Kitchen Activity Detection for Healthcare using a low-power Radar-Enabled Sensor Network

Fei Luo, Stefan Poslad, Eliane Bodanese
School of Electronic Engineering and Computer Science
Queen Mary University of London
London, United Kingdom
Email: {f.luo, stefan.poslad, eliane.bodanese }@qmul.ac.uk

***Abstract*—Human activity detection plays a crucial role in the recognition of activities of daily living (ADLs). In the past ten years, research on activity detection in the home was achieved through the data aggregation from several different sensors (presence sensors, door contacts, appliances tagging, cameras, wearable beacons, mobile phones, etc.). However, the cost of deployment and maintenance of a multitude of sensor devices and the intrusiveness they can infer are quite high. Research on minimal and non-intrusive sensing for recognition of ADLs are vital for the future of remote care. In this paper, we propose a minimal and non-intrusive low-power low-cost radar-based sensing network system that uses an innovative approach for recognizing human activity in the home. We applied our novel approach to the challenging problem of kitchen activity recognition and investigated fifteen different activities. We designed and trained a deep convolutional neural network (DCNN) that classifies different activities based on their distinct micro-Doppler signatures. We achieved an overall classification rate of 92.8% in activity recognition. Most importantly, in nearly real-time, our approach successfully recognized human activities in more than 89% of the time.**

I. Introduction

Human activity detection and monitoring is essential in many applications, such as healthcare monitoring, human-computer interaction, surveillance, occupancy management, intelligent transportation, etc. Specifically, human activity in the home enables the development of important applications for assisted living for the elderly and people in need of rehabilitation and support due to mental health problems (e.g. dementia and depression). For example, statistics agree that in the next 30 years the number of elderly people will increase about a third, therefore, there is an increasing demand for remote healthcare systems for one-person households as it facilitates independent living in a smart home setting and does not require the presence of caregivers at all times. Routine activities such as eating, drinking, dressing, bathing, and toileting are commonly referred to as Activities of Daily Living (ADL), which people tend to do on a daily basis for normal self-care [1]. Real or near real time ADL monitoring can allow in time intervention for medical emergencies, for treatment and rehabilitation support, obesity and diet control or other necessary behavioural health assessment linked to a physical/mental disease or injury. Activities in the kitchen are directly related to dietary health. For example, the indication that a person gets in or out of a kitchen can be used to estimate his/her dinning frequency, time and duration of dining; the activity detection of eating and drinking can provide the information of the intake of food and drink. Kitchen scene context-based activity detection can be a useful method for diet controlling and dietary treatment, also helpful for developing smart kitchen as a part of smart home.

The author sequence of this paper follows the first-last-author-emphasis norm (FLAE).

A wide variety of sensors have been applied in human activity detection indoors, such as cameras, wearable sensors, smart phones, appliance/object tagging, etc. Camera-based approaches require line of sight with enough lighting and may breach users' privacy. Another limitation is that camera based approaches are computation-consuming and memory-intensive. Wearable sensors based approaches require users to carry sensors or attach sensors on their bodies, which is inconvenient sometimes, and some studies have shown that elderly people are more disinclined to use wearable sensors [2] and also may forget to carry them. Appliance and object tagging can be costly to deploy and maintain, and some can be seen as inconvenient or intrusive for certain target users. Differently, radars can provide ambient sensing but they do not require users to carry any sensors or appliance/object tagging. Radars preserve users' privacy, they cannot be affected by light and have through-wall penetration. Radars measure frequency shifts in signals reflected from human bodies to recognize different activities. Human activities also introduce multi-path distortion to the radar signals.

At present, radar-based human activity detection is achieved by using machine learning algorithms to classify radar signals of different activities. Traditional machine learning algorithms need a method to extract features before performing classification. The performance achieved by the machine learning algorithm greatly depends on the features extracted from the input samples and by the method used for the feature extraction. The hierarchical structure of deep learning makes it possible to automatically learn features at multiple levels from unstructured data [3]. So, deep learning is also an automatic feature learning method. As one of the most effective deep learning models, DCNNs have been widely applied in speech recognition and image classification, and achieved superior performance. In this research, the radar signals are transformed into time-frequency spectrograms that can be considered as images, so CNNs are suitable to be applied into human recognition and activity classification.

In this paper, we use radars to detect human activities in a realistic kitchen scenario. We built a radar-based sensor network that consists of two low-power low-cost pulse Doppler radars, which are suitable for long-time continuous human activity detection and easily deployed on a large scale. Combining them with DCNNs, we propose an approach to perform nearly real-time human activity detection in a kitchen scenario. Our contributions can be summarized as follows:

1) We built a non-intrusive and device-free system using a low-cost low-power radar-based wireless network and achieved nearly real-time activity detection.
2) We investigated human activities in a very challenging scenario. Activities in the kitchen are difficult to be separtely distinguished as they are short in duration, they occur in a confined space with multiple appliances and furniture, and the activities may be chained in many different sequences. Our results on the recognition of several different activities in a single confined house room using a device/tag free system is a valuable source for future research and comparison on the field.
3) We designed and implemented a Deep Convolutional Neural Network (CNN) for the target activities that performs automatic feature extraction and classification on radar signals. Our DCNN provided a high rate of classification success, and most importantly, it achieved very good classification results in near-real time.

II. Related work

The methods of human activity detection can be classified into two categories, which are Device-Bound systems and Device-Free systems. Device-Bound systems require users to carry wearable sensors, such as GPS, accelerometer, Bluetooth or Wi-Fi receiver, etc. For example, the authors in [4] investigated eight activities (falling, running, walking, etc.) using one single accelerometer attached on participants. They built and trained a CNN on a large dataset with 31688 samples from 100 subjects, and achieved an accuracy of 93.8%. Smart Phones are very popular wearable sensors applied in recognizing ADL as they have several built-in sensors (accelerometer, gyroscope, magnetometer). In [5], the authors investigated five activities (Staircase Ascend, Staircase Descend, Walking, Running, and Sitting) performed by participants carrying a mobile phone. The classification rate achieved by j48 classification algorithm [6] reached 93%. Device-Free systems do not require users to wear any sensor. They perform remote sensing of activities by using remote sensors, such as camera, Wi-Fi, Lidar, etc. In [7], the authors detected nine cooking activities by using a camera. They built a dynamic SVM-HMM hybrid model trained on tagged temporal video sequence to classify different cooking activities, they achieved an accuracy of 72%. Wi-Fi is increasingly used for human activity recognition due to the widespread availability of deployed Wi-Fi infrastructure. Wi-Fi based activity detection is achieved by measuring Channel State Information (CSI) variations caused by the activities. The authors of [8] proposed the use of Wi-Fi to detect people falling. In [9], the authors used Wi-Fi to recognize spoken words by detecting lip movements. A 2D Lidar was creatively used in [10] to recognize indoor human activities based on location changes of users in a pre-determined room with labelled landmarks.

Human activity detection applications using radars are Device-free systems. Due to its non-intrusive, contactless, light-insensitive and strong penetration characteristics, radar is increasingly applied in object recognition and human activity recognition. In 2000, Chen [11] firstly proposed the concept of micro-Doppler signatures, which promotes the development of radar applications for human activity detection. Micro-Doppler refers to the additional modulations attached on the main Doppler frequency shift. These additional modulations are produced by additional movements of smaller parts of a target relative to a radar. They can be used to distinguish different moving objects and human activities. Researchers have applied micro-Doppler signatures in indoor activity detection [12], differentiation between human and animal [13], heartbeat and respiration detection [14], etc. However, most radars used in these research works are power-consuming and high-cost, which preclude the use of radar in many scenarios. We use low-power low-cost radars to make micro-Doppler based human activity detection more accessible and suitable for long-time continuous human activity detection. In [15], a radar-based network consisting of two short-range radars (25GHz) and two long-range radars (5.8GHz BumbleBee) were used to classify indoor activities. The highest accuracy was achieved when fusing data of the two short-range radars and using a random forest classifier. However, they used a much longer frame length and the short-range radars have higher frequency resolution than the BumbleBee radar. Our DCNN approach, achieves the same performance with only two BumbleBee radars in a more challenging scenario.

This paper presents micro-Doppler signature based human activity detection in a kitchen scenario. The architecture of our proposed approach is shown in Fig. 1. Raw radar signals collected by our radars are firstly transferred into the frequency domain by using Short-Time Fourier Transformation (STFT) which generates frequency spectrograms. The micro-Doppler signatures can be extracted from the frequency spectrograms. We feed the spectrograms into our designed DCNN to extract the features automatically and to classify the different activities.

III. Method

A. Radar micro-Doppler

A moving target relative to a radar sensor induces a frequency shift to the reflected signals as a result of the well-known Doppler Effect. Additional movements of smaller parts of the target will produce additional modulations to the main Doppler frequency shift, known as micro-Doppler effect [16],

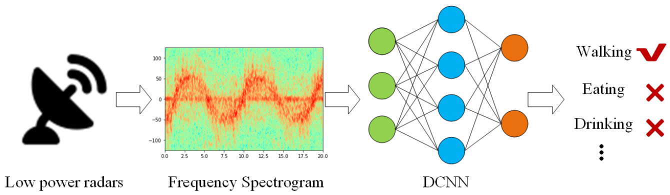

Fig. 1. Architecture of our human activity detection system

[17]. The distinctive characteristics of the observed micro-Doppler effect of an object or a process are called micro-Doppler signatures.

The Doppler frequency shift of a moving target is well-known as

$$f_D = f_0(2v/c), \tag{1}$$

where f_0 is the carrier frequency of the radar and c is the speed of the light, v is the radial velocity of the target relative to the radar.

Given the concept of micro-Doppler signature, a complex object has multiple parts that can move in different speeds and different directions relative to the radar, which will result in multiple time-dependent frequency shifted components. The micro-Doppler signature of such a complex object can be defined as:

$$f_{Dsig}(t) = f_0 \sum_{i=1}^{N} 2v_i(t)/c, \tag{2}$$

where N is the number of parts of the moving target, $v_i(t)$ is the velocity of each part as a function of the time. The analytic signal of the returned echo from such a target is given by:

$$\hat{S}_R(t) = e^{j2\pi f_0 t} e^{j2\pi f_{Dsig}(t)t}, \tag{3}$$

The combination of the received signal $\hat{S}_R(t)$ with the transmitted signal $\hat{S}_T(t)$ as follows:

$$\hat{S}_R(t)\hat{S}_T(t)^* = e^{j2\pi f_{Dsig} t}, \tag{4}$$

it is the component of the signal that contains the micro-Doppler information of the target and it can be used for target or activity recognition and classification.

Micro-Doppler signatures can be represented in a two-dimensional time-frequency space using a Short Time Fourier Transform (STFT):

$$\begin{gathered} STFT(i,K) = \sum_{n=0}^{N-1} x_i(n) e^{-j2\pi(nK/N)}, \\ K = 0, \ldots, N-1 \end{gathered} \tag{5}$$

where $x_i(n)$ is the sliding window with a given length N. The ith window is defined as:

$$x_i(n) = \hat{S}_R(n + i(N/2))w(n), \tag{6}$$

where $w(n)$ is a weighting function.

Fig. 2 shows a frequency spectrogram of an individual walking generated from raw radar signals (in-phase (I) and quadrature-phase (Q) signal components) by STFT. As it can be seen, the fluctuations resulted by the limbs is attached to the main Doppler frequency resulted by the torso. The frequency spectrogram has a two-dimensional structure, which is similar to the structure of an image making CNNs a very suitable classification approach.

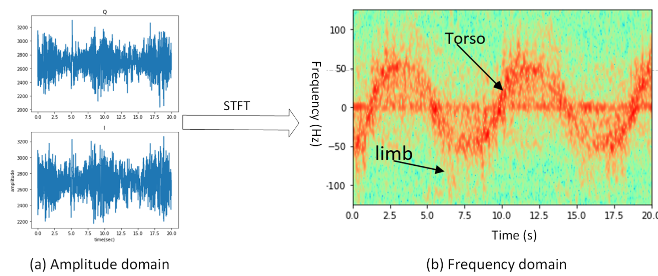

Fig. 2. A frequency spectrogram of walking generated by STFT

B. Deep Convolutional Neural Networks

DCNNs are hierarchical neural networks which consist of multiple convolutional layers that alternate with pooling layers. DCNNs exploit spatially-local correlation by enforcing a local connectivity pattern between neurons of adjacent layers. Generally, there are three typical layers found in CNN architectures: Convolutional Layers, Pooling Layers, and Fully-Connected Layers.

A *convolutional layer* is the main fundamental layer of a CNN. It is designed to learn the feature representation of the input. It operates on an input image/feature map by applying a filter (kernel function) over the image. A convolution applies a set of filters F, whose size is $G \times G$, on an input feature map x of a layer, produces an output feature map $x^{'}$ as

$$x^{'} = \sum_{i=1}^{k} \sum_{r=1}^{G} \sum_{c=1}^{G} x_{rc} \cdot F_i + b^{'} \tag{7}$$

where k is the number of filters F, $b^{'}$ is the bias parameter associated with the feature map $x^{'}$.

A *pooling layer* is a form of non-linear down-sampling. It is designed to reduce the spatial size (dimensionality) of the input, in order to reduce the number of parameters (e.g. neurons and their connectivity) in the CNN. Assuming a pooling layer l and an input layer $l-1$, then the computational formula of l is as follows:

$$x^l = \beta^l down(x^{l-1}) + b^l \tag{8}$$

where $down(\cdot)$ is the down-sampling function. Max pooling is the most common down-sampling function, it seeks the maximum in a $n \times n$ pixel patch. β is the weight value, b is the bias parameter.

Fully connected layers can be seen as regular neural network layers. All neurons in the previous layer are transformed into a one-dimensional vector in a fully connected layer. Fully connected layers are added after all convolution layers and pooling layers.

In training process, a DCNN model is trained to minimize an objective function in terms of the parameters of the network. For a given classification, let C be the number of labeled classes, the following cross-entropy loss function is often used:

$$E_y(y^{'}) = -\sum_{i=1}^{N} y_i \cdot log(y_i^{'}) \tag{9}$$

where E is the loss function evaluated over N samples, y_i is the original label of the i_{th} sample and $y_i^{'}$ is the class score maps of the sample i calculated using a *softmax* activation function:

$$y_j = exp(x_j)/(\sum_{c=1}^{C} exp(x_c)) \tag{10}$$

where y is the softmax score and x is the output layer containing unnormalized class scores.

IV. Implementation

In this section, we detail the characteristics of the radar used in this research, and the structure of our radar sensor network. Then we deployed this network in a kitchen to recognize fifteen human activities. Finally, a DCNN model is built to classify the activities by using the micro-Doppler signatures.

A. Experimental setup

The low-power low-cost radar sensor network implemented in the experiment consists of two BumbleBee radars [18] and three TelosB motes [19] (see Fig. 3). The Bumblebee radar is a low-power Pulsed Doppler Radar operating at 5.8 GHz. Its detection range is up to 8 meters. It only consumes about 12 mAh, so when using typical 1.5v AA alkaline batteries with a capacity of 2400 mAh, it can run at 100% duty cycle for about 8 days. It only costs 100 dollars for each BumbleBee radar. The TelosB mote provides radio communication at low-power consumption (IEEE 802.15.4). It is fully compatible with the open-source TinyOS, an operating system that supports large-scale, self-assembling sensor networks.

Our sensor network contains two nodes (‘Node 1‘ and ‘Node 2‘) and one base station. Each node consists of a BumbleBee radar and a TelsoB mote. The TelosB mote is connected to the BumblebBee radar, and transfers data packets that contain radar signals to the base station. The base station is built with a TelsoB connected to a PC, where the received data packets are processed.

This network is deployed in a kitchen. The plan of the kitchen is shown in Fig. 4(a). In order to cover all kitchen space, two nodes (green rectangles in Fig. 4(a)) are placed in two diagonally opposite corners of the kitchen, with an approximate angle of 30° in relation to their left wall. The height of both nodes is 1.2 meters. Fig. 4(b) presents the bird‘s-eye view of 3D model of the kitchen. As it can be seen, the kitchen contains a long rectangular table, chairs, a cabinet, a microwave oven, a sink, etc.

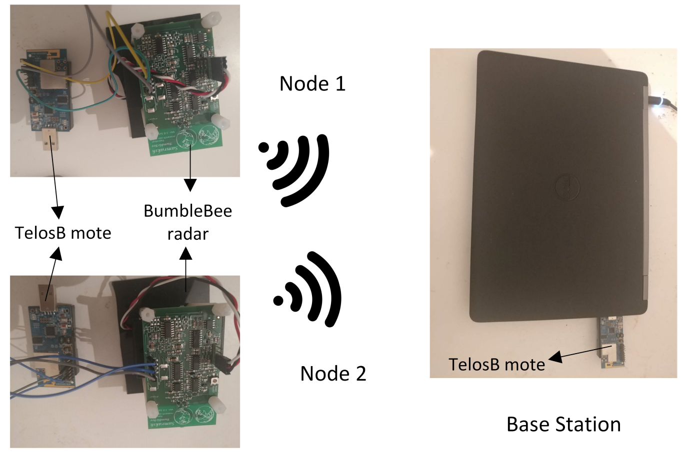

Fig. 3. A wireless radar sensor network

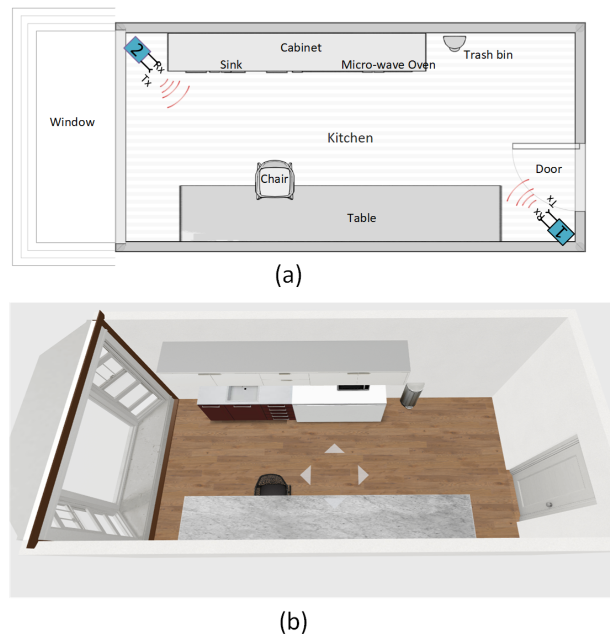

Fig. 4. Kitchen scenario (a) Plan, (b) 3D bird‘s-eye view

B. Human activity detection

We investigated 15 activities of 3 people in the kitchen, including (a) Walking, (b) Eating, (c) Drinking, (d) Sitting, (e) Standing, (f) Washing, (g) Open door and get in, (h) Open door and get out, (i) Open cabinet, (j) Close cabinet, (k) Open oven, (l) Close oven, (m) Open freezer, (n) Close freezer, and (o) No activity. Fig. 5 shows the spectrograms of several targeted activities measured by ‘Node 1‘ and ‘Node 2‘.

The duration for different activity is different. Walking, eating, and drinking take longer time than other activities. So, it is required to select a suitable length for the sliding window of the STFT, in order to make sure the window covers at least one motion cycle of each activity. We use a sliding window of 2.5s to create spectrograms with the interval of 0.5s. As the sampling rate of each node is 250 Hz, a sliding window

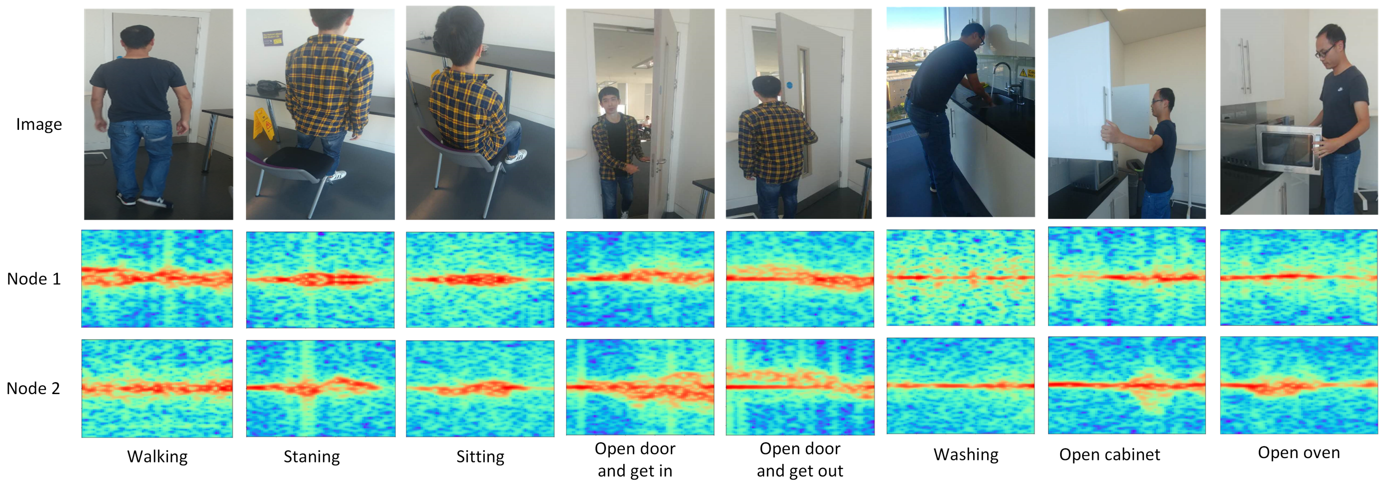

Fig. 5. Spectrograms of human activities in the kitchen

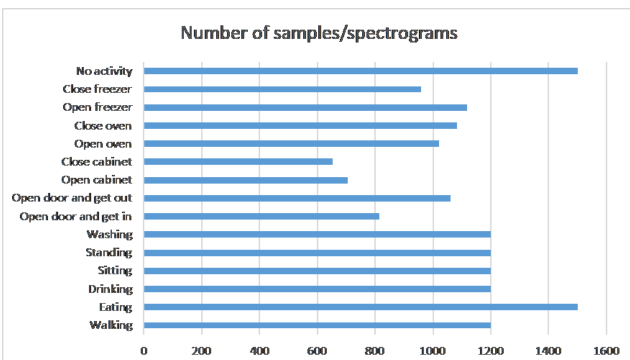

Fig. 6. Number of samples of each type of activity

contains 625 signals.

After radar signal processing, we obtained 15350 spectrograms in total in the experiment. The composition of the samples/spectrograms is shown in Fig 6.

C. DCNN Model

Our radar-based network contains two radar nodes. Each node observes human activities from a different perspective. In order to fuse the signals from the radars, the spectrograms generated from them are firstly down sampled into the size of $80 \times 80 \times 1$, then overlapped together forming a spectrogram with the size of $80 \times 80 \times 2$ (it can be considered as an image with two channels). We built a DCNN model that takes the fused spectrograms as input. The structure of the DCNN is shown in Fig. 7. It contains three convolutional layers (C1, C2, and C3), two Max Pooling layers (M1, M2) and two fully connected layers (F1, F2). All three convolutional layers use 3 kernels to do the convolutionalization. The C1 layer contains 20 feature maps, and the size of each feature map is $78 \times 78 \times 1$. A 2×2 filter is used to perform the Max Pooling on C1, and generate the M1 layer. The C2 layer contains 36 feature maps, and the size of each feature map is $37 \times 37 \times 1$. The M2 layer is

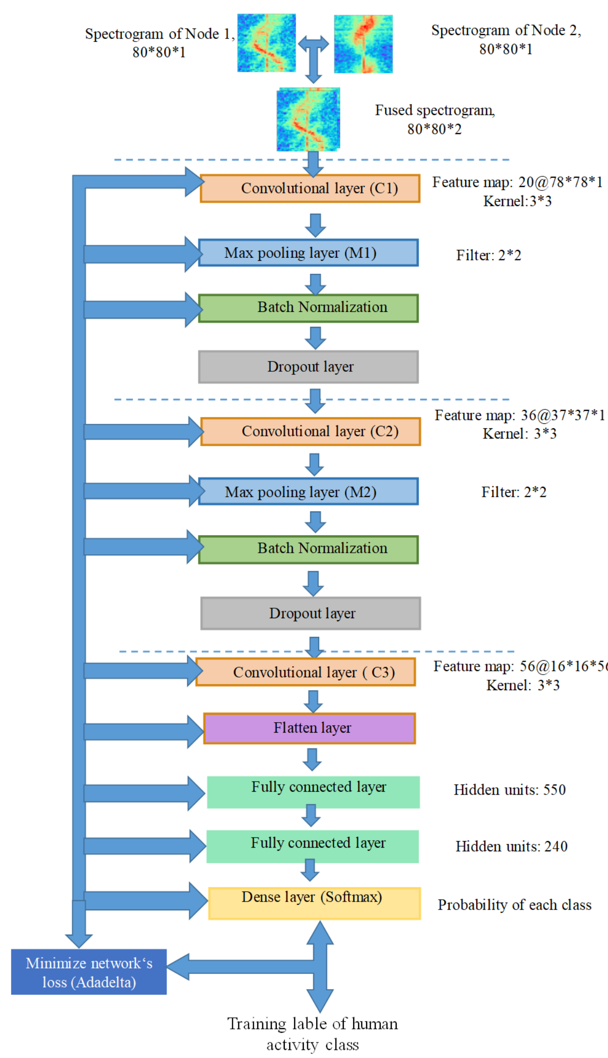

Fig. 7. DCNN model

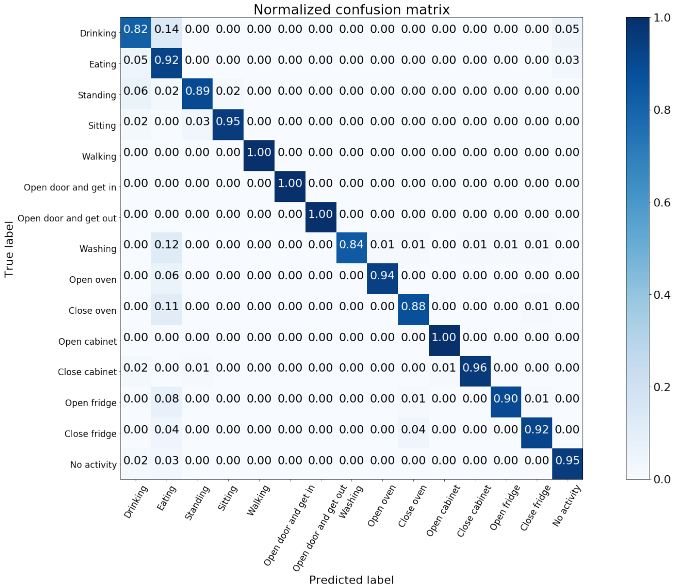

Fig. 8. Normalized confusion matrix

produced by Max Pooling the C2 layer. The C3 layer contains 56 feature maps; each feature map is a $16 \times 16 \times 1$ matrix. The F1 layer contains 550 hidden units, and the F2 contains 240 hidden units. In the training process, dropout [20] has been used to control overfitting with an initial dropout rate of 0.4. And batch normalization [21] has been applied on M1 and M2 as a regularizer to accelerate the convergence.

The use of dropout and batch normalization reduces overfitting and accelerates the training. The optimization function applied is Adadelta, whose initial learning rate is 0.1. Adadelta is an optimization function that can dynamically adapt over time using only first order information and it has minimal computational overhead beyond standard stochastic gradient descent, which is one of the most popular methods used to perform optimization [22].

V. Evaluation and Results

In this section, we evaluated the performance of our DCNN model in kitchen context based human activity detection in three ways.

A. Validation on test dataset

We evaluated this approach on a new dataset which has never been exposed to the DCNN model. The achieved overall accuracy is 92.81%. As shown in the normalized confusion matrix of Fig. 8, the accuracy achieved in 'Drinking' is 82%, which is the lowest. Most of the misclassified samples in 'Drinking' have been classified into 'Eating'. This is because there are a lot of similarities in these two activities. The classification rates of most activities are above 90%. Especially, 'Walking', 'Open door and get in' and 'Open door and get out' achieved the classification rate of 100%.

B. Comparison with SVM

We also compared our DCNN model with Support Vector Machine (SVM). SVM is a popular traditional machine learning algorithm, which has been widely used in image classification. We used Principal Component Analysis (PCA) to extract features as the input of SVM. As shown in Table I, the DCNN is ahead of SVM+PCA about 2 percentages points in all three metrics, which are 92.81% in overall accuracy (OA), 93.14% in Recall, and 93.83% in f1.

TABLE I
Comparison of the DCNN and the SVM+PCA

	OA	Recall	f1
DCNN	92.81%	93.14%	93.83%
SVM+PCA	90.9%	91.56%	91.71%

C. Real-time human activity recognition

We developed a system by combining the radar sensor network and the DCNN model. With this system, We performed real-time human activity recognition in the kitchen to detect human behaviours for 3 minutes. In this period, all targeted activities were performed. Fig. 9 summarized our real-time detection results.The first row is the groundtruth and the second is our automatic recognition. Different color represents different activity. Except 'Close oven', all other activity were recognized. 'Eating' took the longest time and 'Walking' was the most frequently performed activity. By comparing the groundtruth and the prediction, we found human activities were successfully recognized in more than 89% of the time.

VI. Conclusion

In this paper, we proposed a radar sensor network to detect human activity without requiring users to carry any

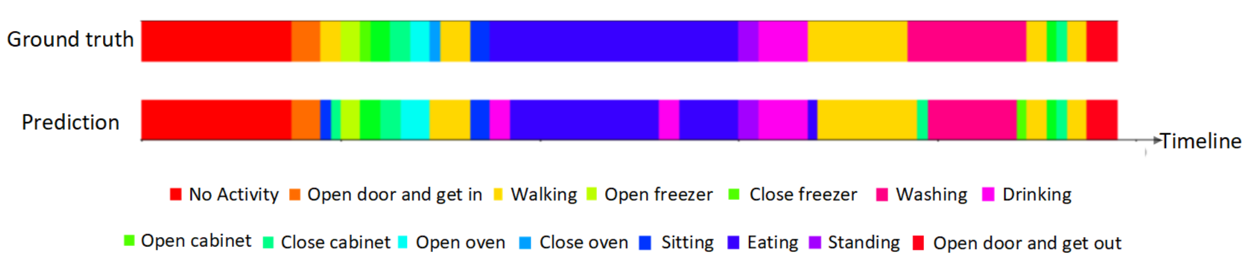

Fig. 9. A timeline visualization of real-time human activity recognition results

sensors. We investigated 15 activities in a kitchen scenario. With the collected radar signals, STFT was used to generate spectrograms that contain micro-Doppler signatures of these human activities in the frequency domain. We built a DCNN to learn the micro-Doppler features automatically and classify the activities. Our DCNN achieved 92.81% overall accuracy in the test, which exceeds the performance of SVM+PCA by around 2%. We further implemented our approach in nearly real-time detection. It successfully recognized human activities in more than 89% of the time. Our work shows the great potential of low-power low-cost radar-enable sensor network in a kitchen context based human activities detection, which is helpful in diet-controlling, patient monitoring and smart kitchen design. Our approach has the advantages of not requiring wearable sensors, it is power-saving and computational efficient. In future, we will further implement our approach to other scenarios and investigate more activities. We will also investigate the use of other low power radars with higher frequency resolution and we expect by using our DCNN approach to achieve greater accuracy of real time activity detection in challenging scenarios.

REFERENCES

[1] M. E. Mlinac and M. C. Feng, "Assessment of activities of daily living, self-care, and independence," *Archives of Clinical Neuropsychology*, vol. 31, no. 6, pp. 506–516, 2016.

[2] D. M. Levine, S. R. Lipsitz, and J. A. Linder, "Trends in seniors use of digital health technology in the united states, 2011-2014," *Jama*, vol. 316, no. 5, pp. 538–540, 2016.

[3] Y. Bengio, "Deep learning of representations for unsupervised and transfer learning," in *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, 2012, pp. 17–36.

[4] Y. Chen and Y. Xue, "A deep learning approach to human activity recognition based on single accelerometer," in *Systems, man, and cybernetics (smc), 2015 ieee international conference on*. IEEE, 2015, pp. 1488–1492.

[5] A. Bansal, A. Shukla, S. Rastogi, and S. Mittal, "Micro activity recognition of mobile phone users using inbuilt sensors," in *2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. IEEE, 2018, pp. 225–230.

[6] T. R. Patil and S. Sherekar, "Performance analysis of naive bayes and j48 classification algorithm for data classification," *International journal of computer science and applications*, vol. 6, no. 2, pp. 256–261, 2013.

[7] S. Bansal, S. Khandelwal, S. Gupta, and D. Goyal, "Kitchen activity recognition based on scene context," in *2013 IEEE International Conference on Image Processing*. IEEE, 2013, pp. 3461–3465.

[8] Y. Wang, K. Wu, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 581–594, 2017.

[9] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with wi-fi!" *IEEE Transactions on Mobile Computing*, vol. 15, no. 11, pp. 2907–2920, 2016.

[10] Ma Z., Bigham J., Poslad S., Wu B., Zhang X., and Bodanese E., "Device-free daily life (adl) recognition for smart home healthcare using a low-cost (2d) lidar."

[11] V. C. Chen and R. D. Lipps, "Time frequency signatures of micro-doppler phenomenon for feature extraction," in *Wavelet Applications VII*, vol. 4056. International Society for Optics and Photonics, 2000, pp. 220–227.

[12] F. Fioranelli, M. Ritchie, and H. Griffiths, "Bistatic human micro-doppler signatures for classification of indoor activities," in *Radar Conference (RadarConf), 2017 IEEE*. IEEE, 2017, pp. 0610–0615.

[13] W. Van Eeden, J. P. De Villiers, R. Berndt, W. A. Nel, and E. Blasch, "Micro-doppler radar classification of humans and animals in an operational environment," *Expert Systems With Applications*, vol. 102, pp. 1–11, 2018.

[14] V. C. Chen, "Advances in applications of radar micro-doppler signatures," in *Antenna Measurements & Applications (CAMA), 2014 IEEE Conference on*. IEEE, 2014, pp. 1–4.

[15] M. A. A. H. Khan, R. Kukkapalli, P. Waradpande, S. Kulandaivel, N. Banerjee, N. Roy, and R. Robucci, "Ram: Radar-based activity monitor," in *INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications, IEEE*. IEEE, 2016, pp. 1–9.

[16] A. Balleri, K. Chetty, and K. Woodbridge, "Classification of personnel targets by acoustic micro-doppler signatures," *IET radar, sonar & navigation*, vol. 5, no. 9, pp. 943–951, 2011.

[17] S. Dura-Bernal, G. Garreau, C. Andreou, A. Andreou, J. Georgiou, T. Wennekers, and S. Denham, "Human action categorization using ultrasound micro-doppler signatures," in *International Workshop on Human Behavior Understanding*. Springer, 2011, pp. 18–28.

[18] Samraksh Company, "Bumblebee user's manual," https://samraksh.com/files/products/Bumblebee%20radar/BumbleBee-UM-v103.pdf, 2008, [Online; accessed 19-July-2018].

[19] J. Polastre, R. Szewczyk, and D. Culler, "Telos: enabling ultra-low power wireless research," in *Information Processing in Sensor Networks, 2005. IPSN 2005. Fourth International Symposium on*. IEEE, 2005, pp. 364–369.

[20] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[21] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[22] M. D. Zeiler, "Adadelta: an adaptive learning rate method," *arXiv preprint arXiv:1212.5701*, 2012.