Information Fusion 55 (2020) 105–115

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Full Length Article

A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare

Md. Zia Uddin[a], Mohammed Mehedi Hassan[b,c,*], Ahmed Alsanad[b], Claudio Savaglio[d]

[a] Department of Informatics, University of Oslo, Oslo, Norway
[b] College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[c] Research Chair of Pervasive and Mobile Computing, King Saud University, Riyadh 11543, Saudi Arabia
[d] Department of Informatics, Modeling, Electronics and Systems (DIMES), University of Calabria, 87036 Rende, CS, Italy

ARTICLE INFO

Keywords:
Body sensor data fusion
Behavior recognition
Deep recurrent neural network
Robust healthcare

ABSTRACT

Recently, human healthcare from body sensor data has been getting remarkable research attentions by a huge range of human-computer interaction and pattern analysis researchers due to its practical applications such as smart health care systems. For example, smart wearable-based behavior recognition system can be used to assist the rehabilitation of patients in a smart clinic to improve the rehabilitation process and to prolong their independent life. Although there are many ways of using distributed sensors to monitor vital signs and behavior of people, physical human action recognition via body sensors provides valuable data regarding an individual's functionality and lifestyle. In this work, we propose a body sensor-based system for behavior recognition using deep Recurrent Neural Network (RNN), a promising deep learning algorithm based on sequential information. We perform data fusion from multiple body sensors such as electrocardiography (ECG), accelerometer, magnetometer, etc. The extracted features are further enhanced via kernel principal component analysis (KPCA). The robust features are then used to train an activity RNN, which is later used for behavior recognition. The system has been compared against the conventional approaches on three publicly available standard datasets. The experimental results show that the proposed approach outperforms the available state-of-the-art methods.

1. Introduction

Body sensors have become very popular these days for many practical applications such as entertainment, security, wellness and healthcare. One major advantage of using body sensors in monitoring people is they can be used to recognize people's vital signs and behavior more accurately compared to ambient sensors (I would suggest rephrase as "compared to ambient sensors, body sensors recognize…more accurately, thus ensuring and improving. Therefore, likewise, personal computers, also body sensors are expected to have a disruptive impact on our life)). Hence, the sensors can be adopted to ensure and improve peoples' sound living. Thus, body sensors seem to be quite impressive to help in revolutionizing our life similar to personal computers. In commercial fields, wearable sensor-based systems are mostly used in the form of emergency buttons to ask for emergency help, and they have been commercially successful so far [1–10]. Wearable sensors can also be adopted to regularly montioring patients with serious diseases such as cardiac-attacks, etc. Wearable sensor can also be quite helpful during the rehabilitation process after operation. The physiologically vital signs and patient's behaviour can be monitored with the help of useful body sensors and provide them audio feedbacks. Besides, wearable technologies can also provide audio feedback to the patients. The health status and behavior of the patient can also be monitored remotely by the family members, caregivers and doctors [3]. Regarding body sensor-based behavior and health monitoring, a huge amount of research is ongoing to develop smart systems, especially fall-detection systems to detect fall of elderly when living alone [4–6]. Body sensors in devices such as heart beat monitoring and smartwatches are rapidly getting more and more famous. The body sensor-based technologies also seem to help in future medical technologies to define doctor-patient relationship and to reduce healthcare cost. The rapid growth of body sensor-based technologies show its continuing in vast sectors of healthcare.

Human behaviour recognition research has been gaining traction in context-aware systems for many applications [7–9], including pervasive computing, surveillance, AAL, etc. Researches related to human behavior analysis is also very relevant in pervasive computing, surveillance, and ambient assistive services. Recognizing postures of human body and body motion are quite important to assist in improving one's health

* Corresponding author at: King Saud University, Riyadh 11543, Saudi Arabia.
E-mail addresses: mdzu@ifi.uio.no (Md.Z. Uddin), mmhassan@ksu.edu.sa (M.M. Hassan), aasanad@ksu.edu.sa (A. Alsanad), csavaglio@dimes.unical.it (C. Savaglio).

https://doi.org/10.1016/j.inffus.2019.08.004
Received 30 March 2019; Received in revised form 3 July 2019; Accepted 3 August 2019
Available online 4 August 2019
1566-2535/

systems. In [7], the authors discussed some medical applications focusing on specially activity pattern analysis for better healthcare and sports. The authors also showed some instances in healthcare diagnosis systems, rehabilitation process, emotional systems and elderly care services. Furthermore, they also discussed several assisted living projects used to help in improving lifestyle, ensuring medical services, better childcare, and people having cognitive problems. In [8], the authors discussed activity monitoring technologies and the links of activities with common diseases. They also focused on daily activity patterns that contributes to enhance the process and diagnosis of neurological and respiratory problems. In [9], the researchers proposed a method based on mobile phone sensors to estimate the energy loss during physical activities in normal living conditions. Thus, assistive technologies for better care of people, can take advantage of body sensor-based behaviour recognition and monitoring system to deliver better personalized services.

For behaviour and activity monitoring, video sensors have been most popular so far for pattern analysis in body movement. The video-based methods are basically based on the single or image sequences acquired by cameras [10–12]. Image quality and parameters in the environment such as illumination variations affects performance of the system quite a lot. Besides, video-based system always carries a big issue of privacy by the users. Non-video-based systems are mostly free from such issues [13–18]. Among the other sensors, body sensors are very popular to precisely provide body data to model different events helpful for useful pattern recognition for different applications such as smart activity recognition for smart healthcare. There are some drawbacks to the body sensors as well such as wearing devices or sensors for a long time seems very cumbersome, but they may be sacrificed by the users for their better healthcare and lifestyle.

In body sensor-based activity recognition systems, robust machine learning algorithms should be applied to handle noisy data. Different body sensor characteristics may vary from subjects to subjects. Therefore, it is quite important to apply a robust machine learning technique to model different behaviour in a smart healthcare system. Regarding machine learning, typical hidden Markov models (HMM) are used to decode and model time-sequential events [19]. Recently, deep learning techniques have been ruling over other pattern recognition techniques for a vast range of applications [20–22]. Amongst the deep learning techniques, Deep Belief Network (DBN) was the first proposed and used for pattern analysis applications [20]. DBN utilizes restricted Boltzmann machines that makes the training process very rapid than typical big neural network. Later on, Convolutional Neural Networks (CNN) came to overpower DBN, mostly to revolutionize image process applications. A huge number of image processing and application researchers quickly adopted CNN as it consisted better discriminative power than DBN. CNN does feature extraction by itself also it uses convolutional stacks to generate a hierarchy of abstract features. The general CNN has five basic parts: convolution, pooling, tangent squashing, rectifier, and normalization [22]. Therefore, CNN is mostly famous and applied to robustly analyse visual patterns such as object modelling, detection, recognition in huge image archives. To model temporally oriented events, CNN does not seem to be applicable as it was targeted with the goal to do analysis on single image-based pattern recognition. However, analysis temporal events in time-sequential applications Recurrent Neural Networks (RNNs) has become very poplar day by day. It can provide discriminative power over other deep learning approaches so far [23–31]. Hence, we adopt RNN in this body sensor-based to train and recognize different activities for better healthcare.

Fig. 1 shows a body sensor-based human activity recognition system for robust healthcare where a person is wearing different body sensors in different body parts i.e., chest, wrist, and ankle. The sensor data is acquired and saved to a computer via the wireless medium. The computer then processes the rest steps such as feature extraction and modelling the activities using deep RNN. Fig. 2 shows the basic flows of the proposed system where there are two basic steps: training and testing.

The contributions of this work are as follows:

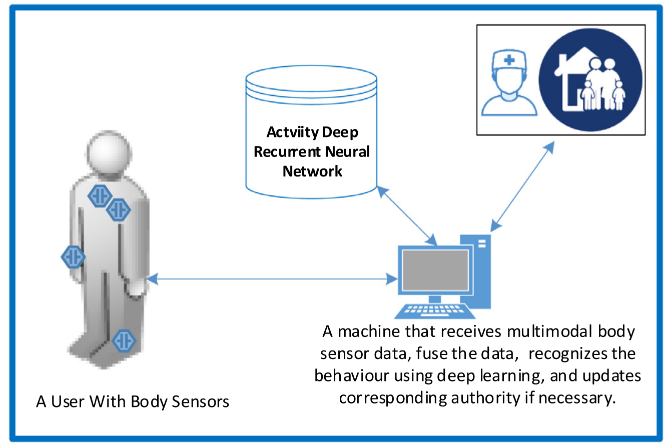

Fig. 1. A schematic setup for body sensor-based human activity prediction system.

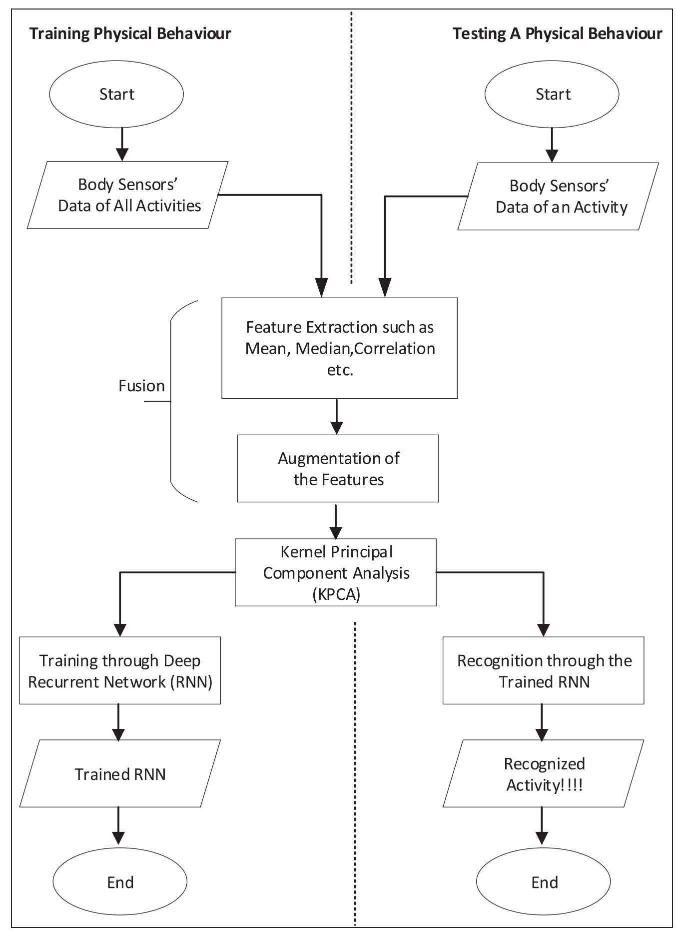

Fig. 2. Basic flows of the proposed body sensor-based physical activity recognition system.

- We propose here a robust deep RNN-based activity recognition system based on efficient features from different wearable body sensor data fusion.
- After extracting the data from the different multimodal sensors, the data is further fused with statistical features of different orders such as mean, variance, standard deviation, skewness, and kurtosis.
- To make the features more robust, they are projected into a nonlinear space with dimension reduction process. For which,

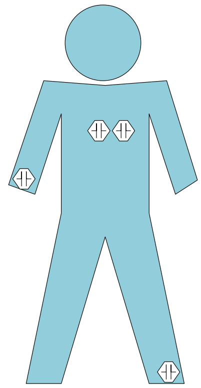

Fig. 3. Sensor placements for MHEALTH dataset.

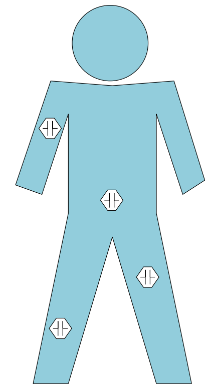

Fig. 4. Sensor placements for PUC-Rio dataset.

kernel-based principal component analysis (KPCA) is applied on the features. In KPCA, the features go via a Gaussian kernel followed by typical PCA to transform the features in nonlinear space, which is better than typical linear feature space-based PCA. The robust features are then trained and tested via a memory-based RNN for activity recognition.

- To check and compare the proposed approach with the traditional ones, we choose three publicly available datasets and adopted different combinations of training and testing methods. The proposed approach shows superior performance on all the datasets.

2. Data and feature processing from MHEALTH dataset

MHEALTH (mobile health) public dataset is first considered for feature analysis and activity recognition [32,33]. The body sensors are placed on the user's body just as shown in Fig. 3.

The places of attaching sensors are chest, right wrist and left ankle. There is an ECG healthcare sensor placed on the chest that has 2-lead ECG measurements of heart data. The other sensors in the system are ralated to motion related data. For instance, the accelerometers provide body acceleration, gyroscopes the rate of turn, and magnetometer magnetic field orientation. The body sensor data are represented as follows. The accelerometer data in chest can be represented as

$$C_C = (A_x, A_y, A_z). \tag{1}$$

ECG sensor data in the chest can be represented as

$$G = g_1 || g_2. \tag{2}$$

The accelerometer data in the left ankle can be represented as

$$C_{LA} = (L_x, L_y, L_z). \tag{3}$$

The gyroscope sensor data in the left-ankle can be represented as

$$Y_{LA} = (R_x, R_y, R_z). \tag{4}$$

The magnetometer data in the left ankle can be represented as

$$N_{LA} = (T_x, T_y, T_z). \tag{5}$$

The accelerometer data in the right wrist can be represented as

$$A_{RW} = (I_x, I_y, I_z). \tag{6}$$

The gyroscope sensor data from the left-wrist can be represented as

$$R_{LW} = (S_x, S_y, S_z). \tag{7}$$

The gyroscope sensor data from the right-wrist can be represented as

$$R_{RW} = (W_x, W_y, W_z). \tag{8}$$

The magnetometer data in the right wrist can be represented as

$$N_{RW} = (E_x, E_y, E_z). \tag{9}$$

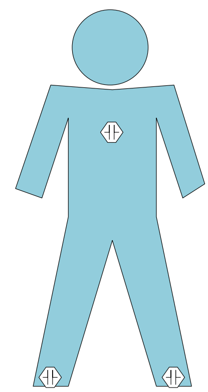

Fig. 5. Sensor placements for AReM dataset.

Furthermore, we extract more features such as mean m, variance v, standard deviation s, and skewness w from the data. Then, the features obtained from a signal of specific time-period for an activity are fused together by augmenting them horizontally as F. Furthermore, we applied kernel principal component analysis (KPCA) on the features. To do that, the features go through a Gaussian kernel and then typical PCA is applied on the transformed features. The KPCA feature space is represented as $K_{mhealth}$. Finally, the fused features $F_{mhealth}$ is applied on the KPCA feature space as

$$P_{mhealth} = F_{mhealth} K^T_{mhealth} \tag{10}$$

3. Data and feature processing from PUC-Rio dataset

The second body sensor database we focused on is PUC-Rio behavior recognition dataset [34]. To record the dataset, accelerometers were placed in four different positions of the body and the positions were waist, left thigh, right ankle, and right arm, respectively. The sensor placement is shown in Fig. 4. Before the data collection process, the accelerometer sensors were calibrated. The calibration process consisted of positioning of the sensors as well as the performance of the reading sensor values. During the data processing time, the read parameters of every axis of the accelerometer sensor were subtracted from the values obtained at prior step of the time of the calibration. To maintain the good standard of data collection process, the body sensors were placed on the top of a flat table in the same spot and then calibration was done. The second type of calibration was done based on the subjects where the accelerometers were calibrated, and data reading was done after placing the sensors in the same places of the bodies of the subjects.

Activitiy	Precision	Recall	F1-Score	Support
A1	1.00	1.00	1.00	1536
A2	1.00	1.00	1.00	1536
A3	1.00	1.00	1.00	1536
A4	1.00	1.00	1.00	1536
A5	1.00	1.00	1.00	1536
A6	1.00	1.00	1.00	1536
A7	1.00	0.97	0.98	1536
A8	0.97	1.00	0.98	1689
A9	1.00	1.00	1.00	1536
A10	0.98	0.99	0.99	1536
A11	0.99	0.98	0.99	1536
A12	0.98	0.97	0.97	538
Avg/Total	**0.99**	**0.99**	**0.99**	**17587**

(a)

Activitiy	Precision	Recall	F1-Score	Support
A1	1.00	1.00	1.00	1536
A2	1.00	1.00	1.00	1536
A3	1.00	1.00	1.00	1536
A4	1.00	1.00	1.00	1536
A5	1.00	1.00	1.00	1536
A6	1.00	1.00	1.00	1587
A7	1.00	1.00	1.00	1664
A8	1.00	1.00	1.00	1715
A9	1.00	1.00	1.00	1536
A10	0.91	0.92	0.92	1536
A11	0.92	0.91	0.91	1536
A12	0.99	0.99	0.99	512
Avg/Total	**0.98**	**0.98**	**0.98**	**17766**

(b)

Activitiy	Precision	Recall	F1-Score	Support
A1	1.00	1.00	1.00	1536
A2	1.00	1.00	1.00	1536
A3	1.00	1.00	1.00	1536
A4	1.00	1.00	1.00	1536
A5	1.00	1.00	1.00	1536
A6	1.00	1.00	1.00	1613
A7	1.00	1.00	1.00	1689
A8	1.00	1.00	1.00	1588
A9	1.00	1.00	1.00	1536
A10	0.99	1.00	0.99	1536
A11	1.00	1.00	1.00	1536
A12	0.99	0.97	0.98	512
Avg/Total	**1.00**	**1.00**	**1.00**	**17690**

(c)

Activitiy	Precision	Recall	F1-Score	Support
A1	1.00	1.00	1.00	1536
A2	1.00	1.00	1.00	1536
A3	1.00	1.00	1.00	1536
A4	1.00	1.00	1.00	1536
A5	1.00	1.00	1.00	1536
A6	1.00	1.00	1.00	1664
A7	1.00	1.00	1.00	1638
A8	1.00	1.00	1.00	1562
A9	1.00	1.00	1.00	1536
A10	1.00	1.00	1.00	1536
A11	1.00	1.00	1.00	1536
A12	1.00	0.97	0.99	512
Avg/Total	**1.00**	**1.00**	**1.00**	**17664**

(d)

Activitiy	Precision	Recall	F1-Score	Support
A1	1.00	1.00	1.00	1536
A2	1.00	1.00	1.00	1536
A3	1.00	1.00	1.00	1536
A4	0.99	1.00	1.00	1536
A5	0.99	0.99	0.99	1536
A6	1.00	1.00	1.00	1382
A7	1.00	1.00	1.00	1434
A8	1.00	1.00	1.00	1357
A9	1.00	1.00	1.00	1536
A10	1.00	1.00	1.00	1536
A11	0.99	0.99	0.99	1536
A12	0.99	0.98	0.98	512
Avg/Total	**1.00**	**1.00**	**1.00**	**16973**

(e)

Activitiy	Precision	Recall	F1-Score	Support
A1	1.00	1.00	1.00	1536
A2	1.00	1.00	1.00	1536
A3	1.00	1.00	1.00	1536
A4	1.00	1.00	1.00	1536
A5	1.00	1.00	1.00	1536
A6	1.00	1.00	1.00	1101
A7	1.00	1.00	1.00	1049
A8	1.00	1.00	1.00	1152
A9	1.00	1.00	1.00	1536
A10	1.00	1.00	1.00	1536
A11	1.00	1.00	1.00	1536
A12	1.00	0.99	0.99	512
Avg/Total	**1.00**	**1.00**	**1.00**	**16102**

(f)

Activitiy	Precision	Recall	F1-Score	Support
A1	1.00	1.00	1.00	1536
A2	1.00	1.00	1.00	1536
A3	1.00	1.00	1.00	1536
A4	1.00	1.00	1.00	1536
A5	1.00	1.00	1.00	1536
A6	1.00	1.00	1.00	1536
A7	1.00	1.00	1.00	1382
A8	1.00	1.00	1.00	1408
A9	1.00	1.00	1.00	1536
A10	1.00	1.00	1.00	1536
A11	1.00	1.00	1.00	1536
A12	1.00	0.99	1.00	512
Avg/Total	**1.00**	**1.00**	**1.00**	**17126**

(g)

Activitiy	Precision	Recall	F1-Score	Support
A1	1.00	1.00	1.00	1536
A2	1.00	1.00	1.00	1536
A3	1.00	1.00	1.00	1536
A4	1.00	1.00	1.00	1536
A5	1.00	1.00	1.00	1536
A6	1.00	0.99	1.00	1075
A7	0.99	1.00	1.00	1511
A8	1.00	1.00	1.00	1280
A9	1.00	1.00	1.00	1536
A10	1.00	1.00	1.00	1536
A11	1.00	1.00	1.00	1536
A12	0.99	0.99	0.99	512
Avg/Total	**1.00**	**1.00**	**1.00**	**16666**

(h)

Activitiy	Precision	Recall	F1-Score	Support
A1	1.00	1.00	1.00	1536
A2	1.00	1.00	1.00	1536
A3	1.00	1.00	1.00	1536
A4	1.00	1.00	1.00	1536
A5	1.00	1.00	1.00	1536
A6	1.00	1.00	1.00	1433
A7	1.00	1.00	1.00	1434
A8	1.00	1.00	1.00	1484
A9	1.00	1.00	1.00	1536
A10	1.00	1.00	1.00	1536
A11	1.00	1.00	1.00	1536
A12	1.00	0.99	1.00	538
Avg/Total	**1.00**	**1.00**	**1.00**	**17177**

(i)

Activitiy	Precision	Recall	F1-Score	Support
A1	1.00	1.00	1.00	1536
A2	1.00	1.00	1.00	1536
A3	1.00	1.00	1.00	1536
A4	1.00	1.00	1.00	1536
A5	1.00	1.00	1.00	1536
A6	1.00	1.00	1.00	1229
A7	1.00	1.00	1.00	1382
A8	1.00	1.00	1.00	1434
A9	1.00	1.00	1.00	1536
A10	1.00	1.00	1.00	1536
A11	1.00	1.00	1.00	1536
A12	1.00	0.99	1.00	512
Avg/Total	**1.00**	**1.00**	**1.00**	**16845**

(j)

Activitiy	Precision	Recall	F1-Score	Support
A1	1.00	1.00	1.00	1536.00
A2	1.00	1.00	1.00	1536.00
A3	1.00	1.00	1.00	1536.00
A4	1.00	1.00	1.00	1536.00
A5	1.00	1.00	1.00	1536.00
A6	1.00	1.00	1.00	1415.60
A7	1.00	1.00	1.00	1471.90
A8	1.00	1.00	1.00	1466.90
A9	1.00	1.00	1.00	1536.00
A10	0.99	0.99	0.99	1536.00
A11	0.99	0.99	0.99	1536.00
A12	0.99	0.98	0.99	517.20
Average	**1.00**	**1.00**	**1.00**	**1429.97**

(k)

Fig. 6. Classification report of subject 1 to 10 (a–j) and the average classification report (k) from MHEALTH dataset.

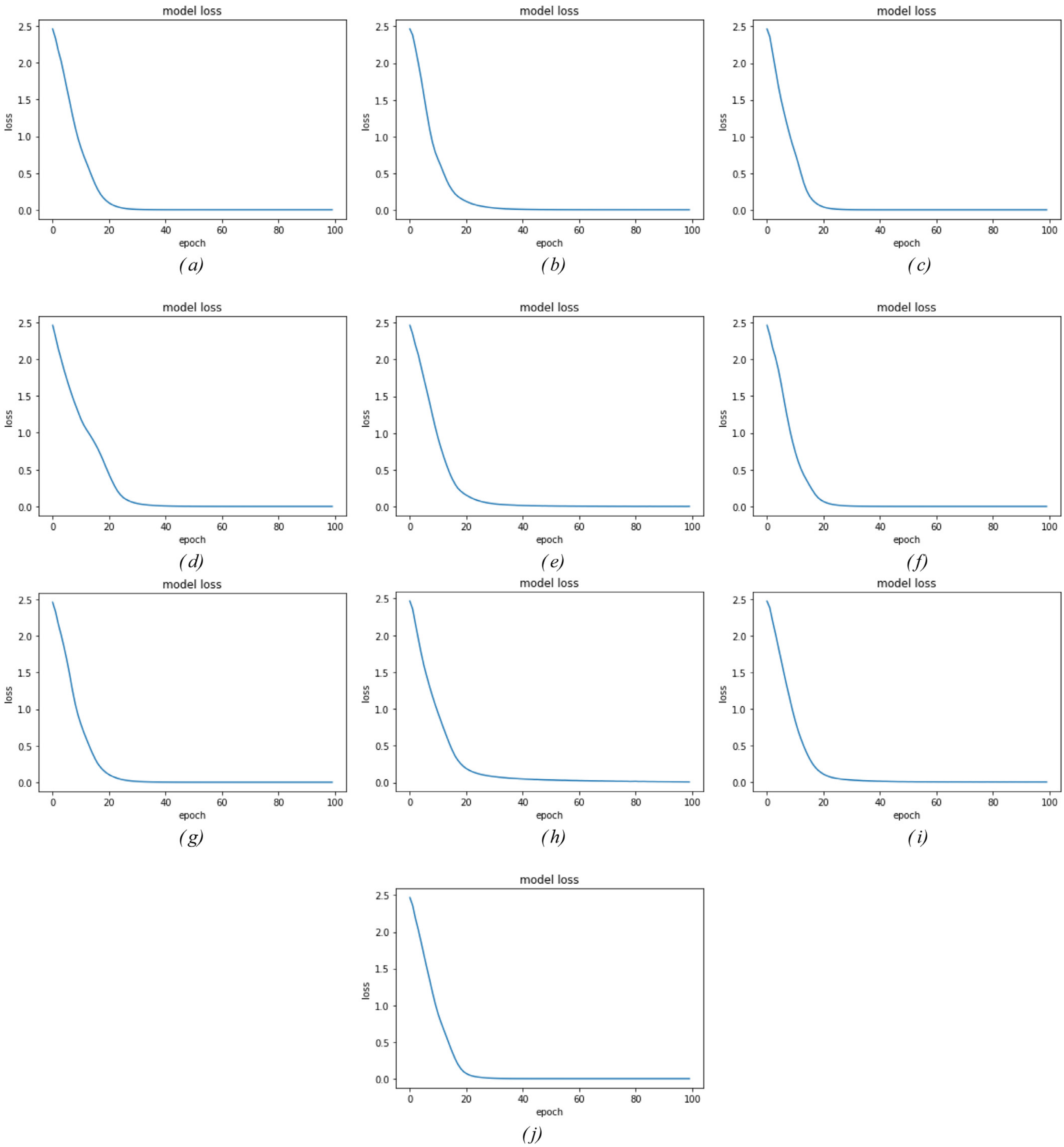

Fig. 7. Training graph of RNN for subject 1–10 (i.e., from (a)–(j)) of MHEALTH dataset.

The sensor data are represented as follows. The accelerometer data in waist can be represented as

$$C_w = (A_x, A_y, A_z). \tag{11}$$

The accelerometer data in the left thigh can be represented as

$$C_{LT} = (L_x, L_y, L_z). \tag{12}$$

The accelerometer data in the right ankle can be represented as

$$W_{RA} = (W_x, W_y, W_z). \tag{13}$$

The accelerometer data in the right upper arm be represented as

$$C_{RA} = (R_x, R_y, R_z). \tag{14}$$

Furthermore, other statistical features such as mean, variance, standard deviation, skewness, and kurtosis are extracted from the data and fusion is done augmenting them with other features. The fused features can be named as U. Then, KPCA is applied on them to find out the direction of variance in nonlinear space. The KPCA feature space is represented as K_{rio}. Finally, the fused features F_{rio} is applied on the KPCA feature space as

$$P_{rio} = F_{rio} K_{rio}^T. \tag{15}$$

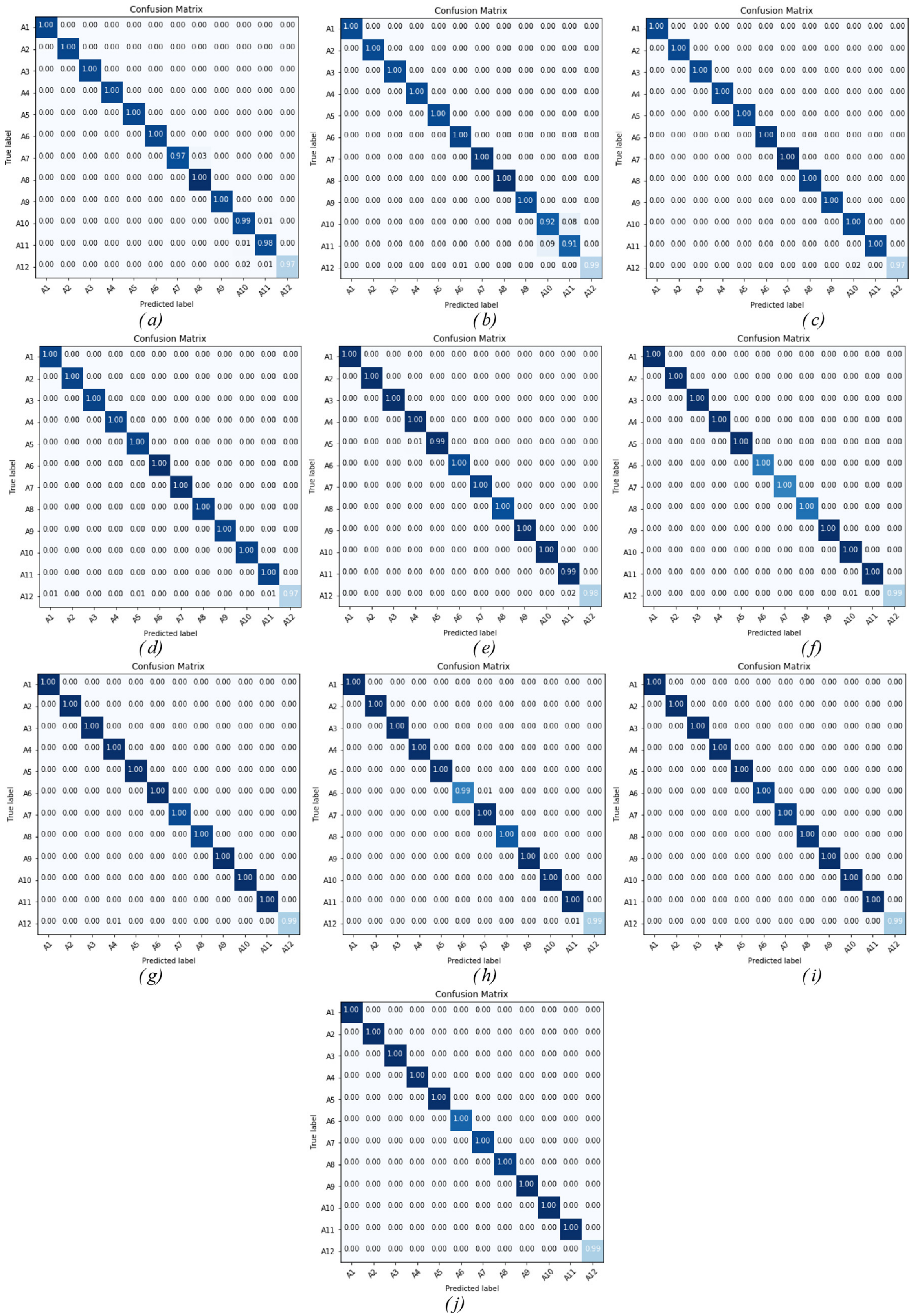

Fig. 8. Confusion matrix using proposed approach for subject 1–10 (i.e., from (a)–(j)) MHEALTH dataset.

4. Data and feature processing from AReM dataset

The third dataset to apply proposed approach that we choose is Activity Recognition system based on Multisensor data fusion (AReM) dataset [35]. The dataset is a benchmark in the area of activity analysis applications. The activity recognition is done from time-series data generated by a Wireless Sensor Network (WSN). In the dataset, the information is used that is coming from the implicit alteration of the wireless channel because of the motion of the user. The sensor devices basically measure the received signal strength (RSS) of the beacon packets during exchanging among themselves in the sensor networks. The sensors are placed on the chest and ankles of the users, as shown in Fig. 5. From the raw data, slightly noise removal is done. In the dataset, an epoch time of 250 ms is chosen. In that time, five RSS samples can be obtained for each of the three couples of body sensor nodes. The features existed in the dataset consists of the mean and variance of the RSS values obtained from the three sensor couples.

The sensor features are augmented as follows:

$$F_{AReM} = (mean\_rss12, var\_rss12, mean\_rss13, var\_rss13, mean\_rss23, var\_rss23). \quad (16)$$

Then, KPCA is applied and feature space is represented as K_{AReM}. Finally, the fused features F is applied on the KPCA feature space as

$$P_{AReM} = F_{AReM} K^{T}_{AReM}. \quad (17)$$

5. Recurrent neural network (RNN) for sequence modeling

As this work focuses on time-sequential body sensor data for human behavior analysis, machine learning model capable of modeling sequence-based information should be adopted in this regard. Amongst the machine learning models applied by researchers in different fields, RNN has got very popular due to its robustness of modeling events lying in time-sequential data [23]. Hence, RNN seems to be most successful deep learning methods to recognize or predict underlying events in the data. It has recurrent communications within the hidden units that basically keeps the relation with history and the present state. Hence, it RNN consists of memory in the system to model a sequence. However, typical RNN has a problem called vanishing gradient that occurs during modeling long sequences. The problem can be defined as Long-Term Dependencies. To overcome that, there was an algorithm proposed as Long Short-Term Memory (LSTM) [24]. An LSTM block has a cell state and three gates. The gates are input, forget, and the output gate. The input gate V_t is determined as

$$V_t = \beta(Y_{PI} P_t + W_{HI} H_{t-1} + \alpha_I) \quad (18)$$

where Y represents the weight matrix, α bias, and β logistic function. The forget gate G is as

$$G_t = \beta(Y_{PG} P_t + Y_{HG} H_{t-1} + \alpha_G). \quad (19)$$

A cell stores the long-term memory in a state vector B that can be represented as

$$B_t = G_t B_{t-1} + V_t \tanh(Y_{PB} P_t + Y_{HB} H_{t-1} + \alpha_B). \quad (20)$$

The output gate O represents the output as

$$O_t = \beta(Y_{PO} P_t + Y_{HO} H_{t-1} + \alpha_O). \quad (21)$$

The hidden state H is represented as

$$H_t = O_t \tanh(B_t). \quad (22)$$

The final output u can be determined as

$$u = \mathrm{softmax}(Y_u H_l + \alpha_U) \quad (23)$$

where l indicates the he final LSTM number in the neural network.

Activity	Precision	Recall	F1-Score	Support
Sitting	1.00	1.00	1.00	25316
Sitting Down	0.98	0.98	0.98	5913
Standing	1.00	1.00	1.00	23685
Standing Up	0.98	0.97	0.97	6208
Walking	0.99	0.99	0.99	21695
Avg/Total	**0.99**	**0.99**	**0.99**	**82817**

(a)

Activity	Precision	Recall	F1-Score	Support
Sitting	1.00	1.00	1.00	25315
Sitting Down	0.98	0.98	0.98	5914
Standing	0.99	1.00	1.00	23685
Standing Up	0.98	0.97	0.98	6207
Walking	0.99	0.99	0.99	21695
Avg/Total	**0.99**	**0.99**	**0.99**	**82816**

(b)

Activity	Precision	Recall	F1-Score	Support
Sitting	1.00	1.00	1.00	25315.50
Sitting Down	0.98	0.98	0.98	5913.50
Standing	1.00	1.00	1.00	23685.00
Standing Up	0.98	0.97	0.98	6207.50
Walking	0.99	0.99	0.99	21695.00
Avg/Total	**0.99**	**0.99**	**0.99**	**82816.50**

(c)

Fig. 9. Classification report using proposed approach on PUC-Rio dataset for (a) fold-1, (b) fold-2, and (c) average of the folds.

Table 1
Twelve physical activities from MHEALTH dataset and their short forms.

Name	Short form
Standing still	A1
Sitting and relaxing	A2
Lying down	A3
Walking	A4
Climbing stairs	A5
Waist bend forward	A6
Frontal elevation of arms	A7
Knees bending	A8
Cycling	A9
Jogging	A10
Running	A11
Jumping front and back	A12

6. Experimental results

For experiments, we conducted experiments on three different public datasets. The datasets are MHEALTH, PUC-RiO, and AReM respectively.

6.1. Experimental results on MHEALTH dataset

To apply the proposed approach and other conventional methods, we first collected the MHEALTH dataset [32,33]. The dataset was targeted to record physiological vital signs and body movement recordings from 10 subjects while they were doing 12 different physical activities. The activities are shortly named as Table 1.

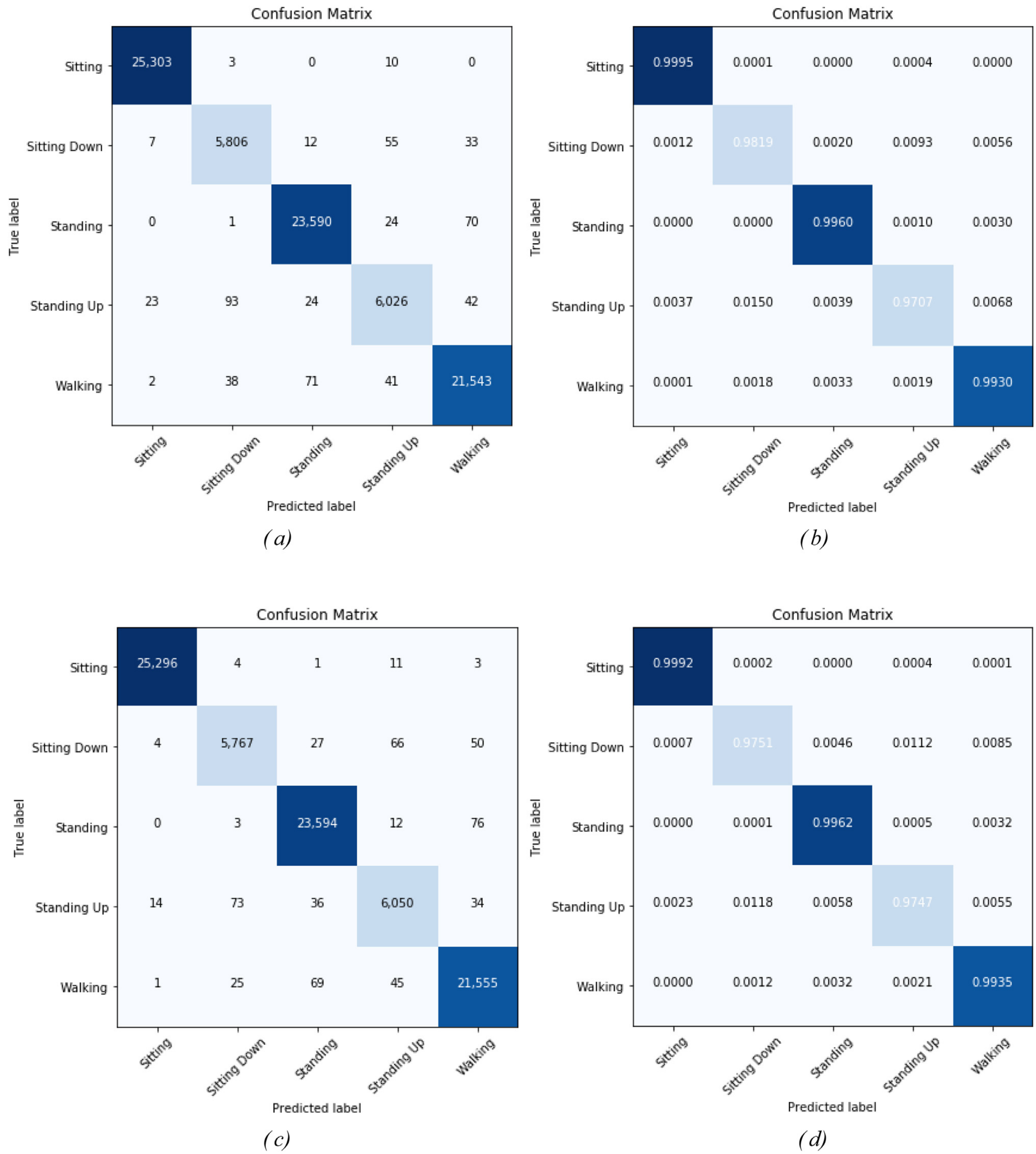

Fig. 10. Confusion matrix using proposed approach on PUC-Rio dataset for fold-1 ((a) by samples and (b) normalized) and fold-2 ((c) by samples and (d) normalized).

As previously mentioned, to collect the dataset, the sensors were placed on the users' chest, right wrist, and left ankle. The data was collected for the dataset was from 2 ECG sensors for heart signal measurement, accelerometers for acceleration during movement, gyroscopes for the rate of turn, and magnetometers to get the orientation of the magnetic field of the subject. The sampling rate used was 50 Hz. We started focusing on individual subject in the dataset for all 12 different activities. We considered 50% of the samples from each subject for training and rest for testing. Regarding subject1, the proposed approach obtained mean recall rate of 0.99. Similarly, for subjects 2 till 10, it obtained as 0.98 , 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, and 1.00, respectively. For decimal places in the numbers, 2 decimal digits after the decimal point were considered. The classification reports obtained for all the subjects indicated convincing precision. Recall, and F1-scores for the corresponding supporting samples as shown in Fig. 6. We also compared the approach with conventional DBN-based approach. The DBN-based approach achieved mean recall rate of 0.93 whereas the proposed approach showed 0.99, as shown in Table 2. The training model loss figures are shown in Fig. 7 where it can be noticed that the loss goes towards zero after some epochs during training for each subject, hence showing a good training tendency of the approach. Besides, Fog. 8 shows the confusion matrices for all the subjects where we can a

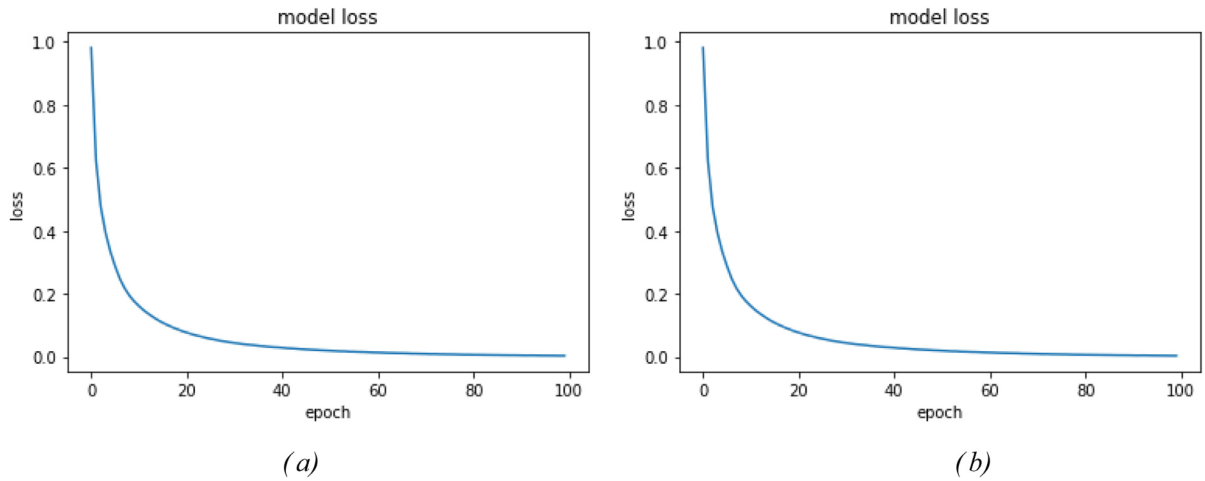

Fig. 11. Training graph using proposed approach on PUC-Rio dataset for (a) fold-1 and (b) fold-2.

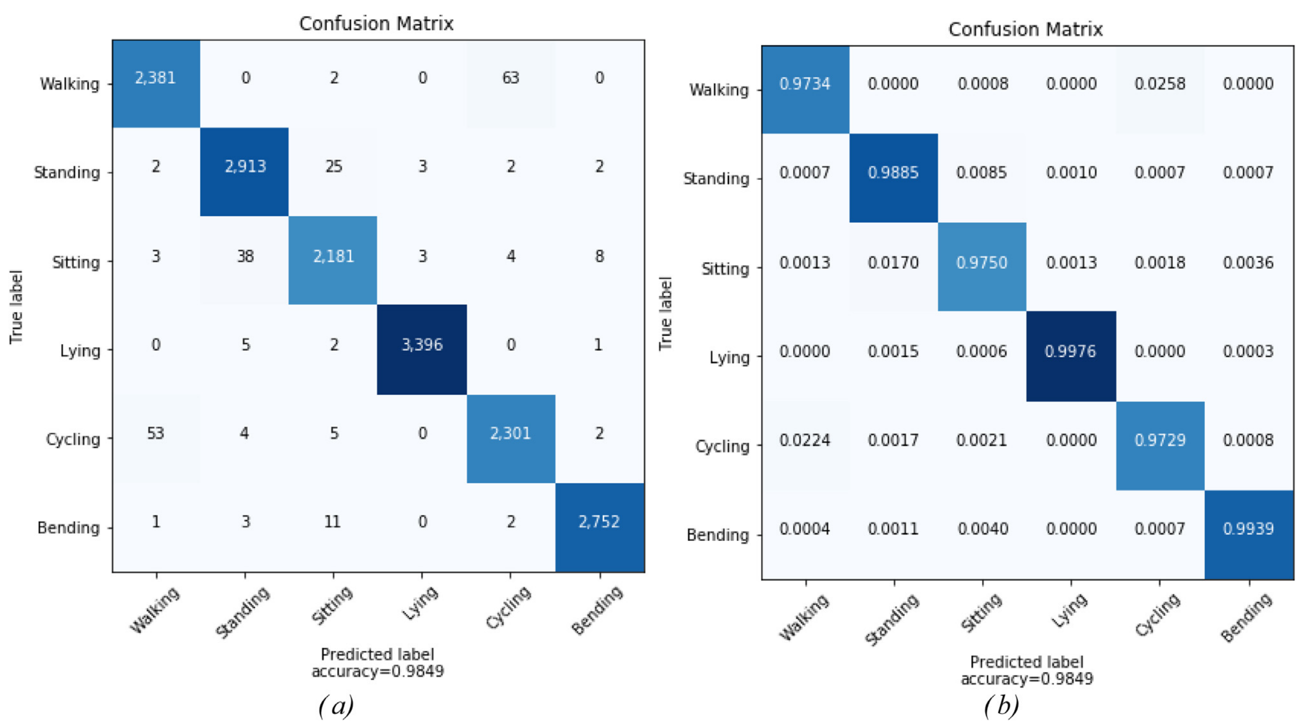

Fig. 12. Confusion matrix using proposed approach on AReM dataset (a) by samples and (b) normalized.

little amount of misrecognition of the samples among different classes (Fig. 8).

6.2. Experimental results on PUC-Rio dataset

We moved to the second public dataset (PUC-Rio Dataset) to check the performance of the proposed system [34]. The dataset consisted of five different activities, namely sitting, sitting down, standing, standing up and walking. The dataset has a total of 165,632 samples for these five activities. A two-fold cross validation was done there. The average precision, recall, and F1-score obtained by the proposed RNN-based systems is 0.99 for each of them as shown in Fig. 9. Like the previous experiments, we chose two decimal places after the decimal point to show the numbers here. Hence, the proposed approach shows a very good performance on the dataset. We also compared the approach with tradition DBN-based method where it achieved an average recall rate of 0.93. Fig. 10 shows the confusion matrices without and with normalized samples where it shows very less confusion within the samples of different classes. Fig. 11 also shows good training condition of the folds. Thus, the proposed approach shows its good performance on the second dataset as well (Table 3).

6.3. Experimental results on AReM dataset

After doing experiments on two public datasets, we continued our experiments to our final dataset called Activity Recognition system based on Multisensor data fusion (AReM). We considered six activities from the dataset, namely walking, standing, sitting, lying, cycling, and bending. We chose random splitting of number of samples for training and testing. So, twenty percent of the whole dataset is considered for testing the approaches. The testing dataset was chosen by random split. For training and testing the approaches we used 1,013,616 and 941,664 samples, respectively. The proposed approach achieved a very good precision, recall, and F1-score of 0.99 as shown in the classification report in Table 4. As like as previous experiments, we have showed 2 decimal places after the decimal points for the results here. The conventional DBN-based ap-

Table 2
Average recall rates by using different approaches to all subjects of MHEALTH dataset.

Activities	DBN-based approach	Proposed RNN-based approach
A1	0.93	1.00
A2	0.94	1.00
A3	0.95	1.00
A4	0.95	1.00
A5	0.90	1.00
A6	0.95	1.00
A7	0.89	1.00
A8	0.94	1.00
A9	0.93	1.00
A10	0.91	0.99
A11	0.93	0.99
A12	0.90	0.98
Mean	**0.93**	**0.99**

Table 3
Average recalls by using different approaches on PUC-Rio dataset.

Activities	DBN-based approach	Proposed RNN-based approach
Sitting	0.94	1.00
Sitting Down	0.95	0.98
Standing	0.93	1.00
Standing Up	0.92	0.97
Walking	0.93	0.99
Mean	**0.93**	**0.99**

Table 4
Classification report using proposed approach on AReM dataset.

Activity	Precision	Recall	F1-Score	Support
Walking	0.98	0.97	0.97	2446
Standing	0.98	0.99	0.99	2947
Sitting	0.98	0.97	0.98	2237
Lying	1.00	1.00	1.00	3404
Cycling	0.97	0.97	0.97	2365
Bending	1.00	0.99	0.99	2769
Average/Total	**0.99**	**0.99**	**0.99**	**16,168**

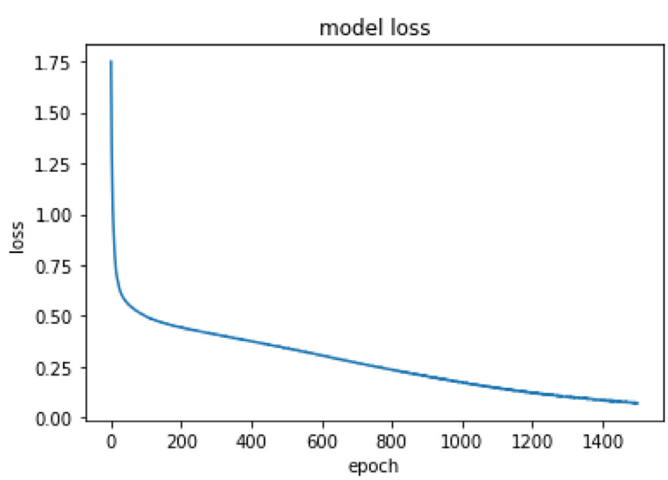

Fig. 13. Training plot proposed approach on AReM dataset.

proach achieved less than 0.90 mean recall rate. Fig. 12 shows confusion matrices by without and with normalized samples of different activity classes. Fig. 13 shows a training graph to train 1,013,616 samples in 1500 epochs. The accuracy achieved via the proposed approach was 0.99 with considering 2 decimal places after the decimal point, shows the superiority and robustness of the approach.

7. Conclusions

Sensor-based user behavior and health status monitoring is getting more and more interest in the huge amount of pattern recognition researchers, with the promise of improving people's wellness, health, and lifetime. Given such goals, smart environments support application which very often demand. Hence, the user care applications in smart environments very often demands continuous observation of the users' activities with the help of an event-driven system. In this work, we proposed a deep RNN-based activity recognition system based on the wearable body sensor fusion data. After extracting the data from the sensors, the multimodal data is fused and followed by KPCA to make it robust. Then, a deep RNN is trained for all the activities, which is later used for testing the data. To check and compare our proposed approach, we have done the experiments on three publicly available datasets. The experimental results have showed good performance by achieving 0.99 precision, recall, F1-score, and accuracies on the big datasets. In the future, the proposed deep RNN-based human activity recognition system can be more analyzed on complex and bigger datasets with more complex activities to get a real-time human behavior monitoring system.

Acknowledgements

This paper was fully financially supported by King Saud University, Saudi Arabia through the Vice Deanship of Research Chairs: Chair of Pervasive and Mobile Computing. Dr. Mohammad Mehedi Hassan is the corresponding author of this paper.

References

[1] H. Aghajan, J.C. Augusto, C. Wu, P. McCullagh, J.-.A. Walkden, Distributed vision-based accident management for assisted living, Proc. Int. Conf. Smart Homes Health Telemat. (2007) 196–205.
[2] S. Bang, M. Kim, S. Song, S.-.J. Park, Toward real time detection of the basic living activity in home using a wearable sensor and smart home sensors, in: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, August, 2008, pp. 5200–5203.
[3] H. Ghasemzadeh, P. Panuccio, S. Trovato, G. Fortino, R. Jafari, Power-aware activity monitoring using distributed wearable sensors, IEEE Trans. Hum. Mach. Syst. 44 (4) (2014) 537–544.
[4] G. Fortino, D. Parisi, V. Pirrone, G. Di Fatta, BodyCloud: a SaaS approach for community body sensor networks, Future Gener. Comput. Syst. 35 (2014) 62–79.
[5] M. Chen, P. Zhou, G. Fortino, Emotion communication system, IEEE Access 5 (2016) 326–337.
[6] G. Fortino, R. Giannantonio, R. Gravina, P. Kuryloski, R. Jafari, Enabling effective programming and flexible management of efficient body sensor network applications, IEEE Trans. Hum. Mach. Syst. 43 (1) (2013) 115–133.
[7] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, P. Havinga, Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: a survey, in: Proceedings of the 23rd International Conference on Architecture of Computing Systems, Hannover, Germany, 22–25 February, 2010, pp. 1–10.
[8] S.J. Preece, J.Y. Goulermas, L.P. Kenney, D. Howard, K. Meijer, R. Crompton, Activity identification using body-mounted sensors-a review of classification techniques, Physiol. Meas. 30 (2009) 1–33.
[9] M.M. Hassan, M.G.R. Alam, M.Z. Uddin, S. Huda, A. Almogren, G. Fortino, Human emotion recognition using deep belief network architecture, Inf. Fusion 51 (2019) 10–18.
[10] P.C. Roy, A. Bouzouane, S. Giroux, B. Bouchard, Possibilistic activity recognition in smart homes for cognitively impaired people, Appl. Artif. Intell. 25 (10) (Nov. 2011) 883–926.
[11] K. Sim, C. Phua, G. Yap, J. Biswas, M. Mokhtari, Activity recognition using correlated pattern mining for people with dementia, in: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Aug. 2011, pp. 7593–7597.
[12] M. Tolkiehn, L. Atallah, B. Lo, G.-Z. Yang, Direction sensitive fall detection using a triaxial accelerometer and a barometric pressure sensor, in: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Aug. 2011, pp. 369–372.
[13] L. Chen, J. Hoey, C.D. Nugent, D.J. Cook, Z. Yu, Sensor-based activity recognition, IEEE Trans. Syst. Man. Cybern. C 42 (2012) 790–808.
[14] W. Ugulino, D. Cardador, K. Vega, E. Velloso, R. Milidiú, H. Fuks, Wearable computing: accelerometers' data classification of body postures and movements, in: Advances in Artificial Intelligence—SBIA 2012, Springer, Berlin, Germany; Heidelberg, Germany, 2012, pp. 52–61.
[15] O.D. Lara, M.A. Labrador, A survey on human activity recognition using wearable sensors, IEEE Commun. Surv. Tutor. 15 (2013) 1192–1209.
[16] P. Dohnálek, P. Gajdoš, P. Moravec, T. Peterek, V. SnáŠel, Application and comparison of modified classifiers for human activity recognition, Prz. Elektrotech. 89 (2013) 55–58.

[17] N. Ravi, N. Dandekar, P. Mysore, M.L. Littman, Activity recognition from accelerometer data, in: Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence (IAAI), Pittsburgh, PA, USA, 9–13 July 2005, 3, AAAI Press, Pittsburgh, PA, USA, 2005, pp. 1541–1546.
[18] A. Mannini, A.M. Sabatini, Machine learning methods for classifying human physical activity from on-body accelerometers, Sensors 10 (2010) 1154–1175.
[19] I. Cohen, N. Sebe, A. Garg, L.S. Chen, T.S. Huang, "Facial expression recognition from video sequences: temporal and static modeling, Comput. Vision Image Underst. 91 (2003) 160–187.
[20] S. Kiranyaz, T. Ince, M. Gabbouj, Real-Time patient-specific ECG classification by 1-D convolutional neural networks, IEEE Trans. Biomed. Eng. 63 (March (3)) (2016) 664–675.
[21] G.E. Hinton, S. Osindero, Y.-.W. Teh, "A fast learning algorithm for deep belief nets, Neural Comput. 18 (7) (2006) 1527–1554.
[22] F. Deboeverie, S. Roegiers, G. Allebosch, P. Veelaert, W. Philips, Human gesture classification by brute-force machine learning for exergaming in physiotherapy, in: Proceedings of IEEE Conference on Computational Intelligence and Games (CIG), Santorini, 2016, pp. 1–7.
[23] A. Graves, A. Mohamed, G Hinton, "Speech recognition with deep recurrent neural networks, in: Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, IEEE, 2013, pp. 6645–6649.
[24] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9.8 (1997) 1735–1780.
[25] W. Zaremba, I. Sutskever, O. Vinyals, Recurrent neural network regularization. arXiv preprint arXiv:1409.2329. 2014.
[26] F.A. Gers, N.N. Schraudolph, J Schmidhuber, "Learning precise timing with LSTM recurrent networks, J. Mach. Learn. Res. 3 (2003) 115–143.
[27] H. Sak, A.W. Senior, F Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling, in: INTERSPEECH, 2014, pp. 338–342.
[28] R.J. Williams, J. Peng, An efficient gradient-based algorithm for on-line training of recurrent network trajectories, Neural Comput. 2.4 (1990) 490–501.
[29] G. Fortino, S. Galzarano, R. Gravina, W. Li, A framework for collaborative computing and multi-sensor data fusion in body sensor networks, Inf. Fusion 22 (2015) 50–70.
[30] G. Fortino, A. Guerrieri, F.L. Bellifemine, R. Giannantonio, SPINE2: developing BSN applications on heterogeneous sensor nodes, SIES (2009) 128–131.
[31] R. Gravina, G. Fortino, Automatic methods for the detection of accelerative cardiac defense response, IEEE Trans. Affect Comput. 7 (3) (2016) 286–298.
[32] O. Banos, R. Garcia, J.A. Holgado, M. Damas, H. Pomares, I. Rojas, A. Saez, C. Villalonga, mHealthDroid: a novel framework for agile development of mobile health applications, in: Proceedings of the 6th International Work-conference on Ambient Assisted Living an Active Ageing (IWAAL 2014), Belfast, Northern Ireland, December 2–5, 2014.
[33] O. Banos, C. Villalonga, R. Garcia, A. Saez, M. Damas, J.A. Holgado, S. Lee, H. Pomares, I. Rojas, Design, implementation and validation of a novel open framework for agile development of mobile health applications, Biomed. Eng. Online 14 (S2:S6) (2015) 1–20.
[34] W. Ugulino, D. Cardador, K. Vega, E. Velloso, R. Milidiú, H. Fuks, Wearable computing: accelerometers' data classification of body postures and movements, in: Brazilian Symposium on Artificial Intelligence, Springer, Berlin, Heidelberg, 2012, pp. 52–61.
[35] F. Palumbo, C. Gallicchio, R. Pucci, A. Micheli, Human activity recognition using multisensor data fusion based on reservoir computing, J. Ambient Intell. Smart Environ. 8 (2) (2016) 87–107.