Sequential Modeling for the Recognition of Activities in Logistics

Zafi Sherhan Syed[1], M. Z. Abbas Shah Syed[1], Muhammad Shehram Shah[1], Aunsa Shah[2]

Abstract:

Activity recognition is an important task in cyber physical system research and has been the focus of researchers worldwide. This paper presents a method for activity recognition in logistics operations using data from accelerometer and gyroscope sensors. A Long Short Term Memory (LSTM) Recurrent Neural Network (RNN), bidirectional LSTM (Bi-LSTM) and a Convolutional LSTM (ConvLSTM) are used to classify between six activities being performed in the logistics operations being carried out. Comparing the performance of the LSTMs to the Conv-LSTM network, the designed Bi-LSTM RNN outperforms the other networks considered in this work. This work will aid in the use of sequential modeling approaches for activity recognition in logistics.

Keywords: *Accelerometer, Gyroscope, LARA, Logistics activity recognition, Sequential Modeling*

1. Introduction

Activity recognition is an important task for multiple applications, including but not limited to physical therapy [1], games [2], fall detection [3] and other health specific causes as in [4, 5]. With the introduction of the internet of things and the Industry 4.0 concept, an interest has developed towards research in to cyber physical systems. Such systems allow for the seamless integration of workers into the factory workflow for optimization of industrial processes, thereby increasing industrial productivity. Moreover, with increasing automation, robots and humans work together on the factory floor; and recognizing activities in the manufacturing process will allow for this to happen in a smooth manner. Furthermore, monitoring workers to look at their movements may also be carried out for various health purposes so as to promote a safer working environment. Activities can be recognized using three different modalities, using ambient sensors [6], imaging/video [7] or using movement sensors such as accelerometer, gyroscopes and magnetometers [8]. Within these three modalities, the use of Inertial Measurement Units (IMUs) is the most popular as they tend to not pose limits to movement, are easy to deploy and use and are cheaper in many application scenarios. In an activity recognition system, the IMUs capture movement data as the subject performs one of several activities to be recognized. This data, after possible suitable processing is sent to a classifier which identifies the activity being performed.

Even though, activity recognition has been of great interest to researchers due to its many applications, this task is nontrivial since

[1]Mehran University of Engineering & Technology, Jamshoro, 76090, Sindh, Pakistan
[2]University of Sindh, Jamshoro, 76080, Sindh, Pakistan
Corresponding Author: zaigham.shah@faculty.muet.edu.pk

the movement patterns for different activities may be very similar, for e.g., walking upstairs or downstairs, jogging or jumping etc. Also, the movement patterns might also differ in between the subject themselves which may confuse the classifier algorithm being used. In this regard, deep learning has fostered an era of improved performance of activity recognition algorithms. In this paper, activity recognition has been performed for logistics utilizing deep learning to perform sequential modeling of inertial sensor data. Data from the publicly available LARA dataset [9] is used which contains accelerometer and gyroscope data for workers performing picking and packing operations typical in logistics. The data gathered from these sensors is used for sequential modeling of these activities in order to differentiate between them. The results indicate that the described method offers promising performance for use in the application of activity recognition in a logistics scenario. The paper is organized as follows, Section 2 provides a discussion of the literature search carried out for this work, Section 3 gives a brief introduction to the LARA dataset, Section 4 discusses the methodology for the current work, Section 5 presents and discusses the results and a conclusion is provided in Section 6 along with directions for future work.

2. Literature Review

As discussed in the previous section, activity recognition has garnered a lot of interest from the research community. We discuss some of the relevant literature in this section.

The authors in [10] perform activity classification for order picking using IMUs. They use three IMUs for their task, one on the torso and the other two on the worker's arm. Various statistical features are first computed from a windowed version of the data, resulting in a feature vector of size 54. This is then passed on to three classifiers Support Vector Machines (SVM), Naïve Bayes and Radom Forests (RF) to determine the activity being performed. They find that Random Forests perform the best for their application. More

work by the authors in [11] also target the recognition of activities in an order picking process. They use data from three inertial measurement sensors placed on both wrists and one on the chest. Sliding windows are then used to extract segments from each IMU and then a Convolutional Neural Network (CNN) [12] is used to determine the activity being performed. In [13], Alexander Diete et.al. develop a sensor setup for workers in logistics picking operations. They design two devices, one worn on the wrist and the other on the head. The sensors used in these devices are inertial sensors, video and depth, pressure and ultrasonic sensors. The authors collect their own data and use a neural network and a Random Forest classifier to differetiate between the various intra-activities that are performed in the picking process. In more work from the same authors [14], the authors perform activity recognition for a grabbing operation in an order picking process. They form their dataset based on inertial sensor data as well as video information captured using four different devices. To perform the task, they extract time-frequency domain features from the IMU data and color/feature descriptor (histogram of oriented gradients) from the video frames. They are able to achieve an F-score of 85% as the best result when comparing three different ML algorithms, the SVM, RF and an Artificial Neural Network (ANN).

The authors in [15], use convolutional neural networks to perform activity recognition on three datasets, Opportunity [16], PAMAP2 [17] and a dataset from an order picking processes. The authors use a temporal convolutional network, first suggested in [12] where the CNN has parallel temporal convolutional branches for each IMU and is called CNN-IMU. Their CNN-IMU consists of four convolutional layers, two pooling layers and two fully connected layers. As a baseline, they compare the performance of their architecture with a typical CNN. Segments are extracted from these datasets using a sliding window approach and fed to two variations of the considered networks, it is found that their network performs the best for

the order picking dataset. They also conclude that max-pooling may not necessarily help in improving network performance. In [18], the authors attempt to perform activity recognition in the industry to assess worker performance. They do this using an accelerometer and a gyroscope placed on a worker's wrists in the meat processing industry and sending the captured data to a cloud for further processing. They also consider meat throughput and use proximity sensors as well. Feature extraction (a total of 86 features) is performed from the raw sensor values and multiple classifiers are used for detecting one of the three activities being performed. They find that Random Forests perform the best for their application. The authors in [19] deploy a wireless sensor network system for biomechanical overload assessment in a material handling process. The sensing modalities they use are inertial measurement sensors and electromyography sensors and they use a multilayer perceptron network for classification.

The authors in [20] utilize an Long Short Term Memory (LSTM) Recurrent Neural Network to perform activity recognition on the WISDM [21] dataset using accelerometer sensor data for six activities. They achieve an accuracy of over 94% using an LSTM network composed of 3 LSTM layers consisting of 64 nodes each. The authors in [22] make use of a bidirectional LSTM network for performing human activity recognition from accelerometer and gyroscope data that they collected themselves. They consider six activities to classify from, these were walking, walking upstairs, walking downstairs, laying, standing and sitting. They were able to achieve an overall accuracy of 92.67% with the worst performing activity being sitting and the best being laying down. The number of LSTM layers and the neurons for each layer are determined through a search grid. Their final model was composed of 3 LSTM layers with 175 nodes in each layer. The authors in [23] use a hierarchical LSTM (H-LSTM) model to perform activity recognition using accelerometer sensor data. They use the model to perform activity recognition on the Human Activity Recognition Using Smartphones dataset [24] and the Heterogeneity Human Activity Recognition (HHAR) [25] dataset. Data from the sensor is first denoised and then feature computation (various time-frequency domain features) is performed before being passed on the H-LSTM. The H-LSTM model is made up of two LSTM layers followed by a softmax classifier. They are able to achieve an accuracy of 99.65% (best result) for this task, which when compared to their baselines, established using a Random Forest and Decision Tree classifier is much higher. The authors in [26] compare the performance of traditional machine learning (TDL) methods against deep learning methods for recognizing human activities using IMUs. A comparison of SVM, K-Nearest Neighbours, RF, RNN and CNN is provided for two different datasets, WISDM and USCHAD [27]. From their experiments, they are able to get an accuracy of 90% for deep learning methods and 87% for TDL methods.

As can be observed from the literature, while there have been several approaches towards performing activity recognition in logistics, work in this domain is still ongoing. To contribute to this body of knowledge, we make use of deep learning based sequential modeling approaches to perform activity recognition.

3. LARa Dataset

The Logistics Activity Recognition challenge (LARa) dataset was developed by the Innovationlab Hybrid Services in Logistics at TU Dortmund University and consists of recorded data for two picking operations and one packing operation performed by fourteen different subjects. Each subject is recorded by three means, an OMoCap system which tracks markers on the workers suit and provides movement measurements as coordinates, the second modality are six IMUs capturing acceleration and gyroscopic information placed on the persons waist, legs and arms and their chest and the third is video recordings. The recording of the IMUs was performed with a sampling frequency of 100Hz and eight activity classes were labeled for multiple trials.

These activity classes are standing, walking, cart, handling (upwards), handling (centered), handling (downwards), synchronization and None. In this work, the last two activity classes are not considered as they are not a choreographed part of the dataset's logistics scenario, synchronization referred to a waving motion at the start of each recording and the class None consisted of unrecognizable parts of the recordings.

4. Methodology

The process of sequential modeling of inertial sensor data for the recognition of activities in logistics follows a two-step methodology. The first is the extraction of segments from the recordings present in the dataset and the next is the use of deep learning architectures to determine the activity being performed. In this section we brief on both steps of the process.

4.1. Windowing

The recordings of the activities contain activity annotations for each signal measurement (sample) captured by the sensors. However, in order to train a deep learning algorithm, it is required that activity samples be sent that characterize a movement pattern associated with the activity being performed. In order to extract samples (at the activity level), that make up the picking/packing operations being performed by workers in the LARa dataset, the IMU recordings are windowed in to segments of duration one second with an overlap of 75%. However, since the segments are drawn continuously, the segments contain sensor measurements which belong to multiple different activities. To establish a single label to every segment, we make use of majority voting as the labeling scheme. A similar method has been used by authors previously performing continuous segmentation of activity signals from IMUs [15].

4.2. Classification

Sequential modeling for activity recognition in this work makes use of two algorithms, these have been based on their popularity for activity recognition.

Sequential modeling using deep architectures is mostly reliant on Recurrent Neural Networks (RNN). Instead of looking at the patterns in the spatial domain as performed by Deep Neural Networks and Convolutional Neural Networks, RNNs also aim to make use of sequential information (time in the case of activity recognition) to better understand the input before producing the output. This sequential learning capability has made RNNs quite useful in applications such as language translation [28], time series prediction [29], speech recognition [30] and more. However, vanilla RNNs suffer from vanishing gradients which hampers the training of such networks. Long Short Term Memory networks are a slightly modified class of RNNs which avoid this problem. An LSTM cell is shown in Fig. 1. In the figure, Xt is the input and ht is the current output. An LSTM network is formed by multiple LSTM cells connected together. In this work, the first network used for sequential modeling is a stacked LSTM. A three-layer LSTM network is used for this purpose which is followed by one dense layer with a softmax to indicate to one of the six output classes. Each of the LSTM layers is made up of 100 units and the first two LSTM layers are followed by a drop out layer. The second network used is a bidirectional LSTM network with identical network parameters as the first simple LSTM network. A bidirectional LSTM can learn from input data in both directions and therefore has the potential to capture a better representation of the data being fed to it. In addition to this, two variants of these networks have been tested, one with an extra dense layer of size 60 with a Relu activation function after the LSTM/Bidirectional LSTM layers. This was done in an attempt to get a better representation of the output LSTM features before sending them to the softmax layer for classification. For the remainder of the discussion, the Stacked LSTM with one Dense layer is denoted as Stacked LSTM-1 and the Stacked LSTM with two Dense layers is denoted as Stacked LSTM-2. Similarly, the Stacked Bidirectional LSTM with one Dense layer is denoted as Stacked Bi-LSTM-1 and the Stacked Bidirectional LSTM with two Dense layers is denoted as Stacked Bi-LSTM-

2. The networks are shown in Fig. 2. One of the dense layers is shown with a dotted border to indicate its absence in the specific network variants.

The other network used for time modeling of the activities is a ConvLSTM network. The ConvLSTM network is used as it has shown promising performance for activity recognition tasks [31] and that this network can make use of convolutional filters to better extract features compared to normal LSTM nodes. The ConvLSTM network used in this work consists of two ConvLSTM layers followed by one dense layer. Each of the ConvLSTM layers uses 64 filters and are followed by a dropout layer before the output is flattened prior to the dense layer. The dense layer consists of six neurons and uses a softmax activation function. Similar to the experiment with Stacked LSTMs, an experiment was also
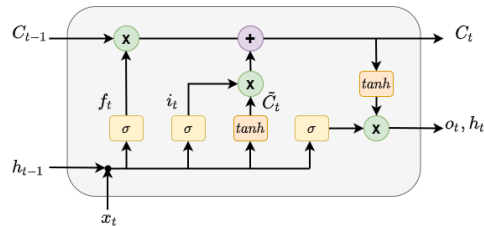

Fig. 1. LSTM Cell

TABLE 1. Parameters for the architectures used for classification

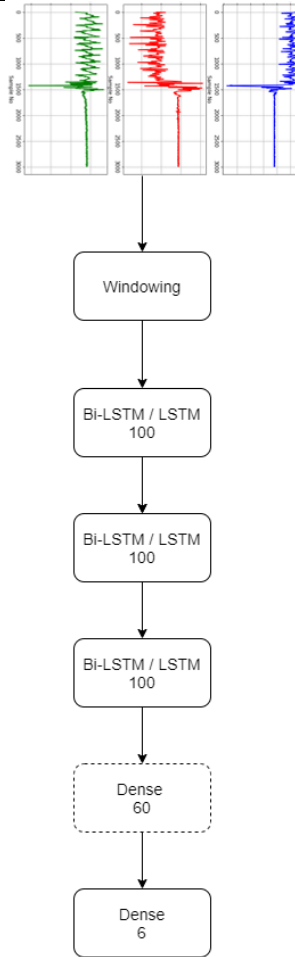| Stacked LSTM / Bi-LSTM		ConvLSTM	
LSTM Layers	3	Conv2D Layers	2
LSTM Layer Size	100	Filters	64
Dense Layers	1/2	Dense Layers	1/2
Dense Layer Size	6/60 and 6	Dense Layer Size	6/60 and 6

Fig. 2. Stacked Bi-LSTM/ LSTM Network

conducted for the ConvLSTM with an extra dense layer before the classification layer, the size of the layer was 60 utilizing a Relu activation function. The first ConvLSTM network with only one dense layer is referred to as ConvLSTM-1 and the ConvLSTM network with two dense layers is referred to as ConvLSTM-2. The ConvLSTM networks used are shown in Fig. 3, a dotted dense layer is shown to indicate its absence in one of the experiments. The details of all these networks are summarized in Table. 1.
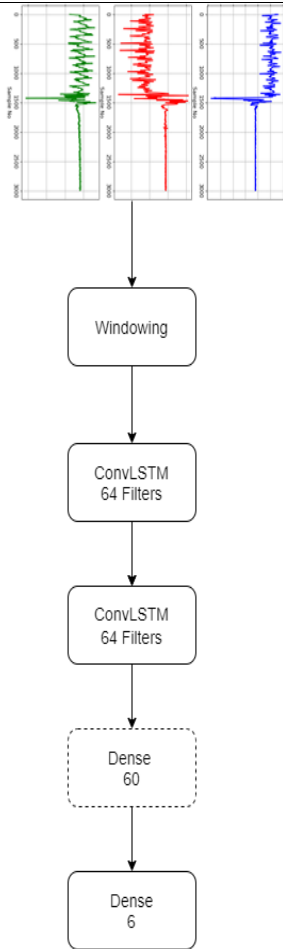
Fig. 3. ConvLSTM Network

5. Experiments

Two experiments were performed, one with using the Stacked LSTM networks and the other using the ConvLSTM network. In both cases, the data was normalized for each sensor and split in to three sets for training, validation and testing. Moreover, a learning rate of 10e-6 with an Adam optimizer was used during the training process and the networks were trained for 50 epochs with early stopping used to determine the best model before overfitting starts. A batch size of 400 was used for these experiments. Results for the experiments have been reported using the F1 score.

5.1. Stacked LSTMs / Bidirectional LSTMs

In this experiment, the Stacked LSTM network was trained using the segmented activity data. Early stopping was used for training all four LSTM networks considered in this work attempting to minimize the unweighted average recall (UAR). Table. 2 lists the epochs for each of the networks to complete training. Moreover, the train and validation loss for each epoch of the training are shown in Fig. 4 - Fig. 7 for the stacked LSTMs / Bi-LSTMs networks respectively.

TABLE 2. Number of epochs for training each LSTM/Bi-LSTM

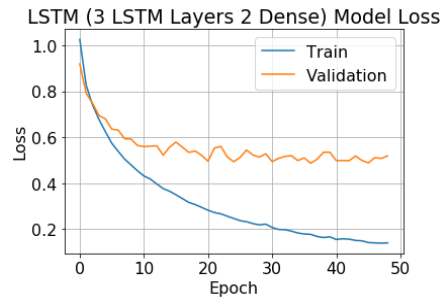| Network	Epochs
Stacked LSTM-1	49
Stacked LSTM-2	47
Stacked Bi-LSTM-1	49
Stacked Bi-LSTM-2	48

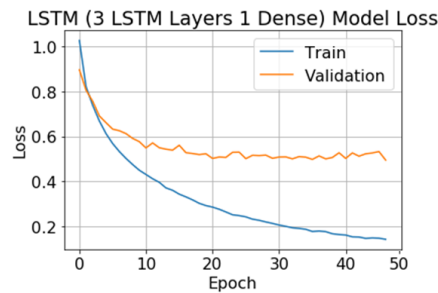**Fig. 4.** Training Loss for Stacked LSTM-1

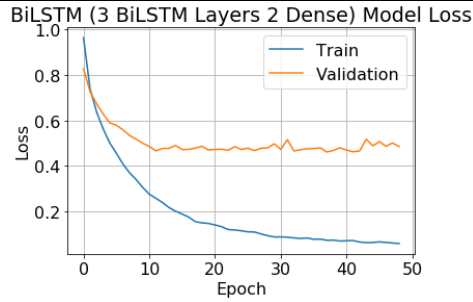
Fig. 5. Training Loss for Stacked LSTM-2

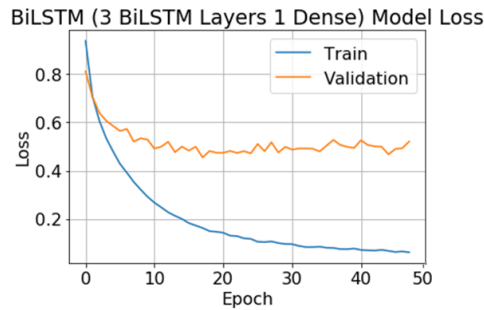**Fig. 6.** Training Loss for Stacked Bi-LSTM-1

Fig. 7. Training Loss for Stacked Bi-LSTM-2

Moreover, the results for activity recognition from this experiment have been shown in Table. 3. The best results have been shown in bold. The test set UARs for the networks were 84.64%, 84.63%, 86.30% and 87.04% for Stacked LSTM-1, Stacked LSTM-2, Stacked Bi-LSTM-1 and Stacked Bi-LSTM-2 respectively.

It can be observed from Table. 3 that the Stacked Bi-LSTM-2 (i.e. with an extra fully connected layer before the classification stage) performs the best overall among the considered LSTM architectures. The stacked LSTM-2 (with an extra dense layer) matches the best performing model for the Cart activity. The worst performing activity was the Stand with the best F1 score of 77%. The best recognized activities were Handling Centered and Cart with an F1 score of 93%. Also, one can note that the addition of an extra dense layer does not provide a performance improvement in the case of stacked LSTMs as opposed to stacked Bi-LSTMs. As far as training is concerned, all four networks do not differ much in the number of epochs required to converge to the best model and the train/validation loss final values are similar as well.

5.2. ConvLSTMs

For this experiment, two stacked ConvLSTM networks were trained which consisted of two ConvLSTM layers and one dense layer for one experiment with an extra dense layer added to for the second one. Early stopping was used for training both stacked ConvLSTM networks considered. The epochs for which each networks' training stopped before overfitting occurred are presented in Table. 4 and training loss plots in Fig. 8 and Fig. 9. The test UARs for the ConvLSTM-1 and ConvLSTM-2 were 72.24% and 72.22% respectively. The results for activity recognition from this experiment have been shown in Table. 5. The best results have been shown in bold.

TABLE 3. F1 Scores for Stacked LSTM/Bi-LSTM

Activity	Stacked LSTM-1	Stacked LSTM-2	Stacked Bi-LSTM-1	Stacked Bi-LSTM-2
	F1 (%)	*F1 (%)*	*F1 (%)*	*F1 (%)*
Stand	72	71	76	**77**
Walking	83	82	**85**	**85**
Cart	91	**93**	92	**93**
Handling Upwards	80	78	81	**83**
Handling Centered	92	91	92	**93**
Handling Downwards	83	83	85	**87**

TABLE 4. Number of epochs for training each ConvLSTM

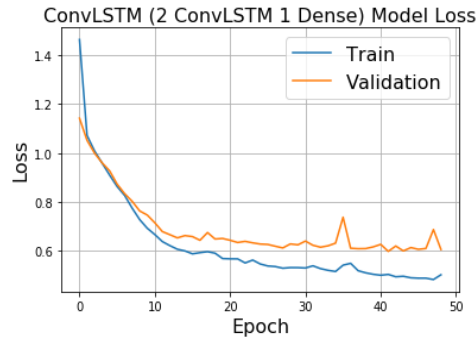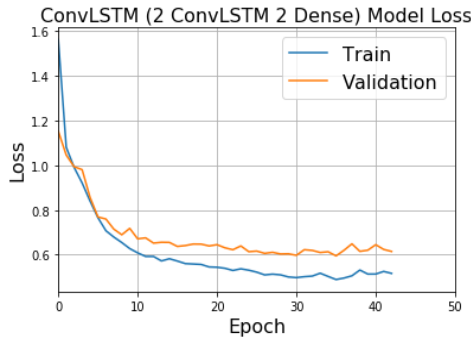| Network	Epochs
Stacked ConvLSTM-1	43
Stacked ConvLSTM-2	49

Fig. 9. Training Loss for ConvLSTM-1

Fig. 10. Training Loss for ConvLSTM-2

It can be observed from Table. 5 that both the ConvLSTM networks have performed nearly equally well. The worst recognized activity, as in the previous experiments, in this case is also Stand. The best recognized activity is Handling Centered as was similar in the experiments of the last section.

When comparing all the experiments performed, it can be observed that the stacked LSTMs in general have performed much better than the stacked ConvLSTM which indicates that the convolutional filters present in the stacked ConvLSTM architecture were not beneficial towards sequential modeling of the activities in the scenario considered. Among all the models experimented with, stacked bi-

directional LSTMs with only one dense layer (classification layer) was found to provide the best results. In order to better understand the performance of this model, the F1 Score, Precision and Recall for each of the activities have been shown in Table 6. Looking at the detailed performance numbers, one can see that the activities Stand and Handling Upwards are the worst recognized activities while Handling Centered, Cart are the best recognized activities.

6. Conclusion

In this paper, activity recognition has been performed for a logistics scenario by using sequential modeling approaches. To achieve this, three networks have been used, Stacked LSTMs, Stacked Bidirectional LSTMs and stacked ConvLSTMs. Moreover, two variants of each of these networks have been utilized. After extracting segments of activity data captured using inertial measurement sensors with sliding windows, the activity data is sent to these three networks for recognition purposes. It was found that, from the networks considered, stacked bidirectional LSTM with two dense layers performs the best. The most poorly recognized activity was Stand whereas the best recognized activity was Handling Centered. Moreover, from the experiments with stacked ConvLSTM architectures, it was determined that with the given experimental settings, the

TABLE 5. F1 Scores for ConvLSTM

Activity	Stacked ConvLSTM-1	Stacked ConvLSTM-2
	F1 (%)	*F1 (%)*
Stand	**49**	48
Walking	**74**	73
Cart	86	87
Handling Upwards	**64**	**64**
Handling Centered	**86**	**86**
Handling Downwards	**77**	**77**

TABLE 6. F1 Score, Precision, Recall for best performing model

Activity	Best Performing Network (Stacked Bi-LSTM-2)		
	F1 (%)	*Precision (%)*	*Recall (%)*
Stand	**77**	76	77
Walking	**85**	86	84
Cart	**93**	90	96
Handling Upwards	**83**	80	86
Handling Centered	**93**	94	92
Handling Downwards	**87**	87	87

convolutional features extracted by the ConvLSTM did not result in better performance compared to a stacked simple LSTM.

This work serves as a basis for further research in the sequential learning for activity recognition in logistics. Further work in his area would be the use of attention mechanisms within the sequential learning scheme, investigating importance of sensors and cross-dataset performance analysis.

AUTHOR CONTRIBUTION

All authors contributed equally to the work.

DATA AVAILABILTY STATEMENT

The dataset used during the current study is publicly available at Logistic Activity Recognition Challenge (LARa) - A Motion Capture and Inertial Measurement Dataset [https://zenodo.org/record/3862782#.YHTlDj 8pBqM].

CONFLICT OF INTEREST

The authors declare no conflict of interest.

FUNDING

Not applicable.

ACKNOWLEDGMENT

Not applicable.

REFERENCES

[1] L. Billiet, T. Swinnen, R. Westhovens, K. de Vlam and S. Van Huffel, Activity recognition for physical therapy: fusing signal processing features and movement patterns. In Proceedings of the 3rd International Workshop on Sensor-based Activity Recognition and Interaction, 2016, p. 1-6

[2] A. Almeida and A. Alves, Activity recognition for movement-based interaction in mobile games. In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, 2017, p. 1-8.

[3] A. Chelli and M. Pätzold, A machine learning approach for fall detection and daily living activity recognition, IEEE Access, 7, 2019, p. 38670-38687.

[4] Wei-Yi Cheng, A. Scotland, F. Lipsmeier, T. Kilchenmann, L. Jin, Jens Schjodt-Eriksen, D. Wolf et al. Human activity recognition from sensor-based large-scale continuous monitoring of Parkinson's disease patients, In 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2017, p. 249-250

[5] K. J. Kim, N. M. Hassan, S. H. Na and E. N. Huh, Dementia wandering detection and activity recognition algorithm using tri-axial accelerometer sensors. In Proceedings of the 4th International Conference on Ubiquitous Information Technologies & Applications, 2009, p. 1-5

[6] D. Singh, E. Merdivan, I. Psychoula, J. Kropf, S. Hanke, M. Geist and A. Holzinger, Human activity recognition using recurrent neural networks. In International Cross-Domain Conference for Machine Learning and Knowledge Extraction, 2017, p. 267-274

[7] T. Dobhal, V. Shitole, G. Thomas and G. Navada, Human activity recognition using binary motion image and deep learning. Procedia computer science, 58, 2015, p. 178-185.

[8] W. Sousa Lima, E. Souto, K. El-Khatib, R. Jalali and J. Gama, Human activity recognition using inertial sensors in a smartphone: An overview. Sensors, 19(14),2019, p. 3213.

[9] F. Niemann, C. Reining, F. M. Rueda, N. R. Nair, J. A. Steffens, G. A. Fink and M. T. Hompel, LARa: Creating a Dataset for Human Activity Recognition in Logistics Using Semantic Attributes. Sensors, 20(15), 2020, p. 4083.

[10] S. Feldhorst, M. Masoudenijad, M. ten Hompel and G. Fink, Motion classification for analyzing the order picking process using mobile sensors. In Proc. Int. Conf. Pattern Recognition Applications and Methods, 2016, p. 706-713.

[11] R. Grzeszick, J. M. Lenk, F. M. Rueda, G. A. Fink, S. Feldhorst, and M. Ten Hompel, Deep neural network based human activity recognition for the order picking process. ACM International Conference Proceeding Series, Part F1319. 2017

[12] Yang, M. N. Nguyen, P. P. San, X. Li and S. Krishnaswamy, S, Deep convolutional neural networks on multichannel time series for human activity recognition. Ijcai, 15, 2013, p. 3995–4001.

[13] A. Diete, L. Weiland, T. Sztyler, and H. Stuckenschmidt, Exploring a multi-sensor picking process in the future warehouse. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, 2016, p. 1755–1758.

[14] A. Diete, T. Sztyler, L. Weiland and H. Stuckenschmidt Recognizing Grabbing Actions from Inertial and Video Sensor Data in a Warehouse Scenario, 14th International Conference on Mobile Systems and Pervasive Computing, 2017, p. 16–23.

[15] F. Moya Rueda, R. Grzeszick, G. A. Fink, S. Feldhorst and M. Ten Hompel, Convolutional neural networks for human activity recognition using body-worn sensors. Informatics, 5(2), 2018, p. 26.

[16] R. Chavarriaga, H., Sagha, A. Calatroni, S. T. Digumarti, G. Tröster, J. D. R. Millán and D. Roggen, D, The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition. Pattern Recognition Letters, 34(15), 2013, p. 2033-2042.

[17] PAMAP2 Physical Activity Monitoring Data Set. Available online: http://archive.ics.uci.edu/ml/datasets/ pamap2+physical+activity+monitoring (accessed: 25 March 2021).

[18] A. R. M. Forkan, F. Montori, D. Georgakopoulos, P. P. Jayaraman, A.Yavari and A. Morshed, An Industrial IoT Solution for Evaluating Workers' Performance Via Activity Recognition. 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), 2019, p. 1393–1403.

[19] P. Giannini, G. Bassani, C. A. Avizzano and A. Filippeschi, Wearable Sensor Network for Biomechanical Overload Assessment in Manual Material Handling. Sensors, 20(14), 2020, p. 3877

[20] S. W. Pienaar and R. Malekian, Human activity recognition using LSTM-RNN deep neural network architecture. In 2019 IEEE 2nd Wireless Africa Conference (WAC), 2019, p. 1-5

[21] WISDM Lab: Dataset. Available online: www.cis.fordham.edu/wisdm/dataset.php#actitrack er (accessed: 25 March 2021).

[22] F. Hernández, L. F.Suárez, J. Villamizar and M. Altuve, Human activity recognition on smartphones using a bidirectional lstm network. In 2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA), 2019, p. 1-5

[23] L. Wang and R. Liu, Human activity recognition based on wearable sensor using hierarchical deep LSTM networks, Circuits, Systems, and Signal Processing, 39(2), 2020, p. 837-856.

[24] J. L. Reyes-Ortiz, D. Anguita, A. Ghio and X. Parra, Human activity recognition using smartphones data set, UCI Machine Learning Repository; University of California, Irvine, School of Information and Computer Sciences: Irvine, CA, USA, 2012

[25] J. L. Reyes-Ortiz, L. Oneto, A. Ghio, A. Samá, D. Anguita and X. Parra, Human activity recognition on smartphones with awareness of basic activities and postural transitions. In International conference on artificial neural networks, 2014, p. 177-184

[26] C. Hou, A study on IMU-Based Human Activity Recognition Using Deep Learning and Traditional Machine Learning, In 2020 5th International Conference on Computer and Communication Systems (ICCCS) , 2020, p. 225-234

[27] M. Zhang and A. A. Sawchuk, USC-HAD: a daily activity dataset for ubiquitous activity recognition using wearable sensors. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, 2012, p. 1036-1043.

[28] K. Cho, B. Van Merriënboer, C. Gulcehre, D. BahdanauF. Bougares, H. Schwenk and Y. Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:,2014, 1406.1078.

[29] Z. Che, S. Purushotham, K. Cho, D. Sontag and Y. Liu, Recurrent neural networks for multivariate time series with missing values. Scientific reports, 8(1), 2018, p. 1-12.

[30] Y. Miao, M. Gowayyed and F. Metze, F, EESEN: End-to-end speech recognition using deep RNN models and WFST-based decoding. In 2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), 2015, p. 167-174

[31] Zili Li, Liu Yixin, Guo Xuerong and Ji Zhang. Multi-convLSTM neural network for sensor-based human activity recognition, In Journal of Physics: Conference Series, 1682(1), 2020, p. 012062.