

DEEP ARCHITECTURES FOR HUMAN ACTIVITY RECOGNITION USING SENSORS

Zartasha Baloch

Mehran University of Engineering & Technology, Jamshoro (Pakistan)
E-mail: zartasha.baloch@faculty.muet.edu.pk

Faisal Karim Shaikh

Mehran University of Engineering & Technology, Jamshoro (Pakistan)
E-mail: faisal.shaikh@faculty.muet.edu.pk

Mukhtiar Ali Unar

Mehran University of Engineering & Technology, Jamshoro (Pakistan)
E-mail: mukhtiar.unar@faculty.muet.edu.pk

Recepción: 05/03/2019 **Aceptación:** 15/03/2019 **Publicación:** 17/05/2019

Citación sugerida:

Baloch, Z., Shaikh, F. K. y Unar, M. A. (2019). Deep Architectures for Human Activity Recognition using Sensors. *3C Tecnología. Glosas de innovación aplicadas a la pyme. Edición Especial, Mayo 2019*, pp. 14–35. doi: <http://dx.doi.org/10.17993/3ctecno.2019.specialissue2.14-35>

Suggested citation:

Baloch, Z., Shaikh, F. K. & Unar, M. A. (2019). Deep Architectures for Human Activity Recognition using Sensors. *3C Tecnología. Glosas de innovación aplicadas a la pyme. Special Issue, May 2019*, pp. 14–35. doi: <http://dx.doi.org/10.17993/3ctecno.2019.specialissue2.14-35>

ABSTRACT

Human activity recognition (HAR) is a renowned research field in recent years due to its applications such as physical fitness monitoring, assisted living, elderly-care, biometric authentication and many more. The ubiquitous nature of sensors makes them a good choice to use for activity recognition. The latest smart gadgets are equipped with most of the wearable sensors i.e. accelerometer, gyroscope, GPS, compass, camera, microphone etc. These sensors measure various aspects of an object, and are easy to use with less cost. The use of sensors in the field of HAR opens new avenues for machine learning (ML) researchers to accurately recognize human activities. Deep learning (DL) is becoming popular among HAR researchers due to its outstanding performance over conventional ML techniques. In this paper, we have reviewed recent research studies on deep models for sensor-based human activity recognition. The aim of this article is to identify recent trends and challenges in HAR.

KEYWORDS

Deep Learning models, Sensors, Human activity recognition.

1. INTRODUCTION

Recent years have shown significant progress in the use of smart gadgets and sensor-enabled devices. The reduced cost of these devices and ease of use makes them a perfect choice to use for human activity recognition (HAR). HAR is trending research field with its various applications including smart homes, sports, health monitoring, emergency services, and lifelogging (Chan, EstèVe, Fourniols, Escriba & Campo, 2012; Lara & Labrador, 2013).

Initially, activity recognition task was successfully done through video recordings but video-based systems are location specific and it somewhat interfere one's personal life. For the reason, sensor-based activity recognition is gaining widespread acceptance. In sensor based HAR systems, low cost wearable sensors are deployed which reduces interference in daily activities. Another recent development in HAR is use of smartphones as these latest cell phones are equipped with many sensors. The unobtrusive nature of smartphones makes them appropriate for HAR.

Activity recognition systems most often use classification algorithms to classify activities as class labels. Like other time-series data, the first step in sensor based HAR is to segment data into time frames and then to extract time and frequency domain feature from those data segments. In conventional machine learning algorithms feature extraction is often done by manually using heuristic methods, in contrast deep Learning provides automatic feature extraction. It also helps in mining complex knowledge from massive amount of unsupervised data. Plötz, Hammerla, and Olivier (2011) used deep learning for the first time for feature extraction and compared the results with principal component analysis. After that a number of researchers worked on deep learning approaches for automatic feature extraction in human activity recognition (Twomey, *et al.*, 2018; Ronao, Charissa & Cho, 2015; Alsheikh, *et al.*, 2016). The main contribution of this research is to review latest trends in human activity recognition using deep architectures. This paper reviews and analyses recent research articles on deep learning based HAR using sensors.

The rest of the paper is organized as follows: section 2 discusses some of the deep learning architectures and section 3 elaborates some publically available datasets used for HAR. In section 4, recent studies on deep learning based HAR are presented. Section 5 presents research challenges in activity recognition field. Finally, section 6 concludes the article.

2. DEEP LEARNING ARCHITECTURES

Deep learning (DL) is a renowned field of research and successfully implemented in image and voice recognition problems. Generally, there are three categories of deep models, which are generative, discriminative and hybrid models (Deng & Jaitly, 2016). A generative model learns the true distribution of training data and makes some variations to generate new samples which follow same probabilistic distribution. Some generative DL methods include Restricted Boltzman Machine (RBM), Deep Autoencoders, and Sparse Coding. A discriminative method directly estimates the probability of the output given an input i.e. $p(y|x)$ by approximating posterior distribution classes. Most commonly used discriminative models in activity recognition are Convolutional Neural Network (CNN) and Recurrent Neural Networks (RNN) (McDaniel & Quinn, 2018). Many research studies have combined discriminative and generative methods to extract more effective features. The combination of generative and discriminative models is known as hybrid model. In most studies CNN is used along with other generative or discriminative methods for HAR. This section explores some of the deep learning models used in sensor-based HAR.

2.1. CONVOLUTIONAL NEURAL NETWORK

The Convolutional Neural Network (CNN) learns internal representations of raw sensor data without domain expertise in feature engineering (Ronao & Cho, 2015). For the reason, it is most widely used method for data analysis and activity recognition. In CNN convolution operation is performed on sensor data through many hidden layers. The components of CNN include convolutional

layer, pooling layer, dense (fully connected) layer and softmax layer (Ignatov, 2018). **Convolutional layer** detects distinct features from input by performing convolution operation on data. The first convolution layer identifies low level features whereas next convolutional layers detect higher level features (Namatēvs, 2017). The convolutional layers then introduce nonlinearities to the model through using activation functions such as tanh, sigmoid and rectified linear unit (ReLU) (Albelwi & Mahmood, 2017). **Pooling layer** is used to downsample the dimensionality of the feature map. It compresses features and reduces network's computational complexity (Affonso, Rossi, Vieira & Ferreira, 2017). Most frequently used pooling algorithm is max pooling which is robust to small changes (Kautz, *et al.*, 2017). The last component of CNN is **dense layers or fully connected layers**. These layers are fused with softmax classifier to perform classification on extracted features. So far CNN is the most widely used deep model in activity recognition and feature learning.

Zeng, *et al.* (2014) proposed a CNN based approach for HAR which automatically extract discriminative patterns and captures local dependencies of a sensor signal. They used partial weight sharing method to accelerometer data for performance improvement. Yang, Nguyen, San, Li, and Krishnaswamy (2015) also presented a CNN model for multichannel time-series data for HAR. The convolution and pooling layers of the proposed model capture the salient features, which are systematically unified among multiple channels and then mapped into activity classes.

2.2. RESTRICTED BOLTZMANN MACHINE

Restricted Boltzmann Machine (RBM) is a stochastic deep model, which learns a probability distribution on its input dataset using a layer of binary hidden units. The meaningful features are automatically extracted from input labelled and unlabeled data. It is most commonly used for dimensionality reduction and complex feature learning problems. It is a type of shallow neural network that learns to reconstruct data by itself in an unsupervised manner. There are two variations in RBM, one is Deep Belief Networks (DBN) and other is Deep Boltzmann Machine (DBM).

The concept of deep belief networks was first conceived by Hinton, Osindero, and Teh (2006) as a replacement of backpropagation. In terms of network structure, a DBN is very similar to multilayer perceptron but their training process is entirely different. In fact, the difference in training method is key factor that enables DBN to outperform the shallow counterpart. A deep belief network consists of multiple hidden layers. The layers are connected with each other but the units in each layer are not connected. To make learning easier the connectivity is restricted i.e. there is no connection between hidden units. DBNs can be divided in two major parts. The first one consists of multiple layers of RBMs to pre-train the network, while the second one is a feed-forward backpropagation network that will further refine the results from the RBM stack. Alsheikh, *et al.* (2016) proposed a DBN based model which is trained on greedy layer-wise training of RBMs. The proposed model provides better recognition accuracy of human activities and avoids expensive design of handcrafted features. Bhattacharya and Lane (2016) used RBM-based pipeline for activity recognition and have shown their approach outperforms other modeling alternatives.

2.3. AUTOENCODERS

Autoencoders are deep neural networks to perform data compression using machine learning. An autoencoder learns compressed distributed representation of input data for dimensionality reduction (Nweke, Teh, Al-Garadi & Alo, 2018). It applies back propagation i.e. the output values will be set as the input. Principle Component Analysis (PCA) does the same for linear functions whereas autoencoder can perform non-linear transformations. An autoencoder also gives a representation as to the output of each layer and having multiple representations of different dimensions is always useful. So an autoencoder uses pre-trained layers from other models to apply transfer learning to prime the encoder of the decoder. There are three components of an autoencoder; encoder, code, and decoder. **The encoder** compresses/encodes the input data into a latent space representation. **The code** represents the compressed input that is fed to the decoder. **The decoder** reconstructs/decodes the input from the latent space representations.

Different variations in autoencoders include Sparse Autoencoder (SAE), Denoising Autoencoder (DAE) (Nweke, *et al.*, 2018). Almaslukh, AlMuhtadi and Artoli (2017) proposed stacked autoencoder based model for better recognition accuracy along with reduced recognition time.

2.4. RECURRENT NEURAL NETWORK

Recurrent Neural Network (RNN) is a deep model with cyclic connections, which empowers it to capture correlations between time series data. RNN is successfully used in handwriting recognition and speech recognition applications (Wang, Chen, Hao, Peng & Hu, 2018). RNN is a network with a loop in it allowing information to persist. The iterative nature of RNN enables data to be passed starting with one stage of the network to the next. RNN can be considered as numerous replicas of the same network, each network passes information to the next. RNN is a very flexible and powerful network which does not require additional data labelling and works well for modelling short-term memory. This makes it a good choice to easily model sequence learning or time-related problem where the output of one layer acts as an input to the next layer. There are two variations of recurrent neural networks, one is Long Short Term Memory (LSTM) and the other is Gated Recurrent Unit (GRU). Such networks make use of different gates and memory cells to store time series sequences (Graves, 2013). Murad and Pyun (2017) used unidirectional, bidirectional and cascaded deep RNN on five public datasets. They proposed three novel LSTM-based deep RNN architectures which extract discriminative features using deep layers and provide performance improvements. M. Inoue, S. Inoue and Nishida (2018) proposed an RNN based approach to provide better recognition accuracy with reduced recognition time.

2.5. HYBRID MODELS

Hybrid models are a combination of generative and discriminative models. Many researchers have implemented hybrid models in activity recognition as well as in other fields. For instance, Murahari and Plötz (2018) used deep convolutional LSTM model to explore the temporal context in activity recognition. Lee, Grosse, Ranganath, and Ng (2009) proposed a convolutional deep belief network that used the probabilistic max-pooling technique for visual recognition tasks. In some studies, RNN and CNN are combined together where CNN captures spatial relationships and RNN uses temporal relationships. Ordóñez and Roggen (2016) presented a deep convolutional LSTM recurrent neural network for multimodal wearable sensors. The deep CNN is used for automated feature extraction and LSTM recurrent unit captures temporal dynamics of activities. Yao, Zhao, Hu, and Abdelzaher (2018) also introduced a CNN and RNN based framework which designs a self-attention module for estimating input quality by exploiting its temporal dependencies. More research in these models is expected in future.

Deep Models for Activity Recognition

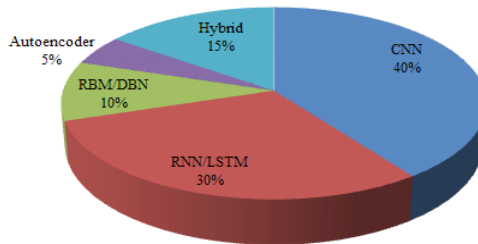


Figure 1. Deep Models for Activity Recognition.

The Figure 1 shows a pie chart showing the percentage of the deep learning methods used in activity recognition. This percentage only represents deep models used in studies presented in this research article.

The mostly used deep model here is CNN which is 40%, this is due to the success of CNN in the image processing field. CNN also gives an outstanding performance in sensor-based HAR due to its discriminative feature extraction

capabilities. Other models also perform well in activity recognition and are gaining popularity.

3. DATASETS

Validating a new human activity recognition approach on the new or self-created dataset is a challenging task. The effectiveness of such approaches can be achieved by testing them on some standard datasets where the researchers have already tested their results. This section gives a brief description of some publically available benchmark datasets which paid a remarkable contribution in HAR research. All these datasets are sensor based and are summarized in Table 1.

Table 1. Datasets for sensor-based HAR (ADL= Activities of Daily Life, A= Accelerometer, G= Gyroscope, M= Magnetometer, HR= Heart Rate, AM= Ambient Sensors, O= Object Sensors, L= Light Sensors, S= Sound Sensor, ECG= Electrocardiogram, EEG= Electroencephalogram, EOG= Electro-Oculogram, GPS= Global Positioning System, MF= Magnetic Field Sensor).

Dataset	Sensors used	Activities	Sampling rate (Hz)	No. of Instances	No. of Subjects	Application
DaLiAc (Leutheuser, Schulthaus & Eskofier, 2013)	A, G	Sitting, Lying, Standing, Washing, dishes, Vacuuming, Sweeping, Walking outside, Ascending stairs, Descending stairs, Treadmill running (8.3 km/h), Bicycling (50 watt), Bicycling (100 watt), Rope jumping	200	—	19	ADL
UCI HAR (Anguita, Ghio, Oneto, Parra & Reyes-Ortiz, 2013)	A, G	Walking, Walking upstairs, Walking downstairs, Sitting, Standing, Laying	50	10,299	30	ADL
PAMAP2 (Reiss & Stricker, 2012)	A, G, M, HR	Lying, Sitting, Standing, Walking, Running, Cycling, Nordic Walking, Watching Tv, Computer Work, Car Driving, Ascending Stairs, Descending Stairs, Vacuum Cleaning, Ironing, Folding Laundry, House Cleaning, Playing Soccer, Rope Jumping, Other (Transient Activities)	100	3,850,505	9	ADL
WISDM (Kwapisz, Weiss & Moore, 2010)	A	Walking, Jogging, Upstairs, Downstairs, Sitting, Standing	20	1,098,207	29	ADL

Dataset	Sensors used	Activities	Sampling rate (Hz)	No. of Instances	No. of Subjects	Application
Actitracker (Lockhart, <i>et al.</i> , 2011)	A	Walking, Jogging, Stairs, Sitting, Standing, Lying down	20	2,980,765	36	ADL
OPPORTUNITY (Roggen, <i>et al.</i> , 2010)	A, G, M, O, AM	17 activities including ADL run and Drill run	32	701366	4	ADL
STISEN (Stisen, <i>et al.</i> , 2015)	A, G	Biking, Sitting, Standing, Walking, Stair Up, Stair Down	100~200	43930257	9	ADL
GAIT (Ngo, <i>et al.</i> , 2014)	A, G	walking on a flat surface, walking up the slope, walking down the slope, descending stairs and ascending stairs	100	—	744	Gait Analysis
Sleep-EDF (Goldberger, <i>et al.</i> , 2000)	EEG, EOG, EMG	Sleep stages i.e. Awake, Stage-1, Stage 2, Stage 3, Stage 4, REM	100	—	20	Sleep Analysis
RealWorld HAR (Sztyley & Stuckenschmidt, 2016)	A, G, GPS, L, MF, S	Climbing Downstairs, ClimbingUpstairs, Jumping, Lying, Standing, Sitting, Running/Jogging, And Walking	50	944,356	15	ADL

4. DISCUSSION

Although conventional machine learning algorithms have shown remarkable performance in recognition of human activities, these algorithms require domain expertise to develop robust features for high dimensional complex real-world data. However, this is time consuming and expensive task. This captivated researchers towards the use of deep learning. In deep architectures, the layers of feature representations are stacked together to extract more complex features in data. Recent studies have shown the incredible performance improvement of deep learning in HAR. Feature extraction plays a significant role in recognition process as it extracts features from sensor data which helps in reducing computational complexity and improves classification accuracy (Abidine, Fergani, Fergani & Oussalah, 2018). Conventional approaches use hand-crafted feature engineering, whereas in deep learning features are automatically learned through the deep network. Another challenge is most ML algorithms require a good amount of labelled data for model training but the data in real-time applications is mostly unlabeled. Deep learning works well with unlabeled data too (Almaslukh, *et al.*, 2017). Table 2 provides recent research on deep learning based human activity recognition models.

Table 2. Deep learning based recent research for HAR.

Study	Dataset	Deep Learning Method	Application	Description
(Alsheikh, <i>et al.</i> , 2016)	WISDM, Daphnet Gait, Skoda	DBN	ADL, Parkinson	DBNs are trained on greedy layer-wise training of RBMs. The proposed model provides better recognition accuracy of human activities and avoids the expensive design of handcrafted features.
Hammerla, Halloran & Plötz, 2016)	OPPORTUNITY, PAMAP2, Daphnet Gait	DNN, LSTM, CNN	ADL, smart home, Gait	They introduced a novel regularization approach. Three deep learning approaches DNN, RNN and CNN are explored across three benchmark datasets.
(Ordóñez, <i>et al.</i> , 2016)	OPPORTUNITY, Skoda	Deep Convolutional LSTM	ADL	They presented a deep convolutional LSTM recurrent neural network for multimodal wearable sensors. The deep CNN is used for automated feature extraction and LSTM recurrent unit captures temporal dynamics of activities.
(Bhattacharya, <i>et al.</i> , 2016)	OPPORTUNITY, Self-generated	RBM	ADL, Gesture, Transportation	They used RBM-based pipeline for activity recognition which outperforms for other modelling alternatives.
(Murad, <i>et al.</i> , 2017)	UCI-HAD, USC-HAD, Daphnet FoG, OPPORTUNITY, Skoda	DRNN	ADL, Gait	They used unidirectional, bidirectional and cascaded DRNN on five public datasets. They proposed three novel LSTM-based DRNN architectures which extract discriminative features using deep layers and gives performance improvement.
(Ravi, Wong, Lo & Yang, 2017)	ActiveMiles, WISDM, Skoda, Daphnet FoG	CNN	ADL	They used shallow and deep features for activity classification and it resolves the issues related to on-node computations.
(Münzner, <i>et al.</i> , 2017)	PAMAP2, Robert Bosch Hospital (RBK)	CNN	ADL	They evaluated the influence of normalization techniques and explored the change in classification accuracy by using early and late fusion techniques.
(Almaslukh, <i>et al.</i> , 2017)	UCI HAR	Stacked Autoencoder (SAE)	ADL	They proposed stacked autoencoder based model which provides better recognition accuracy with reduced recognition time.
(Yao, <i>et al.</i> , 2018)	STISEN	Deep Convolutional RNN	ADL	They introduced QualityDeepSense framework for IoT applications. It estimates input sensing quality by using temporal dependencies.
(Radu, <i>et al.</i> , 2018)	STISEN, GAIT, Sleep-Stage, Indoor-Outdoor	Multimodal-CNN, Multimodal-DNN	ADL, Gait, Sleep analysis	Four distinct multimodal CNN architectures have been proposed for activity and context recognition.
(Murahari, <i>et al.</i> , 2018)	OPPORTUNITY, PAMAP2, Skoda	DeepConvLSTM	ADL	An Attention model has been proposed in activity recognition research as a data-driven approach to explore temporal context. Attention layers have been added to DeepConvLSTM model.

Study	Dataset	Deep Learning Method	Application	Description
(Almaslukh, Artoli & Al-Muhtadi, 2018)	RealWorld HAR	CNN	ADL	Deep convolution neural network model for position-independent activity recognition.
(Khan, Roy & Misra, 2018)	Self-generated	Heterogeneous Deep CNN	ADL	A CNN-based HAR approach for transfer learning with automatic model learning across different domains while requiring minimum labelled data.
(Zhu, Chen & Yeng, 2018)	UCI HAR	Deep LSTM	ADL	A semi-supervised model using a DeepLSTM based approach with temporal ensembling for activity recognition using inertial sensors.
(Xi, <i>et al.</i> , 2018)	OPPORTUNITY, PAMAP2	CNN, RNN	ADL	They used dilated convolutional layers to automatically extract inter-sensor and intra-sensor features. They also proposed a novel dilated SRU (Simple Recurrent Unit) approach to capture the latent time dependencies among features.
(Ignatov, 2018)	WISDM, UCI HAR	CNN	ADL	A CNN based approach to provide user-independent human activity recognition with small recognition intervals (1s) and almost no preprocessing and feature engineering required.
(Inoue, <i>et al.</i> , 2018)	HASC corpus, UCI HAR	RNN	ADL	They used an RNN based approach to provide better recognition accuracy with reduced recognition time.
(McDaniel & Quinn, 2018)	UCI HAR	LTSM	ADL	They proposed LTSM based pipeline which can directly process raw data without extensive preprocessing and gives outstanding performance.

5. RESEARCH CHALLENGES

Human activity recognition is a trending research field with many challenges that need to be addressed. Although, HAR is a well-researched field still these challenges need to be further investigated for the effective realization of HAR systems. These research challenges include:

Sensor placement: The position of the sensor plays an important role in recognition accuracy. Different placement positions include right/left arm, ankle, foot, hip and chest etc. Sensor signal readings vary at different positions for the same activity.

Sensor modalities: Sensor modality can be classified into wearable sensors, ambient sensors and object sensor. In most HAR systems, wearable sensors are successfully used but in only few research studies these sensor modalities are

combined to improve the recognition accuracy and to infer high-level activities such as having coffee with RFID tag on the cup.

Compatibility with real-world data: Real world data is often different from laboratory data with the constrained environment. Most of real-world data come in streams and are unlabeled; therefore the HAR systems should be robust enough for real-world scenarios.

Context Awareness: As HAR systems are designed to provide analysis on user's activity and behavior, so it is necessary that the system must be aware of user's behavior, age, gender and physical condition and environment. For example, running the signal of a 75 years old patient might be equivalent to the walking of a young user. In such situations context information is vital.

Overlapping activities: Most of the HAR systems recognize single activity at a time such as walking, standing, sitting or brushing teeth, but generally there can be some overlapping activities like having coffee while watching TV or walking while drinking water. Few types of research have been done in this direction but still, there are good research opportunities in this direction.

Hyper-parameter setting: accuracy of deep models heavily rely on adjustment of network parameters such as learning rate, dropout, filter size, kernel reuse, no. of units and deep layers, regularization etc. In most of the research, these parameters are set using heuristic methods (Liu, *et al.*, 2017), so there is a need to use optimization algorithms to adjust these hyper-parameters.

Sensor Fusion: In sensor based HAR systems, it is crucial to choose which sensors need to be fused together to improve the recognition process. Münzner, *et al.* (2017) presented four data fusion techniques including early fusion, sensor-based late fusion, channel-based late fusion and shared filters hybrid fusion. Chowdhury, Tjondronegoro, Chandran and Trost (2017) also presented a fusion technique namely posterior adapted class based fusion.

6. CONCLUSION

This article discussed recent developments in sensor-based human activity recognition using deep architectures. The goal of this article is to identify recent trends and challenges in HAR. Recent research studies on HAR are compared with respect to sensor type, the dataset used, deep learning models and its applications. Some basic deep models are also discussed, which are successfully implemented in HAR. The paper also presented some publically available sensor based datasets for activity recognition. In the end, various research challenges are discussed which may be addressed to make HAR systems more robust and implementable in real-world scenarios.

ACKNOWLEDGEMENTS

This work has been performed under IICT, Mehran University of Engineering and Technology, Jamshoro and funded by ICT Endowment Fund for sustainable development.

REFERENCES

- Abidine, B. M. H., Fergani, L., Fergani, B. & Oussalah, M.** (2018). The joint use of sequence features combination and modified weighted SVM for improving daily activity recognition. *Pattern Analysis and Applications*, 21(1), pp. 119–138. doi: <http://dx.doi.org/10.1007/s10044-016-0570-y>
- Affonso, C., Rossi, A. L. D., Vieira, F. H. A. & de Leon Ferreira, A. C. P.** (2017). Deep learning for biological image classification. *Expert Systems with Applications*, 85, pp. 114–122. doi: <http://dx.doi.org/10.1016/j.eswa.2017.05.039>
- Albelwi, S. & Mahmood, A.** (2017). A framework for designing the architectures of deep convolutional neural networks. *Entropy*, 19(6), p. 242. doi: <http://dx.doi.org/10.3390/e19060242>
- Almaslakh, B., AlMuhtadi, J. & Artoli, A.** (2017). An effective deep autoencoder approach for online smartphone-based human activity recognition. *International Journal of Computer Science and Network Security*, 17, p. 160.
- Almaslakh, B., Artoli, A. & Al-Muhtadi, J.** (2018). A Robust Deep Learning Approach for Position-Independent Smartphone-Based Human Activity Recognition. *Sensors*, 18(11), p. 3726. doi: <http://dx.doi.org/10.3390/s18113726>
- Abu Alsheikh, M., Selim, A., Niyato, D., Doyle, L., Lin, S. & Tan, H. P.** (2016). Deep activity recognition models with triaxial accelerometers. In *AAAI Conference on Artificial Intelligence: Workshop on Artificial Intelligence Applied to Assistive Technologies and Smart Environments* (Vol. WS-16-01 – WS-16-15, pp. 8–13). [WS-16-01] Pheonix, United States: AI Access Foundation.
- Anguita, D., Ghio, A., Oneto, L., Parra, X. & Reyes-Ortiz, J. L.** (2013). A public domain dataset for human activity recognition using smartphones. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning* (ESANN).

- Bhattacharya, S. & Lane, N. D.** (2016). From smart to deep: Robust activity recognition on smartwatches using deep learning. In *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pp. 1–6. IEEE. doi: <http://dx.doi.org/10.1109/PERCOMW.2016.7457169>
- Chan, M., EstèVe, D., Fourniols, J. Y., Escriba, C. & Campo, E.** (2012). Smart wearable systems: Current status and future challenges. *Artificial intelligence in medicine*, 56(3), pp. 137–156. doi: <http://dx.doi.org/10.1016/j.artmed.2012.09.003>
- Chowdhury, A. K., Tjondronegoro, D., Chandran, V. & Trost, S. G.** (2017). Physical activity recognition using posterior-adapted class-based fusion of multi-accelerometers data. *IEEE Journal of Biomedical and Health Informatics*, (99), pp. 1–1. doi: <http://dx.doi.org/10.1109/JBHI.2017.2705036>
- Deng, L. & Jaitly, N.** (2016). Deep discriminative and generative models for speech pattern recognition. In *Handbook of pattern recognition and computer vision*, pp. 27–52. doi: http://dx.doi.org/10.1142/9789814656535_0002
- Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., et al.** (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23), pp. e215–e220.
- Graves, A.** (2013). *Generating sequences with recurrent neural networks*. arXiv preprint arXiv: <https://arxiv.org/abs/1308.0850>
- Hammerla, N. Y., Halloran, S. & Plötz, T.** (2016). *Deep, convolutional, and recurrent models for human activity recognition using wearables*. arXiv preprint arXiv: <https://arxiv.org/abs/1604.08880>
- Hinton, G. E., Osindero, S. & Teh, Y. W.** (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), pp. 1527–1554. doi: <http://dx.doi.org/10.1162/neco.2006.18.7.1527>

- Ignatov, A.** (2018). Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. *Applied Soft Computing*, 62, pp. 915–922. doi: <http://dx.doi.org/10.1016/j.asoc.2017.09.027>
- Inoue, M., Inoue, S. & Nishida, T.** (2018). Deep recurrent neural network for mobile human activity recognition with high throughput. *Artificial Life and Robotics*, 23(2), pp. 173–185. doi: <http://dx.doi.org/10.1007/s10015-017-0422-x>
- Kautz, T., Groh, B. H., Hannink, J., Jensen, U., Strubberg, H. & Eskofier, B. M.** (2017). Activity recognition in beach volleyball using a Deep Convolutional Neural Network. *Data Mining and Knowledge Discovery*, 31(6), pp. 1678–1705. doi: <http://dx.doi.org/10.1007/s10618-017-0495-0>
- Khan, M. A. A. H., Roy, N. & Misra, A.** (2018). Scaling human activity recognition via deep learning-based domain adaptation. In *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 1–9. doi: <http://dx.doi.org/10.1109/PERCOM.2018.8444585>
- Kwapisz, J. R., Weiss, G. M. & Moore, S. A.** (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, 12(2), pp. 74–82. doi: <http://dx.doi.org/10.1145/1964897.1964918>
- Lara, O. D. & Labrador, M. A.** (2013). A survey on human activity recognition using wearable sensors. *IEEE communications surveys & tutorials*, 15(3), pp. 1192–1209. doi: <http://dx.doi.org/10.1109/SURV.2012.110112.00192>
- Lee, H., Grosse, R., Ranganath, R. & Ng, A. Y.** (2009). Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In *Proceedings of the 26th annual international conference on machine learning*, pp. 609–616. ACM. doi: <http://dx.doi.org/10.1145/1553374.1553453>
- Leutheuser, H., Schuldhaus, D. & Eskofier, B. M.** (2013). Hierarchical, multi-sensor based classification of daily life activities: comparison with state-of-the-art algorithms using a benchmark dataset. *PloS one*, 8(10), p.e75196. doi: <http://dx.doi.org/10.1371/journal.pone.0075196>

- Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E.** (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, 234, pp. 11–26. doi: <http://dx.doi.org/10.1016/j.neucom.2016.12.038>
- Lockhart, J. W., Weiss, G. M., Xue, J. C., Gallagher, S. T., Grosner, A. B. & Pulickal, T. T.** (2011). Design considerations for the WISDM smart phone-based sensor mining architecture. In *Proceedings of the Fifth International Workshop on Knowledge Discovery from Sensor Data*, pp. 25–33. ACM. doi: <http://dx.doi.org/10.1145/2003653.2003656>
- McDaniel, C. & Quinn, S.** (2018). Developing a Start-to-Finish Pipeline for Accelerometer-Based Activity Recognition Using Long Short-Term Memory Recurrent Neural Networks, pp. 31–40. doi: <http://dx.doi.org/10.25080/Majora-4af1f417-005>
- Münzner, S., Schmidt, P., Reiss, A., Hanselmann, M., Stiefelhagen, R. & Dürichen, R.** (2017). CNN-based sensor fusion techniques for multimodal human activity recognition. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pp. 158–165. ACM. doi: <http://dx.doi.org/10.1145/3123021.3123046>
- Murad, A. & Pyun, J. Y.** (2017). Deep recurrent neural networks for human activity recognition. *Sensors*, 17(11), p. 2556. doi: <http://dx.doi.org/10.3390/s17112556>
- Murahari, V. S. & Plötz, T.** (2018). On attention models for human activity recognition. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers*, pp. 100–103. ACM. doi: <http://dx.doi.org/10.1145/3267242.3267287>
- Namatēvs, I.** (2017). Deep convolutional neural networks: Structure, feature extraction and training. *Information Technology and Management Science*, 20(1), pp. 40–47. doi: <http://dx.doi.org/10.1515/itms-2017-0007>

- Ngo, T. T., Makihara, Y., Nagahara, H., Mukaigawa, Y. & Yagi, Y.** (2014). The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication. *Pattern Recognition*, 47(1), pp. 228–237. doi: <http://dx.doi.org/10.1016/j.patcog.2013.06.028>
- Nweke, H. F., Teh, Y. W., Al-Garadi, M. A. & Alo, U. R.** (2018). Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. *Expert Systems with Applications*, 105, pp. 233–261. doi: <http://dx.doi.org/10.1016/j.eswa.2018.03.056>
- Ordóñez, F. & Roggen, D.** (2016). Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1), p. 115. doi: <http://dx.doi.org/10.3390/s16010115>
- Plötz, T., Hammerla, N. Y. & Olivier, P. L.** (2011). Feature learning for activity recognition in ubiquitous computing. In *Twenty-Second International Joint Conference on Artificial Intelligence*.
- Radu, V., Tong, C., Bhattacharya, S., Lane, N. D., Mascolo, C., Marina, M. K., et al.** (2018). Multimodal deep learning for activity and context recognition. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(4), p. 157. doi: <http://dx.doi.org/10.1145/3161174>
- Ravi, D., Wong, C., Lo, B. & Yang, G. Z.** (2017). A deep learning approach to on-node sensor data analytics for mobile or wearable devices. *IEEE journal of biomedical and health informatics*, 21(1), pp. 56–64. doi: <http://dx.doi.org/10.1109/JBHI.2016.2633287>
- Reiss, A. & Stricker, D.** (2012). Introducing a new benchmarked dataset for activity monitoring. In *2012 16th International Symposium on Wearable Computers*, pp. 108–109. IEEE. doi: <http://dx.doi.org/10.1109/ISWC.2012.13>

- Roggen, D., Calatroni, A., Rossi, M., Holleczeck, T., Förster, K., Tröster, G., et al.** (2010). Collecting complex activity datasets in highly rich networked sensor environments. In *2010 Seventh international conference on networked sensing systems (INSS)*, pp. 233–240. IEEE. doi: <http://dx.doi.org/10.1109/INSS.2010.5573462>
- Ronao, C. A. & Cho, S. B.** (2015). Evaluation of deep convolutional neural network architectures for human activity recognition with smartphone sensors. In *proceeding of the KIISE Korea Computer Congress*, pp. 858–860.
- Stisen, A., Blunck, H., Bhattacharya, S., Prentow, T. S., Kjærgaard, M. B., Dey, A., et al.** (2015). Smart devices are different: Assessing and mitigating mobile sensing heterogeneities for activity recognition. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, pp. 127–140. ACM.
- Sztyler, T. & Stuckenschmidt, H.** (2016). On-body localization of wearable devices: An investigation of position-aware activity recognition. In *2016 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 1–9. IEEE.
- Twomey, N., Diethe, T., Fafoutis, X., Elsts, A., McConville, R., Flach, P., et al.** (2018). A comprehensive study of activity recognition using accelerometers. In *Informatics*, 5(2), p. 27. Multidisciplinary Digital Publishing Institute.
- Wang, J., Chen, Y., Hao, S., Peng, X. & Hu, L.** (2019). Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119, pp. 3–11. doi: <http://dx.doi.org/10.1016/j.patrec.2018.02.010>
- Xi, R., Li, M., Hou, M., Fu, M., Qu, H., Liu, D., et al.** (2018). Deep dilation on multimodality time series for human activity recognition. *IEEE Access*, 6, pp. 53381–53396. doi: <http://dx.doi.org/10.1109/ACCESS.2018.2870841>

- Yang, J., Nguyen, M. N., San, P. P., Li, X. L. & Krishnaswamy, S.** (2015). Deep convolutional neural networks on multichannel time series for human activity recognition. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.
- Yao, S., Zhao, Y., Hu, S. & Abdelzaher, T.** (2018). QualityDeepSense: Quality-Aware Deep Learning Framework for Internet of Things Applications with Sensor-Temporal Attention. In *Proceedings of the 2nd International Workshop on Embedded and Mobile Deep Learning*, pp. 42–47. ACM. doi: <http://dx.doi.org/10.1145/3212725.3212729>
- Zeng, M., Nguyen, L. T., Yu, B., Mengshoel, O. J., Zhu, J., Wu, P., et al.** (2014). Convolutional neural networks for human activity recognition using mobile sensors. In *6th International Conference on Mobile Computing, Applications and Services*, pp. 197–205. IEEE. doi: <http://dx.doi.org/10.4108/icst.mobicase.2014.257786>
- Zhu, Q., Chen, Z. & Yeng, C. S.** (2018). A Novel Semi-supervised Deep Learning Method for Human Activity Recognition. *IEEE Transactions on Industrial Informatics*. doi: <http://dx.doi.org/10.1109/TII.2018.2889315>