scientific data

Check for updates

OPEN A multi-sensor dataset with annotated activities of daily living DATA DESCRIPTOR recorded in a residential setting

Emma L. Tonkin¹, Michael Holmes¹, Hao Song¹, Niall Twomey¹, Tom Diethe^{1,2}, Meelis Kull³, Miquel Perello Nieto¹, Massimo Camplani², Sion Hannuna¹, Xenofon Fafoutis⁶, Ni Zhu⁵, Przemysław R. Woznowski⁶, Gregory J. L. Tourte¹, Raúl Santos-Rodríguez 1, Peter A. Flach 1 & Ian Craddock 1

SPHERE is a large multidisciplinary project to research and develop a sensor network to facilitate home healthcare by activity monitoring, specifically towards activities of daily living. It aims to use the latest technologies in low powered sensors, internet of things, machine learning and automated decision making to provide benefits to patients and clinicians. This dataset comprises data collected from a SPHERE sensor network deployment during a set of experiments conducted in the 'SPHERE House' in Bristol, UK, during 2016, including video tracking, accelerometer and environmental sensor data obtained by volunteers undertaking both scripted and non-scripted activities of daily living in a domestic residence. Trained annotators provided ground-truth labels annotating posture, ambulation, activity and location. This dataset is a valuable resource both within and outside the machine learning community, particularly in developing and evaluating algorithms for identifying activities of daily living from multi-modal sensor data in real-world environments. A subset of this dataset was released as a machine learning competition in association with the European Conference on Machine Learning (ECML-PKDD 2016).

Background & summary

Obesity, depression, stroke, falls, cardiovascular and musculoskeletal diseases are some of the most significant health issues and fastest-rising categories of healthcare costs. The financial expenditure associated with these conditions is widely regarded as ineffective and unsustainable. The impact on quality of life is also felt by millions of people around the world each day.

Smart technologies can unobtrusively quantify activities of daily living, and these can provide long-term behavioural patterns that are objective, insightful measures for clinical professionals and caregivers. The EPSRC-funded Sensor Platform for HEalthcare in Residential Environment (SPHERE) Interdisciplinary Research Collaboration (IRC)¹⁻³ project has designed a multi-modal sensor system driven by data analytics requirements. The system is currently under test in over 50 homes across Bristol (UK). The data sets collected will be made available to researchers in a variety of communities. This paper describes a particular dataset focusing on the task of activity recognition using machine learning approaches.

To collect this dataset, the overall sensing system includes the following three sensing modalities:

- wrist-worn accelerometer;
- RGB-D cameras (video with depth information); and
- . passive environmental sensors.

With these sensors, the system is capable of capturing information related to most indoor daily living activities. It is then possible to learn patterns of behaviour and track the deterioration/recovery of persons that suffer or

¹University of Bristol, Bristol, UK. ²Amazon, Bellevue, USA. ³University of Tartu, Tartu, Estonia. ⁴Technical University of Denmark, Lyngby, Denmark. ⁵China Mobile International, Beijing, China. ⁶BJSS, Manchester, UK. ^{\inference}e-mail: e.l.tonkin(a) bristol.ac.uk

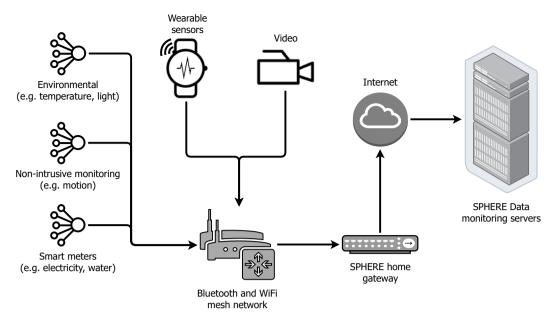


Fig. 1 The SPHERE sensor network. Note that the environmental and smart meter data are excluded from this dataset, as they are not relevant to the scripted tasks.

recover from various medical conditions. To achieve this design, sensor data in this dataset is also accompanied by annotations of domestic human behaviour to facilitate classification of:

- activities of daily living (tasks such as meal preparation, watching television);
- posture/ambulation (e.g., walking, sitting, transitioning); and
- room-level indoor location.

On the machine learning side, these data have value in training, optimising and evaluating activity recognition and localisation classification algorithms, as well as generic time series models. The data can also be used as a source dataset for exploring sensor fusion methods and multi-modal learning approaches. For example, these data have been used in *The SPHERE Challenge*⁴ competition by students and researchers to model and classify activities of daily living, and the wining approach is documented by Liu *et al.*⁵. In a related paper⁶, the authors propose an unsupervised approach to learn the sensor layout for a multi-modal sensor network within a residential environment. The overall ML system construction and applications are discussed for the SPHERE project⁷. On the healthcare side, these data support further development of methods for Smart Home and e-health applications. Recent research shows that the classification of natural human movement can indicate the recovery level from hip and knee replacement surgery⁸.

The remainder of this paper is structured as follows: Methods describes the methods used to collect these data. Data Records. provides detail on the data records, structures and format, Technical Validation describes the method and results of technical validation and finally, Usage Notes highlights notes for use of these data. Code availability includes details on code availability.

Methods

The following section describes recruitment, data collection, data extraction and processing for this dataset, as well as the validation processes undergone for label annotations.

Recruitment. Ten healthy volunteer participants (two female, eight male) were recruited from the student and faculty community at the University of Bristol to participate in a data collection activity. Eight participants were in the 18–29 category, two in 30–39. Participants filled out a health questionnaire prior to participation, and no health issues were recorded.

The data was collected in the SPHERE house (Fig. 2), a test-bed property in Bristol, UK, which has been fitted with the SPHERE sensor network (Fig. 1). Ethical approval was secured from the University of Bristol's ethics committee to conduct data collection, and informed consent was obtained from all volunteers. Volunteers were not compensated for participation.

Data collection. Data were collected within the SPHERE House. All recording sessions occurred in timeslots between 10 am and 4 pm over ten working days in the late summer (Aug–Oct). Overall length of session ranged from 23 min 57 s to 36 min 46 s, mean 00:29:20, stdev 00:02:59. Time of day of data collection has been excluded from the dataset as a participant privacy measure. Figure 2 show the floor plan of the ground and first floors of the

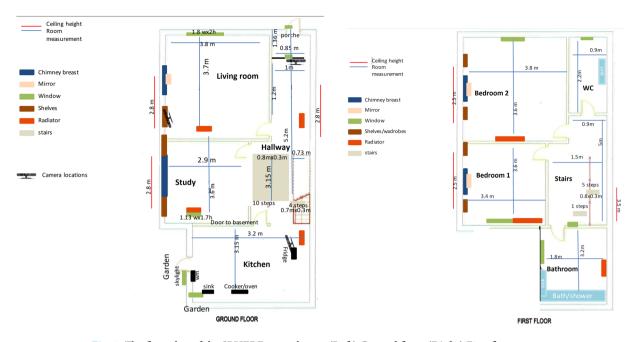


Fig. 2 The floorplan of the SPHERE smart home. (Left) Ground floor. (Right) First floor.

smart environment respectively. The SPHERE House is designed to be a unique environment with controllable experimental conditions. Valuable smart-home data can be collected on multiple participants without the prohibitive time and cost burdens associated with installation and removal of a large number of sensors in multiple residential locations.

Figure 1 shows the SPHERE sensor network with constellations of sensors sending measurements back to the Home Gateway. The SPHERE sensor network incorporates multiple modalities of data: environmental, video and wearable-sensor streams. Data from sensors were streamed via MQTT to a Mongo database on the Home Gateway. All sensors are synchronised with network time protocol (NTP). Following experimentation, data from the home gateway database was exported for annotation and analysis.

As introduced above, the dataset captures scripted activities performed by multiple participants – one participant at a time – within the SPHERE house. The experimental script for activities can be found at https://raw.githubusercontent.com/IRC-SPHERE/sphere-challenge/master/documents/data_collection_script.pdf. The script includes behaviours and activities such as room and floor transitions, posture and ambulation changes, interaction with domestic appliances and simulated activities of daily living. Each participant undertook activities individually with the supervision of a SPHERE project researcher. While the order of activities was enforced, the pace is not scheduled and could be different for every participant.

Data extraction and processing. The following subsections describe the sensing modalities that are found in the smart home and explain the overall data extraction and processing steps.

Accelerometers. Participants wore the SPHERE wearable on their dominant wrist (the device used was the first generation SPW-1, as described by Fafoutis *et al.*⁹), attached using a strap. The SPHERE wearable is an acceleration-based activity sensor. The device is equipped with two ADXL362 accelerometers¹⁰ and wirelessly transmits data using the Bluetooth Low Energy (BLE) standard to several access points (receivers) positioned within the house¹¹. The outputs of these sensors are a continuous numerical stream of the accelerometer readings (units of *g*, i.e., approximately 9.81 m s⁻²). Accompanying the accelerometer readings are the received signal strength indications (RSSI) that were recorded by each access point (in units of dBm), and these data will be informative for indoor localisation. The accelerometers record data at 20 Hz (12-bit resolution), and the accelerometer ranges are between $\pm 8g$. RSSI values are also recorded at 20 Hz, and values are no lower than -110 dBm.

For privacy concerns, all the residential environments are assumed to be hard-to-access, hence should require minimum to no maintenance for long-term operations. Therefore, the system is optimised for low energy consumption. To that end, the communication between the wearable sensor and the smart house is performed via undirected connectionless BLE advertisements. Although data reliability can be addressed at the receiver¹¹, this communication approach does not provide delivery guarantees and, thus, there may be missing packets from the data. Recent accelerometer work done by SPHERE researchers on activity recognition with accelerometers includes^{12,13}. Data from the SPHERE wearable have also been used for the validation of a privacy-preserving algorithm for wearable embedded systems¹⁴.

RGB-D cameras. Video recordings were taken using ASUS Xtion PRO RGB-Depth (RGB-D) cameras. Automatic detection of humans was performed using the OpenNI library, and false-positive detections were

Туре	Labels
Locations	bath, bed1, bed2, hall, kitchen, living, stairs, study, toilet
Ambulation	a_ascend, a_descend, a_jump, a_loadwalk, a_walk
Posture	p_bent, p_kneel, p_lie, p_sit, p_squat, p_stand
Transition	$t_bend, t_kneel_stand, t_lie_sit, t_sit_lie, t_sit_stand, t_stand_kneel, t_stand_sit, t_straighten, t_turn$

Table 1. Annotation labels for locations, activities, postures and transitions.

.....

manually removed by visual inspection. Three RGB-D cameras are installed in the SPHERE house, and these are located in the living room, hallway, and the kitchen. No cameras are located elsewhere in the house due to privacy considerations.

In order to preserve the anonymity of the participants, the raw video data are not stored. Instead, several data points are stored describing the detected individual: the coordinates of the 2D bounding box enclosing them, the 2D centre of mass, the 3D bounding box enclosing the individual and the 3D centre of mass are collected into the dataset. These are generated using OpenNI and OpenCV on the basis of RGB-D data. The limitations of this data include occlusion (e.g., if a participant is partially occluded by a wall or other object, the bounding box will be truncated) and a risk of false-negative or false-positive detection. Imperfect detection of object boundaries is also a possibility. Occlusion was minimised by the experimental script and by a careful choice of recording angles. Data was screened for missing modalities or false positives.

The units of 2D coordinates are in pixels (i.e., number of pixels down and right from the upper left-hand corner) from an image of size 640×480 pixels. The coordinate system of the 3D data is axis-aligned with the 2D bounding box, with an extra dimension that projects from the central position of the video frames. The first two dimensions specify the vertical and horizontal displacement of a point from the central vector (in millimetres), and the final dimension specifies the projection of the object along the central vector (again, in millimetres).

RGB-D tracking data is very valuable in indoor environments as it can facilitate improvement in accuracy for fundamental computer vision tasks such as tracking¹⁵ while also enabling specific analysis at higher levels. RGB-D data collected in the SPHERE house has been used for example for specific action recognition¹⁶ and action quality estimation¹⁷.

Environmental sensors. The environmental sensing nodes are built on development platforms (Libelium, with CE marking), powered by batteries or/and 5 V DC converted from mains. Passive Infra-Red (PIR) sensors are employed to detect presence in the data. Values of 1 indicate that motion was detected, whereas values of 0 mean that no motion was detected. Several methods that deal with environmental sensor data are detailed in the following^{6,18}.

Annotation. The annotation processes use a simplified taxonomy based on the SPHERE annotation ontology¹⁹. This ontology was developed with reference to various existing sources, notably the taxonomy published by BoxLab. A team of 12 annotators was recruited and trained to annotate the set of locations and activities (Table 1). Every data sequence was annotated by either 2 or 3 annotators to avoid bias in the ground truth. Four sessions were annotated by three annotators. As the pairwise Cohen's kappa calculated remained reasonably stable between annotators in most cases, the decision was taken to reduce the number of annotators to two for the remaining sessions (see Table 3). To support the annotation process, a head-mounted camera (Panasonic HX-A500E-K 4 K Wearable Action Camera Camcorder) recorded 4 K video at 25 FPS to an SD-card. This data is only used to facilitate the annotation process and is removed after the annotation and therefore is not shared in this dataset. Synchronisation between the NTP clock and the head-mounted camera was achieved by focusing the camera on an NTP-synchronised digital clock at the beginning and end of the recording sequences. An annotation tool called ELAN²⁰ was used for annotation. ELAN is a tool for the creation of complex annotations on video and audio resources, developed by the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands.

Annotators labelled the data set to indicate the location of the participant, using the labels shown in Table 1 (note that only one bedroom is in use for this experiment). Figure 3 shows location annotations for one participant.

In addition to location, ambulation (e.g., walking, climbing the stairs), posture (e.g., sitting, standing, lying), transitions between postures, activity (e.g., brushing teeth, preparing a toast with jam and a cup of tea) annotations involving actions using the participant's hands were also recorded, in particular: using, holding, grabbing and releasing an object for both left and right hand. Alongside scripted activities, participants were also asked to complete a brief set of jumps to ensure synchronisation of wearable data. The annotations are defined in Table 2.

Posture transitions were performed in groups of 5, e.g., stand-sit-lay-sit-stand performed by a participant 5 times in a row. Table 1 shows a list of all activity annotation labels. Figure 4 shows annotated activities in locations for one participant.

The annotation process resulted in millisecond-accuracy labels that are described by two timestamps: one for the start time and one for the end time. For a typical supervised learning setting, these labels are transformed into probabilistic labels for each one-second time window. For any time window that contains only one label, we assign a probability vector with one on the corresponding label and zero elsewhere.

In contrast, if a time window contains more than one type of label, we calculate the time length of each label, from which we obtain the relative proportion of each label, and assign it as the resulted probability vector. While such an approach can naturally deal with multiple labels within a given time window, it can be further used to

Label	Ontology Definition	Description
Ambulation		
a_ascend	ascending stairs	Ascending stairs
a_descend	descending stairs	Descending stairs
a_jump	jumping	A standing vertical jump
a_walk	walking	Walking without objects in one's hands
a_loadwalk	Walking with load	walking whilst holding objects (e.g. a bowl).
Posture		
p_bent	bend over	Static stand with the torso bent forward at any angle.
p_kneel	kneeling	Crouching with knees bent, resting on the knees.
p_lie	lying	Lying down in any position (e.g. supine, prone)
p_sit	sitting on chair/bed/ sofa	Any seated posture other than those defined elsewhere.
p_squat	squating	Crouching with knees bent and balancing on the feet
p_stand	standing	Any standing posture.
Transition		
t_bend	bending	Bending the torso forward, defined to include partial rotation.
t_straighten	straightening	Straightening the torso.
t_turn	turning/pivoting	A turn or pivot of the torso, defined to include partial rotation.
t_stand_kneel	stand to kneel	Transition from a standing position to a kneeling position.
t_kneel_stand	kneel to stand	Transition from a kneeling position to a standing position.
t_lie_sit	lying to sit	Transition from lying down (typically in a supine position) to a standing position
t_sit_lie	sit to lying	Transition from a seated position to lying down.
t_sit_stand	sit to stand	Transition from a seated position to a standing position.
t_stand_sit	stand to sit	Transition from a standing position to a seated position.

Table 2. Ontology definitions and detailed description of the activity labels found in the dataset.

.....

solve disagreements among multiple annotators. We can use the time length annotated by each annotator to calculate the final proportion and probability vector. Figure 4 shows the annotated activities in locations for one participant.

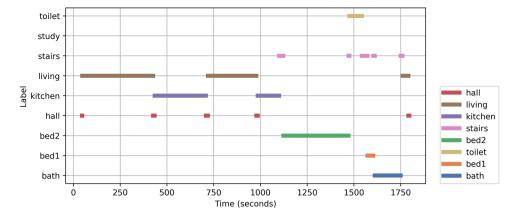
Interrater reliability was calculated using Cohen's kappa²¹, using the implementation provided in the scikit-learn package. Table 3 shows per-annotation, per-annotator-pair scores calculated across all sessions in which two or more annotation sets were available. Across all annotations, the majority of pairs achieved moderate agreement as defined by Cohen and McHugh²² (pairwise mean > 0.61). However, performance varied significantly on individual annotations. Some annotations were employed inconsistently, notably 'walking with load (loadwalk)', the posture 'squatting' and the location 'toilet', which are seldom used; this results in low agreement (loadwalk $\bar{\kappa} > 0.22$; squatting: $\bar{\kappa} > 0.47$; toilet: $\bar{\kappa} > 0.33$. There are also inconsistencies in the annotations for bedroom, where some differentiate between the different bedrooms, and other do not. For the purpose of the interrater table (Table 3), the different bedroom labels have been merged into a single 'bedroom' label. There is fair agreement on turns ($\bar{\kappa} > 0.39$) and moderate agreement on bending ($\bar{\kappa} > 0.54$), which may reflect the brief duration of these actions (average duration 1.09 s in both cases); difficulty in annotating turns may also result from the use of a head-mounted camera, and could potentially be reduced by supplementary information from additional, fixed cameras. All other annotations see moderate, substantial or in some cases almost perfect agreement.

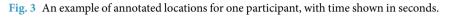
Data Records

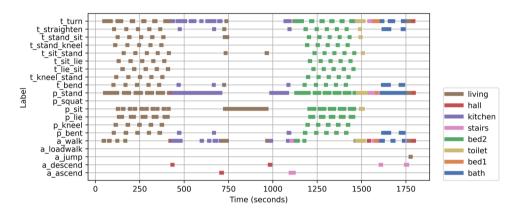
The SPHERE multi-sensor dataset is published as an openly available set of CSV files at the University of Bristol institutional data repository, with the DOI: https://doi.org/10.5523/bris.2h0wyctxrd69j2oqccsi45hy1p²³. This dataset is available under the CC-BY licence. This section outlines the format of the data records available in this dataset.

Directory structure. This dataset provides three views of the data: raw, sphere-challenge-2016 and sphere-challenge-complete-2022, each of which can be found under the data directory. The raw view consists of the full, complete and un-shuffled data (in interval and tabular form) and labels. sphere-challenge-2016 is a replication of the SPHERE Challenge that took part in ECML 2016⁴, but with more consistent cross-modality naming conventions and activity/location labels of the test set. Train and test data are created in such a way that every user from the test set can be found in the train set. The data for this view was shuffled and randomly sliced in order to force the participants of the challenge to learn activity and location recognition models rather than memorising the script. Finally, the sphere-challenge-complete-2022 has preserved the train and test splits from the SPHERE Challenge and the sequences are provided in un-sliced format. This means there are fewer sequences in this view, but each sequence is long (approximately 30 mins). The data provided in the sphere-challenge-2016 and sphere-challenge-complete-2022 views are provided in dense, CSV format (i.e., raw interval data are not provided), and the code that produces these views can be found in main.py.

A-B A-C B-C anding stairs 0.80 0.78 0.93 ading stairs 0.90 0.89 0.86 ing with load 1.0 0.0 0.89 0.86 ing with load 1.0 0.0 0.0 0.0 0.0 ing with load 1.0 0.72 0.65 0.65 0.65 0.65 ing 0.72 0.67 0.73 0.74 0.74 0.74 ing 0.73 0.43 0.44 0.72 0.65 0.65 0.65 0.44 infred sofa 0.91 0.91 0.91 0.72 0.43 0.44 <th>D-F 0.71 0.74 0.74 0.74 0.74 0.74 0.74 0.74 0.74 0.74 0.74 0.74 0.75 0.64 0.65 0.666 0.667 0.768 0.13 0.0</th> <th>E-F 0.77 0.82</th> <th></th> <th>I-J I-</th> <th></th> <th></th> <th>0-N</th> <th>D-0</th> <th>R-S</th> <th>Т.Т</th> <th>T-V</th> <th>V-U</th> <th><u>X-V</u></th> <th><u>k</u></th> <th>Usage per</th> <th>(an com) -</th> <th>Average</th> <th></th>	D-F 0.71 0.74 0.74 0.74 0.74 0.74 0.74 0.74 0.74 0.74 0.74 0.74 0.75 0.64 0.65 0.666 0.667 0.768 0.13 0.0	E-F 0.77 0.82		I-J I-			0-N	D-0	R-S	Т.Т	T-V	V-U	<u>X-V</u>	<u>k</u>	Usage per	(an com) -	Average	
ing stairs 0.80 0.78 0.93 g stairs 0.90 0.89 0.86 withload 0.39 0.91 0.8 withload 1.0 0.0 0.0 0.75 0.75 0.45 0.75 0.72 0.65 0.65 0.44 0.73 0.73 0.43 0.44 0.73 0.73 0.73 0.73 irbed sofa 0.89 0.71 0.78 irbed sofa 0.89 0.71 0.73		0.77 0.82			t	t		1		2-1					session	o (usage)	duration (S)	σ (s)
gstairs 0.90 0.89 0.86 0.86 withload 0.89 0.91 0.8 withload 1.0 0.0 0.0 0.73 0.73 0.73 0.79 0.73 0.73 0.44 0.73 0.73 0.73 0.43 0.44 0.73 0.73 0.43 0.74 0.73 0.73 0.43 0.44 0.73 0.91 0.92 0.44 0.73 0.91 0.92 0.43 irbed sofa 0.89 0.71 0.73 0.00 0.00 0.00 0.43		0.82	0.85 (0.78 0.57	50 0.95	5 0.63	0.87	0.81	0.73	0.87	0.85	0.94	0.79	7.50	3.22	3.35	1.22
0.89 0.91 0.8 with load 1.0 0.0 0.0 0.72 0.65 0.65 0.65 0.72 0.73 0.73 0.79 0.73 0.73 0.74 0.74 0.73 0.73 0.74 0.74 0.73 0.73 0.73 0.74 0.73 0.73 0.73 0.74 0.73 0.73 0.73 0.74 0.73 0.73 0.73 0.74 irbed sofa 0.91 0.91 0.72 0.00 0.00 0.43 0.43			0.91 (0.72 0.3	0.93 0.78	8 0.92	2 0.91	0.85	0.75	0.80	06.0	0.86	0.60	0.81 6	6.19	3.29	4.39	3.55
with load 1.0 0.0 0.0 0.72 0.65 0.65 0.65 0.67 0.73 0.79 0.79 0.73 0.73 0.73 0.74 0.73 0.91 0.91 0.72 ir bed sofa 0.91 0.91 0.92 ir bed sofa 0.89 0.71 0.43		0.67	0.77 (0.75 1.0	.0 0.75	5 0.73	3 0.8	1.0	0.67	1.0	0.83	0.83	1.0	0.80	4.19	2.20	1.99	0.67
0.72 0.65 0.65 0.67 0.73 0.79 0.73 0.43 0.44 0.73 0.91 0.92 irbed sofa 0.89 0.71 0.78		0.0	1.0	0.47 0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.56	0.0	0.22	5.44	5.70	3.58	2.80
0.67 0.73 0.79 0.73 0.43 0.44 0.73 0.91 0.92 0.91 0.91 0.92 chairbed sofa 0.89 0.71 0.78		0.6	0.67	0.63 0.3	0.53 0.56	6 0.61	0.67	0.58	0.62	0.61	0.65	0.64	0.57	0.62	116.38	54.21	2.38	1.56
0.73 0.43 0.44 0.91 0.91 0.92 0.89 0.71 0.78 0.0 0.00 0.43		0.57	0.72 0	0.65 0.	0.79 0.71	1 0.76	5 0.79	0.82	0.53	0.46	0.57	0.68	0.45	0.67	40.94	20.12	3.37	5.07
0.91 0.91 0.92 0.89 0.71 0.78 0.0 0.0 0.43		0.66	0.83	0.34 0.3	0.52 0.46	6 0.72	2 0.89	0.88	0.23	0.66	0.73	0.79	0.82	0.63	15.00	9.01	3.36	0.60
0.89 0.71 0.78 0.0 0.0 0.43		0.86	0.83 (0.91 0.	0.82 0.8	87 0.85	5 0.93	0.73	0.75	0.91	0.9	0.92	0.91	0.87	50.56	21.18	4.82	1.52
0.0 0.43	0.0	0.12	0.93 (0.18 0.	0.84 0.25	25 0.87	7 0.91	0.75	0.76	0.89	0.85	0.93	0.91	0.7	46.00	21.34	11.89	33.60
000 000		-0.01	1.0	0.0 1.	1.0 0.0	0.1.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.47	2.81	3.21	3.83	0.85
Posture: standing 0.82 0.82 0.82 0.75	0.8	0.74	0.83	0.71 0.	0.77 0.78	78 0.82	2 0.87	0.83	0.79	0.76	0.79	0.81	0.79	0.79	204.75	94.77	4.95	7.34
Transition: lying to sit 0.8 0.76 0.8 0.52	0.86	0.66	0.62	0.61 0.	0.61 0.88	38 0.82	2 0.86	0.55	0.24	0.88	0.86	0.9	0.72	0.72	20.13	8.55	1.61	0.51
Transition: sit to lying 0.81 0.68 0.75	0.72	0.7	0.9	0.72 0.	0.65 0.68	68 0.49	9 0.8	0.53	0.33	0.86	0.77	0.82	0.75	0.7	19.94	8.50	1.58	0.51
Transition: kneel to stand 0.76 0.36 0.35 0.51	0.62	0.66	0.68	0.51 0.	0.45 0.33	33 0.71	1 0.93	0.74	0.15	0.72	0.68	0.78	0.8	0.6	15.00	9.18	1.40	0.30
Transition: Stand to kneel 0.72 0.22 0.46	0.54	0.58	0.9	0.35 0.	0.6 0.26	26 0.73	3 0.75	0.8	0.0	0.51	0.43	0.87	0.68	0.54	15.56	8.78	1.68	0.46
Transition: sit to stand 0.68 0.74 0.68 0.37	0.51	0.47	0.69	0.64 0.	0.47 0.6	5 0.83	3 0.66	0.74	0.66	0.68	0.6	0.71	0.51 0	0.62 2	25.31	11.76	1.29	0.46
Transition: stand to sit 0.73 0.55 0.55 0.57	0.55	0.45	0.64 (0.59 0.5	0.59 0.71	1 0.68	8 0.79	0.77	0.64	0.66	0.63	0.7	0.66	0.64 2	25.38	11.63	1.52	0.65
Transition: bending 0.6 0.5 0.4 0.39	0.52	0.51	0.7 (0.48 0.4	0.48 0.51	61 0.52	2 0.78	0.6	0.51	0.54	0.54	0.56	0.54	0.54	47.69	26.47	1.09	0.48
Transition: straightening 0.7 0.58 0.69 0.39	0.65	0.46	0.65 (0.56 0.	0.59 0.61	0.46	5 0.72	0.61	0.54	0.53	0.58	0.58	0.53	0.58	46.75	25.22	1.08	0.37
Transition: turning pivoting 0.55 0.56 0.52 0.36	0.3	0.27	0.42 (0.25 0.3	0.31 0.28	8 0.4	0.55	0.45	0.41	0.4	0.38	0.18	0.41	0.39	313.63	163.27	1.19	0.60
Bathroom 0.8 1.0 0.81 0.78 1.0	0.78	0.77	0.75 (0.75 0.3	0.99 0.8	1.0	0.81	0.82	0.99	0.82	0.83	0.8	0.85	3.0	3.00	1.37	101.28	54.86
Bedroom 1.0 1.0 0.99	0.99	1.0	1.0	1.0 1.0	0 1.0	1.0	1.0	0.86	1.0	0.99	1.0	1.0	1.0	66.0	3.63	1.71	166.02	149.78
Hallway 0.78 0.59 0.9	0.88	0.95	0.44 (0.86 0.3	0.87 0.92	2 0.88	3 0.88	0.77	0.93	0.88	0.89	0.93	0.94	0.83	18.38	7.96	13.72	37.41
Kitchen 1.0 0.99 0.98	0.98	1.0	0.72	0.99 0.0	0.99 0.99	9 1.0	66.0	0.99	0.99	0.99	0.99	1.0	0.99	0.98	3.75	1.61	158.11	63.02
Living room 0.99 1.0 0.99 1.0	1.0	1.0	1.0 (0.99 0.0	0.99 0.99	9 1.0	1.0	1.0	0.69	1.0	1.0	1.0	1.0	0.98	5.75	2.79	212.02	126.14
Toilet 0.0 0.0 1.0 0.0	0.98	0.0	0.0	0.0 0.0	0.0 1.0	0.0	0.96	1.0	0.0	1.0	0.0	0.0	0.0	0.33 (0.81	0.75	53.3	0.0
Pairwise mean 0.72 0.63 0.66 0.57	0.64	0.56	0.71	0.55 0.	0.63 0.6	5 0.68	8 0.77	0.72	0.52	0.71	0.67	0.7	0.64					









Taking the sphere-challenge-2016 view, two sub-folders, for the training set and testing set respectively can be found. The training set contains 10 folders with long sequences of recorded data, with each sequence lasting around 20 to 30 minutes. The testing set contains 872 folders with short sequences of recorded data, and most sequences are about 30 seconds.

All recorded data are marked with unique codes (each recording will be referred to as a 'sequence folder'). Timestamps are re-based to be relative to the start of the sequences, i.e., each sequence always starts from t=0. However, individual CSV files begin with an offset that reflects the time between the timestamp at which the sequence began and the datetime at which the first subsequent data point was generated: for example, accelerometer data may begin some milliseconds following t = 0, while video data, which is event-driven, may begin long after t = 0 or not appear in a certain sample at all.

Each sequence folder contains the following files for sensor data:

- pir.csvand/orpir raw.csv
- acceleration.csv
- rssi.csv
- rgbd hall.csv
- rgbd living.csv
- rqbd kitchen.csv
- meta.json

Each folder furthermore contains a set of files for the annotations with the following name formats:

- activity.csv .
- location.csv
- per_ann_activity_*.csv and/or activity_*_raw.csv
 per_ann_location_*.csv and/or location_*_raw.csv

Here * will be a positive integer indicating corresponding annotator. That is, having two files as per_ann_ activity 1.csv and per ann activity 2.csv means there are two different annotators for the activity labels. The activity.csv and location.csv files contain the merged ground-truth from all annotators.

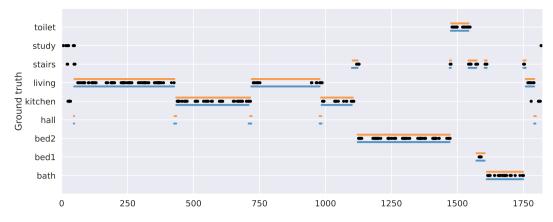


Fig. 5 Example PIR for training record 00001. The black lines indicate the time duration where a PIR is activated. The blue and green wide horizontal lines indicate the room occupancy labels as given by the two annotators that labelled this sequence.

Sensor data files. pir raw.csv

This file contains the start time and duration for all PIR sensors in the smart environment. PIR sensors are located in the following nine rooms: *bath*, *bed1*, *bed2*, *hall*, *kitchen*, *living*, *stairs*, *study*, *toilet*. The columns of this CSV file are:

- start: the start time of the PIR sensor (relative to the start of the sequence)
- end: the end time of the PIR sensor (relative to the start of the sequence)
- name: the name of the PIR sensor being activated (from the pir locations list)
- index: the index of the activated sensor from the pir locations

Example PIR sensor activation and ground truth locations are overlaid in Fig. 5. Note, the PIR sensors can be noisy as they are based on infra-red technologies, which can be affected by strong natural light.

pir.csv

This file contains the PIR data processed and sampled at a 0.1 second period (for raw view) or 1.0 seconds (for the sphere-challenge-2016 and sphere-challenge-complete-2022 views). The file consists of dense binary data for each PIR location (bath, bed1, bed2, hall, kitchen, living, stairs, study and toilet) and one supplementary column containing time.

acceleration.csv

The acceleration file consists of four columns:

- t: this is the time of the recording (relative to the start of the sequence)
- x/y/z: these are the acceleration values recorded on the x/y/z axes of the accelerometer.

Sample acceleration and RSSI with overlaid annotations are shown in Fig. 6. The top Fig. 6 shows the acceleration signals overlaid with activity labels.

rssi.csv

The RSSI file consists of five columns:

- t: this is the time of the recording (relative to the start of the sequence)
- **kitchen/living/study/stairs**: these specify the RSSI signal as received by each access point. Empty values indicate that the access point did not receive the packet.

The bottom Fig. 6 shows the RSSI signal information with the overlaid room occupancy.

rgbd *.csv

The following sixteen columns are found in the rgbd hallway.csv,

rgbd_kitchen.csv and rgbd_living.csv files:

- **t**: The current time (relative to the start of the sequence)
- **centre_2d_x/centre_2d_y**: The x- and y-coordinates of the centre of the 2D bounding box.
- **bb_2d_br_x/bb_2d_br_y**: The x and y coordinates of the bottom right (br) corner of the 2D bounding box
- **bb_2d_tl_x/bb_2d_tl_y**: The x and y coordinates of the top left (tl) corner of the 2D bounding box
- centre_3d_x/centre_3d_y/centre_3d_z: the x, y and z coordinates for the centre of the 3D bounding box

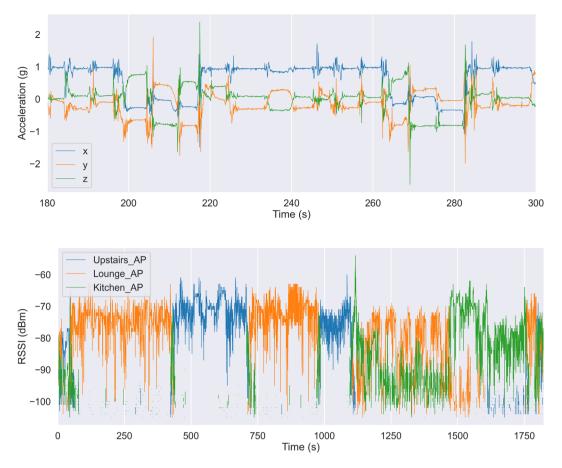


Fig. 6 Example acceleration and RSSI signals for training record 00001. The line traces indicate the accelerometer/ RSSI values recorded by the access points. The horizontal lines indicate the ground-truth as provided by the annotators (two annotators annotated this record, and their annotations are depicted by the green and blue traces respectively). (**Top**) Acceleration signal trace shown over a 5 minute time period. Annotations are overlaid. (**Bottom**) RSSI values from the four access points. Room occupancy labels are shown by the the horizontal lines.

- **bb_3d_brb_x/bb_3d_brb_y/bb_3d_brb_z**: the x, y, and z coordinates for the bottom right back (brb) corner of the 3D bounding box
- bb_3d_flt_x/bb_3d_flt_y/bb_3d_flt_z: the x, y, and z coordinates of the front left top (flt) corner of the 3D bounding box.

Example 3D centre of mass data is plotted for the hallway, living room and kitchen cameras in Fig. 7. Room occupancy labels are overlaid on these, where we can see strong correspondence between the detected persons and room occupancy.

Annotation data files.

The following two sets of files need not be used for the challenge, but are included to facilitate users that wish to perform additional modelling of the sensor environment. For example, indoor localisation can be modelled with the locations.csv file²⁴.

activity_*_raw.csv

This file provides the individual annotations as provided by the annotators. The target variables are the same as for assets/activity_labels.csv. The following 20 activities are annotated: {*a_ascend*, *a_descend*, *a_jump*, *a_loadwalk*, *a_walk*, *p_bent*, *p_kneel*, *p_lie*, *p_sit*, *p_squat*, *p_stand*, *t_bend*, *t_kneel_stand*, *t_lie_sit*, *t_sit_lie*, *t_sit_stand*, *t_stand_kneel*, *t_stand_sit*, *t_straighten*, *t_turn*}.

As before, the prefix 'a_' indicates an ambulation activity (i.e., an activity consisting of continuing movement), 'p_' annotations indicate static postures (i.e., times when the participants are stationary), and 't_' annotations indicate posture-to-posture transitions.

The file activity_*_raw.csv contains the following four columns:

- **start**: the start time of the activity (relative to the start of the sequence)
- end: the end time of the activity (relative to the start of the sequence)

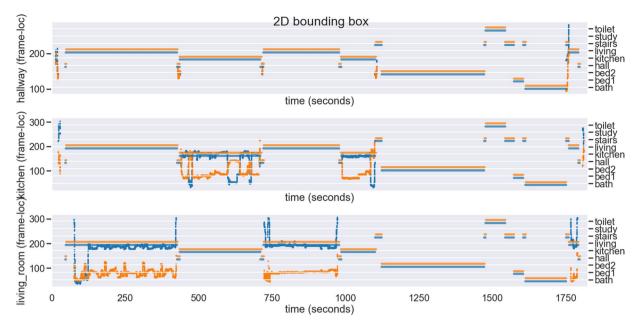


Fig. 7 Example centre_3d for training record 00001. The horizontal lines indicate annotated room occupancy. The blue, green and red traces are the *x*, *y*, and *z* values for the 3D centre of mass.

- **name**: the name of the label
- index: the index of the label name in activity_labels

location * raw.csv

This file provides the annotation labels for room occupancy. The same nine rooms are labelled as seen with the PIR sensor: {*bath, bed1, bed2, hall, kitchen, living, stairs, study, toilet*}.

The file assets/location_label.csv contains the following four columns:

- start: the time a participant entered a room (relative to the start of the sequence)
- end: the time the participant left the room (relative to the start of the sequence)
- name: the name of the room (from the location_labels list)
- index: the index of the room name starting at 0

activity.csv/location.csv

The activity.csv and location.csv files contains the aggregated ground-truth from multiple annotators on the activities and locations. As briefly introduced in the Methods section, the annotations are separated into a one-second time window. Within each window, the annotated length of each activity from each annotator is calculated, the final ground-truth is given as the normalised total annotated time for every activity (e.g., the rows sum to one).

The activity.csv files are dense CSV with columns corresponding to time and the labels listed in location_*_raw.csv. Similarly, the location.csv files are also in dense CSV format, but with columns corresponding to those in location * raw.csv.

per ann activity *.csv/per ann location *.csv

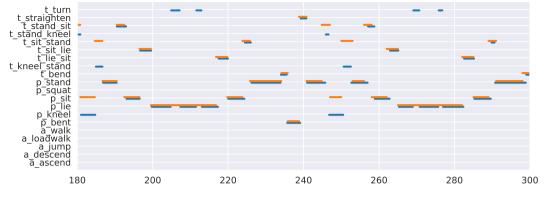
These files provide a dense binary CSV file of each annotator for activity and location. Averaging over the set of annotation files will produce the data in activity.csv/location.csv. The columns of per_ann_activity_*.csv are the activity labels from activity_*_raw.csv, and the columns of per_ann_location_*.csv are the location labels from location_*_raw.csv.

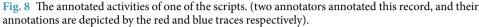
Supplementary data files. meta.json

This file contains the metadata of the file including the sequence start and end times, the ID of the annotators as a list of integers, and the ID of the volunteer as an integer.

Programatic loading. For convenience, two helper classes are provided to load data from raw.

```
from sphere_challenge import ProcessedSequence, RawSequence
# Load a raw view
raw=ProcessedSequence.load_from_path("data/raw/001")
# Load sphere-challenge-2016 test
sc2016=ProcessedSequence.load_from_path("data/sphere-challenge-2016/
test/00011")
```





.....

Load sphere-challenge-2016 test
sc2022 = ProcessedSequence.load_from_path("data/
sphere-challenge-complete-2022/test/00011")

The following member variables are exposed for RawSequence objects instantiated to a variable called obj:

- obj.meta: an object with additional member variables of start, end, annotators, user id
- obj.acceleration: a Pandas dataframe with t (time), x, y, z columns
- obj.rssi: a Pandas dataframe with
- obj.pir:a Pandas dataframe with t, bath, bed1, bed2, hall, kitchen, living, stairs, study, toilet
- obj.rgbd: an object with additional fields of hall, kitchen, and living. Each sub-field is a Pandas dataframe with columns indicated in rgbd *.csv.
- obj.labels.activity: a Pandas dataframe with the activity labels as columns (see activity_*_ raw.csv)
- obj.labels.location: a Pandas dataframe with the location labels as columns (see location_*_raw.csv)

Alternatively, to load the raw data from file, one may call <code>RawSequence.load_from_path</code> in the same manner as above.

Technical Validation

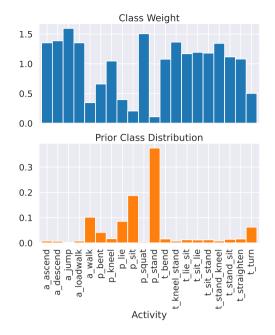
In this section we provide a set of technical validation results on the dataset. As mentioned above, the overall aim of the sensor platform is to capture different behaviour patterns that are potentially linked to certain health conditions. As the sensors are equipped within residential homes, people's behaviours can hence be quantified through these three variables: time, location, and activity. In this section, we demonstrate the dataset via a fundamental application: activity recognition with supervised machine learning.

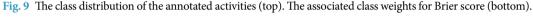
Since for each scripted experiment we allocated two annotators to provide the ground-truth for the activities, there can potentially be disagreements between the two annotators at a certain time point. On the other hand, for a given time range (e.g., one second), it might contain multiple annotations according to the original precision. Figure 8 shows an example from two annotators on a particular script. As indicated, while both annotators agree on most activities, some styling differences can still be observed. For instance, one of the annotators records the activity of consecutive jumping with three segments, but the other annotator treats them together as one. With such differences in mind, we formalised the annotations into a normalised vector as a frequentist view of probabilistic classification. For a given second, we calculate the overall time length given by all the annotators, then normalise these time intervals into a probability vector over all the activities/locations.

Here we demonstrate a simple approach to achieve the task described above, which uses some basic features together with a probabilistic nearest neighbour classifier, with the implementation provided by the scikit-learn package²⁵. For the features, we extract the mean, minimal value, maximum value, median, and standard deviation from each modality (Acceleration and RSSI values from the wearable device, PIR sensor values, bounding box locations from the three camera locations.), and combine them into a feature vector for that particular second. After obtaining the features, a probabilistic nearest neighbour can be used to model the probability of each activity on a given second, based on a number of closest feature vectors. The overall steps are shown by a flowchart in Fig. 10.

Figure 12 shows the results on a particular scripted experiment, while the model is trained on other nine further experiments (using 128 nearest neighbours). As indicated by the figure, a simple model described above is able to capture the major activities like stand and sit, while other minor activities and postures can be still improved by considering advanced feature and model learning approaches.

In terms of performance, since both the ground-truths and model predictions are probability vectors, one of the most common evaluation measures here is the Brier score, which equals to the sum of squared errors among





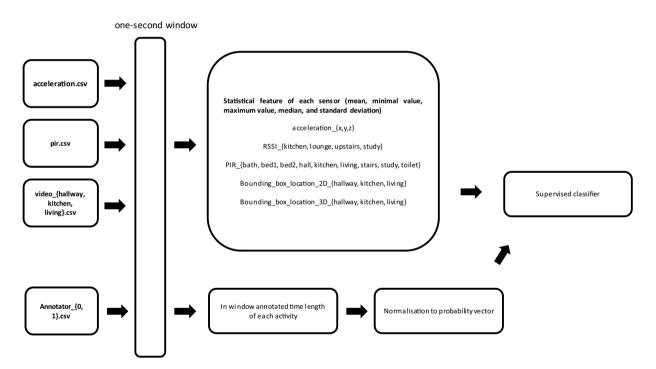


Fig. 10 The flowchart of the data (sensor readings and annotations) processing steps for supervised classifier training (best viewed zoomed in).

the probability vectors^{26,27}. Since the class distribution is highly unbalanced in this dataset (e.g., the activity of jump only takes a few seconds while the posture of standing can take minutes), it is common to further adjust the weights of the Brier score of each class as in the case of cost-sensitive classification²⁷. Here we consider to set the weights of each class as the reciprocal of their marginal class probability value. An example is given in Fig. 9, and we can see that minority classes such as 'jump' and 'squat' are assigned top weights for the score evaluation.

Figure 11 shows the Brier score on different number of neighbours considered by the model described above. For this particular task, it can be seen the performance increases when the number of neighbours grows larger. While the models are very simple in this case, it is sufficient to demonstrate the datasets are of great value when modelling the activities within a residential home.

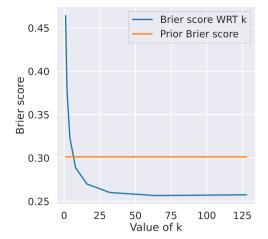
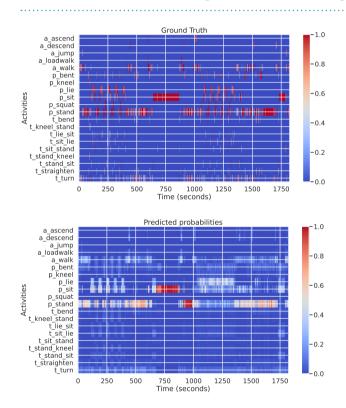
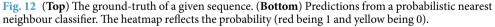


Fig. 11 The effect of the number of neighbours, evaluated by the Brier score. The performance is obtained by leave-one-out validation (train with sequence 01 to 09, tested on sequence 10).





Usage Notes

In this section we provide some additional notes for this dataset, as well as introducing other related work on health-care and sensor-based behaviour modelling. Diethe *et al.*⁷ provides a general view of the SPHERE project in terms of applied machine learning and data mining. Diethe *et al.*²⁸ introduces HyperStream, which is a generic stream processing software developed under the SPHERE project. Diethe *et al.*²⁹ discusses different approaches to fuse multi-modal streams towards smart home applications. Various approaches on activity recognition have been further documented^{19,30,31}, generating discussion of some of the results on monitoring long-term health conditions^{32,33}.

Code availability

A git repository is publicly available at https://github.com/IRC-SPHERE/sphere-challenge-sdata/. In this repository a number of scripts for visualisation, bench marking and data processing are available. (All subsequent sensor images were generated using these scripts).

Received: 2 September 2021; Accepted: 13 February 2023; Published online: 23 March 2023

References

- Woznowski, P. et al. A multi-modal sensor infrastructure for healthcare in a residential environment. In 2015 IEEE International Conference on Communication Workshop (ICCW), 271–277, https://doi.org/10.1109/ICCW.2015.7247190 (2015).
- Zhu, N. et al. Bridging e-health and the internet of things: The SPHERE project. IEEE Intelligent Systems 30, 39–46, https://doi.org/ 10.1109/MIS.2015.57 (2015).
- 3. Diethe, T. & Flach, P. A. Smart-homes for eHealth: Uncertainty management and calibration. In NIPS 2015 Workshop on Machine Learning in Healthcare (2015).
- Twomey, N. et al. The SPHERE challenge: Activity recognition with multimodal sensor data. CoRR abs/1603.00797, https://doi.org/ 10.48550/arXiv.1603.00797 (2016).
- Liu, X., Liu, L., Simske, S. J. & Liu, J. Human daily activity recognition for healthcare using wearable and visual sensing data. In 2016 IEEE International Conference on Healthcare Informatics (ICHI), 24–31 (IEEE, 2016).
- Twomey, N., Diethe, T., Craddock, I. & Flach, P. A. Unsupervised learning of sensor topologies for improving activity recognition in smart environments. *Neurocomput.* 234, 93–106, https://doi.org/10.1016/j.neucom.2016.12.049 (2017).
- Diethe, T. et al. Releasing ehealth analytics into the wild: lessons learnt from the sphere project. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 243–252 (2018).
- Holmes, M. *et al.* Analysis of patient domestic activity in recovery from Hip or Knee replacement surgery: modelling wrist-worn wearable RSSI and Accelerometer data in the wild. In *The 3rd International Workshop on Knowledge Discovery in Healthcare Data @* IJCAI. (CEUR Workshop Proceedings, 2018).
- 9. Fafoutis, X. et al. SPW-1: A Low-Maintenance Wearable Activity Tracker for Residential Monitoring and Healthcare Applications. In Proc. Int. summit on eHealth, 294–305 (2016).
- 10. Analod Devices. ADXL362 Micropower, 3-Axis, $\pm 2g/\pm 4g/\pm 8g$, Digital Output MEMS Accelerometer, Rev. B (2013).
- Fafoutis, X. et al. A residential maintenance-free long-term activity monitoring system for healthcare applications. EURASIP Journal on Wireless Communications and Networking 2016, https://doi.org/10.1186/s13638-016-0534-3 (2016).
- Diethe, T., Twomey, N. & Flach, P. A. Active transfer learning for activity recognition. In European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (2016).
- Twomey, N., Diethe, T. & Flach, P. A. Bayesian active learning with evidence-based instance selection. In Workshop on Learning over Multiple Contexts, European Conference on Machine Learning (ECML'15) (2015).
- Fafoutis, X. et al. Privacy leakage of physical activity levels in wireless embedded wearable systems. IEEE Signal Processing Letters 24, 136–140, https://doi.org/10.1109/LSP.2016.2642300 (2017).
- Camplani, M. *et al.* Real-time rgb-d tracking with depth scaling kernelised correlation filters and occlusion handling. In Xie, X., Jones, M. W. & Tam, G. K. L. (eds.) *Proceedings of the British Machine Vision Conference (BMVC)*, 145.1–145.11, https://doi.org/ 10.5244/C.29.145 (BMVA Press, 2015).
- Tao, L. et al. A comparative home activity monitoring study using visual and inertial sensors. In 2015 17th International Conference on E-health Networking, Application Services (HealthCom), 644–647, https://doi.org/10.1109/HealthCom.2015.7454583 (2015).
- Tao, L. et al. A comparative study of pose representation and dynamics modelling for online motion quality assessment. Computer Vision and Image Understanding (2015).
- Twomey, N. & Flach, P. Context modulation of sensor data applied to activity recognition in smart homes. In Workshop on Learning over Multiple Contexts, European Conference on Machine Learning (ECML'14) (2014).
- Woznowski, P., Tonkin, E. L. & Flach, P. A. Activities of daily living ontology for ubiquitous systems: Development and evaluation. Sensors 18, 2361, https://doi.org/10.3390/s18072361 (2018).
- Max Planck Institute for Psycholinguistics, The Language Archive. Elan (version 6.4) [computer software]. retrived from https:// archive.mpi.nl/tla/elan (2022).
- Cohen, J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20, 37–46, https://doi.org/ 10.1177/001316446002000104 (1960).
- 22. McHugh, M. L. Interrater reliability: the kappa statistic. Biochemia medica 22, 276-282 (2012).
- 23. Twomey, N. *et al.* Sphere house scripted dataset: A multi-sensor dataset with annotated activities of daily living recorded in a residential setting v2.0. *University of Bristol* https://doi.org/10.5523/bris.2h0wyctxrd69j2oqccsi45hy1p (2022).
- 24. Fafoutis, X. et al. An rssi-based wall prediction model for residential floor map construction. In Proceedings of the 2nd IEEE World Forum on Internet of Things (WF-IoT) (2015).
- 25. Pedregosa, F. et al. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12, 2825-2830 (2011).
- Brier, G. W. Verification of forecasts expressed in terms of probability. *Monthly Weather Review* 78, 1–3, https://doi.org/ 10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2 (1950).
- 27. Flach, P. A. Machine learning: the art and science of algorithms that make sense of data (Cambridge University Press, 2012).
- 28. Diethe, T. et al. Hyperstream: a workflow engine for streaming data. arXiv preprint abs/1908.02858 (2019).
- 29. Diethe, T., Twomey, N., Kull, M., Flach, P. A. & Craddock, I. Probabilistic sensor fusion for ambient assisted living. arXiv preprint abs/1702.01209 (2017).
- Twomey, N., Diethe, T. & Flach, P. A. On the need for structure modelling in sequence prediction. *Machine Learning* 104, 291–314, https://doi.org/10.1007/s10994-016-5571-y (2016).
- Bi, H., Perello-Nieto, M., Santos-Rodriguez, R. & Flach, P. Human activity recognition based on dynamic active learning. *IEEE Journal of Biomedical and Health Informatics* https://doi.org/10.1109/JBHI.2020.3013403 (2020).
- Holmes, M. et al. Modelling patient behaviour using iot sensor data: a case study to evaluate techniques for modelling domestic behaviour in recovery from total hip replacement surgery. *Journal of Healthcare Informatics Research* 4, 238–260, https://doi.org/ 10.1007/s41666-020-00072-6 (2020).
- 33. Yamagata, T. et al. Model-based reinforcement learning for type 1 diabetes blood glucose control. arXiv preprint abs/2010.06266 (2020).

Acknowledgements

This work was performed under the SPHERE IRC funded by the UK Engineering and Physical Sciences Research Council (EPSRC), Grant EP/K031910/1. NT is currently supported by 'Continuous Behavioural Biomarkers of Cognitive Impairment' project funded by the UK Medical Research Council Momentum Awards under Grant MC/PC/16029.

Author contributions

E.L.T. is the data manager of the project. E.L.T., M.H., H.S and G.J.L.T. prepared the manuscript and manage the submission process. N.T., T.D. and M.K. initiated and developed related software and assisted with experiment design, data collection and management. N.T. and T.D. organised the *SPHERE challenge* workshop in ECML-PKDD 2016. E.L.T., M.H., H.S, N.T., T.D., M.K. and M.P.N. are members of the data mining and data fusion team, and contributed to the general analysis and maintenance of the dataset. M.C. and S.H. assisted the data collection with the vision system including the bounding box algorithms. X.F. assisted the data collection with the wearable

platform and wireless infrastructure. Z.N. and P.W. assisted the data collection with the IoT and infrastructure system. P.W. recruited and consented the experiment's participants and conducted the scripted data collection. R.S.R., P.F. and I.C. supervised the overall project and individual team members.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to E.L.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2023