Intelligent Human Systems Integration (IHSI 2023), Vol. 69, 2023, 399–407
https://doi.org/10.54941/ahfe1002859

Recognition Model for Activity Classification in Everyday Movements in the Context of Dementia Diagnostics – Cooking

Sergio Staab, Lukas Bröning, Johannes Luderschmidt, and Ludger Martin

RheinMain University of Applied Sciences, Wiesbaden 65195, Germany

ABSTRACT

By monitoring movements and activities, the progression of neurological diseases can be detected. The documentation required for this is associated with a high level of effort, which is hardly possible in view of the increasing shortage of nursing staff. In order to gradually relieve the nursing staff, we are developing an approach to automate documentation in cooperation with two dementia residential communities. The aim of this work is to facilitate everyday life of caregivers. Previous research results from this working group show that everyday activities of dementia patients can be recognized well by combining smartwatch sensor technology and machine learning. However, the state of research has gaps when it comes to recognize activities consisting of a variety of movement patterns. In this paper, we present an approach to classify the activity of cooking. We divide this activity into several sub-activities each consisting of a distinct motion pattern that a recurrent network recognizes. This is followed by a model for calculating the probability that cooking actually occurred based on the different sub-activities recognized. We show the advantages of different smartwatch sensor combinations and compare the different approaches of our model with the prediction accuracy of the classification. This model can later be integrated into the care documentation of the residential communities in addition to the activities that are easier to recognize.

Keywords: Human motion analysis, Machine learning, Dementia

INTRODUCTION

The time-consuming documentation of patients activities is one important task of the caregivers especially in nursing homes for people with dementia. The documentation should help to understand what stage of the disease patients are in. Approaches exist allowing to achieve a digital nursing documentation with an online tool in which activities can be gathered by ticking boxes next to a list of activities (Staab and Martin, 2021). Although such an online tool simplifies the work of caregivers, it still has to be done manually. Furthermore, there are chances that due to high workload the staff overlooks certain activities. Building on four years of cooperation with two self-managed dementia living communities, in which we try to relieve the caregivers step by step, in this work we offer an approach to automate monitoring. The monitoring in the two dementia living communities refers to the documentation of daily activities of the patients over three shifts per day. The documentation should contribute to the understanding of which stage of the disease the patients are in.

This paper presents an approach to automatically recognize certain activities using smartwatches in combination with machine learning algorithms. More specifically, this work focuses on recognizing the arm movements of people wearing the watch and associating certain patterns of arm movements with activities using a recurrent neural network called Long Short-Term Memory (LSTM). We have been able to recognize everyday activities that consist of one distinct motion pattern such as eating, drinking, blowing one's nose, making phone calls in various previous works, for instance (Staab et al., 2021) or (Staab et al., 2022).

However, based on the feedback from the care teams, further important activities exist that do not consist of a clear pattern of movement. Activities like cooking, walking, baking, handicrafts, painting, playing games, memory training, helping with the laundry etc. are composed of a variety of movement patterns. Because of this, such activities are clearly more demanding to identify. In this paper, we want to present a model for recognizing the activity of cooking. For this, we divide the activity into nine sub-activities, which must be classified before a probability matrix decides if a person does cook or not.

In this framework, the paper makes the following contributions:

- An approach to recognize specific arm movement sequences using smartwatches.
- A way to classify and compare the recognition accuracy of motion sequences.
- Approach to recognize sub-activities to compound activity
- Matrix that recognizes compound activities
- Model for the recognition of cooking

This paper is structured as follows. Chapter "RELATED WORK" introduces related work in the field of health information technologies, human activity recognition, and recurrent neural networks. Chapter "PROJECT STRUCTURE" describes the project structure that includes the watchOS application, the sensory system, the data tracking setup, the aggregation of the sensor data as well as the machine learning tool, and the LSTM. Chapter "ACTIVITY RECOGNITION" shows the application and describes our approach. The results of the tracking process and the evaluation of the tracking process can be found in chapter "RESULTS AND DISCUSSION". Chapter "CONLUSION" summarizes the work and a look into the future.

RELATED WORK

Cooking, according to (Alia et al., 2020), is an activity that indicates people's cognitive health and their ability to live independently. For this reason, the authors consider the monitoring of cooking to be an important use case.

In this context, cooking is a complex activity that usually consists of several sub-activities. Looking at previous works dealing with the classification of cooking, (Dernbach et al., 2012), for instance, define the activity cooking precisely as a simple movement process of getting water from a refrigerator and putting it into a bowl to be heated in the microwave. This shows that cooking is too simplistically broken down to one process, which does not do justice to its complex structure and diversity. The general application of models to classify cooking is useful in practice only if many different sub-activities subordinate to cooking are considered.

An approach that goes into more detail about cooking and its sub-activities was developed by (Duong et al., 2009). The authors divide cooking into 12 sub-activities and examine three preparation processes to be classified, including one breakfast and two lunches. According to the authors, the three cooking activities differ little; they share the same sequence of sub-activities but differ in the duration of their tasks. All three activities follow twelve fixed sequential steps that include various tasks in cooking such as gathering ingredients from the refrigerator, taking them to the stove, washing vegetables, cooking, setting the table, eating, cleaning the stove, washing dishes, and so on. Duong (Duong et al., 2009) classify cooking based on both the typical duration of sub-activities and their inherent hierarchical structures and temporal dependencies using variants of the Hidden Markov Model. The authors argue to achieve high and comparable accuracy and outperform baseline models for the task of activity segmentation.

In their summary of the Cooking Activity Recognition Challenge, (Alia et al., 2020) provide an overview of possible approaches to classify the activity of cooking. According to the authors, cooking was divided into the three macro-activities preparing sandwich, fruit salad, and cereal and nine micro-activities (cutting, taking, mixing, adding, miscellaneous, pouring, opening, peeling, washing). The micro-activities are very similar, which, according to the authors, increases the difficulty of identifying the correct macro-activities. The creation of the dataset was done using acceleration data from each of two human-mounted smartphones and smartwatches, as well as using data from a motion capture system and a pose detection system. Multi-class classification was used predominantly for macro-activity classification and multi-label classification for micro-activities. The average accuracy achieved by the teams is in the range of 32 % to 92 %. For macro-activities, the highest accuracy was 100 %, and for micro-activities, the highest accuracy was 84 %. (Alia et al., 2020) conclude that micro-activity detection is easier, which leads to higher average accuracy. Regarding the models used, the authors summarize that machine learning algorithms were superior to deep learning models in both training and testing.

Previous work shows that existing approaches can hardly be transferred to other use cases. A more general classification of cooking has to be established, which is not designed for the recognition of specific recipes, but enables the recognition of cooking through the recognition of various general sub-activities.

To the best of our knowledge, no work has yet taken the approach of using a single smartwatch to provide sensor data for cooking classification.

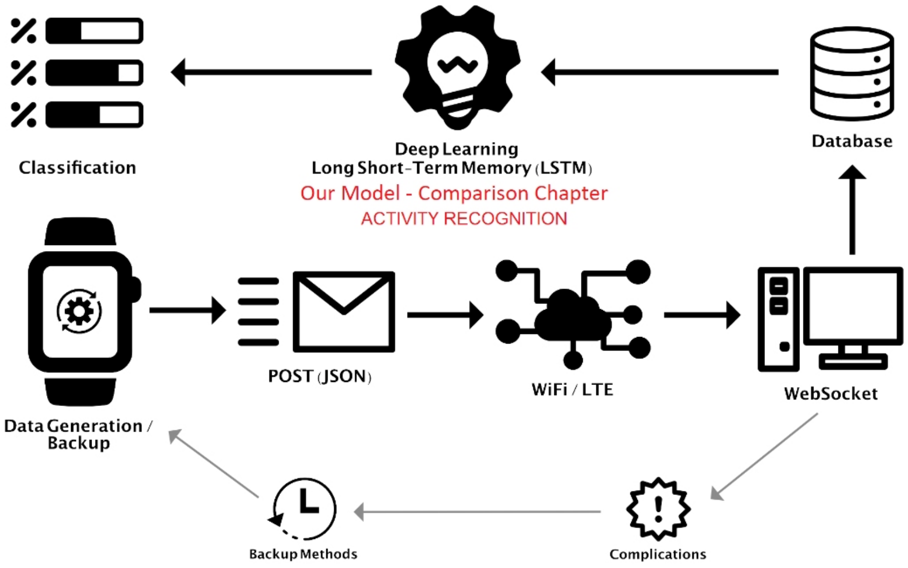

Figure 1: Overview of the developed Human-Activity-Recognition system, which includes data generation, data transfer and storage, visualization, and classification of data.

However, smartwatches are particularly well suited for integrating tracking into the daily lives of people in need of care, as we have found in consultation with our collaborating dementia living communities. Our work addresses these issues by developing a holistic LSTM-based activity recognition system. Tracking is done with smartwatches, which can be comfortably worn like regular watches in everyday life. In addition, the focus is on the automatic documentation of fine-grained activities, in this case cooking, performed by people with dementia in everyday life. Accordingly, our work is an extension of the aforementioned work. In the following, we will explain our prototype in more detail.

PROJECT STRUCTURE

This chapter gives an overview of the structure of our project. Figure 1 illustrates the data generation using a smartwatch, the communication between the watch and the server and the further processing of the data using machine learning.

The process shown in Figure 1 begins with an Apple Watch Series 7 collecting and providing movement data via an interface. The watch, equipped with advanced sensor technology, forms the basis of our work. The accelerometer (x,y,z, in g), the gyroscope (rotational speed, in rad/s,), the attitude sensor (pitch, roll and yaw as specific attitude angles, in rad), gravity (x,y,z, in 9.81 m/s^2), and heart rate (pulse per minute, as an integer) are used for fine-grained motion data collection. Since the goal of this work is a continuous data stream, a continuous query is logically implemented. The continuous polling of the movement data takes place in an individually adjustable time

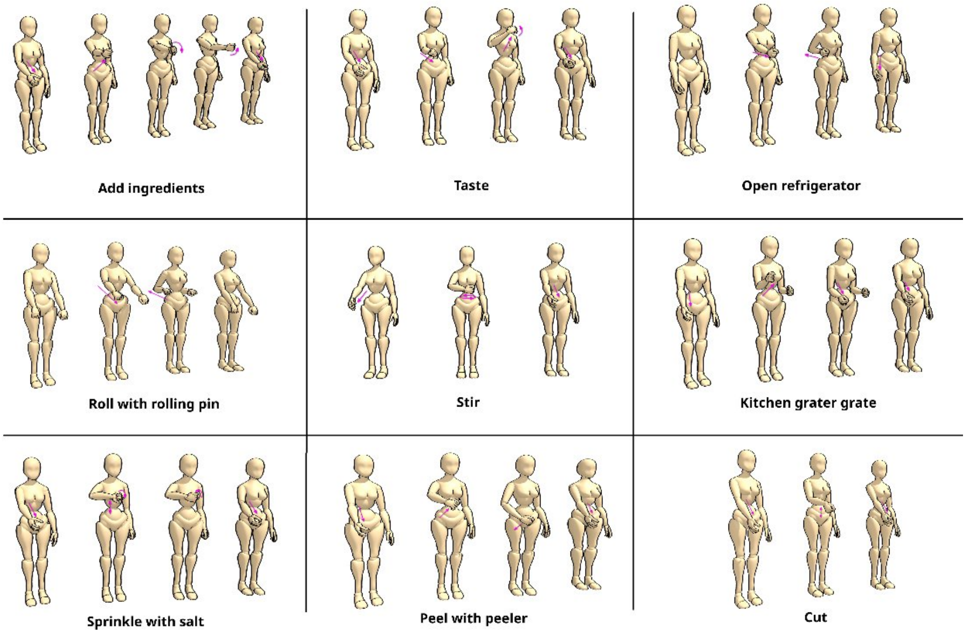

Figure 2: Nine sub-activities that subdivide the activity of cooking into different movement patterns.

interval, the frequency. The maximum frequency is 100 hertz (Hz) when using an Apple Watch Series 7, which means that a maximum of 100 motion data objects can be generated per second. At this point, it must be questioned which frequency of the query meets the requirements of this work. If more objects are generated than needed, this not only leads to redundancies, but also to a noticeable increase in the load on the CPU and battery due to the increased number of queries. In this work, a standard frequency of 20 Hz is therefore initially set, especially since, according to (Dadafsha, 2014), the natural movements of a human do not exceed 12 to 20 Hz. Thus, one object with 13 data sets (12 movements and heart rate) is generated per Hz, which corresponds to 260 (13*20) sensor data per second entering the system. Data generation is followed by the machine data analysis on which our model is based.

ACTIVITY RECOGNITION

The aim of this work, as already mentioned, is simple to formulate. How do we recognize that a person is cooking, that is, performing the activity of cooking? The activity of cooking can be reflected in different patterns of movement and does not follow a perpetual pattern. Therefore, the task of recognition is very challenging. Our first step is to divide the activity of cooking into nine sub-activities. These are tasting, adding ingredients, grating, opening the fridge, rolling with a rolling pin, stirring, sprinkling salt, cutting, and peeling with a peeler. Figure 2 shows the movement patterns of the individual sub-activities.

In the following, the individual sub-activities are briefly explained from top left to bottom right.

Adding ingredients: The measured ingredient is grasped with the right hand and held in front of the chest with the arm extended. The ingredient is tipped into a container by turning the hand 90 degrees. The hand is then rotated back to the starting position and the arm lowered again.

Tasting: Tasting begins by picking up the spoon with the right hand. With a lowering movement, the right arm guides the spoon into the food and then to the mouth, where it remains for two seconds. With a tilting movement of the hand, the food is picked up and the arm is lowered again.

Opening the refrigerator: To begin, the right arm is stretched away from the body and pulled towards the chest. Then the arm is lowered again.

Rolling with a rolling pin: The rolling pin is held by the hand pieces with both hands and repeatedly moved away from and towards the body.

Stirring: The stirring sub-activity consists of reaching forward with the right hand to pick up the utensil, making a smooth circular counterclockwise movement of the arm and then lowering the arm to put the utensil down.

Kitchen grater: The right arm is extended to reach for the food. The hand is raised and placed on the upper part of the grater. Then the arm is repeatedly moved down and back to the upper part of the grater. Finally, the right arm is lowered again.

Salting: First the salt shaker is grasped with the right hand. Then the arm is raised to chest level and the hand is rotated about 90 degrees. Then the hand is jerked up and down. Finally, the hand is turned back to the starting position and lowered.

Peeling with a peeler: The peeler is grasped with the right hand and held at chest level. The arm is then moved down and up again several times. As the arm is lowered, the peeler is finally set down.

Cutting: To begin, the right arm reaches forward for the knife. The right hand is now moved up and down again several times from the cutting board. Then the knife is put down with the lowering of the arm.

We performed these nine activities with nine subjects 20 times each for 10 seconds. Because of the 13 features of the sensor system and the 20 Hz, the training is therefore based on 13 (figures) * 10 (seconds) * 20 (Hz) * 20 (executions) * 9 (participants) * 9 (sub-activities) = 4,212,000 data sets. Table 1 shows which sensor combinations achieve which validation accuracy with the trained LSTM with respect to all sub-activities. The combination of Acceleration, Attitude, Gyro and Gravity achieves the best result with 61,94 % prediction accuracy.

It should be noted that the heart rate with poor predictions has no added value for the classification of the cooking activity, so this feature was excluded from the table. Our model assumes that the activity cooking involves several sub-activities in a time frame of 10 minutes. The further approach is thus a classification pattern built over a zenith distance with weights of probabilities and predictive accuracies of the sub-activities towards the statement cooking, yes or no. Our model is an approximation in terms of direct food handling and processing. For this purpose, the sub-activities are divided into two categories. The first category includes sub-activities that require direct contact

Table 1. Comparison of different sensor combinations in terms of the predictive accuracy of the LSTM with respect to nine different human activities.

Sensor combination	Validation accuracy (LSTM)
Acceleration	11.93 %
Attitude	44.40 %
Gyro	11.93 %
Gravity	49.91 %
Acceleration, Attitude	42.02 %
Acceleration, Gravity	46.33 %
Acceleration, Gyro	11.93 %
Attitude, Gyro	46.67 %
Attitude, Gravity	47.41 %
Gyro, Gravity	42.93 %
Acceleration, Attitude, Gyro	46.30 %
Acceleration, Attitude, Gravity	47.61 %
Acceleration, Gyro, Gravity	51.46 %
Attitude, Gyro, Gravity	45.58 %
Acceleration, Attitude, Gyro, Gravity	61.94 %

with food or that have to be performed over a longer period of time compared to the other sub-activities. These include the following sub-activities: Rolling with a rolling pin, stirring, grating, slicing and peeling with a peeler. The second category includes the sub-activities that are also related to cooking but do not require any interaction with the food or do not require a long interaction with the food compared to the first category. These include the other sub-activities: Opening the fridge door, sprinkling salt, seasoning and adding ingredients. To identify the cooking activity from the previously defined sub-activities, the activity is compared over a period of 10 minutes (Hwan-Hee, 2015). Within the 10 minutes of active cooking, the percentage of certainty with which partial activities are recognized is listed in 10-second intervals. It follows that over a 10-minute period, 60 of these listings are produced. In order to make a general statement about the entire period, the number of individually detected partial activities is divided by the sum of their probabilities over all 60 listings. The threshold for recognizing a partial activity is 60 %, which was also demonstrated in the latter work (Staab et al., 2022).

RESULTS AND DISCUSSION

The result is a list of all detected partial activities with the associated probability over the entire 10 minutes. From this list, it can now be derived whether it is a cooking activity. As already described, this definition is based on the assumption that cooking aims at the direct handling and processing of food. For this, at least one sub-activity of the first category must be recognized >= 90 percent and any two further sub-activities, i.e. at least 60 %. The high recognition probability of at least one partial activity from the first category ensures that it is active cooking, i.e. the direct handling of food. The recognition of two further sub-activities minimizes the danger that one sub-activity is

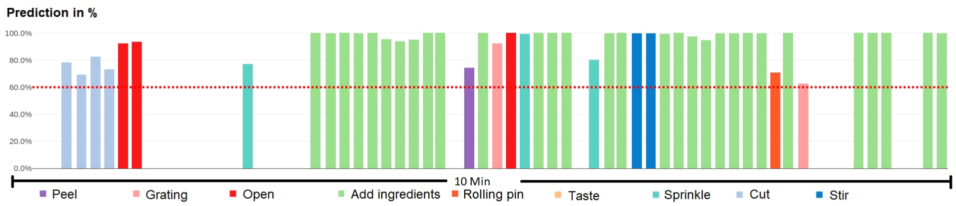

Figure 3: Activities Threshold at a prediction value of 60 % over 10 min cooking.

often falsely recognized with a high probability and thus falsifies the categorization. In addition, the activity of cooking often consists of several composite sub-activities. Figure 3 shows the activity of the cooking over a period of 10 minutes, with a prediction threshold of 60%.

Cooking can also consist of only one sub-activity, which, however, as already described, increases the risk of incorrect categorization. Therefore, only cooking processes that consist of several sub-activities are considered in this work.

CONCLUSION

In this work we have dealt with problems of motion recognition of everyday activities that do not consist of a single motion pattern. We have taken on the challenging classification of the activity of cooking that we now can classify automatically with our system. This system is based on a neural network called Long Short Term Memory, for which the best sensor data combinations of a smartwatch were evaluated for classification using cross-validation. We divided the cooking activity into nine sub-activities, based on which an algorithm was then developed over a 10-minute time interval to weight, select and rank the sub-activities according to their own prediction probabilities. Our model detected the activity of cooking 9 times out of 10, giving our system a prediction accuracy of 90%. Currently, the system can only say yes or no. In a next step, we will add another main activity and test the system for stability.

REFERENCES

Alia, S. S., Lago, P., Takeda, S., Adachi, K., Benaissa, B., Ahad, M. A. R., & Inoue, S. (2020). Summary of the Cooking Activity Recognition Challenge. In Human Activity Recognition Challenge (pp. 1–13). Springer Singapore. https://doi.org/10.1007/978-981-15-8269-1_1

Dadafshar, M. (2014). Accelerometer and gyroscopes sensors: operation, sensing, and applications. Journal: Maxim Integrated.

Dernbach, S., Das, B., Krishnan, N. C., Thomas, B. L., & Cook, D. J. (2012). Simple and Complex Activity Recognition through Smart Phones. In 2012 Eighth International Conference on Intelligent Environments. 2012 8th International Conference on Intelligent Environments (IE). IEEE. https://doi.org/10.1109/ie.2012.39

Duong, T., Phung, D., Bui, H., & Venkatesh, S. (2009). Efficient duration and hierarchical modeling for human activity recognition. In Artificial Intelligence (Vol. 173, Issues 7–8, pp. 830–856). Elsevier BV. https://doi.org/10.1016/j.artint.2008.12.005

Hwan-Hee, K., Effects of experience-based group therapy on cognitive and physical functions and psychological symptoms of elderly people with mild dementia. In: Journal of Physical Therapy Science 27.7 (2015), S. 2069–2071. DOI: 10.1589/jpts.27.2069

Staab, S., Luderschmidt J., Martin L. (2021). Recognition of Usual Similar Activities of Dementia Patients via Smartwatches using Supervised Learning, International Conference on Progress in Informatics and Computing, In: *2021 IEEE International Conference on Progress in Informatics and Computing (PIC)*. IEEE, 2021. S. 397–402.

Staab, S., Martin L. (2021). Informationsplattform INFODOQ, 56. Jahrestagung der Deutschen Gesellschaft für Sozialmedizin und Prävention, 56. Wissenschaftliche Jahrestagung der DGSMP 2021–digital. *Gesundheitswesen*, 2021, 83. Jg., S. 566–567.

Staab, S., Bröning L. Luderschmidt J., Martin L. (2022). Live Activity Recognition in Dementia Patients with Smartwatch Sensor Technology Using Long Short Term Memory, In: 24th International Conference on Human-Computer Interaction. Purdue University, USATsinghua University, PR Chinaand University of Central Florida, USA, HCII, Gothenburg Sweden. 2022. S. 315–322.