Aileen Krabbe, Roman Seidel and Gangolf Hirtz
Faculty of Electrical Engineering and Information Technology, Chemnitz University of Technology, Germany

Detection of activities of daily living with decision trees through a technical assistance system

https://doi.org/10.1515/cdbme-2023-1057

Abstract: To cope with the increasing demands in care due to the aging society and the simultaneous lack of professional caregivers, a technical assistance system can help to monitor elderly people in their own homes and to support professional caregivers and caring relatives. The technical assistance system consists of a smart sensor with an omnidirectional camera on the ceiling of each room and additionally, other smart home sensors in an apartment for elderly persons. Based on smart sensor data, the positions, poses, and activities of the patient are detected with the help of machine learning (ML) techniques. In this work a temporal behaviour model of the patient is developed to recognize activities of daily living (ADL) such as "eating", "sleeping" or "emergency". For this, the actions (e.g., walking, sitting) and the location in the current room of the patient, as well as the data of other sensors in the apartment are used. This input data is fed into a trained decision tree of depth 14, which ultimately determines the patient’s activity. The accuracy for detecting activities of daily living with the decision tree is 96.47%, where the activities can be detected in real-time.

Keywords: ADL detection, dementia, longer independent life, machine learning, elderly care, decision tree

1 Introduction

According to estimates by the world health organization (WHO) and Alzheimer's Disease International, 50 million people worldwide are currently affected by dementia, and by the year 2050, 0.8 to 1.2 million more sufferers are expected [1, 2]. To face this challenge, technical assistance systems can help to monitor elderly people in their own homes and to support professional caregivers and dementia patients as well as their relatives. This work is part of the development of a technical assistance system, AUXILIA, which uses a smart sensor that was attached to the ceiling of a room and is combined with other sensors in an apartment for elderly, for example on electrical devices such as the oven or on the doors. These smart sensors can determine the patient's pose and location in the room.

The activities of daily living (ADL) can be used for a behavioural classification of patients with incipient dementia. The ADLs can be detected using the status of the smart home sensors and the included object and person detection, as well as other sensors in the apartment. ADLs can include cooking, cleaning or sleeping, which we determine by a machine learning method, namely the decision tree.

With this, a temporal behaviour model is created and based on this model a daily routine of the patient is documented. In order to be able to better assess dangerous situations in the course of Alzheimer's disease, the deviations from the norm of a daily course can be determined in the next step.

The contribution of this paper is twofold:

- The creation of a dataset, which consists of data from multiple sensors and corresponding activity labels.
- Developing a real-time system based on a decision tree for the detection of activities on smart-sensor data.

The remainder of the paper is structured as follows: After this introduction, related works for person detection are pointed out, due to their relevance as a basis for the dataset. The second part of the related work chapter is about introducing several methods for ADL detection. Chapter 3 explains the technical assistance system for the elderly, wherein this work is embedded. The ADL detection and its qualitative and quantitative evaluation is shown in chapter 4 before the paper ends with the last chapter, chapter 5 with a conclusion and potential future work.

2 Related Work

Our work is related to two active research fields. The detection of top-view persons from a wide-angled, fisheye camera and the derivation of human behaviour from a smart home environment, here referred as ADL detection.

The basics for the ADL detection are reflected in the person detection, object detection and a subsequent tracking. Person detection is a widespread field, where here are shown approaches, which use adapted versions of person detectors for top-view omnidirectional person detection.

Person detection: The person and object detection use YOLOv3 or version 2, respectively, where the cameras are mounted on the ceiling of a room [3, 4, 5]. One of the major

***Aileen Krabbe:** Chemnitz University of Technology, Reichenhainer Str. 70, Chemnitz, Germany, e-mail: aileen.krabbe@etit.tu-chemnitz.de
Roman Seidel, Gangolf Hirtz: Chemnitz University of Technology, Chemnitz, Germany

Open Access. © 2023 The Author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

goals of this work is the detection of human falls [3, 4]. The authors of [6] are using an omnidirectional camera and YOLOv2 and a non-maxima suppression to detect persons. As well, the posture of the person can be detected using the YOLO algorithm with the help of the silhouette of a person [5]. Multiple persons can be tracked in a room with an omnidirectional camera based on the world coordinates and local rectifications, as shown in [16].
ADL detection: According to Naeem et al. [7], several Hidden Markov Models (HMM) can be used for ADL detection. An HMM consists of a random process， which expresses the number of hidden states that control a second random process. As a result, an observation is determined from a state-dependent probability distribution for each point in time. This technique includes data mining techniques and ontologies in which unknown sensors can be used for activity determination. Ontology describes the course of activity determination. Segmented tasks (“preparing tea or toast”) and finally the ADL (“breakfast”) are derived from the sensor data. The dependent probability is used to determine the task under the condition of an active or inactive sensor [7]. The hierarchy of ADLs begins with the components responsible for capturing the sensory events in the living environment. The second level is task identification. This allows a sensor to be linked to an event, for example, the dementia patient is making tea. The hidden Markov models perform this identification. In the higher levels, further (partial) goals of the patient to be monitored are set, each of which corresponds to an ADL [7]. Another way to determine ADL is to use a lookup table (LUT) that contains the center points of all objects in the home and a threshold value. This threshold indicates the maximum distance to the patient (center point) to be able to interact with the object. In this model, each object is assigned a specific relevance for the detection [8].

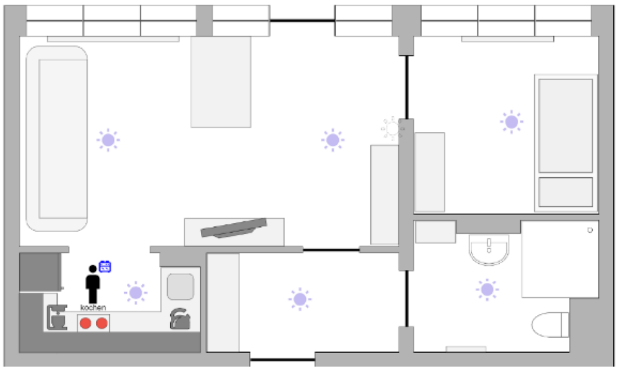

Figure 1: A floorplan of a 2-room apartment that is equipped with smart sensors. Here the ADL cooking is derived from a decision tree based on different sensor data. The person is standing in the kitchen and the stove is turned on where the ADL cooking is detected.

Figure 2: Example Frame of the omnidirectional camera in the apartment

3 Technical Assistance System

The technical assistance system, which is used to collect the data, contains a smart sensor with an omnidirectional camera in each room of an apartment (living room, kitchen, bathroom, bedroom, hall), as well as multiple on/off-sensors at the windows, doors, kitchen appliances, tv or sink/basin. The apartment, which is equipped with these smart sensors, is shown in Figure 1. The images of the omnidirectional cameras were used to detect the face, person, skeleton, and object of the patient. One input image can be seen in figure 2.
The face detection used MobileFaceNet [9], because to only get the activities of the person in need of care. This ensures that the detected pose and activity is assigned to the patient and not to a guest.
In the assistance system, a deep neural network was used for the person recognition including YOLOv2 with a perspective transformation and soft NMS with Gaussian smoothing [10]. For this training phase PASCAL VOC [11] and COCO [12] dataset was used, as well as a synthetic omnidirectional dataset THEODORE [13]. The quantitative evaluation of the MS COCO, THEODORE and Bomni DB is with the Faster R-CNN leads to a mean average precision of 0,740.

The skeleton detection a top-down method and therefore uses the resulting bounding box [17]. The skeleton points follow the MSCOCO notation. Based on these key points, actions like sitting, laying, walking and fall can be recognized. Our detection of a skeleton is described by Yu et al. [18].
The object detection was using an anchorless convolutional neural network, because of the lower calculation time. During its quantitative evaluation trained on the THEODORE dataset the Faster R-CNN results in average precision for the rollator 0.596, armchair 0.148, chair 0.141, person 0.873, table 0.98 and TV 0.943. Using SSD or R-FCN give slightly better results by some objects, for example SSD for chair and person or R-FCN for rollator and armchair.
With tracking of the person in an apartment, the current room can be determined and written in the database. For this, a

prediction with the Kalman Filter was used for tracking the person in an apartment.

4 ADL Detection

Section 2 describe state-of-the-art models for estimating ADLs, which are not applicable in real time.

ADLs contain a mixed variety of daily activities such as showering, relaxing, eating, watching TV, dishwashing, cooking, emergency (fall), sleeping, defecating, and washing that are estimate in this work by a decision tree.

Two ADL datasets were used in Wang et al. for abnormal behaviour detection in a series of activities. MobiAct uses a smartphone-sensor for action detection and SIMADL, which contains activities from different sources [19].

Various activities can be derived from the individual status of all sensor data. The action of the person, as well as the room and the area, in which the patient is located, are also available sensors. For example, if the patient is recognized by the stove being turned on in the kitchen, the ADL "cooking" is issued. In Figure 1, a floor plan of the ADL “cooking” activity is shown with an avatar at a stove.

4.1 Functional Structure

The status of a selection of sensors in the home of the patient is used to determine ADLs. The used dataset was created in a test apartment with a smart sensor integrated in an omnidirectional camera on the ceiling of multiple rooms like living room, kitchen, or bedroom. The dataset was collected by 3 different staff of our professorship, where they were for example relaxing on the couch, washing hands in the kitchen, drinking in the living room, or just laying in bed. For each action around 1 min of data from each contestant was collected and cleaned, where the sensors status wasn’t updated yet in the database. In the smart sensor, the object, person, and sleep detection were executed. Furthermore, there were more sensors attached to the coffee machine, stove, windows, doors, or fridge, where the current status was on/off or open/closed. The existing dataset from the sensors, which also includes the position and action of the dementia patient, was unevenly distributed. To equalize the distribution of the data, the algorithm SMOTE (Synthetic Minority Over-sampling Technique) is used, which can generate synthetic data samples of a minority class. In this method, synthetic examples are created from at least two of the same ADL with similar status values. In the newly generated activity, a sensor index is calculated by a random amount between the two samples [14], so that a new dataset with uniformly distributed data points is generated. The uniformly distributed dataset is used to train a decision tree with the CART algorithm. The CART algorithm was used, because it is more accurate und precise than the ID3 or MARS algorithm by classification problems. The decision tree was chosen because other machine learning methods deliver worse (e.g., logistic regression, multi-layer perceptron) or equally good results (e.g., k-nearest-neighbor algorithm) and require more storage space. For example, k-Nearest-Neighbor or Support Vector Machines or Multilayer Perceptron deliver better results than other ML-methods like Naive Bayes or Self-Organizing Maps (SOM). Different combinations of these methods in an ensemble learning method (bagging or boosting) do not produce better results for this dataset either. Bagging is an ensemble learning method that learns all ML-methods parallel instead of sequential like Boosting. In Boosting, the input data are all parameters plus the output of the previous ML method. These can be arranged in different ways [15]. Nevertheless, the ADL detection with Boosting was not delivering better results than the best ML-method.

Furthermore, the decision tree makes it easier to understand how the model determined an ADL. If there is no current status of a sensor, it is filled with a constant or its last known status value to determine an ADL reliably.

Decision trees are easier to interpret than other machine learning methods and can be used in real time up to a certain size, as in this case with 44 nodes and a depth of 14. The tree follows the right side when the sensor status index is less than a certain value and the left side when the index is greater than the value. For example, in Figure 3: If the index of the status of the TV sensor is greater than 19.5 (on) and the sleep sensor is greater than 0.005, then the patient is resting (ADL “relaxing”). The status index will not change if the model is used in a new environment, but the index’ will change if the decision tree would be new trained with other data from another environment.

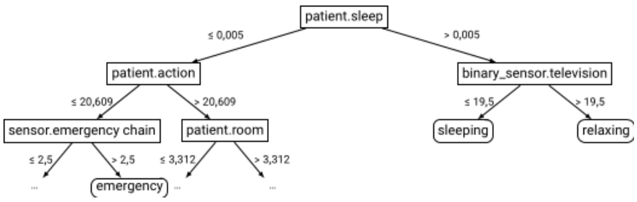

Figure 3: The first layers of the 4-of-14 depth decision tree for determination of ADL is shown.

In our experiments, where the root nodes are randomly chosen, their influence on activity prediction is not the highest. The reason for this is the even distribution of the dataset. Usually, the rule applies that the closer the sensor is to the root of the decision tree, the greater its influence on the prediction. Our observation is that the action and location of the patient have the biggest influence on activity detection. However, the output highly depends on the accuracy of determining the action and location.

4.2 Classification Rates of Test Data Set

All the selected ADLs are listed in table 1. These are the most common activities of daily life [7,8].

The used dataset has a total size of 781 data points. After the decision tree was trained and built with 80% of the entire dataset, a testing phase was performed with the remaining 20% of the dataset to finally determine the performance of the model. Table 1 shows the classification rates for each ADL. The classifications with a precision greater than 0.94 and an

F1 value greater than 0.94 give the best results. For example, "cooking", "sleeping" or "eating".

Less good results are “relaxing” with 0,88 precision and “watching TV” (0,93). In the confusion matrix (figure 4) is readable, that “watching TV” was labeled as “relaxing” 3 times. The reason for this is the similarity between those two activities, the main difference is that the TV is on, when the person is for example sitting or lying on the couch in the living room. Another example is the similarity of dishwashing and cooking because the person is standing in the kitchen and either the stove or the sink is on. There could be a misleading, when none of these is on and the person is just standing in the kitchen for a moment.

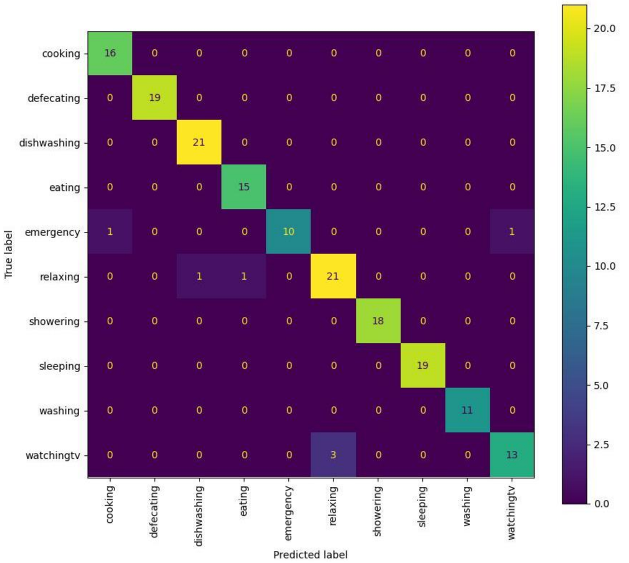

Figure 4: Confusion Matrix of CART decision tree for ADL detection

Table 1: Classification rates of activities of daily living.

ADL	Precision	Recall	F1-score	Number of data points
Showering	1,00	1,00	1,00	18
Relaxing	0,88	0,91	0,89	23
Eating	0,94	1,00	0,97	15
Watching TV	0,93	0,81	0,87	16
Dishwashing	0,95	1,00	0,98	21
Cooking	0,94	1,00	0,97	16
Emergency	1,00	0,83	0,91	12
Sleeping	1,00	1,00	1,00	19
Defecating	1,00	1,00	1,00	19
Washing	1,00	1,00	1,00	11

The accuracy for detecting activities of daily living with the decision tree is 96.47%. This model is real-time capable, as it takes less than 0.001 ms to predict an ADL from all sensor data. For this prediction, a Lenovo laptop with a 16GB RAM, Intel Core i7 processor and a 64-bit system was used.

5 Conclusions and Future Work

In this work, a decision tree was developed to detect human activities of daily life (ADLs). To prevent falls and help diagnose, a technical assistance system can detect falls and ADLs. With multiple days of collected ADLs, forgotten activities can be shown, which could be an indicator for health issues. For detection of activities of daily life, a dataset with a total size of 781 data points was collected. For the detection multiple machine learning methods, as well as combined in ensemble learning were applied to the dataset. It turned out that the decision tree gives better results than the ML-methods like SOM or k-Nearest-Neighbor, as well as Bagging or Boosting with these ML-methods. The detection of ADLs, such as “cooking” or ”washing”, was performed by using a decision tree with a depth of 14 layers. The advantages of using the decision tree are the quality of the results and the real-time capability. Furthermore, an activity can also be determined if status values are missing. The dataset used here had only 781 data points, which is very small, and more data points should be collected in the future. With this new data, the decision tree model can be retrained with only the smart home sensor (omnidirectional camera) or other detailed/additional activities like taking medicine or walking/exercising. This new model could be easier and faster to install in a patient’s apartment. It is also less computationally intensive.

For further processing, the ADL determination can be used to detect anomalies in the everyday routines of the patient or for long-term analyses to identify dementia at an early stage. As post-processing of an anomaly detection, reminders such as taking pills can also be given. As a future goal of our work, the detected ADLs can be automatically documented and relieve the documentation work of professional caregivers in nursing homes.

References

[1] Alzheimer’s Disease International. Numbers of people with dementia worldwide https://www.alzint.org/resource/numbers-of-people-with-dementia-worldwide/. [Zugriff 29. Oktober 2021].

[2] Deutsche Alzheimer Gesellschaft. Selbsthilfe Demenz - Die Häufigkeit von Demenzerkrankungen https://www.deutsche-alzheimer.de/fileadmin/Alz/pdf/factsheets/infoblatt1_haeufigkeit_demenzerkrankungen_dalzg.pdf. [Zugriff 29. Oktober 2021].

[3] Lafuente-Arroyo, S., Martin-Martin, P., Iglesias, C. Maldonado-Bascon, S., Acevedo-Rodriguez, F. RGB camera-based fallen person detection system embedded on a mobile platform (2022)

[4] Brulin, D.; Benezeth, Y.; Courtial, E. Posture recognition based on fuzzy logic for home monitoring of the elderly. IEEE Trans. Inf. Technol. Biomed. 2012, 16, 974–982

[5] Yu, M.; Rhuma, A.; Naqvi, S.M.; Wang, L.; Chambers, J. A posture recognition-based fall detection system for monitoring an elderly person in a smart home environment. IEEE Trans. Inf. Technol. Biomed. 2012, 16, 1274–1286.

[6] Seidel, R. Apitzsch, A., Hirtz, G.: Improved Person Detection on Omnidirectional Images with Non-maxima Suppression, VISAPP 2019 (2019)

[7] Naeem, U., Bigham, J. & Wang, J. Recognising activities of daily life using hierarchical plans in European Conference on Smart Sensing and Context (2007), 175–189.

[8] Richter, J., Wiede, C., Dayangac, E., Heß, M. & Hirtz, G. Activity Recognition based on High-Level Reasoning in Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods (2016), 415–422.

[9] Chen, S., Liu, Y., Gao, X. & Han, Z. MobileFaceNets: Efficient CNNs for accurate real-time face verification on mobile devices in Chinese Conference on Biometric Recognition (2018), 428–438.

[10] Billestrup, J., Stage, J., Nielsen, L. & Nielsen, K. S. The use of personas in software development: Advantages, obstacles and experiences in practice in Proceedings IRIS36: August 11-14 2013 at Gran, Norway (Department of Informatics, University of Oslo, 2014).

[11] Everingham, M., Van Gool, L., Williams, C. K., Winn, J. & Zisserman, A. The pascal visual object classes (voc) challenge. International journal of computer vision 88, 303–338 (2010).

[12] Lin, T.-Y. u. a. Microsoft COCO: Common Objects in Context in European conference on computer vision (2014), 740–755.

[13] Scheck, T., Seidel, R. & Hirtz, G. Learning from THEODORE: A Synthetic Omnidirectional Top-View Indoor Dataset for Deep Transfer Learning in 2020 IEEE Winter Conference on Applications of Computer Vision (WACV) ISSN: 2642-9381 (März 2020), 932–941.

[14] Brownlee, J.: 8 Tactics to Combat Imbalanced Classes in Your Machine Learning Dataset. Machine Learning Mastery (2015).

[15] Rocca, J.: Ensemble methods: bagging, boosting and stacking - Towards Data Science. Towards Data Science (2019).

[16] Nishimura, H., Makibuchi, N., Tasaka, K., Kawanishi, Y., Murase, H.: Multiple Human Tracking Using an Omnidirectional Camera with Local Rectification and World Coordinates Representation; PAPER - Special Section on Machine Vision and its Applications (2020)

[17] Zhou, X., Wang, D. & Krähenbühl, P. Objects as points. arXiv preprint arXiv:1904.07850 (2019).

[18] Yu, J., Scheck, T., Seidel, R., Adya, Y., Nandi, D. & Hirtz, G., Human Pose Estimation in Monocular Omnidirectional Top-View Images, Computer Vision and Pattern Recognition, arXiv:2304.08186 (2023)

[19] Wang, Y.; Wang, X.; Arifoglu, D.; Lu, C.; Bouchachia, A.; Geng, Y.; Zheng, G. A Survey on Ambient Sensor-Based Abnormal Behaviour Detection for Elderly People in Healthcare. *Electronics* **2023**, *12*, 1539. https://doi.org/10.3390/electronics12071539