A smartphone-enabled fall detection framework for elderly people in connected home healthcare

Mohammad Mehedi Hassan, *Senior Member*, *IEEE*, Abdu Gumaei, Gianluca Aloi, Member, IEEE, Giancarlo Fortino, *Senior Member, IEEE* , Mengchu Zhou, Fellow, IEEE

Abstract— In recent years, connected home healthcare, which involves multiple technologies such as wearable sensors, audio and video technology, and pervasive computing, has drawn attention for its ability to improve quality of life for elderly people. One of its most important services is fall detection. Falls represent a significant threat to the health and independence of adults of older than 65 years. However, commercial fall detection devices are expensive and charge a monthly fee for their use. A more cost-effective, adaptable, and reliable fall detection system is necessary to detect falls and send alarms to an appropriate authority. This paper introduces a Mobile-Enabled Fall Detection (MEFD) framework that detects falls by elderly people and allows family members and caregivers to help by immediately locating them. In the proposed framework, real time data retrieved from an accelerometer sensor on a smartphone are processed and analyzed by an online fall-detection system running on the smartphone itself. The system sends an indoor sound alert to family members through a wireless access point at home or an outdoor SMS alert to a hospital or caregiver through a mobile network base station. A hybrid deep learning model is used to detect falls. The model is trained offline using a public dataset called MobiAct. Experimental results show that the proposed framework can detect falls from real-time streaming data with high accuracy compared to state-of-the-art approaches.

Index Terms—Connected home healthcare, fall detection, mobile technology, hybrid deep learning

I. INTRODUCTION

Falls are a serious public health problem for people older than 65 years. A fall is defined as an event in which a person moves rapidly from a higher to a lower level without control. According to the World Health Organization, in most countries, almost 40% of injury-related deaths result from falls in the elderly population [1]. Thus, falls are the second-leading cause of death after traffic accidents. Falls account for 90% of hip and wrist cracks and 60% of head wounds for older individuals [2]. In addition, frequent fall incidence can negatively impact the minds of elderly people, which in turn crumbles their trust in independent living.

Older people are encouraged to stay at home as long as possible, unless they suffer from life-threatening conditions that require them to move to a hospital. This approach is natural and practical for elderly people only if an effective system is in place to remotely check their wellbeing at home.

Recently, connected home healthcare [3] is offering an effective solution to remotely monitor elderly people at home. It utilizes different technologies, such as cameras, mouthpieces, infrared sensors, pressure sensors, and wearable sensors to progressively monitor the wellbeing status of older individuals and convey this status to specialists and caregivers [4-6]. It facilitates better disease management and ensures the wellbeing of the elderly at home. Fall detection is one of the most important services in connected home healthcare. However, commercial fall detection devices have high installation and maintenance costs with limited coverage. In addition, video-based fall monitoring systems have privacy problems. Therefore, it is vital to develop a more cost-effective, automated, adaptable, and reliable fall detection system.

Elderly people commonly carry smartphones with various on-board sensors, such as accelerometers and gyroscope. It is therefore now possible to detect falls by analyzing data from these sensors [7]. Such sensors are versatile and work for all intents and purposes, anywhere. Elderly people are usually not comfortable with wearable sensors being attached to their bodies, but it would be relatively simple to persuade them to carry a smartphone.

Therefore, this paper proposes a Mobile-Enabled Fall Detection framework (MEFD) to detect falls in real time and enable calls for assistance in a connected home healthcare platform. In the proposed framework, real-time data retrieved from the accelerometer sensor on a smartphone are processed and analyzed by an online fall detection system running on the phone itself. The system sends an indoor sound alert to family members through a wireless access point at home or an outdoor SMS alert to a hospital or caregiver through a mobile network base station.

Mohammad Mehedi Hassan and Abdu Gumaei are with College of Computer and Information Sciences and Research Chair of Pervasive and Mobile Computing,, King Saud University, Riyadh, Saudi Arabia (e-mail: mmhassan@ksu.edu.sa, abdugumaei@gmail.com)

G. Aloi and G. Fortino are with University of Calabria, Italy (e-mail: gianluca.aloi@unical.it , g.fortino@unical.it).

Mengchu Zhou is with New Jersey Institute of Technology, Newark, NJ, USA (e-mail: zhou@njit.edu).

Smartphone-based fall recognition systems have been widely studied [8-10]. Prior studies have utilized threshold-based decision methods to detect falls through features collected from smartphone sensors [11]. Although threshold-based decision algorithms require less computation and are easy to program or implement, adjusting appropriate threshold values requires a tradeoff between the number of falls that remain undetected (false negatives) and the number of daily living activities classified as falls (false positives). It is difficult to obtain ideal threshold values that guarantee consistent execution for everyone.

Recently, machine learning techniques, such as support vector machines (SVMs) [8] and artificial neural networks (ANNs) [9], have been used to classify falls from daily living activities. Although the results are satisfactory, more profound results can be achieved by using deep learning models with the large amount of data that may be collected from mobile sensors. Therefore, we develop a hybrid deep learning model in the proposed MEFD framework, combining Convolutional Neural Network (CNN) [12] with a Long Short-Term Memory (LSTM) network [13]: a deep CNN-LSTM model. We use CNN to extract local, representative features from the accelerometer sensor embedded in smartphones and use LSTM to learn the dependencies among features extracted from the sensor data. We also use a dropout technique, randomly ignoring some neurons during training to prevent overfitting. Because of the variety and limited capability of smartphones, the proposed model is trained offline and then transferred to smartphones for online, real-time fall detection. We validate MEFD for real-time fall-detection use cases by analyzing the accelerometer data collected from an open, published dataset, MobiAct [14]. Our experimental results show that the proposed framework can classify falls and non-falls with better accuracy than state-of-the-art approaches.

The remainder of this article proceeds as follows. We first present the proposed fall detection framework for elderly people. Then, we describe each component in this framework. Next, we present the proposed hybrid deep learning–based online fall-detection approach. Finally, experimental results are presented, showing the performance of the proposed hybrid approach compared to existing work.

II. Proposed Framework

Figure 1 shows the main components of MEFD framework. Real-time data retrieved from the accelerometer, gyroscope, and orientation sensors of a smartphone are processed and analyzed by an online fall detection system within the smartphone itself. The phone sends an indoor sound alert to family members through a wireless access point at home or an outdoor SMS alert to a hospital or caregiver through a mobile network base station. The following subsections explain the three modules of the system in more detail.

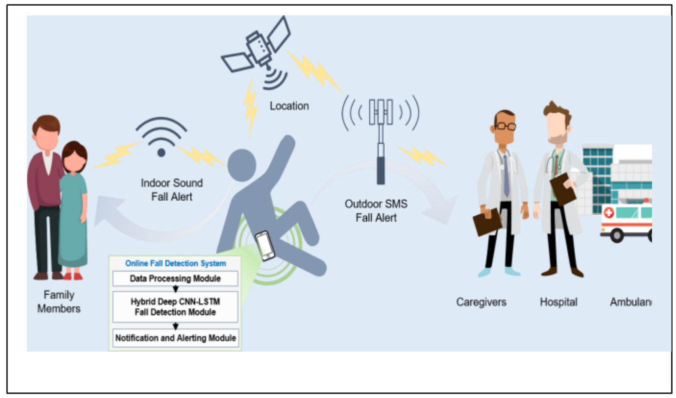

Figure 1. The proposed MEFD framework

A. *Data Processing Module*

In this module, the real-time signal of human activity sensed by the smartphone sensors is processed and normalized directly, with a window size of 200 records, which is sufficient to analyze each activity. Fifteen features are computed from each axis of the accelerometer sensor, generating 45 features across the three axes. Then, ten features are calculated from theta and magnitude from the accelerometer signal. Additionally, three features are computed for the slop value, absolute slop value, and average acceleration rate of the accelerometer signal. Hence, there are 58 values for 20 statistical features computed for each window frame. Table 1 demonstrates and summarizes these features. Table 1 shows the kurtosis—sharpness of the frequency distribution curve of the peak signal; skewness is the asymmetry of the probability distribution for the real-valued signal. Zero-crossing rate is the number of times in a given window or frame that the signal amplitude passes through zero value. The output data of this module are used as input for the next module.

B. *Hybrid Deep CNN-LSTM-based Fall Detection Module*

The hybrid deep CNN-LSTM–based fall detection module is responsible for recognizing falls based on the data coming from the data processing module. It contains a trained deep CNN-LSTM model, which is a hybrid of deep a Convolutional Neural Network and a Long Short-Term Memory Network. CNN [12] is a one-dimensional convolution network which uses the

Table 1. Statistical features computed from the accelerometer data of Mobile

No.	Feature Name	Applied on	Number of features
1	Absolute value	x-axis, y-axis, and z-axis	3
2	Mean value	x-axis, y-axis, z-axis, magnitude of the accelerometer signal, and theta of the magnitude	5
3	Absolute mean value	x-axis, y-axis, and z-axis	3

4	Median value	x-axis, y-axis, and z-axis	3
5	Absolute median value	x-axis, y-axis, and z-axis	3
6	Standard deviation value	x-axis, y-axis, z-axis, magnitude of the accelerometer signal, and theta of the magnitude	5
7	Absolute standard deviation value	x-axis, y-axis, and z-axis	3
8	Slope value	accelerometer signal	1
9	Absolute slope value	accelerometer signal	1
10	Skewness value	x-axis, y-axis, z-axis, and theta of the magnitude	4
11	Absolute skewness value	x-axis, y-axis, and z-axis	3
12	Kurtosis value	x-axis, y-axis, z-axis, and theta of the magnitude	4
13	Absolute kurtosis value	x-axis, y-axis, and z-axis	3
14	Minimum value	x-axis, y-axis, z-axis, and magnitude of the accelerometer signal	4
15	Absolute minimum value	x-axis, y-axis, and z-axis	3
16	Maximum value	x-axis, y-axis, z-axis, and magnitude of the accelerometer signal	4
17	Absolute maximum value	x-axis, y-axis, and z-axis	3
18	Zero-crossing rate value	magnitude of the accelerometer signal	1
19	Difference of minimum and maximum value	magnitude of the accelerometer signal	1
20	Average acceleration rate value	accelerometer signal	1
	Total number of features values		58

weight sharing concept for effective feature representation and extraction. Moreover, it is relatively fast to train because the number of weights is less than fully connected networks, such as multilayer perceptron or deep neural networks. LSTM [13] is a Recurrent Neural Network, comprising an input gate, an output gate, and a forget gate. It has sufficient ability to learn the dependencies that traditional RNN cannot. Furthermore, it can efficiently solve the gradient vanishing problem in the back-propagation process and can also learn input features with longer and shorter time-steps. Combining the advantages of CNN and LSTM together, we build a hybrid model comprising two one-dimensional convolution layers and one LSTM layer. Moreover, a max-pooling layer is added after the first and second convolution layers to improve system efficiency. The activation function used in both convolution layers is a Rectified Linear Unit (ReLU) to avoid the gradient vanishing problem when training the model. The output of the max pooling inputs to the LSTM, proceeding through a fully-connected layer with a softmax activation function to output the final recognition result. Figure 2 shows the architecture of the proposed model. The recognition resulting from this module is used by the next module of the proposed framework.

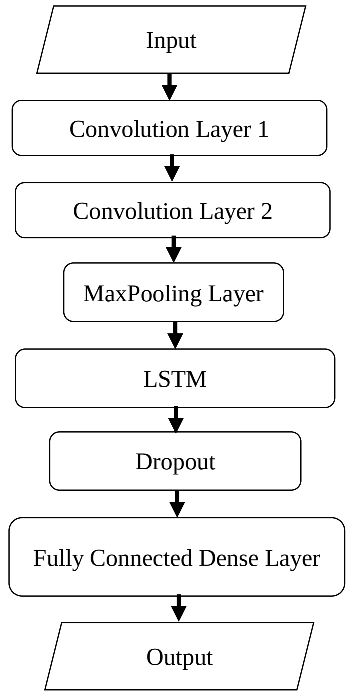

Figure 2. Architecture of the proposed CNN-LSTM model.

C. *Notification and Alerting Module*

This module takes the results of the recognition produced by the previous module and, if the fall recognition results are positive, generates one of two types of alert: (1) an indoor sound alert to family members through a wireless access point at home or (2) an outdoor SMS alert to a hospital or caregiver through a mobile network base station.

III. Experiments and Discussion

Experiments were conducted on a laptop with a i7-4510U CPU (2.0 GHz), 8 GB RAM, and Windows 10 operating system (x64), implemented in the Python programming language. Three main measures are used to evaluate the performance of the proposed framework, the F1 score, and the recognition time:

$$Accuracy = (TP + TN)/(TP + FP + TN + FN) \quad (1)$$
$$Precision = TP / (TP + FP) \quad (2)$$
$$Recall = TP / (TP + FN) \quad (3)$$

$$\textbf{\textit{F1-score}} = \frac{2\times(Recall\times Precision)}{(Recall+Precision)} \quad (4)$$

where TP and FP are true and false positive rates, respectively, and TN and FN are true and false negative rates, respectively.

Table 2: Falls and daily living activities in the new version of MobiAct dataset

No.	Label	Activity	Trials	Duration	Description
1	FOL	Forward-lying	3	10s	Fall Forward from standing, use of hands to dampen fall.
2	FKL	Front-knees-lying	3	10s	Fall forward from standing, first impact on knees.
3	BSC	Back-sitting-chair	3	10s	Fall backward while trying to sit on a chair.
4	SDL	Sideward-lying	3	10s	Fall sideward from standing, bending legs.
5	STD	Standing	1	5min	Standing with subtle movements.
6	WAL	Walking	1	5min	Normal walking.
7	JOG	Jogging	3	30s	Jogging.
8	JUM	Jumping	3	30s	Continuous jumping.
9	STU	Stairs up	6	10s	Stairs up (10 stairs).
10	STN	Stairs down	6	10s	Stairs down (10 stairs).
11	SCH	Stand to sit (sit on chair)	6	6s	Transition from standing to sitting.
12	SIT	Sitting on chair	1	1min	Sitting on a chair with subtle movements.
13	CHU	Sit to stand (chair up)	6	6s	Transition from sitting to standing.
14	CSI	Car-step in	6	6s	Step in a car.
15	CSO	Car-step out	6	6s	Step out a car.
16	LYI	Lying	12	-	Activity taken from the lying period after a fall.

Simply put, accuracy is the ratio of correctly recognized samples to all samples in the testing set. Precision measures the ratio of correctly recognized true positive samples to the total number of true positive and false positive samples. Recall measures the ratio of correctly recognized true positive samples to the total number of recognized true positive and false negative samples in the testing set. Finally, the F1 score represents the weighted average of recall precision. In the following subsections, we explain all materials included in the experiments, such as the dataset used, and describe our results, including comparisons and discussion.

A. MobiAct Dataset

We used the new version of the MobiAct dataset in our experiment as a case study to evaluate the proposed MEFD framework. This public dataset was collected by the Biomedical Informatics and Health Laboratory for educational purposes and non-commercial research [14]. Data in the set were collected by placing a Samsung Galaxy S3 smartphone in participants' pockets and recording the raw data from the smartphone's gyroscope, accelerometer, and orientation sensors. The raw data of orientation sensors are generated based on software that uses data from the geomagnetic and accelerometer sensors. The SENSOR-DELAY-FASTEST parameter enabled collecting data at the highest possible sampling rate. The new MobiAct dataset contains four types of falls made by 67 subjects, eleven types of daily living activities recorded by 59 subjects, and a lying activity ("LYI") taken from a lying period after a fall recorded by 67 subjects. The experimental subjects included both men and women of ages ranging between 20 and 47 years (26 on average), height from 160 to 189 cm (175 on average), and weight from 50 to 120 kg (76 on average). Table 2 presents information about falls and daily living activities.

In our case study, we conducted two experiments on a subset of the raw data containing four different types of falls and two types of daily activities (standing and lying) that occur during the inactivity period after a fall. The first experiment to recognize a fall and the two types of daily living activities was applied to 7,671 samples, which were randomly divided into 5,370 samples for training and 2,301 samples for testing. The second experiment to recognize falling from non-falling daily living activities (Non-FLI) was performed on 4,000 samples, which were randomly divided into 2,800 samples for training and 1,200 samples for testing.

B. Hyper-parameter Settings

When building deep learning models, intensive effort is required to select the hyper-parameter settings that attain the best results. There are three common techniques to tune a model's parameters: manually, using experience in the deep learning field; randomly, by taking random values; and grid search, searching in an exhaustive set of values. In this work, we used grid search, starting with an initial range of coarse hyper-parameters that is narrowed based on our experience in the field. After applying this technique, the learning rate was set to 0.001, the number of epochs was set to 125, the output size of the LSTM was set to 30, the number of convolution filters was set to 64, and the kernel size was set to 3.

C. Experimental Results

This subsection presents the experimental results to validate the performance of our online fall detection system. Experiments 1 and 2 were performed by training and testing the CNN-LSTM model on the respective training and testing sets. Both experiments were evaluated as shown in Table 3.

Table 3. Evaluation Results for experiments 1 and 2 (F=Fall, N=Non-Fall, S=Standing, L=Lying, WA= Weighted average, TR = Time to recognize a sample).

Measures	**Experiment 1**				**Experiment 2**		
	Classes			**WA**	**Classes**		**WA**
	F	S	L		F	N	
Precision	0.90	0.98	0.95	0.97	0.98	0.96	0.97
Recall	0.92	0.99	0.92	0.97	0.95	0.98	0.97
F1-score	0.91	0.99	0.93	0.97	0.97	0.97	0.97
Accuracy	97.00%				96.75%		
TR	**0.394 (millisecond)**				**0.0414 (millisecond)**		

The proposed model achieved accuracy of 97% and 96.75% for experiments 1 and 2, respectively as depicted in Table 3. Moreover, the model attained a high weighted average F1 score (97%) for both experiments. The high precision and recall scores for falls and non-falls (98%) in experiment 2 prove that the system has a high recognition rate, recognizing falls as falls and non-falls as non-falls. In other words, if the precision is not high, family members or caregivers would be alarmed falsely. During the experiments, we found that the model takes very little time to recognize one sample, as shown in Table 3. Fig. 3(a) shows that the gap between training loss and testing loss is very small, which means that the model works properly, without any overfitting. The loss value indicates how well the model performs after each epoch of optimization, while the accuracy represents how accurately the model recognizes the true test set after setting its parameters.

To validate the model's performance, we computed the Receiver Operating Characteristic (ROC) curve, shown in Fig. 3(b) for experiment 1 and in Fig. 3(c) for experiment 2. Clearly, the *y* coordinate in Figs. 3(b)-(c) represents the true positive rate (TPR), and the *x* coordinate in Figs. 3(b)-(c) represents the false positive rate (FPR) and used to compute the area under curve (AUC) value for fall detection.

We compare our result with [15], which presented work related to the proposed model, proposing a fall detection model using LSTM. The overall accuracy for recognizing falling, standing, and lying were 79.7% and 95.22% for recognizing falls and non-falls, which is lower than the accuracy achieved by the proposed CNN-LSTM model. This means that the proposed model outperforms state-of-the-art methods for detecting falls in the new version of the MobiAct dataset.

IV. Conclusions

Accurate and real-time fall recognition is important for connected home healthcare for elderly adults. This paper presented a smartphone-enabled fall-recognition framework that recognizes falls by automatically and continuously observing data generated by the smartphone's sensors. We developed a hybrid deep-learning model for online fall recognition in real-time, trained offline using a public dataset called MobiAct. Test results demonstrate that the proposed system can identify falls and non-falls with higher precision than can state-of-the-art approaches. In the future, we intend to combine other sensor data, such as heartbeat information before and after falling, and GPS data to track people's activities.

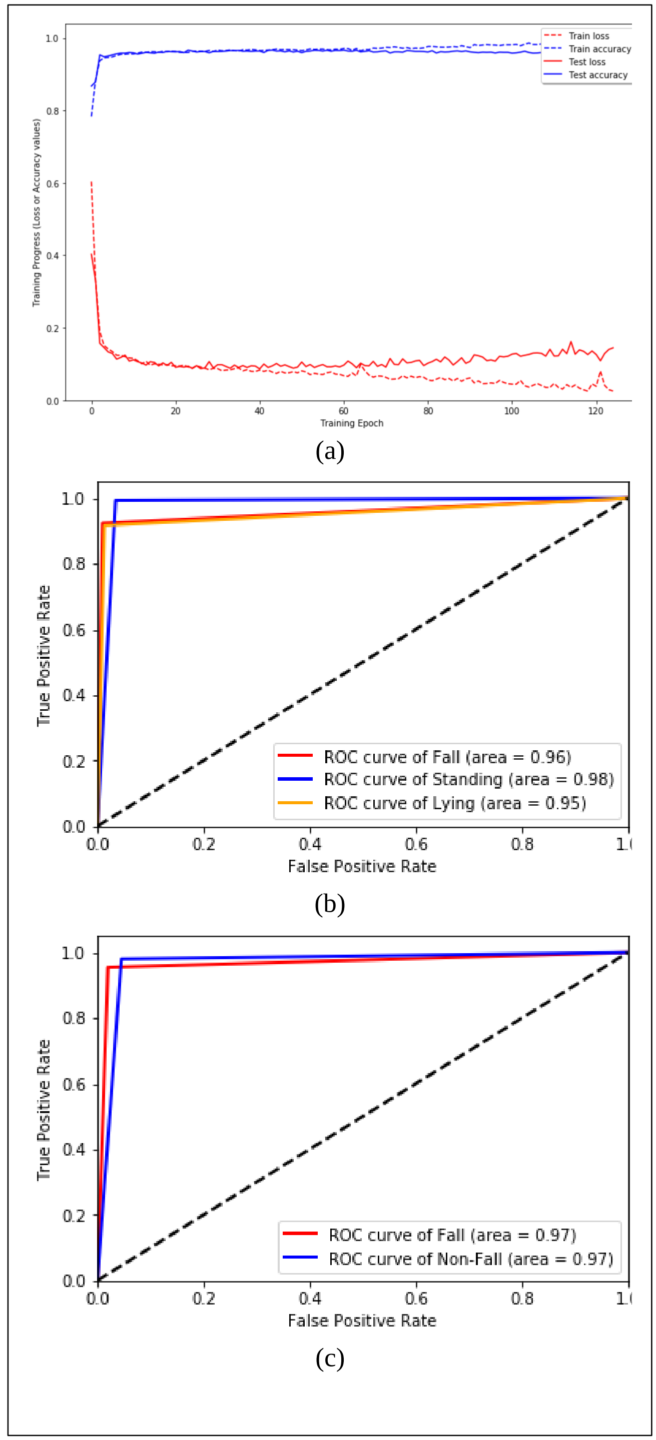

Figure 3. (a) Progress of loss and accuracy for training and testing over 125 epochs, (b) ROC curve of experiment 1, and (b) ROC curve of experiment 2.

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] C. Todd, D. Skelton, and World Health Organization, What are the Main Risk Factors for Falls Amongst Older People and What are the Most Effective Interventions to Prevent These Falls?World Health Organization, Geneva, Switzerland, 2004

[2] Fortino G, Gravina R. Fall-MobileGuard: A smart real-time fall detection system. In Proceedings of the 10th EAI International Conference on Body Area Networks, Sep 28 (pp. 44-50). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2015

[3] Kvedar, J., Coye, M. J., & Everett, W. (2014). Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth. Health Affairs, 33(2), 194-199.

[4] Wan, J., Zou, C., Ullah, S., Lai, C. F., Zhou, M., & Wang, X. (2013). Cloud-enabled wireless body area networks for pervasive healthcare. IEEE Network, 27(5), 56-61.

[5] M. S. Hossain, and G. Muhammad, “Emotion Recognition Using Deep Learning Approach from Audio-Visual Emotional Big Data,” Information Fusion, vol. 49, no. 2019, pp.69-78. September 2019.

[6] M. S. Hossain, G. Muhammad, and A. Alamri, “Smart Healthcare Monitoring: A Voice Pathology Detection Paradigm in Smarter Cities,” ACM/Springer Multimedia Systems, 2017, DOI: 10.1007/s0053

[7] J. S. Lee and H. H. Tseng, “A preliminary study on using smartphones to detect falling accidents,” IARIA International Conference onSystems and Networks Communications (ICSNC), Athens, Greece, October 2017, pp. 42-47

[8] M. V. Albert, K. Kording, M. Herrmann, andA. Jayaraman,“Fall classification by machine learning using mobile phones,”PloS One,vol.7,no.5, p. e36556, 2012

[9] H. Kerdegari, S. Mokaram, K. Samsudin, and A. R. Ramli, “A pervasive neural network based fall detection system on smart phone,” J. Ambient Intell. Smart Environ., vol. 7, no. 2, pp. 221–230, 2015.

[10] Hassan, M. M., Uddin, M. Z., Mohamed, A., & Almogren, A. A robust human activity recognition system using smartphone sensors and deep learning. Future Generation Computer Systems, 81, 307-313, 2018

[11] D. Yu and J. Li. Recent progress in deep learning based acoustic models. IEEE/CAA Journal of Automatica Sinica, 4(3), 396-409, 2017

[12] Hsieh, Y. Z., & Jeng, Y. L. Development of Home Intelligent Fall Detection IoT System Based on Feedback Optical Flow Convolutional Neural Network. IEEE Access, 6, 6048-6057, 2018

[13] Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R. and Schmidhuber, J., 2017. LSTM: A search space odyssey. IEEE transactions on neural networks and learning systems, 28(10), pp.2222-2232.

[14] G. Vavoulas, C. Chatzaki, T. Malliotakis, M. Pediaditis, and M. Tsiknakis, The MobiAct Dataset: Recognition of Activities of Daily Living using Smartphones," International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE), no. Ict4awe, pp. 143{151, 2016.

[15] Dharmitha Ajerla, “A Real-time Patient Monitoring Framework for Fall Detection”, Masters Dissertation, Queen's University Kingston, Ontario, Canada, August 2018.