RF SENSORS FOR MEDICAL AND CYBER-PHYSICAL INTELLIGENCE

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Zijing Zhang

May 2023

RF SENSORS FOR MEDICAL AND CYBER-PHYSICAL INTELLIGENCE

Zijing Zhang, Ph. D.

Cornell University 2023

My research has focused on continuous and non-invasive sensing of physiological signals including respiration, muscle activities, heartbeat dynamics, and other biological signals. I seek to establish a touchless RF sensor that can be implemented as wearables on users, or integrated into the furniture to become invisible to the user. Such sensor can greatly enhance data continuity, comfort and convenience to enable many healthcare applications, especially for at-home continuous diagnosis and prognosis, with less reliance on subjective self report. My research utilized machine-learning (ML) algorithms that can take the physiological data from our sensors to provide holistic diagnostics and prognosis. This sensor has been applied to pulmonary diseases including COVID-19 and chronic obstructive pulmonary diseases (COPD) to help identify dyspneic exacerbation, leading to early intervention and possibly improving outcome. The sensor has also been applied to prevalent sleep disorders such as apnea and hypopnea.

Another aspect of my research focuses on muscle monitoring. Conventional electromyography (EMG) measures the neural activity during muscle contraction, but lacks explicit quantification of the actual contraction. I proposed radiomyography (RMG), a novel muscle wearable sensor that can non-invasively and continuously capture muscle contraction in various superficial and deep layers. Continuous

monitoring of individual skeletal muscle activities has significant medical and consumer applications, including detection of muscle fatigue and injury, diagnosis of neuromuscular disorders such as the Parkinson's disease, assessment for physical training and rehabilitation, and human-computer interface (HCI) applications. I verified RMG experimentally on a forearm wearable sensor for extensive hand gesture recognition, which can be applied to various applications including assistive robotic control and user instructions. I also demonstrated a new radiooculogram (ROG) for non-invasive eye movement monitoring with eyes open or closed. ROG is promising for gaze tracking and study of sleep rapid eye movement (REM).

BIOGRAPHICAL SKETCH

Zijing Zhang was born in Wuhan, Hubei, China in 1997, the year marked by the handover of Hong Kong from the United Kingdom to China. Thus, she was named after the national flower of Hong Kong, Bauhinia Blakeana, in Chinese. She received a B.Eng. degree in optoelectronic engineering from Huazhong, University of Science and Technology (HUST), Wuhan, China, in 2019. During her undergraduate period, she was an exchange student to University of California, San Diego in 2017 Fall and a research intern in Georgia Institute of Technology in 2018 Summer. She started pursuing her Ph.D. degree at Cornell University since Fall 2019. In the summer of 2021, she was a signal processing and machine learning intern at Analog Devices, Wilmington, MA. Zijing Zhang is now a Ph.D. candidate in the School of Electrical and Computer Engineering at Cornell University. She works with Prof. Edwin Kan and focuses on radio-frequency systems for biosensing, machine learning in healthcare, and human computer interface.

This dissertation is dedicated to

my parents and

people supporting my journey.

ACKNOWLEDGMENTS

This thesis and research work has been a result of support and encouragement from many wonderful people.

First of all, I would like to express my sincere gratitude to my advisor Professor Edwin C. Kan, for his consistent support and guidance. Prof. Kan has encouraged me with high patience especially during the first year so I can gradually make transition to a research mindset at the very beginning. I would also like to thank my PhD committee members, Prof. Peter Doerschuk and Vikram Krishnamurthy, for their valuable advice and insights. I would also hope to extend my appreciation to my mentoring collaborators, Dr. Ana Krieger, Dr. Veerawat Phongtankuel, and Prof. Huiju Park.

I would also like to thank my group mates in Kan's lab, Dr. Pragya Sharma, Dr. Xiaonan Hui, Dr. Jianlin Zhou, Guoyi Xu, Thomas Conroy, Aakash Kapoor, Kapil Gangwar, Upekha Delay, Brian Hsu, and Hannah Goldstein. My senior colleagues of Pragya Sharma, Xiaonan Hui, Jianlin Zhou, and Guoyi Xu, and Thomas Conroy have provided a lot of guide and advices during my PhD in the group.

Last but not least, I would like to express my gratitude to my parents, Lianfang Jiang and Wusheng Zhang. They have been supportive to every decision I made and they gave me full freedom to choose the direction I want.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTOIN

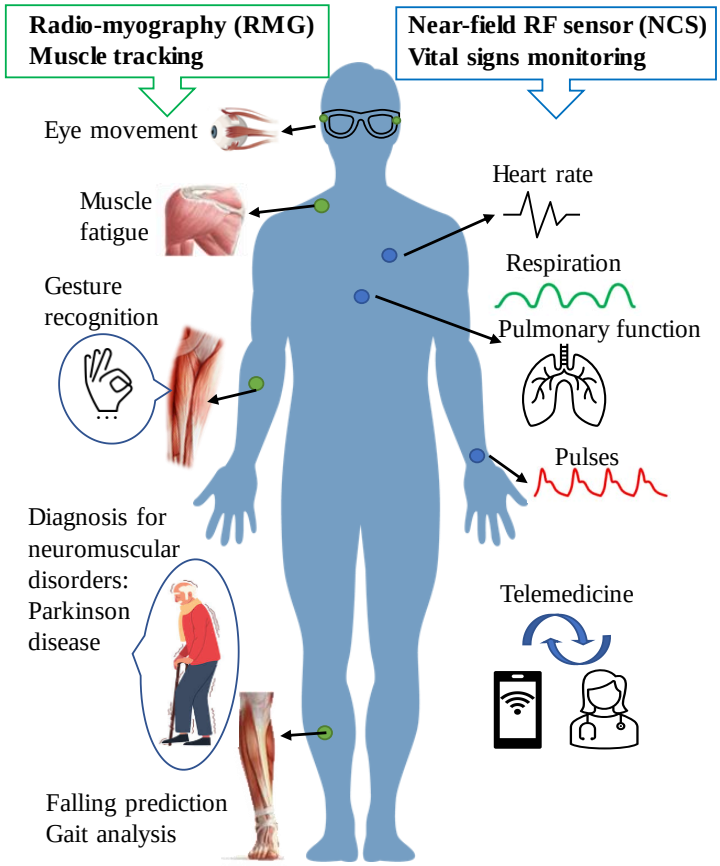


Fig. 1.1 Overview of my research topics on biosensors.

The world in the next century will face many unprecedented challenges including, but not limited to, over population, global warming, sustainable health care, mental disorders, pandemic countermeasure, and equality with diversity. However, our generation also accumulates an unprecedented amount of knowledge and technology that can potentially enable us to help ward off such eminent threats. I hope my engineering career would make an impact to provide solutions, or tools to

solutions, to a small subset of these global challenges.

Thesis organization

This dissertation presents two main contributions in the area of human sensing technologies focusing on 1. vital signs and 2. muscle monitoring.

1. Sensing vital signs for telemedicine and medical diagnosis:

In chapter 2, I presented details on the first sensing system I developed for vital signs monitoring: “Furniture-integrated respiration sensors by notched transmission lines”. In this chapter, I focused on the principle of RF near-field sensing, sensor configurations, signal model, and system performance on Breath rate and Heart rate monitoring. In this work, I established a touchless and continuous system utilizing RF sensors that can be implemented as wearables, or integrated to the furniture to become invisible to the user. This can greatly enhance accuracy, comfort and convenience to enable many healthcare applications, especially for at-home continuous diagnosis. Chapter 1 laid the foundation of sensing hardware for future applications in medical diagnosis.

In Chapter 3, I demonstrated a new approach to invisibly and precisely identify prevalent sleep disorders using our RF sensors: “ Identification and prediction of sleep disorder by covert bed integrated RF sensors”. Respiratory disturbances during sleep are a prevalent health condition that affects a large adult population. Current sleep disorder symptoms are mostly scored by human operators during sleep studies, which is time-consuming. This work is to develop an autonomous system to detect and predict respiratory events reliably based on real-time covert sensing. Predictive

warning of the sleep disorders in advance can intervene serious apnea, especially for infants, servicemen, and patients with chronic conditions.

In Chapter 4, I presented an important application in pathological diagnostics for respiratory diseases using RF near-field sensors: “Objective scoring of physiologically induced dyspnea by non-invasive RF sensors.” Modern medicine has made great progress in the molecular level, but most devices of symptomatic screening and pathological diagnostics are still difficult to operate continuously as wearables or outside clinics. For example, dyspnea, pain, and physical/emotional discomfort still rely heavily on self-report, which is not only subjective and unreliable, but also inapplicable to people who cannot communicate or unwilling to cooperate. This chapter focuses on dyspnea, also known as the patient’s feeling of difficult or labored breathing, is one of the most common symptoms for respiratory disorders. My work developed a learning-based model that can evaluate the correlation between the self-report Borg score and the respiratory metrics for dyspnea. The method can formulate a baseline for clinical dyspnea assessment and help caregivers track dyspnea continuously, especially for patients who cannot report themselves.

In Chapter 5, I extended the dyspnea scoring model in Chapter 4 for clinical study on COVID-19 and presented the result as “Objective dyspnea evaluation on COVID-19 patients learning from exertion-induced dyspnea scores”. Overnight (~16h) respiratory waveforms were collected on 12 COVID-19 patients, and a benchmark on 13 healthy subjects. This paper validates the viability to use our objective dyspnea scoring for clinical dyspnea assessment on COVID patients. The proposed system can help the identification of dyspneic exacerbation in conditions

such as COVID, leading to early intervention and possibly improving their outcome.

2. Muscle Monitoring for human computer interaction and biomedical application:

In Chapter 6, for the second part of my research, I developed a novel muscle sensing technique by RF sensors and explored applications to hand gesture recognition. Direct muscle sensing can potentially bring forth a paradigm shift for biomedical applications and human computer interface (HCI), as such continuous sensing does not exist before. Conventional electromyography (EMG) measures the continuous neural activity during muscle contraction, but lacks explicit quantification of the actual contraction. Here we propose a novel radiomyography (RMG) for continuous muscle actuation sensing that can be wearable and touchless, capturing both superficial and deep muscle groups. We verified RMG experimentally by a forearm wearable sensor for detailed hand gesture recognition. RMG can be used with synchronous EMG to derive stimulation-actuation waveforms for many potential applications in kinesiology, physiotherapy, rehabilitation, and human-machine interface.

In Chapter 7, I extended the RMG to a different application, “Radiooculogram (ROG) for eye movement sensing with eyes closed”. This work presents radiooculogram (ROG), a novel sensor for non-invasive eye movement (EM) monitoring with eyes closed. Compared with biopotential-based sensors, ROG has higher user comfort due to touchless operation and can capture direct muscle activity even in deep tissues. This work on voluntary EM sensing can serve as the baseline implementation for eventual sleep rapid EM monitoring.

In Chapter 8, I concluded my overall intellectual contribution of my PhD thesis and touched on the future work.

I have a strong belief that my research in RF sensors can ultimately contribute to enhance healthcare practice and delivery, providing highly personalized and precisely targeted diagnostics as well as improving the well-being of the public without overburdening the social resources.

CHAPTER 2

FURNITURE-INTEGRATED RESPIRATION SENSORS BY NOTCHED TRANSMISSION LINES

Introduction

Monitoring respiration dynamics has many applications [1], including in wellness evaluation and diagnosis of respiratory disorders such as asthma [2], coughing, dyspnea, and chronic obstructive pulmonary diseases (COPD) [3]. It is also valuable for prognostic assessment of prevalent sleep disorders including central sleep apnea (CSA) and obstructive sleep apnea (OSA) [4]. Irregular breathing patterns can be an early indicator of cardiac arrest [5] and acute myocardial infarction [6]. Besides clinical applications, respiration monitoring can also be a useful gauge of cognitive load [7], emotional stress, and physical efforts during sports and exercises [8].

Current methods to monitor respiration efforts include spirometry [9], nasal airflow probe [10], capnography [11], strain gauges [12], respiratory inductance plethysmography (RIP) [13], pulse oximetry [14], skin strain sensors [15], optical and far-field RF sensing [16][17], and ballistocardiogram (BCG) [18]. Spirometry measures the volume and speed of inhaled and exhaled air by the lung, but requires attentive user participation. Nasal airflow probes utilize mouthpieces or facemasks, which are cumbersome and disruptive to the normal breathing patterns. Capnography needs similar conspicuous mouth or nasal cannula. Strain gauges and RIP by chest belts can measure the chest and abdominal wall movement during respiration, but the

belt tension is required to capture the full breathing cycle which causes discomfort. The skin strain sensor involves tight skin contact which is also uncomfortable. Pulse oximetry monitors the oxygen saturation (SpO2) and offers mixed observation of respiration and circulation. Optical and RF sensing methods require the user to be stationary and the reader to be in a direct line-of-sight (LoS) to the front torso. BCG-based sensors often assume the user body weight as the quiescent point of operation, and can only be installed under the bed but not to the chair. Posture variation and movement on bed are problematic as well.

Transmission-line sensors have been used in biological and chemical sensing as well as health monitoring. A splitter-combiner structure can measure minute dielectric property changes in the near-field region [19]. A coplanar waveguide loaded with symmetric resonators can measure angular motion [20] and dielectric properties [21]. A resonator coupled to a transmission line was used to estimate glucose concentration [22]. The sensing modality is typically the variation of the resonance frequency, phase, and/or quality factor. The transmission-line structure is highly sensitive to small motion and composition changes, and hence has been rarely applied to the large motion of respiration. The highest sensitivity may not be favorable here, as tolerance to user variation and the capture of the full waveforms are also important in practical applications.

The notched transmission-line sensor proposed in this work is adapted from the previous antenna-based near-field coherent sensing (NCS) [23], as shown in Fig. 1. In the near-field region, the dielectric boundary movement by internal organs and body parts will couple to the leaked electromagnetic (EM) energy from the notched part of

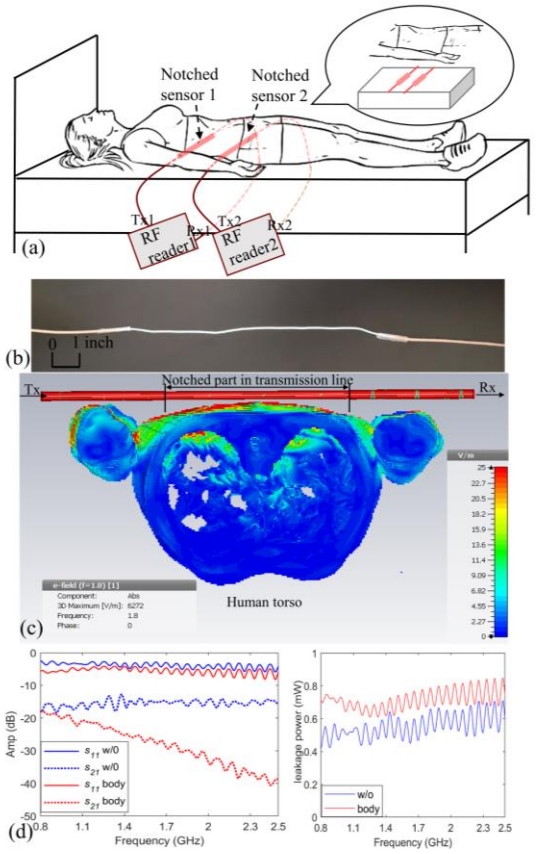


Fig. 1 Notched-cable respiration sensor setup: (a) System schematic with bed integration; (b) A photo of the notched cable; (c) Cross section of the CST simulation of a human torso with the NCS frequency at 0.9GHz; (d) Left:Measured S11 and S21 results in 0.8 – 2.5 GHz after de-embedding. Right: Estimated leakage power from the notched sensor with 1mW input.

the transmission line, and hence affect the signals at the transmitting (S_{11}) and receiving (S_{21}) ends. To connect to the sensing notch with high ambient isolation, we opt to use a miniature coaxial RF cable, as shown in Fig. 1(b), which is more feasible than the coplanar or slotline waveguides. The metal shield of the middle part is removed to allow the EM energy coupled into the nearby user body. Simulation by CST Microwave Studio [24] with the notched sensor on the human torso is shown in Fig. 1(c). The simplified torso model only includes the sweatshirt, skin, muscle, body fat, bones and lungs to save computational time. RF signals were emitted from the transmitter (Tx) and propagated down the transmission line. The EM leakage from the

notched structure was coupled into the human body, and finally received by the receiver (Rx). We can clearly observe that the electric field was strongly coupled into the layers of skin, fat, muscle, and lungs.

The notch length L = 9 in is designed to accommodate the position uncertainties for different users on beds and chairs. As shown in Fig. 2, the notched cable was sewn on the bedsheet for the bed setup, while it was fixed at the back of the chair for the chair setup. Intermediate layers of thin fabric, plastic protector and foam can be added for comfort and aesthetic with little influence on performance. In comparison with either far-field 17 or near-field 2526 effects of the sensing antenna structure, the present design can operate in a much broader bandwidth with a given hardware as indicated by the measured S_{21} in Fig. 1(d). As no RF radiation is intended and the Rx power can be anywhere between –5 dB to –50dB lower than the Tx power, the Bode-Fano limit on percentage bandwidth and impedance match 27 and spectral regulation are much more relaxed. Also, in contrast to wearable NCS sensors 232526, the user may not need to tend to the sensor at all, which greatly reduces the comfort and conspicuousness concerns. We further demonstrated the capability of accurate respiration monitoring under various postures, including supine, recumbent and sitting. Different respiratory routines including voluntary deep, fast and tidal breathing, as well as breath holding and airway blockage exercises, can be identified accurately. Our sensor can also retrieve the heart rate in one setup, although the cardio waveform is clearest during breath holding with reduced interference by the large respiratory motion. We benchmarked our sensor against the chest-belt respiration measurements with high consistency. The validity of our system has been confirmed

on 10 healthy adults.

Experimental Setup and Signal Processing

Notched sensor configuration

To demonstrate the broadband operability, continuous-wave (CW) signals between 0.8 and 2.5 GHz were tested by the network analyzer (Keysight E5063A) in the bed setup. The EM energy emitted from Port 1 will partially leak out from the notch and couple to the surface and internal body motion of breathing and heartbeats, which will change the signal at Port 2 as represented by the scattering parameter S_{21}. Some of the traveling wave will be reflected back to Tx, and can be potentially evaluated as S_{11} or standing-wave ratios, as shown in Fig. 1(d) when the human body was lying on the bed (red line) and when the bed is empty (blue line). The RF carrier power is set at 0 dBm, and the leakage power from the notched structure can be estimated by $(1 - |S_{21}|^2 - |S_{11}|^2)\times$1mW after de-embedding the cable loss, as shown in Fig. 1(e).

Two sets of the notched-cable sensors underneath the thorax and abdomen regions can capture the separate motion during breathing. Multiplexing by frequency or chip codes is optional, as the cross coupling of the thorax and abdomen sensors is weak due to small far-field radiation. Two software-defined radios (SDR) by National Instrument Ettus B200mini were connected to two notched cables and then interfaced with the host computer through USB (Universal Serial Bus). Alternatively, one SDR by the two ports of Ettus B210 can be employed. When the participant is lying on the bed or sitting on the chair, the centers of the sensor notches have a horizontal distance

d to the mediastinum, and the vertical distances h_1 and h_2 to the umbilicus for the thorax and abdomen sensors, as shown in Fig. 2(a). Figs. 2(b)(c) are the photos for the bed- and chair-integrated setups. In realistic deployment, we can also sew the sensor beneath the bed sheet and protective layers to facilitate routine changes, disinfection and aesthetic design. For verification, we benchmarked the NCS results with the BIOPAC tension-belt sensors (BIOPAC Systems, Goleta, CA). Two torso belts, SS5LB and PTM SS11LB, were placed at thorax and abdomen, with the vertical positions right above the NCS sensors. A photo is shown in Fig. 2(d). The torso belts can measure the change in local tension due to the geometrical change during respiration. BIOPAC data are down-sampled to 500Hz. The synchronization between NCS and BIOPAC is achieved by buffering the NCS and BIOPAC data at approximately the same time within a few milliseconds.

In the SDR Tx signal chain, the digital baseband goes through the digital-to-analog converter (DAC) and is then mixed with the carrier frequency f_{RF}. The RF signal from the notched structure is coupled into internal body motion in the near-filed, received by Rx, and then demodulated and sampled by the analog-to-digital converter (ADC) to retrieve the baseband. Together with a baseband tone f_{BB} in the quadrature scheme, the NCS vital-sign signal is represented by the amplitude modulation on the quadrature signal as

$$NCS(t) = \sqrt{I_{Rx}(t)^2 + Q_{Rx}(t)^2} \quad (1)$$

$$I_{Rx}(t) = A(t)\cos(2\pi f_{BB}t + \theta_0) \quad (2)$$

$$Q_{Rx}(t) = A(t)\sin(2\pi f_{BB}t + \theta_0) \quad (3)$$

where θ_0 is the phase offset accumulated from the Tx-Rx signal chains. The

baseband frequency is set at f_{BB} = 51 kHz for B200mini. When B210 is used for both channels, the two basebands are set at f_{BB1} =355 kHz and f_{BB2} = 440 kHz. f_{RF} is selected to be at one of 900MHz, 1.8GHz and 2.4GHz. As the respiration waveform is a low-frequency analog signal, we chose the superheterodyne scheme of converting the frequency twice by f_{BB} and f_{RF} to minimize the effects of the Flicker noise in f_{RF} local oscillators (LO) and to implement subcarrier multiplexing from f_{BB1} and f_{BB2} when needed. We have varied a few f_{BB} to substantiate our choices. The data converter has a sampling frequency of 1M samples per second (Sps) to enable the digital baseband processing, which is performed by the host computer. The demodulated respiration waveform is further down-sampled to 500 Sps, the same as the BIOPAC data. The allowable broad choice of f_{RF} is not only a unique feature over the antenna-based NCS approach, but can also facilitate system tuning after furniture integration so that f_{RF} can yield strong coupling to the user body with higher ambient tolerance.

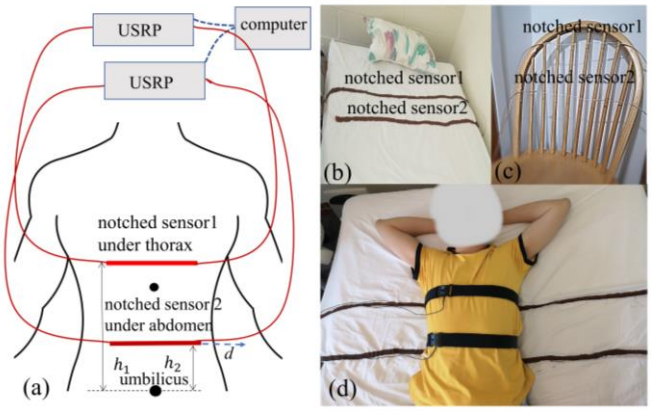


Fig. 2 Experimental setup of the notched sensors: (a) Sensor position variation with respect to the user body; (b) Bed integrated and (c) chair integrated setup; (d) A photo with the participant lying on the bed-integrated sensor and wearing commercial torso belts for benchmark.

Breath rate (BR) estimation

For estimation of the breath rate (BR), the waveform was first bandpass-filtered from 0.05Hz to 1Hz to remove the DC drift and high-frequency noises by post-processing in MATLAB. For peak detection to identify the inhalation and exhalation periods, we utilized the moving average-crossing algorithm [28], which is suitable to process broad range of varying frequencies, as we expect BR ranging from nearly 5 breaths per minute (BPM) to 40 BPM. A moving-average curve is first calculated at each time point in a given window length, which is around one respiration cycle and will be constantly updated. The points when the moving average curve crosses the original signal are marked as up-crossing points for positive slopes in the original signal or down-crossing points for negative. Local maximum is labelled as the maximal point between two up-down crossing points, and local minimum as the minimal point between two down-up crossing points. BR is finally calculated by counting the number of detected breathing cycles over a window size of 15s. One breathing cycle includes an inhalation peak (maximum) and the trailing exhalation (minimum) peak. The number of cycles is calculated as the number of inhalation peaks minus 1, and the total period is between the two adjacent inhalation peaks.

Signal quality and bit depth

The signal quality can be assessed by the acquired bit depth of the respiratory waveform, which represents the signal resolution and provides insights into the amount of RF energy coupled inside the internal organ. In this work, we define the bit depth of the signal quality Q as,

$$Q = \log_2 \frac{NCS_{pp}}{NCS_{DC}} + N - 1 \quad (4)$$

where NCS_{pp} is the peak-to-peak value in the quadrature-demodulated amplitude, NCS_{DC} is the DC amplitude at the local point, and N is the number of bits in ADC. NCS_{pp} can be retrieved by the peak detection algorithm, and NCS_{DC} can be obtained by the DC signal amplitude before filtering. Q has the meaning of the number of significant bits in the NCS signal. As the respiration waveform is cyclic around an equilibrium point, the larger the signal strength, the more bit depth can be retrieved by

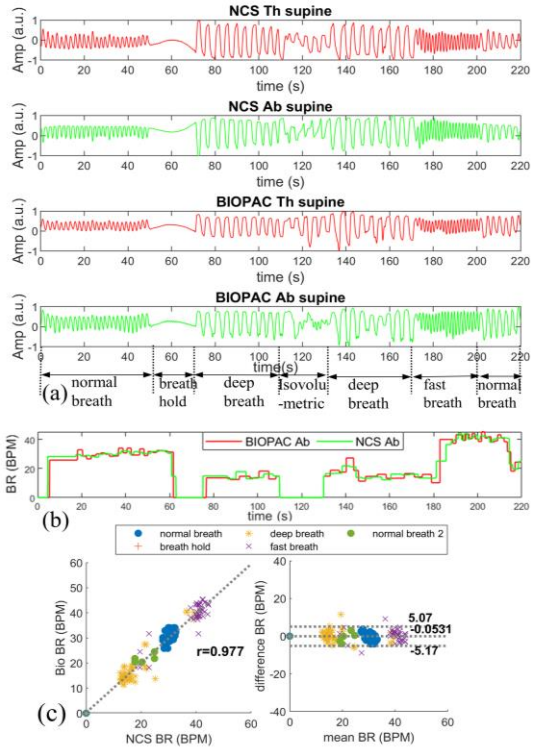


Fig. 3 Respiration monitoring in the supine posture with fRF = 900MHz: (a) NCS and BIOPAC amplitude waveforms in thorax and abdomen during the whole breathing protocol. (b) Th breath rate (BR) calculated from NCS (green) and BIOPAC (red). (c) Correlation and agreement between NCS and BIOPAC BR data. Left: The scatter plot with denoted Pearson's correlation coefficient. Right: The Bland-Altman plot showing the mean difference m at the center dotted line and the corresponding limits of agreement (LoA) at the upper and lower dotted lines given by m±1.96σ.

ADC, and hence the higher the signal resolution and the less quantization noises.

Signal-to-noise ratio (SNR) is another important parameter to evaluate the hardware system setup. We estimate SNR during the normal breathing period when BR is roughly constant. The fundamental breathing frequency can be determined from the largest nonzero spectral component. The 3-dB bandwidth of the breathing frequency includes all adjacent frequency components that decrease monotonically away from the maximum. The noise level is estimated by using the total power outside the 3-dB regions of the fundamental breathing frequency, its harmonics and the DC component.

Signal quality and SNR under different postures, breathing patterns, positions and frequencies can give guidance to the hardware design and signal processing to accommodate inevitable user variations in realistic applications. Larger signal quality often indicates higher SNR from the NCS signal strength, but in a less controlled ambient, SNR still offers additional useful information on ambient noises.

Isovolumetric breathing detection during airway blockage

One of the most common apneas is the obstructive sleep apnea (OSA). OSA often occurs when the throat muscles intermittently relax and block the airway during sleep. As the lung volume will remain nearly constant during OSA, the thorax and abdomen move asynchronously to generate a paradoxical motion. During our experiments, this symptom was simulated by the isovolumetric breathing exercise. By blocking the airway voluntarily with a nose pinch, abdomen contraction was performed with paradoxical expansion of the thorax and vice versa. To detect this

paradoxical movement, we can use the slope-product of thorax and abdomen waveforms, which will be negative during the occurrence of isovolumetric breathing [29]. We use the tanh() function to rescale the slope-product to [-1,1] with a moving average over a 3s window [29].

Experimental Results

Benchmark with BIOPAC tension belts

The NCS notched-cable sensor was benchmarked against the reference BIOPAC tension-belt sensors. Notice that both sensors can have errors in respiration monitoring, but due to the different signal transduction, the errors from the two methods should be reasonably uncorrelated. Vertical positions of the sensors are recorded by the distance to the umbilicus with h_{1NCS} = 8.5 in, h_{2NCS} = 2.5 in, $h_{1BIOPAC}$ = 9 in, and $h_{2BIOPAC}$ = 3.5 in. The breathing protocol in each posture has a length of 220 s, including 0 – 50 s: normal tidal breathing; 50 − 70 s: breath hold; 70 – 110 s: deep breathing; 110 – 130 s: isovolumetric exercise; 130 − 170 s deep breathing; 170 – 200 s: fast breathing; 200 – 220 s: normal breathing.

Fig. 3(a) shows the normalized NCS and BIOPAC amplitude waveforms in thorax and abdomen with the supine position on bed during the whole protocol. Both NCS and BIOPAC data can distinguish different breathing patterns clearly and show a good agreement. Fig. 3(b) shows BR calculated from NCS and BIOPAC abdomen amplitude, as the geometrical change from diaphragm and abdomen is higher than that of thorax during respiration. When the negative slope product was detected during the isovolumetric exercise, BR was truncated to 0. BR estimation matched well with the

breathing patterns, but a delay in time was observed due to the chosen 15-s epoch for BR calculation. Fig. 3(c) shows the correlation between NCS and BIOPAC. The BR data was preprocessed by getting rid of the outliers that were three standard deviations further away from the mean. The outliers only occurred at the beginning or ending of a specific breathing pattern due to the instability of a transition. These outliers were not included for correlation calculation. High correlation with a Pearson coefficient r = 0.977 is achieved. As the two sensors have uncorrelated errors, correlation is one of the possible indicators for inter-sensor consistency and overall BR accuracy. The Bland-Altman plot on the right presents the agreement by the mean (m) and limits of agreement (LoA). The X axis is the average of the two data, and the Y axis is the difference. The mean difference is m = –0.306 BPM, as shown in the middle dotted line. LoA within which 95% of the differences is estimated by $m \pm 1.96\sigma$, assuming a normal distribution. Low m and narrow LoA indicate that the errors in NCS and BIOPAC are uncorrelated and small. In Fig. 3, f_{RF} is set at 900MHz.

Posture variation

The NCS notched-cable sensor can be applied to different postures, including supine and recumbent on bed, and sitting on chair, which can be readily adapted for sleep study, driver seat and waiting-room chairs. In Fig. 4, we demonstrated the results from the recumbent and sitting postures. The breathing protocol was the same as in Fig. 3. Waveforms were based on the normalized NCS amplitude after bandpass filtering of 0.05 − 1 Hz. For the recumbent posture in Figs. 4 (a)(b), h_1 = 8.5inches, h_2 = 2.5inches, and f_{RF} = 2.4GHz. For the sitting posture in Figs. 4 (c)(d), h_1 = 9.5inches,

h_2 = 2inches, and f_{RF} = 1.8GHz. BR during the 20-s isovolumetric exercise was also truncated to zero once the negative slope product was detected. Data quality from the recumbent posture is sufficient for BR estimation, but slightly worse than that in the supine posture, which is likely due to less interaction area with the RF signal. As can be observed in the first 50-s normal breathing with reduced magnitude, the sitting posture suffered more noises and showed less clear breathing pattern. This may be due to possible slouching, which will seriously affect the abdomen waveforms. However, various breathing patterns can still be identified in Figs. 4 (a)(c), including tidal breathing, breath hold, deep breathing, and fast breathing. BR in Figs. 4 (b)(d) shows a good match between the thorax and abdomen sensors.

Position variation

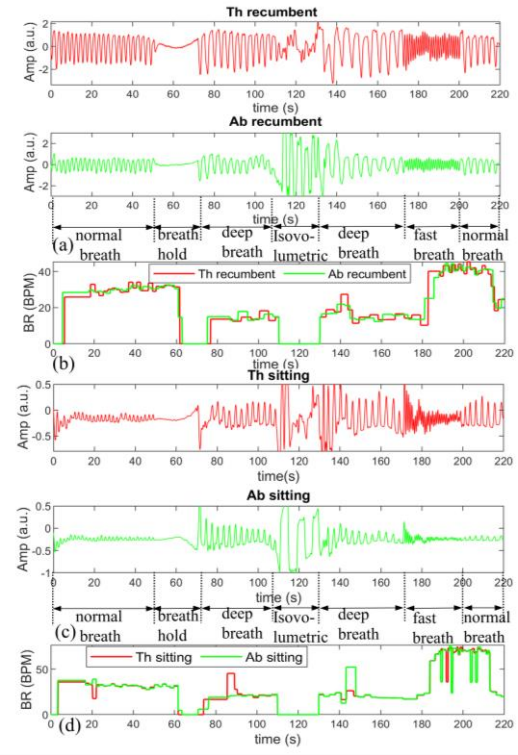


Fig. 4 Respiration monitoring in the recumbent and sitting postures: (a) Respiration waveforms from thorax and abdomen NCS sensors in the recumbent posture on bed with f_{RF} = 2.4GHz. (b) BR estimated from (a). (c) Respiration waveforms from thorax and abdomen NCS sensors in the sitting posture on chair with f_{RF} = 1.8GHz. (d) BR estimated from (c).

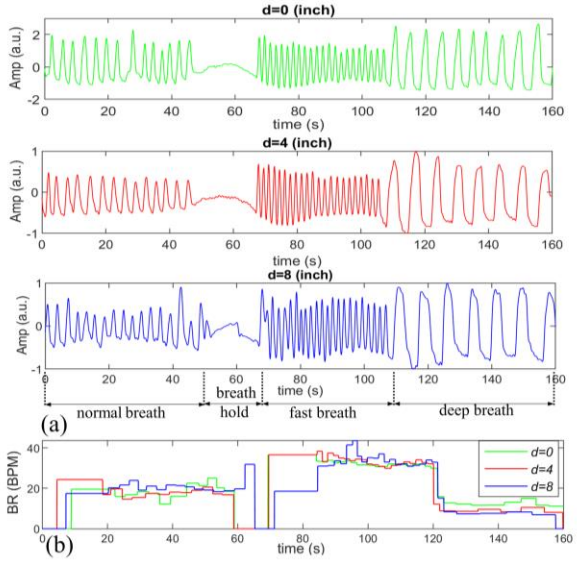


Fig. 5 (a) Respiration waveforms with horizontal shift of d = 0, 4, 8 inches for the abdomen sensor in the supine posture on bed. (b) BR estimated from the respiration waveforms for d = 0 (green); d = 4 in (red); d = 8 in (blue).

One of the requirements for furniture-integrated sensors different from the wearable ones is the large placement tolerance to accommodate user variation in size and position. We tested the sensor performance when a horizontal shift *d* was introduced in the supine posture, which can commonly happen in realistic conditions. We confirmed that our notched cable system can tolerate up to *d* = 8 in, which should be sufficient in consideration of user uncertainties. The participants in our test have an average waist width around 12 in. To speed up the study, the breathing protocol was shortened to 160s, including 0 – 50 s: normal breathing; 50 – 70 s: breath hold; 70 – 110 s: fast breathing; 110 – 160 s: deep breathing. f_{RF} was set to 2.4GHz. Fig. 5(a) shows the normalized NCS waveforms after filtering with the horizontal shift of *d* = 0, 4, 8 in from the abdomen sensor. Different respiration patterns can still be clearly identified and the transition remained evident. Fig. 5(b) shows BR estimated from the

respiration waveforms, where normal BR is around 20 BPM, fast BR around 35 BPM, and deep BR around 10 BPM. For d = 8 in, BR around the breath hold period is not as accurate as d = 0 and 4 in due to the reduced signal quality. We also tested the vertical position variation of h_1 = 1, 2.5, 4 in for abdomen sensor with the same protocol.

Isovolumetric exercises

The isovolumetric exercise was represented by paradoxical motion between thorax and abdomen. The 270-s breathing protocol was customized to increase the total period of isovolumetric motion including 0 − 40 s: normal breathing; 40 – 60 s: breath hold; 60 – 80 s normal breathing. Four isovolumetric exercises was then performed roughly during 80 – 100 s; 120 – 140 s; 160 – 180 s; 200 – 220 s, with normal breathing in between and afterwards.

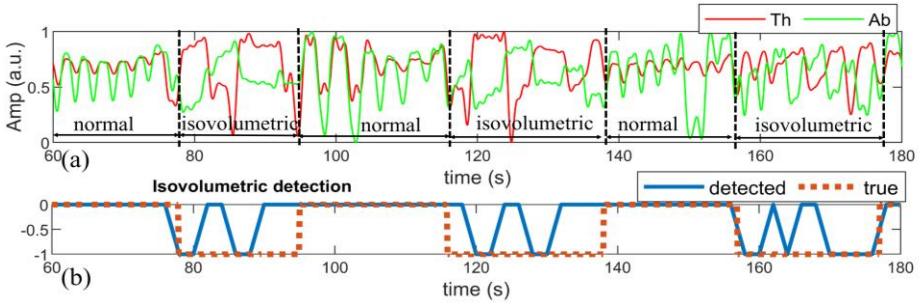


Fig. 6. Isovolumetric detection: (a) Respiration waveforms from thorax (red line) and abdomen (green line) sensors during three isovolumetric breathing exercises. (b) Detection by the slope-product method: The blue solid line by the NCS sensor and the orange dotted line from protocol annotation, where –1 indicates detection.

Fig. 6(a) shows an example of the waveforms from thorax and abdomen NCS signals during three isovolumetric cycles. f_{RF} was set to 900 MHz. As the waveforms indicate, in normal breathing periods, the thorax and abdomen moved synchronously; in the isovolumetric periods, the two motions were paradoxical to each other. The

slope-product method was implemented in Fig. 6(b), where −1 indicates positive detection of isovolumetric motion and 0 indicates no occurrence. The annotated instruction to perform the isovolumetric exercise is shown as the orange dotted line and the NCS detection result as the blue solid line.

Isovolumetric exercises can be missed or falsely detected by NCS or BIOPAC. False statistics against the user instruction are summarized for 10 volunteers in Table I. NCS could detect a large portion of airway blockage successfully and performed slightly better than BIOPAC. Possible reasons for erroneous detection include: 1) The notched sensors couple into the motion of associated muscles whose movement during airway blockage can be complex instead of merely paradoxical; 2) Some participants had difficulty performing the isovolumetric exercise during airway blockage, or had introduced body motion artifacts during their attempts; 3) The thorax sensor was too low in position and had coupled in part of the abdomen motion as well.

TABLE 1 CONFUSION MATRICES OF ISOVOLUMETRIC BREATHING DETECTION DURING AIRWAY BLOCKAGE BY NCS AND BIOPAC.

NCS			
TOTAL NUMBER OF EPOCHS IN 10 PARTICIPANTS: 180		DETECTION	
		POSITIVE	NEGATIVE
AIRWAY BLOCKAGE INSTRUCTION	POSITIVE	16	4
	NEGATIVE	10	150
ACCURACY		0.92	
SENSITIVITY		0.8	
BIOPAC			
TOTAL NUMBER OF EPOCHS IN 10 PARTICIPANTS: 180		DETECTION	
		POSITIVE	NEGATIVE
AIRWAY	POSITIVE	11	9

BLOCKAGE INSTRUCTION	NEGATIVE	11	149
ACCURACY		0.89	
SENSITIVITY		0.65	

Heart rate (HR) detection

With the same setup, the thorax NCS sensor can also retrieve the heart rate (HR) with reasonable accuracy. The test protocol was further simplified to normal breathing of 15 s and breath hold of 15 s, the latter of which had clearer heartbeat signals due to the lack of respiration interferences. The signal was first filtered by a bandpass filter between 0.1 − 5 Hz to remove the DC component and high-frequency noise. Fig. 7(a) shows the NCS waveform during normal breathing in time (left) and frequency (right) domains. Fig. 7(b) shows the breath-hold counterpart. The waveform has a typical characteristic of a main peak and a small recoil. HR can be observed as the fundamental, second and third harmonics in the spectrum. BR at 0.266 Hz and its harmonics can also be observed in Fig. 7(a).

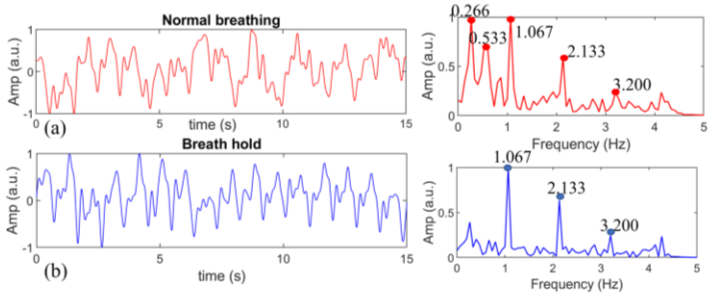


Fig .7. (a) Left: The heartbeat time-domain waveforms from the thorax sensor during normal breathing. Right: The corresponding spectrum with three HR peaks and two BR peaks indicated by the annotation above. (b) Left: The heartbeat waveforms during breath hold. Right: The corresponding spectrum with three peaks.

Signal quality and SNR

We have calculated the signal quality *Q* for different scenarios in Fig. 8. First,

the breathing patterns over time can cause variation in Q. For the same setup, the deep breathing usually has the highest Q, and the breath hold the lowest. The position variation of d = 0, 4, 8 in is shown in Fig. 8(a) with f_{RF} = 2.4 GHz. When d becomes larger, most often Q decreases. We compare the thorax and abdomen sensors in Fig. 8(b) when they operated together in the supine position with f_{RF} = 2.4 GHz. Abdomen often has larger motion from the diaphragm and associated muscles, and overall larger Q. Using measurements in Fig. 8(c), we compared Q in the supine, recumbent, and sitting postures from the abdomen sensor with f_{RF} = 1.8GHz. The supine signal has larger Q than those by recumbent and sitting due to higher coupling strength between the RF signal and the body. Finally, for frequency variation in Fig. 8(d) at f_{RF} = 0.9, 1.8, and 2.4GHz, we used the abdomen sensor in the supine position for illustration. Q has a more complex relation with respect to f_{RF}, although 1.8 GHz usually performed the best. All frequencies can be employed to retrieve respiration waveforms, but different Tx and Rx gains need to be chosen in our adaptive-gain implementation in SDR.

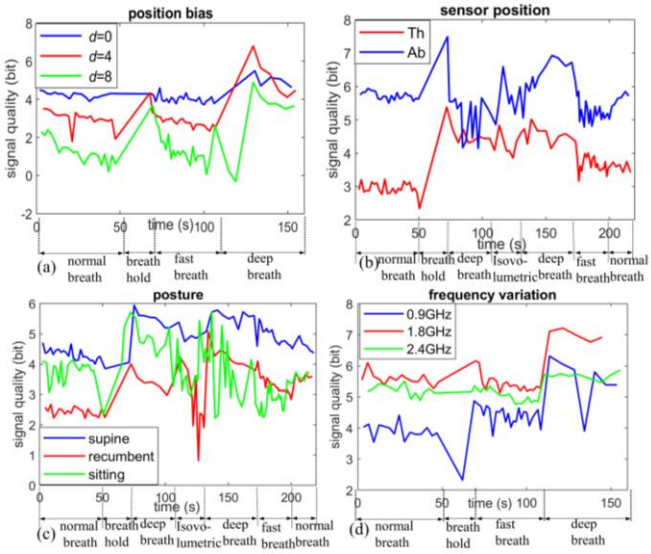


Fig. 8. Signal quality for (a) Position variation: d = 0 (green), 4 ins (red), and 8 ins (blue); (b) Sensor placement: thorax (blue) and abdomen (red); (c) Postures: sitting (red), supine (green), and recumbent (blue); (d) Carrier frequency: f_{RF} = 0.9 GHz (blue), 1.8 GHz (red), and 2.4 GHz (green).

The corresponding SNRs calculated in different conditions in Fig. 8 are shown in Table II. SNR is calculated during the normal breathing period of 0 – 50s when BR is roughly constant. The comparison of SNR under different conditions show similar trends to *Q* in our controlled ambient without additional interferences.

TABLE 2.2 SNR CALCULATED IN DIFFERENT CONDITIONS

(A) POSITION BIAS	SNR (DB)	(B) SENSOR PLACEMENT	SNR (DB)
$D = 0$	16.2	SUPINE THORAX	10.7
$D = 4$ IN	6.60	SUPINE ABDOMEN	14.3
$D = 8$ IN	5.92		
(C) POSTURE	SNR (DB)	(D) FREQUENCY	SNR (DB)
SUPINE	19.3	0.9GHZ	8.41
RECUMBENT	14.6	1.8GHZ	12.2
SITTING	8.05	2.4GHZ	9.69

Comparison with antenna-based NCS sensors

We also compared the performance of the notched NCS sensor and the wearable antenna-based NCS sensor [26]. The wearable sensor consists of ultra-high frequency (UHF) monopole antennas used at both Tx and Rx. Two wearable sensors were separately attached to the thorax and abdomen in the front torso in the similar location of the BIOPAC chest belts. Two notched sensors were integrated to the chair setup in the back. One SDR by B210 was connected to two notched sensors and another B210 was connected to two wearable sensors. f_{RF} was set at 1.8GHz. The breathing protocol was the same as in Fig. 5. In Fig. 9(a), we demonstrated the waveforms of the wearable sensor (red) and the notched sensor (green) both at the

abdomen position. In Figs. 9(b)(c), the corresponding BR and *Q* were extracted. BR from the two sensors are well matched, but the wearable sensors have a higher *Q* indicating better coupling strength than the notched sensor. This is reasonable because the wearable sensors were more closely attached to the body in the front torso. Also, the wearable sensors are monopole antennas with intentional radiation into the body, but the notched sensors depend only on near-field leakage by the evanescent mode.

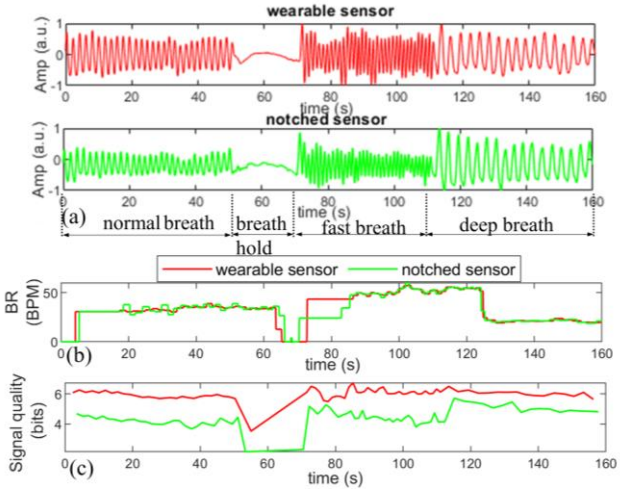


Fig. 9. (a) Waveforms of the wearable sensor (red) and the notched sensor (green) both at the abdomen position. (b)(c) The corresponding estimated BR and Q.

Two-tone measurements

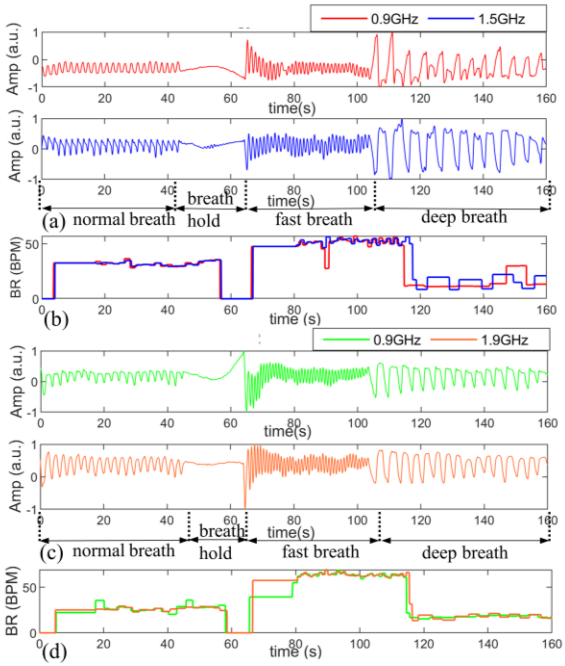


Fig. 10. (a) Two-tone Strategy 1: Respiration waveforms for 0.9 GHz (upper) and 1.5 GHz (lower); (b) The corresponding BR estimation. (c) Two-tone Strategy 2: Respiration waveforms for 0.9 GHz (upper) and 1.9 GHz (lower); (d) the corresponding BR estimation.

We also confirmed the two-tone measurement capability in the notched sensor, which can enhance reliability by frequency diversity. Different f_{RF} will have different body penetration depth. In comparison with the antenna-based NCS systems, dual-band antennas are more complicated to design, especially in the consideration of body antenna detuning. For the notched sensor, two f_{RF} can be applied simultaneously to one sensor, and one or two f_{RF} can be applied to two collocated cables with reasonable spatial isolation. In Fig. 10, two-tone Strategy 1 has a frequency multiplexing setup by two SDRs feeding into the same RF cable at the same time. Broadband 3dB splitters are utilized to combine two RF signals into one notched sensor and then splitting back to the Rx of the respective carrier. Two-tone Strategy 2 has two closely collocated

notched cables. Two SDRs with different frequencies separately fed into two notched sensors. At the same position, we have two sensors by two f_{RF} with minimal interference to each other.

The 160-s protocol was the same as Fig. 4 for position variation testing in the supine posture with the abdomen sensor. In Figs. 10(a)(b), Strategy 1 of frequency multiplexing has f_{RF} at 0.9 GHz and 1.5 GHz. In Figs. 10(c)(d), Strategy 2 of the collocated sensor cables has f_{RF} at 0.9 GHz and 1.9 GHz. BR extracted from different f_{RF} is consistent for both strategies. Notice that the two-tone measurements can be employed for signal quality check or for coupling into different body depth. Optimal choices of multiple f_{RF} in different scenarios will need further study.

Validation on multiple participants

Finally, we have validated our system on 10 different participants. The routine was the same as that in Fig. 3 with the notched sensors on bed and commercial torso-belt sensors on the user body. The retrieved BR from 10 participants compare well with the synchronous chest-belt sensors in all breathing routines. The correlation and agreement plots between NCS and BIOPAC BR data of all 10 people are shown in Fig. 11. The correlation and B&A statistics of individual participants are shown in Table III. For most cases, the correlation is relatively high, m is low and LoA is narrow. Only a small number of cases have a relatively lower correlation. Our present sample size is too small to draw reliable conclusion on the effects from body characteristics. However, it remains safe to say that the NCS system can accurately detect BR on different people, demonstrating high consistency with commercial torso-

belt sensors.

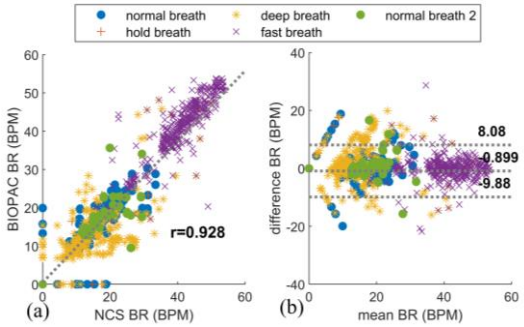


Fig.11. Correlation and agreement between NCS and BIOPAC BR data from 10 participants. (a): The scatter plot with denoted Pearson's correlation coefficient. (b): The Bland-Altman plot.

TABLE 3.3 CORRELATION AND B&A STATISTICS FOR EACH PARTICIPANT

CASE NO.	R	M (RPM)	Σ (RPM)
1	0.927	–1.65	4.50
2	0.987	0.586	2.24
3	0.951	–0.465	4.36
4	0.999	–0.0103	0.288
5	0.978	–0.210	2.12
6	0.805	–4.81	7.13
7	0.883	0.552	5.37
8	0.954	–0.631	4.22
9	0.955	0.131	3.27
10	0.913	–1.24	5.62

Conclusion

In this work, a new notched-cable sensor based on NCS for respiration monitoring was demonstrated. The sensor was integrated to a bed and a chair under

layers of fabrics, and is hence highly comfortable, stable, accurate, and cost-effective. In comparison with the wearable sensors attached to the body, our system is invisible to the user with the least concerns of comfort and conspicuousness. Confirmed by experiments, our system has the following advantages: 1) Tolerance of large position variation to accommodate user uncertainty; 2) Posture applicability including supine and recumbent on bed, as well as sitting on chair; 3) Availability of HR estimation in one setup; 4) Identification of multiple breathing patterns including deep, fast, tidal, held and blocked breathing; 5) Broad bandwidth of operations to facilitate multiplexing and signal quality improvement in different scenarios.

There are still remaining issues for future work. The accuracy of paradoxical motion detection during isovolumetric exercise should be further improved. Also, larger variation on people size, body figure and breathing habits should be included in broader user study to further establish heuristic improvement from known body characteristics.

REFERENCES

1 C. Massaroni, A. Nicolo, D. Lo Presti, M. Sacchetti, S. Silvestri and E. Schena, "Contact-based methods for measuring respiratory rate," Sensors, vol. 19, no. 4, Feb. 2019, art. 908, doi: 10.3390/s19040908.

2 D. Oletic and V. Bilas, "Energy-efficient respiratory sounds sensing for personal mobile asthma monitoring," IEEE Sens. J., vol. 16, no. 23, pp. 8295-8303, Dec. 2016, doi: 10.1109/jsen.2016.2585039.

3 D. Naranjo-Hernández, A. Talaminos-Barroso, J. Reina-Tosina, L.M. Roa, G. Barbarov-Rostan, P. Cejudo-Ramos, E. Márquez-Martín and F. Ortega-Ruiz, "Smart vest for respiratory rate monitoring of COPD patients based on non-contact capacitive sensing," Sensors, vol. 18, no. 7, Jul. 2018, art. 2144, doi: 10.3390/s18072144.

4 Oldenburg, B. Lamp, L. Faber, H. Teschler, D. Horstkotte and V. Topfer, "Sleep-disordered breathing in patients with symptomatic heart failure: a contemporary study of prevalence in and characteristics of 700 patients," Eur. J. Heart. Fail., vol. 9, no. 3, pp. 251-257, Mar. 2007, doi: 10.1016/j.ejheart.2006.08.003.

5 M. A. Cretikos, R. Bellomo, K. Hillman, J. Chen, S. Finfer and A. Flabouris, "Respiratory rate: The neglected vital sign," Med. J. Austral., vol. 188, no. 11, pp. 657-659, Jun. 2008.

6 P. Barthel, R. Wensel, A. Bauer, A. Müller, P. Wolf, K. Ulm, K.M. Huster, D.P. Francis, M. Malik and G. Schmidt, "Respiratory rate predicts outcome after acute myocardial infarction: A prospective cohort study," Eur. Heart. J., vol. 34, no. 22, pp. 1644-1650, Jun. 2013, doi: 10.1093/eurheartj/ehs420.

7 M. Grassmann, E. Vlemincx, A. von Leupoldt, J. M. Mittelstadt and O. Van den Bergh, "Respiratory changes in response to cognitive load: A systematic review," Neural. Plast., vol. 2016, 2016, Art. no. 8146809, doi: 10.1155/2016/8146809.

8 A. Nicolo, C. Massaroni and L. Passfield, "Respiratory frequency during exercise: The neglected physiological measure," Front. Physiol., vol. 8, Dec. 2017, Art. no. 922, doi: 10.3389/fphys.2017.00922.

9 W. J. DePaso, R. H. Winterbauer, J. A. Lusk, D. F. Dreis and S. C. Springmeyer, "Chronic dyspnea unexplained by history, physical examination, chest roentgenogram, and spirometry. Analysis of a seven-year experience," Chest, vol. 100, no. 5, pp. 1293-1299, Nov. 1991, doi: 10.1378/chest.100.5.1293.

10 F. Q. Al-Khalidi, R. Saatchi, D. Burke, H. Elphick and S. Tan, "Respiration rate monitoring methods: A review," Pediatr. Pulmonol., vol. 46, no. 6, pp. 523-529, Jun. 2011, doi: 10.1002/ppul.21416.

11 A. Abid, "Model-based estimation of respiratory parameters from capnography, with application to diagnosing obstructive lung disease". IEEE. Trans. Biomed. Eng., vol. 64, no. 12, pp. 2957–2967, May. 2017.

12 S. Gong, W. Schwalb, Y. Wang, Y. Chen, Y. Tang, J. Si, B. Shirinzadeh, and W. Cheng, "A wearable and highly sensitive pressure sensor with ultrathin gold nanowires," Nat. Commun., vol. 5, Feb. 2014, Art. no. 3132, doi: 10.1038/ncomms4132.

13 S. Stick, E. Ellis, P. LeSouëf and P. Sly, "Validation of respiratory inductance plethysmography ("Respitrace"®) for the measurement of tidal breathing parameters in newborns," Pediatr. Pulmonol., vol. 14, no. 3, pp. 187-191, Nov. 1992.

14 N. Netzer, A. H. Eliasson, C. Netzer and D. A. Kristo, "Overnight pulse oximetry for sleep-disordered breathing in adults: A review," Chest, vol.120, no. 2, pp. 625–633, Aug. 2001.

15 E. Nemati, M. J. Deen, and T. Mondal, "A wireless wearable ECG sensor for long-term applications," IEEE Commun. Mag., vol. 50, no. 1, pp. 36-43, Jan. 2012.

16 B. A. Reyes, N. Reljin, Y. Kong, Y. Nam and K. H. Chon, "Tidal volume and instantaneous respiration rate estimation using a volumetric surrogate signal acquired via a smartphone camera," IEEE J. Biomed. Health Inform., vol. 21, no. 3, pp. 764-777, May 2017, doi: 10.1109/JBHI.2016.2532876.

17 C. Li, V. M. Lubecke, O. Boric-Lubecke and J. Lin, "A review on recent advances in doppler radar sensors for noncontact healthcare monitoring," IEEE Trans. Microw. Theory Tech., vol. 61, no. 5, pp. 2046-2060, Apr. 2013, doi: 10.1109/tmtt.2013.2256924.

18 K. S. Park, S. H. Hwang, H. N. Yoon and W. K. Lee, "Ballistocardiography for nonintrusive sleep structure estimation," in Conf. Proc. 36th IEEE Eng. Med. Biol. Soc., Aug. 2014, pp. 5184-5187.

19 C. Song and P. Wang, "A radio frequency device for measurement of minute dielectric property changes in microfluidic channels", Appl. Physc. Lett., vol. 94, no. 2, Jan. 2009, Art. no. 023901.

20 J. Naqui and F. Martin, "Transmission lines loaded with bisymmetric resonators and their application to angular displacement and velocity sensors," IEEE Trans. Microw. Theory Tech., vol. 61, no. 12, pp. 4700-4713, Oct. 2013, doi: 10.1109/tmtt.2013.2285356.

21 A. Ebrahimi, J. Scott and K. Ghorbani, “Differential sensors using microstrip lines loaded with two split-ring resonators”, IEEE Sens. J., vol. 18, no. 14, pp. 5786-5793, May 2018, doi: 10.1109/jsen.2018.2840691.

22 S. Harnsoongnoen and A. Wanthong, “Coplanar waveguide transmission line loaded with electric-LC resonator for determination of glucose concentration sensing”, IEEE Sens. J., vol. 17, no. 6, pp. 1635-1640, Jan. 2017, doi: 10.1109/jsen.2017.2652121.

23 X. Hui and E. C. Kan, "Monitoring vital signs over multiplexed radio by near-field coherent sensing," Nat. Electron., vol. 1, no. 1, pp. 74-78, Jan. 2018, doi: 10.1038/s41928-017-0001-0.

24 CST Microwave Studio, Computer Simulation Technology, (2020) [Online]. Available: http://www.cst.com

25 P. Sharma and E. C. Kan, "Sleep scoring with a UHF RFID tag by near field coherent sensing," in 2018 IEEE/MTT-S International Microwave Symposium (IMS), Philadelphia, PA, June 10 – 15, 2018, pp. 1419-1422.

26 P. Sharma, X. Hui and E. C. Kan, "A wearable RF sensor for monitoring respiratory patterns," in 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, July 23 – 27, 2019: IEEE, pp. 1217-1223.

27 R. M. Fano, "Theoretical limitations on the broadband matching of arbitrary impedance, " J. Franklin Institution, vol. 249, no. 1, pp. 57–83, Jan. 1950.

28 W. Lu, M.M. Nystrom, P.J. Parikh, D.R. Fooshee, J.P. Hubenschmidt, J.D. Bradley and D.A. Low, "A semi-automatic method for peak and valley detection in free-breathing respiratory waveforms," Med. Phys., vol. 33, no. 10, pp. 3634-3636, Oct. 2006, doi: 10.1118/1.2348764.

29 P. Sharma, X. Hui, J. Zhou, T. B. Conroy and E. C. Kan, “Robust cardiopulmonary monitoring using wearable radio-frequency sensors”, NPJ Digit. Med., vol. 3, doi: 10.1038/s41746-020-0307-6, July 2020.

CHAPTER 3

DETECTION AND PREDICTION OF SLEEP DISORDERS BY COVERT BED-INTEGRATED RF SENSORS

Introduction

Sleep disorders are a major public health problem, and 50 to 70 million Americans chronically suffer from the consequences from sleep disorders [1]. Obstructive sleep apnea (OSA) is the most common sleep-related breathing disorder [2][3], with a prevalence in the adult population ranging from 6% to 17%, explicated by the apnea-hypopnea index (AHI) greater than 15 events per hour. OSA can be as high as 49% in geriatrics [4], and can still be under-diagnosed due to the inconvenience of the present monitoring setup [5]. OSA is characterized by repeated episodes of partial or complete obstruction of the respiratory passages during sleep, and may result in sleep fragmentation and non-restorative sleep. The consequences of OSA include excessive day-time sleepiness, insomnia, and increased risks of stroke, obesity, pulmonary hypertension and heart attack [6]. Missed identification of sleep disorders can be especially serious for young children with concerns of sudden infant death [7], and servicemen whose circadian rhythm are difficult to maintain but continuous vigilance is frequently required [8].

Currently, the gold standard of sleep disorder diagnosis is an overnight sleep study by polysomnography (PSG) [9], which records the breath airflow, respiratory torso movement, oxygen saturation (SpO2), body motion, electroencephalogram (EEG), electrooculogram (EOG), electromyogram (EMG), and electrocardiogram

(ECG) by a plethora of sensors. Sleep disorders are most often scored by trained sleep technicians considering various PSG waveforms. Despite of high accuracy by PSG, disadvantages include the high cost and scarce availability in clinics, and the uncomfortable experience for users. Several methods have been put forward as PSG alternatives for sleep disordered breathing (SDB) detection, such as the at-home systems using portable devices [10] which often has inadequate accuracy [11].

Electrocardiogram (ECG) is one of the most extensively scrutinized signals for sleep study [12][13]. Together with the chest belts for respiratory patterns, SDB can be reasonably detected. Apnea alarm systems also usually employ SpO2 from the pulse oximeter to provide warning when SpO2 falls below a predefined threshold [14]. However, the high rate of false alarms can be triggered by motion artifacts and poor sensor contact [15].

Many methods and algorithms have been proposed to detect sleep disorders autonomously [9][16]-[18], including various machine-learning (ML) models of neural networks, regression, and ensemble learning. Nevertheless, in the past decades, fewer studies have explored the prediction capability [19][20]. Predictive warning of the SDB events in advance can potentially improve the effectiveness of therapy. Currently, the gold-standard treatment for OSA is the continuous positive airway pressure (CPAP) [21][22] by blowing air into the nose. Though effective, the use of a single pressure and cumbersome equipment could cause pressure intolerance and reduce long-term acceptance. To improve user comfort, auto-titrating continuous positive airway pressure (APAP) [23] was later developed [24]. Recently the COVID-19 pandemic has cast a spotlight on the ventilators [25] as life-support machines

providing intensive ventilatory support. Other innovative methods have also been proposed to intervene sleep disorders after detecting abnormal events [26]. However, interventions can be best applied with apneic prediction, as detection may be too late for intervention after 30s. Predictive warning of SDB for an advance of 30 – 90 s might be critical to improve therapeutic outcomes and reduce the impact on oxygen levels and sleep structure. Motivated by the unmet needs of reliable prediction of sleep disorders with minimal user disturbance, we developed a bed-integrated system with predicative SDB warning up to 90 s. Our system is based on the near-field coherent sensing (NCS) of ultra-high frequency (UHF) electromagnetic (EM) waves to monitor the dielectric boundary movement of internal organs and body parts [26]-[28]. It can be invisible to users and requires no personal setup time, especially considering occasional leave from beds such as restroom visits. Comfort and convenience are of critical importance for overnight sleep monitoring because apnea is a sparse event in the long recording of sleep in different stages. If users decide to take the monitor off and do not take the trouble of re-installation, serious apneic incidences can be missed.

We have further developed a learning-based algorithm for detection and prediction. We recruited 27 patients with sleep disorders from the Weill Cornell Center for Sleep Medicine for overnight recording. The data from the NCS bed sensor was processed for feature extraction to feed into the random-forest ML model, which gave accurate SDB detection. Apneic events can be further predicted up to 90 s in advance based on the present respiratory features. We could also determine the correlation between respiratory features and SDB to identify the most critical physiological factors for detection and prediction of SDB episodes.

Experimental setup and protocol

Experimental setup

As shown in Fig. 1(a), the bed-integrated sensor consisted of a notched miniature coaxial radio-frequency (RF) cable, where the metal shield of the middle 5 inches was removed to allow a small amount of EM energy leaking into the nearby user body [28]. The notch length was designed to accommodate positional variations for different patients. In the near-filed region, the dielectric boundary movement by lungs and associated muscles would couple to the leaked EM energy, and hence affect the signals between the transmitter (Tx) and the receiver (Rx) [29]. Two notched-cable sensors underneath the approximate thorax and abdomen positions were adopted to capture the motion in separate regions during breathing [27]. Software-defined radio (SDR) was used to connect the notched sensors and then interface with the host computer through USB (Universal Serial Bus). As shown in Fig. 1(c), SDR was implemented by the National Instrument Ettus B210 with two Tx/Rx ports. The notched cables were sewn on the bottom side of the mattress pad, as shown in Fig. 1(d), which were placed under an incontinence protector and a fitted bedsheet for comfort and aesthetic. No apparent performance degradation was observed for the additional layers, as evident from our prior studies for sensing over several layers of fabrics [30]. The schematic of the experimental setup was summarized in Fig. 1(e). In our previous work, vital-sign monitoring by this setup has been benchmarked against the strain-based chest belts and ECG with various sleep postures and large position variation [28].

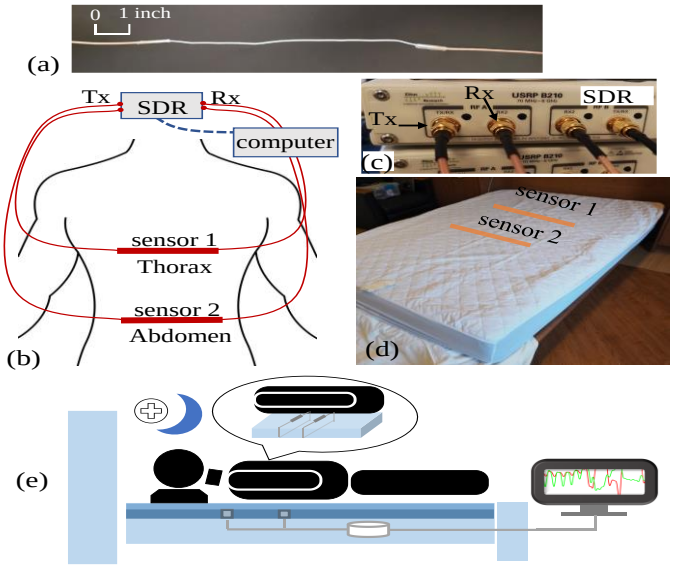


Fig. 1. The RF sensor setup: (a) A photo of the notched transmission line sensor; (b) NCS sensor positions in relation to the body; (c) The photo of the SDR transceiver connected to two NCS sensors; (d) The position of the notched sensors under the mattress pad; (e) Schematics of the overall experimental setup (PSG not shown).

For the SDR Tx signal chain, the digital baseband went through the digital-to-analog converter (DAC) and was then mixed with the carrier frequency f_{RF}. The RF power is less than –10 dBm or 0.1 mW, well under the safety limits set by occupational safety and health administration (OSHA) in the UHF band. The RF signal leaked from the notched structure is coupled into internal organ motion, received by Rx, and then demodulated and sampled by the analog-to-digital converter (ADC) to retrieve the baseband. We employed the quadrature scheme as the baseband tone f_{BB}, and the NCS signal can be represented by the magnitude and phase modulation on the quadrature signal as

$$NCS_{amp}(t) = \sqrt{I_{Rx}(t)^2 + Q_{Rx}(t)^2} \quad (1)$$

$$NCS_{ph}(t) = unwrap(\tan^{-1}\frac{Q_{Rx}(t)}{I_{Rx}(t)} - 2\pi f_{BB} - \theta_0) \quad (2)$$

$$I_{Rx}(t) = A(t)\cos(2\pi f_{BB}t + \theta_0) \quad (3)$$

$$Q_{Rx}(t) = A(t)\sin(2\pi f_{BB}t + \theta_0) \quad (4)$$

where θ_0 is the phase offset accumulated from the Tx-Rx signal chains. f_{RF} was selected at 1 GHz, and any choice between 0.9 and 2.4 GHz should have similar performance [27]. Two RF notched sensors, one approximately at the thorax and the other at the abdomen, were operated at two distinctive basebands of f_{BB1} =355 kHz and f_{BB2} = 440 kHz, which were implemented by two Tx-Rx channels synchronized in one SDR to reduce cross interference. Both channels were sampled at 10^6 samples per second (Sps), and were further down-sampled to 500 Sps after demodulation.

Subjects and Data Collection

Overnight PSG was performed at the Weill Cornell Center for Sleep Medicine at approximately the participants' regular sleep time, and included recording of EEG, EOG, submental and anterior tibialis EMG, two-lead ECG, chest and abdominal movement by inductive plethysmography, body position, SpO2 pulse oximetry, and nasal pressure respiratory flow monitoring. Scoring of SDB and sleep stages was performed by a registered PSG technician licensed in the State of New York. All events were scored according to the recommended rules by the American Academy of Sleep Medicine [31]. The human study was performed under the approved protocol of Weill Cornell Medical Center IRB# 19-12021223.

Data Processing

After gathering overnight recording of patients, we first processed our NCS data along with PSG respiratory data, then extracted the respiratory features, and finally fed the data into the ML classifier for SDB detection and prediction. Notice that NCS and PSG respiratory channels had the same signal processing procedure for fair comparison. We used MATLAB for signal processing and feature extraction.

Signal processing

We included the full overnight NCS recording of 27 patients with the duration of 7 – 8 hours. PSG was mainly used for feature validation and model comparison, and had 3 respiratory channels of the airflow, thorax belt, and abdomen belt. SpO2 from PSG was also collected as an additional input apart from the respiratory motion waveforms. Two NCS sensors from the thorax and abdomen positions produced respective magnitude and phase as four individual inputs.

The feature extraction contained 8 steps:

1. Down sample NCS and PSG to 25 Hz.
2. Synchronize NCS and PSG signals (precision to 1s).
3. Perform bandpass filtering and smoothing (0.05 Hz to 2 Hz).
4. Segment waveform into epochs of T_{epoch} = 40 s and a sliding window of T_{slide} = 15 s.
5. Label operator annotation in epoch.
6. Normalize waveform and extract features in epoch.
7. Select epochs by signal quality for NCS and PSG.

8 Output features and annotation to the ML model.

The bandpass filter in Step 3 was implemented in MATLAB by the digital infinite impulse response (IIR). Another Savitzky-Golay finite impulse response (FIR) smoothing filter [32] with 4^{th} polynomial order was further employed to rid of high-frequency noises. The operator annotations were adapted to give epoch-based references. In each epoch, if any annotation has a time duration > 40% × T_{epoch} = 16 s, the current epoch will be labelled accordingly. The choice of T_{epoch} between 10 – 20 s will not affect the end result significantly. Epoch labels include normal, snore, hypopnea, OSA, mixed apnea, and CSA (central sleep apnea). If the annotated disorder event lasts < 16 s, the epoch will be labeled as normal. If the time duration with no annotation > 20 s, current epoch will be removed from dataset. This is often due to insufficient evidence of proper PSG monitoring, such as patients going to restroom or taking some sensors off.

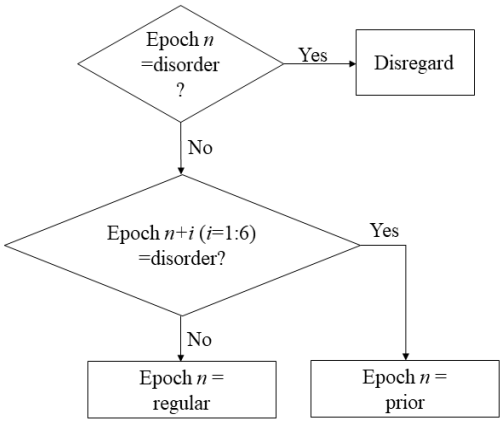


Fig. 2. The prediction labelling criterion.

The annotation was directly used as the ground truth for disorder detection training and validation. For prediction, labelling criterion should be modified. In each epoch, if the current label was a disorder event, we would remove it from the dataset

because prediction was based on the normal period before the disorder events. Aiming to predict disorders 0 – 90s in advance, we labelled the epochs containing six sliding windows before the forthcoming disorder event as the disorder precedence “prior”, and the other epochs as “regular”. A simple flow chart of prediction labelling is shown in Fig. 2.

Feature extraction

After epoch segmentation, waveforms were normalized to [–1,1] in each epoch. To extract respiratory features, we first implemented the peak detection algorithm [33]. A moving-average curve was first calculated at each time point in a given period, which was around one respiration cycle and then constantly updated. The points when the moving-average curve crossed the original signal were marked as up-crossing points for positive slopes in the original signal or down-crossing points for negative. Local maximum was labelled as the maximal point between two up-down crossing points, and local minimum as the minimal point between two down-up crossing points. Fig. 3 presents sample epochs of respiratory waveforms, annotated with different labels of (a) normal, (b) OSA, and (c) hypopnea. The red waveform was derived from NCS and the green ones from PSG. Solid magenta upward-pointing triangles marked the maximum peaks detected by the algorithm and blue downward-pointing triangles mark the minimum peaks. We could observe distinctive patterns in different events. Waveforms showed a more regular respiratory pattern in normal epochs, while irregular patterns more frequently indicated disorder epochs.

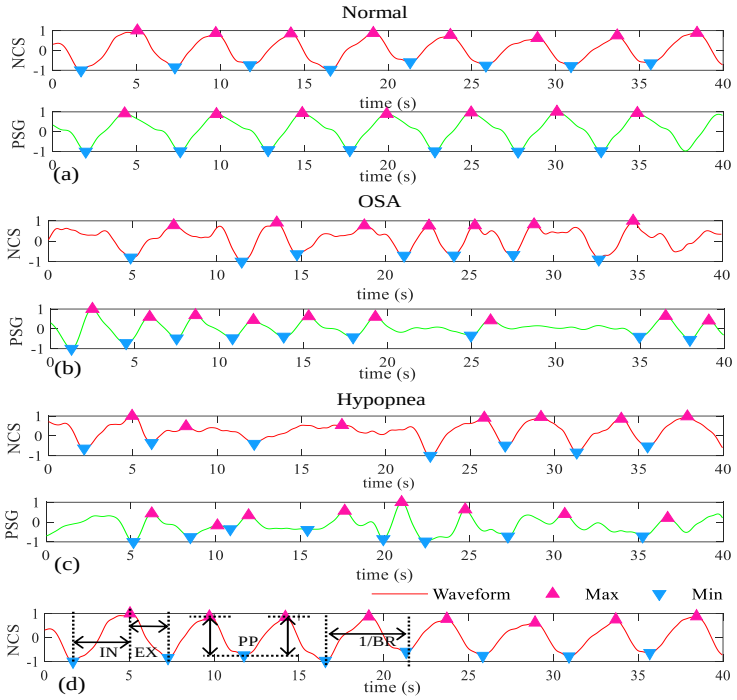


Fig. 3. Waveform examples from NCS and PSG in the epochs labelled as (a) normal; (b) OSA; (c) hypopnea. NCS channels are from (a): thorax phase; (b): thorax phase; (c): abdomen amplitude. PSG channels are from (a): chest; (b): airflow; (c): airflow. (d) Examples of the peak detection and feature extraction process. The red line is the respiratory waveform from (a). The magenta and blue triangles are the detected maximum and minimum peaks. Features of IN, EX, PP and BR are marked.

TABLE I. RESPIRATORY FEATURES I (16).

μ_{BR}	μ_{PP}	μ_{IN}	μ_{EX}
Σ_{BR}	Σ_{PP}	Σ_{IN}	Σ_{EX}
CoV_{BR}	CoV_{PP}	SKEW	KURT
N	ENTR	η_{P0}	T_{HOLD}

TABLE II. RESPIRATORY FEATURES II (21).

η_{P1}	η_{P2}	η_{P3}	η_{P4}
$F1_{MAX}$	$F2_{MAX}$	$F3_{MAX}$	$F4_{MAX}$
COR_{BR}	COR_{PP}	COR_{IN}	COR_{EX}
SD_{BR}	SD_{PP}	SD_{IN}	SD_{EX}
MAX_{IN}	MAX_{EX}	MAX_{BR}	
MIN_{BR}	MIN_{PP}		

TABLE III. SPO2 FEATURES (4).

μ_{SpO2}	Σ_{SpO2}	η_{SpO2}	MIN_{SpO2}

After identifying respiratory cycles using peak detection, we could first extract the 4 respiratory parameters in each breath cycle to represent the instantaneous respiratory characteristics, including BR (breath rate in BPM), PP (peak-to-peak in arbitrary units as an estimate of the lung volume), IN (inhalation interval in s), and EX (exhalation interval in s). Several examples are shown in Fig. 3(d) for the extraction process of respiratory parameters.

After gathering respiratory cycles and parameters, we extracted 16 respiratory features which would function as the epoch features fed into the ML classifier for detection as listed in Table I. The first 8 features were the mean (μ) and the standard deviation (σ) of the above 4 respiratory parameters. Because BR and PP were two significant factors representing the respiratory pattern, we added 2 more features as the coefficient of variation (CoV) of BR and PP,

$$CoV = \left(\frac{\sigma}{\mu}\right)^2 \quad (4)$$

CoV showed the extent of variability in relation to the mean. Additionally, Skew and kurt measured the tailedness and asymmetry of each respiratory cycle, and were averaged over all cycles within the epoch. Apart from features derived from respiratory parameters, we appended four supplemental features including 1) the total number of detected respiratory cycles n; 2) the total randomness or entropy of the waveform entr; 3) the power in the lower frequency band ((0.05,0.5) Hz) divided by the total power in all frequencies η_{p0}; 4) the time duration when no peak was detected within the epoch T_{hold}.

Other than 16 respiratory features in Table I, we added 21 respiratory

features in Table II for the prediction classifier, which had 37 respiratory features in total. Augmentation of features can enhance the performance of the ML model before overfitting becomes dominant. η_{pi} and fi_{max} (i = 1~4) represented the power in specific bandwidth divided by the total power in all frequencies and the frequency with the maximum power density within the bandwidth, respectively. The four chosen bandwidths were f_1 = (1, 2) Hz; f_2 = (2, 5) Hz; f_3 = (5, 8) Hz; f_4 = (8, 12.5) Hz. Cor was the autocorrelation in a time lag of one respiratory cycle to measure the successive similarity of a given respiratory parameter. SD representing the successive difference was defined as the mean absolute difference between adjacent cycles. At last, we added the maximum of IN, EX and BR within the epoch and the minimum of BR and PP. The choice of these features is based on the physiological reasoning that in the events of disorder, or in the anticipation of the events, there would be larger variation in PP and BR within the epoch. η_{pi} and fi_{maxc} contained the regularity of IN and EX, and possibly some tissue vibration characteristics during the disorder events below the audible range.

Beyond respiratory features derived from the NCS and PSG waveforms, we also added features representing oxygen saturation as listed in Table III: 1) μ_{SpO2}: the mean SpO2 level; 2) σ_{SpO2}: the standard deviation of SpO2 level; 3) η_{SpO2}: the percentage of time when SpO2 < threshold (92%); 4) min_{SpO2}: the minimum level of SpO2.

After segmentation and feature extraction, we added an extra step for NCS epoch selection. Signal quality cannot be guaranteed during the entire course of overnight recording because patients may have random motion lying on the bed or

leave the bed for restroom visits. Various other factors such as ambient movement might bring about noises to cause SNR (signal-to-noise ratio) degradation. Therefore, we opted to remove the epochs with very low SNR by pre-determined thresholds, i.e., epochs with $\eta_{p0} < Th_{\eta p0} = 70$ % and $\sigma_{PP} > Th_{\sigma PP} = 0.3$ will be removed from the dataset.

Machine learning Models

Data composition

For output datasets, Table IV shows NCS and PSG dataset composition for detection and prediction, respectively. Labels for epochs were divided into 7 classes, namely, normal, snore, arousal, hypopnea, OSA, mixed apnea, and CSA. Labels of normal, snore, and arousal were further grouped into the binary classification of "normal", while labels of hypopnea, OSA, mixed apnea, and CSA into "disorder". We studied the performances of our ML model by both 7 classes and 2 classes, although the main focus was on the binary classes of normal and disorder. The first 7 rows in Table IV presented the total number of epochs annotated with these labels in NCS and PSG. For the last 2 rows, the disorder ratio was the proportion of disorder epochs within all epochs, and the epoch selection ratio was the ratio between the total duration of selected epochs and overall recording time. Because prediction only included normal epochs for disorder precedence, NCS prediction dataset has a relatively smaller ratio than detection. As for the PSG dataset, we used the epochs derived from the same time periods as selected from NCS for fair comparison. Note that though PSG and NCS shared the same recording time, PSG only utilized one optimal channel

out of three respiratory channels for each epoch, while NCS may include more than one channel with acceptable signal quality within each epoch.

TABLE IV. NCS DATASET COMPOSITION OF EVENTS AND PRECEDENCIES.

	NCS Detection	PSG Detection	NCS Prediction	PSG Prediction
Normal	23574	15334	13350	9538
Snore	621	244	265	126
Arousal	548	274	578	334
Hypopnea	7674	3483	3221	1689
OSA	1902	852	1452	560
Mixed Apnea	63	34	7	4
CSA	401	188	405	157
Disorder ratio	0.289	0.223	0.307	0.231
Epoch selection ratio	0.413	0.413	0.254	0.254

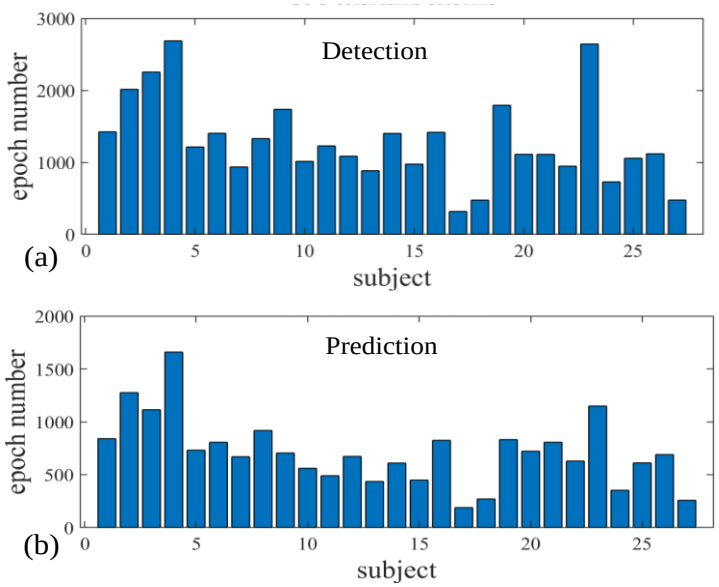


Fig. 4. The numbers of the selected NCS and PSG epochs from each patient in the dataset of (a) detection and (b) prediction.

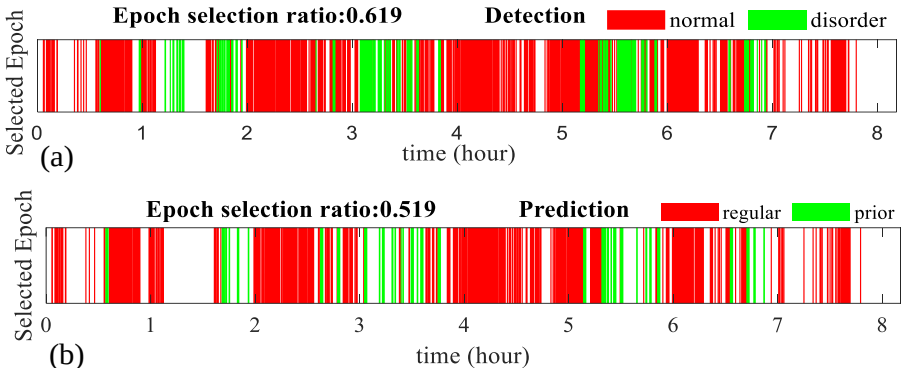


Fig. 5. An example of the selected NCS epoch distribution within the whole overnight recording in the (a) detection and (b) prediction datasets.

TABLE V. COMPARISON OF THE MEAN AND STANDARD DEVIATION OF NORMAL AND DISORDER EPOCHS OF SELECTED FEATURES.

AVG± DEV	μ_{BR}	Σ_{BR}	μ_{PP}	Σ_{PP}	η_{P0}
NCS: NORMAL	15.6± 3.12	2.62 ± 1.65	0.992 ± 0.206	0.147 ±0.050	87.2 ± 7.44
NCS: DISORDER	17.8± 3.64	4.61 ± 2.91	0.859 ± 0.220	0.206 ±0.079	86.1 ± 7.37
PSG: NORMAL	15.6± 3.23	1.97 ± 1.67	1.171 ± 0.208	0.138 ±0.085	91.7 ± 8.52
PSG: DISORDER	16.7± 3.47	2.73 ± 1.88	1.032 ± 0.209	0.213 ±0.096	89.6 ± 8.98

AVG± DEV	CoV_{BR}	CoV_{PP}	μ_{SpO2}	Σ_{SpO2}	
NCS: NORMAL	0.164 ± 0.094	0.159 ±0.074	93.5± 5.52	0.620 ±1.85	
NCS: DISORDER	0.252 ± 0.141	0.266 ±0.145	91.6± 3.59	1.70 ± 1.34	
PSG: NORMAL	0.122 ± 0.089	0.135 ±0.115	93.5± 5.67	0.622 ±1.82	
PSG: DISORDER	0.162 ± 0.101	0.228 ±0.137	91.7± 3.61	1.68 ± 1.29	

Fig. 4 showed the selected NCS epoch number for each patient in the detection and prediction datasets. The NCS epoch selection had a large variation mainly due to subject variation. An example of the NCS epoch selection during the whole overnight recording was shown in Fig. 5 for detection and prediction datasets from one representative patient. For detection in Fig. 5(a), the red bars represented the selected normal epochs and the green bars represented the selected disorder epochs. For prediction in Fig. 5(b), the normal epochs in the detection dataset are further divided into regular and prior according to whether an abnormal event will happen in the forthcoming 90 s. Note that the PSG dataset had the same epoch time distribution with NCS. Table V presented the comparison of the mean and standard deviation of the selected features in normal and disorder epochs in the NCS and PSG detection datasets which had dominant significance in the ML model in the following section. Disorder epochs had higher standard deviation for all respiratory features, indicating disorder epochs were less stable and tend to fluctuate more. Meanwhile, disorder epochs also have distinctively higher CoV_{BR}, and CoV_{PP} in comparison with normal ones, which were important factors to distinguish the two classes as well. For oximetry, disorder epochs usually had smaller μ_{SpO2} and higher σ_{SpO2}.

The Random-forest Model

We chose the random-forest classifier [34] as the ML model, which was an ensemble learning method for classification that constructed a multitude of decision trees during training, and then output the class selected by most trees. There are two advantages in the tree-based ML models: 1) Straightforward to interpret as a white-

box model, which can help us understand the intrinsic relationship between respiratory features and sleep disorders; 2) Non-parametric without assumption on the data distribution or linearity. Random forest can reduce overfitting in a single decision tree but keep the advantages of the decision tree.

However, we faced the problem of class imbalance as shown in Table IV, where the number of normal epochs was much larger than those of the disorder epochs. In other words, a bias or skewness would shift towards the normal event present in the dataset. For remediation, we added class weights for statistical amplification, which assigned different weights to the normal and disorder labels (normal:disorder = 1:3). The model thus penalized the misclassification made on the minority class of disorder. This practice achieved higher sensitivity to disorder detection effectively.

Class weighting can improve sensitivity, while outlier removal can improve specificity. Before constructing the ML model, we first cleaned the dataset by removing the normal epochs that were distinctively deviant from the majority. In our study, we assumed that the normal epochs in the dataset should have a relatively regular and similar respiratory pattern and thus data should form a dense cluster. Abnormal observations that are far from the majority ones within normal epochs were removed as outliers that were most likely due to noisy or wrong data. Here, we used the isolation forest algorithm[35], an unsupervised anomaly detection method based on random forests, as the outlier detection method. The outlier removal ratio was set at 0.2, which efficiently promoted specificity by eliminating noisy data in normal epochs.

Results for disorder detection

The k-fold and leave-one-participant-out cross-validations (CV) were employed as model verification, where k-fold CV tested the skill of the model on new data, and leave-one-participant-out CV tested the robustness to unseen patients. For k-fold CV of k = 5, we divided the whole dataset (N cases) into separate training (0.8N cases) and testing (0.2N cases), and the process was repeated 5 times until all cases had been tested as unseen data. For leave-one-participant-out CV, the model was trained on the data sets from all patients except one, whose data were then used as testing. The CV process was reinitialized and repeated for each patient as the testing case.

Fig. 6 shows the overall confusion matrices for detection using k-fold random forest, while Table VI further presents the statistics for k-fold. The binary class of normal and disorder achieved better accuracy than the full seven classes. NCS + SpO2 resulted in the best performance for disorder detection with 88.9% accuracy, 88.6% sensitivity and 89.0% specificity. The top three important features were SpO2 deviation σ_{SpO2}, peak-to-peak deviation σ_{PP}, and breath rate deviation σ_{BR}. When only NCS datasets were used, the sensitivity has significantly degraded to 63.6%, indicating SpO2 was an important factor for apnea identification apart from respiratory patterns. The top three important features became σ_{PP}, σ_{BR}, and μ_{BR}.

Note that the PSG detection here only utilized one optimal channel out of three respiratory channels for each epoch. Overall, the accuracy using different PSG respiratory channels was similar.

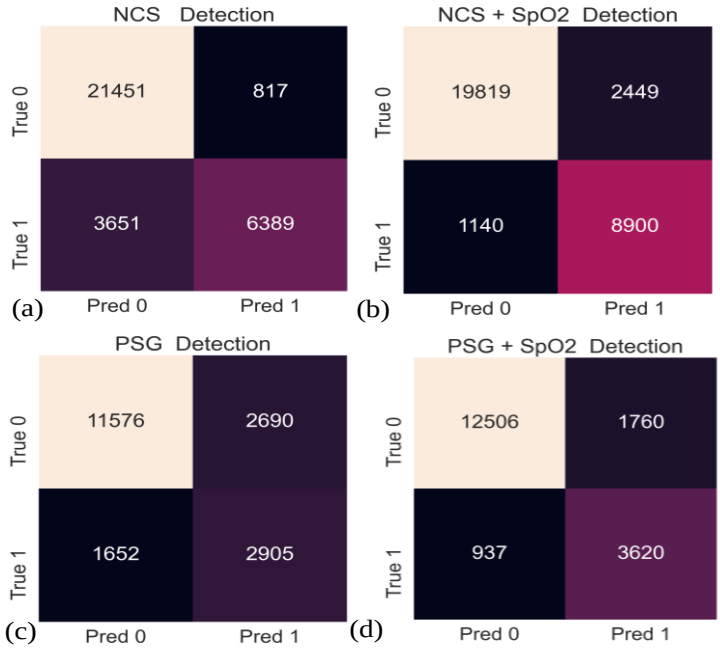


Fig. 6. The confusion matrices showing Normal (0) and Disorders (1) detection by the random forest model using the features from (a) NCS; (b) NCS +SpO2; (c) PSG; (d) PSG +SpO2. The cells list the number of epochs in each category. A 5-fold CV was tested on the entire data.

Table VI. Comparison of the detection devices by 5-fold CV.

DATA SET	NCS	NCS + SPO2	PSG	PSG + SPO2
ACCURACY	86.2%	88.9%	76.9%	85.2%
SENSITIVITY	63.6%	88.6%	63.7%	78.3%
SPECIFICITY	96.3%	89.0%	81.1%	87.4%
FEATURE IMPORTANCE	Σ_{PP} (0.50) Σ_{BR} (0.27) μ_{BR} (0.05) CoV_{PP} (0.03)	Σ_{SpO2} (0.58) Σ_{PP} (0.14) Σ_{BR} (0.08) CoV_{PP} (0.04)	CoV_{PP} (0.42) T_{HOLD} (0.10) μ_{PP} (0.05) μ_{EX} (0.05)	Σ_{SpO2} (0.36) η_{SpO2} (0.05) μ_{BR} (0.05) Σ_{PP} (0.05)

We further compared multiple classifiers including k-nearest neighbor (kNN), support vector machine (SVM), decision tree, hybrid model and random forest, as

presented in Table VII. The hybrid model consisted of the voting classifier ensembled from SVM, kNN and decision tree altogether. kNN had the lowest sensitivity to disorder detection, although the specificity was very high. Random forest resulted in highest accuracy for NCS + SpO2 dataset, and also achieved high sensitivity and specificity. The overall difference among SVM, decision tree and random forest was relatively small.

TABLE VII. ALGORITHM COMPARISON FOR DETECTION BY 5-FOLD CV.

ALGORITHM	CV ACCURACY (%)		SENSITIVITY (%)		SPECIFICITY (%)	
	NCS	NCS + SPO2	NCS	NCS + SPO2	NCS	NCS + SPO2
KNN	82.9	87.0	47.1	62.0	99.1	98.3
SVM	71.8	87.1	80.3	90.3	68.0	85.7
DECISION TREE	77.3	87.5	72.1	88.3	79.7	87.2
HYBRID*	81.5	89.7	69.6	86.8	86.9	91.0
RANDOM FOREST	86.2	88.9	63.6	88.6	96.3	89.0

HYBRID* IS THE VOTING CLASSIFIER ENSEMBLED FROM SVM, KNN AND DECISION TREE.

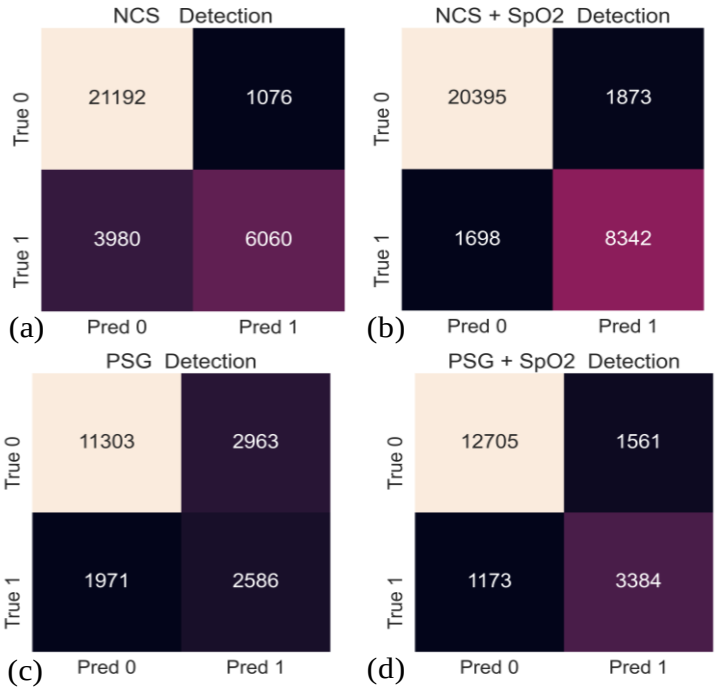


Fig. 7. The confusion matrices showing Normal (0) and Disorders (1) detection by the random forest model using the features from (a) NCS; (b) NCS +SpO2 (c) PSG; (d) PSG +SpO2 by the leave-one-participant-out.

TABLE VIII. COMPARISON OF THE DETECTION DEVICES BY LEAVE-ONE-PARTICIPANT-OUT CV.

DATA SET	NCS	NCS + SPO2	PSG	PSG + SPO2
ACCURACY	84.4%	88.9%	73.8%	85.5%
SENSITIVITY	60.4%	83.1%	56.7%	74.3%
SPECIFICITY	95.2%	91.6%	79.2%	89.1%
FEATURE IMPORTANCE	ΣPP (0.50) ΣBR (0.28) μBR (0.05) COVPP (0.02)	ΣSPO2 (0.62) ΣPP (0.15) ΣBR (0.08) COVPP (0.03)	COVPP (0.43) THOLD (0.10) μEX (0.05) μBR (0.05)	ΣSPO2 (0.70) COVPP (0.03) ηSPO2 (0.03) KURT (0.02)

The results above were tested from k-fold CV, and the leave-one-participant-out CV for unseen patients was shown in Fig. 7. High performance of 88.9% accuracy, 83.1% sensitivity and 91.6% specificity was maintained using NCS + SpO2 features. As shown in Figs. 7 (c)(d), for PSG dataset, sensitivity to disorder events remain slightly lower than NCS. The accuracy and feature importance using k-fold and leave-one-participant-out CV were similar, as shown in Table VIII.

The misalignment in timing between the thorax and abdomen waveforms in the epoch can be a potential feature for OSA detection from the paradoxical breathing patterns [29]. T_{lag} was calculated by the shifted time lag of the abdomen channel that gave the highest cross-correlation between the thorax and shifted abdomen channels. Phase and amplitude channels were separately compared. However, the accuracy results after adding the T_{lag} feature were not much improved, probably because the OSA event was already represented in other respiratory features. Therefore, we did not include T_{lag} as a feature in the other benchmarks.

Results for disorder prediction

In this section, we presented the accuracy statistics for SDB prediction using the waveforms in the normal epochs preceding the disorder epoch by 0 – 90 s. Similar CV tests were performed on prediction datasets as in detection.

Fig. 8 shows the overall confusion matrices using the k-fold random forest model, and Table VIII further presents the statistics. Similar to detection, prediction of individual events from 7 classes had lower accuracy in comparison with the binary class. In contrast to the detection results when NCS + SpO2 was better than NCS

alone, we can find that the performance of using only NCS was comparable to NCS + SpO2, meaning that NCS respiratory information alone can function as a predictor for disorders. Physiologically speaking, low SpO2 was the result of the apneic event, and therefore was useful in detection, but not in prediction. Sensitivity to disorder precedence epochs (81.3%) was relatively lower than those of detection (88.6%), which was also understandable because disorder precedence has less evident changes in respiratory features than the actual disorder events.

In comparison with PSG, NCS had distinctively higher accuracy and sensitivity for SDB prediction. This was likely because the important feature ηp0, representing the signal power dominance within the bandwidth of (0.05, 0.5) Hz, was not well represented in PSG, as NCS had a unique capability to extract motion characteristics of a broad bandwidth[38][39]. The important features for prediction in NCS included ηp0, σPP CoVPP, and σBR according to Table VIII. To identify forthcoming abnormality, the feature ηp0 captured whether the waveform was monotonic in the fundamental BR or contained more high-frequency attributes. The features CoVPP and σPP represented the peak-to-peak variations, which corresponded to the lung volume. σBR, representing the breath rate variation, can be important too.

TABLE IX. COMPARISON OF THE PREDICTION DEVICES BY 5-FOLD CV.

DATA SET	NCS	NCS + SPO2	PSG	PSG + SPO2
ACCURACY	81.9%	81.9%	74.1%	76.1%
SENSITIVITY	74.6%	81.3%	55.8%	64.8%
SPECIFICITY	84.9%	82.1%	78.9%	79.1%

Feature importance	η_{P0} (0.31) Σ_{PP} (0.30) CoV_{PP} (0.19) Σ_{BR} (0.13)	Σ_{PP} (0.35) η_{P0} (0.28) Σ_{SpO2} (0.18) Σ_{BR} (0.09)	Σ_{PP} (0.21) MIN_{PP} (0.17) CoV_{PP} (0.10) μ_{PP} (0.09)	Σ_{SpO2} (0.49) η_{SpO2} (0.05) μ_{PP} (0.05) μ_{PP} (0.05)

TABLE X. Algorithm comparison for prediction by 5-fold CV.

Algorithm	CV Accuracy (%)		Sensitivity (%)		Specificity (%)	
	NCS	NCS + SpO2	NCS	NCS + SpO2	NCS	NCS + SpO2
SVM	77.4	80.4	76.7	80.1	77.8	80.5
kNN	83.0	83.4	46.3	47.6	99.6	99.5
Decision tree	79.7	79.7	81.0	81.0	79.2	79.2
Hybrid*	79.3	82.7	73.9	77.7	81.5	84.7
Random forest	81.9	81.9	74.6	81.3	84.9	82.1
Hybrid* is the voting classifier ensembled from SVM, kNN and decision tree.						

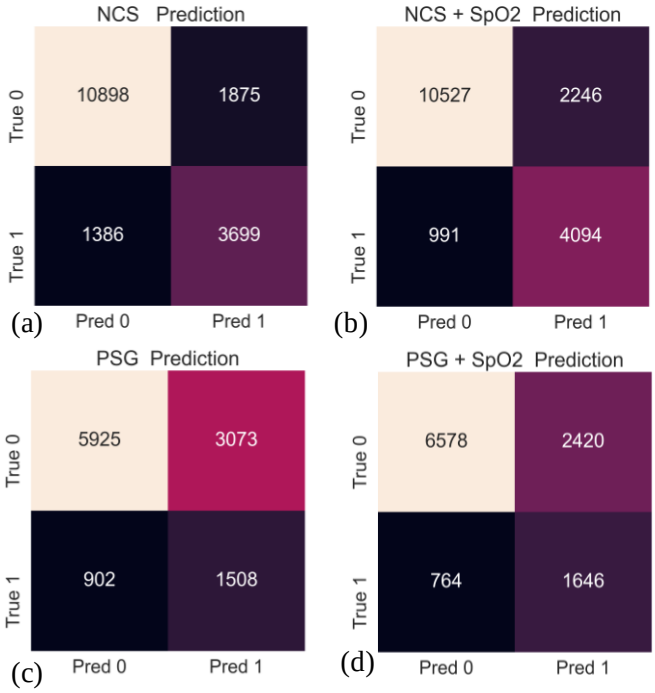


Fig. 9. The confusion matrices showing Regular (0) and Prior (1) prediction by the random forest model using the features from (a) NCS; (b) NCS +SpO2 (c) PSG; (d) PSG +SpO2 by the leave-one-participant-out CV.

TABLE XI. COMPARISON OF THE PREDICTION DEVICES BY LEAVE-ONE-PARTICIPANT-OUT CV.

ACCURACY	81.7%	81.9%	65.2%	72.1%
SENSITIVITY	72.7%	80.5%	62.6%	68.3%
SPECIFICITY	85.3%	82.4%	65.8%	73.1%
FEATURE IMPORTANCE	η_{P0} (0.32) Σ_{PP} (0.27) COV_{PP} (0.21) Σ_{BR} (0.14)	Σ_{PP} (0.36) η_{P0} (0.28) Σ_{SpO2} (0.18) Σ_{BR} (0.10)	MIN_{PP} (0.22) Σ_{PP} (0.15) μ_{PP} (0.11) COV_{PP} (0.08)	Σ_{SpO2} (0.51) μ_{PP} (0.06) MIN_{SpO2} (0.05) μ_{SpO2} (0.05)

Different algorithms for prediction were compared regarding classification performance in Table IX. kNN lacks a reasonable sensitivity to disorder events.

Random forest has a good performance on specificity and achieves reasonably high sensitivity and overall accuracy. SVM and decision tree can also generate results with nearly similar performance.

In addition to k-fold CV, leave-one-participant-out CVs were performed to validate the robustness of the prediction system on unseen patients and the results are shown in Fig. 9. The result reached 81.7% accuracy, 72.7% sensitivity and 85.3% specificity using only the NCS features. The accuracy remained high for unseen patients and showed a good match with k-fold CV results. Similar results from the PSG dataset were presented in Figs. 9 (c)(d). The accuracy and feature importance using leave-one-participant-out CV were shown in Table XI.

We also compared the results for prediction in advance of different time length ranging from 30s to 120s as shown in Fig. 10. Accuracy decreased when the prediction time length increased, and sensitivity degraded significantly when the prediction time exceeded 90 s. We selected 90 s as the final choice to obtain a reasonably high accuracy as well as a longest feasible warning time.

In Fig. 10(b), we presented additional comparison for the different combinations of epoch duration T_{epoch} and sliding window T_{slide} using (upper) NCS and (lower) NCS +SpO2 features. Accuracy was hardly affected by the choices of T_{epoch} and T_{slide} . We chose T_{epoch} = 40 s and T_{slide} = 15 s in the main analysis mostly for convenience and a relatively large number of epochs.

The above results of our SDB detection and prediction were based on the feature extraction followed by the ML model. We also experimented on convolutional neural network (CNN) as the ML model, eliminating the feature extraction process.

For CNN, we used the waveform from NCS as the direct input and constructed the network consisting of 5 convolution layers and 3 linear layers. Dataset was divided into training (80%) and testing (20%) parts and the accuracy for the unseen testing data was estimated. In this study, CNN had inferior performance to the approaches using feature extraction followed by the classic ML models. This was likely due to the insufficient SNR in the raw waveforms, where the feature extraction in the epoch duration could provide some data smoothing and selection effects.

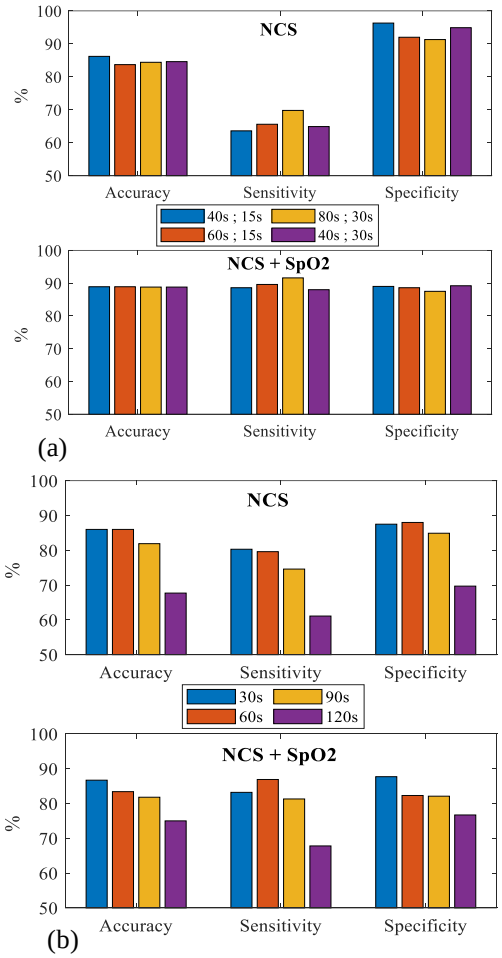


Fig. 10. (a) Comparison of NCS prediction accuracy of the different time length ranging from 30s to 120 s using features from (upper) NCS and (lower) NCS +SpO2. (b) Comparison of detection accuracy using different combinations of epoch duration T_{epoch} and sliding window T_{slide} in the (upper) NCS and (lower) NCS +SpO2 datasets.

Results for AHI classification

AHI, calculated by the number of apnea and hypopnea events per hour of sleep, is an important feedback to the patient to indicate the severity of sleep apnea [36].

In addition to epoch-based apnea detection and prediction, we also evaluated the performance for AHI classification from the NCS inputs only. We first divided all participants into binary classes of AHI $\leqslant$ 5 as "Normal" and AHI > 5 as "OSA Present" [37]. Using our NCS detection results for normal and disorder, we extracted overall features for each participant including the total epoch numbers of normal and disorder and the NCS F selection rate. We then adopted a simple random forest model by these features from each participant as the input, and estimated the AHI class. The resulting confusion matrix of the two-class AHI classification between annotation and NCS output is shown in Fig. 11 (a). In Fig. 11 (b), we also presented the three-class confusion matrix from AHI $\leqslant$ 5 as "Normal"; 5 < AHI $\leqslant$ 15 as "Mild OSA"; 15 < AHI as "Moderate OSA" [37]. Our NCS estimation achieved accuracy of 0.93 for binary AHI classification, and 0.70 for three classes. Our AHI classification performance is limited due to the present small sample size. Alternatively a random-forest regressor ML model can produce a continuous AHI score [27]. Our AHI accuracy can likely make further improvement in future studies when data from more patients with broader distribution become available.

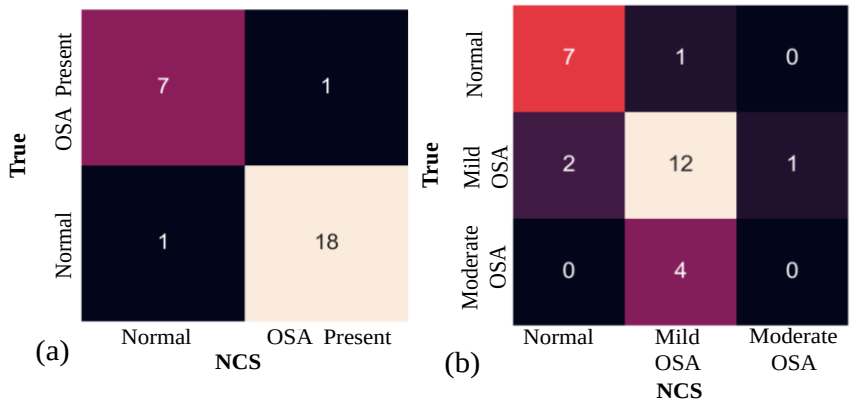


Fig. 11. The Confusion matrix of the (a) two-class and (b) three-class AHI classification between annotation and NCS output from 27 patients. In (a), the two AHI classes are: AHI ≤ 5: Normal; 5 < AHI : OSA Present, with an accuracy = 0.93; sensitivity = 0.95; specificity = 0.88. In (b), the three AHI classes are: AHI ≤ 5: Normal; 5 < AHI ≤ 15: Mild OSA; 15 < AHI < 30: Moderate OSA, with an accuracy = 0.70.

Discussion

Remaining Challenges

Challenges to construct a clinically acceptable sleep apnea detection and prediction platform still remained:

1) Sensitivity was only above 70% for prediction at the current stage. On the brighter side, the NCS deployment can be invisible to patients throughout the monitoring, and was hence ultimately comfortable and convenient. Our present system can still be a reasonable complement to guide interventions [21]-[26].

2) Our current system cannot classify different respiratory events with sufficient accuracy. Though the binary-class (normal/disorder) results had a relatively high accuracy, our system did not perform well for identifying OSA, CSA, and hypopnea individually. This may be due to the much more hypopnea events than OSA

and CSA ones in the dataset. The system thus had limited learning for OSA and CSA, and tend to classify new disorder events as hypopnea.

3) SNR of the waveform needed further improvements. In this study, we relied on NCS epoch selection to eliminate noisy episodes. Sensor improvement for higher SNR and higher tolerance to subject variation and motion interference should be investigated.

4) The snoring event needed features from higher frequency. In this study, the snoring event was not included in the disorder class. However, snoring was an important sleep abnormality in need of more comprehensive investigation. Our NCS sensing technology can couple to low-frequency motion like respiration [27] as well as high-frequency motion of internal tissues [38]. Snoring can be detected with minimal ambient interference by an additional NCS probe on the jugular or submental area, similar to cough sensing [39].

Future Improvements from Expanded Scope

The potential extension in future research studies includes:

1) We will expand future studies by adding more severe cases with AHI $\geqslant$ 30. During the execution period of our study protocol in the Weill Cornell Center for Sleep Medicine, we recruited participants with suspected sleep apnea that turned out to have mild and moderate conditions, while others were normal. Future study including patients with wider range of AHI can also help build a more comprehensive and mature learning model.

2) Broadening the demographic groups to include high-risk patients for SDB,

including opioid addicts, COPD, or infants at risk for severe apneic events and respiratory arrest, especially those who are born preterm [7][40]. Reliable sleep apnea detection and prediction in these high-risk patients would help improve outcome and prevent fatality.

3) Developing a system that uses SDB prediction to guide real-time intervention including user warning and ambient stimulation, followed by overall effectiveness assessment. The benefit of intervention will likely depend on the prediction accuracy and reliability. By integrating detection, prediction and intervention, we would hopefully improve diagnosis, prognosis and therapy for SDB.

4) Extending the study to include patients with more severe cases of OSA and the associated risk in comorbidities by examining the possible correlation [41]. We will also try to include patients with CSA to further explore the clinical utility of the proposed technology, although the number of patients needs to be much larger due to the infrequent occurrence of CSA. The detection model will also need broader dimension of pathological features in order to achieve a higher confidence level.

5) Improving the learning model to achieve higher accuracy and reliability. Improvements can include additional preprocessing for feature extraction, better noise reduction algorithms, and more complex ML models such as gain-adversarial network (GAN) [42].

Conclusion

In this work, we reported a hardware-software co-designed system that can detect and predict SDB. This system was based on a covert bed-integrated RF sensor

by NCS, which can be non-invasive and invisible to user. SDB detection for considering apneas and hypopneas together achieved a sensitivity and specificity up to 88.6% and 89.0% for k-fold validation, and 83.1% and 91.6% for subject-independent validation, respectively. Subsequent apneic events can be predicted up to 90 s in advance based on the present respiratory features. Disorder prediction achieved a sensitivity and specificity up to 81.3% and 82.1% for k-fold validation, and 80.5.0% and 82.4% for subject-independent validation, respectively. By the random forest ML model, the most significant physiological symptoms before and during the SDB episodes can also be revealed.

The current sleep apnea diagnosis platform was mostly based on PSG, which remained expensive in terms of hardware and operators, uncomfortable from body electrodes, and time-consuming for deployment. The ability to predict upcoming SDB events by PSG was also limited. In the future, our covert detection and prediction system could expedite intervention, and improve diagnosis and therapy for respiratory disturbance during sleep.

REFERENCES

1 H. R. Colten and B. M. Altevogt, Eds., *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem*. National Academies Press, 2006.

2 T. Young, M. Palta, J. Dempsey, J. Skatrud, S. Weber, and S. Badr, “The occurrence of sleep-disordered breathing among middle-aged adults,” *New Engl. J. Med.*, vol. 328, no. 17, pp. 1230-1235, 1993.

3 T. Young, L. Evans, L. Finn, and M. Palta, “Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women,” *Sleep*, vol. 20, no. 9, pp. 705-706, 1997.

4 C. Senaratna et al., “Prevalence of obstructive sleep apnea in the general population: A systematic review,” *Sleep Med. Rev.*, vol. 34, no. 1, pp. 70–81, Aug. 2017.

5 A. Roebuck et al., “A review of signals used in sleep analysis,” *Physiol. Meas.*, vol. 27, no. 3, pp. 320–331, 2015.

6 E. A. Phillipson, "Sleep apnea--a major public health problem," *New Engl. J. Med.*, vol. 328, no. 17, pp. 1271-1273, 1993.

7 E. S. Katz, R. B. Mitchell, and C.M. D'Ambrosio, "Obstructive sleep apnea in infants," *Am. J. Respir. Crit.*, vol. 185, no. 8, pp. 805-816, Apr. 2012.

8 V. Mysliwiec et al., "Sleep disorders in US military personnel: a high rate of comorbid insomnia and obstructive sleep apnea," *Chest*, vol. 144, no. 2, pp. 549-557, Aug. 2013.

9 F. Mendonça, S. S. Mostafa, A. G. Ravelo-García, F. Morgado-Dias and T. Penzel, "A review of obstructive sleep apnea detection approaches," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 2, pp. 825-837, Mar. 2019.

10 W. W. Flemons et al., "Home diagnosis of sleep apnea: A systematic review of the literature," *Chest*, vol. 124, no. 4, pp. 1543-1579, Oct. 2003.

11 O. C. Ioachimescu et al., "Performance of peripheral arterial tonometry–based testing for the diagnosis of obstructive sleep apnea in a large sleep clinic cohort," *J Clin. Sleep Med.*, vol. 16, no. 10, pp. 1663-1674, Oct. 2020.

12 O. Faust, U. R. Acharya, E. Ng, and H. Fujita, "A review of ECG-based diagnosis support systems for obstructive sleep apnea," *J. Mech. Med. Biol.*, vol. 16, no. 01, p. 1640004, Feb. 2016.

13 G. D. Clifford, F. Azuaje, and P. McSharry, *Advanced Methods and Tools for ECG Data Analysis*. Artech house Boston, 2006.

14 U. J. Magalang et al., "Prediction of the apnea-hypopnea index from overnight pulse oximetry," *Chest*, vol. 124, no. 5, pp. 1694-1701, Nov. 2003.

15 V. Monasterio, F. Burgess, and G. D. Clifford, "Robust classification of neonatal apnoea-related desaturations," *Physiol. Meas.*, vol. 33, no. 9, p. 1503, Aug. 2012.

16 N. Pombo, N. Garcia, and K. Bousson, "Classification techniques on computerized systems to predict and/or to detect Apnea: A systematic review," *Comput. Methods Programs Biomed.*, vol. 140, pp. 265-274, Mar. 2017.

17 B. Xie and H. Minn, "Real-time sleep apnea detection by classifier combination," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 469-477, May 2012.

18 A. Burgos, A. Goni, A. Illarramendi, and J. Bermudez, "Real-time detection of apneas on a PDA,", *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 4, pp. 995-1002, Nov. 2009.

19 J. Bock and D. A. Gough, "Toward prediction of physiological state signals in sleep apnea," *IEEE. Trans. Biomed. Eng.*, vol. 45, no. 11, pp. 1332-1341, Nov. 1998.

20 J. A. Waxman, D. Graupe, and D. W. J. A. j. o. r. Carley, "Automated prediction of apnea and hypopnea, using a LAMSTAR artificial neural network," *Am. J. Respir. Crit.*, vol. 181, no. 7, pp. 727-733, Apr. 2010.

21 N. B. Kribbs et al., "Objective measurement of patterns of nasal CPAP use by patients with obstructive sleep apnea," *Am. Rev. Respir. Dis.*, vol. 147, no. 4, pp. 887-895, Apr. 1993.

22 R. G. Andrade et al., "Nasal vs. oronasal CPAP for OSA treatment: A meta-analysis," *Chest*, vol. 153, no. 3, pp. 665-674, Mar. 2018.

23 R. B. Berry, J. M. Parish, and K. M. Hartse, "The use of auto-titrating continuous positive airway pressure for treatment of adult obstructive sleep apnea," *Sleep*, vol. 25, no. 2, pp. 148-173, Jan. 2002.

24 V. Patruno et al., "Fixed and auto-adjusting continuous positive airway pressure treatments are not similar in reducing cardiovascular risk factors in patients with obstructive sleep apnea," *Chest*, vol. 131, no. 5, pp. 1393-1399, May 2007.

25 A. Esteban et al., "How is mechanical ventilation employed in the intensive care unit? An international utilization review," *Am. J. Respir. Crit. Care Med.*, vol. 161, no. 5, pp. 1450-1458, May 2000.

26 “Sleep smart, smart pillow,” ZEREMA. [Online]. Available: https://www.zerema.co/. [Accessed: 06-Oct-2021].

27 Z. Zhang, P. Sharma, T. B. Conroy, V. Phongtankuel, and E. C. Kan, "Objective scoring of physiologically induced dyspnea by non-invasive RF sensors," *IEEE. Trans. Biomed. Eng.,* July 2021, doi: 10.1109/TBME.2021.3096462.

28 Z. Zhang, P. Sharma, J. Zhou, X. Hui, and E. C. Kan, "Furniture-integrated respiration sensors by notched transmission lines," *IEEE Sens. J.*, vol. 21, no. 4, pp. 5303-5311, Feb. 2021.

29 P. Sharma, X. Hui, J. Zhou, T. B. Conroy, and E. C. Kan, "Wearable radio-frequency sensing of respiratory rate, respiratory volume, and heart rate," *NPJ Digit. Med*., vol. 3, p. 98, July 2020.

30 X. Hui, P. Sharma, and E. C. Kan, "Microwave stethoscope for heart sound by near-field coherent sensing," *Proc. IEEE MTT-S International Microwave Symposium (IMS)*, 2019, pp. 365-368.

31 R. B. Berry et al., "Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events: Deliberations of the sleep apnea definitions task force of the American Academy of Sleep Medicine," *J. Clin. Sleep Med.*, vol. 8, no. 5, pp. 597-619, Oct. 2012.

32 R. W. Schafer, "What is a Savitzky-Golay filter? [lecture notes]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 111-117, July 2011.

33 W. Lu et al., "A semi-automatic method for peak and valley detection in free-breathing respiratory waveforms," *Med. Phys.*, vol. 33, no. 10, pp. 3634-6, Oct. 2006, doi: 10.1118/1.2348764.

34 T. K. Ho, "The random subspace method for constructing decision forests *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832-844, Aug. 1998.

35 F. T. Liu, K. M. Ting, and Z. H. Zhou, "Isolation-based anomaly detection," *ACM Trans. Knowl. Discov. Data (TKDD)*, vol. 6, no. 1, pp. 1-39, Mar. 2012.

36 V. Hoffstein and J. Szalai, "Predictive value of clinical features in diagnosing obstructive sleep apnea," *Sleep*, vol. 16, no. 2, pp. 118-122, Mar. 1993.

37 W. R. Ruehland, P. D. Rochford, F. J. O’Donoghue, R. J. Pierce, P. Singh, and A. T. Thornton, "The new AASM criteria for scoring hypopneas: impact on the apnea hypopnea index," *Sleep*, vol. 32, no. 2, pp. 150-157, Feb. 2009.

38 X. Hui, T. B. Conroy, and E. C. Kan, "Near-field coherent sensing of vibration with harmonic analysis and balance signal injection," *IEEE Trans. Micro. Theory Tech.*, vol. 69, no. 3, pp. 1906-1916, May 2021.

39 X. Hui, J. Zhou, P. Sharma, T. B. Conroy, Z. Zhang and E. C. Kan, “Wearable RF near-field cough monitoring by frequency-time deep learning”, *IEEE Trans. Biomed. Circuits & Sys*, vol. 15, no. 4, pp. 756 – 764, Aug. 2021, doi: 10.1109/TBCAS.2021.3099865.

40 D. R. Halloran and G. R. Alexander, “Preterm delivery and age of SIDS death,” *Ann Epidemiol.*, vol. 16, no. 8, pp. 600 – 606, Aug. 2006, doi: 10.1016/j.annepidem.2005.11.007.

41 J. A. Pinto, D. K. Ribeiro, A. F. d. S. Cavallini, C. Duarte, and G. S. Freitas, "Comorbidities associated with obstructive sleep apnea: A retrospective study," *Int. Arch. Otorhinolaryngol.*, vol. 20, pp. 145-150, 2016.

42 F. Zhou, S. Yang, H. Fujita, D. Chen, and C. Wen, "Deep learning fault diagnosis method based on global optimization GAN for unbalanced data," *Knowledge-Based Systems*, vol. 187, p. 104837, Oct. 2020.

CHAPTER 4

OBJECTIVE SCORING OF PHYSIOLOGICALLY INDUCED DYSPNEA BY NON-INVASIVE RF RESPIRATORY SENSORS

Introduction

The symptom of dyspnea, or so called difficult or labored breathing, defined as a "patient's subjective awareness of uncomfortable or distressing breathing" [1], can be caused by heavy exertion, deficiency of ambient oxygen, increased airway resistance, and respiratory disorders. Dyspnea can be extremely distressing for patients with serious illness [2], such as asthma, heart failure, COVID-19 [3], and chronic obstructive pulmonary diseases (COPD) [4], [5], [6], leading to a poor quality of life [7]. The prevalence of dyspnea is common, and ranges from 33% to 76% in critically ill patients, while 85% of patients with heart failure and up to 95% of patients with COPD report dyspnea. In addition to the possible physiological consequence in cardiopulmonary functions such as low oxygen saturation levels in SaO2 and SpO2 [8], the repercussion of dyspnea can also include the sense of suffocation, distress, fear, panic or anxiety. Although dyspnea has been associated with the intertwined physiological, psychological, and social-demographic contributors [9], [10], no present theory was able to encompass all causes of dyspnea reliably. Our study design focuses on the physiological factor of dyspnea that can be induced from exertion and increased airway resistance, and we hypothesize that this kind of dyspnea will have a high correlation between the self-report and the measurable respiratory features.

In present clinical practices, dyspnea is most often assessed by different kinds

of scales self- reported by patients [11]. The most popular scales include the following scores conducted during patient interviews for specific purposes.

1) The perceived disability scale, such as the Medical Research Council (MRC) dyspnea scale, describes the breathlessness sensation out of the exercise capacity in a score of 1 – 5[12], where Score 1 denotes that the participant is not troubled by dyspnea except on strenuous exercise, and Score 5 is for patients who are too breathless to leave the house or during dressing/undressing.

2) The experiential history scale, such as the baseline and transition dyspnea index (BDI/TDI), measures changes in the three domains of dyspnea severity from functional impairment, magnitude of task, and magnitude of effort [13]. In BDI, the scale is from 0 – 12, which is the sum of 0 – 4 in each of the three domains of dyspnea severity. Score 0 denotes the severe dyspnea and Score 4 for unimpaired functions. The TDI scale ranges from –3 (major deterioration) to +3 (major improvement) in each domain from previous BDI. The popular use of BDI/TDI also illustrates the importance of continuous evaluation of dyspnea episodes.

3) The psychophysical scale, such as the Borg scale or visual analog scale (VAS), assesses symptom intensity in response to a specific stimulus such as exercise. In the Borg scale, patients report their feeling of discomfort from 0 – 10 at the moment, where 0 corresponds to the sensation of normal breathing or absence of dyspnea, and 10 corresponds to the maximum possible sensation of dyspnea [14], [15]. This is similar to the other popular Borg scale for pain.

Regardless of the scale of choice, self-reported dyspnea is subjective and variable for each person on each day, and can be challenging to assess for those who

refuse to cooperate or cannot communicate due to medical issues such as stroke, dementia, and delirium. Frequent quires to patients for continuous dyspnea evaluation are not only tedious, but can also cause stress and discontent, introducing a bias to the self-reported score. To supplement the subjective patient self-report, the association between dyspnea and respiratory metrics, such as the breath rate (BR), lung volume (LV), and inhalation/ exhalation patterns, has been carefully examined in previous works [16]. Studies of the association have been however cross sectional, lacking the ability to account for continuous metrics and their changes, often due to the discomfort or inconvenience of sensor deployment on patients [17]. Despite these limitations in past studies, it is of critical importance to ameliorate the objective methods of dyspnea evaluation, which can help clinicians make better decision on treatment and triage in patients with chronic lung and heart diseases, can enhance screening effectiveness in pandemic situations, and can enable caregivers in palliative and hospice medicine to provide timely service to those patients who are unable to communicate due to cognitive impairment, loss of language ability, or delirium.

However, current methods of respiration monitoring [18] such as respiratory inductance plethysmography (RIP)[19], strain gauge (SG) [20], spirometer [21], pneumotachometer [22], and capnography [23] are seldom practical for use outside of clinical settings due to the need of operator assistance and patient cooperation. Many of these devices are uncomfortable or require long-time connection to immobile machines, which are thus not feasible for broad deployment, especially on chronical patients with advanced lung diseases. RIP and tension-based chest belts can measure respiratory waveforms from the thoracic and abdominal movement. However, the

required tight belt tension to respond to the full waveform inevitably causes discomfort. Body movement to relax the belts to the least constraining position also brings additional artifacts. The SG sensor requires tight skin contact which is also uncomfortable for long-term monitoring. Spirometry can measure the volume and speed of inhaled and exhaled air, but involves attentive user participation and strenuous breathing protocol. The pneumotachometer is the gold-standard device for measuring airflow by detecting the pressure drop against a small resistive field, but is cumbersome to deploy and uncomfortable for users. Capnography measures the amount of CO_2 in exhaled air to derive the respiratory effort and distress, but the device is bulky and patients need to be intubated. Radar-based sensing [24], [25], [26] as another alternative respiration monitoring method can assess body surface motion as a result of breathing, but requires a reader in the direct line of sight (LoS) to the front chest of the user. Ambient and body motion can cause severe interference, although BR during tidal breathing can be reasonably derived through careful filtering.

In this work, we opt to use the wearable radio-frequency (RF) sensors [27], [28] based on near-field coherent sensing (NCS) to monitor respiratory features over layers of clothing or furniture fabrics to enhance user comfort and convenience. In the near-field region, the dielectric boundary movement by internal organs and muscles will be modulated on the antenna characteristics, and then be received in a multiplexed channel [28] in a non-invasive manner. Due to its touchless operation, many subjects under test will not notice the on-going sensing activities at all [29].

The sensation of dyspnea can be voluntarily and objectively induced from either physical exertion or increase in airway resistance, which is the assumption

behind the MRC scale as the perception of the exertion capabilities. We began our study by collecting data from 32 healthy participants wearing the respiratory sensors including the NCS sensor and the commercial chest belts for sensor validation. Participants were instructed to report dyspnea scores D_{self} frequently in the Borg scale under various exertion and breathing exercises following a given protocol. Scales based on patient perception or experience such as MRC and TDI were less feasible for real-time exertion-based studies. Four respiratory features of BR, LV, and inhalation and exhalation intervals were extracted from the continuous waveforms. We defined 15 respiratory metrics out of the 4 respiratory features, and constructed the machine-learning (ML) models based on the decision tree [30] and random forest [31] to investigate the correlation between the respiratory metrics and D_{self} during the entire protocol. The resulting ML model can produce an objective dyspnea score D_{obj} in cross validation with D_{self}, and can also identify the individual importance factors of respiratory metrics in determining D_{obj}.

The main contribution of this work is to provide a new way to generate the objective score for physiologically induced dyspnea, using a comfortable and continuous respiratory sensor and an established ML model which can simultaneously consider multiple factors with different importance weighting. As far as we know, this is the first study to examine the association between dyspnea sensation and continuous respiratory metrics that account for changes in respiratory behavior over a period of time under exertion and increased airway resistance. The objective dyspnea score D_{obj} can potentially complement or substitute the self-report dyspnea score D_{self}. When more comprehensive clinical data in established patient population are available for

training in the future, this model can potentially assist clinicians and caregivers in more reliable diagnosis and treatment of dyspnea.

Methods

Experimental setup

NCS is based on the near-field coupling of ultra-high frequency (UHF) electromagnetic (EM) waves with the dielectric boundary movement of internal organs and body parts. Fig. 1(a) shows the NCS deployment together with the reference commercial sensors by BIOPAC (BIOPAC Systems, Inc., Goleta, CA), including electrocardiogram (ECG) and tension-based torso belts. Fig. 1(b) shows the photo of two software-defined radios (SDR) by National Instrument Ettus B210 to implement the NCS transceivers [28], [29]. We utilized both chair-integrated notched transmission lines [32] and wearable antennas [27] as the NCS sensing elements, as shown in Figs. 1(c)(d). For the wearable sensor, we used an antenna pair as the transmitter (Tx) and receiver (Rx). UHF radiation can penetrate dielectrics in the near-field of the sensor, so the dielectric composition will modulate the EM distribution, and be exhibited in the cross-coupling scattering parameter S_{21} of the antenna pair. Alternatively, the notched sensor was constructed by a miniature coaxial RF cable with the metal shield of the middle part removed. One end will be connected to Tx and the other to Rx. The dielectric boundary movement by lungs and associated muscles will be coupled to the leaked EM energy from the notched part of the RF cable, and hence can be detected either at Tx as signal reflection S_{11} or at Rx as signal transmission S_{21}. Notice that the SDR setup can use a digitally modulated

superheterodyne signal to improve the channel isolation from ambient RF signals [27], [32], and high-quality measurements do not rely on anechoic RF chambers.

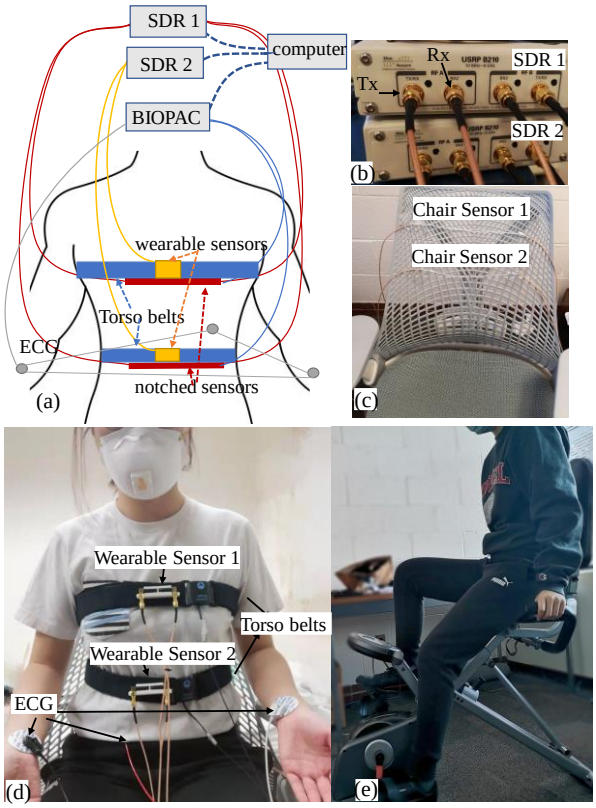


Fig. 1. The experimental system: (a) Schematics of NCS and BIOPAC sensors and data flow; (b) The photo of software-defined radio (SDR) transceivers that were connected to NCS sensors;(c) Chair-integrated NCS sensor setup; (d) A participant wearing NCS and BIOPAC torso belt sensors in the sitting position during breathing through a facemask. Two NCS sensors and two belts were deployed at the thorax and abdomen position. (e) A participant undergoing physical exertion. Written informed consent was obtained from the participants to publish their photos.

The two wearable NCS sensors were placed on the chest and the abdomen in the front torso, and the two notched sensors were integrated to the back of a chair behind the thorax and abdomen. The two sets of NCS sensors can be operated individually or at the same time. A participant wearing all sensors, including NCS, the torso belts and ECG in the sitting posture, is shown in Fig. 1(d) when breathing

through an N95 facemask. BIOPAC and two sets of NCE sensors were included to verify that the conclusion on association between D_{obj} and D_{self} would be sensor independent. The exercise bike used in the exertion protocol is shown in Fig. 1(e). The synchronization between SDR and BIOPAC recordings is achieved by buffering both data at approximately the same time within a few milliseconds.

Both SDR and BIOPAC were connected to the host computer through USB (Universal Serial Bus). In the SDR Tx signal chain, the digital baseband went through the digital-to-analog converter (DAC) and was then mixed with the carrier frequency f_{RF}. The RF power is less than −10 dBm or 0.1 mW, well under the safety limits set by occupational safety and health administration (OSHA) in the UHF band. The RF signal emitted from Tx was then coupled into the internal body motion within the near field. The modulated signal was received by Rx, and then demodulated and sampled by the analog-to-digital converter (ADC) to retrieve the motion characteristics in the baseband. The quadrature baseband signals $I_{TX}(t)$ and $Q_{TX}(t)$ were presently implemented by sinusoidal monotones with the NCS signal as amplitude modulation [27]:

$$I_{Rx}(t) = NCS(t)\cos(2\pi f_{BB} t + \theta_0) \quad (1)$$

$$Q_{Rx}(t) = NCS(t)\sin(2\pi f_{BB} t + \theta_0) \quad (2)$$

$$NCS_{amp}(t) = \sqrt{I_{Rx}(t)^2 + Q_{Rx}(t)^2} \quad (3)$$

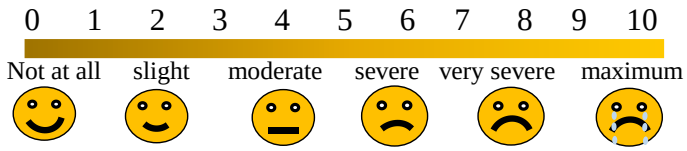
where θ_0 was the phase offset accumulated from the Tx-Rx signal chains. f_{RF} was selected at 900 MHz for the wearable sensors, and 1GHz for the notched sensors. Each B210 was used for two Tx-Rx channels with the two basebands at f_{BB1} =355 kHz

and f_{BB2} = 440 kHz, both sampled at 1M samples per second (Sps). The demodulated respiration waveform was further down-sampled to 500 Sps, which was the same as the BIOPAC data.

Human Study Protocol

The experimental setup described in Sec. II.A was applied to evaluating the respiratory waveforms of 32 healthy participants. The human study has been approved by Cornell Institutional Review Board (IRB) Protocol ID #1812008488. Written informed consent to take part in the study was obtained from all participants. Participants were instructed to follow a sequence of routines as documented in Table I. The breathing exercise includes: 0 – 30s fast breathing, 30 – 60s slow breathing, 60 – 90s normal breathing, 90 – 120s fast breathing, 120 – 150s slow breathing, and 150 – 180s normal breathing to build a library of various breathing patterns. The participant followed a voice instruction at the beginning of each section, and reported a dyspnea score D_{self} in the Borg VAS scale after each routine as shown in Fig. 2 [14]. The participant sat on a chair for all routines except during Routines 4 and 6 of exertion. Dyspnea was induced by aerobic rope jumping and exercise biking, as well as by wearing an N95 mask to increase the airway resistance. Fig. 3 presents examples of NCS respiratory waveforms during the study protocol. When the routines in dyspnea (red curves) were compared with routines without dyspnea (green curves), in Routine 3, BR decreases due to partial airway obstruction and breath-to-breath variation also decreases; in Routines 5 and 7, BR increases and breath-to-breath variation decreases due to post-exertion. We will further analyze these changes by extracting quantitative

metrics in the respiratory waveforms in the later sections.



0: Relaxed as in sleeping or watching TV.
1-3: Can maintain for hours.
4-6: Can hold short conversation. But noticeable challenging.
7-9: Difficult to maintain exercise. Can barely speak a sentence.
10: Impossible to keep going. Unable to talk.

Fig. 2. Description of the self-reported Borg visual analog scale (VAS) for dyspnea evaluation [14].

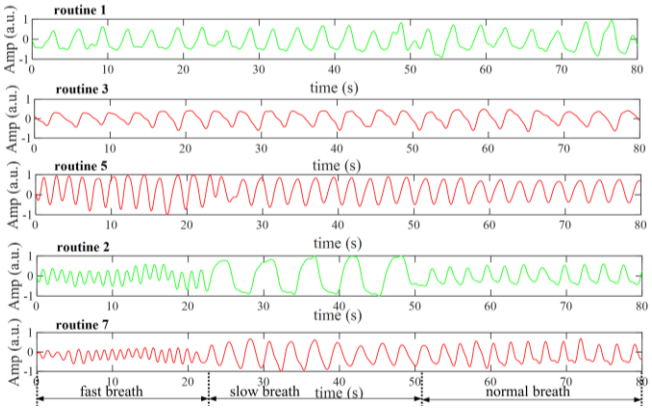


Fig. 3. An example of NCS respiratory waveforms during the study protocol. Routine 1: Normal breathing; Routine 3: Normal breathing with a facemask; Routine 5: Normal breathing after physical exertion; Routine 2: Breathing exercise; Routine 7: Breathing exercise after physical exertion. The illustration period is truncated from 10 to 90 s of each routine. Green curves indicate absence of dyspnea, and red curves indicate some degrees of dyspnea.

Data processing

Physiological manifestation of dyspnea

The main purpose of respiration is to supply oxygen to body cells by circulation, with the auxiliary functions of making sound, sniffling, and clearing of

airway by coughing and sneezing. Respiration can be initiated involuntarily and voluntarily, and the voluntary part can be trained. When the blood oxygen saturation (SaO2) is low or CO_2 high, the breathing action will be triggered for more lung ventilation. However, when the body cannot respond fast enough due to various reasons such as airway obstruction, insufficient ambient oxygen supply, weakened respiratory muscles, or voluntary control for speaking, singing or holding, the feeling of dyspnea will arise quickly. To increase lung ventilation, often BR and LV will increase by panting or deep breathing. Alternatively, the inhalation and exhalation intervals will be adjusted depending on the muscle condition, airway obstruction, and ambient factors. As the respiratory reaction to dyspnea can be trained to reduce the uncomfortable feeling, similar to experiential avoidance coping of pain, another common physiological reaction to dyspnea is the reduction of variability in successive breaths [16] together with speaking restraint, when the body tries hard to use the best known breathing cycle to reduce the discomfort of dyspnea.

Therefore, we propose to use the respiratory features of BR, LV, and inhalation and exhalation intervals to correlate to dyspnea manifestation. The mean and variation of these features within a chosen epoch as well as the variation between successive breaths will be extracted from the breathing waveforms for further data processing. Notice that here we will not complicate our protocol with speaking and coughing, as they can be separately identified from their high-frequency characteristics [33]. For future extension to realistic continuous monitoring, both speaking and coughing will need to be accounted for.

Feature extraction

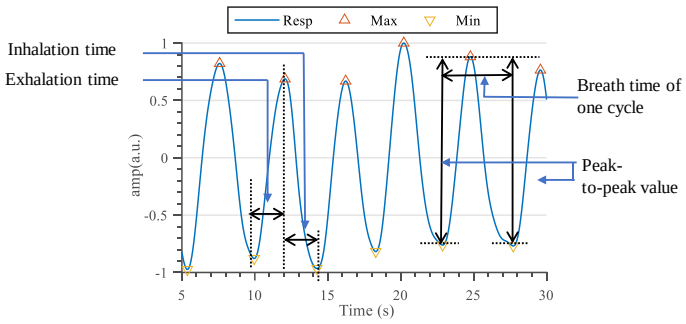


Fig. 4. An example of feature extraction from the NCS respiratory waveforms. The blue line labeled as 'resp' is the raw NCS waveform after bandpass filtering of 0.05 – 1 Hz, and the red and yellow triangles are detected peaks of maximum and minimum by the moving-average crossing method. The breath rate, peak-to-peak value, inhalation interval and exhalation interval can be estimated out of the detected peaks.

The retrieved respiratory waveforms from 4 NCS and 2 BIOPAC torso-belt sensors were first bandpass-filtered from 0.05Hz to 1Hz to remove the DC drift and high-frequency noises. Various sensor combinations will be further studied in the dyspnea recognition below. The filtering processing was implemented in MATLAB by the digital infinite impulse response (IIR). We then utilized the moving average-crossing algorithm [34] to detect peaks of the breathing waveform. A moving-average curve was first calculated at each time point in a given window length, which was around one respiration cycle and would be constantly updated. The points when the moving-average curve crossed the original signal were marked as up-crossing points for positive slopes in the original signal or down-crossing points for negative. Local maximum was labelled as the maximal point between two up-down crossing points, and local minimum as the minimal point between two down-up crossing points. As shown for an example in Fig. 4, the blue line was the filtered respiratory waveform from the wearable NCS sensor, and the red triangles and yellow triangles marked the

maximum and minimum peaks detected by the algorithm. Then, we can extract the 4 respiratory features in each breath cycle to represent the instantaneous respiratory characteristics: 1) BR was calculated from the inverse of the interval between two neighboring minima; 2) The peak-to-peak (PP) value representing LV [27] was estimated by the signal difference in successive peaks; 3) The inhalation interval (IN) was evaluated by the time difference between one minimum and the following maximum; 4) The exhalation interval (EX) between one maximum and the following minimum. The respiratory waveform from the torso belts was processed in the same way. The peak-detection algorithm is of critical importance for accurate feature extraction and subsequent processing.

Respiratory metrics

TABLE I. RESPIRATORY METRICS FROM RESPIRATORY FEATURES.

	BREATH RATE (BPM)	PEAK-TO-PEAK (A.U.)	INHALATION INTERVAL (S)	EXHALATION INTERVAL (S)
COEFFICIENT OF VARIATION	CoV_{BR}	CoV_{PP}	CoV_{IN}	CoV_{EX}
MEAN	μ_{BR}		μ_{IN}	μ_{EX}
AUTOCORRELATION	$R1_{BR}$	$R1_{PP}$	$R1_{IN}$	$R1_{EX}$
SUCCESSIVE DIFFERENCES	$R2_{BR}$	$R2_{PP}$	$R2_{IN}$	$R2_{EX}$

After calculating the above 4 respiratory features, we defined 15 metrics that serve as the input to the ML model as shown in Table I. The first 4 metrics were the coefficient of variation (*CoV*) of the above 4 respiratory features, which was defined as

$$CoV = \left(\frac{\sigma}{\mu}\right)^2 \quad (4)$$

where σ denotes the standard deviation and μ denotes the mean. At each sampling point, *CoV* was calculated over an epoch of T = 15s. We first found all breath cycles in the previous 15s and calculated BR of every cycle, whose μ_{BR} and σ_{BR} were estimated to obtain CoV_{BR}. CoV_{BR} were then averaged over all epochs in the given routine. *CoV* for PP, IN and EX were derived in the same way. These respiratory metrics can represent the breath variability within the epoch. Mean values were estimated as μ_{BR}, μ_{IN} and μ_{EX} for the BR, inhalation and exhalation intervals during each routine. As PP was normalized and contained the bias from personal deployment, μ_{PP} was excluded from the respiratory metrics.

To further capture variability between adjacent breaths, we used autocorrelation in a time lag of one breath cycle to measure the successive similarity of a given respiratory feature. The autocorrelation function *R1* is defined as,

$$c_i = \frac{1}{N}\sum_{n=1}^{N-1} y_n \cdot y_{n+i} \quad (5)$$

$$R1 = \frac{c_1}{c_n} \quad (6)$$

where n is the breath index within a total of N breaths in the epoch of T = 15s during metrics evaluation, y_n is the discrete breath-by-breath measurement of the selected respiratory feature, and y_{n+i} is the same feature lagged by i breaths. We also define a similarity measure of *R2* as the mean absolute difference between adjacent breaths:

$$R2 = \frac{1}{N}\sum_{n=1}^{N-1} \frac{|y_{n+1} - y_n|}{y_n} \quad (7)$$

Fig. 5 shows a representative sample analysis using the NCS recording for two routines. Fig. 5(a) is from the routine with D_{self} = 0, i.e., no sense of dyspnea, while Fig. 5(b) from a routine with D_{self} = 7 after heavy exercises. For the higher D_{self}, μ_{BR} is higher but CoV_{BR} remains similar, μ_{PP} is higher with smaller CoV_{PP}, and μ_{IN}, CoV_{IN}, μ_{EX} and CoV_{EX} are all reduced due to the faster breathing with more regularity. Increase in autocorrelation R1 and decrease in successive differences R2 are consistently observed for all 4 respiratory features in our data, implicating reduction in variability between adjacent breaths.

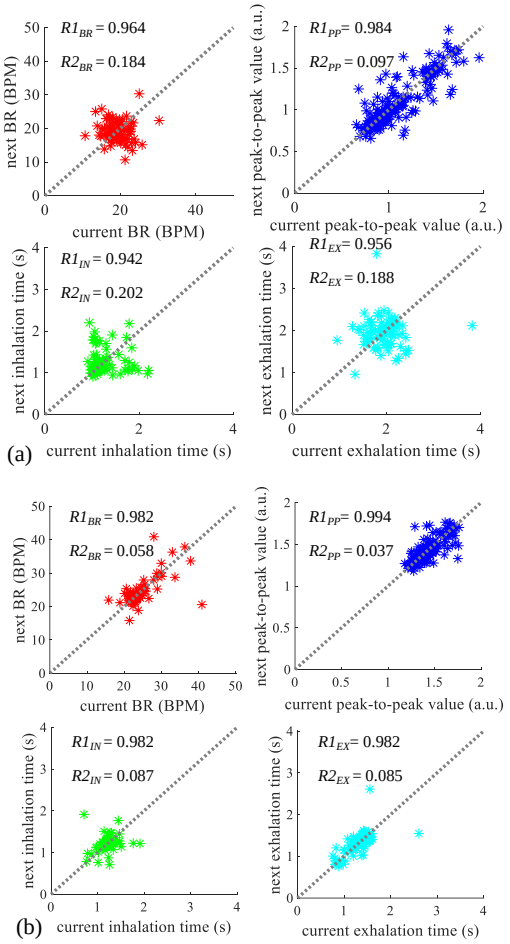


Fig. 5. Analysis of the breath rate (BR), lung volume (LV) by peak-to-peak magnitude, inhalation interval and exhalation interval between adjacent breaths by NCS during 5-min recordings. (a) Normal breathing under a self-report dyspnea score of 0; (b) Normal breathing

after exercises with a self-report dyspnea score of 7. Decrease of variability in the respiratory features can be observed by increasing R1 and decreasing R2.

Sensor fusion considerations

To investigate the error from the sensor hardware, we compare the analyses using the wearable and chair-integrated NCS as well as the BIOPAC chest belts. To expand the dyspnea model, we also added the cardiogram features from NCS bandpass-filtered between 0.5 and 3 Hz or BIOPAC ECG to the combined model learning. Three heartbeat metrics were included: 1) The mean heart rate (HR) μ_{HR}, 2) the standard deviation of NN intervals σ_{NN}, [35] where NN denotes the normal RR distance in the QRS cardiogram complex, and 3) the root mean square of the successive differences between adjacent NN RMS_{NN}. When both NCS and BIOPAC data were used together, the number of cases would be doubled. The heartbeat signals added 3 additional metrics to the original 15 metrics of respiration. Notice that ideally the NCS and BIOPAC sensors would measure the same cardiopulmonary features, but each had its own noises and errors during measurements that can cause inaccurate prediction of D_{obj}. Combination of sensors measuring the same physiological features may or may not improve the overall accuracy, as inconsistent derivation of intended metrics can aggravate the ambiguity, unless the noise can be assumed to be totally uncorrelated in the fusion of a large number of sensors. The combination of two sets of NCS and BIOPAC sensors cannot guarantee such an assumption. To understand whether the wearable or chair-integrated NCS sensor can be used alone in the applications, we compare each individual data set and the combination as well.

Results

Personal calibration

We extracted the respiratory and heartbeat metrics from the breathing and cardiac waveforms of 32 healthy participants using a protocol with various levels of exertion and airway obstruction in Table II. The breathing exercises in Routines 2 and 7 contain specified inhalation and exhalation instructions for the participant to follow, which cover multiple periods of fast, slow and hold breathing. Each routine in the protocol contains a self-reported dyspnea score D_{self} and all measured respiratory metrics. Because the tidal breathing pattern varied from person to person, we opted to first calibrate out the personal difference by subtracting the respiratory metrics in the normal breathing routine (D_{self} = 0 as the baseline) from the exertion routines (after exercise or wearing a facemask with D_{self} = 1 – 9) for the same participant.

TABLE II. ROUTINES IN THE HUMAN STUDY PROTOCOL.

ROUTINE	DURATION (MINUTES)	ROUTINE CONTENT	RESPIRATORY MONITORING
1	5	NORMAL BREATHING	ON
2	3	BREATHING EXERCISES	ON
3	5	NORMAL BREATHING WITH FACEMASK	ON
4	3 OR 10	ROPE JUMPING OR EXERCISE BIKING	OFF
5	5	NORMAL BREATHING	ON
6	3 OR 10	ROPE JUMPING OR EXERCISE BIKING	OFF
7	5	BREATHING EXERCISE	ON

This personal calibration against the normal breathing routine reduced the individual biases in D_{obj}, but the case of D_{obj}= 0 would be excluded from the ML model output. It is important that the eventual model can also give reliable D_{obj}= 0 for

negative dyspnea cases. To remedy the case of D_{self} = 0, we imputed the input with the new cases where the respiratory metrics in the first half of normal breathing was subtracted from the second half. The difference between the two halves of the first normal breathing routine was used as the training under D_{self} = 0. Extreme dyspnea scenario (D_{self} = 10) was not included here due to safety concerns in the human study. Extrapolated respiratory metrics for complete obstructive and central apnea have been attempted by data imputation, but with only limited success. Hence, the extreme cases of D_{self} = 10 will be left for future clinical studies when experimental observation can be available.

The decision-tree and random-forest models

To produce D_{obj} from the measured respiratory metrics, we chose the decision-tree regressor [30] as the ML model for the following reasons. (1) Decision tree is a white-box model, so the physical explanation for the result can be observable through the tree structure, which can help us understand the physiological correlation between the respiratory metrics and D_{self}. (2) Decision tree helps dominant feature selection in multitudinous respiratory metrics. Irrelevant respiratory metrics will be assigned a less importance weight to evolve with the dominant features, and the importance factors can be part of the model output for physiological reasoning. (3) The regressor model is preferred over the classifier because the output can be a continuous quantity of the predicted D_{obj}. Although D_{self} in the Borg scale is discrete for subjective convenience, a continuous D_{obj} can reduce the ambiguity between fine discrete levels. To illustrate the advantage of decision-tree regressor against the popular discrete classification

methods such as the principal component analysis (PCA), we showed the scattered plots of the two dominant features of D_{self} found by the decision-tree model. It can be observed that clusters of D_{self} of similar values cannot be identified in any reasonable hyperplane separation. As PCA can only capture linear correlation, the inability of classification by PCA indicates that the correlation between D_{self} and respiratory metrics is more convoluted and ambiguous. In comparison, the decision tree can incorporate nonlinear relationship into the model with reasonable tolerance of ambiguous contributors.

Table III. data composition in the ML model.

TRAINING SET (95 CASES):		
NUMBER OF CASES	D_{SELF}	RESPIRATORY METRICS
32	1 – 9	CALIBRATED NORMAL BREATHING AFTER EXERCISE IN ROUTINE 5
31	1 – 9	CALIBRATED NORMAL BREATHING WITH FACEMASKS IN ROUTINE 3
32	0	IMPUTED NORMAL BREATHING BY TWO HALVES OF ROUTINE 1
TESTING SET (30 CASES):		
23	1 – 9	CALIBRATED BREATHING EXERCISE AFTER EXERCISE IN ROUTINE 7
7	0	IMPUTED NORMAL BREATHING BY TWO HALVES OF ROUTINE 1

After training in the 95 cases in Table III, the decision-tree model can predict D_{obj} on the unseen 30 testing cases based solely on the respiratory metrics, which can then be compared with D_{self} to assess the model accuracy. In Table III, the 95 training cases consist of measurements from Routine 1, 3 and 5 with observation after exertion

and during facemask wearing, as well as cases of D_{self} = 0 for calibration and imputation. One participant opted out of the facemask wearing routine during the study.

TABLE IV. PREDICTION ACCURACY FROM DATA SETS IN TABLE III.

DATA SET	NCS + TORSO-BELT + ECG	NCS	TORSO-BELT
MODEL	DECISION TREE	DECISION TREE	DECISION TREE
MEAN η BY K-FOLD CROSS-VALIDATION	0.876	0.825	0.818
FEATURE IMPORTANCE	μBR=0.402 R2IN=0.149 COVEX=0.076 RMSNN =0.068	μBR=0.402 R2BR=0.151 R2EX=0.075 R1IN =0.071	μBR=0.330 R2IN=0.290 R2BR=0.078 COVEX=0.074
MEAN η BY LEAVE-ONE-PARTICIPANT-OUT CROSS-VALIDATION	0.841	0.864	0.854
FEATURE IMPORTANCE	μBR=0.463 COVEX=0.126 R2IN=0.111 RMSNN =0.083	μBR=0.443 R2BR=0.198 COVPP=0.054 R2EX=0.045	μBR=0.330 R2IN=0.262 COVEX=0.074 R2BR=0.056
η FOR TESTING DATA	0.872	0.871	0.834
DATA SET	NCS + TORSO-BELT + ECG	NCS	TORSO-BELT
MODEL	RANDOM FOREST	RANDOM FOREST	RANDOM FOREST
MEAN η BY K-FOLD CROSS-VALIDATION	0.884	0.866	0.848
FEATURE IMPORTANCE	μBR=0.280 R2IN=0.150 R2BR=0.103 COVEX=0.054	μBR=0.232 R2BR=0.164 R2EX=0.103 R1IN =0.082	μBR=0.223 R2IN=0.184 R2BR=0.130 COVBR=0.068
MEAN η BY LEAVE-ONE-PARTICIPANT-OUT CROSS-	0.874	0.881	0.866

VALIDATION			
Feature importance	μ_{BR}=0.332 $R2_{IN}$=0.161 $R2_{BR}$=0.068 CoV_{Ex}=0.056	μ_{BR}=0.232 $R2_{BR}$=0.164 $R1_{IN}$=0.106 $R2_{EX}$=0.080	μ_{BR}=0.229 $R2_{IN}$=0.183 $R2_{BR}$=0.128 CoV_{Ex}=0.087
η for testing data	0.903	0.907	0.873

In order to estimate the skill of our model on new data, I first performed the procedure of *k*-fold and leave-one-participant-out cross-validations on the 95 training cases[36]. *K*-fold cross-validation can investigate the robustness to unseen data, and leave-one-participant-out cross-validation can test the robustness to unseen participants. For *k*-fold cross-validation, we divided the whole training set of 95 cases into separate training (76 cases) and testing (19 cases). We chose k = 5 and the model was trained using 4 folds as the training data and the resulting model is validated on the remaining fold as the testing data. For leave-one-participant-out cross-validation, the model was trained on the data sets from 31 participants excluding one participant, who was then used as testing by generating D_{obj} to compare with D_{self}. The validation process was repeated for each participant as the testing case.

Because we used the regressing estimator, the predicted D_{obj} can be a continuous number from 0 to 9. The upper limit of D_{obj} = 9 is due to the lack of training cases with D_{self} = 10. We define a prediction accuracy η of the ML model as:

$$\eta = 1 - \frac{|D_{obj} - D_{self}|}{9} \quad (8)$$

where the error is the normalized absolute distance of D_{obj} to D_{self}. The maximum value of η for perfect prediction is 1. In the *k*-fold cross-validation procedure, mean η for the validation set ranges from 0.818 to 0.881. In the leave-one-participant-out cross validation, mean η fell within similar ranges. When testing on the 30 unseen data set as shown in Table IV, η ranges from 0.834 to 0.907. The 30 unseen testing set contains Routine 7 recording from 23 participants, as 9 participants out of 32 cannot follow the breathing exercise after exertion, as well as 7 participants who had repeated Routine 1 in different study dates. The accuracy and the metrics of importance for all three testing scenarios are summarized in Table IV. The metric with a higher importance factor has a higher correlation with D_{self}. The sum of the importance factors from all features was normalized to 1 in each method. To understand the magnitude of mean η better, I performed total random guesses of D_{obj} for the unseen 30 testing cases in Table III, where η would range from 0.566 to 0.677. A fixed guess of D_{obj} in all dyspnea prediction will render η ranging from 0.396 (D_{obj} = 9) to 0.766 (D_{obj} = 4). When more cases are available with homogeneous distribution across all possible values of D_{self}, η will approach 0.5 for random and fixed guesses.

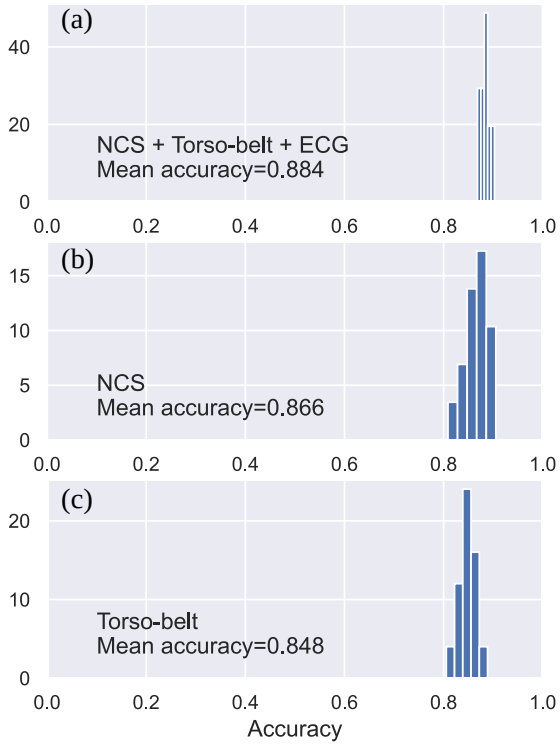


Fig. 6. Accuracy distribution by *k*-fold cross-validation using the random forest model. (a) Training data from wearable NCS + torso-belt +ECG; (b) Wearable NCS only; (c) BIOPAC torso-belt only.

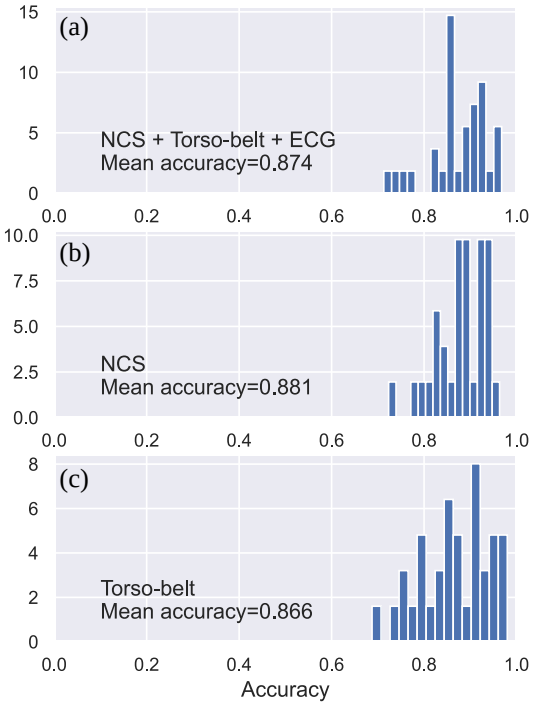


Fig. 7. Accuracy distribution by leave-one-participant-out cross-validation using the random forest model. (a) Training data from wearable NCS + torso-belt + ECG; (b) Wearable NCS only; (c) BIOPAC torso-belt only.

In Table IV, I added the random-forest method composed by an ensemble of decision trees [31] as the ML model. I first used the *k*-fold cross-validation method to choose multiple decision trees that have high prediction accuracy in the 95 training sets, and then their average would be used to predict D_{obj} for the 30 unseen testing data. I also listed the top four most important metrics in generating the final random-forest model under consideration. As observed from Table IV, our model can predict the dyspnea score in the unseen testing data with reasonable accuracy. Random forest can only improve η marginally, but tends to have smaller variation, as will be seen in the later Bland-Altman (B&A) plots. This was possibly because the set of optimal decision trees was reasonably similar, as can be observed from the dominant feature importance with the same input. The accuracy difference between NCS and NCS + BIOPAC was also trivial, which suggested that the NCS sensor has captured the respiratory metrics with sufficient accuracy, and the torso belts in BIOPAC do not offer new information.

Fig. 6 presents the accuracy distribution during *k*-fold cross-validation using the random forest model. Fig. 6(a) used the training data from NCS + torso-belt + ECG sensors, Fig. 6(b) from only NCS sensors, and Fig. 6(c) from only torso-belt sensors. Fig. 7 presents the similar accuracy distribution by leave-one-participant-out cross-validation. By separating one independent participant's data as the testing set and estimating the accuracy on each participant, our model has been evaluated to have a good performance to predict an unseen participant. In general, NCS performs

slightly better than the torso belt in cross validation, but the difference is probably within error ranges. Notice that the two sensors have similar signal-to-noise ratios (SNR) [27].

Additional scattered plots for all columns in Table IV are shown in the Fig. 8(a)(b) where the error in different D_{self} values can be more clearly observed. I also investigated the effectiveness of the wearable and chair-integrated NCS in Table V, as these two setups can be applied to different clinical applications, for example, the wearable sensor for patients in the pulmonary ward and the chair-integrated sensor in the observation room. The accuracy difference between wearable and chair-integrated NCS was insignificant. I can also conclude that the touchless NCS sensors alone can generate the ML model with high validity in either the wearable or chair setup. Scattered and B&A plots for Table V are shown in Figs. 8(c)(d) and 9, where NCS in both setups are shown to produce D_{obj} with reasonable limits of agreement (LoA) and no significant systematic bias m. The random-forest model has tighter LoA than the decision-tree model, probably due to the variation reduction during ensemble averaging.

TABLE V. COMPARISON OF NCS WEARABLE AND CHAIR SETUPS.

DATA SET	WEARABLE NCS	CHAIR NCS	WEARABLE NCS	CHAIR NCS
MODEL	DECISION TREE	DECISION TREE	RANDOM FOREST	RANDOM FOREST
MEAN η BY K-FOLD CROSS-VALIDATION	0.822	0.825	0.865	0.866
MEAN η BY LEAVE-ONE-	0.861	0.865	0.883	0.883

PARTICIPANT-OUT CROSS-VALIDATION				
η FOR UNSEEN TESTING DATA	0.855	0.879	0.906	0.903

TABLE VI. TESTING RESULTS BY SEPARATING EXERTION AND FACEMASK.

DATA SET	NCS + TORSO-BELT + ECG	NCS	TORSO-BELT
MODEL	DECISION TREE	DECISION TREE	DECISION TREE
MEAN η BY *K*-FOLD CROSS-VALIDATION	0.937	0.868	0.900
MEAN η BY LEAVE-ONE-PARTICIPANT-OUT CROSS-VALIDATION	0.940	0.917	0.941
η FOR UNSEEN DATA	0.798	0.773	0.764
DATA SET	NCS + TORSO-BELT + ECG	NCS	TORSO-BELT
MODEL	RANDOM FOREST	RANDOM FOREST	RANDOM FOREST
MEAN η BY *K*-FOLD CROSS-VALIDATION	0.944	0.878	0.902
MEAN η BY LEAVE-ONE-PARTICIPANT CROSS-VALIDATION	0.939	0.905	0.929
η FOR UNSEEN DATA	0.802	0.828	0.790

To further test the transferrable capability of the ML model, I also shifted the 31 cases of normal breathing with facemask in Routine 2 from the training set to the testing set. Facemask only increases the lung elasticity, and causes the dyspnea sensation in a different manner from exertion. The result was presented in Table VI. The accuracy for the cross validation becomes slightly higher as the data were collected from more consistent routines. The accuracy for the unseen data from

Routine 2 was lower than those in Table IV, partially because less training data was provided and the testing data came from totally unseen routines. However, the accuracy remains reasonably high, which showed the model extendibility to predict dyspnea scores in unseen routines of different dyspnea contributors, possibly even in patients with various respiratory disorders. The corresponding scattered plots for the columns of the unseen data in Table VI are shown in Figs. 8 (e)(f), where both torso-belt and NCS had non-negligible numbers of cases where D_{self} was high but D_{obj} was close to 0. As the dyspnea after exertion is used for training and the dyspnea during facemask wearing for testing, these cases indicated that the respiratory features in Routine 2 of facemask wearing was more similar to normal breathing and less to dyspneic episodes after exertion, even though the participant reported a high D_{self}. However, D_{obj} from 'NCS + torso-belt + ECG' had much fewer such cases. Our conjecture is that the cardiac information can help during model transference.

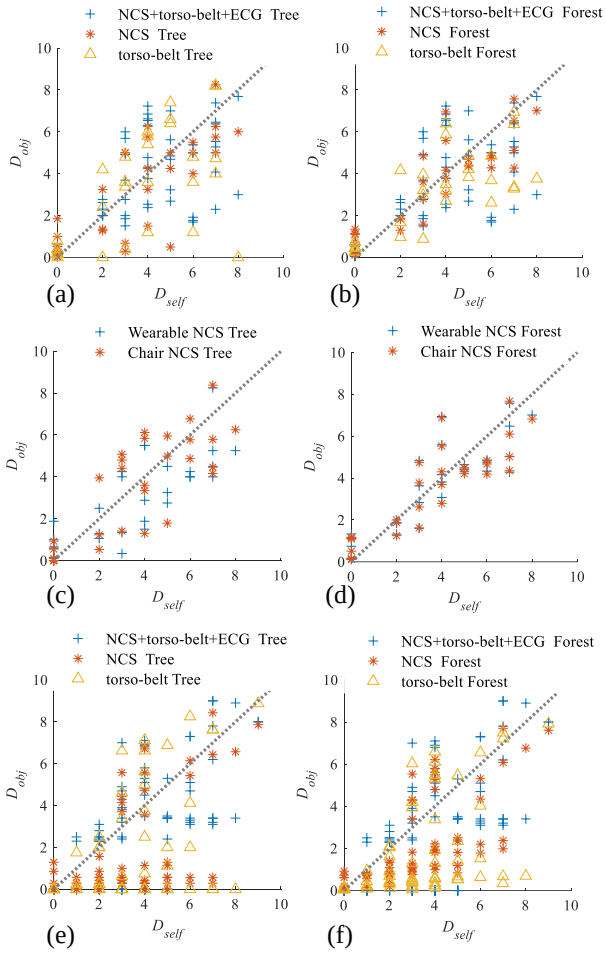


Fig. 8. The scatter plots between D_{self} and D_{obj} for all columns in (a) (b) Table IV, (c) (d) able V, and (e) (f) Table VI.

Further observation from Table IV indicates that the additional information from ECG boosts the prediction accuracy in general, although not by much, especially in the case of the random-forest model. RMS_{NN}, a form of the heart-rate variation (HRV), is the fourth most important metric for the sensor fusion with ECG, although the importance factor is much smaller than the successive breath variation. During dyspnea sensation, people are often under some degrees of psychological stress, which can then indirectly influence HRV [37], [38] to become a contributor for dyspnea recognition. For the two cases in Table VI where the increased airway resistance is in

the testing data but not in the training data, ECG can sometime causes overall accuracy degradation, as in the case of random forest between 'NCS + torso-bet + ECG' and NCS alone. This is also explainable as mask wearing can cause certain degrees of stress for some people who had to make efforts to breath in sufficient air. However, during the COVID-19 pandemic era when the human study had been done, many people were used to wearing masks without any physical or psychological stress. Thus, the heartbeat information becomes an inconsistent and ambiguous contributor in the facemask wearing cases that can degrade the dyspnea prediction.

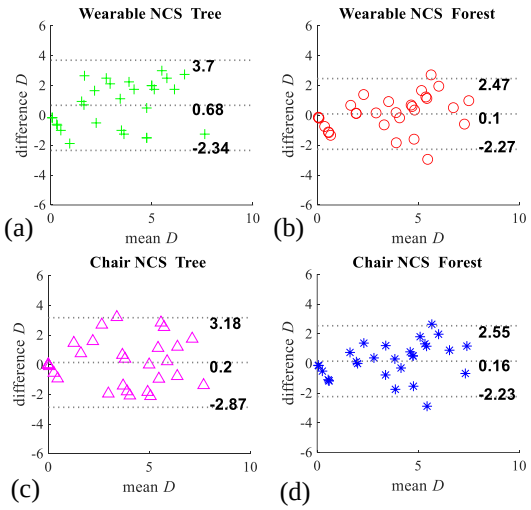


Fig. 9. The Bland-Altman (B&A) plots between D_{self} and D_{obj} for all columns in Table V. Plots show the mean difference m at the center dotted line and the corresponding limits of agreement (LoA) at the upper and lower dotted lines given by $m \pm 1.96\sigma$.

Discussion

Barriers to accomplish a mature D_{obj} model still remain. For example, D_{self} has individual biases and tolerances due to the multiple factors in the cause and outcome of dyspnea even for the same person in different days of testing, so the training data might not be repeatable and inevitably contained subjective variations, which will be

trickling down to the training model as ambiguity. Nevertheless, the Borg scale has been designed with small intervals, and small variations in D_{self} can be acceptable from large data set. To establish the simplified baseline of objective dyspnea scoring, our present study only includes healthy participants with physiologically induced dyspnea by exertion and facemasks in a prescribed study protocol. The respiratory features by psychological contributors or in patients with chronic cardiopulmonary diseases have not been examined, and can be different from dyspnea induced by physical exertion. It will be critical to extend to large-scale patient studies on real patients with various cardiopulmonary disorders to establish the true effectiveness of the proposed ML model. However, the present study offers a useful starting point for the future studies of specific respiratory disorders, especially towards the patient population who refuse to cooperate or cannot communicate due to deterioration or loss of mental functions, where the D_{obj} model has to be achieved through transference as individual D_{self} cannot be available.

Conclusion

In conclusion, dyspnea is an important common symptom for respiratory disorder screening, diagnosis and prognosis. Objective dyspnea scores to complement self-report are especially important for patients with serious illness and at the end of life when communication or cooperation is compromised due to dementia, delirium, anesthesia, and other restraining procedures. In this work, I show that the non-invasive NCS sensors can continuously capture useful respiratory features during various levels of dyspnea physiologically induced by exertion and increased airway resistance. We

have performed human study on 32 healthy participants and constructed a learning-based model to identify the correlation between continuous respiratory metrics and self-reported dyspnea score, and hence can predict an objective dyspnea score induced by physiological reason with reasonable accuracy. In future clinical study, there can be additional intertwined contributors to dyspnea in patients under different disorders and conditions, but our present study can provide a baseline of physiological analyses and a useful reference to the eventual prediction model

REFERENCES

1 M. L. Campbell, "Respiratory distress: a model of responses and behaviors to an asphyxial threat for patients who are unable to self-report," Heart & Lung, vol. 37, no. 1, pp. 54-60, Jan. 2008.

2 A. H. Kamal, J. M. Maguire, J. L. Wheeler, D. C. Currow, and A. P. Abernethy, "Dyspnea review for the palliative care professional: assessment, burdens, and etiologies," J. Palliat. Med., vol. 14, no. 10, pp. 1167-1172, Oct. 2011.

3 D. Wang et al., "Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China," Jama, vol. 323, no. 11, pp. 1061-1069, Mar. 2020.

4 J. P. Solano, B. Gomes, and I. J. Higginson, "A comparison of symptom prevalence in far advanced cancer, AIDS, heart disease, chronic obstructive pulmonary disease and renal disease," J. Pain Symptom Manag., vol. 31, no. 1, pp. 58-69, June 2006.

5 D. C. Currow, J. L. Plummer, A. Crockett, and A. P. Abernethy, "A community population survey of prevalence and severity of dyspnea in adults," J. Pain Symptom Manag., vol. 38, no. 4, pp. 533-545, 2009.

6 D. C. Currow, J. Smith, P. M. Davidson, P. J. Newton, M. R. Agar, and A. P. Abernethy, "Do the trajectories of dyspnea differ in prevalence and intensity by diagnosis at the end of life? A consecutive cohort study," J. Pain Symptom Manag., vol. 39, no. 4, pp. 680-690, Apr. 2010.

7 N. Karnani, G. Reisfield, and G. R. Wilson, "Evaluation of chronic dyspnea," Am.

Fam. Physician, vol. 71, no. 8, pp. 1529-1537, Apr. 2005.

8 M. Tsao, E. Barnes, and E. Chow, "The relationship between dyspnea and blood oxygen saturation," J. Pain Symptom Manag., vol. 12, no. 2, pp. 109-114, 2019.

9 K. Tanaka, T. Akechi, T. Okuyama, Y. Nishiwaki, and Y. Uchitomi, "Factors correlated with dyspnea in advanced lung cancer patients: organic causes and what else?," . J. Pain Symptom Manag., vol. 23, no. 6, pp. 490-500, June 2002.

10 T. Janssens et al., "Dyspnea perception in COPD: Association between anxiety, dyspnea-related fear, and dyspnea in a pulmonary rehabilitation program," Chest, vol. 140, no. 3, pp. 618-625, Sep. 2011.

11 E. Crisafulli and E. M. Clini, "Measures of dyspnea in pulmonary rehabilitation," Multidiscip. Respir. Med., vol. 5, no. 3, p. 202, Jun. 2010.

12 J. Bestall, E. Paul, R. Garrod, R. Garnham, P. Jones, and J. Wedzicha, "Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease," Thorax, vol. 54, no. 7, pp. 581-586, July 1999.

13 D. A. Mahler, D. H. Weinberg, C. K. Wells, and A. R. Feinstein, "The measurement of dyspnea: contents, interobserver agreement, and physiologic correlates of two new clinical indexes," Chest, vol. 85, no. 6, pp. 751-758, June 1984.

14 R. C. Boshuizen, A. D. Vincent, and M. M. van den Heuvel, "Comparison of modified Borg scale and visual analog scale dyspnea scores in predicting re-intervention after drainage of malignant pleural effusion," Supportive Care in Cancer, vol. 21, no. 11, pp. 3109-3116, July 2013.

15 G. A. Borg, "Psychophysical bases of perceived exertion," Med. Sci. Sports Exerc., vol. 14, no. 5, pp. 377-381, 1982.

16 T. Brack, A. Jubran, and M. J. Tobin, "Dyspnea and decreased variability of breathing in patients with restrictive lung disease," Am. J. Respir. Crit. Care Med., vol. 165, no. 9, pp. 1260-1264, May. 2002.

17 N. Wolkove, E. Dajczman, A. Colacone, and H. Kreisman, "The relationship between pulmonary function and dyspnea in obstructive lung disease," Chest, vol. 96, no. 6, pp. 1247-1251, Dec. 1989.

18 F. Q. AL-Khalidi, R. Saatchi, D. Burke, H. Elphick, and S. Tan, "Respiration rate monitoring methods: A review," Pediatr. Pulmonol., vol. 46, no. 6, pp. 523-529, Jan. 2011.

19 G. Brullmann, K. Fritsch, R. Thurnheer, and K. E. Bloch, "Respiratory monitoring

by inductive plethysmography in unrestrained subjects using position sensor-adjusted calibration," Respiration, vol. 79, no. 2, pp. 112-20, Apr. 2010, doi: 10.1159/000212117.

20 M. Chu et al., "Respiration rate and volume measurements using wearable strain sensors," NPJ Digit. Med., vol. 2, p. 8, Feb. 2019, doi: 10.1038/s41746-019-0083-3.

21 G. T. Ferguson, P. L. Enright, A. S. Buist, and M. W. Higgins, "Office spirometry for lung health assessment in adults: a consensus statement from the National Lung Health Education Program," Chest, vol. 117, no. 4, pp. 1146-1161, Apr. 2000.

22 M. P. Yeh, R. M. Gardner, T. D. Adams, and F. G. Yanowitz, "Computerized determination of pneumotachometer characteristics using a calibrated syringe," J. Appl. Physiol., vol. 53, no. 1, pp. 280-285, July 1982.

23 K. Bhavani-Shankar, A. Kumar, H. Moseley, and R. Ahyee-Hallsworth, "Terminology and the current limitations of time capnography: A brief review," J. Clin. Monit., vol. 11, no. 3, pp. 175-182, May 1995.

24 C. Li, V. M. Lubecke, O. Boric-Lubecke, and J. Lin, "A review on recent advances in Doppler radar sensors for noncontact healthcare monitoring," IEEE Trans. Microw. Theory. Tech., vol. 61, no. 5, pp. 2046-2060, May 2013.

25 W. Massagram, V. M. Lubecke, and O. Boric-Lubecke, "Microwave non-invasive sensing of respiratory tidal volume," in Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009, pp. 4832-4835.

26 C. Gu et al., "Accurate respiration measurement using DC-coupled continuous-wave radar sensor for motion-adaptive cancer radiotherapy," IEEE. Trans. Biomed. Eng., vol. 59, no. 11, pp. 3117-3123, June 2012.

27 P. Sharma, X. Hui, J. Zhou, T. B. Conroy, and E. C. Kan, "Wearable radio-frequency sensing of respiratory rate, respiratory volume, and heart rate," NPJ Digit. Med., vol. 3, p. 98, July 2020, doi: 10.1038/s41746-020-0307-6.

28 X. Hui and E. C. Kan, "Monitoring vital signs over multiplexed radio by near-field coherent sensing," Nat. Electron., vol. 1, no. 1, pp. 74-78, Nov. 2017, doi: 10.1038/s41928-017-0001-0.

29 X. Hui and E. C. Kan, "No-touch measurements of vital signs in small conscious animals," Sci. Adv., vol. 5, no. 2, art. eaau0169, Feb. 2019.

30 L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees. Belmont, CA: Wadsworth (International Group). 1984, pp. 151-166.

31 T. K. Ho, "The random subspace method for constructing decision forests IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 8, pp. 832-844, Aug. 1998.

32 Z. Zhang, P. Sharma, J. Zhou, X. Hui, and E. C. Kan, "Furniture-Integrated Respiration Sensors by Notched Transmission Lines," IEEE Sens. J., vol. 21, no. 4, pp. 5303-5311, Feb. 2021.

33 P. Sharma, X. Hui, and E. C. Kan, "A wearable RF sensor for monitoring respiratory patterns," in 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, July 23 – 27, 2019, pp. 1217-1223.

34 W. Lu et al., "A semi-automatic method for peak and valley detection in free-breathing respiratory waveforms," Med. Phys., vol. 33, no. 10, pp. 3634-6, Oct. 2006, doi: 10.1118/1.2348764.

35 R. R. Singh, S. Conjeti, and R. Banerjee, "A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals," Biomedical Signal Processing and Control, vol. 8, no. 6, pp. 740-754, Nov. 2013.

36 G. Vanwinckelen and H. Blockeel, "On estimating model accuracy with repeated cross-validation," in Benelearn 2012: Proc. 21st Belgian-Dutch Conf. Machine Learning, 2012, pp. 39-44.

37 W. von Rosenberg, T. Chanwimalueang, T. Adjei, U. Jaffer, V. Goverdovsky, and D. P. Mandic, "Resolving ambiguities in the LF/HF ratio: LF-HF scatter plots for the categorization of mental and physical stress from HRV," Front. Physiol., vol. 8, p. 360, June 2017.

38 K. H. Kim, S. W. Bang, and S. R. Kim, "Emotion recognition system using short-term monitoring of physiological signals," Med. Biol. Eng., vol. 42, no. 3, pp. 419-427, May 2004.

CHAPTER 5

OBJECTIVE DYSPNEA EVALUATION ON COVID-19 PATIENTS LEARNING FROM EXERTION-INDUCED DYSPNEA SCORES

Introduction

Dyspnea , also known as the patient's feeling of difficult or labored breathing, is a clinical symptom nearly as important as pain, affecting a quarter of the general population and half of seriously ill patients [1][2]. Dyspnea can be a prevalent manifestation in conditions such as chronic obstructive pulmonary diseases (COPD), bronchitis, asthma,COVID-19, pneumonia, heart failure, and panic disorders [3]. Dyspnea can be further divided into acute onset and chronic dyspnea: the latter, by definition, has been present for more than four weeks. COVID-19 caused by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has rapidly spread across the globe since 2020. Over 30% of patients with COVID have experienced chronic dyspnea [4]. Dyspnea typically sets in between the 4th and 8th day of illness. Timing of dyspnea may be one of the most important hallmarks of more significant COVID infection, especially for clinicians seeing patients in an ambulatory setting [1]. Studies found that dyspnea, rather than fever [8], was significantly associated with higher mortality in COVID patients [6][7]. The initial days after the onset of dyspnea are critical for identifying progressively worsening conditions [1]. In addition to the dyspnea experienced by many patients during the acute phase of COVID infection, dyspnea may also be found in association with the longer-term sequelae post-COVID, or so called long COVID, which is thought to affect 10–50% of COVID survivors [9].

Dyspnea is also a frequent symptom of post-COVID syndrome (PCS) [10][11]. Some patients with persistent dyspnea after recovering from COVID also have documented decrease and/or dysfunction in myocardial performance [12].

Nevertheless, in present clinical practices, dyspnea is mainly assessed by self-reports from patients. Subjective dyspnea can be assessed in person or remotely [14] using patient interview, and augmented by surrogate measures such as the Medical Research Council (MRC) Dyspnea Scale [15] and Borg Scale [16]. Studies indicated that subjective dyspnea measures have inadequate accuracy in high-risk COVID patients, not only because the sensation is gradual and varied with time, but also the patients can become nervous after encountering positive test results [17], both of which can contribute to biases in the self-report. The subjective dyspnea score can also vary for each person based on emotion and tolerance, and can be challenging to assess for those who refuse to cooperate or cannot communicate due to medical issues such as stroke, dementia, and sudden loss of speech. Frequent queries to patients for continuous dyspnea evaluation are tedious and stressful, and essentially impractical for timely prognosis and diagnosis. Shortness of breath highly correlates with pulmonary functions [18]. Surrogate measures of respiratory function can help determine dyspnea severity, however existing techniques have limitations. Pulmonary function tests can only capture respiratory measures at discrete point in time [19], and require patients' adequate effort and cooperation. Chest Computed Tomography (CT) can provide high-resolution images of the lung, however it is expensive, requiring specialized equipment that is not always available, providing only discrete lung snapshots in a dedicated clinical setup, with limited utility for continuous assessment of lung

function [20]. Respiratory inductance plethysmography (RIP) [21], strain gauge (SG) [22], and spirometer [23], can measure lung function however often can be uncomfortable, requiring connection to immobile machines, operator assistance, and patient cooperation. Thus they are not feasible tests for clinical settings, especially in conditions that require repetitive or continuous monitoring.

Considering the limitations of the current methods to monitor respiratory function and assess dyspnea, we propose to augment our previous study on the objective scoring of physiologically induced dyspnea [24], which used a non-invasive and wearable respiratory sensor and a machine learning (ML) model, to provide real-time objective dyspnea scores for COVID patients based on continuous respiratory metrics.

In this work, we collected overnight clinical data of patients admitted to the hospital due to acute COVID (N = 12), using wearable respiratory sensors from the Weill Cornell Center for Sleep Medicine. These patients had confirmed pulmonary involvement based on radiological imaging. To benchmark the results, we also performed a separate human study (N = 13) on healthy participants using exactly the same experimental setup. We analyzed the statistic distribution of respiratory metrics from COVID patients and healthy controls, and demonstrated a high similarity between dyspnea on COVID patients and the physiologically induced dyspnea on healthy subjects. The features associated with breathing changes due to physical exertion were similar to those from pulmonary disorders. By training on our previous objective dyspnea scoring model (N = 32) on healthy subjects with induced dyspnea from exertion and airway blockage [24], we can further output continuous dyspnea

scores of COVID patients using respiratory waveforms captured from wearable sensors to evaluate dyspnea severity for prognosis of dyspneic episodes, as well as rehabilitation after recovering from COVID. In the future, our respiratory sensor and objective dyspnea scoring system can be potentially applied to symptomatic evaluation of dyspnea in patients of asthma, pneumonia, and COPD.

Experimental setup and protocol

Experimental setup

To monitor the respiration in hospitalized COVID patients with confirmed lung infection by chest imaging, we built an all-in-one wearable radio-frequency (RF) near-field sensor on a 4-layer printed circuit board (PCB), as shown in Fig. 1(a). The block diagram of the near-field RF sensing system is shown in Fig. 1(b). Two SimpleLink modules (Texas Instrument CC1310) with sub-1 GHz ultra-low-power wireless microcontrollers were used as the sensing transmitter (Tx) and receiver (Rx). The transceivers employed quadrature I/Q modulation, where two channels of 12-bit I and Q samples were sampled at 2 kHz and accumulated into one 400-bytes cyclic buffer at the sensor Rx. Once the I/Q buffer was filled, the Rx radio core would bundle the I and Q samples with the readings of temperature, accelerometer, and gyroscope to a micro secure digital (SD) card through the serial peripheral bus (SPI). The temperature sensor (Texas Instrument TMP112) and inertial measurement unit (Bosch Sensortec BMI160) were connected by the inter-integrated-circuits (I2C) serial protocol. The battery provided the system power through a low-dropout (LDO) module.

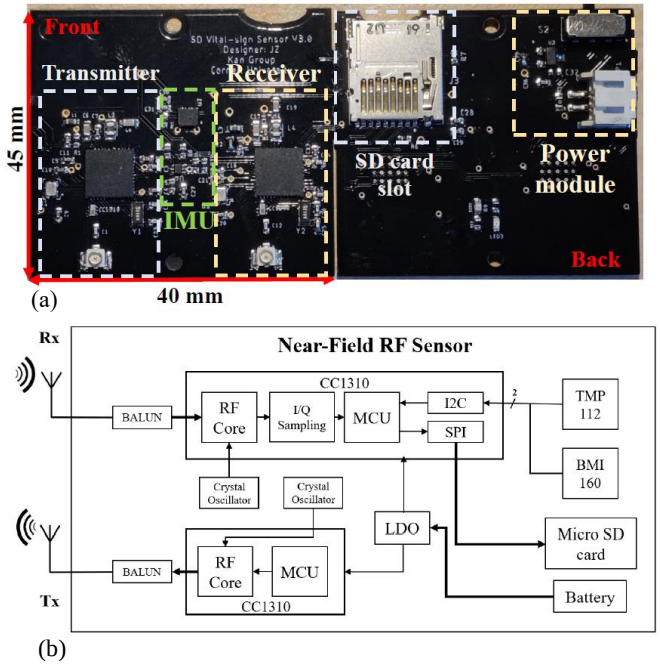


Fig. 1. The all-in-one wearable respiratory sensor on PCB: (a) The front and back photo; (b) The block diagram for the RF sensor.

Data collection from COVID patients

Respiratory data acquisition from COVID patients took place at Weill Cornell Medicine July 2021 and March 2022. Patients admitted to New York Presbyterian Hospital with COVID-19 symptoms that had lung imaging studies were offered participation in the study. All participants signed an informed consent form. The study protocol was reviewed and approved by the Weill Medical Center Institutional Review Board (IRB Protocol #: 20-06022181).

Upon enrollment, patient demographic, health status, and baseline vital-sign data were gathered. Health status information included hypertension, obstructive sleep apnea, cancer history, asthma, COPD, and other chronic lung diseases. Baseline vital-sign data included heart rates, breathing rates (BR), temperature, and oximetry SpO_2.

Other recorded information included medications administered or the presence of ventilation or supplemental oxygen during recording. After the admission information was gathered, the medical staff applied the ApneaLink device (Resmed ®) with two NCS sensing units on the patient's chest for overnight monitoring. The two sensors were synchronously powered on to begin recording. Patients wore the sensor for an average of 14.3 hours overnight.

The setup could be removed by the medical staff at any point to allow for appropriate medical care. This active medical environment created uncertainty in sensor positioning as not all sensor removals were recorded and thus information was based on the recording from the sensor data. Additionally, chest belt tension and positioning were not recorded upon each mounting and removal, so the exact sensor conditions cannot be known. However, the uncertainty in an active clinical environment also implies the resulting data analyses can be representative of typical clinical or at-home use where strict sensor and behavior restrictions cannot be enforced, and therefore should enhance the generalizability of our findings.

Healthy participant study protocol

TABLE I. ACQUISITION OF DIFFERENT DATASETS

	PARTICIPANTS	RECORDING TIME	SENSORS
COVID	12 COVID PATIENTS	CONTINUOUS 14 HOURS	PORTABLE NCS SENSORS WITH ACCELEROMETERS.
EXP. 1	13 HEALTHY SUBJECTS	NORMAL (30 MINS) POST-EXERCISE (5 MINS)	PORTABLE NCS SENSORS WITH ACCELEROMETERS.
EXP. 2 [24]	32 HEALTHY SUBJECTS	NORMAL (5 MINS) POST-EXERCISE (5 MINS)	WEARABLE NCS SOFTWARE-DEFINED RADIOS.

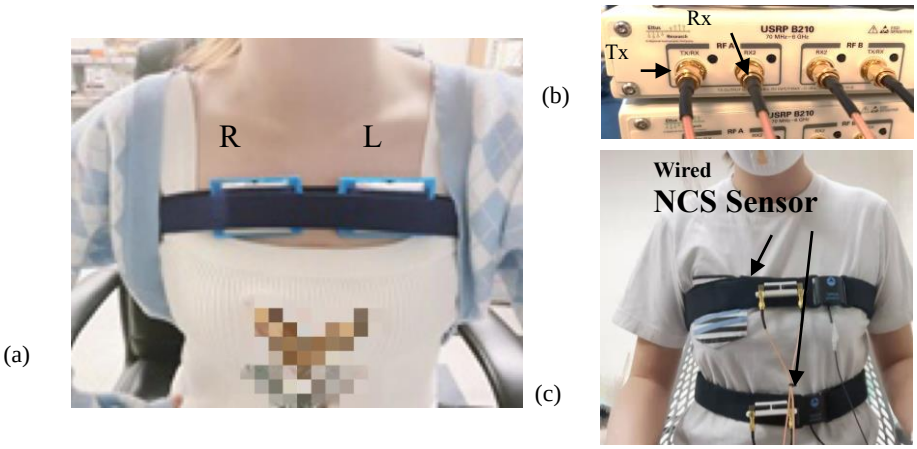


Fig. 2. Experimental setup: (a) Body deployment of two wireless NCS units by a chest belt; (b) SDR transceivers used for wired NCS sensors. (c) Wired wearable NCS sensors that were connected by cables to the sensing antennas on the chest and abdomen of a participant.

To further investigate the correlation between the dyspnea of COVID patients and physiologically-induced dyspnea on healthy subjects, we conducted another human study as indicated in Table I (Exp 1) on 13 healthy participants reporting dyspnea scores and measuring respiratory behaviors with dyspnea induced by exercise. For fair comparison without concerns on hardware difference, we used the same wearable respiratory sensors as in the COVID data collection. Fig. 2(a) shows the experimental setup with the participant wearing two sensors on left and right. The vertical position of two sensors is at the level of the sternum, roughly between the 3rd and 7th ribs. The human study has been approved by Cornell Institutional Review Board (IRB) Protocol ID #1812008488. Written informed consent to take part in the study was obtained from all participants. Participants were instructed to follow a set of routines as listed in Table I. The participant first sat on a chair in a relaxed mode for normal breathing of 30 mins. To induce dyspnea, the participant would follow a 5-min

cardio exercise video[25]. The participant would then sit back to the chair and be recorded for 5-min post-exercise breathing. The participants were asked to report subjective dyspnea scores D_{self} several times in the Borg visual analog scale (VAS), as shown in Fig. 3 [26], during the transition points of the study –. The Borg scale is widely used in clinical assessement for dyspnea: 0 represents no dyspnea sensation at all, while 10 indicates maximum level of dyspnea.

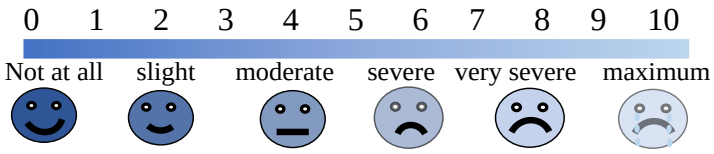


Fig. 3. Description of the self-reported Borg visual analog scale (VAS) for dyspnea evaluation.

As shown in Table I and Figs. 2(b)(c) [24], we also adopted our previous dyspnea study for comparison, denoted as data from Exp 2. In this dyspnea study, we utilized the software-defined radio (SDR, Ettus B210) to the Tx/Rx antennas as shown in Fig. 2(b). Two wired NCS sensors were placed on the chest and the abdomen in the front torso. In this human study, participants first recorded 5-min normal breathing sitting on a chair, then used aerobic exercise to introduce dyspnea, and then recorded another 5 mins of post-exercise breathing. The dyspnea was also induced through facemask to change the lung elastance. The Borg dyspnea score was reported several times throughout the routine.

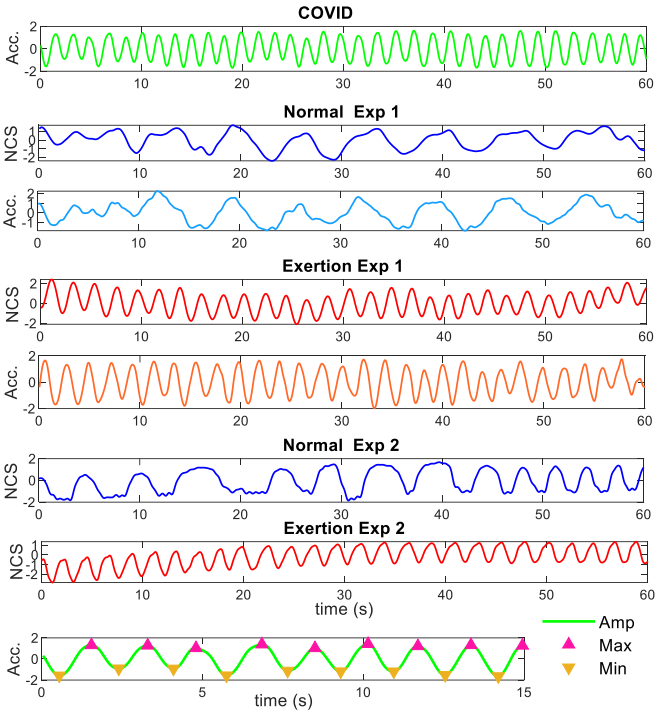


Fig. 4. Waveform examples: (a) COVID patients; (b) Healthy normal baseline breathing in Exp 1; (c) Healthy post-exertion breathing in Exp 1; (d) Healthy normal breathing in Exp 2; (e) Healthy post-exertion breathing in Exp 2; (f) Min-max peak detection for respiratory parameter extraction.

Fig. 4 presents examples of respiratory waveforms we acquired from different datasets. Y-axes are individually normalized in different channels. For the COVID dataset, we utilized the accelerometer channel for respiration monitoring. For Exp 1, we demonstrated NCS and accelerometer channels for both Normal and Exertion. For Exp 2, we demonstrated NCS recording for the same routines. Comparing normal breathing (b)(d) with exertion (c)(e) in Exp 1 and Exp 2, we can observe a consistent change in BR, with an increase in rate and a decrease in breath-to-breath variation post-exertion. We can also observe that COVID patients had higher BR than healthy participants during normal breathing, and the COVID waveforms were similar in frequency and amplitude to those acquired post-exertion in healthy participants.

Respiratory parameters were extracted from the waveforms after min-max peak detection [24] as shown in Fig. 4(f), including inter-breathing intervals (IBI), inspiration intervals (IN), expiration intervals (EX) and peak-to-peak magnitude (PP).

Data processing

Physiological analysis of dyspnea

The main purpose of respiration is to supply oxygen to body cells through the pulmonary circulation, with the auxiliary functions of producing sound, sniffing, and clearing of airway by coughing and sneezing. Respiration can be initiated involuntarily and voluntarily, and the voluntary part can be trained. When the blood oxygen saturation (SaO2) is low or CO_2 high, the breathing will be triggered to increase ventilation. However, when the body cannot respond properly due to various reasons such as airway obstruction, insufficient ambient oxygen, weakened respiratory muscles, cardiopulmonary disorders, or voluntary control for speaking, singing or breath holding, the sensation of dyspnea arises. To increase lung ventilation, often BR and lung volume (LV) will increase by panting or deep breathing. Alternatively, the inhalation and exhalation intervals will be adjusted depending on the muscle condition, airway obstruction, and ambient factors. As the respiratory reaction to dyspnea can be trained to reduce the uncomfortable feeling, similar to experiential avoidance for coping with pain, another common physiological reaction to dyspnea is the reduction of variability in successive breaths together with speaking restraint, when the body tries hard to use the best known breathing cycle to reduce the discomfort from dyspnea.

One major symptom for COVID patients is the dyspnea, which relies primarily on self-report sensation at the present practice. The dyspnea sensation is often derived from the decreased ventilation efficiency caused by pneumonia or related bronchitis. A distinct phenotype in long COVID is that patients have reduced exercise tolerance and experience exertional dyspnea more easily, even though major pulmonary parenchymal and airway abnormalities cannot be identified with chest imaging [27][28].

Data preprocessing

After gathering overnight recording of COVID patients and performing comparison dyspnea study on healthy participants, we pre-processed our datasets, and then extracted the respiratory features to feed into the ML algorithms for dyspnea classification and scoring. We used MATLAB for signal processing and feature extraction, and python for ML algorithms.

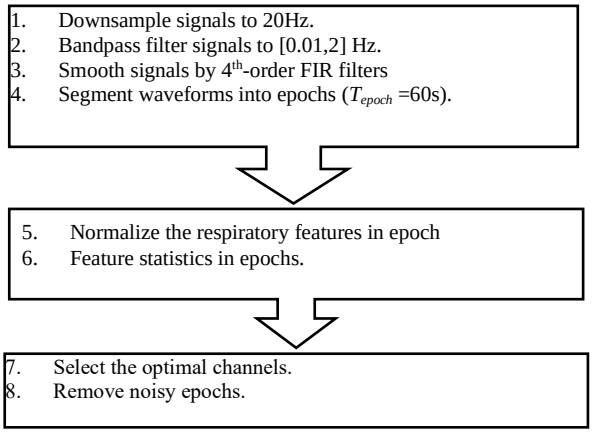


Fig. 5. Processing procedures of respiratory datasets in COVID patients and healthy participants.

For datasets in COVID and Exp 1 described in Table I, we retrieved respiratory waveforms from 1 NCS (amplitude) and 6 accelerometer (translational and rotational) channels. For Exp 2, we utilized multiple NCS channels from thorax and abdomen. Different channels and different datasets went through the same signal processing procedure for fair comparison. As shown in Fig. 5, we first down-sampled all datasets to 20 Hz, and then bandpass-filtered waveforms from 0.05 Hz to 2 Hz to remove the DC drift and high-frequency noises. Savitzky-Golay 4^{th}-order finite impulse response (FIR) smoothing filter [29] was further employed to perform a local polynomial regression to smooth the waveform. For the long recording in each dataset, we opted to segment waveforms into short epochs of T_{epoch} = 60 s and a sliding window of T_{slide} = 30 s for feature extraction.

Respiratory features

After epoch segmentation, waveforms were normalized to center at 0 with standard deviation of 1.0 in each epoch for every channel. Then we extracted features in each epoch for data analysis and constructed the ML model in the next section. We implemented the peak detection algorithm [30] by tracing a constantly updated moving-average curve in a given window. Then local maximum and minimum were accordingly labelled for parameter extraction. An example was shown in Fig. 4(f), where the green line was the filtered respiratory waveform from the COVID patient, and the red and yellow triangles marked the maximum and minimum peaks detected by the moving-average algorithm. The false peaks caused by the noise were mostly avoided. Respiratory parameters in each breath cycle were extracted to represent the

instantaneous respiratory characteristics, as shown in Table II.

After gathering respiratory cycles and parameters, we extracted 37 respiratory features as shown in Table III. The first three features were: 1) mean (μ); 2) standard deviation (σ); 3) coefficient of variation (*CoV*) of the respiratory parameters in Table II, where *CoV* was defined as

$$CoV = \left(\frac{\sigma}{\mu}\right)^2 \quad (1)$$

showing the extent of variability in relation to the mean.

To further capture variability between adjacent breaths, $\Re$ was the autocorrelation in a time lag of one respiratory cycle to measure the successive similarity of the given respiratory parameter. ς representing the successive difference was defined as the mean absolute difference between adjacent cycles. Additionally, *Skew* and *kurt* measured the tailedness and asymmetry of each respiratory cycle, and were averaged over all cycles within the epoch. *Cycle* denoted the total number of detected respiratory cycles in the epoch, and *entropy* denotes the total randomness or entropy of the waveform.

TABLE II. INSTANTANEOUS RESPIRATORY PARAMETERS (7)

EXTRACTED PARAMETERS	DESCRIPTION
BREATH RATE (*BR*)	INVERSE OF THE INTERVAL BETWEEN TWO NEIGHBORING MINIMA.
PEAK-TO-PEAK (*PP*)	LUNG VOLUME REPRESENTED BY SIGNAL DIFFERENCE IN SUCCESSIVE PEAKS.
INHALATION INTERVAL (*IN*)	TIME DIFFERENCE BETWEEN ONE MINIMUM AND THE FOLLOWING MAXIMUM.
EXHALATION INTERVAL (*EX*)	TIME DIFFERENCE BETWEEN ONE MAXIMUM AND THE FOLLOWING MINIMUM.
INTER-BREATH INTERVAL (*IBI*)	INTERVAL BETWEEN TWO NEIGHBORING MAXIMA.
IN- EX RATIO (*IER*)	INHALATION/EXHALATION INTERVAL RATIO.
IN- EX VOLUME RATIO (*IEPP*)	INHALATION/EXHALATION VOLUME RATIO.

TABLE III. RESPIRATORY FEATURES (37)

μ_{BR}	μ_{PP}	μ_{IN}	μ_{EX}	μ_{IBI}	μ_{IER}	μ_{IEPP}
Σ_{BR}	Σ_{PP}	Σ_{IN}	Σ_{EX}	Σ_{IBI}	Σ_{IER}	Σ_{IEPP}
CoV_{BR}	CoV_{PP}	CoV_{IN}	CoV_{EX}	CoV_{IBI}		
$\Re_{BR}$	$\Re_{PP}$	$\Re_{IN}$	$\Re_{EX}$	$\Re_{IBI}$	$\Re_{IER}$	$\Re_{IEPP}$
Σ_{BR}	Σ_{PP}	Σ_{IN}	Σ_{EX}	Σ_{IBI}	Σ_{IER}	Σ_{IEPP}
μ_{SKEW}	μ_{KURT}	*ENTROPY*	*CYCLE*			

TABLE IV. FREQUENCY FEATURES (14).

η_{F1}	η_{F2}	η_{F3}	η_{F4}	η_{F5}
p_{F1}	p_{F2}	p_{F3}	p_{F4}	p_{F5}
F_{BR}	F_{HR}	SNR_{BR}	SNR_{HR}	

Apart from 37 respiratory features extracted from the time domain, we added 14 features extracted from the frequency domain as shown in Table IV. η_{fi} and p_{f1} (i = 1~5) represented the power in specific bandwidth divided by the total power in all frequencies and time-averaged power density (dB/Hz), respectively. The five chosen bandwidths were f_1 = (0, 0.4) Hz (mainly breathing frequency range); f_2 = (0.4, 1) Hz; f_3 = (1, 2) Hz; f_4 = (f_{BR} – 0.15, f_{BR} + 0.15) Hz; f_5 = (f_{HR} – 0.15, f_{HR} + 0.15) Hz. f_{BR} and f_{HR}

were first estimated from the average BR and heart rate (HR) provided by hospital reports for every patient, and then further refined to be the local BR and HR in every epoch by finding the maximal energy in the possible frequency band. Signal-to-noise ratios (SNR) in BR and HR were denoted by SNR_{BR} and SNR_{HR} which were calculated by the maximal energy on the f_{BR} and f_{HR} divided by the estimated noise power.

Channel and epoch selection

After segmentation and feature extraction, we selected the optimal channel and epoch from the datasets according to the estimated signal quality. For the accelerometer, we had 6 channels consisting of X, Y, and Z translational (acc) and rotational (gyro) motions, as shown in Fig. 6 for an example. Feature extraction was performed on every channel, and the optimal channel was decided by the least variation of respiration parameters. We can observe from the waveforms that most channels can get similar BR =35, but the channels with smaller σ_{BR} and σ_{PP}, such as 'gyro X' and 'gyro Y', have more stable respiratory waveforms. Therefore, we opted to use the optimal channel by the minimum mean of all covariation features in BR, PP, IN, EX and IBI. In Fig. 6, the optimal channel is 'gyro Y'.

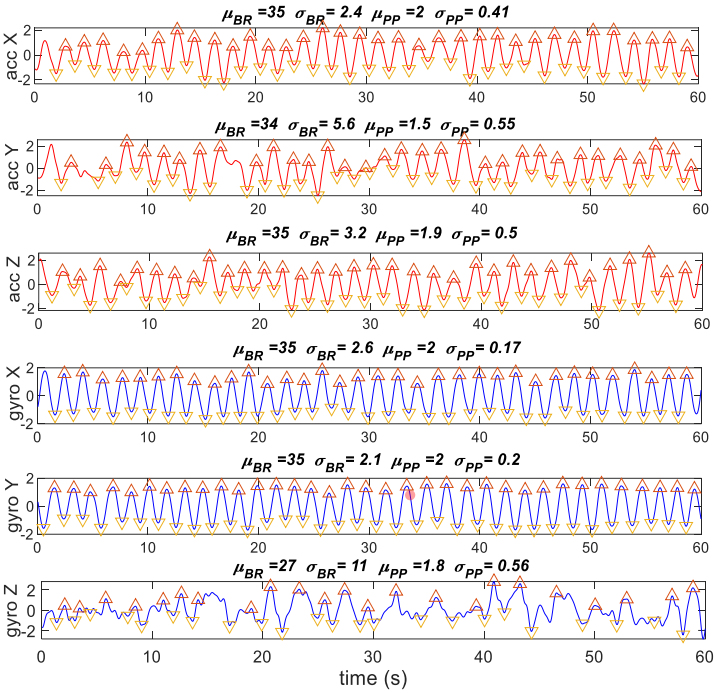


Fig. 6. An example of channel selection for the accelerometer. The optimal channel is 'gyro Y'.

Signal quality cannot be guaranteed during the entire course of overnight recording because patients may have random motion lying on the bed or leave the bed for restroom visits. Various factors such as ambient movement might bring about noises to cause SNR degradation. The position of the wearable sensor to the patient clothing might sometimes move during long or deep breathing, and brought further noise to the signal. Therefore, we opted to remove the epochs with low SNR by pre-determined thresholds. We selected the threshold to be mean of all covariation features to be smaller than 0.4. Table V provides the selection ratio for every dataset and the final cases we have collected after all the signal processing procedures. The datasets from Exp 1 and Exp 2 have higher quality because of the better controlled lab environment during data collection.

TABLE V. STATISTIC COMPARISON OF COVID AND DIFFERENT DATASETS

	COVID ACC.	NORM. NCS EXP 1	EXER. NCS EXP 1	NORM. ACC. EXP 1	EXER. ACC. EXP 1	NORM. NCS EXP 2	EXER. NCS EXP 2
CASES	10131	1049	188	918	231	256	240
RATIO (%)	30.2	74.0	77.7	64.7	95.5	100	100

Results

Demographic information

TABLE VI. DEMOGRAPHICS IN HUMAN STUDY

DATASETS	GENDER	NUMBER	BMI ($\mu \pm \Sigma$)	AGE ($\mu \pm \Sigma$)
COVID	MALE	8	30.1 ± 7.3	-
	FEMALE	4	27.8 ± 6.3	-
EXP 1	MALE	7	22.6 ± 2.5	29 ± 12
	FEMALE	6	21.4 ± 3.7	21 ± 2
EXP 2	MALE	14	23.3 ± 2.5	28 ± 9
	FEMALE	18	20 ± 1.3	24 ± 2

This study involved three distinct datasets: COVID patients, and two dyspnea human studies on healthy subjects. The study population information of each dataset is shown in Table VI. Age information was not gathered for the COVID dataset.

Feature Analysis and Comparison

After acquiring all datasets, we first evaluated the similarity of respiratory features between COVID patients and healthy subjects where dyspnea was physiologically induced by exercise. The changes of breathing features in exertion-induced dyspnea and acute short-term dyspnea from pulmonary disorders can have

correlation with important implications. We first examined a few representative respiratory features and presented the scatter plots in Fig. 7 from 3 datasets: 1) COVID patients (accelerometer); 2) Healthy subjects during normal breathing in Exp. 1 (NCS); 3. Healthy subjects breathing after exertion in Exp. 1 (NCS). Respiratory features collected from NCS and accelerometer in our human study have high similarity, so the difference using two different sources were mainly determined by SNR considerations. In Fig. 7(a), the X and Y axes represented respiratory features $\Re_{BR}$ and $\Re_{PP}$, while in 6(b) represented ς_{IBI} and $\Re_{IBI}$. To better compare the feature distribution for different datasets, we used the Gaussian kernel smoothing function to estimate the returned probability density in top and right lines. The dataset from healthy subjects during normal breathing had a much broader range of distribution compared with the other two datasets. For a better visual demonstration, we set the X-Y limits to only show all of points from dataset 1 and 3, and some points from dataset 2 were out of range. We can observe that COVID patients had a higher similarity of breathing features to healthy subjects after exertion. In Fig. 7(a), for $\Re_{BR}$ and $\Re_{PP}$, and in Fig. 7(b), for $\Re_{IBI}$, both COVID patients and healthy exertional breathing had higher values closer to 1, indicating higher autocorrelation of neighboring breathing cycles. In Fig. 6(b), for $\varsigma_{IBI,}$, COVID patients and healthy exertional breathing were concentrated on smaller successive differences, while healthy normal breathing had broader spread to higher variation.

For a more comprehensive comparison of similarity in different datasets, we calculated the Kullback–Leibler (KL) divergence between the dataset of COVID patients and the other datasets in Table VII. The KL divergence, also called the

relative entropy, is a type of statistical distance between two probability distributions [31]. We first transformed our discrete datasets to smoothed continuous Gaussian distributions just like the top and right lines in Fig. 7, where the KL divergence represents a natural dissimilarity by the mixture of Gaussians. In Table VII, we presented 8 representative respiratory feature statistics. In both Exp 1 and Exp 2 with healthy subjects, the normal breathing features had larger KL divergence to those of COVID patients than the exertional breathing features. In Table VIII, we also examined the dissimilarity between NCS and accelerometer in the same experiment of Exp 1. The small KL divergence between the two measurements showed the similarity and interchangeability for respiration measurements. Therefore, in the following sections, we will mainly use NCS datasets in Exp 1 for calculation.

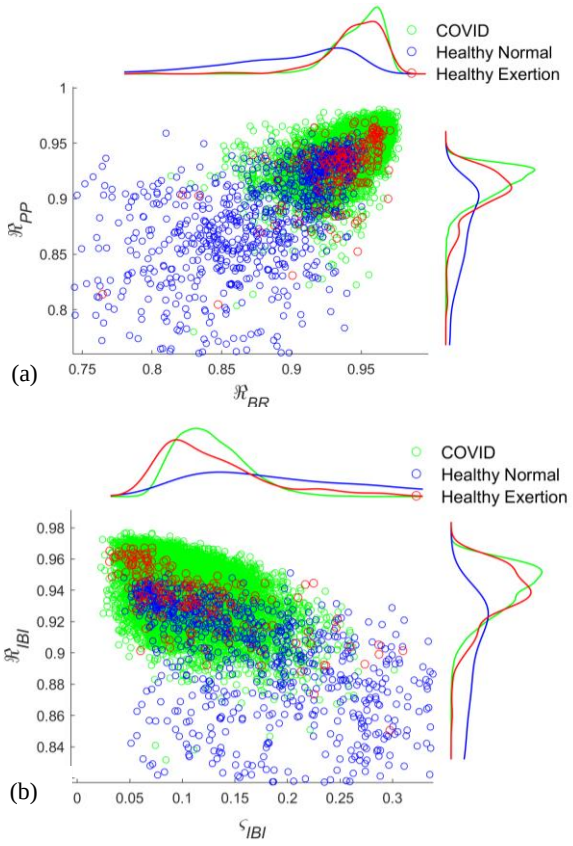


Fig. 7. Scatter plots of chosen respiratory features from COVID and human study datasets. Top and right lines are smoothed continuous distribution by Gaussian kernels. (a): ℜBR and ℜPP; (b): ς IBI and ℜIBI.

TABLE VIII. KL-DIVERGENCE OF NCS AND ACCELEROMETERS IN EXP1.

	μ_{BR}	Σ_{BR}	CoV_{BR}	CoV_{IBI}	
NORM.	0.05	0.09	0.12	0.13	
EXER.	0.01	0.13	0.11	0.06	
	$\Re_{BR}$	$\Re_{PP}$	Σ_{IBI}	Σ_{IER}	AVG.
NORM.	0.04	0.05	0.12	0.21	0.10
EXER.	0.08	0.04	0.22	0.08	0.09

TABLE VII. KL-DIVERGENCE OF COVID TO OTHER DATASETS

	NORM. NCS EXP 1	EXER. NCS EXP 1	NORM. ACC EXP 1	EXER. ACC EXP 1	NORM. NCS EXP 2	EXER. NCS EXP 2
μ_{BR}	2.14	0.17	2.62	0.16	1.44	0.30
Σ_{BR}	1.71	0.69	0.68	0.72	0.49	0.91
CoV_{BR}	3.91	0.76	2.75	0.65	1.17	1.03
CoV_{IBI}	4.79	0.99	3.37	1.08	1.66	1.33
$\Re_{BR}$	3.42	0.16	2.97	0.16	1.52	0.50
$\Re_{PP}$	2.96	0.21	2.34	0.08	0.98	0.21
Σ_{IBI}	3.87	0.44	2.61	0.68	1.32	0.73
Σ_{IER}	1.82	0.26	1.34	0.16	0.43	0.17
AVG	3.08	0.46	2.34	0.46	1.13	0.65

Dyspnea Classification Model

After comparing individual respiratory features, we can find high similarity between pneumonia-induced dyspnea in COVID patients and exertional dyspnea in healthy subjects. We now adopted our previous dyspnea model derived from Exp 2

[24] on 32 healthy subjects as the training model to evaluate the COVID patients and the healthy subjects with the same sensor setup in Exp 1. By utilizing the ML algorithm in [24], we can evaluate the dyspnea score from all respiratory features as a whole and validate the feasibility of our objective dyspnea reporting system in the clinical setting. As it is impractical to ask patients constantly to self report their dyspnea scores, this objective dyspnea report can be of high value to provide a gauge of dyspnea in COVID patients continuously, especially during inconvenient periods such as patients sleeping or going through treatment.

The first model we built was the binary classification model, namely normal = 0 and dyspnea = 1. For the training dataset we adopted from our previous dyspnea study on healthy subjects, all normal breathing epochs were labelled as normal = 0, while all exertional breathing epochs as dyspnea = 1. By training on the dataset to build a dyspnea classifier, we can output the dyspnea classification results for all COVID datasets and the reference cases from Exp 1. We utilized k-nearest neighbor classifier as the model here with neighbor number k = 40. Before feeding the dataset into the model, all features were preprocessed with a standard scaler to regularize features by removing the mean and scaling to unit variance. As Table IX showed, almost all of COVID patients' cases were classified into dyspnea, while 73.6% of exertional breathing cases of healthy subjects were classified into dyspnea. In comparison, only 4.24% of the normal breathing cases in healthy subjects were classified into dyspnea. The dyspnea classification results for COVID patients further corroborated our hypothesis that COVID dyspnea had high similarity to exertional dyspnea in healthy subjects.

TABLE IX. CLASSIFICATION RESULTS OF DYSPNEA FOR COVID PATIENTS AND HEALTHY SUBJECTS (IN EXP 1).

	COVID	HEALTHY NORMAL	HEALTHY EXERTION
PERCENTAGE OF DYSPNEA	98.05 %	4.24 %	73.63 %

Dyspnea Scoring Model

In this section, we built a regressor model for objective dyspnea scoring in the Borg scale (0 - 10) D_{obj} for COVID patients. In our previous work, we built a similar dyspnea scoring system and achieved high accuracy for generating D_{obj} for exertional dyspnea on healthy subjects in comparison with self report. In this work, we used our previous dyspnea study as the training model to build the scoring system, and treated the COVID patients as testing cases. We implemented the k-nearest neighbor regressor as the main model here. Since we had overnight recording for COVID patients, we first reported the epoch-wise dyspnea scores for all datasets, and then we averaged all dyspnea scores as the final score for every patient.

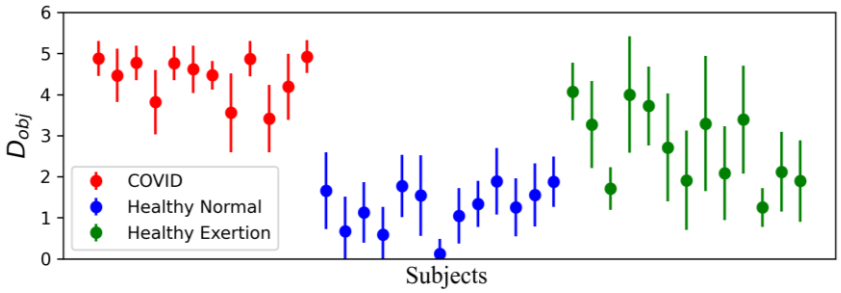


Fig. 8. Dyspnea scoring results for COVID patients and healthy subjects (Exp 1). The average of Dobj : COVID = 4.39; Healthy Normal = 1.26; Healthy Exertion = 2.72.

TABLE X. T-TEST STATISTICS FOR DYSPNEA SCORING RESULTS

	COVID VS. HEALTHY NORMAL	COVID VS. HEALTHY EXERTION	HEALTHY NORMAL VS. HEALTHY EXERTION
T-STATISTIC	14.60	5.47	–4.82
P-VALUE	4.61×10^{-13}	2.75×10^{-5}	1.1×10^{-4}

Fig. 8 presents the results for dyspnea scoring of COVID patients using exertion induced dyspnea on healthy subjects in Exp 1. Results for benchmark Exp 2 on healthy subjects are also shown. For average D_{obj} reported from different datasets, COVID = 4.39; Healthy Normal = 1.26; Healthy Exertion = 2.72. As observed from Fig. 8, D_{obj} for COVID patients were more concentrated around 4-5, while normal breathing for healthy participants were mostly below 2. Exertional breathing for healthy participants had more participants with higher dyspnea score, but subject variation was also evident, possibly because different subjects had variation in physical conditioning after the same cardio exercise. D_{obj} for COVID patients was less dispersed probably due to the more uniform manifestations of the underlying pulmonary disease.

We further preformed T-tests for dyspnea scoring results on different datasets as shown in Table X. The calculated T-statistic is positive when the sample mean of the first dataset is greater than the second dataset, and negative otherwise. As the T-statistic showed, the dyspnea scores for COVID patients were distinctively higher than normal breathing in healthy subjects, while differences with exertion breathing were smaller. The very small p-value between COVID patients and normal breathing indicated that they had distinctively different distributions for dyspnea scores. For the dataset of healthy exertion, the p-values to COVID and healthy normal were also

sufficiently small to suggest high distinguishability among the datasets.

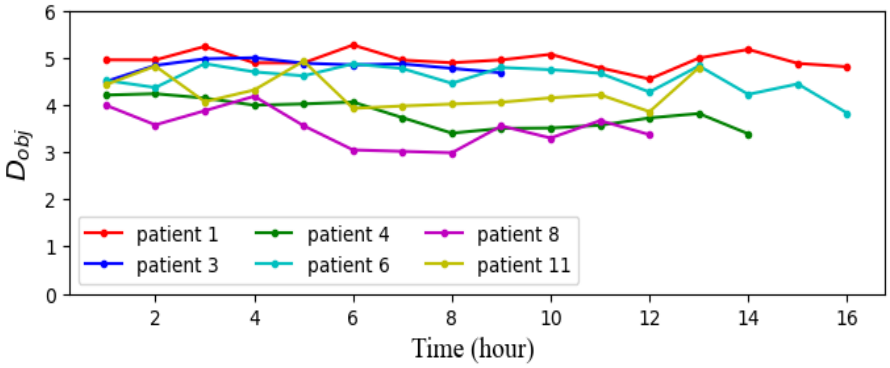


Fig. 9. An example of continuous monitoring of objective dyspnea scores for COVID patients for every hour.

Apart from reporting average D_{obj} to give a general objective evaluation, our system can also output continuous real-time D_{obj} to give an indication to the infection progression. COVID and other important pulmonary diseases like COPD are frequently accompanied by dyspnea sensation from reduced lung function. It is thus critical to continuously monitor patients because the dyspnea often develops insidiously over a period of time. Frequent self report is inconvenient and less accurate for long-term tracing of the symptoms. Fig. 9 shows an example of continuous monitoring of dyspnea score for 6 COVID patients. In the whole recoding of 12-16 hours, D_{obj} was reported every hour, to align with the clinical data recorded every hour.

Conclusion

Dyspnea is a key symptom for patients with COVID-19 and many other respiratory disorders. Existing clinical evaluation of dyspnea currently depends on self-report, which is subjective and challenging for continuous monitoring. In this

work, we used an innovative approach to continuously monitor respiratory features using wireless and wearable respiratory sensors to develop an objective dyspnea scoring system derived from exertion routines on healthy subjects. We then tested this model without further learning on COVID patients and control subjects under the same sensor setup. We found high similarity between pneumonia-induced dyspnea of COVID patients and physiologically induced dyspnea on healthy subjects, suggesting that changes of respiratory features from physical exertions could be representative of the dyspnea found in pulmonary disorders. We also demonstrated the unique capability to continuously report objective dyspnea scores during 16 hours for COVID patients. Our system can be a promising tool for diagnosis and prognosis of COVID, offering warning of possible worsening dyspnea and respiratory function, as well as the degree of recovery. This work validates the feasibility of our objective dyspnea scoring for clinical dyspnea assessment, and can be applied to symptomatic evaluation of dyspnea in patients with similar conditions including asthma, pneumonia [32], and COPD [33].

REFERENCES

1 R. W. Lansing, R. H. Gracely, and R. B. Banzett, "The multiple dimensions of dyspnea: Review and hypotheses," *Respir. Physiol. Neurobiol.* , vol. 167, no. 1, pp. 53-60, 2009.

2 D. A. Mahler, D. H. Weinberg, C. K. Wells, and A. R. Feinstein, "The measurement of dyspnea: Contents, interobserver agreement, and physiologic correlates of two new clinical indexes," *Chest*, vol. 85, no. 6, pp. 751-758, June 1984.

3 D. Berliner, N. Schneider, T. Welte, and J. Bauersachs, "The differential diagnosis of dyspnea," *Dtsch. Arztebl. Int.* , vol. 113, no. 49, p. 834, 2016.

4 R. da Rosa Mesquita *et al.*, "Clinical manifestations of COVID-19 in the general population: Systematic review," *Wien. Klin. Wochenschr.* , vol. 133, no. 7, pp. 377-382, 2021.

5 P. A. Cohen, L. E. Hall, J. N. John, and A. B. Rapoport, "The early natural history of SARS-CoV-2 infection: Clinical observations from an urban, ambulatory COVID-19 clinic," in *Mayo Clin. Proc.* , 2020, vol. 95, no. 6, pp. 1124-1126: Elsevier.

6 Z. Zheng *et al.*, "Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis," *J. Infect.*, vol. 81, no. 2, pp. e16-e25, 2020.

7 F. Zhou *et al.*, "Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study," *Lancet*, vol. 395, no. 10229, pp. 1054-1062, 2020.

8 L. Shi, Y. Wang, Y. Wang, G. Duan, and H. Yang, "Dyspnea rather than fever is a risk factor for predicting mortality in patients with COVID-19," *J. Infect.*, vol. 81, no. 4, pp. 647-679, 2020.

9 G. Y. Lam *et al.*, "Exertional intolerance and dyspnea with preserved lung function: an emerging long COVID phenotype?," *Respir. Res.*, vol. 22, no. 1, pp. 1-4, 2021.

10 A. U. Trillig, M. Ljuslin, J. Mercier, M. Harrisson, and P. Vayne-Bossert, "“I am not the same man…”: A case report of management of post-COVID refractory dyspnea," *J .Palliat. Med.*, doi: 10.1089/jpm.2021.0605, Mar. 2022.

11 K. J. Wirth and C. Scheibenbogen, "Dyspnea in post-COVID syndrome following mild acute COVID-19 infections: Potential causes and consequences for a

therapeutic approach," *Medicina,* vol. 58, no. 3, p. 419, 2022.

12 M.-L. Luchian *et al.*, "Subclinical myocardial dysfunction in patients with persistent dyspnea one year after COVID-19," *Diagnostics,* vol. 12, no. 1, p. 57, 2021.

13 C. Curci *et al.*, "Early rehabilitation in post-acute COVID-19 patients: Data from an Italian COVID-19 Rehabilitation Unit and proposal of a treatment protocol," *Eur. J Phys. Rehabil. Med.*, pp. 633-641, 2020.

14 T. Greenhalgh, G. C. H. Koh, and J. Car, "Covid-19: A remote assessment in primary care," *BMJ,* vol. 368, art. m1182, 2020.

15 J. Bestall, E. Paul, R. Garrod, R. Garnham, P. Jones, and J. Wedzicha, "Usefulness of the Medical Research Council (MRC) dyspnea scale as a measure of disability in patients with chronic obstructive pulmonary disease," *Thorax*, vol. 54, no. 7, pp. 581-586, July 1999.

16 R. C. Boshuizen, A. D. Vincent, and M. M. van den Heuvel, "Comparison of modified Borg scale and visual analog scale dyspnea scores in predicting re-intervention after drainage of malignant pleural effusion," *Support. Care Cancer*, vol. 21, no. 11, pp. 3109-3116, July 2013.

17 L. Berezin *et al.*, "The diagnostic accuracy of subjective dyspnea in detecting hypoxemia among outpatients with COVID-19," *medRxiv,* 2020.

18 M. Balbi *et al.*, "Post-discharge chest CT findings and pulmonary function tests in severe COVID-19 patients," *Eur. J Radiol.*, vol. 138, p. 109676, 2021.

19 J. You *et al.*, "Abnormal pulmonary function and residual CT abnormalities in rehabilitating COVID-19 patients after discharge," *J. Infect.,* vol. 81, no. 2, pp. e150-e152, 2020.

20 M. D. Hope, C. A. Raptis, A. Shah, M. M. Hammer, and T. S. Henry, "A role for CT in COVID-19? What data really tell us so far," *Lancet,* vol. 395, no. 10231, pp. 1189-1190, 2020.

21 G. Brullmann, K. Fritsch, R. Thurnheer, and K. E. Bloch, "Respiratory monitoring by inductive plethysmography in unrestrained subjects using position sensor-adjusted calibration," *Respiration*, vol. 79, no. 2, pp. 112-20, Apr. 2010, doi: 10.1159/000212117.

22 M. Chu et al., "Respiration rate and volume measurements using wearable strain sensors," *NPJ Digit. Med.*, vol. 2, p. 8, Feb. 2019, doi: 10.1038/s41746-019-0083-3.

23 G. T. Ferguson, P. L. Enright, A. S. Buist, and M. W. Higgins, "Office spirometry

for lung health assessment in adults: A consensus statement from the National Lung Health Education Program," *Chest*, vol. 117, no. 4, pp. 1146-1161, Apr. 2000.

24 Z. Zhang, P. Sharma, T. B. Conroy, V. Phongtankuel, and E. C. Kan, "Objective scoring of physiologically induced dyspnea by non-invasive RF sensors," *IEEE. Trans. Biomed. Engr.*, vol. 69, no. 1, pp. 432-442, 2021.

25 BowflexStories, “Bowflex® bodyweight workout | Five-minute cardio blast workout,” YouTube, 21-Oct-2015. [Online]. Available: https://www.youtube.com/watch?v=BR0jT6JxH-o. [Accessed: 06-Sep-2022].

26 R. C. Boshuizen, A. D. Vincent, and M. M. van den Heuvel, "Comparison of modified Borg scale and visual analog scale dyspnea scores in predicting re-intervention after drainage of malignant pleural effusion," *Support Care Cancer.*, vol. 21, no. 11, pp. 3109-3116, July 2013.

27 J. Spiesshoefer *et al.*, "Diaphragm dysfunction as a potential determinant of dyspnea on exertion in patients 1 year after COVID-19-related ARDS," *Respir. Res.*, vol. 23, no. 1, pp. 1-6, 2022.

28 R. Núñez-Cortés, G. Rivera-Lillo, M. Arias-Campoverde, D. Soto-García, R. García-Palomera, and R. Torres-Castro, "Use of sit-to-stand test to assess the physical capacity and exertional desaturation in patients post COVID-19," *Chronic Respir. Dis.* , vol. 18, p. 9205, 2021.

29 R. W. Schafer, "What is a Savitzky-Golay filter? [lecture notes]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 111-117, July 2011.

30 W. Lu et al., "A semi-automatic method for peak and valley detection in free-breathing respiratory waveforms," *Med. Phys.*, vol. 33, no. 10, pp. 3634-6, Oct. 2006, doi: 10.1118/1.2348764.

31 S. Kullback, "The Kullback-Leibler distance," *Am. Stat.* , vol. 41, no. 4, pp. 340 – 341, 1987.

32 M. N. Lutfiyya, E. Henley, L. F. Chang, and S. W. Reyburn, "Diagnosis and treatment of community-acquired pneumonia," *Am. Fam. Physician* , vol. 73, no. 3, pp. 442-450, 2006.

33 A. Anzueto and M. Miravitlles, "Pathophysiology of dyspnea in COPD," *Postgrad. Med.*, vol. 129, no. 3, pp. 366-374, 2017.

CHAPTER 6

NOVEL MUSCLE SENSING BY RADIOMYOGRAPHY (RMG) AND ITS APPLICATION TO HAND GESTURE RECOGNITION

Introduction

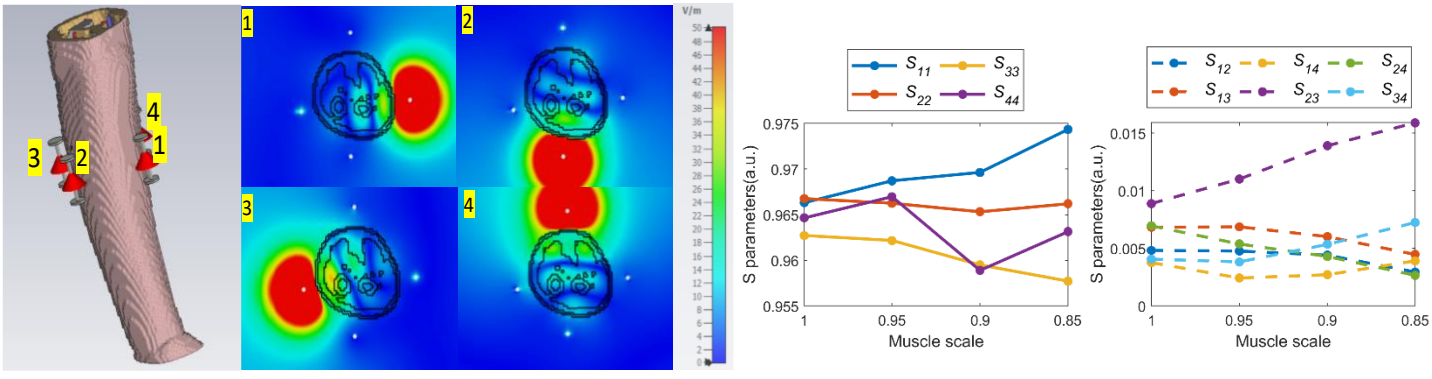
Muscle Monitoring

Monitoring skeletal muscle activities has significant medical and commercial applications, including detection of muscle fatigue and injury, diagnosis of neuromuscular disorders, assessment for physical training and rehabilitation [1]-[3], human-computer interface (HCI) [4], and robotic control [5]. Electromyography (EMG) is presently the prevalent continuous muscle sensing technique, which employs intramuscular needle electrodes or epidermal electrodes on the bare skin [6] to record the neural signals during muscle contraction, though the electrical pathways and neuronal depolarization can depend on the muscle tissue condition. Mechanomyography (MMG) and accelerometers record the mechanical motion on the body surface, but lacks information in deep muscle layers [7]. Ultrasound monitoring requires body probes with surface preparation for impedance match [8]. Magnetic resonance imaging (MRI) and computer tomography (CT)-scan can obtain high-resolution muscle imaging, but is expensive and immobile, providing only short snapshots of muscle motion in a dedicated clinical setup [9].

Hence, a direct muscle activity sensor that can accurately and continuously track muscle contraction in the superficial and deep layers with high user comfort is a critical complement to EMG in many biomedical and HCI applications.

Hand gesture recognition by muscle actuation detection

Hand gestures are controlled by complex muscle groups, many of which are extended from the forearms. Wearable sensors on wrist or arm bands for hand gesture recognition (HGR) [10][11] is of high interest to facilitate HCI [12][13] and a myriad of other applications [14][15]. Current HGR techniques however have many limitations. Vision-based system requires off-body line-of-sight (LoS) cameras, and is vulnerable to self and ambient obstruction [16]-[18]. Depth-perception methods demand excessive geometry reconstruction computation [19][20]. Gloved-based sensors can hinder hand motion, and are often inconvenient and uncomfortable [21]. Accelerometers can only transduce surface motion, and are impractical and cumbersome if deployed on fingers or phalanges [22]. The HGR radar, such as Google Soli [23], has to be in the LoS path to the target hand and can also suffer from path noises [24][25].



(a) (b) (c) (d)

Fig. 1. Demonstration of the near-field coupling principle in RMG by electromagnetic simulation. (a) The human forearm phantom in CST Bio Extension 4.0 library with 4 dipole antennas around the arm circumference. (b) Electric field distribution in the cross section of forearm with excitation by antennas 1 – 4. (c) The normalized scattering (S) parameters of the four antennas self channels for different muscle scales, and (d) S parameters of the cross channels.

By monitoring muscle activities on the forearm, an HGR system can be built with user convenience for many gestures, such as the conventional surface electromyography (sEMG) tracking the neural stimulation of forearm muscles [26]-[30]. However, sEMG is limited by the ambiguity in the skin potential, vital-sign interference, and requirement of numerous electrodes with high-quality contact to the bare skin, sometime raising health concerns in long-term wearing and suffering low user acceptance [31][32]. Electrical impedance tomography (EIT) recovers the interior impedance geometry of arm muscles, but resolution is limited and cross-user generalization is often questionable [33].

RMG for muscle activity sensing

In this work, we propose radiomyography (RMG), a novel skeletal muscle sensor that can non-invasively and continuously capture muscle contraction in various layers. RMG uses multiple-input-multiple-output (MIMO) near-field coherent sensing (NCS) radio signals [34][35] to measure the dielectric property change and dielectric boundary movement of nearby muscle groups. NCS couples ultra-high frequency (UHF) electromagnetic (EM) waves inside the body and reads out the internal organ and tissue motion signals as modulated antenna characteristics or scattering matrices [34]. As the near-field coupling is nonlinear and convoluted in the capture volume from different observation points, we explored spatially diverse channels to distinguish the detailed muscle action. MIMO provides N2 observation channels in 3D from N sensing units on or above the body surface to enhance this critical spatial diversity [35].

Radio-frequency (RF) signals in the UHF band, especially in the near-field region, will penetrate most dielectrics effectively without requiring direct skin contact. Therefore, our RMG system can be wearable over clothing or installed in a nearby off-body apparatus such as armrests and wrist pads.

To demonstrate that RMG can monitor the superficial and deep muscles, we carried out continuous recording of the complex forearm muscle contraction during various hand gestures by the MIMO channels, which provided rich information to decode the convoluted muscle activities. To validate HGR by RMG, we performed a human study of 8 participants with 23 gestures, including 8 basic gestures of fingers, palm, and wrist with multiple degrees of freedom (DoFs) and speeds. Various sensor modalities and forearm positions were also tested. After segmenting data from the RMG sensor, we transformed the 1D time waveforms to 2D time-frequency spectrograms using the short-time Fourier transform (STFT) and continuous wavelet transform (CWT). For gesture classification, we adopted vision transformer (ViT) as the deep-learning model [36] to compare with conventional convolutional neural networks (CNN). To provide a baseline comparison, we also benchmarked RMG with simultaneous sEMG for correlation and timing verification. To broaden the application scope, we also investigated leg RMG and radiooculogram (ROG), which tracked leg and eye muscle activities, respectively.

RMG can be applied to numerous applications. Gesture recognition and eye movement detection can be used as a middleware for HCI, such as virtual reality (VR) control and cybersickness reduction. In clinical applications, RMG can be used as assessment for voluntary and evoked muscle contraction, diagnosis of muscle

disorders, and rehabilitation. RMG can be integrated with sEMG for possible diagnosis of neuromuscular disorders including Parkinson's disease, as well as for feedback control of electromyostimulation (EMS). ROG can be further applied to rapid eye movement (REM) monitoring with eyes closed during sleep.

Operational Principles

Challenges for deep-layer muscle tracking

Muscles in a forearm are divided into anterior and posterior muscles, both containing superficial and deep layers. Hand gestures by the superficial muscle layers can be captured by motion sensors like accelerometers with tight skin contact. However, deep-layer muscles are critical for HGR but can raise ambiguity for surface-based sensors. The forearm muscles actuating the hand gestures are listed in Table 1. For example, flexor digitorum profundus is the only major muscle that can flex the distal interphalangeal joints of the fingers, and four of deep posterior muscles are important for thumb and index finger movements. Hence, forearm muscle sensors for hand and wrist gestures will be able to differentiate hand gestures reliably only if all muscle groups, not just the surface ones, are included in the sensor readout. Here, RMG provides a new solution to detect muscle actuation in superficial as well as deep layers for accurate HGR.

NCS: Near-field coupling inside the muscles

From near-field EM coupling, NCS directly modulates the superficial and deep muscle motion onto multiplexed radio signals. Previous radar-based systems often

operated in the far field when the EM energy is mostly reflected at the body surface, so only the surface motion would be captured [37]. In comparison, NCS has more EM energy directed inside the body so the modulated signal from internal tissues and organs is significantly larger. In our previous studies, NCS has been validated for deep coupling into human body to monitor heart valve motion [35][38], femoral pulses [39], and diaphragmatic breathing [40][41].

RMG adopted the sensing principle of NCS for muscle monitoring. In the near-field region of the forearm, the sensing antenna is designed to couple more EM energy into the muscles with high signal-to-noise ratio (SNR). The dielectric change of the internal muscle groups during the manipulation of hand gestures will be represented as the RF channel characteristics in terms of the scattering (S) parameters. Owing to the high penetration capability of UHF in the near field, RMG can potentially monitor all muscle groups in the forearm.

MIMO: Rich N^2 channel by N points

In this work, we adopt MIMO to incorporate N^2 usable channels from N observation points [35]. MIMO is a mature RF technique where different transmitters (Tx) can be well isolated by either frequency or code multiplexing. Similar techniques can be employed by colors in vision and subcarriers in ultrasound, but RF MIMO offers higher channel isolation than optical or acoustic waves with much lower cost thanks to the mature wireless industry. N Tx can then be simultaneously received and demodulated by N receivers (Rx) to accomplish N^2 synchronous channels to fulfill the spatial diversity requirement to observe complex 3D geometry and motion. Due to

tissue dispersion and near-field nonlinearity, the channel by Tx1–Rx2 would represent different information from Tx2–Rx1. Our RMG prototype utilized $N = 4$ sensing points around the forearm, and collected signals from 16 channels, and can be extended to even more channels with modest system cost.

Electromagnetic simulation for RMG

We further demonstrated the sensing principle of RMG using a numerical simulation in CST Microwave Studio [42]. The human forearm model for EM simulation was constructed from the Tom anatomical model in the CST Bio Extension 4.0 library. This voxel-based forearm model has a resolution of 0.5 mm^3 and contains accurate dielectric properties of the biological tissues in the UHF band including skin, muscle, blood, fat, and other tissues. Four dipole Tx antennas, as shown in Fig. 1(a), were deployed on the arm circumference without direct contact, as was done in the box RMG experiment in later sections. Each Tx was driven with a 1-W input source at 1 GHz and 50-Ω source impedance. The power here was selected for normalization convenience, as the radiated power level was less than 1 mW in actual experiments. Fig. 1(b) shows the electric field magnitude originating from each of the four Tx antennas at the cross section of forearm. We can clearly observe that the electric field was strongly coupled into the layers of skin, fat, and muscles. The sensing locality of the RMG system can be observed by different antenna coupling into different nearby muscle groups, providing high spatial diversity.

To show the change in the antenna characteristics by the NCS principle during the different muscle contraction phases, we performed a mock muscle contraction in

CST by isotropic muscle scaling of 1, 0.95, 0.9, and 0.85. Fig. 1(c) shows the normalized antenna reflections as the backscattering S parameters of S_{11}, S_{22}, S_{33} and S_{44}, which have small percentage changes demanding differential extraction or direct-path cancellation [43]. Fig. 1(d) presented cross channels between 4 antennas. The magnitude of the cross channels is smaller, but contains distinctive features during muscle scaling.

System design

Experimental setup

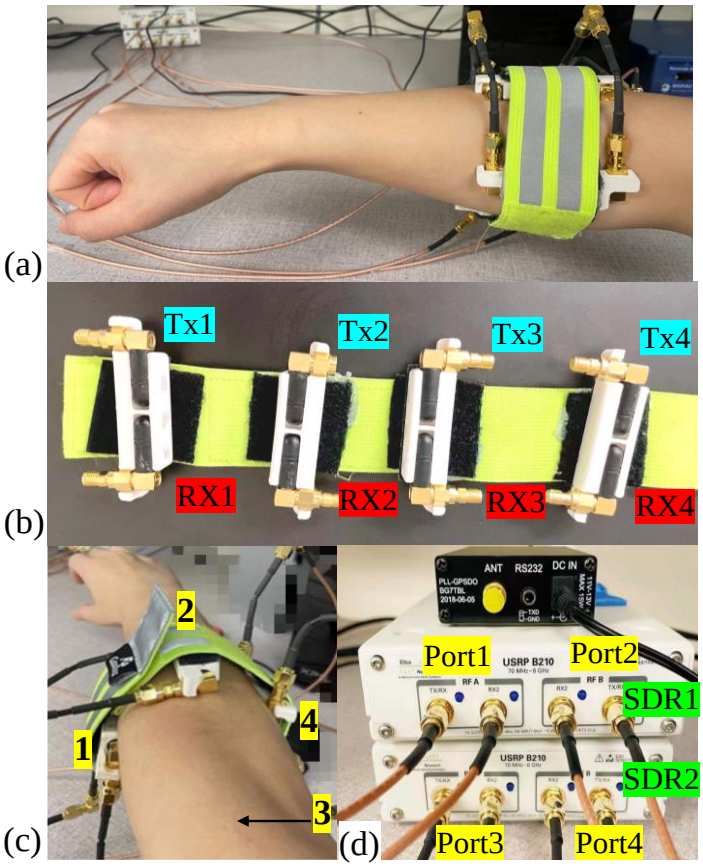


Fig. 2. (a) Forearm placement of a wearable RMG by SDR. (b) Four sensing antenna pairs attached to the armband. (c) The cross section view of RMG placement of 4 measuring points.

The third probe is on the anterior side of the forearm. (d) The transceiver s setup by two software-defined radios.

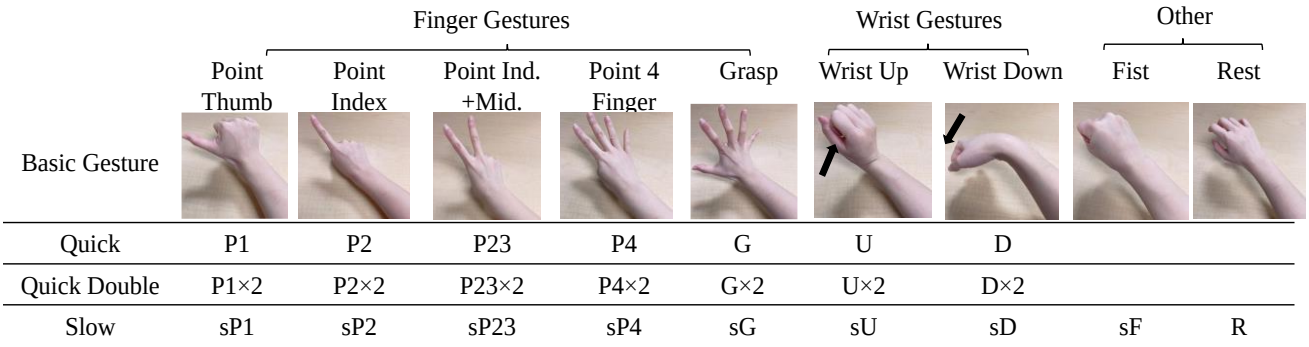
The first RMG prototype employed four pairs of the sensing antennas attached to a wearable armband on the middle forearm as shown in Fig. 2(a). Each sensing unit

TABLE I

Major muscle groups generating the hand gestures

Basic Gesture	Step 1	Major Muscles	Step 2	Major Muscles
Grasp	Extend 5 fingers	ED, EPL, EDM	Flex 5 fingers	FDP, FDS, FPL
Point Thumb	Extend thumb	EPL	Flex thumb	FPL
Point Index	Extend index	EI	Flex index	FDP, FDS
Point Ind.+Mid.	Extend ind.+mid.	EI, ED	Flex ind.+mid.	FDP, FDS
Point 4 Fingers	Extend 4 fingers	EI, ED, EDM,	Flex 4 fingers	FDP, FDS
Fist	Flex 5 fingers	FDP, FDS, FPL	Rest	
Wrist Up	Extend wrist	ECU, ECRL,ECRB	Flex wrist	FCU, FCR
Wrist Down	Flex wrist	FCU, FCR	Extend wrist	ECU, ECRL,ECRB

Muscle groups: ECU: Extensor Carpi Ulnaris; ECRL: Extensor Carpi Radialis Longus; ECRB: Extensor Carpi Radialis Brevis; FCU: Flexor Carpi Ulnaris; FCR: Flexor Carpi Radialis; ED: Extensor digitorum; EDM: Extensor digiti minimi; FDS: Flexor Digitorum Superficialis; EPL: Extensor Pollicis Longus; EI: Extensor Indicis; FDP: Flexor Digitorum Profundus; FPL: Flexor Pollicis Longus.
Green font: Superficial; Blue font: Intermediate; Red font: Deep.



Quick	P1	P2	P23	P4	G	U	D		
Quick Double	P1×2	P2×2	P23×2	P4×2	G×2	U×2	D×2		
Slow	sP1	sP2	sP23	sP4	sG	sU	sD	sF	R

Fig. 3. 23 hand gestures used in the study protocol.

consisted of two monopole whip antennas (Taoglas TG.19.0112) mounted on a 3D-printed holder, and had a dimension of 69 × 17 × 11 mm. As shown in Fig. 2(b), the antennas were aligned in parallel to the forearm muscle for enhanced coupling. Unit 1 was placed close to extensor pollicis longus and flexor pollicis longus, which controlled extension and flexion of the thumb. Unit 2 was placed close to the extensor muscles, which produced extension at the wrist and fingers. Unit 3 was placed close to the flexion muscles, which were associated with pronation of the forearm, as well as

flexion of the wrist and fingers. Unit 4 was close to flexor digitorum profundus which flexed the four fingers except the thumb. Multi-channel observation can help decode the convoluted muscle motion in various hand gestures by sensor proximity and rich MIMO channels.

The RMG transceiver was prototyped by two synchronized software defined radios (SDR, Ettus B210). The two SDRs were synchronized by an external local oscillator (LO, BG7TBL-GPSDO) with 10 MHz reference and 1 PPS (pulse per second) baseband synchronization. The SDRs were connected to a host computer through universal serial bus (USB), and the control software was implemented in LabVIEW. Each port of the MIMO system consisted of one Tx and one Rx, which was then connected to one sensing antenna pair. Each SDR supported two synchronous ports. Note that the present RMG on the armband were connected by cables to off-body SDR for fast and flexible prototyping of RF transceiver parameters. An all-in-one wireless unit of RMG can be a straightforward extension in the future [44], and further implementation by integrated circuits (IC) and custom packaging can make use of the present findings for product development with reduced size, power and cost.

The digital baseband tone f_{BB} of each Tx went through the digital-to-analog converter (DAC) and was then mixed with the carrier frequency f_{RF} in a standard quadrature scheme. The RF power was less than −10 dBm or 0.1 mW, well under the safety limits set by occupational safety and health administration (OSHA) in the UHF band [45]. The Tx signal was coupled into the forearm muscle groups, received by all Rx, and then demodulated and sampled by the analog-to-digital converter (ADC) to

retrieve the baseband. We employed the quadrature scheme as the baseband tone f_{BB}, and the NCS signal can be represented by the amplitude and phase modulation on the quadrature signal as

$$NCS_{am}(t) = \sqrt{I_{Rx}(t)^2 + Q_{Rx}(t)^2} \quad (1)$$

$$NCS_{ph}(t) = unwrap(\tan^{-1}\frac{Q_{Rx}(t)}{I_{Rx}(t)} - 2\pi f_{BB} - \theta_0) \quad (2)$$

$$I_{Rx}(t) = A(t)\cos(2\pi f_{BB}t + \theta_0) \quad (3)$$

$$Q_{Rx}(t) = A(t)\sin(2\pi f_{BB}t + \theta_0) \quad (4)$$

where θ_0 was the phase offset accumulated from the Tx–Rx signal chain and was not constant among different channels or setups. The antenna pair here can operate around 900 MHz and 1.8 GHz. Lower frequency often provided stronger penetration into human body and better signal coupling. Therefore, f_{RF} was selected at 900 MHz. The multiple Tx channels utilized frequency-division multiple access (FDMA) by setting f_{BB} =10, 25, 40, and 125 kHz, respectively, for Tx1–Tx4.

We configured the dual SDR as 4 self-channels and 12 cross channels. For example, Tx1 can be received by Rx1 as self backscattering, which was most affected by the muscle changes around Unit 1 to detect the extension and flexion of the thumb. Tx1 can also be received by Rx2–Rx4 as cross channels to collect information on the individual paths. All 16 channels are sampled at 10^6 samples per second (Sps) to implement Tx FDMA, and further down-sampled to 500 Sps to retrieve NCS magnitude and phase.

Human study protocol

RMG was tested on 8 healthy participants. The human study was approved by Cornell Institutional Review Board (IRB) under Protocol ID #1812008488, and conducted with the written consent of the participants. We designed 23 gestures including finger, palm, and wrist motions with various speeds and multiple DoFs as shown in Fig. 3. We had 8 basic dynamic gestures and 1 static resting gesture. Basic gestures were extended to three versions including quick, double quick, and slow, except that the gesture 'Fist' only had the slow version. These gestures are chosen for their common rendition in HGR, as well as for confirmation of deep muscle sensing. Every gesture was performed in a fixed time window of 5 s. All gestures excluding 'Rest' were dynamic and comprised two steps as described in Table 1. For the quick version, step 2 was performed immediately after step 1, while the slow version had a holding time around 2 s between steps 1 and 2. For each dynamic gesture, after step 2, the hand would relax back to the 'Rest' gesture. The on-off timing for each gesture motion inside the 5-s window was not fixed due to subject difference and variation in the response time for different repetitions. Each gesture was repeated around 30 times for each participant. The study procedures were divided into several repetitions of 5-min routines of two kinds. Routine 1 contained 16 finger-based gestures with 3 repetitions; Routine 2 contained 6 wrist-based gestures with 8 repetitions. The 'Rest' gesture was inserted between routines. Total recording time for each participant was around 1 hour. Participants occasionally made mistakes on the instructed gesture, and were suggested to report their mistakes after each 5-min routine. Routines with reported mistakes were removed from the datasets.

Data Processing

After collecting data on multiple participants, we processed the raw data before feeding the output to the machine learning (ML) models for classification. The signal processing before learning helped de-noise the dataset and avoid overfitting, which will become apparent when comparison was made against the ML model based on raw waveforms. The data flow is shown in Fig. 4.

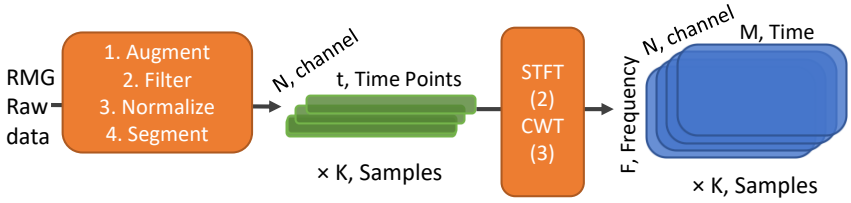


Fig. 4. Schematic for data pre-processing to feed spectrogram into machine learning.

Multi-channel augmentation

From the MIMO configuration in RMG, we obtained 16 channels on a forearm. Each channel contained the baseband phase $NCS_{ph}(t)$ and amplitude $NCS_{am}(t)$ in the quadrature scheme on f_{BB}. In addition to employing phase and amplitude, we also augmented the original complex number as part of the information to retain the intricate relation between $NCS_{am}(t)$ and $NCS_{ph}(t)$ [43]. Therefore, from 16 MIMO channels, we had 48 temporal series in total.

Filtering and segmentation

The 48 1D waveforms in time was then processed by:

1) Bandpass filtering (0.1 Hz to 5 Hz) to eliminate the noises in the higher frequency.

2) Waveform normalization with center 0 and standard deviation 1.

3) Waveform segmentation into individual segments of T_{seg} = 5 s. Each segment now contained one gesture guided by voice instruction.

4) Waveform detrending by subtracting the best-fit linear line from the data within T_{seg}.

5) Annotation of the instructed gesture for each segment.

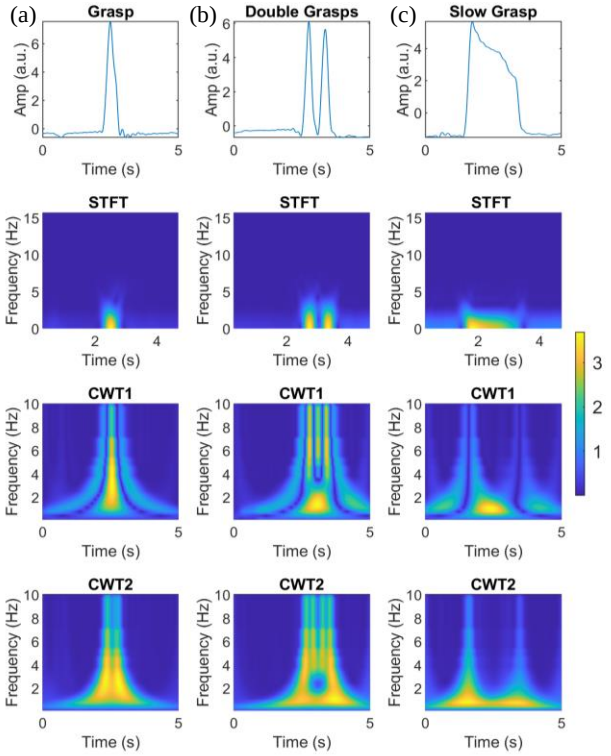


Fig. 5. Examples of 1D waveforms and transformed 2D spectrograms.

1D waveforms to 2D spectrograms

We employed STFT and CWT to generate 2D spectrograms to feed into the ML model. Transformation of 1D time waveforms to 2D time-frequency spectrograms would bring forth significant improvement in accuracy. The ensemble of five 2D spectrograms from different transforms were incorporated into ML for classification.

We explored two STFT outputs with different window lengths (T_{win1} = 0.6 s and T_{win2} = 1 s). Two window lengths can allow us to acquire information with different time and frequency resolutions [4]. CWT [5] takes advantage of multi-resolution analysis (MRA), which can effectively mitigate time-frequency resolution tradeoff. We used three different mother wavelets to capture different patterns and extend the feature diversity: 1) Ricker; 2) Gaussian; 3) Morelet.

In Fig. 5, three columns represent gestures of (a) fast grasp, (b) double grasps, and (c) slow grasps in the 5-s time windows. In each column, the first row shows the RMG time waveform, and the second to fourth rows are STFT (T_{win1} = 0.6 s), CWT1 (Ricker) and CWT2 (Gaussian) spectrograms. Note that STFT requires a short time window for n-point Fourier transform, so the starting and ending times of the spectrogram is truncated to 0.3 and 4.7 s.

Classification by vision transformer

Though classical ML models can be computationally less expensive, algorithmic and hardware improvements in recent years have facilitated complex neural networks on embedded systems efficiently [46]. We implemented vision transformer (ViT) as the classification ML model.

ViT has a deep-learning architecture inherited from the transformer model in natural language processing (NLP) [47] and is now gaining popularity in computer vision. To benchmark ViT performance, we also built a conventional CNN classifier.

Over the ViT architecture, patches of the 2D input image (size = 5) were constructed from the time-frequency spectrogram, and were then linearly embedded

with dimension = 512. Position embedding was added, and the resulting vector sequences were fed to a standard transformer encoder. Inside the encoder, we had 6 transformer blocks, 16 heads in the multi-head attention layer, 64 dimensions of the multi-layer perceptron (MLP) (feed-forward) layer, and the dropout rate was set to 0.1.

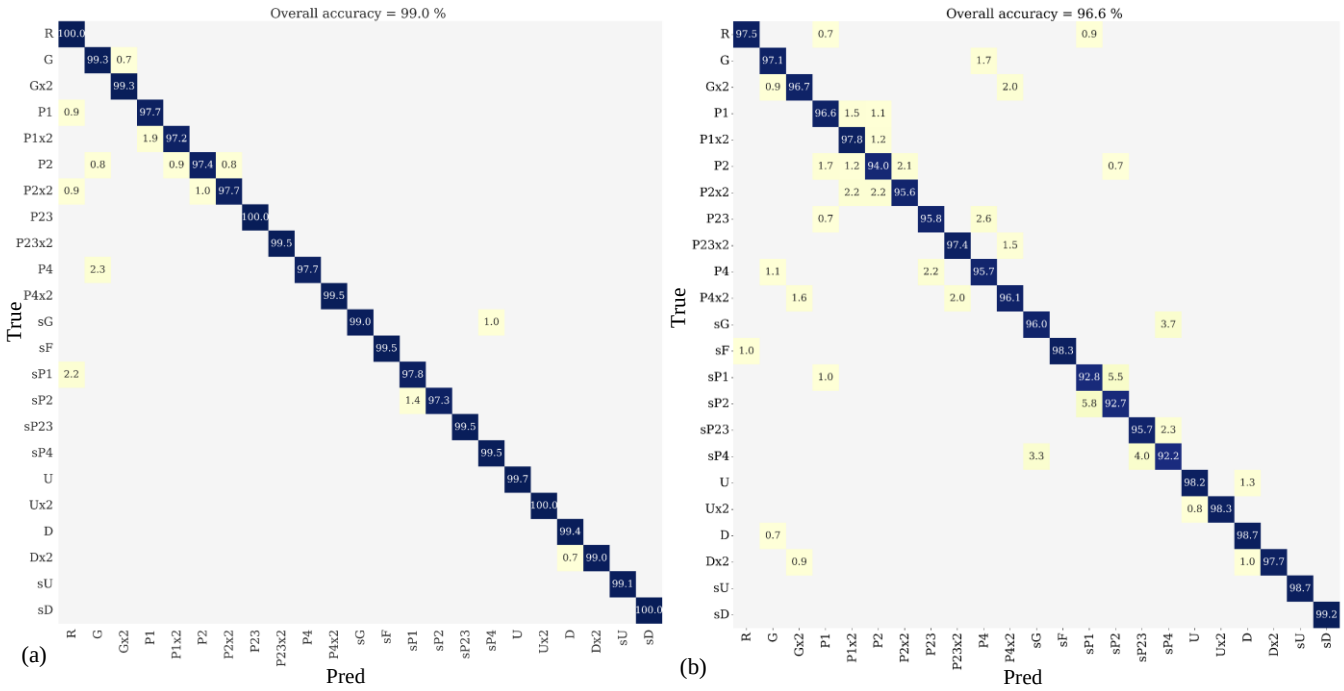


Fig. 6. The confusion matrices showing the overall accuracy on all participants using (a) The personal training model; (b) Transfer learning on the unseen participant by 1/5 of new data.

In CNN, each convolution layer was followed by a BatchNorm layer, and then 2 linear layers. Adam optimizer was used for both ViT and CNN.

Results and Analyses

Dataset composition

The final output dataset from all 8 participants consisted of 5,847 samples of

23 gestures. The ‘Rest’ gesture had 461 samples, each wrist-based gesture had around 283 – 288, each grasp-based gesture had 293 – 294, and each finger-based gesture had 215 – 222. Data exclusions were mostly due to reported mistakes by the participants, such as failing to follow the instruction on time and performing the wrong gestures.

The personal training model

We evaluated the classification accuracy of RMG by different cross-validation (CV) methods, feature sets and deep learning models. First, we built the personal training model for each participant. From individual person’s dataset, each gesture was repeated around 30 times, and the total sample number was around 700 – 800. K-fold (k = 7) CV was performed to estimate the mean accuracy for each participant. An overall accuracy was averaged on results from all participants. Fig. 6(a) shows the overall confusion matrix of ViT by the personal training model, which is normalized to the percentage of samples. RMG achieved an overall accuracy of 99.0% ± 0.48% for 23 gestures, which employed the ensemble method by majority voting of all feature sets from 2 STFT and 3 CWT versions. Fig. 7(a) shows an example of the trend of training and testing losses during model training on one subject. 6/7 of the overall data from one subject was trained, and 1/7 of the data was tested as unseen cases. We chose the learning rate of 10^{-4} and the epoch number of 20, and each epoch iteration has a batch size of 16. In the first 5 epochs, both training and testing losses decreased rapidly. In the following 5-20 epochs, training and testing losses both tended to be stable in low values, which indicated that the testing loss had a highly correlated decreasing pattern with the training loss and their eventual values at the end

did not have distinctive difference. Overfitting during training can often be spotted when the error on the training data decreased to a small value but the testing error increased in a reverse trend.

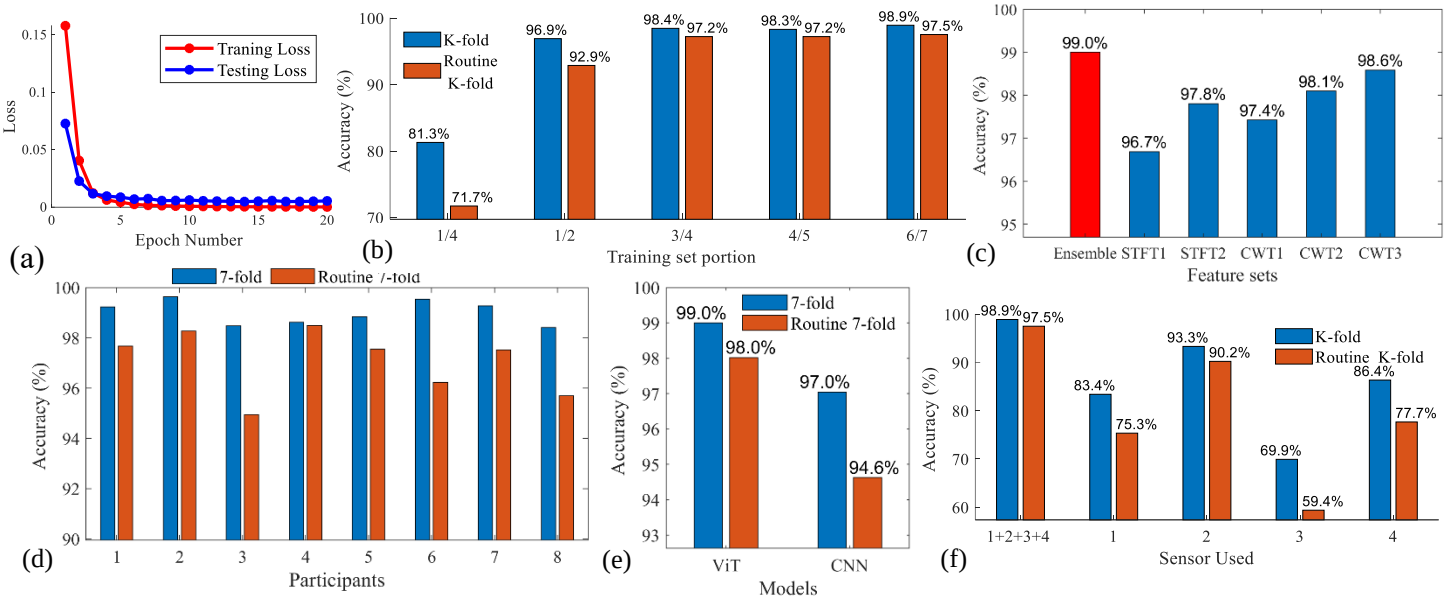


Fig. 7. Performance of the personal training model. (a) Example of training and testing loss during model training (Epoch number=20, learning rate=10-4). (b) Accuracy using different portion of total dataset. (c) Accuracy using different transforms, (d) Accuracy for individual participants, (e) Accuracy by different ML models. (f) Accuracy using all RMG sensor units vs. individual sensor.

Fig. 7(b) then showed the accuracy when different portion of the dataset was chosen for training and testing. We changed to train on 4/5 (k = 5), 3/4 (k = 4), 1/2 (k = 2), and 1/4 (k = 4 with training and testing swapped). Though training on ¼ data had lower accuracy, when the model had at least half of the overall data to learn, the accuracy maintained at 96.9%. We can observe that the model performance did not degrade much even with limited training cases. These results indicated that there was no apparent overfitting in our classification model. We also compared the results using different feature sets individually and presented the results in Fig. 7(c). STFT2 used a

longer time window length and thus had a higher frequency resolution and accuracy than STFT1. CWT features generally outperformed STFT. CWT3 by Morelet wavelet achieved the highest accuracy of 98.6% among all individual feature sets. The ensemble method with the flexibility to choose among all alternatives achieved the highest accuracy.

Our data were collected based on 5-min routines (around 60 repetitions of each gesture in total) in the hour-long study for each participant to allow some rest and sensor adjustment. Each subject had around 12 routines. Between the 5-min routines, the hardware would be reset, and the subjects would take off the sensor to rest, followed by redeployment. Small sensor position variation around 1-2 cm or rotation around 5-15° would occur between the routines. Due to the hardware restart and sensor position alteration, signals collected by different routines can have more variations. Therefore, we presented another CV process, where all routines were independently tested, i.e., the gestures in the same routine were never divided between training and testing. As shown in Fig. 7(d), routine-independent CV still achieved a high accuracy of 97.0% ± 1.27% for all 8 subjects. Compared with the random shuffle k-fold, routine-independent CV had lower mean accuracy and higher standard deviation across different subjects. This observation of maintaining high accuracy in routine-independent CV corroborated the system robustness against the hardware reboot and small position variation in practical scenarios.

ViT was compared to CNN in Fig. 7(e), where the accuracy dropped from 99.0% in ViT to 97.0% in CNN for personal training CV, and from 98.0% in ViT to 94.6% in CNN for routine-independent CV. To illustrate the advantage of 2D

spectrogram, we also built a 1D-CNN model using the time waveforms directly. Accuracies dropped from 97.0% in 2D CNN to 88.5% in 1D CNN for personal training CV, and similarly from 94.6% to 88.0% for routine-independent CV. As ViT is often computationally more expensive than CNN, this comparison can be regarded as a tradeoff between accuracy and computing resources.

We adopted the MIMO setup for RMG in order to collect both self and cross channels. We also tested the recognition accuracy using only self-backscattering channels as the input for the ViT model. Accuracy degraded from 99.0% to 95.0% in the personal training CV. Currently, we have 4 RF antennas functioning as sensing units on the armband. We further analyzed the accuracy degradation in Fig. 7(f) when the number of the sensing units was reduced. We chose one subject as an example where MIMO achieved higher accuracy than any other individual sensor used alone. Sensor 2 on the anterior side had the highest accuracy of 93.2%, while sensor 3 on the posterior side had the lowest accuracy of 69.9%. Large accuracy variation from different sensors was probably due to the different coupling to various muscle groups. In summary, the channel spatial diversity in MIMO played a critical role in HGR accuracy.

Transfer learning for unseen users

The HGR system must be robust to various practical conditions, especially for subject variation. Not only people perform hand gestures differently, but also the forearm size and muscle conditions have considerable distribution. Here, we adopted conventional transfer learning (TL) [48] where we leveraged a pre-trained model with

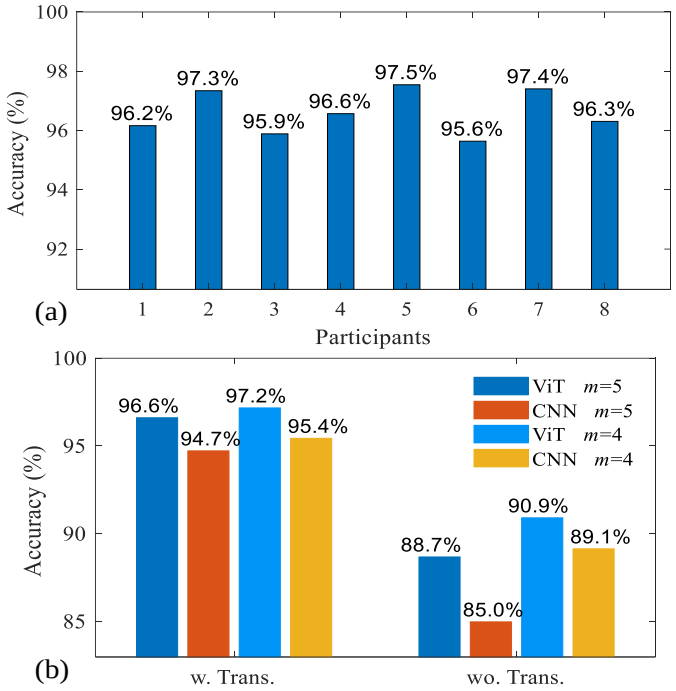


Fig. 8. Performance of transfer learning on unseen participants. (a) Accuracy using TL with m = 5 on individual participants. (b) Accuracy with and without TL for m = 4 or 5 by ViT and CNN.

large amount of data from multiple users to test on a new user with a small amount of individual training data. TL also has been widely used by other HGR systems with high generalization and low training burden [46][48]. We first generated the pre-trained model using all data from 7 participants. We then fine-tuned the model with *1/m* data from the new participant as short personal calibration. The final model was tested on the rest (*1 – 1/m*) data. This CV process is similar to *k*-fold, but only one fold is for training, and (*m – 1*) folds are for testing. Fig. 8(a) shows the accuracies for all participants rotating as the new test case by the above TL strategy with *m* = 5. The model was entirely reset for each rotation. The averaged accuracy is 96.6%±0.736%, and the normalized confusion matrix is presented in Fig. 6(b).

ViT also outperformed CNN for our TL strategy[36]. Fig. 8(b) shows accuracies for *m* = 4 or 5 with and without TL in the ViT and CNN models. ViT achieved higher accuracy than CNN in every scenario. Direct learning from *1/m* data

without TL had much lower accuracy than the pre-trained model by TL. When the personal training set increased from ⅕ to ¼, accuracy also noticeably increased, indicating the trade-off between high accuracy and the amount of personal training data.

Variations in experimental designs

TABLE II EVALUATION FOR POSITION AND DESIGN VARIATIONS

	D = 3CM WITH TL	*D* = 3CM WITHOUT TL	NOTCH	BOX	WRIST
ACCURACY (%)	97.2	87.2	99.0	97.4	95.8

Apart from subject dependence, accuracy degradation can also be induced from the sensor placement on the forearm. To test the adaptability against large sensor position variation, we performed another test on one participant with the same protocol but with the sensor position moved to a higher position by d = 3 cm. We used the same TL strategy to achieve the result in the first two columns of Table II. After adopting TL, accuracy was boosted from 87.2% to 97.2%.

For the human study above, we used the antenna-based sensing unit on the forearm. We also explored more sensor design variations. The first design in Fig. 9(a) was notch RMG, where the muscle motion was coupled to an RF coaxial cable with an open notch leaking out a portion of the EM energy[49]. The notch RMG has the potential to miniaturize the sensor size, and can be readily adapted to flexible wearables. The second design in Fig. 9(b) is a non-contact square box with the antenna sensors attached to the inside walls. Forearm can be placed into the box freely without direct contact. The third design in Fig. 9(c) is by the same sensing antenna, but placed

on the wristband, which can be convenient for integration into the smart watch as a new input method. Present user interface by fingers on the smart watch display has been impeded by the small screen size, and hand gestures can be a promising alternative [22]. Table II presents the HGR accuracies using the above three design variations. Notch RMG showed the highest accuracy and can be favorable in certain applications. Box RMG can still attain reasonable accuracy by the non-contact setup, which further enhance the design flexibility over clothing or in armrests. Wrist RMG showed lower accuracy than the forearm placement because tendon and ligament motion had less dielectric contrast than the muscle motion.

To validate that strong RF coupling was from the forearm muscles and not from the direct hand motion in the radar mode, we conducted measurements with the hand inside a radar-absorption-material (RAM) box, as shown in Fig. 9(d), where minimal difference in collected waveforms and achievable HGR accuracy was observed.

BENCHMARK WITH sEMG

We performed RMG with synchronous sEMG for the baseline comparison and physiological correlation. Notice that our sEMG setup had only one or two channels and was implemented by a commercial device without optimization. The performance of our sEMG study is expected to lag behind many state-of-the-art multi-channel implementations 27. Nevertheless, the two sensing schemes can be complementary in operation to establish the complete physiological sequence of stimulation and

actuation, as well as to study the neuromuscular disorders in the future.

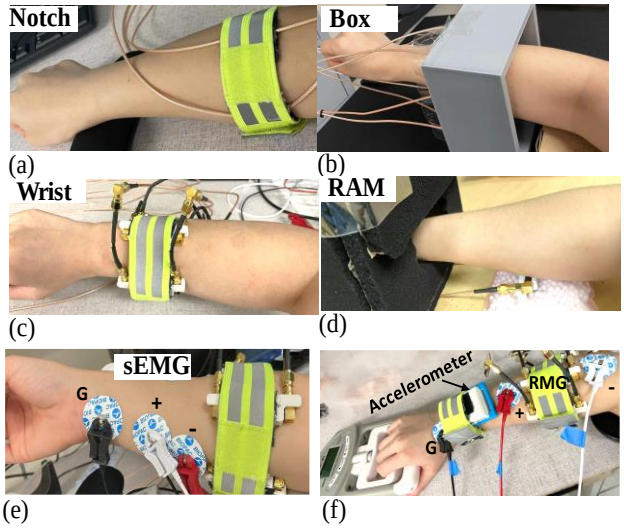


Fig. 9. Experimental setups of various designs: (a) A notch RMG; (b) A non-contact box; (c) A wristband; (d) Verification by the hand inside an RAM box; (e) Benchmark with sEMG with short +/– separation; (f) Slow grasp strength testing with sEMG and accelerometer.

RMG and sEMG placement

For the reference sEMG setup, we used BIOPAC MP36R with the SS2LB leads set and EL503 electrodes (BIOPAC Systems, Goleta, CA). Fig. 9(e) shows the experimental setup with RMG and EMG both on the forearm. Each EMG channels has

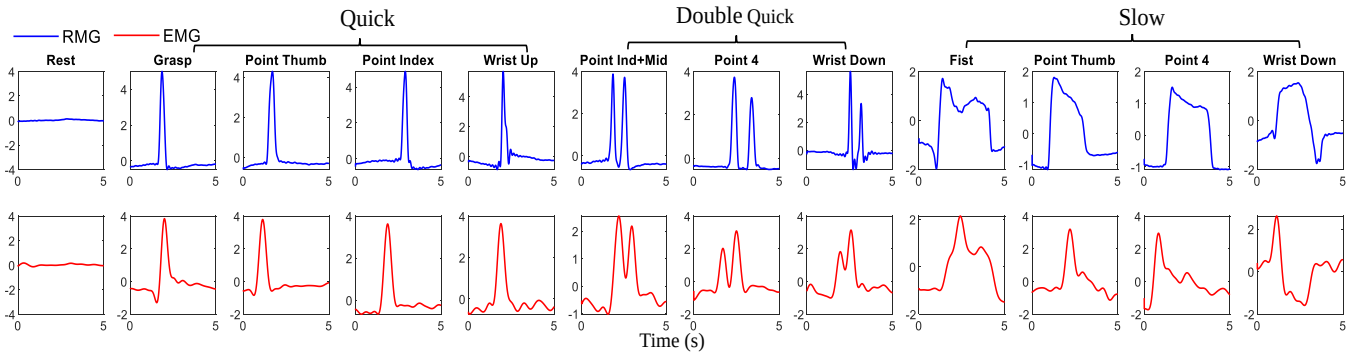


Fig. 10. RMG and sEMG waveforms for various gestures by DTW averaging on all samples.

3 electrodes on skin as +, −, and ground. We used 2 sEMG channels on the anterior and posterior sides of the forearm. The ground electrodes for two sEMG channels were both placed close to a wrist spot with minimal muscles. RMG and sEMG channels were synchronized on Labview. We performed the same study protocol on two participants as Exp1 and Exp2 in Table III. The two participants had the same sEMG placement, where + and − electrodes were on the two longitudinal sides of the RMG armband to capture more differential signals with a large distance. Exp3 was the same participant as Exp1, but had a different sEMG placement where the + and − electrodes were on the lower position from the RMG armband. The smaller distance would measure only the muscles close to the electrodes with less voltage resolution.

TABLE III

ACCURACY COMPARISON OF RMG VS. SEMG

	EXP1	EXP2	EXP3	MEAN
RMG	99.0%	98.5%	98.7%	98.7%
SEMG	68.2%	70.8%	66.7%	68.6%

RMG and sEMG waveform comparison

As our sEMG waveforms were noisy during the hand gestures, we added two pre-processing procedures: Enveloping the raw data by spline interpolation over local maxima, and smoothing by moving average 28. The subsequent signal transformation and learning models were the same for sEMG and RMG. The overall HGR accuracy by 7-fold CV is shown in Table III. Accuracy of sEMG was relatively low in comparison with RMG in our setup, which may be caused by the small number of sEMG channels under the large number of gesture classes. Our sEMG implementation

was mainly for comparative purposes and was far from ideal. A more comprehensive comparison with the literature results will be presented in next section.

As shown in Fig. 10, we also compared the averaged waveforms of different gestures obtained from RMG and sEMG using global dynamic time warping (DTW) [50]. Each gesture had a time segment of 5 s, while the y-axis was the normalized amplitude. The RMG waveforms were examples from Tx2-Rx2 and the sEMG from channel 2, both positioned on the posterior side of the forearm. For quick and double quick gestures, both RMG and sEMG presented sharp peaks corresponding to the fast muscle motion. However, compared with RMG, sEMG signals had longer duration of pulse waveforms and showed more tailing after the gesture motion terminated. For slow gestures, RMG showed a more consistent square-wave pattern from the holding period. The sEMG signal showed a shorter pulse duration for gestures that do not require continuous myoelectrical simulation such as the point-finger gestures. For other gestures that require continuous efforts to maintain the position such as the wrist up/down, the sEMG pulse duration were extended. During 'Rest' and between gestures with no intended hand motion, sEMG had more interference and ambiguity due to either hardware sources such as inconsistent electrode contact resistance or from biological sources such as the neural signals from vital signs[32]. In comparison, RMG is less susceptible to vital signs or noises from electrode contacts.

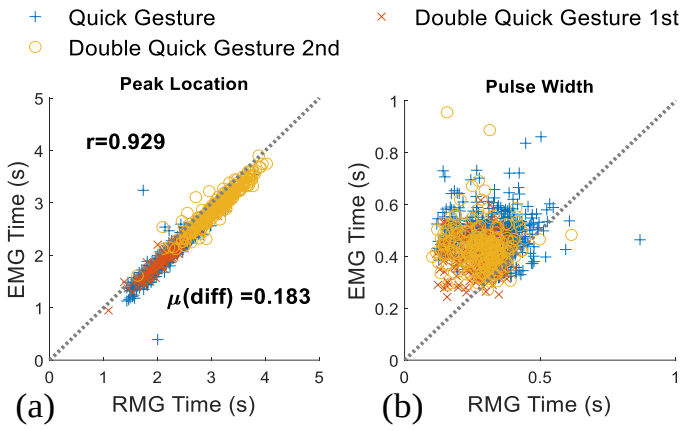


Fig. 11. Scatter plots of RMG and sEMG for peak location and pulse width during quick gestures.

To compare the waveform features further, we performed peak detection in the 14 quick gestures. Fig. 11(a) shows the scatter plot of peak locations of quick gestures in synchronous RMG and sEMG in all samples, where the Pearson correlation coefficient r = 0.929 and the mean time difference is a delay of 0.183 s, i.e., RMG and sEMG have a high temporal correlation and a consistent time lag. This delay may indicate the time offset between neural stimulation and muscle actuation. Fig. 11(b) compares the feature of the pulse width, computed as the time duration between the points to the left and right of the half peak magnitude. Most data points are scattered above $y = x$ line, which indicates that RMG waveforms have sharper peaks with less spreading during the quick gestures. Note that the few outliers are probably due to peak detection errors caused by the cases of questionable signal quality. For slow gestures, peak detection is not an appropriate comparison because the waveform features are not always consistent.

Timing and latency of RMG

RMG has ultra-low latency with the sampling rate readily over 10^5 samples per

second (Sps), which is important for dynamic HGR. Here, we performed the high-speed gesture tracking by RMG and sEMG. The participant followed a metronome of 150 beats/minute and performed the gesture of 'point index and middle fingers' with equal strength at each beat. The sensor setup was the same as Fig. 9(d). The waveforms from one of each RMG and sEMG channels are shown in Fig. 12(a). We can observe from the time waveforms that RMG had a consistent signal pattern corresponding to the quick motions, while sEMG had more fluctuations.

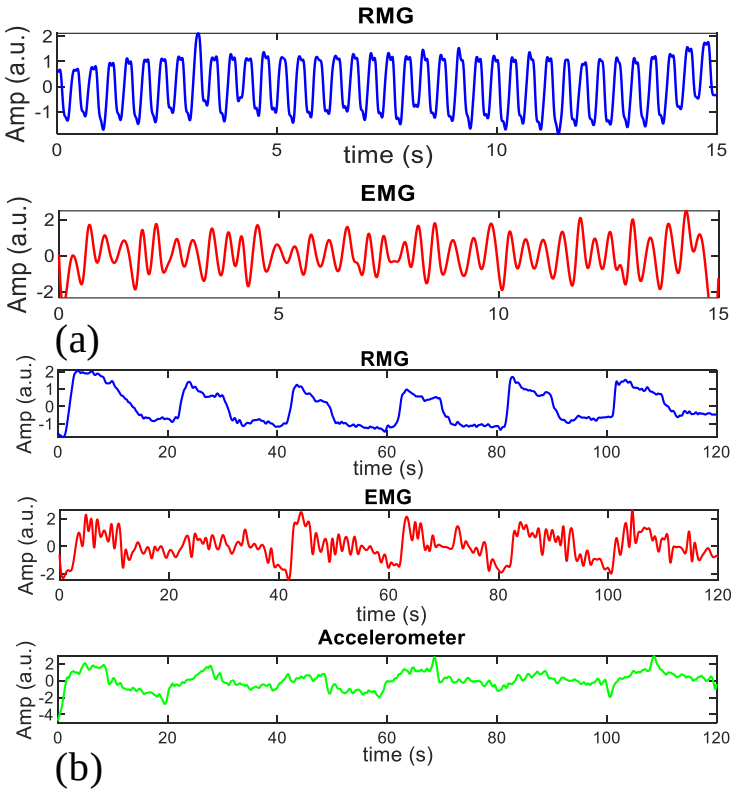


Fig. 12. Waveforms recorded from RMG, sEMG, and accelerometer for (a) fast finger motion; (b) slow grasps in 3 times/min with equal strength.

Compared with surface-motion based sensors including MMG and accelerometers, RMG possesses the unique capability to capture deep muscle contraction. To further corroborate this claim, we tested a slow grip strength detection by RMG together with accelerometers and sEMG. As shown in Table I, during the grip motion, the main muscles include the flexor digitorum superficialis (intermediate), flexor digitorum profondus (deep) and the flexor policus longus (deep) [51]. The participant performed firm holds on the hand dynamometer with a speed of 3 times/minute in equal strength as shown in Fig. 9(f). The waveforms from one of each RMG, sEMG, and accelerometer channels are shown in Fig. 12(b). RMG had a clear and stable signal pattern reflecting the strong and slow grip motion, while sEMG showed some ambiguity and the accelerometer presented even more noisy patterns. This is likely due to the different coupling strength to the deep muscle groups by different sensors.

Fig. 13. Setup of the ROG and leg RMG systems. (a) One ROG sensor unit by a notched transmission line; (b) Four ROG sensor units on a mask; (c) ROG on a participant's face; (d) EOG setup for baseline comparison; (e) One lower leg RMG sensor unit by whip antennas together with one EMG; (f) Four leg RMG sensor units on two legs.

Extension to eye and leg RMG

To validate the general applicability of RMG to different skeletal muscles, we

further extended the setup to wearable radiooculogram (ROG) on eyes and RMG on legs.

As shown in Figs. 13(a)(b), the ROG system integrated four notch RMG to a facemask around the eyes. A participant wearing ROG and electrooculorgram (EOG) was shown in Figs. 13(c)(d). In a human study of 5 subjects, participants were instructed to move eyes in four directions (up, down, left, and right) with eyes closed, all of which had 2 versions of moving once and twice. Hence, we had 8 distinctive eye movements, and each motion was performed in a time segment of T_{seg} = 5 s with around 24 repetitions for each participant. Then the training model within each participant was built and 7-fold CV was performed to estimate the mean accuracy. ROG achieved an overall accuracy of 94.2%. ROG can monitor fine eye muscle activities with eyes open or shut. In the future, ROG can be applied for sleep REM and dream stage monitoring[52], and facilitate HCI applications using eye motion control.

Another extension is for monitoring lower leg muscles. We implemented 2 RMG sensing units on each leg with sEMG for reference, as shown in Figs. 13(e)(f). We tested 7 postures: 1) tiptoe standing; 2) tiptoe sitting; 3) reverse tiptoe standing; 4) reverse tiptoe sitting; 5) tiptoe sitting with only the right foot; 6) tiptoe sitting with only the left foot; 7) squat. Each posture was also performed in a time segment of T_{seg} = 5 s with around 34 repetitions. Leg RMG achieved accuracy of 100% for one participant using 7-fold CV. RMG on lower legs can monitor body postures and can be applied for balance training and fall warning[53].

DISCUSSION

Comparison to previous HGRworks

A comparison of RMG to previous HGR systems is presented in Table IV. Li et al. [17] can achieve high accuracy of 98.5% for 8 classes, but requires off-body line-of-sight (LoS) cameras, and is vulnerable to change of light and background. Zhang et al. [33] used Electrical Impedance Tomography (EIT) to recover the interior

TABLE IV

COMPARISON TO PREVIOUS WORKS

	Li 2019 [17]	Zhang 2016 [25]	Zhang 2015 [33]	McIntosh 2016 [30]	Savur 2016 [14]	Qi 2020 [29]	Côté-Allard 2019 [46]	Moin 2021 [27]	**This work**
Class	8	8	5	15	27	9	7/18	13/21	**23**
Subject	5	4	10	12	1	-	17/10	2	**8**
Sensor	Camera	FMCW Radar	EIT	sEMG+ pressure	sEMG	sEMG	sEMG	sEMG	**RMG**
Model	CNN	CNN	SVM	SVM	Ensemble	GRNN	ConvNet	Neural	**ViT**
Accuracy	98.5%	96.0%	97% (hand) 87% (pinch)	95.8%	79.4%	95.3%	98.3% (7) 69.0% (18)	97.1% (13) 92.9% (21)	**99.0%**

impedance distribution of the tested arm, which was similar to RMG because both techniques monitored interior muscle activities. However, EIT needs many electrodes and suffers from reproducibility and accuracy degradation across users. Many previous efforts employed sEMG for HGR, which required direct skin contact and a large number of electrodes to achieve high accuracy. McIntosh et al. [30] successfully integrated pressure sensor with sEMG, but needed 8 wet electrodes and 4 pressure

sensors. The recent work from Moin et al. [27] showed 92.9% for 21 gestures with high in-sensor adaptability, but required 64 electrodes integrated on the armband. In comparison, RMG only utilized 4 sensing units to achieve 16 channels by MIMO. In this work, we demonstrated the competitive RMG to recognize 23 gestures (including 8 basic gestures in three different speeds) with accuracy up to 99.0%. Notice that the number of hand gestures was different in various works due to the intended applications, and the larger number did not directly indicate higher sensor capability except for the increased complexity in the classification algorithm. Moreover, the wearable and armrest RMG setups without requiring direct skin contact or restricting the capture volume offer inherent operational advantages over sEMG and camera-based systems. Our choice of ViT classification on spectrograms also shows better performance than traditional ML models adopted in previous works.

Potential future improvements

A. Sensor hardware improvement

In future hardware implementations, we should be able to miniaturize RMG into convenient and comfortable packages as all-in-one wireless wearables, because the expected power consumption and data bandwidth are both very low in view of modern RF devices. The notch RMG offers a promising design path to reduce cost, form factors, and complexity, especially for integration with a wristwatch.

B. Real-time classification for HCI

For HCI in robotic and gaming control, real-time HGR with minimal latency is an important feature. Embedded learning capability with local signal processing and

accurate HGR output of RMG will be attractive to many applications. With a given pre-trained ViT model, an inference on a single gesture presently took less than 1 ms for the execution time in a modest PC gaming console. Further custom hardware acceleration and algorithmic optimization can be applied to enable future real-time HGR.

C. Fusion with sEMG

sEMG can estimate neural stimulation of muscle actuation, and RMG can directly detect the actual muscle change. Thus RMG should not be viewed as a competition of sEMG, but the two sensors can be combined for a fuller physiological interpretation. Our consistent observation of the RMG delay from sEMG possibly indicated the non-trained muscle actuation without the participation of proprioceptive neurons, which can be promising for neuromuscular disorder diagnosis with RMG and sEMG fusion.

D. Closed-loop EMS control

A closed-loop control of EMS is another possible future application. EMS has long been employed to either supplement or substitute voluntary muscle stimulation in many settings of rehabilitation and electroceuticals [54]. However, inadequate EMS due to personal and daily differences can cause confusion of antagonistic and synergistic coordination of the muscle groups, and even induce serious spasm. For a more precise control on EMS, RMG can give feedback on actual muscle actuation to control the EMS signal with higher adaptability to personal and conditional variations.

CONCLUSION

We have reported a novel muscle monitoring technique, named as radiomyography (RMG), which can directly measure the muscle motion by coupling RF energy to superficial and deep internal muscles. Operation over clothing without direct skin touch enables convenient setup and comfortable operation.

The MIMO approach enriches the collected information with a relatively small number of sensing points. We implemented RMG as a wearable forearm sensor to accurately track forearm muscles. For the HGR purpose, we adopted ViT as the classification model and effectively boosted the accuracy up to 99.0% for 23 hand gestures tested on 8 participants. We further adopted TL to address cross-subject and operational variations. For HGR systems, RMG has lower cost, lower complexity, lower latency and less privacy issues than camera-based devices, as well as higher user comfort and accuracy than contact-based devices.

RMG has the unique advantage to monitor internal muscles non-invasively. In the future, RMG and sEMG can be fused together to derive the closed-loop information of stimulation and actuation. RMG can potentially lead to new methods for assessment of muscle functions, monitoring of muscle fatigue, and diagnosis of neuromuscular disorders. RMG is also promising for future HCI applications including exoskeleton robotic control, virtual reality interface, and in-air gesture capture.

REFERENCES

1 M. Cifrek, V. Medved, S. Tonković, and S. Ostojić, "Surface EMG based muscle fatigue evaluation in biomechanics," *Clin. Biomech.*, vol. 24, no. 4, pp. 327-340, 2009.

2 D. Leonardis et al., "An EMG-controlled robotic hand exoskeleton for bilateral rehabilitation," *IEEE Trans. Haptics*, vol. 8, no. 2, pp. 140-151, 2015.

3 P. K. Jamwal, S. Hussain, Y. H. Tsoi, and S. Q. Xie, "Musculoskeletal model for path generation and modification of an ankle rehabilitation robot," *IEEE Trans. Hum. Mach. Sys.*, vol. 50, no. 5, pp. 373-383, 2020.

4 I. Moon, M. Lee, J. Chu, and M. Mun, "Wearable EMG-based HCI for electric-powered wheelchair users with motor disabilities," in *Proc. IEEE Intl. Conf. Robotics and Automation*, 2005, pp. 2649-2654.

5 Q. Qian, X. Hu, Q. Lai, S. C. Ng, Y. Zheng, and W. Poon, "Early stroke rehabilitation of the upper limb assisted with an electromyography-driven neuromuscular electrical stimulation-robotic arm," *Front. Neurol., Clinical Trial*, vol. 8, Sept. 2017.

6 M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of EMG signal analysis: detection, processing, classification and applications," *Biol. Proced.*, vol. 8, no. 1, pp. 11-35, 2006.

7 M. O. Ibitoye, N. A. Hamzaid, J. M. Zuniga, and A. K. A. Wahab, "Mechanomyography and muscle function assessment: A review of current state and prospects," *Clin. Biomech.*, vol. 29, no. 6, pp. 691-704, 2014.

8 C. D. Lee, Y. Song, A. C. Peltier, A. A. Jarquin-Valdivia, and P. D. Donofrio, "Muscle ultrasound quantifies the rate of reduction of muscle thickness in amyotrophic lateral sclerosis," *Muscle Nerve*, vol. 42, no. 5, pp. 814-819, 2010.

9 E. Mercuri, A. Pichiecchio, J. Allsop, S. Messina, M. Pane, and F. Muntoni, "Muscle MRI in inherited neuromuscular disorders: past, present, and future," *J. Magn. Reson. Imaging*, vol. 25, no. 2, pp. 433-440, 2007.

10 J. S. Sonkusare, N. B. Chopade, R. Sor, and S. L. Tade, "A review on hand gesture recognition system," in *2015 International Conference on Computing Communication Control and Automation*, 2015, pp. 790-794.

11 R. Z. Khan and N. A. Ibraheem, "Hand gesture recognition: A literature review," *Int. J. Artif.*, vol. 3, no. 4, p. 161, 2012.

12 J. DelPreto and D. Rus, "Plug-and-play gesture control using muscle and motion sensors," in *Proc. ACM/IEEE Intl. Conf. Human-Robot Interaction*, 2020, pp. 439-448.

13 H. Liu and L. Wang, "Gesture recognition for human-robot collaboration: A review," *Int. J. Ind. Ergon.*, vol. 68, pp. 355-367, 2018.

14 C. Savur and F. Sahin, "American Sign Language Recognition system by using surface EMG signal," in *IEEE Intl. Conf. Systems, Man, and Cybernetics (SMC)*, 2016, pp. 002872-002877.

15 K. A. Smith, C. Csech, D. Murdoch, and G. Shaker, "Gesture recognition using mm-wave sensor for human-car interface," *IEEE Sens. Lett.*, vol. 2, no. 2, pp. 1-4, 2018.

16 S. S. Rautaray and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey," *Artif. Intell. Rev.*, vol. 43, no. 1, pp. 1-54, 2015.

17 G. Li et al., "Hand gesture recognition based on convolution neural network," *Cluster Comput.*, vol. 22, no. 2, pp. 2719-2729, 2019.

18 R. H. Venkatnarayan and M. Shahzad, "Gesture recognition using ambient light," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 1, pp. 1-28, 2018.

19 J. Suarez and R. R. Murphy, "Hand gesture recognition with depth images: A review," in *IEEE RO-MAN: 21st IEEE Intl. Symp. Robot and Human Interactive Communication*, 2012, pp. 411-417.

20 B. Feng *et al.*, "Depth-projection-map-based bag of contour fragments for robust hand gesture recognition," *IEEE Trans. Hum. Mach. Sys.*, vol. 47, no. 4, pp. 511-523, 2016.

21 D. Mazumdar, A. K. Talukdar, and K. K. Sarma, "Gloved and free hand tracking based hand gesture recognition," in *1st Intl. Conf. Emerging Trends and Applications in Computer Science*, 2013, pp. 197-202.

22 C. Xu, P. H. Pathak, and P. Mohapatra, "Finger-writing with smartwatch: A case for finger and hand gesture recognition using smartwatch," in *Proc.16th International Workshop on Mobile Computing Systems and Applications*, 2015, pp. 9-14.

23 J. Lien et al., "Soli: Ubiquitous gesture sensing with millimeter wave radar," *ACM Trans. Graphics*, vol. 35, no. 4, pp. 1-19, 2016.

24 X. Gao et al., "Barcode based hand gesture classification using AC coupled quadrature Doppler radar," in *IEEE MTT-S Intl. Microwave Symp. (IMS)*, 2016,

pp. 1-4.

25 Z. Zhang, Z. Tian, and M. Zhou, "Latern: Dynamic continuous hand gesture recognition using FMCW radar sensor," *IEEE Sens. J.*, vol. 18, no. 8, pp. 3278-3289, 2018.

26 K. S. Krishnan, A. Saha, S. Ramachandran, and S. Kumar, "Recognition of human arm gestures using Myo armband for the game of hand cricket," in *IEEE Intl. Symp. Robotics and Intelligent Sensors (IRIS)*, 2017, pp. 389-394.

27 A. Moin et al., "A wearable biosensing system with in-sensor adaptive machine learning for hand gesture recognition," *Nat. Electron.*, vol. 4, no. 1, pp. 54-63, 2021.

28 U. Côté-Allard, C. L. Fall, A. Campeau-Lecours, C. Gosselin, F. Laviolette, and B. Gosselin, "Transfer learning for sEMG hand gestures recognition using convolutional neural networks," *IEEE Intl. Conf. Systems, Man, and Cybernetics (SMC)*, 2017, pp. 1663-1668.

29 J. Qi, G. Jiang, G. Li, Y. Sun, and B. Tao, "Surface EMG hand gesture recognition system based on PCA and GRNN," *Neural. Comput. Appl.*, vol. 32, no. 10, pp. 6343-6351, 2020.

30 J. McIntosh, C. McNeill, M. Fraser, F. Kerber, M. Löchtefeld, and A. Krüger, "EMPress: Practical hand gesture classification with wrist-mounted EMG and pressure sensing," in *Proc. CHI Conf. Human Factors in Computing Systems*, 2016, pp. 2332-2342.

31 N. Amrutha and V. Arul, "A review on noises in EMG signal and its removal," *Int. J. Sci. Res. Publ*, vol. 7, no. 5, pp. 23-27, 2017.

32 K. S. Türker, "Electromyography: some methodological problems and issues," *Phys. Ther.*, vol. 73, no. 10, pp. 698-710, 1993.

33 Y. Zhang and C. Harrison, "Tomo: Wearable, low-cost electrical impedance tomography for hand gesture recognition," *Proc. 28th Annual ACM Symposium on User Interface Software & Technology*, 2015, pp. 167-173.

34 X. Hui and E. C. Kan, "Monitoring vital signs over multiplexed radio by near-field coherent sensing," *Nat. Electron.*, vol. 1, no. 1, pp. 74-78, 2018.

35 X. Hui, T. B. Conroy, and E. C. Kan, "Multi-point near-field RF sensing of blood pressures and heartbeat dynamics," *IEEE Access*, vol. 8, pp. 89935-89945, 2020.

36 A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint* arXiv:2010.11929, 2020.

37 Y. Kim and B. Toomajian, "Application of Doppler radar for the recognition of hand gestures using optimized deep convolutional neural networks," *11th European Conf. Antennas and Propagation (EUCAP)*, 2017, pp. 1258-1260.

38 X. Hui, P. Sharma, and E. C. Kan, “Microwave stethoscope for heart sound by near-field coherent sensing,” *IEEE MTT-S Intl. Microwave Symp. (IMS)*, pages 365-368. IEEE, 2019.

39 X. Hui and E. C. Kan, “Seat integration of RF vital-sign monitoring,” *IEEE MTT-S International Microwave Biomedical Conference (IMBioC)*, pages 1-3, 2019.

40 Z. Zhang, P. Sharma, T. B. Conroy, V. Phongtankuel, and E. C. Kan, "Objective scoring of physiologically induced dyspnea by non-invasive RF sensors," *IEEE. Trans. Biomed. Engr*., vol. 69, no. 1, pp. 432-442, 2021.

41 P. Sharma, X. Hui, J. Zhou, T. B. Conroy, and E. C. Kan, "Wearable radio-frequency sensing of respiratory rate, respiratory volume, and heart rate," *NPJ Digit. Med*., vol. 3, p. 98, July 2020.

42 “CST Studio Suite 3D EM simulation and analysis software.” https://www.3ds.com/products-services/simulia/products/cst-studio-suite/

43 J. Zhou, T. B. Conroy, G. Xu and E. C. Kan, “Morphology transformation and content selection of near-field RF sensing by complex vector injection”, *IEEE J. Electromagn. RF Microw. Med. Biol.*, vol. 6, no. 4, pp. 555 – 565, Dec. 2022, doi: 10.1109/JERM.2022.3199615

44 J. Zhou, P. Sharma, X. Hui, and E. C. Kan, "A wireless wearable RF sensor for brumation study of chelonians," *IEEE J. Electromagn. RF Microw. Med. Biol*., vol. 5, no. 1, pp. 17-24, 2020.

45 R. E. Fields, "Evaluating compliance with FCC guidelines for human exposure to radiofrequency electromagnetic fields," *OET bulletin*, vol. *65, no.* 10, 1997.

46 U. Côté-Allard et al., "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Trans. Neural Syst*., vol. 27, no. 4, pp. 760-771, 2019.

47 A. Vaswani *et al.*, "Attention is all you need," *Adv. Neural. Inf.*, vol. 30, 2017.

48 X. Chen, Y. Li, R. Hu, X. Zhang, and X. Chen, "Hand gesture recognition based on surface electromyography using convolutional neural network with transfer learning method," *IEEE J. Biomed. Health. Inform.*, vol. 25, no. 4, pp. 1292-1304, 2020.

49 Z. Zhang, P. Sharma, J. Zhou, X. Hui, and E. C. Kan, "Furniture-integrated respiration sensors by notched transmission lines," *IEEE Sens. J.*, vol. 21, no. 4,

pp. 5303-5311, Feb. 2021.

50 F. Petitjean, A. Ketterlin, and P. Gançarski, "A global averaging method for dynamic time warping, with applications to clustering," *Pattern recognition*, vol. 44, no. 3, pp. 678-693, 2011.

51 R. W. Bohannon, "Muscle strength: Clinical and prognostic value of hand-grip dynamometry," *Curr. Opin. Clin. Nutr. Metab. Care*, vol. 18, no. 5, pp. 465-470, 2015.

52 N. Cooray, F. Andreotti, C. Lo, M. Symmonds, M. T. Hu, and M. De Vos, "Detection of REM sleep behaviour disorder by automated polysomnography analysis," *Clin. Neurophysiol.*, vol. 130, no. 4, pp. 505-514, 2019.

53 T. Isezaki *et al.*, "Sock-type wearable sensor for estimating lower leg muscle activity using distal EMG signals," *Sensors*, vol. 19, no. 8, p. 1954, 2019.

54 S. Zhang, X. Zhang, S. Cao, X. Gao, X. Chen, and P. Zhou, "Myoelectric pattern recognition based on muscle synergies for simultaneous control of dexterous finger movements," *IEEE Trans. Hum. Mach. Sys.*, vol. 47, no. 4, pp. 576-582, 2017.

CHAPTER 7

RADIOOCULOGRAM (ROG) FOR EYE MOVEMENT SENSING WITH EYE CLOSED

Introduction

Eye movement (EM) measurement can derive profuse information in emotion perception [1], neurodegenerative diseases [2], and monitoring of sleep and dream stages [3]. It can also facilitate human-computer interface (HCI) and virtual reality (VR) applications [4][5]. Current eye tracking systems with eyes open by camera-based methods can achieve high accuracy, though still have concerns of privacy, complexity, and occlusions [6][7]. Sensing EM with eyes shut under low ambient light can be even more difficult for cameras. EM sensing with eyes closed during sleep is important for the detection of rapid eye movement (REM), a sleep phase characterized by random rapid EM with an inclination of vivid dreaming. REM as an important sleep stage can be an indicator of health and cognitive performance, such as brain maturation [8], memory consolidation [9], and learning facilitation [10]. Existing methods for REM recording during sleep mainly used biopotential signals from electrooculogram (EOG) [11] and electroencephalography (EEG) [12], as parts of the clinical polysomnography (PSG) [13]. However, the electrode-based sensor can be limited by low user comfort and skin irritation as well as ambiguity and interferences due to skin potentials and leaky neural signals.

Here we propose radiooculogram (ROG), a novel EM sensor based on radio-frequency (RF) signals that can accurately and non-invasively monitor internal eye

muscle activities with eyes open or shut. We validated accurate measurement of EM frequencies and directions by a human study of 5 participants with selected longitudinal experiments. We further benchmarked ROG with synchronous EOG as the baseline comparison and physiological correlation. The main advantage of the proposed ROG system can be summarized as:

Improved user comfort. EOG and EEG measurements demand numerous electrodes around the eye region with stable electrical contact, which are inconvenient, uncomfortable, and prone to face motion interference. ROG can operate without direct skin contact.

Unmediated sensing of directional EM. While the biopotential-based sensors such as EOG, EEG, and electromyography (EMG) measure neural stimulus for muscle activity, ROG directly measures the muscle motion by coupling RF energy to deep internal muscles. EOG and ROG can be used together to derive the closed loop of stimulation and actuation.

Baseline for sleep REM detection. While camera-based methods are difficult to use for sleep REM, ROG has the flexibility to operate when eyes are open or closed without privacy concern. This work can formulate a validation baseline for future sleep REM monitoring.

Sensor setup and experiement protocol

Sensor Setup

ROG is based on the near-field coherent sensing (NCS) [14][15] of ultra-high frequency (UHF) RF signals to monitor the dielectric boundary change of internal

muscles during EM. As shown in Fig. 1(a), one ROG sensing unit consisted of a notched miniature coaxial RF cable, where the metal shield of the 1-inch middle part was removed to allow a small amount of RF energy leaking into the user's upper face region [16]. The ROG system integrated four sensing units attached to an eye mask around the eyes, as indicated in Fig. 1(b). The ROG RF transceiver was implemented by software-defined radios (SDR) to drive the notched sensors and to interface with the host computer through USB. Two National Instrument Ettus B210 were used, each of which had two transmitter/receiver (Tx/Rx) ports as shown in Fig. 1(c). The two SDRs were synchronized by an external local oscillator (LO, BG7TBL-GPSDO) with 10 MHz reference and 1 PPS (pulse per second) baseband synchronization. The experimental setup on a user's face was in Fig. 1(d) for ROG and in Fig. 1(e) for EOG. The ROG system can be alternatively implemented in wireless active [17] and passive [14]. units, although the present prototype is a wired system for convenient benchmarking.

In the near-filed region, the dielectric boundary change of associated eye muscles during EM would couple into the leaked RF energy, and hence affected the signals between Tx and Rx. Four sensors at different positions around the eyes provided more observation diversity to improve the amplitude and direction resolution. We adopted the multiple-input multiple-output (MIMO) strategy to explore N^2=16 coupling channels from N=4 sensing units to further enhance the spatial diversity [18].

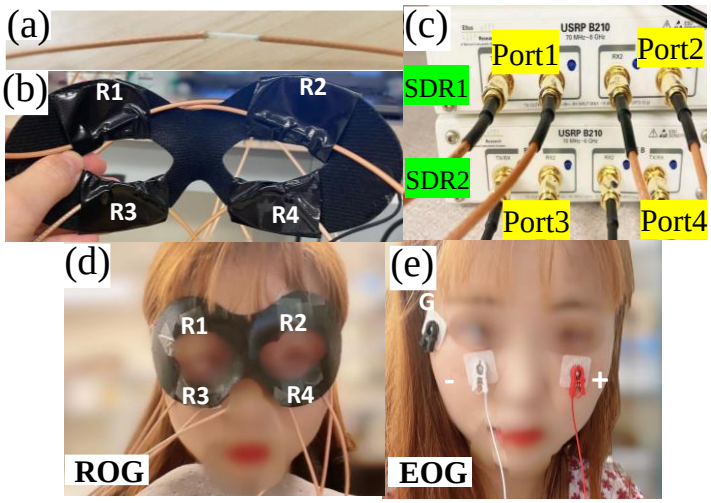


Fig. 1. The ROG system. (a) One ROG sensor unit by a notched transmission line; (b) Four ROG sensor units on a mask; (c) The SDR transceiver; (d) ROG on a participant's face; (e) EOG setup for baseline comparison.

The digital baseband of each Tx went through the digital-to-analog converter (DAC) and was then mixed with the carrier frequency f_{RF}, selected at 1 GHz. The RF power was less than −10 dBm or 0.1 mW, well under the safety limits set by Occupational Safety and Health Administration (OSHA). The RF signal leaked from the notched structure is coupled into internal muscle motion, received by Rx, and then demodulated and sampled by the analog-to-digital converter (ADC) to retrieve the baseband. The quadrature scheme was employed as the baseband tone f_{BB}. The multiple Tx channels utilized frequency-division multiple access (FDMA) by setting f_{BB} =10, 25, 40, and 125 kHz, respectively, for Tx1–Tx4. The system were configured as 4 self and 12 cross channels, which were all sampled at 10^6 samples per second (Sps), and further down-sampled to 500 Sps after demodulation.

Human Study Protocol

Two routines of human study on 5 volunteers were executed when eyes were

closed. The ROG signals were similar with open eyes, but with interference from blinking. Routine 1 was for EM frequency detection when the participant followed voice instructions and exercised EM with 10, 15, 20, 30, and 60 beats per minute (BPM). The eye exercise in each frequency had a duration of 30s with eyes moving left and right. Figs. 2(a)-(b) presented several examples of ROG (Tx3 – Rx3) and EOG waveforms. Participants were then instructed to move eyes in four directions in Routine 2. All directions had 2 versions of moving once and twice. Hence, we had 8 distinctive EMs, and each motion was performed in a time window of T_{win} =5s with around 24 repetitions. Unlike gaze localization with open eyes [19], the ground truth of EM direction and voluntary control of eyeball rotation were less precise when eyes were closed.

Benchmark with EOG

The reference EOG setup was by BIOPAC MP36R with the three EL513 electrodes around the eyes as + (under right eye), − (under left eye), and ground (left to left eye). ROG and EOG channels were synchronized in Labview and transferred to the host computer by USB. The same study protocol was performed on two participants with longitudinal iterations.

Signal processing

EM Frequency Estimation

For the EM frequency testing in Routine 1, the signal was first bandpass-filtered from 0.05Hz to 2Hz to remove the DC drift and high-frequency noises. Then

we utilized the moving average-crossing algorithm to first extract a moving-average curve in a given window length, and then label local maximum and minimal points [15]. The EM rate is estimated by counting the number of detected cycles over an epoch of 10s. One EM cycle includes moving eyeballs from left to right and then back to left. We have collected 16 quadrature channels from ROG, and each channel has amplitude and phase separately. We selected the channel with minimum covariance of the EM rate to output the final estimates. EOG was processed in a similar way, although there was only one channel in our setup.

EM Direction Estimation

For the EM direction testing in Routine 2, after obtaining 1D time waveforms from 32 channels, we first applied bandpass filtering (0.05 Hz to 10 Hz) and normalization. The waveforms were then segmented into motion-based windows of T_{win} =5s, each containing one instructed EM. We transformed the 1D waveforms to 2D spectrograms using continuous wavelet transform (CWT) by Morelet and Gaussian mother wavelets. Finally, the 2D image-like data was fed into the deep learning network as the classifier to differentiate all EM directions. We adopted vision transformer (ViT) [20], a deep learning model in natural language processing (NLP) and computer vision, for classification.

Results and analyses

EM frequency estimation

Fig. 2(c) shows the EM rate in beats per minute (BPM) calculated from EOG

(blue) and ROG (red) in comparison with the ground truth (green) from instruction in Routine 1. Fig. 2(d) shows the correlation of the EM rate from ROG (red markers) and EOG (blue markers) against the ground truth. In the left figure, both ROG and EOG achieved high correlation to the ground truth with Pearson coefficients denoted as r_{ROG} = 0.99 and r_{EOG} = 0.98. In the right figure, the Bland-Altman plot presents the agreement by the mean (*m*) and limits of agreement (LoA). The X axis is the average of the estimation and ground truth, and the Y axis is the difference. Both EOG and ROG achieved low *m* and narrow LoA. Note that *m* is positive for both sensors, which implies that the ground-truth EM rate is higher than the estimated results. Notice that the participant may not perfectly follow the instruction especially for very fast EM at 60 BPM. Fig. 2(e) further presents the correlation between ROG and EOG in a similar format. In comparison with Fig. 2(d), correlation between ROG and EOG is higher than that to the ground truth, indicating higher consistency between the two sensors. Table I summarizes all correlation and B&A statistics of the EM rate using only ROG across 5 participants.

We further compared the temporal correlation between ROG and EOG waveforms. We extracted the optimal time lag that can maximize the cross-correlation between the two waveforms. When the time lag = 0.032s, cross-correlation achieves the maximum value of 0.96, which indicates that the ROG waveform has a time delay following EOG events, as EOG detects the neural stimulation of EM and ROG detects the actual EM. Table II presents the statistics of time lag and correlation between EOG and ROG in two subjects during longitudinal tests, where $r_{R\&E}$ is the Pearson

coefficient. Table III presents the correlation and B&A statistics of EM rate estimation using ROG and EOG in comparison with the ground truth in the longitudinal tests, where both ROG and EOG remain highly accurate.

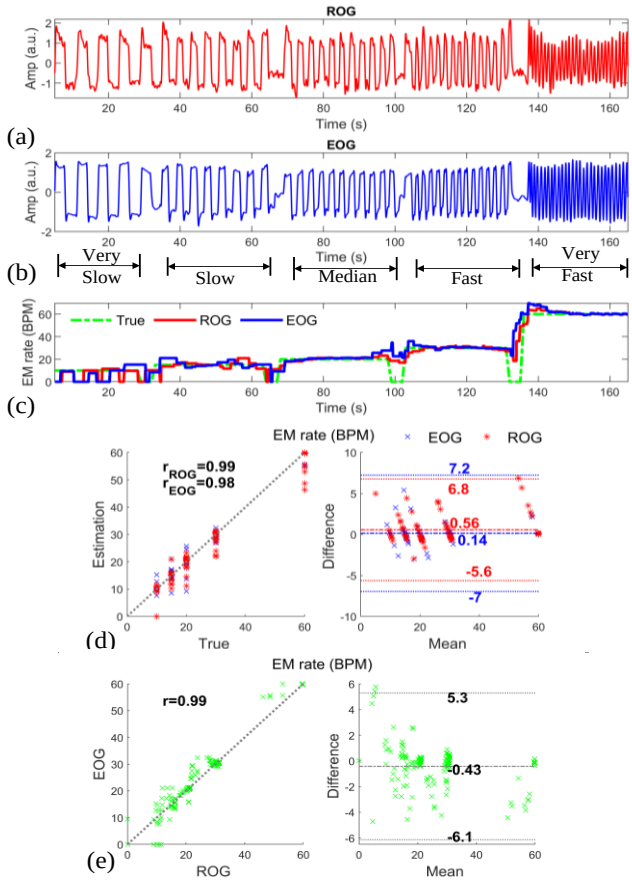


Fig. 2. EM frequency estimation. (a) ROG amplitude from Tx3- Rx3 and (b) EOG waveform samples. (c) EM rate in BPM. Correlation and agreement (d) between the ground truth and estimation from ROG and EOG, and (e) between ROG and EOG.

TABLE I. CORRELATION AND B&A STATISTICS OF THE EM RATE ESTIMATION USING ROG FOR EACH SUBJECT

SUBJECT NO.	RROG	M± Σ (BPM)
1	0.987	0.70 ± 2.90
2	0.986	0.75 ± 3.02
3	0.985	1.04 ± 3.21
4	0.982	1.00 ± 3.11
5	0.984	1.46 ± 3.26
MEAN	0.985	0.99 ± 3.10

TABLE II. TIME LAG AND CORRELATION BETWEEN ROG AND EOG

SUBJECT NO.	ITERATION NO.	TIME LAG (S)	MAX CORR.	$R_{R\&E}$	$M \pm \Sigma$ (BPM)
1	2	0.090	0.88	0.99	0.28 ± 2.24
2	2	0.054	0.97	0.95	−1.75 ± 6.02
2	3	0.052	0.92	0.99	0.31 ± 3.00
2	4	0.032	0.96	0.99	−0.43 ± 2.91
MEAN		0.057	0.93	0.98	−0.40 ± 3.54

TABLE III. CORRELATION AND B&A STATISTICS OF EM RATE ESTIMATION USING ROG AND EOG AGAINST THE GROUND TRUTH

SUBJECT NO.	ITER. NO.	R_{ROG}	$M \pm \Sigma$ (ROG)	R_{EOG}	$M \pm \Sigma$ (EOG)
1	2	0.985	0.11 ± 3.13	0.992	0.39 ± 2.26
2	2	0.979	0.95 ± 4.10	0.950	2.75 ± 5.68
2	3	0.970	1.43 ± 4.50	0.979	1.74 ± 4.07
2	4	0.986	0.56 ± 3.16	0.983	0.14 ± 3.61
MEAN		0.980	0.76 ± 3.72	0.976	1.26 ± 3.90

EM direction estimation

Routine 2 for EM direction estimation include 5 subjects with 947 samples of 8 classes of EM, namely 4 directions (up, down, right, and left) and 2 instances (once and twice). We built the training model within each participant and performed *k*-fold (k = 7) cross validation to estimate the mean accuracy for each participant. An overall accuracy was averaged on results from all participants. Fig. 3 shows the normalized confusion matrix by ViT. ROG can achieve high accuracy for distinguishing different directions. Note that the class 'D' (down) has relatively lower accuracy than other classes. It may be difficult for participants to follow the instruction to move eyeballs downwards in a consistent way with eyes closed. We also collected data from EOG in

the benchmark experiment using the same protocol and signal processing procedures. Accuracy by one EOG channel only achieved 57.3%.

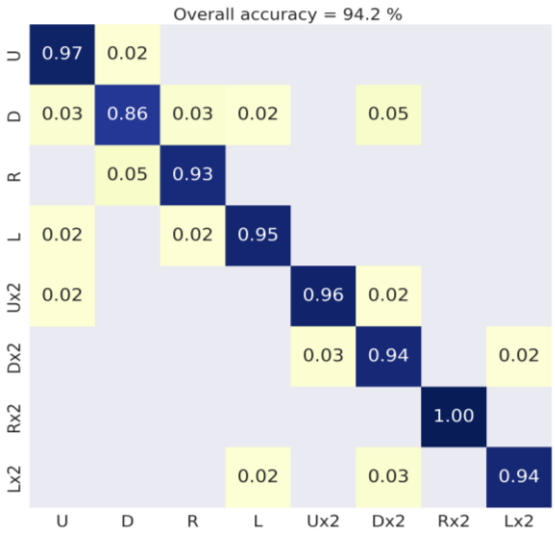


Fig. 3. The confusion matrix showing the overall accuracy of 94.2% for EM direction detection on all 5 subjects.

Conclusion

In this work, we present a new non-invasive and touchless radiooculogram (ROG) for EM monitoring with eyes closed. ROG can accurately detect the EM frequencies in a broad range and recognize different EM directions. In comparison with conventional EOG, ROG has high accuracy and improved user comfort without requiring direct skin contact. ROG can capture direct muscle actuation during EM with less ambiguity and interference. The consistent delay of ROG trailing EOG events indicates the lag of muscle actuation after the neural stimulation. In the future, ROG can be a promising alternative for sleep REM monitoring in clinical studies.

REFERENCES

1 Y. Wang, Z. Lv, and Y. Zheng, "Automatic emotion perception using eye movement information for E-healthcare systems," *Sensors,* vol. 18, no. 9, p. 2826, 2018.

2 T. J. Crawford *et al.*, "Inhibitory control of saccadic eye movements and cognitive impairment in Alzheimer's disease," *Biological Psychiatry,* vol. 57, no. 9, pp. 1052-1060, 2005.

3 S.-F. Liang *et al.*, "Development of an EOG-based automatic sleep-monitoring eye mask," *IEEE Transactions on Instrumentation and Measurement,* vol. 64, no. 11, pp. 2977-2985, 2015.

4 S. Z. Homayounfar *et al.*, "Multimodal smart eyewear for longitudinal eye movement tracking," *Matter,* vol. 3, no. 4, pp. 1275-1293, 2020.

5 A. Poole and L. J. Ball, "Eye tracking in HCI and usability research," in *Encyclopedia of Human Computer Interaction*: IGI Global, 2006, pp. 211-219.

6 R. A. Naqvi, M. Arsalan, G. Batchuluun, H. S. Yoon, and K. R. Park, "Deep learning-based gaze detection system for automobile drivers using a NIR camera sensor," *Sensors,* vol. 18, no. 2, p. 456, 2018.

7 N. Valliappan *et al.*, "Accelerating eye movement research via accurate and affordable smartphone eye tracking," *Nature Communications,* vol. 11, no. 1, pp. 1-12, 2020.

8 M. S. Knoop, E. R. de Groot, and J. Dudink, "Current ideas about the roles of rapid eye movement and non–rapid eye movement sleep in brain development," *Acta Paediatrica,* vol. 110, no. 1, pp. 36-44, 2021.

9 S. Diekelmann and J. Born, "The memory function of sleep," *Nature Reviews Neuroscience,* vol. 11, no. 2, pp. 114-126, 2010.

10 M. Tamaki *et al.*, "Complementary contributions of non-REM and REM sleep to visual learning," *Nature Neuroscience,* vol. 23, no. 9, pp. 1150-1156, 2020.

11 A. Boukadoum and P. Ktonas, "EOG-Based Recording and Automated Detection of Sleep Rapid Eye Movements: A Critical Review, and Some Recommendations," *Psychophysiology,* vol. 23, no. 5, pp. 598-611, 1986.

12 M. A. Cruz-Aguilar, I. Ramírez-Salado, M. Hernández-González, M. A. Guevara, and J. M. Del Río, "Melatonin effects on EEG activity during non-rapid eye movement sleep in mild-to-moderate Alzheimer´ s disease: a pilot study,"

International Journal of Neuroscience, vol. 131, no. 6, pp. 580-590, 2021.

13 N. Cooray, F. Andreotti, C. Lo, M. Symmonds, M. T. Hu, and M. De Vos, "Detection of REM sleep behaviour disorder by automated polysomnography analysis," *Clinical Neurophysiology,* vol. 130, no. 4, pp. 505-514, 2019.

14 X. Hui and E. C. Kan, "Monitoring vital signs over multiplexed radio by near-field coherent sensing," *Nat. Electron.*, vol. 1, no. 1, pp. 74-78, 2018,

15 P. Sharma, X. Hui, J. Zhou, T. B. Conroy, and E. C. Kan, "Wearable radio-frequency sensing of respiratory rate, respiratory volume, and heart rate," *NPJ Digital Medicine*, vol. 3, no. 1, pp. 1-10, 2020.

16 Z. Zhang, P. Sharma, J. Zhou, X. Hui, and E. C. Kan, "Furniture-integrated respiration sensors by notched transmission lines," *IEEE Sensors Journal,* vol. 21, no. 4, pp. 5303-5311, 2020.

17 X. Hui, J. Zhou, P. Sharma, T. B. Conroy, Z. Zhang and E. C. Kan, "Wearable RF near-field cough monitoring by frequency-time deep learning", *IEEE Trans. Biomed. Circuits & Sys*, vol. 15, no. 4, pp. 756 – 764, 2021

18 X. Hui, T. B. Conroy, and E. C. Kan, "Multi-point near-field RF sensing of blood pressures and heartbeat dynamics," *IEEE Access*, vol. 8, pp. 89935-89945, 2020.

19 Y. Wang, X. Ding, G. Yuan, and X. Fu, "Dual-cameras-based driver's eye gaze tracking system with nonlinear gaze point refinement," *Sensors*, vol. 22, no. 6, p. 2326, 2022.

20 A. Dosovitskiy *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

CHAPTER 8

CONCLUSION AND FUTURE WORK

Conclusion of my contribution

My present research interests focus on non-invasive sensing of physiological signals including respiratory efforts, muscle activities, heartbeat dynamics, and tissue properties using RF sensor. I established smart healthcare systems powered by machine learning that can be implemented as wearables, or invisibly integrated to the furniture to enable many medical applications, especially for remote continuous diagnosis. The summary of major contributions is as follows:

1. I presented a new respiration sensor integrated into a bed or a chair by modifying a radio-frequency (RF) coaxial cable structure with a designed notch. Non-invasive respiration sensors integrated into furniture can be invisible to the user and greatly enhance comfort and convenience to facilitate many applications.
2. I provided a new way to generate the objective score for physiologically induced dyspnea, using a comfortable and continuous respiratory sensor and an established ML model which can simultaneously consider multiple factors with different importance weighting. As far as we know, this is the first study to examine the association between dyspnea sensation and continuous respiratory metrics that account for changes in respiratory behavior over a period of time under exertion and increased airway resistance. The method can potentially formulate a baseline for clinical

dyspnea assessment and help caregivers track dyspnea continuously, especially for patients who cannot report themselves.

3. I developed a bed-integrated RF sensor can covertly and reliably detect and predict apneic events. Respiratory disturbances during sleep are a prevalent health condition that affects a large adult population. Predictive warning of the sleep disorders in advance can intervene serious apnea, especially for infants, servicemen, and patients with chronic conditions.
4. I validated the viability to use my objective dyspnea scoring for clinical dyspnea assessment on COVID patients. The proposed system can help the identification of dyspneic exacerbation in conditions such as COVID, leading to early intervention and possibly improving their outcome. This approach can be potentially applied to other pulmonary disorders such as asthma, emphysema, and pneumonia.

I also developed novel muscle sensing technique for biomedical application including muscle assessment for Parkinson's disease and human computer interface (HCI) including gesture recognition and biometric authentication.

1. I proposed a novel radiomyography (RMG) for continuous muscle actuation sensing that can be wearable or touchless, capturing both superficial and deep muscle groups. I verified RMG experimentally by a forearm wearable sensor for detailed hand gesture recognition. RMG can be used with synchronous EMG to derive stimulation-actuation waveforms for many potential applications in kinesiology, physiotherapy, rehabilitation, and human-machine interface.

2. I presented radiooculogram (ROG), a novel sensor for non-invasive eye movement (EM) monitoring with eyes closed. I have experimentally demonstrated accurate measurements of EM frequency and directions for 5 participants and benchmarked ROG with electrooculogram (EOG). Compared with biopotential-based sensors, ROG has higher user comfort due to touchless operation and can capture direct muscle activity even in deep tissues. This work on voluntary EM sensing can serve as the baseline implementation for eventual sleep rapid EM monitoring.

Future work

Apart from my Ph.D. research in the last 4 years with substantial supporting evidences, I will introduce research work that can be extended in the future.

Sensing Vital Signs and Internal Tissues

Promotion of remote healthcare

As evident from the COVID-19 pandemic, telemedicine has seen great promises to transform healthcare delivery. The COVID-19 pandemic has driven rapid growth in telemedicine use for urgent care and primary care far beyond previous baseline periods. In the future, telemedicine has the potential to further become a standard service offered across all primary care settings. Telemedicine can dramatically change the interaction between consumers and clinicians, but can at the same time improve patient engagement and experience. While telemedicine has the potential to drive significant values in many areas, nevertheless, there remain

significant challenges to developing and scaling the virtual-care platform, such as lack of physical and digital infrastructure, limited choices of at-home diagnostic sensors with acceptable reliability, and a suboptimal user experience.

My present research focuses on developing the non-invasive platform of vital-sign monitoring that can hopefully address these challenges. The proposed radio-frequency (RF) sensor can be integrated into an apparel, a bed, or a chair, hidden behind layers of fabrics. The senor can be cost-effective, compact, stable and compatible to various digital wireless protocols. It can further interact with various accessories such as body-electrode based sensors, cameras, accelerometers, and smart phones. Integrating our sensing hardware with telemedicine software can provide physicians with real-time continuous vital-sign data, while maintaining high user comfort and avoiding unnecessary public exposure. The sensing technology I plan to develop can potentially help virtual healthcare gain momentum and become a core component in the overall clinical infrastructure, as well as to improve early diagnosis and proper choices of specific therapy.

Preventive and chronic care delivery

Today, chronic disease management has been, and will continue to be, one of the biggest burdens for medical care in terms of cost and patience satisfaction. Nearly half of all adults in the US have a chronic disease, almost 33% of the population is living with more than one chronic diseases, which eventually are responsible for seven out of 10 deaths in the U.S., or approximately 1.7 million Americans each year. Heart diseases, chronic obstructive pulmonary diseases (COPD), and strokes are among the

top chronic diseases with high prevalence. Identifying patients at risk and getting access to useful data with least privacy invasion can be crucial for preventive care towards these chronic diseases.

My research on non-invasive and continuous vital-sign sensing can potentially provide an effective solution and transform how preventive care will be delivered in the future. By implementing the proposed sensor system, remote physicians and care providers can access the full record of the patient's vital signs or examine the analysis processed by artificial intelligence, surely under the appropriate protection of security and privacy. Obtaining accurate readings from unbiased and objective data, in addition to the patient self report, can help develop an improved strategy for proper care. The wearable sensors can track respiratory patterns, heart rates and blood pressures, which are the significant indicators for these chronic diseases. Mobile app can be further integrated with the hardware to keep long-term and real-time recording data and provide real-time feedback for early warning and diagnosis, which are the keys to successful preventive care. Another aspect of my research is to provide automated prognosis with high accuracy, leveraging data-driven decision making. After tracking and recording long-term continuous vital signs associated with chronic cardiopulmonary diseases, pre-warning for the upcoming symptoms can be provided to ameliorate disease deterioration. Sophisticated human machine interfaces with convenience and ease of use can further enhance patient engagement for compliance to medical treatment.

For instance, my current research develops a new approach to invisibly and precisely identify prevalent sleep disorders, including central sleep apnea (CSA) and

obstructive sleep apnea (OSA). Sleep apnea is very common with more than 3 million cases per year in US. Current sleep disorder symptoms are mostly scored by human operators during sleep studies, which is time-consuming. Real-time identification of the sleep disorders with predicative warning in a few minutes in advance can be extremely beneficial as episodes of serious apnea during sleep can be very dangerous, especially for infants and senior citizens.

Diagnosis powered by artificial intelligence

Another aspect of my research is to utilize machine-learning (ML) algorithms that can take the physiological data from our sensors to provide holistic diagnostics and to assist therapeutics to people at home or clinics. My previous work has developed new hardware-software co-design systems that can continuously track the cardiac and respiratory waveforms for accurate symptomatic evaluation through conventional signal processing and ML. The method can formulate a baseline for assessment of wellness and disorders, and provide caregivers with prognosis information, especially for patients who cannot communicate or cooperate themselves.

The current practice of diagnosis of cardiopulmonary disorders, such as asthma, cardiac arrest, COVID-19, dyspnea, and COPD, heavily replies on physical examinations by experienced doctors or self-reports by patients, which can be costly for doctor's time, inconvenient during patient rest time, as well as inaccurate and subjective from individual variations. My goal is to develop cost-effective ML-based diagnosis platforms that can complement or substitute visits to doctor's office with improved accuracy and quality of service to patients. The most common problem for

ML-based diagnosis today is the lack of rigorous data and continuous evaluation. Many studies have not used real patient data, or do not have the continuous data from patient monitoring. Robust, real-world studies are of high priority to train and validate the ML algorithm. I believe the use of my hardware-software co-design system in cardiopulmonary disorders will potentially transform the diagnostic system to become available to more people regardless of their financial means and accessible infrastructure.

Internet of beings (IoB)

The Internet of Things (IoT) has been regarded as the emerging technology sector where all things around us can be digitally connected with unique item-level ID. At the same time, integration of the man-made items and the living ecosystem now becomes possible to achieve a long-term sustainable future, where IoT can be morphed into the Internet of Beings (IoB) to bring forth broader impacts to overall wellness and happiness. The emerging biotechnology will herald an exciting and promising new era, in which biology and technology will be finally merged into one synthetic system. Our sensing technology has been able to measure vital signs in human as well as animals without having to shave, sedate, or force for basic veterinary care. The sensor can be implemented as a passive tag, or integrated into the habitat and feeder. Our technology can improve our understanding of the living beings, and hence bring forth benefits, instead of harms, to the broad ecosystem in the long run.

Microwave imaging of internal body organs

Imaging the interior body is essential for disease diagnosis and precision

treatment in clinical practice. Current X-ray radiography, MRI, and ultrasound usually require bulky devices and provide snapshots. In the future, our RF system has the potential to revolute 3D internal imaging by massive multiple-input-multiple-output (MIMO) sensors. RF signals can permeate through the body and reveal internal dielectric boundaries with the least concern for harmful dosage. The RF imaging system can be highly economical in comparison with existing devices for medical imaging. Wearable RF devices can be developed in the future so that continuous disease diagnosis and precision medicine can be realized remotely. Present RF imaging for pulmonary edema and breast tumors suffers insufficient resolution and still requires a lot of research works to serve as a useful diagnostic tool.

Muscle sensing by Radiomyography (RMG) and its application in biomedical research

Continuous monitoring of skeletal muscle activities has significant medical and commercial applications, including detection of muscle fatigue and injury, diagnosis of neuromuscular disorders, assessment for physical training and rehabilitation, human-computer interface (HCI), and robotic control. Conventional electromyography (EMG) measures the neural activity during muscle contraction, but lacks explicit quantification of the actual contraction. Mechanomyography (MMG) and accelerometers only measure body surface motion, while ultrasound, CT-scan and MRI are restricted to in-clinic snapshots. For the first time, I proposed radiomyography (RMG), a novel muscle sensor that can non-invasively and continuously capture muscle contraction in various layers [5]. RMG uses MIMO

near-field coherent sensing (NCS) signals to measure the dielectric property change and boundary movement of nearby muscle groups. As shown in Fig. 2, I verified RMG experimentally on a forearm wearable sensor for detailed hand gesture recognition, and further demonstrated monitoring of eye and leg muscles with high accuracy of eye movement and posture tracking. In the following, I will present several directions worth of future endeavor.

Study of Parkinson's disease

Parkinson's disease (PD) is a brain disorder that causes unintended or uncontrollable body movements, such as shaking, stiffness, and difficulty with balance and coordination. Many studies have used EMG to detect the abnormalities in electrical signals produced by muscles such as increased muscle tone, abnormal posture, gait and tremor. Compared with EMG measuring neural stimulation, RMG can directly detect the actual muscle change. The two sensors can be combined for a fuller physiological interpretation. In my previous work [], consistent observation of the RMG delay from synchronized EMG indicated the non-trained muscle actuation without the participation of proprioceptive neurons which can be transferable to PD detection.

In the future, I believe RMG and EMG fusion can be used as a powerful combination to diagnosis and prognosis of PD, including 1) gait abnormality, where RMG and EMG can be combined to measure leg muscle activities to detect abnormal gaits, and 2) Tremor detection, where wearable RMG and EMG can be attached to wrist and arm to monitor the resting tremor with differentiation of tremor pattern and

severity. The signal correlation between RMG and EMG can also help investigate the control of muscle actuation by the central nervous system (CNS).

Sleep rapid eye movement (REM) monitoring

Eye movement sensing with eyes closed during sleep is important for the detection of rapid eye movement (REM), a sleep phase characterized by random rapid eye movement with an inclination of vivid dreaming. REM as an important sleep stage can be an indicator of health and cognitive performance, such as brain maturation, memory consolidation, and learning facilitation. Sensing eye movement with eyes shut under low ambient light can be difficult for cameras. Biopotential signals from electrooculogram (EOG) and electroencephalography (EEG) are other alternatives. However, these electrode-based sensor can be limited by low user comfort and skin irritation as well as ambiguity and interferences due to skin potentials and leaky neural signals. I demonstrated a new radiooculogram (ROG) for non-invasive eye movement monitoring with eyes open or closed. In my future work, ROG will be used to further study sleep REM in Sleep Clinic.

Assessment for muscle fatigue, pain, and rehabilitation

RMG can directly and non-invasively measure the muscle motion by coupling RF energy to superficial and deep internal muscles. In the future, my research can be extended to a wide range of studies for muscle function assessment. For example, fatigue is a common non-specific symptom experienced by many people related to difficulty in performing voluntary tasks. Currently, muscle fatigue is mainly evaluated by EMG with limited capability. I will explore new methods using RMG along with

EMG to have a more comprehensive interpretation of dynamic fatigue in muscle contraction. The study will be from acute short-term muscle fatigue stimulated on healthy subjects to long-term chronic fatigue syndrome (CFS) experienced by patients. I will develop a novel technology in rehabilitation research by providing quantifiable information on the myoelectric output of a muscle. Physical rehabilitation aims to restore functional ability from pain or disability in muscle or nerve damage, where RMG and EMG can provide quantifiable and continuous information on the muscle contraction.

Closed-loop Electrical muscle stimulation (EMS) control

A closed-loop control of EMS is another possible future application. EMS has long been employed to either supplement or substitute voluntary muscle stimulation in many settings of rehabilitation and electroceuticals. However, inadequate EMS due to personal and daily differences can cause confusion of antagonistic and synergistic coordination of the muscle groups, and even induce serious spasm. For a more precise control on EMS, I will develop RMG to give feedback on actual muscle actuation to EMS with higher adaptability to personal and conditional variations.

Intelligent Human Computer Interface (HCI) based on muscle tracking

Gesture interface

Gestures provide effective non-verbal communication and can help deliver intuitive interactions to machines. There has been high interest to develop new technologies in gesture recognition to facilitate HCI. Gesture interface can be applied

to various applications including 3D virtual reality/augmented reality (VR/AR) control; sign-language detection for hard of hearing persons; assistive robotic control; communication in hostile environment such as fire and covert operations; non-contact navigation and infotainment in smart cars. In my future work, I will continue work on RMG for an effective method of gesture interface to provide: 1) High recognition accuracy for a large diversity of gestures; 2) Robustness against subject difference, background noise, and scenario variations; 3) Free movement without obstruction and discomfort; 4) Convenience to configure and deploy; 5) Low computing and network loads, and real-time response with ultra-low latency.

Eye tracking

I have demonstrated the capability using ROG, a novel sensor for non-invasive and accurate eye movement monitoring. Current eye tracking systems with eyes open by camera-based methods can achieve high accuracy, though still have concerns of privacy, complexity, and occlusions. ROG has higher user comfort due to touchless operation and can capture direct muscle activity even in deep tissues. It can facilitate various HCI and VR/AR applications. In the future, I will explore solutions using ROG for device interaction, including car navigation, interface control replacing the mouse, and VR/AR training. ROG can also be potentially used for analyzing driver's attention, cybersickness, and emotion extraction.

Biometric authentication based on muscle recognition

Handwriting recognition is an essential form of verification system for security and privacy. However, handwriting can often be faked and matching handwriting is

not always accurate. I will explore the new dynamic air signature system by writing in the air and recognizing the unique muscle behavior pattern using wearable forearm RMG. Compared with traditional handwriting, this new technique requires dynamic recording of muscle activities which can be an important next-generation marker that is mostly immune to presentation and eavesdropping attacks.

PUBLICATION LIST

1 **Z. Zhang**, and E. C. Kan, " Novel muscle monitoring by radiomyography (RMG) and its application to hand gesture recognition ", submitted to *IEEE Sens. J.*

2 **Z. Zhang,** J. Zhou, T. B. Conroy, S. Chung, J. Choi, P. Chau, D. B. Green,A. C. Krieger and E. C. Kan, " Objective dyspnea evaluation on COVID-19 patients learning from exertion-induced dyspnea scores," submitted to *Sensors*

3 **Z. Zhang,** T. B. Conroy, A. C. Krieger and E. C. Kan, " Identification and prediction of sleep disorder by covert bed integrated RF sensors," *IEEE. Trans. Biomed. Engr.*, 2022, doi: 10.1109/TBME.2022.3212619.

4 **Z. Zhang**, and E. C. Kan, " Radiooculogram (ROG) for eye movement sensing with eyes closed", in *21st IEEE Conf. on Sensors*, Dallas, TX, Oct. 30 – Nov. 2, 2022

5 **Z. Zhang**, G. Xu, and E. C. Kan, "Outlooks for RFID-based autonomous retails and factories", *IEEE J. Radio Frequency Identification (RFID),* 2022, doi: 10.1109/JRFID.2022.3211474

6 **Z. Zhang**, G. Xu, and E. C. Kan, "3D geometry recognition by RFID Box based on deep learning", in *16th Intl. Conf. on RFID,* Las Vegas, NV, May 16 – 19, 2022.

7 **Z. Zhang**, P. Sharma, T. B. Conroy, V. Phongtankuel, and E. C. Kan, "Objective scoring of physiologically induced dyspnea by non-invasive RF sensors," *IEEE. Trans. Biomed. Engr.*, vol. 69, no. 1, pp. 432-442, 2021.

8 **Z. Zhang,** P. Sharma, J. Zhou, X. Hui and E. C. Kan, "Furniture-integrated respiration sensors by notched transmission lines," *IEEE Sens. J.*, vol. 21, no. 4, pp. 5303-5311, 2021

9 P. Sharma, **Z. Zhang**, T. B. Conroy, X. Hui, and E. C. Kan, "Attention Detection by Heartbeat and Respiratory Features from Radio-Frequency Sensor," *Sensors*, vol. 22, no. 20, p. 8047, 2022.

10 X. Hui, J. Zhou, P. Sharma, T. B. Conroy, **Z. Zhang** and E. C. Kan, "Wearable RF near-field cough monitoring by frequency-time deep learning", *IEEE Trans. Biomed. Circuits & Sys*, vol. 15, no. 4, pp. 756 – 764, 2021

11 **Z. Zhang**, et al., "Wideband and continuously-tunable fractional photonic Hilbert transformer based on a single high- birefringence planar Bragg grating," *Opt. Express*, vol. 26, pp. 20450-20458, 2018.

12 **Z. Zhang**, et al., "Design of a broadband achromatic dielectric meta-lens for linear polarization in the near-infrared spectrum," *OSA Contin.*, vol. 1, pp. 882-890, 2018.

13 **Z. Zhang**, et al., "Micro-machining for TE/TM mode phase matching in high-birefringence planar waveguide and implementation in continuously-tunable fractional Hilbert transform," *Intl. Photonics & Optoelectronics Mtg., OSA Tech. Dig.*, OT4A.2, 2018.

14 H. Sun, W. Zhou, **Z. Zhang** and Z. Wan. "A MEMS variable optical attenuator with ultra-low wavelength-dependent loss and polarization-dependent loss," *Micromachines*, vol. 9, no. 12, p. 632, 2018.