Fusion of Radar Data Domains for Human Activity Recognition in Assisted Living

Le Kernec, Julien; Fioranelli, Francesco; Romain, Olivier; Bordat, Alexandre

DOI
10.1007/978-3-030-98886-9_7

Publication date
2022

Document Version
Final published version

Published in
Sensing Technology - Proceedings of ICST 2022

Citation (APA)
Le Kernec, J., Fioranelli, F., Romain, O., & Bordat, A. (2022). Fusion of Radar Data Domains for Human Activity Recognition in Assisted Living. In N. K. Suryadevara, B. George, K. P. Jayasundera, J. K. Roy, & S. C. Mukhopadhyay (Eds.), *Sensing Technology - Proceedings of ICST 2022* (pp. 87-100). (Lecture Notes in Electrical Engineering; Vol. 886). Springer. https://doi.org/10.1007/978-3-030-98886-9_7

Fusion of Radar Data Domains for Human Activity Recognition in Assisted Living

Julien Le Kernec, Francesco Fioranelli, Olivier Romain, and Alexandre Bordat

Abstract Radar has long been considered an important technology for indoor monitoring and assisted living. As ageing has become a worldwide problem, it causes a huge burden on the government's healthcare expenses and infrastructure. Radar-based human activity recognition (HAR) is foreseen to become a widespread sensing modality for health monitoring at home. Conventional radar-based HAR task usually adopts the amplitude of spectrograms as input to a convolutional neural network (CNN), which can limit the achieved performances. A hybrid fusion model is here proposed, which can integrate multiple radar data domains. The result shows that the proposed framework can achieve superior classification accuracy of 92.1% (+2.5% higher than conventional CNN) and a lighter computational load than the state-of-the-art techniques with 3D-CNN.

Keywords Radar · Human activity recognition · Fusion · Machine learning

1 Introduction

The growing population worldwide causes a huge burden on the government's healthcare expenses and health infrastructure. Radar has long been considered an important technology for indoor monitoring and fall detection in assisted living. Compared with other competing technologies such as camera monitoring or ultrasonic sounding,

J. Le Kernec (✉)
James Watt School of Engineering, University of Glasgow, Glasgow, UK
e-mail: Julien.lekernec@glasgow.ac.uk

F. Fioranelli
Department of Microelectronics, TU Delft, Delft, The Netherlands

O. Romain · A. Bordat
ETIS Lab, CY University, Cergy-Pontoise, France
e-mail: olivier.romain@cyu.fr

A. Bordat
e-mail: alexandre.bordat@ensea.fr

N. K. Suryadevara et al. (eds.), *Sensing Technology*, Lecture Notes in Electrical Engineering 886, https://doi.org/10.1007/978-3-030-98886-9_7

radar has the advantages of non-intrusive sensing, insensitivity to lighting conditions, privacy preservation [1], and safety. These make radar very attractive in fields like human–machine interface, site surveillance and assisted living [2].

Radar data domains include but are not limited to raw data, range-time (RT), range-Doppler (RD), Doppler-time (DT) [3]. Traditional radar-based human activity recognition (HAR) focuses on a single radar domain, usually Doppler-time, and typically adopts a single-input network structure. Although this method can achieve a good accuracy, when activities are close to each other in Doppler-time representations, such as falling and tying a shoe lace, or if the signatures are different with elderly and young people in the training/testing data samples, the performance can significantly be impacted. Furthermore, only the amplitude is exploited and not the phase information, and other data domains can help distinguish activities with range spread for example.

2 Literature Review

2.1 Radar-Based Human Activities Recognition

Radar can be used in indoor monitoring largely because of the progress of machine learning technology and the speed of graphic processing units [3]. Based on those technological advances, the use of deep neural networks (DNN) is now feasible for radar-based HAR. Usually, a set of handcrafted features from the micro-Doppler signature will be extracted from the radar signal, such as the Doppler bandwidth, or the centroid (torso frequency) for the DT domain. Then, statistical learning techniques, like support vector machine (SVM) [4] and random forest classifier [5] are used for classification. The DT signature is commonly used in radar-based HAR. In [4], the authors have classified human activities using SVM and a set of handcrafted features extracted form radar DT signatures. However, the effectiveness of this approach is largely dependent on some operational and situational factors [6], such as the transmit and pulse-repetition frequencies, dwell time and signal-to-noise ratio.

Researchers gradually shifted their focus from statistical learning to deep learning, which extracts features automatically and performs classification simultaneously. In [7], the authors investigated the feasibility of using convolutional neural network (CNN) to categorise the radar DT signatures. Considering that the training of deep learning network requires a large amount of data samples, at the same time, radar signal processing and computer image processing often have strong similarities. Craley et al. [8] realised this problem and used transfer learning to pre-train the CNN aquatic activity classification to improve accuracy. In [9] a sparse auto-encoder was proposed to process the radar DT and RT signatures in parallel. Besides, [10, 11], both developed a hybrid network structure (convolutional and recurrent) to achieve a better recognition accuracy.

However, all of the studies mentioned above have limited their scope to two-dimensional domains and only combined the disjoint features inside the classifier. There are two innovative methods [12, 13] which integrate radar data from 2D multi-channels into a 3D single channel. Both of them intend to transform the radar echoes into a 3D range-Doppler time (RDT) signature. But their RDT signatures lack sufficient resolution to extract sufficient information from micro-motions. To resolve this limitation, a geometric deep learning method based on the points formed by micro-motion signatures was proposed in [14]. This method obtained a higher performance in both classification accuracy and noise robustness.

2.2 Multi-domain/Multi-modal Fusion

Our experience of the world is multimodal, we see objects, through our five senses [15]. In general, "modal" refers to the way things happen or exist, and a research problem is characterised as multimodal when it includes multiple modalities.

For the purpose of enabling the artificial intelligence to understand the world around it, we need to teach them to observe and to interpret the multimodal information like a human being. Recent research works are mainly dealing with sound, image and text multi-modal learning in such applications as speech recognition (audio + image) with challenges related to:

- Representation—to find some unified representation of multi-modal information,
- Translation [16]—mapping one typical modality to another modality,
- Alignment [17]—finding the relationships between the modal sub-components,
- Fusion—obtaining more cross-features by integrating the multi-modal information,
- Co-learning—use information-rich modalities to assist information-poor modalities.

To sum up, there are mainly two advantages for implementing multimodal fusion into machine learning.

1. Information complementarity: Multi-modal fusion can complete the missing information of single mode to ensure the integrity of information to improve the model performance.
2. Information crossing: Multi-modal fusion can fully explore the information interaction between different modalities, so even more abundant feature information can be obtained through fusion.

Modal fusion can be divided into four categories. They occur at different stages of the network training:

- Signal/Pixel level fusion is the most intuitive fusion method. The data is pre-processed before entering the machine learning network and therefore, can fuse

the smallest particle of data. Some related research that implement pixel level fusion are [18–20].

- Feature level fusion, for instance [21, 22], includes Early and Late Fusion. These two fusion types occurr at different stages consists in the fusion of features extracted from the machine learning network. The difference is that early fusion refers to the fusion of features through concatenation, element-wise sum, element-wise average and then input to the full connection layer for processing, while late fusion mainly occurs in the full connected layer. Two advantages of fusing data at the feature level are that it can fully exploit the cross information between different modalities and also improve the robustness of the network.
- Decision level, where the output of the classifier is combined to make a final decision. Since the output of different classifiers corresponds to different modalities, the results of different classifiers tend to have strong independence. Therefore, some problems of misclassification caused by the shortcoming of a single modality can be avoided. Common decision level fusion methods include max-fusion, average-fusion, Bayes' rule based fusion and ensemble learning, among others [23].
- Hybrid combination of the three aforementioned fusion methods, so as to combine the advantages of different fusion methods.

Since each single 2D radar data domain can provide supplementary information for other domains, recent studies have combined representationsin radar-based HAR. In [24], the authors have proposed an innovative architecture which implements a stacked auto-encoder (SAE) to fuse the multi-dimensional data at the feature level. In [15], the idea of pixel-level fusion to process the radar signal into a 3D point cloud was adopted. Combined with the PointNet [25, 26], they have achieved enhanced performance in HAR.

Despite those existing improvements in radar-based HAR, the recognition ability can still be improved further by fully utilising the radar data domains. In this paper, we propose a novel multimodal fusion framework fusing radar information during network training and show improved performances compared to the state of the art.

3 Methodology

We used the University of Glasgow (UoG) Human Activities dataset [27, 28] to validate our proposed network structure to benchmark performances against existing research. All data samples were collected from 72 volunteers (23 females, 49 males) aged between 25 and 98 years, and different locations from lab spaces to retirement homes. Each volunteer was asked to perform six activities: walking back and forth (A01), sitting down (A02), standing up (A03), picking up an object (A04), drinking water (A05), and simulated frontal fall (A06). The data was captured with a FMCW

radar operating at 5800 MHz with 400 MHz instantaneous bandwidth, a pulse repetition frequency of 1 kHz, a transmitted power of 100 mW and Yagi antennas for transmit and receive gain of 17 dBi.

Radar data domains include RT, DT, RD. Traditional radar-based HAR usually only retains the magnitude information and discards the phase information, i.e. only the magnitude of the information is finally used for the network training.

From the prospective of physics, any slight motion of the target will give the radar echo a micro-Doppler shift. Taking the range-time phase information as a source of network input will undoubtedly increase the computation burden and complexity of our system, and therefore reduces the versatility. This is also one of the primary reasons why few researchers have utilised the phase information of radar signals to train the network.

Inspired by the mechanism of "Attention", we put forward a method of magnitude masking to combine the range information and phase information together. The core steps of this algorithm can be summarized as follows:

Phase Unwrapping—One important thing that needs to be mentioned is that, during the process of phase information extraction, all phase information is automatically encapsulated in the range of $[-\pi; \pi]$, which greatly limits the information continuity (as time is a continuous variable). With such a method, we can eliminate the phase discontinuity and protect the time-varying phase information.

Threshold Filtering—The threshold allows to only focus on strong signals of interest and discard the noisy parts of the radar data domain representations below this threshold by setting them to 0.

This attention-based magnitude masking is also used to identify the region of interest in the phase data for the RT data domain.

The essence of this operation is to perform a pixel-level fusion on the phase and magnitude of RT domain. In addition, through threshold filtering, we only retain the most critical information in the RT representation. Therefore, during the fusion, we can give more weight to the important information in the phase map, so that our network can pay more "attention" in learning the information of the emphasised part during network training. Figure 1 summarizes the flow chart of the data preprocessing for the RT and DT domains as discussed in this section.

4 Network Training

4.1 Overall Network Structure

Our network can be divided into 3 parts as shown in Fig. 2.

It is composed of:

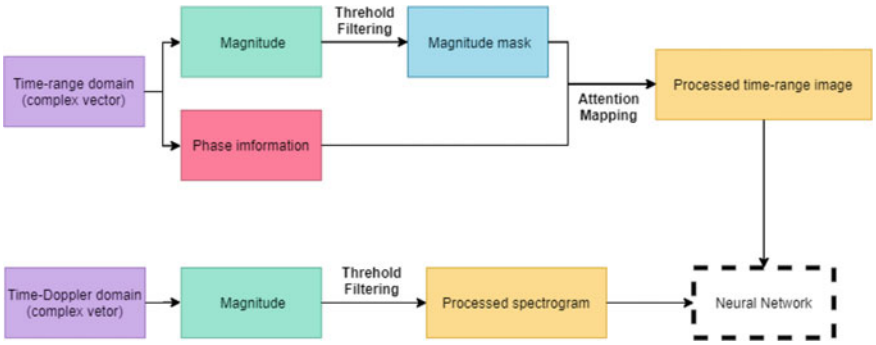

Fig. 1 The extraction process of the network input from the radar data

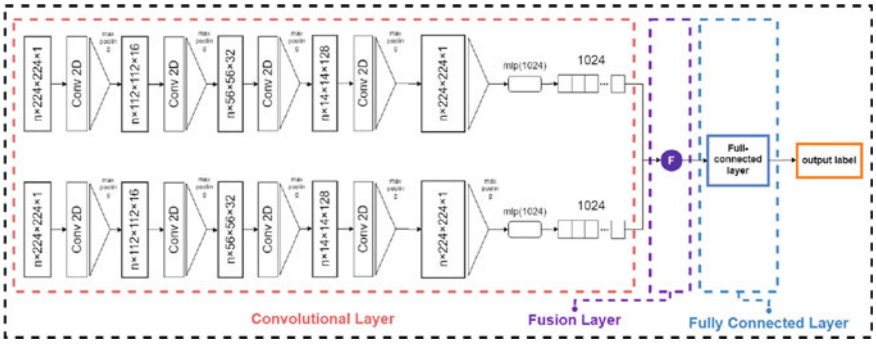

Fig. 2 Overall network structure

- A convolutional layer—this CNN consists of three structures, which are convolution, activation, and pooling. Together, they can automatically extract the feature map from the input data.
- A fusion layer—it implements a feature-level fusion with typical modal fusion element (e.g. element_wise sum, element_wise multiplication, concatenation) as shown in Fig. 3.
- A fully connected layer—it completes the mapping from feature information to label set (e.g. traditional fully connected layer, weight-shared fully connected layer).

Our network has two inputs, each connected to a series of symmetric and identical convolution layers. In the actual operation, the processed RT and DT signatures are taken as the two inputs of the network respectively. Since high data resolution will burden our network and cause unnecessary waste of computation resources, both our input data are down-sampled to a 224×224 input image.

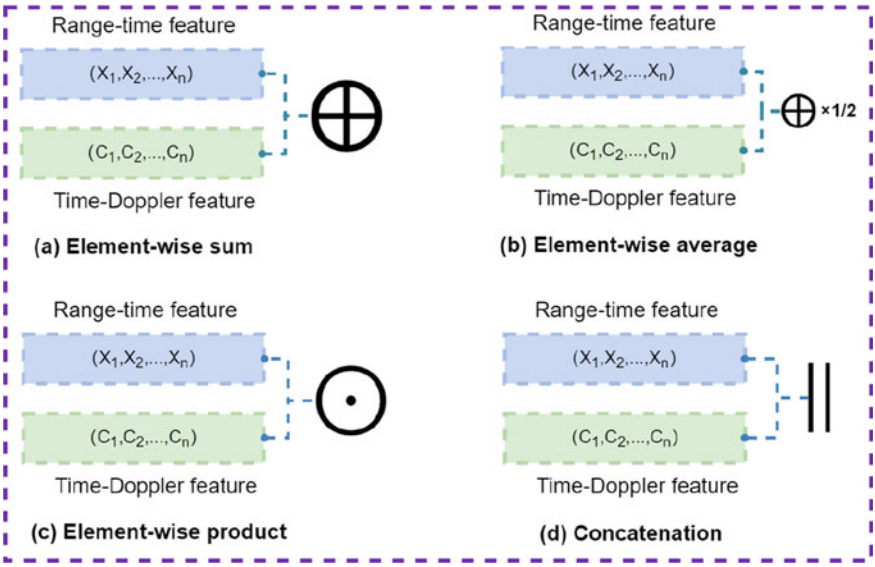

Fig. 3 Four Fusion Elements

Firstly, the feature information is extracted by a series of 2D convolution and max-pooling layers. The output are the global features (with a size of 1×1024) each for RT and DT signatures. The next step is fusion.

4.2 Fusion Layer

After obtaining the global features, we implement 4 different feature-level fusion elements to fully explore the capabilities of modal fusion, for instance, element-wise sum, element-wise product, element-wise average and concatenation as shown in Fig. 3.

In most cases, multimodal feature-level fusion occurs just after the convolution layer. But in this project, we also tried to propose a fusion mechanism in the full connection layer (FCL), called deep fusion.

As shown in Fig. 4, deep fusion is combined by three pairs of shared-weight fully connected layers and four fusion elements. In practice, those two structures alternate in the FCL of deep fusion, which enable more interactions among two sets of features by feeding back the average error in the fusion network to both branches of our fusion network.

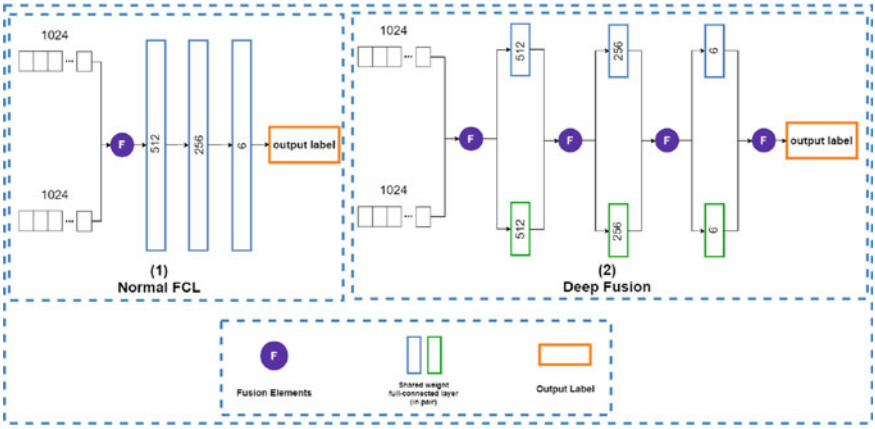

Fig. 4 Two FCL structures

5 Results

In this section, the experiment parameter settings and results of our proposed methodologies will be presented.

5.1 Dataset Separation

The UoG dataset contains over 1700 radar signatures of six human activities. Among them, 1736 data sample were randomly picked to form two training sets and 192 samples divided into three separate validation sets. What needs to be stated here, is that those three validation sets will not participate in the network training, that is to say, we can use them to verify the effectiveness of the network we have trained.

Young/Old Separation—Our dataset is relatively small compared to the huge data sets adopted in other deep learning tasks. Besides, the data set implemented in this project also contains a certain number of data samples collected from older people. Typically, there are certain differences between the elderly and the young participants in motion, posture and speed. For example, some elderly people need to use crutches, a cane or a walker for assistance in walking, which resulted in a few abnormal data samples for the elderly in our dataset.

Based on the above situation, we infer that compared with the data set of the young volunteers, the data set of the elderly is more likely to have anomalies. Therefore, if we simply mix the data sets of the elderly and the young together, and then pack them randomly to form a validation set and a training set, we will not be able to guarantee that those abnormal data samples can be evenly distributed for the validation and training sets.

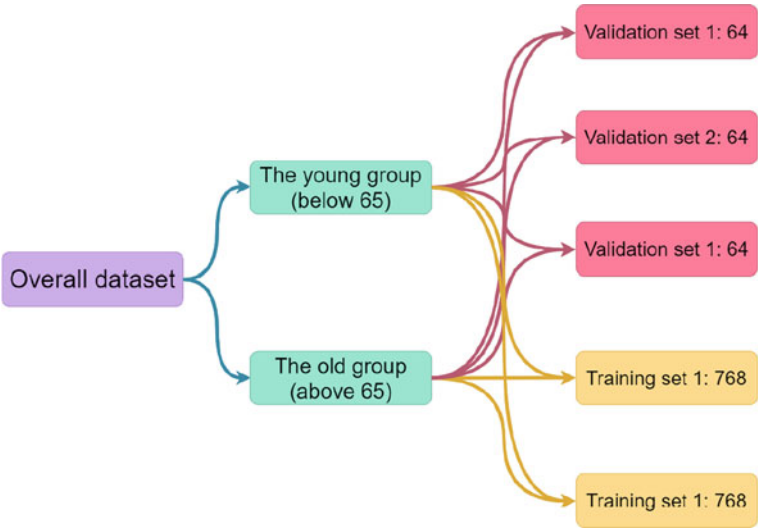

Fig. 5 The data separation procedure

Once these abnormal data samples are concentrated in a certain set, the stability and robustness of our proposed model will be inevitably affected. To solve such a problem, we manually divided the UoG dataset into Young (below 65) and Old (above 65) groups according to the samples' annotation. After that, we can randomly pick data from those two sets of data and compose a new data set. This separation process can be visualised as shown in Fig. 5.

The proportion of people in the two age stages is evenly distributed in every one of the data sets. That is to say, in any training set or validation set, the proportion of samples of the elderly and the young is the same. This approach not only avoids the problems of system instability, but also allows our network to have a more comprehensive understanding of the movement characteristics of the elderly and the young people.

5.2 Data Preparation

The signal processing consists of a 128-point Fast Fourier Transform (FFT) per sweep (1 ms) to obtain the range profiles from the raw I/Q data with a Hamming window. A 4th order Butterworth moving target indicator is implemented to remove static clutter. Then, 300 range profiles are accumulated to perform a zero-padded 1200-point FFT for every range bin in the slow-time direction to obtain a range-Doppler map with an overlap factor of 0.90.

After getting the RT and DT signatures, the threshold filtering and attention-based magnitude masking is applied with thresholding factor set at 0.6 of the maximum

value of the signatures. And finally, those processed data were resized to 224 × 224 and packed into 5 independent data sets which include three validation sets and two training sets.

5.3 Training Detail

In the following sub-sections we will cover the details of network training, which includes the network parameter settings, comparison between different fusion elements, analysis toward the classification result and robustness analysis.

Network Parameter Settings: Our network was trained with a batch size of 32 for 50 epochs. To make our network converge faster in the early stage and be more stable in the later training stage, we have introduced a dynamic learning rate into the training process. This dynamic learning rate will decay exponentially from 0.01 to 0.00001, and has a decay rate of 0.7 and decay step of 200,000.

Fusion Element Analysis: In order to fully explore the influence of different fusion methods on the final classification accuracy of our network, we have implemented the four different fusion elements and demonstrate the evaluation results and the confusion matrices in Table 1 and Fig. 6.

Similarity—The four fusion models have high recognition accuracy for A01, A02, A03 and A06. However, for A05 and A06, although our model improves the recognition accuracy of A05 and A06 compared with ordinary CNN networks, there is

Table 1 Human activity classification accuracy using different fusion models

	Avg acc (%)	Avg class acc (%)	A01 (%)	A02 (%)	A03 (%)	A04 (%)	A05 (%)	A06 (%)
Element-wise average	92.1	91.6	100.0	94.6	96.9	85.7	72.1	100.0
Element-wise product	88.5	90.5	100.0	91.9	96.8	85.2	69.6	100.0
Element-wise sum	91.4	91.3	100.0	97.3	96.8	67.8	86.0	100.0
Concatenation	89.1	87.8	94.7	96.9	91.8	74.2	81.25	100.0
Element-wise Average + Deep Fusion	92.2	92.7	100.0	97.1	91.7	75.8	91.4	100.0
PointNet [24]	92.2	92.8	96.6	96.7	86.7	83.3	83.3	100.0
CNN (no fusion applied)	89.6	90.4	97.2	97.1	88.9	93.3	65.7	100.0

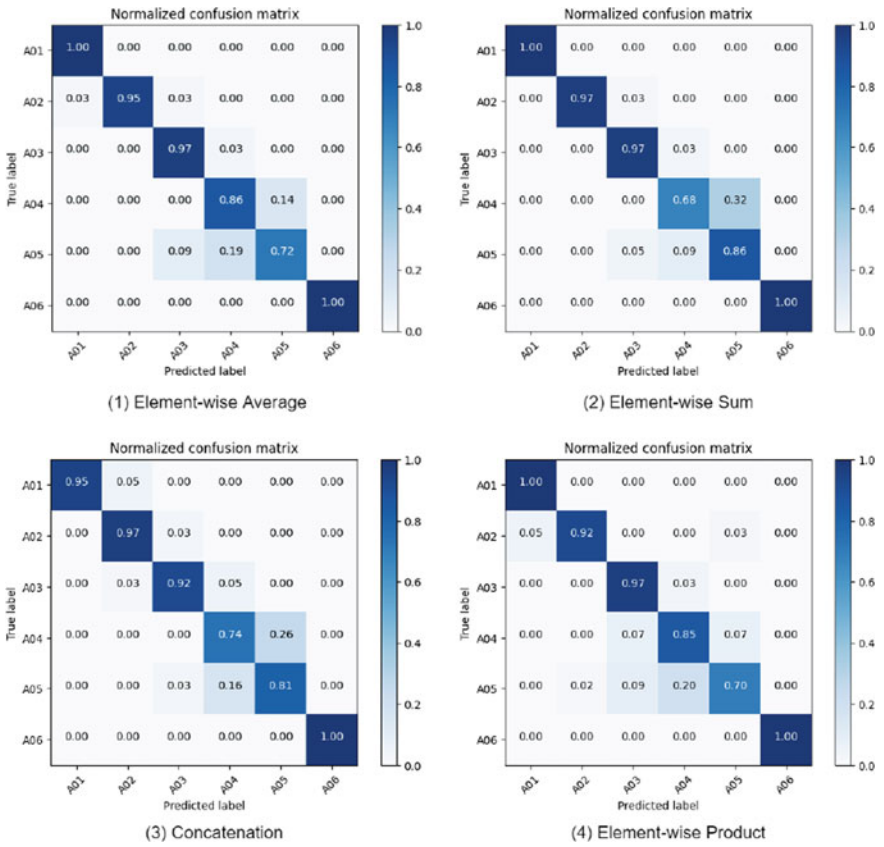

Fig. 6 Confusion matrices using four different fusion models

still an obvious confusion. This is probably caused by the high similarity between the picking up objects and drinking water activities.

Difference Analysis—By comparing the above four basic fusion methods, element-wise average and element-wise sum have relatively higher classification accuracy compared to element-wise product and concatenation. On the one hand, this is because the two methods are very similar, and on the other hand, the two methods allow for deeper integration of radar multimodal data, and therefore can achieve higher recognition accuracy. But for method like concatenation, which leads to a high dimension intermediate layer. Such operation will not only increase the computation complexity but also decrease the interactions between two independent radar modalities. That is why it has the worst recognition accuracy.

Since element-wise average has the best performance (92.1%), we have introduced a hybrid fusion model which can combine the element-wise average and deep fusion (Fig. 4) together. This method, as illustrated, yields quite good results, especially for identifying A01, A02, and A06. However, this method has a drawback, the weight

sharing of the fully connected layer will increase the computational load for network training (twice the computation complexity). Element-wise average seems to be a better choice as the performance decrease is only 0.1% with half the required computation. In addition, we also tried the method of extracting point clouds from RD sequence inspired from [15], and then do the classification on the point cloud using PointNet [25] and reported in [26]. We use the results of this experiment as a benchmark to compare with our proposed method. Compared with PointNet, our multimodal fusion method achieves similar recognition accuracy with element-wise average with fewer calculations.

6 Conclusion

This paper investigated a method of radar multimodal fusion, which can enable CNN to improve the recognition of six kinds of human activities.

In the data pre-processing stage, we put forward a data pre-processing pipeline, which can integrate the phase and range information at the pixel level and reduce the unwanted noise in the raw radar data to a certain extent. We believe that this method can effectively reduce the computing burden of the system and improve the robustness of the system at the same time.

During the network training, we adopted four different fusion elements and two different FCL structures. Compared to traditional single-input CNN (89.6%), our fusion network has stronger HAR capability, especially when using element-wise average method (92% accuracy) to do the feature-level fusion. We have shown that we can achieve similar performances to more complex deep learning methods [26] with a lighter implementation by exploiting element-wise average feature-level fusion. The light implementation is in part achieved thanks to the reduced the number of network inputs through'Attention' fusion.

The next level of classification algorithms will need to integrate the complex form as a native data format as input as exemplified in [29] for example. The exploitation of the phase information has been shown in [26] to accelerate convergence in training as well as improving performances.

Acknowledgements The authors would like to thank the British Council 515095884 and Campus France 44764WK—PHC Alliance France-UK, and PHC Cai Yuanpei—41457UK for their financial support. The authors would also like to give special thanks to Mr. J.G. for his valuable contribution to the elaboration of this article.

References

1. Li, X., He, Y., Jing, X.: A survey of deep learning-based human activity recognition in radar. Remote Sens. **11**(9), 1068 (2019)
2. Clemente, C., Balleri, A., Woodbridge, K., Soraghan, J.J.: Developments in target micro-doppler signatures analysis: radar imaging, ultrasound and through-the-wall radar. EURASIP J. Adv. Sig. Process. **2013**(1), 1–18 (2013)
3. Gurbuz, S.Z., Amin, M.G.: Radar-based human-motion recognition with deep learning: promising applications for indoor monitoring. IEEE Sig. Process. Mag. **36**(4), 16–28 (2019)
4. Kim, Y., Ling, H.: Human activity classification based on micro-doppler signatures using a support vector machine. IEEE Trans. Geosci. Remote Sens. **47**(5), 1328–1337 (2009)
5. Smith, K.A., Csech, C., Murdoch, D., Shaker, G.: Gesture recognition using mm-wave sensor for human-car interface. IEEE Sens. Lett. **2**(2), 1–4 (2018)
6. Gürbüz, S.Z., Erol, B., Cagliyan, B., Tekeli, B.: Operational assessment and adaptive selection of micro-doppler features. IET Radar, Sonar Navigat. **9**(9), 1196–1204 (2015)
7. Kim, Y., Moon, T.: Human detection and activity classification based on microdoppler signatures using deep convolutional neural networks. IEEE Geosci. Remote Sens. Lett. **13**(1), 8–12 (2015)
8. Craley, J., Murray, T.S., Mendat, D.R., Andreou, A.G.: Action recognition using micro-Doppler signatures and a recurrent neural network. In: 2017 51st Annual Conference on Information Sciences and Systems (CISS), pp. 1–5. IEEE (2017)
9. Jokanović, B., Amin, M.: Fall detection using deep learning in range-doppler radars. IEEE Trans. Aerosp. Electron. Syst. **54**(1), 180–189 (2018)
10. Wang, S., Song, J., Lien, J., Poupyrev, I., Hilliges, O.: Interacting with soli: exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum. In: Proceedings of the 29th Annual Symposium on User Interface Software and Technology, pp. 851–860 (2016)
11. Wang, M., Zhang, Y.D., Cui, G.: Human motion recognition exploiting radar with stacked recurrent neural network. Digit. Sig. Process. **87**, 125–131 (2019)
12. Erol, B., Amin, M.G.: Radar data cube analysis for fall detection. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2446–2450. IEEE (2018)
13. Erol, B., Amin, M.G.: Radar data cube processing for human activity recognition using multisubspace learning. IEEE Trans. Aerosp. Electron. Syst. **55**(6), 3617–3628 (2019)
14. Du, H., Jin, T., Song, Y., Dai, Y., Li, M.: A three-dimensional deep learning framework for human behavior analysis using range-doppler time points. IEEE Geosci. Remote Sens. Lett. **17**(4), 611–615 (2019)
15. Baltrušaitis, T., Ahuja, C., Morency, L.-P.: Multimodal machine learning: a survey and taxonomy. IEEE Trans. Pattern Anal. Mach. Intell. **41**(2), 423–443 (2018)
16. Hunt, A.J., Black, A.W.: Unit selection in a concatenative speech synthesis system using a large speech database. In: 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings, vol. 1, pp. 373–376. IEEE (1996)
17. Karpathy, A., Fei-Fei, L.: Deep visual-semantic alignments for generating image descriptions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3128–3137 (2015)
18. Li, S., Kang, X., Fang, L., Hu, J., Yin, H.: Pixel-level image fusion: a survey of the state of the art. Inform. Fus. 33, 100–112 (2017)
19. Rockinger, O.: Pixel-level fusion of image sequences using wavelet frames. In: Proceedings of 16th Leeds Annual Statistical Research Workshop, Citeseer, pp. 149–154 (1996)
20. Naidu, V., Raol, J.R.: Pixel-level image fusion using wavelets and principal component analysis. Def. Sci. J. **58**(3), 338 (2008)
21. Chen, X., Ma, H., Wan, J., Li, B., Xia, T.: Multi-view 3d object detection network for autonomous driving. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1907–1915 (2017)

22. Kor, S., Tiwary, U.: Feature level fusion of multimodal medical images in lifting wavelet transform domain. In: The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 1, pp. 1479–1482. IEEE (2004)
23. Li, H., Mehul, A., Le Kernec, J., Gurbuz, S.Z., Fioranelli, F.: Sequential human gait classification with distributed radar sensor fusion. IEEE Sens. J. **21**(6), 7590–7603
24. Jokanovic, B., Amin, M., Ahmad, F.: Radar fall motion detection using deep learning. In: 2016 IEEE Radar Conference (RadarConf), pp. 1–6. IEEE (2016)
25. Qi, C.R., Su, H., Mo, K., Guibas, L.J.: Pointnet: deep learning on point sets for 3D classification and segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 652–660 (2017)
26. Guo, J., Shu, C., Zhou, Y., Wang, K., Fioranelli, F., Romain, O., Le Kernec, J.: Complex field-based fusion network for human activities classification with radar. In: IET International Radar Conference 2020, Chongqing City, China, 4–6 November 2020, pp. 68–73
27. Fioranelli, F., Shah, S.A., Li, H., Shrestha, A., Yang, S., Le Kernec, J.: Radar signatures of human activities (2019)
28. Fioranelli, F., Shah, S.A., Li, H., Shrestha, A., Yang, S., Le Kernec, J.: Radar sensing for healthcare. Electron. Lett. **55**(19), 1022-10
29. Gao, J., Deng, B., Qin, Y., Wang, H., Li, X.: Enhanced radar imaging using a complex-valued convolutional neural network. IEEE Geosci. Remote Sens. Lett. **16**(1), 35–39 (2019)