Radar Signal Processing for Sensing in Assisted Living

Julien Le Kernec[1,3,4*], Senior Member, IEEE, Francesco Fioranelli[1], Member, IEEE, Chuanwei Ding[2], Student, IEEE, Heng Zhao[2], Student, IEEE, Li Sun[2], Member, IEEE, Hong Hong[2], Member, IEEE, Jordane Lorandel[3], Member, IEEE, Olivier Romain[3], Member, IEEE,

Abstract: This article presents radar signal processing for sensing in the context of assisted living. This is covered through 3 example applications: human activity recognition for activities of daily living, respiratory disorder and Sleep Stages classification. The common challenge of classification is discussed within a framework of measurements/pre-processing, feature extraction, and classification algorithms for supervised learning. Then, the specific challenges of the 3 applications from a signal processing standpoint are detailed in their specific data processing and ad-hoc classification strategies, focusing on recent trends in the field of activity recognition (multi-domain, multi-modal and fusion) and healthcare applications based on vital signs (super-resolution techniques) and commenting on outstanding challenges. To conclude, this paper explores the challenge of the real-time implementation of signal processing/classification algorithms.

Index terms: Radar, Classification, Neural Networks, Healthcare, Assisted Living

I. INTRODUCTION

Recently, vital sign detection and activity recognition have captivated the research community especially in assisted living (AL) [1-4], for aging in place (Figure 1) and hospitals to monitor in/out-patients e.g. pulmonary obstruction disorder (POD). Aging population is a global challenge, with an increasing proportion of older/vulnerable people living alone and at high risk of falling (30%) [1]. Thus, societies will have to adapt to rising health challenges from managing chronic/cognitive diseases and preserving mobility after a stroke, to the need for rehabilitative technologies.

F. Fioranelli and J. Le Kernec (JLK) are with the School of Engineering, University of Glasgow, UK. JLK is also affiliated with the School of Communication and Information, University of Electronic Science and Technology of China, Chengdu (julien.lekernec@glasgow.ac.uk). C. Ding, H. Zhao, L. Sun and H. Hong are with the Nanjing University of Science and Technology, China. J. Lorandel, JLK and O. Romain are with ETIS-ASTRE lab at University Cergy-Pontoise, France. They acknowledge support from the UK EPSRC EP/R041679/1, PHC/CSC Cai Yuanpei - 41457UK, No. 201806070002, and NSFC Grant No. 61871224.

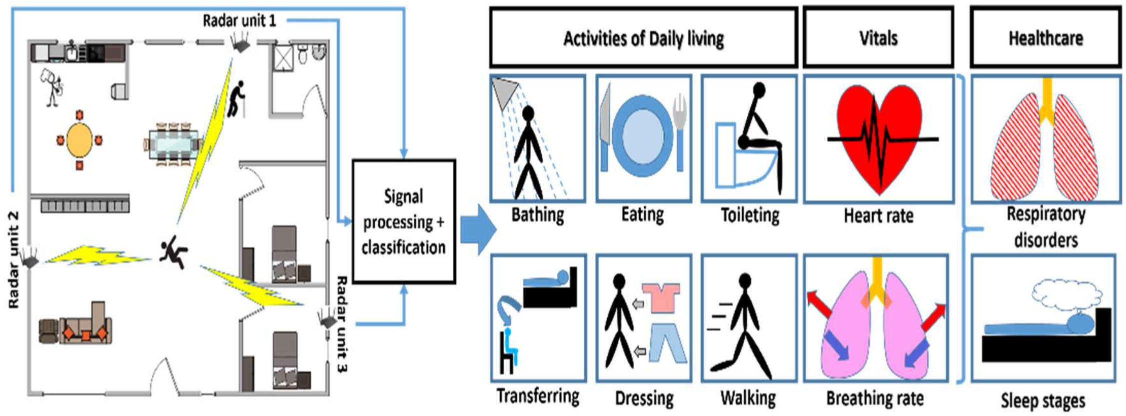

Figure 1: In-home sensing scenario and extracted information for AL.

Radar has gained popularity as a sensing technology to support innovative/personalized healthcare services [1-4]. Why Radar? Many sensors have been investigated for AL [5], ranging from wearable sensors (e.g. accelerometers, gyroscopes, optical for heart rate (HR)/oximetry), sensors embedded in the built environment (e.g. pressure, acoustic, infrared sensors), to cameras based on visible/infrared light, or depth information. Amongst these, the attractiveness of radar in AL lies in its contactless nature, versatility, and privacy preservation. People under observation need not wear/interact with devices or comply with instructions that would change their routines. Radar would not be perceived as a stigma, as it can be beautified with a personalized radome fixed on a wall that could look like a painting of the person's choosing.

This article considers the aging-in-place scenario where the common challenge in those applications is classification, which is introduced in section II for supervised learning. Activities of daily living (ADLs) through human activity recognition (HAR) are monitored, as ADLs are basic self-care activities and part of an assessment to maintain a person living independently at home. Vital signs are also monitored for healthcare applications: namely breathing disorder (BD) and sleep stage (SS) recognition. Section III and IV respectively present the specific challenges pertaining to the applications of HAR for AL, and RD/SS recognition respectively. Finally, section V covers open challenges for real-time implementation of classification algorithms. Radar can also be applied in different healthcare applications [4] (e.g. cane-assisted walking analysis, monitoring of bed-ridden patients), but these are beyond the scope of this article.

II. SUPERVISED CLASSIFICATION PROBLEM

What is supervised classification, i.e. recognition of patterns? Given a new, unknown test data sample, this will be assigned to a class/category predefined by the researcher/engineer/clinician [6]. The ever-increasing availability of computational power with GPUs, FPGAs and CPUs has generated a plethora of classification techniques. It becomes apparent from the literature that there is to date no specific "one-size-fits-all" supervised classification algorithm.

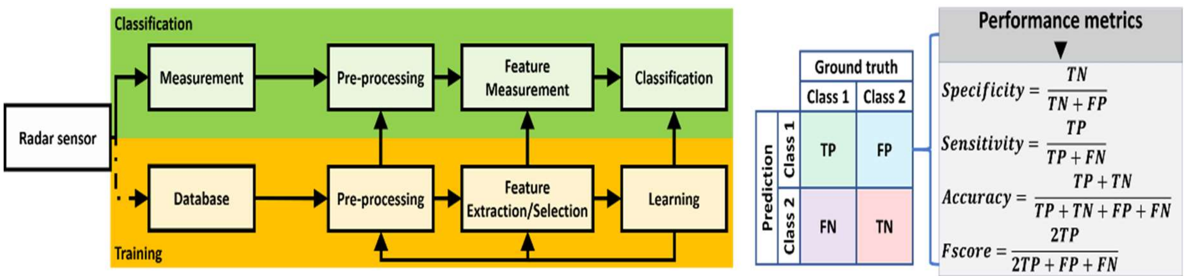

Figure 2: right) Common high-level framework of classification applications – adapted from [6]; left) Classification example for a binary class problem and performance metrics for class 1.

A common framework (Figure 2) encompassing statistical learning (SL) and deep learning (DL) approaches for supervised learning, comprises three steps: data acquisition/signal pre-processing, feature extraction, and classification. The pre-processing is linked with the nature of the problem, the sensing modalities, how the data is represented and inputted in classification algorithms. The "optimal" data representation is still being intensely researched to find the best classification performance and evaluate its "portability" across scenarios/applications. The problem is how to differentiate between classes, given a set of intra-/inter-class information that can lead to minimal variations within each class, and sufficiently significant variations between different classes to define boundaries for simple classification strategies.

In SL and DL approaches, one point with N features in a N-dimensional space is representative of a class. The goal is to find the features that will allow a simple decision boundary between classes. The boundary to categorize new samples is determined by the probability distributions of each class that is learned via training on a database of samples. It can be defined differently based on the adopted SL, for example Bayes' theory (NB), K-Nearest Neighbors (K-NN), Support Vector Machines (SVM), Decision Trees, Neural Networks (NN) all described in [7]; or based on DL such as Stacked-Auto Encoders (SAE), Convolutional Neural Networks (CNN), Recurrent NN (RNN) including Gated Recurrent Units (GRU), and Long Short-Term Memory (LSTM) listed in [1]. Essentially, all methods implement a form of cost/loss function minimization of a new sample x belonging to a class c_i. In supervised learning, the performance of a classifier is then assessed in terms of the sensitivity, specificity, accuracy and/or F1-score (Figure 2).

DL has been recently proposed for radar data classification to leverage the disruptive effect that its adoption had in image/video/audio signal processing [8]. Its attractiveness lies in the possibility to bypass the need of "expert knowledge and operators" for the pre-processing, feature extraction and/or selection stages, as the NNs decide the most salient features autonomously for classification.

In both SL/DL, the researcher is confronted with the "curse of dimensionality", i.e. how much input data is necessary for robust classification algorithms, and specifically how many features (and which ones) are significant and not over-fitted to a specific scenario or dataset. In DL, an added degree of freedom is the adjustment of hyperparameters such as learning rate, dropout rate, iterations, the number of hidden units/layers, the activation functions. A large number of data/features/hyperparameters values have the potential to make the computational requirements explode, which will be discussed in Section V.

Regarding features, selecting the most salient ones can have a greater impact on accuracy than the selected SL/DL algorithm, while reducing its complexity and computational requirements. Feature selection methods [4] include Principal Component Analysis (PCA), "wrapper" methods which test different combinations with a specific classifier to select the best performing features, and "filter" methods which rank them based on inter/intra-class information metrics e.g. entropy and T-tests. In DL, convolutional filters are used in CNN and CAE as general-purpose feature extractors creating a sparse representation of the input data accounting for their structure. SAE is similar but restructures input data as a column vector. The feature selection process is hidden, i.e. the algorithm will automatically extract/select the salient features to minimize the cost function. The dropout rate can force more stringent feature selection in the performance optimization.

III. Human activity recognition for AL

Research in HAR for AL using radar have risen considerably due to the challenge of aging population that the developed/developing countries are now facing [1,4]. This section offers a quick overview of the most recent research trends and outstanding challenges. In radar, classification work is still primarily performed on spectrograms (Figure 3), i.e. micro-Doppler signatures showing Doppler-time representations of the received signals based on Short-Time Fourier Transform (FT) [1,4]. This approach offers simple and effective data representation for problems presenting large interclass variations. However, when "*confusers*" (activities with low interclass variations) are introduced or high-frequency accuracy is required for data representation, then spectrograms show their limitations in time/frequency resolution. Spectrograms were popular, because only CW radar were affordable for research and the computational requirements for FT-based algorithms were low compared to other time-frequency (TF) transforms. However, Short-time FT (STFT) suffers from TF resolution trade-off – they cannot be jointly optimized. More generalized TF distributions were proposed to address this issue and are discussed in [4].

There is therefore an interest in exploring other radar domains for classification and the Cadence Velocity Diagram (CVD) is an easily achievable one (even for CW radar) from Spectrograms through a simple FT, which reveals the average walking speed and stride rate for example. For spectrograms and CVD, a specific subset of features is chosen for a given application and dataset. How to make this selection given the different operational conditions such as radar parameters (e.g. Pulse Repetition Frequency, carrier frequency), deployment and geometry constraints (aspect angle, obscuring of body parts, Doppler spread), and SNR is an outstanding research problem. Furthermore, dynamic changes to the operational conditions may require the radar to dynamically adapt its features ("feature diversity", Chapter 8 [4]). Table I summarizes findings from simulations on the robustness of different features based on operational parameters for classification, which may provide guidance in selecting among a more extensive list of features - physical, transform-based, speech inspired, and non-parametric - found in the literature for gait classification.

Experimentally, PRF plays an important role in classification accuracy and should not exceed too much the Doppler spread of the received signals at the risk of degrading performances.

Since the democratization of sub-metric accuracy radar stemming from the automotive industry for anti-collision radar and software-defined radio platforms, accompanied by an increase in computational power, more information domains are available from which to extract features for classification (Figure 3) ranging from raw data [1], range-time [1,9], range-Doppler [9,10], Cadence Velocity Diagrams [4] or Cepstrogram, Mel-frequency [4] or a combination such as composite range-Doppler maps [9-11]. Activities can now be distinguished in terms of distance, power spectrum, Doppler and Cadence to further discriminate between low inter-class variations. However, this poses also challenges regarding the search for the salient features across different domains and the implementation of fusion techniques between multiple domains [9-11], other radar nodes [12], or different sensing modalities complementing radar [13-14].

Table I. Features robustness in classification based on operational parameters, Chapter 12 [4]

Features	Overall	Physical	LPC	DCT	Cepstrum
f_c (GHz)	N/A	Stable > 10	Stable > 10	Best at 10-20	Varies
PRF	Stable	Idem overall	Idem overall	>1.5 kHz	Idem overall
SNR	Follows SNR degradation	more affected	Idem overall	Idem overall	more affected
Aspect angle	Degradation from 0 º to 90º	Idem overall	Most affected	Least drop	Most affected
Dwell time t_d	Better performance with longer t_d	Most affected	Idem overall	Good even with short t_d	Idem overall

*f_c: Operational frequency, PRF: pulse repetition frequency, SNR: Signal-to-Noise Ratio

Information fusion is used to overcome limitations of one domain/sensor by merging complementary information from different sensors at different abstraction levels: signal, feature and/or decision [15]. Fusion generally improves the overall classification performances compensating the shortcomings of a single sensor and can increase sensitivity and specificity related to a specific class of interest, for example falls in human activity recognition [13-14]. Three examples using multi-domain information for classification are summarized in Table II.

In [9], a combination of range-time, spectrogram and integrated range-Doppler (IRD) information (see Figure 3) are processed with SAEs to extract features that are then classified by a

Softmax layer for each of the 3 inputs. A vote takes place then to define the activity label (walk, fall, sit, bend). This method displays an overall improvement of 3% in accuracy, 2% in specificity and ~6% in sensitivity. However, the sensitivity in detection of falls was reduced compared to individual inputs, indicating that the voting strategy should be revised to maximize fall detection.

In [10], the range-Doppler cube is pre-processed with an extended CLEAN algorithm to eliminate unwanted noise/distortions while enhancing the signal. A multi-dimension PCA (MPCA) approach is proposed reaching 97.88% accuracy outperforming handcrafted features and PCA-based classification by 10 and 11.8 % respectively. MPCA maintains the variations in the projected data, minimizes reconstruction error, and is equivalent to PCA or 2D-PCA when M = 1,2. This is a very interesting technique to explore the variety of domains now open where M is greater than 2.

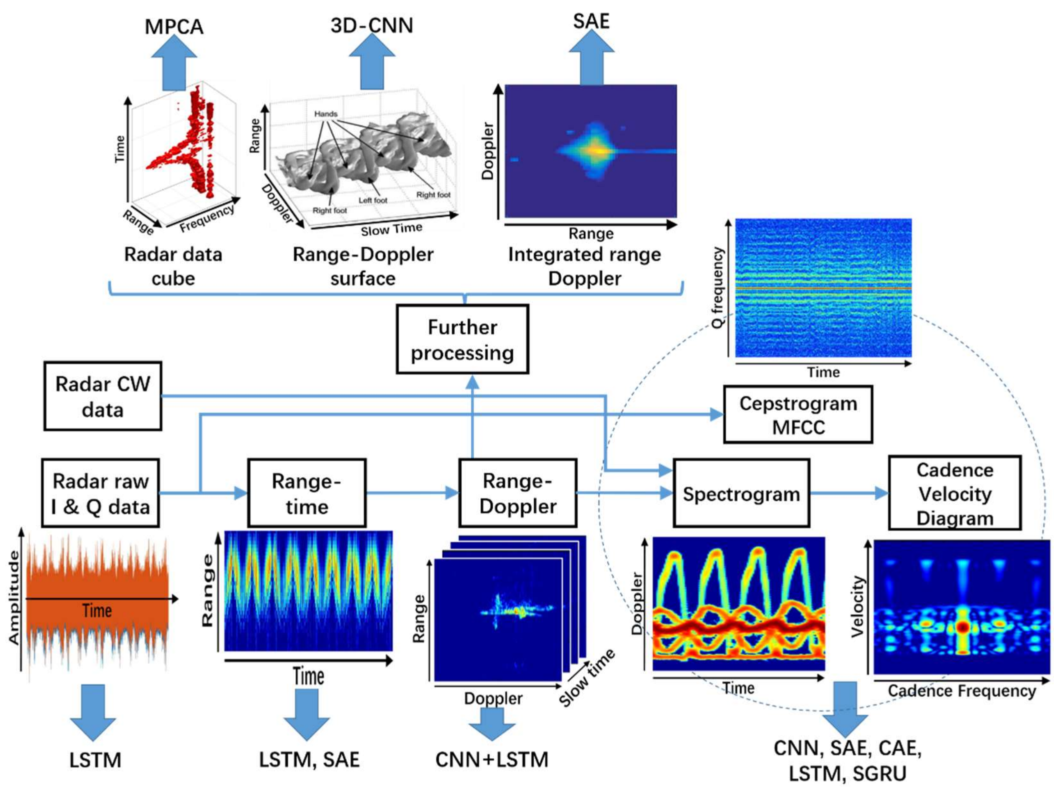

Figure 3: Information domains of radar data used for human activity classification in the literature, from raw IQ data to more complex integrated 3D cubes and range-Doppler surfaces

In [16], the range information is used to perform a binary classification between in-situ and non-in-situ activities based on the number of activated range bins with 99.9% accuracy. After performing weighted range TF transform, 2 PCA-based approaches were used yielding respectively

98.1% accuracy (non-in-situ, bagged trees classifier) and 98.5% (in-situ, subspace KNN classifier). The centroid features were abandoned for in-situ motion since they were not relevant whereas they are very salient for walk. This demonstrates a novel hierarchical multi-domain classification and that different features perform better depending on the nature of the activity.

Table II: Summary of performance of sample papers for multi-domain classification [9,10,16]

Metric	**Multi-domain**	**Spectro-gram**	**Range time**	**IRD**	**expert features**	**PCA**	**MPCA**	**In situ vs non in situ**	**In situ**	**Non-in-situ**
Sensitivity	96.43	89.8	88.7	90.62	92.88	91.14	99.12	99.74	94.36	95.38
Specificity	98.81	96.6	96.23	96.88	83.91	82.03	96.7	100	98.87	99.08
Accuracy	98.21	94.9	94.35	95.31	87.87	86.01	97.88	99.87	98.12	98.46
# of subjects	3				4			13		
# of motions	4				4			12		
# of samples	96				385			780		

Table III: Classification performance of sample papers for fusion for the activities

	Feature level fusion SFS [14]			**Feature level fusion – no SFS [14]**			**Fusion: rad+inertial sensor [13]**			**FL fusion with SFS**		
Metric	**Mag**	**rad**	**Fusion**	**rad**	**Rad +acc.**	**Rad+ acc+Ki**	**Rad**	**FL**	**DecL**	**CW**	**FMCW**	**Fused**
Sensitivity	92.75	91.61	96.95	68.75	80.63	90	82.37	97.04	97.78	72.14	74.04	87.17
Specificity	99.24	99.07	99.69	96.32	97.85	98.96	97.98	99.63	99.71	97.12	97.23	98.69
Accuracy	98.54	98.32	99.39	93.75	96.13	98	96.47	99.41	99.56	94.39	94.78	97.43
# of subjects	20			16			9			20		
# of motions	10			10			10			10		
# of samples	600			480			270			600		

*SFS: sequential feature selection – wrapper method, Ki: Kinect, acc: accelerometer; Mag: Magnetic, Rad: Radar; FL: Feature level; DEcL: Decision Level

Some results of multi-modal fusion are reported in Table III considering radar combined with wearable sensors and radar nodes (also shown in Figure 4, left). The accuracy of fall detection decreased slightly with fusion (FMCW+CW) to 95.28 % compared to 96.23 % with CW. Choosing the right fusion strategy while maintaining a high level of accuracy for specific classes remains challenging. The best combinations of sensors varied across subgroups of people based on physical traits and age. Recently, DL examples [1, 12] have fused information in the deep layers by pooling information from different sensors into a further NN before classification. In [12], a multistatic-DCNN (Figure 4 right) and a Monostatic-DCNN with voting are proposed in a multistatic radar scenario showing robust performances and low variance for different training schemes ranging from 20 to 50% of the data consistently outperforming standalone monostatic DCNN. They also implemented weight sharing between different layers across nodes to avoid having too many parameters to optimize. This evolution opens new horizons for the integration of multimodal, multi-domain information from heterogeneous algorithms to maximize performances.

Finally, classification algorithms will need to be adaptive and cope with the diversity of people and environments without losing generality. Although this aspect is not thoroughly investigated as yet, advances in reinforcement learning (RL), transfer learning, and cognitive radar framework might provide inspiration to address this challenge.

An interesting concluding observation is that some research communities organize contests to benchmark different algorithms on a set task to validate the performances of proposed algorithms (e.g. the Data Fusion contest in IGARSS for remote sensing). It may be time to adopt this approach in HAR for AL within the radar community to reliably compare the performances of algorithms.

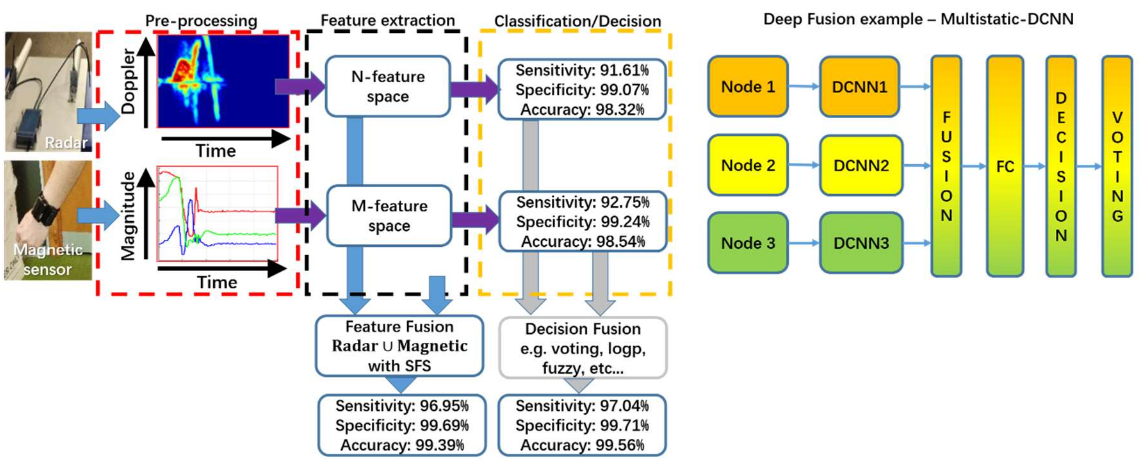

Figure 4: left) Example of multimodal fusion between radar and magnetic sensor for feature level fusion and decision level fusion [13,14]; right) Deep fusion for multi-static radar [12].

IV. Heart/Respiration monitoring for respiratory disease and sleep cycle classification

A. NOVEL SIGNAL PROCESSING TECHNIQUES

Signal processing to estimate health parameters is paramount for AL applications. Existing techniques for noncontact health parameters estimation are mainly based on FT analysis and spectral estimation algorithm (e.g. MUSIC, RELAX) [17-19]. However, due to smearing/leakage problems caused by the limited data length, FT typically is “plagued” with limited resolution. In spectral estimation, the determination of key parameters is a great challenge which influences the estimation performance. To obtain the appropriate resolution, the required window length for the above algorithms is too long to observe the time-varying characteristics of vital signs, which are crucial in intensive care, patients' post-operatory recovery, and severe trauma treatment.

Recently, more accurate/efficient approaches were proposed to overcome these challenges for radar. These include Sparse Reconstruction (SR) to achieve superior short-time processing performance on limited window lengths [20], stepwise atomic norm minimization (StANM) to solve aliasing problems of respiration harmonics when estimating HR [21], and Synchro-Squeezing Transformation (SST) to estimate the instantaneous respiration rate (RR) and HR [22].

SR theory has received a lot of attention in noncontact health parameters estimation [20]. Exploiting their "sparse" characteristic, physiological signals can be reconstructed by solving an undetermined model with an l_0-norm minimization. Thus, short-time processing and super-resolution spectral estimation can be performed. Furthermore, due to the weak magnitude of HR signal (~one-tenth of RR), conventional SR algorithms based on single measurement vector (SMV) model have a high probability of false estimation. To enhance robustness, a rearrangement scheme of the demodulated data is proposed based on the multiple MV (MMV) model. Then, the Regularized MMV FOCal Underdetermined System Solution (RMFOCUSS) algorithm is applied to recover RR/HR frequencies. Both simulated and experimental results have demonstrated the superior short-time processing performances with a window length of ~5 s.

To solve aliasing problems in HR estimation caused by RR harmonics, a StANM is proposed to accurately assess the RR/HR frequencies with limited data [21]. First, the RR frequency is estimated by the conventional ANM. Then, the RR harmonics are generated based on the inherent relationship between the fundamental and harmonics. Finally, with the pre-estimated frequencies, the HR frequency can be located by solving a modified ANM problem. Simulations and experiments showed accurate RR/HR frequency estimations from 6.5s of raw data sampled at 4 Hz.

To extract the instantaneous health parameters, a novel TF analysis algorithm, SST is proposed [22]. It is a special case of reassignment technique while the reconstruction is enabled. SST was first proposed in [23]. Assuming the radar-detected RR/HR signals are based on AM-FM model, the SST calculates the TF distribution W_s by using the Wavelet transform which will spread out over a region around its frequency, which has a low resolution. Thus, the SST reallocates the WT distribution and squeezes it along the frequency axis. Then according to the ranges of RR/HR components, the instantaneous RR/HR frequencies are extracted by searching the peaks.

As an example, Figure 5 compares the estimation performances of FFT, MUSIC, ANM and stANM. In the experiment, a 6.5s-long signal was randomly chosen. The subject sat 2m away from the antenna. All algorithms performed well in RR estimation. Only StANM estimated the HR accurately (64bpm/1.06Hz) from RR harmonics, which matched the JP2000-09 readings used for ground truth.

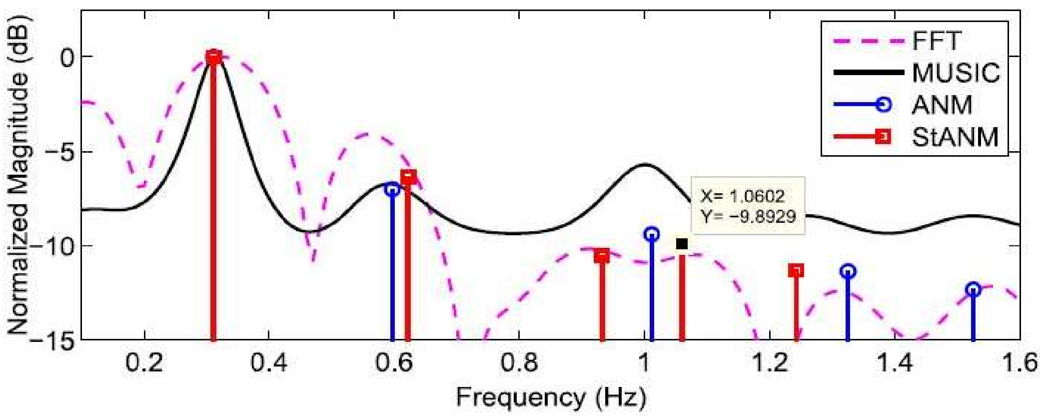

Figure 5: The vital sign estimation results derived from FFT, MUSIC, ANM and stANM [21]

B. RESPIRATION/HEART MONITORING APPLICATIONS

Radar-based accurate time-varying RR/HR estimations are crucial for respiration/heart monitoring applications. They include breathing disorders (BD), sleep stages (SS) monitoring [24-25], potential blood pressure values and their variability, and finally human emotion identification. Radar sensing frees subjects from the discomfort of having probes attached to them, enables to extend the monitoring period, and avoids biases due to psychological stress or physical constraints due to the measuring equipment. These make radar very suited for AL. This section focuses on two possible biomedical applications: BD and SS recognition.

Breathing is an important human vital sign and strongly correlates to health conditions. Unlike regular parameters (e.g RR, tidal volume), BDs are difficult to capture accurately. Notably, they provide significant information for health monitoring and disease diagnosis/prognosis [26]. Generally, BDs are characterized as irregular breathing patterns, and can be induced by an injury of respiratory centers, usage of narcotic medications, metabolic imbalances, respiratory muscle weakness, stroke, and heart failure [26]. Common methods for BD monitoring devices are costly and need to be worn which is inconvenient. They can hardly be used for long-term applications such as overnight breathing monitoring during sleep, or in-home monitoring of POD patients.

For noncontact long-term BD monitoring, a CW radar-based BD recognition system has been proposed [24]. As shown in Figure 6 (bottom), the proposed system consists of a CW radar and a machine-learning-based BD recognition module. A custom-designed 2.4-GHz CW digital-IF radar accurately captures the time-domain breathing waveform. Then, the recognition module is designed to identify different BDs with selected features using ECOC-based multiclass SVM. 7 out of 13 extracted features are selected using the ReliefF algorithm.

Clinical experiments were also carried out with real patients in the Nanjing Chest Hospital to assess the system performance. A patient diagnosed with nocturnal BDs was involved in the experiment. The radar was 2 m above the chest wall with its radiation beam pointing down. The experiment was approved by the institutional review board of the hospital and performed with the patient's informed consent. Figure 6 (top-left) illustrates the subject's hourly distribution of Cheyne-Stokes breathing and variant breathing during sleep. These two BDs occur mostly in the 6^{th}, 7^{th}, and 8^{th} hours. The references were manually labeled by a respiratory physician. The error in each hour is less than 4 except in the 1^{st} hour, when the patient just fell asleep, and random body movements distorted the breathing signal making the results inaccurate.

Furthermore, the recognition of 6 breathing patterns namely normal breathing, Cheyne-Stokes breathing, Cheyne-Stokes variant breathing, Dysrhythmic breathing, Biot's breathing, and Kussmaul's breathing was carried out. After removing the body movements, 832 valid epochs were used for testing. Figure 6 (top-right) shows the classification results for 8h of breathing signals. NSVM means correct predictions by the SVM classifier, and NHand is the hand-labeled predictions.

The accuracy is defined as the ratio NSVM/NHand. During his sleep, 5 breathing patterns were found except for Biot's breathing. The average algorithm accuracy was 86.5%.

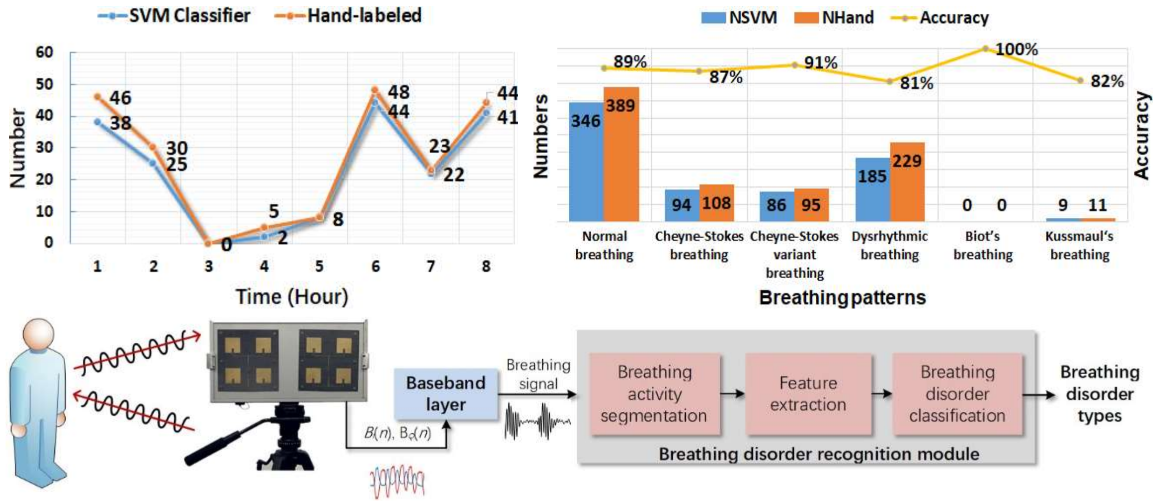

Figure 6: top-right) The hourly distribution of the Cheyne-Stokes breathing and variant breathing during sleep; top-left) The classification results of the 6 breathing patterns during sleep; bottom) The block diagram of the non-contact BD recognition system [24]

Sleep is a basic human need, like eating, drinking, and breathing. As other needs, sleep is vital for good health and well-being throughout a person's lifetime. There is a growing recognition of the adverse effects of poor sleep quality as it is usually associated with physical and mental health problems, injuries, loss of productivity, and even a greater risk of death. Sleep quality analysis is gaining increasing attention and used for the diagnosis of various health problems. In sleep medicine, polysomnography (PSG) is considered the gold standard to estimate SS and evaluate sleep quality [27]. However, PSG requires sleep technicians in a specialized laboratory, which is difficult to access in many communities. To enable sleep monitoring under more natural conditions, sleep-related physiological signals, RR/HR mean, variance, amplitude difference accumulation, and entropy, body movements and its derived features deep sleep and rapid eye movement (REM) are measured based on the CW radar system with subspace K-NN to improve accuracy.

Figure 7 (bottom) shows the overall system architecture of the noncontact SS estimation system. First, the CW radar captures the sleep-related physiological signals and generates the baseband I/Q signals. Then, the radar demodulates the I/Q signals and extracts the physiological signals. Lastly, the SS estimation layer extracts the features from the physiological signals, and classify the SS. To evaluate the quality of sleep, we analyze the relationship between SS and physiological signals measured by the CW radar, a 6h long experiment was conducted to test classification. Figure 7 (top-left) shows the CW radar located over a subject to record physiological signals. A PSG (SOMNOCheck 2 RK) was set up to record ground truth. Then, 11 features from RR/HR and body movement signals were extracted and then classified with Subspace K-NN. The SS estimation framework divided the 6h signals into successive epochs of 60s. Compared with the PSG, the proposed method classified 4 SS including “Wake”, “REM”, “LightSleep” and “DeepSleep” with accuracy rates of 89.1%, 75.8%, 87.0% and 82.0%, respectively. (cf. Figure 7 top-right).

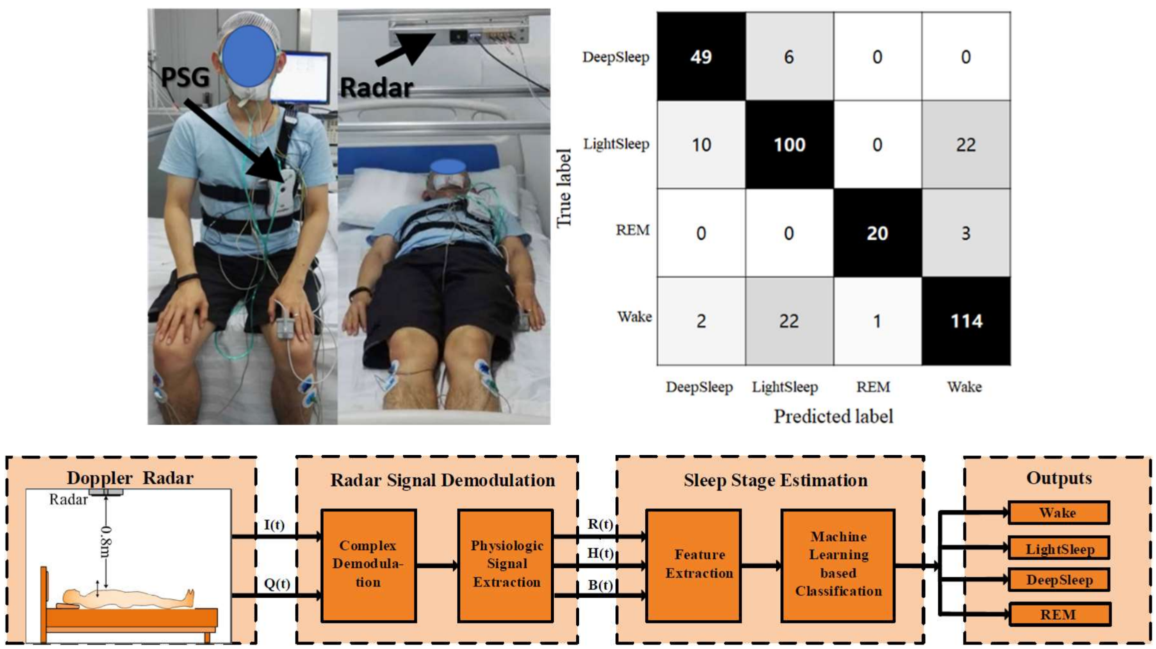

Figure 7: top-left) Experimental setup: the subject lying in bed with the CW radar located 0.8 m over the subject. The PSG provides the ground truth; top-right) SS classification confusion matrix; bottom) Synoptic of the noncontact SS estimation system [25].

V. CHALLENGES IN REAL-TIME IMPLEMENTATIONS

Radar sensing applications for assistive living can require real-time implementation of the signal processing. The challenge lies in designing parallel architectures under the constraints of low footprints (power consumption, area, and heterogeneous resources), and high input data throughput and processing time. A typical radar-based classification system can be broken down in 3 steps:

data acquisition/pre-processing, feature extraction, and classification. For data pre-processing, signal acquisition is a critical step strongly dependent on the ADC performance level, which can reach high sample rates (GS/s) for high resolution radar. Then, TF transforms have to be computed, requiring significant computational power. The number of Multiply-ACcumulate per Second (MACS) can be used as a performance metric counted in GMACS/TMACS, and uses fixed-point rather than floating-point, which has a cascade effect on the performance of the following elements.

Massively parallel architectures can help achieve real-time performance, while approximate computing could be a solution to retain a good level of performance. For feature selection/classification, DL techniques using CNN or other Deep NN (DNN) architectures are promising because of their regular architectures which matches well the FPGA structure. While GPUs are the most widely used platforms to implement CNNs due to their processing power (up to 11 TFLOP/s), FPGA is a real alternative for real-time radar data analysis in terms of power consumption (vs GPUs), rapid prototyping, and massively parallel computing capabilities at different data rates [28]. As a result, numerous FPGA-based CNN accelerators have been proposed, targeting both High Performance Computing data-centers and embedded applications [29,30]. Some works target the design of DNN or RNN units on FPGA showing significant acceleration and power efficiency against conventional processor architecture [31,32]. To reduce the resources required for implementation, a time sharing based parallel implementation of CNN could be used [33] or optimized network operators [34]. For a deeper analysis, refer to the online tool [35] that compares NN accelerators on several hardware platforms (FPGA, ASIC and GPU) in terms of speed and power consumption. From Figure 8, FPGA-based NN accelerators are good contenders to their main counterparts achieving high throughput with a relatively low power consumption. Recently, Xilinx presented DNN engine implemented on the programmable Alveo accelerators cards which can process over 4000 images per second [36]. A complete software library is currently available for integration with machine learning frameworks e.g. Caffe, TensorFlow, or MxNet. Such example highlights the benefit of using programmable architectures to develop very efficient computational systems for radar applications.

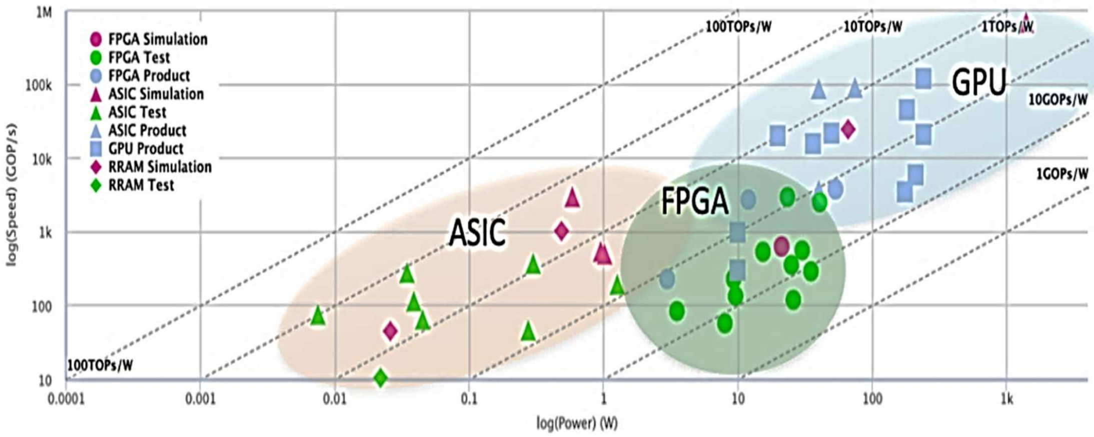

Figure 8: Comparison of neural network accelerators for FPGA, ASIC and GPU devices in terms of speed and power consumption; results obtained from [35].

Further challenges in implementing active learning techniques such as Reinforcement Learning on FPGA to make processing and computation adaptive, are still open and discussed below. As mentioned in previous sections, RL and multimodal detection methods are likely to be widespread used in HAR in the near future to leverage on the increasing number of smart sensors in home environment and to enable robust classification algorithms across people and scenarios.

Regarding multimodal sensing, different types of sensing, computing and communication in embedded systems are required. The nature of the collected data has a significant impact on the computing architecture as well as if the sensing is continuous (gait analysis, video application for fall detection) or periodic e.g. triggered on time or events (sensors for smart home applications). Providing a fast and efficient reconfigurable architecture to adapt the processing locally or to self-organize the learning array [37], at low cost, with respect to the environment seems unavoidable.

Reconfigurable devices have proven to deliver energy efficient systems, by reducing the processing time. SRAM-based FPGAs are still suffering from power leakage, but low-power FPGAs built around FLASH technology deliver good performance with minimal power consumption. In addition, several proven power management strategies such as Dynamic Voltage Frequency Scaling, power and clock gating can be implemented to further reduce the power consumption. Analog and digital converters also have a significant impact on power consumption when large bandwidths and high frequencies are required. Adaptive sampling for example allows a reduction in power consumption by dynamically adjusting the sampling frequency. New System on Chip (SoC) architectures dedicated to high complexity radio frequency (RF) signal processing were recently introduced by Xilinx, called RF SoC. The aim is to shift part of the analog signal processing to the digital domain onto the FPGA, including converters [46], improving the flexibility, and simplifying the filtering requirements, thus providing a fully integrated RF signal processing front-end. These proposed accelerator designs revolve around achieving high parallelism, increasing on-chip data reuse and efficient memory hierarchy, thus enabling to meet the high computational requirements of CNNs. The communications in these massively parallel accelerators architectures also require novel custom Network-on-Chip strategies [38], which are anticipated to be the performance bottleneck. All the different degrees of freedom offered by SoC FPGA make possible the development of complex signal processing for AL applications.

SoC FPGAs also provide a large amount of I/Os and various interfaces. This is of further importance for multimodal sensing where several heterogeneous sensors have to be connected to the FPGA, enabling to gather and process all the collected data (compression, fusion, filtering, classification, etc.). High performance interfaces offered by FPGA, in particular with memories, enable the transfer of large amounts of data. FPGAs also have high performance DSP blocks and floating-point units, delivering much more processing power than conventional CPUs.

Implementing learning techniques onto a reconfigurable architecture is a promising approach. However, note that offline training is often realized and the algorithm performances are evaluated using floating-point precision. Regarding real-time implementation, fixed-point arithmetic is privileged, limiting the performances and significantly decreasing the accuracy [39]. However, half-precision floating-point format seems to be interesting to address future implementations on FPGA as well as approximated computing to maintain good energy-performance trade-offs [30].

Furthermore, the training and prediction time increase significantly for DNN implemented on GPU. Even using heterogeneous platforms with CPU, GPU, FPGA or ASIC with enough processing power, several issues remain such as thermal dissipation, efficient programming of such platforms to achieve the required performance, as well as communications between the cores. In [29], the same issues for embedded vision systems were raised, adding to the fact that most algorithms are not designed for heterogeneous platforms.

VI. CONCLUSIONS

This article introduced radar-based HAR and vital signs monitoring in the context of AL through three classification applications: Activities of Daily Living (ADL), Breathing Disorders (BDs), and (Sleep Stage) SS estimation, and considered the challenge of real-time implementation. The common problem of supervised learning classification for the 3 applications was first presented with a general framework and discussed the concepts of feature extraction, selection, the choice of the algorithms (SL or DL), and the "curse of dimensionality" from a general standpoint. Specific challenges relating to HAR provided an overview of the various radar information domains available for feature extraction and then illustrated with ADL classification with low interclass variations with multi-domain and multi-modal fusion examples. Several proposed algorithms exist in supervised learning, but their relative performances are hard to assess because there is no common benchmarking dataset available to date. Then, the extraction of health parameters for respiration/heart monitoring applications was introduced looking at signal processing techniques for RR/HR estimation, including FT, MUSIC, RELAX, ANM, stANM, and SST techniques. This was followed by two biomedical applications for BD/SS classification using a custom designed CW radar exploiting the estimated RR/HR and body movements to better inform medical prognosis showing over 80% accuracy compared to gold standard ground truth.

In many real-world applications and particularly in healthcare applications, the original data sets used for the learning must evolve to refine classifier performance. One of the recent challenges in active learning lies on the partial reconfiguration on FPGA on the fly. Massively parallel architectures based on FPGA are a good opportunity to support complex learning techniques, but some research innovations have to be proposed to enable efficient real-time implementation that match offline performances. Bio-inspired and neuromorphic computing can be a source of innovations in this field. Recently, some breakthroughs have taken place in the area of near-memory and in-memory computing. 3D memory can offer an order of magnitude higher bandwidth and significantly lower power consumption compared to 2D memory, such as Hyper Memory Cube (HMC) proposed by Micron, which uses through silicon vias (TSV) to stack the dynamic random-access memory (DRAM) on top of the logic circuit. These new computing paradigms represent a serious alternative for FPGA, in a near future.

Other outstanding challenges in classification include but are not limited to temporality (features, successive actions), how to isolate actions from a continuum, unsupervised learning, multi-occupancy scenarios, clutter and interference mitigation but are beyond the scope of this article and will be the subject of further work. Much is to be learned from other fields such as autonomous vehicles since they include sensor suites including radar for object tracking, and geoscience and remote sensing e.g. from the data fusion contest organized regularly in IGARSS and the inclusion of semantic representations in classification approaches.

REFERENCES

[1] J. Le Kernec *et al.*, "Radar for Assisted Living in the Context of Internet of Things for Health and Beyond," *26th IFIP/IEEE Int. Conf. Very Large Scale Integr. VLSI-SOC*, Verona, 2018, pp. 1-5.

[2] C. Li *et al.*, "A Review on Recent Progress of Portable Short-Range Noncontact Microwave Radar Systems," *IEEE Trans. Microw. Theory Tech.,* vol. 65, no. 5, pp. 1692-1706, 2017.

[3] L. Ren *et al.*, "Phase-Based Methods for Heart Rate Detection Using UWB Impulse Doppler Radar," *IEEE* Trans. Microw. Theory Tech., vol. 64, no. 10, pp. 3319-3331, 2016.

[4] M. G. Amin, Ed., Radar for indoor monitoring, 1st edition. CRC Press, 2017.

[5] V. Nathan et al., "A Survey on Smart Homes for Aging in Place: Toward Solutions to the Specific Needs of the Elderly," IEEE Signal Process. Mag., vol. 35, no. 5, pp. 111-119, 2018.

[6] A. K. Jain et al., "Statistical pattern recognition: a review," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 1, pp. 4-37, 2000.

[7] P. V. Klaine *et al.*, "A Survey of Machine Learning Techniques Applied to Self-Organizing Cellular Networks," *IEEE Commun. Surveys Tuts.,* vol. 19, no. 4, pp. 2392-2431, 2017.

[8] I. Goofellow *et al.*, *Deep learning*. MIT Press, 2016

[9] B. Jokanovic *et al.*, "Multiple joint-variable domains recognition of human motion," *IEEE Radar Conf.*, Seattle, USA, 2017, pp. 0948-0952.

[10] B. Erol, M. G. Amin, "Radar Data Cube Analysis for Fall Detection," *IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, 2018, pp. 2446-2450.

[11] Y. He *et al.*, "Range-Doppler surface: a tool to analyse human target in ultra-wideband radar," *IET Radar, Sonar & Navigation,* vol. 9, no. 9, pp. 1240-1250, 2015.

[12] Z. Chen *et al.*, "Personnel Recognition and Gait Classification Based on Multistatic Micro-Doppler Signatures Using Deep Convolutional Neural Networks," *IEEE Geosci. Remote Sens. Lett.,* vol. 15, no. 5, pp. 669-673, 2018.

[13] H. Li *et al.*, "A Multisensory Approach for Remote Health Monitoring of Older People," *IEEE J. of Electromagn., RF and Microw. in Med. and Biol.,* vol. 2, no. 2, pp. 102-108, 2018.
[14] H. Li *et al.*, "Magnetic and Radar Sensing for Multimodal Remote Health Monitoring," *IEEE Sensors J.,* pp. 1-1, 2018.
[15] F. Castanedo, "A Review of Data Fusion Techniques," *Scientific World J.,* 2013.
[16] C. Ding *et al.*, "Non-contact Human Motion Recognition based on UWB Radar," in *IEEE J. on Emerging and Sel. Topics in Circuits and Syst.*, ed, 2018.
[17] Y. She *et al.*, "Group Iterative Spectrum Thresholding for Super-Resolution Sparse Spectral Selection," *IEEE Trans. Signal Process.,* vol. 61, no. 24, pp. 6371-6386, 2013.
[18] C. Li *et al.*, "Accurate Doppler Radar Noncontact Vital Sign Detection Using the RELAX Algorithm," *IEEE Trans. Instrum. Meas.,* vol. 59, no. 3, pp. 687-695, 2010.
[19] P. Bechet *et al.*, "A non-contact method based on multiple signal classification algorithm to reduce the measurement time for accurately heart rate detection," *Review of Scientific Instruments,* vol. 84, no. 8, pp. 84707-84709, 2013.
[20] L. Sun *et al.*, "Super-resolution spectral estimation in short-time non-contact vital sign measurement," *Review of Scientific Instruments,* vol. 86, no. 4, 2015.
[21] L. Sun *et al.*, "Noncontact Vital Sign Detection based on Stepwise Atomic Norm Minimization," *IEEE Signal Process. Lett.,* vol. 22, no. 12, pp. 2479-2483, 2015.
[22] H. Zhao *et al.*, "Noncontact Physiological Dynamics Detection Using Low-power Digital-IF Doppler Radar," *IEEE Trans. Instrum. Meas.,* vol. 66, no. 7, pp. 1780-1788, 2017.
[23] F. Auger *et al.*, "Time-frequency reassignment and synchrosqueezing: An overview," *IEEE Signal Process. Mag.,* vol. 30, no. 6, pp. 32-41, 2013.
[24] H. Zhao *et al.*, "A Noncontact Breathing Disorder Recognition System Using 2.4-GHz Digital-IF Doppler Radar," *IEEE J. of Biomed. and Health Informat.,* 2018.
[25] H. Hong *et al.*, "Noncontact Sleep Stage Estimation Using a CW Doppler Radar," in *IEEE J. on Emerging and Sel. Topics in Circuits and Syst.*, ed, 2018.
[26] G. Yuan *et al.*, "Respiratory rate and breathing pattern," *McMaster University Med. J.,* vol. 10, no. 1, 2013.
[27] D. Chang *et al.*, "A Versatile Wireless Portable Monitoring System for Brain–Behavior Approaches," *IEEE J. Emerging Sel. Topics Circuits Syst.,* vol. 1, no. 4, pp. 440-450, 2011.
[28] S. Abdelhedi *et al.*, "Fall Detection FPGA-Based Systems: A Survey," *Int. J. of Aut. and Smart Technol.,* vol. 6, no. 4, 2016.
[29] S. Mittal, "A survey of FPGA-based accelerators for convolutional neural networks," *Neural Comput. and Appl.,* 2018/10/06 2018.
[30] K. Abdelouahab *et al.*, "Accelerating CNN inference on FPGAs: A Survey," 2018, Available: http://arxiv.org/abs/1806.01683.
[31] C. Wang *et al.*, "DLAU: A Scalable Deep Learning Accelerator Unit on FPGA," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.,* vol. 36, no. 3, pp. 513-517, 2017.
[32] Y. Guan *et al.*, "FPGA-based accelerator for long short-term memory recurrent neural networks," *22nd Asia and South Pacific Design Autom. Conf. (ASP-DAC)*, 2017, pp. 629-634.
[33] S. A. Dawwd,B. S. Mahmood, "Time Sharing Based Parallel Implementation of CNN on Low Cost FPGA," *Int. Conf. on Advanced Sci. and Eng. (ICOASE)*, 2018, pp. 54-59.
[34] Y. Yang *et al.*, "Synetgy: Algorithm-hardware Co-design for ConvNet Accelerators on Embedded FPGAs," arXiv:1811.086342018.
[35] Y. Wang *et al.* (2018, 13 Dec). *Neural Network Accelerator Inference*. Available: https://nicsefc.ee.tsinghua.edu.cn/projects/neural-network-accelerator/
[36] "Accelerating DNNs with Xilinx Alveo Accelerator Cards," Xilinx14 Oct 2018, Available: https://www.xilinx.com/support/documentation/white_papers/wp504-accel-dnns.pdf.
[37] J. Starzyk, Y. Guo, "Dynamically Self-Reconfigurable Machine Learning Structure for FPGA Implementation," *Proc. Int. Conf. Eng. Reconfig. Syst. & Algo.*, Las Vegas, 2003, pp. 296-299.
[38] C. Anthony, "White Paper: Direct RF-Sampling," 2017.
[39] E. Unlu *et al.*, "An OFDMA based RF interconnect for massive multi-core processors," *8th IEEE/ACM Int. Symp. on Netw.-on-Chip (NoCS)*, Ferrara, Italy, 2014, pp. 182-183.
[40] A. H. Do *et al.*, "Brain-computer interface controlled functional electrical stimulation device for foot drop due to stroke," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2012, pp. 6414-6417.