Wearable RF Near-Field Cough Monitoring by Frequency-Time Deep Learning

Xiaonan Hui, *Student Member, IEEE*, Jianlin Zhou, Pragya Sharma, *Student Member, IEEE*, Thomas B. Conroy, Zijing Zhang and Edwin C. Kan, *Senior Member, IEEE*

***Abstract*— Coughing is a common symptom for many respiratory disorders, and can spread droplets of various sizes containing bacterial and viral pathogens. Mild coughs are usually overlooked in the early stage, not only because they are barely noticeable by the person and the people around, but also because the present recording method is not comfortable, private or reliable for long-term monitoring. In this paper, a wearable radio-frequency (RF) sensor is presented to recognize the mild cough signal directly from the local trachea vibration characteristics, and can isolate interferences from nearby people. The sensor operates at the ultra-high-frequency band, and can couple the RF energy to the upper respiratory track by the near field of the sensing antenna. The retrieved tissue vibration caused by the cough airflow burst can then be analyzed by a convolutional neural network trained on the frequency-time spectra. The sensing antenna design is analyzed for performance improvement. During the human study of 5 participants over 100 minutes of prescribed routines, the overall recognition ratio is above 90% and the false positive ratio during other routines is below 2.09%.**

***Index Terms*— Biomedical signal processing; machine learning; microwave sensors; near-field radiation pattern; wearable devices.**

I. Introduction

PATIENTS with respiratory disorders caused by infectious pathogens and tissue dysfunctions often manifest coughing as a typical symptom for diagnosis and prognosis [1]. To evaluate the degree of disease severity, physicians most often rely on the symptomatic self-report from patients and limited observation in the clinic. Other indicators, such as blood or urine examination and liver function tests, can offer additional references to caregivers, but have delays and thus less contributive to the progress evaluation. The advanced medical imaging systems such as magnetic resonance imaging (MRI) [2] and X-ray computed tomography (CT) [3] can provide rich information of tissue lesion, but are difficult to be the long-term monitoring methods for cost [4] and radiation exposure concerns [5]. The auscultation stethoscope [6], [7] is a standard practice for the respiratory illness, which is a powerful clinical instrument operated by the well-trained doctors. However, the requirement of acoustic impedance matching with direct skin contact limits the system integration as a wearable device for long-term monitoring. The ultrasound system also suffers similar constraints on the body probe [8]. The passive acoustic sensor can be vulnerable to the ambient noise and coughing from nearby people, which increases the recognition algorithm complexity and decreases the overall system reliability. Additionally, using a microphone [9]–[11] for long-term recording will also involve privacy concerns. The inertial measurement unit (IMU) sensor [12], [13] can also detect the body motion during coughing, but a stable body contact is required to couple to its proof mass, while other body movements can interference with the targeted cough signals. Moreover, IMU detects the body surface motion instead of the trachea directly.

Due to the advantage of touchless operation, the radio frequency (RF), microwave and millimeter-wave systems were applied to sensing respiration and phonation from various previous works, including adapting the implementation of Doppler radar [14]–[18], frequency modulated continuous wave (FMCW) radar [19], [20], ultra-wide band (UWB) radar [21], [22] and radio-frequency identification (RFID) [23]. In order to increase the sensitivity to mild coughs, some of the approaches employed very short wavelength in the millimeter-wave and sub-Terahertz bands [24]–[26]. The off-body reader design and high power consumption of the radar-like system raise further concerns for long-term sensing. The reflection of the far-field electromagnetic (EM) wave from the front chest will either have difficulty in tracking the specific body location or suffer from motion of other body parts and in the ambient. Hence, a convenient, comfortable, reliable and non-intrusive wearable sensor system [27]–[30] that can perform long-term and nearly continuous cough monitoring is still missing.

By applying the near-field coherent sensing (NCS) principle [23], [31], we propose to couple efficiently a small amount of EM energy in the ultra-high-frequency (UHF) band to the trachea region to accomplish cough sensing. Because NCS will only detect the local vibration within the near field of the sensing antenna [32], the interferences from other motion, including coughing from nearby persons, will be minimized. As

This project is supported by NIH R21 DA049566-01 and NSF RAPID 2033838.

X. Hui, J. Zhou, P. Sharma, T. B. Conroy, Z. Zhang and E. C. Kan are with School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA. The corresponding author is X. Hui, E-mail: xh273@cornell.edu.

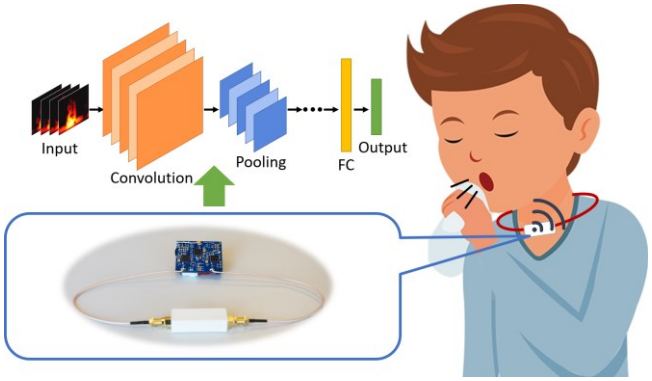


Fig. 1. The wearable RF near-field cough monitoring system.

shown in Fig. 1, the sensor is designed as a wearable prototype in the position of a choker necklace to capture the tissue vibration related to coughing. The modulated sensing signal will be transmitted to the console device such as a smartphone or a computer through Bluetooth for further signal processing and recognition based on a convolutional neural network (CNN). The transmitter (Tx) power is below –10 dBm or 0.1 mW within a narrow band, much lower than the other radar-based systems in terms of both power and spectral cost, which enables the ultra-low-power wearable system with long operation time. In comparison with the Tx power in the cell phone around 30 dBm and WiFi around 20 dBm, the NCS device only transmits less than one-thousandth of RF power to relieve any safety concerns.

In Section II, the NCS principle will be briefly introduced and analyzed with the EM simulation in the trachea region. In Section III, the system design and antenna placement analysis will be discussed. Section IV illustrates the experimental setup and results, followed by the cough signal features with the corresponding recognition by frequency-time deep learning. Finally, the discussion and conclusion are presented in Section V.

II. The NCS Principle and EM Simulation

The NCS operation [23], [31] can be briefly introduced as the following steps: When the NCS sensing antenna (denoted as Tx) is placed close to the skin within the antenna near-field region, a significant amount of the RF energy will be coupled into the tissue. The tissue with different permittivity from the surrounding body will backscatter some part of the impinging RF signal. The motion and vibration caused by the vital signs and other internal tissue movement such as heartbeats [33], [34], wrist and femoral pulses [31], [34], respiration [36], [37] and vocal cord vibration [23] will modulate the phase of the backscattered signal by permittivity boundary movement. Another part of the Tx signal is not modulated and serves as the reference signal. The modulated backscattered signal and non-modulated reference signal are from the same source with different phases and are combined at the receiver (Rx) antenna to cause an amplitude modulation. Thus, the internal tissue vibration can be demodulated accordingly at Rx, analogous to the interferometer model. To adapt to various applications, the NCS system can be implemented by different configurations, such as the passive RFID system [23], [31], [33], the active

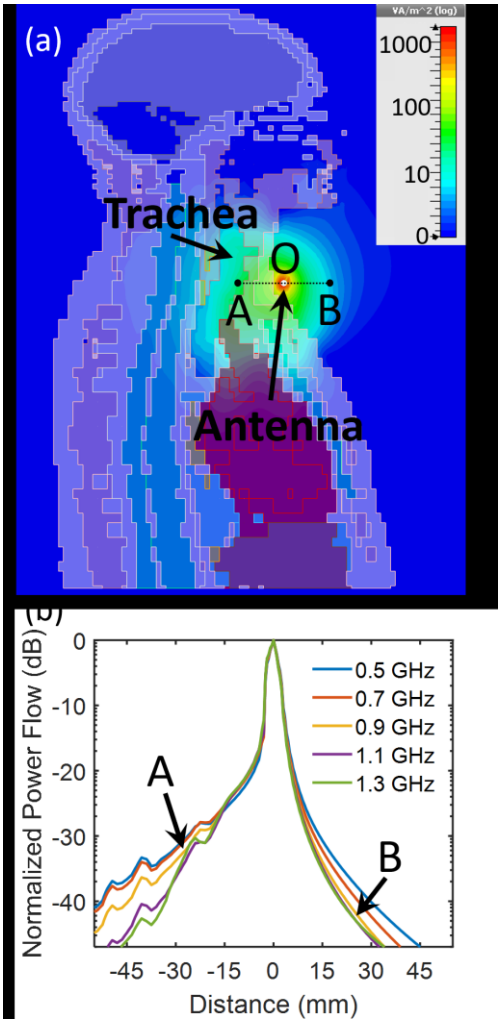


Fig. 2. (a) The CST simulation when the NCS Tx antenna is placed close to the trachea area. (b) The normalized power flow along AB.

antenna-pair transceiver system [32], [34], [36]–[38], and the furniture-integrated system [39], [40]. For cough sensing, it is important to couple the EM energy efficiently to the trachea area. The choker necklace placement of the sensing antenna can be effective as shown by the EM simulation in Fig. 2(a). The human EM model is built in CST Microwave Studio [41], and a dipole antenna is placed at Point O. The color contours show the power flow of the EM field on the sagittal cross section of the human model, when the antenna is fed by a 1-W excitation signal at 900 MHz. We used 1 W in the simulation mainly for the convenience of normalization. From the power flow contour, we can see that a significant amount of the EM energy is coupled into the desirable body parts due to the antenna near-field effect. As the near-field coupling decays much faster than the propagation far field, the EM energy coupled to the brain and eye regions is at least 40 dB lower. Furthermore, because of the high permittivity of the human tissue, the power flow in the body is also higher than that in the air at the same distance to the antenna, which are indicated by Points A and B in Fig. 2(a). The two monitoring points of A and B are both at the distance of 28 mm from the dipole antenna at Point O, with Point A approximately at the permittivity boundary of the trachea. The power flows at Points A and B are

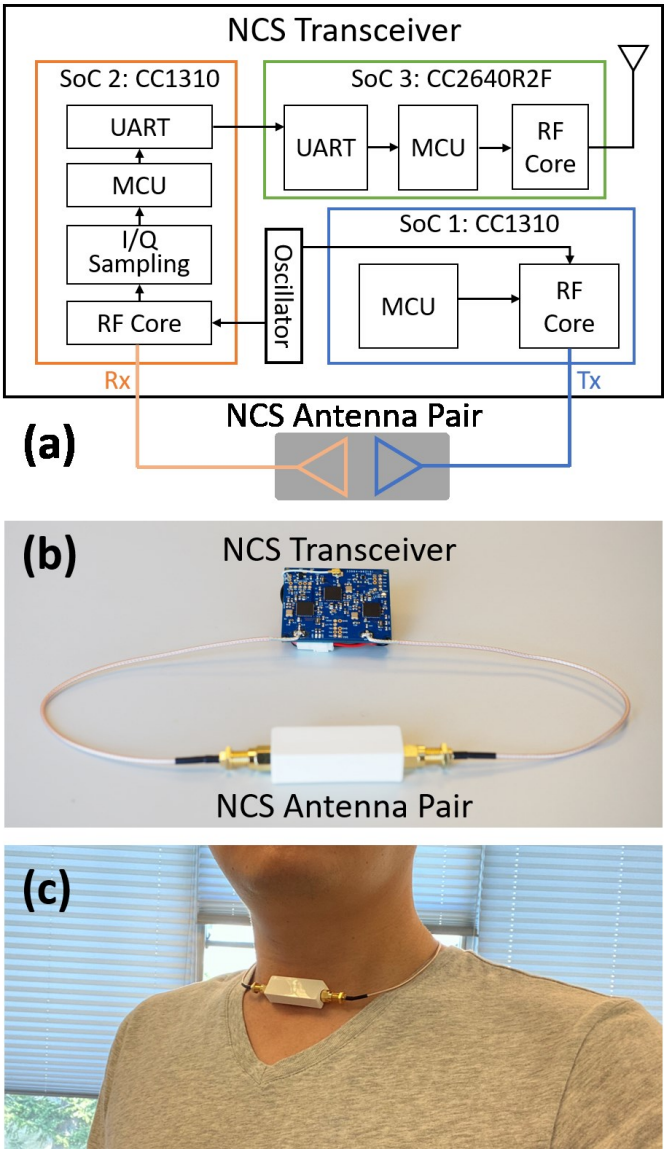


Fig. 3. (a) The schematic of the NCS cough sensor. (b) The PCB prototype of the sensor. (c) The sensor deployed as a choker necklace.

39.3 W/m^2 and 2.48 W/m^2, respectively, where Point A is 12 dB stronger than Point B. The power flow distribution along the line AB at different frequencies from 0.5 GHz to 1.3 GHz with a 0.2 GHz step is shown in Fig. 2(b). Each curve is normalized to its own peak value to eliminate the performance difference of the antenna within the frequency band. From Fig. 2(b), the power flow in the air decays faster than that in the human tissue and the lower frequency signal has better penetration into the body. With more power coupled into the sensing area, the signal strength of the backscattered signal from the permittivity boundary will be stronger, and eventually increase the sensor signal-to-noise ratio (SNR). We can also see that when a dipole antenna was deployed, whose radiation pattern is symmetric in the free space, the sensing power is concentrated within the area of the trachea and a relatively small area out of the body. Hence, other body movement and ambient noises will only cause limited interference to the target sensing signal. When a receiver antenna is also deployed close to the Point O, the interference signals can be further isolated.

III. System Design

Fig. 3(a) shows the schematic of the NCS cough sensor, and a picture of the sensor is in Fig. 3(b). The NCS transceiver is designed with 3 systems-on-a-chip (SoC), including 2 Texas Instrument CC1310s and 1 CC2640R2F. SoC 1: CC1310 in Fig. 3(a) is configured as the RF Tx, where its MCU controls the RF core to generate the sensing signal that is fed to the Tx antenna of the NCS antenna pair. The modulated NCS signal by the tissue vibration is then received by the Rx antenna of the NCS antenna pair. The received signal is further processed by SoC 2: CC1310, which will forward the demodulated NCS signal to SoC 3: CC2640R2F through a universal asynchronous Rx-Tx (UART) module. SoC 3 is configured as the BLE (Bluetooth low energy) transceiver to relay the digitized NCS signal to a computer or cellphone for postprocessing [42]. As shown in Fig. 3(b), with the appropriate length of the RF cables, the cough sensor can be worn as a choker necklace where the antenna pair is close to the trachea area as shown in Fig. 3(c). The Tx and Rx antennas in the inline configuration are held by a 3D-printed case as shown in Fig. 4(a), which can be represented by the small dipole model in Fig. 4(c). The current I in the Tx (blue dipole) and Rx (orange dipole) antennas is aligned with the z axis. The length of the monopole antenna is denoted by Δl, and the distance between the two antenna phase centers by d. Another possible configuration is shown in Figs. 4(b) and (d), where the two antennas are deployed in parallel with a separation of d. The magnetic potential $\mathbf{A}$ of the Tx antenna can be written as

$$\mathbf{A} = \mathbf{z}_0 \frac{\mu I \Delta l e^{-jkr}}{4\pi r} \tag{1}$$

where $\mathbf{z_0}$ is the unit vector in the z axis, μ is the permeability of vacuum, r denotes the radial coordinate of the observation point, and k is the wavenumber of the EM wave. The magnetic field $\mathbf{H}$ generated by the Tx antenna can be expressed as

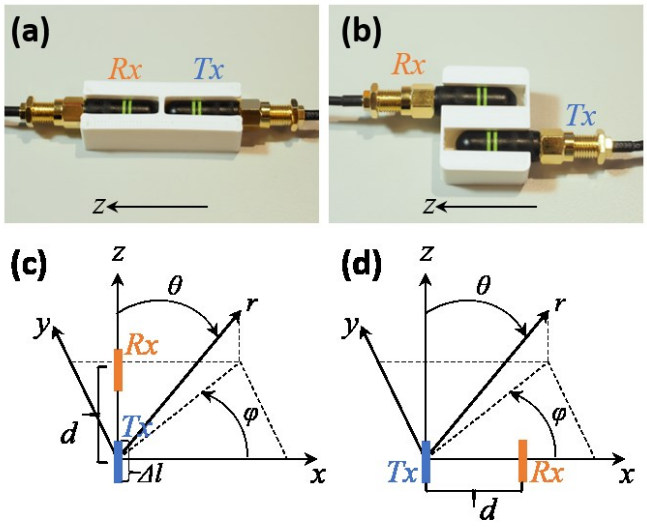


Fig. 4. The NCS sensing antenna pairs. (a) and (c): The picture and analytical model of the inline configuration. (b) and (d): The picture and analytical model of the parallel configuration.

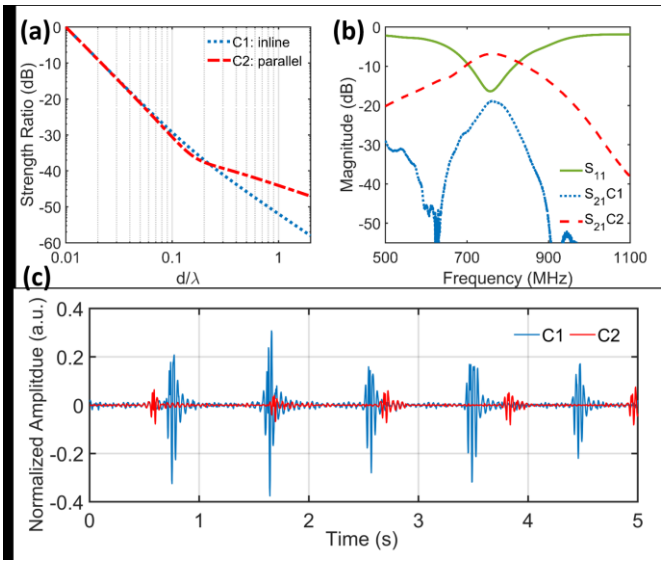


Fig. 5. (a) The electric-field strength ratio versus the distance of d in C1: inline and C2: parallel. The electric-field strength is normalized to 1 at d = 0.01λ for the respective configuration. (b) The S parameters measured from the two pairs of Tx and Rx antennas. (c) The NCS cough signals from C1 (blue) and C2 (red).

$$\mathbf{H} = \frac{1}{\mu}\nabla\times\mathbf{A} = \boldsymbol{\varphi}_0 \frac{jkI\Delta l e^{-jkr}}{4\pi r}\left(1+\frac{1}{jkr}\right)\sin\theta \tag{2}$$

and the electric field $\mathbf{E}$ is

$$\mathbf{E} = \frac{1}{j\omega\varepsilon}\nabla\times\mathbf{H} = \sqrt{\frac{\mu}{\varepsilon}}\frac{jkI\Delta l e^{-jkr}}{4\pi r}\left\{\begin{array}{l}\mathbf{r}_0\left[\frac{1}{jkr}+\frac{1}{(jkr)^2}\right]2\cos\theta \\ +\boldsymbol{\theta}_0\left[1+\frac{1}{jkr}+\frac{1}{(jkr)^2}\right]\sin\theta\end{array}\right. \tag{3}$$

where $\mathbf{r}_0$, $\boldsymbol{\varphi}_0$ and $\boldsymbol{\theta}_0$ are the unit vectors in the spherical coordinates. When observation by Rx is at the far field ($kr \gg 1$) of the Tx antenna, the time-average Poynting power density (<**S**>) can be calculated as

$$<\mathbf{S}>=\frac{1}{2}\mathrm{Re}\left\{\mathbf{E}\times\mathbf{H}^*\right\} \approx \mathbf{r}_0\frac{1}{2}\sqrt{\frac{\mu}{\varepsilon}}\left(\frac{kIl}{4\pi r}\right)^2\sin^2\theta \tag{4}$$

For the inline configuration as shown in Fig. 4(c), $\theta = 0$ in Eq. (4), and there will be no transmission from Tx to Rx directly. Meanwhile, the backscattered signal modulated with the internal tissue vibration will be co-polarized to be received at Rx. However, in the parallel configuration of Fig. 4(d), not only the backscattered signal will be received, but also the direct non-modulated Tx signal will be received by the Rx, which can potentially saturate the receiver front end due to the over-biased signal [23, 43].

When the Rx antenna is much closer to the Tx antenna with the distance of d smaller than the wavelength, the electric field of the Tx antenna at the inline Rx location with $\theta = 0$ can be calculated as

$$\mathbf{E}_{Rx1} = \sqrt{\frac{\mu}{\varepsilon}}\frac{jkI\Delta l e^{-jkd}}{2\pi d}\left[\frac{1}{jkd}+\frac{1}{(jkd)^2}\right]\mathbf{r}_0 \tag{5}$$

Similarly, the electric field at the parallel Rx location in the Fig. 4(d) can be calculated as

$$\mathbf{E}_{Rx2} = \sqrt{\frac{\mu}{\varepsilon}}\frac{jkI\Delta l e^{-jkd}}{4\pi d}\left[1+\frac{1}{jkd}+\frac{1}{(jkd)^2}\right]\boldsymbol{\theta}_0 \tag{6}$$

Fig. 5(a) shows the results of the electric-field strength ratios versus the distance of d normalized to the wavelength. The electric-field strength is normalized to 1 (0 dB) at 0.01λ in each configuration. From Fig. 5(a), when $d < 0.2\lambda$, the field-strength ratios are similar in the near-field region. However, the ratio of the inline configuration (C1) will be much smaller when d further increases. Fig. 5(b) shows the S-parameters measured from the network analyzer (Keysight E5063A) when the antenna pair is placed at the choker necklace position of a person. The green curve is the measured reflection S_{11} curve of the inline configuration, where the center frequency is detuned from around 900 MHz in air to around 800 MHz due to the nearby human body and Rx antenna. The Tx antenna of the antenna pair has the reflection of –16.4 dB at 780 MHz and the bandwidth of 110 MHz where S_{11} is less than –10 dB. The blue-dot curve shows S_{21} in the inline configuration (S_{21}C1), while the red-dash curve is S_{21} in the parallel configuration (S_{21}C2). As S_{21}C1 is much lower than S_{21}C2, the direct Tx-to-Rx coupling is weaker by using the inline configuration, where a reasonable sensing dynamic range and signal resolution can be better achieved before Rx saturation [23, 43]. The improved signal transduction can be observed in Fig. 5(c), where the blue curve (C1) is the cough signal obtained from the inline configuration and the red curve (C2) from the parallel configuration. As the C2 Rx can be more easily saturated by the non-modulated signal, both Tx and Rx gains have to be set lower than those in C1. The simulated coughing in Fig. 5(c) by the same healthy participant was not precisely timed in two separate recordings, where the participant was instructed to make 5 similar mild coughs within 5 seconds in the routine.

IV. Experimental Results

A. Experimental setup and study protocol

Based on the principle and system design presented in Sections II and III, the NCS cough signals were recorded and demodulated using the inline configuration [31], [32]. The system was set at the carrier frequency of 820 MHz and –10 dBm Tx power. The received signal was then amplified by the Rx circuitry and digitized by a 12-bit analog-to-digital converter (ADC). The signal was sampled at 2 kSps (kilo-samples per second) and then transferred to the host computer for postprocessing.

The healthy participant under test wore the sensor as a choker necklace and mimicked the mild cough every 1 – 5 seconds. Each recording period was about 12 minutes. After a 5-minute rest, the participant took the next period of recording. All participants went through 5 periods of recordings, and were in the average body condition, both genders and an age span from 22 to 35. The recorded data were used for cough signal feature analysis and recognition. The second session was designed for the interference activities. The participant was asked to wear the sensor again to take another 5 periods of recording. Each period was about 3 minutes, when the participant was asked to perform the interference routines of normal breathing, speaking, body motion with hand gestures, mock swallowing, and mock

sneezing. The recorded data were used to analyze the false positive rate.

B. Cough signal features

The demodulated signal after the bandpass filter (20 – 1000 Hz) is shown as the blue curves in Fig. 6(a). Six mild coughs within 14 seconds were recorded by the sensor system. The insertions in Fig. 6(a) are the magnified waveform of the 2nd, 4th and 6th coughs from top to bottom, respectively. We can see that the waveform is very different from those of other signal sources detected from the NCS system such as the heartbeat [32], breath [35] and vocal cord tunes [23]. The signal packet caused by coughing was much shorter, where the main part of the signal burst and decayed within 50 ms. The signal was also analyzed by the short-time Fourier transform (STFT), where the frequency-time spectra of the 6th cough is shown in Fig. 6(b). The signal happened around 100 ms with a lower frequency and intensity, and quickly evolved to the higher frequency around 250 Hz within 10 ms as indicated by the black solid arrow. The 2nd and 3rd harmonics were also present in the spectra. Then the signal decayed in both frequency and intensity in 100 ms, as indicated by the dashed arrow. This frequency-time contour reflected the signal relaxation feature after the main peak in the time domain, showing the vocal cord vibration after the strong air burst caused by coughing. The other cough signals were analyzed in the same manner and the statistic results from about 120 cough signals are shown in Table I. We can observe that the NCS signals have distinctive features in both time and frequency domains, which means that signal analysis relying purely on the time domain or frequency domain may not be sufficient for high recognition accuracy. For example, because the single cough signal is not a steady signal, when it is converted to the frequency domain for spectral composition, the time domain information will be lost. On the other hand, a time-domain signal recognition method such as dynamic time warping (DTW) [44], [45] in speech recognition can miss the fine spectral features. Furthermore, DTW utilizes the warping to adapt to the template to accommodate individual differences, which can however distort the spectral features.

TABLE I
OCCURRENCE PROBABILITY OF COUGH RECOGNITION FEATURES

	Percentage of the features being observed
Feature 1: Fast burst of the main part in time	100 %
Feature 2: Fast decay of the main part in time	100 %
Feature 3: Fast rise in frequency and time	98.3 %
Feature 4: Slow decay in frequency and time	97.5 %
Feature 5: 2nd and 3rd harmonics	95.8 %

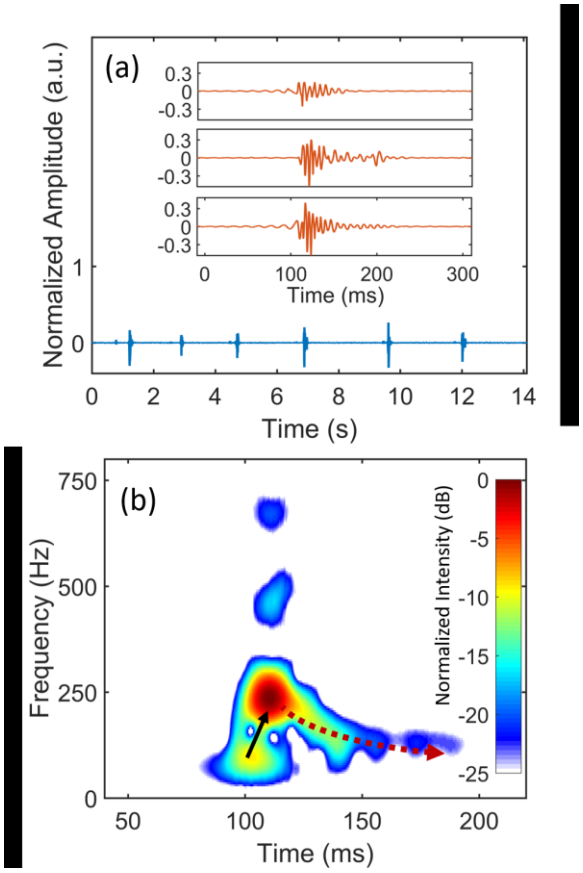


Fig. 6. The vocal-cord NCS recording of coughs. (a) The demodulated time-domain signal; (b) The frequency-time characteristics.

C. CNN based cough recognition

In order to recognize the cough signal automatically, the recorded signal should be properly segmented, where the algorithm will first locate the signal peaks in the time domain, and then make the indices for the frequency-time spectra segmentation. Because NCS is only sensitive to the local tissue vibration and a bandpass filter of 20 – 1000 Hz was applied as mention above, the noise signal caused by other motion interferences is usually weak. Hence, adequate peak readings mainly from the tissue vibration near the trachea area can be readily recognized. The peak detection threshold is set as 0.2 of the averaged peak height from the cough signal recording, and the minimum peak separation is set as 300 ms. More sophisticated peak detection algorithms such as difference between moving average and actual signals can also be helpful to improve the detection accuracy in practical scenarios in the future. After the time-domain signal is processed by STFT, the frequency-time spectrum is segmented based on the signal peak indices. As shown in Fig. 7(a), spectra of 36 mild cough signals

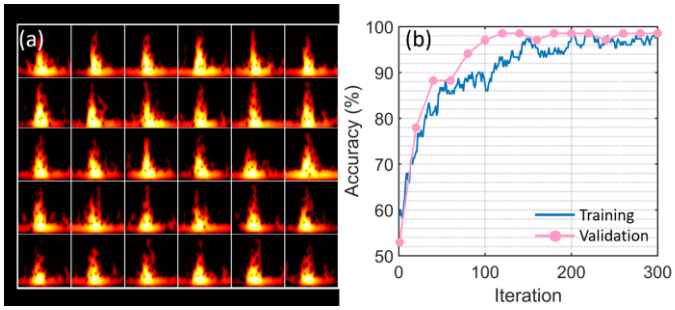


Fig. 7. (a) The randomly picked representative cough signal frequency-time spectra. (b) The training and validation accuracies curves of the CNN applied in this work.

were randomly selected from about 200 signal segments. The dimension of the spectra is 512×512. For cough signal recognition, a convolutional neural network (CNN) by SqueezeNet [46] was trained and verified by the frequency-time spectra extracted from the NCS recording. SqueezeNet has 18 layers in CNN depth. For imaging recognition, under equivalent accuracy, SqueezeNet has fewer parameters, which provides several advantages such as high efficiency for distributed training, reduced overhead when exporting new models to the updating device, and feasibility for embedded deployment. As the eventual wearable sensor will have limited computing power, SqueezeNet can be a good candidate for signal recognition if there is no major compromise on accuracy.

During the training phase, the images of the frequency-time spectra were shifted randomly within several pixels before feeding to the networks to further enlarge the training set by simulating the frequency and timing shifts of the cough signals. Fig. 7(b) shows the training (blue) and validation (pink) accuracy curves of each cross-validation iteration for the cough signals collected from participant 1 (P1). 70% of the samples were chosen for training, and the remaining 30% were configured for the validation. The overall validation accuracy was 97.53%.

D. Human study

The CNN parameters trained from the dataset of one person can be directly transferred to signal recognition of other persons. Fig. 8 shows the recognition accuracies of 5 recording periods from 5 participants (P1 – P5), where the accuracy of each recording period is denoted as the square marker. In each period of recording, around 200 cough signals are extracted from the NCS signals for analysis. The accuracy of the first person in period 1 was the validation accuracy of the training data, which is 97.53%. Then we applied the data from each person as the training data set to obtain 5 trained networks, and validated each person's recording through each network, as shown in Table II. Training data sets 1 – 5 indicate the networks trained by the data from participants 1 – 5, respectively. The overall averaged accuracy is 93.26%, and the minimum accuracy is 90.33%, where the network was trained by participant 4's data and validated by participant 1. Furthermore,

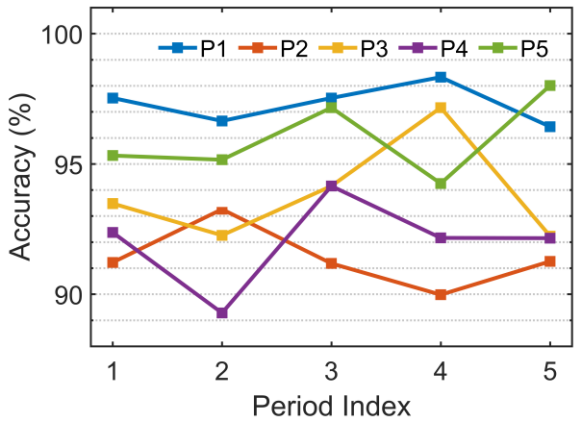


Fig. 8. The recognition accuracies of 5 people (P1 to P5) in 5 recording periods.

TABLE II
AVERAGED COUGH RECOGNITION ACCURACY WITH EACH PARTICIPANT'S DATA FOR TRAINING TO BE TESTED ON ALL PARTICIPANTS

		Participant				
		1	2	3	4	5
Training data set	1	97.32%	91.41%	93.25%	93.42%	93.82%
	2	92.21%	96.32%	93.32%	95.46%	92.15%
	3	92.56%	93.24%	94.12%	93.31%	91.28%
	4	90.33%	93.25%	91.04%	93.45%	92.32%
	5	93.27%	92.34%	93.25%	92.57%	96.46%

we trained the network with only one of the periods from a participant. In Table III, the training data sets 1 – 5 are the period 1 of participant 1, period 2 of participant 2, and so on. The validation data are the periods 1 – 5 of the same participant that trained the network. We can see that the accuracies are high in all cases, because the training data and validation data are from the same participant but different recording periods.

The lower limit of the detection accuracy depends on the cough signal strength. When the strength decreases, the received NCS signal will be weaker with lower SNR and SIR (signal to interference ratio), which can possibly be improved from the NCS transceiver optimization [23] and the antenna design similar to the analysis in Section 3. Alternatively, the signal strength can be increased by deploying the NCS antennas closer to the cough source area to make stronger near-field coupling and higher isolation against the interferences by other motions.

The NCS sensor detects the local permittivity boundary motion. The interference signals not from trachea vibration will be minimally coupled to the system, while speaking and other sounds made by the person under test will be differentiated by the frequency-time spectra. Within the routines in our human study, the 5 people under test were instructed to sit and breathe normally (Routine 1), read aloud normally (Routine 2), perform body motion and hand gestures (Routine 3), mimic swallowing

TABLE III
AVERAGED COUGH RECOGNITION ACCURACY WITH DIFFERENT TRAINING DATA PERIODS TO BE TESTED ON THE SAME PARTICIPANTS

		Period				
		1	2	3	4	5
Training data set	1	97.47%	96.81%	97.53%	98.22%	96.48%
	2	95.12%	97.54%	96.12%	97.44%	96.43%
	3	97.35%	96.29%	97.93%	95.72%	96.78%
	4	96.85%	97.12%	95.64%	98.15%	96.62%
	5	98.64%	97.17%	95.73%	96.43%	97.43%

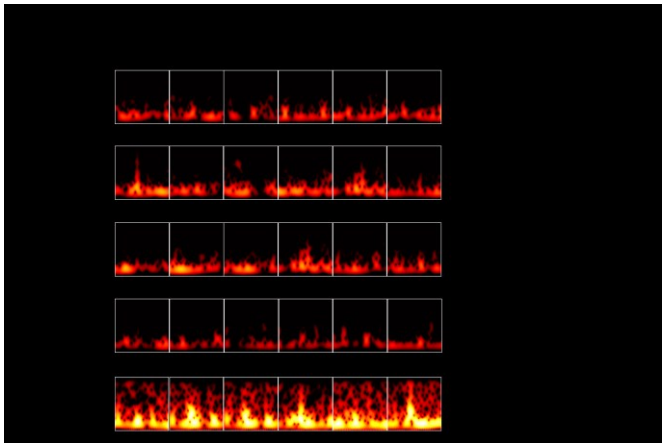


Fig. 9. The interference signals and their false positive rates in 5 routines: Routine 1: Normal breathing; Routine 2: Speaking; Routine 3: Body motion with hand gestures; Routine 4: Mock swallowing; Routine 5: Mock sneezing. The middle column shows 6 randomly picked spectra in each routine.

(Routine 4), and mimic sneezing (Routine 5). The recorded raw data are processed by the same signal segmentation and CNN spectrum recognition algorithms with exactly the same parameters. 300 spectral images in each interference routine from 5 people were collected. The false positive rate in each routine is shown in Fig. 9, together with the randomly selected frequency-time spectra. Routine 4 with the mock swallowing has the least false positive at 0.22%. Not surprisingly, mock sneezing with local trachea vibration by the sneezing sound and the associated body motion has the highest false positive at 2.09%. The false negative rate can be readily derived from the accuracy listed in Tables II and III.

V. Discussion and Conclusion

Comparing to other CNN architectures [47]–[49], SqueezeNet provided sufficient accuracy with much fewer parameters in this application. As shown in Table IV, the size of the SqeezeNet (4.8 MB) is significantly smaller than the other networks, and can be further compressed to 0.47 MB [46]. The other networks, AlexNet, GoogLeNet and ResNet were trained and verified with the same data set as used for SqeezeNet, and the resulting accuracies were at the similar level. The training time of the SqueezeNet was also on the shorter side.

TABLE IV
Performance Comparison with Different Networks

CNN name	SqueezeNet	AlexNet	GoogLeNet	ResNet
Accuracy	97.53%	98.58%	98.63%	96.67%
GPU training time (sec)	119	98	252	399
CPU training time (sec)	665	672	1169	2716
CNN size (MB)	4.8	227	27	96

In this work, we applied a PCB RF transceiver as the NCS sensor for wearable cough monitoring, where Tx and Rx were operating at the same frequency band. To further improve the sensitivity, the harmonic RF transceiver [23], [34], [50] can be applied to the NCS system. The sensing antenna pair can also be designed in other configurations [32], or even manufactured with embroidered conductive fiber [31], [51] and printed conductive ink [52], so that the devices can be integrated with garments. The microcontroller on the PCB board demodulated the received RF signal to be either sent to a host computer through Bluetooth or stored locally in an SD (secure digital) card. Although the memory requirement of the CNN algorithm is relatively small, it was not implemented on the present PCB board in order to reduce the power consumption. The present 350 mAh lithium battery on PCB can provide over 20-hour continuous operation. The current system design and experimental tests were validated only on the average body conditions. Further complications to the sensor performance can be caused by special human factors such as obesity, large body mass around the neck, and glottis disorders, which should be further investigated in the future clinical research.

The experimental results show that this wearable RF sensor with the NCS technology is capable of long-term, continuous monitoring of the mild cough. A CNN algorithm was employed to recognize the frequency-time spectra of coughing, where the true positive rate was above 90% and the false positive rate was below 2.09% when the person under test was in various common interference routines.

References

[1] R. S. Irwin, M. J. Rosen and S. S. Braman, "Cough: A comprehensive review," *Archives of Internal Medicine*, vol. 137, no. 9, pp. 1186–1191, Sep. 1977.

[2] M. Wielpütz and H. U. Kauczor, "MRI of the lung: State of the art," *Diagnostic and Interventional Radiology*, vol. 18, no. 4, pp. 344, July 2012.

[3] H. Matsumoto *et al.*, "Airway wall thickening in patients with cough variant asthma and nonasthmatic chronic cough," *Chest*, vol. 131, no. 4, pp. 1042–1049, Apr. 2007.

[4] M. M. Salvatore, R. C. Go and M. A. Pernia M., *Chest CT for Non-Radiologists: A Practical Guide, Springer*, 2018.

[5] R. Smith-Bindman, J. Lipson, R. Marcus, K. P. Kim, M. Mahesh, R. Gould, A. B. de Gonzalez and D. L. Miglioretti, "Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer", *Arch Intern Med.*, vol. 169, no. 22, pp. 2078–2086, 2009.

[6] R. Lethbrideg and M. L. Everard, "The stethoscope: Historical considerations," in *Breath Sounds* (eds. K. N. Priftis, L. J. Hadjileontiadis and M. L. Everard), pp. 15–31, Springer, Switzerland, 2018.

[7] H. Pasterkamp, S. S. Kraman and G. R. Wodicka, "Respiratory sounds: Advances beyond the stethoscope," *American J. Respiratory and Critical Care Medicine*, vol. 156, no. 3, pp. 974–987, Sep. 1997.

[8] M. Blaivas, "Lung ultrasound in evaluation of pneumonia," *J. Ultrasound in Medicine*, vol. 31, no. 6, pp. 823–826, June 2012.

[9] S. Subburaj, L. Parvez and T. G. Rajagopalan, "Methods of recording and analysing cough sounds," *Pulmonary Pharmacology*, vol. 9, no. 5-6, pp. 269–279, Oct. 1996.

[10] E. C. Larson *et al.*, "Accurate and privacy preserving cough sensing using a low-cost microphone," *Proc. 13th International Conference on Ubiquitous Computing*, pp. 375-384, Sep. 2011.

[11] J. Amoh and K. Odame, "Deep neural networks for identifying cough sounds," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 5, pp. 1003–1011, Sep. 2016.

[12] M. Roudjane *et al.*, "Smart T-shirt based on wireless communication spiral fiber sensor array for real-time breath monitoring: Validation of the technology," *IEEE Sensors Journal*, vol. 20, no. 18, pp. 10841–10850, May 2020.

[13] T. Elfaramawy *et al.*, "A wireless respiratory monitoring system using a wearable patch sensor network," *IEEE Sensors Journal*, vol. 19, no. 2, pp. 650–657, Oct. 2019.

[14] C. Chang, C. Lin, S. Chang, C. Lin, Z. Jiang and S. Yu, "Glottal Doppler radar system and its applications to communication and speaker recognition," in *41st Euro. Microw. Conf.,* Manchester, UK, Oct. 2011.

[15] H.-C. Kuo *et al.*, "A fully integrated 60-GHz CMOS direct-conversion Doppler radar RF sensor with clutter canceller for single-antenna noncontact human vital-signs detection," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 4, pp. 1018–1028, Apr. 2016.

[16] 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 Techn.* vol. 61, no. 5, pp. 2046–2060, Apr. 2013.

[17] E. L. Chuma and Y. Iano, "A movement detection system using continuous-wave Doppler radar sensor and convolutional neural network to detect cough and other gestures," *IEEE Sensors Journal*, vol. 21, no. 3, pp. 2921–2928, Feb. 2021.

[18] M. C. Huang *et al.*, "A self-calibrating radar sensor system for measuring vital signs," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 2, pp. 352–363, May 2015.

[19] E. Miller, N. Banerjee and T. Zhu, "Smart homes that detect sneeze, cough, and face touching," *Smart Health*, vol. 19, art. 100170, Mar. 2021.

[20] G. Wang, C. Gu, T. Inoue and C. Li, "A hybrid FMCW-interferometry radar for indoor precise positioning and versatile life activity monitoring," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 11, pp. 2812–2822, Sep. 2016.

[21] Y. H. Shin and J. Seo, "Towards contactless silent speech recognition based on detection of active and visible articulators using IR-UWB radar," *Sensors*, vol. 16, pp. 1812, Nov. 2016.

[22] P. Birkholz *et al.*, "Non-invasive silent phoneme recognition using microwave signals," *IEEE/ACM Trans. Audio, Speech, and Language Processing*, vol. 26, pp. 2404–2411, 2018.

[23] X. Hui, T. B. Conroy and E. C. Kan, "Near-field coherent sensing of vibration with harmonic analysis and balance signal injection," *IEEE Trans. Microw. Theory Techn.*, 2021, doi: 10.1109/TMTT.2021.3053978

[24] S. Li *et al.*, "A 94-GHz millimeter-wave sensor for speech signal acquisition," *Sensors*, vol. 13, no. 11, pp. 14248–14260, 2013.

[25] M. Geiger, D. Schlotthauer and C. Waldschmidt, "Improved throat vibration sensing with a flexible 160-GHz Radar through harmonic generation," in *Proc. Microw. Symp. (IMS), 2018 IEEE MTT-S Int.*, Philadelphia, PA, USA, June 2018.

[26] D. Rodriguez and C. Li, "Sensitivity and distortion analysis of a 125-GHz interferometry radar for submicrometer motion sensing applications," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 12, pp. 5384–5395, Dec. 2019.

[27] J. Min *et al.*, "Wearable electrochemical biosensors in North America," *Biosens. Bioelectron.*, vol. 172, pp. 112750, Jan. 2021.

[28] A. Nag, S. C. Mukhopadhyay, J. Kosel, "Wearable flexible sensors: A review," *IEEE Sensors Journal*, vol. 17, no.13, pp. 3949–3960, July 2021.

[29] N. Golestani and M. Moghaddam, "Human activity recognition using magnetic induction-based motion signals and deep recurrent neural networks," *Nat. Commun.*, vol. 11, no. 1, pp. 1–11, Mar. 2020.

[30] Y. Chuo *et al.*, "Mechanically flexible wireless multisensor platform for human physical activity and vitals monitoring," *IEEE Trans. Biomed. Circuits Syst.*, vol. 4, no. 5, pp. 281–294, Sep. 2010.

[31] 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, Jan. 2018.

[32] 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, May 2020.

[33] X. Hui, P. Sharma and E. C. Kan, "Microwave stethoscope for heart sound by near-field coherent sensing," in *Proc. Microw. Symp. (IMS), 2019 IEEE MTT-S Int.*, Boston, MA, USA, June 2019.

[34] 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.

[35] X. Hui and E. C. Kan "Seat integration of RF vital-sign monitoring," in *Proc. 2019 IEEE Int. Microw. Bio. Conf. (IMBioC),* Mar. 2019.

[36] 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 2019.

[37] P. Sharma, X. Hui, J. Zhou, T. B. Conroy and E. C. Kan, "Robust cardiopulmonary monitoring using wearable radio-frequency sensors," *Nature Partner Journal (NPJ), Digital Medicine*, doi: 10.1038/s41746-020-0307-6, Aug. 2020.

[38] X. Hui, J. Zhou and E. C. Kan, "High-volume parallel mouse vital-sign monitoring with near-field coherent sensing," *IEEE Antennas Wirel. Propag. Lett.*, vol. 19, no. 7, pp. 1152 – 1156, May 2020.

[39] Z. Zhang, *et al.*, "Furniture-integrated respiration sensors by notched transmission lines," *IEEE Sensors Journal*, vol. 21, no. 4, pp. 5303–5311, Feb. 2021.

[40] X. Hui and E. C. Kan, "Seat integration of RF vital-sign monitoring," in *2019 MTT-S Int. Microw. Bio Conf. (IMBIOC)*, Nanjing, China, May 2019.

[41] *CST Studio Suite*, retrieved June 20, 2020. [Online]. Available: https://www.3ds.com/products-services/simulia/products/cst-studio-suite/

[42] J. Zhou *et al.*, "A wireless wearable RF sensor for brumation study of chelonians," *IEEE J. Electromagn., RF, Microw. Med. Biol.*, vol. 5, no. 1, pp. 17–24, Mar. 2021.

[43] D. Tang, J. Wang, W. Hu, Z. Peng, Y. Chiang and C. Li, "A DC-coupled high dynamic range biomedical radar sensor with fast-settling analog DC offset cancelation," *IEEE Trans. Instrumentation and Measurement*, vol. 68, no. 5, pp. 1441–1450, May 2019.

[44] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. Acoust. Speech Signal Process,* vol. 26, no. 1, pp. 43–49, Feb. 1978.

[45] J. F. Lichtenauer, E. A. Hendriks and M. J. T. Reinders, "Sign language recognition by combining statistical DTW and independent classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 11, pp. 2040–2046, May 2008.

[46] F. N. Iandola *et al.*, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size," *arXiv preprint*, arXiv:1602.07360, Feb. 2016.

[47] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks", *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.

[48] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Boston, MA, USA, June 2015.

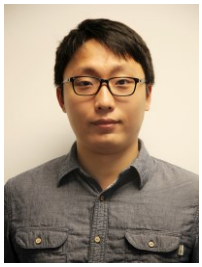
[49] K. He *et al.*, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, June 2016.

[50] X. Hui and E. C. Kan, "Radio ranging with ultrahigh resolution using a harmonic radio-frequency identification system," *Nat. Electron.*, vol. 2, no. 3, pp. 125, Mar. 2019.

[51] Z. Wang, L. Zhang, Y. Bayram, J. L. Volakis, "Embroidered conductive fibers on polymer composite for conformal antennas," *IEEE Trans. Antennas Propag.*, vol. 60, no. 9, pp. 4141–4147, Sep. 2012.

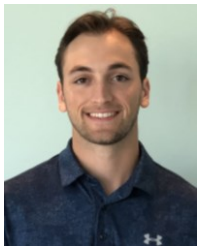
[52] J. Sidén, M. K. Fein, A. Koptyug and H.-E. Nilsson, "Printed antennas with variable conductive ink layer thickness," *IET Microw. Antennas Propag.*, vol. 1, no. 2, pp. 401–407, Apr. 2007.

Xiaonan Hui received the B.S. degree from Northeastern University, China in 2012, and the M.S. degree from Zhejiang University, China in 2015, all in electrical engineering. And he received the Ph.D. degree in electrical and computer engineering at Cornell University, Ithaca, NY, USA.



Jianlin Zhou received the B.S. degree in electronics and information engineering from Nanjing University of Science and Technology, Nanjing, China in 2013, and the M.S. degree in electrical and computer engineering from Northeastern University, Boston, MA in 2016. He is currently pursuing the Ph.D. degree in electrical and computer engineering at Cornell University, Ithaca, NY, USA.

Pragya Sharma received the B.Tech. degree in electrical engineering from Indian Institute of Technology (IIT) Kharagpur, India, in 2015. She is currently pursuing the Ph.D. degree in electrical engineering at Cornell University, Ithaca, NY, USA.



Thomas Conroy received the B.S. degree from Clarkson University, Potsdam, New York in 2019 in electrical engineering. He is currently pursuing the Ph.D. degree in electrical and computer engineering at Cornell University, Ithaca, NY, USA.

Zijing Zhang received a B.Eng. degree in optoelectronic engineering from the Huazhong, University of Science and Technology (HUST), China, Wuhan, in 2019. She is currently pursuing a Ph.D. degree in Electrical Engineering at Cornell University, Ithaca, NY, USA.

Edwin C. Kan received the B.S. degree from National Taiwan University in 1984, and the M.S. and Ph.D. degrees from University of Illinois, Urbana-Champaign in 1988 and 1992, respectively, all in electrical engineering. He has served in the Air Force, Taiwan, R. O. China as a second lieutenant from 1984 – 1986. In January 1992, he joined Dawn Technologies as a Principal CAD Engineer. He was then with Stanford University as a Research Associate from 1994 to 1997. In 1997, he joined the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY as an assistant professor, where he is now a professor. His research interests include RF biosensors, RFID, RF sensors for smart buildings, CMOS technology, Flash memory, molecular sensors, and technology computer-aided design (TCAD).