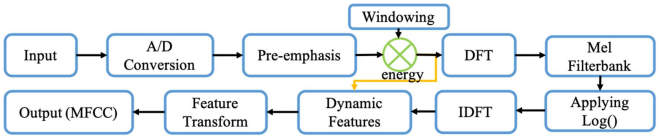

Fig. 14. MFCC FE technique.

skewness. Overall, eight statistical features were obtained. Then in the features selection step, we run an ANOVA F -value test [57], [58] to choose the best statistical features among them. The ANOVA F-value test can be defined by $F = \frac{between-groupvariability}{within-groupvariability}$ where *"between-group variability"* is

$$\sum_{i=1}^{\hat{K}} \frac{\hat{n}_i(\bar{Y}_{i.} - \bar{Y})^2}{(\hat{K}-1)} \tag{13}$$

where $\bar{Y}_{i.}$ denotes the sample mean in the ith group, $\hat{n}_i$ is the number of observations in the ith group, $\bar{Y}$ denotes the overall mean of the data, and $\hat{K}$ denotes the number of groups. *"within-group variability"* is,

$$\sum_{i=1}^{\hat{K}} \sum_{j=1}^{\hat{n}_i} \frac{(Y_{ij} - \bar{Y}_{i.})^2}{(\hat{N}-\hat{K})} \tag{14}$$

where Y_{ij} is the jth observation in the ith group out of $\hat{K}$ groups, and $\hat{N}$ is the overall sample size. This F-statistic follows the F-distribution with degrees of freedom $\hat{d}_1 = \hat{K} - 1$ and $\hat{d}_2 = \hat{N} - \hat{K}$ under the null hypothesis. The statistic result will be large if the *"between-group variability"* is larger relative to the *"within-group variability"*, which is unlikely to happen if the population mean of the groups all have the same value [59]. From the ANOVA F-value test result, we decided to choose only six features for our statistical features, namely mean, median, maximum, standard deviation, kurtosis, and skewness.

2) *MFCC:* Previously in [30] we proposed MFCC FE. MFCC method [60]–[62] involves windowing technique, applying the Discrete Fourier Transform (DFT), getting the magnitude log, then distorting the frequencies on the mel scale, followed by applying the Inverse Discrete Fourier Transform (IDFT) as shown in Fig. 14.
3) *MFCC-Stat:* Prior to training, we plotted our data (which had been extracted via FE) in a two-dimensional graph of linear discriminant analysis (LDA). LDA is a generative probabilistic model that can reduce the dimensions [63]. LDA attempts to project the sample onto a straight line with interclass projection points as close as possible and intraclass projection points as far apart as possible. Thus, we used LDA to see the effect of using different FE in Fig. 15. Based on the observation in Fig. 15, the scattering point of the data with the combination of both statistical and MFCC FE has fewer numbers of overlapping classes and is better separated than the one with MFCC FE only. Thus, we decided to use this approach for our proposed system and applied it to our data set.

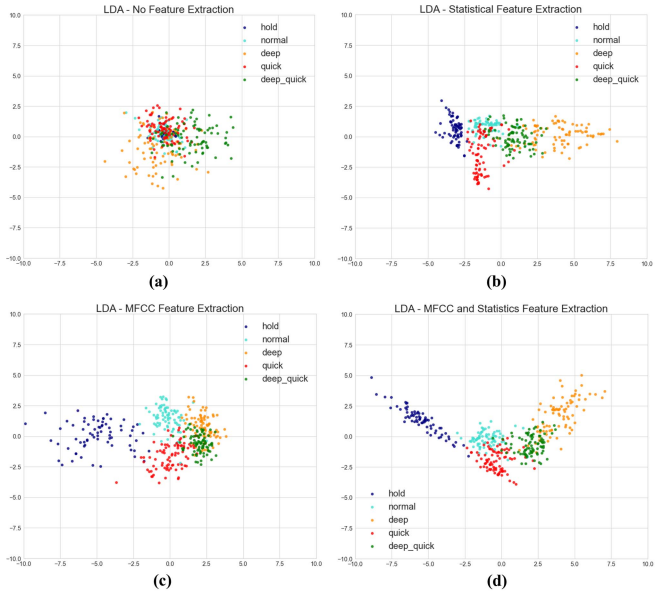

Fig. 15. LDA scattering point for (a) raw data, (b) data with statistics FE, (c) data with MFCC FE, and data with the combination of MFCC and statistics FE.

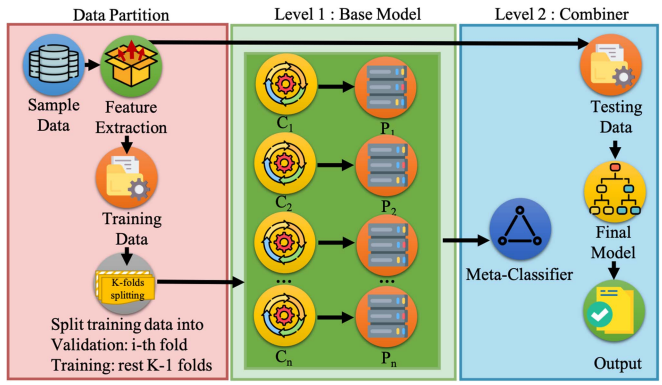

Fig. 16. Stacking-based ensemble learning framework for breathing waveform classification.

3) Classification Techniques: XGB classifier was used in our previous study to classify breathing waveforms in a real-time system [30]. However, we proposed a two-level classification model based on stacked ensemble learning in this study. In ensemble learning, several algorithms were utilized, and two-level predictions were implemented to get the best performance. Stacking technique learns high-level classifiers on top of the base classifiers [34], [64]–[66]. This technique can be regarded as a meta-learning approach. The base classifiers are called first-level classifiers. The second-level classifiers learns to combine the first-level classifiers. Thus, a framework of the proposed model is given in Fig. 16. The framework consists of a stack with the first level as the base model, which is comprised of several individual learners, and the second level as the meta-classifier (combiner), which takes those individual learners' predictions to get the final output prediction. Some classifiers used in this study are described in the following explanation.

TABLE II
PARAMETER SETTINGS FOR BASE AND META-CLASSIFIER OF THE PROPOSED STACKED ENSEMBLE MODELS

Model	Parameter	
	First-level classifiers	Meta-classifier
SEC	MLR DT RF SVM MLP	LGBM
BTSC	XGB LGBM CB	MLR
NSEM	XGB LGBM CB	MLP

1) Individual Classifier
In the stacked ensemble learning method, the first-level classifiers that consist of multiple individual classifiers are trained parallelly. Thus, to build a robust ensemble model, we utilize and explore several individual classifiers from single model algorithms such as MLR, DT, and SVM; homogeneous ensemble models such as RF, XGB, LGBM, and CB; to a neural network algorithm such as MLP.
2) Meta-classifier (Combiner)
Meta-classifier or combiner [67] trained multiple classifiers, either in a serial or parallel manner. To construct a stacked ensemble learning, we employ a meta-classifier in the second-level prediction. Meta-classifier will learn how to combine the output predictions made by individual classifiers in the first level. Thus, meta-classifier in the second level requires the presence of other learning algorithms that have already been trained on data. See Fig. 16 and Table II as a reference on how we built the meta-classifier for this study.

In this study, we combined the idea of stacking with K-fold cross-validation (CV) to avoid the over-fitting problem that commonly occurs in ensemble learning. K-fold CV is a widely used technique to evaluate classification performance. The modern form of stacking that uses internal K-fold CV was the idea of Breiman in [66]. Since then, stacking has been transformed into an algorithm known as the Super Learner [68], which learns the ideal combination of base learner fits. Algorithm 1 [64] demonstrates the general stacking procedure, which consists of three major steps:

- *Step 1:* Use K-fold CV to train first-level classifiers and prepare a training set for the second-level classifiers. Instead of using all the training example $\mathcal{D}$, we split $\mathcal{D}$ into K equal-size subset and then run the learning algorithm K times. For each subset, we learned a classifier from $K-1$ subsets and used the learned model to predict the remaining subset. The final prediction accuracy is averaged over K runs. The final prediction of the first-level classifiers will then be used as the input feature space for the second-level classifiers.
- *Step 2:* Construct a new data set based on the results of the first-level classifiers. Assume that each examples in $\mathcal{D}_k$ is $\{x_i, y_i\}$. We created a corresponding example $\{x'_i, y_i\}$ in the new data set, where $x'_i = \{h_{k1}(x_i), h_{k2}(x_i), \ldots, h_{kT}(x_i)\}$. Using the newly constructed data set, we train a second-level classifiers.
- *Step 3:* Use the second-level classifiers to combine the first-level classifiers. For an unobserved example x, the predicted class label of stacking is $h'\ (h_1(x), h_2\ (x), \ldots, h_T\ (x))$, where $\{h_1, h_2, \ldots, h_T\}$ are the first-level classifiers and h' is the second-level classifiers.

Algorithm 1: Stacked ensemble learning with K-fold CV [64], [66], [68]. $\mathbb{R}^n$ is the features space. $\mathcal{D}$ is the training data where $\{x_i, y_i\}$ is each example in $\mathcal{D}$. $\mathcal{D}_k$ is a data partition $\mathcal{D}$ in k subset of K-fold CV procedure. $\mathcal{Y}$ is the class label set. m is the number of training examples. T represents classifiers learned from $\mathcal{D}$. H is ensemble classifier, where h is the base classifier and h' is the second-level classifiers (meta-classifier).

input : Training Data $\mathcal{D} = \{x_i, y_i\}_{i=1}^{m}\ (x_i \in \mathbb{R}^n,\ y_i \in \mathcal{Y})$
output: An ensemble classifier H

Step 1: Adopt CV approach in preparing a training set for second-level classifiers;
Randomly split $\mathcal{D}$ into K equal-size subsets: $\mathcal{D} = \{\mathcal{D}_1,\ \mathcal{D}_2, \ldots,\ \mathcal{D}_K\}$;
for $k \leftarrow 1$ **to** K **do**
  Step 1.1: Learn first-level classifiers;
  for $t \leftarrow 1$ **to** T **do**
    Learn a classifier h_{kt} from $\mathcal{D} \setminus \mathcal{D}_k$;
  end
  Step 1.2: Construct a training set for second-level classifiers;
  for $x_i \in \mathcal{D}_k$ **do**
    Get record $\{x'_i, y_i\}$,where $x'_i = \{h_{k1}(x_i), h_{k2}(x_i), \ldots, h_{kT}(x_i)\}$;
  end
end
Step 2: Learn second-level classifiers;
Learn a new classifier h' from the collection of $\{x'_i, y_i\}$;
Step 3: Re-learn first-level classifiers;
for $t \leftarrow 1$ **to** T **do**
  Learn a classifier h_t based on $\mathcal{D}$;
end
return $H(x) = h'\ (h_1(x), h_2\ (x), \ldots, h_T\ (x))$

4) Hyperparameters Optimization: This study uses the search space, a dictionary where the hyperparameter arguments and values are used for the random search CV. As the classifier's performance is dependent on the hyperparameters, all candidate classifiers are optimized prior to selection. The optimization method used in this study is a random search with 10-fold CV where random combinations of hyperparameters are utilised to discover the optimal solution for each classifier. The algorithm assumes that not all hyperparameters are of equal significance. Random combinations of parameters are explored in each iteration of this search pattern. Due to the random search pattern, where the model may end up being trained on optimized parameters without aliasing, the probabilities of discovering the optimal parameter are considerably higher. The final values for each classifier based on the hyperparameter configurations that had the best performance across 10-fold CV are presented in Table III.

5) Evaluation Metrics: To assess the model performance, we calculate four different measurements of multiclass classification problem [69] [70]: (1) **Accuracy** $= \frac{TP+TN}{TP+FP+FN+TN}$, i.e., the proportion of accurately predicting the breathing waveform classes to all observed breathing waveform classes; (2)

TABLE III
HYPERPARAMETER SEARCH SPACE FOR EACH ALGORITHM. THE FINAL VALUE IS GENERATED FROM RANDOM SEARCH WITH 10-FOLD CV PROCEDURE

Models	Hyperparameter	Values	Final Value
MLR	C	0.1, 1.0, 10	10
	penalty	'l1','l2', 'elasticnet'	l2
	max_iter	100, 200,. . . ,1000	500
DT	max_features	'auto', 'sqrt', 'log2'	auto
	ccp_alpha	1e-1, 1e-2, 1e-3	1e-3
	max_depth	1,2,..., 10	10
	criterion	'gini', 'entropy'	entropy
	min_samples_split	2, 5, 10	2
	min_samples_leaf	1, 2, 4	1
RF	max_features	'auto', 'sqrt', 'log2'	sqrt
	ccp_alpha	1e-1, 1e-2, 1e-3	1e-3
	max_depth	3, 4, ..., 10	8
	criterion	'gini', 'entropy'	entropy
	min_samples_split	2, 5, 10	5
	min_samples_leaf	1, 2, 4	1
	n_estimators	100, 200, ...,1000	300
SVM	C	0.1, 1, 10, 100	10
	kernel	'linear', 'sigmoid', 'rbf'	linear
	gamma	1, 0, 1e-1, ..., 1e-4	1e-1
MLP	hidden_layer_sizes	(50,50), (50,100,50), (100,)	(50,100,50)
	max_iter	100, 200,...,1000	300
	activation	'logistic', 'tanh', 'relu'	tanh
	solver	'lbfgs', 'sgd', 'adam'	adam
	alpha	0.0001, 0.05	0.05
	learning_rate	'constant', 'adaptive'	adaptive
XGB	learning_rate	0.01, 0.2, 0.3, 0.5	0.2
	subsample	0.5, 0.6, ..., 1	1
	colsample_bytree	0.5, 0.6, ..., 1	1
	max_depth	3, 4, ..., 8	7
	n_estimators	100, 200, ..., 1000	700
	alpha	0.7, 1, 1.3	0.7
	gamma	0, 0.5, 1	0
LGBM	eta	1e-1, 1e-2, 1e-3	1e-1
	min_child_weight	1e-5, 1e-3, 1e-2, 1e-1, 1, 1e1, 1e2, 1e3, 1e4	1e-5
	subsample	0.1, 0.2,... , 1	0.4
	colsample_bytree	0.1, 0.2,... , 1	0.7
	max_depth	3, 4, 5, 6, 7, 8	5
	n_estimators	100, 200,... , 1000	900
	num_leaves	10, 20,..., 100	10
CB	learning_rate	0.03, 0.001, 0.01, 0.1, 0.2, 0.3	0.1
	depth	3, 4, ..., 10	8
	iterations	100, 200, ..., 1000	400
	l2_leaf_reg	1, 3, 5, 10, 100	1

Recall $= \frac{TP}{TP+FP}$, i.e., the proportion of correctly predicted positive breathing waveform class in actual class compared to all observations in a predicted class; (3) **Precision** $= \frac{TP}{TP+FN}$, i.e., the proportion of correctly predicting the positive breathing waveform class to the total predicted positive observations; and (4) **F1 score** $= 2 \times \frac{Precision \times Recall}{Precision + Recal}$, i.e., the weighted average of precision and recall.

Where TP, FP, TN, FN are

- True positive (TP): correct positive prediction occurs when the system *correctly* predicts that the patient *does* have a disease related to the measured class.
- False positive (FP): incorrect positive prediction (falsely predict positive) occurs when the system *incorrectly* predicts that the patient *does* have a disease related to the measured class. In actual condition, the patient does not have the related disease.
- True negative (TN): correct negative prediction occurs when the system *correctly* predicts that the patient *does not* have a disease.
- False negative (FN): incorrect negative prediction (falsely predict negative) occurs when the system *incorrectly* predicts that the patient *does not* have a disease related to the measured class. In actual condition, the patient does have the related disease.

In the medical problem, there is a critical miss classification case that requires specific attention, namely FN. FN indicates a condition when the system does not detect the class correctly even though the actual condition is positive for the measured class. An example case related to this study: The system does not detect the presence of COVID-19 suffered by the patient; instead, the system detects that the patient is healthy or suffered from other diseases such as asthma. This condition will be dangerous because the patient may not get proper treatment.

FP cases need to be considered as well although they might not be as dangerous as FN cases. FP indicates a condition when the system detects positive to the measured class, but the actual condition is negative to the measured class. FP cases are more tolerable because it is acceptable to classify non-COVID-19 sufferers as positive COVID-19 and follow up with more medical tests. However, for the FN case, it is clear that we should not miss in identifying a COVID-19 patient or classifying a COVID-19 patient as a non-COVID-19 sufferer. This will endanger the person itself and those around him.

These FP and FN cases can be observed from the precision and recall values of each class. Therefore, it is important to mention both the precision and recall for each class to ensure that the used algorithm is not only good in terms of accuracy but also

must be suitable for precision and recall values in each class. A system with a low recall has significant FN cases, while a system with low precision has an immense FP value. Since FN is more dangerous than FP, the recall has a higher priority than precision for evaluating the classification algorithm.

High recall with low precision implies that most classes have been detected, but most detections are false (high FP). High precision with low recall indicates that most predicted classes are correct, but most classes have not been detected (high FN). High precision with high recall implies an ideal classification algorithm that classifies all the classes correctly. In contrast, low precision with low recall means a poor classification algorithm that does not detect the most fundamental truth class (high FN), and most detections are false (high FP).

III. Experiment Results and Discussion

A. Machine Learning Model

The experiment was conducted by selecting the base models for the first-level classifier and followed by choosing the second-level classifier. In this step, we employ individual classifiers described in Section II, Subsection C on machine learning development. The purpose is to build an ensemble model based on stacking and select a second classifier as a meta-classifier. In building a good ensemble model, the diversity of base models and the meta-classifier selection plays an important role in achieving the best accuracy and generalization [71]–[73]. It is crucial to choose the most suitable classifier, especially in heterogeneous ensemble learning techniques such as stacking, as the meta-classifier parameter can aid in selecting a classifier that best fits the output of the base models. Therefore, the selection of base and meta-classifier combinations is always a matter of concern during the design of stacked ensemble architecture. In this study, we utilized different classifiers to select the most suitable classifier configuration for the system. We also proposed three stacked ensemble-based models using various combinations of base learner and meta-learner as explained in Table II with the following considerations:

- **SEC**: This model employs MLR, DT, RF, SVM, and MLP as our first-level classifiers based on the study and experiment in [74]–[76] and LGBM as the meta-classifier. Using various ML classifiers is suggested by [77]. In [77], they stated that stacking is effective when multiple machine learning models exhibit distinct abilities or have a low correlation. Thus, employing various models that make different assumptions about how to solve predictive modelling tasks as the base models can improve the quality of the stacking model. However, the idea to employ LGBM as a meta-classifier in our proposed model was influenced by the studies in [78], [79], and [80]. Using LGBM as a meta-classifier, the one-sided gradient algorithm can filter out samples with small gradients, eliminating unnecessary calculation and computing and optimizing features to accelerate the parallel computation. In addition, LGBM has outperformed the existing boosting frameworks in terms of efficiency and accuracy with much-reduced memory usage in numerous comparison studies using public datasets [81].
- **BTSC**: The idea of selecting XGB, LGBM, and CB as base models is because of their high performance in [82], [83]. With a more robust heterogeneous baseline model, we believe that we can achieve higher performance. Furthermore, [82], [83] use boosting algorithm which has proven to be successful with high performance for the base learner model. Hence, for the meta-classifier, we employed MLR. Based on [84] formulation, we could process meta-learning using stacking with a logistic regression (LR) classifier trained on the probabilistic output of several first-level classifiers. Several prior research have also utilized LR as a meta-classifier [85]–[88], and it has been proved that it can prevent the overfitting problem [84]. Although we could use other classifiers, the simple LR as a meta-classifier helps the system avoid overfitting conditions and lead to high-performance results. In addition, according to [89], the LR coefficients can be interpreted intuitively as the significant weight of each first-level classifiers. However, because this study is a multi-class classification problem, we used MLR as the extension of the LR model.
- **NSEM**: In this model, the meta-classifier is replaced with MLP. However, the decision to use MLP is motivated by the experiment of [82]. In their experiment, a boosting-based model combined with MLP as a meta-classifier outperformed both existing techniques and hybrid models for stacked ensemble learning. Furthermore, the research of [90] using MFCC FE demonstrated that an MLP as a meta-classifier improved the overall performance of the models. In addition, [91] also used MLP as one of their meta-classifier. Thus, we employed an MLP as our meta-classifier. MLP is a classification method based on neural networks that have been proven to be adaptable in modelling a complex problem [92]. MLP iteratively adjusts its parameters during training using the loss function's fractional derivatives. The chance of data samples matching a specific breathing waveform pattern was calculated using the cross-entropy loss function.

We performed the experiment in two steps: (1) we fed the MFCC-stat features data to these individual classifiers to evaluate our proposed stacked ensemble models' performance; (2) we proceeded with the experiments starting from SEC to NSEM. The experiment used the hyperparameter settings in Table III. These settings were obtained by performing a random search hyperparameter optimization to find the most optimal setting for our system. The experimental results for these experiments can be found in Table IV with its visualization in Fig. 17.

Fig. 17 depicts a visual comparison of the performance of all breathing classification methods. The box plot presents the performance of each model based on a 10-fold CV procedure with average accuracy and variance. The upper part of the box represents the highest data point, and the lower part represents the lowest data point. A horizontal red line within the rectangle represents the median of all values. The white diamond indicates the mean value. In Fig. 17, outliers did not appear for all algorithms. The more significant box-plot length indicates that the distribution of MLR, DT, and SVM values is more

TABLE IV
PERFORMANCES OF THE BREATHING CLASSIFICATION TASK (*MEAN* ± *STD* WITH 10-FOLD CV)

Models	Accuracy
MLR	0.885 ± 0.015
DT	0.845 ± 0.020
RF	0.920 ± 0.108
SVM	0.896 ± 0.018
MLP	0.963 ± 0.008
XGB	0.958 ± 0.006
LGBM	0.966 ± 0.010
CB	0.965 ± 0.010
Proposed stacked ensemble-based model	
SEC	0.967 ± 0.007
BTSC	0.967 ± 0.007
NSEM	**0.971 ± 0.007**

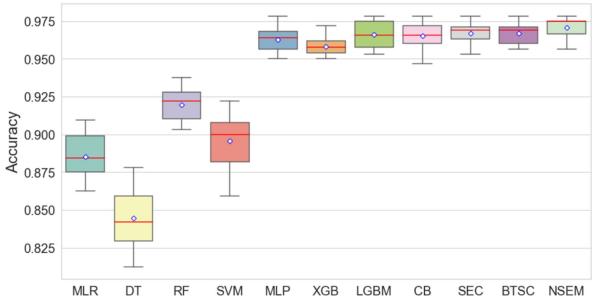

Fig. 17. Box plot of the accuracy for all algorithms with a 10-fold CV procedure.

varied than the other algorithms. Almost all high-performance algorithms have a narrow distribution of values, suggesting that the distribution contains values close to each other. However, in general, the mean accuracy of the NSEM model is the best among all algorithms, with the median value of the distribution located in the third quartile. It indicates that seventy-five percent of the values are below the upper quartile. Table IV summarizes the performance of each algorithm, including the mean accuracy and its standard deviation. The model's best performance is highlighted in **bold**.

Fig. 17 and Table IV show that even using a single classifier, the system yielded promising performance. The obtained accuracy using a single model such as MLR, DT, and SVM is quite good because these models can classify each class with an accuracy rate of more than 80%. On the other hand, using an ensemble-based model produces a better result than a single model algorithm. We can see those ensemble-based models such as RF, XGB, LGBM, and CB can achieve an accuracy rate of more than 90%. Additionally, MLP achieves comparable accuracy to ensemble-based models due to the complexity of multilayer networks.

Our three proposed models perform pretty well. SEC and BTSC models achieved a similar average accuracy of 96.7%, and the NSEM model achieved the best results with an average accuracy of 97.1%. This result outperforms the previous experiment in [30]. We show the detailed evaluation performance of these three proposed models using the confusion matrix in Fig. 18 followed by the classification report in Table V. As shown in Fig. 18, in all three models, the majority of the miss classifications were caused by incorrect predictions of deep quick breathing becomes quick breathing, normal breathing becomes quick breathing, normal breathing becomes deep quick breathing, and conversely. These same miss classifications also occurred in the previous study [30]. However, the intensity of the error was decreased in this experiment.

In medical problem, the cost of FN is greater than the cost of FP, which means that miss classifying a healthy person as a COVID-19 patient is less risky than miss classifying a positive COVID-19 patient as a healthy person, especially when the patient's life is at stake. As a result, recall is more critical than precision. As shown in Table V, all classes in the three proposed models have varied precision and recall performance, except for the hold breathing class, which has a perfect recall, precision, and F1-score performance. This ideal performance might happen because the breathing waveform of the holding class is less complex. Thus it is easier for the system to detect. COVID-19 patients often suffer quick and short breathing at unexpected times. This condition is associated with a quick breathing class. Therefore, the NSEM model is suitable for detecting COVID-19 patients due to its higher recall than its precision in recognizing quick/short breaths.

Overall, the results demonstrate that the classification performance of an ensemble-based model is superior to the single model algorithm. The performances of our three proposed models reflect an improvement of 0.5% to 1.3% compared to the homogeneous ensemble models (XGB, LGBM, and CB) and MLP. Furthermore, the result of this study achieved better performance compared to our previous study [30]. The combination of MFCC and statistic FE greatly contributes to this improvement. We can conclude that the proposed ensemble models well performed to classify five classes of our breathing waveform data.

B. Real-Time Measurement

The experiment was carried out in a room with the size of 3×3 m. In the beginning, a subject was asked to sit in front of the radar without any movement during the measurement, and the chest displacement was detected, as presented in Fig. 19(a). After going through some signal processing steps as described in Section 2B, the breathing waveform was obtained and shown in Fig. 19(b).

After the breathing signal was detected, the subject was asked to imitate the breathing pattern of each class, and the system recorded the signal. Fig. 20 presents five breathing pattern measurements: holding the breath, normal breath, quick breath, deep and quick breath, and deep breath. It can be seen that when the subject imitates a different breathing pattern, the system could detect a different breathing waveform as well, appropriate to the breathing characteristics of each class.

ML model was then used to classify the breathing pattern of multi-person in front of the radar. In this experiment, the subjects were asked to sit in front of the radar at different distances and angles, as shown in Fig. 21, and were asked to imitate different breathing waveforms simultaneously.

In the next experiment, we performed the detection of four subjects. The user interface of the system is shown in Fig. 22. By applying 2D-FFT, the system is able to provide the range-azimuth map, as shown in Fig. 22. It can be seen on the azimuth

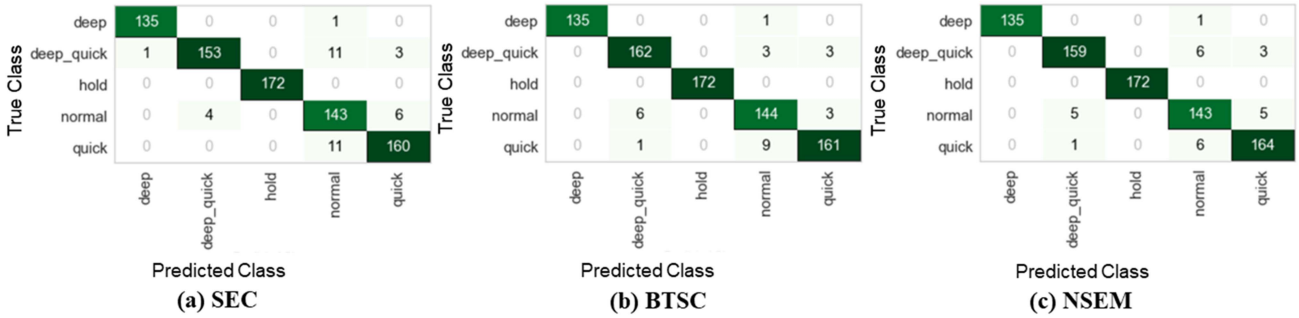

Fig. 18. The confusion matrix shows how the test values predicted classes compared to their actual classes.

TABLE V
THE CLASSIFICATION REPORT FOR THE PROPOSED ENSEMBLE-BASED MODEL

Class	SEC			BTSC			NSEM		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score	Precision	Recall	F1-Score
Deep	0.993	0.965	0.965	1	0.993	0.996	0.986	1	0.993
Deep quick	0.969	0.954	0.933	0.948	0.970	0.959	0.975	0.940	0.958
Hold	1	1	1	1	1	1	1	1	1
Normal	0.912	0.935	0.952	0.928	0.928	0.928	0.927	0.915	0.921
Quick	0.965	0.965	0.989	0.958	0.942	0.950	0.932	0.965	0.948

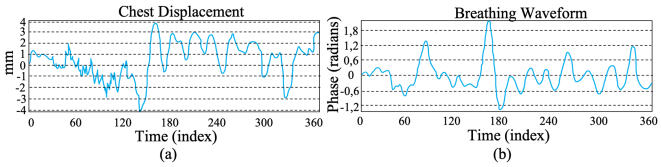

Fig. 19. (a) Chest displacement, (b) Breathing waveform.

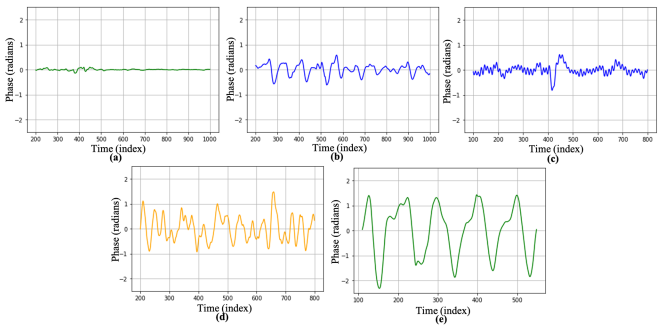

Fig. 20. Breathing waveform in the time domain for (a) holding the breath;(b) normal breath; (c) quick breath; (d) deep and quick breath; (e) deep breath recorded by TI-IWR1443.

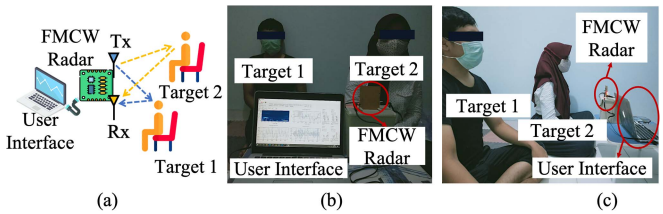

Fig. 21. (a) Experiment scenario, (b) photo of experiment environment from the front side and (c) from the right side.

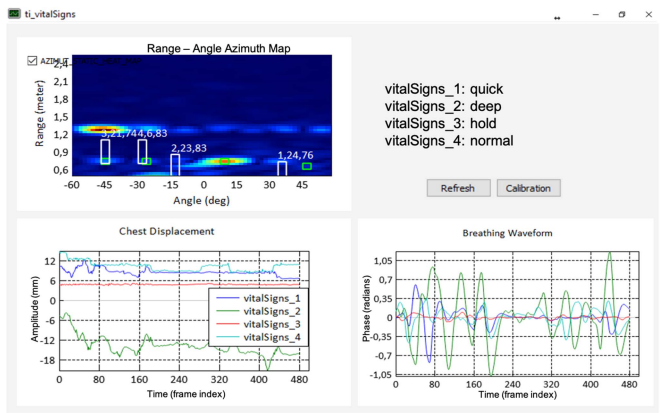

Fig. 22. Real-time detection for supervising four people simultaneously.

map that the four subjects were located in front of the radar at different distances of $0.8, 0.9, 0.6$ and 0 m with the angles of $-45^o, -30^o, -15^o$, and 39^o from the radar.

Furthermore, our user interface also presents the real-time measurement of chest displacement and breathing waveform. Based on Fig. 22, four subjects with each different type of breathing can be detected simultaneously in the system. The first subject at 0 m with an angle of 39^o in front of the radar was detected as having quick breathing. The second subject located at 0.6 m and -15^o in front of the radar was detected as having deep breathing. The third subject located at 0.8 m and -45^o in front of the radar was detected as imitating holding breath. The fourth subject, located at 0.9 m and -30^o in front of the radar, was detected as having normal breathing. The experiments demonstrate that the system successfully detects four subjects with different breathing waveforms simultaneously. Thus, the system can assist medical practitioners in simultaneously monitoring several patients with breathing disorders in a room.

The system's latency was evaluated by measuring the delay time as shown in Fig. 23 . The time takes for the system to activate the transmitter until displaying the breathing waveform

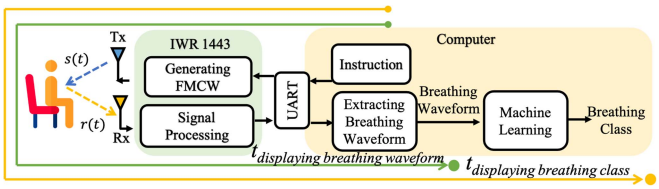

Fig. 23. Delay for displaying breathing waveform and breathing class.

TABLE VI
DELAY TIME OF THE SYSTEM

Delay Type	Time (second)
Displaying breathing waveform for the first time	0.113 ± 0.023
Updating breathing waveform	0.055 ± 0.007
Displaying breathing class for the first time	5.079 ± 0.039
Updating or classifying breathing class	0.059 ± 0.011

for the first time is defined as the displaying breathing waveform delay. It is defined as

$$t_{\text{displaying breathing waveform}} = t_{\text{UART}} + t_{\text{generating FMCW}} + t_{\text{FMCW travelling}} + t_{\text{signal processing}}. \quad (15)$$

The time takes for the system to activate the transmitter until classifying the breathing class for the first time is defined as displaying breathing class delay. It is defined as

$$t_{\text{displaying breathing class}} = t_{\text{UART}} + t_{\text{generating FMCW}} + t_{\text{FMCW travelling}} + t_{\text{signal processing}} + t_{\text{classification}} \quad (16)$$

In classifying the breathing waveform, each data in the data set was recorded for 5 seconds. Then we divided it into 85 windows. For the first classification result, the system needs to wait for the first 85 window sizes to be collected. Thus, the first classification will take longer than the rest since the system needs to wait for 85 window size for the first time.

In measuring the delay time in Table VI, we used an Intel(R) Core($^{\text{TM}}$) i7-4700HQ CPU with a speed of 2.4 GHz, 64-bit operating system, 8 GB of installed RAM with the speed of 1600 MHz. The measured delay might yield different results due to the radar hardware and computer specifications.

Based on the measurement, it took 5.079 ± 0.039 seconds for the system to display the breathing class for the first time. This delay happened only once at the beginning when the system started the process. In the next classification updates, the system uses the last 84 and 1 current window size. For updating the breathing class, it took 0.059 ± 0.011 seconds.

IV. CONCLUSION

This study proposed a non-contact monitoring and classification system for supervising patients with respiratory disorders, such as COVID-19 patients. For multiple-object detection, the FMCW radar calculated the range and angle of each object for tracking the object's position. FFT and some signal processing such as unwrapping phase, noise removal, and 2-nd order IIR cascaded bi-quad BPF were used to get the breathing signal of the subjects. Some important information on the respiratory signal was extracted by combining MFCC and statistical FE. Based on the test result, the combination of statistical and MFCC FE yielded better accuracy than the performance of each FE. Several classification algorithms such as MLR, DT, RF, SVM, XGB, LGBM, CB, MLP, and three proposed stacked ensemble models, namely SEC, BTSC, and NSEM, were used. Based on the results, the three proposed models well performed, with the SEC and BTSC models achieving an average accuracy of 96.7%, and the NSEM model achieving the best results with an average accuracy of 97.1%. Moreover, these results outperform 9.7% higher than the previous experiment that uses MFCC FE and XGB algorithm. In addition, all classifiers in this experiment perform satisfying results.

In a real-time implementation, multiple subjects with different breathing characteristics could be identified simultaneously, and the system was able to identify five different types of classes. The experimental results show that the system can be used to monitor the respiratory conditions of several patients in a room simultaneously without physical touch. It is hoped that this system can help medical practitioners to monitor the patient's condition. The development of this system can also be used to search and detect the presence of humans trapped under the rubble of buildings. This will be useful for saving people's lives in a disaster.

ACKNOLEDGMENT

The authors thank Prof. Trio Adiono and Dr Eng. Ayu Purwarianti for their valuable feedback, and Diyah Widiyasari for the help on simulations. We also thank Ahmad Wisnu Mulyadi from the Department of Brain & Cognitive Engineering, Korea University, for his advice during the revision process of this paper.

REFERENCES

[1] L. Pan et al., "Clinical characteristics of COVID-19 patients with digestive symptoms in Hubei, China: A descriptive, cross-sectional, multicenter study," *Amer. J. Gastroenterol.*, vol. 115, pp. 766–773, 2020.
[2] N. Poyiadji, G. Shahin, D. Noujaim, M. Stone, S. Patel, and B. Griffith, "COVID-19 associated acute hemorrhagic necrotizing encephalopathy: Imaging features," *Radiology*, vol. 296, no. 2, pp. E119–E120, 2020.
[3] J. Hellewell et al., "Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts," *Lancet Glob. Health*, vol. 8, no. 4, pp. e488–e496, 2020.
[4] C. Massaroni, D. Lo Presti, D. Formica, S. Silvestri, and E. Schena, "Non-contact monitoring of breathing pattern and respiratory rate via RGB signal measurement," *Sensors*, vol. 19, no. 12, 2019, Art. no. 2758.
[5] J. Cai, W. Sun, J. Huang, M. Gamber, J. Wu, and G. He, "Indirect virus transmission in cluster of COVID-19 cases, Wenzhou, China, 2020," *Emerg. Infect. Dis.*, vol. 26, no. 6, pp. 1343–1345, 2020.
[6] D. Dong et al., "The role of imaging in the detection and management of COVID-19: A review," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 16–29, 2020.
[7] P. Fonseca, X. Long, M. Radha, R. Haakma, R. Aarts, and J. Rolink, "Sleep stage classification with ECG and respiratory effort," *Physiol. Meas.*, vol. 36, no. 10, pp. 2027–2040, 2015.
[8] K. McClure, B. Erdreich, J. Bates, R. McGinnis, A. Masquelin, and S. Wshah, "Classification and detection of breathing patterns with wearable sensors and deep learning," *Sensors*, vol. 20, no. 22, pp. 6481–6493, 2020.
[9] Y. Wang, M. Hu, Q. Li, X. Zhang, G. Zhai, and N. Yao, "Abnormal respiratory patterns classifier may contribute to large-scale screening of people infected with COVID-19 in an accurate and unobtrusive manner," *arXiv.org*, Dec. 21, 2020. Accessed: Jul. 27, 2022. [Online]. Available: https://arxiv.org/abs/2002.05534

[10] S. Yoo et al., "Radar recorded child vital sign public data set and deep learning-based age group classification framework for vehicular application," *Sensors*, vol. 21, no. 7, pp. 2412–2427, 2021.

[11] J. Mühlsteff, J. Thijs, and R. Pinter, "The use of a two channel Doppler radar sensor for the characterization of heart motion phases," in *Proc. IEEE Int. Conf. Eng. Med. Biol. Soc.*, 2006, pp. 547–550.

[12] H. Zhao, H. Hong, L. Sun, Y. Li, C. Li, and X. Zhu, "Noncontact physiological dynamics detection using low-power digital-IF Doppler radar," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1780–1788, Jul. 2017.

[13] B. Schleicher, I. Nasr, A. Trasser, and H. Schumacher, "IR-UWB radar demonstrator for ultra-fine movement detection and vital-sign monitoring," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 5, pp. 2076–2085, May 2013.

[14] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart homes that monitor breathing and heart rate," in *Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst.*, 2015, pp. 837–846.

[15] A. Ahmad, J. C. Roh, D. Wang, and A. Dubey, "Vital signs monitoring of multiple people using a FMCW millimeter-wave sensor," in *Proc. IEEE Radar Conf.*, 2018, pp. 1450–1455.

[16] S. Ji, H. Wen, J. Wu, Z. Zhang, and K. Zhao, "Systematic heartbeat monitoring using a FMCW mm-wave radar," in *Proc. IEEE Int. Conf. Consum. Electron. Comput. Eng.*, 2021, pp. 714–718.

[17] K. Yamamoto, K. Toyoda, and T. Ohtsuki, "CNN-based respiration rate estimation in indoor environments via MIMO FMCW radar," in *Proc. IEEE Glob. Commun. Conf.*, 2019, pp. 1–6.

[18] J. Tu and J. Lin, "Fast acquisition of heart rate in noncontact vital sign radar measurement using time-window-variation technique," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 112–122, Jan. 2016.

[19] S. Pisa, E. Pittella, and E. Piuzzi, "A survey of radar systems for medical applications," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 31, no. 11, pp. 64–81, Nov. 2016.

[20] Y. Wang, W. Wang, M. Zhou, A. Ren, and Z. Tian, "Remote monitoring of human vital signs based on 77-GHz mm-wave FMCW radar," *Sensors*, vol. 20, no. 10, 2020, Art. no. 2999.

[21] F. Jin et al., "Multiple patients behavior detection in realtime using mmWave radar and deep CNNs," in *Proc. IEEE Radar Conf.*, 2019, pp. 1–6.

[22] L. Ding, M. Ali, S. Patole, and A. Dabak, "Vibration parameter estimation using FMCW radar," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2016, pp. 2224–2228.

[23] M. He, Y. Nian, and Y. Gong, "Novel signal processing method for vital sign monitoring using FMCW radar," *Biomed. Signal Process. Control*, vol. 33, pp. 335–345, 2017.

[24] M. C. Guerrero, J. S. Parada, and H. E. Espitia, "EEG signal analysis using classification techniques: Logistic regression, artificial neural networks, support vector machines, and convolutional neural networks," *Heliyon*, vol. 7, no. 6, 2021, Art. no. e07258.

[25] N. S. Bastos, B. P. Marques, D. F. Adamatti, and C. Z. Billa, "Analyzing EEG signals using decision trees: A study of modulation of amplitude," *Comput. Intell. Neurosci.*, vol. 2020, 2020, Art. no. 3598416.

[26] R. Rohan and L. V. R. Kumari, "Classification of sleep apneas using decision tree classifier," in *Proc. 5th Int. Conf. Intell. Comput. Control Syst.*, 2021, pp. 1310–1316.

[27] A. Ejupi and C. Menon, "Detection of talking in respiratory signals: A feasibility study using machine learning and wearable textile based sensors," *Sensors*, vol. 18, no. 8, 2018, Art. no. 2474.

[28] D. Miao, H. Zhao, H. Hong, X. Zhu, and C. Li, "Doppler radar-based human breathing patterns classification using support vector machine," in *Proc. IEEE Radar Conf.*, 2017, pp. 0456–0459.

[29] S.-H. Kim and G.-T. Han, "1D CNN based human respiration pattern recognition using ultra-wideband radar," in *Proc. Int. Conf. Artif. Intell. Inf. Commun.*, 2019, pp. 411–414.

[30] A. T. Purnomo, D.-B. Lin, T. Adiprabowo, and W. F. Hendria, "Non-contact monitoring and classification of breathing pattern for the supervision of people infected by COVID-19," *Sensors*, vol. 21, no. 9, Apr. 2021, Art. no. 3172.

[31] A. Paleczek, D. Grochala, and A. Rydosz, "Artificial breath classification using XGBoost algorithm for diabetes detection," *Sensors*, vol. 21, no. 12, 2021, Art. no. 4187.

[32] M. Bahoura, "Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes," *Comput. Biol. Med.*, vol. 39, no. 9, pp. 824–843, 2009.

[33] T. L. Nguyen and Y. Won, "Frequency features selection using decision tree for classification of sleep breathing sound," in *Proc. Inform. Sci. Appl.*, 2016, vol. 376, pp. 1375–1380.

[34] D. H. Wolpert, "Stacked generalizatin," *Neural Netw.*, vol. 5, pp. 241–259, 1992.

[35] S. Dzeroski and B. Zenko, "Is combining classifiers with stacking better than selecting the best one?," *Mach. Learn.*, vol. 54, no. 3, pp. 255–273, 2004.

[36] R. Maclin and D. W. Opitz, "Popular ensemble methods: An empirical study," *J. Artif. Intell. Res.*, vol. 11, pp. 169–198, 1999.

[37] T. G. Dietterich, "Ensemble methods in machine learning," in *Proc. 1st Int. Workshop Mult. Classifier Syst.*, 2000, pp. 1–15.

[38] D. Sarkar and V. Natarajan, *Ensemble Machine Learning Cookbook*. Birmingham, U.K.: Pack Publishing, 2019.

[39] G. Yuan, N. Drost, and R. McIvor, "Respiratory rate breathing pattern," *McMaster Univ. Med. J.: Clin. Rev.*, vol. 10, no. 1, pp. 23–25, 2013.

[40] Y. S. Lee, P. N. Pathirana, C. L. Steinfort, and T. Caelli, "Monitoring and analysis of respiratory patterns using microwave doppler radar," *IEEE J. Transl. Eng. health Med.*, vol. 2, pp. 1–12, 2014.

[41] L. Lum, "Hyperventilation syndromes in medicine and psychiatry: A review," *J. Roy. Soc. Med.*, vol. 80, no. 4, pp. 229–231, 1987.

[42] L. Whited and D. D. Graham, "Abnormal respirations". StatPearls Publishing - NCBI bookshelf, Jan. 2021. Accessed: Jul. 27, 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK470309

[43] W. C. Moore et al., "Characterization of the severe asthma phenotype by the national heart, lung, and blood institute's severe asthma research program," *J. Allergy Clin. Immunol.*, vol. 119, no. 2, pp. 405–413, 2007.

[44] J. E. Fergeson, S. S. Patel, and R. F. Lockey, "Acute asthma, prognosis, and treatment," *J. Allergy Clin. Immunol.*, vol. 139, no. 2, pp. 438–447, 2017.

[45] M. Rehman et al., "Improving machine learning classification accuracy for breathing abnormalities by enhancing data set," *Sensors*, vol. 21, no. 20, 2021, Art. no. 6750.

[46] G. M. Brooker, "Understanding millimetre wave FMCW radars," in *Proc. 1st Int. Conf. Sens. Technol.*, 2005, vol. 1, pp. 152–157.

[47] "Texas instrument IWR1443 single-chip 76-GHz to 81-GHz mmWave sensor integrating MCU and hardware accelerator," 2018. [Online]. Available: https://www.ti.com/product/IWR1443

[48] J. Guetlein, A. Kirschner, and J. Detlefsen, "Calibration strategy for a TDM FMCW MIMO radar system," in *Proc. IEEE Int. Conf. Microw., Commun., Antennas Electron. Syst.*, 2013, pp. 1–5.

[49] G. R. V. Matias, "Radar interferometry: 2D phase unwrapping via graph cuts," dissertation, Universidade Técnica De Lisboa Instituto Superior Técnico, Lisbon, Portugal, 2006.

[50] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.

[51] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J. Big Data*, vol. 6, pp. 1–54, 2019.

[52] T. Barros, P. Neto, I. Silva, and L. Guedes, "Predictive models for imbalanced data: A school dropout perspective," *Educ. Sci.*, vol. 9, pp. 1–17, 2019.

[53] M.-C. Huang et al., "Inconspicuous on-bed respiratory rate monitoring," in *Proc. ACM Int. Conf. Proc. Ser.*, 2013, pp. 1–8.

[54] D. Sankar, K. S. Durairaj, and K. Seethalaksmi, "Energy based feature extraction for classification of respiratory signals using modified threshold based algorithm," *Int. J. Eng. Sci. Technol.*, vol. 2, pp. 5488–5496, Oct. 2010.

[55] W. Caesarendra and T. Tjahjowidodo, "A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing," *Machines*, vol. 5, no. 4, 2017, Art. no. 21.

[56] B. Esmael, A. Arnaout, R. K. Fruhwirth, and G. Thonhauser, "A statistical Feature-based approach for operations recognition in drilling time series," *Int. J. Comput. Inf. Syst. Ind. Manage. Appl.*, vol. 4, no. 6, pp. 100–108, 2012.

[57] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn.*, vol. 3, pp. 1157–1182, Mar. 2003.

[58] N. Gayatri, S. Nickolas, and A. V. Reddy, "ANOVA discriminant analysis for features selected through decision tree induction method," in *Proc. Int. Conf. Comput. Commun. Syst.*, 2012, pp. 61–70.

[59] "F-test, Wikipedia, Wikimedia Foundation," Accessed on: Sep. 9, 2021. [Online]. Available: https://en.wikipedia.org/wiki/F-test

[60] M. B. Alsabek, I. Shahin, and A. Hassan, "Studying the similarity of COVID-19 sounds based on correlation analysis of MFCC," in *Proc. Int. Conf. Commun., Comput., Cybersecur., Inform.*, 2020, pp. 1–5.

[61] G. Chaudhari et al., "Virufy: Global applicability of crowdsourced and clinical data sets for AI detection of COVID-19 from cough," *arXiv.org*, Jan. 9, 2021. Accessed: Jul. 27, 2022. [Online]. Available: https://arxiv.org/abs/2011.13320

[62] J. Laguarta, F. Hueto, and B. Subirana, "COVID-19 artificial intelligence diagnosis using only cough recordings," *IEEE Open J. Eng. Med. Biol.*, vol. 1, pp. 275–281, 2020.

[63] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, 2003.

[64] Y. Li, J. Gao, Q. Li, and F. Wei, "Data classification: Algorithms and applications," *Data Mining Know. Discov. Ser.*, pp. 37–64, 2014.

[65] L. Kuncheva and C. Whitaker, "Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy," *Mach. Learn.*, vol. 51, pp. 181–207, 2003.

[66] L. Breiman, "Stacked regressions," *Mach. Learn.*, vol. 24, no. 1, pp. 49–64, 1996.

[67] M. A. Ganaie, M. Hu, M. Tanveer, and P. N. Suganthan, "Ensemble deep learning: A review," *arXiv.org*, Mar. 8, 2022. Accessed: Jul. 27, 2022. [Online]. Available: https://arxiv.org/abs/2104.02395

[68] M. J. van der Laan, E. C. Polley, and A. E. Hubbard, "Super Learner," *Stat. Appl. Genet. Mol. Biol.*, vol. 6, no. 1, 2007, Art. no. 25, doi: 10.2202/1544-6115.1309.

[69] M. Grandini, E. Bagli, and G. Visani, "Metrics for multi-class classification: An overview," *arXiv.org*, Aug. 13, 2020. Accessed: Jul. 27, 2022. [Online]. Available: https://arxiv.org/abs/2008.05756

[70] A. Zheng, *Evaluating Machine Learning Models*. O'Reilly Media, Inc., 2015.

[71] Z. Zhi-Hua, *Ensemble Methods: Foundations and Algorithms*. London, U.K.: Chapman and Hall Press, 2012.

[72] A. M. Canuto et al., "Investigating the influence of the choice of the ensemble members in accuracy and diversity of selection-based and fusion-based methods for ensembles," *Pattern Recognit. Lett.*, vol. 28, no. 4, pp. 472–486, 2007.

[73] L. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*. Hoboken, NJ, USA: Willey, 2004.

[74] J. Prusa, T. M. Khoshgoftaar, and D. J. Dittman, "Using ensemble learners to improve classifier performance on tweet sentiment data," *Proc. IEEE Int. Conf. Inf. Reuse Integration*, 2015, pp. 252–257.

[75] S. E. Seker and I. Ocak, "Performance prediction of readheaders using ensemble machine learning techniques," *Neural Comput. Appl.*, vol. 31, no. 4, pp. 1103–1116, 2019.

[76] B. Y. Tama, L. Nkenyereye, S. M. R. Islam, and K. S. Kwak, "An enhanced anomaly detection in web traffic using a stack of classifier ensemble," *IEEE Access*, vol. 4, pp. 1–15. 2020.

[77] M. Lanes, P. F. Schiavo, S. F. Pereira Jr., E. N. Borges, and R. Galante, "An analysis of diversity on stacking supervised classifiers," in *Proc. 19th Int. Conf. Enterprise Inf. Syst.*, 2017, vol. 1, pp. 233–240.

[78] J. Zhou et al., "Automatic sleep stage classification with single channel EEG signal based on two-layer stacked ensemble model," *IEEE Access*, vol. 8, pp. 57283–57297, 2020, doi: 10.1109/ACCESS.2020.2982434.

[79] M. Massaoudi, H. Abu-Rub, S. S. Refaat, I. Chihi, and F. S. Oueslati, "An effective ensemble learning approach-based grid stability assessment and classification," in *Proc. IEEE Kansas Power Energy Conf.*, 2021, pp. 1–6, doi: 10.1109/KPEC51835.2021.9446197.

[80] N. Ke, G. Shi, and Y. Zhou, "Stacking model for optimizing subjective well-being predictions based on the CGSS database," *Sustainability*, vol. 13, no. 21, pp. 1–17, 2021.

[81] G. Ke et al., "LightGBM: A highly efficient gradient boosting decision tree," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 3149–3157.

[82] M. Massaoudi, S. S. Refaat, I. Chihi, M. Trabelsi, F. S. Oueslati, and H. A. Rub, "A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for short-term load forecasting," *Sustainability, Energy*, vol. 214, 2020, Art. no. 118874.

[83] X. Guo, Y. Gao, D. Zheng, Y. Ning, and Q. Zhao, "Study on short-term photovoltaic power prediction model based on the stacking ensemble learning," *Energy Rep.*, vol. 6, pp. 1424–1431, 2020.

[84] K. M. Ting and I. Witten, "Issues in stacked generalization," *J. Artif. Intell. Res.*, vol. 10, pp. 271–289, 1999.

[85] S. L. Javan, M. M. Sepehri, M. L. Javan, and T. Khatibi, "An intelligent warning model for early prediction of cardiac arrest in sepsis patients," *Comput. Methods Programs Biomed.*, vol. 178, pp. 47–58, 2019.

[86] J. Niestroy, J. Han, J. Luo, R. Zhao, D. E. Lake, and A. Flower, "Prediction of decompensation in patients in the cardiac ward," in *Proc. Syst. Inf. Eng. Des. Symp.*, 2019, pp. 1–6, doi: 10.1109/SIEDS.2019.8735602.

[87] M. L. Williams, W. P. James, and M. T. Rose, "Variable segmentation and ensemble classifiers for predicting dairy cow behaviour," *Biosyst. Eng.*, vol. 178, pp. 156–167, 2019.

[88] N. T. Cockroft, X. Cheng, and J. R. Fuchs, "STarFish: A stacked ensemble target fishing approach and its application to natural products," *J. Chem. Inf. Model.*, vol. 59, no. 11, pp. 4906–4920, 2019.

[89] A. Altmann et al., "Comparison of classifier fusion methods for predicting response to anti HIV-1 therapy," *PLoS ONE*, vol 3, no. 10, pp. 1–9, 2008.

[90] N. AbaeiKoupaei and H. Al Osman, "A multi-modal stacked ensemble model for bipolar disorder classification," *IEEE Trans. Affect. Comput.*, to be published, doi: 10.1109/TAFFC.2020.3047582.

[91] H. Ehya, T. Skreien, and A. Nysveen, "Intelligent data-driven diagnosis of incipient inter-turn short circuit fault in field winding of salient pole synchronous generators," *IEEE Trans. Ind. Inform.*, vol. 18, no. 5, pp. 3286–3294, May 2022.

[92] A. Meyer-Baese and V. Schmid, *Pattern Recognition and Signal Analysis in Medical Imaging*. New York, NY, USA: Academic Press, 2014.

[93] C. Massaroni, A. Nicolò, D. Lo Presti, M. Sacchetti, S. Silvestri, and E. Schena, "Contactbased methods for measuring respiratory rate," *Sensors*, vol. 19, no. 4, 2019, Art. no. 908.

[94] F.-T.-Z. Khanam et al., "Remote monitoring of vital signs in diverse non-clinical and clinical scenarios using computer vision systems: A review," *Appl. Sci.*, vol. 9, no. 20, 2019, Art. no. 4474.

[95] S. Wang et al., "A novel ultra-wideband 80 GHz FMCW radar system for contactless monitoring of vital signs," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2015, pp. 4978–4981.

[96] J. Herron, A. Hay-David, A. Gilliam, and P. Brennan, "Personal protective equipment and COVID 19-a risk to healthcare staff?," *Brit. J. Oral Maxillofac. Surg.*, vol. 58, no. 5, 2020, Art. no. 500.

[97] A. Green, "A tribute to some of the doctors who died from COVID-19," *Lancet*, vol. 396, no. 10264, 2020, Art. no. 1720.

[98] H. Pratiwi, M. R. Hidayat, A. Pramudita, and F. Y. Suratman, "Improved FMCW radar system for multi-target detection of human respiration vital sign," *Jurnal Elektronika dan Telekomunikasi*, vol. 19, no. 2, pp. 38–44, 2019.

[99] Y. Li and S. O'Young, "Method of doubling range resolution without increasing bandwidth in FMCW radar," *Electron. Lett.*, vol. 51, no. 12, pp. 933–935, 2015.

[100] C. Gu, "Short-range noncontact sensors for healthcare and other emerging applications: A review," *Sensors*, vol. 16, no. 8, 2016, Art. no. 1169.

[101] C.-H. Hsieh, Y.-F. Chiu, Y.-H. Shen, T.-S. Chu, and Y.-H. Huang, "A UWB radar signal processing platform for real-time human respiratory feature extraction based on four-segment linear waveform model," *IEEE Trans. Biomed. circuits Syst.*, vol. 10, no. 1, pp. 219–230, Feb. 2016.

[102] C. Büyükhan, S. Karamzadeh, M. Orhan, A. Caliskan, and I. Cilesiz, "Experimental study of human respiratory detection using UWB GPR," in *Proc. 4th Int. Conf. Elect. Electron. Eng.*, 2017, pp. 314–317.

[103] A. Lazaro, D. Girbau, and R. Villarino, "Analysis of vital signs monitoring using an IR-UWB radar," *Prog. In Electromagn. Res.*, vol. 100, pp. 265–284, 2010.

[104] K. Ramasubramanian, "MmWave radar for automotive and industrial applications – TI training," *mmWave Radar Automot. Ind. Appl.*, Dec. 7, 2017. Accessed: Jul. 27, 2022. [Online]. Available: https://training.ti.com/sites/default/files/docs/Mmwave_webinar_Dec2017.pdf

[105] S.-H. Kim, Z. W. Geem, and G.-T. Han, "A novel human respiration pattern recognition using signals of ultra-wideband radar sensor," *Sensors*, vol. 19, no. 15, 2019, Art. no. 3340.