Nurse care activity recognition challenge: summary and results

Conference Paper · September 2019

DOI: 10.1145/3341162.3345577

11 authors, including:

Paula Lago
Kyushu Institute of Technology

François Charpillet
National Institute for Research in Computer Science and Control

Sozo Inoue
Kyushu Institute of Technology

Sayeda Shamma Alia
Kyushu Institute of Technology

Nurse Care Activity Recognition Challenge: Summary and Results

Paula Lago*
Sayeda Shamma Alia*
(paula|alia)@sozolab.jp
Kyushu Institute of Technology
Kitakyushu, Fukuoka, Japan

Shingo Takeda
takeda@sozolab.jp
Kyushu Institute of Technology
Kitakyushu, Fukuoka, Japan

Tittaya Mairittha
fon@sozolab.jp
Kyushu Institute of Technology
Kitakyushu, Fukuoka, Japan

Nattaya Mairittha
fah@sozolab.jp
Kyushu Institute of Technology
Kitakyushu, Fukuoka, Japan

Farina Faiz
farina@sozolab.jp
Kyushu Institute of Technology
Kitakyushu, Fukuoka, Japan

Yusuke Nishimura
nishimura@sozolab.jp
Kyushu Institute of Technology
Kitakyushu, Fukuoka, Japan

Kohei Adachi
adachi@sozolab.jp
Kyushu Institute of Technology
Kitakyushu, Fukuoka, Japan

Tsuyoshi Okita
tsuyoshi.okita@gmail.com
Kyushu Institute of Technology
Kitakyushu, Fukuoka, Japan

François CHARPILLET
francois.charpillet@inria.fr
Université de Lorraine, CNRS, Inria,
LORIA, F-54000 Nancy, France

Sozo Inoue
sozo@brain.kyutech.ac.jp
Kyushu Institute of Technology /
RIKEN AIP
Kitakyushu, Fukuoka, Japan

ABSTRACT

Although activity recognition has been studied for a long time now, research and applications have focused on physical activity recognition. Even if many application domains require the recognition of more complex activities, research on such activities has attracted less attention. One reason for this gap is the lack of datasets to evaluate and compare different methods. To promote research in such scenarios, we organized the *Open Lab Nursing Activity Recognition Challenge* focusing on the recognition of complex activities related to the nursing domain. Nursing domain is one of the domains that can benefit enormously from activity recognition but has not been researched due to lack of datasets. The competition used the CARECOM Nurse Care Activity Dataset, featuring 7 activities performed by 8 subjects in a controlled environment with accelerometer sensors, motion capture and indoor location sensor. In this paper, we summarize the results of the competition.

*Both authors contributed equally to this research.

UbiComp/ISWC '19 Adjunct, September 9–13, 2019, London, United Kingdom

ACM ISBN 978-1-4503-6869-8/19/09…$15.00
https://doi.org/10.1145/3341162.3345577

CCS CONCEPTS

• Computing methodologies → Machine learning; **• Applied computing → Health care information systems**.

KEYWORDS

datasets, activity recognition, motion capture, accelerometer, nurse care

ACM Reference Format:
Paula Lago, Sayeda Shamma Alia, Shingo Takeda, Tittaya Mairittha, Nattaya Mairittha, Farina Faiz, Yusuke Nishimura, Kohei Adachi, Tsuyoshi Okita, François CHARPILLET, and Sozo Inoue. 2019. Nurse Care Activity Recognition Challenge: Summary and Results. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct), September 9–13, 2019, London, United Kingdom.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3341162.3345577

1 INTRODUCTION

Although activity recognition has been studied for a long time now, research and applications have been mainly focused on physical activity recognition [13][1] [2]. Many commercial products can now recognize when we walk, cycle or run, count steps and give a general overview of how active we have been on a day. However, research in applications to domains like healthcare, although advertised and highly regarded as important [5], is not very mature. One reason is that the activities are more complex and hard to analyze. Research in these areas is even more difficult due to the lack of publicly data that can be used by researchers to propose and compare the performance of different methods [8, 10].

Figure 1: CARECOM Nurse Care Challenge

Following the successful challenge for locomotion activities [6], we organized the Nurse Care Activity Recognition Challenge as part of the HASCA Workshop. With this challenge, we aimed at bridging the gap for nurse care activity recognition. Complex activity recognition in health care applications has been focusing on recognition of patient activities or at-home activities [4, 12, 16] and overlooked nurses and caregivers. Recognition of nurses activities can have many applications like automatic record creation to reduce documentation time, checking compliance with care routines for a given patient and identification of risk activities that require special, for example hand washing after taking blood samples.

Nurse care activity recognition is challenging because, unlike other activity settings in which the user is doing an activity, nurses usually perform some activity *to* a patient. For example, they give a drink to the patient instead of just drinking. This feature introduces challenges that are not studied in other activity recognition applications. For example, the activities can be performed differently by the same nurse depending on the patient receiving the care. In this case, intra-class variability depends not only on the subject, as in other domains, but also on the receiving patient.

To study the feasibility and challenges of implementing activity recognition using movement and position for this domain, we collected a a multimodal dataset in a controlled environment [9] for this challenge. In this paper, we present the collection experiment details, a summary of the collected data and we summarize the results of the activity recognition challenge.

2 NURSE CARE ACTIVITY RECOGNITION DATASET

In this section, we describe the nursing activities selected for the experiment, describe the experimental settings under which the data was collected and summarize the dataset. Detailed information about the dataset can be found in [15]

2.1 Nursing activities in the Dataset

The CARECOM Nurse Care Activity dataset [9] contains data about 6 nursing activities, related to hospital patient care. In Table1, we describe each activity.

2.2 Experimental settings

The dataset was recorded in the Smart Life Care Unit of the Kyushu Institute of Technology in Japan (Figure 1). In this experiment, 8 subjects participated. Each participant performed 5 repetitions of each activity, yielding about 240 activity sequences and 407 recorded minutes. We recorded each activity using accelerometers, meditag data (bluetooth based localization and pressure sensor), and motion capture data including 25 body markers. Using this dataset, we organized the "Nurse Care Activity Recognition Challenge". Data collection was done in three days total. Each day, two or three professional nurses came to the laboratory and were asked to perform the six activities of Table 1. All nurses agreed to the sensor data collection for this experiment.

For each activity, instructions about the procedure were given, so that the activity comprises all necessary steps. One person acted as patient so the activity is performed in naturalistic way. Each nurse performed the activity in her/his preferred way, pace and order, where applicable. The room was equipped with hospital bed, desk,

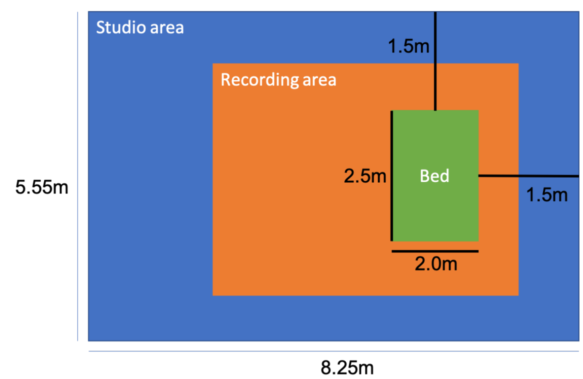


Figure 2: Recording Studio area plan

and wheeled cart for the nurse to take the necessary equipment. All instruments like drip, gauze, diaper, etc. were provided and used by the nurses during data collection. A plan of the studio area where the recordings were done is shown in Figure 2 with corresponding measurements.

2.3 Sensor modalities

We collected data from motion capture, meditag and accelerometer sensors. The description of each data modality is given below.

Motion Capture: We use the motion capture system form Motion Analysis Company[1]. The setup used consists of 29 body markers located as in Figure 3. The markers are tracked using 16 infrared cameras. The three dimensional position of each marker is recorded with a frequency of 100Hz. The markers may be labeled incorrectly in some cases due to the complex setting.

Meditag Sensor: The meditag is a bluetooth based in-door localization sensor. Four receivers were installed in the room and the nurse carried a bluetooth beacon on her right chest pocket. It measures two dimensional position (x and y) of the beacon in meters and air pressure in mHg with a sampling rate of approximately 20 Hz.

Accelerometer sensor: We used a Freetel Priori 3 smartphone carried in the right chest pocket of the nurse in upright position.

[1] http://motionanalysis.com/movement-analysis/

Table 1: Description of the activities

Id	Activity Name	Activity Purpose
2	Vital signs measurements	Confirm the signs of human life by measuring respiration, pulse, temperature and blood pressure. In the experiment, these four signs were measured in a sequence for this activity.
3	Blood Collection	Collect blood from the body to know the progress of diagnosis and treatment of disease. In the experiment, blood was not extracted, instead acting of the steps was done. This might have influenced the duration of the activity.
4	Blood Glucose Measurement	Measure blood sugar to control it. This procedure is indispensable for diseases that require sugar control such as diabetes.
6	Indwelling drip retention and connection	The purpose of indwelling placement is: 1. Refill the body with insufficient fluid. 2. Correct and maintain electrolyte balance. 3. Prosperity for patients who can not take orally. Provide support. 4. Administer the necessary medications.Etc.
9	Oral care	Keep the mouth clean to keep the health of the whole body as well as in the mouth. Other purposes are: 1. Promote the secretion of saliva. 2. Prevent infection and fever. 3. Prevent dementia. 4. Prevent aspiration pneumonia. 5. Prevent deterioration of oral function
12	Diaper exchange and cleaning of area	Wash to keep the genital area clean for patients who can not take a bath or for patients wearing diapers. If a bladder catheter is indwelling, it is done as prevention of retrograde urinary tract infection.

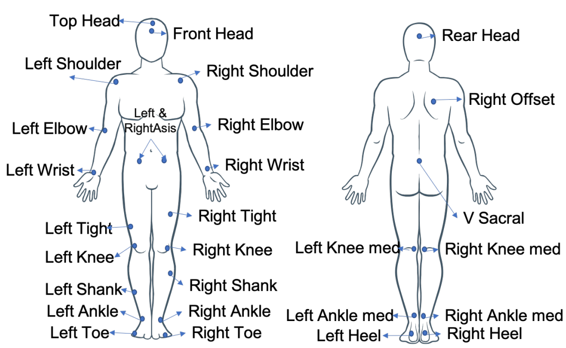


Figure 3: Motion capture markers

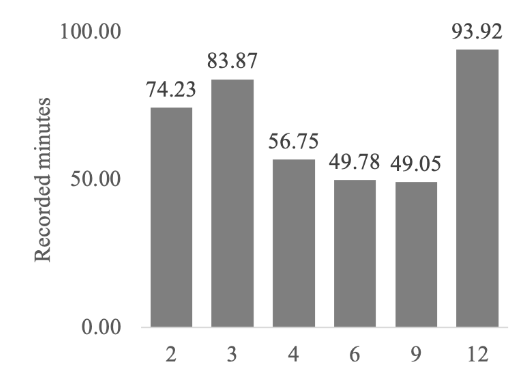


Figure 4: Recorded minutes of each activity. Different activity duration raise an imbalance in time distribution even if the number of samples for each activity is the same.

The sampling rate is approximately 4 Hz. This sensor measures three dimensional inertial acceleration in m/s^2.

3 DATA DESCRIPTION

In this section, we summarize the data collected as a result of the experiment. In total, we recorded 240 activity sequences and 407 minutes. The distribution of recorded minutes for each activity is shown in Figure 4. Even if we collected 40 samples (5 per user) of each activity, the duration of the activities is not the same, so the time distribution is not equal. We show the distribution of the duration of each activity in Figure 5. From the figure, we can see that the duration of all activities is between 1 and 4 minutes in average. Notice that the activities with longer duration (Activity 3 and 12) also have the highest number of recorded minutes.

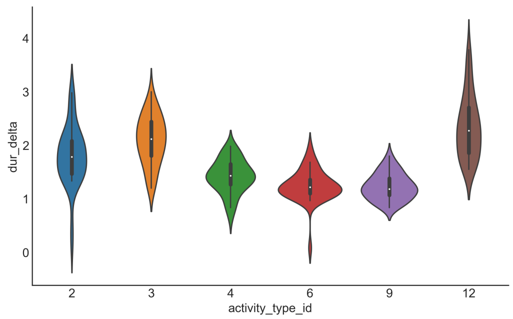


Figure 5: Distribution of the duration of each activity class.

However, the recorded minutes for each nurse is almost equally distributed (Figure 6). The difference in recording times is due to variations in activity performance, some nurses taking more time than others, and not due to differences in number of samples per activity. The number of observations by each sensor is shown in

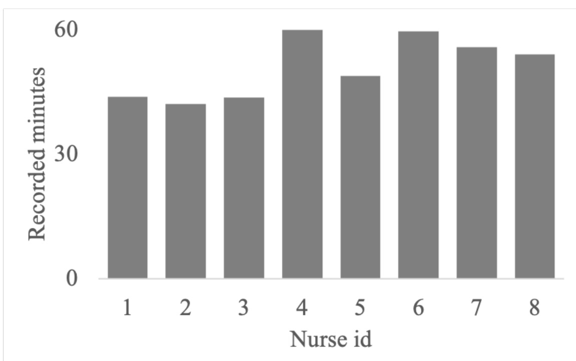


Figure 6: Recorded minutes by user.

Figure 7. Due to the big sampling rate differences, there are more measurements for the motion capture than for other sensors.

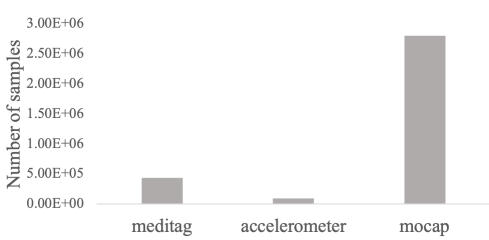


Figure 7: Number of observations by each sensor source.

For this challenge, each activity recording was segmented into 1-minute segments. The goal of the challenge was to classify each segment's activity. Segments were given random identifiers, so their sequence is unknown. Due to the different activity durations, some segments may be shorter than 1-minute, although those shorter than 20 seconds were removed from the data.

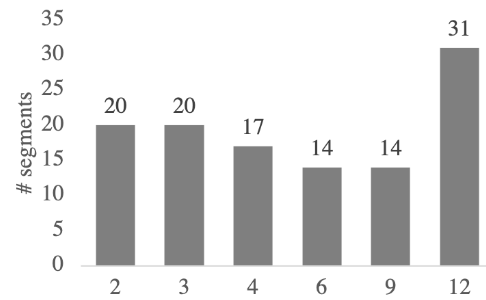


Figure 8: Number of 1-minute segments per class in the test data

The dataset was divided into two parts: one labeled dataset which contained data from 6 users and one unlabeled dataset containing data from the remaining 2 users. The unlabeled dataset was used as test data for the challenge, and the correct labels are used to evaluate the accuracy. The main challenge is to create a user-independent model to recognize the activities using any combination of the sensor data provided. The number of segments per class in the test dataset is shown in Figure 8

4 CHALLENGE RESULTS

Four teams participated in the challenge: *Team IITDU* [11], *Team TDU-DSML* [3], *Team Dark Shadow* [7] and *Team DataDigger* [14]. In addition, we prepared two baseline methods: one using acceleration data and Random Forest and the other using all modalities and a deep learning approach. Each team had one month to prepare their submission (May 24 - June 30). In this Section, we summarize and compare the approaches and results of the challenge for all teams and both baselines. An overview of all approaches is shown in Table 2

We now summarize the approaches of the top two teams. *Team IITDU* [11] used motion capture and meditag sensors. From motion capture they extracted three types of features: trajectory for each joint, joint angles, joint distances and features to represent hand movement. From meditag sensor they use mean, minimum, maximum and standard deviation of both axis of the position (X and Y). They also extract the block where the person stays the longest during the segment and the distance traveled in both directions. Finally they use the air pressure. *Team TDU-DSML* [3] used a graph to represent the skeleton in a spatio-temporal graph. Close points in the skeleton are considered to be neighbors in this graph.

We now analyze results by class for each submission. We omit the analysis for team *Data Digger* [14] due to the low accuracy. The low accuracy is a result of the classifier over-fitting to class 2 and predicting almost all segments as belonging to this class. Precision, Recall and F1-Score by class for each team are shown in Figures 9, 10 and 11 respectively.

From the results we highlight that motion capture was the most commonly used data source. Only one team didn't use motion capture sensor, with low accuracy in the test data. However, the baseline using only acceleration data achieved a better performance.

Table 2: Summary of submissions made to challenge ordered by accuracy in the test dataset

Team	Sensors used	Method	Training Accuracy	Test Accuracy
Team IITDU[11]	Motion Capture and Meditag data	KNN	66.11%	80.2%
TDU-DSML[3]	Motion Capture	Spatio-temporal Graph Convolutional Network	57%	64.6%
Baseline DL	All modalities	CNN	100%	46.5%
Baseline ML	Acceleration	Random Forest	60%	43.1%
Dark Shadow[7]	Motion Capture and Meditag	Gated Recurrent Unit	66%	29.3%
Data Digger[14]	Acceleration	Random Forest	82.86%	18.1%

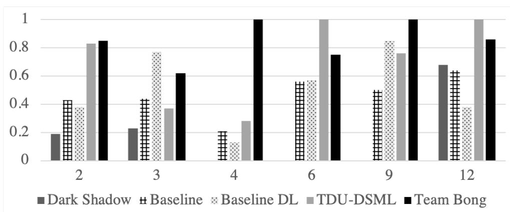


Figure 9: Precision by class of each team on the test data

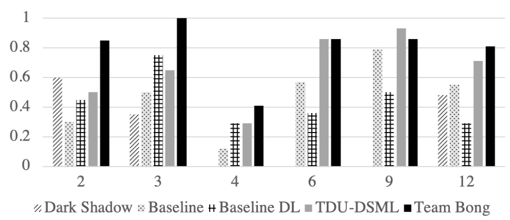


Figure 10: Recall by class for each team on the test data

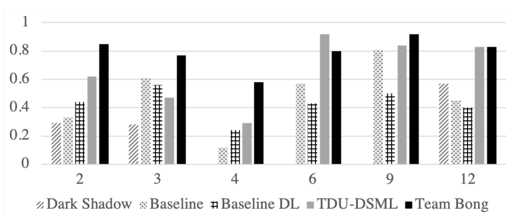
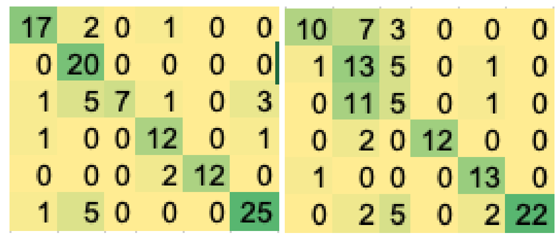


Figure 11: F1 Score by class for each team on the test data



17	2	0	1	0	0	10	7	3	0	0	0
0	20	0	0	0	0	1	13	5	0	1	0
1	5	7	1	0	3	0	11	5	0	1	0
1	0	0	12	0	1	0	2	0	12	0	0
0	0	0	2	12	0	1	0	0	0	13	0
1	5	0	0	0	25	0	2	5	0	2	22

Figure 12: Confusion Matrix of the top two teams: Team IITDU (left) and Team TDU-DSML (right)

Clearly, using motion capture data with features related to movement can increase the accuracy.

The top two teams, teams *IITDU* [11] and *TDU-DSML* [3] used motion capture, however, the winning team combined it with meditag sensor. The difference in their results is mainly due to the low precision and recall for activity 4 (Blood Glucose measurement) that *TDU-DSML* [3] team obtained. In fact, this activity was confused with 'Blood Collection' as can be seen in the confusion matrix (Figure 12). Both activities are very similar activities, so this confusion is not surprising. Team *IITDU* [11] had a better discrimination, probably due to the combination with location data. The location data might be useful due to the different durations of both activities. While blood collection is longer activity, the glucose measurement is very short and its location may be different because of this.

Another interesting insight comes from comparing team *IITDU* [11] and team *Dark Shadow* [7]. Both teams used the same modalities similar features for both modalities, yet obtained different results. Team *Dark Shadow* [7] used a deep learning approach while team *IITDU* [11] used a nearest neighbor approach. We believe that due to the amount of data, the deep learning approach didn't obtain a good performance. Team *TDU-DSML* [3] also used a deep learning approach with good results. They divided the segments into shorter windows to create more samples and overcome this limitation.

Finally, we see that both *Team Data Digger* [14] and *Baseline ML* used Random Forest and only acceleration data, but obtained very different results (Table 2). One reason behind this is that, *Team Data Digger* used the raw data for their classification while *Baseline ML* used basic statistical features. We can observe that raw data can not be much of help during classifying these complex activities. It can be seen that stronger features and possibly information about movement of other body parts can improve this result significantly.

With respect to programming language, all of the teams used Python for implementing their module. One reason can be it has becoming more popular both for Machine learning and Deep Learning. Also for its easy syntax and English-like commands.

From the results, we see that Deep Learning approaches had a disadvantage, mainly due to the small amount of data available.

5 CONCLUSIONS

In this paper, we have summarized the submissions received to the Nurse Care Activity Recognition Challenge. One of the main goals of the challenge and of the experiment was to prove the feasibility of recognizing complex nurse care activities based on movement. The results of the challenge confirm that, among the six activities chosen, an 80% accuracy can be achieved. However, such results use motion capture data, that is whole body data. Using only accelerometer data and simple features, the accuracy achieved was 43%. The dataset was collected in a controlled-environment and the activities were clearly segmented. However, we used 1-minute segments and removed the order of these segments, so the recognition is based only on movement features. In real-life this can be completed with knowledge about the duration of the activities and their timing. The most challenging aspects were finding a good representation for the data and finding robust algorithms for inter-person variability. Additional experiments and research are required to see if the results hold in real field data and how to transfer knowledge to sensors available in real field. Nevertheless, this is a first step in nurse care activity recognition. As future work, we will continue to create datasets from both the open lab and real field to advance research in this area.

ACKNOWLEDGEMENTS

The dataset and this research were done in collaboration with CARECOM Company.

REFERENCES

[1] Stefano Abbate, Marco Avvenuti, Francesco Bonatesta, Guglielmo Cola, Paolo Corsini, and Alessio Vecchio. 2012. A smartphone-based fall detection system. *Pervasive and Mobile Computing* 8, 6 (2012), 883 – 899. https://doi.org/10.1016/j.pmcj.2012.08.003 Special Issue on Pervasive Healthcare.

[2] Ling Bao and Stephen S. Intille. 2004. Activity Recognition from User-Annotated Acceleration Data. In *Pervasive Computing. Pervasive 2004. Lecture Notes in Computer Science, vol 3001.*, A. Ferscha and F. Mattern (Eds.). Springer, Berlin, Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-24646-6_1 arXiv:9780201398298

[3] Xin Cao, Wataru Kudo, Chihiro Ito, Masaki Shuzo, and Eisaku Maeda. 2019. Activity Recognition with ST-GCN for 3D Motion Data. In *HASCA '19 (Ubicomp '19)*.

[4] Macarena Espinilla, Javier Medina, and Chris Nugent. 2018. UCAmI Cup. Analyzing the UJA Human Activity Recognition Dataset of Activities of Daily Living. *Proceedings* 2, 19 (2018). https://doi.org/10.3390/proceedings2191267

[5] J. Favela. 2013. Behavior-Aware Computing: Applications and Challenges. *IEEE Pervasive Computing* 12, 3 (July 2013), 14–17. https://doi.org/10.1109/MPRV.2013.44

[6] H. Gjoreski, M. Ciliberto, L. Wang, F. J. Ordonez Morales, S. Mekki, S. Valentin, and D. Roggen. 2018. The University of Sussex-Huawei Locomotion and Transportation Dataset for Multimodal Analytics With Mobile Devices. *IEEE Access* 6 (2018), 42592–42604. https://doi.org/10.1109/ACCESS.2018.2858933

[7] Nazmul Haque, Mahir Mahbub, Hasan Tarek, Lutfun Nahar Lota, and Amin Assan Ali. 2019. Nurse Care Activity Recognition: A GRU-based approach with attention mechanism. In *HASCA '19 (Ubicomp '19)*.

[8] Sozo Inoue, Paula Lago, Tahera Hossein, Tittaya Mairittha, and Nattaya Mairittha. 2019. Integrating Activity Recognition and Nursing Care Records: the System, Deployment, and a Verification Study. *IMWUT* (2019). In Press.

[9] Sozo Inoue, Paula Lago, Shingo Takeda, AliaShamma, Farina Faiz, Nattaya Mairittha, and Tittaya Mairittha. 2019. Nurse Care Activity Recognition Challenge. https://doi.org/10.21227/2cvj-bs21

[10] Sozo Inoue, Naonori Ueda, Yasunobu Nohara, and Naoki Nakashima. 2015. Mobile Activity Recognition for a Whole Day: Recognizing Real Nursing Activities with Big Dataset. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 1269–1280. https://doi.org/10.1145/2750858.2807533

[11] Eusha Kadir, Pritom Saha Akash, Sadia Sharmin, Amin Ahsan Ali, and Mohammad Shoyaib. 2019. Can a Simple Approach Identify Complex Nurse Care Activity?. In *HASCA '19 (Ubicomp '19)*.

[12] Paula Lago, Fréderic Lang, Claudia Roncancio, Claudia Jiménez-Guarín, Radu Mateescu, and Nicolas Bonnefond. 2017. The ContextAct@A4H Real-Life Dataset of Daily-Living Activities. In *Modeling and Using Context*, Patrick Brézillon, Roy Turner, and Carlo Penco (Eds.). Springer International Publishing, Cham, 175–188.

[13] Jonathan Lester, Tanzeem Choudhury, and Gaetano Borriello. 2006. A Practical Approach to Recognizing Physical Activities. In *Pervasive Computing*, Kenneth P. Fishkin, Bernt Schiele, Paddy Nixon, and Aaron Quigley (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 1–16.

[14] Nazmus Sakib Patwary, Protap Kumar Saha, and Ifthakhar Ahmed. 2019. Nurse Care Activity Recognition Challenge Using A Supervised Methodology. In *HASCA '19 (Ubicomp '19)*.

[15] Lago Paula, Takeda Shingo, Sayeda Shamma Alia, Mairittha Tittaya, Mairittha Nattaya, and Inoue Sozo. 2019. *Open Lab Nursing Activity Recognition Challenge.* Technical Report 12. Kyushu Institute of Technology. https://ipsj.ixsq.nii.ac.jp/ej/index.php?active_action=repository_view_main_item_detail&page_id=13&block_id=8&item_id=196715&item_no=1

[16] M. Stikic, D. Larlus, S. Ebert, and B. Schiele. 2011. Weakly Supervised Recognition of Daily Life Activities with Wearable Sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 12 (Dec 2011), 2521–2537. https://doi.org/10.1109/TPAMI.2011.36