IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 19, 2019, accepted December 22, 2019, date of publication January 13, 2020, date of current version January 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2966126

Respiration and Activity Detection Based on Passive Radio Sensing in Home Environments

QINGCHAO CHEN[1], (Member, IEEE), YANG LIU[1], BO TAN[2], (Member, IEEE), KARL WOODBRIDGE[3], (Senior Member, IEEE), AND KEVIN CHETTY[3], (Member, IEEE)

[1]Department of Engineering Science, University of Oxford, Oxford OX1 2JD, U.K.
[2]Faculty of Information Technology and Communication Sciences, Tampere University, 33100 Tampere, Finland
[3]Department of Electronics and Electrical Engineering, University College London, London WC1E 6BT, U.K.

Corresponding author: Bo Tan (bo.tan@tuni.fi)

The work of BO TAN was supported by the Profi 2 funded by the Academy of Finland under Grant 326654. The work of KARL WOODBRIDGE and KEVIN CHETTY was supported by the U.K. Engineering and Physical Sciences Research Council (EPSRC) OPERA Project under Grant EP/R018677/1.

ABSTRACT The pervasive deployment of connected devices in modern society has significantly changed the nature of the wireless landscape, especially in the license free industrial, scientific and medical (ISM) bands. This paper introduces a deep learning enabled passive radio sensing method that can monitor human respiration and daily activities through leveraging unplanned and ever-present wireless bursts in the ISM frequency band, and can be employed as an additional data input within healthcare informatics. Wireless connected biomedical sensors (Medical Things) rely on coding and modulating of the sensor data onto wireless (radio) bursts which comply with specific physical layer standards like 802.11, 802.15.1 or 802.15.4. The increasing use of these unplanned connected sensors has led to a pell-mell of radio bursts which limit the capacity and robustness of communication channels to deliver data, whilst also increasing inter-system interference. This paper presents a novel methodology to disentangle the chaotic bursts in congested radio environments in order to provide healthcare informatics. The radio bursts are treated as pseudo noise waveforms which eliminate the requirement to extract embedded information through signal demodulation or decoding. Instead, we leverage the phase and frequency components of these radio bursts in conjunction with cross ambiguity function (CAF) processing and a Deep Transfer Network (DTN). We use 2.4GHz 802.11 (WiFi) signals to demonstrate experimentally the capability of this technique for human respiration detection (including through-the-wall), and classifying everyday but complex human motions such as standing, sitting and falling.

INDEX TERMS Machine learning, deep transfer networks, opportunistic wireless networks, signs-of-life detection, human activity monitoring, micro-Doppler signature, phase-sensitive detection.

I. INTRODUCTION

Human activities have been a indispensable component of the healthcare informatics in both clinical and daily contexts because of its value in ambient assistant living (AAL) and detection of early symptoms of disease. In-home human activity detection has attracted increasing attention because of its indexing effect for understanding health, emotion, psychology [1] and behavioural patterns [2] in residents. Thus, detection and recognition of physical activities via direct and indirect sensing data is a growing area of research. Within the framework of the Internet of Things (IoT) and smart homes, healthcare sensors are experiencing a transition from hospitals, care houses and laboratories into ordinary homes. The SPHERE Project [3], developed by the University of Bristol in the UK, is an exciting showcase that demonstrates the detection of residential activity via the connected home sensors such as cameras, wearable devices, environmental sensors (measuring light, temperature, air quality, ambient sound, humidity etc). The capability of the integrated sensor platform for activity detection shows significant promise. However, two problems must be addressed to ensure the feasibility of real-world deployment for the technology. Firstly, the intrusive nature of interacting and/or the data output from

The associate editor coordinating the review of this manuscript and approving it for publication was Zheng Xiao.

these technologies must be balanced e.g. the use of wearable devices versus passive contactless sensors, or the accuracy (e.g. high definition video stream) versus more private environmental and meta data in the residential environments. Second, the large-scale rollout of wireless connected sensors in the ISM bands results in overcrowded spectrum that impacts on interference control, media access scheduling and Quality of service (QoS). Therefore, we introduce a non-invasive and accurate technology that also maintains user privacy for human activity detection (including respiration) based on passive radio sensing.

Different from the regular sensing approaches which modulate the coded sensory data onto radio carriers for transmission, the concept in this paper makes use of radio bursts that are already present in the environment. Rather than flooding the existing RF environment with additional bursts, our work aims to extract information from properties inherent to the radio signals without demodulation and decoding. It is well known that the properties of the radio signals are impacted by the environmental characteristics during propagation, therefore a high correlation between radio signal properties and the human activities (part of the environmental) is implicit. However, extracting the human activity information from the uncontrolled environmental radio signals is non-trivial. We leverage both signal processing approaches and deep learning to achieve this goal. A software defined radio (SDR) prototype system is employed to demonstrate the concept. The core signal processing technique used in this paper is termed the cross ambiguity function (CAF) [4] which is able to differentiate between the transmitted wireless signal and its reflected counterparts in both time and frequency domains. We postulate that the post CAF processing signal contains two types of physical activities. ***type_i:*** activity that can be direct correlated with one of the radio signal properties. In this paper, we refer to respiration rate which has high visible correlation with the phase variations of the signals reflected by a subject under test. ***type_ii:*** activity for which interpretation is beyond the scope of a direct observation and must therefore rely on computational inference. In this paper we therefore propose a novel Deep Transfer Network (DTN) for learning the relationship between the measured micro-Doppler (μ-D) signature and everyday human activities. The classification methods employed in the network are the Sparse representation classifier (SRC) and support vector machine (SVM). The DTN [5] we propose is pre-trained by the ImageNet vision dataset [6] but fine-tuned using a much smaller dataset containing micro-Doppler signatures (μ-DS) relating to everyday activities. Both cases are tested experimentally using a software defined radio (SDR) system and unmodified 802.11 signals from a commodity WiFi access point. We also validate our results and explain the mechanisms involved with reference to the wireless network spatial geometries. Moreover, we argue that the highly congested overcrowded radio-frequency (RF) spectrum could actually be a beneficial resource for monitoring people in wireless enabled environments. These proof-of-principle results could lead to a wide range low-impact, non-cooperative human activity and respiration data collection in normal residential environments for healthcare informatics.

The rest of the paper is organised as follow: Section II gives a comprehensive review of recent related work. Section III introduces the enabling cAF signal processing for extracting phase information from signals, frequency offset estimation, and our deep transfer network for classifying μ-DS information. The experimental results and related discussion are presented in Section IV. Finally, Section V concludes the work and provides some perspective on possible future directions for the technology.

II. RELATED WORK

Using sensors to monitor human activities in home environment is not a brand new idea. Various direct and indirect approaches have been tested or applied in the laboratory or more practical scenarios, including auxiliary (or wearable) sensors (accelerometers, gyroscopes, RFID etc.), environmental sensors (camera, LiDAR, PIR, temperature, humidity, smart meters etc.) and the radio sensors or indicators. In this section, outline the types of sensors currently used for activity monitoring within a healthcare context, and provide a rationale for the approach taken in our work.

A. AUXILIARY SENSORS

Auxiliary sensors are often used to monitor physical human activity [7] and these include wearable technologies [8]–[10], mobile phones [11] and radio frequency identification (RFID) [12]. Among these, sensors embedded in wearables and mobile phones such as accelerometers, gyroscopes and gravity sensors are able to provide accurate movement information that can be used to estimate activity levels, classifying the activities and calibrating the physical gestures. However, wearable technologies are associated with physical discomfort and low rates of acceptance, especially among the elderly. RFID based devices employ complex transmitters and receivers, and require pre-planning in order to optimally site the positions of the nodes [7]. Also, this approach is often limited by short transmission distance of RFID tags.

B. ENVIRONMENTAL SENSORS

Different from auxiliary sensors, environmental sensors have the advantage of providing contactless activity detection. We categorise these sensors into two types according to the high-definition (HD) and low-definition (LD) environmental sensors according the definition level of the data provided by the sensor. The typical HD environmental sensors include camera and LiDAR that usually can provide detailed pose and gesture information by using the rich spatio-temporal relations between the pixels. Vision based sensors [13], [14] such as MS Kinect and Intel RealSense have been investigated in some healthcare projects [3]. However, in general, the video camera systems suffer from poor contrast and their use in home environments raises many privacy issues. The fact that

high volume real-time HD data streams require a high quality radio network or wired installation also limits deployment in residential homes. The LD environmental sensors, for example passive infrared (PIR), door/window sensors, temperature, humidity meters etc lay at the other end of the scale. These sensors require very limited network resources, but only provide human activity information which also relies heavily on annotation. PIR [15], [16] for example, is capable of outputting coarse-grained room level existence [17], but suffers from the high false alarms and is affected by varying levels of brightness.

C. RECEIVED SIGNAL STRENGTH AND CHANNEL STATE INFORMATION

Measurements based on the received signal strength indication (RSSI) and the channel state information (CSI) - read from network interface card (NIC) drivers - have become popular in recent years for human activity estimation. Patwari *et al* [18] employ a tomographic approach to achieve accurate localization performance using RSS measurements from distributed radio transceivers. Additionally, the authors in [19]–[21] discuss using variations in the RSSI of both WiFi and Bluetooth signals within indoor scenarios for human presence detection, activity and location estimation. The WiGest system described in [22] has even been shown to recognize arm and hand gestures using RSSI. Other systems such as WiDraw [23], Wi-See [24], CARM [25] and Wi-Key [26] use dedicated commercial off-the-shelf (COTS) 802.11 NIC to estimate angular and frequency components from the recorded CSI, and successfully demonstrate in human body poses and recognition of hand gestures. We highlight here the reliance of these approaches on communications data provided by the drivers of commodity network devices.

D. ACTIVE RADIO SENSING

Aside from using the commodity wireless network devices for activity sensing, various other studies have examined the use of dedicated radio transceivers that enable higher-frequency carriers and bespoke waveform designs optimised for human activity sensing. These active systems operate with improved range and Doppler resolutions permitting the detection of human physiology such as the beating of the heart. Wi-Track for example [27] and [28] transmits a wideband frequency-modulated continuous-wave (FMCW) signal to achieve accurate indoor human target tracking, even for multiple targets. [29] employs the similar FMCW signal to monitor breathing and heart beats from stand-off distances. In [8] and [30], the authors demonstrate the ability of an ultra-wideband (UWB) radar on 60 GHz for tracking the movements and activities of personnel. However, disadvantages associated with active systems include the increased cost of having a dedicated transmitter, inherent coverage limits from large attenuation and narrow directional beams of high carrier frequencies, and the requirement to obtain a spectrum licence to operate within various bands.

E. PASSIVE RADIO SENSING

To overcome the limitations associated with active RF transmitters, human activity detection and classification based on passive sensing has attracted significant research attention [31] and [32] as it offers a promising non-cooperative and non-invasive solution for indoor sensing. In passive operation, the sensing system captures the transmitted radio signal from an illuminator of opportunity (e.g. a WiFi AP) as well as reflected radio signals from the environment (in separate receive channels). The passive system identifies human activity information by differentiating the signals from two channels in frequency and time domains. In the context of the human activity recognition, the frequency shift (Doppler) is more popular as its direct correlation with the dynamic movements. Use-cases include through-the-wall person detection [31], [32], body gesture recognition [33], [34], and respiration detection using Doppler [35] and small movement detection based on phase shifts from indoor wireless signals [36].

This paper focus on passive radio sensing because of its ability to provide non-invasive, highly accurate, low privacy sensing in residential areas. But without control of signal waveform and system geometry, passive systems often show large variance on signal representation and performance. This fact brings challenges on interpreting the detailed or inconspicuous human physical movement and recognition rate stability in passive radio sensing. Thus, we work on advanced signal processing and deep learning technology on the passive captured radio signals to push the passive radio sensing technology towards real-world deployments.

III. PROCESSING AND MACHINE LEARNING FRAMEWORK

In this paper we adopt an end-to-end framework that integrates **i)**. a highly sensitive phase extraction technique for respiration detection and **ii)**. a deep transfer network for daily activity recognition, based on the CAF processing (time-frequency differentiating). The signal and data processing framework is illustrated in Fig. 1. Our phase-sensitive instantaneous Doppler technique described in Section III-B1 identifies very a small physical movement. Similar sensitivity would require tens or even hundreds of milliseconds sampling duration in the traditional CAF [4] or short-time Fourier transform (STFT) based approaches. Furthermore, we leverage a DTN architecture that is pre-trained on ImageNet dataset for μ-DS data classification. To facilitate the model training, we proposed a novel μ-D data processing and an augmentation pipeline to fine-tune the DTN (described in in Section III-C).

A. CROSS AMBIGUITY FUNCTION

In this paper, a CAF is first applied to raw I/Q samples from the direct and reflection channels (equivalient to reference and surveillance signals in passive radar). The CAF carries out the preliminary Difference analysis of reference and

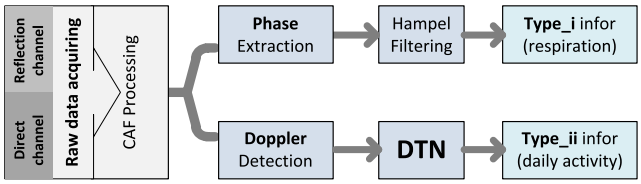


FIGURE 1. Signal and data processing framework.

surveillance signals in the time and frequency domains. This enables further phase extraction for respiration detection or feature fine-graining for activity recognition. The reference signal $x_{ref}(t)$, can be described as a linear combination of the "clean" transmitted Wi-Fi signal $x_{source}(t)$ and reflections from static objects which present as copies of $x_{source}(t)$ associated with delay τ_p and the corresponding complex magnitude A_p^{ref} of p_{th} static objects:

$$x_{ref}(t) = \sum_p A_p^{ref} \cdot x_{source}(t - \tau_p) + n_{ref} \tag{1}$$

where n_{ref} is the noise on reference receiver. Similarly, the measured signal in the surveillance channel $x_{sur}(t)$, is composed of echoes from all moving targets in the area of interest (defined by the signal coverage area of the Wi-Fi AP) which can be characterized as the q_{th} path delay τ_q, its Doppler shift $f_{d,q}$ and relevant complex magnitude A_q^{sur}:

$$x_{sur}(t) = \sum_q A_q^{sur} \cdot x_{source}(t - \tau_q) e^{j2\pi f_{d,q} t} + n_{sur} \tag{2}$$

where n_{sur} is the noise on surveillance receiver. It is assumed that the reference and surveillance channel are separated via the spatially distributed directional antennas. In general, a target can be identified by taking Fourier transform of cross correlation of the reference and surveillance signals. The processing can be represented by the CAF $\mathcal{X}(\tau, f)$ (3) in [37], [38]. The $\mathcal{X}(\tau, f)$ spans a surface. The corresponding delays τ and frequency shifts f of the peaks on $\mathcal{X}(\tau, f)$ surface indicates the distance and moving velocity of the detected targets.

$$\mathcal{X}(\tau, f) = \int_0^t e^{-j2\pi ft} \cdot x_{ref}^*(t - \tau) \cdot x_{sur}(t)\, dt \tag{3}$$

In practice, the signals $x_{ref}(t)$ and $x_{sur}(t)$ will present as the time domain discrete I/Q samples captured via reference and surveillance channels. Thus, the equations (1) to (3) can be written in time discrete form.

B. RESPIRATION DETECTION

From the equation (3), we can see that the integration time determines the achievable Doppler resolution. To capture the small body motions like chest wall movements caused by respiration, longer integration time will be required. However, the expected effect may not be observed since the longer integration time may also mix multiple body movements into one CAF calculation. This may results in an incomprehensible Doppler observation in practice. Thus, this paper targets on taking instantaneous Doppler measurements, with which we can obtain the phase information which is sensitive enough to discern tiny movements smaller than one wavelength. In the following sections, we first discuss the limitations of the conventional in the CAF for respiration detection, and subsequently propose a phase-sensitive processing technique that is embedded on the CAF. Secondly, we use the Hampel filter to guarantee the phase stability.

1) PHASE-SENSITIVE PROCESSING

In normal human breathing, the chest moves slowly with relatively low amplitudes. Detecting this motion is therefore limited by three factors. The first factor is the bandwidth of the Wi-Fi burst which dictates that the achievable range resolution is 17 meters [39], [40] which is too coarse for respiration detection applications. The second factor is the extended integration time which is required to resolve the Doppler shift. Finally, the direct signal interference (DSI) leakage from the Wi-Fi transmitter to surveillance channel may mask small Doppler values which are close to zero. Here, we analyze breathing detection in detail and denote the range of chest movements as $d(t)$. Accordingly, the phase caused by the chest movement that is contained in the received echo signal can be represented by function of the time $\phi(t)$:

$$\phi(t) = \frac{2\pi \times d(t)}{\lambda}. \tag{4}$$

Due to the coarse range resolution of the Wi-Fi signal, we can assume that the respiration motion remains in the same range bin which contains maximum power and can identified by cross correlation option in the equation (1). The correlation result of the m^{th} slow-time sample $x[m], m \in [0, 1, \ldots, M-1]$ with respect to the maximum power range bin l_{max} is then given by:

$$x[m] = \sum_{n=0}^{N_m - 1} x_{ref}^*[i_m + n - l_{max}] \times x_{sur}[i_m + n] \tag{5}$$

where M is the number of batches the signal is divided into and $*$ is the Hermitan operator. N_m is the number of data samples in m^{th} batch and i_m is the starting sample index of each batch. Batch processing [4] is computationally efficient and enables real-time time-frequency analysis. In an indoor environment, given that there are echoes from various other reflectors, the phase of $x[m]$ can be represented as:

$$\phi[m] = \frac{2\pi \times d[m]}{\lambda} + \phi_{static}[m] \tag{6}$$

where $\phi_{static}[m]$ is the sum of echo phases from static reflectors. The time varying trend of The phase of $x[m]$ will present high consistency with time varying of $\phi[m]$ because both varying caused by the the quasi-periodical chest surface displacement of respiration. Thus, we use the phase of $x[m]$ to indicate the respiration. Since the motion caused by breathing is relatively slow and small, $\phi_{static}[m]$ cannot be larger than half the wavelength ϕ_{static} is time-invariant between two consecutive CAF operations in equation (3), we are able to

instantly monitor the phase variation corresponding to the chest movement, without having to consider the distortion from DSI and the long integration times involved in the Fourier transform (FFT if discrete form is used) operation.

2) PHASE STABILITY

Extracting stable and continuous phase outputs from the cross-correlation processing is a challenging task because of the modulation characteristics of Wi-Fi signals. The performance may further degrades when the Wi-Fi AP is working in the beacon mode because of the low duty cycles of the beacon signals. When $x_{source}(t)$ is sparse in the time domain, discontinuous phase samples are frequently observed and the output is subject to significant distortions. In this situation, phase noise in the receiver and the clutter will dominate the result. Thus, in this paper a Hampel filter is used as a post-processing step to eliminate outliers caused by the discontinuous phase output [41]. Suppose we have obtained a series of phase samples, denoted as ϕ_k:

$$\phi_k = [\phi_{k-\omega}, \phi_{k-\omega+1}, \ldots, \phi_k, \ldots, \phi_{k+\omega-1}, \phi_{k+\omega}], \quad (7)$$

The task is to check if ϕ_k is an outlier. First, the median absolute deviation (MAD) scale estimation $\tilde{\phi}_k$ within a window size ω is calculated by following equation (8):

$$\tilde{\phi}_k = median\{|\phi_k - median\{\phi_k\}|\}. \quad (8)$$

Then, with the assumption that the phase data within the window are sampled according to the following normal distribution:

$$\phi_k \sim \mathcal{N}(\mu_k, \sigma_k) \quad (9)$$

the estimated deviation of the distribution, σ_k can be approximately given by the criterion from [41]:

$$\sigma_k = 1.4826 * \tilde{\phi}_k, \quad (10)$$

where 1.4826 is the calculated in the Gaussian distributions [42]. Given a threshold T, we can classify the phase as an outlier if $\|\phi_k\| \geq T \times \sigma_k$. In this case, the outlier will be set as the median of the series of phase samples. Although outliers can be eliminated by the methods outlined above, phase shifts caused by the background static reflections still remain. Due to the time invariant naure during consecutive CAF processing, we can subtract the mean of phase outputs as background elimination, represented by the following equation:

$$\phi_{true,k} = \phi_k - \frac{1}{L} \sum_{k-L+1}^{k} \phi_k, \quad (11)$$

where ϕ_k is the k_{th} slow-time sequence of phase information, L is the window length and $\phi_{true,k}$ is the phase sequence after background elimination (mean subtractions).

After applying our phase sensitive processing and Hampel filtering after the CAF processing, we notice that the rate of the phase varies and this correlates directly with the respiration rate of the human subject (**type_i** information). The detailed results are presented in Section IV-B and includes a discussion the impact of the bi-static detection geometry.

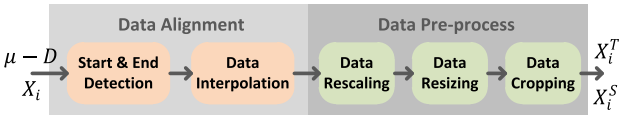


FIGURE 2. Pre-processing steps for activity recognition using passive radio sensing.

C. DTN ENABLED ACTIVITY RECOGNITION

The **type_ii** information often refers to body gestures or daily activities, it is non-trivial to determine a direct correlation between the processed wireless signals and various activities being examined. Taking the micro Doppler (μ-D) traces in Fig 3 for example, we can observe that different activities correspond to different μ-D patterns but it is beyond the scope of human intuition to relate these μ-D patterns and activities. Our approach therefore is employ a deep transfer network (DTN) to enable activity recognition using opportunistic wireless signals.

1) PROBLEM FORMULATION

The previous description of Section III outlines the passive radio sensing signal model and the CAF processing to obtain the range-Doppler surface. A more detailed description of the processing can be found in [4]. The μ-D provides a temporal trace of frequency vectors at a specific delay induced by a moving target. These vectors are concatenated along the time axis to generate the Doppler history (trace), which are considered to be the basis of the μ-D dataset. More specifically, suppose that the original μ-D dataset $\mathbf{D_{PWS}} = \{(X_i, y_i)\}_{i=1}^{N_{PWS}}$ is obtained with N_{PWS} samples, $X_i \in \mathbb{R}^{N_f \times N_t}$ indicates the i^{th} time-frequency representation and $y_i \in \mathbb{R}$ indicates its n_{cls}-class label. X_i is the 2-D time-frequency representations of dynamic size N_t in the time domain. Different motions may induce different time periods. Thus, we normalize it to the range of $[0, 1]$ using well-defined techniques for each recording X_i. The aim of PWS based activity recognition is to classifiy activities using the μ-DS dataset $\mathbf{D_{PWS}}$.

In this section we describe preparatory processing for enabling our μ-D based activity recognition methodology. It includes two main steps: (i) alignment of the μ-DS trace, and (ii) pre-processing methods as illustrated in Figure 2.

2) PRE-PROCESSING i: μ-D ALIGNMENT

It is essential that the data samples have the same length when preparing the μ-D dataset for the deep transfer network. It is straightforward to ensure the same number of frequency bins in X_i within our CAF processing. However, due to the large temporal variations in the duration of different human activities, it is not possible to ensure a constant number of time bins for each μ-D sample. To tackle this, and simultaneously minimize the impact of the DSI, we adopt the following two

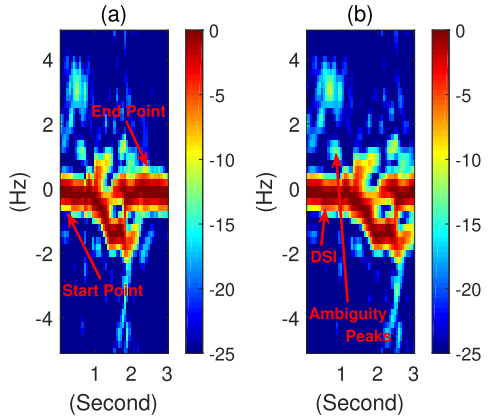


FIGURE 3. (a) Start and end point detection; (b)The useful part of the μ-D is adjusted into the same size, with the DSI and the ambiguity peak examples.

procedures which are implemented in time domain [34] as step 1 and 2 in FIGURE 2:

- Automatic start and end point detection;
- Bi-cubic interpolation to adjust the data sample size [43].

To illustrate the steps, we show an example in Figure 3 (a) and (b). The technique starts from selecting the smallest and largest time bin as the start and end point respectively in the μ-D. After the interpolation, we obtain the transformed data denoted as $X_{i,Fix} \in \mathbb{R}^{N_{freq} \times N_{time}}$, where N_{freq} is the number of maximum selected Doppler bins and N_{time} is the interpolated number of time bins. In fact, we prefer to represent the well-aligned dataset using the following Set-format, fitting the DTN classification method:

Set-Format: the original set $\mathbf{D_{PWS}} = \{(X_i, y_i)\}_{i=1}^{N_{PWS}}$ is transformed to another set: $\mathbf{D_{Fix,PWS}} = \{(X_{i,Fix}, y_i)\}_{i=1}^{N_{PWS}}$, where $X_{i,Fix} \in \mathbb{R}^{N_{freq} \times N_{time}}$.

3) PRE-PROCESSING ii: RESCALING, RESIZING AND CROPPING

Our DTN adopts an end-to-end training pipeline which is composed of feature extraction and classifier design. We represent the interpolated μ-D dataset using Set-format $\mathbf{D_{Fix,PWS}}$. The three-step pre-processing method consists of data *rescaling*, *resizing* and *cropping*, as illustrated in FIGURE 2:

Step_1-Data Re-Scaling: since the original DTN is designed for RGB images, this step linearly re-scales original values of $X_{i,Fix}$ in the range $[0, 255]$.

Step_2-Data Re-Sizing: Suppose that DTN requires the input size of $N_{height} \times N_{wid} \times 3$, we first interpolate the μ-DS data $X_{i,Fix}$ to the size of $N_{height} \times N_{wid,ext} \times 3$, where $N_{wid,ext} > N_{wid}$. Both N_{wid} and $N_{wid,ext}$ indicate the number of time bins and N_{height} indicates the number of frequency (Doppler) bins. Taking AlexNet as an example, the input size is $227 \times 227 \times 3$ so we interpolate $X_{i,Fix}$ using the bi-linear method to $227 \times 256 \times 3$ and assign the other two channels using the same μ-D.

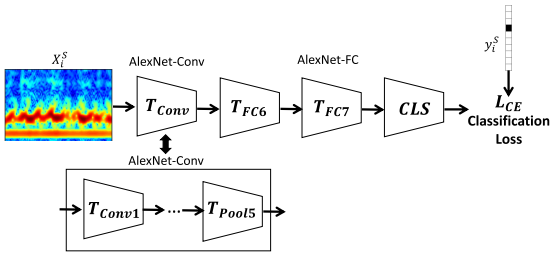


FIGURE 4. DTN for passive radio sensing based μ-D classification using AlexNet.

Step_3-Data Cropping: In the training, we randomly crop the time domain input from $N_{height} \times N_{wid,ext} \times 3$ (the dimension in AlexNet is $227 \times 256 \times 3$) to $N_{height} \times N_{wid} \times 3$ ($227 \times 227 \times 3$ in AlexNet). For the test data, we centrally crop them to the size $N_{height} \times N_{wid} \times 3$ ($227 \times 227 \times 3$ in AlexNet). The reason for not directly interpolating to $227 \times 227 \times 3$ in the AlexNet case is that randomly cropping from the 256 to 227 time bins increases the training data diversity and the generalization capability of the model. This augmentation technique is useful to extract features that are invariant to time-domain shifts and perturbations.

Finally, we denote the output of the pre-processing method as the training and test sets: $\mathbf{D^T_{PWS}} = \{X_i^T, y_i^T\}_{i=1}^{N_T}$ and $\mathbf{D^S_{PWS}} = \{X_i^S, y_i^S\}_{i=1}^{N_S}$ respectively.

4) TRANSFERRING VISION KNOWLEDGE TO μ-D CLASSIFICATION

The aim to use DTN for passive radio sensing μ-D classification is two-folds: firstly, low-level features and kernel filters used in computer vision tasks are transferable to μ-D data given that they are both 2D representations and the complexity of μ-D is less than the RGB image (there are no fine-grained texture features in RGB images); secondly, making use of pre-trained weights and fine-tuning from the ImageNet allows us to extract the hierarchical and high-level features using a very small number of training samples. To improve the results of non-deep methods, we apply the most conventional and light-weight version Alexnet for μ-D classification using pre-trained weights from ImageNet. This section focuses on applying and fine-tuning AlexNet for μ-D classification.

The simplified AlexNet architecture is shown in Figure 4, assuming the AlexNet takes μ-D X_i^S and the label y_i as inputs. The convolutional and pooling layers in AlexNet aim to extract convolutional features of the μ-D data, denoted as $T_{Conv}(X_i^S)$. Next, to predict the class categories explicitly, these spatial sensitive convolutional features are transformed to $Logits(X_i^S)$ where $Logits(X_i^S) = CLS \circ T_{FC7} \circ T_{FC6} \circ T_{Conv}(X_i^S)$ by two Fully-Connected (FC) layers T_{FC6}, T_{FC7} and the classifier CLS.

Applying a soft-max function (shown in equation (12)) on $Logits(X_i^S)$, the i_{th} element of the output $Logits_{soft}$ indicates

the probability that the input X_i^S belongs to the i^{th} class. Since the ground-truth label y_i is available, the cross-entropy (CE) loss L_{CE} in equation (13) is used to train the networks with the $L2$ regularization terms on network weights, where $[j]$ is operation to select the j_{th} element of the vector.

$$Logits_{soft} = \frac{exp(Logits)}{\sum_{j=1}^{n_{class}} exp(Logits[j])} \quad (12)$$

$$L_{CE} = -\sum_{j=1}^{n_{class}} y_i[j] \times log(Logit_{soft}[j]) \quad (13)$$

$$\min_{T_{Conv},T_{FC},CLS} L_{Total} = L_{CE} + L_{Reg} \quad (14)$$

$$i_{class} = \arg\max_j \ Logits_{soft}[j] \quad (15)$$

5) FINE-TUNING TECHNIQUES IN THE DTN

Finally, we describe the fine-tuning technique adopted in our pipeline which includes three key stages: ***first***, we initialize the network using the weights trained on ImageNet, except the *CLS* network (the weight of *CLS* is initialized based on the technique described in [5]); ***second***, we update the networks T_{FC7} and *CLS* through back-propagation but close-off the paths to the T_{Conv} layers and stop updating them until convergence; ***third***, we update the whole network using a reduced learning rate. With our fine-tuned DTN, the complex signatures can be interpreted in terms of human activity recognition. In Section IV we design and implement a series of experiments to prove how the **type_i** and **type_ii** human activity information is captured without demodulating and decoding the wireless signals in a typical smart home environment.

IV. EXPERIMENTS AND RESULTS

In this section we first describe the hardware architecture of a passive radio sensing system and the signal processing approach that permits a real-time detection output (Section IV-A). Then, we summarize our experimental testing which makes use of four bistatic geometries in line-of-sight (LoS) and through-wall scenarios, and under experimental conditions that involve using both Wi-Fi data and Wi-Fi beacon transmissions in Section IV-B. Finally, we evaluated the performance of DTN based activity recognition result in Section IV-C using a μ-DS datset of recorded activities containing six classes. The result is compared with the SRC features.

A. SYSTEM DESIGN AND IMPLEMENTATION

1) HARDWARE SYSTEM

The passive radio sensing system is built based on a software defined radio (SDR). Two synchronized Ettus^{TM}universal software radio peripherals (USRP) (model N210) are used as reference and surveillance channels to down-convert the Wi-Fi signals centered at 2.462GHz, as illustrated in Figure 5 (a) and (b). In the USRP, a 14-bit ADC is used for digitizing the intermediate frequency analog signal and followed by a Xilinx^{TM}(mode: Spartan 3A-DSP 3400) field-programmable gate array (FPGA) for digital down-converting (DDC) (Figure 5 (c)). The output of DDC is the I/Q samples is transferred into a laptop via gigabit Ethernet port for real time processing which includes CAF processing. The reference and surveillance receiver channels used in the indoor LoS experiments employ log-periodic PCB antennas with a 5 dBi gain and 60 degree beam-width [44]. In the through-wall experiments, we use two Yagi antennas with 15 dBi gain a beamwidth of 32 degree [45]. An Edimax EW-7416APn commodity Wi-Fi AP with two omni-directional antennas of 3 dBi gain was used to act as the opportunistic transmitter.

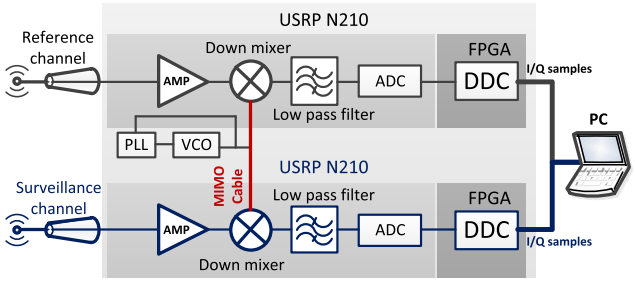


FIGURE 5. The SDR system for passive radio sensing.

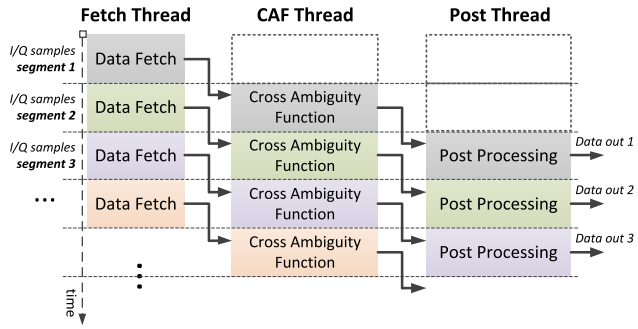


FIGURE 6. Multi-thread parallel data processing.

2) SOFTWARE FLOW

The software is mainly to handle the data flow and implemented in NI LabVIEW^{TM}. For both beacon and data-transmitting AP working modes, we sample the Wi-Fi signal at 2 MHz (which would prohibit demodulation) on both reference and surveillance channels and use a 0.5 second integration time (means each I/Q samples segment in Figure 6 has 0.5 second duration) with a 0.4 second overlap to facilitate CAF processing. These parameters are chosen based on empirical value and trade-off between the integration time (Doppler resolution) and the output frame rate. The batch processing methodology described in [4] is used to optimize the data throughput for real-time operation. The technique includes pipeline and multi-thread processing. While in LabVIEW, a parallel processing framework can be set up as shown in Figure 6. To increase the throughput of the system, the whole data flow is divided into three sub-flows which are allocated to three threads respectively. Inside of the CAF

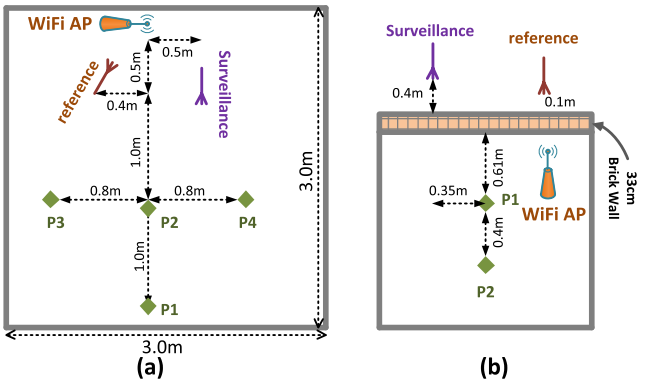


FIGURE 7. (a) Line-of-sight (LoS) experiment scenario; (b) Through-wall experiment scenario.

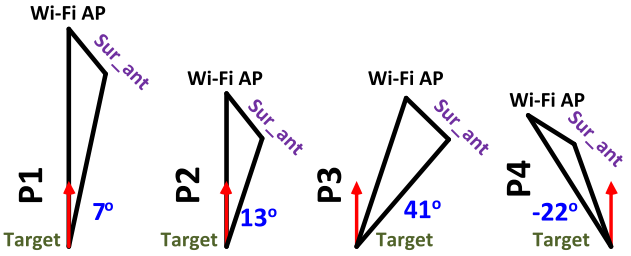


FIGURE 8. Bi-static triangles and the bi-static aspect angles (θ_{aspect} in degree) in the four test positions in the LoS scenario Figure 7 (a).

thread of Figure 6, the batch processing is (the details can be found in [4]) applied to accelerate the processing speed. When the AP works in beacon mode, we use 10 batches, while 50 batches are chosen if AP is working in data transmission mode.

B. RESPIRATION DETECTION

1) LINE OF SIGHT MEASUREMENTS

The indoor line-of-sight (LoS) experimental set up is shown in Figure 7 (a). There are four test positions (from P1 to P4) each of which define a different bistatic detection geometries, and are sketched-out in Figure 8. During the experiments, the same subject sat on a chair and breathed normally for all testing positions.

Figure 9 shows the phase extraction results for breathing detection when the Wi-Fi AP transmits beacon signals only. The phase variations caused by the chest-wall movements can be clearly observed in Figure 9. Examination of the output trace shows that inhaling causes a decrease of the relative phase, while increasing phase is observed during exhaling. It can also be seen that approximately 8 to 9 respiration cycles are detected within around 40 seconds for all testing positions. These correspond to normal human respiration rates and have been validated through video recordings. The subjects were required to breath at normal pace and magnitude in order to achieve consistent observation results. However, some of the phase variation magnitudes shows a high correlation with positions. Results from P2&3 show explicit phase variation magnitudes that are matched

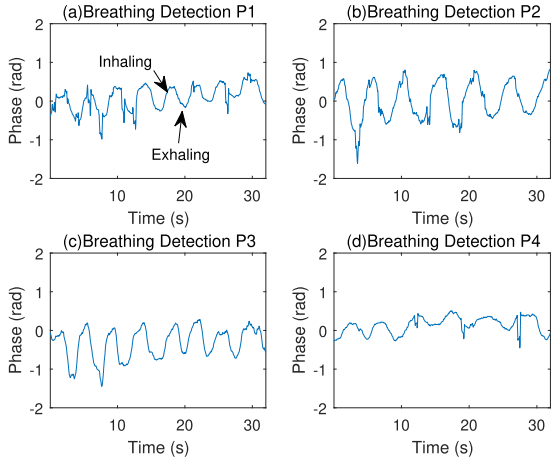


FIGURE 9. Time domain respiration capturing using beacon signals at four positions; (a), (b), (c) and (d) correspond to results obtained from P1 to P4 in Figure 7 (a) respectively.

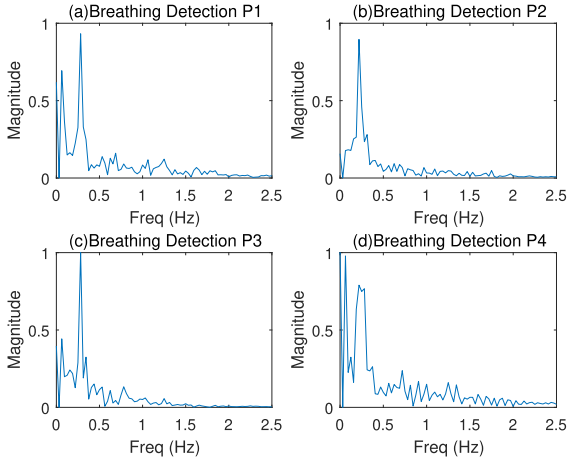


FIGURE 10. Respiration rate estimation by using beacon signals at four positions; (a), (b), (c) and (d) correspond to results obtained from P1 to P4 in Figure. 7 (a) respectively in frequency domain.

with respiration. The results conducted from P1 is in the middle, while P4 appears unfavorable for respiration detection. A number of potential reasons for this phenomenon are described in the following. First, the wavelength of Wi-Fi signal is around 12 cm so small geometrical changes due to body movements will change the phase variation. In Figure 9 (a) and (b) for example, the difference between phase deviations is approximately 1 radian, which equates to 1 cm difference in chest movement. Second, different parts of the chest may be detected which may alter the phase perturbation. Finally, although the targets are breathing normally, the chest movement is unlikely to be absolutely regular from one cycle to the next and bulk body movements will cause larger immediate phase changes, such as the red curves in Figure 9 (a) and 11 (c). By taking the Fourier transformation of the time-domain phase detection in Figure 9, we can see that observations from all four positions show peak about 0.3 Hz which is matched with the normal respiration tempo.

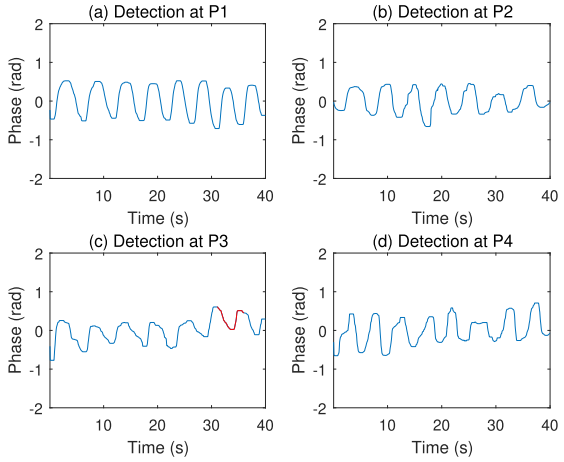


FIGURE 11. Time domain respiration capturing when AP working in data transmission mode; (a), (b), (c) and (d) correspond to results obtained from P1 to P4 in Figure. 7 (a) respectively; red curve in (c) indicates the phase variations caused by irregular chest-wall movements.

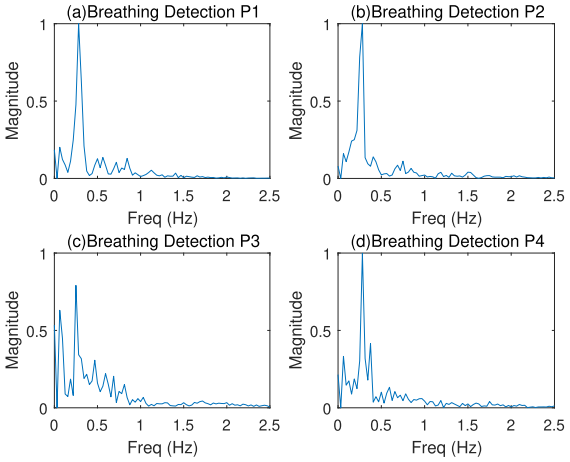


FIGURE 12. Respiration rate estimation when AP working in data transmission mode; (a), (b), (c) and (d) correspond to results obtained from P1 to P4 in Figure. 7 (a) respectively in frequency domain.

Figure 11 shows the similar LoS reparation detection when Wi-Fi AP working in data transmitting mode. In general, the results are similar to those found using the beacon signals. However, by comparing Figure 9 (a) and 11 (a) we observe that the phase variations are smoother when using data transmissions. Based on these experimental results, we estimate the average maximum phase variation within a breathing cycle for the four positions in Table.1 and match it with its corresponding aspect angle in Figure 8. These values are calculated manually based on estimating the difference between every two consecutive maxima and minima phase values. In general, it seems that increasing the bi-static aspect angle decreases the phase variations, which is in agreement with bi-static Doppler theory. More specifically, phase variation at position P1 is the largest and corresponding to the smallest aspect angle which is 7 degrees.

2) THROUGH-WALL MEASUREMENTS

The through-wall respiration detection experiments were setup in a house with a standard bricked cavity wall of 33cm thickness (Figure 7(b)). Both the reference and surveillance antennas were placed outside the house 10 and 40 cm away from the wall respectively. The Wi-Fi AP was set to data transmission mode and was located inside the room at a height of 1.15m. Two test positions (P1 and P2) were 61 cm and 101 cm respectively from the wall in the bore sight of the surveillance and reference antennas. During experiments, the subject was instructed to remain stationary and breath normally. The detection results are shown in Figure 13. The periodical phase variation patterns are similar to those seen in Figure 9 and 11 and the breathing pattern can clearly be seen. However, when target is at P2 the phase variation pattern is not as stable as that of P1, which is likely to be because of the weaker signal strength caused by the increased range of P2. We also plot the Fourier transformation of the through-wall respiration detection. More frequency scattered components can be observed from results because of the low signal to noise ratio (SNR) fact.

TABLE 1. Phase variations of the four test positions from P1 to P4 based on Figure 11 with the corresponding aspect angle and cosine value of the aspect angles.

Positions	Average Phase Variation (rad)	Aspect Angle θ_{aspect} (deg)	$cos(\theta_{aspect})$
P1	1.05	7	0.993
P2	0.96	13	0.974
P3	0.80	41	0.755
P4	0.86	22	0.927

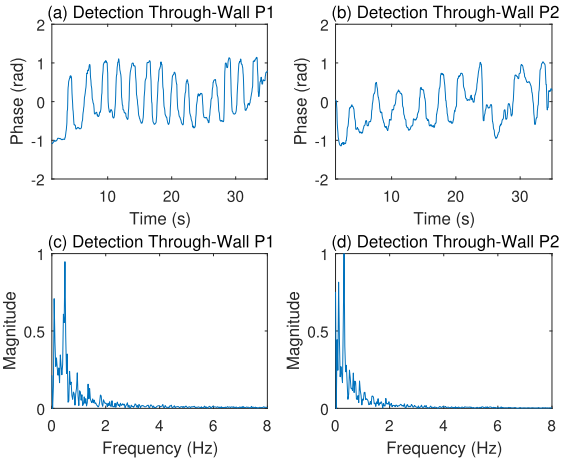


FIGURE 13. Time domain respiration capturing and rate estimation results in the through-wall scenario (a), (c) at P1 (b), (d) at P2 in Figure. 7 (b), both in time and frequency domains.

C. ACTIVITY RECOGNITION

In this section, we first introduce and illustrate the μ-DS generated by our passive radio sensing system, then show the results of "start-end" point detection method. Three experimental settings are implemented, which use 20%, 40% and 60% of the dataset are used for training respectively. Finally, to ensure fair comparisons, we compare the DTN

TABLE 2. Methods under evaluation for μ-D classification and their descriptions.

Methods	Description
PCA Feature + SVM	PCA feature and apply SVM classifier.
PCA Feature + SRC	We extract PCA feature and apply SRC classifier.
AlexNet + FC + FT (DTN 1)	Fine-tune the final FC layers and classifier freeze the Conv layer weights.
AlexNet + WholeNet + FT (DTN 2)	Fine-tune the whole network using the strategy in Section III-C5.

TABLE 3. Dataset description of the six activities in PWS experiments. M1(20) indicates the first motion including 20 μ-DS samples.

Activity Index	Description
M1(20)	Subject pickes up from the ground and stand up
M2(20)	Subject sits down on a chair
M3(20)	Subject stands up from a chair
M4(57)	Subject falls down onto the mattress
M5(62)	Subject stands up after falling
M6(10)	Subject lies on a mattress first then gets out of it

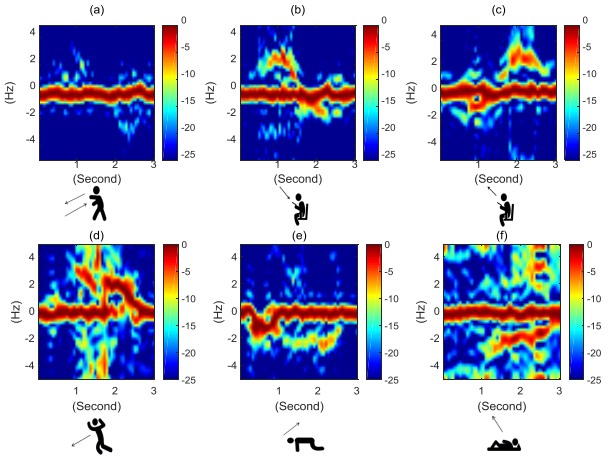


FIGURE 14. Overview of PWS μ-Ds in dB scale. (a) to (f) are the μ-Ds of motion 1 (M1) to 6 (M6) in Table 3.

based performance with PCA feature based SVM and SRC classifications, which are listed in Table 2.

1) PASSIVE RADIO SENSING μ-D ANALYSIS

In this section, a μ-D dataset is collected for 6 classes of typical everyday human activities as described in Table 3. We also show the corresponding μ-DS in Figure 14.

The six μ-Ds exhibit different patterns and their visual discriminative characteristics are outlined as: ***i)***, The maximum Doppler shift; ***ii)***, Time duration of the μ-D; ***iii)***, Switches between negative to positive or vice-versa, if any; ***iv)***, The magnitude of the zero Doppler line caused by the DSI or multipath.

In general, the maximum Doppler frequencies of these six μ-DS range from 2Hz to 4.5Hz. The second discriminative feature relates to the relative direction of motion, indicated by the sign (positive or negative) of the Doppler shift frequency: some motions induce Doppler frequencies that transverse from positive to negative, (e.g. M1 and M2), while others induce only positive or negative Doppler frequencies (M3, M4, M5 and M6). Although M1 and M2 both have the similar patterns (from positive to negative), the time duration of each signature segment increases the discriminative characterictics of the feature, such as the shorter duration of the positive Doppler frequency in M2 compared to the positive Doppler frequency in M1. The final distinguishable feature is the presence of the zero Doppler line during the motion. A clear example is the comparison between the output for M5 and M6, where the μ-Ds patterns are similar, but the latter has a stronger zero Doppler line. The reason why M5 exhibits no zero Doppler line can be attributed to a shielding effect on the DSI when the subject gets up from the floor. Although these selected empirical features agree closely with the intuitive visual interpretation, obtaining them accurately requires complex feature selection methods such as detecting accurate Doppler patterns. In addition, these methods are prone to errors which in turn can distort the outcome of the classifier. For the classification scheme using SRC, we utilize the reduced-dimensionl data vectors after the PCA operation. For selecting features for DCNN, we simply input the original μ-D according to methods introduced in Section III-C4.

TABLE 4. Activity recognition results for different features and classifiers, percentage in (%).

Feature + Classifier	Train on 20%	Train on 40%	Train on 60%
PCA + SVM	32.9	57.0	60.0
PCA + SRC	82.0	81.0	88.0
AlexNet+FC+FT	80.3	90.5	93.5
AlexNet+WholeNet+FT	86.3	94.7	98.7

2) ACTIVITY RECOGNITION

The recognition results shown in Table 4 shows that the PCA feature based SRC outperforms average by 28% on SVM with same features. Compared with the PCA feature based SRC, the *AlexNet+FC+FT* approach achieves an improved performance when using 40% and 60% of the data for training. While, PCA feature based SRC outperforms *AlexNet+FC+FT* approach by around 1.7% if 20% of the dataset for training. In addition, other DTN methods e.g. *AlexNet+WholeNet+FT* fine-tuning the whole networks achieve at least 11% improvement than SRC. This result suggests that shallow methods, like SRC, is suitable for handling small training datasets, while the DTN framework achieves a superior performance compared to shallow methods only if the whole network is fine-tuned. Furthermore, the result indicates that adapting and changing local features in *Conv* layers is essential for the DTN.

The example confusion matrix of *AlexNet+WholeNet+FT* based methods in Table 5 shown that we can still achieve comparable high recognition rate even with only 20% data for training. As can be seen from the confusion matrix that the most misclassifications happen from M2→M3 and between M5 & 6. It is not difficult to interpreting this result by

TABLE 5. Confusion matrix of *AlexNet+WholeNet+FT* based activities classification performance using 20% data for training, percentage in (%).

		Prediction					
		M1	M2	M3	M4	M5	M6
Actual	M1	94.1	5.9	0.0	0.0	0.0	0.0
	M2	0.0	68.8	31.2	0.0	0.0	0.0
	M3	0.0	0.0	100.0	0.0	0.0	0.0
	M4	6.5	0.0	2.1	84.8	2.1	4.3
	M5	0.0	0.0	0.0	2.0	86.0	12.0
	M6	0.0	0.0	0.0	0.0	12.5	87.5

observing the μ-Ds in Figure 14. The μ-Ds of M2 and M3 are very similar in shape, just present in reversed order. Both M5 and M6 contain the action of getting up from lying down pose. This observation also indicates that introducing the time sequential analysis in the models of the μ-Ds traces will further improve the recognition rate in future works.

V. CONCLUSION AND FUTURE WORK

In this paper, we have examined two use-cases to demonstrate that wireless signal bursts in ISM band can be exploited to collect signs-of-life and human activity information without actually performing any type of signal demodulation or decoding. A summary of the key research achievements for these two cases is given below:

Respiration Detection: A novel phase-sensitive processing system that works in real-time to mitigate the requirement for long coherent integration times in conventional passive radio sensing technology is proposed and implemented. We utilized the instantaneous Doppler which is the phase output of the CAF processing. This method was evaluated and discussed with respect to measured data from various geometrical layouts and LoS conditions.

Activity Recognition: A DTN approach is proposed and tested based on of the six daily activities μ-Ds dataset. We observed that our fine tuning DTN approach outperforms the PCA feature based shallow classifiers like SVM, SRC by 9% on average.

Our overall conclusions are that, using the methods we have develped, human respiration can be observed directly using phase measurements from wireless bursts. In addition, an interpreting strategy like a DTN is required in order to classify very small fine-tuned features and the irregular μ-Ds patterns generated from daily activities. Passive radio sensing is technology that is only now beginning to mature and this work takes it a step further towards real-world applications by proving its ability to identify signs-of-life and recognize everyday human activities by using ambient wireless communications signals without demodulation or decoding. We expect future work in this area to be directed towards: **i).** data processing with multiple wireless sources; **ii).** multiple targets discrimination; **iii).** performance optimisation by assessing the impact of burst frequency and waveform characteristics; **iv).** cross disciplinary approaches for large scale data collection to improve machine learning performance.

REFERENCES

[1] C. L. Hogan, L. I. Catalino, J. Mata, and B. L. Fredrickson, "Beyond emotional benefits: Physical activity and sedentary behaviour affect psychosocial resources through emotions," *Psychol. Health*, vol. 30, no. 3, pp. 354–369, Mar. 2015.
[2] R. R. Pate, G. W. Heath, M. Dowda, and S. G. Trost, "Associations between physical activity and other health behaviors in a representative sample of US adolescents.," *Amer. J. Public Health*, vol. 86, no. 11, pp. 1577–1581, Nov. 1996.
[3] N. Zhu, T. Diethe, M. Camplani, L. Tao, A. Burrows, N. Twomey, D. Kaleshi, M. Mirmehdi, P. Flach, and I. Craddock, "Bridging e-Health and the Internet of Things: The SPHERE project," *IEEE Intell. Syst.*, vol. 30, no. 4, pp. 39–46, Jul. 2015.
[4] B. Tan, K. Woodbridge, and K. Chetty, "A real-time high resolution passive WiFi Doppler-radar and its applications," in *Proc. Radar Conf.*, Dec. 2014, pp. 1–6.
[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
[6] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
[7] M. Chan, D. Estève, C. Escriba, and E. Campo, "A review of smart homes—Present state and future challenges," *Comput. Methods Programs Biomed.*, vol. 91, no. 1, pp. 55–81, Jul. 2008.
[8] C. D. Kidd, R. Orr, G. D. Abowd, C. G. Atkeson, I. A. Essa, B. MacIntyre, E. Mynatt, T. E. Starner, and W. Newstetter, "The aware home: A living laboratory for ubiquitous computing research," in *Proc. Int. Workshop Cooperat. Buildings*. Berlin, Germany: Springer, 1999, pp. 191–198.
[9] M. Perry, A. Dowdall, L. Lines, and K. Hone, "Multimodal and ubiquitous computing systems: Supporting independent-living older users," *IEEE Trans. Inf. Technol. Biomed.*, vol. 8, no. 3, pp. 258–270, Sep. 2004.
[10] A. Adami, T. Hayes, and M. Pavel, "Unobtrusive monitoring of sleep patterns," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Jun. 2004. pp. 1360–1363.
[11] S. Das, D. Cook, A. Battacharya, E. Heierman, and T.-Y. Lin, "The role of prediction algorithms in the MavHome smart home architecture," *IEEE Wireless Commun.*, vol. 9, no. 6, pp. 77–84, Dec. 2002.
[12] Y. Isoda, S. Kurakake, and H. Nakano, "Ubiquitous sensors based human behavior modeling and recognition using a spatio-temporal representation of user states," in *Proc. 18th Int. Conf. Adv. Inf. Netw. Appl. (AINA)*, Jun. 2004, pp. 512–517.
[13] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale, and S. Shafer, "Multi-camera multi-person tracking for EasyLiving," in *Proc. 3rd IEEE Int. Workshop Vis. Surveill.*, Nov. 2002, pp. 3–10.
[14] D. Riedel, S. Venkatesh, and W. Liu, "Spatial activity recognition in a smart home environment using a chemotactic model," in *Proc. Int. Conf. Intell. Sensors, Sensor Netw. Inf. Process.*, 2005, pp. 301–306.
[15] N. Barnes, D. Rose, P. Garner, and N. Edwards, "Lifestyle monitoring–technology for supported independence," *Comput. Control Eng. J.*, vol. 9, no. 4, pp. 169–174, Aug. 1998.
[16] M. Chan, C. Hariton, P. Ringeard, and E. Campo, "Smart house automation system for the elderly and the disabled," in *Proc. 21st IEEE Int. Conf. Syst., Man Cybern. Intell. Syst.*, Nov. 2002, pp. 1586–1589.
[17] T. Yamazaki, "The ubiquitous home," *Int. J. Smart Home*, vol. 1, no. 1, pp. 17–22, 2007.
[18] J. Wilson and N. Patwari, "Radio tomographic imaging with wireless networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 5, pp. 621–632, May 2010.
[19] Y. Gu, F. Ren, and J. Li, "PAWS: Passive human activity recognition based on WiFi ambient signals," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 796–805, Oct. 2016.
[20] B. Mrazovac, B. Todorović, M. Bjelica, and D. Kukolj, "Device-free indoor human presence detection method based on the information entropy of RSSI variations," *Electron. Lett.*, vol. 49, no. 22, pp. 1386–1388, Oct. 2013.
[21] M. Scholz, T. Riedel, M. Hock, and M. Beigl, "Device-free and device-bound activity recognition using radio signal strength," in *Proc. 4th Augmented Human Int. Conf.*, New York, NY, USA, 2013, pp. 100–107.
[22] H. Abdelnasser, M. Youssef, and K. A. Harras, "WiGest: A ubiquitous WiFi-based gesture recognition system," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2015, pp. 1472–1480.

[23] L. Sun, S. Sen, D. Koutsonikolas, and K.-H. Kim, "WiDraw: Enabling hands-free drawing in the air on commodity WiFi devices," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, New York, NY, USA, 2015, pp. 77–89.

[24] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proc. 19th Annu. Int. Conf. Mobile Computing Netw.*, New York, NY, USA, 2013, pp. 27–38.

[25] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, New York, NY, USA, 2015, pp. 65–76.

[26] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Keystroke recognition using WiFi signals," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, New York, NY, USA, ACM, 2015, pp. 90–102.

[27] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3D tracking via body radio reflections," in *Proc. 11th USENIX Conf. Netw. Syst. Design Implement.*, Berkeley, CA, USA, 2014, pp. 317–329.

[28] F. Adib, Z. Kabelac, and D. Katabi, "Multi-person localization via RF body reflections," in *Proc. 12th USENIX Conf. Networked Syst. Design Implement.*, Berkeley, CA, USA, 2015, pp. 279–292.

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

[30] Y. Zeng, P. H. Pathak, Z. Yang, and P. Mohapatra, "Poster Abstract: Human tracking and activity monitoring using 60 GHz mmWave," in *Proc. 15th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2016, pp. 1–2.

[31] K. Chetty, G. E. Smith, and K. Woodbridge, "Through-the-wall sensing of personnel using passive bistatic WiFi radar at standoff distances," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1218–1226, Apr. 2012.

[32] B. Tan, K. Woodbridge, and K. Chetty, "Awireless passive radar system for real-time through-wall movement detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 5, pp. 2596–2603, Oct. 2016.

[33] W. Li, B. Tan, Y. Xu, and R. J. Piechocki, "Log-likelihood clustering-enabled passive RF sensing for residential activity recognition," *IEEE Sensors J.*, vol. 18, no. 13, pp. 5413–5421, Jul. 2018.

[34] Q. Chen, B. Tan, K. Chetty, and K. Woodbridge, "Activity recognition based on micro-Doppler signature with in-home Wi-Fi," in *Proc. IEEE 18th Int. Conf. e-Health Netw., Appl. Services*, Sep. 2016, pp. 1–6.

[35] W. Li, B. Tan, and R. J. Piechocki, "Non-contact breathing detection using passive radar," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[36] Q. Chen, K. Chetty, K. Woodbridge, and B. Tan, "Signs of life detection using wireless passive radar," in *Proc. IEEE Radar Conf.*, May 2016, pp. 1–5.

[37] J. E. Palmer, H. A. Harms, S. J. Searle, and L. Davis, "DVB-T passive radar signal processing," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 2116–2126, Apr. 2013.

[38] P. M. Woodward, *Probability and Information Theory, with Applications to Radar: International Series of Monographs on Electronics and Instrumentation*, vol. 3. Amsterdam, The Netherlands: Elsevier, 2014.

[39] P. Falcone, F. Colone, C. Bongioanni, and P. Lombardo, "Experimental results for OFDM WiFi-based passive bistatic radar," in *Proc. IEEE Radar Conf.*, May 2010, pp. 516–521.

[40] F. Colone, P. Falcone, C. Bongioanni, and P. Lombardo, "WiFi-based passive bistatic radar: Data processing schemes and experimental results," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1061–1079, Apr. 2012.

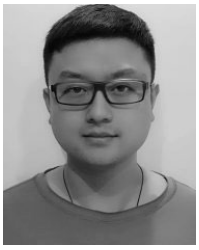
[41] R. Pearson, "Outliers in process modeling and identification," *IEEE Trans. Control Syst. Technol.*, vol. 10, no. 1, pp. 55–63, Aug. 2002.

[42] J. D. Emerson and D. C. Hoaglin, *Understanding Robust and Exploratory Data Analysis*. Hoboken, NJ, USA: Wiley, 1983.

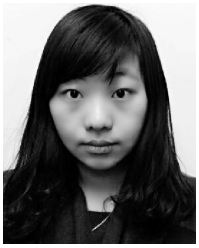
[43] C. De Boor, C. De Boor, and E.-U. Mathématicien, *A Practical Guide to Splines*, vol. 27. New York, NY, USA: Springer-Verlag, 1978.

[44] E. Research. (2009). *LP0965 Antenna Specification*. [Online]. Available: https://www.ettus.com/product/details/LP0965

[45] S. W.-F. Network. (2009). *Yagi Antenna Specification*. [Online]. Available: https://www.simplewifi.com/products/yagi



QINGCHAO CHEN (Member, IEEE) received the B.Sc. degree in telecommunication engineering from the Beijing University of Post and Telecommunication and the Ph.D. degree from the University College London. He is currently a Postdoctoral Researcher with University of Oxford, U.K. His current researches focus on computer vision and machine learning, radio-frequency signal processing and system design, and biomedical multimodality data analysis.



YANG LIU received the B.Sc. degrees in telecommunication engineering from the Beijing University of Posts and Telecommunications, Beijing, China, and the Ph.D. degree in computer science from the Computer Laboratory, University of Cambridge, Cambridge, U.K. She is a Postdoctoral Researcher with University of Oxford, U.K. Her current research interests include pattern recognition, computer vision, and applied machine learning.

BO TAN (Member, IEEE) received the Ph.D. degree in digital communication from the University of Edinburgh, U.K. He was Postdoctoral Researcher of radar signal processing and machine learning of wireless signals for healthcare with University College London (UCL) and the University of Bristol, U.K. He is currently an Assistant Professor (Tenure Track) with Tampere University, Finland. His research focuses on wireless communications and sensing. He was the Co-Investigator and a participant of U.K., EPSRC, Innovate U.K., EU ERC Projects. He has over 40 academic publications in leading the IEEE, IET, and ACM venues, also the patented technological innovation in passive wireless sensing.

KARL WOODBRIDGE (Senior Member, IEEE) is currently an Emeritus Professor of electronic and electrical engineering with University College London. He is also an Honorary Professor with the School of Engineering, University of Birmingham, U.K. His research interests include multistatic and software-defined radar systems, passive wireless surveillance, and Doppler classification using machine learning methods, development of passive wireless based sensors for activity detection and classification with application to Healthcare, and the IoT and Security. He has published or presented over 200 journal article and conference papers in the areas of semiconductors, photonics, and RF sensor systems. He is also a Fellow of the IET and U.K. Institute of Physics.

KEVIN CHETTY is currently an Associate Professor with University College London. He is the author of over 60 peer reviewed publications. His research expertise includes the field of radio frequency sensing, and radar signal processing using machine learning. Major achievements in RF sensing include demonstrating the first detections of personnel targets using passive WiFi radar, and proving the ability of these systems to perform through the wall sensing at standoff distances. His works include radar micro Doppler signatures, indoor mapping, target tracking, and high-throughput data processing. He has been an Investigator on grants funded by Government and industry. He is also sits on the Technical Committee for the International Security and Crime Science conference. He is also a Reviewer for a number of the high-esteem IEEE and IET journals.