Convolutional Neural Networks as Means to Identify Apposite Sensor Combination for Human Activity Recognition

Muhammad U. S. Khan
Department of Computer Science
COMSATS University Islamabad,
Abbottabad Campus, Pakistan
E-mail: ushahid@ciit.net.pk

Assad Abbas
Department of Computer Science
COMSATS University Islamabad,
Islamabad Campus, Pakistan
E-mail: assadabbas@comsats.edu.pk

Mazhar Ali
Department of Computer Science
COMSATS University Islamabad,
Abbottabad Campus, Pakistan
E-mail: mazhar@ciit.net.pk

Muhammad Jawad
Department of Electrical Engineering
COMSATS University Islamabad,
Lahore Campus, Pakistan
E-mail: mjawad@ciitlahore.edu.pk

Samee U. Khan
Department of Electrical and Computer Engineering
North Dakota State University
Fargo, ND, USA
E-mail: samee.khan@ndsu.edu

***Abstract*— Accurate classification of human activities is significant particularly for remote monitoring of patients requiring orthopedic treatment and therapies in case of various injuries. Previously, we presented a Local Energy-based Shape Histogram (LESH) based approach that considers the energy expenditure for various activities and differentiates among the activities on the basis of energy level. Although the approach effectively recognized various activities, the recognition accuracy for the activities, such as walking forward, walk forward left circle, and walk forward right circle was significantly low. In this paper, we present a Convolutional Neural Network (CNN) based approach for recognizing human activities with sufficiently high accuracy. The experiments are conducted on the Wearable Action Recognition Database (WARD) dataset. Experimental results demonstrate that the CNN based approach (in general) not only achieves high recognition accuracy for all of the activities but also performed extremely well for the activities requiring frequent inter-posture transitions. Another important contribution of this research is that the CNN based approach is capable of finding a single combination of sensors to identify thirteen different activities with sufficiently high accuracy.**

Keywords*— *activity recognition, convolutional neural networks, sensor data, wearable devices

I. INTRODUCTION

With the widespread availability of low-cost sensing devices in the recent years, the research on the human activity recognition using wearable devices has significantly increased [1]. The human activity recognition through wearable sensors is beneficial for various types of daily life activities, such as jogging, personal fitness, and elderly care. Moreover, the Body Area Networks (BANs) and Internet-of-Things (IoT) have also revolutionized the healthcare monitoring as the ubiquity of healthcare services has increased manifold [2]. Several types of wearable devices, such as sensors and smart jewelry are being used on the human body to measure several vital signs, for example blood pressure, heart rate, oxygen saturation, and energy expenditure due to their physical activities [3].

Particularly, in the recent past the research on human activity recognition has attained significant attention due to the development of applications meant for patient monitoring and fall detection in elderly people. The aforesaid applications and several others utilize wearable sensors for monitoring activities, such as walking, standing, going upstairs, and going downstairs. The accurate recognition of human activities is highly dependent on the location/position of sensors on human body. Therefore, it is important to place the sensors at the appropriate positions on the human body to correctly identify the activities to support monitoring.

Keeping in view how a human activity is performed in real-time, each activity is a set of many elementary and continuous movements. Conventionally, any human activity lasts for only few seconds (or less) and in these seconds a combination of basic activities can be involved. In light of sensor data, these continuous basic movements correspond to the flat signals. However, the transition between these basic movements is presented by a significant change in the signal value. Therefore, only an effective and complex feature extraction technique can be applied to such applications that have the capability to capture any salience feature in the dataset. In our previous work in [4], we presented a cloud based framework to investigate the effects of sensor location on the accuracy of human activity recognition. The aforementioned framework utilizes a Local Energy-based Shape Histogram (LESH) approach for transformation of human activities data into a feature space that subsequently is used by machine learning algorithms for classification of activities. Several classifiers were used to recognize the activities and in fact none of the classifiers performed very well particularly for the activities, such as walk forward, walk forward left circle, and walk forward right circle. In this paper, we improve the accuracy of human activity recognition by employing a Convolutional Neural Network (CNN) based approach for feature extraction. An important attribute of the CNNs is that they do not require excessive domain knowledge to achieve reasonable performance and area also capable of achieving the maximum performance by fine tuning the architecture.

In the previous LESH based approach [4], there was not a single combination of sensors that could identify all of the activities with high accuracy. Therefore, an important question was whether the CNN based model has the capability to identify all of the activities with a single combination of sensors. To evaluate the performance of the CNN based approach for human activity recognition, we conducted experiments on Wearable Action Recognition Database (WARD) [5] dataset comprising of the data from 20 subjects for thirteen activities using five sensors on the human body. For the details of the location of sensors on human body, interested readers are encouraged to refer to [4]. With the presented CNN based approach, the recognition accuracy for all of the activities in general and particularly for the three activities, such as walk forward, walk forward left circle, and walk forward right circle that in fact was significantly low using the LESH based approach is significantly improved. Moreover, with the CNN based approach all of the thirteen activities were accurately recognized with a single combination of sensors, which was not possible with the LESH based approach.

- We successfully employed the CNN for human activity recognition with high accuracy. We observe that activities, for example Rest at lying, walk left circle, walk right circle, turn right, go upstairs, jog, and jump can be identified with 100% precision using the CNN.
- The performance of the proposed CNN based approach in terms of recognition accuracy was observed significantly better than several classifiers, such as the SLR, Naïve Bayes, and the SMO. On an average, the CNN has shown 26% higher F-measure score for recognition of thirteen different human activities than different classifier used with LESH based approach in [4].
- The activities like walk forward, walk circle left, and walk right circle can now be identified with 97%, 99%, and 99% F-measure score, respectively using the CNN despite very similar nature of the activities.
- In contrast to the previous LESH based approach, the CNN based approach is also capable of recognizing all of the activities with a single combination of sensors. The combination of all five sensors in a body area network has shown 3.4 % higher F-measure score on an average than the combination of 3 sensors and individual sensors.

The rest of the paper is organized as follows. Section II discusses the related work. The presented CNN approach for human activity recognition is presented in Section III. Section IV presents experimental results and discussion on tuning of the hyperparameters whereas Section V finally concludes the paper and highlights the direction for future work.

II. Related Work

This section overviews the works related to human activity recognition. The previous works have focused only on the recognition of human activities. In addition to the recognition of human activities, our presented approach is capable of identifying all activities with a single combination of sensors (using all five sensors).

A deep learning methodology for activity recognition through on-Node sensors is proposed in [6]. In the proposed methodology, the authors have utilized both the deep and shallow features obtained though the sensors. Selection of both of the aforesaid features reduces the number of hidden layers in the deep learning process. Reduced number of layers in turn reduces the training time. Raw data is collected from the sensors in the form of the signals. The n samples are sent to the process (dealing with deep features) with segment length of 4—10 seconds. Authors in [6] utilized segments instead of single data point to improve the classification accuracy. Length of the segment depends upon the kind of application and sensor used. Deep features are extracted using spectrogram. Spectrogram is utilized due to the fact that spectrogram captures more interpretable features. For predefined shallow features, authors have utilized a vector of six manually selected features. After selection of deep and shallow features, they are passed through connected and soft-max layer. Unified deep neural networks are utilized to train both of the aforementioned features together. The training process calculates the weights of the hidden layers by backward propagation. To improve the weight assignment, concepts of weight decay, momentum, and dropout also utilized. The proposed methodology is evaluated on five different publicly available datasets. One of the datasets is developed by the authors themselves during the process and is made public. Precision and recall results obtained after experimentation show significant improvement after employing deep learning methodologies.

Zeng et al. [1] proposed human activity recognition model based on Convolutional Neural Networks (CNN). The model is applied on the data obtained through mobile sensors. Important feature of the proposed methodology is the fact that it captures scale invariance and local dependencies of the input signal. Capturing of local dependencies and scale invariance results in effective determination of features in presence of varied postures of the same activity. Moreover, no domain knowledge is required for features extraction as the process is carried out automatically by the CNN layers. The aforesaid is performed through local connectivity of the adjacent layers. The max-pooling layer computes weights and divides the features into multiple partitions. Weights are regularized by using weight decay, momentum, and drop out processes. Afterwards, local features are mapped onto global sets. Training process is carried out through forward and backward propagations. Unlike traditional CNNs, the proposed methodology contains only one pair of convolutional and max-pooling layers along with two fully connected neural networks. Three publicly available data sets (Skoda, Opportunity, and Actitracker) are used to evaluate the proposed methodology. Seven different activities were evaluated. Results show improved precision and recall than the traditional methodologies. A methodology for collective recognition of human activity and sensor location is proposed in [7]. The base of the aforesaid work is sparse signal theory. Authors in [7] employed traditional methodologies of manual feature extraction and Bayesian sparse signal classification. Same authors have also jointly recognized the human activity and sensor location in [8], where the authors have reconstructed the sparse signal by using compressed sensing theory. The reconstructed signals were then used to recognize the activity and sensor location on human body. Though the presented methodologies improved the results from the traditional technologies in few cases, the limitation of robust automatic feature extraction existed.

Authors in Ref. [9] made use of time series data to recognize human activities. The authors claim to recognize

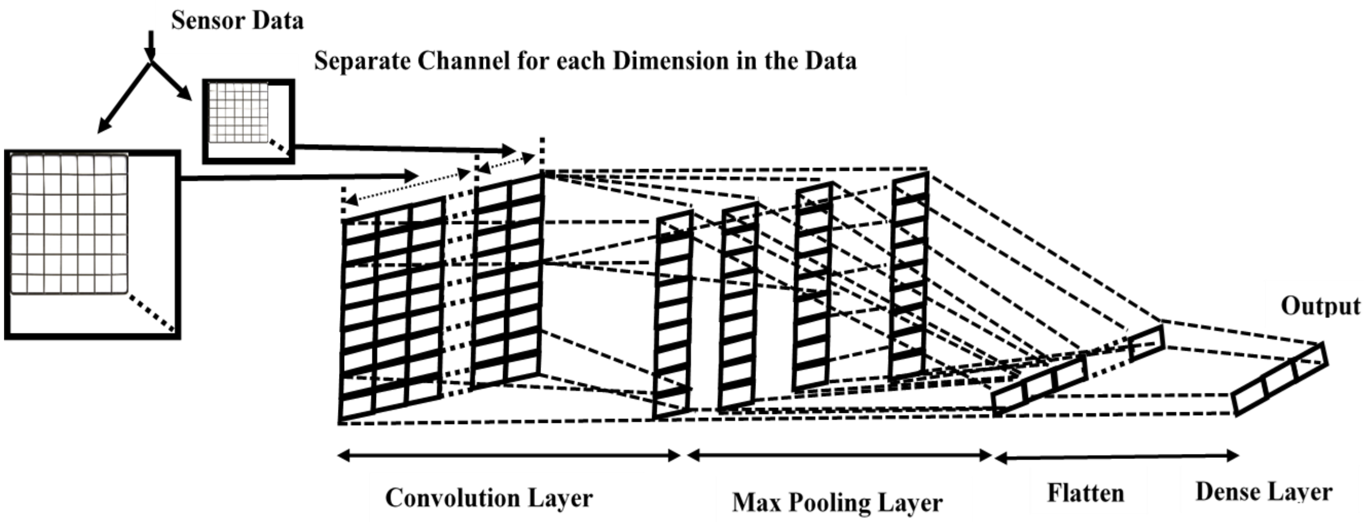

Fig. 1: The CNN architecture for human activity recognition from sensor data

both atomic and complex activities. The proposed methodology converts the atomic and complex activities into shapelets. Later on the shapelets are compared with the recorded signal to classify the activities. However, the proposed methodology is not able to capture variations in the signal obtained from the same activity. A similar approach based on shapelets was also proposed in [10].

III. THE CNN BASED APPROACH FOR HUMAN ACTIVITY RECOGNITION

The convolution neural network (CNN) is a class of deep feed-forward neural networks. The CNN architecture consists of multilayer perceptron with variations that require minimal pre-processing. Therefore, the CNN is also known as Space Invariant Artificial Neural Networks (SIANN). The CNN decomposes the big and complex problems into small problems. Therefore, it gathers variance in the signals for the same activity but on different subjects. Our CNN model is built upon that of presented in [11] that was originally built for sentence classification and we adapt the model for the WARD dataset for human activity recognition. The conventional design of the CNN has an input-output layer and multiple hidden layers in between. The hidden layers consist of (a) convolution layer, (b) pooling layer, (c) flatten layer, and (d) dense layer. Fig. 1 presents the CNN architecture for the sensor based human activity recognition whereas Fig. 2 shows the position of sensors on human body.

The CNN input is the sensor data of length $m \times n$, where $m = 5$ represents the dimensions of the sensor data from 3-axis accelerometer and 2-axis gyroscope and n represents the length of trail sequence (each dimension) of the sensor data. Let $s_i \epsilon \mathbb{R}^m$ corresponds to the i-th sensor attached on the human body with five dimensions. The sensor data is represented as:

$$s_i=[s_1, s_2, s_3\ s_4 \dots \dots s_h] , \qquad (1)$$

where $s_i \epsilon \mathbb{R}^m$ corresponds to the i-th sensor attached on the human body and h is the number of values per window.

In the first step, the matrix s_i is fed into the convolution layer for extracting high level features. In convolution operation, a filter w of window size h is applied on the matrix s_i. In response, a feature c_i is generated from the window by using Eq. 2.

$$c_i = \mathfrak{f}\left(w . s_{h,5} + b\right), \qquad (2)$$

where b is the bias term associated with each window h and $\mathfrak{f}$ is a non-linear function. The filter w is applied on each possible window made on the length of each trail sequence to produce a feature map as:

$$c = [c_1, c_2, \dots , c_k] , \qquad (3)$$

where, k is the total number of features extracted from the total windows made.

In the second step, we apply the max-over-time pooling operation on the feature map given in Eq. 3 [12]. The purpose of the pooling layer is to further abstract the features that were extracted from the convolution layer. We, then take the maximum value $c_{max} = max\{c\}$ as the dominating/ important feature among all the abstract features for a particular filter.

A. *Classification*

In the third step, the max pooling score of each applied filter is flattened to get a feature of a single dimension. The flatten feature is represented as:

$$p = [p_1, p_2, \dots , p_j], \qquad (4)$$

where, j is the total number of applied filters and p_j is the maximum pooling score of the jth filter. The dense layer performs the classification on the features extracted in the convolutional layer and down sampled by the pooling layer. The last layer is a dense layer. The dense layer represents the matrix multiplication. The values in the matrix are trainable parameters and output is the m dimension vector where m represents the number of output classes.

We have used forward propagation method of the dense layer in our model. The forward propagation model comprises of three inputs: (i) input features, (ii) weights, (iii) and bias. The output of the forward propagation consists of only one vector. Finally, the output y is computed as:

$$y = (w \circ p) + b , \qquad (5)$$

where, $\circ$ is the element by element multiplication operator.

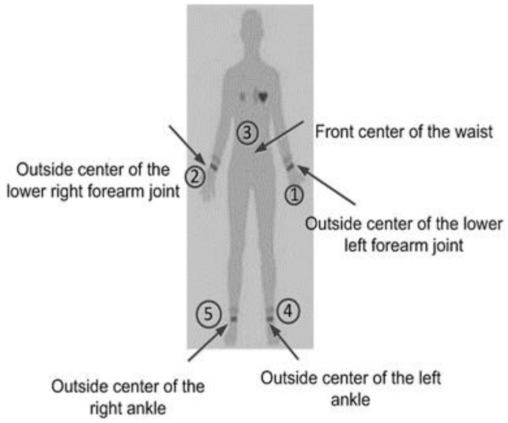

Fig. 2. Position of on body sensors [4]

IV. Experimental Results

We carried out experiments on Amazon EC2 [14] to meet the requirements of high computational resources for the implementation of the CNN. We used the WARD dataset [5] to evaluate the human activity recognition using convolutional neural network. The dataset contains the data collected from sensors located at five different locations on 20 different people. The data was recorded for thirteen different activities presented performed by the people and are presented in Table 1. For convolutional neural network, we have used Keras and Tensorflow libraries of Python programming language. The performance is measured using Precision, Recall, and F-measure [13]. We used the technique of splitting the dataset into three portions, training, development, and test datasets. We have used 60% dataset as training data, 20% as development data, and 20% as test data.

Table 1. Activities in the WARD Dataset

Activity	Action Class	Activity	Action Class
1	Rest at Standing	8	Turn right
2	Rest at Sitting	9	Go upstairs
3	Rest at Lying	10	Go downstairs
4	Walk Forward	11	Jog
5	Walk left circle	12	Jump
6	Walk right circle	13	Push Wheelchair
7	Turn left		

Table 2. Hyperparameter values after tuning

	Parameter	Value
Conv2D		
	# of neurons in a layer	1000
	Kernal Size	(1 60)
	Activation	Relu
MaxPool2D	Pool Size	(1 20)
Dense	Activation	Sigmoid
Compile	Optimizer	Adam
Fit	Epochs	10
	Batch Size	30
Input data	Window Size	90

A. *Hyperparameters' Tuning*

The convolutional neural networks are notoriously difficult to configure because there are a lot of parameters that must be set for the efficient working of the models. We have used the grid search capability from the scikit-learn library of Python programming language to find the best values for the parameters being used in the model. Moreover, the input data reshaped to fixed size windows. The reshaped data is taken as input to the convolutional neural network model. A separate channel is used for each dimension. The data from each sensor consists of five dimensions. Therefore, when data from all five sensors is used, 25 channels are required to feed data to CNN model. Table 2 presents the values of the parameters that we obtained after hyperparameters' tuning of our CNN model. The hypertuning of the parameters was performed on the development dataset. Fig. 2 shows shape of the output layers of our model. The model used 1,514,013 parameters in the training of the model.

Table 3: Output shapes and no. of parameters

Layer (type)	Output Shape	Param #
Conv2d_1 (Conv2D)	(None, 1, 31, 1000)	1501000
Max_pooling2d_1 (MaxPooling2)	(None, 1, 1, 1000)	0
flatten_1 (Flatten)	(None, 1000)	0
dense_1(Dense)	(None, 13)	13013
Total parameters: 1,514,013 Trainable params: 1,514,013 Non-trainable params: 0		

We evaluated the recognition accuracy of thirteen different human activities listed in Table 1 using the data from: (i) one sensor, (ii) different combinations of three sensors, and (iii) all five sensors at the same time. Due to space limitations, the F-measure scores for activity recognition using the CNN methodology are presented in Table 4. The highest values in the results are presented with bold-font in the tables. Results demonstrate that the convolutional neural network is extremely accurate in recognizing the human activities. Table 2 shows the output shape and the number of parameters in different layers. The precision, recall, and F-measure scores have remained above 90% in most of the cases. Seven out of thirteen activities can be recognized with 100% precision with different combination of sensors. These activities include activity number 3 (Rest at Lying), activity number 5 (Walk left circle), activity number 6 (walk right circle), activity number 8 (Turn right), activity number 9 (Go upstairs), activity number 11 (Jog), and activity number 12 Remaining 5 activities can also be recognized with 99% precision with different combination of sensors. Looking only at the highest values, it appears that in recognition for each activity, we will require a separate combination of sensors. For example, turning right that is the activity number 8 in Table I can be recognized with 100% precision using the data from the combination of Sensor 1 (located at lower left forearm), Sensor 2 (located at lower right forearm), and Sensor 5 (located at right ankle). Similarly, Walk right circle that is activity number 6 can be recognized with 100% precision using the data from sensor at location 3 (waist). However, going through all the tables, it can be observed that difference between highest and other values is too small. Therefore, we took the average of each combination over all of the activities to observe that which combination acquires the highest average. The averaged results of precision, recall, and F-measure are presented in Fig. 3, Fig. 4, and Fig. 5, respectively. The results show that the combination of all five

TABLE 4. F-MEASURE OF HUMAN ACTIVITY RECOGNITION USING THE CNN BASED APPROACH

	F-measure															
Act	S1	S2	S3	S4	S5	S123	S124	S125	S134	S135	S145	S234	S235	S245	S345	All 5
1	0.92	0.89	0.82	0.8	0.84	0.96	0.97	**0.98**	0.95	0.94	0.98	0.96	**0.98**	0.91	0.94	**0.98**
2	0.81	0.78	0.79	0.65	0.75	0.97	0.98	**0.99**	0.97	0.98	0.98	**0.99**	0.98	0.97	0.96	**0.99**
3	0.82	0.86	0.99	0.92	0.95	0.99	0.99	0.99	0.99	0.99	0.99	**1**	**1**	0.99	0.98	0.99
4	0.84	0.85	0.9	0.83	0.89	0.9	0.96	0.94	0.95	0.96	**0.97**	0.95	0.96	0.94	0.94	0.96
5	0.93	0.94	0.98	0.98	0.98	0.94	0.97	0.97	0.96	0.97	0.96	0.96	0.98	0.96	**0.99**	0.97
6	0.9	0.95	0.99	0.97	0.98	0.98	**0.99**	0.98	0.98	0.98	0.98	**0.99**	**0.99**	0.98	**0.99**	**0.99**
7	0.96	0.97	0.98	**0.99**	0.97	0.98	**0.99**	0.97	0.98	0.98	0.97	**0.99**	0.97	0.98	0.97	0.98
8	0.94	0.95	**0.99**	0.98	**0.99**	**0.99**	0.97	**0.99**	**0.99**	**0.99**	**0.99**	**0.99**	0.97	0.96	**0.99**	**0.99**
9	0.69	0.85	0.93	0.97	0.96	0.88	0.95	0.92	0.97	0.96	0.97	0.96	0.97	0.97	**0.98**	0.97
10	0.84	0.93	0.97	0.96	0.95	0.9	0.96	0.95	0.96	0.92	0.96	**0.98**	0.96	0.94	0.97	0.97
11	0.87	0.94	0.95	0.93	0.96	0.95	0.93	0.91	0.96	**0.99**	0.95	0.95	0.97	0.95	0.97	0.97
12	0.94	0.96	0.98	0.95	0.96	0.98	0.97	0.98	0.98	**0.99**	0.97	0.97	**0.99**	0.98	0.98	**0.99**
13	0.91	0.88	0.89	0.82	0.87	0.97	0.98	0.98	0.98	0.98	0.97	0.98	**0.99**	0.98	0.94	**0.99**

TABLE 5. COMPARISON OF F-MEASURE OF CNN AND LESH METHODOLOGY

	[4]			
Activity #	Classifier	Combination of Sensors	F-Score	CNN F-score
1	Naïve Bayes	1,2,3	0.95	0.98
2	Naïve Bayes	1,2,5	0.78	0.99
3	SMO	2,3,4	0.92	1
4	SMO	All 5 Sensors	0.67	0.97
5	SLR	1,3,5	0.68	0.99
6	SLR	1,4,5	0.6	0.99
7	SLR	3,4,5	0.71	0.99
8	SLR	3,4,5	0.74	0.99
9	SLR	3,4,5	0.92	0.98
10	Naïve Bayes	1,3,4	0.77	0.98
11	SMO	1,3,4	0.81	0.99
12	SMO	1,2,4	0.83	0.99
13	SMO	1,2,3	0.92	0.99

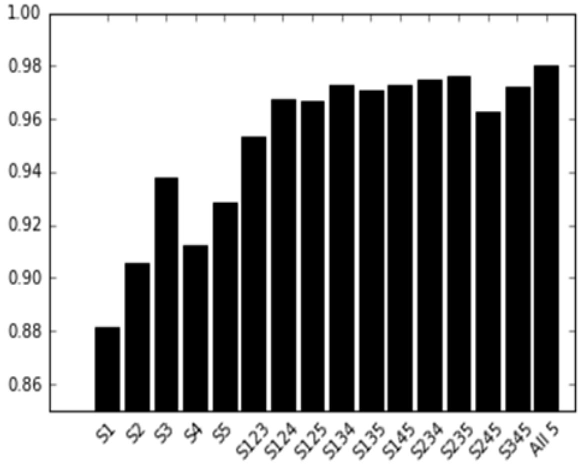

Fig. 3. Average precision for combination of sensors

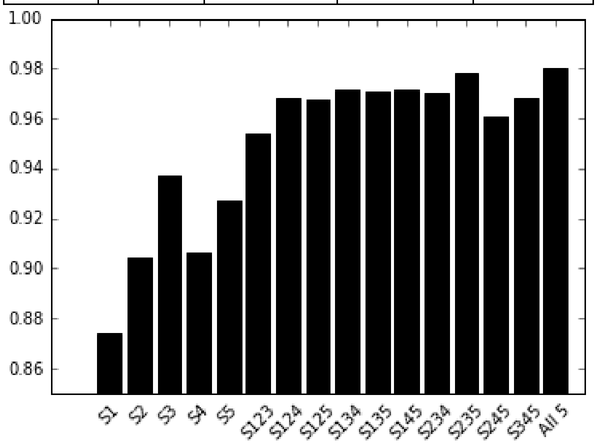

Fig. 4. Average recall for combination of sensors

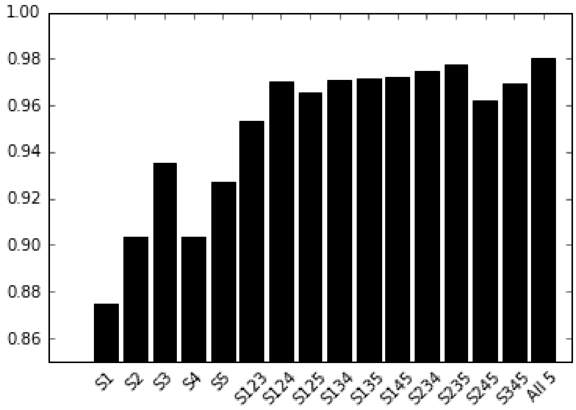

Fig. 5. Average F-measure for combination of sensors

sensors exhibits better results than any other combination. On an average, the combination of all five sensors obtained 98% scores in precision, recall, and F-measure. Another important observation that we have made is that the results from combination of Sensor 2, Sensor 3, and Sensor 5 are slightly behind the combination of all five sensors. The two combinations, one of Sensor 2, Sensor 4, and Sensor 5 and the other of Sensor 3, Sensor 4, and Sensor 5 performed worst in combinations of more than one sensor. Moreover, although the individual sensors showed high scores but they did not perform significantly as better as the combination of sensors. Therefore, we can conclude that the CNN model when applied on the data from combination of sensors gives better results than data from any single sensor. We compared the results of methodology with the best results obtained in our previous study [4] that utilizes the LESH feature descriptor and different classifiers to recognize the human activities. In our previous study, among different classifiers, Simple Linear Regression (SLR), Naïve Bayes, and Sequential Minimal Optimization (SMO) classifiers had shown best performance in activity recognition with different combination of sensors. Table 5 shows the comparison of our model using all five sensors with the best results of our previous study against F-measure scores. The results have shown that the CNN has outclassed the previous study in recognition of all the thirteen activities. Overall, the combination of all five sensors has shown 3.4 % higher F-measure score than the other combinations. The key observation we have made in our comparison is the significant gain in F-measure for the activity 4 (Walk forward), activity 5 (walk left circle), and activity 6 (walk right circle) using the CNN. These activities are problematic in recognition in all other recognition methods due to their very similar nature. For example, the F-measure score for walking forward is improved from 60% to 99% using the CNN model. On taking the average of the percentage increase using the CNN over the classifier used with LESH methodology, we can see a 26% higher F-measure score than all other classifiers used with LESH methodology. Therefore, we can conclude that the CNN methodology outperformed the LESH methodology with a significant margin. Therefore, it can be concluded that we can use single combination of sensors (all five sensors at a time) with CNN for recognition of thirteen different human activities with very high accuracy and 3.4 % higher F-measure score than other combinations and 26% higher F-measure score than the other classifiers.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we evaluated the effectiveness of the CNN for recognition of human activities using the data from sensors placed at different locations. Data from different combination of sensors was also used to identify best combination of sensors that can be used to recognize thirteen different human activities. Results demonstrate that the presented CNN approach achieved high recognition accuracy as compared to the previous LESH based approach. Moreover, the CNN based approach also demonstrated that a single combination of sensors (all five sensors) is sufficient for recognition of all thirteen activities with significantly high accuracy. An important direction for future research can be to explore the relationship of the depth of the model (number of hidden layers) and the location of individual sensors.

ACKNOWLEDGMENT

Samee U. Khan’s work is supported by (while serving at) the National Science Foundation. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Moreover, the research is partially supported by Higher Education Commission, Pakistan’s startup research grant no. 1604.

REFERENCES

[1] M. Zeng, L. T. Nguyen, B. Yu, O. J. Mengshoel, J. Zhu, P. Wu, and J. Zhang, “Convolutional neural networks for human activity recognition using mobile sensors,” In *6th IEEE International Conference on Mobile Computing, Applications and Services (MobiCASE)*, 2014, pp. 197-205.

[2] M. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, and C.-H. Youn, “Wearable 2.0: Enabling human-cloud integration in next generation healthcare systems," *IEEE Communications Magazine* 55, no. 1, 2017, pp.54-61.

[3] M. Quwaider and Y. Jararweh, “An efficient big data collection in body area networks,” in *5th IEEE International Conference on Information and Communication Systems*, 2014, pp. 1–6.

[4] M. U. S. Khan, A. Abbas, M. Ali, M. Jawad, S. U. Khan, K. Li, and A. Zomaya, “On the Correlation of Sensor Location and Human Activity Recognition in Body Area Networks (BANs),” *IEEE Systems Journal*, vol..12, no. 1, pp. 82-91.

[5] A. Y. Yang, R. Jafari, S. S. Sastry, and R. Bajcsy, “Distributed recognition of human actions using wearable motion sensor networks,” *Journal of Ambient Intelligence and Smart Environments* 1, no. 2, 2009, pp. 103-115.

[6] D. Ravi, C. Wong, B. Lo, GZ Yang, “A deep learning approach to on-node sensor data analytics for mobile or wearable devices,” *IEEE journal of biomedical and health informatics*, vol. 21, no. 1, 2017, pp. 56-64.

[7] W. Xu, M. Zhang, A. A. Sawchuk, and M. Sarrafzadeh, “Robust human activity and sensor location corecognition via sparse signal representation,” *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 1, 2012 pp. 3169–3176.

[8] W. Xu, M. Zhang, A. A. Sawchuk, and M. Sarrafzadeh, “Co-recognition of human activity and sensor location via compressed sensing in wearable body sensor networks,” *in Proceedings of 9th IEEE International Conference on Wearable Implantable Body Sensor Networks,* 2012, pp. 124–129.

[9] L. Liu, Y. Peng, S. Wang, M. Liu, and Z. Huang, “Complex activity recognition using time series pattern dictionary learned from ubiquitous sensors,” *Information Sciences*, vol. 340–342, pp. 41–57, 2016

[10] L. Liu, Y. Peng, M. Liu, and Z. Huang, “Sensor-based human activity recognition system with a multilayered model using time series shapelets,” *Knowledge-Based Systems*, vol. 90, pp. 138–152, 2015.

[11] Y. Kim, “Convolutional neural networks for sentence classification,” In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1746–1751, October 25-29, 2014

[12] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, “Natural Language processing (almost) from scratch,” *Journal of Machine Learning Research,* 2011, vol. 12, 2493—2537.

[13] A. Abbas, M. Ali, M. U. S. Khan, and S. U. Khan, “Personalized Healthcare Cloud Services for Disease Risk Assessment and Wellness Management using Social Media,” *Pervasive and Mobile Computing,* vol. 28, pp. 81-90, 2016.

[14] “Amazon EC2”https://aws.amazon.com/ec2/, accessed on May 2, 2018.